Thesis submitted for the degree of Doctor of Philosophy to

the Department of Chemistry, University of Sheffield

Spectroscopic evidence for catalytically-required FEN1-mediated DNA conformational change; a novel strategy for FEN1 inhibition

Supervisor: Prof. Jane Grasby

Jack Exell

February 2015

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.

Jack Exell

Acknowledgements

Abstract

Flap endonuclease 1 (FEN1) is a member of the 5'-nuclease superfamily of DNA replicative and repair proteins. It acts as a sequence-independent nuclease whose divalent metal ion-dependent phosphodiesterase activity specifically cleaves 5'-bifurcated nucleic acids at the first nucleotide in the duplex region. FEN1's integral role as a replicative enzyme, combined with upregulation in cancer cells, makes it an attractive chemotherapeutic target.

FEN1 has been proposed to select the correct phosphate diester for reaction by using a double-nucleotide-upairing (DNU) mechanism. This precise reaction site selection is key to allowing DNA repair and replication process to continue seamlessly. Evidence for this mechanistic step is largely inferred from structural data. This thesis describes characterisation of mutations of conserved residues that may be involved in the FEN1 reaction mechanism. Combined kinetic and biophysical data revealed that these residues were important to reaction but not substrate binding. A spectroscopic study of the DNU process was undertaken using DNAs containing two tandem 2-APs, to directly observe the local conformation of DNA in solution using exciton coupling CD. A clear change in signal was observed upon hFEN1-Ca^{2+} binding to substrate and product DNAs containing this spectral probe. Gross-conformational changes of nucleotides at the site of incision were dependent on the presence of active-site divalent metal ions. Conserved residues, Y40, K93, D181A and R100, were proposed to be important for DNU, their involvement was assessed by creating alanine variants, identifying them as important for product, but not substrate, upairing. These catalytic residues reside at the base of the helical arch which traverses the active site. The introduction of leucine to proline mutations at several intervals in the arch sequence more distal from the active site revealed that the structural integrity of the top of the arch was crucial for DNU, suggesting proper ordering of the arch is essential to bring about a DNA conformational change.

Recent screening identified a *N*-hydroxy urea series whose FEN1 inhibitory activity is thought to derive from a metal-chelating 'warhead'. Isothermal titration calorimetry confirmed active-site magnesium ion dependent nanomolar affinity of these inhibitors. Kinetic analyses permitted discrimination between reversible inhibition models. It was found that a mixed inhibition model fitted the data with overwhelming probability, indicating the formation of an enzyme-substrate-inhibitor (ESI) complex. Fluorescence anisotropy and FRET experiments independently confirmed that DNA and *N*-hydroxy urea inhibitors are bound to the same hFEN1 molecule, consistent with the kinetic data that implies inhibitor bound in the active-site only moderately

decreases the affinity of the substrate for hFEN1. Finally, ECCD of the ESI complex confirmed that these inhibitors prevent reaction by sterically blocking the DNU step.

Publications arising from work covered in this thesis

Finger, L. D., Patel, N., Beddows, A., Ma, L., **Exell, J. C.**, Jardine, E., Jones, A. C., Grasby, J. A. (2013). Observation of unpaired substrate DNA in the flap endonuclease-1 active site. Nucleic Acids Research *41*, 9839-9847.

Patel, N., Exell, J. C., Jardine, E., Ombler, B., Finger, L. D., Ciani, B., Grasby, J. A. (2013) Proline scanning mutagenesis reveals a role for the flap endonuclease-1 helical cap in substrate unpairing. The Journal of Biological Chemistry *288*, 34239-34248.

Contents

Abbreviations

K_D	dissociation constant
K_M	Michaelis constant
K_i	inhibition constant
K_{ic}	competitive inhibition constant
K_{iu}	uncompetitive inhibition constant
k_{ST}	single-turnover rate constant
k_{cat}	catalytic rate constant
2-AP	2-aminopurine
aa	amino acid
AIC	Akaike's Information criterion
Ala / A	alanine
Arg / R	arginine
Asp / D	aspartic acid
BER	base excision repair
bp	base pair
BSA	bovine serum albumin
CD	circular dichroism
Cm	chloramphenicol
DDR	DNA damage repair
dHPLC	denaturing high pressure liquid chromatography
DMSO	dimethyl sulfoxide
DNAP	DNA polymerase
dNTP	Deoxynucleoside triphosphate
double flap	double flap
DR	direct repair
dRP	deoxyribose phosphate
ds	double stranded
DTT	dithiothreitol
E	enzyme, macromolecule
ECCD	exciton coupling circular dichroism
EDTA	ethylenediaminetetraacetic acid
EMSA	electrophoretic mobility shift assay
EP	enzyme-product
ES	enzyme-substrate
ESI	enzyme-substrate-inhibitor complex
ETDM	electronic transition dipole moments
FA	fluorescence anisotropy
FAM	fluorescein
FRET	fluorescence resonance energy transfer
HEPES	hydroxyethylpiperazineethane sulphonic acid
His / H	histidine
HJ	Holliday junction
HR	homologous recombination

I	inhibitor
IMAC	immobilised metal ion affinity chromatography
IPTG	isopropyl β D-I-thiogalactopyranoside
ITC	isothermal titration calorimetry
Kan	kanamycin
LB	Luria-Bertani broth
Leu / L	leucine
LFER	linear free energy relationship
LpBER	long-patch base excision repair
Lys / K	lysine
MALDI-TOF	matrix assisted laser desorption ionisation-time of flight
MMR	mismatch repair
MWCO	molecular weight cut off
n	number of identical binding sites per macromolecule
NaOH	sodium hydroxide
NER	nucleotide excision repair
nt(s)	Nucleotide(s)
OD	optical density
P, Q	product
PARP1	poly (ADP-ribose) polymerase 1
PCNA	proliferating cell nuclear antigen
Phe / F	phenylalanine
Pro / P	proline
pY	pseudo Y
r	anisotropy
RRB	reduced reaction buffer
RT	room temperature
SA	streptavidin
SAR	structure activity relationship
SDS-PAGE	sodium dodecyl sulphate- polyacrylamide gel electrophoresis
sm	Single molecule
SOC	super optimal broth with catabolite repression
SpBER	short-patch base excision repair
ss	single stranded
ST	single-turnover
TAMRA	tetramethyl rhodamine
TBE	TrisborateEDTA
TBSE	Tris borate saline EDTA
TOF-ESI	time of flight electrospray ionisation
Tris	tris(Hydroxymethyl)aminomethane
Tyr	tyrosine
Tyr / Y	tyrosine
WT	wild-type

Chapter 1: introduction

1.1: Phosphodiester hydrolysis

The hydrolysis of phosphate diester bonds contained within nucleic acids in processes such as DNA replication, recombination and repair, RNA maturation and apoptosis is essential for all life. Phosphodiesterases that catalyse phosphate diester bond cleavage in nucleic acid polymers are generally referred to as nucleases but their sites of action or the specific cellular function of the hydrolyase can vary. Thus, nucleases can be endonucleolytic or exonucleolytic, and they can be DNases or RNases, topoisomerases, recombinases, helicases, ribozymes or RNA splicing enzymes. Historically nucleases have been classified according to reaction mechanism, metal-ion dependence, topology or a combination of these categories (Horton, 2008; Yang, 2011).

1.1.1 Chemistry of Phosphoryl-transfer reaction

In DNA and RNA structures, there are two possible P-O bonds that can be cleaved. Hydrolysis of one of the two bridging P-O bonds dictates the specific products. Either a 5'-phosphate monoester and 3'-hydroxyl or a 3'-phosphate monoester and 5'-hydroxyl products are generated. The majority of nucleases that act on DNA cleave the 3'-O-P bond, resulting in a deoxynucleoside 5'-monophosphate and a free 3' hydroxyl. The reason being is that the 5'-phosphate is an ideal substrate for DNA ligation which completes replication, recombination and repair processes.

Phosphorous is able to exist in either tri-coordinate, tetra-coordinate or penta-coordinate states. Phosphoryl transfer reactions are substitution reactions at phosphorus and can occur by three mechanisms (*figure 1a*). One is a dissociative (S_N1) mechanism, where there is first elimination of a leaving group to form an unstable tri-coordinate metaphosphate intermediate followed by rapid addition of a nucleophile (*figure 1a*). The other two mechanisms are associative (S_N2). Both possibilities have been demonstrated to occur in nonenzymatic phosphoryl-transfer reactions (Kirby and Jencks, 1965). The associative (S_N2) mechanism can be either a concerted or multi-step process. In a purely associative mechanism, usually involving formation of a covalent enzyme-phosphate intermediate, a highly negatively charged pentacovalent intermediate, called a phosphorane is formed, first through addition of a nucleophile, followed by elimination of the leaving group; also known as an addition-elimination pathway. In the concerted mechanism bond formation and fission of the nucleophile and leaving group respectively occur simultaneously in the transition state. The concerted mechanism is

complicated further depending on whether the transition state is deemed as 'loose', 'tight' or 'synchronous'. A transition state that is close to the dissociative limit is denoted as 'loose' while a 'tight' transition state close to a purely associative mechanism making a 'synchronous' transition state in the middle of loose and tight, shown graphically in a More-O'Ferrall-Jencks diagram (*figure 1.1b*) (Jencks, 1972; Oferrall, 1970).
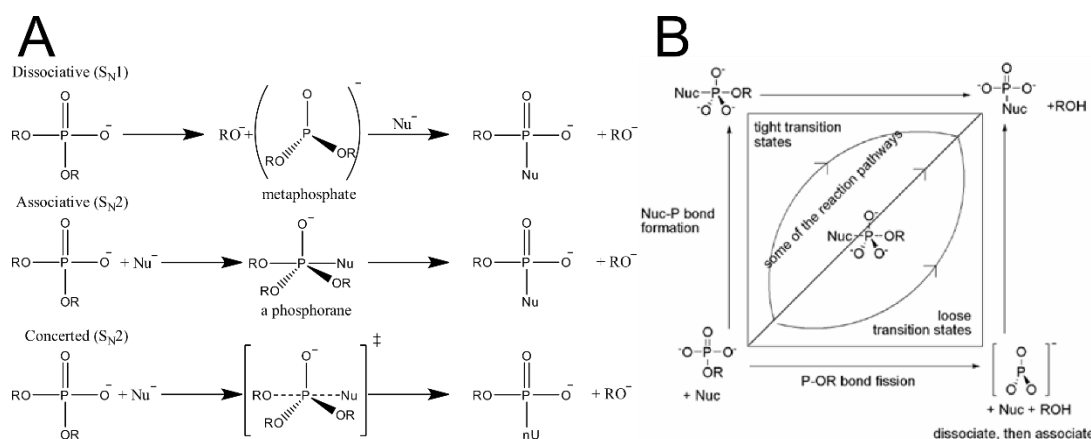
Figure 1.1: Transition state structures of phosphate diester hydrolysis (A) Associative, dissociate and concerted phosphate diester hydrolysis reaction mechanism. (B) More O'Ferrall-Jencks diagram with phosphorus-leaving group bond fission shown along the horizontal axis and phosphorus-nucleophile bond formation along the vertical axis.

1.1.2 Tools for Mechanistic Analysis

The experimental differentiation between a reaction pathway via a metaphosphate intermediate and a penta-coordinate transition state is extremely difficult. Gathering evidence from several experimental techniques is therefore vital in order to elucidate the reaction mechanism. Following the stereochemical course of a reaction is one such technique. Extensive work using chiral phosphoryl compounds in conjunction with ^{31}P NMR has been carried out to characterise the reactions of phosphodiesterases (Connolly et al., 1984; Eckstein, 1985). The majority of nuclease catalysed reactions proceed with inversion at phosphate, ruling out the formation of covalent enzyme-substrate intermediates (Gerlt et al., 1983). Specific examples include P1 nuclease (Potter et al., 1983), Klenow fragment of *Escherichia. Coli* (*E.coli*) DNA Polymerase I (Gupta and Benkovic, 1984), and Type II restriction endonucleases (Grasby and Connolly, 1992). Although inversion of the stereochemical centre is possible through a S_N1 mechanism since enzymes control the spatial arrangement of the substrate, formation of an

unstable metaphosphate intermediate is unlikely. Additionally, the widespread presence of metal ions in nuclease active-sites (Steitz and Steitz, 1993), which will be discussed in detail later on (*section 1.2*), would favour formation of a penta-coordinate transition state. Nucleases that catalyse reactions with retention of configuration at phosphorus are generally are thought to use a double S_N2 displacement mechanism by forming a covalent-enzyme intermediate that is subsequently attacked by water.

Linear free energy relationships (LFER), in particular Brønsted plots is/are another technique of physical organic chemistry which has been used to study phosphoryl-transfer reactions (Williams, 1992). The rate limiting step in a S_N1 mechanism is fission of the leaving group bond, whereas in a concerted S_N2 mechanism it is a combination of leaving group bond fission and attacking group bond fusion. Studies suggest that enzyme catalyzed phosphoryl transfer reactions generally proceed through transition states that are similar to their solution counter parts (Zalatan and Herschlag, 2006). The nonenzymatic reactions of a wide range of nucleophiles with a series of aryl methyl phosphate diesters were found to show sensitivity to both the nucleophile and the leaving group (Hengge, 2005; Liao et al., 2001; Williams et al., 1998). The observation of a significant dependence of the rate on both leaving group and nucleophile suggests a concerted S_N2 mechanism for phosphate diester hydrolysis in solution. When compared with solution studies the phosphodiesterase activity of alkaline phosphatase and the reaction catalysed by *Tetrahymena* ribozyme are predicted to go through a synchronous enzymatic transition state, with bond breaking and forming occurring simultaneously (Zalatan and Herschlag, 2006).

Collectively these data allow for a generic phosphodiester hydrolysis reaction mechanism to be proposed for the majority of nucleases, whereby inversion is achieved through a concerted S_N2 mechanism which involves formation of a synchronous transition state through direct in-line attack of the phosphate by water, with no need to invoke an enzyme-substrate covalent intermediate. However, there is still uncertainty surrounding the timing of the proton transfer in phosphodiester hydrolysis (McWhirter et al., 2008). It has been postulated that protonation of the leaving group accompanies deprotonation of the phosphoryl group because of the high concentration of water molecules in the active site (Kirby and Varvogli.Ag, 1967).

1.2: Metal-ion catalysis

1.2.1. Nuclease catalysis

Catalysis of phosphodiester hydrolysis can be achieved by deprotonation of the attacking group, neutralisation of charge of the electrophilic phosphate diester and its corresponding transition state or intermediate, and ultimately protonation of the leaving group. Considering the tools enzymes have to their disposal this type of catalysis can be achieved in a number of ways. Rate enhancement can be brought about by the direct participation of amino acid side chains acting as general acid and general base catalysts. Metal ions also have the potential to participate in enzymatic phosphodiester hydrolysis. They can also act as a general acid or general base by providing metal-bound water and metal-bound hydroxide, respectively. In addition, metal ions are able to lower the activation energy of the reaction by acting as a Lewis acid catalyst or by leaving group activation. In reality metal ion coordination to two or more of the reacting groups usually means that they perform a combination of general acid-base catalysis, electrophilic catalysis and leaving group activation.

There are a few examples of metal-ion independent RNases and DNases. RNase T1 and T2 are two large families along with RNase A which utilise His/Glu and two His residues respectively to perform general acid-base catalysis (Deshpande and Shankar, 2002). DNases found to be metal-ion independent are generally part of the phospholipid D (PLD) superfamily characterised by a HKD motif. DNAse II and Nuc of *Salmonella typhimurium* represent two enzymes of the PLD superfamily (Gottlin et al., 1998). Topoisomerases and site-specific recombinases also fall into this category. A larger proportion of nucleases display a one metal ion-dependency, whereby the metal ion acts as an electrophilic catalyst and stabilizes the leaving group and the role of the general base is still occupied by an amino acid side chain. Nucleases belonging to this category tend to be DNA endonucleases (Yang, 2011). Larger still is the class of two metal ion-dependent nucleases. There are over a dozen superfamilies which fall into this category and they encompass sequence specific and non-specific nucleases, endo- and exonucleases with either 3'-5' or 5'-3' polarity.

One family of enzymes that have an absolute requirement for divalent metal ion cofactors are the flap endonucleases (FENs) (*section 1.3*). FENs catalyse the hydrolysis of nucleic acids at key steps during DNA replication and repair, and are present in organisms as diverse as bacteriophage and humans. Members of the FEN1 family have a 5' to 3' exonuclease activity in addition to an endonuclease activity (Harrington and Lieber, 1994; Lyamichev et al., 1993). It is

not surprising that FENs are capable of this dual activity because of the similarity in chemistry of exo- and endo- phosphodiester hydrolysis. This review will concentrate largely on two metal ion-dependant DNA nucleases involved in replication and repair.

1.2.2 Chemical properties of divalent cations

Divalent cations exist in living organisms at differing concentrations. However, Mg^{2+} has by far the highest intracellular concentration (Lagomarsino et al., 2010), and is most frequently associated with nucleic acid enzymes (Cisneros et al., 2009; Yang, 2008). The preferred binding of Mg^{2+} is perhaps because of its solubility, redox stability, strict coordination geometry and its high charge density ratio in addition to its availability. In particular, Ca^{2+} and Zn^{2+} can have an octahedral coordination but both have coordination freedom, Zn^{2+} is frequently found in a tetrahedral geometry, and Ca^{2+} bound by 7, 8 or 9 ligands. Magnesium's coordination chemistry contrasts with this as it is almost always found to have six inner-sphere ligands arranged in an octahedral geometry.

The fundamental property of divalent metal ions exploited by nucleases is their high charge density making them able to neutralize the highly negatively charged phosphodiester backbone. The stringent coordination geometry and high charge to density ratio of Mg^{2+} means that water molecules preferential bind over larger bulkier ligands. Thus, water molecules are often retained upon complex formation with enzymes and provide the nucleophilic active site water/hydroxide molecule required for phosphodiester hydrolysis. Direct binding of metal ions through protein side chains, typically those containing carboxylate groups is classed as inner sphere coordination. Alternatively, amino acids residues are able to coordinate Mg^{2+} through hydrogen bonding networks by indirect interactions with metal bound waters (outer-sphere binding). Similarly, activation of the nucleophile and stabilisation of the substrate can be mediated through inner- or outer-sphere binding (*figure 1.2*).
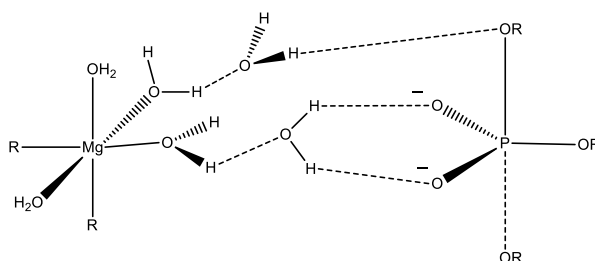
Figure 1.2: Transition state stabilisation by outer sphere coordination shell of hydrated magnesium ion. *R represents carboxylate residues of enzyme active site.*

1.2.3 A two metal ion phosphoryl transfer mechanism

The majority of metallonucleases are proposed to catalyze reactions using a two metal-ion mechanism analogous to that suggested for the Klenow fragment of DNA Polymerase I (Beese and Steitz, 1991), although this is not universally accepted. DNA Polymerase I (Pol I) is a member of the DNA replication and repair proteins catalysing template-directed DNA synthesis, 3'-5' proofreading hydrolysis and 5'-3' hydrolysis through three active sites each requiring divalent metal ions such as Mn^{2+} and Mg^{2+} for activity. Removal of the 5'-3' nuclease domain corresponds to the large proteolytic (Klenow) fragment of the enzyme (Klenow and Overgaar.K, 1970).

Several crystallographic studies of enzyme substrate complexes containing DNA bound at the 3'-5'-exonuclease site of the Klenow fragment of Pol I have provided insight into the divalent metal cation dependent catalysis mechanism. The active-site is proposed to contain two-metal ion binding sites whose occupancy is dependent on the presence of nucleoside monophosphate (NMP) (Derbyshire et al., 1988; Ollis et al., 1985). The two bound metal ions are separated by 3.9 Å in the wild-type protein and nucleoside monophosphate crystal structures, central to the proposed mechanisms of two-metal-ion catalysis.

Beese and Steitz (1991) refined earlier proposed two-metal ion mechanisms based on previous structures using a high resolution crystal structure image of the Klenow fragment complexed with product in the 3'-5'-exonuclease site (*figure 1.3*) (Beese and Steitz, 1991; Freemont et al., 1988). The mechanism is essentially a S_N2 mechanism whereby the attacking and leaving groups take up apical positions, implying an associative in-line mechanism, resulting in inversion of configuration at phosphorous. This is consistent with previous biochemical data (Gupta and Benkovic, 1984). A pentacoordinate transition state can be inferred in which the one non-bridging oxygen is contacted by both metal ions thereby stabilising the increasing negative charge. One of the metal ions acts as a Lewis acid catalyst facilitating the formation of a putative attacking hydroxide ion by lowering the pKa of bound water and also binds to one of the non-bridging oxygens acting as an electrophilic catalyst (Derbyshire et al., 1991). The other metal ion, in addition to acting as an electrophilic catalyst, facilitates departure of the 3' hydroxyl group.
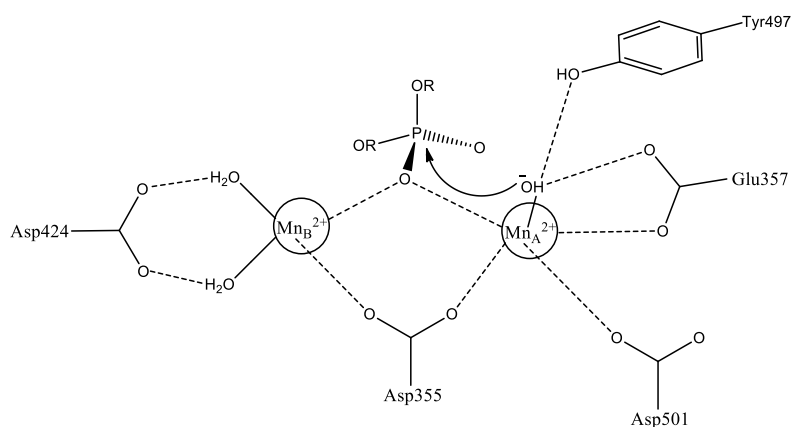
Figure 1.3: Active site of the 3'-5' exonuclease domain of the Klenow fragment, *depicting the essential active site residues and metal ions required for 2-metal-ion mechanism of phosphate diester hydrolysis. The high affinity metal ion site A (Mn_A^{2+}) and metal ion site B (Mn_B^{2+}) are coordinated by the carboxylate residues (Asp424/355/501) and the orientation of the attacking hydroxyl is mediated by hydrogen bonding to Glu357 and also Tyr497.*

This 'two-metal ion mechanism' has now become a template for other enzymes whose active site metal cofactors are responsible for rate enhancement. Mechanistic proposals for two metal ion catalysis are largely based of crystallographic studies as described for the Klenow fragment above and other examples include, EcoRV endonuclease (Kostrewa and Winkler, 1995), RNase H (Nowotny et al., 2005), and the double strand break repair nuclease Mre11 (Hopfner et al., 2001), although in some cases variants of the two metal ion mechanism are invoked.

1.2.4 Substrate dependent metal ion binding

As mentioned earlier despite there being considerable experimental support for the aforementioned mechanism there is still considerable debate regarding whether one, two or three metal ions are required, the apurinic/apyrimidinic endonuclease 1, APE1 exemplifies this particular argument (Tsutakawa et al., 2013). The one factor which there is agreement on is that there is a complicated relationship between substrate binding and metal coordination. Several crystal structures of Type II endonucleases (MutH, BamHI and BgIII) in the presence of their cognate substrates show that each enzyme binds two metal ions in the active site that each coordinate the scissile phosphate (Yang et al., 2006). One theory in support of substrate dependent binding nature of metal ions is that perhaps the stringent coordination geometry and charge requirements of Mg^{2+} require the presence of DNA in the active site. Mg^{2+} has a very high charge density therefore binding may only occur under physiological conditions if there is also a negatively charged phosphate ester present within the active site. In addition, functional groups

of the substrate involved in metal binding may fulfil coordination requirements. This concept is supported by data that showed that Mg^{2+} ion binding to the Klenow fragment is impeded if scissile phosphate diester bridging oxygen is replaced by sulphur (Brautigam et al., 1999). However, there is a fine balance between binding Mg^{2+} in its preferred geometry in the enzyme-substrate-Mg complex because binding Mg^{2+} in a distorted octahedral geometry which switches to a genuine octahedron upon the formation of penta-coordinate transition state and this is believed to contribute to lowering of the activation energy. This is presumably part of the reason why replacement of Mg^{2+} with Ca^{2+} generally inhibits nuclease reactions because Ca^{2+} does not need to be in a specific coordination (Rosta et al., 2014). Therefore, there is no gain in energy upon switching between coordination states brought about by formation of the transition state. There could also be other reasons for the inert cofactor status of Ca^{2+} too e.g. pKa of Ca^{2+} bound to water, ionic radius etc.

1.2.5 Flap Endonuclease (FEN) Metal-ion requirement

Typifying the controversy that surrounds some nuclease mechanisms the flap endonuclease family have been the topic of debate regarding the number of metal ions needed for catalysis. The active site of FENs comprise of a large number of acidic amino acids, seven of which are positionally conserved. The carboxylate groups of these residues coordinate two divalent metal ions which have been crystallographically observed. The metal ions are designated as metal ion 1 (M1) and metal ion 2 (M2), the position of M1 is conserved across the FEN family but the position of M2 varies (*figure 1.4*).
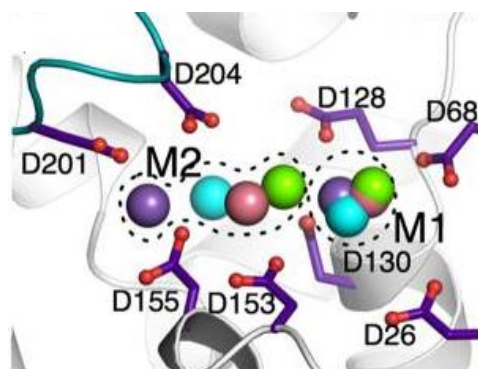
Figure 1.4: The active site structure of T5FEN with the position of metal ions 1 and 2 superimposed from structures of T4FEN, MjFEN and hFEN. *Depicting the secondary structural elements of the T5FEN active site (1UT5, gray with purple carboxylate residues) including the loop of the helix-three-turn-helix (H3TH) (teal) motif and seven active site carboxylates present in similar positions in all FENs and the eighth carboxylate (D201) present in the active sites of bacteriophage and bacterial enzymes. Two metal ions, M1 and M2 (purple) are bound with a separation of 8 Å. Metal ions observed in structures of T4FEN (1TFR, cyan), MjFEN (*Methanococcus jannaschii, *1A77, rose), hFEN (1UL1x, green), positioned by overlay, are shown. Although all FENs conserve seven active site carboxylates, the position of M2 observed in each structure is variable. Taken from Syson, K. et al. JBC (2008) 283, 28741-28746.*

Metal ions play a catalytic role in FEN reactions. Evidence suggesting this comes from pH-rate profiles of T5 Flap endonuclease (Pickering et al., 1999), which show that amino acid residues contained within the active site of FENs cannot account for the pH-rate profile of the enzyme. The effect of varying the metal ion cofactor in T5FEN on the pH rate profiles supports a catalytic role for metal ions in the reaction (Tock et al., 2003), as an increase in the Lewis acidity of the metal cofactor causes an increase in the maximal single turnover rate under pH dependent conditions. On this basis, a two-metal ion mechanism analogous to that of the 3'-5' exonuclease domain of the Klenow fragment has been proposed for FENs. However, in all but one of the active site structures (human FEN1) the distance between the two metal ions is greater than the 4 Å required for a two-metal ion mechanism (*figure 1.5*).

A study into the metal ion dependence of T5FEN cleavage suggests the participation of at least three metal ions (Syson et al., 2008). The paper proposes a model, wherein the major rate acceleration is achieved using a two-metal-ion mechanism, and a third metal ion binds in a substrate dependent manner and is involved in stabilizing various enzyme-substrate conformations. Superposition of hFEN1 and T5FEN structures demonstrate that the third metal ion implicated in this study could be accommodated in the T5FEN active site bound to Asp-130 and in doing so becomes equivalent to M2 in hFEN1, therefore supporting a two-metal ion mechanism. Mutation of Asp-130 and its equivalent in hFEN1 knocks out the catalytic ability of these enzymes demonstrating that these residues are vital for metal ion coordination (Feng et al., 2004; Shen et al., 1997).
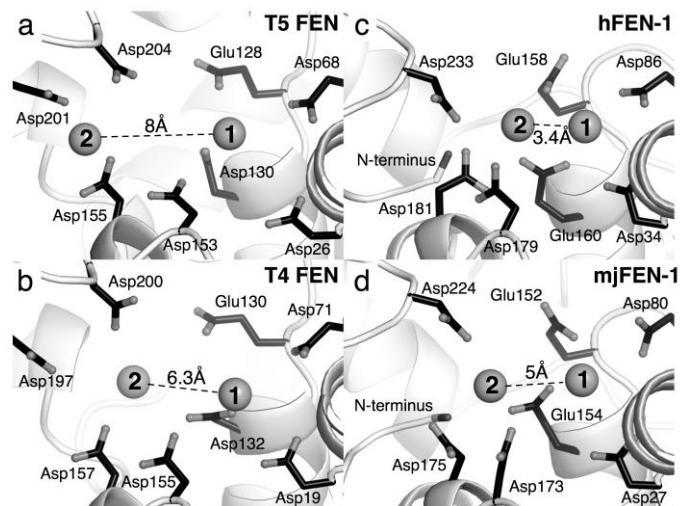
Figure 1.5: The active sites of flap endonucleases. *Comparison of active sites from T5FEN (a; 1UT5), T4FEN (b; 1TFR), hFEN (c; 1UL1:X), and MjFEN (d; 1A77) illustrating the positional conservation of active site carboxylates (black) and metal ions (1 & 2, grey).The spacing between metal 1 and metal 2 in the structures are (a) T5FEN 8 Å,(b) T4FEN 6.3 Å, (c) hFEN 3.4 Å ,(d) MjFEN 5 Å. Taken from Tomlinson, C. G. et al. Biochem Soc Trans. (2010) 38, 433-437.*

1.3: Flap Endonuclease-1

1.3.1 FEN superfamily Fold

Flap endonuclease 1 (FEN1) is a structure specific 5ʹ-endo/exonuclease. Structure-specific phosphodiesterases are required in all areas of DNA metabolism. Nucleases belonging to the FEN superfamily are specifically involved in DNA replication and repair pathways (*table 1.1*). Their importance in these pathways is highlighted by the fact that mutations in their gene sequence results in cellular stress and genome instability and, for FEN and xeroderma pigmentosum complementation group G (XPG), increased cancer susceptibility (Zheng et al., 2011). This is discussed in more detail in *section 1.8*.

Table 1.1: SAM/PIN-like fold, with corresponding structure and function. [a]Mismatch repair. [b]Double-strand break repair. [c]Nucleotide excision repair. [d]Holliday junction

Superfamily Fold	Conserved motifs	Structure	Biological Pathway	Function	Ref^
SAM/PIN-Domain-like	Rossman (α/β) structure	hFEN1	Lagging Strand Replication, lpBER	5'-Endonucleoytic cleavage of DNA double flap	(Grasby et al., 2012; Tsutakawa et al., 2011)
		EXO1	[a]MM/ [b]DSB repair	5'-Exonucleoytic cleavage of DNA nicks, gaps, blunt ends	(Nimonkar et al., 2008; Orans et al., 2011)
		GEN1	DSB repair	[d]HJ resolvase	(Rass et al., 2010)
		XPG	[c]NER	5'-Endonucleoytic cleavage of DNA bubble	(Aboussekhra et al., 1995; Schaerer, 2008)

FEN superfamily architecture has been conserved throughout evolution. A structural comparison study between human flap endonuclease-1 (hFEN1) and two FENs from the hyperthermophilic archaea, *Methanococcus jannaschii* and *Pyrococcus furiosus*, shows that the archaeal FENs have considerable sequence similarity to hFEN1. FENs from bacteriophage T4 and T5, T4RNaseH and T5'-exonuclease, and the prokaryotic paralogue of hFEN1, exonuclease-1 (EXO1) and XPG also exhibit a conserved sequence and structure to hFEN1. Although at present

there is no crystal-structures of gap endonuclease-1 (GEN1) available it shows high sequence homology to hFEN1 in the nuclease core domain.

To date thirteen crystal structures of FENs are available from a wide range of organisms including bacteriophages (Ceska et al., 1996; Devos et al., 2007; Mueser et al., 1996), archaea (Hosfield et al., 1998b; Hwang et al., 1998; Matsui et al., 2002; Shah et al., 2015), thermophilic bacteria (Harrington and Lieber, 1994; Kim et al., 1995), yeast (Chapados et al., 2004), and humans (Orans et al., 2011; Sakurai et al., 2005; Tsutakawa et al., 2011). hFEN1 has a two domain saddle like structure, formed form the N and C terminal regions of the protein that are connected through an intermediate (I) domain (*figure 1.6*). The nuclease core domain, which contains the active site carboxylate residues that bind the divalent metal ions, is formed from the N and C domains. The nuclease core domain represents the minimal protein topology required to perform FEN catalysed phosphate diester hydrolysis of DNA intermediates found on the specific replication and repair pathways. In addition to the nuclease core domain, EXO1, XPG and GEN1 contain large C-terminal extensions which through protein:protein interactions tailor each nuclease to their specific biological function. The core domain is folded into a Rossman (α/β) structure, whereby a seven stranded twisted β sheet forms a central groove packed against α helical structure on both sides. The α helices are connected by the I domain which comprises of a helical arch ($\alpha4$ - $\alpha5$), that spans the active site groove (*figure 1.6*). This configuration forms a hole in the 'back' of the protein just above the active site. The role of the helical arch in substrate recognition and catalysis has been a matter of some debate from the start (Sayers and Artymiuk, 1998), and is discussed in detail in *section 1.5*. More recently, to help clarify the role of the helical arch, the I domain has been separated further into a gateway and cap (*figure 1.7*). The cap region of GEN1 is considerably smaller than the 70 amino acid (aa) of FEN1 and EXO1, whereas this region in XPG is 600 aa in length (Odonovan et al., 1994). Unlike the cap region, the gateway at the base of the helical arch is structurally conserved and also contains active site residues which are positionally conserved across the superfamily (*section 1.5*).
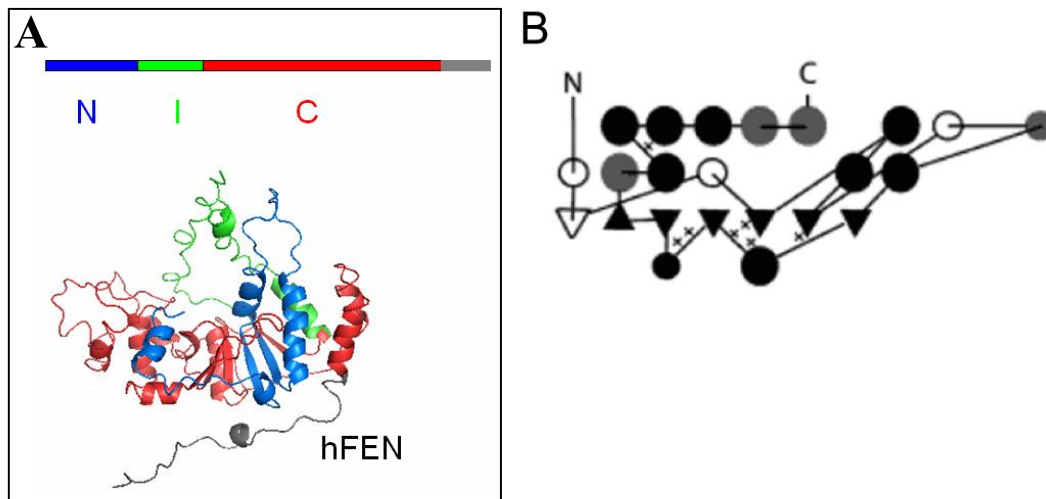
Figure 1.6: The overall structure of human flap endonuclease-1 (hFEN1). **(A)** *Structural model of hFEN1 depicting the helical arch (green) which bridges the active site. The nuclease core domain comprises of the C-terminal domain (red) and the N-terminal domain (blue), higher organisms FENs possess a C-terminal tail (grey) which interacts with other DNA binding proteins.* **(B)** ***Topology diagram*** *shaded by common features shared within family (darkest=most conserved, lightest=least conserved). The positions of active site residues are shown as crosses. Taken from Horton, N. C. DNA nucleases, 2008.*
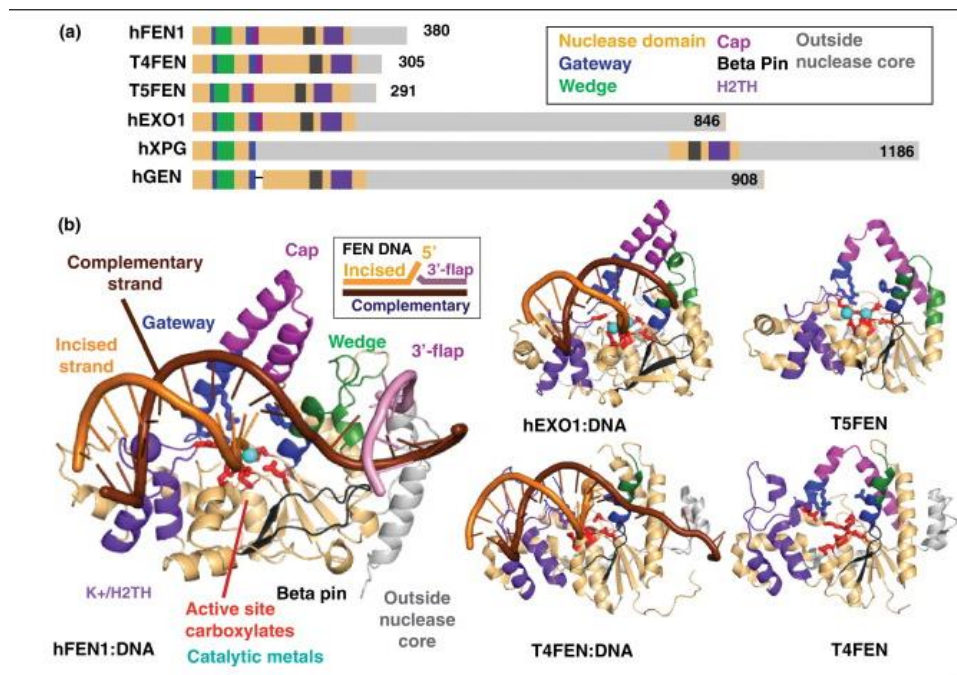
Figure 1.7: FEN superfamily architecture **(A)** *The four major domains which interact with DNA have a conserved architecture as determined by sequence alignment. Unlike these domains the helical cap and C-terminal domain are variable in length.* **(B)** *Comparison of FEN superfamily structures coloured as in (A) to highlight the conservation of the catalytic core and showing how the superfamily recognises the substrates. The structures shown are: hFEN1 (residues 2–336) with DNA and active site metals (PDB: 3Q8K), hEXO1 (residues 2–346) with DNA and active site metals (PDB: 3QEB), T5FEN with active site metals (PDB: 1UT5), T4FEN with DNA and without active site metals (PDB: 2INH) and T4FEN without active site metals (PDB: 3H8J). Enzyme*

interactions are mainly with the complementary DNA strand, coordinated through H2TH/K+ and helical wedge/β-pin motifs, which force the DNA to bend 100° at the ss/ds junction. A superfamily-conserved gateway forms the base of the helical arch and selects for 5'-termini. In FENs and EXO1, a helical cap tops the gateway, partly disordered in T4FEN with substrate. Taken from Grasby, J. A. et al, TIBS, (2012) 37(2), 74-84.

1.3.2 Human flap endonuclease-1

hFEN1 is the archetypal 5'-nuclease superfamily member and catalyses the endonucleolytic removal of 50 million RNA/DNA primers during the Okazaki fragment maturation required to replicate a single mammalian cell (Debrauwere et al., 2001). hFEN1 has further responsibility in genomic fidelity as it also removes 5'-DNA flaps that are generated during long-patch base excision repair (lpBER) in an analogous strand-displacement manner (*section 1.3.4*). Birfucated intermediates processed by hFEN1 comprise of a displaced single stranded (ss) 5'-flap flanked by upstream and downstream double-stranded (ds) DNA regions. hFEN1 selects, positions and cleaves these types of structures with a high specificity at diffusion controlled rates (Sengerova et al., 2010; Tock et al., 2003).

The ability to remove 5'-flap structures was first identified as a 5'-3'-exonuclease domain of *E.coli* DNA Pol I (Lyamichev et al., 1993). The eukaryotic counterpart, FEN1, an individual 43 kDa protein was initially purified from murine pre-B cell extracts (Harrington and Lieber, 1994). Originally, Harrington and Lieber were looking for a nuclease capable of removing 5'-flaps formed during non-homologous end joining (NHEJ). Although this was a reasonable proposal at the time, no further evidence has emerged for the involvement of FEN1 in NHEJ. Soon after, the true function of FEN became apparent when two previously characterised enzymes calf 3'-5'-endo/exonuclease (Siegal et al., 1992; Turchi et al., 1994) and DNaseIV (Robins et al., 1994), which were found to cooperate with Polε to complete lagging strand synthesis, were later re-identified as flap endonucleases. Moreover, FEN directly interacts with proliferating cell nuclear antigen (PCNA) (Chapados et al., 2004; Craggs et al., 2014; Sakurai et al., 2005; Wu et al., 1996), whose function is to orchestrate protein interactions with DNA during replication in addition to others. Since this point, FEN activity has been shown to be evolutionarily conserved across all domains of life (Habraken et al., 1993; Hosfield et al., 1998a; Ishikawa et al., 2004; Kaiser et al., 1999; Sayers and Eckstein, 1990; Shen et al., 1998), and has been shown to be essential to life (Zheng et al., 2011).

1.3.3 hFEN1 involvement in Lagging strand synthesis

Replication of double stranded DNA involves opening of the helix to form a replication fork (*figure 1.8*). The replication fork is made up of two antiparallel single strands of DNA, and because polymerase can only catalyse phosphodiester bond formation with a 5' to 3' directionality only the leading strand can be replicated continuously. For DNA polymerase to be able to replicate the lagging strand RNA primers are required. The lagging strand is synthesised as short fragments (Okazaki fragments) in a discontinuous fashion (Ogawa and Okazaki, 1980). Synthesis of the lagging strand is completed when these fragments are joined together. In eukaryotes the length of these fragments varies from 100-150 nucleotides (Anderson and Depamphilis, 1979).
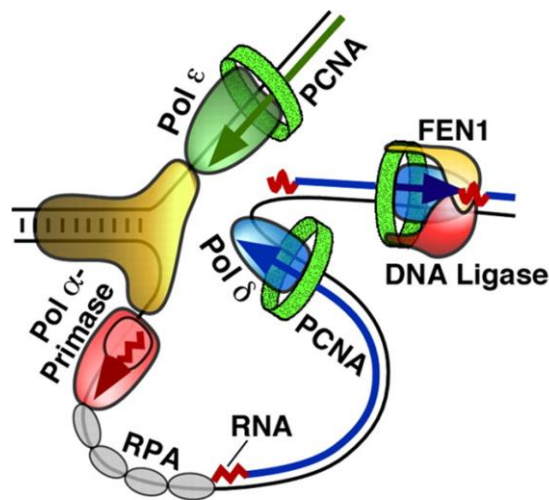
Figure 1.8: Model of proteins organised around the replication fork. *Replication of the leading strand can proceed continuously by Pol ε. Replication of the lagging strand occurs discontinuously and is coupled to movement of the replication fork by looping of the exposed strand. Adapted from Burgers, P. M. J. 2009 JBC 284, 4041-4045.*

Lagging strand synthesis is initiated by the primase domain of a DNA polymerase α/primase complex (pol α) which synthesises oligoribonucleotide primers, usually 8-12nt in length. The polymerase then extends the primer by adding an additional 20 nucleotides. On completion of primer synthesis the PCNA replaces the pol α and encircles the DNA duplex at the 3'-terminus of the primer (Garg and Burgers, 2005). PNCA is able to recruit DNA polymerase δ, and in doing so it initiates DNA polymerisation of the full length Okazaki fragment in a highly processive manner (Bambara et al., 1997). When synthesis of one fragment encounters the next, the RNA primer and adjuvant DNA are displaced by DNA polymerase δ creating an unannealed 5'-flap. The 5'-flap must be removed in order for lagging strand synthesis to be completed. FEN1 endonucleolytically hyrolyses the displaced 5'-flap leaving a nicked DNA duplex which is ligated by DNA ligase (Lig1) (Tomkinson et al., 2006). This mechanism removes the need for additional

steps in replication. The length of the flap removed by FEN1 can either be short (1-2 nucleotides (nt)) or longer (2-6 nt) depending on three distinct maturation pathways. Current opinion suggests that the pathway taken is as a result of the equilibrium between the polymerase and 3'-exonuclease activities of Polδ (Burgers, 2009). Flaps longer than

1.3.4 Base excision repair

Base excision repair (BER) is the primary DNA repair pathway that removes modifies bases that arise due to oxidative, alkylation, deamination and depurinatiation/depyrimidination damage. Currently it is believed that BER facilitates the repair of damaged DNA via two general pathways; short patch and long patch (*figure 1.9*) (Robertson et al., 2009). BER is a multistep pathway requiring the function of four proteins, DNA glycosylases, an apurinic/apyrimidinic AP endonuclease or AP lyase, DNA Polymerase and DNA ligase (Kubota et al., 1996).

The first step in BER is initiated by DNA glycosylase, which upon recognition of the appropriate damaged base catalyses the cleavage of the N-glycosidic bond (McCullough et al., 1999). This action completely removes the damaged base and generates an AP site. DNA glycosylases are able to recognize a multitude of damaged bases and many, but not all, belong to the HhH superfamily; defined by the helix-hairpin-helix binding motif (Hardeland et al., 2007). The reason why BER is the most frequently used repair mode in nature is because of the vast number of glycosylases, each specific for a type of DNA damage, but they all generate the same AP site which can be process by a handful of downstream enzymes (Dianov and Huebscher, 2013).

The next step of the reaction pathway involves cleavage of the DNA backbone. AP endonuclease or AP lyase, an activity present in DNA Polymerase β and some glycosylases, cleave 5' or 3' to the AP site, respectively, generating a single stranded nick (McCullough et al., 1999). Glycosylases with intrinsic AP lyase activity cleave the phosphodiester bond 3' to the AP site, thus generating a 3'-unsatutated aldehydic sugar and a 5'-phosphate end. The unsaturated aldehydic sugar is then incised on the 5'-side by AP endonuclease generating an appropriate substrate for strand resynthesis. For monofunctional glycosylases not possessing AP lyase activity, the glycosylase dissociates permitting AP endonuclease access to the AP site. Cleavage by AP endonuclease 5'-to the AP site, generating a 5'-deoxyribose phosphate (dRP) and 3'-OH. The dRP is displaced by DNA Polymerase β permitting subsequent gap filling. The displaced dRP is cleaved by Polβ on its 5'-side allowing for efficient ligation by DNA ligase. Polβ is capable of using its lyase activity to cleave the sugar phosphate backbone 3' to either an intact AP site or one which has been previously incised by AP endonuclease. The mechanism of Polβ lyase activity involves the of a schiff base intermediate (Longley et al., 1998), requiring the AP site deoxyribose

to be in equilibrium between the open (aldehyde) form and the closed furanose form. The electrophilic centre of the aldehyde form can be lost if the dRP is oxidised and the sugar is no longer in equilibrium. This is believed to be a major factor in the switch between spBER and the more energetically costly lpBER pathway.

During lpBER, Polβ, Polδ, PCNA, FEN1 and the replicative DNA ligase are sequestered to the AP site, contrasting recruitment of only Polβ in spBER. In a PCNA-dependent manner Polβ/δ performs strand-displacement polymerization of tracts of DNA longer than one nucleotide (2-13 nts) (Robertson et al., 2009). Displacement of not only the dRP but other nucleotides 5' to this site generates a flap structure which cannot be cleaved by Polβ. Therefore, lpBER has an absolute requirement FEN1 which cleaves flap structures specifically creating a ligatable nick(Kim et al., 1998). LpBER can also be divided into two sub-groups, the first being the mechanism described above and the other is a 'hit-and-run' mechanism (Liu et al., 2005) (*figure 1.9*) involving iterative rounds of Polβ synthesis and FEN1 cleavage. This 'hit-and-run' mechanism leads to the replacement of 2-11 nts (long patch) and is terminated upon ligation of the newly synthesised strand and existing strand. Additional factors in the switch between sp and lpBER mechanisms are scaffolding proteins initially recruited to the nick generated by AP endonuclease. X-ray cross-complementary protein 1 (XRCC1) is one of the first proteins to bind the damaged site in spBER. XRCC1 having no known enzyme activity is believed to coordinated spBER (Caldecott, 2003). On the other hand PCNA, which plays no role in spBER, has been shown to stimulate lpBER (Klungland and Lindahl, 1997).
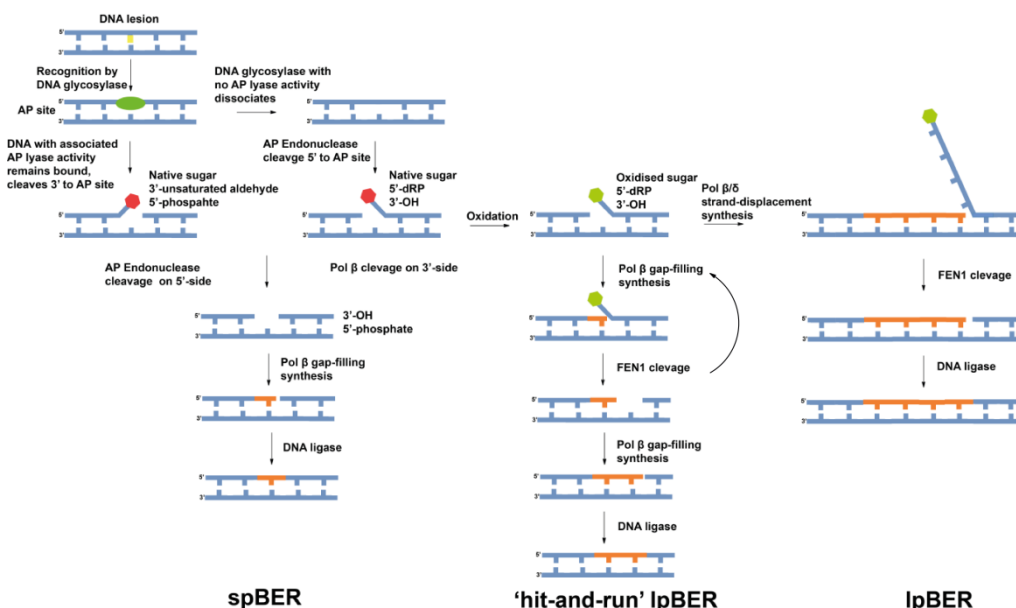
Figure 1.9: Schematic representation of the two different mechanism of BER, *short patch and long patch. Highlighting the various 'switch points' along the pathways. The damaged base*

(yellow) is recognised by the appropriate DNA glycosylase. Bifunctional glycosylases containing AP lyase activity cleave 3' to the AP site setting off a casecade of events involving APE1 and Polδ in a process known as spBER. Monofunction glycosylase dissociate form the AP site and depending on the oxidation state of the dRP (native=red and oxidised=green) the downstream pathway can be either spBER or lpBER. Newly incorporated nucleotides catalysed by Polδ are shown in orange. lpBER can be further divided into an iterative process (hit and run) or a more stepwise process (lpBER) similar to Okazaki fragment maturation.

1.3.5 FEN substrate specificity

Unlike bacteriophage FENs which show a preference for single flap/pseudo Y substrates (Bhagwat et al., 1997), higher organism FENs prefer substrates with duplex DNA upstream and downstream relative to the cleavage site in order to perform catalysis (*figure 1.10a*). Bifurcated DNA structures make contact with hFEN1 through several positively charged residues that extend across the length of the enzyme. The specific nucleic acid side chain interactions made by hFEN1 are shown in *figure 1.10b*. hFEN1:DNA binding is mediated by four separate regions 1) the Helix-2-Turn-Helix (H2TH) motif with bound K+ ion and charged side chains that bind downstream dsDNA, 2) the 3'-flap binding pocket interacting with upstream dsDNA, 3) the hydrophobic wedge composed of α2 and the α2-α3 loop at the junction between the upstream and downstream dsDNA portions of the two-way junction, and 4) the two-metal ion active site that binds the 5'-flap strand; each binding site will be discussed in detail below.
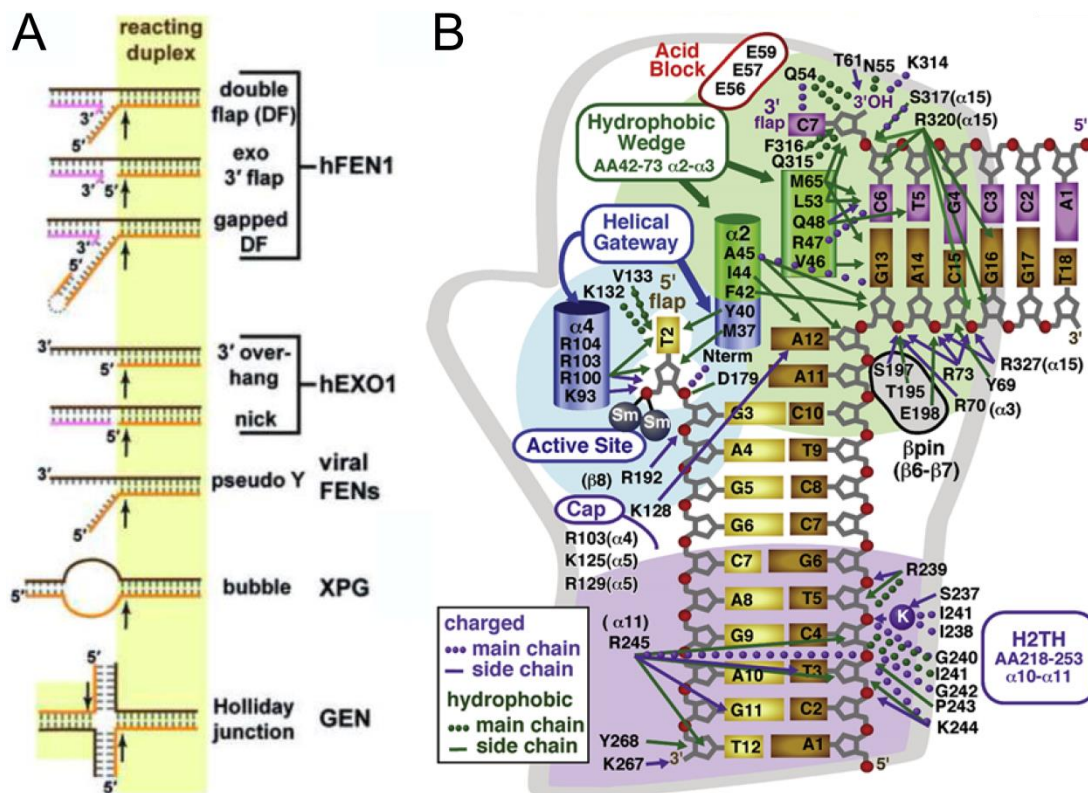
*Figure 1.10: (A) **FEN substrate topology.** The structural requirements of the substrates for each FEN superfamily member is depicted also showing the major site of phosphate diester hydrolysis (arrow) upon the reacting strand (yellow). Taken from Grasby, J. A. et al, TIBS, (2012) 37(2), 74-84. (B) **Schematic representation of FEN showing the specific interactions between the DNA double flap substrate and residues contained within the four major binding regions.** Here, FEN is represented as a boxing glove, FEN makes contacts with the DNA through acidic residues extending the length of the palm and the helical arch is symbolized by the thumb of the glove. Taken from Tsutakawa, S.E., et al. (2011). Cell 145, 198-211.*

1.3.6 FEN Downstream duplex binding

The H2TH:K ion motif, α-helices 10 and 11 in hFEN1, is a classic DNA-binding motif found in a large range of replication and repair enzymes. The K+ ion is coordinated by Ile-238 and Ile-241 backbone carbonyl oxygens and Ser-237 hydroxyl; these residues reside on the loop between α10 and 11 (Tsutakawa et al., 2011). This motif in combination with four basic residues protruding from the protein surface (Lys-267, Arg-239, Arg-245, Lys-244) creates a furrow in which the minor grove can slide into (*figure 1.11*) (Qiu et al., 2004). Contacts with the H2TH motif and a hydrophobic track formed by α2/3 and a β-pin to the DNA strand are thought to align the rest of the substrate with the active site. Beyond this point the DNA arcs away from the protein surface, only to re-join at the active site (*figure 1.11*). The position of these two binding sites, a helical-turn apart (Orans et al., 2011; Tsutakawa et al., 2011), is conserved

throughout evolution. The majority of enzyme-substrate interactions occur at these sites and the majority are with the DNA strand complementary to the reaction site, enforcing specificity for dsDNA substrates. This binding mode can also explain how FEN has the ability to cleave RNA when it is contained within the 5'-flap but not when the complete reacting strand (5'-flap and connecting downstream region) is RNA (Harrington and Lieber, 1994). The reason behind this is presumably because the shorter, A-conformer, RNA:DNA heteroduplex means that the duplex cannot contact the HT2H and active site, simultaneously.

The importance of this motif towards the contribution to the binding affinity of human and phage FENs as well as being characterised crystalographically (Devos et al., 2007) has also been shown biochemically. Disruption of the electrostatic 'rail' in the K244A/R245A double variant reduces activity by half (Liu et al., 2006). Similarly, in T5FEN, were it is actually a H3TH motif and no K+ ion has been observed (Hosfield et al., 1998b), mutation of the extended basic residues (K215A and R216A) results in significantly impaired catalytic efficiency and binding (Dervan et al., 2002).
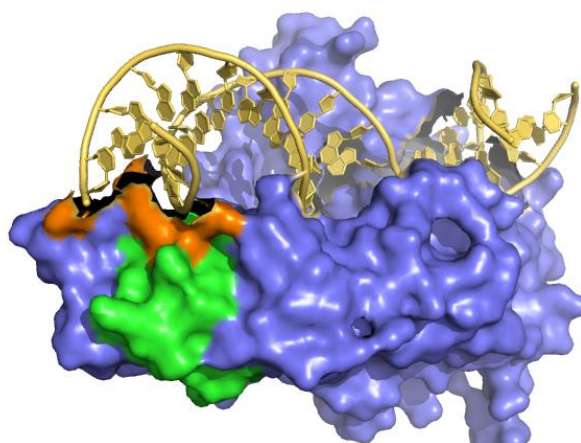
Figure 1.11: H2TH forms the major downstream DNA binding site in hFEN1. *Space-filling representation of hFEN1 bound to DNA (PDB: 3Q8K), the H2TH (green) and residues contained with on this motif, Lys267, Arg239, Arg245, Lys244 (orange) form rails which the minor groove and fit into.*

1.3.7 FEN Upstream duplex binding and the 3'-flap pocket

Early biochemical evidence highlighted the importance of dsDNA in the upstream binding site to the ability of higher organism FENs perform efficient catalysis (Harrington and Lieber, 1994). The length of ds DNA in the upstream region was found to be directly linked to the affinity *in vitro* substrates (Harrington and Lieber, 1995). Harrington and Lieber performed binding and cleavage assays using double flap (DF) substrates, where the upstream and downstream regions both had an adjuvant flap. These structures comprised of a 19 nt ss 5'-flap and either a 1 or 10 nt

extrahelical 3'-flap. These double flap substrates were cleaved in the same position to yield the major product of cleavage upon single 5'-flap structure. More importantly, the rate of phosphodiester hydrolysis was much faster upon a double flap structure with the 1 nt 3'-flap compared to the single 5'-flap structure. The 5'-3' cleavage specificity of FENs mean that 3'-flap cannot be cleaved, but they do possess a role in binding although the significance was not realised until much later. This same rate enhancement, in addition to an increased cleavage specificity, has been observed in bacteria (Xu et al., 2001), archaea (Kaiser et al., 1999), and eukaryotes (Kao et al., 2002).
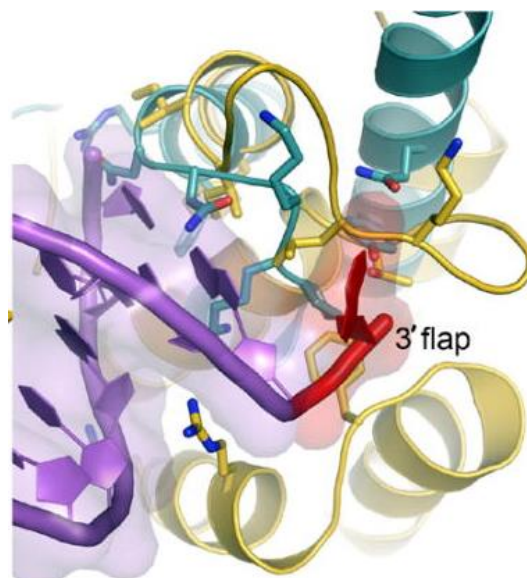
Figure 1.12: Structural comparisons of the 3'-Flap Binding Pocket of T5FEN and AfFEN proteins. The residues of T5FEN and AfFEN are represented as blue and yellow chains respectively. T5FEN lacks the extended loop and helical regions of AfFEN which form the binding site for the 3' flap (red) and associated upstream DNA duplex (purple). Taken from R. Williams, R. et al. J. Mol. Biol (2007) 371, 34-48.

The molecular basis for the preference for 3'-flaps first was revealed by a crystal structure of AfFEN, in complex with dsDNA with a 3'-overhang (*figure 1.12*). Analysis of the structure revealed a pocket neighbouring the upstream dsDNA binding site that contacted the 3'-flap through van der Waals and hydrogen bonding interactions with the nucleotide ribose (Chapados et al., 2004). Sequence alignment analyses reveals that residues contained within the 3'-binding pocket are highly conserved in archaea and humans but not in phage. The 3'-flap is contacted nine times through van der Waals and hydrogen bonding interactions and comprises 1/8 of the hFEN1:DNA binding surface (Tsutakawa et al., 2011). No other contacts are made to this strand, but hFEN1 makes multiple contacts with the upstream template strand analogous to the binding

mode in the downstream region. The presence of a 3'-flap results in a significant increase in both the single and multiple turnover rates of reactions catalysed by hFEN1 . However, the mechanism by which the 3'-flap contributes to rate enhancement is not yet fully understood.

Equilibrating flap structures as shown in *figure 1.13* resemble intermediates on the lagging strand and lpBER pathways. FEN is believed to 'capture' this optimal substrate and endonucleolytically cleave one nucleotide into the downstream duplex. Aided by, steric constraints that are enforced on the substrate through interactions made with the 3'-flap positioning the scissile phosphate in the active site gives the specific product resulting from reaction 1nt into the downstream duplex. The correct orientation of the DF substrate is proposed to be selected for by the presence of an 'acid block' (*figure 1.10b*), the collection of negatively charged residues prevents the 3'-flap form being any longer than 1nt. With this the 3'-flap binding pocket can be viewed as an anchoring point for the whole DF structure.

Single nucleotide gap - longer 5'-flap   Nicked DNA - single 5'-flap   Nicked DNA - double 3'/5'-flap
'optimal substrate'
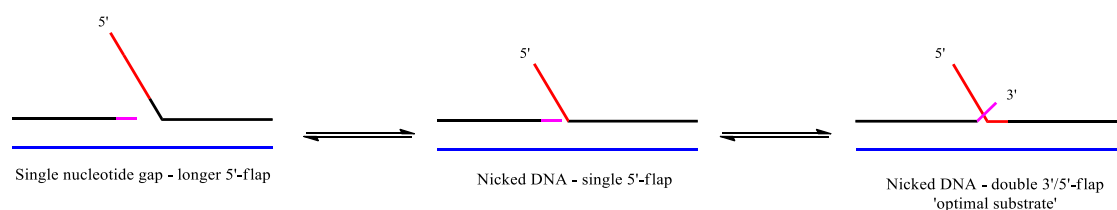
Figure 1.13: Equilibrating forms of the FEN structure. *The ss 5'-flap is shown in red and the 3'-flap nucleotide in pink this base is complementary to the template strand (blue) as is the first base of the 5'-flap. So, when the optimal substrate is cleaved one nucleotide into the downstream region, the 3'-flap nucleotide base pairs with the template strand to leave a ligatable nick.*

Bacteriophage FENs do not possess a 3'-flap binding pocket, and show no enhanced cleavage site specificity on substrates containing a 1 nt 3'-flap compared to substrates without a 3'-flap (Williams et al., 2007a). Thus, the 3'-flap completely confers cleavage specificity in higher organism FENs; this is biologically important as it negates the need for DNA Polymerase after cleavage of DNA by FEN. The 3'-flap nucleotide is complementary to the template strand and therefore yields a nicked DNA duplex, which is immediately ligatable. Mutation of conserved residues in the 3'-flap binding pocket result in reduced substrate binding affinity and also loss of cleavage site specificity (Friedrich-Heineken and Hubscher, 2004). Early studies showed that on equilibrating flaps structures FEN1 selected for conformers with a 1 nt 3'-flap (Kao et al., 2002). However, *in vitro* studies have predominantly used double flap substrates that are not strictly

speaking equilibrating as the 3'-flap nt is non-complementary to the opposite strand. This is to 'guarantee' cleavage one nt into the downstream region making data analysis less complicated.

Global conformational changes of FEN proteins on DNA binding are required to transmit binding of the correct substrate into the catalytically active form of the enzyme. Small angle X-ray scattering studies suggest that there is a global conformational change in a catalytically inactive variant of hFEN1 on binding to 'product' oligonucleotide in the presence of magnesium ions (Kim et al., 1999). Similarly, Fourier transform infra-red difference spectra show a conformational change in hFEN1 on binding to a substrate oligonucleotide (Kim et al., 2001). The 3'-flap binding pocket is believed to undergo such a conformational change upon binding of the extrahelical nt. Part of the C-terminal region of hFEN1 and a portion of the α2/3 loop participate in formation of the 3'-flap binding pocket, and ordering of this region has been proposed to cause a distal/remote conformational change, specifically ordering of the helical arch (α4/5) resulting in the correct positioning of key residues involved in catalysis (Kim et al., 2001; Qiu et al., 2004). Moreover, upon ordering of the α2/3 loop due to 3'-flap binding, Arg-47, which is part of the α2/3 loop, is positioned to contact upstream nucleotides via its peptide backbone, while the side chain is positioned to C-cap α2 and make van der Waals contacts with Lys-128 side chain on α5, thereby supporting the idea that occupation of the 3'-flap site translates to ordering of the helical arch. The hFEN1 R47A mutation affects the cleavage site specificity on substrates without a 3'-flap and the reaction rate on substrates with both a 5' and 3'-flap (Qiu et al., 2002).

1.3.8 FEN Hydrophobic wedge

In an analogous manner to the 3'-flap binding pocket, a comparison of crystal structures of the free enzyme and enzyme substrate complex showed the hydrophobic wedge (α2/3) and positively charged residues provided by the β-pin (β6/7) undergo a conformational change upon substrate binding,. The combination of these motifs pinches the duplex at the ds-ds or ss-ds junction causing the defining 90-100˚ bend in the bound DNA. In support of this, recent single-molecule fluorescence resonance energy transfer (sm-FRET) studies observed that substrates possessing only a 5ʹ-flap were bound less tightly (Craggs et al., 2014), and also bent sub-optimally (Sobhy et al., 2013). These studies confirmed findings from previous ensemble FRET experiments suggesting that the kinked conformation of DF structures is stabilised by hFEN1 binding (Chapados et al., 2004).
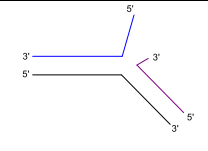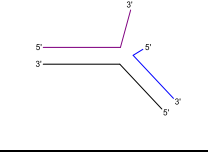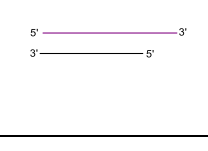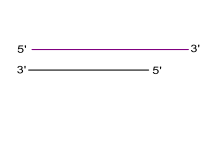
In the most recent crystal structure of hFEN1:DNA it can be seen how this distortion of the duplex is achieved. The hydrophobic wedge is parallel to the downstream bases and butts up

against the terminal bp diverting the 5'-flap strand towards the active site. The interaction with the upstream region is different, α2/3 and connecting loop are end on with the upstream dsDNA and sit within the dimensions of the dsDNA against the first bp after the 1nt 3'-flap. This binding interaction is thought to be a major recognition element as only junctions capable of undergoing such a large conformational change will be able to bind in all four binding sites. Moreover, complementary strand binding and bending, forces the trajectory of the 5'-flap strand toward the superfamily-conserved active site (Grasby et al., 2012). However, although all substrates of FEN superfamily members are manipulated in the same manner, DNA bubble and Holliday junction (HJ) structures do not have a free 5'-end and yet are still endonuceolytically cleaved by XPG and GEN1, respectively. A consequence of blocking the path of the 5'-flap and forcing it to point towards the active site is that the hydrophobic wedge prises open the duplex between the 5'-flap strand and complementary strand causing sub-optimal basepairing of the nts either side of the scissile phosphate. This unique conformation has been proposed to be stabilised by interactions with PCNA (Craggs et al., 2014), this distortion of the duplex and its importance will be described below

1.4 DNA unpairing

Structure-specific nucleases are involved in nearly every nucleic acid replication, recombination and repair pathway. The ability of these multifunctional enzymes to recognise and process replication and repair intermediates is essential to the cell. Although the substrates, binding modes, and mechanistic details vary, it is becoming more apparent that these phosphodiesterases share a common mechanism for scissile phosphate diester selection. Upon selecting the correct structure, they are then tasked with the responsibility of strictly assuring hydrolysis occurs at the correct position so that the product can be used for downstream processes. As discussed in *section 1.2*, most nucleases achieve catalysis through metal ion coordination. The general consensus is that for a two metal ion mechanism there must be direct (inner sphere) coordination of the scissile phosphate nonbridging oxygen to both metal ions (Yang, 2011). It has been postulated that nucleotides intended for cleavage are transferred from an initial binding site to a distal metal occupied active site when specific aspects of the nucleic acid substrate have been verified (*table 1.2*).

Table 1.2: Structure specific nucleases proposed to utilise DNA unpairing. ᵃHR = homologous recombination.

Nuclease	DNA metabolic pathway	Substrate structure	Substrate nomenclature	Activity	RNA vs DNA	Reviewed in

FEN1	BER, Replication		Double Flap	5'-endo / exo	RNA/DNA	(Grasby et al., 2012)
Mus81-Eme1/Mms4 and XPF-ERCC1	HR, Replication		Double Flap	3'-endo	DNA	(Ciccia et al., 2008)
MRE11	HR		3'-overhang, Blunt end	3'-5' exo, 3'-endo	DNA	(Shibata et al., 2014)
λ exonuclease / RecE	HR		Blunt end	3'-5' Exo 3'-Endo	RNA/DNA	(Zhang et al., 2011)

1.4.1 FEN1 active site residues involved in double nucleotide unpairing

In the absence of divalent metal ions, the phosphate diester corresponding to the major site of FEN reaction in the T4FEN:DNA complex is stacked in the duplex at a distance of 7Å from the closest metal ion site (Devos et al., 2007). In other words, the binding motifs of FEN described above position the targeted phosphodiester, one nt into the downstream duplex, near the active site but not quite close enough for hydrolysis to occur. Therefore, a base unpairing mechanism had been proposed, wherein the bases flanking the scissile phosphate (+1 and -1, defined in reference to the scissile phosphate) must unpair from the template strand to place the phosphate diester bond between the metal ion sites (*figure 1.14*) (Syson et al., 2008; Tomlinson et al., 2010). More recently, crystal structures of EXO1 and hFEN1 enzyme substrate (ES) and product (EP) complexes supported a double nucleotide unpairing (DNU) mechanism (Orans et al., 2011; Tsutakawa et al., 2011) (*figure 1.14*). The cofactor occupying both metal ion sites in the hFEN1 crystal structure was Sm^{3+}, which supports binding but inhibits cleavage. Even with the negative charge of the active site neutralised, the +1 and -1 nt of the uncleaved substrate was >6 Å form either metal ion. In the product structures, however, the -1 nt that processed a terminal 5'-phospahte monoester, was contacting the two Sm^{3+} ions. The gap had been bridged because the terminal nucleotide had unpaired. For this to have happened during the reaction, there was likely some enzyme-substrate intermediate in which the two nts were unpaired to place the scissile phosphate on the metal ions. Consistent with the proposed intermediate,

prevention of double nucleotide unpairing using interstrand DNA cross-links severely decreased the hFEN1 reaction rate and abolished cleavage site specificity (Beddows et al., 2012).

Figure 1.14: Double Nucleotide Unpairing mechanism conserved across the FEN superfamily.
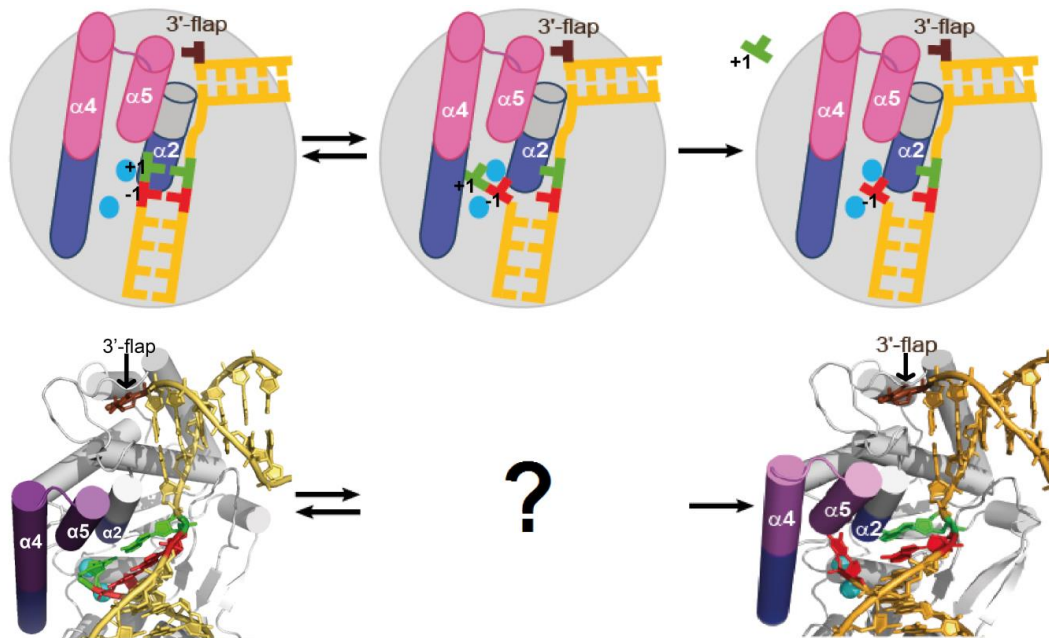
(Top panel) Schematic representation of DNU illustrating how DNU positions the target phosphate in the active site to allow the scissile phosphate diester to interact with the requisite divalent metal ions. (Bottom panel) hFEN1: substrate (PDB: 3Q8L), and hFEN1: product (PDB: 3Q8K), crystal structures provide support for this mechanism. In the ES structure the two nucleotides flanking the scissile phosphate, +1 (green) and -1 (red), are paired with the complementary strand. The -1 with 5'-phosphate monoester is unpaired in the EP structure. Thus, a putative ES intermediate has been proposed in which both the -1 and +1 nucleotides are unpaired. The helical gateway (base of α4 and α2), coloured blue and the helical cap (top of α4 and α5), coloured magenta.

From crystallographic data a model of the unpairing process has been proposed, in which formation of the unpaired state is thought to be facilitated in part through residues provided by the helical gateway, comprised of the bottom portion of α4 and α2 (*figure 1.15*). Although the helical cap is not conserved, the helical gateway is found across the FEN superfamily. Lys-93 and Arg-100 (hFEN1 numbering), located at the base of α4, are positionally conserved across the whole FEN superfamily (*figure 1.15*). Mutational analyses demonstrated the importance of these two residues (Sengerova et al., 2010; Storici et al., 2002; Tsutakawa et al., 2011). These two residues extend into the active site (*figure 1.16*), and in the enzyme-product (EP) complex contact the 5'-phosphate monoester of the terminal unpaired nucleotide. It has been suggested that K93 acts as an electrostatic catalyst by stabilizing the negatively-charged penta-coordinate transition state, not as a general acid which was initially suspected (Sengerova et al., 2010). A

similar role for R100 was also hypothesized. The position of K93 and R100A suggests that these residues may also be involved in stabilization of an unpaired state, as they would be able to neutralise the negatively charged phosphate diester moving closer to the protein surface during unpairing.

The other helix of the gateway, α2 habours a stacking residue, Tyr-40. However, unlike Lys-93 and Arg-100, Tyr-40 is not conserved across the superfamily (*figure 1.15*). Notably, Y40 stacked against the 5'-face of the +1 nt in the enzyme substrate complex and the 3'-face of the-1nt in the enzyme product complex (*figure 1.17*). To have achieved these two states, in total Tyr-40 must have undergone a 25˚ rotation during the reaction mechanism. The importance of Tyr-40 was further emphasised by a 20-fold reduction in the multiple turnover rate (Tsutakawa et al., 2011). In the crystal structure of Exo1 complex with product, the terminal 5'-base was involved in the same π-stacking interactions with Tyr-32. However, mutation of His36 had a more detrimental effect on the reaction rate. Therefore, His-36 and not Tyr-32 was suspected to be involved in a DNU mechanism. XPG and GEN1 contain, asparagine and cysteine, respectively, in the equivalent positions which could both potentially interact with an unpaired base stabilising an intermediate state in the same way.
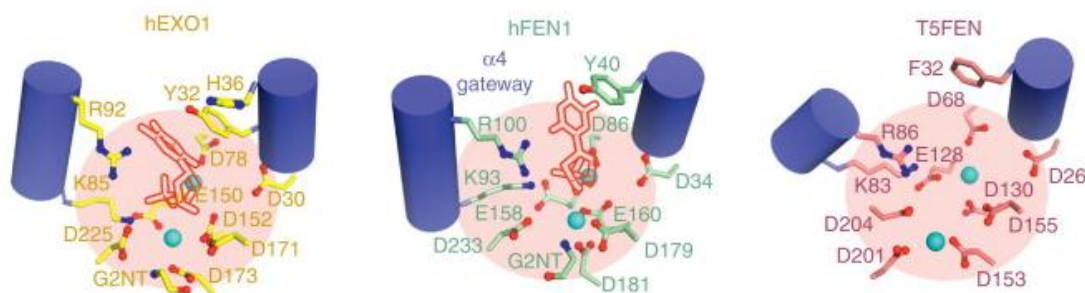
Figure 1.15: Comparison of residues on the gateway helices conserved throughout the 5'-nuclease superfamily. *For structures with DNA, the terminal bases of the 5'-strand are shown (outlines) to indicate that the target phosphate diester must enter the active site for the divalent metal ions and catalytic basic residues to be within the correct distance of the target phosphate diester. Key conserved catalytic side chains are shown as sticks and divalent metal ions as cyan spheres. The gateway formed by two α helices (α2 and α4) (cylinders) prevents intimate association of catalytic metal ions and residues with dsDNA. A residue with a side chain that can stack with the terminal base is always present. In hEXO1, two different residues, His36 and Tyr32, stack with terminal nucleotide in the substrate and product complexes, respectively. Taken from Grasby, J. A. et al, TIBS, (2012) 37(2), 74-84.*
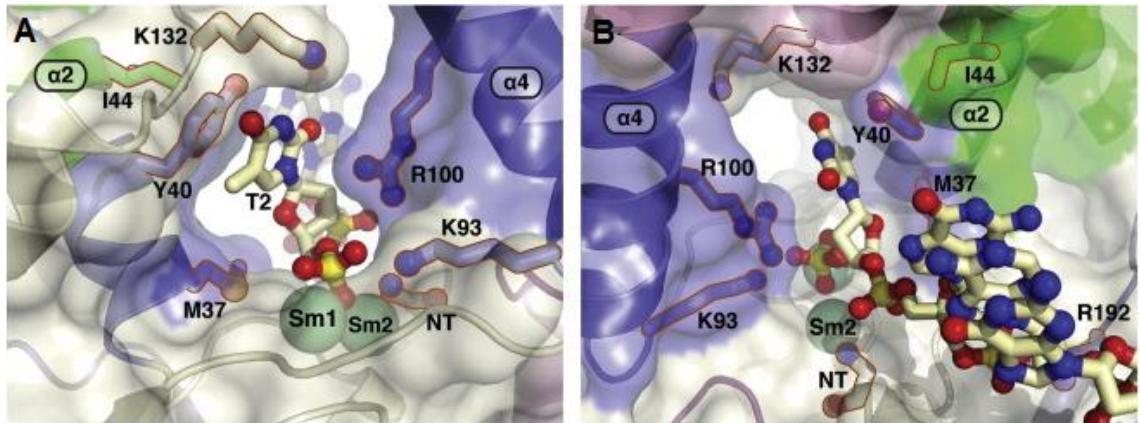
Figure 1.16: (A) Rear view of the active site. *The helical gateway (blue) forms a hole to allow the 5'-flap strand to exit the active site. The product DNA and key catalytic residues are shown as sticks. K93 and R100 are oriented towards the scissile phosphate* ***(B) Front view of the active site****. The 5'-flap strand approaches the active site and its trajectory suggests passage through the helical arch. Taken from Tsutakawa, S.E., et al. (2011). Cell 145, 198-211.*
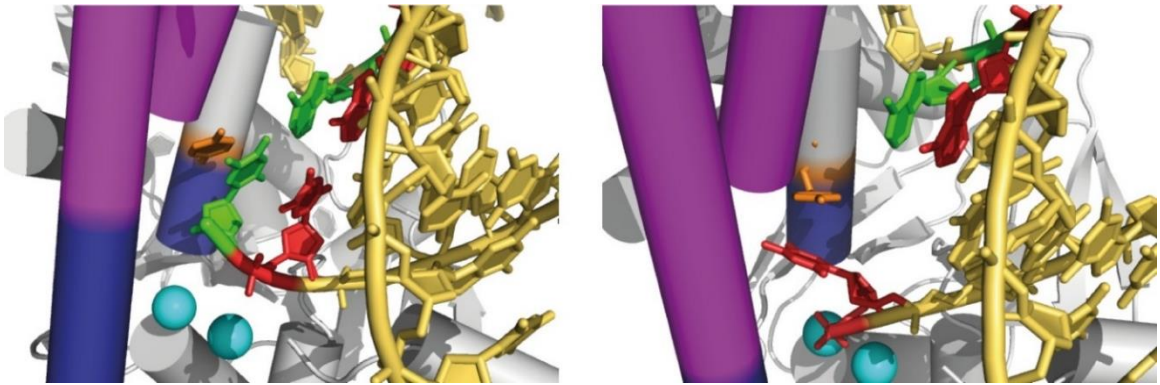
Figure 1.17: Stacking of the Tyrosine 40 residue with multiple nucleotides during catalysis in hFEN1*. Tyr-40 (orange) stacks against the 5'-face of the +1 nt (green) relative to the sciscle phosphate in the hFEN1: substrate (PDB: 3Q8L), crystal structure. In the hFEN1: product (PDB: 3Q8K), crystal structure Tyr-40 is stacked against the 3'-face of the -1 (red) nucleotide with terminal monophosphate ester, stabilizing the unpaired state of this nucleotide.*

1.4.2 Mus81/XPF are putative DNA unpairing enzymes

Mus81 and XPF, members of the MUS/XPF endonuclease family (Schwartz and Heyer, 2011), and associate with noncatalytric partners, Eme1/Mms4 and ERCC1, respectively. The resultant heterodimers (Mus81-Eme1/Mms4 and XPF-ERCC1) display structure specific endonuclease activity and target similar branched structures but in different DNA metabolic pathways. Mus81-Eme1/Mms4 acts on 3'-flap structures formed during stalled replication forks and HJ resolution (Schwartz and Heyer, 2011), and XPF-ERCC1 is the 3'-endonuclease cutting the bubble structure on 5'-side of lesion during nucleotide excision repair (de Laat et al., 1998).

FEN1 and Mus81/XPF bind at the ss-ds junction but cleave flaps with opposite polarity. The binding mode of FEN1 bears striking similarities with Mus81 and XPF (figure 1.18a and b). Like FEN1, structural data indicate that XPF family members also display substantial conformation changes between unbound and DNA bound forms (Gwon et al., 2014; Newman et al., 2005). As well as protein restructuring, DNA conformational changes were also shown to be important for substrate discrimination. Flap structures were observed to be sharply bent in Mus81 (Gwon et al., 2014), and XPF (Hutton et al., 2010) complexes (*figure 1.18b*). Unlike FEN1, XPF (Roberts and White, 2005), and Mus81 (Bastin-Shanower et al., 2003) cleaved further into the duplex region adjacent to the flap (*figure 1.18c*). However, like FEN, it was argued that bending alone could not account completely for the substrate rearrangements required to place the scissile phosphate in the active site. From crystallographic data it was shown that the two bent duplex regions are connected by a short stretch of ssDNA on the non-reacting strand (*figure 1.18c*). As a result, nucleotides adjacent to the flap on the incision strand were proposed to be unpaired in agreement with footprinting data (Nishino et al., 2005).
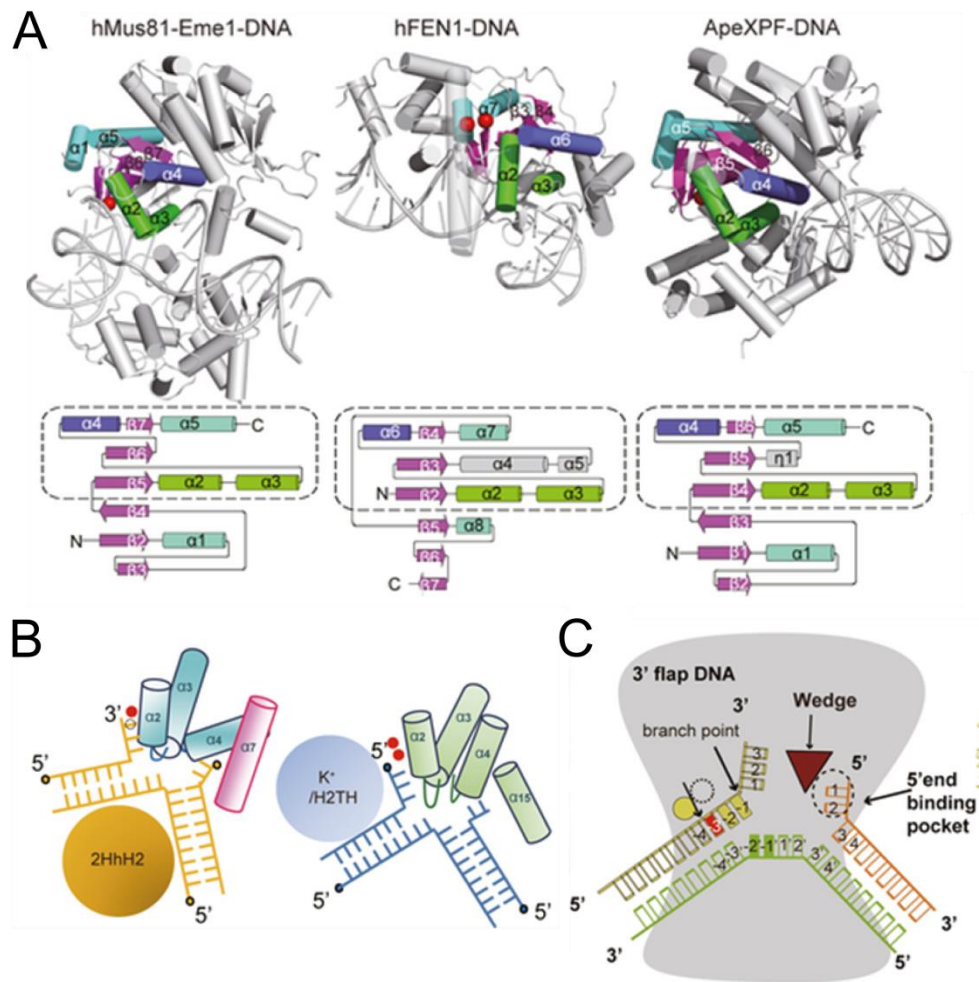
Figure 1.18: Comparison of the binding modes of 5'- and 3'-endonucleases (A) From left to right, architecture of hMus81-Eme1 (4P0P), hFEN1 (3QE9), and ApXPF (2BGW) with bound DNA, topologies shown below each structure coloured to represent secondary structure similarity. (B) Schematic models of the hMus81-Eme1 bound to the 3' flap DNA (left) and the FEN1/hExo1—5'-flap DNA (right) are shown to highlight the common structural features. (C) Schematic representation of Mus81-Eme1 unpairing mechanism. The nucleotides (yellow) adjacent to the branch point/ss-ds junction (marked with arrow) are proposed to un-base-pair with complementary strand (green) placing the scissile phosphate (5'-side of red nucleotide) next to Mg^{2+} ion (yellow) and putative Mg^{2+} ion (dotted circle). Adapted from Gwon, G., H. et al. (2014) EMBO J. 33, 1061-1072.

1.4.3 Mre11 as a DNA unpairing enzyme

Mre11 functions as a homodimer (*figure 1.19a*) at a variety of forks and breaks that materialize due to DSBs and stalled replication forks (Schlacher et al., 2011; Williams et al., 2007b). *In vitro*, Mre11 possess both ds <u>exo</u>nuclease and ss <u>endo</u>nuclease activities (Paull and Gellert, 1998; Trujillo and Sung, 2001). An unpairing-like mechanism was suggested to be responsible for linking the two activities.

The structure of Mre11 bound to branched DNA eluded to the ds exonuclease mechanism (Williams et al., 2008). Protein conformational changes have been linked to DNA melting at the terminus of the substrate (*figure 1.19d*). The precise mechanism is unknown, but these architectural rearrangements in combination with a wedge that penetrates into the minor groove (*figure 1.19b*) causes disruption of terminal base pairing interactions (Williams et al., 2008). The next step in the mechanism is proposed to involve a His residue-dependent phosphate rotation two nucleotides into the duplex region tilting the adjuvant 3'- flap of the incised strand toward the active site and is only permitted on dsDNA in which the terminus is positioned close enough, mediated by unpairing.

Much less is known about mechanism of ss endonuclease activity. It has been argued that the ssDNA is able to access the common active site by binding to a channel only wide enough for ssDNA (Hopfner et al., 2001). This alternative binding mode negated the need for a substrate conformational rearrangement and was consistent with data suggesting mutation of the His responsible for the phosphate rotation of one of the dsDNA strands abolishes exo activity, but only had a moderate effect on endo activity (Williams et al., 2008). Taken together, these models suggested that dsDNA cannot access the active site, but Mre11 used the malleability of a duplex terminus to facilitate unpairing, making the substrate essentially single-stranded. Mre11 does not possess ds endonuclease activity because the ss-ds junction acts as a road-block determining how much of the 'reacting strand' can enter the active site.
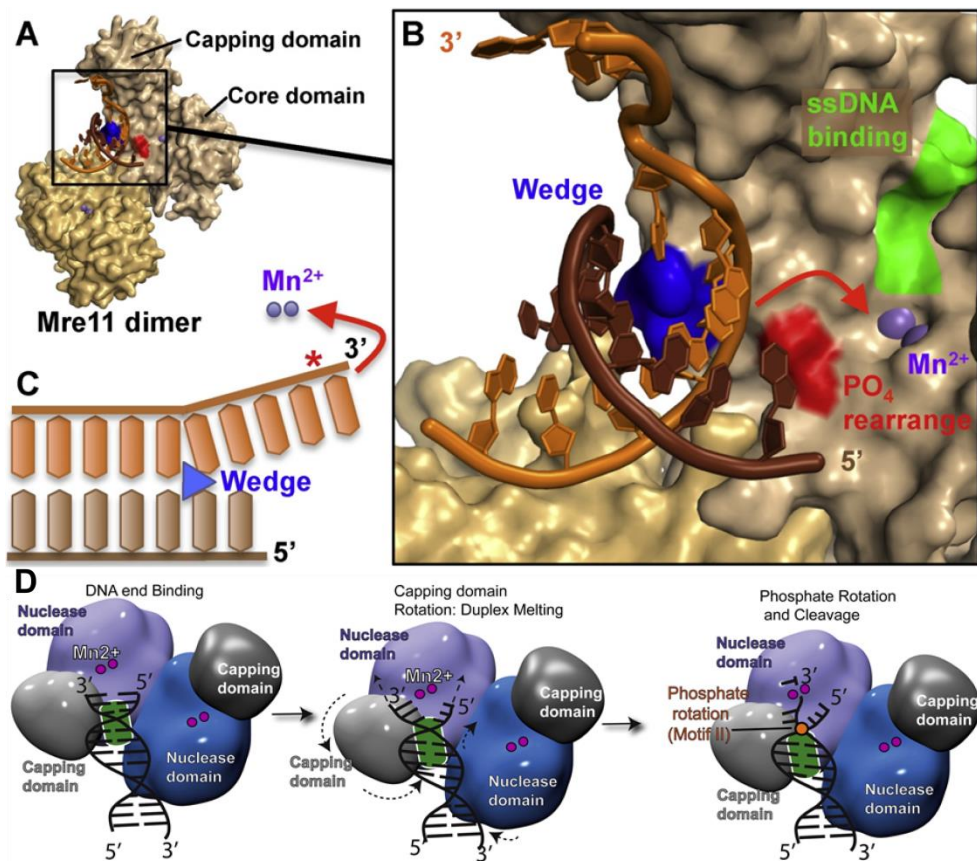
Figure 1.19: Exonucleotytic mechanism of Mre11, involves DNA conformational changes at the DNA terminus. *(A) Mre11 dimer bound to branched DNA (3DSD). (B) Magnified view of active site with ss-ds junction of bound DNA shown, uncleaved strand (brown) and incised strand (orange). The phosphate rotation, and inferred unpairing on bases required to associate with bound Mn^{2+} ions (purple) depicted by red arrow. The helical wedge (blue) sits in the minor grove of substrate. Also shown is the ssDNA binding site (green) proposed to be important for ds endonuceoytic activity of Mre11. (C) Schematic representation substrate sculpting mechanism placing the scissile phosphate (red star) in the active site. Taken from Tsutakawa, S. E. et al. (2014) DNA Repair 19, 95-107. (D) Overall view of mechanistic steps involved to afford exonucleoytic catalysis of blunt DNA by $(Mre11)_2$ (3DSD). Adapted from Williams, S. R. et al. (2008) Cell 135, 97–109.*

1.4.4 λ exonuclease as a DNA unpairing enzyme

The λ exonuclease recognizes and incises on the 5'-strand of blunt end DNA in the resection process of bacteriophages (Carter and Radding, 1971). DNA bound and free structures have shown that λ exonuclease forms a tordial homotrimer (*figure 1.20a*) (Kovall and Matthews, 1997; Zhang et al., 2011). The three active sites are equidistantly positioned around a central channel, but only one is used at any one time (*figure 1.20a*). In the substrate bound crystal structure, the terminus is approximately 15 Å away from two Mg^{2+} ions bound in the active site (Zhang et al., 2011). The ds DNA binds to the central channel, but access to active site is only

wide enough for ss DNA. Therefore, the authors argue that the two terminal nucleotides must unpair to allow placement of the reacting strand in the active site (*figure 1.20c*). This process is driven by the attraction between the negatively charged phosphate and the metal ions. In agreement with the structure of a non-5'-phosphorylated substrate not in the unpaired form, the scissile phosphate is not fully inserted in the active site (Zhang et al., 2011). Evidence also suggests that Arg-28 at the back of the active site pocket contacts the phosphate helping to drive this DNA conformational change (*figure 1.20b*), consistent with data showing that mutation of this residue disrupts activity (Zhang et al., 2011). Like Mre11, λ exonuclease employs a hydrophobic wedge (four apolar residues been called DE loop) to aid strand separation and Lue-78 of the wedge is inserted between bases at the ss-ds junction (*figure 1.20c).*

Unlike other structure specific nucleases discussed in this section, λ exonuclease is proposed to be processive (Subramanian et al., 2003). However, there is no evidence to suggest that large protein structural changes are responsible for driving movement along DNA. Therefore, a 'ratchet' mechanism has been proposed involving successive double nucleotide unpairing events to propel the enzyme forward (*figure 1.20c*). This is in agreement with biochemical data reporting a direct link between the rate of digestion and stability of the terminal base pair (van Oijen et al., 2003). The electrostatic attraction between the terminus and active site is predicted to help pull the enzyme along the DNA, whereas the Lue-78 residue prevents backward movement and slippage off the substrate. A similar active site topology is observed for RecE (Zhang et al., 2009) and Xrn1 (Jinek et al., 2011); these two enzymes are also believed to be processive. XRN1 is believed to be a member of the FEN-superfamily. XRN1 is a exoribonuclease. Investigating the mechanism of how these enzymes deal with both DNA and RNA may shed light how substrate specificity is mediated by distortion.
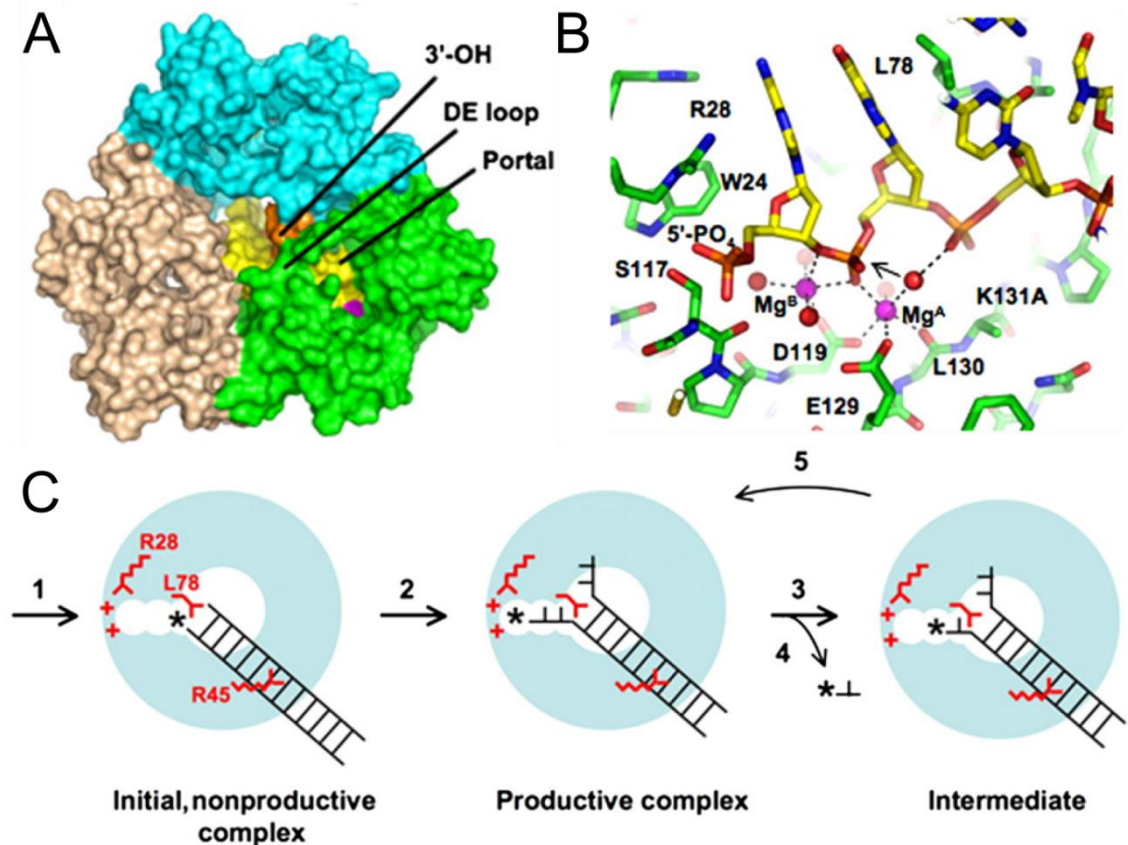
***Figure 1.20: The λ exonuclease mechanism involves parting of the 3'- and 5'- strands of blunt
DNA to place the 5'-terminal phosphate in positively charge active site. (A)*** *Surface view of
homotrimer, depicting the DE loop, Val-73, Ala-75, Ala-77, and Leu-78 which is predicted to split
the two strands. Also shown is a possible exit tunnel for the cleaved mononucleotides.* ***(B)***
*Magnified view of active site highlighting important residues and catalytic Mg^{2+} ions (magenta
spheres). The arrow indicates the angle attack of hydrolytic water molecule on the scissile
phosphate.* ***(C)*** *Proposed ratchet mechanism and requirement for double nucleotide unpairing
for λ exonuclease processivity. The 5'-phosphate group on the DNA represented as an asterisk
(*). The positively charged pocket at the end of the active site cleft contains metal ions (+) and
arginine residue. Step 1: initial capture of substrate. Step 2: 5'-phosphate attracted to metal
ions, simultaneous insertion of hydrophobic wedge into base-pairs. Step 3: hydrolysis. Step 4:
release of single nucleotide. Step 5: Electrostatic interactions pull the enzyme along the DNA and
in combination with the wedge force two nucleotides to unpair. Lue-78 hydrophobic wedge,
inserts in-between second and third bases of duplex. R28 contacts the terminal phosphate. Arg-
45 has a central role in minor groove binding. Taken from Zhang, J. et al. PNAS (2011) 108, 11872-
11877.*

1.4.5 DNA Polymerases

The highly fidelity of DNA polymerases (DNAP) is not only a consequence of their ability to select
the correct deoxynucleoside triphosphate (dNTP), but also due to removal of misincorporated
nucleotides by an internal 3'-5' exonuclease activity (Kunkel, 2004). Structural studies of the
DNAPs revealed that the polymerase (*pol*) and 3'-5' exonuclease (*exo*) domains are spatially in

different enzyme domains raising the question of how these activities work together to provide a proofreading function. The special separation of these two catalytic centres is maintained in other bacterial homologues, like the pol I family (also known as family A) (Doublie et al., 1998), B-family polymerases despite significant structural differences (Shamoo and Steitz, 1999; Swan et al., 2009), and C-family polymerases (Evans et al., 2008).

DNAPs have been crystallised in both the suspected 'extending' and 'editing' conformation (Beese et al., 1993; Freemont et al., 1988; Li et al., 1998). Comparison of these structures revealed that in the *pol* conformation base-pairing was maintained at the primer-template junction, but in the *exo* conformation the 3'-terminus primer-template was unpaired. Thus, it has been proposed that on the rare occasion in which DNAPs incorporate a mismatched nucleotide, primer-template strand unpairing transfers the 3'-terminus of the primer strand to the exonuclease site allowing the proofreading activity to excise the incorrect nucleotide. Spectroscopic studies have revealed that both sites can be sampled by the same substrate, but partitioning between the two sites is dependent on the state of the primer-terminus. A completely complementary primer-template junction is bound predominantly at the pol site, observed by sm-FRET (Tsoi et al., 2003). The introduction of mismatches at the primer-template junction results in a faster rate of transfer from the *pol* to *exo* site (Lamichhane et al., 2013).

Many steady-state and time-resolved fluorescent measurements have been carried out with bacteriophage polymerases using 2-AP-containing DNA (Fidalgo da Silva et al., 2002; Frey et al., 1995; Hariharan and Reha-Krantz, 2005; Subuddhi et al., 2008). The fluorescent analogue of adenine, 2-aminopurine (2-AP) has long been used as a probe of DNA structure, distortion and dynamics owing to its emission properties being dictated by its environment (Guest et al., 1991; Jean and Hall, 2001). 2-AP has reasonable fluorescence signal in ssDNA but is strongly quenched by base-stacking interactions when it is present in double-stranded DNA. In general, binding of DNAPs enhances the steady-state fluorescence of 2-AP primer-templates, and time-resolved experiments demonstrate a reduction in the population of shorter-lifetime components. Furthermore, both phenomena are exaggerated by presence of a mismatch. These results were interpreted as evidence for increased strand separation on formation of exonuclease-competent complexes.

Connolly and co-workers have used similar approach to study the uracil and hypoxanthine recognition capabilities of archaeal family B-polymerases (Richardson et al., 2013). The results showed increased primer-template separation upon encountering DNA containing deaminated base in the template strand a mechanism to prevent replication past the deaminated base. In

support of this, a crystal structure of *Thermococcus gorgonarius* (*Tgo*-Pol) bound to DNA containing hypoxanthine showed unpairng at the primer-template junction (*figure 1.21*).
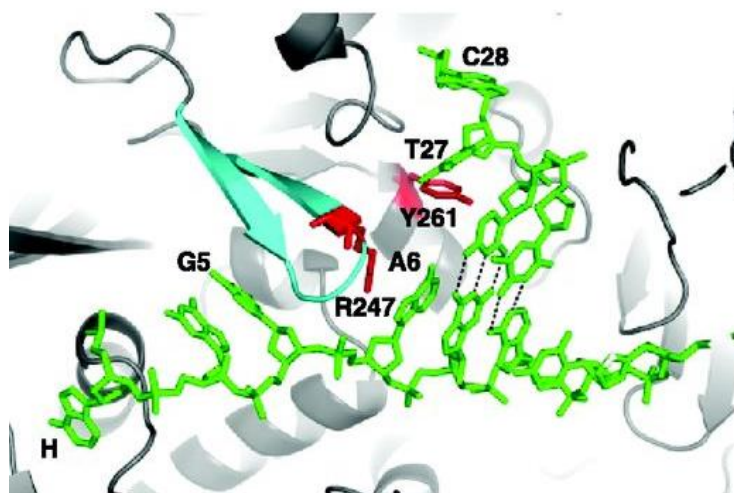
Figure 1.21: Structure of the DNA separation region of Tgo-Pol bound to a primer-template mimic containing hypoxanthine at the +2 position. *The two bases at the extreme 3' end of the primer, C28 and T27, are single stranded and not base paired with their template partners, G5 and A6, respectively. Thus, the base pairs nearest to the 3' end of the primer are C26–G7 and T25–A8 (shown with hatched lines indicating Watson–Crick hydrogen bonds rather than being individually labelled). The position of hypoxanthine (H), which is buried in the deaminated base–binding pocket, is also shown. The β-hairpin motif is shown in cyan. Tyr261 and Arg247 are shown in red. Taken from Richardson, T. T. et al. (2013) 1- 13.*

Von Hippel and co-workers developed a related approach termed exciton-coupled circular dichroism (ECCD), that involves monitoring of changes in the low-energy CD spectrum of two tandem 2-APs substituted in the primer strand, to characterize the solution conformations of DNA when bound to polymerases (Datta et al., 2009; Johnson et al., 2004, 2005b). A theoretical description regarding the background of this technique is described in *section 4.2*. The advantage of this technique over fluorescence measurements was that it is able to *directly* report on the relative conformations of the DNA at the position of the 2-AP probe. In fluorescent experiments, the 2-AP had to be placed in the template strand, and thus, the position of the primer 3'-terminus was ultimately being indirectly reported on. Therefore, ECCD complimented steady-state and time-resolved fluorescence measurements because it was able to report on DNA conformation of the primer strand and did not have to rely on structural information to interpret the results.

The major findings were consistent with previous studies suggesting that the terminal bases of the primer strand bind in an extended conformation in the *exo* site, while remaining stacked when bound at the *pol* site of DNAPs (Datta et al., 2009). Specifically, spectroscopic studies of 2-AP dimers in the context of ssDNA suggested that the four 3'-terminal bases of the ssDNA are significantly unstacked consistent with the crystallographic studies of a DNAP-ssDNA complex (Freemont et al., 1988). Like λ exonuclease, divalent metal ions were required for producing a DNA conformational change, suggesting the 3'-end of the DNA does not interact with the *exo* site when no metals are present. Collectively, these studies and previous studies suggest that stabilisation of an unpaired state may have been a way of halting replication in the presence of errors that may have evolved in exonuclease-mediated proofreading capabilities.

This technique has been extended to study RNA Polymerases and other components of the transcription complex (Datta et al., 2006; Datta and von Hippel, 2008; Johnson et al., 2005b), in addition to DNA helicases (Jose et al., 2012). It has been postulated that dsDNA binding – unpairing – ssDNA incision may be the hallmark of the FEN superfamily and even structurally unrelated nucleases discussed above. Most of the evidence for DNA unpairing by structure-specific nucleases is derived from structural data. Therefore, there is a need to observe these possible mechanistic steps in solution. Central to understanding of the unpairing mechanism is to determine what is necessary for these enzymes to sense DNA structures able to unpair, and then be able to promote/capture the unpaired state.

1.5 FEN Selectivity for free 5'-termini

Initially the 5'-flap of FEN substrates was believed to play an important role in binding, however, this conclusion was based on experiments conducted without knowledge of the importance of the 3'-flap. Later reports have demonstrated that the 5'-flap length does not affect the binding affinity when there is a 1nt 3'-flap is present (Patel et al., 2012). Moreover, total removal of the 5'-flap to create a nicked DNA substrate with a 3'-flap has very little effect on the binding of these substrates to FENs (Williams et al., 2007a).

Other early reports concentrated on the importance of a free 5'-flap terminus. The cap region of the I-domain determines the dimensions of the helical arch of FEN superfamily members. When structured, the helical arch of hFEN1 bridging the active site is only large enough to accommodate ssDNA and not dsDNA (*figure 1.22)*. Bulky additions (cisdiamminedichloroplatin (CDDP) and streptavidin (SA)) to the 5'-flap prevented FEN reaction (Barnes et al., 1996; Murante et al., 1995). A subsequent tracking model was proposed, in which, FEN initially recognises the 5'-end of the flap and tracks down the ss-DNA pulling itself towards the ds/ss junction using the

structured helical arch as a winch (Murante et al., 1995). In total there have been three major models of FEN recognition proposed. A clamping model in which the 5'-flap does not pass through the helical arch but instead the helical arch 'folds' over the 5'-flap allowing intimate association with the catalytic residues contained on the gateway (Chapados et al., 2004). Finally, a third model proposing binding of the upstream and downstream ds regions first, followed by flap threading (threading model) (Ceska et al., 1996).
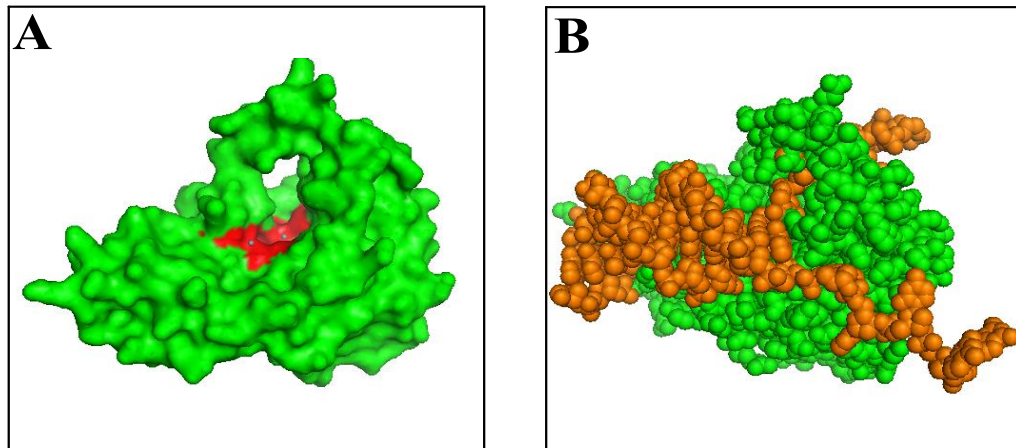
Figure 1.22: Space-filling models showing that single stranded DNA is able to pass under the ordered helical arch. *(**A**) Free enzyme complex, the helical arch bridges the active site groove (red). (**B**) Enzyme substrate complex, DNA (gold) bound to T4FEN (J. M. Devos, Journal of Biological Chemistry, 2007, **282**, 31713-31724) with the X-ray structure of T5FEN with no substrate bound. Taken from Ceska, T. A. et al. Nature (1996) 382, 90-93.*

FEN activity on 5'-flaps containing regions of duplex, so called GAP-substrates (*figure* 1.23) typifies the uncertainty surrounding correct model selection. FEN has robust GAP-endonuclease (GEN) activity with similar rates to a ss 5'-flap (Liu et al., 2006). Relating this *in vitro* GEN activity to a biologically relevant *in vivo* activity, FEN1 may be important in resolution of trinucleotide repeat (TNR) secondary structures (Singh et al., 2007). FEN1 has been proposed to endonucleolytically cleave the fold-back hairpin structures formed by $(CTG)_n$ and $(GAA)_n$ repeats (Zheng et al., 2011), with FEN1 deletion causing TNR expansions and contraction (Navarro et al., 2007). Although it may turn out to be the case that FEN1 is not involved in TNR resolution, mechanistically knowing FEN can cleave a hairpin in the flap is important.

Two independent studies have shown that 5'-fold-back hairpin flaps (known as gapped flaps) are excised *in vitro* as a duplex DNA products and exclusively by direct endonucleoytic cleavage, disfavouring a tracking model (*figure* 1.23) (Patel et al., 2012). Bulky adducts cannot fit through a structured or unstructured arch. However, theoretically the dsDNA contained within the 5'-flap of GAP-substrates could pass through a disordered helical arch, which then reorders after

threading around the single stranded region of the flap. The disordered helical arch is ~34 residues in the DNA-free structures, and assuming disordered residues may stretch ~3.8 Å each, the unfolded arch could have three ~42 Å sides, which is sufficient to accommodate ds DNA flaps. The arch could then reorder around the ss-region (gap) of the 5'-portion of the substrate, which is in agreement with data showing that FENs cannot process duplex Y-DNA junctions but their activity increases with increasing length of the ssDNA region before the hairpin (*figure 1.24*).
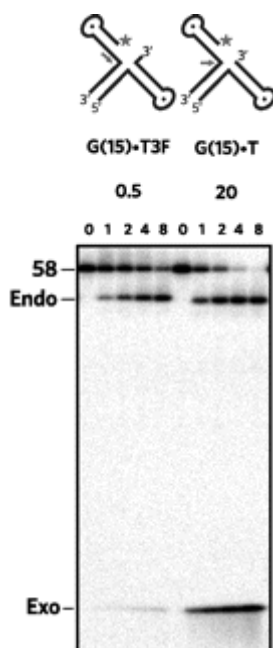
Figure 1.23: Double-flap substrates, containing 5'-fold back hairpins in the flap strand (GAP-ENDO substrate) and the products of reaction with hFEN1 shown on SDS Page gel. *These two substrates are identical in sequence but differ in the placement of the FAM label (star). One substrate has the FAM label on the 3'-end of the flap strand (left), the other on the 5'-end of the flap strand (right). Analysis of the hFEN1 cleavage products of these substrates shows that the DNA is incised at the same base (indicated by an arrow) ruling out progressive EXO cleavage at the 5'-terminus of the flap. Taken from L. D. Finger, L. D. et al. JBC (2009) 284, 22184-22194.*
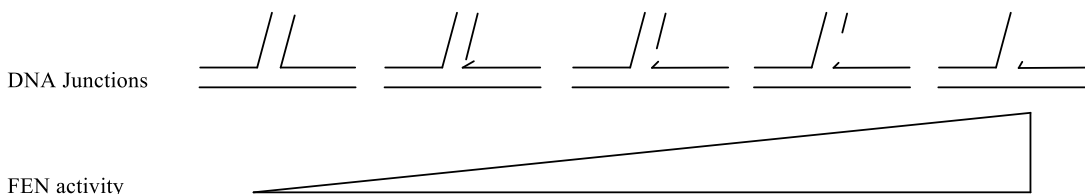
Figure 1.24: Schematic representation of DNA junctions with varying gap between the ds downstream and upstream regions and the ds DNA contained within the 5'-flap with a dramatization of hFEN1 activity below. *Relative FEN activity on Y-junctions in 0% and on ss 5'-flaps with a one nt 3'-flap is 100% with increasing gap length corresponding to increasing endonucleoytic cleavage.*

Equally, the GAP-endonuclease activity of FEN1 can be explained by invoking the clamping model. The helical arch could fold over the ss-region of the 5'-flap, placing the key residues near the target phosphodiester, and simply by-passing the need to transverse the hairpin of the 5'-flap. However, the previous data presented whereby blocking the 5'-end prevents cleavage suggests that this model is not the major mechanism of endonucleoytic cleavage. Additionally, there is data to suggest that the 5'-flap can be trapped threaded through the helical arch (Patel et al., 2012). Preincubation of the enzyme and a 5'-biotinylated substrate in the presence of catalytically inert divalent metal ions then addition of streptavidin allows capture of the flap in a threaded position. Moreover, 5'-flap containing hairpin substrates can also be trapped using the same technique (Patel et al., 2012). In this state, the substrate cannot be competed out of the complex. In support of this, the dissociation rate of DF substrates was slow after substantial substrate conformational changes upon FEN binding (i.e. threading) (Craggs et al., 2014). Furthermore, the dissociation rate was shown to significantly decrease in relation to increasing 5'-flap length (Gloor et al., 2010; Sobhy et al., 2013), presumably because substrate release is more difficult when the 5ʹ-flap was threaded. However, a variation of the clamping model has been proposed to explain hEXO1 activity on DNA nicks; highlighting the controversy of this particular topic. This alternative model suggests that the 5'-flap passes either side of the gateway and not through α4 and α2 (Orans et al., 2011).

1.6 FEN1 Recognition of the ds-junction

The majority of the contacts made in the FEN1 enzyme-substrate and product complexes are with the dsDNA either side of the junction. There are very little contacts made with the 5'-flap, supported by data showing that FEN can bind blocked 5'-flap substrates with a similar affinity to the analogous unmodified substrate (Gloor et al., 2010; Hohl et al., 2007). The rate of substrate conformational rearrangements is slower than the diffusion limited association predicted for FEN1 (Craggs et al., 2014). This suggests that a step after initial substrate binding is occurring, possibly accommodation of the 3'- and/or 5'-flap (Craggs et al., 2014). Collectively these suggest that FEN1 targets the ds junction in DF substrates.

The importance of the H2TH and 3'-binding pocket with regards to both binding and activity have been discussed. The idea that FENs initially recognise the ds junction unifies the mechanism of all the family members as they do not all catalyse the phosphate diester hydrolysis of substrates with free 5ʹ-flaps (Grasby et al., 2012). Specifically, Okazaki maturation and lpBER intermediates, whose ability to bend upon hFEN1 binding and the combination of the 3ʹ-binding pocket, double nucleotide unpairing, helical gateway and cap select for the free 5'-terminus.

The requirement for the helical arch to disorder to accommodate certain 5'-flap structures and then restructure positioning essential catalytic residues had led to a proposed disorder-to-order mechanism (Grasby et al., 2012; Tsutakawa et al., 2011) (*figure 1.25*). Binding of FENs to the dsDNA two-way junction, and subsequent bending it thought to position the 3'-flap in the hydrophobic pocket and direct the 5'-flap towards the active site. The helical gateway has been proposed to act to ensure that the 5'-flap is threaded through the disordered helical arch. As previously described, binding of the 3'-flap and ordering of the pocket causes an allosteric ordering of the helical cap. Double nucleotide unpairing is then thought to occur, positioning the target phosphate diester bond between the two metal ions. The reverse conformational change of the helical cap is thought to cause product release, and studies which have deliberately disrupted the proper formational of the helical arch have shown this to severely retard the activity of FEN (Hwang et al., 1998; Storici et al., 2002).

This disorder-to-order model is proposed not only to explain how dsDNA within the 5'-flap can be accommodated but also is part of a self-regulation mechanism of hFEN1. Unlike EXO1, where the C-terminal domain contains potential inhibitory domains that could prevent ordering of the helical cap, hFEN1 must internally regulate its activity. The efficient 5'-structure-specific endonuclease activity of hFEN1 is dependent on all of the structural elements of the bifurcated DNA being present. After 'one-round' of cleavage FEN would not be capable of processive cleavage since the double-flap structure is unable to form. Cleavage one nt into the downstream duplex would yield a one nt gap product after the 3'-flap had bound to the complementary template strand. This gap structure would be unable to form a double equilibrating flap structure required for endonucleoytic by hFEN1.
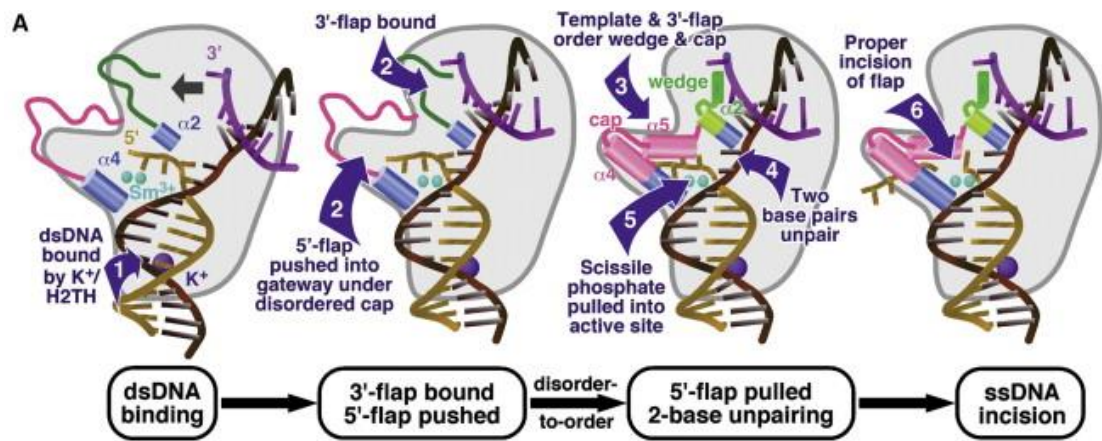
Figure 1.25: Cartoon representation of the disorder-to-order mechanism employed by FENs to position the 5ʹ-flap in the active site and stimulate product release. FEN1 recognises the dsDNA flanking the ss/dsDNA junction interacting through the K+:H2TH and Hydrophobic wedge/β-pin. The relative position of these two binding motifs sharply bend ~100° the DNA. Consequently, the 5ʹ-ss flap is directed under the disordered cap domain and through the helical gateway selecting for ssDNA nearing the active site. Assembly of the 3ʹ-flap site, and subsequent cap ordering positions residues which promote double base unpairing positioning the scissile bond and two-metal ion catalysed hydrolysis reaction of the specific phosphate diester one nucleotide into the downstream dsDNA. Taken from Tsutakawa, S.E., et al. (2011). Cell 145, 198-211.

1.7 Current cancer treatments: synthetic lethality

Normally hereditary mutations are prevented as cell cycle progression is tightly coupled to DNA damage response (DDR) pathways. DNA repair pathways including direct repair (DR), BER, NER, mismatch repair (MMR), the fanconi anemia (FA) pathway and homologous recombination (HR) are encompassed under the umbrella of DDR. The pathways are redundant for certain types of DNA damage, and interact with each other to regulate DNA repair (Sancar et al., 2004). However, failure to repair DNA damage by any of these pathways can result in cell cycle arrest, cell death, or cell senescence. In contrast, activated oncogenes allow cancer cells to circumvent these cell-cycle checkpoints inducing replicative stress generating highly proliferating, highly unstable cells.

Alkylation-induced DNA damage is one of the most efficacious anti-cancer therapeutic strategies (Zhang et al., 2012). Chemotherapeutic agents exploit the inherent weakness of most tumour types to perform all DDR functions to the same level as normal non-cancerous cells. However, cancer cells are normally very good at recognizing damage and upregulating the relevant DNA repair pathways to compensate for the impairment of other DDR pathways which is an important mechanism for therapeutic resistance (Longley and Johnston, 2005). In an attempt to combat the adaptability of cancer cells, drug target strategies have been designed based on the concept of synthetic lethality (Wright and Dobzhansky, 1946). These approaches do not

41

specifically target the oncogene per se, but rather exploit the defect by targeting a second unlinked gene partner (Kaelin, 2005). Thus, a cancer cell harbouring an oncogenic mutation may be susceptible to loss of function of another gene that is not essential for normal cells. For example, if the two gene products lie on parallel pathways dealing with DNA alkylation damage then in principle, only cancer cells harbouring defects in one pathway will be susceptible to targeted inhibition of the alternative pathway, and thus all normal cells will remain largely unaffected by chemotherapeutic exposure. The challenge is to identify which pathways to target (Sajesh et al., 2013).

Knowledge of the genetic differences between normal and cancerous cells paves the way for personalized cancer therapy treatments based on distinguishing between which DDR pathways are activated and targeting them using a synthetically lethal strategy (Furgason and Bahassi, 2013). Invariably the situation is complicated by intratumor heterogeneity and developed resistance (Gerlinger et al., 2012). Therefore, targeting of single ongogenes is often not sufficient, and combinatorial anticancer strategies are realistically required for long-term disease control, and potentially even cures (Arnedos et al., 2014).

1.7.1 The role of base excision repair in tumour survival

Biochemical, preclinical and clinical studies have established that the BER pathway is a promising target in anticancer treatment (Illuzzi and Wilson, 2012). The most well-known and characterised example of a synthetically lethal strategy at work, exploiting the DNA repair capabilities of the BER pathway, is the interaction observed between BRCA1 (or BRCA2) and PARP1 (poly ADP-ribose polymerase 1). Mutations in *BRCA1* or *BRCA2* give predisposal to developing breast and ovarian cancer (Miki et al., 1994). BRCA1 and BRCA2 both function in HR-mediated repair of double-strand breaks (Moynahan et al., 1999). PARP1 is receptive to ssDNA and functions in one of the sub-classes of BER (El-Khamisy et al., 2003). The cross-over point of the two pathways occurs because ssDNA breaks (efficiently repaired by BER) can be converted to dsDNA breaks during replication (Kuzminov, 2001). Therefore, PARP1 inhibition results in an accumulation of double-strand breaks in *BRCA1* and *BRCA2* defective cells, ultimately leading to cell death (Farmer et al., 2005). There are various inhibitors of PARP1 at various stages of clinical trials (Papeo et al., 2013), and similar strategies are now being applied to other BER proteins.

The efficacy of the most widely used methylating agent, Temozolomide (TMZ) classically has been attributed to the formation of the highly toxic O^6-methylguanine (O^6-MeG) lesion (Wedge et al., 1996a; Wedge et al., 1996b). However, O^6-MeG only accounts for approximately 20 % of the DNA damage induced by TMZ, and the majority of the *N*-methylated base adducts are

processed efficiently by BER (*figure 1.26*) (Liu and Gerson, 2004; Samson et al., 1991). Because MMR deficient cells, which is common in some cancers, can replicate in the presence of high levels of O^6-MeG (*figure 1.26*). MMR deficient cells are especially sensitive to methylating agents in combination will PARP1 inhibitors (Liu et al., 1999; Tentori et al., 2002). Another example of a synthetically lethal interaction; thus, other proteins involved in BER are potential therapeutic targets.
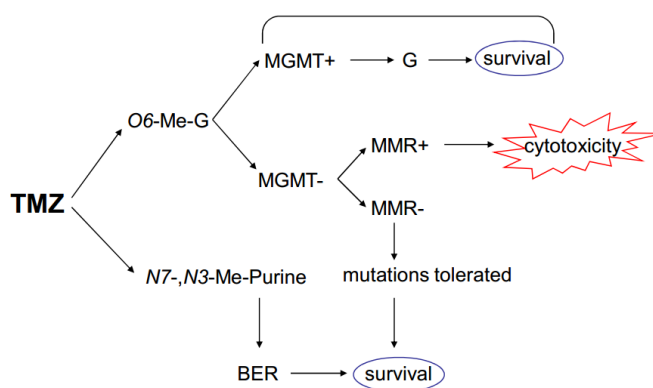
Figure 1.26: The interplay between DR, BER, and MMR Predicting the cellular survival after TMZ exposure. *Taken from Zhang, J. et al Current Molecular Pharmacology (2012) 5, 102-114.*

As discussed in *section 1.3.4*, AP endonuclease activity sits at the branch-point between sp and lpBER pathways. APE1 is the major human AP endonuclease, and is over-expressed in several types of tumours (Kakolyris et al., 1998; Moore et al., 2000; Silber et al., 2002). Thus, APE1 has been targeted based on chemosentization (Silber et al., 2002; Walker et al., 1994), and synthetically lethal interactions (Fung and Demple, 2005; Lau et al., 2004). Small molecule inhibitors of APE1 with chemotherapeutic potential have been published, and their mode of action are discussed in *section 1.8.1* below.

FEN1 activity is essential for lpBER (*section 1.3.4*) and FEN1 knockdown or inhibition has been demonstrated to increase sensitivity to TMZ in glioblastoma (Nikolova et al., 2009), and colorectal (McManus et al., 2009; van Pel et al., 2013) cancer cell lines. Although FEN1 is involved in only one of the sub-classes of BER, it remains a broad target for therapeutics due to its role in DNA replication. Furthermore, FEN1 is proposed to be involved in the alternative lengthening of telomeres pathway (ALT) (Saharia and Stewart, 2009), utilised by some cancers (Durant, 2012). In established tumours FEN1 overexpression has been characterized in breast (Singh et al., 2008), prostate (Lam et al., 2006), and brain (Nikolova et al., 2009) cancer types. Importantly,

preclinical evidence suggests that increased FEN1 expression levels sustain cancer survival and aggressiveness (Abdel-Fatah et al., 2013; Kim et al., 2005; Krause et al., 2005; Sato et al., 2003). Proof-of-concept, exploiting FEN1 as in a synthetically lethal approach was shown in *CDC4* deficient cells (van Pel et al., 2013). Mutations in *CDC4* are commonly associated with certain cancer types; complete or partial loss of *CDC4* in model organisms sensitizes cells to FEN1 inhibition. In summary, published data in pre-clinical studies suggests that FEN1 expression levels and activity are significant in cancers; thus blocking FEN1 activity with small molecule inhibitors would be predicted to potentiate the cytotoxicity of therapeutic agents (*section 1.8.4*).

1.8 Metal-binding Inhibitors

1.8.1 APE1 inhibitors

APE1 catalyses the hydrolysis of the DNA backbone on the 5$'$-side of AP sites in a Mg^{2+} dependent manner (Boiteux and Guillet, 2004; Dianov et al., 2003). APE1 has been successfully crystallised in its apo form and complexed with either substrate or product DNA. The protein has a two layered beta sheet core surrounded by helices, and binds to both major and minor grooves of DNA (*figure 1.27a*) (Mol et al., 2000b). The majority of the protein:DNA surface is mediated through the damaged DNA strand and the APE1-bound substrate is distorted 35° to a regular B-DNA helix. Specificity is afforded by the protein surface adjacent to the active-site forming shallow DNA grooves sterically selecting DNA lacking a bulky-base. The DNA backbone is flipped ~180° at the AP site positioning the scissile phosphate appropriately for in-line nucleophilic attack by an activated water molecule (*figure 1.27a*) (Mol et al., 2000b). The flipped out abasic deoxyribose is stabilised in a hydrophobic pocket (Gorman et al., 1997; Mol et al., 2000a). Initially, APE1 was proposed to utilise a two metal-ion mechanism (Beernink et al., 2001). However, more recent crystallographic data of free (Manvilla et al., 2013), and product bound (Tsutakawa et al., 2013), structures predict the presence of only one Mg^{2+} ion (*figure 1.27b*), consistent with recently published mutational studies (Kanazhevskaya et al., 2014).

Several compound library screens have identified potent inhibitors of APE1 (Bapat et al., 2010; Madhusudan et al., 2005; Seiple et al., 2008; Simeonov et al., 2009; Zawahir et al., 2009). A selection of the most potent is shown in *figure 1.28*, and most share an aromatic hydrophobic core with ionisable substituents, frequently a carboxylate. The first small molecule inhibitor of was CRT0044876 (7-Nitro-indole-2 carboxylic acid, *figure 1.28*), reported to have an IC$_{50}$ value of 3 μM (Madhusudan et al., 2005). More recently, improved lead compounds were identified in a screen of well-characterized drug-like molecules, which identified 6-hydroxy-dl-DOPA and myricetin (*figure 1.28*) as direct inhibitors of APE1 activity (Simeonov et al., 2009). *In silico*

modelling of 6-hydroxy-dl-DOPA and CRT0044876 suggested that they bind in a very similar manner (*figure 1.27c*). The carboxylate group is predicted to bind to the sequestered Mg^{2+} ion and the aromatic groups sit in the hydrophobic abasic pocket. DNA binding causes very little active-site conformational change (Gorman et al., 1997), and these types of inhibitor have been proposed to prevent reaction by occupying the hydrophobic pocket preventing flipping of the abasic nucleotide. Other APE1 inhibitors that have been identified whose activities are believed to be partly due to DNA-mimetic properties are aurintricaroxylic acid (Simeonov et al., 2009) (*figure 1.28*), and a series of arylstibonic acids (Seiple et al., 2008), resembling phosphonate derivatives; as a result these may have significant off-target effects.
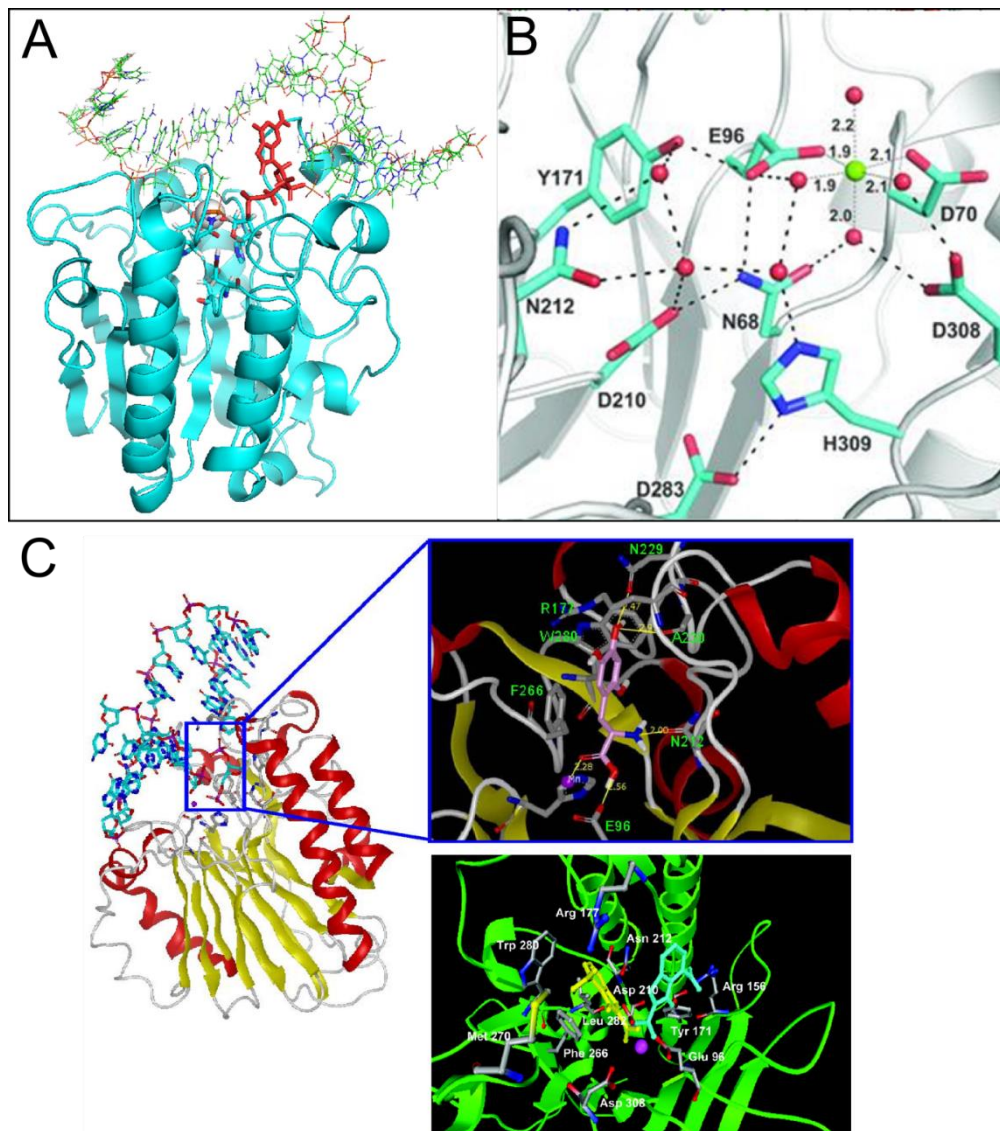
*Figure 1.27: APE1 structure and docking of inhibitors into the active site. (**A**) The atomic structure of human APE1 (PDB: 1DEW) containing a 15-mer DNA duplex with 2-AP residue flipped out into the active site (red sticks), taken from Kanazhevskaya, L. Y., et al. (2014). DNA Repair 21, 43-54. (**B**) Close-up view of the active site (PDB: 4IND), showing Mg^{2+} (green sphere) and several ordered water molecules (red spheres) Also shown are the important catalytic residues. (**C**) Left panel: ribbon representation of APE1 structure with DNA bound (PDB: 1DE9); the active site is marked within the box. Upper right panel: docking pose of 6-hydroxy-dl-DOPA in APE1 active site (PDB: 1DE9); potential hydrogen-bonds are labelled in yellow lines with distances, Mg^{2+} (magenta sphere) selected active site amino acid residues are indicated. Taken from Simeonov, A., et al. (2009). PLoS ONE 4(6): e5740. Lower right panel: docking pose of CRT0044876 in APE1 active site (PDB: 1DE9); CRT0044876 was found to dock on to the active site of APE1 in two different low energy conformations, (yellow and blue sticks), Mg^{2+} (magenta sphere). Taken from Madhusudan. S., et al. (2005). NAR 33, 4711-4724.*

CRT0044876 (3.0 µM)   6-hydroxy-DL DOPA (0.11 µM)   Myricetin (0.32 µM)

NSC-13793 (0.017 µM)   Aurintricaroxylic Acid (0.055 µM)
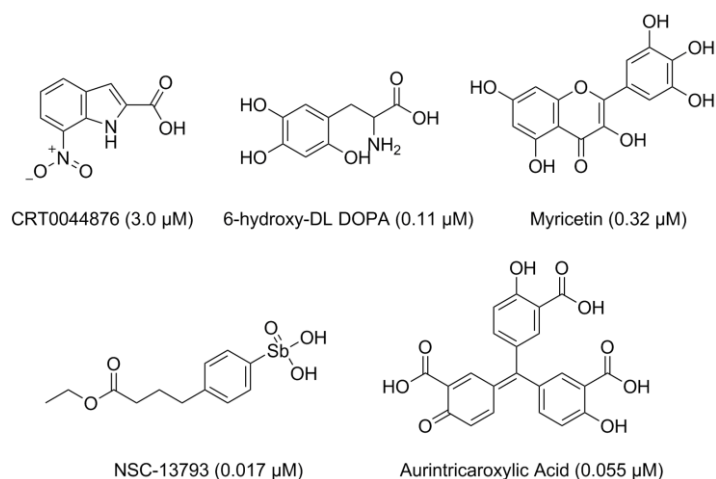
Figure 1.28: Identified inhibitors of APE1 endonuclease activity. Shown are the chemical structures and concentrations of compound reported to inhibit 50% of APE1 incision activity in vitro (in parentheses).

1.8.2 Mre11 Inhibitors

As discussed in *section 1.4.3*, Mre11 is reported to have both ss exo- and ds endo-nuclease activities; their roles in resection are unclear. However, both activities are proposed to be important for the regulation of pathway choice during the repair of DSBs (i.e switching from NHEJ to HR) (Shibata et al., 2014). The separate activities can be individually inhibited by the use of related compounds: mirin and PFM39 primarily block exonuclease and PFM03 and PFM01 inhibit endonuclease activity (*figure 1.29a*) (Dupre et al., 2008). These inhibitors can be used as tools to probe Mre11's activity and suggest Mre11 employs distinct mechanisms for exo- and endo-nuceloytic cleavage. As previously mentioned Mre11's ability to open (unpair) the dsDNA to allow access to the active site is crucial for its exonuclease activity. The crystal structure of DNA free Mre11 containing two Mn^{2+} ions in its active site complexed with each of these inhibitors shows that exonuclease and endonuclease-inhibitors bind at distinct sites. Neither are proposed to interact directly with the catalytic divalent metal ions but prevent reaction by blocking access to them. The endonuclease inhibitor sits in a previously proposed ssDNA binding channel (*figure 1.29b*). In contrast, the exonuclease inhibitor sits in a pocket between this ssDNA channel and the essential His-residue (*section 1.4.3*) (*figure 1.29b*). Therefore, mirin and PFM39 presumably block dsDNA unpairing, a requisite for association of the scissile phosphate and Mn^{2+} ions.
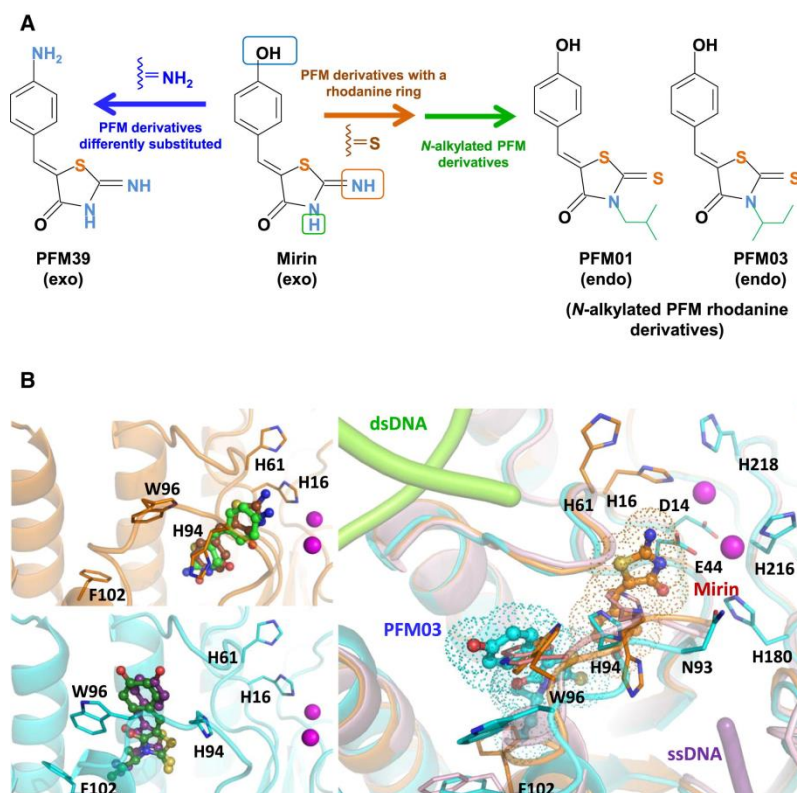
Figure 1.29: Exo- and endonuclease inhibitors of Mre11. (A) *Functionalisation strategy from known exonuclease inhibitor mirin, to novel exonuclease inhibitor PFM39 and endonuclease inhibitors PFM01 and PFM03. **(B)** The top left panel shows an overlay of cocrystal structures of Mre11 (cartoon, coloured brown) in complex with the exonuclease inhibitors, mirin (sticks, carbon atoms colored brown), and PFM39 (green). The bottom left panel shows an overlay of the cocrystal structures of Mre11 (cyan) in complex with the endonuclease inhibitors PFM01 (purple) and PFM03 (dark green). The right hand panel shows an overlay of the unliganded Mre11 (cartoon, colored pink) with mirin (brown) and PFM03 (cyan). Structures are of Themotoga maritima Mre111 (TmMre11), unliganded-complex (PDB 3T1I). The relative position of each inhibitor is highlighted by a dotted representation of its van der Waals surface. The proposed routes for ss and ds DNA into the active site have been modelled into the structures. The exonuclease inhibitor blocks dsDNA (light green) unpairing, the route for ssDNA (purple) to the metal ions is blocked by the endonuclease inhibitor. Nitrogen, oxygen, and sulphur atoms are coloured blue, red, and yellow, respectively. Active site metal ions are shown as magenta spheres. Taken from Shibata, A. et al. Mol Cell (2014) 53, 7-18.*

1.8.3 HIV-IN inhibitors

Integrase (IN) is one of three enzymes which are essential for the replication of HIV-1 (Lafemina et al., 1992), the other two being reverse transcriptase and protease, but until only relatively recently HIV-1 IN has been an unexploited target for antiretroviral therapy.

IN catalyses two reactions that are required for insertion of the reverse-transcribed viral genome into the host DNA: 3'-processing and strand transfer (Asante-Appiah and Skalka, 1999; Delelis et al., 2008). IN is a 32 kDa protein (288 aa) that contains three domains, the catalytic core, the N-

terminal domain and the C-terminal domain. The catalytic core (51-212) has a RNase-H like topology and contains the conserved DDE motif (Asp-65, Asp-116, Glu-152) that coordinates the divalent metal cofactor(s) required for phosphodiester bond cleavage/formation. Apart from the substrates being unique to each reaction the actual chemistry in both steps is very similar and the same catalytic residues are utilised. The two reactions are thought to be distinguished by the oligomerization state of the protein (Alian et al., 2009).

The development of a virologically relevant assay turned out to be the breakthrough in the discovery of IN inhibitors for clinical use (Hazuda et al., 2000). A screen was conducted using this assay, and the most potent and active compounds proved to be 4-aryl-2,4-diketones containing a distinct 1,2-diketo acid (DKA), such as pyrrole derivative (L-731,988) (*figure 1.30a*) (Hazuda et al., 2000). The subsequent design of DKAs isosters that include the diketo acid pharmacophore within a heterocyclic rings have led to the first commercially available antiretroviral agent to target HIV IN, Raltegravir (*figure 1.30b*) (Anker and Corales, 2008; Summa et al., 2008).
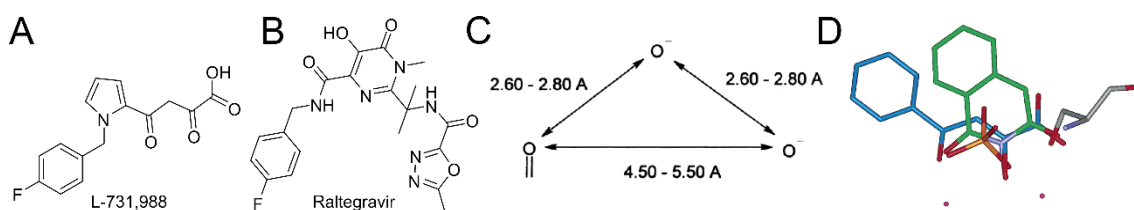
*Figure 1.30: The structure of DKA pharmacophore based HIV integrase inhibitors and proposed minimal pharmacophore for inhibitors of integrase. (**A**) L-731,918 was the first published HIV IN inhibitor, (**B**) Raltegravir. (**C**) Minimal pharmacophore, defined by spacing of three oxygen atoms. On or two oxygen atoms may be ionized. (**D**) Overlay of N-hydroxyimide and DKA compounds Taken from Parkes, K. E. B. et al. Journal of Medicinal Chemistry (2003) 46, 1153-1164.*

The mechanism of inhibition is proposed to be derived from the ability of DKAs to coordinate the essential divalent cofactors (Marchand et al., 2003), supported by evidence suggesting that mutation of two aspartate residues of the catalytic DDE motif prevented L-731,988 binding (Espeseth et al., 2000). Influenza endonuclease can also be targeted by DKAs based on the positionally conserved divalent cofactor(s) (Doan et al., 1999). Parkes *et al*, proposed a minimal pharmacaphore for inhibitors of influenza endonuclease (*figure 1.30c*) (Parkes et al., 2003). DKAs satisfy this specific spatial arrangement of oxygen atoms and a screen identified a range of metal-binding influenza endonuclease inhibitors among which were *N*-hydroxyimide compounds which also displayed the required minimal pharmacophore. An overlay of the DKA

and *N*-hydroxyimide pharmacophore highlights the similarity of the specific spatial arrangement of oxygen atoms (*figure 1.30d*). Furthermore, if the distances between the oxygen atoms are disrupted by changing the ring size of the *N*-hydroxyimide pharmacophore this also weakens the activity of the inhibitor (*figure 1.31*).These studies imply that successful inhibition requires the presence of a 'war-head' group which coordinates to the metal ions. More recently, raltegravir was modelled into the active site of the prototype foamy virus (PFV) IN:DNA complex (a suitable model of HIV-IN) (Krishnan et al., 2010). The results predicted that the war-head of raltegravir was able to coordinate the metal ions when 3′-processing product was present, and as proposed for Mre11 (above), it was this association that blocks the acceptance of host DNA, preventing strand transfer reaction (*figure 1.32*).
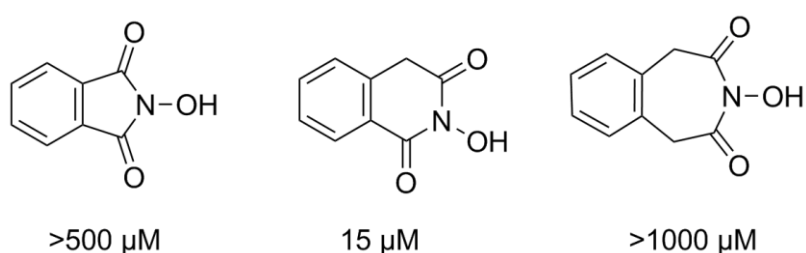
Figure 1.31: (A) N-Hydroxyimide inhibitors of influenza endonuclease. *The presence of the hydroxyimine is key for activity. The pharmacophore needs to comply with the spatial arrangement of the oxygen atoms in order to be able to directly contact the divalent cofactors. The numbers below each compound denote the IC_{50} value. Taken from Parkes, K. E. B. et al. J. Med. Chem. (2003) 46, 1153-1164.*
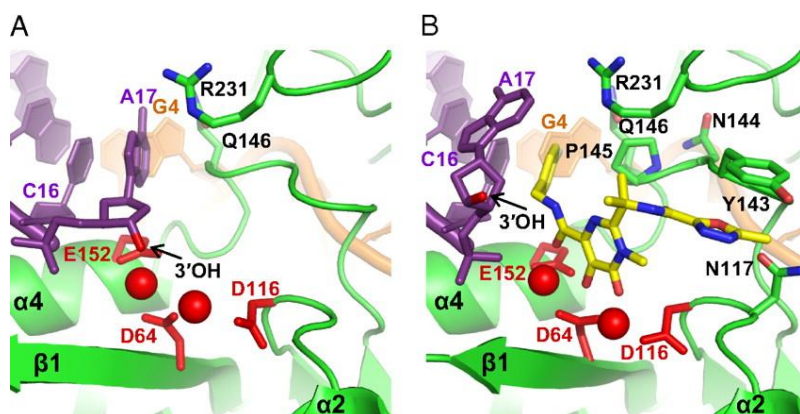
Figure 1.32: PFV IN active site with cognate DNA, Mg^{2+} ions (A) free and (B) raltegravir bound. *Protein (green) and DNA (purple) are cartoons, DNA bases and the side chains of indicated amino acids as sticks. Drug atoms are coloured: yellow, C; blue, N; red, O; grey, F; green, Cl. The 3'-OH of the terminal adenosine is labelled. Red spheres, Mg ions. Taken from Krishna, L. et al. PNAS (2010) 107, 15910-15915.*

1.8.4 FEN1 Inhibition

Potentially, small compounds containing a DKA or *N*-hydroxyimide pharmacophore could be used to inhibit any 2-metal ion dependent endonuclease. Therefore, this inhibition of influenza endonuclease by *N*-hydroxyimide compounds has been extended to hFEN1. Tumey *et al* (2005) designed a structure activity relationship (SAR) strategy to develop potent and selective inhibitors of hFEN1 and the final series of compounds had an aromatic core with *N*-hydroxy urea functionalisation (*figure 1.33*) (Tumey et al., 2005). Since then two separate large screens have identified inhibitors of hFEN1 (Abdel-Fatah et al., 2014; van Pel et al., 2013). Some of the most potent inhibitors identified were also *N*-hydroxy urea compounds (*figure 1.33*) and this class of inhibitors are approximately 100-fold more potent than any endonuclease inhibitor reported to date. The problem with specificity is highlighted by the fact that the two most recent screens also established that APE1 inhibitors myricetin and aurintricaroxylic acid (*figure 1.28*) also act on hFEN1. Although no direct evidence has been reported, *N*-hydroxy urea compounds are proposed to interact with active-site metal ions, therefore they too could potentially inhibit other metal dependent nucleases. Selectivity for FEN1 over XPG was reported but the reason for this bias was not investigated (Tumey et al., 2005). In summary, metal binding pharamcophores have emerged as an effective way of targeting metal dependent phosphodiesterases, and represent a good starting strategy to find potential inhibitors of nucleases. However, there remains a question over their selectivity.
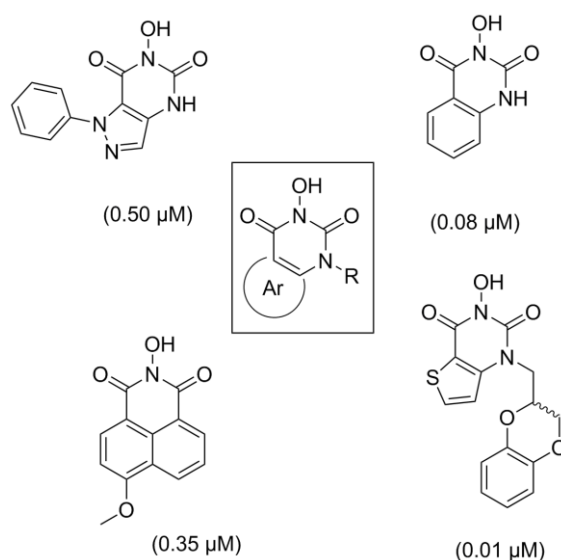
Figure 1.33: N-Hydroxyurea inhibitors of flap endonuclease. *Proposed minimal pharmacophore. (boxed). Left hand side the most potent inhibitors discovered from a large screen were N-Hydroxyurea compounds, taken from Abdel-Fatah, T. M. A., Molecular Oncology (2014) 8, 1326-1338. Right hand side most portent hFEN1 inhibitors reported to date, taken from Tumey, L. N. et al. Bioorg. Med. Chem. Lett. (2005) 15, 277–281. The numbers below each compound denote the IC_{50} value.*

1.9 Aims

The task of DNA structure-specific nucleases to manipulate their substrates to bring about efficient and at the same time specific catalysis is a challenging feat. This is proposed to occur in FEN1 using a double-nucleotide unpairing mechanism. The work detailed in this thesis aims to better clarify this mechanism using solution studies to directly observe this process. The technique adopted has previously been used to successfully determine the dynamics of nucleotides within larger structures and protein-engaged complexes. The role of hFEN1 active-site and helical-arch in this potential requiste reaction step will be investigated. To understand their effect on DNU it will be necessary to characterise the binding and catalytic activity of these variants. hFEN1 is a potential chemotherapeutic target, and inhibitors have been published but how they prevent hFEN1 from performing catalysis is unknown. Therefore, this work aims to determine their mode of inhibition and appreciation of the FEN molecular mechanism will aid this process.

Chapter 2: Materials and methods

2.1 Expression and purification of hFEN1

2.1.1 Media and buffers

LB (1L)	10 g tryptone, 5 g yeast extract, 10 g NaCl
TB (1L):	12 g tryptone, 24 g yeast extract, glycerol 4ml (v/v)
SOB media (1L)	20 g tryptone, 5 g yeast extract, 0.5 g NaCl, 10mL 250 mM KCl, 10mL 1M $MgCl_2$, 10 mL 1M $MgSO_4$
SOC media:	SOB media supplemented with 20 mM D- glucose
IMAC FF Buffer A:	20 mM Tris pH = $7.0^{25^{\circ}C}$, 1M NaCl, 0.02% NaN_3, 5 mM Imidazole, 5 mM 2-mercaptoethanol
IMAC FF Buffer B:	20 mM Tris pH = $7.0^{25\ ^{\circ}C}$, 500 mM NaCl, 0.02% NaN_3, 40mM Imidazole, 0.1% Tween20, 5 mM 2-mercaptoethanol
IMAC FF Buffer C:	250 mM Imidazole pH = $7.2^{25^{\circ}C}$, 500 mM NaCl, 0.02% NaN_3, 5 mM 2-mercaptoethanol
Hitrap Heparin HP Buffer A:	50 mM MES pH = $6.0^{25^{\circ}C}$, 1 mM EDTA, 0.02% NaN_3, 20 mM 2-mercaptoethanol
Hitrap Heparin HP Buffer B:	50 mM MES pH = $6.0^{25^{\circ}C}$, 1 mM EDTA, 0.02% NaN_3, 1M NaCl, 20 mM 2-mercaptoethanol
HiTrap Phenyl Sepharose HP Buffer A:	20 mM Tris pH = $7.4^{25^{\circ}C}$, 1.5 M $(NH_4)_2SO_4$, 1 mM EDTA, 0.02% NaN_3, 20 mM 2-mercaptoethanol
HiTrap Phenyl Sepharose HP Buffer B:	20 mM Tris pH = $7.4^{25^{\circ}C}$, 1 mM EDTA, 0.02% NaN_3, 20 mM 2-mercaptoethanol
TBSE	20 mM Tris pH = $7.6^{25^{\circ}C}$, 150 mM NaCl, 1 mM EDTA, 20 mM 2-mercaptoethanol
Anion Exchange Buffer A:	20 mM Tris pH = $8.0^{25^{\circ}C}$, 1 mM EDTA, 0.02% NaN_3, 20 mM 2-mercaptoethanol
Anion Exchange Buffer B:	20 mM Tris pH = $8.0^{25^{\circ}C}$, 1M NaCl, 1 mM EDTA, 0.02% NaN_3, 20 mM 2-mercaptoethanol
Sephacryl GF Buffer A:	100 mM HEPES, pH = $7.5^{25^{\circ}C}$, 200 mM KCl, 0.04% NaN_3, 20 mM 2-mercaptoethanol

Storage Buffer: 50% glycerol, 1 mM $CaCl_2$, 50 mM HEPES-NaOH pH=7.5, 100 mM KCl, 0.02% NaN_3, 10 mM dithiothreitol

2.1.2 Protein expression

This stage is analogous for full-length and truncated forms of hFEN1, with the exception that for full-length forms of hFEN1 the expression vector pET-28b-hFEN1 (Kan^R) was used. For clarity this process has only been described for hFEN1Δ336.

The expression vector pET-29b-hFEN1Δ336 (Kan^R), from which the truncated form of human FEN1 can be produced, was transformed into chemically competent (Method of Inoue) Rosetta(DE3) (F^- ompT $hsdS_B$(rB^- mB^-) gal dcm lacY1(DE3) $pRARE^6$ (Cm^R)) according to protocol. Briefly, approximately 50 ng of the vector was added to a 50 µl aliquot of cells that had been thawed on ice. After incubating on ice for 45 minutes, the cells were heat shocked at 42 ˚C for 90 seconds, and then, placed on ice for 5 minutes. The transformed cells were rescued by addition of 800 µl SOC media and then, incubated at 37 ˚C for one hour. The cells were spread on LB agar plates supplemented with 34 µg/ml chloramphenicol (Cm) and 25 µg/ml kanamycin (Kan). The plates were incubated at 37˚C for 14-16 hours. Transformants, which were designated Rosetta-29b-hFEN1Δ336, were observed on plates as single colonies.

A single colony of Rosetta-29b-hFEN1Δ336 was inoculated in 25 ml of LB media supplemented with 34 µg/ml Cm and 25 µg/ml Kan and allowed to grow for 14-16 hours overnight at 37 ˚C. Four 2L culture flask each containing 500mL TB media supplemented with 34 µg/ml Cm and 25 µg/ml Kan were each inoculated with 5mL of the overnight Rosetta-28b-hFEN1 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.

In the morning, the cells were pelleted by centrifugation (6000 × g, 30 min, 4 ˚C) in two 1L centrifuge bottles, and the supernatant was removed. 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 was removed. The cell pellets were weighed. Two litres of TB media resulted in ~8g of wet cell paste. The two cell pellets (4 g per 50mL Falcon tube) were each suspended in 40mL 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 frozen at -20 ˚C.

To prepare the cell lysate for protein purification, the suspensions were incubated in cold tap water until they were 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.

2.1.3 Protein purification, truncated forms of hFEN1 (hFEN1∆336)

All protein purification steps were conducted in the cold room using an AktaPURE (GE Lifesciences). A Chelating Sepharose Fast Flow (GE Lifesciences) column (1.6cm ID, 10cm 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 using a 50mL superloop (GE Lifesciences) 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, it was applied directly to HiTrap Heparin HP (3X 5mL 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. Protein containing fractions 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 calculate 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 5mL 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 collected in 10mL fractions. Protein containing fractions were pooled and concentrated by ultrafiltration using a 50mL Amicon Ultrafiltration cell with 10,000 MWCO PES membrane (Millipore) pressurized with N_2 (40 PSI). Once the volume of pooled fractions was reduced to less than 10mL, the retentate was applied directly to desalting column (4X 5ml in tandem) using TBSE buffer. The protein was eluted isocratically using 20CV of TBSE buffer and collected as 1.5 mL fractions. The fractions were pooled, and the protein concentration was determined by absorbance at 280 nm using a

Nanoview spectrophotometer and the calculated extinction coefficient (22,920 M^{-1} cm^{-1}). 1 unit of precision protease (PP) was then added for every mg of protein and left overnight at 4˚C. Complete cleavage of the precision protease tag was confirmed by SDS page. Removal of the PP was performed using 200 μl of GST beads, which were repeatedly washed and aspirated in 1X PBS and then applied to the truncated protein. The GST beads were removed from the hFEN1Δ336 solution by filtration through a 0.22 μm syringe filter. For anion exchange chromatography to be performed, the salt concentration of the TBSE buffer (125 mM NaCl) was reduced (<25 mM NaCl) by appropriate dilution of the truncated protein using MilliQ H_2O. The protein was then applied to HiTrap Q HP column (2X 5ml in tandem) using Anion exchange Buffer A and eluted using a 20CV linear gradient using Anion exchange Buffers A and B. Protein containing fractions were pooled and concentrated by ultrafiltration using a 250mL Amicon Ultrafiltration cell with 10,000 MWCO PES membrane (Millipore) pressurized with N_2 (40 PSI). Once the volume of pooled fractions was reduce to less than 20mL, the retentate was equally transferred to two 15 mL Vivaspin (10,000 MWCO) spin columns (4000 x g, 30 min, 4 ˚C) to further reduce the combined volume of the sample to less than 1mL. The retentate from the spin columns was then applied to a Sephacryl S-100 (1.6c m ID x 60 cm) size exclusion column (GE Lifesciences), isocratically eluted using Sephacryl GF buffer and collected in 1.5mL fractions. Protein containing fractions were pooled and concentrated using a Vivaspin (10,000 MWCO) spin columns (4000 x g, 30 min, 4 ˚C).

2.1.4 Protein purification, full-length forms of hFEN1 (hFEN1, hFEN1Y40A, hFEN1K93AR100A, hFEN1R100A, hFEN1K93AR100A, hFEN1D181A, hFEN1L97P, hFEN1L111P, and hFEN1L130P)

All protein purification steps were conducted in the cold room using an AktaFPLC (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 using a 50mL superloop (GE Lifesciences) 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 it was applied directly to HiTrap Heparin HP (3X 5mL 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. Protein containing fractions 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 5mL 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 collected in 10mL fractions. Protein containing fractions were pooled and concentrated by ultrafiltration using a 50mL Amicon Ultrafiltration cell with 10,000 MWCO PES membrane (Millipore) pressurized with N_2 (40 PSI). The retentate from the spin columns was then applied to a Sephacryl S-100 (1.6cm ID x 60cm) size exclusion column (GE Lifesciences) and isocratically eluted using Sephacryl GF buffer and collected in 1.5mL fractions. Protein containing fractions were pooled and concentrated using a Vivaspin (10,000 MWCO) spin columns (4000 x g, 30 min, 4 ˚C).

2.1.5 Determination of protein concentration

Protein concentration was determined by absorbance at 280 nm using a Nanoview spectrophotometer and the calculated extinction coefficient (hFEN1: 22,920 M^{-1} cm^{-1}). The volume of the sample was adjusted with the appropriate volume of glycerol, 1 M dithiothreitol (DTT), and Sephacryl GF buffer supplemented to correct the protein concentration to 100μM, the glycerol concentration to 50% (v/v), the DTT concentration to 10 mM, and the Sephacryl GF buffer components to 50 mM HEPES pH=7.5, 100 mM KCl, 0.02% (w/v) NaN_3 (i.e., storage buffer). The concentration of enzyme in storage solution (~4mg/mL) was confirmed using the Bradford microtiter plate assay using bovine serum albumin (BSA) as a standard curve according to the manufacturer's protocol, and the protein purity was assessed by SDS PAGE (*figure 2.1*).
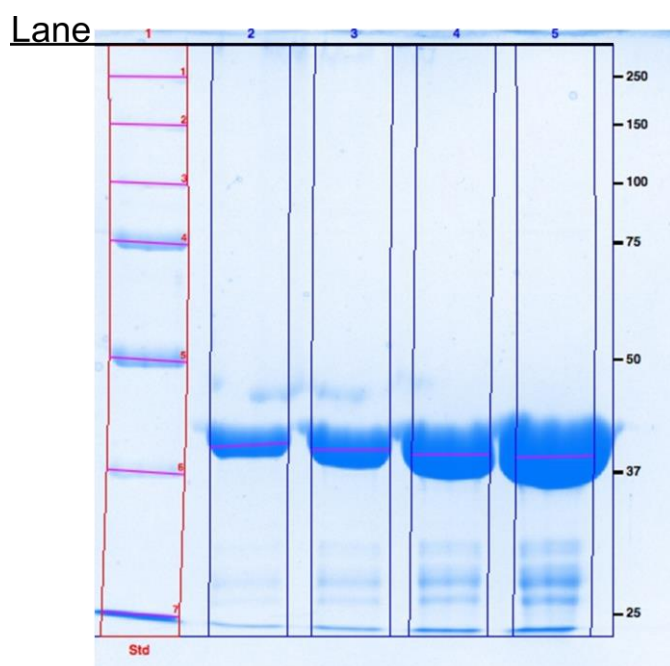
Figure 2.1: SDS-PAGE to show purity of WThFEN1 after purification steps. Lane 1: protein ladder (bands 1, 2, 3, 4, 5, 6 and 7 representing 250, 150, 100, 75, 50, 37 and 25 kDa markers), lane 2: 1 µg/ µL, lane 3: 2 µg/ µL, lane 4: 4 µg/ µL, and lane 5: 8 µg/ µL.

2.2 Synthesis and purification of substrates for hFEN1

All oligonucleotide substrates were ordered purified using reverse-phase HPLC and synthesised by DNA Technology (Risskov, Denmark) using standard phosphoramidites chemistry. Oligonucleotides used in either kinetic or FA, FRET, EMSA, and ECCD assays are shown in *table 2.1*, *table 2.2*, *table 2.3*, and *table 2.4* respectively.

Fluorescent oligonucleotides used to study hFEN1 were labelled either on the 5ʹ-terminus or internally using 5ʹ-fluorescein-CE-phosphoramidite (*figure 2.2a*), 5ʹ-tetramethylrhodamine-CE-phosphoramidite (*figure 2.2b*), C6-dT-fluorescein-phosphoramidite (*figure 2.3a*), or C6-dT-tetramethylrhodamine -phosphoramidite (*figure 2.3b*). Oligonucleotides biotinylated on the 5'-terminus of the 5'-flap were synthesised using the biotin-precursor shown in *figure 2.4*. Oligonucleotides containing 2-aminopurine substitution were synthesised using 2-aminopurine-CE phosphoramidite (*figure 2.5*).
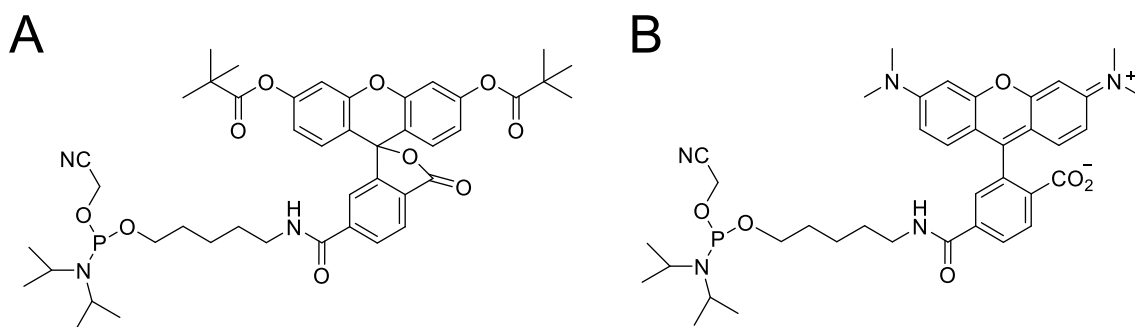
Figure 2.2: Fluorophore precursors used to synthesis 5'-terminally labelled (A) 5'-fluorescein and (B) 5'-tetramethylrhodamine oligonucleotides.
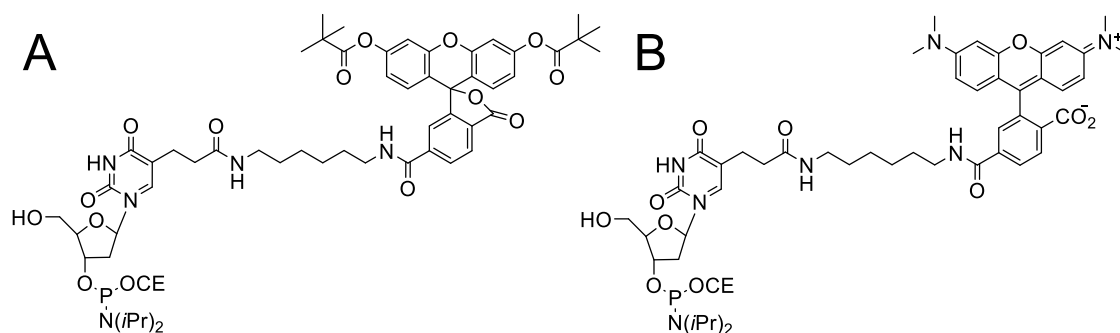
Figure 2.3: Fluorophore precursors used to synthesis internally labelled (A) 5'-fluorescein and (B) 5'-tetramethylrhodamine oligonucleotides.
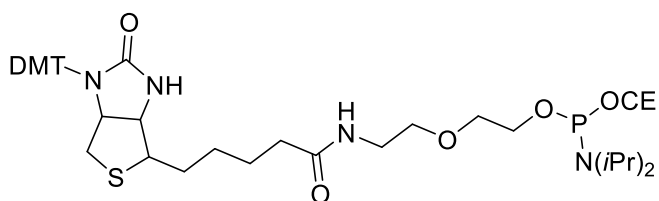
Figure 2.4: Biotin-linker-phosphoramidite used in oligonucleotide synthesis
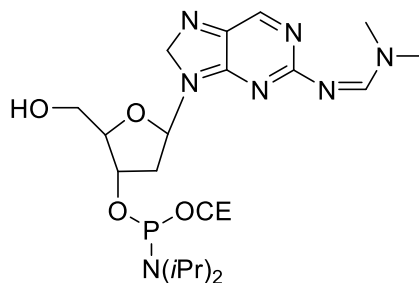
Figure 2.5: 2-aminopurine-CE phosphoramidite used in oligonucleotide synthesis

Stability prediction software (MFold, Zuker 2003) was used to predict the overall folded structure of all bimolecular and trimolecular constructs used. Sequences of the individual strands are shown in *table 2.5*. Purified oligonucleotides were desalted using NAP-10 columns and subject to MALDI-TOF or TOF-ESI mass spectrometry. Residual divalent metal ion contaminants were removed by treatment with Chelex resin. Experimental MWs were all within 5 Da of calculated (*table 2.6*). The concentration of the purified oligonucleotides were determined by measuring the absorbance on the Nanoview spectrophotometer until it entered the linear range and using Beer-Lambert law with calculated extinction coefficients. The ε_{260} value of the individual oligonucleotides (*table 2.6*) was calculated on the 'Integrated DNA Technologies®' website (http://eu.idtdna.com/Analyzer/Applications/Instructions/Default.aspx?AnalyzerDefinitions=true#ExtinctionCoefficient) from the following equation (equation 2.1):

$$\varepsilon_{260} = 2 \times \left(\sum_{1}^{N-1} \varepsilon_{Nearest\cdot Neighbor} \right) - \sum_{2}^{N-1} \varepsilon_{Individual} + \sum_{1}^{N} \varepsilon_{Modification} \quad equation\ 2.1$$

where $\varepsilon_{Nearest\cdot Neighbor}$ is the nearest neighbour coefficient for a pair of bases, $\varepsilon_{Individual}$ is the coefficient for an individual base, and N is the length of the oligonucleotide (Cantor,C.R. et al.,*Biopolymers*, **9**, 1059-1077., Cavaluzzi,M.J. and Borer,P.N. *Nucleic Acids Res*., **32**, e13).

Table 2.1: Flap substrate structure, sequence and modification positions used in hFEN1 kinetic and FA assays. [a]FAM coupled through C6 linker to the terminal phosphate. Melting temperatures (*Tm*) are for the heteroduplex portion of the substrate and are calculated (M. Zucker, 2003) for DNA at a final concentration of 2.5 nM at 100 mM KCl. Sequences of the individual strands are shown in table 2.5.

Modification	X	Y	Z	Individual oligonucleotides used to form construct	Tm (°C)
3' C T^TTTT-Z 5' 3' GAG GTG TCG GGT CCA GCG CTG CCA A • A 5' CTC CAC AGC CCA GGT CGC GAC GGT G					
DF-5	-	-	5'FAM[a]	F1 + T1	56.2
3'C T^TTTT-Z 5' 3'CAC AGC TCG TCA GGA ACA CTGCTG A • A 5'GTG TCG AGC AGT CCT TGT GACGAC G					
DF-5*	-	-	5'FAM[a]	F1* + T1*	51.5
UDF-5*	-	-	-	UF1* + T1*	

Table 2.2: Flap substrate structure, sequence and modification positions used in hFEN1 FRET assays. [a]Coupled to the terminal phosphate. Melting temperatures (*Tm*) are for the heteroduplex portion of the substrate and are are calculated (M. Zucker, 2003) for DNA at a final concentration of 10 nM, at 100 mM KCl. [b]Tm for 5'-flap (5′F) and template strand. [c]Tm for 3'-flap (3′F) and template strand. [d]Tm for 3'-flap (5′F3) and template strand. [e]Tm for 3'-flap (3′F2) and template strand. Sequences of the individual strands are shown in table 2.5.

Modification	X	Y	Z	Individual oligonucleotides used to form construct	Tm (°C)
3'G T^TTT-Z 5' 3' AC CTG CCC ACC GCA ATT CCA ATC CGA-X 5' 5' Y-TG GAC GGG TGG CGT TAA GGT TAG GCT 3'					
DF-4 (unlabelled)	-	-	-	3′F + 5′F + T3	
DF-4 (donor-only-labelled)	5'FAM[a]	-	-	3′Fd + 5′F + T3	38.2[b] and 42.4[c]
DF-4 (acceptor-only-labelled)	-	5'TAMRA[a]	-	3′F + 5′F + T3a	

Name	Donor	Acceptor	Biotin	Composition	Tm
DF-4 (donor-acceptor-labelled)	5'FAM[a]	5'TAMRA[a]	-	3'Fd + 5'F + T3a	
3'G T^{TTTTTTTTT}-Z 5' 3' AC CTG CCC ACC GCA ATT CCA ATC CGA-X 5' 5' Y-TG GAC GGG TGG CGT TAA GGT TAG GCT 3'					
DF-B4 (unlabelled)	-	-	5'Biotin[a]	3'F + 5'F2 + T3	
DF-B4 (donor-only-labelled)	5'FAM[a]	-	5'Biotin[a]	3'Fd + 5'F2 + T3	
DF-B4 (acceptor-only-labelled)	-	5'TAMRA[a]	5'Biotin[a]	3'F + 5'F2 + T3a	38.2[b] and 42.4[c]
DF-B4 (donor-acceptor-labelled)	5'FAM[a]	5'TAMRA[a]	5'Biotin[a]	3'Fd + 5'F2 + T3a	
3 C T^{TTTTT}-Z 5' 3' CCA GGA TGA GAC GGA GTT CTG CCA GAC GAC GXG ACC 5' 5' GGT CCY ACT CTG CCT CAA GAC GGT CTG CTG CAC TGG 3'					
DF-B6 (unlabelled)	-	-	5'Biotin[a]	3'F2 + 5'F3 + T4	
DF-B6 (donor-only-labelled)	C6-dT-FAM	-	5'Biotin[a]	3'F2d + 5'F3 + T4	
DF-B6 (acceptor-only-labelled)	-	C6-dT-TAMRA	5'Biotin[a]	3'F2 + 5'F3 + T4a	50.0[d] and 56.4[e]
DF-B6 (donor-acceptor-labelled)	C6-dT-FAM	C6-dT-TAMRA	5'Biotin[a]	3'F2 + 5'F3 + T4a	

Table 2.3: Flap substrate structure, sequence and modification positions used in hFEN1 EMSA assays.
[a]Coupled to the terminal phosphate. Melting temperatures (Tm) are for the heteroduplex portion of the substrate and are calculated (M. Zucker, 2003) for DNA at a final concentration of 1 nM, at 100 mM KCl. Sequences of the individual strands are shown in table 2.5.

Modification	X	Y	Z	Individual oligonucleotides used to form construct	Tm (°C)
3'C ᴛTTTTTTTTT-Z 5' 3' CT ACT GCT CGT CAG GAT TGA CCT TTA GAT CGA GAC ACC TG A ◆ A 5' GA TGA CGA GCA GTC CTA ACT GGA AAT CTA GCT CTG TGG AC G					
DF-B10	-	-	5'Biotin[a]	F2 + T2	50.1
DF-10	-	-	-	F3 + T2	

Table 2.4: Flap substrate structure, sequence and modification positions used in hFEN1 ECCD assays. T_m stated are calculated (M. Zucker, 2003) for DNA at a final concentration of 10 µM, at 100 mM KCl. Sequences of the individual strands are shown in table 2.5.

Construct	Oligonucleotide composition	Tm (°C)
P$_{-1-2}$	CF$_{-1-2}$ + T5	48.6
HO-P$_{-1-2}$	HO-CF$_{-1-2}$ + T5	48.6
S$_{-1-2}$	F$_{-1-2}$ + T5	49.6
Sb$_{-1-2}$	Fb$_{-1-2}$ + T5	49.6
MM$_{+1}$S$_{-1-2}$	MM$_{+1}$F$_{-1-2}$ + T5	48.9
S$_{+1-1}$	F$_{+1-1}$ + T6	52.5
S$_{-8-9}$	F$_{-8-9}$ + T7	53.6

Oligonucleotide Code	Sequence
Table 2.5: Sequences of oligonucleotides used in this work. *FAM = fluorescein, TAMRA = tetramethylrhodamine, B = 5'-biotin moiety, 5PHOS = 5´-phosphate, HO = 5´-hydroxyl, and 2AP = 2-aminopurine. Calculated and experimental masses, along with calculated extinction coefficient values shown in table 2.6.*	
F1	5'-FAM TTT TTG CGA CCT GGG CTG TGG AG-3'
T1	5'-CTC CAC AGC CCA GGT CGC GAC GGT GAA ACC GTC C-3'
F1*	5´-FAM-TTT TTA CAA GGA CTG CTC GAC AC-3´
T1*	5´-GTG TCG AGC CCT TGT GAC GAC GAA GTC GTC C-3´
UF1*	5´-TTT TTA CAA GGA CTG CTC GAC AC-3´
F2	5'-B-TTT TTT TTT TAG TTA GGA CTG CTC GTC ATC-3'
T2	5'-GAT GAC GAG CAG TCC TAA CTG GAA ATC TAG CTC TGT GGA CGA AGT CCA CAG AGC TAG ATT TCC C-3'
F3	5'- TTT TTT TTT TAG TTA GGA CTC GTC ATC-3'
3´F	5'-AGC CTA ACC TTA ACG G -3'
3´Fd	5'-FAM-AGC CTA ACC TTA ACG G -3'
5´F	5'-TTT TCC ACC CGT CCA -3'
5´F2	5'-B-TTT TTT TTT TCC ACC CGT CCA -3'
T3	5'-TGG ACG GGT GGC GTT AAG GTT AGG CT -3'
T3a	5'-TAMRA-TGG ACG GGT GGC GTT AAG GTT AGG CT -3'
3´F2	5'-CCA GTG CAG CAG ACC GTC C -3'
3´F2d	5'-CCA G/FAM-dT/G CAG CAG ACC GTC C -3'
5´F3	5'-B-TTT TTT TTG AGG CAG AGT AGG ACC -3'
T4	5'-GGT CCT ACT CTG CCT CAA GAC GGT CTG CTG CAC TGG-3'
T4a	5'-GGT CC/TAMRA-dT/ ACT CTG CCT CAA GAC GGT CTG CTG CAC TGG-3'
CF₋₁₋₂	5'-/5PHOS//2AP//2AP/A GGC AGA GTG -3'
HO-CF₋₁₋₂	5'-HO-/2AP//2AP/A GGC AGA GTG -3'
F₋₁₋₂	5'-TTT TTG /2AP//2AP/A GGC AGA GTG-3'
Fb₋₁₋₂	5'-B-TTT TTT TTT TG/2AP/ /2AP/AG GCA GAG TG-3'

MM$_{+1}$F$_{-1-2}$	5'-TTT TTC /2AP//2AP/A GGC AGA GTG -3'
T5	5'-CAC TCT GCC TTT CGA CAG CGA AGC TGT CC -3'
F$_{+1-1}$	5'-TTT TT/2AP/ /2AP/GA GGC AGA GTG -3'
T6	5'-CAC TCT GCC TCT TGA CAG CGA AGC TGT CC -3'
F$_{-8-9}$	5'-TTT TTA AGA GGC G/2AP//2AP/ GTG -3'
T7	5'-CAC TTC GCC TCT TGA CAG CGA AGC TGT CC -3'

Table 2.6: Details of all the oligonucleotides synthesised and purified for work with hFEN1.			
Oligonucleotide Code	Calculated Mass (gmol^{-1})	Actual Mass (gmol^{-1})	ε_{260} (Lmol^{-1}cm^{-1})
F1	7651.64	7651	230460
T1	10358.73	10353	313700
F1*	7536.1	7536	236760
T1*	10474.8	10475	323200
UF1*	6988.6	6999	215800
F2	9719.6	9720	271200
T2	19728.9	19730	620200
F3	9150.0	9150	271200
3ˈF	4850.2	4850	154900
3ˈFd	5387.7	5388	175860
5ˈF	4438.9	4439	129300
5ˈF2	6833.7	6833	183500
T3	8138.3	8138	252900
T3a	8712.3	8714	283200
3ˈF2	5758.8	5759	176300
3ˈF2d	6271.3	6271	190000
5ˈF3	7819.3	7807	231500
T4	11004.1	11004	322900
T4a	11570.8	11571	352000
CF$_{-1-2}$	3823.5	3845	108700
HO-CF$_{-1-2}$	3743.5	3743	108700

F_-1-2	5593.7	5593	158400
Fb_-1-2	7684.3	7684	198900
MM_+1F_-1-2	5553.7	5554	155700
T5	8789.7	8788	258100
F_+1-1	5593.7	5594	158400
T6	8789.7	8788	257500
F_-8-9	5593.7	5593	158400
T7	8789.7	8788	258100

2.3 Characterising DNA binding affinities of FEN using fluorescence anisotropy

Three compounds provided by AstraZeneca were denoted inhibitor **1**, **2**, and **3** (*figure 2.6*). The compounds were provided as a solid and dissolved in diethyl sulfoxide (DMSO) to make 100 mM stocks and stored at -20 °C. All dilutions required were made from these stocks. The appropriate dilutions were made so that the final buffer composition contained 1 % DMSO (v/v) to aid solubility.
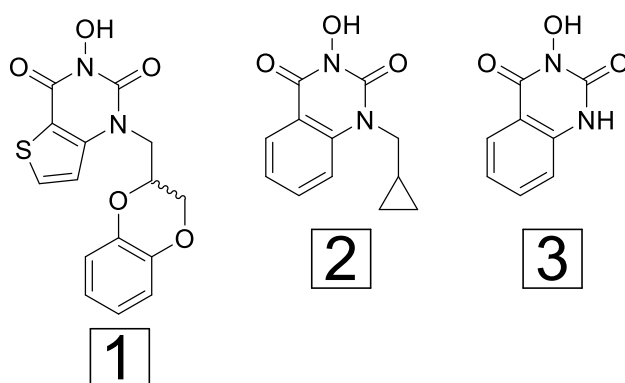
Figure 2.6: Structure of inhibitors used in this study.

2.3.1 Measuring hFEN1-substrate binding by fluorescence anisotropy

Dissociation constants for free enzyme and the enzyme:inhibitor complex for the DNA substrate (DF-5/DF-5*) under equilibrium conditions were measured by fluorescence anisotropy using a Horiba Jobin Yvon FluoroMax-3® fluorometer with automatic polarizers. The program Fluoresence (Horiba Scientific, Middlesex, UK) was used to control the spectrometer. The excitation wavelength was 490 nm (excitiation slit width 5 nm) with emission detected at 510 nm (emission slit width 5 nm).

Samples contained 10 mM $CaCl_2$ or 8 mM $MgCl_2$ or 2 mM EDTA and 10 nM DF-5/DF-5*, 110 mM KCl, 55 mM HEPES-NaOH, pH 7.5, 0.1 mg/ml BSA, 1mM DTT and 100 μM inhibitor **1/2/3** or 1 % DMSO. Reaction mixtures were placed in a 0.5 mL quartz cuvette and kept at a constant temperature of 37 ˚C. The substrate was incubated at 37 ˚C for a minimum of ten minutes before the first measurement at 0 nM protein with subsequent readings taken on the cumulative addition of enzyme in a matched buffer, with corrections made for dilution.

Prior to each experiment, DF-5/DF-5* stock solution was prepared in 100 mM KCl, 50 mM HEPES-NaOH pH=7.5, and made to a final concentration of 50 μM and used for all the experimental repeats (containing a ratio of 1:1.1 of flap: template strands). The substrate was pre-folded at 95 ˚C for 2 minutes, and allowed to cool to 37 ˚C before addition to the reaction mixture. Before each condition tested, 1 μM and 10 μM enzyme solutions were made and kept on ice until required. Enzyme dilutions were performed using a buffer composition identical to that of the initial reaction mixture (including inhibitors where appropriate), so that, the buffer concentration remained constant throughout the titration.

 Six measurements of $I_{VH}, I_{VV}, I_{HV},$ and I_{HH} were recorded over 1 minute with the first measurement being taken prior to the addition of enzyme. Subsequent readings were taken on the cumulative addition of the appropriate enzyme. Calculated molar equivalents of enzyme were added in 0.5-20 μl quantities (0.1 – 3000 nM of appropriate enzyme). The measurements were used to calculate anisotropy (*r*) to *equation 2.2* below:

$$r = \frac{I_{VV} - GI_{VH}}{I_{VV} - 2GI_{VH}} \qquad equation\ 2.2$$

, where *r* is anisotropy, I_{VV} and I_{VH} are the intensities of the vertical and horizontal components of the emitted light using vertically polarised excitation. G is a correction factor for the difference in response of the detector to vertical and horizontal polarized light, and it is given by $G = I_{HV}/I_{HH}$, where I_{HV} and I_{HH} are the intensities of the vertical and horizontal components of the light emitted using horizontally polarised excitation. In addition to measuring the signal through polarisers, the total fluorescence, I_{TOT} (without polarisers) was measured after each addition.

Data were modelled by nonlinear least squares regression in KaleidaGraph 4.0 (Kaleidagraph, Synergy Software, Reading, PA, USA) using *equation 2.3*,

$$r = r_{min} + (r_{max} - r_{min}) \frac{\left[([S] + [E] + K_D - \sqrt{([S] + [E] + K_D)^2 - 4[S][E]} \right]}{2[S]} \quad equatopn \ 2.3$$

, where *r* is the measured anisotropy at a particular total concentration of enzyme ([E]) and fluorescent substrate ([S]), with r_{min} giving the minimum anisotropy, (the anisotropy of free oligonucleotide in the absence of any enzyme) and r_{max} the maximum anisotropy, (the anisotropy of the saturated substrate). The dissociation constant is given by K_D. Each measurement was independently repeated in triplicate, and samples were taken after completion of the titration and analysed by dHPLC to determine the amount of product produced.

2.3.2 Competition Experiments

Samples were prepared and anisotropy readings taken as described for the protein-DNA equilibrium binding measurements above. 10 nM DF-5 in buffer containing 110 mM KCl, 55 mM HEPES-NaOH pH=7.5, 8 mM MgCl2, 0.1 mg/ml BSA, 1 mM DTT, 100 μM inhibitor **1** was incubated at 37 ˚C for a minimum of ten minutes, then hFEN1R100A was added in a cumulatively up to ~ 80 % saturation of the substrate (DF-5*). At this point unlabelled DNA (UDF-5*) was added in a stepwise manor with readings taken after each addition of the competitor until the anisotropy value reached that of oligonucleotide in the absence of any protein.

2.4 Measurement of single turnover rates of FEN-catalysed reactions

2.4.1 hFEN1, hFEN1Δ336, and hFEN1Y40A

Rapid quench experiments were carried out at 37 ˚C using a RQF-63 quench flow device (Hi-Tech Sci Ltd, Sailsbury, UK). An 80 μl aliquot of enzyme in <u>r</u>educed <u>r</u>eaction <u>b</u>uffer (<u>RRB</u>, 8 mM MgCl₂, 110 mM KCl, 55 mM HEPES-NaOH pH=7.5, 0.1 mg/ml BSA and 1 mM DTT) was mixed with an equal volume of oligonucleotide substrate (annealed as described in *section 2.3.1*) in the same buffer conditions. Enzyme was used at a final concentration of $10 \times K_M$, and FAM-labelled oligonucleotide substrates were used at a final concentration of $1/40 \times K_M$, down to a minimum of 2.5 nM to avoid signal detection limits (*table 2.7*). The single-turnover reaction of hFEN1Y40A was conducted at a final enzyme concentration of $10 \times K_D$, because the steady-state catalytic parameters of Y40A had yet to be determined. WThFEN1 was also repeated at its respective dissociation constant value (*table 2.7*). Coincidentally, the K_D and K_M, values were approximately equal for hFEN1Δ336. Solutions of substrate, enzyme and quench were held in a temperature controlled water bath set to 37 ˚C within the instrument during the reactions. The following time-points were taken from the initiation of the reaction: 4.5ms, 9ms, 12ms, 19ms,

31ms, 42ms, 59ms, 84ms, 126ms, 142ms, 0.24s, 0.44s, 0.84s, 1.64s, 3.2s, 6.4s, 12.8s, 25s, and 52s.

The reaction was stopped by addition of 80 µl of quench (1.5 M NaOH and 80 mM EDTA). The quench reaction mixtures were recovered and then analysed using dHPLC (Wave® fragment analysis dHPLC fitted with a fluorescence detector; Transgenomic, Glasgow). A gradient (*figure 2.7*) made using WAVE buffers A and B (*table 2.8*) was used to separate product from starting material on a DNAsep® (Transgenomic, Glasgow) column using the FAM fluorescent tag for detection (excitation wavelength 494 nm, emission, 525 nm). All reactions were independently repeated three times. The first-order rate of reaction was determined by plotting the appearance of product against time and by non-linear regression fitting to *equation 2.4* below:

$$Q_t = Q_\infty \left(1 - exp^{-k_{ST}t}\right) \qquad\qquad equation\ 2.4$$

where Q_t is the amount of product at time t, Q_∞ is the amount of product at time ∞, and k_{ST} is the single turnover rate of reaction. The percentage product was constrained to zero at t = 0 s.

Table 2.7: Final enzyme and substrate concentrations *used to determine the single-turnover rate (k_{ST}) for WThFEN1, hFEN1Δ336 and hFEN1-variants catalysed reactions of synthetic substrates (section 2.2) in the absence and presence of inhibitors. [a]Performed with inhibitor 1 and 2. [b]Performed with inhibitor 1. [c]Performed with inhibitor 1, 2, and 3. Inhibitor structures are shown in figure 2.6.*

Enzyme	Construct name	Substrate (nM)	Enzyme (nM)	Inhibitor 1/2/3 (μM)
hFEN1	DF-5	2.5	1000	-
	DF-5*	2.5	1000	-
	DF-5*	2.5	200	-
	DF-5*	2.5	1000	100[a]
hFEN1Δ336	DF-5*	3.78	1500	-
hFEN1Y40A	DF-5	5	1000	-
hFEN1K93A	DF-5	5	1000	-
hFEN1K93A	DF-5	5	1000	100[b]
hFEN1K93AR100A	DF-5	5	1000	-
hFEN1R100A	DF-5	5	1000	-
hFEN1R100A	DF-5*	5	1000	-
hFEN1R100A	DF-5*	5	1000	100[c]
hFEN1D181A	DF-5*	5	1000	-

Table 2.8: WAVE buffers used to generate oligonucleotide separation gradient on dHPLC

Buffer A	0.1% Acetonitrile, 1 mM EDTA,, 2.5 mM tetrabutyl ammonium bromide – dissolved in water
Buffer B	85% Acetonitrile, 1 mM EDTA,, 2.5 mM tetrabutyl ammonium bromide – dissolved in water
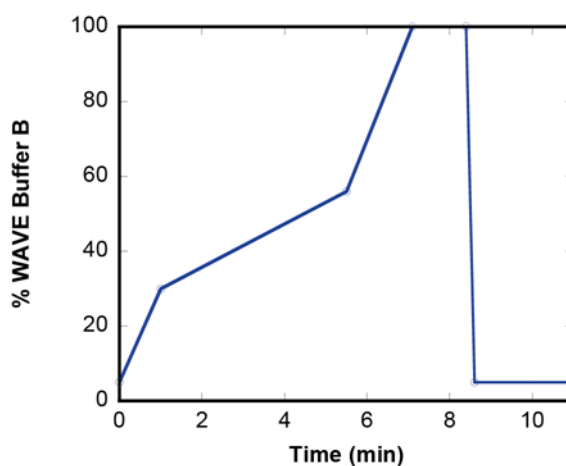
Figure 2.7: Denaturing HPLC gradient used to separate the reaction products from substrate.
The specific gradient switch points are as follows: 0min 5%B 95%A, 1min 30%B 70% A, 5.5min 56%B 44% A, 7.1min 100%B, 8.4min 100%B, 8.6min 5%B 95%A, 11min 5%B 95%B.

2.4.2 hFEN1-variants

Due to the slow reaction experiments involving hFEN1-variants (K93A, R100A, K93AR100A, and D181A), the maximal single turnover rates of reaction for these were performed on the bench-top. The single-turnover rate for the hydrolysis of DF-5/DF-5* was measured by mixing enzyme and substrate diluted in RRB, pre-incubated at 37 ˚C in a 1:1 volume ratio, followed by quenching of 100 µl reaction samples in 100 µl quench (1.5 M NaOH and 80 mM EDTA). The steady-state catalytic parameters of any of the variants have yet to be determined with either substrate. Therefore, the final enzyme and substrate concentration was set to 1 µM and 5 nM, respectively (*table 2.7*). This was to ensure that the maximal single turnover rate was being measured. The following time-points were taken from the initiation of the reaction: 0.5min, 1min, 1.5min, 2min, 2.5min, 3min, 4min, 5min 6min, 8min, 10min, 12min, 15min, 20min, 25min, 30min, 45min, 60min. As a consequence of the R100A, K93AR100A, and D181A mutations being more detrimental to reaction rate than the K93A mutation, extra time points had to be taken: 90min, 120min, 180min, 300min, 360min.

2.4.3 Measuring the decay of enzyme-substrate-inhibitor complexes

For the single-turnover rates of DF-5* hydrolysis in the presence of inhibitors, reactions were performed analogously to that described above, with substrate and enzyme dilutions performed in 8 mM MgCl₂, 110 mM KCl, 55 mM HEPES-NaOH pH=7.5, 0.1 mg/ml BSA 1 mM DTT, and 100 µM of the appropriate inhibitor (*table 2.7*).

Specifically, to sample the entirety of the for WThFEN1 reactions with inhibitors **1** and **2** , time-course reactions had to be performed in part on the quench-flow apparatus and in part on the bench top. On the quench flow, the following time-points were taken from the initiation of the reaction: 4.5ms, 9ms, 12ms, 19ms, 31ms,42ms, 57ms, 82ms, 124ms, 140ms , 440ms, 1.6s, 3.2s, 6.4s, 12.8s. On the bench top, the following time-points were taken from the initiation of the reaction: 0.5min, 1min, 2min, 3min, 5min, 8min, 10min, 20min, 30min, 60min, 120min, and 180min

As a consequence of the slower reaction rates for K93A and R100A with inhibitors (*table 2.7*), reactions were performed exclusively on the bench top. The following time-points were taken from the initiation of the reaction: 0.5min, 1min, 1.5min, 2min, 2.5min, 3min, 4min, 5min 6min, 8min, 10min, 12min, 15min, 20min, 25min, 30min, 45min, 60min.

2.4.5 Measuring the single turnover rate using fluorescence anisotropy

The single-turnover rate of DF-5 hydrolysis of hFEN1R100A was measured by monitoring the anisotropy change from release of the ss-5'-flap (for which hFEN1 has no affinity) from a saturated ds substrate. The substrate (annealed as described in *section 2.3.1*) was used at a final concentration of 10 nM and incubated at 37 ˚C in RRB. Reaction was initiated by addition of the enzyme to a final concentration of 1000 nM. Enzyme dilutions were made from a 100 µM stock (in storage buffer, *section 1.2*), in RRB. Anisotropy values were measured, as described in *section 2.3.1*, and data points were collected every 12.5 s for a total of 200 minutes. The data were fit to a single-exponential decay to determine the rate constant.

2.5 Electrophoretic Mobility Shift Assay (EMSA)

2.5.1 Determination of the hFEN1 binding constant by EMSA

The assay is based on that DNA-protein complexes migrate slower than double-stranded oligonucleotides in a native polyacrylamide gel, resulting in a "shift" in migration of the labelled DNA band. The detection of bands was by chemiluminescent nucleic acid detection module kit (Thermo Scientific) that uses a non-isotopic method to detect DNA-protein interactions. 5'-Biotin-labelled DNA substrate (DF-B10, *table 2.3*) at 1 nM or 10 nM was incubated at 37 ˚C for 10 minutes with 0.5-2500 nM hFEN1 in buffer containing 10 mM $CaCl_2$, 110 mM KCl, 55 mM HEPES-NaOH, pH 7.5, 0.1 mg/ml BSA, and 1mM DTT.

After the equilibration period, the DNA-protein complexes were electrophoresed on 6 % polyacrylamide gels (29:1) containing 0.5 × Tris-borateEDTA (TBE). Moreover, the gels were loaded while running at 200 V at RT and then further electrophoresed at 200 V until the

bromphenol blue indicator in a separate well reached the bottom of the gel. The gels were then electroblotted onto Biodyne B nylon membranes (Thermo Scientific) at 4 °C in 0.5×Tris-borate-EDTA at 380 mA for 30 min using a Mini Trans-Blot®module (Bio-Rad). The membranes were removed from the blot sandwich and were exposed face-down to a transilluminator equipped with 312-nm bulbs for 10 minutes to cross-link all complexes. The blots were then incubated in nucleic acid blocking buffer (ThermoScientific) for 30 minutes, and subsequently, incubated in 1:300 dilution of stabilized streptavidin-horseradish peroxidase (SAHRP) in nucleic acid blocking buffer for 30 minutes. The blots were washed five times in 20 mL of 1×wash buffer and once in 20 mL of substrate equilibration buffer (Thermo Scientific). Enhanced chemiluminescence (ECL) detection was accomplished using chemiluminescence substrate (Thermo Scientific) and a ChemiDoc XRS system (Bio-Rad). The percentage of unbound and bound biotinylayted substrate was measured using ImageLab 5.0 software. The dissociation constant was determined by modelling of the data using nonlinear least squares regression in GraphPad Prism (version 6.00 for Windows, GraphPad Software, San Diego California USA) to *equation 2.5*,

$$\% \, Bound = \frac{(\% \, Bound_{max} \cdot [protein])}{(K_D + [protein])} \qquad equation \; 2.5$$

2.5.2 EMSA Streptavidin Blocking/Trapping Experiments

All FEN1-substrate binding reactions were prepared with 1 nM of the biotinylated substrate DF-B10 in 50 mM HEPES-KOH, pH 7.5, 100 mM KCl, 8 mMCaCl$_2$, 0.25 mg/ml BSA, 2 mM DTT, and 5 % glycerol. Where appropriate, the indicated protein (50 nM final concentration), SA (Roche Diagnostics; 0.2 μg), and/or the nonbiotinylated competitor DNA (DF-10; 250 nM final concentration) were added in specific orders to create blocked or trapped conditions. Blocked complexes refer to binding reactions that were prepared by incubating the substrate with SA (5 molar excess) for 10 min at RT prior to the addition of the indicated hFEN1 protein. Trapped complexes refer to binding reactions that were prepared by incubating the substrate with the indicated hFEN1 protein for 10 min at RT before the addition of SA (5 molar excess) to binding reactions. After the addition of the second component in both blocked and trapped binding reactions, the reactions were incubated for a further 10 min. For reactions where FEN1-substrate binding was challenged with DNA substrate lacking biotin, DF-10 was added after the second incubation (5-fold excess with respect to enzyme concentration), and the mixture was allowed to incubate for 10 min at 37 °C. Samples were then loaded onto a 6 % polyacrylamide gels (29:1) containing 0.5 × TBE while running at 200 V, and the resultant species were visualised analogously to that described above.

2.6 Fluorescence resonance energy transfer

2.6.1 Characterising DNA binding affinities of FEN using an ensemble FRET assay

FRET energy transfer efficiencies (FRET$_{eff}$) were determined using the (ratio)$_A$ method (Clegg, 1992), by measuring the enhanced acceptor fluorescence. The steady state fluorescent spectra of 10 nM doubly-labelled, donor-only labelled and non-labelled trimolecular DNA substrates were recorded using a Horiba Jobin Yvon FluoroMax-3® fluorometer. For direct excitation of the donor (fluorescein) or acceptor (rhodamine), the sample was excited at 490 nm or 560 nm (2 nm slit width) and the emission signal collected form 515 – 650 nm or 575 – 650 nm (5 nm slit width). The emission spectra were corrected for buffer and enzyme background signal by subtracting the signal form the non-labelled DNA sample. In total, the emission spectra of three separate samples were collected for each FRET$_{eff}$ calculation. As stated, non-labelled and donor-only- and donor-acceptor-labelled DNA constructs were consecutively titrated with protein at the indicated concentrations. Reaction mixtures were placed in a 0.5 mL quartz cuvette and kept at a constant 37 ˚C. The substrate was incubated at 37 ˚C for a minimum of ten minutes before the first measurement. In addition to 10 nM of the appropriate DNA construct, samples contained 10 mM $CaCl_2$ or 8 mM $MgCl_2$ or 2 mM EDTA and 110 mM KCl, 55 mM HEPES-NaOH pH 7.5, 0.1 mg/ml BSA, 1mM DTT and 100 µM inhibitor **1/2** or 1 % DMSO. The first measurement was taken prior to the addition of protein with subsequent readings taken on the cumulative addition of enzyme, with corrections made for dilution.

The (ratio)$_A$ value at each protein concentration was calculated using the following equation (*equation 2.6*),

$$(ratio)_A = \left(F_{DA}(\lambda^D_{EX}, \lambda^A_{EM}) - N \cdot F_D(\lambda^D_{EX}, \lambda^A_{EM}) \right) / F_{DA}(\lambda^A_{EX}, \lambda^A_{EM}) \quad equation\ 2.6$$

where F$_{DA}$ and F$_D$ is the fluorescent signal of the doubly-labelled DNA and donor-only labelled DNA at the given wavelengths, respectively. The fluorescence intensity was an average, taken 5 nm either side of the emission maximum wavelength. The value of *N* represents a donor emission normalization factor and can be calculated using *equation 2.7*, which is the donor emission when excited at the donor excitation wavelength of the donor-acceptor-labelled sample divided by the donor-only-labelled sample.

$$N = F_{DA}(\lambda^D_{EX}, \lambda^D_{EM}) / F_D(\lambda^D_{EX}, \lambda^D_{EM}) \quad equation\ 2.7$$

The energy transfer efficiencies (FRET$_{eff}$) can then be calculated using the (ratio)$_A$ value and experimentally determined molar absorption coefficients from *equation 2.8*.

$$FRET_{eff} = (ratio)_A / \left(\frac{\varepsilon^D(490)}{\varepsilon^A(560)} \right) - \left(\frac{\varepsilon^A(490)}{\varepsilon^A(560)} \right) \qquad equation\ 2.8$$

ε^D and ε^A are the molar absorption coefficients of donor and acceptor at the given wavelengths. $\varepsilon^D(490)/\varepsilon^A(560)$ and $\varepsilon^A(490)/\varepsilon^A(560)$ are determined from the absorbance spectra of doubly - labelled molecules and the excitation spectra of rhodamine-only-labelled molecules, respectively (*table 2.9*).

Table 2.9: Molar absorption coefficients ratios calculated from the absorption spectra of the different oligonucleotides used in the FRET studies.

Substrate	DF-4	DF-B4	DF-B6
$\left(\dfrac{\varepsilon^D(490)}{\varepsilon^A(560)} \right)$	0.55 (± 0.01)	0.52 (± 0.01)	0.79 (± 0.01)
$\left(\dfrac{\varepsilon^A(490)}{\varepsilon^A(560)} \right)$	0.08 (±0.01)	0.07 (± 0.01)	0.06 (± 0.01)

Data were modelled by nonlinear least squares regression in KaleidaGraph 4.0 (Kaleidagraph, Synergy Software, Reading, PA, USA) using *equation 2.9*,

$$FRET_{eff} = FRET_{eff(min)} + (FRET_{eff(max)} - FRET_{eff(min)}) \frac{[([S] + [E] + K_D - \sqrt{([S] + [E] + K_D)^2 - 4[S][E]}]}{2[S]} \quad equation 2.9$$

, where *FRET$_{eff}$* is the calculated energy transfer efficiency at a particular total concentration of enzyme ([E]) and fluorescent substrate ([S]), with $FRET_{eff(min)}$ giving the minimum energy transfer efficiency (i.e., the energy transfer efficiency of free oligonucleotide in the absence of any enzyme) and $FRET_{eff(max)}$ the maximum energy transfer efficiency (i.e., the energy transfer efficiency of the saturated substrate). The dissociation constant is given by K_D. In some situations, it was necessary to normalize the data to globally determine the dissociation constant of replicates. This was performed by fitting each replicate to *equation 2.9* to determine the $FRET_{eff(max)}$ value and then dividing through by that value. The separate replicates were then plotted on the same graph and re-fit to the same equation (*equation 2.9*).

2.6.2 Measuring the affinity of 'blocked' enzyme substrate complexes.

FRET experiments of substrate bound to streptavidin and then bound to enzyme were measured by incubating 5 molar equivalents of streptavidin and biotinylated substrate (10 nM DF-B4, *table 2.2*) in buffer containing 10 mM CaCl$_2$, 55 mM HEPES-NaOH pH 7.5, 110 mM KCl, 1 mg/ml BSA and 1 mM DTT for 10 minutes at RT. This was procedure was performed for all three FRET DF-4 constructs (non-, donor-only- and donor-acceptor-labelled). Samples were then incubated at 37

˚C for a further 10 minutes before the first measurement and were recorded and analysed analogously to experiments described in *section 2.6.1* above. SA (same stock used in *section 2.5.2*) was diluted in 50 mM HEPES-NaOH pH 7.5, 100 mM KCl, the same stock was used for all experiments.

2.6.3 Substrate conformational changes as a function of KCl

The FRET change of the substrate (DF-4, *table 2.2*) was measured at a range of KCl concentrations (10 - 500 mM) in a continuous titration. Samples containing 10 nM of either non-, donor-only-, or donor-acceptor-labelled DF-4 constructs, 10 mM KCl, and 50 mM HEPES-NaOH pH 7.5, were prepared with subsequent acquisition of the fluorescent emission spectra at 37 ˚C. The fluorescence intensity of the three samples, at the indicated KCl concentrations were then processed analogously to experiments described in *section 2.6.1* above.

2.6.4 Substrate conformational changes as a function of MgCl$_2$

Samples containing various concentrations of MgCl$_2$ (0.1 – 100 mM) were prepared in 100 mM KCl, and 50 mM HEPES-NaOH pH 7.5, and 10 nM of either non-, donor-only-, or donor-acceptor-labelled DF-4 constructs. The fluorescent emission spectra of the samples at 37 ˚C were collected and analysed analogously to experiments described in *section 2.6.1* above.

2.7: CD spectroscopy

2.7.1. 2-aminopurine Exciton-coupled Circular Dichroism (ECCD) Spectroscopy

Samples containing 10 μM of the indicated DNA construct, 55 mM HEPES-NaOH pH 7.5, 110 mM KCl, 1 mM DTT and, where appropriate, 12.5 μM protein and either 10 mM CaCl$_2$, or 8 mM MgCl$_2$ + 100 μM inhibitor **1/2** were prepared with subsequent acquisition of CD spectra (300–480 nm) at 20 ˚C using a JASCO J-810 CD spectrophotometer. Low-energy CD spectra were collected using a quartz cuvette with a 5 mm pathlength. In samples containing either inhibitor **1** or **2,** the enzyme was pre-incubated with the inhibitor before addition of the substrate. After acquisition of the spectra, EDTA was added to a final concentration of 23.8 mM and CD spectra were re-collected. Dilutions of the appropriate proteins were made from 100 μM stocks in storage buffer (*section 2.1.1*).

CD spectra are an average of five scans recorded in 0.5 nm steps (0.5 s response time) that were baseline corrected using spectra recorded on samples containing the same components, but lacking DNA. The baseline-subtracted spectra were then corrected by smoothing using the Means–Movement option with a convolution width of five in the JASCO Spectra Analysis

software (version 1.53.07). The signal was normalized to account for base-line shift. To normalize for baseline shifts, the signal was averaged between 400-480 nm, a region in which 2-aminopuirine has no absorbance (Johnson et al., 2004), and was subtracted from the associated spectrum. The CD spectra were plotted as Δε per mol 2-AP residue versus wavelength. Each measurement was independently repeated (typically in triplicate) and gave identical results

2.7.2 Determining the affinity of hFEN1 and hFEN1-variants for ECCD oligonucleotides

Samples containing 10 µM of S$_{-1-2}$ or P$_{-1-2}$ (*table 2.4*), were incubated with 12.5 µM of the indicated protein at 20 °C for 10 minutes. The buffer conditions were identical to that used in the collection of low-energy CD spectra (*section 2.7.1*); therefore, samples were either in 10mM CaCl$_2$ alone or 10mM CaCl$_2$ and 40mM EDTA in addition to 110 mM KCl, 55 mM, HEPES pH=7.5, and 1 mM DTT. The samples (20 µl) were electrophoresed on 6 % polyacrylamide gels (29:1) containing 0.5 × TBE and 5 % glycerol. The gels were loaded while running at 200 V at 4 °C and then further electrophoresed at 200 V until the bromphenol blue indicator in a separate well reached the bottom of the gel. DNA was visualized using SYBR® Green according to manufactures protocol and a ChemiDoc^{TM} Imaging System (BioRad).

2.7.3 Measuring unpairing of 'trapped' and 'blocked' enzyme substrate complexes.

Low-energy CD spectra of 'blocked' enzyme substrate complexes were collected by incubating 10 µM biotinylated substrate (Sb$_{-1-2}$, *table 2.4*) with 2 molar equivalents of streptavidin in buffer containing 10 mM CaCl$_2$, 55 mM HEPES-NaOH pH 7.5, 110 mM KCl, 1 mg/ml BSA and 1 mM DTT for 10 mins at RT. 12.5 µM of protein (in storage buffer, *section 2.1.1*) was added to the sample, and spectra were collected analogously to that described in *section 2.7.1*.

'Trapped' enzyme substrate complexes were formed by incubating 10 µM of the biotinylated construct (Sb$_{-1-2}$, *table 2.4*) with 12.5 µM of protein for 10 minutes at RT in buffer containing 10 mM CaCl$_2$, 55 mM HEPES-NaOH pH 7.5, 110 mM KCl, 1 mg/ml BSA and 1 mM DTT. 2 molar equivalents of streptavidin were added to the sample, and spectra were collected analogously to that described in *section 2.7.1*. Dilutions of the appropriate proteins were made from 100 µM stocks in storage buffer (*section 2.1.1*). The same streptavidin stock used for FRET experiments (*section 2.6.2*) was used to from the 'blocked' and 'trapped' species for ECCD analysis.

2.8 Isothermal Titration Calorimetry

ITC was employed to measure the binding affinities of wildtype, truncated and variant hFEN1 for inhibitor **1**, inhibitor **2** and inhibitor **3** using a NANO-ITC machine (TA Instruments). The protein was exchanged from storage buffer into buffer containing 8 mM MgCl$_2$ or 10 mM CaCl$_2$

and 100 mM KCl, 50 mM HEPES (pH 7.5), 1mM DTT using a desalting column. Subsequently, the protein was dialyzed overnight at 4 ˚C against the same buffer composition using a D-Tube Dialyzer Midi (3.5 kDa MWCO) according to the manufacturer's protocol (Novagen). In all cases, the dialysate was used to make the final protein concentration 18 µM and the final inhibitor concentration 200 µM both with 1 % DMSO. Twenty five injections were performed with 180 s spacing time at 25 ˚C. The titration traces were integrated by NITPIC (Keller et al., 2012), and the resultant data were globally fit by SEDFIT to a single site binding model (Houtman et al., 2007). The figures were prepared using GUSSI (http://biophysics.swmed.edu/MBR/software.html).

2.9 Determining the steady-state kinetic parameters of the WThFEN1 and hFEN1Δ336 catalysed reaction of DF-5*

Reaction mixtures containing a range of concentrations of FAM-labelled DF-5* substrate were prepared in RRB and incubated at 37 ˚C for 10 minutes. DF-5* stock solution (annealed as described in *section 2.3.1*) was prepared in 100 mM KCl, 55 mM HEPES pH=7.5, made to a final concentration of 50 µM and used for all the experimental repeats (containing a ratio of 1:1.1 of flap: template strands). Reactions were initiated by the addition of enzyme in RRB. The final substrate and enzyme concentrations used are shown in *table 2.10*. Reactions were sampled, at eight time intervals between 0 and 20 min and quenched with excess EDTA (250 mM). Reaction progress being determined as above (*section 2.4.1*).

Table 2.10: Final enzyme and substrate concentration ranges *used to determine the steady-state kinetic parameters (k_{cat} and K_M) for DF-5* hydrolysis by WThFEN1 and hFEN1Δ336.*

DF-5* (nM)	WThFEN1 (pM)	hFEN1Δ336 (pM)
5000	300	350
2500	250	250
1000	150	100
750	65	65
500	65	50
250	30	40
100	15	15
75	15	15
50	5	15
25	5	10
10	3	10
5	3	5

All reactions were independently repeated at least three times. Initial rates (v_0, nM.min[-1]) were determined by linear regression of plots of the concentration product *versus* time for the first 10% of product formed during the reaction. Normalized initial rates ($v_0/[E]_0$, min[-1]) were obtained from the quotient of initial rate and enzyme concentration used. Kinetic parameters k_{cat} and K_M were determined by generalized nonlinear least squares using a Michaelis-Menten model (MM, *equation 2.10*), from which plots of normalized initial rate as a function of substrate concentration were generated. The error distribution was assumed to be normal, but to account for the increasing variances as a function of substrate concentration the data were weighted to $1/Y^2$. All graph fitting and statistics analyses were conducted using GraphPad Prism (version 6.00 for Windows, GraphPad Software, San Diego California USA).

$$\frac{v}{[E]_0} = \frac{k_{cat}[S]}{K_M + [S]} \qquad equation\ 2.10$$

2.10 Steady-state Inhibition Studies

2.10.1: Estimation of K_i

To determine the concentration range of each inhibitor required to perform a complete kinetic analyses, an estimation of their inhibition constant (K_i) was determined. The initial rate of reaction was determined at various concentrations of inhibitor **1**, **2**, or **3** (*table 2.11*) and 150 nM substrate (DF-5*, *table 2.1*), the approximate K_M value for hFEN1Δ336. The final DMSO concentration was kept constant (1 % (v/v) across the range of inhibitor concentrations [I] tested by creating appropriate 100 X stocks. The reaction mixture at 0 nM inhibitor also contained 1 % DMSO (v/v). Reaction mixtures were pre-incubated before initiation by the addition of hFEN1Δ336. Aliquots of each reaction mixture were removed and quenched with 250 mm EDTA at eight time intervals between 0 and 20 min. The quenched reaction mixtures were analyzed as described above for steady-state analyses (*section 2.9*). The normalized initial rates (v_0/[E]$_0$ (min^{-1})) in the absence of inhibitor (k_{obs}) and at each inhibitor concentration (k_0) were fit to *equation 2.11* using KaleidaGraph 4.0 (Kaleidagraph, Synergy Software, Reading, PA, USA).

$$\frac{k_{obs}}{k_0} = \frac{K_i}{K_i + [\text{I}]} \qquad\qquad equation\ 2.11$$

Table 2.11: Final inhibitor and hFEN1Δ336 concentrations used to determine the initial rate at 150 nM DF-5*. Inhibitor structures shown in figure 2.6.

Inhibitor (nM)	1	2	3
	Enzyme (pM)		
5000	2000	2500	1250
1000	750	500	250
500	400	300	150
100	150	100	50
50	100	40	40
10	40	30	30
0	15	15	15

2.10.2 Determining the steady-state kinetic parameters of the hFEN1Δ336 catalysed reaction of DF-5* as a function of inhibitor 1 and 2

The steady-state kinetic parameters of hFEN1Δ336 with DF-5* were determined at various concentrations of either inhibitor **1** or inhibitor **2** (0, 5, 10, 50, 100, 500, 1000 nM). For each inhibitor concentration, normalized initial rates of reaction ($v_o/[E]_o$, min^{-1}) were measured in triplicate at six different concentrations of DF-5* (10, 50, 100, 500, 1000, 5000 nM) using hFEN1Δ336 concentrations ranging from 20 to 3500 pM. The enzyme concentration never exceeded more than 1 % of the total substrate concentration, but was varied to achieve 10 % product in approximately 10 minutes across the inhibitor concentration range. The final concentrations of all buffer components were 8 mM MgCl$_2$, 110 mM KCl, 55 mM HEPES-NaOH pH=7.5, 0.1 mg/ml BSA and 1 mM DTT with 1 % DMSO (v/v). All reactions were independently repeated at least three times. Reactions were assayed and normalized initial rates were determined as described for steady-state analyses (*section 2.9*). Kinetic parameters k_{cat} and K_M were determined globally for the for simplest types of reversible linear inhibition (mixed, non-competitive, competitive and uncompetitive; *equations 2.12 – 2.15*) by non-linear regression plots of normalized initial rates ($v_o/[E]_o$, min^{-1}) versus the substrate concentration (nM) for each concentration of inhibitor. The same weighting ($1/Y^2$) was applied in each case due to the heteroskedastic nature of the data. Akaike information criteria (AIC) was employed as statistical tests to aid model selection (e.g. non-competitive versus competitive); statistical analyses were done using GraphPad Software Inc. Unless the more complicated model gave a difference in AIC of more than -6 (95 % probability), the less complicated model was chosen as the appropriated model (i.e. law of parsimony). This type of analyses penalises the more parameterized model unless the sum-of-squares is significantly reduced. As an additional check that the correct model was selected, the residuals from both the non-competitive and mixed inhibition models were inspected.

Uncompetitive: $$\frac{v}{[E]_0} = \frac{k_{cat}[S]}{K_M + \left(1 + \frac{[I]}{K_{iu}}\right)[S]}$$ *equation* 2.12

Competitive: $$\frac{v}{[E]_0} = \frac{k_{cat}[S]}{K_M\left(1 + \frac{[I]}{K_{ic}}\right) + [S]}$$ *equation* 2.13

$$\text{Non}-\text{competitive:} \qquad \frac{v}{[E]_0} = \frac{k_{cat}[S]}{K_M\left(1+\frac{[I]}{K_{ic}}\right)+\left(1+\frac{[I]}{K_{iu}}\right)[S]} \qquad equation\ 2.14$$

$$\text{Mixed:} \qquad \frac{v}{[E]_0} = \frac{k_{cat}[S]}{K_M\left(1+\frac{[I]}{K_{ic}}\right)+\left(1+\frac{[I]}{K_{iu}}\right)[S]} \qquad equation\ 2.15$$

2.10.3 Secondary Inhibition plots

Normalized initial rates ($v_o/[E]_o$, min^{-1}) were calculated at each substrate and inhibition concentration tested by inputting the catalytic parameters (k_{cat} and K_M) and inhibition constants (K_{ic} and K_{iu}) determined by non-liner regression fitting to the relevant inhibition models (*section 2.10.2*) into a competitive, non-competitive and mixed inhibition model (*equation 2.13-2.15*). These calculated rates were used to plot a best fit line on a graph of $S/(v_o/[E]_o)$ against inhibition concentration. The experimentally determined normalized initial rates ($v_o/[E]_o$, min^{-1}) at the same substrate and inhibitor concentrations were then plotted on the same graph to assess their goodness of fit and aid in model-selection.

2.11 Expression and purification of T5FEN

WT T5FEN was overexpressed and purified to homogeneity according to the protocol described in Sengerova, 2010.

2.12 Synthesis and purification of substrates for T5FEN

The oligonucleotide used for experiments with T5FEN is detailed in *table 2.12* below. Stability prediction software (MFold, Zuker 2003) was used to predict the overall folded structure. Oligonucleotides were synthesised and purified according to the protocol described in Patel, 2012. Characterisation of oligonucleotides was carried out using parameters in *table 2.13*.

Table 2.12: Sequences of each oligonucleotide used for T5FEN. F=FAM. Calculated and experimental masses along with calculated extinction coefficient values shown in table 2.13.

Oligonucleotide Code	Sequence
pY-7	5'-FAM-CGC TGT CGA ACA CAC ACC GCT TGC GGT GTG TGT TCC ACA AC-3'

Table 2.13: Details of all the oligonucleotides synthesised and purified for work with T5FEN. Naming system of the oligonucleotides pY = pseudo Y. T_m stated are calculated (M. Zucker, 2003) for DNA at a final concentration of 10 nM, at 50 mM KCl for pseudo Y substrates designed for T5FEN.

Oligonucleotide Code	Calculated Mass (gmol⁻¹)	Actual Mass (gmol⁻¹)	ε_{260} (Lmol⁻¹cm⁻¹)	T_m
pY-7	13059.6	13059	374900	73.3

2.13 Determining the steady-state kinetic parameters of the T5FEN catalysed reaction of pY-7 in the presence and absence of inhibitor 1

The normalized initial rate of reaction ($v_0/$[E]$_0$, min⁻¹) of T5FEN with pY-7 was measured at a substrate concentration equal to a previously reported K_M value (35 nM) and at an appropriate enzyme concentration (5 pM) to obtain ~10% product in 20 minutes. A 1 µM pY-7 stock solution was prepared in 50 mM KCl and 25 mM potassium glycinate pH=9.3. This stock was then used for all the experimental repeats. The unimolecular T5FEN substrate was annealed by heating to 95 °C for 2 minutes and cooled on ice for 5 minutes. Reactions concerning T5FEN were treated identically to that describe for hFEN1; however, final solution concentrations were 27.5 mM potassium glycinate pH=9.3, 55 mM KCl, 0.01 mg/ml BSA, 1 mM DTT, 10 mM MgCl$_2$, and 100 µM inhibitor **1** or 1 % DMSO (v/v). In addition, a different gradient using WAVE buffers (*table 2.8*) was used to separate the substrate and product (*figure 2.8*).
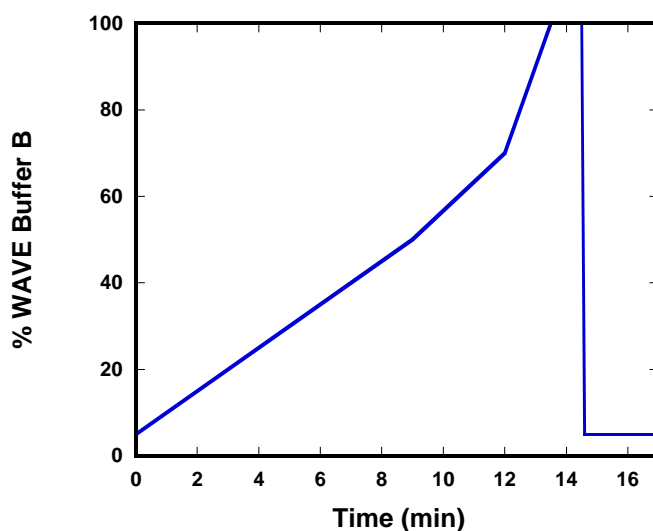
Figure 2.8: Denaturing HPLC gradient used to separate the reaction products from substrate. The specific gradient switch points are as follows: 0min 5%B 95%A, 5min 30%B 70% A, 9 min 56%B 44% A, 12 min 70%B, 13.5 min 100%B, 14.5 min 100% B, 14.6 min, 5%B 95%A, 17min 5%B 95%B.

2.13.1 Determining the steady-state kinetic parameters of T5FEN and hFEN1Δ336 with inhibitor 1 as a function of pH

To rule out the prospect of pH affecting the ability of inhibitor **1** to inhibit the T5FEN catalysed hydrolysis of pY-7, reactions were performed analogously to those described above but at pH=7.5. As a control, the hFEN1Δ336 catalysed reaction with DF-5* was performed at pH=9.3 and at its respective K_M (150 nM). The steady-state catalytic parameters were determined as described above, expect now the final T5FEN solution concentrations were 110 mM KCl, 55 mM HEPES pH=7.5, 8 mM MgCl$_2$, 0.1 mg/ml BSA, 1 mM DTT, and 100 µM inhibitor **1** or 1 % DMSO (v/v). The final hFEN1Δ336 solution concentrations were 27.5 mM potassium glycinate pH=9.3, 55 mM KCl, 0.01 mg/ml BSA, 1 mM DTT, 10 mM MgCl$_2$, and 100 µM inhibitor **1** or 1 % DMSO (v/v). All reactions were independently repeated at least three times. To account for the non-ideal reaction conditions, the enzyme concentrations were altered to obtain ~10% product in 20 minutes. The final enzyme concentrations were 50 pM and 500 pM for T5FEN and hFEN1Δ336, respectively.

2.14 Expression and purification of XRN1

The sequence corresponding to residues 1-1245 of *Kluyveromyces lactis* Xrn1 (K.lactisXrn1) was amplified from a full-length cDNA clone (courtesy of Chang, J. H. *et al*. Nat Struct Mol Biol (2011) 18, 270-276) and subcloned in frame with a C-terminal hexahistidine tag into the expression vector pET-26b-XRN1 (Kan^R). Plasmids were isolated using a Qiagen miniprep kit, and verified

using DNA sequencing (GATC-Biotech) and the visualisation software Serial Cloner 2.1 (Serial Basics, Freeware); data can be found in the Appendices, *figure A10*. Successful sub-clones were transformed into Rosetta (DE3) BL21(Kan[R]; Cm[R]) *E.Coli* cells according to the protocol described in *section 2.12*. These cells were plated on LB (Kan (25 μg/ml), Cm (34 μg/ml)), and single colonies were used for overexpression and purification. The expression and purification methodology was identical to that described for hFEN1 in *section 2.1.2*, except when an OD_{600} of 0.6 was achieved, protein expression was induced by adding IPTG to a final concentration of 0.4mM to each of the cultures. Then, the cultures were then incubated overnight at 20 ˚C. In addition, the storage buffer differed to that of hFEN1, 20 mM Tris pH = 7.5, 200 mM NaCl, 2 mM DTT and 50 % glycerol.

2.15 Synthesis and purification of substrates for XRN1

RNA and DNA oligonucleotides used in XRN1 assays were ordered purified using reverse-phase HPLC and synthesised by DNA Technology (Risskov, Denmark), using standard phosphoramidites. Reactions were performed with both single-stranded and double-stranded constructs. The composition of the bi-molecular constructs is shown in *table 2.14* with the sequence of all oligonucleotides shown in *table 2.15*.

Table 2.14: RNA:DNA constructs used in XRN1 kinetic assays. T_m stated are calculated (M. Zucker, 2003) for RNA at a final concentration of 500 nM at 100 mM KCl. Sequences of the individual strands are shown in table 2.15.

Construct	Oligonucleotide composition	Tm (˚C)
RNA1/DNA1	RNA1 + DNA1	55.4
RNA1/DNA2	RNA1 + DNA2	81.3
RNA1/DNA3	RNA1 + DNA3	77.8

Table 2.15: Sequences of oligonucleotides for work with XRN1. *FAM = fluorescein, TAMRA = tetramethylrhodamine, 5PHOS = 5′-phosphate, and HO = -hydroxyl. Calculated and experimental masses along with calculated extinction coefficient values are shown in table 2.16.*

Oligonucleotide Code	Sequence
RNA1-OH	5'-HO-ACU CAC UCA CUC ACC AAA AAA AAA AAA ACC-FAM-3'
RNA1	5'-/5PHOS/ACU CAC UCA CUC ACC AAA AAA AAA AAA ACC-FAM-3'
DNA1	5'-TAMRA-GGT TTT TTT TTT TTT GG-OH-3'
DNA2	5'-TAMRA-GGT TTT TTT TTT TTT GGT GAG TGA GTG AGT-OH-3'
DNA3	5'-TAMRA-GGT TTT TTT TTT TTT GGT GAG TGA GTG AGT CCC CCC CCC CCC C-OH-3'

Table 2.16: Details of all the oligonucleotides synthesised and purified for work with XRN1.

Oligonucleotide Code	Calculated Mass (gmol^{-1})	Actual Mass (gmol^{-1})	ε_{260} (Lmol^{-1}cm^{-1})
RNA1-OH	10074.5	10074	326500
RNA1	10154.5	10154	326500
DNA1	5783.5	5885	178000
DNA2	9915.2	10017	312800
DNA3	13674.5	13776	406700

2.16 XRN1 activity assays

XRN1 assays were performed in buffer containing 30 mM Tris-HCl pH = 8.0, 2 mM MgCl$_2$, 50 mM NH$_4$Cl, 20 µg/µl BSA, and 0.5 mM DTT. Samples containing 500 nM of the indicated RNA:DNA construct were incubated at 28 °C (optimal growth temperature of *Kluyveromyces lactis*), and reactions were initiated by addition of XRN1 in the same buffer composition. The final enzyme concentration for each particular substrate is shown in *table 2.17*. Prior to each experiment, single-stranded and double-stranded substrates stock solutions were prepared in 50 mM NH$_4$Cl, 30 mM Tris-HCl pH = 8.0, made to a final concentration of 10 µM, and used for all experimental repeats. Bi-molecular substrates contained a ratio of 1:1.1 of RNA: DNA strands and were pre-folded at 95 °C for 30 seconds, and allowed to cool to RT before addition to the reaction mixture. Aliquots of the reaction mixture were removed at the indicated time intervals and quenched by

addition of an equal volume of 2 × formamide loading buffer containing EDTA (80 % formaide, 500 mM EDTA). Separation of substrate and products was achieved using a 20 % polyacrylamide gels (19:1) containing 7M urea and 1 × TBE. The gel was electrophoresed at 25 W at RT. In total, 20 ng of nucleic acid was loaded onto the gel. RNA and DNA strands were visualised by detecting FAM and TAMRA fluorescence, respectively, using a ChemiDoc XRS system (Bio-Rad).

Table 2.17: Final enzyme concentrations used to determine activity of XRN1 on range of substrates. *All reactions were performed at a final substrate concentration of 500 nM, sequences given in table 2.15.*

Substrate	XRN1 (nM)
RNA1-OH	50
RNA1	50
RNA1/DNA1	50
RNA1/DNA2	500
RNA1/DNA3	500

2.16.1 XRN1 inhibition assay

The activity of XRN1 on RNA1/DNA2 was determined at an inhibitor **1** concentration of 100 μM. The presence of the inhibitor meant that the final buffer conditions were similar to that described above, 30 mM Tris-HCl pH = 8.0, 2 mM $MgCl_2$, 50 mM NH_4Cl, 20 μg/μl BSA, and 0.5 mM DTT, but also contained 1 % DMSO. The final enzyme and substrate concentrations were 50 and 500 nM, respectively. The reaction set-up and analysis was no different to that described above, except XRN1 was incubated on ice with 100 μM inhibitor **1** in buffer containing, 30 mM Tris-HCl pH = 8.0, 2 mM $MgCl_2$, 50 mM NH_4Cl, 20 μg/μl BSA, and 0.5 mM DTT, for 10 minutes prior to addition to the reaction mixture.

2.16.2 Partial alkaline hydrolysis of RNA1

For partial alkaline hydrolysis of RNA1, approximately 20 μl of a 50 μM solution was incubated with 80 μl of 50 mM Na_2CO_3/$NaHCO_3$ buffer (pH 10), for 20 minutes at 90 °C.

Chapter 3: The effect of human flap endonuclease active site mutations upon substrate binding and catalysis

3.1 Mutational Analyses

Structural and evolutionary conservation analyses in combination with *in vivo* toxicity screens have identified residues important for FEN1 catalysis (Storici et al., 2002; Tsutakawa et al., 2011). Two basic residues of α4, Lys-93 and Arg-100 and the Tyr-40 residue of α2 are implicated in the unpairing mechanism of hFEN1 (*section 1.4.1*) (Finger et al., 2013). The stacking residue (Tyr-40) is proposed to stabilise the unpaired state through π-π stacking interactions with the unpaired nucleobase and is conserved in FEN1s from mammals, yeast and archaea, but is replaced with an alternative stacking residue in bacteriophage flap endonucleases and other superfamily members (*figure 3.1*). Lys-93 and Arg-100 are strictly conserved across the entire 5'-nuclease superfamily. In studies with T5FEN the equivalent residue of Lys-93 has been proposed to act as an electrophilic catalyst (Sengerova et al., 2010). In hFEN1 this residue is also proposed to assist with capture of the unpaired scissile phosphate; a similar role has been proposed for Arg-100 (*figure 3.1*) (Tsutakawa et al., 2011).

FEN1 has been proposed to utilise one variation of the two-metal-ion mechanism (*section 1.2*), and therefore acidic residues within the active site are crucial for the binding of Mg^{2+} ions which provide the catalytic enhancement for phosphate diester hydrolysis (*figure 3.1*). Mutation of the aspartate to alanine at position 181 in the amino acid sequence (D181A) reduces the affinity of FEN1 for M^{2+} to such an extent that in the D181A:substrate complex crystal structure no metal ions were observed in the active site despite being present in the crystallisation buffers (Tsutakawa et al., 2011). This variant was first characterised in hFEN1 using flow cytometry and was reported to be inactive with respect to phosphate diester hydrolysis activity but retain substrate binding ability (Nolan et al., 1996).

3.1.1 Production of mutated hFEN1 proteins

Four active site residues, Tyr-40, Lys-93, Arg-100 and Asp-181 were targeted for site-directed mutagenesis. cDNAs encoding variant FENs were engineered so that each of the targeted residues was substituted with alanine, and mammalian cell expression constructs were generated by subcloning the variant cDNAs into the vector pET28b. The mutagenesis required to create these proteins was performed by David Finger (Tsutakawa et al., 2011). Here, the same methodology was employed to create a double hFEN1 variant (K93AR100A), a previously uncharacterised mutation (*Appendices, figure A1*).
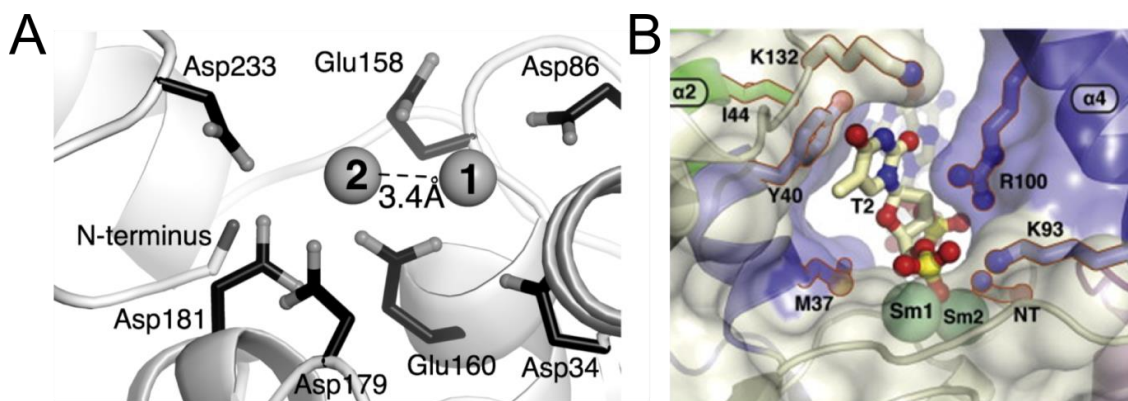
Figure 3.1: Positions of mutated residues in hFEN1 and hFEN1-complexes. (*A*) *Substrate free active site of hFEN1 (PDB: 1UL1) illustrating the conservational position of carboxylate residues including Asp-181 and metal ions (1 and 2, grey). Taken from Tomlinson, C. G. et al. BIOCHEM. SOC. T (2010) 38, 433-437.* (*B*) *Active site structure of hFEN1 with bound DNA product (PDB:38QK). The cleaved 5'-phosphate is coordinated by Lys-93, Arg-100 and two Sm^{3+} ions. The base of the terminal nucleotide (T2) stacks against Tyr-40. K93/R100 and Y40 reside on α4 and α2, respectively. Helices of the helical arch (blue). NT indicates the N-Terminal region which stabilises the bound DNA in the active site. Taken from Tsutakawa, S.E., et al. (2011). Cell 145, 198-211.*

To investigate the effects these mutations had upon the proposed double nucleotide unpairing mechanism of hFEN1, wild type hFEN1 (WT) and variant FENs were all purified following expression in BL21 (DE3) *E.coli* expression using the procedure described in *section 2.1*. Proteins were purified using five different chromatography phases (Ni^{2+} immobilization affinity, anion-exchange, heparin, phenyl sepharose and size exclusion). Firstly, IMAC was used as a capture step to separate the protein of interest. To enable this, all proteins were labelled with a C-terminal (His)₆-tag. It was discovered that anion-exchange chromatography was necessary to remove nucleic acid contamination, and the subsequent chromatographic steps were necessary to purify the protein to a level acceptable for kinetic analyses. Prior to experiments designed to investigate the unpairing mechanism (*chapter 4*), the effect these mutations had upon substrate binding and catalysis was investigated. Apart from the double hFEN1 variant K93AR100A, all of these mutations have been studied previously and are known to be detrimental towards catalysis (Finger et al., 2013; Patel et al., 2012; Patel et al., 2013; Qiu et al., 2004; Storici et al., 2002; Tsutakawa et al., 2011). Here, we sought to study each mutation under identical conditions to allow effective quantitative comparison of catalytic and binding parameters and to underpin further work investigating DNA conformational change.

3.2 Substrate design

Early literature characterised the minimal substrate for hFEN1 *in vitro* to be a bifurcated ds molecule with a ss 5'-flap (Pseudo-Y) (Kaiser et al., 1999). Binding and catalysis have both been shown to be more efficient when substrates have ds regions either side of the flap (Harrington and Lieber, 1994; Turchi et al., 1994). Reaction rate and specificity have been shown to be increased by the presence of a one-base overlap at the site of the nick resulting in a local double-flap structure (*figure 3.2*) (Kao et al., 2002). The one nt 3ʹ-flap corresponds to the last nucleotide added by pol δ during strand displacement synthesis (Finger et al., 2012). hFEN1-DNA structures show that hFEN1 recognises a single nucleotide 3'-flap, and it has been proposed that this recognition event marks the substrate as being eligible for phosphate diester hydrolysis .

In vivo the 3' unpaired nucleotide and upstream region of hFEN1 substrates are one Okazaki fragment, whereas the 5'-flap and downstream region are the previous Okazaki fragment. Because both Okazaki fragments are synthesised using the same lagging strand template during replication, double-flap structures generated *in vivo* are 'equilibrating' (Liu et al., 2004). As a result the 5'- and 3'- flap length can vary giving rise to a number of different conformations (*figure 3.2*), although hFEN1 action on equilibrating double flaps always produces the same products presumably by recognising a single conformer with a 3'-single nt flap. *In vitro* 'non-equilibrating' (static) flap substrates, with non-complementary 5'- and 3'- flaps, are commonly used to simplify analysis (*figure 3.2*) and undergo reaction at the same site as equilibrating double flaps (Kao et al., 2002). Substrates used throughout this thesis were static, double-flap (DF) constructs, comprising of a displaced ss 5'-flap of a known length with a one nt 3'-flap.

Equilibrating flaps

polymerase DNA   FEN DNA

Static flaps

Single flap   Double flap
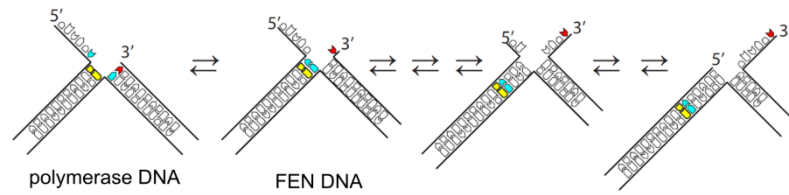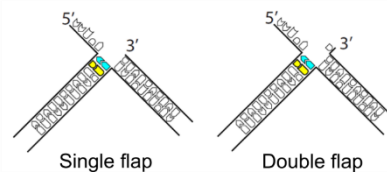
Figure 3.2: Flaps can potentially migrate in vivo, but static flaps are commonly used in vitro (courtesy of David Finger). *Discontinuous synthesis of the lagging stand produces bifurcated structures that can potentially migrate and theoretically form multiple structures; structures relevant to polymerase and FEN have been highlighted. Single- and double flap structures that are commonly used in vitro, the 5'- and 3'- flaps are non-complementary to the template strand and therefore are static and do not change length.*

3.3 Studying FEN-substrate interactions using fluorescence polarisation

Fluorescence anisotropy (FA) is commonly used to accurately measure the dissociation constants (K_D) of fluorescently labelled DNA-protein complexes as it is able to quantify the amount of bound and free substrate while in equilibrium. Fluorescence anisotropy measures the tumbling rate of fluorophores in solution that have been excited with plane polarized light of appropriate wavelength. Only fluorophores which have absorption transition moments aligned along the electric vector of the incident light will be preferentially excited (Lundblad et al., 1996). The term anisotropy (*r*) describes the extent to which the emission is polarised. This is 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.

The slower a molecule tumbles, the more the light will remain polarized. The faster a molecule tumbles, the less the light will remain polarized. Binding of an enzyme to a substrate molecule will cause the tumbling rate to decrease and this effect will be reflected in the *r* value, which will become increasingly large. Eventually, the point will be reached were all the substrate molecules are bound by enzyme and the *r* value should become independent of the enzyme concentration.

3.3.1 The majority of the interface between hFEN1 and DNA is distant from the active site

A DF substrate used to characterise the binding affinity of hFEN1 and variant FENs was prepared as bi-molecular construct. Two oligonucleotides were annealed, one which formed the 5'-flap strand (F1) and the other which acted as a template (T1) and contained a fold-back hairpin providing the single nt 3'-flap (*figure 3.3*).The DF substrate possessed a 5 nt 5'-flap (DF-5) and was labelled with a 5'-fluorescien (FAM) (*figure 3.3*).

The conditions of experiments were kept as close as possible to those that gave optimal rates of reaction (*section 2.3*) (37 ˚C, 110 mM KCl, 55 mM HEPES pH=7.5, 1 mM DTT and 0.1 μg/μl BSA), whilst at the same time preventing catalysis. As described previously (*section 1.2.4*), Ca^{2+} ions do not support catalysis therefore in FA experiments $MgCl_2$ was substituted for $CaCl_2$. The concentration of $CaCl_2$ was chosen based on previous data which states that 10 mM Ca^{2+} results in full occupation of the T5FEN1 active site (Tomlinson et al., 2011). Dissociation constants were also determined in the absence of divalent metal ions. EDTA was used to chelate any contaminating divalent metal ions, a concentration of 2 mM EDTA was used so that the ionic strength matched FA experiments with 10 mM Ca^{2+}.

After FA titrations were complete, 100 μl aliquots were removed and quenched in 1.5 M NaOH and 80 mM EDTA. This was then analysed using dHPLC equipped with a fluorescence detector to check for reaction. Samples analysed in the presence of EDTA had not reacted. In the presence of Ca^{2+} ions and WThFEN1 ≤ 5% of the substrate present had reacted due to unavoidable Mg^{2+} contamination of samples.
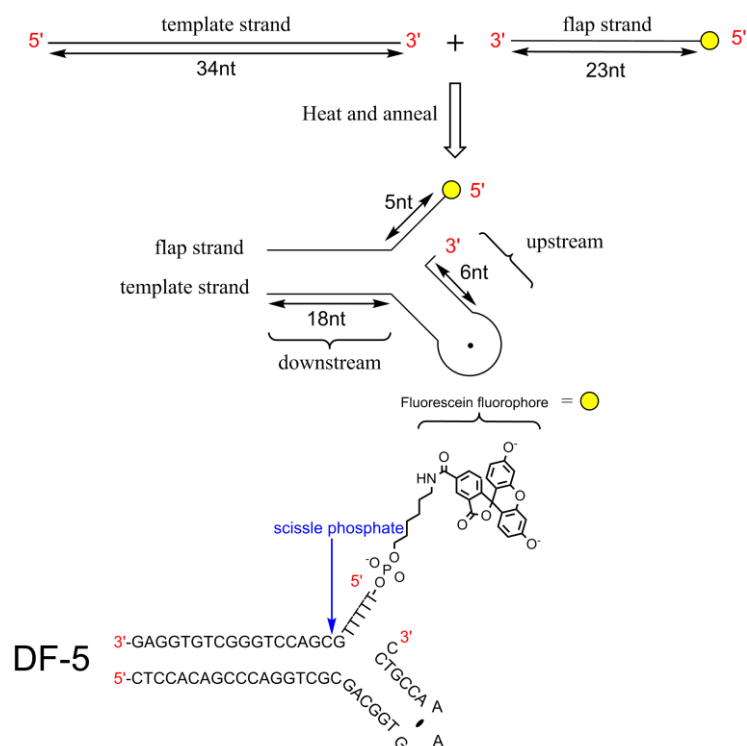
Figure 3.3: Structure of static double-flap oligonucleotide substrate DF-5, *a bi-molecular construct prepared from a 34 nt template strand (T1) and 23 nt flap strand (F1). The template strand forms a fold-back hairpin stabilised by a GAA tri-loop and provides a 1 nt non-complementary 3'-flap. The flap strand, possessing a 5'-FAM, anneals with the template strand and provides a 5 nt non-complementary 5'-flap.*

Steady-state and time resolved anisotropy are the two main methods used to measure fluorescence anisotropy. The experiments discussed in this section utilised anisotropy measurements where continuous excitement with vertically polarised light leads to excitation of fluorophores, which are aligned with the plane of polarised light. The excited state population has anisotropy (*r*), and is calculated by measuring the difference in intensity of fluorescence in the plane parallel (I_{VV}) and perpendicular (I_{VH}) to the vertical polarisation plane, and is described by *equation 3.1*.

$$r = \frac{I_{VV} - GI_{VH}}{I_{VV} + 2GI_{VH}} \qquad equation\ 3.1$$

where *G* is an instrument specific correction factor accounting for the wavelength dependence of radiation transmission by the monochromator (Azumi and McGlynn, 1962). *G* can be estimated comparing the emission intensities from the sample in planes parallel and perpendicular to the plane of a horizontally polarized excitation beam (I_{HH} and I_{HV} respectively), as stated in *equation 3.2*.

$$G = \frac{I_{HV}}{I_{HH}} \qquad\qquad equation\ 3.2$$

Dissociation constants were measured directly by the step-wise addition of the respective enzyme to labelled substrate. Values of I_{VV}, I_{VH}, I_{HV}, I_{HH} and G were measured using a Horiba Jobin Yvon FluoroMax-3® fluorimeter.

The value of r was calculated using the equations described above and data were fitted using non-linear regression fitting (Kaleidagraph, Synergy Software, Reading, PA, USA), to *equation 3.3*.

$$r = r_{min} + (r_{max} - r_{min}) \frac{\left[([S] + [E] + K_D - \sqrt{([S] + [E] + K_D)^2 - 4[S][E]} \right]}{2[S]} \quad equation\ 3.3$$

where r is the measured anisotropy at a particular total concentration of hFEN1 protein (*[E]*) and fluorescent substrate (*[S]*), with r_{min} giving the minimum anisotropy, (the anisotropy of free oligonucleotide in the absence of any protein) and r_{max} the maximum anisotropy, (the anisotropy of the enzyme-saturated substrate). The dissociation constant is given by K_D. The solution to this equation 3.3 is a quadratic to account for the various states of the ligand during the titration as the K_D values measured are close to the experimentally known total substrate concentration. *Table 3.1* shows the K_D values obtained for WThFEN1 and mutated FENs with DF-5 as the substrate.

Table 3.1: K_D parameters for 10 nM DF-5 bound to WThFEN1, Y40A, K93A, R100A, K93AR100A and D181A determined by fluorescence anisotropy. *All measurements were made at 37 ˚C in 110 mM KCl, 55 mM HEPES-NaOH pH = 7.5, 1 mM DTT, 0.1 mg/ml BSA and 10 mM Ca^{2+} or 2 mM EDTA as indicated. The dissociation constants derived from equation 3.3 are reported with the standard errors in parentheses. The minimum and maximum anisotropy values, r_{min} and r_{max}, respectively, for triplicate experiments are reported with the standard errors in parentheses. n.d. = not determined. Accompanying graphs are shown in figure 3.4.*

Substrate	DF-5											
	5 nt 5'-flap, 1nt 3'-flap											
Enzyme	**WT**		**Y40A**		**K93A**		**R100A**		**K93AR100A**		**D181A**	
Buffer	Ca^{2+}	EDTA	Ca^{2+}	EDTA	Ca^{2+}	EDTA	Ca^{2+}	EDTA	Ca^{2+}	EDTA	Ca^{2+}	EDTA
K_D **(nM)**	46 (± 6)	50 (± 4)	175 (±25)	n.d.	58 (± 8)	93 (± 6)	92 (± 10)	111 (± 8)	78 (± 8)	106 (± 9)	180 (± 32)	n.d.
r_{min}	0.05 (± 0.001)	0.03 (± 0.001)	0.04 (± 0.002)	n.d.	0.05 (± 0.001)	0.03 (± 0.001)	0.04 (± 0.001)	0.03 (± 0.001)	0.05 (± 0.001)	0.04 (± 0.001)	0.03 (± 0.001)	n.d.
r_{max}	0.10 (± 0.001)	0.11 (± 0.001)	0.13 (± 0.003)	n.d.	0.12 (± 0.002)	0.14 (± 0.001)	0.12 (± 0.001)	0.13 (± 0.001)	0.12 (± 0.002)	0.14 (± 0.002)	0.13 (± 0.003)	n.d.
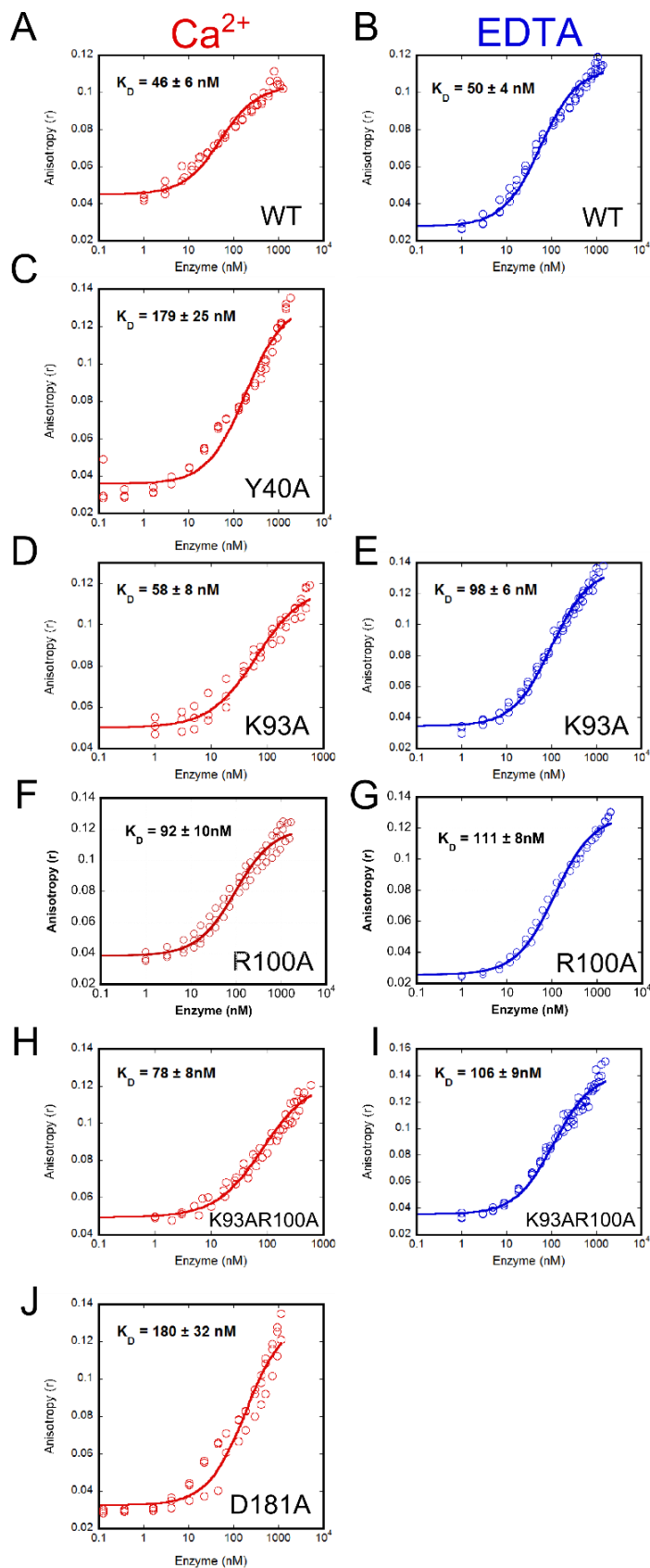
Figure 3.4: DF-5 binding curves for (A) hFEN1 / 10 mM CaCl₂, (B) hFEN1 / 2 mM EDTA, (C) hFEN1Y40A / 10 mM CaCl₂, (D) hFEN1K93A / 10 mM CaCl₂, (E) hFEN1K93A /2 mM EDTA, (F) hFEN1R100A / 10 mM CaCl₂, (G) hFEN1R100A / 2 mM EDTA, (H) hFEN1 K93AR100A / 10 mM CaCl₂, (I) hFEN1K93AR100A / 2 mM EDTA, (J) hFEN1D181A / 10 mM CaCl₂. *Apart from the specified conditions, assays were performed in 110 mM KCl, 55 mM HEPES pH=7.5, 0.1 μg/μl BSA, 1 mM DTT and at 37 ˚C. The data points were fitted to equation 3.3. The K_D value for each condition is shown in the insert of each graph. Each K_D is representative of at least three repeats, and standard error is given.*

The measured r_{min} and r_{max} values (*table 3.1*) are similar to those obtained in previous FA assays of DNA-protein interactions using FAM as the fluorophore (*table 3.1*) (Fogg et al., 2001; Richards et al., 2008). Previous work conducted by Nikesh Patel demonstrated that the starting *r* value varies with 5'-flap length (Patel 2012). The low r_{min} values in the absence of any protein imply that the rotation of the fluorophore is not restricted on the terminus of the 5 nt 5'-flap. However, there is a subtle variation in the r_{min} value depending on the presence or absence of Ca^{2+} ions. The r_{min} values are approximately 0.045 and 0.035 in the buffers containing Ca^{2+} and EDTA, respectively. This suggests that Ca^{2+} ions are interacting with either the highly negatively charged DNA backbone or FAM itself and thereby restricting the fluorophore's rotation, albeit to a small extent. The r_{max} observed for these titrations was typically between 0.11 and 0.13. These values are far from the fluorphore-specific theoretical maximum value (r_0), which for fluorescein is 0.35 (Johnson, 2005; Lakowicz, 1999). This suggests that the binding mode does not severely restrict the movement of the fluorphore and implies that FEN does not bind directly to the 5'-flap terminus.

In each case addition of protein to the DNA resulted in a curve that approached a saturation plateau. However, titrations did not reach a point where the *r* value was completely independent of the enzyme concentration. To investigate possible light scattering effects, enzyme was added to an unlabelled version of DF-5. In this case the *r* value remained constant throughout the titration. A similar result was obtained for the same control experiment with T5FEN (Sengerova, 2009). It is thought that this phenomenon may be an indication of initial tight binding, followed by other, non-specific binding events taking place. This is a complicating factor in this assay and is discussed in more detail in *section 3.5*. However, fits to all curves yielded similar K_D values (46 – 180 nM), regardless of the presence of mutation.

The substrate binding affinity measured by FA for WThFEN1 was in the low nanomolar range in accordance with K_D values measure by gel shift assays (Gloor et al., 2010; Kim et al., 2001), and recently published single-molecule experiments (Craggs et al., 2014). The dissociation constant of WThFEN1 for DF-5 is 46 ± 6 nM and 50 ± 4 nM in the presence and absence of Ca^{2+}, respectively. Thus, the affinity for the substrate measured by FA was identical ± Ca^{2+} within experimental error and suggested that the strength of the interaction between FEN and DNA was independent of the presence of Ca^{2+} ions within the active site. Furthermore, the dissociation constants in the presence of absence of Ca^{2+} ions are similar for all variant FENs tested, with only a slight decrease in affinity brought about by removal of the Ca^{2+} ions (*table 3.1*). In agreement with the hFEN1:substrate crystal structure, these FA studies indicate that the major FEN:DNA binding interface is between the ds-regions of the DNA and the protein

(including the 3'-flap) rather than the 5'-flap (*figure 3.5*). Therefore, neutralisation of the highly negatively charged active site by divalent metal ions appears to have little effect on the on complex stability.

The variant FENs tested maintained substrate binding capabilities. In the presence of 10 mM Ca^{2+} the dissociation constants were 58 ± 8 nM and 92 ± 10 nM for K93A and R100A, respectively. This modest decrease in substrate dissociation constant implies that removal of either basic residue is not severely detrimental to the interaction between the substrate and FEN1. Removal of both these residues has the same effect as removal of only one residue; the dissociation constant measured for K93AR100A in the presence of Ca^{2+} ions was 78 ± 8 nM. This suggests that the potential positive charge delivered by either residue is not contributing significantly to the overall affinity. The dissociation constants for D181A and Y40A are of the same order of magnitude to WThFEN1. However, both K_D values for D181A and Y40A are ~3-fold higher compared to WThFEN1 (*table 3.1*). This data shows that the mutations made do not dramatically change the affinity for a common substrate and all proteins are comparable to WThFEN1.
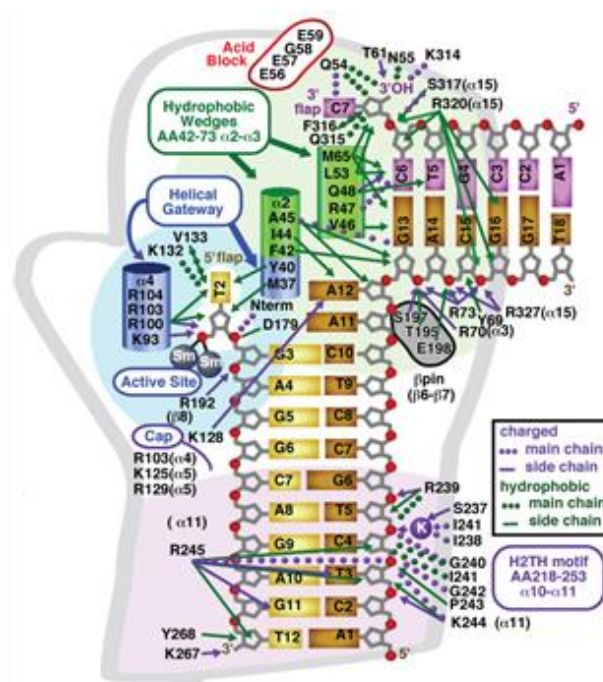
Figure 3.5: Schematic representation of FEN showing the specific interactions between the DNA double flap substrate and residues in the major binding regions. *Extensive contacts are made with the downstream ds DNA through the H2TH motif and bound K+ ion. In addition, the 3'-flap upstream ds DNA are tightly bound through the combined effect of multiple resides. However, there are very few interactions between hFEN1 and the 5'-flap strand near the active site. Taken from Tsutakawa, S.E., et al. (2011). Cell 145, 198-211.*

3.4 Investigating the effect of active site mutations on the hFEN1 mechanism using single turnover kinetics

hFEN1 catalysed reaction of a DF substrate produces two products, a dsDNA product (P) and a short ssDNA fragment (Q). Product P is nicked DNA *in vivo*, but with a static DF, P resembles gapped DNA as the 3'-flap is non-complementary (*figure 3.6*). It is suspected that the ssDNA product (Q) is released immediately after phosphate diester hydrolysis. However, release of the dsDNA product (P) is rate limiting under substrate-saturating multiple turnover (MT) conditions ($[E] < [S] > K_M$) (*figure 3.6*) (Williams et al., 2007a). It is also noteworthy that the catalytic efficiency (k_{cat}/K_M = 1.5×10^8 ± 1.8×10^7 s^{-1} M^{-1}) of hFEN1 acting on static double flap substrates approaches enzyme:substrate association rates (10^9 s^{-1}M^{-1}). Indeed under substrate limiting conditions ($[E] < [S] < K_M$; k_{cat}/K_M conditions) a bacteriophage FEN has been shown to undergo diffusion controlled reaction (Sengerova et al., 2010). Studies of T5FEN under single-turnover (ST) conditions have shown a physical step in the reaction cycle (protein or DNA conformational change, k_{conf}), rather than chemistry, is rate limiting (Sengerova et al., 2010). This is likely to be the case for eukaryotic FENs *in vitro* as well (Finger et al., 2012). Therefore, assuming that chemistry is not rate limiting for WT hFEN1 then the single-turnover rate (k_{ST}) for hydrolysis of a DF substrate is approximately equivalent to k_{conf}, the rate of one or more protein and/or DNA conformational changes (*figure 3.6*).

Previously, the catalytic efficiency of WT and mutated FENs has been compared under k_{cat}/K_M conditions (Tsutakawa et al., 2011). Because substrate-limiting rates are likely diffusion controlled and saturating conditions are rate limited by product release, multiple turnover comparisons of WT and mutated proteins may underestimate the impact of protein alteration. However, maximal ST rates are more likely to better reveal the impact of protein mutations. Thus, the experiments described below measure the rates of reaction after enzyme substrate complex formation and before product release (*figure 3.6*).

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 substrate construct, DF-5, which was used to measure K_D values for WT and mutated FENs (*section 3.3*), was used to determine the single-turnover values (k_{ST}) for the same enzymes. In all ST experiments presented in this chapter the final *[E]* and *[S]* concentrations were 1000 nM and 2.5 nM, respectively. Therefore, single turnover reactions measured the decay of the enzyme-substrate complex ($d[ES]/dt$).
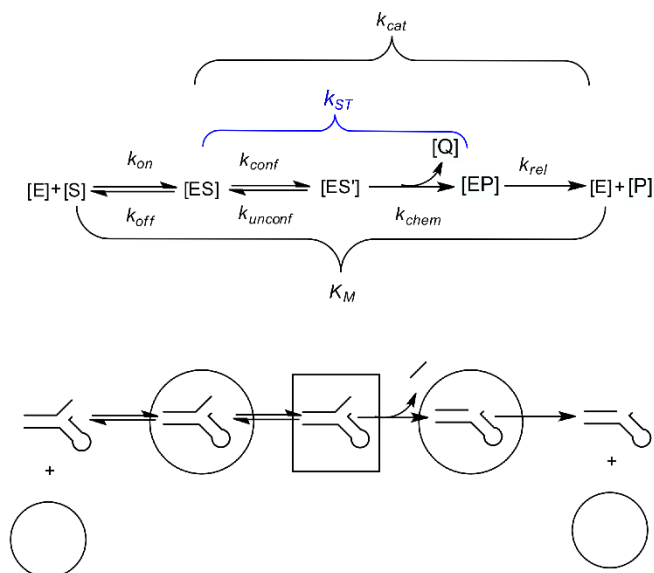
Figure 3.6: Schematic representation of the proposed reaction mechanism of FEN upon DF substrates. Top panel: *The relevant rate constants used are enzyme ($[E]$) and substrate ($[S]$) association (k_{on}), ES complex dissociation $\left(k_{off}\right)$, reversible conformational change $\left(k_{conf} \text{ and } k_{unconf}\right)$, the chemistry step ($k_{chem}$) and product release step (k_{rel}). The processes included in the Michaelis-Menten paratmeters (k_{cat} and K_M) are represented and the steps measured under single-turnover conditions (k_{ST}) is also indicated.* ***Bottom panel:*** *The enzyme represented as a circle, the alternative enzyme conformation is represented as a square, and the substrate is represented as sticks.*

The cleavage product was observed using denaturing HPLC equipped with a fluorescence detector (*figure 3.7*), which allowed separation of FAM labelled substrate and product P. Cleavage site was assessed by coinjection with oligonucleotide standards of known size (*figure 3.7*). In all cases, with DF-5, hFEN1 hydrolysed the phosphate diester located one nucleotide into the duplex of the 5'-flap strand to yield a 6 nt single stranded DNA (Q) and a 23 nt double-stranded DNA (P) (*figure 3.7*). The amount of product produced against incubation time of the reaction mixture was determined using *equation 3.4*, where *[Q]* is the total product concentration at a particular time point and *[S]* is the initial substrate concentration at time 0. $Q area$ represents the total area under the product peak and $S_0 area$ the same for the substrate peak.

$$[Q] = \left(\frac{Q_0 area}{Q_0 area + S_0 area}\right) \times [S] \qquad equation \ 3.4$$

Plots of percentage product (%Q) versus time were produced and the k_{ST} values were determined by fitting the data in *figure 3.8* to a single exponential model (*equation 3.5*) and are summarised in *table 3.2*.

$$Q_t = Q_\infty \left(1 - exp^{-k_{ST}t}\right) \qquad equation\ 3.5$$
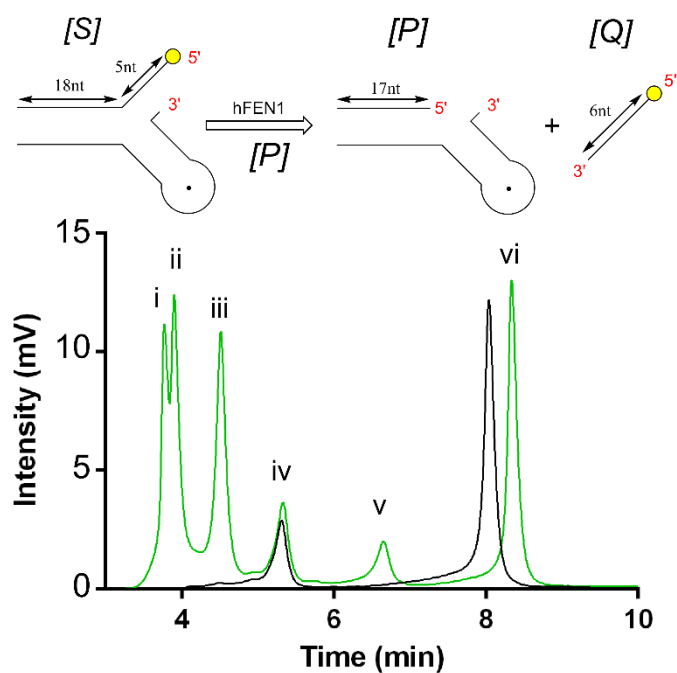
Figure 3.7: Monitoring hFEN1 catalysed cleavage of DF-5 by dHPLC*. Separation of oligonucleotide standards (green) indicating single nucleotide resolution can be achieved with the buffers and gradients used (section 2.4.1). Oligonucleotide standards (i) 1 mer, (ii) 2 mer, (iii) 4 mer, (iv) 6 mer, (v) 10 mer, (vi) 25 mer. Retention times: 1mer, 3.6 min, 2mer, 3.85 min, 4mer, 4.64 min, 6mer, 5.59 min, 10mer, 6.87 min, 25mer, 8.35 min. hFEN1 catalysed cleavage of DF-5 ([S]) produces ssDNA ([Q]) and dsDNA ([P]) products. Separation of products by dHPLC (black), indicating a single six nucleotide product (5.53 min) and twenty three nucleotide substrate (7.93 min), quantitative detection by fluorescently labelled substrate, with the yellow ball representing fluorescein.*

Table 3.2: Single-turnover reaction parameters of DF-5 hydrolysis. *Reactions were performed at 37 ˚C in 8 mM MgCl₂, 110 mM KCl, 55 mM HEPES-NaOH pH =7.5, 1 mM DTT and 0.1 mg/ml BSA. The single-turnover rate constants derived from equation 3.5 are reported with the standard errors in parentheses. The data in the right-hand column were taken from Tsutakawa, S.E., et al. (2011). Cell 145, 198-211. Accompanying graphs are shown in figure 3.8.*

Enzyme	k_{ST} (min^{-1})	$t_{1/2}$ (s)	Rate decrease relative to WThFEN1	Rate decrease relative to WThFEN1. K_{cat}/K_M conditions
WThFEN1	740 (± 32)	$5.6×10^{-2}$	1	1
K93A	$3.6×10^{-1}$ (± $2×10^{-2}$)	116	2060	400
R100A	$1.0×10^{-1}$ (± $4×10^{-3}$)	416	7400	400
Y40A	7.91 (± $1.9×10^{-1}$)	5.3	94	20
K93AR100A	$1.6×10^{-1}$ (± $1X10^{-2}$)	265	5700	N/A
D181A	$7.5×10^{-2}$ (± $3×10^{-3}$)	555	9900	800
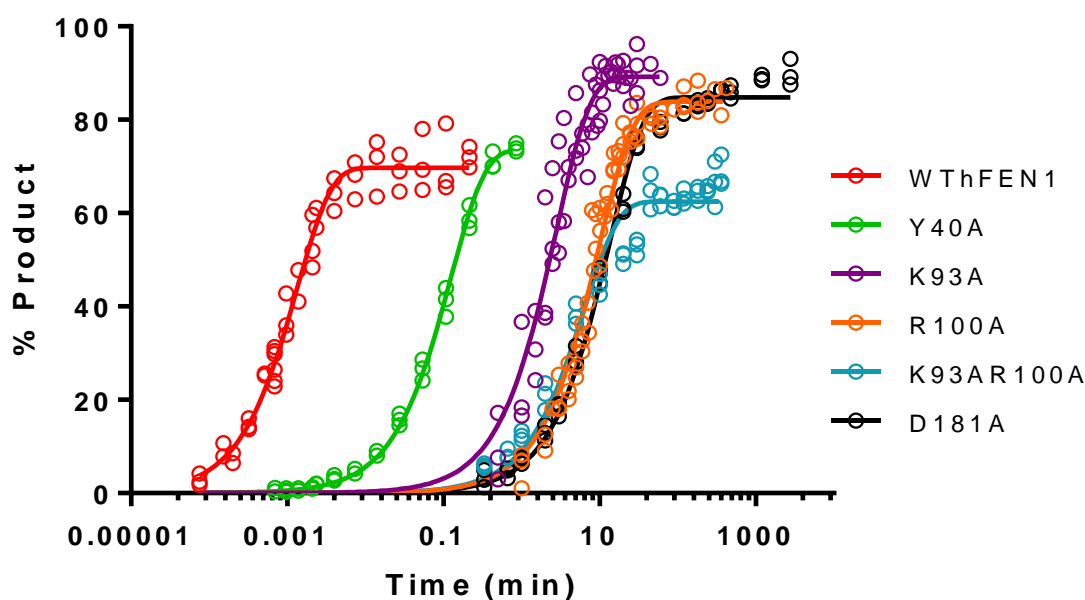
Figure 3.8: The single turnover rate profiles for DF-5 cleavage by hFEN1 and variant hFEN1 proteins. WThFEN1: k_{ST} = 740 ± 32 min^{-1}, $t_{1/2}$ = 5.6×10^{-2} s, data points-open circles (red), line of best fit- solid (red). **Y40A**: k_{ST} = 7.91 ± 1.9×10^{-1} min^{-1}, $t_{1/2}$ = 5.3 s, data points-open circles (green), line of best-fit- solid (green). **K93A**: k_{ST} = 3.6×10^{-1} ± 2×10^{-2} min^{-1}, $t_{1/2}$ = 116 s, data points-open circles (purple), line of best fit- solid (purple). **R100A**: k_{ST}= 1.0×10^{-1} ± 4×10^{-3} min^{-1}, $t_{1/2}$ = 416 s, data points-open circles (orange), line of best fit- solid (orange). **K93AR100A**: k_{ST} = 1.57×10^{-1} ± 1×10^{-2} min^{-1}, $t_{1/2}$ = 265 s, data points-open circles (cyan), line of best fit- solid (cyan). **D181A**: k_{ST} = 7.5×10^{-2} ± 3×10^{-3} min^{-1}, $t_{1/2}$ = 555 s, data points-open circles (black), line of best fit- solid (black). Reactions were performed in 8 mM MgCl₂, 110 mM KCl, 55 mM HEPES-NaOH pH = 7.5, 1 mM DTT and 0.1 µg/µl BSA and at 37 ˚C. Data were fitted to equation 3.5, and each single turnover rate constant is representative of at least three repeats.

Comparing the single turnover rates in *table 3.2* it is clear that the mutations introduced severely retard the decay of ES complex. As anticipated, the magnitude of reduction when mutations were introduced was always significantly greater than that observed under substrate-limiting multiple turnover conditions (*table 3.2*). Mutation of Tyr-40 to Ala was the least detrimental, reducing the k_{ST} value by approximately 100-fold compared to WT (*table 3.2*). There is a 2000-fold reduction in the maximal single turnover rate for K93A, and a 10000-fold reduction in the maximal single turnover rate for D181A with respect to the WT enzyme (*table 3.2*). The effect of mutating either Lys-93 or Arg-100 under steady-state conditions at limiting substrate concentrations is essentially indistinguishable (*table 3.2*). However, the single-turnover rate for R100A is 3-fold slower than for K93A (*table 3.2*). Despite K93A and R100A independently reducing the single-turnover rate considerably, the double variant K93AR100A is no more detrimental to catalysis than R100A alone (*table 3.2*). This implies cooperation between the basic residues of α4. For example, Arg-100 could be involved in the conformational step allowing Lys-93 to participate in the chemical step.

3.5 An alternative approach to measure substrate binding

3.5.1 Electromobility shift assays indicate the presence of higher order species.

One possible issue with the use of fluorescence anisotropy to determine equilibrium dissociation constants is that the signal will be affected by non-specific binding (Weinberg et al., 2004). Non-specific binding of FEN1 to a DF substrate could generate higher order complexes which would cause the r value to keep increasing beyond saturation with a single substrate molecule. This could result in a false (higher) K_D value when measured by FA. To examine this possibility, an electomobility shift assay (EMSA) was employed to investigate whether relatively high concentrations of FEN protein resulted in complexes where more than one enzyme was bound to the same substrate molecule through non-specific interactions.

To visualise low concentrations of the substrate without the need for radiolabeling, unbound and bound DNA was detected by chemiluminescence using a modified southern blotting approach described in detail in experimental methods *section 2.5.* The substrate used for this assay was a bi-molecular static DF construct which differed slightly from the substrate used for FA. Instead of a 5 nt flap, the substrate used for EMSA had a 10 nt flap that was biotinylated at the 5'-termius. This newly designed substrate was denoted as DF-B10 (*figure 3.9*). Biotinylation allowed the substrate to be visualised by chemiluminescence using a streptavidin-HRP conjugate. Work conducted by Nikesh Patel has shown that biotinyaltion of the 5'-flap does not change the binding interaction of hFEN1 and DF substrates (Patel 2012). In addition to the

increased 5'-flap length, the downstream and upstream duplex DNA of DF-B10 were longer than DF-5 to ensure efficient transfer of the substrate, especially when in an enzyme-substrate complex, from the gel onto the nylon membrane. Both DF-5 and DF-B10 substrates had a one nt 3'-flap and despite the slight difference in design were expected to display similar binding behaviour.
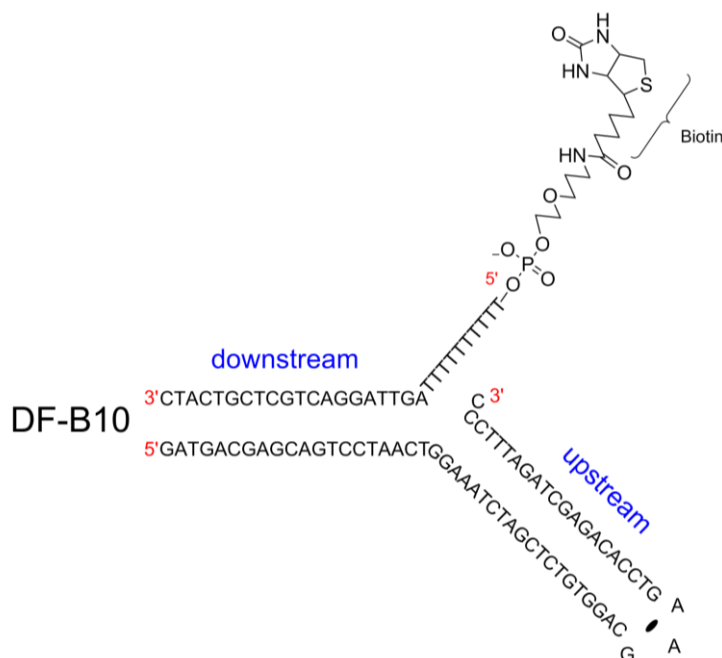
Figure 3.9: Schematic structure of the DFB-10 substrate, a bi-molecular construct prepared from a 64 nt template strand (T2) and 30 nt flap strand (F2). The template strand forms a fold-back hairpin stabilised by a GAA tri-loop and provides a 1 nt non-complementary 3´-flap. The flap strand, possessing a 5´-Biotin, anneals with the template strand and provides a 10 nt non-complementary 5´-flap. Together the two strands form a 20 nt ds upstream and 20 nt ds downstream region.

Comparison of binding constants determined by two different approaches required FA conditions to be replicated as faithfully as possible during EMSA experiments. Therefore, all buffer conditions were identical (37 ˚C, 110 mM KCl, 55 mM HEPES pH=7.5, 1 mM DTT, 0.1 µg/µl BSA and 10 mM CaCl$_2$), apart from it was necessary to have a final glycerol concentration of 5 % (w/v) in EMSA to facilitate gel loading. Initially, EMSA was performed at the same substrate concentration (10 nM) found to obtain a detectable signal in anisotropy and the protein concentration was varied (0 – 2500 nM) to reflect the titration points used in a typical FA experiment. The data clearly shows that separation of free substrate and bound substrate was achieved (*figure 3.10a*). The intensity of the band representing bound substrate increased

(middle band, *figure 3.10a*) as the as the enzyme concentration was increased. This species is proposed to be a 1:1 hFEN1:DF-B10 complex. However, evidence for a more slowly migrating species was observed with 125 nM hFEN1 (*figure 3.10a, lane 9*), and approximately 20 % of the total substrate is associated with more than one enzyme molecule at concentration greater than 500 nM hFEN1 (*figure 3.10a, lanes 11, 12 and 13*). The longer upstream and downstream duplex regions of the EMSA substrate compared to the one used in FA may encourage non-specific binding. However, due to the large amount of aberrant species detected, it is unlikely to be completely eliminated by the use of a more minimal substrate. The percentage of bound substrate, in what was assumed to be a stoichiometric enzyme:substrate complex, was fit to *equation 3.6*,

$$\% \, Bound = \frac{(\% \, Bound_{max} \cdot [protein])}{(K_D + [protein])} \qquad equation \; 3.6$$

A binding constant of 10 ± 3 nM was calculated based on a hyperbolic curve fit (*figure 3.10b*), in good agreement with previous values of K_D determined by EMSA for wild-type FEN1 (Gloor et al., 2010; Hutton et al., 2008). However, the binding constant measured by EMSA in this initial experiment was approximately equal to the substrate concentration. Therefore to further explore the binding affinity of hFEN1 for DF-B10 using EMSA, the experiment was repeated at 1 nM substrate. Both free and bound substrate were able to be detected at this lowered total substrate concentration (*figure 3.10c*) and fitting of the binding curves to *equation 3.6* yielded a dissociation constant of 6 ± 0.8 nM (*figure 3.10d*). The amount of band smearing was reduced compared to the experiment conduced at 10 nM DF-B10, and evidence for higher order species was detected, but only at the highest protein concentration at very low amounts (*figure 3.10c, lane 13*). The K_D values measured at a final substrate concentration of 10 nM and 1 nM are virtually undistinguishable. This suggests that the dissociation constant of approximately 10 nM determined for the binding of hFEN1 to DF-B10 is a true binding constant, and hFEN1 binds tightly in a structure-specific manner and weakly in a non-specific manner. Therefore, the binding constant for WThFEN1 determined by FA and EMSA differed by approximately 5-fold. The lower K_D value determined by EMSA may be because, unlike anisotropy, true enzyme substrate complex and aberrant complexes are separated. Therefore, K_D values measured by FA, as the assay stands, are more likely a mixture of specific and non-specific interactions. Attempts to abolish the non-specific binding during FA experiments, by increasing the ionic strength and introducing an unlabelled ssDNA mimic, heparin, were unsuccessful (data not shown). One possible solution to try and improve the FA assay based on these results would be

to lower the starting substrate concentration. This would require the design of a new substrate bearing a fluorophore with a higher quantum yield.
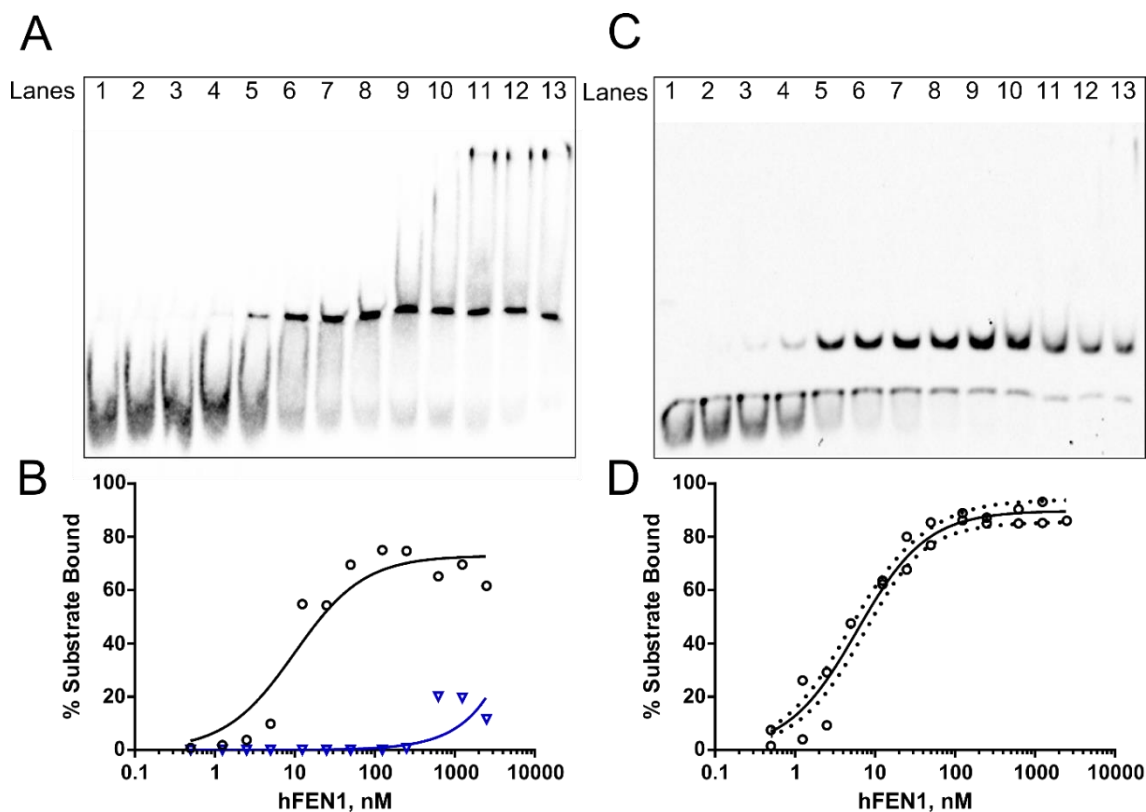
Figure 3.10: Discrimination between specific binding and non-specific binding by hFEN1 to a DF substrate. *FEN1 binding was measured by EMSA as described in section 2.5. Binding was measured by incubating increasing concentrations of hFEN1 with (**A**) 10 nM and (**C**) 1nM DF-B10 for 5 min in the presence of catalytically inert Ca^{2+} ions. The specific concentrations of hFEN1 were 0, 0.5, 1.25, 2.5, 5, 12.5, 25, 50, 125, 250, 625, 1250, and 2500 nM (lanes 1-13). Quantitation of the percent experimental substrate bound is shown in figures (B) and (D) for 10 nM and 1 nM, respectively. Suspected stoichiometric enzyme:substrate complex is indicated by open circles, black and higher order species by open inverted triangles, blue. The percentage of bound substrate was fit to equation 3.6 (black line).*

3.5.2 Using hFEN1-induced substrate conformation change to measure a true binding constant: A fluorescence resonance energy transfer analysis

As discussed above, a change in anisotropy can be produced by various interactions of hFEN1 with the DF substrate. Thus, the FA assay gives very little information on the binding mode of hFEN1. A 1:1 interaction of a FEN protein with the DNA in the conformation observed in the crystal structure with the two duplex regions bound will give the same signal in FA as FEN interacting with just one of these regions or, even though unlikely, with the 5'-flap. Therefore,

an assay which only gives a response when the substrate is bound by FEN in the correct conformation would yield a more reliable K_D value. In addition, it would be unlikely that this specific DNA conformational change could be produced by non-specific interactions and therefore these would not complicate binding analyses.

Fluorescence resonance energy transfer (FRET) involves a nonradiative transfer of energy from an excited state donor fluorophore to a nearby acceptor (Forster, 1948; Stryer and Haugland, 1967). The efficiency of energy transfer (FRET_{eff}) is directly related to the distance R separating a given donor and acceptor pair by the following equation,

$$FRET_{eff} \ = \ \frac{1}{[1 + (R/R_0)^6]}$$

R_0 is typically 10 -100 Å for donor-acceptor pairs and the specific value depends on the extent of overlap between the donor emission and the acceptor spectra, the absorption coefficient for the acceptor, the quantum yield of the donor, and the relative orientation of the donor and acceptor (Clegg, 1992; Fairclough and Cantor, 1978). Therefore, intuitive placement of an appropriate donor-acceptor pair within a DNA construct affords dynamic information on the possible conformations the constructs can adopt and their equilibrium position (Klostermeier and Millar, 2002; Yang and Millar, 1997). The conformation of DNA bound to FEN1 has previously been studied by comparing the efficiency of FRET between donor and acceptor molecules attached at various positions on a DF substrate (Chapados et al., 2004; Craggs et al., 2014; Sobhy et al., 2013). FRET has been used to confirm crystallographic data suggesting FENs enforce a kink at the ss-ds junction of DF structures, resulting in the two ds regions being bent 90-100° relative to each other. In this study, the donor and acceptor molecules were attached to the duplex arms of a DF substrate as it has been established that this gives a detectable change in FRET efficiency upon FEN1 binding.

The substrate used in this study was identical to that used in a previous study to determine the binding constant of *Saccharomyces cerevisiae* FEN (ScFEN) (Chapados et al., 2004). The substrate was tri-molecular, comprising of a 5'-flap strand, a 3'-flap strand and a template strand. The donor and acceptor molecules were placed on complementary strands allowing doubly- and singly-labelled constructs to be formed (*figure3.11*). As with all the substrates in this thesis this substrate was a static DF construct but this particular substrate had a 4 nt 5'-flap (DF-4) rather than the 5 nt and 10 nt flaps used previously. As mentioned above, the fluorophores were placed on the 5'-terminus of the DF ds regions; tetramethylrhodamine (TAMRA) was attached to the 5'-thymine of the template strand and FAM was attached to the 5'-adenine of 3'-flap strand (*figure*

3.11). The separation between the TAMRA and FAM is predicted to be approximately 88 Å when the duplex arms of the DF substrate are coaxially stacked, and approximately 64 Å when bound to FEN assuming a 100˚ angle (*figure 3.11*). When $R = R_0$, the FRET$_{eff}$ is halved and when $R = 2R_0$ no FRET is observable. R_0 for the TAMRA-FAM pair is approximately 55 Å. This value will change depending on linker length and sequence context. However, this demonstrates that these dyes are suitable for this assay.
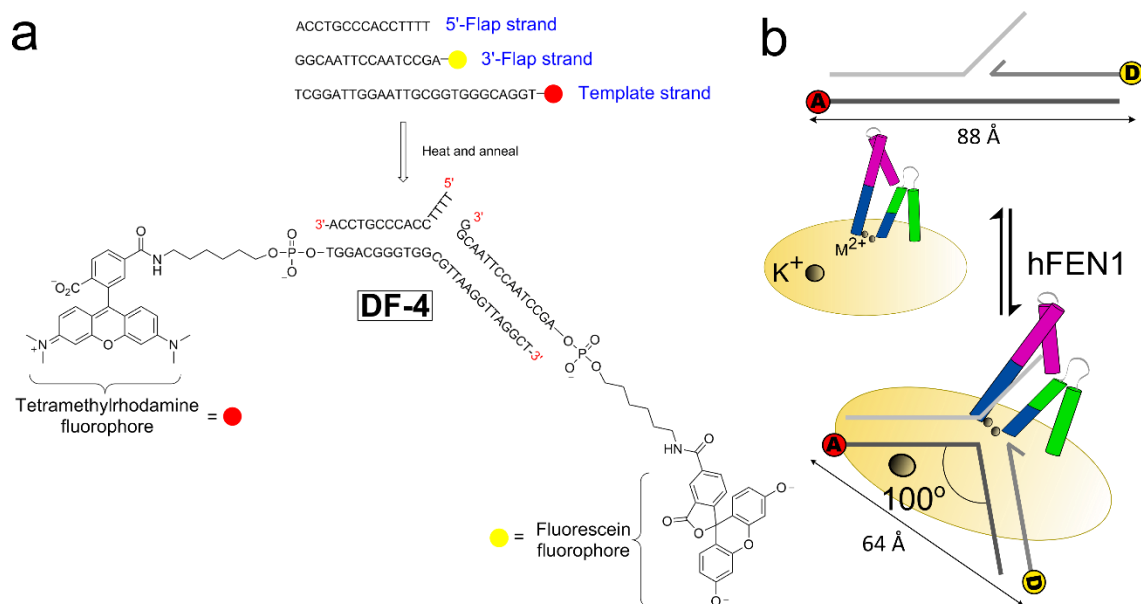
Figure 3.11: Schematic of substrate design for FRET measurements with hFEN1. (A) The substrate, DF-4, is tri-molecular formed from an unlabelled 5′-flap strand, a 3′-flap strand and a complemenatry template strand/ As depicted here, a doubly-labelled construct is prepared by annealing the 5′-flap strand with a 3′-flap strand labelled with a fluorescein (FAM) and a template strand labelled with a tetramethylrhodamine (TAMRA). Unlabelled versions of the 3′-flap strand and template strand were designed to allow the formation of singly-labelled substrate constructs, either with FAM only, or TAMRA only and completely unlabelled constructs (table 2.3). (B) When in B DNA the predicted distance from the donor (FAM) and the acceptor (TAMRA) is 88 Å, assuming a 100 bend of the substrate bound to hFEN1, this distance decreased to 64 Å. Therefore, an increase in the FRET efficiency should be observed upon FEN addition.

The efficiency of energy transferred (FRET$_{eff}$) was measured from the induced emission of the acceptor, and calculated using the (ratio)$_A$ method (Clegg, 1992). This allows the FRET efficiency to be determined from the fluorescent spectra of a donor-only labelled molecule and a donor-acceptor labelled molecule at two appropriate wavelength pairs. In this study, to account for background noise and any artefacts associated with specific buffer conditions, additional spectra

of an unlabelled construct were measured with excitation at the donor and acceptor wavelengths and subtracted from the appropriate spectra of labelled constructs.

Steady-state fluorescence measurements were made by exciting the donor (FAM) at 490 nm and observing the emission at 510 – 650 nm, and exciting the acceptor (TAMRA) at 565 nm and observing the emission at 570 nm – 650 nm. Because the excitation and emission spectra of TAMRA and FAM cannot be separated completely, the acceptor emission (λ_{EM}^A) of a doubly-labelled molecule (DA) following excitation of the donor (λ_{EX}^D), the FRET signal($F_{DA}(\lambda_{EX}^D, \lambda_{EM}^A)$), is actually a composite of three fluorescent processes:

(1) Emission from the acceptor due to FRET
(2) Acceptor emission from direct excitation
(3) Donor emission from direct excitation

Therefore, the contribution of the donor emission must be subtracted from the FRET signal. This is achieved by measuring the emission of the donor in the context of a donor only-labelled construct, when excited at the donor wavelength ($F_D \lambda_{EX}^D, \lambda_{EM}^D$). This step is necessary because although FRET measurement only applies to weakly coupled fluorophores, there will be some quenching of the donor excitation due to the proximity of the acceptor. Prior to subtraction, the donor emission must be normalised to match the donor peak in the doubly labelled molecule. The extracted spectrum, consisting of purely acceptor emission, from both FRET and direct excitation, is also normalized by the maximum emission of the acceptor when directly excited in a doubly-labelled molecule $F_{DA}(\lambda_{EX}^A, \lambda_{EM}^A)$. This normalization gives a measure of the total number of acceptor molecules in the sample. These steps are used to calculate (ratio)$_A$ are summarized in *equation 3.7*,

$$(ratio)_A = \frac{F_{DA}(\lambda_{EX}^D, \lambda_{EM}^A) - N \cdot F_D \lambda_{EX}^D, \lambda_{EM}^D}{F_{DA}(\lambda_{EX}^A, \lambda_{EM}^A)} \qquad equation\ 3.7$$

where *N* is the normalization factor ($N = F_{DA}(\lambda_{EX}^D, \lambda_{EM}^D)/F_D \lambda_{EX}^D, \lambda_{EM}^D$).

The FRET$_{eff}$ is functionally dependent on (ratio)$_A$ and the extinction coefficients of the donor and acceptor molecules at the excitation and emission wavelengths as shown in *equation 3.8*,

$$(ratio)_A = FRET_{eff} \cdot \frac{\varepsilon^D(490)}{\varepsilon^A(560)} + \frac{\varepsilon^A(490)}{\varepsilon^A(560)} \qquad equation\ 3.8$$

, where ε^D and ε^A are the molar absorption extinction coefficients at the indicated wavelengths of the donor and acceptor, respectively.

The local environment of a protein surface can affect the fluorescence intensity of the donor and/or the acceptor (Steinmetzer et al., 2002). The potential sensitization of the acceptor fluorescence induced by protein binding is accounted for in the (ratio)$_A$ method. For the donor this same effect was accounted for by a simultaneous titration of FEN protein into a donor-only labelled construct and as described above this can then be subtracted from the FRET response at each individual protein concentration.

FRET measurements can be conducted without the need for changes in the acceptor and donor quantum yield to be calculated, as these do not interfere with the calculation of the FRET efficiency using the (ratio)$_A$ method (Craggs et al., 2014). In addition, the percentage labelling of the acceptor does not enter into the equation due to the normalization procedure (Clegg, 1992). However, the errors in the percentage labelling of the donor directly affect the specific value of FRET$_{eff}$, but not the measurement (Clegg et al., 1994). Therefore, quantitative comparison of FRET$_{eff}$ values requires the percentage labelling of the donor to be known. Because the main focus of this investigation was not to calculate dye-to-dye distances but use the signal change to determine binding constants of WThFEN1 and hFEN1 variants, this method allows for a relatively straightforward approach to determine FRET efficiencies.

The data observed with DF-4 shown in *figure 3.12*, reported on the change in kink angle between the two duplex arms. FEN1 binding decreases this angle, causing a decrease in the average end-to-end distance of the substrate indicated by an increase in the observed FRET$_{eff}$ (*figure 3.12*). The FRET$_{eff}$ values measured were in good agreement with the previous study using the same oligonucleotide construct (*figure 3.12*).
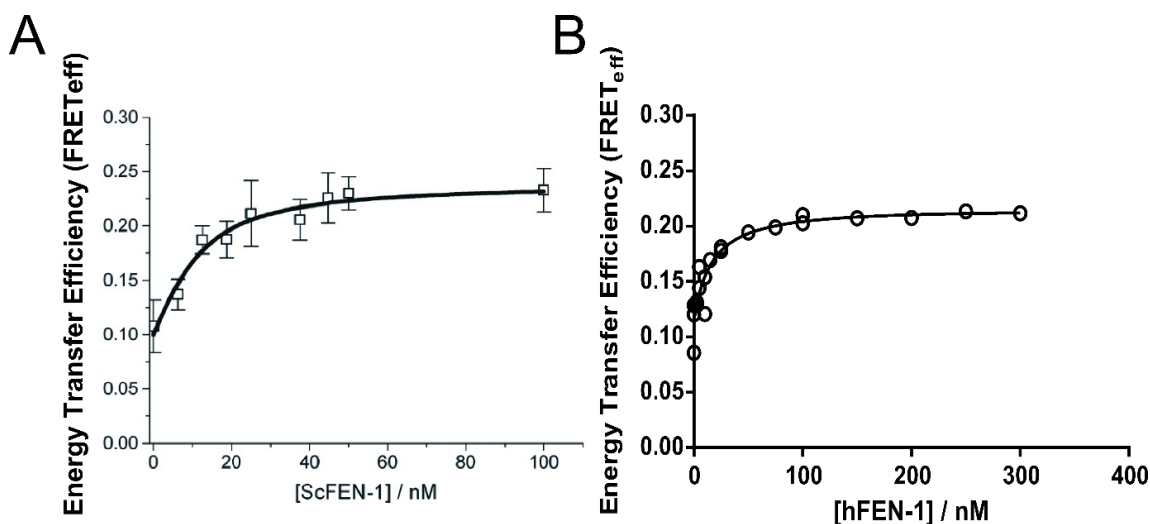
Figure 3.12: Comparison of FEN1 induced FRET efficiency changes of a DF substrate bearing fluorophores on the duplex arms by, (A) Saccharomyces cerevisiae FEN (ScFEN) (taken from Chapados, B.R., et al. (2004). Cell 116, 39-50) and (B) human FEN (hFEN1). Measurements were performed in different buffer conditions, in the absence of metal ions in (A) and in the presence of 10 mM Ca^{2+} in (B).

3.5.3 Unbound substrate FRET measurements

Previously reported evidence suggests that DF-DNA structures are in equilibrium between coaxially stacked and bent conformations, and this equilibrium can be shifted depending on the buffer conditions (Craggs et al., 2014). Here, an assessment of salt variation with intra-molecular $FRET_{eff}$-values shows that the distance between the two duplex regions of a DF-substrate decrease with increasing ionic strength (*figure 3.13a*). Conversely, changes in magnesium ion concentration at a constant ionic strength had no detectable effect on the conformation of the DF construct (*figure 3.13b*). The tri-molecular substrate is predicted to have poor stability at the lowest salt concentrations, thus changes in $FRET_{eff}$ may also be measuring substrate annealing effects. However, at high salt concentrations (100 mM KCl and above) at which the substrate is more stable, an increase in $FRET_{eff}$ is still detected suggesting a conformational response to ionic strength.

It was noticed that the starting $FRET_{eff}$-value, for unbound DNA, varied over the course of this study. Therefore, to compare different titrations, the $FRET_{eff}$ was normalised to the maximum $FRET_{eff}$-value for that particular titration. Individual titrations were fit to the quadratic binding equation (*equation 2.9*), and the endpoint was then used to normalise the energy transfer efficiency value so that replicates could be plotted on the same graph. Replicates were then fit

to the same binding equation using the newly generated normalised energy transfer efficiency values. A schematic representation of the normalization procedure is provided in *figure 3.14*. The binding constants derived from the individual fits of raw FRET$_{eff}$-values and collective fits of normalized FRET$_{eff}$-values were in good agreement (data not shown).
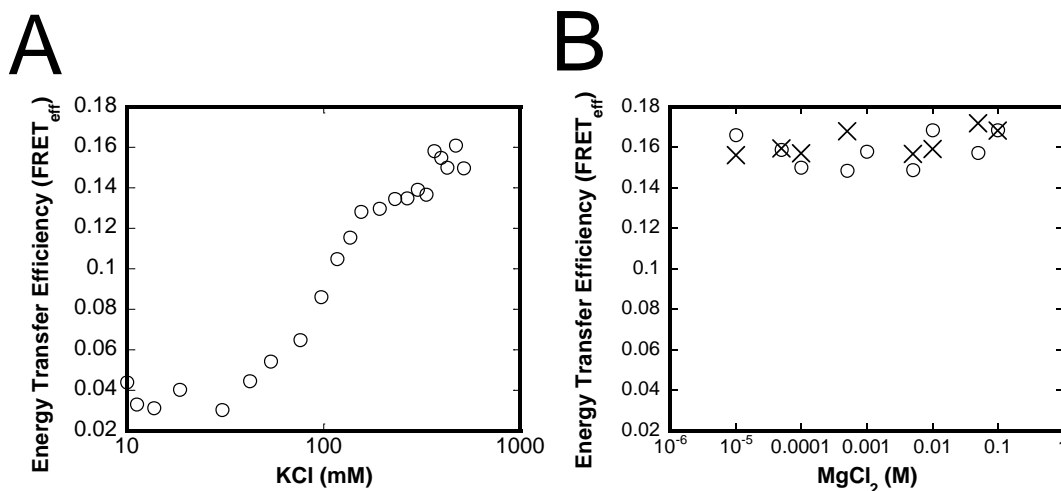
Figure 3.13: FRET changes of DF substrate, *in the presence of increasing (**A**) KCl and (**B**) Mg^{2+} ion concentration. (A) FRET$_{eff}$ of 10 nM DF-4 measured upon addition of KCl in continuous titration, starting buffer conditions were 10 mM KCl, 50 mM HEPES-NaOH, pH=7.5, 1 mg/ml BSA and 1 mM DTT. (B) FRET$_{eff}$ of 10 nM DF-4 measured at various Mg^{2+} ion concentrations in a discontinuous manner for two replicates (open circles and crosses). Apart from the indicated MgCl$_2$ concentration the buffer conditions were, 100 mM KCl, 50 mM HEPES-NaOH, pH=7.5, 1 mg/ml BSA and 1 mM DTT.*
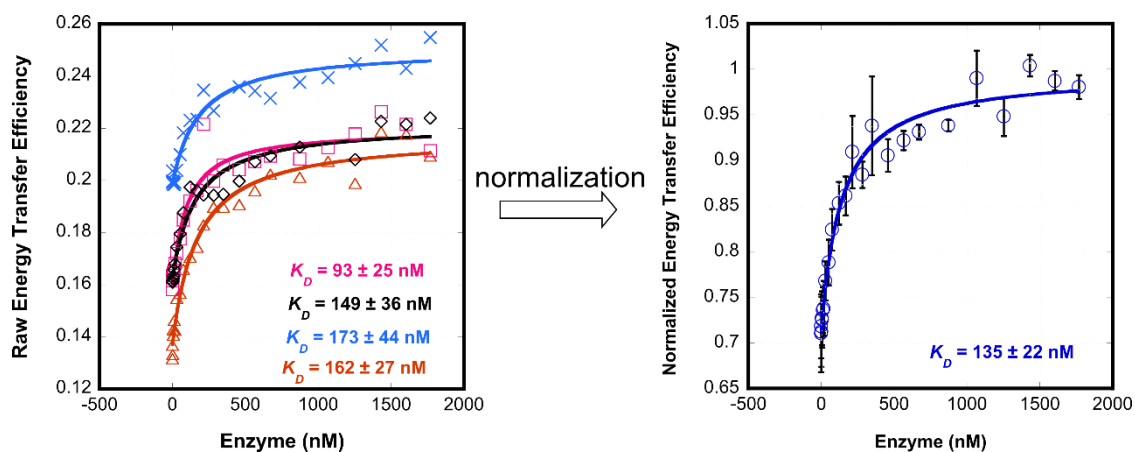
Figure 3.14: Schematic representation of normalization procedure employed for replicate of FRET titration curves. *The steps involved are described in the main body of the text.*

3.5.4 hFEN1-induced substrate conformational change

The dissociation constants derived from FRET measurements have been termed $K_{D-bending}$ to discriminate them from previous dissociation constants reported in this chapter. For WThFEN1, $K_{D-bending}$ was 20 ± 3 nM in the presence of Ca^{2+} ions (*table 3.3*). This is consistent with the value of 14 ± 3 nM reported for *Sulfolobus solfataricus* FEN using similar ensemble FRET experiments (Craggs et al., 2014). Analogous to the binding constant determined by EMSA, this value is lower than the value obtained from FA by approximately 3-fold. This suggests that a binding constant determined by a FEN-induced FRET change may be just that, true binding constants, and a more reliable approach for measuring and comparing K_D values. It should be pointed out that the phrase FEN-induced might be misleading, as it suggests that the increase in $FRET_{eff}$ is only brought about by FEN binding (*figure 3.15a*). A recent single-molecule (sm) study on DF substrates established that they are able to sample a range of conformers, including the bent conformation observed in the FEN:substrate crystal structure. The reported evidence implied that FEN binding simply stabilises rather than induces the bent conformation (Craggs et al., 2014).

The substrate is still capable of bending when bound to hFEN1 containing no metal ions in the active site (*figure 3.15b*). The dissociation constant measured for hFEN1 binding to DF-4 in buffer containing EDTA was 59 ± 11 nM (*table 3.3*). The dissociation constants measured by FA were identical in Ca^{2+} and EDTA. In contrast, there is a 3-fold difference in the binding affinity of hFEN1 for DF-4 when measured by FRET. Previously published sm-FRET and gel-shift data, studying the effects of flap length on the dissociation constant (Gloor et al., 2010; Sobhy et al., 2013), both reported lower K_D values for longer flap substrates, and rationalised this phenomenon by suggesting that the dissociation rate was slower for longer flap substrates as it is more difficult for a threaded flap to withdraw from the helical arch. In addition, the same sm-FRET study and a complementary sm-FRET investigation, measured slow dissociation rate constants after a conformational change (suspected to be threading) (Craggs et al., 2014; Sobhy et al., 2013). The slightly higher $K_{D-bending}$ value in EDTA may be due to a faster dissociation rate, as evidence suggests the threading process is dependent on metal ions (Patel, 2012). Therefore, in EDTA the substrate might dissociate more easily because the 5'-flap is not 'trapped' under the helical arch. The same trend is observed for R100A (*table 3.3*).

Table 3.3: $K_{D-bending}$ parameters for 10 nM DF-4 bound to WThFEN1, Y40A, K93A, and R100A determined by FRET. *All measurements were made at 37 ˚C in 110 mM KCl, 55 mM HEPES-NaOH pH = 7.5, 1 mM DTT, 0.1 mg/ml BSA and 10 mM Ca^{2+} or 2 mM EDTA as indicated. The dissociation constants derived from equation 3.3 are reported with the standard errors in parentheses. The change in raw $FRET_{eff}$-values from triplicate experiments are reported with the standard errors in parentheses. [a]n.d. = not determined. Accompanying graphs are shown in figure 3.15.*

Substrate	DF-4							
	4 nt 5'-flap, 1 nt 3'-flap							
Enzyme	**WT**		**Y40A**		**K93A**		**R100A**	
Buffer	Ca^{2+}	EDTA	Ca^{2+}	EDTA	Ca^{2+}	EDTA	Ca^{2+}	EDTA
$K_{D-bending}$ (nM)	20 (± 3)	59 (± 11)	70 (± 10)	n.d.[a]	71 (± 11)	n.d.[a]	135 (± 22)	289 (± 84)
$\Delta FRET_{eff}$	0.11 (± 0.002)	0.05 (± 0.01)	0.09 (± 0.01)	n.d.[a]	0.13 (± 0.01)	n.d.[a]	0.07 (± 0.01)	0.09 (± 0.01)
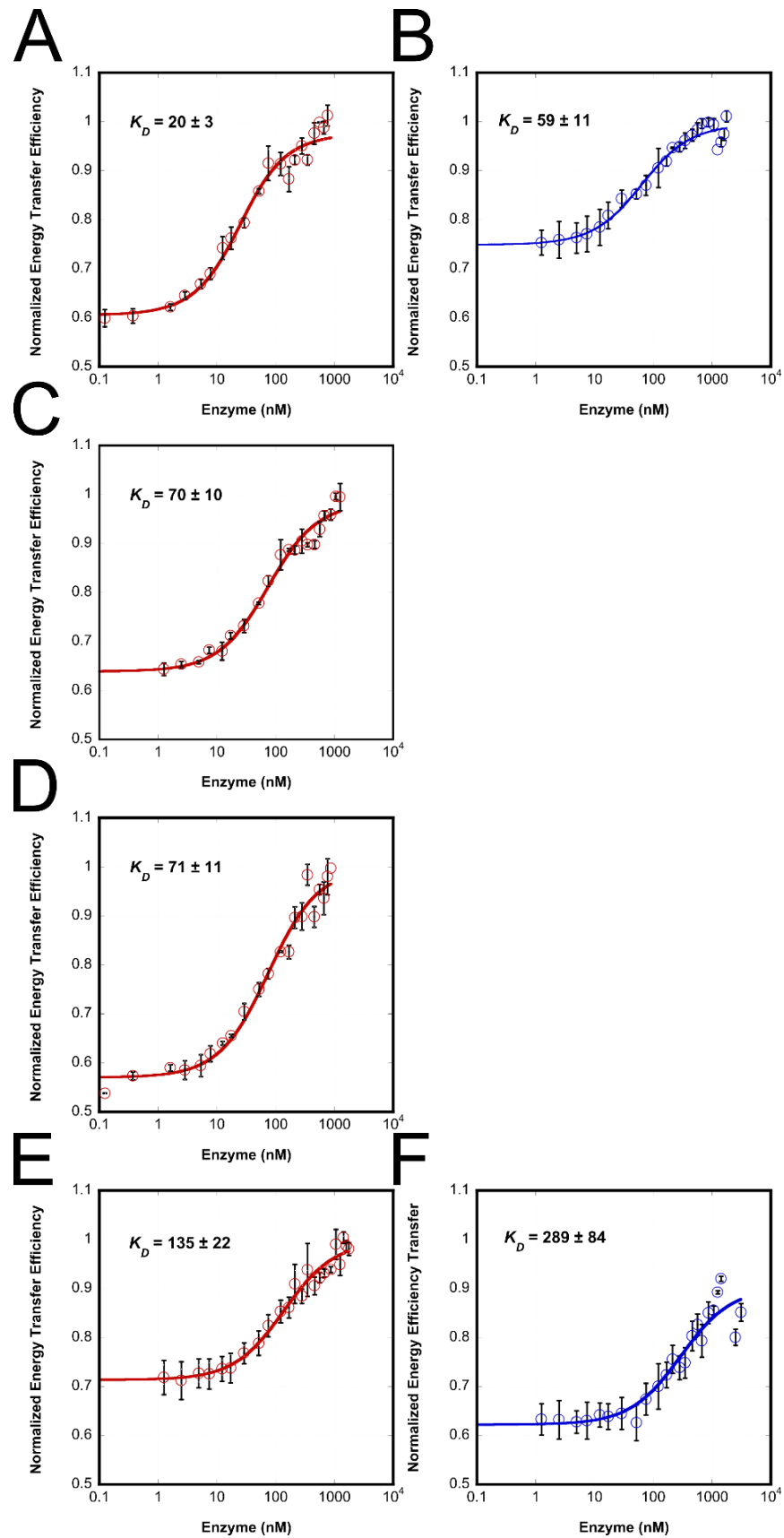
Figure 3.15: DF-4 binding curves for (A) hFEN1 / 10 mM CaCl₂, (B) hFEN1 / 2 mM EDTA, (C) hFEN1Y40A / 10 mM CaCl₂, (D) hFEN1K93A / 10 mM CaCl₂, (E) hFEN1R100A / 10 mM CaCl₂, (F) hFEN1R100A / 2 mM EDTA. *Apart from the specified conditions, assays were performed in 110 mM KCl, 55 mM HEPESpH=7.5, 0.1 μg/μl BSA, 1 mM DTT and at 37 ˚C. The data points were fit to equation 2.9. The K_D value for each condition is shown in the insert of the relevant graph, each K_D is representative of at least three repeats, and errors shown are standard error.*

115

The interpretation of $K_{D-bending}$ values can be simplified if it is assumed that the substrate is in the same conformation when bound to WThFEN1 and mutated FENs. The limitations of the assay as it stands means it is difficult to compare the raw FRET$_{eff}$-values at protein saturation. The change in raw FRET$_{eff}$-values (from unbound to saturation) was similar for all proteins tested (*table 3.3*), suggesting that the substrate is in the same conformation when bound to WThFEN1 and mutated FENs. Therefore, it is reasonable to assume that the assay is reporting on the affinity for the substrate in the same state/conformation. Resolution of the problems associated with the assay would allow this type of analysis with ensemble FRET measurements. However, being confident about the end-point value is not necessarily very helpful as the percentage of the population that is giving that specific signal is unknown in bulk-average / steady state FRET measurements, sm techniques would be required to probe the heterogeneity (Rothwell et al., 2003). Single molecule experiments would be very informative for this type of analysis as they could report not only the range of end-to-end distances for each protein but also the population of each state.

3.5.5 Mutated FENs bring about a substrate conformational change

The DNA conformational-dependent FRET response suggests that this is more of a functional assay than either gel-shift or anisotropy assays as initial association and substrate manipulation are examined simultaneously. Therefore, this technique lends itself to investigating the effect of single-point mutations on the ability of the substrate to be in a bent conformation when bound to variant FENs. On addition of each of the mutated FENs tested, Y40A, K93A and R100A, the DF-4 flap construct exhibited a significant increase in the efficiency of energy transferred indicative of a global reorganization of the flap structure to a bent conformation (*figure 3.15*). The $K_{D-bending}$ values for Y40A, K93A, and R100A were 70 ± 10, 71 ± 11, and 135 ± 22 nM, respectively. These values are essentially identical within experimental error. However, unlike anisotropy results, all three variants display slightly weaker binding than WThFEN1 (*table 3.3*). Previously it was believed that the binding affinity of these particular mutations was comparable to WTFEN (Patel et al., 2013; Qiu et al., 2004). These data suggest that this may not be the case, and when dissociation constants have been determined previously the nuances of mutated FENs have been disguised in the 'signal-change' by non-specific binding. Hence, the data presented here suggests that this approach is an improvement on previous ensemble-binding measurements and if extended to other variants could provide useful information to help gain insights into binding mechanism of FEN1.

3.5.6 Assay development

The efficiency of energy transfer is dependent on the orientation of the fluorophores (Dale et al., 1979). If the average orientation of the fluorophores changes in the absence and presence of protein then the assumptions of the (ratio)$_A$ method do not hold. The mobility of the dyes can be determined indirectly from their anisotropy.

Donor and acceptor steady-state anisotropy values of singly-labelled constructs were measured over the course of a typical titration. TAMRA attached to the terminus of the downstream duplex has an anisotropy of 0.13 (*figure 3.16a*). In contrast, FAM attached to the terminus of the upstream duplex has an anisotropy of 0.04 (*figure 3.16b*), suggesting that fluorescein must undergo considerably faster rotational diffusion than rhodamine (the maximum anisotropy possible for both dyes is almost 0.4) (Clegg et al., 1992). The anisotropy for the TAMRA-only labelled substrate increases with increasing hFEN1 concentration, to a maximum anisotropy of 0.19, and when fit to the quadratic binding equation (*equation 3.3*) yields a reasonable K_D-value of 29 ± 4 nM (*figure 3.16a*). Whereas the anisotropy of the FAM-only labelled substrate construct does not plateau during the course of the titration, FAM bound to the 5'-flap of a similar construct has anisotropy of 0.12 at saturation (*section 3.3*). This implies that binding of hFEN1 to the donor-only labelled substrate construct restricts the rotation of the FAM fluorophore itself, most likely be means of a non-specific interaction with the FAM.

Another issue with this particular construct was that the change in energy transfer efficiency was relatively small on transition for unbound to the bound state ($\Delta FRET_{eff} = 0.05 - 0.11$). This meant that the data were relatively noisy and one of the reasons why normalization was necessary. The FRET response is linear between 0.5 and 1.5R, and increases considerably when the R_0 distance is crossed (not true for these fluorophores on this construct). Therefore, a new substrate (DF-B6) was designed with these two points in mind (*figure 3.17*) (personal communication: Tim Craggs). The substrate was designed using the same FRET-pair. The fluorophores were modelled into B-DNA attached to an internal thymine at the C6 position so that they would extend out into the major groove (*figure 3.17*). The stretches of B-DNA were then modelled onto the upstream and downstream regions of the hFEN1:product structure so that they would be outside the limits of the hFEN1 protein surface
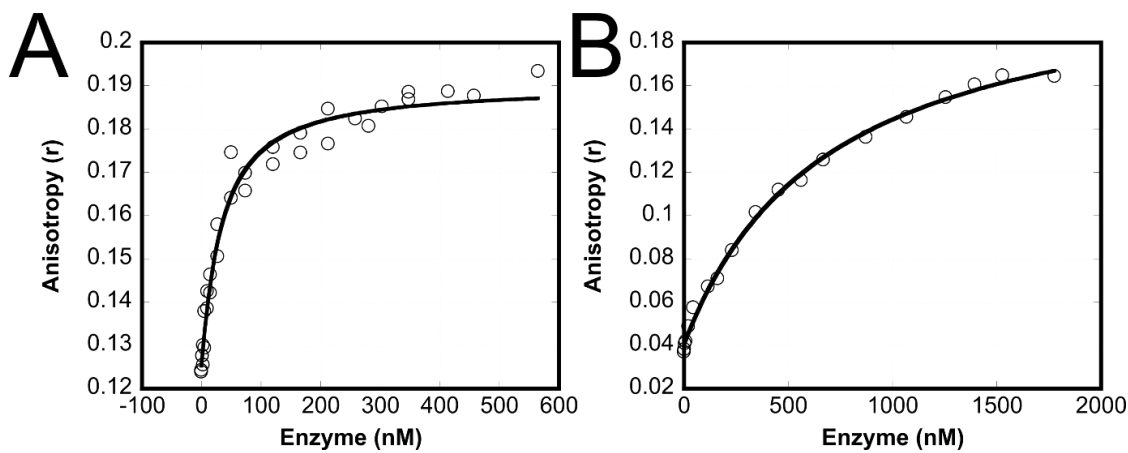
Figure 3.16: Anisotropy changes for two variations of the DF-4 substrate (A) the acceptor-only construct (TAMRA) and (B) the donor-only construct (FAM) with increasing hFEN1 concentration. Assays were performed in 10 mM CaCl₂, 110 mM KCl, 55 mM HEPESpH=7.5, 0.1 µg/µl BSA, 1 mM DTT and at 37 ˚C. The data points were fit to equation 3.3.
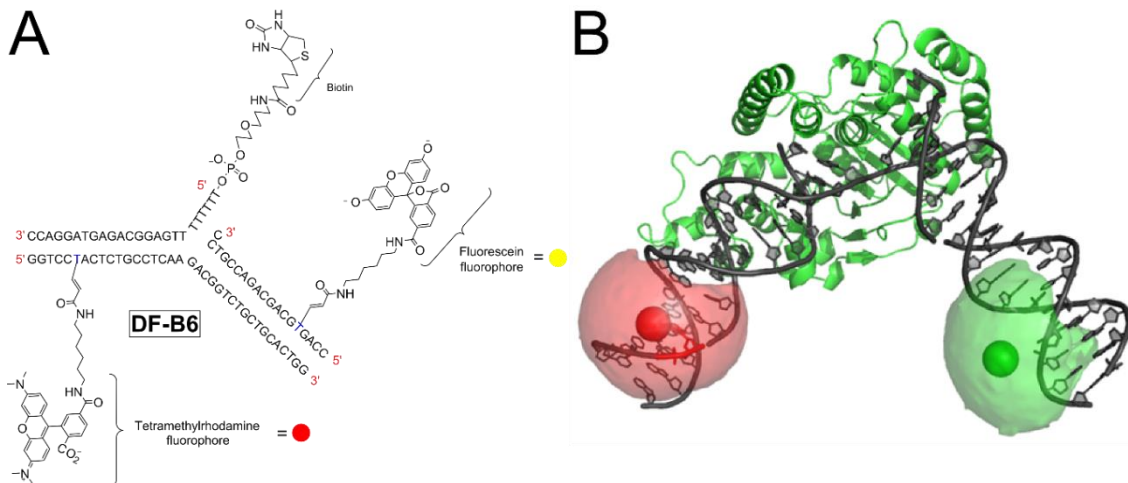
Figure 3.17: Model of DF-B6 FRET construct bound to FEN, with FAM and TAMRA attached to C6 of thymine (blue T) in upstream and downstream duplex regions, respectively. Assuming $R_0 = 55$ A and 100° angle when bound to FEN should produce a FRET efficiency change of 0.34. Courtesy of Tim Craggs.

As predicted from the model, repetition of the anisotropy experiments using the singly-labelled constructs showed that both the donor-only and acceptor-only labelled substrate produced a hyperbolic response upon the addition of hFEN1. This suggests that both of the fluorophores had free rotational freedom when the protein was bound to the DNA in the correct conformation (*figure 3.18*). In addition, as expected, the change in FRET$_{eff}$ was significantly increased from the unbound to the saturated state (*figure 3.19*). Initial experiments supported

the results collected with the older FRET substrate but the data was considerably less noisy due to the larger change in signal (*figure 3.19*).
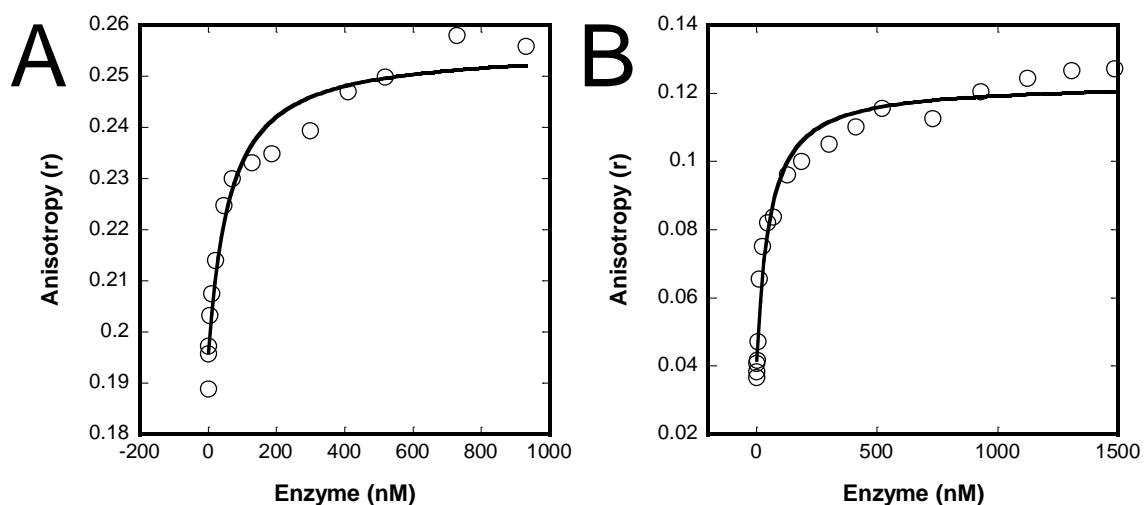
Figure 3.18: Anisotropy changes for two variations of the DF-B6 substrate (A) the acceptor-only construct (TAMRA) 56 ± 14 nM and (B) the donor-only construct (FAM) 48 ± 10 nM with increasing hFEN1 concentration. *Assays were performed in 10 mM CaCl$_2$, 110 mM KCl, 55 mM HEPESpH=7.5, 0.1 µg/µl BSA, 1 mM DTT and at 37 ˚C. The data points were fit to equation 3.3.*
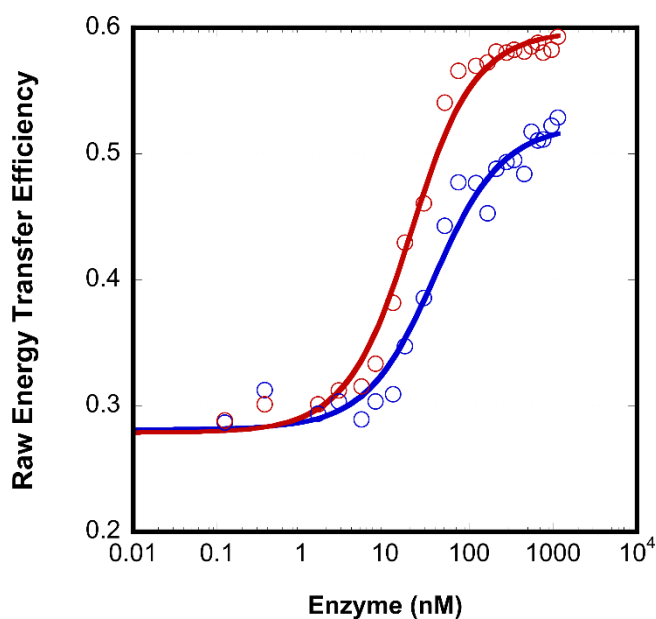
Figure 3.19: DF-B6 binding curves with wild-type hFEN1. *Red data points in 10 mM Ca^{2+} and blue data points in 2 mM EDTA. The data points were fit to equation 2.9 to yield respective K$_D$ values of 16 ± 2 and 34 ± 7 nM for Ca^{2+} and EDTA, respectively. Apart from the specified conditions, assays were performed in 110 mM KCl, 55 mM HEPESpH=7.5, 0.1 µg/µl BSA, 1 mM DTT and at 37 ˚C. Spectra were collected by SIA and EJ.*

3.6 Summary and discussions

The importance of the residues investigated is inferred from their conservation throughout evolution and across the hFEN1-superfamily. Previous biochemical data revealed that as expected these mutations were detrimental towards catalysis, but the work detailed in this chapter showed that their removal has a much greater effect that initially appreciated. Mutation of the stacking residue, Tyr-40 is the least detrimental towards catalysis. Crystallographic studies have suggested that this residue stacks with the unpaired nucleotide of the product, but it may not be important for stabilisation of the unpaired substrate. This is investigated further in *chapter 4*. Here, creation of a double-variant (K93AR100A) allowed the relationship between Lys-93 and Arg-100 to be investigated. Mutation of each of these residues reduces the activity 2000 and 7500-fold, respectively, compared to WThFEN1. However, the double-variant had the same level of activity as the single R100A variant. Together, these data suggest that there is likely a concerted involvement of Lys-93 and Arg-100 in the hFEN1 reaction mechanism. Further evidence is presented in *chapter 4*. Unsurprisingly, the most detrimental mutation studied in this chapter was D181A, emphasising the importance of divalent metal ions in the catalysis of phosphate diester hydrolysis by hFEN1.

Collectively, the different binding assays described in this chapter all suggest that hFEN1 has nanomolar affinity for DF substrates, as summarized in *table 3.4*. Although differences were seen depending on which biophysical approach was used, the affinities measured were all of the same order of magnitude and in accordance with previously reported dissociation constants. Additionally, despite differences in the length of the duplex regions and fluorophore positions between the two FRET substrates, DF-4 and DF-B6 used in this study, the dissociation constants ($\pm$ Ca^{2+}) for hFEN1 were very similar; giving validation to this type of assay for measuring FEN1 binding. The data suggests that the reduced catalytic rate of the variants studied is not a consequence of impaired substrate binding affinity, but it does suggest that there are subtle differences between themselves and the wild-type counterpart. The data presented here suggests that the best approach in the future to investigate these effects would be to use a FRET assay.

Interestingly, although divalent metal ions bound in the active site have no effect on overall affinity, their presence may determine the conformation of DNA much further away from the active-site. Preliminary data with the new FRET substrates suggest that the substrate may not bend as much in EDTA (lower FRET end-point). Understanding of this difference may have further implications as to the order of substrate manipulations (bending/unpairing), performed

by hFEN1. Nevertheless it is clear from the data obtained to date that bending can occur without active site metal ions (in EDTA), but whether there is a greater reduction in the duplex angle in their presence deserves further investigation.

Table 3.4: Summary of K_D parameters and standard errors determined by fluorescence anisotropy, electromobility shift assay and fluorescence resonance energy transfer for WThFEN1 and mutated FENs. All measurements were made at 37 ˚C in 110 mM KCl, 55 mM HEPES-NaOH pH = 7.5, 1 mM DTT, 0.1 mg/ml BSA and 10 mM Ca^{2+} or 2 mM EDTA as indicated. The dissociation constants derived from equation 3.3 are reported with the standard errors in parentheses. The minimum and maximum anisotropy values, r_{min} and r_{max}, respectively, for triplicate experiments are reported with the standard errors in parentheses. n.d. = not determined.

	Anisotropy		EMSA		FRET	
Substrate Code	**DF-5**		**DF-B10**		**DF-4**	
Dissociation constant	K_D		K_D		$K_{D-bending}$	
Enzyme	Ca^{2+}	EDTA	Ca^{2+}	EDTA	Ca^{2+}	EDTA
WT	46 (± 6)	50 (± 4)	10 (± 3)	n.d.	20 (± 3)	59 (± 11)
Y40A	175 (± 25)	n.d.	n.d.	n.d.	70 (± 10)	n.d.
K93A	58 (± 8)	93 (± 6)	n.d.	n.d.	71 (± 11)	n.d..
R100A	92 (± 10)	111 (± 8)	n.d.	n.d.	135 (± 22)	289 (± 84)
K93AR100A	78 (± 8)	106 (± 9)	n.d.	n.d.	*n.d.*	n.d.
D181A	180 (± 32)	n.d.	n.d.	n.d.	*n.d.*	n.d.

Chapter 4: Investigating local conformation changes of nucleotides in the human flap endonuclease active site by low energy circular dichroism of 2-aminopurine

4.1 5'-structure specific nuclease mechanism

As discussed in the introduction (*section 1.4*), the current double nucleotide unpairing (DNU) mechanism proposed for hFEN1 is largely based on structural data (Devos et al., 2007; Tsutakawa et al., 2011). Comparison of the substrate and product crystal structures (*figure 4.1*) suggested that there was likely an enzyme-substrate intermediate in which the first two nucleotides of the duplex region adjacent to the 5'-flap were extra-helical. The use of spectroscopic techniques to directly observe this proposed intermediate is discussed.
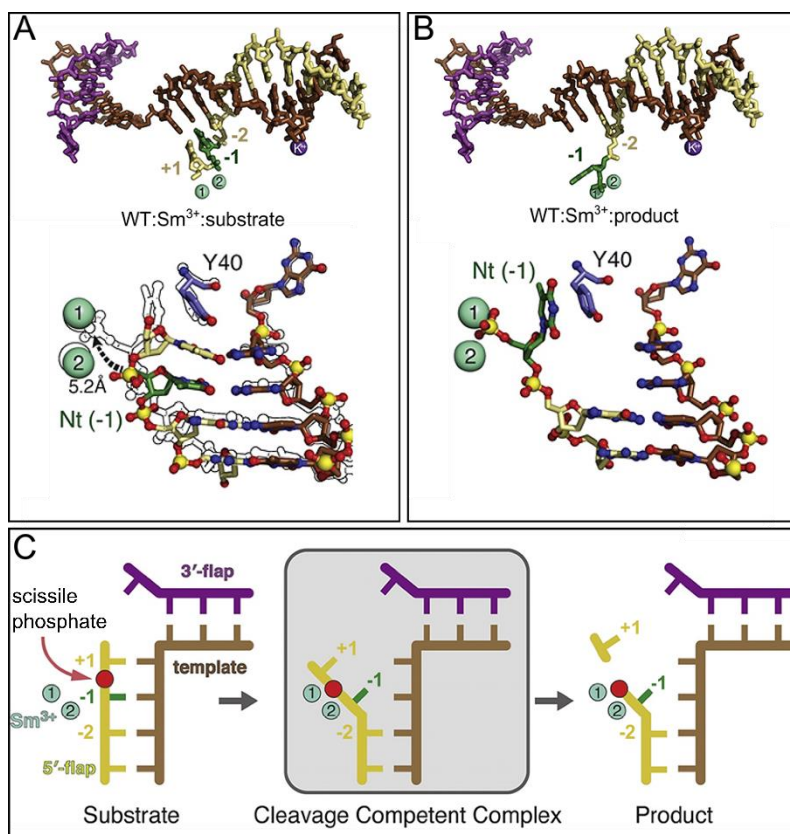
Figure 4.1: Substrate and product complex DNA structure comparison reveals a double-base unpairing mechanism for scissile phosphate placement. *(**A**) DNA from WT:substrate complex show the +1 (sand) and −1 (green) nts are paired with the complementary strand (brown) and a close up of the active-site (below) showing the distance required for the scissile phosphate of the -1 nt to move in order to be positioned within the catalytic radius of the metal ions (cyan spheres). (**B**) Close-up view of the DNA in the active site of the WT:Sm^{3+}:product DNA complex. Unlike the substrate complex, the +1 nt has departed, the −1 nt is unpaired, and Tyr40 stacks with the −1 nt. (**C**) Model of double nt unpairing to move substrate into a catalytic position for incision. Adapted from Taken from Tsutakawa, S.E., et al. (2011). Cell 145, 198-211.*

4.2 2-aminopurine substitution in enzymology

Circular dichroism (CD) is used to monitor conformational changes of nucleic acid structures, however, many important DNA metabolic processes occur at specific junctions, therefore only the conformational changes of nucleic acids at these junctions are of relative importance. The CD of a polynucleotide effectively comes from the sum of all base-base interactions (Johnson et al., 2004). Therefore the signal originating from the residues of interest will be masked by the total signal. Exciton coupling circular dichroism (ECCD) offers a way to deconvolute CD spectra by using 2-aminopurine (2-AP) as a spectral probe, more specifically adjacent 2-AP residues, to report on the local DNA conformations (Datta et al., 2009; Johnson et al., 2004).

2-AP is a structural isomer of adenine, and although the hydrogen bonding pattern is changed by repositioning of exocyclic amino group a 2-AP:T base-pair does not significantly alter the structure or the stability of dsDNA (*figure 4.2*) (Eritja et al., 1986; Sowers et al., 1986), and 2-AP-containing dNTPs can serve as polymerase substrates in vivo and in vitro (Fidalgo da Silva et al., 2002; Goodman et al., 1977; Mandal et al., 2002; Purohit et al., 2003; Wang et al., 1981).
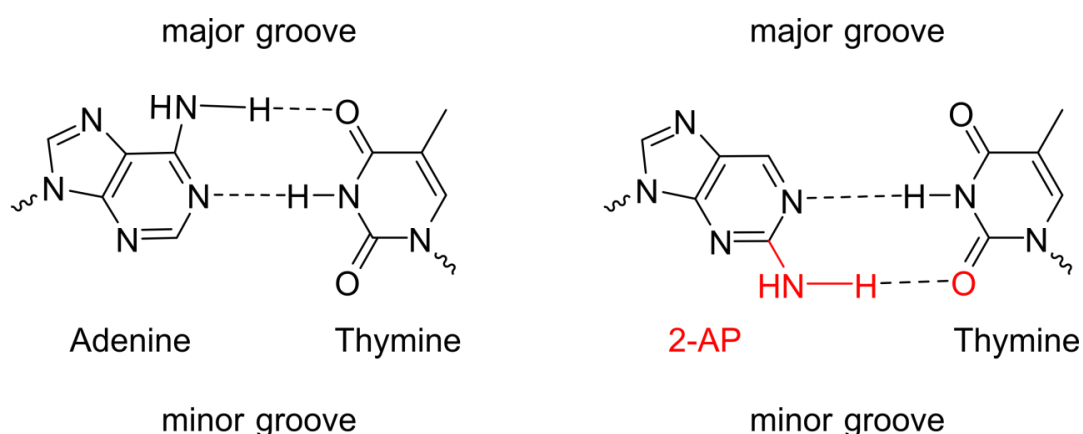
Figure 4.2: Comparison of the hydrogen bonding scheme between adenine:thymine and 2-aminopurine:thymine. *The position of the hydrogen bonds is altered in the 2-AP substituted base pair but the Watson-Crick hydrogen bonding pattern is maintained.*

2-AP contains an electronic transition dipole moment (ETDM), the magnitude and orientation of which can be modelled (*figure 4.3a*) (Holmen et al., 1997). In the absence of coupling, two 2-AP nucleotides possess ETDMs that have degenerate electronic transition energy levels. In other words, a dimer of two non-interacting monomers (such as two 2-APs separated by a single canonical base) will have the same absorption frequency as that of the monomer, except that the intensity will be twice as large. On the other hand, two tandem or adjacent 2-APs form an

exciton-pair with two singly excited electronic transitions of unequal energies (*figure 4.3b*), contributing to the uneven oscillator strengths (absorption intensities) resulting in a bisignate signal at wavelengths >300 nm in a CD spectrum.
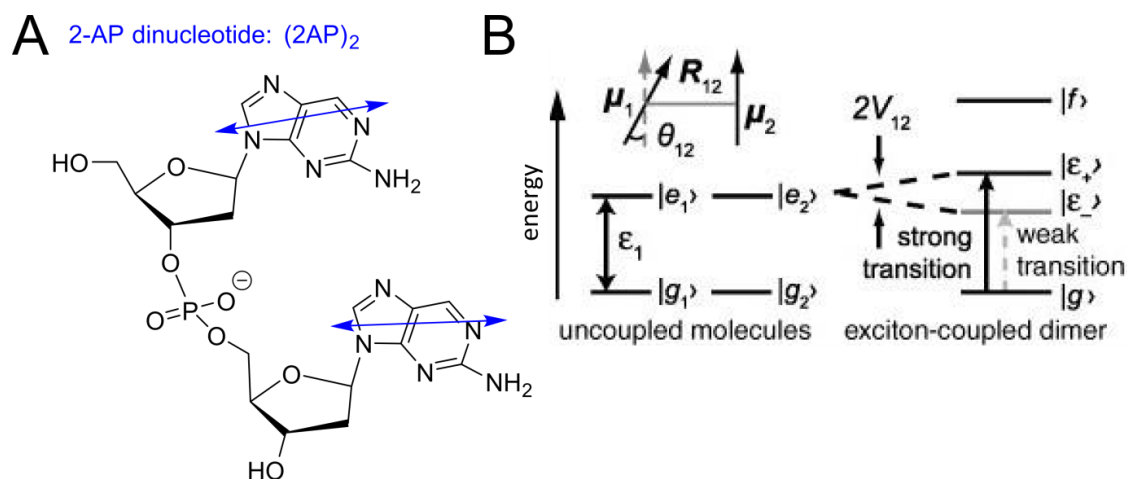
*Figure 4.3: (A) **The molecular structure of the 2-AP dinucleotide.** The lowest-energy electronic transition dipole moments (EDTMs) are indicated as blue arrows. (B) **Energy level diagram of two chemically identical two-level molecules.** An example configuration with relative transition dipole angle θ_{12}. Electronic interactions result in an exciton-coupled four-level dimer with a single ground state (g), two singly excited states (ε_+ and ε_-), and a single doubly excited state (f). The conformation of the dimer determines the energy level spacing and the strengths of the transitions. V_{12} is the coupling strength, R_{12} is the vector connecting their centres and μ_1 and μ_2 are the magnitudes of the electric transition dipole moments. Taken from Widom, J. R. et al. New J Phys, (2013), 15(2).*

The monomers of adjacent 2-APs interact due to the fluctuating electric fields in one monomer interacting with the other causing a transfer of excitation. As a result, the two singly excited states are no longer degenerate, but one is displaced above the energy level of the monomer, the other below; and the energy gap is not necessarily equal (Cantor and Schimmel, 1980). The electronic interaction between adjacent 2-AP residues and therefore the exact splitting of the energy levels is dominated by the separation but also by the relative orientation of the 2-AP electric dipole transition moments *and* by the relative phase of their excitation (Widom et al., 2013). Whether these interactions are constructive or unfavourable will result in a red-shift or blue-shift, respectively, in the absorption maximum. In addition, the exciton-coupling will result in splitting of the peak, however, usually the two spectral bands are unable to be resolved and this effect is manifested in a broadening of the peak. Moreover, a function of two interacting identical chromophores is that the relation between the absorption signal and absorption state

(orientation of dimer) is degenerate, in that different conformations can give rise to identical spectra. However, as stated above, CD can be used to enhance and exaggerate the exciton coupling effect considerably as two spectral bands of opposite sign are produced. Additionally, the low energy CD spectra not only reflect the asymmetry of the conformational relationships between 2-AP residues, but immediately adjacent stacked (or unstacked) bases or base pairs. ECCD is particularly useful for studying protein:DNA complexes because these transitions occur in a region of the CD spectra where protein and unmodified DNA are transparent.

The spectra of an isolated dinucleoside monophosphate (2-AP)$_2$ molecule, and in the context of ss and ds DNA, in addition to two 2-AP separated by an interviewing base have been characterised previously (Johnson et al., 2004). The low energy CD spectrum of an isolated dinucleoside monophosphate (2-AP)$_2$ molecule exhibits a bisignate signal with a positive cotton effect (*figure 4.4a*). The negative and positive bands display maxima at 220 and 317 nm, respectively. The cross over point, where there is no CD intensity (left and right circularly polarised light absorbed evenly) occurs at the same wavelength as the maximum absorption of unpolarised light ~270 nm.
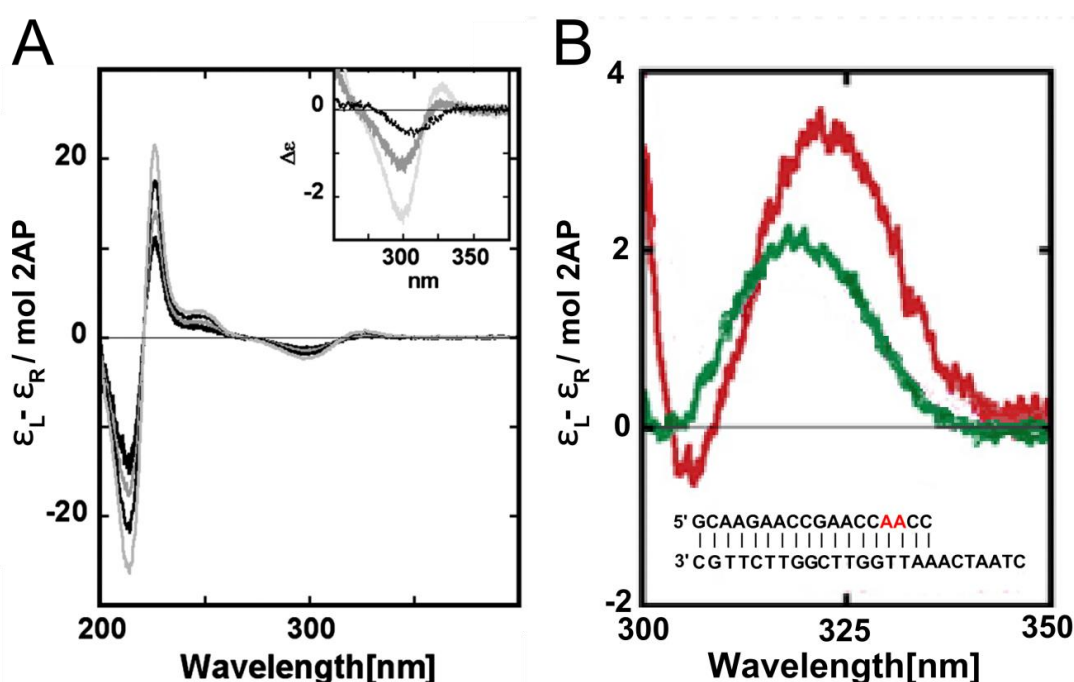
Figure 4.4: (A) CD spectrum of isolated (AP)$_2$. *The CD of (AP)$_2$ per mol of AP in 20 mM phosphate buffer, pH 7.5, at various temperatures. The spectra, in order from greater to lesser intensity, were obtained at 20, 40, 60, and 80°C. (Inset) The molar CD of (AP)$_2$ in phosphate buffer at 20°C (highest intensity) and 80°C (intermediate intensity), and in absolute EtOH (20°C) (lowest intensity). Taken from Johnson, N. P. et al. PNAS, (2004), 101(10), 3426-3431.* **(B) Comparison of CD spectrum of 2AP dimer in ssDNA (green) and dsDNA (red).** *Insert, specific oligonucleotide sequence of construct used showing 2AP nucleotide in red positioned two bases from the 3ʹ-terminus in a ssDNA construct and two bases from the 3ʹ-terminus in a 5ʹ-overhang dsDNA construct. Adapted from Datta, K. et al. JBC, (2009), 284(25), 17180-17193.*

It has been demonstrated that the intensity of this transition is sensitive to thermal and chemical denaturation. Under denaturing conditions the exciton coupling contribution to the CD signal decreases as the two chromophores become increasingly unstacked (*figure 4.4a*). In the same report, a hyper- and bathochromic shift was observed when the 2-AP dimer was buried within a 15-mer oligonucleotide. The shape of the absorption band remains bisgnate, indicative of a right-handed helical conformation for the (AP)$_2$ molecule (Johnson and Tinoco, 1969), but the maximum of the positive band is red-shifted to 326-328 nm and the negative portion of the signal is masked by the 'bleed-over' from the CD signal of the canonical DNA bases at ~260 nm. The shift in absorption maximum is associated with an increase in the CD signal. This reflects the constrained conformation of the dimer probe within the context of a ssDNA polymer and increased base stacking that results in stabilization of the excited state. This behaviour was further emphasised when the 15-mer oligonucleotide was annealed with a complementary strand. Specifically, within the context of dsDNA the absorption maximum remains at 326-328 nm but has an intensity of $\Delta\varepsilon = 3.2 \pm 0.4$ (M AP)$^{-1}\cdot$cm^{-1}. This intensity is 2.5 and 6-fold larger than in the ssDNA and isolated dinucleotide, respectively (*figure 4.4b*). This demonstrates that (2-AP)$_2$ can act as a spectral probe to report on the local conformation of DNA depending on where the dimer is placed within a construct, providing information on the base stacking environment and indirectly on the base pairing of the nucleotide analogue. As a footnote, because the CD spectral properties derive from the absorption properties degenerate CD signals can be produced, in that different conformations can give rise to identical spectra. Nevertheless, this technique has proven useful in determining the equilibrium dynamics of nucleic acid structures (*section 1.4.5*).

4.3 Using 2-aminopurine substitution to study hFEN1

4.3.1 Substrate design

The aim of the work described in this chapter was to probe specific conformational changes of bases and neighbouring bases at the ss-ds junction of hFEN1 substrates and products. Low-energy CD spectra of hFEN1:substrate complexes were recorded using minimal, static, double flap constructs that comprise of a displaced 5 nt 5'-flap and a 1 nt 3'-flap (*figure 4.5*), analogous to the substrate construct (DF-5) used in kinetic and biophysical studies (*Chapter 3*). 2-AP dimer was placed at various positions within the flap strand. The static nature of the constructs was important as it restricted the nucleic acid structure to one conformer allowing the spectrum to be categorically assigned to a particular position of the 2-AP dimer within the DNA.

Substrate constructs were bimolecular, formed from a single stranded flap strand (F) and a template strand (T). For example, F_{-1-2} was annealed with T5 to form the ds substrate construct S_{-1-2}. The numbering refers to the position of the 2-AP dimer relative to the scissile phosphate (*figure 4.5*). The equivalent ds product construct (P_{-1-2}) was prepared from a single stranded cleaved flap (CF_{-1-2}) and the same template strand (T5) (*figure 4.5*). The product construct, P_{-1-2}, had a 5'-terminal phosphate on the cleaved flap strand. An additional product construct was created lacking the 5'-terminal phosphate. This was formed from a cleaved flap strand in which the phosphate monoester was replaced with a hydroxyl group ($HO\text{-}CF_{-1-2}$). When $HO\text{-}CF_{-1-2}$ was annealed with the same template strand T5, this formed $HO\text{-}P_{-1-2}$, analogous in sequence to P_{-1-2} but lacking a 5'-terminal phosphate (*figure 4.5*). Additional substrate constructs were prepared to investigate DNA conformational change when mismatches were present or at alternative positions in the substrate. A substrate with a C:C mis-match at the ss-ds junction was designed to test the ability of hFEN1 to deal with a genetically relevant substrate which might be produced *in vivo* by a fallible replication process. This was created from a flap strand which had $(2\text{-AP})_2$ at the -1 and -2 positions and a cytosine at the +1 position ($MM_{+1}F_{-1-2}$). When annealed with the same template strand, T5, this created a ds substrate construct ($MM_{+1}S_{-1-2}$) with a C:C mis-match at the +1 position (*figure 4.5*). Finally, in order to investigate the local DNA conformation at other positions within the FEN:substrate complex the 2-AP dimer was also moved to the +1-1 and -8-9 positions. So that the complementary thymine would be positioned opposite 2-AP two different template strands were designed, T6 and T7, which when annealed with F_{+1-1} and F_{-8-9} formed S_{+1-1} and S_{-8-9}, respectively (*figure 4.5*).
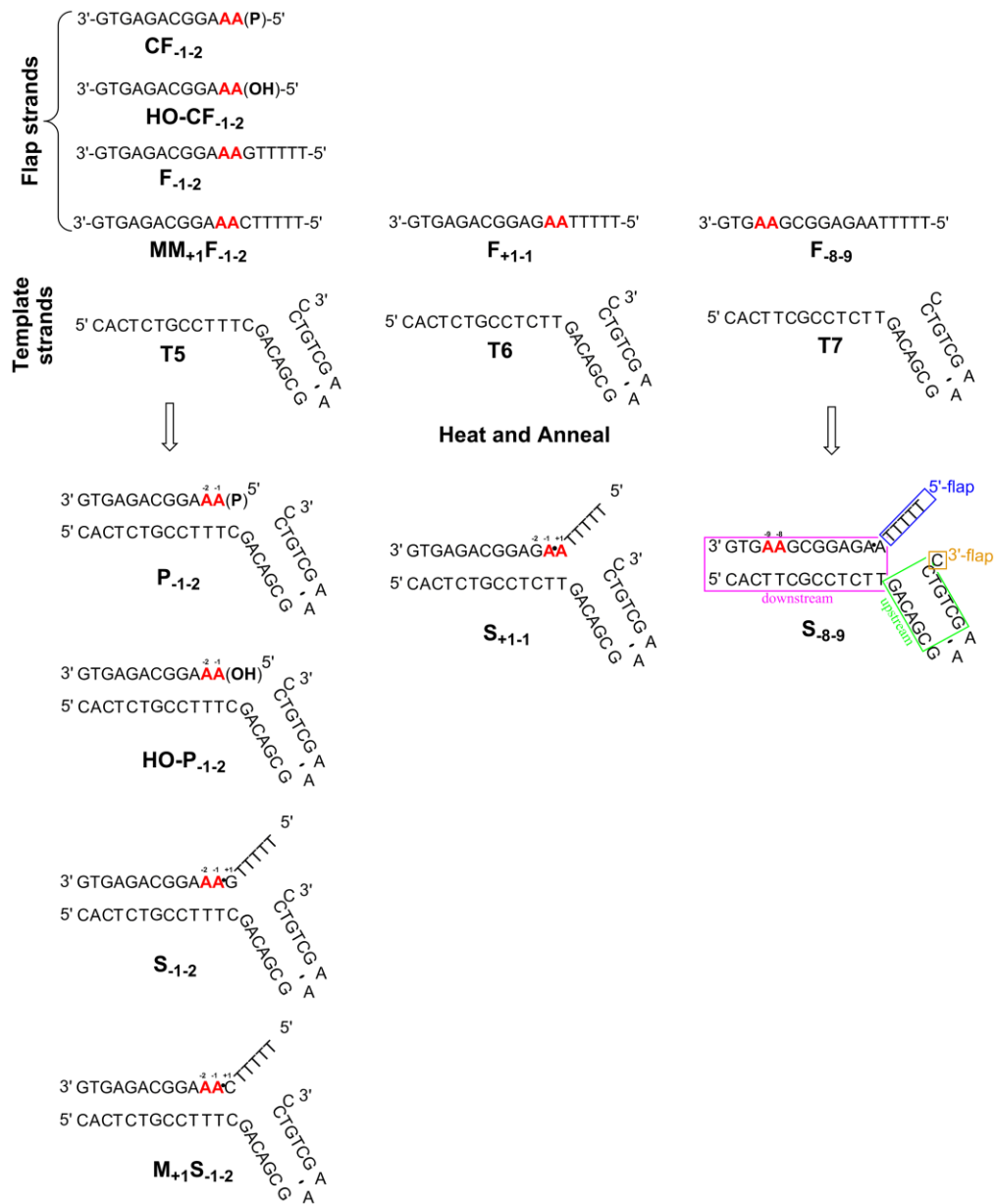
Figure 4.5: A schematic representation of the synthetic oligonucleotides used to investigate the unpairing mechanism of hFEN1. *The constructs are heteroduplexes formed from a template and flap strand (sequences: table 2.5). The template strand forms a fold-back hairpin and is capped with a [GAA] tri-loop. The scissile phosphate (black circle) connects the +1 and -1 nts. Bold P represents a 5ʹ-monophosphate, bold OH represents a 5ʹ-hydroxyl and 2-aminopurine nucleotides are depicted in red.*

4.4 Low-energy CD spectra of unbound product and substrate constructs

4.4.1 CD Spectra of $(2\text{-}AP)_2$ containing single-strands

The CD spectra of $(2AP)_2$ inserted into ssDNA (CF_{-1-2}, F_{-1-2}, F_{+1-1}, $MM_{+1}F_{-1-2}$, F_{-8-9}) resemble previously published spectra of ssDNA (*figure 4.6*). For example, the low-energy spectrum of F_{-1-2} shows two bands of opposite sign, with crossover at 310 nm, the positive band reaches a maximum intensity of 2.49 ± 0.22 (M AP)$^{-1}$·cm^{-1} at 326 nm; this was in good agreement with earlier spectra of ssDNA containing $(2\text{-}AP)_2$, albeit with a different sequence (*figure 4.6a*) (Johnson et al., 2004). The difference in sequence is important to highlight because the magnitude of the CD signal is sequence dependent. The optical activity shares a common physical origin with fluorescent properties; static quenching of 2-AP is more effective with neighbouring purines, especially guanine (Jean and Hall, 2001; Johnson et al., 2005a). Comparison between the intensity at 326 nm of the substrate flap strands (F_{-1-2}, $MM_{+1}F_{-1-2}$, F_{+1-1} and F_{-8-9}) where the 2-APs are in similar position within the oligonucleotide shows a correlation between the observed exciton-coupling and the neighbouring bases. The intensity was lowest when there were guanines on both the 3'- and 5'-side of the 2-AP dimer, this was observed in the F_{-8-9} oligonucleotide (*figure 4.6e*). The intensity increased when there was only one guanine neighbouring the $(2\text{-}AP)_2$ (F_{-1-2} and F_{+1-1}), and increased further when there were no directly adjacent guanines, $MM_{+1}F_{-1-2}$ (*figure 4.6e*).

Unlike the substrate flap strands where the $(2\text{-}AP)_2$ is buried in the middle of the sequence, the spectral probe was at the 5'-terminus of the product cleaved flap strand (ssCF_{-1-2}). Nevertheless, the low-energy CD spectra of the ss product flap strand displayed a bisignate signal reaching a maximum at 326 nm (*figure 4.6a*), and the intensity of the 326 nm spectra band was experimentally identical to, F_{-1-2} and F_{+1-1} (*figure 4.6e*). This suggested that even at the terminus of a ss oligonucleotide (CF_{-1-2}) 2-APs were in a relatively stacked environment.
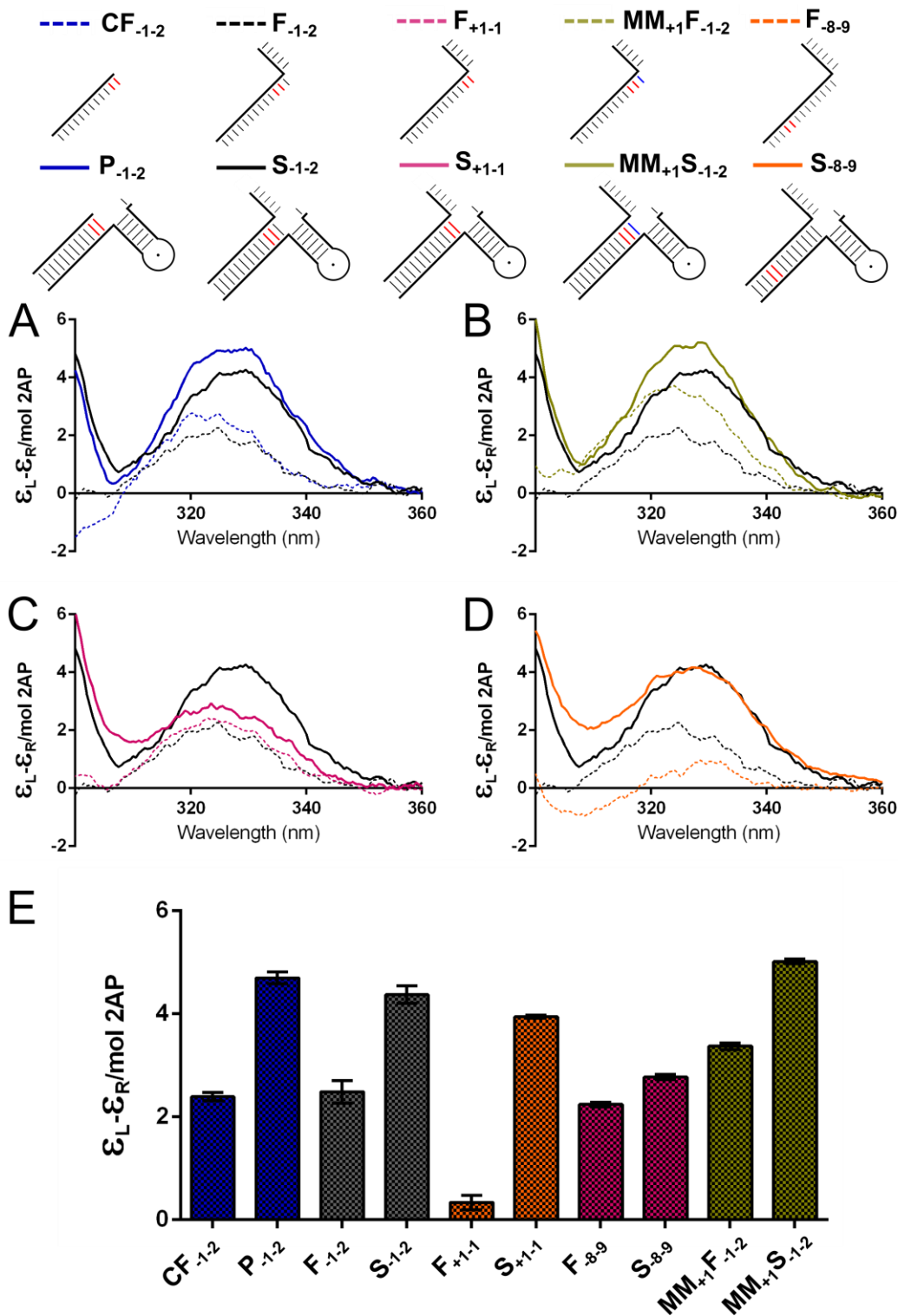
Figure 4.6: Low-energy CD spectra of unbound 2-AP-containing DNA product and substrate constructs. *(**A**) Unbound P$_{-1-2}$ (blue), the corresponding single strand (CF$_{-1-2}$, cyan) and unbound S$_{-1-2}$ (purple), the corresponding single strand (F$_{-1-2}$, magenta). (**B**) Unbound S$_{-1-2}$ (purple), the corresponding single strand (F$_{-1-2}$, magenta) and unbound MM$_{+1}$S$_{-1-2}$ (black), the corresponding single strand (MM$_{+1}$F$_{-1-2}$, gray). (**C**) Unbound S$_{-1-2}$ (purple), the corresponding single strand (F$_{-1-2}$, magenta) and unbound S$_{+1-1}$ (dark green), the corresponding single strand (F$_{+1-1}$, light green). (**D**) Unbound S$_{-1-2}$ (purple), the corresponding single strand (F$_{-1-2}$, magenta) and unbound S$_{-8-9}$ (orange), the corresponding single strand (F$_{-8-9}$, yellow). (**E**) Ellipticity changes at 326 nm for constructs containing 2-AP dimer probes. Ellipticity changes are per mol 2-AP. All data sets are two or more experimental repeats, standard errors are shown. All spectra were recorded in the presence of 10 mM Ca^{2+}, samples were buffered at pH =7.5, and measurements performed at 20 ˚C; specific conditions given in section 2.7. Each measurement was independently repeated and gave equivalent results, spectra were normalized as described in experimental section 2.7. DNA constructs are illustrated schematically (red 2-AP) with sequence shown in figure 4.5.*

130

4.4.2 CD Spectra of (2AP)$_2$ containing duplexes in the context of FEN1 product and substrate

Transition from ss to ds for all the oligonucleotides tested produced an increase in the CD intensity at 326 nm (*figure 4.6e*). Careful inspection of these CD spectra also showed that a small redshift of the peak to 328 nm accompanied the transition from ss to ds. As mentioned in *section 4.2*, these spectroscopic behaviours suggest increased stacking of the spectral probe in the duplex environment. The enhanced excition-coupling between the 2-APs suggests that the nts are restricted to fewer conformations in a DF and product constructs compared to a ss oligonucleotide. When the bases were fully stacked in ds DNA, S$_{-8-9}$, the CD peak shifted no further. Based on these observations, CD results are presented as elipticity per mol 2-AP at 326 nm for 2-AP dimers, alongside typical traces in *figure 4.6.* Although spectra were recorded from 300 – 480 nm, no absorbance was detected above 360 nm (*Appendices figure A2*) for any experiment presented in this chapter and thus spectra shown display the signal from 300 -360 nm.

Nicked, gapped and branched nucleic acid structures are highly dynamic and able to adopt a multitude of global conformations (Kahn et al., 1994; Lin et al., 2009; Sabir et al., 2011). It is generally accepted that recognition, in part at least, of DNA replication and repair intermediates is mediated through this conformation diversity relative to limited flexibility of intact DNA (Lin et al., 2009). Relating this to FEN1, bifurcated structures can adopt two extremes. Substrates can either be in a fully extended conformation whereby the two duplex arms coaxially stack on one another or bent at the ss-ds junction removing the coaxial stacking. The difference in conformations is important to highlight because previous data shows that bases at the ss-ds junction of different nucleic acid structures have different tendencies to unpair (Ramstein and Lavery, 1988).

The CD spectrum of S$_{-1-2}$ has an exciton-coupling intensity of $\Delta\varepsilon$ = 4.37 ± 0.17 (M AP)$^{-1}\cdot$cm^{-1}, considerably enhanced from that of the corresponding single strand (F$_{-1-2}$). The DF construct S$_{-1-2}$ would be expected to readily adopt coaxial stacked structures, at least some of the time. In the ds product construct (P$_{-1-2}$) the exciton-coupling at 326 nm was considerably enhanced (4.70 ± 0.12 (M AP)$^{-1}\cdot$cm^{-1}) compared to the ss flap (CF$_{-1-2}$) and was of a similar magnitude to that observed with S$_{-1-2}$. It is interesting to find that the signal associated with P$_{-1-2}$ so closely resembles S$_{-1-2}$. In our so-called "static" product the 3'-flap is not complementary to the template DNA and the structure might be expected to more closely resemble a one nucleotide gapped DNA. This might be expected to exhibit significant "breathing" of the terminal nucleotide which

would predict a lowered signal compared to S$_{-1-2}$. However, the comparable magnitude of the ECCD signals of S$_{-1-2}$ and P$_{-1-2}$ suggested that there was a similar amount of stacking in both DNA constructs. It is possible that the non-complementary 3'-flap nt adopts an intrahelical conformation in P$_{-1-2}$ allowing coaxial stacking, and/or that the DNA adopts a slightly bent conformation filling the "gap".

The ds substrate construct with a C:C mis-match the ss-ds junction (MM$_{+1}$S$_{-1-2}$) has the greatest CD intensity of any of the oligonucleotides used in this study, $\Delta\varepsilon$ = 5.02 ± 0.04 (M AP)$^{-1}$·cm^{-1}, modestly increased from S$_{-1-2}$ (*figure 4.6b*). As discussed above this may be a sequence dependent change in the signal as it is also enhanced in MM$_{+1}$F$_{-1-2}$.

As mentioned above, the two isolated flap single-strands, F$_{-1-2}$ and F$_{+1-1}$, had identical spectroscopic behaviour (*figure 4.6c*). However, the double stranded substrate construct, S$_{+1-1}$, only had an intensity of $\Delta\varepsilon$ = 2.77 ± 0.05 (M AP)$^{-1}$·cm^{-1}, compared to $\Delta\varepsilon$ = 4.37 ± 0.17 (M AP)$^{-1}$·cm^{-1}, for the S$_{-1-2}$ construct. Previously reported 2-AP ECCD and fluorescence data suggest that significant breathing of base-pairs located at ssDNA–dsDNA junctions such as forked and overhang DNAs occurs and is detected by these techniques (Jose et al., 2009). This effect would be expected to be most marked in the S$_{+1-1}$ substrate. The change in exciton-coupling implies that the +1 nt is more unstacked with the -1 nt than the -1 nt is with the -2 nt. Thus, free in solution the first base-pair into the downstream region of a double-flap substrate is not as strong as those buried further into the ds downstream region, and therefore, the +1 nt is solvent exposed more of the time (*figure 4.6c*). As previously mentioned, the ds-ss junction of a double flap substrate is in equilibrium between a coaxial stacked and bent conformation. Therefore, depending on the position of that equilibrium the 5'-face of the 2-AP dimer at the +1-1 position will be solvent exposed and have no stacking partner.

The low-energy CD spectrum of the substrate construct, S$_{-8-9}$, in which the (2-AP)$_2$ was located furthest away from the ss-ds junction was also enhanced compared to single strand (*figure 4.6d*). The DNA of a DF construct in the downstream ds region is B-form. Nts in this region are buried in the helical axis of DNA and well-stacked with one another. Indicated by an intensity of $\Delta\varepsilon$ = 3.95 ± 0.05 (M AP)$^{-1}$·cm^{-1} for the S$_{-8-9}$ spectral band at 326 nm.

The spectroscopic studies discussed in this chapter have been performed at either 10 mM Ca^{2+} or 2 mM EDTA. The data present in this section for unbound ss and ds oligonucleotides has only been presented for spectra collected in buffer containing 10 mM CaCl$_2$, as the traces are identical in either buffer (*Appendicies figure A2*). Divalent metal ions have been reported to affect the DF equilibrium between coaxially stacked and bent conformations (Craggs et al., 2014). However,

as traces of constructs measured in the presence of 10 mM Ca^{2+} or 2mM EDTA overlay this implies that the same conformations are accessible in both conditions and any changes in equilibria brought about by divalent metal ions are not reflected in our measurements. Initially, exciton-coupling experiments were conducted by Nikesh Patel, David Finger, Emma Jardine, and myself. The data presented here represents independent repeats collected by myself unless stated otherwise.

4.5 Low-energy CD spectra of WT- and mutated-FEN product complexes

As a starting point for the analysis of unpairing of FEN1-DNA complexes we analysed P_{-1-2} complexes. Structural analyses have shown that in hFEN1:product complexes the terminal nucleotide (-1) of the 5'-flap strand is extra-helical (figure 4.7), whereas the neighbouring nucleotide (-2) remains base paired with the template strand; although the sugar is slightly distorted due to the wide dihedral angle of the backbone between the -1 and -2 nts. Therefore, replacement of the -1 and -2 nts with a 2-AP dimer in P_{-1-2} (*figure 4.5*) should allow a signal to be established for the hFEN1 induced changes of nucleotide conformation.
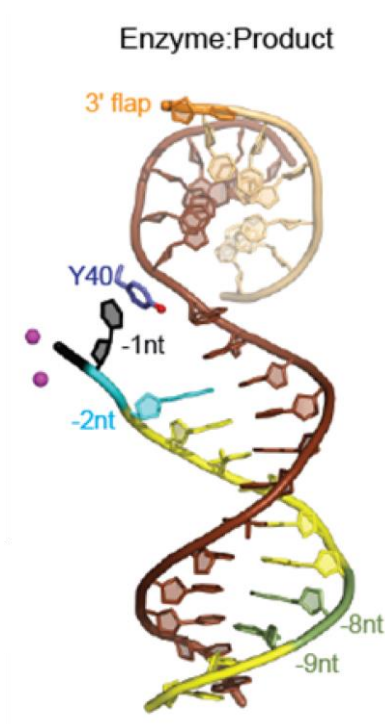
Figure 4.7: hFEN1 product complex. *-1 nt (black) is extrahelical and not involved in stacking interaction with the -2 nt (blue), the backbone is slightly distorted but the -2 nt remains base paired with the opposite strand in this structure. Nucleotides at the -8 and -9 positions (green) reside within B-DNA. Taken from Tsutakawa, S.E., et al. (2011). Cell 145, 198-211.*

Low-energy CD spectra were typically recorded using 10 μM oligonucleotide and 12.5 μM protein. As dissociation constants of FEN1-substrate complexes are in the nanomolar range in buffers containing 10 mM Ca^{2+} or 2 mM EDTA (*table 3.4*) it was suspected that all the DNAs would be fully bound by protein during the CD measurement. This was verified by EMSA results performed at the concentrations used in CD experiments. In all cases, regardless of whether WT or mutated FEN1s are used or whether substrate or product was present, at least 95% of the substrate and product DNA was complexed with hFEN1 (*figure 4.8*).

Samples of 2-AP-containing DNA product and substrate constructs were initially prepared separately in 10 mM Ca^{2+} and 2 mM EDTA. However, it was realised that an excess of EDTA could be added to the calcium sample to remove any divalent metal ions and save on expensive 2-AP oligonucleotide. Samples prepared separately were identical to those prepared by the additive method (data not shown).
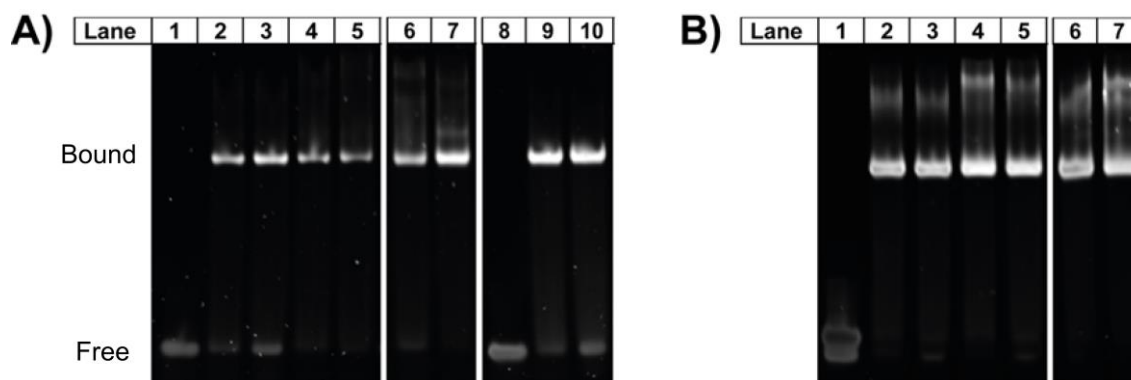
Figure 4.8: Electrophoretic mobility shift assays (EMSA) of product and substrate heteroduplexes with WT and variant hFEN1s under conditions analogous to CD experiments. Final enzyme and substrate concentrations were 12.5 μM and 10 μM respectively, in buffer (110 mM KCl, 55 mM, HEPES pH=7.5, 1 mM DTT with either 10mM $CaCl_2$ alone or 10mM $CaCl_2$ and 40mM EDTA). DNA was visualized using SYBR® Green and a ChemiDoc^{TM} Imaging System (BioRad). (A) Lane 1: P_{-1-2} (Ca^{2+}), Lane 2: P_{-1-2} + WThFEN1 (Ca^{2+}), Lane 3: P_{-1-2} + WThFEN1 (EDTA), Lane 4: P_{-1-2} + Y40A (Ca^{2+}), Lane 5: P_{-1-2} + Y40A (EDTA), Lane 6: P_{-1-2} + R100A (Ca^{2+}), Lane 7: P_{-1-2} + R100A (EDTA), Lane 8: HO-P_{-1-2} (Ca^{2+}). Lane 9: HO-P_{-1-2} + WThFEN1 (EDTA), Lane 10: HO-P_{-1-2} + WThFEN1 (Ca^{2+}). All samples were electrophoresed on the same gel. (B) Lane 1: S_{-1-2} (Ca^{2+}), Lane 2: S_{-1-2} + WThFEN1 (Ca^{2+}), Lane 3: S_{-1-2} + WThFEN1 (EDTA), Lane 4: S_{-1-2} + Y40A (Ca^{2+}), Lane 5: S_{-1-2} + Y40A (EDTA), Lane 6: S_{-1-2} + R100A (Ca^{2+}), Lane 7: S_{-1-2} + R100A (EDTA)⁺). All samples were electrophoresed on the same gel.

4.5.1 Product unpairing is metal dependent

The exciton-coupling for the unbound P_{-1-2} construct has an intensity of $\Delta\varepsilon = 4.37 \pm 0.17$ (M AP)⁻¹·cm⁻¹, in the absence of divalent metal ions. Upon formation of the WThFEN1:P_{-1-2} complex in

EDTA the exciton band at 326 nm was modestly reduced to, $\Delta\varepsilon = 3.07 \pm 0.36$ (M AP)$^{-1}$·cm^{-1} (*figure 4.9*). Both these values were greater than the signal observed for the corresponding single strand CF$_{-1-2}$ From FRET binding data (*section 3.5.2*) the substrate (at least) is forced into a bent conformation upon the addition of hFEN1 and the substrate is bent to a similar extent in the absence and presence of divalent metal ions. Thus binding to FEN1 would break the co-axial stacking. This 25% decrease in intensity suggests that DNA at the (2-AP)$_2$ remains in largely in duplex form on FEN1 in EDTA even if it is bent.

hFEN1:P$_{-1-2}$ complex was then formed in the presence of catalytically inert Ca^{2+} ions. This produced a dramatic reduction in the amplitude at 326 nm to near zero, $\Delta\varepsilon = 0.58 \pm 0.08$ (M AP)$^{-1}$·cm^{-1} (*figure 4.9*) indicating that the exciton coupling between the 2-APs has nearly been abolished. The maximum intensity of the P$_{-1-2}$ bound to WThFEN1(Ca^{2+}) was considerably lower than the unbound ss CF$_{-1-2}$ ($\Delta\varepsilon = 2.39 \pm 0.08$ (M AP)$^{-1}$·cm^{-1} at 326 nm). Firstly, this suggests that the -1 and -2 nts are considerably unstacked from each other when bound to hFEN1 in the presence of Ca^{2+} ions. Secondly, this highlights the considerable secondary structure of ss DNA. This result can be rationalised by remembering that in the hFEN1:product crystal structures the -1 nucleotide is not base paired with the complementary strand and is extrahelical, completely unstacked with the -2 nucleotide, in this case the 3'-2-AP. Moreover, together the data imply that there is a dependence on divalent metal ions for conformational change of the -1 and -2 nt. This is in-line with crystal structures of hFEN1 bound to product DNAs, where the terminal phosphate monoester is inner sphere coordinated to active site metal ions.
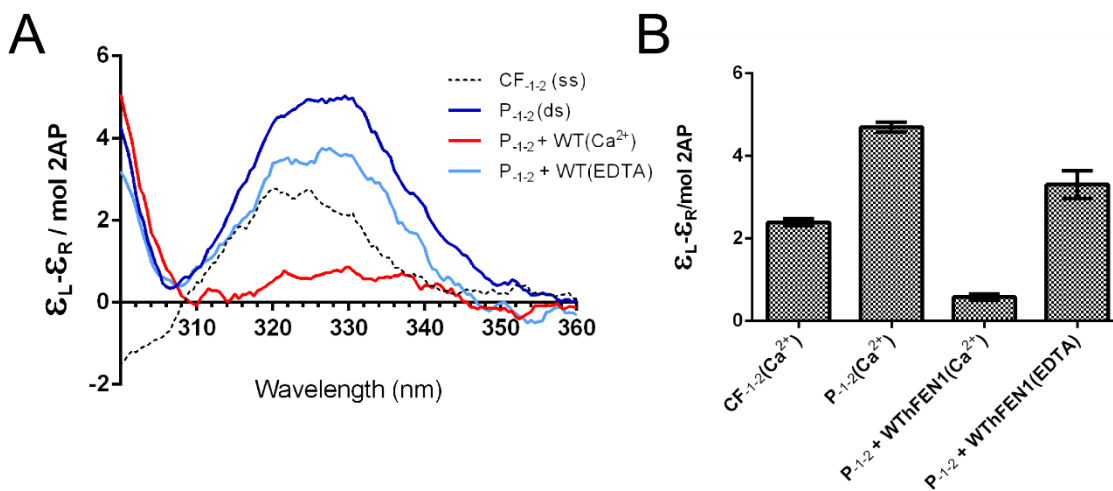
Figure 4.9: Product unpairing by WThFEN1 protein (A): Low energy CD spectra of 2-AP containing DNAs and WT-DNA complexes. Unbound P-1-2 (blue), the corresponding single strand (CF-1-2, gray) and P-1-2 bound to WThFEN1 (red) all in Ca^{2+} containing buffer. P-1-2 bound to hFEN1 in buffer containing no available divalent metal ions (25 mM EDTA, cyan). Samples were buffered at pH =7.5, and measurements performed at 20 °C; specific conditions give in section 2.7. Each measurement was independently repeated and gave equivalent results, spectra were normalized as describe in experimental section 2.7. (B) Ellipticity changes at 326 nm for constructs containing 2-AP dimer probes. Ellipticity changes are per mol 2-AP. All data sets are two or more experimental repeats, standard errors are shown. Spectra were initially taken by NP, LDF, EJ and repeated by JCE.

4.5.2 A 5'-terminal phosphate is required for unpairing

To test for a key role for the active site metal ion phosphate interaction, a product construct was designed lacking the 5'-phosphate monoester (HO-P-1-2, *figure 4.5*). Unbound HO-P-1-2 and its constituent single strand HO-CF-1-2 both behaved analogously to P-1-2 and CF-1-2. At the risk of sounding redundant, the 2-AP dimer inserted into the single stranded oligonucleotide (HO-CF-1-2) at >300 nm displays a bisignate signal with no CD intensity at the absorption maximum of 2-AP and reaches a maximum CD absorption at 323 nm (*figure 4.10*). When the flap strand was annealed with the template strand creating the ds product, HO-P-1-2, there was an increase in CD intensity and concurrent red-shift and enhanced ECCD signal on duplex formation. This suggests that removal of the 5'-phosphate monoester does not significantly alter the dynamics at the ss-ds junction.

However, the hFEN1:HO-P-1-2 complex in the presence of divalent metal ions did not produce the same response as the product complex containing the 5'-phosphate monoester. HO-P-1-2 complex in Ca^{2+} was associated with a slight blue-shift but no change in intensity (*figure 4.10*). Although by using this technique alone the exact conformation of the -1 and -2 nts cannot be determined, this data suggests that the two nts at the ss-ds junction remain base-paired but in

a slightly different conformation than free DNA. Furthermore, the low-energy CD spectra of P$_{-1-2}$ and HO-P$_{-1-2}$ in complex with hFEN1 are identical in buffer containing EDTA. Thus, unpairing is dependent on both divalent metal ions and the 5'-phosphate monoester, consistent with the crystallographic data where the terminal phosphate monoester is bound to the active-site metal ions in the product complex.
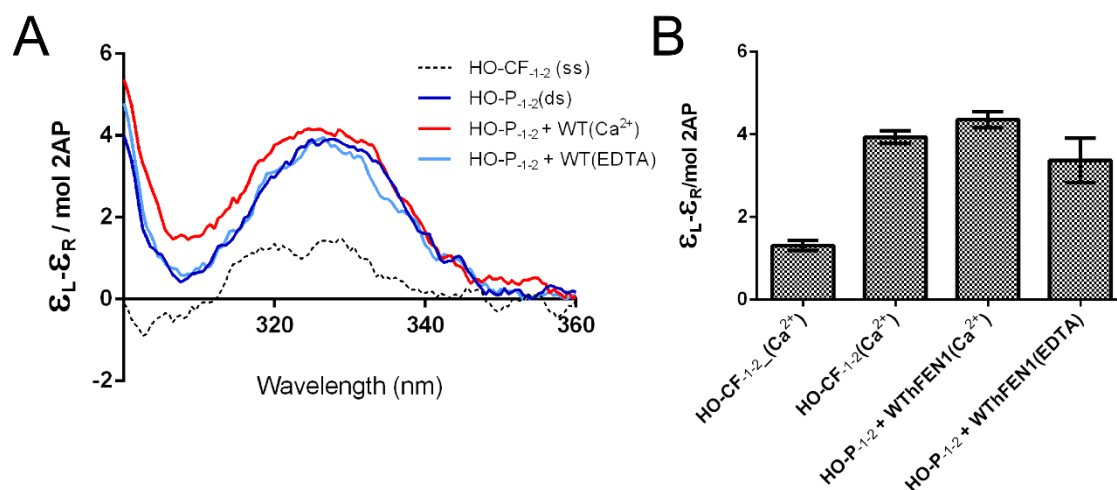
Figure 4.10: Product lacking the 5'-phosphate monoester (HO-P$_{-1-2}$) is not unpaired by WThFEN1 protein (A): Low energy CD spectra of 2-AP containing DNAs and WT-DNA complexes. Unbound P$_{-1-2}$ (blue), the corresponding single strand (HO-CF$_{-1-2}$, gray) and HO-P$_{-1-2}$ bound to hFEN1 (red) all in Ca^{2+} containing buffer. HO-P$_{-1-2}$ bound to hFEN1 in buffer containing no available divalent metal ions (25 mM EDTA, cyan). Samples were buffered at pH =7.5, and measurements performed at 20 °C; specific conditions give in section 2.7. Each measurement was independently repeated and gave equivalent results, spectra were normalized as describe in experimental section 2.7. (B) Ellipticity changes at 326 nm for constructs containing 2-AP dimer probes. Ellipticity changes are per mol 2-AP. All data sets are two or more experimental repeats, standard errors are shown. Spectra were initially taken by NP, LDF, EJ and repeated by JCE.

4.5.3 The behaviour of D181A P$_{-1-2}$ complexes does not mimic those of WT FEN1

The active site carboxylate D181A hFEN1 variant was selected to study the dependence of product unpairing on metal ions. Previous work has suggested that mutation of Asp-181 to Ala alters the sequestration of M^{2+} within the active site, but not DNA affinity (Greene et al., 1999). Crystallisation of D181A-DNA with Ca^{2+} revealed no Ca^{2+} ions within the active site. However, other FEN1 superfamily members have been crystallised with the equivalent mutation and active site metal ions have been observed. (Chang et al., 2011; Orans et al., 2011).

In EDTA, D181A:P$_{-1-2}$ produced a similar CD response to WThFEN1 (*figure 4.11*). However, in contrast to the decrease in signal observed for WThFEN1 upon binding P$_{-1-2}$ at 326 nm in 10mM Ca^{2+}, a large increase in the exciton-coupling accompanied with a blue-shift to 322 nm was

137

observed for D181A complexation (*figure 4.11*). In addition, the height of the trough at 310 nm was substantially increased. This large increase in intensity at 326 nm was unique to D181A and has been observed for additional ECCD constructs (Sana Algasaier personal communication). When bound to D181A in the presence of Ca^{2+} ions the substrate construct CD intensity has increased to above that observed for unbound dsDNA. This may reflect reduced thermal fluctuations and the increased stability of the stacked 2-AP dimers. In unbound dsDNA the 2-APs will be sampling a multitude of conformations, some stacked, some unstacked, and the signal is an average of these, but in D181A the nucleotides at the junction may be forced into one particular conformation and this happens to be where to 2-APs are constrained in a well stacked environment.

The data presented here suggests that the correct occupancy of the active site by divalent metal ions found in the WT proteins is required for the -1 nt of product to unpair. Moreover, if it is assumed that there are no divalent metal ions in the active site of hFEN1 in the presence of EDTA then at least one Ca^{2+} ion must be present in the active site of D181A with Ca^{2+} as the D181A enzyme in complex with the P_{-1-2} construct displays different exciton-coupling behaviour depending on the presence of available Ca^{2+} ions.
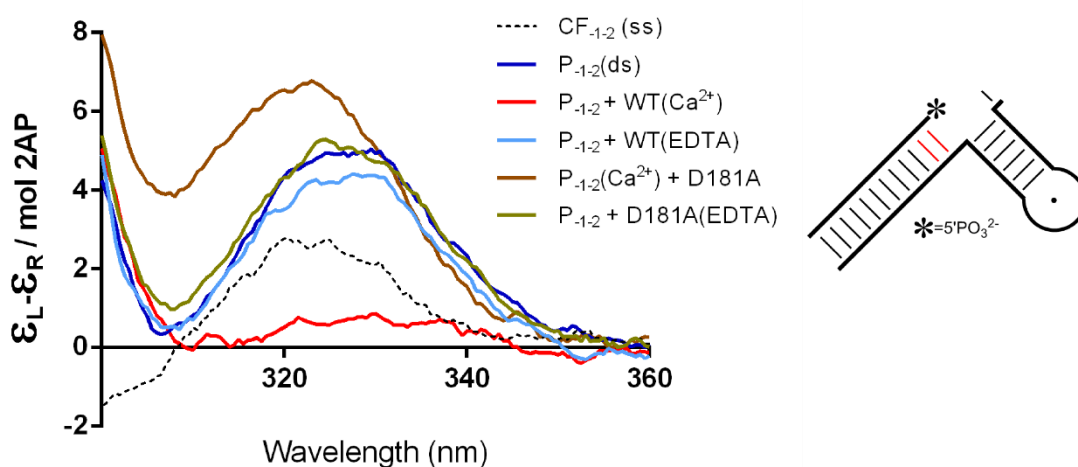
Figure 4.11: Low-energy CD spectra of 2AP-containing DNA product construct and hFEN1 WT- and D181A-DNA product complexes. (**A**) *Unbound P_{-1-2} (blue), the corresponding single strand (CF$_{-1-2}$, gray), P_{-1-2} bound to WThFEN1 (red), and P_{-1-2} bound to D181A (black)all in buffer containing Ca^{2+}.P_{-1-2} bound to WThFEN1 (teal), and P_{-1-2} bound to D181A (orange) all in buffer containing no available divalent metal ions (10 mM Ca^{2+} + 25 mM EDTA).Samples were buffered at pH =7.5, and measurements performed at 20 ˚C; specific conditions given in section 2.7. Each measurement was independently repeated and gave equivalent results, spectra were normalized as describe in section 2.7. Product construct is illustrated schematically (red 2AP) with sequence shown in figure 4.5.*

4.5.4 Mutation of hFEN1 active site residues impairs catalysis and product unpairing

In agreement with crystallographic data, the spectroscopic data above implies that the divalent metal ions in the active site contact the 5'-monophosphate of the terminal (-1) nt in the hFEN1:product complex. However, this is not only interaction believed to be important for promoting the unpairing process. Contacts were also observed to the phosphate of the unpaired nucleotide by Lys-93 and Arg-100, which interact via salt bridges. In addition in the unpaired product structure, the nucleobase of the -1 nt is stacked with Try-40. Mutation of any of these residues is detrimental towards catalysis (*section 3.4*) and this has been proposed to be, in part at least, due to their involvement in the unpairing mechanism.

The signal for the WThFEN1:P_{-1-2}(Ca^{2+}) complex cannot be replicated in the presence of either of the mutated FENs K93A or R100A (*figure 4.12a*). Assembly of the product complex with either K93A or R100A in the presence of Ca^{2+} ions produced only a very slight decrease in the exciton-coupling. This indicated that the -1 nt remained well-stacked with the -2 nt despite being bound to K93A or R100A which were assumed to have Ca^{2+} ions bound in the active site. Therefore, clearly the integrity of these basic residues are very important for producing the conformation observed in the hFEN1:product crystal structure.

The signal at 326 nm associated with Y40A:P_{-1-2} decreases further than either K93A or R100A in the presence of Ca^{2+} ions but not as dramatically as WThFEN1, along with a simultaneous increase in the signal at 310 nm (*figure 4.12a*). This may reflect differences in a partition between paired and unpaired forms and/or a different orientation of the 2-AP nucleobases in the Y40A complex. The latter interpretation is suggested by time resolved fluorescence measurements, where a significant proportion of the 2-AP decays analogously to 2-AP nucleoside in free solution (Finger et al., 2013).

Overall the equilibrium position of the (2-AP)$_2$ is similar for WT, K93A, R100A and Y40A product complexes in the presence of EDTA (*figure 4.12b*). However, the WT:P_{-1-2}(EDTA) signal cannot be replicated faithfully. This suggests that nucleotide conformational changes flanking the reaction site in the absence of M^{2+} are likely a result of the dsDNA binding motifs distinct form the active site, although very subtle changes can be brought about by mutation of conserved basic and stacking residues.
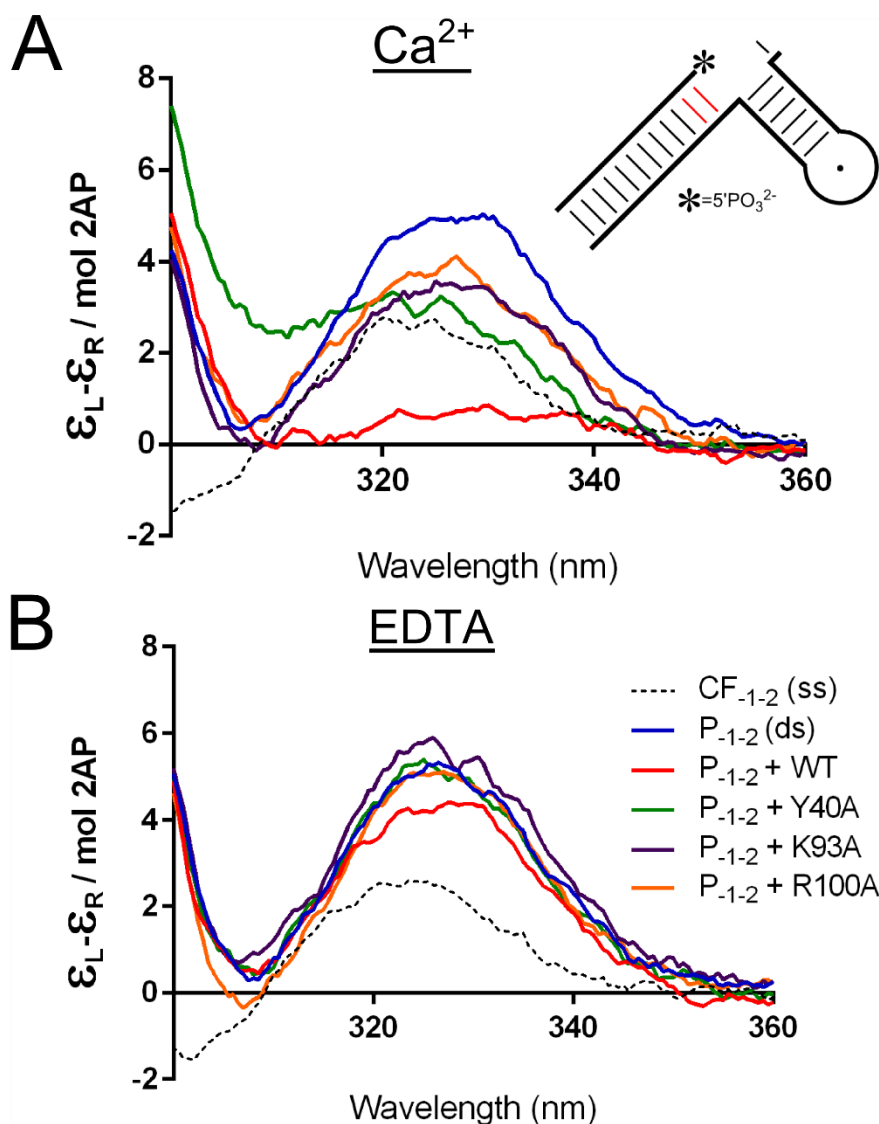
Figure 4.12: Low-energy CD spectra of 2AP-containing DNA product construct and hFEN1 WT-and variant-DNA product complexes. (**A**) *Unbound P$_{-1-2}$ (blue), the corresponding single strand (CF$_{-1-2}$, gray) and P$_{-1-2}$ bound to WThFEN1 (red), Y40A (green), K93A (purple), and R100A (orange) all in Ca^{2+} containing buffer.* (**B**) *Unbound P$_{-1-2}$ (blue), the corresponding single strand (CF$_{-1-2}$, gray) and P$_{-1-2}$ bound to WThFEN1 (red), Y40A (green), K93A (purple), and R100A (orange) all in buffer containing no available divalent metal ions (10 mM Ca^{2+} + 25 mM EDTA). Samples were buffered at pH =7.5, and measurements performed at 20 ˚C; specific conditions given in section 2.7. Each measurement was independently repeated and gave equivalent results, spectra were normalized as describe in section 2.7. Product construct is illustrated schematically (red 2AP) with sequence shown in figure 4.5. Spectra were initially taken by NP, LDF, EJ and repeated by JCE.*

4.4.5 Summary and conclusions of P$_{-1-2}$ data

The low-energy CD spectra of unbound and bound (2-AP)$_2$ containing FEN1 product DNA shows that information about the local DNA conformation can be collected by judicious placement of this spectral probe. Therefore, this technique lends itself to studying the proposed unpairing mechanism of hFEN1. The low-energy CD spectra of ss and ds oligonucleotides containing two tandem 2-APs resemble previously published data. The exciton-coupling between the 2-APs is affected by their sequence context, and neighbouring guanines are most debilitating to the CD intensity. It has previously been established that the transition from ss to ds DNA is associated with an increase and red-shift in the maximum CD intensity. This phenomenon is also observed for two tandem 2-APs at the ss-ds junction of DF structures. Although comprehensive data is already available on the spectroscopic behaviour of 2-AP within various ds DNA structures this represents the first study to characterise the low-energy CD spectra of a 2-AP dimer at the ss-ds junction of a double-flap construct. Surprisingly, the bases are well stacked at the terminus of a nicked-like DNA structure, but destabilised by the presence of a 5'-flap. Furthermore, the CD intensity varies with respect to the distance of the spectral probe from the ss-ds junction of DF structures. The first nucleotide of the ds region adjacent to the 5'-flap is significantly more 'wobbly' than even the next nt buried further into the duplex.

The 5'-terminal nucleotide of the flap strand is extrahelical in the product structure. Therefore, the two terminal nts of the cleaved flap strand were replaced with 2-AP to allow the dynamics at this position to be investigate when the ds product was bound to hFEN1. The results show that product unpairing is dependent on both the 5'-phosphate diester of the cleaved flap strand and bound divalent metal ions in the hFEN1 active site. In addition, active site residues Y40, K93, and R100A have a significant influence on the equilibrium position of the -1 and -2 nts in the presence of Ca^{2+} ions. The results presented here support the view that Y40, K93, R100 are required to unpair the terminal nucleotide of the product in an analogous fashion to the WT protein. In summary, a distinct and significant change in exciton-coupling is observed when the 2-AP dimer is placed at a position known to undergo conformational change. Therefore, the change in intensity can be thought of as an unpairing signal, and can be used to investigate the unpairing of substrate DNA proposed to be an important intermediate the molecular mechanism of FEN1.

4.6 Low-energy CD spectra of WT- and mutated-FEN substrate complexes

4.6.1 DNA conformational in $(2AP)_2$ S_{-1-2} brought about by WT and mutated FENs

Substrate unpairing was also monitored in the presence and absence of divalent metal ions. The functional requirement for DNU along the FEN reaction course was first revealed by nucleobase cross-linking experiments (Beddows et al., 2012). It was shown that the ability of the nucleotides at the ss-ds junction to 'breath' is important for activity and reaction site specificity. To gain insights into nucleotide dynamics at the base of the 5'-flap the spectral $(2\text{-}AP)_2$ probe was substituted into a substrate construct at equivalent positions (-1 and -2) to the product construct (*figure 4.5*).

A dramatic decrease in CD intensity was associated with binding of WThFEN1 to the substrate construct in the presence of divalent metal ions. The absorption band at 326 nm had an intensity of $\Delta\varepsilon = -0.01 \pm 0.08$ (M AP)$^{-1}\cdot$cm^{-1}, compared to $\Delta\varepsilon = 4.73 \pm 0.17$ (M AP)$^{-1}\cdot$cm^{-1} for dsDNA, implying a loss of exciton-coupling consistent with unstacking between the -1 and -2 nts. The reduction in intensity of the exciton-coupling at 326 nm for the S_{-1-2}:WThFEN1:Ca^{2+} complex was similar to that obtained for the product complex when compared to the dsDNA product (P_{-1-2}, $\Delta\varepsilon = 4.70 \pm 0.12$ (M AP)$^{-1}\cdot$cm^{-1}). Notably however, the absorption from 300 - 360 nm was not equivalent in both complexes (compare red trace in *figures 4.12a and 4.13a*). There was a distinctive trough reaching a minimum at 313 nm for the low-energy CD spectrum of the S_{-1-2}:WThFEN1:Ca^{2+} complex. This indicated that the position of the -1 and -2 nts relative to each other were inequivalent in the two complexes.
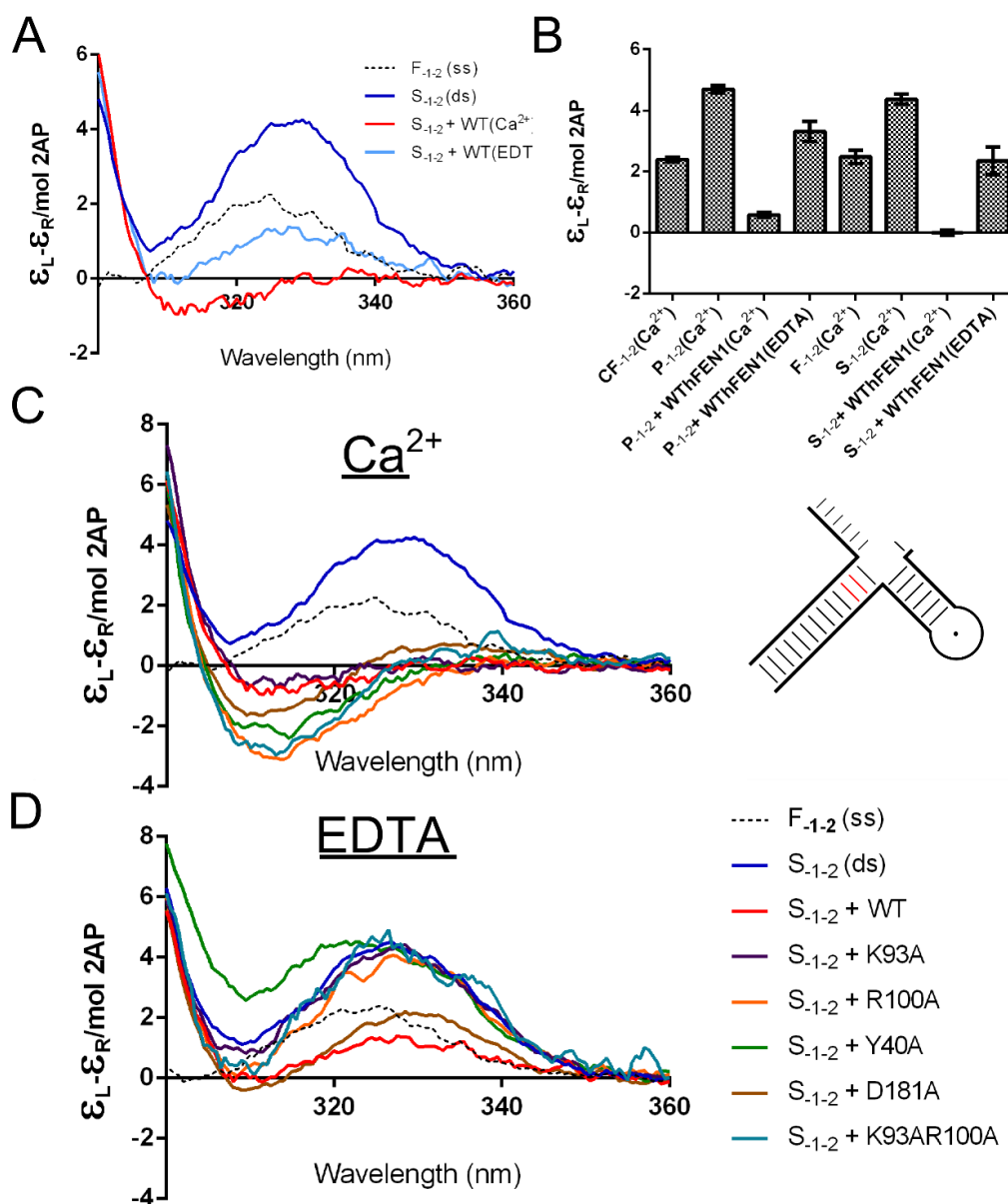
Figure 4.13: Low-energy CD spectra of 2AP-containing DNA substrate construct and hFEN1 WT- and variant-DNA substrate complexes. (A) *Unbound S$_{-1-2}$ (blue), the corresponding single strand (F$_{-1-2}$, gray) and S$_{-1-2}$ bound to WThFEN1 (red) all in Ca^{2+} containing buffer. S$_{-1-2}$ bound to WThFEN1 in buffer containing no available divalent metal ions (10 mM Ca^{2+} + 25 mM EDTA) (cyan).* **(B)** *Ellipticity changes at 326 nm for constructs containing 2-AP dimer probes. Ellipticity changes are per mol 2-AP. All data sets are two or more experimental repeats, standard errors are shown.* **(C)** *CD spectra recorded in buffer containing 10 mM Ca^{2+} ions, unbound S$_{-1-2}$ (blue), the corresponding single strand (F$_{-1-2}$, gray), S$_{-1-2}$ bound to WThFEN1 (red), Y40A (green), K93A (purple), R100A (orange), D181A (black) and K93AR100A (teal). CD intensity at 326 nm ((M AP)$^{-1}$·cm^{-1}): unbound S$_{-1-2}$ (4.10), unbound F$_{-1-2}$, (2.06), S$_{-1-2}$ bound to WThFEN1 (-0.12), Y40A (-0.49), K93A (-0.07), R100A (-1.60), D181A (0.36) and K93AR100A (-0.13).* **(D)** *CD spectra in buffer containing no available divalent metal ions (10 mM Ca^{2+} + 25 mM EDTA) same colouring as in (C). Samples were buffered at pH =7.5, and measurements performed at 20 ˚C; specific conditions given in section 2.7. Each measurement was independently repeated and gave equivalent results, spectra were normalized as describe in section 2.7. Substrate construct is illustrated schematically (red 2AP) with sequence shown in figure 4.5. Spectra were initially taken by NP, LDF, EJ and repeated by JCE.*

Based on crystallographic data, active site residues Lys-93, Arg-100 and Tyr-40 have been proposed to be important for the double nucleotide unpairing mechanism of hFEN1 (Tsutakawa et al., 2011). The CD spectrum of S_{-1-2} bound to K93A in comparison with WT in the same buffer conditions was essentially identical, implying K93 was not required for positioning of the nucleotides flanking the scissile phosphate of substrate DNA (*figure 4.13c*). Therefore the recent proposal that this basic residue was important for promoting the unpairing process is incorrect.

R100A and Y40A display similar ECCD behaviour, produce a spectroscopic change similar to WT when bound to substrate (S_{-1-2}) (*figure 4.13c*). However, although the intensity at 326 nm for WT and the mutated FENs was essential zero, the depth of the trough a 313 nm was considerably deeper for R100A and Y40A (*figure 4.13c*). This implied that the -1 and -2 nts were unstacked with each other when bound to R100A or Y40A in the presence of Ca^{2+} ions but the equilibrium position was different to that in the WT complex. Interestingly, the addition of K93AR100A in the presence of Ca^{2+} ions produced a decrease in the exciton pair signal, which indicated that the substrate had undergone an analogous conformational change to that when bound to R100A(Ca^{2+}) (*figure 4.13c*). This implied that when bound to K93AR100A the -1 and -2 nts were in the same position as with R100A.

Given that D181A gave a unique result with P_{-1-2} in the presence of Ca^{2+} ions, it was worthwhile to determine what effect this mutation had on the exciton-coupling between the 2-APs in the substrate construct. Examination of the D181A:S_{-1-2}:Ca^{2+} complex CD spectrum (*figure 4.13c*) shows no unexpected absorption bands, the CD response was equivalent to WThFEN1 and the other mutated FENs. Firstly, this showed that the ability of hFEN1 to unpair the _substrate_ was unaffected by any perturbation of the metal ion sequestration in D181A. Secondly, this validated the D181A:P_{-1-2}(Ca^{2+}) result proving that increase in intensity of the trough at 310 nm was a *bona fide* response.

Addition of EDTA to the S_{-1-2}:WThFEN1(Ca^{2+}) sample returned the signal near to that observed for free ss F_{-1-2} (*figure 4.13b*). The intensity of the exciton-coupling for the S_{-1-2}WT(EDTA) complex was 2.35 ± 0.49 (M AP)$^{-1}\cdot$cm^{-1} , the equivalent product complex had an intensity of 3.31 ± 0.33 (M AP)$^{-1}\cdot$cm^{-1}. This suggests that the 5'-flap and/or any potential interaction hFEN1 has with this part of the substrate affects the dynamics of the dsDNA adjacent to the junction, shifting the equilibrium slightly towards the unpaired state.

Reminiscent of the product construct (*figure 4.12b*), the signal for substrate was unique when bound to WThFEN1 compared to mutated FENs in the absence of Ca^{2+} ions. The signal produced for S_{-1-2} bound to either K93, R100A, Y40A or K93AR100A was identical to the unbound S_{-1-2}

construct (*figure 4.13d*). This implied that when the residues proposed to be important for unpairing were present, even without divalent metal ions, the 2-APs could unpair, albeit to a lesser extent. Consistent with this conclusion, the D181A mutation was much more similar to WT than the other mutated FENs in the absence of metal ions (*figure 4.12d*). In EDTA the D181A mutation is essentially *invisible* because D181A, just as WT, has the same residues implicated in the unpairing mechanism.

4.6.2 Low energy CD of 2-AP probes located in the double-stranded downstream region of substrates

The 2-APs were moved to a position in the DNA construct which was known from crystal structures not to undergo a conformational change upon hFEN1 binding at the -8 and -9 position relative to the scissile phosphate (*figure 4.5*). The control substrate was an isomer of the S_{-1-2} substrate, and had had the same 5'- and 3'-flaps and also the same length downstream and upstream regions. Instead of the 2-APs being flanked by guanine on both sides it was flanked by guanine on the 5′-side and alanine on the 3'-side, as a result of maintaining the G:C content. The low energy CD spectra of S_{-8-9} was consistent with $(2\text{-AP})_2$ buried within B-form dsDNA (*figure 4.14*). This CD response does not significantly change on complexation with WThFEN1 either in the presence of absence of Ca^{2+} ions. This suggests that the adjacent 2-APs remain stacked with one another when located far from the ss-ds junction of a double flapped substrate.
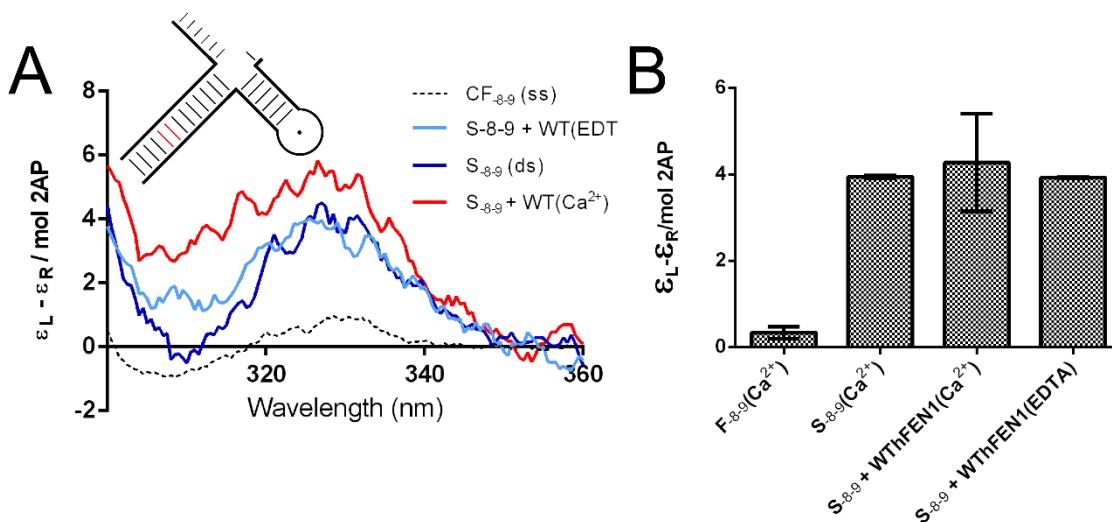
Figure 4.14: Low energy CD spectra of substrate containing 2-AP dimer in a region away from proposed site of conformational change. (A) Unbound S_{-8-9} (blue), the corresponding single strand (CF_{-8-9}, gray) and S_{-8-9} bound to WThFEN1 (red) all in Ca^{2+} containing buffer. S_{-8-9} bound to hFEN1 in buffer containing no available divalent metal ions (10 mM Ca^{2+} + 25 mM EDTA, cyan). (B) Ellipticity changes at 326 nm for constructs containing 2-AP dimer probes. Ellipticity changes are per mol 2-AP. All data sets are two or more experimental repeats, standard errors are shown. Samples were buffered at pH =7.5, and measurements performed at 20 ˚C; specific conditions given in section 2.7. Substrate construct is illustrated schematically (red 2AP) with sequence shown in figure 3.5. Each measurement was independently repeated and gave equivalent results, spectra were normalized as describe in section 2.7.

4.6 Investigating the unpairing situation at the +1 and -1 positions using 2-aminopurine

The tandem 2-APs were initially placed at the -1 and -2 positions so that the ECCD signal could be related to the nucleotide conformations in the hFEN1:product complex crystal structure. In the interpretation of CD experiments, thus far, it has been assumed that as the position of -1 nt is changing relative to the -2 nt, and looks unstacked when bound to hFEN1. Less well characterised is the position of the +1 nt, it has been assumed that the +1 nt is also unpaired and in a similar local environment to the -1 nt. Using ECCD, the position of the +1 nt in the S_{-1-2} construct is spectroscopically invisible. Therefore, to gain more information about the relative stacking between the two nts proposed to unpair the S_{+1-1} construct was designed (*figure 4.5*).

This substrate had identical dimensions to the substrate used previously, a 13 nt ds downstream region, 6 nt ds upstream region capped by a GAA hairpin, 1 nt 3'-flap and a 5 nt ss 5'-flap. The G:C content of the downstream region was kept constant to maintain the stability, a consequence of this meant that the neighbouring guanine was switched from the 5'-side to the 3'-side. The major difference was that the 2-AP dimer had been moved from the -2-1 position to

the +1-1 position. As previously discussed (*section 4.4*), the low-energy CD spectrum of unbound S_{+1-1} compared to unbound S_{-1-2} suggests that the first nucleotide of ds DNA is less stacked than others in the duplex either in the presence or absence of Ca^{2+} ions (*figure 4.15*).

The low-energy CD spectrum of the WT:$S_{+1-1}(Ca^{2+})$ did not produced the dramatic decrease of CD intensity as previously observed for $(2\text{-}AP)_2$ containing constructs bound to hFEN1 in the presence of divalent metal ions (*figure 4.16a*). When complexed with WT(Ca^{2+}) the double stranded substrate had the same CD intensity at 326 nm as the unbound single stranded oligonucleotide (*figure 4.16b*). It has been established in this study and previously using exciton-coupling, and using complementary techniques, that there is considerable stacking between the nts in ss DNA. Thus, the data presented here implies the +1 and -1 nts when bound to hFEN1 are based stacked with one another; this does not mean they are not unpaired. The data with -1-2 strongly implies that the -1 nt is unpaired therefore taken together, if +1 nt is stacked with -1 nt then this too must be unpaired. This implies that hFEN1 binding stabilizes the double-nucleotide unpaired state preferentially over the single-nucleotide unpaired conformation (*figure 4.15*). However, as previously appreciated measurement of exciton-coupling by this manner is an equilibrium technique. Therefore, the resolution of ECCD is insufficient to observe an initial unpairing event at the +1 position followed by unpairing of the -1 nt.

A modest but statistically significant increase in the intensity was observed for the WT:S_{+1-1} complex in the absence of Ca^{2+} ions (*figure 4.16b*). This again suggests that the divalent metal ions are helping to neutralise the highly negatively charged active site and in doing so assists nucleotide unpairing. Due to time constraints but also because changes in the signal for S_{+1-1} are more subtle than S_{-1-2} it was decided that it would not be advantageous to characterise the unpairing behaviour of all hFEN1 variants. In addition, because the -1-2 doesn't look like the +1-1 there was no advantage in moving the 2-AP dimer further into the downstream region as -2 nucleotide is not unpairing so presumably will not see a change for a S_{-2-3} complex.
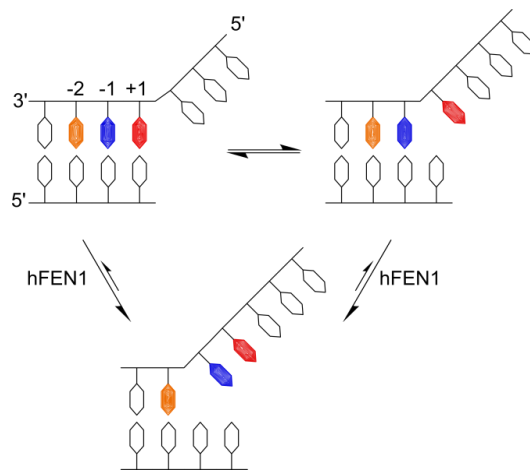
Figure 4.15: A double base unpairing in the FEN mechanism is suggested by the low-energy CD spectra of -1-2 and +1-1 (2-AP)₂ WT substrate complexes. *In solution the nucleotides at the ss-ds junction are in equilibrium between base-paired (top left) and a single nt (+1) being unpaired (top right), hFEN1 binding to either conformations shifts the equilibrium towards a double-nucleotide unpaired conformation.*
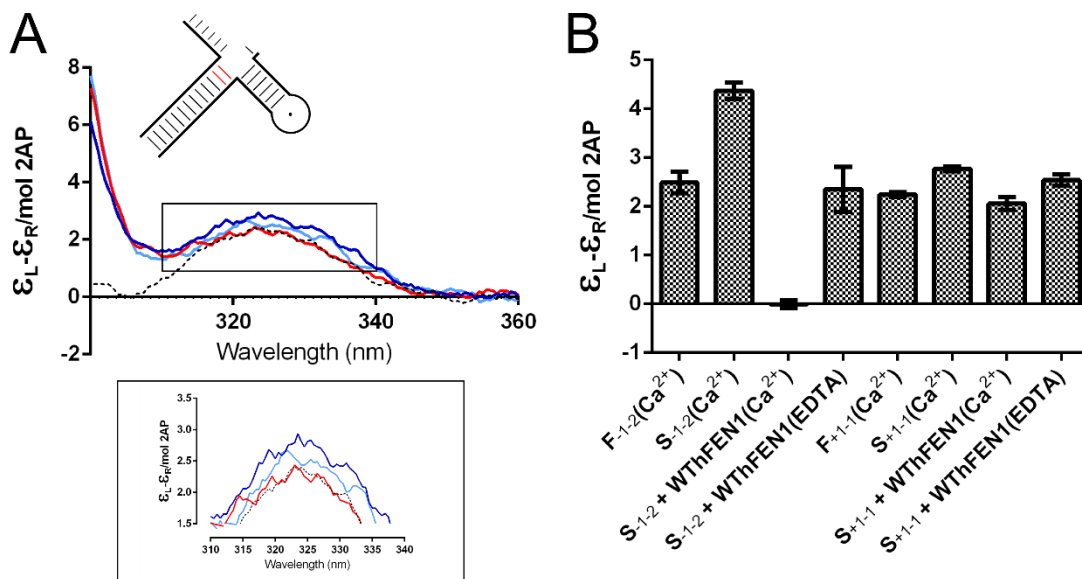
Figure 4.16: Low-energy CD spectra of DNA substrate construct containing 2-AP at the +1 and -1 positions and hFEN1 WT- complexes. (**A**) *Unbound S_{+1-1} (blue), the corresponding single strand (F_{+1-1}, gray) and S_{+1-1} bound to WThFEN1 (red) all in Ca^{2+} containing buffer. S_{+1-1} bound to WThFEN1 in buffer containing no available divalent metal ions (10 mM Ca^{2+} + 25 mM EDTA) (cyan). Insert magnified view of low-energy CD traces at region of maximum intensity. (**B**) Comparison of ellipticity changes at 326 nm for constructs containing 2-AP dimer probes at the -1-2 and +1-1 positions, respectively. Ellipticity changes are per mol 2-AP. All data sets are two or more experimental repeats, standard errors are shown. Substrate construct is illustrated schematically (red 2AP) with sequence shown in figure 4.5, note position of 2-AP dimer has moved to the +1 and -1 position.*

4.7 Mismatches disrupt the unpairing mechanism

Incorporation of base-pair mismatches as a result of DNA replication and recombination events can reduce genome integrity. Despite the high fidelity of replication polymerases, the chance of an incorrectly incorporated base can result in ~1000-mismatched per cell division (Lee et al., 2014). This level of mutation obviously cannot be tolerated, therefore there is a requirement for mismatches to be identified and dealt with by the relevant DNA repair pathway. To investigate the ability of hFEN1 to process a substrate which mimics the product of a mis-incorporation event a double-flap construct containing a mismatch at the terminal base-pair of the flap strand was created. The nomenclature adopted is consistent with the previous substrate constructs, a C:C mismatch at the +1 position (MM_{+1}) and 2-AP at the -1 and -2 positions ($MM_{+1}S_{-1-2}$, *figure 4.5*)

Upon assembly of the $WT:MM_{+1}S_{-1-2}(Ca^{2+})$ complex the positive band reached a maximum intensity of 3.51 ± 0.18 (M AP)$^{-1}\cdot$cm^{-1} at 326 nm (*figure 4.17b*), compared to -0.01 ± 0.08 (M AP)$^{-1}\cdot$cm^{-1} for the $WT(Ca^{2+})$ complex with the fully Watson-Crick base-paired substrate. This implied that the position of the -1 nt relative to the -2 nt was not the same in both complexes. Furthermore, binding of hFEN1 to $MM_{+1}S_{-1-2}$ in the presence of Ca^{2+} produced a similar phenomenon to that observed for $WT(Ca^{2+})$ binding to the S_{+1-1} construct. Although the spectral band remained slightly blue-shifted the height of the peak was essentially identical to the corresponding single strand ($MM_{+1}F_{-1-2}$) (*figure 4.17a*). This modest decrease in excition-coupling can be explained using the same principles inferred for the S_{+1-1} construct. Thus, suggesting that the -1 nt and -2 nt remain stacked, but again not necessarily base-paired. Conceivably, the situation could occurred whereby FEN 'mistakes' the mismatched base-pair as being part of the 5'-flap, therefore shifting the position of the scissile phosphate so now it is between the -1 nt and -2 nt. Previous data has shown that at mismatch at the same position can be tolerated; however, the rate of reaction was considerably reduced compared to the equivalent substrate containing a complementary base pair at the +1 position (Beddows et al., 2012). In addition, the reaction site specificity was also reduced, and the production of shorter and longer 5'-reaction products was observed. Longer products are a result of cleaving further into the ds downstream region and this could be rationalized by unpairing of the +1, -1 and -2 nts. This proposal could be tested in future by creating a substrate construct with a mismatch at the same +1 position but moving the 2-AP dimer to the -2-3 positions. If the -2 nt was unpairing due to the mismatch then a change in the excition coupling would be observed for this substrate, similar to that observed for S_{-1-2}, upon hFEN1 binding.
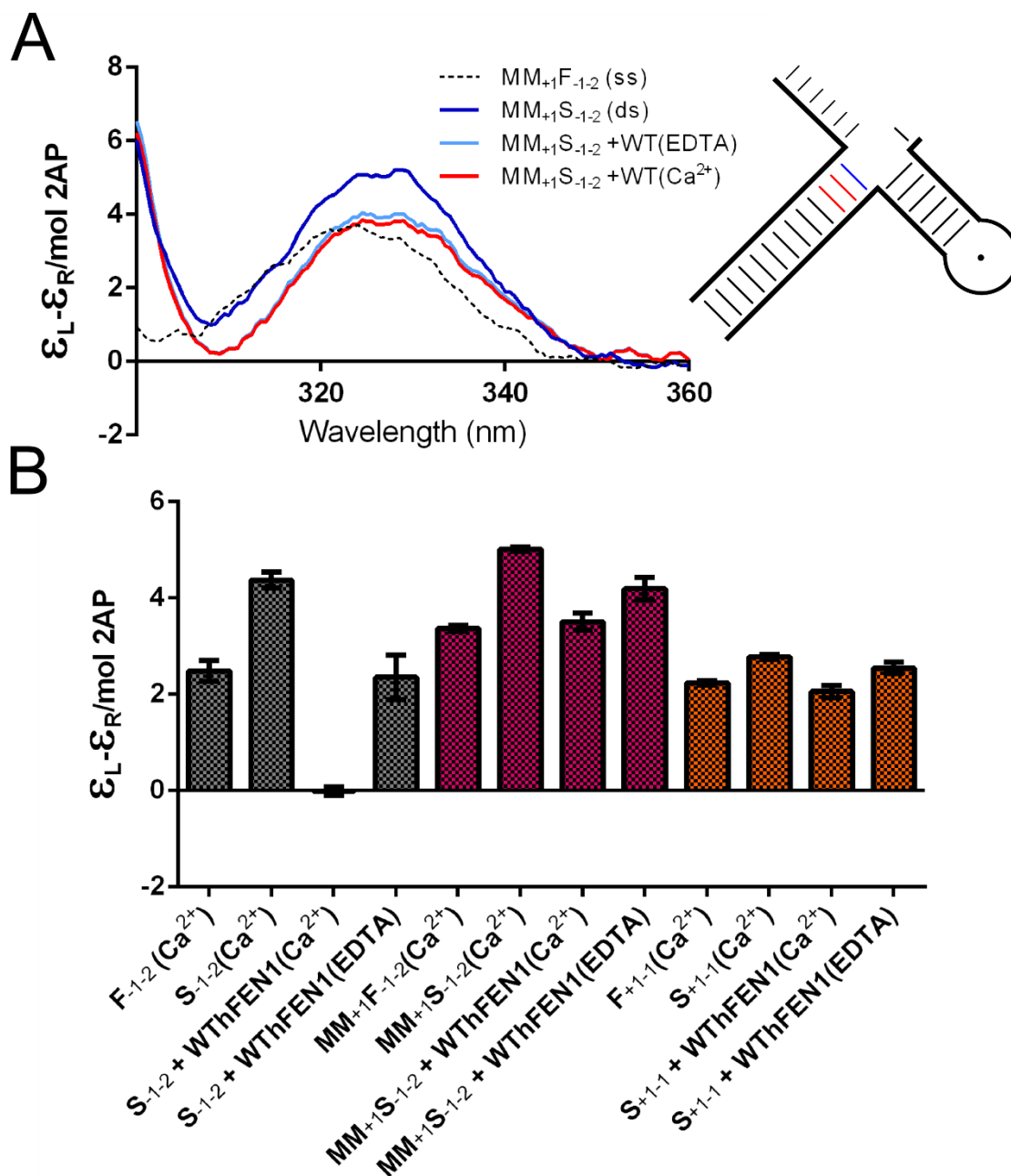
Figure 4.17: Low-energy CD spectra of DNA substrate construct containing 2-AP at the -1 and -2 positions and a C:C mismatch at the +1 position unbound and in hFEN1 WT- complexes. (**A**) *Unbound $MM_{+1}S_{-1-2}$ (blue), the corresponding single strand ($MM_{+1}F_{-1-2}$, gray) and $MM_{+1}S_{-1-2}$ bound to WThFEN1 (red) all in Ca^{2+} containing buffer. $MM_{+1}S_{-1-2}$ bound to WThFEN1 in buffer containing no available divalent metal ions (10 mM Ca^{2+} + 25 mM EDTA) (cyan). Insert magnified view of low-energy CD traces at region of maximum intensity. (**B**) Comparison of ellipticity changes at 326 nm for constructs containing a 2-AP dimer at the +1-1 (orange) and -1-2 with (magenta) and without (gray) a mismatch at the +1 position. Ellipticity changes are per mol 2-AP. All data sets are two or more experimental repeats, standard errors are shown. Substrate construct is illustrated schematically (red 2-AP, blue C:C mismatch) with sequence shown in figure 4.5.*

4.8 Summary and conclusions of ECCD data

The 2-aminopurine replacement studies described in this chapter were carried out with the aim of investigating the unpairing mechanism of the hFEN1-catalysed reaction. Product and substrate constructs containing $(2\text{-AP})_2$ can reveal the conformation of these nucleotides within DNA when bound to hFEN1. The low-energy CD spectra of 2-AP residues inserted into DNA reflect base conformations at the sites of the probes. This is a well characterised technique, and has been used to study the dynamics of nucleotides in a variety of nucleic acid structures (Johnson et al., 2004; Jose et al., 2009). The advantage of this technique is that it can be exploited to study the conformation of specific nucleotides in the context of large protein:DNA complexes (Datta et al., 2009; Datta et al., 2010; Jose et al., 2012). To date the focus has largely concentrated on polymerases but work in this chapter has shown that this technique can be extended to a different class of nucleic acid enzymes hopefully leading to wider application of this methodology.

In this study of hFEN1, 2-AP substitution of nts in substrate and product constructs revealed how FEN is able to select for, then place the scissile phosphate within the active site using a double nucleotide unpairing mechanism. The extent of unpairing can be quantified by the amount of exciton-coupling. Comparison of ss and ds spectra collected here, and with previously published data shows that the nts at the ss-ds junction are in relatively well stacked environment. The results show that the hFEN1 induced bending a DF substrate shifts the equilibrium position of the nts at the ss-ds junction towards an unpaired state. However, substantial unpairing of the targeted nts requires divalent metal ions bound in the active site of hFEN1. The position of the nucleotides either side of the scissile phosphate in the substrate can be inferred from the product crystal structure. The CD spectra of substrate and product DNA with the $(2\text{-AP})_2$ in equivalent positions are essentially identical. The terminal nt of the product flap strand is pinned on the active site metals in the crystal structure, thus presumably so is the equivalent nucleotide in the substrate when bound to $hFEN1(Ca^{2+})$.

The data presented in this chapter represents one of the first directly observed unpaired hFEN1:substrate intermediates. A complementary study conducted by Amanda Beddows used time resolved fluorescence (TRF) spectroscopy of similar product/substrate constructs containing a single 2-AP substitution to investigate the ability of hFEN1 to unpair DNA (Finger et al., 2013). However, the interpretation of results with WThFEN1 was complicated by the interference of the Tyr-40 residue which is capable of quenching the AP-florescence through stacking interactions, a phenomenon which has previously been observed for other nucleic acid

binding enzymes studied using this technique (Neely et al., 2009). The involvement of Tyr-40 was also believed to complicate the analysis of the product/substrate unpairing when complexed with hFEN1 variants. Here, low-energy CD spectroscopy, which is not affected by charge-transfer quenching by Y40, has been used to investigate the effect of active site mutations on the ability to unpair.

5'-nuclease superfamily strictly conserved (K93, R100) and semi-conserved residues (Y40), whose involvement in the hFEN1 catalyzed reaction are required for reaction to occur at a biologically relevant rate have been proposed to be involved in the unpairing mechanism. However, data presented here suggests that these residues are not essential for 'crude' substrate unpairing and active site divalent metal ions alone are sufficient to coax nucleotides out of duplex DNA. It has been proposed that in addition to catalysis, active-site metal ions provide specificity through positioning of the scissile phosphate diester in nucleases containing two (or more) metal ions (Yang et al., 2006), and FEN1 appears to be an example of this. However, exact positioning of the nucleotides does appear to be sensitive to the presence of R100A and Y40A but not K93A. The difference between the conformation of the selected nucleotides when bound to K93A and R100A is intriguing. Single-turnover kinetic data suggested Lys-93 and Arg-100 act in a concerted fashion. Reaction cannot occur optimally without the input of K93, based on a T5FEN study, K93 has been proposed to act as an electrophilic catalyst. The position of the nucleotides appears to be similar in the R100A and K93AR100A variants based on the presented ECCD data. Therefore, taken together these studies imply that Arg-100 is partly responsible for the nucleotide positioning permitting the involvement of Lys-93.

In stark contrast, both Lys-93 and Arg-100 are absolutely required to stabilize the single unpaired nucleotide observed in the wild-type product complex. This likely hints at unpaired intermediates further along the reaction course which the product mimics more closely than the substrate, an intermediate positioned correctly to react. Alternative unpaired states are evident in TRF data, most clearly seen in the ability of Tyr-40 to interact with both +1 and -1, likely requiring a conformational change. Partial unpairing of the product DNA was observed with Y40A-Ca^{2+}, and the absence of Y40 appears to alter the partition between paired and unpaired forms of the substrate and/or the relative conformation of these nts, which is further evidence for this stacking residue stabilizing the unpaired state in both complexes. Placement of the spectral probe at both positions proposed to undergo a conformational change (+1 -1) opposed to at the paired/unpaired divide (-1 -2) suggests that the first two nucleotides unpair in a concerted manner.

Of all the point mutations discussed in this section, the D181A mutation is the most detrimental towards catalysis (*section 3.4*). However, due to its proposed role in active-site metal ion sequestration and the prerequisite for there to be bound in the active site to generate unpaired nucleotides, the ability of D181A to successfully unpair the substrate is somewhat surprising. The active site of D181A is proposed to be partly occupied with M^{2+} ions. In support of this D181A cannot unpair the product DNA. The 5'-terminal phosphate of the product DNA has a -2 charge, whereas the charge on the scissile phosphate in the substrate is -1, thus the suspected partial M^{2+} occupancy of the D181A active site may be able to compensate for the charge imbalance of the DNA backbone of the substrate in the active site but not the product. Furthermore, maybe one divalent metal ion is sufficient for initial substrate unpairing but two or more are required for catalysis.

A mismatch was introduced into the substrate to determine if hFEN1 unpairing was specific for properly replicated Okazaki fragments. The relative orientations of the -1 and -2 nts when bound to hFEN1(Ca^{2+}) differ depending on the stability of the base-pair immediately at the ss-ds junction. Paradoxically a more stable bp at the base of the flap translates into a larger conformational change upon FEN1(Ca^{2+}) binding. It is possible that FEN1 uses the unpairing mechanism to recognize replication errors. The nucleotides in a fully Watson-Crick based paired substrate settle in a different position than nucleotides adjacent to a mismatch. Armed with the knowledge that the rate of reaction is slower for a similar mismatch substrate would suggest the reason for this is that it is not positioned optimally for reaction (Beddows et al., 2012). Therefore, FEN1 selects for stable DF structures, decreasing the chance of processing a potentially mutagenic mismatch.

The obvious next line of inquiry would be to use ECCD to determine if double-nucleotide unpairing is conserved across the FEN superfamily. A double-nucleotide unpairing mechanism is an attractive model because it unifies how different family members which act upon very different nucleic acid structures recognise their respective substrates because of their ability to unpair at the ss-ds junction. The low-energy CD experiments were first conceived based upon the differences between the substrate and product complexes. These complexes were crystallised with the trivalent metal ion samarium (Sm^{3+}) in the active site, therefore, for completeness it would be interesting to see if the same unpairing equilibrium was reached when replacing Ca^{2+} with Sm^{3+} as the cofactor.

Chapter 5: An Unexpected Role for the Helical Arch in FEN1 Substrate Unpairing

5.1 The possible roles of the FEN1 helical arch

The so-called helical arch (α4 – α5) straddles the active site of FENs. The role of this portion of FEN1 has remained elusive. Different models have been proposed for the role of the helical arch in the FEN1 reaction mechanism (Bornarth et al., 1999; Murante et al., 1995; Orans et al., 2011; Patel et al., 2012). An earlier substrate recognition model where flapped DNA initially became threaded through the helical arch allowing the protein to slide along until it contacted duplex is now discredited (Liu et al., 2004). Instead, both structural and biophysical studies suggest the main interface between FEN1 and its substrates are to the dsDNA parts (Orans et al., 2011; Tsutakawa et al., 2011). Another suggestion is that the arch acts as a clamp to force the presumably unpaired substrate into a reactive conformation. In this clamping model the flap portion of substrates would have to wind round one of the arch helices to leave the active site (*figure 5.1a*, route 2) (Hosfield et al., 1998a; Liu et al., 2006). Alternatively, in another model, the substrate departs from the active site through the helical arch (Patel et al., 2012). This so-called threading model would explain the selectivity of FEN1 for substrates that have free 5'-termini (*figure 5.1a*, route 1).

More recent analyses have divided the helical arch into two regions, a helical gateway and a helical cap (*figure 5.1b*). The helical gateway comprises of the base of α4, harbouring the superfamily conserved basic residues (K93 and R100 in hFEN1), with the other "post" of the gateway formed by the base of α2, which provides the Tyr stacking residue for unpaired DNA. A more detailed review of FEN structure is given in *section 1.3*, but the gateway is ascribed the role of ensuring that only single stranded (unpaired) DNA reaches the active site. The helical cap, is not conserved across the superfamily and is composed of the C-terminal region of α4 and the entirety of α5; (~30 aa). The helical cap is present in the FEN1 family member hEXO1, but is not sequence conserved. It is this feature of the arch that is thought to confer specificity for substrates with 5'-termini (Patel et al., 2012). In other superfamily members XPG and GEN1 that process continuous DNA bubbles and four way junctions the connection between the base of α4 and α5 is either much longer (700 aa) or much shorter, respectively. In substrate-free hFEN1 structures the helical cap is not visible due to disorder. It has been noted that in ordered form only single stranded DNA could pass through the arch, but in disordered form the dimensions of the arch would allow passage of a duplex.

Although there is still controversy over the role of the helical arch, support for the threading mechanism has come from substrates where the 5'-flap DNA has been biotinylated to allow conjugation of streptavidin (SA). Using methodology that exploits the SA-biotin interaction, two distinct complexes can be formed (Gloor et al., 2010; Murante et al., 1995; Patel et al., 2012; Sobhy et al., 2013) that have distinct differences in binding equilibria and reaction kinetics. Using SA capture the substrate is accommodated in a way that prevents exchange of the bound DNA with an unlabelled competitor and readily undergoes hydrolysis when reaction is initiated with Mg^{2+}; this is termed a *trapped* complex. To form a trapped complex SA is added to 5'-biotinylated substrates in complex with FEN(Ca^{2+}) (*figure 5.3*). Alternatively, SA can be added to the substrate before it is bound to FEN(Ca^{2+}), termed a *blocked* complex (*figure 5.3*). In a blocked complex the substrate can be easily competed away and reaction is drastically slowed. It is assumed that in a trapped complex the 5'-flap is threaded through the helical arch. However, streptavidin is too large to pass through the helical arch even if the arch is disordered. Therefore, once threaded and conjugated with SA the 5'-flap will not be able to pass back through under the helical arch. However, in a blocked complex the 5'-flap DNA would not have been able to pass through the arch.
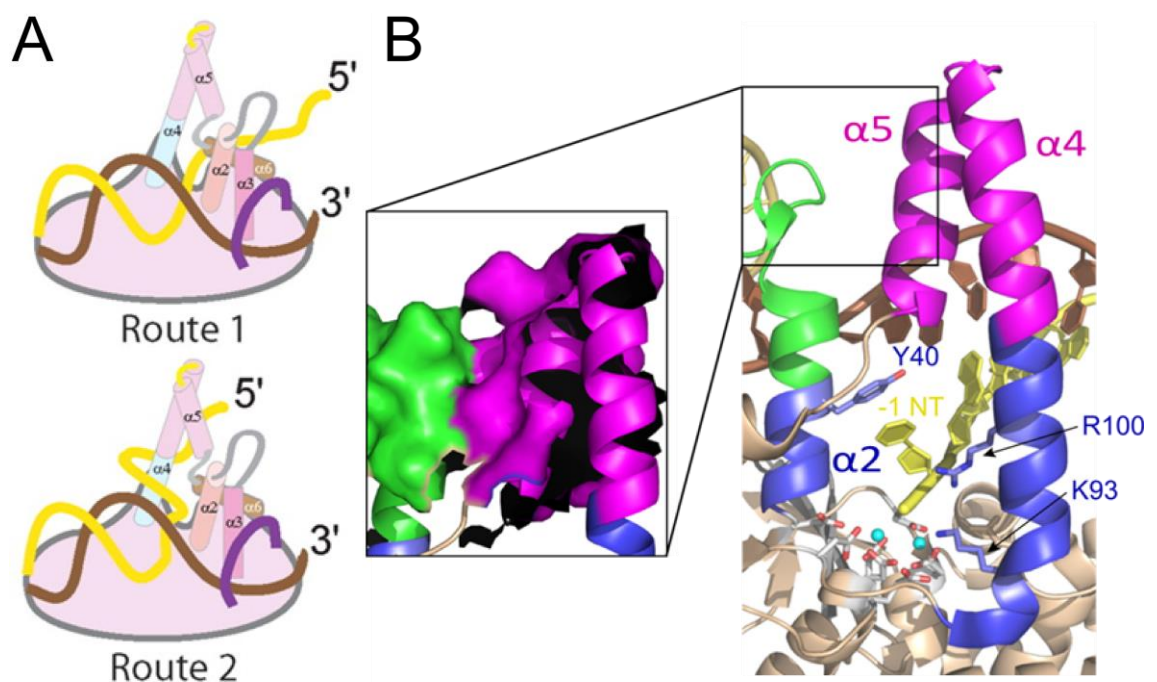
Figure 5.1: hFEN1 helical arch/cap architecture. *(**A**) Schematic models for departure of the 5'-portion of hFEN1 substrates from the active site involving threading (route 1) or clamping. Adapted from Patel, N. et al. (2012) NAR 40, 4507-4519 (**B**) hFEN1 product complex (3Q8K.pdb) showing magnified view of helical arch (α4-α5) which has been divided into two regions, the helical gateway (blue) and the helical cap (magenta). The α2-α3 loop is shown in green and the packing interactions with the helical cap are represented by surface projections of these regions (insert). Active site metal ions (cyan spheres), conserved carboxylates (red), and the gateway α4 K93 and R100 sidechains are shown (blue). Also highlighted is the α2 Y40 (blue) which stacks with the terminal (-1) nucleotide (yellow) in the product complex.*

155

5.2 Streptavidin does not affect enzyme and substrate association

The work detailed in this chapter investigated differences in trapped and blocked hFEN1:DNA complexes, therefore the FRET assay was implemented to determine association of biotin and streptavidin affected the affinity of hFEN1 for the substrate. A biotinylated version of the original DF-4 substrate was used to form a blocked complex. The substrate DF-B4 (*figure 5.2a*) was biotinylated on the 5'-terminus of the 5'-flap, and as with all oligonucleotides synthesised with a biotin the flap length was increased (from 5 to 10 nts) to minimize any interaction between hFEN1 and SA. The binding affinity of hFEN1 for a blocked substrate was measured in buffer containing calcium. The dissociation constant for DF-B4-SA was 53 ± 9 nM, only moderately decreased than compared with the uncomplexed substrate c.f. 20 ± 3 nM (*figure 5.2b*). The experiment also revealed that the bending of the substrate was not affected by the presence of SA. Moreover, the substrate could still bend presumably when the 5'-flap was not threaded through the helical arch. This is important because it shows threading is not a prerequisite for bending which is most likely a prerequisite for efficient unpairing at the ss-ds junction.
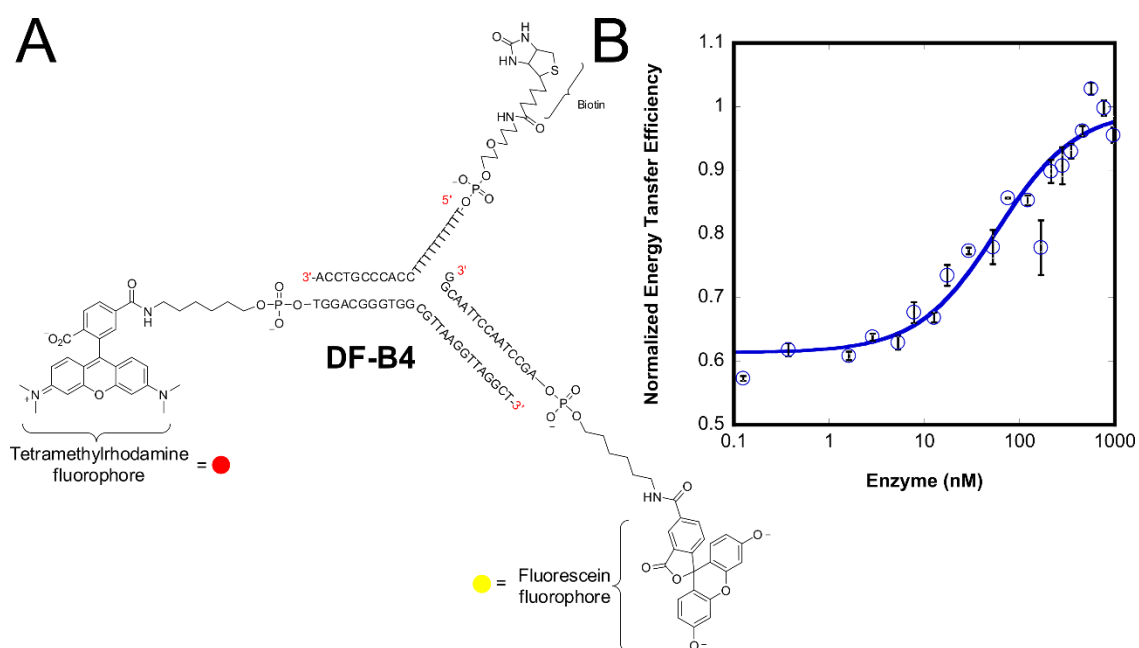
Figure 5.2: Threading is not a prerequisite for bending. **(A)** *Bitoinylated FRET substrate used to produce blocked complex.* **(B)** *Titration of hFEN1 into DF-B4 after pre-incubation with 5 molar excess SA. The titration was performed in 10 mM CaCl₂, 110 mM KCl, 55 mM HEPES pH=7.5, 0.1 μg/μl BSA, 1 mM DTT and at 37 ˚C. The data points were fit to equation 3.3 yielding a K_D value of 53 ± 9 nM and is representative of three repeats.*

5.3 Flap accommodation is a pre-requisite for unpairing

As detailed above, the two types of FEN1-substrate complexes that can be formed with streptavidin ("trapped" and 'blocked") have different properties. The aim of this section was to investigate whether these complexes are capable of unpairing their respective DNA substrates.

A 5'-biotinylated version of the substrate used in the ECCD experiments described in *chapter 4* was created (Sb$_{-1-2}$, *figure 5.3*), where 2AP is present at positions -1 and -2. The low-energy CD spectrum of the biotinylated (AP)$_2$ doped substrate construct in the presence of a 2-fold molar equivalent of streptavidin (Sb$_{-1-2}$-SA) is consistent with the spectroscopic behaviour of a 2-AP dimer in a stable B-DNA environment (*figure 5.4a*). EMSA was used to show that a 2-fold molar ratio of SA was sufficient to saturate the Sb$_{-1-2}$ construct (*Appendices, figure A3*). Moreover, the average CD intensities of S$_{-1-2}$ and its biotinylated-SA conjugated counterpart (Sb$_{-1-2}$-SA) were identical within experimental error, suggesting that the dynamics of the ss-ds junction are not altered when SA is bound at the terminus of the 5'-flap (*figure 5.4b*). Furthermore, comparison of the WT FEN1 complexes (± Ca^{2+}) with the biotinylated and non-biotinylated substrates without SA showed that there are no discernable differences in CD spectra, emphasized by identical per 2-AP residue ellipticity differences at the indicator wavelength (326 nm) (*figure5.4a*). This implies that the 5'-biotin moiety had no detectable effect on the base stacking interactions between the tandem 2-APs at -1 and -2 positions and does not interfere with the way in which hFEN1 interacts with the substrate. These results are in line with previous observations where 5'-biotinylated substrates have been shown to have no effect on the FEN catalysed reaction rate or on the affinity of the substrate when measured by fluorescence anisotropy (Patel et al., 2012). In summary, the ECCD results with the 5'-biotinylated substrate are identical to previous ECCD experiments on S$_{-1-2}$ and demonstrate the same dramatic hFEN1-Ca^{2+} dependent change in the base stacking interactions of the two 2-APs.

Trapped complex was formed by addition of streptavidin to WThFEN1:Sb$_{-1-2}$(Ca^{2+}). The trapped WThFEN1:Sb$_{-1-2}$(Ca^{2+}) complex and the unmodified WT:Sb$_{-1-2}$(Ca^{2+}) complex yielded very similar CD spectra (*figure5.4*). Moreover, these strongly resembled the spectra obtained with substrate without 5'-modification (S$_{-1-2}$) bound to FEN1(Ca^{2+}) described in detail in *chapter 4*. In both cases the ECCD signal was decreased to near zero at 326 nm. Thus, in WThFEN1:Sb$_{-1-2}$(Ca^{2+}) and trapped WT:Sb$_{-1-2}$(Ca^{2+}) complexes the DNA had undergone a conformational change that was likely the unpaired conformer. Moreover, like unmodified substrate complexes, the DNA conformational change was strictly dependent on the presence of divalent metal ions. In the presence of EDTA, both complexes had an appreciable ECCD signal at 326 nm. In contrast, when

proper accommodation of the 5'-flap was prevented in the blocked complex, the signal for the WT:substrate complex could not be replicated (± Ca^{2+}). Instead, there was a modest decrease in the CD intensity upon hFEN1 binding the substrate conjugated to SA *regardless of the* presence or absence of metal ions. (*figure 5.4a and b*). This implied that the DNA conformational change brought about by the FEN1-Ca^{2+} protein could not take place when the substrate was "blocked". Previously, the trapped complex had been interpreted as one where the flap is threaded through the helical arch, and the blocked complex is one where threading has been prevented. Based on this, the data strongly suggested that threading must occur before and is a pre-requisite for DNA unpairing.

The difference in CD signal between the trapped and blocked complexes in buffer containing EDTA is intriguing, if a little difficult to interpret. Evidently, the base-stacking between the 2-APs in the trapped complex was perturbed to a greater extent than in the blocked complex (both in the absence of divalent metal ions). One possibility is that despite the lack of divalent metal ions bound in the active site the +1 nt may still be capable of unpairing when the 5ʹ-flap is threaded having a knock-on effect of destabilising the -1 nt. This proposal could be investigated by repeating these experiments using the substrate with the 2-AP residues at the +1 and -1 positions (*section 4.6*).

As touched on briefly in *section 4.5.5*, a similar difference in EDTA was observed when the substrate was bound to WT and mutated FENs. Taken together these results imply that having a properly accommodated 5'-flap and the residues which are important for unpairing the product are sufficient to perturb the DNA structure substantially. Moreover, nucleobases at the +1 and -1 positions only feel the influence of helical arch residues (K93, R100 and Y40) and the active site metals if the 5'-flap is threaded through the helical arch. It is likely that the helical arch needs to be ordered to push down on the 5'-flap and 'force' the residues selected for unpairing into the active site, this is discussed in more detail in *section 5.4*.

The aspect of 5'-flap accommodation is not addressed in great deal by this thesis and more detail is given in the introduction (*section 1.5*). The method used to trap and block the 5'-flap and subsequent analyses could also be explained if the helical arch is used as a clamp, although the competition results suggest this process would need to be irreversible. As mentioned above, in this alternative model the flap passes close to the base of α4 but bends away, and never traverses the arch, then it has been proposed that the helical arch folds over the flap. The data present here are consistent with a threading mechanism and previously published data seem to favour this mechanism. However, it would be useful to repeat these experiments using

alternative techniques to probe threading rather than relying on a single biotin-SA interaction and as is typically the case until a crystal structure of a threaded complex is solved the debate will continue.
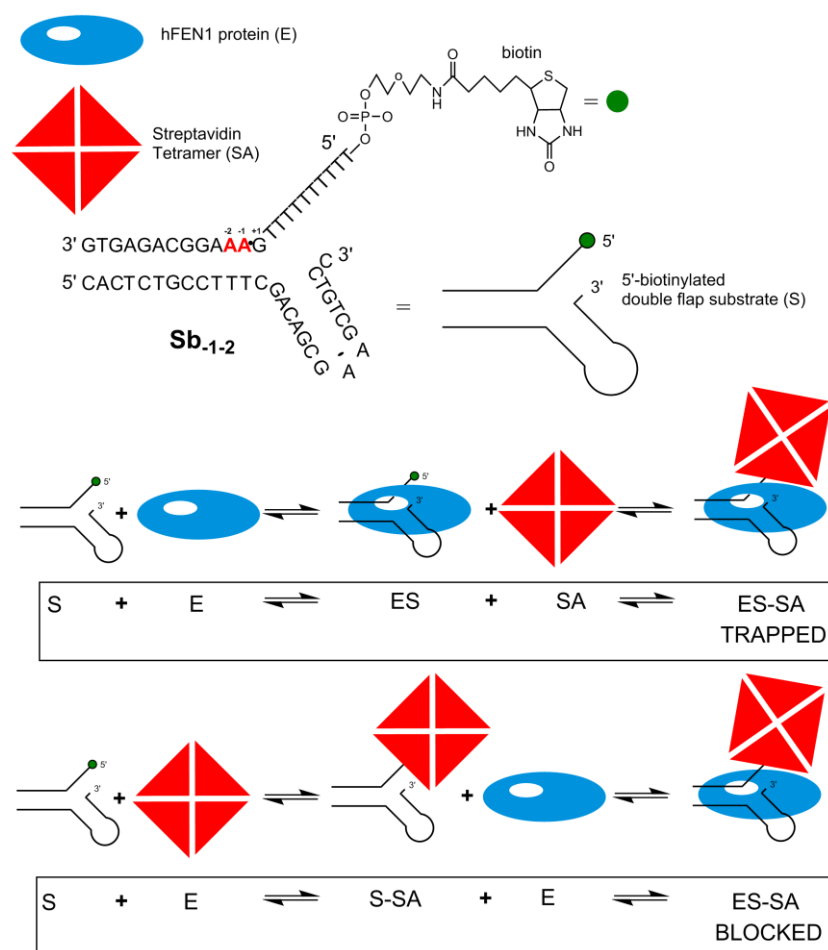
Figure 5.3: Structure of double flap oligonucleotide Sb$_{-1-2}$ and schematic showing the procedures involved in forming trapped and blocked complexes with ECCD substrate. *The substrate Sb$_{-1-2}$ is a heteroduplex formed from a template and flap strand (table 2.5). The template strand forms a fold-back hairpin and is capped with a [GAA] tri-loop. The scissile phosphate (black circle) connects the +1 and -1 nts. The biotin moiety is represented by a green circle and 2-aminopurine nucleotides are depicted in red. A trapped complex was formed by incubating enzyme and substrate together for 10 mins in buffer containing Ca^{2+}, then SA was added and incubated for a further 10 mins. Incubation of substrate and SA for 10 mins prior to the addition of enzyme, and subsequent incubation for 10 mins, led to the formation of a blocked complex. For ECCD experiments all incubations were performed on ice to reduce the chance of phosphate diester hydrolysis occurring.*
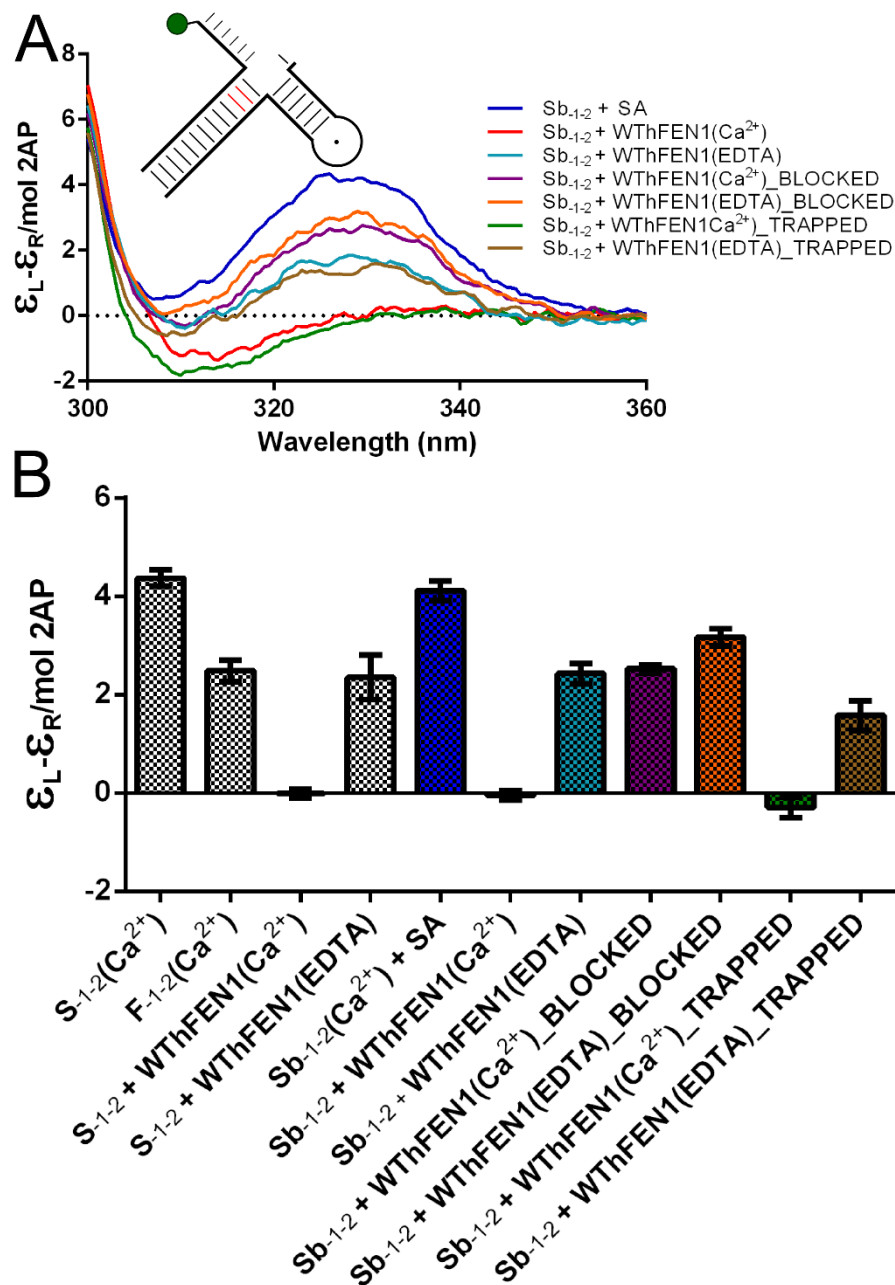
Figure 5.4: Low-energy CD spectra of trapped and blocked hFEN1 complexes. (A): *Unbound Sb-1-2 in presence of 5-fold molar ratio SA (blue), Sb-1-2 bound to hFEN1 (red) Sb-1-2 in blocked hFEN1 complex (purple) and Sb-1-2 in trapped hFEN1 complex (green) all in Ca²⁺ containing buffer. Sb-1-2 bound to hFEN1 (cyan) Sb-1-2 in blocked hFEN1 complex (orange) and Sb-1-2 in trapped hFEN1 complex (gold) all in buffer containing no available divalent metal ions (10 mM CaCl₂ + 25 mM EDTA). Samples were buffered at pH =7.5, and measurements performed at 20 ˚C; specific conditions given in section 2.7.3. Each measurement was independently repeated and gave equivalent results, spectra were normalized as described in section 2.7.1. (**B**) Comparison of ellipticity changes at 326 nm with non-biotinylated and biotinylated substrate constructs containing 2-AP dimer probes. Ellipticity changes are per mol 2-AP. All data sets are two or more experimental repeats, standard errors are shown. Substrate construct is illustrated schematically (red 2AP and green biotin) with sequence in figure 5.3.*

5.3 Investigating the effect of helical arch variants on catalysis and substrate binding

To investigate the role of the helical arch a series of proline mutations (L97P, L111P, L130P) were introduced within the cap and gateway regions. Insertion of a proline residue into an alpha helical region can potentially disrupt the secondary structure of the protein due to the unique secondary amine. The conformational rigidity and lack of hydrogen bonding potential causes a kink (45-50˚) in the polypeptide chain that can be accommodated with the first four alpha helix residues but beyond that a proline residue will potentially break the secondary structure (Sauer et al., 1992). The selected mutations (L97P, L111P, L130P) were of particular interest because they were initially identified from a dominant negative hFEN1 mutation screen in yeast (Storici et al., 2002). Each of the leucine residues are semi conserved across eukaryotic FEN1 enzymes (*figure 5.5*) (Patel, 2012). L97 is located at the base of α4 between the two superfamily conserved basic residues K93 and R100, whereas L111 and L130 are located in the helical cap region at the top of α4 and end of α5, respectively (*figure 5.5*).

The nature of the proline mutations that were employed as helix breakers meant it was important to determine if they had grossly altered the protein structure. CD spectra of each proline variant in the presence of M^{2+}, closely resembled that of WThFEN1 implying that there are no significant changes in the overall secondary structure (Patel, 2012). This result was to be expected because the helical arch of apo FEN is thought to be disordered. In addition, tryptophan fluorescence denaturation titrations, performed by Nikesh Patel, could be fit to a two-state unfolding transition and showed that L97P, L111P and L130P were all as stable as, or more stable than, WThFEN1.
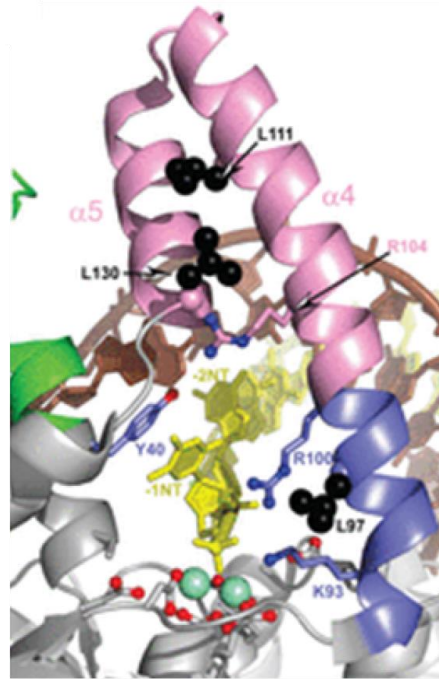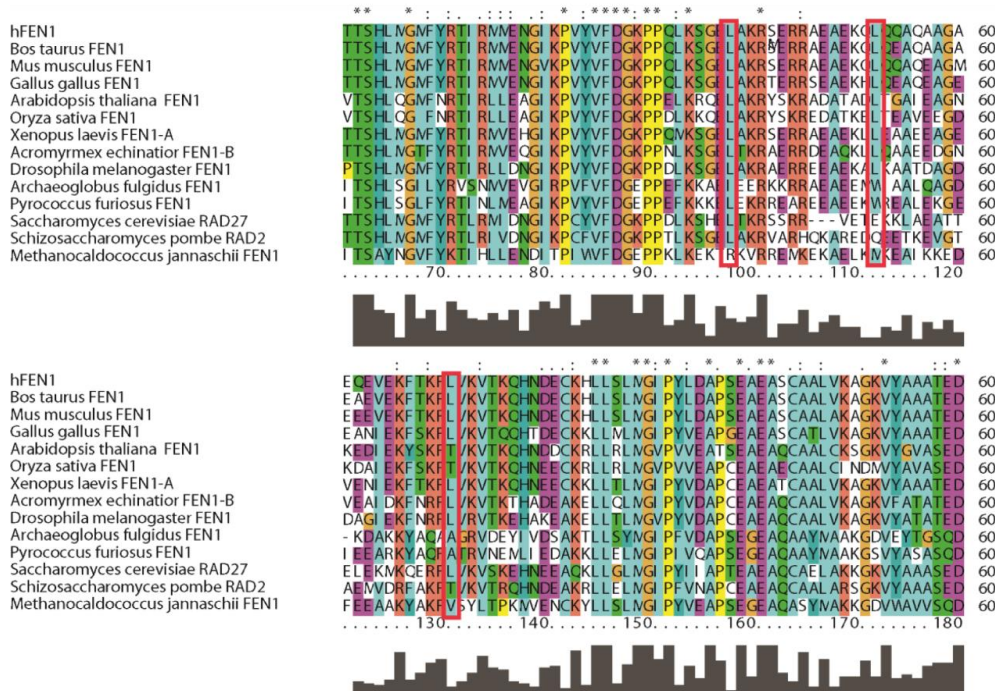
Figure 5.5: The position of semi-conserved leucine residues in the helical arch/cap hFEN1 structure. (A) *Position of the Leu residues (black) mutated in this study (L97P, L111P and L130P) in hFEN1 product structure (3Q8K.pdb). Same colouring as in figure 5.1. **(B)** Sequence alignment of residues 60-180 (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. Courtesy of Nikesh Patel.*

As part of his PhD thesis work Nikesh Patel characterised the *in vitro* effect these mutations had on the maximal single turnover rate using optimal static double flap substrates (*table 5.1*). Unsurprisingly, the mutation which had the most dramatic effect was L97P, located between

162

the two superfamily conserved basic residues. L97P decreased the rate of reaction >100,000-fold with respect to the WT hFEN1 catalysed reaction of a double flap substrate. This result could be partly rationalised by preventing the participation of Lys-93 and Arg-100 in catalysis. To test this I created the K93AR100A double variant (*Chapter 3*). However, K93AR100A only decreased the rate of reaction by 6000-fold implying that the incorrect positioning of K93 and R100 cannot account entirely for the decrease in rate by L97P. The L111P and L130P mutations are distally located from the residues proposed to be important for catalysis. However, they still significantly decrease the rate of reaction by ~15,000 and ~3000-fold, respectively. This implies that the structural integrity of the helical cap is some-how important for catalysis, a piece of information not fully appreciated until this study was performed.

The dissociation constants of the proline variants in the presence and absence of divalent metal ions for a minimal, static DF substrate (DF-5) were comparable to WT (*table 5.2*). After the crystal structures of substrate and product complexes were determined it is widely accepted that the helical arch is not important for initial substrate binding and these results confirm this theory.

Table 5.1: Single-turnover reaction parameters of DF-5 (JMA) hydrolysis. *Reactions were performed at 37 ˚C in 8 mM MgCl$_2$, 110 mM KCl, 55 mM HEPES-NaOH pH =7.5, 1 mM DTT and 0.1 mg/ml BSA. Data were fit to equation 3.5. *reactions performed by Nikesh Patel.*

Enzyme	k_{STmax} (min^{-1})	t$_{1/2}$ (s)	Rate decrease relative to WThFEN1
WThFEN1	740 ± 32	5.6X10^{-2}	1
L97P	*5.6X10^{-3} ± 4X10^{-4}	7426	132,000
L111P	*5.0X10^{-2} ± 3X10^{-3}	831	15,000
L130P	*2.6X10^{-1} ± 1.3X10^{-2}	160	2850

Table 5.2: K_D parameters and standard errors for 10 nM DF-5 bound to WThFEN1, L97P, L111P and L130P determined by fluorescence anisotropy. *All measurements were made at 37 ˚C in 110 mM KCl, 55 mM HEPES, 1 mM DTT, 0.1 mg/ml BSA and 10 mM Ca^{2+} or 2 mM EDTA as indicated. Data were fit to a quadratic equation shown in section 3.3. Accompanying graphs are shown in figure 5.6.*

Substrate	DF-5 (JMA)							
Enzyme	WT		L97P		L111P		L130P	
Buffer	Ca^{2+}	EDTA	Ca^{2+}	EDTA	Ca^{2+}	EDTA	Ca^{2+}	EDTA
K_D (nM)	46 (± 6)	50 (± 4)	47 (± 3)	64 (±7)	41 (± 4)	61 (± 5)	63 (± 5)	46 (± 4)
r$_{min}$	0.05 (± 0.001)	0.03 (± 0.001)	0.05 (± 0.001)	0.04 (± 0.001)	0.04 (± 0.001)	0.03 (± 0.001)	0.04 (± 0.001)	0.03 (± 0.001)
r$_{max}$	0.10 (± 0.001)	0.11 (± 0.001)	0.13 (± 0.001)	0.14 (± 0.002)	0.15 (± 0.001)	0.14 (± 0.002)	0.13 (± 0.002)	0.13 (± 0.001)
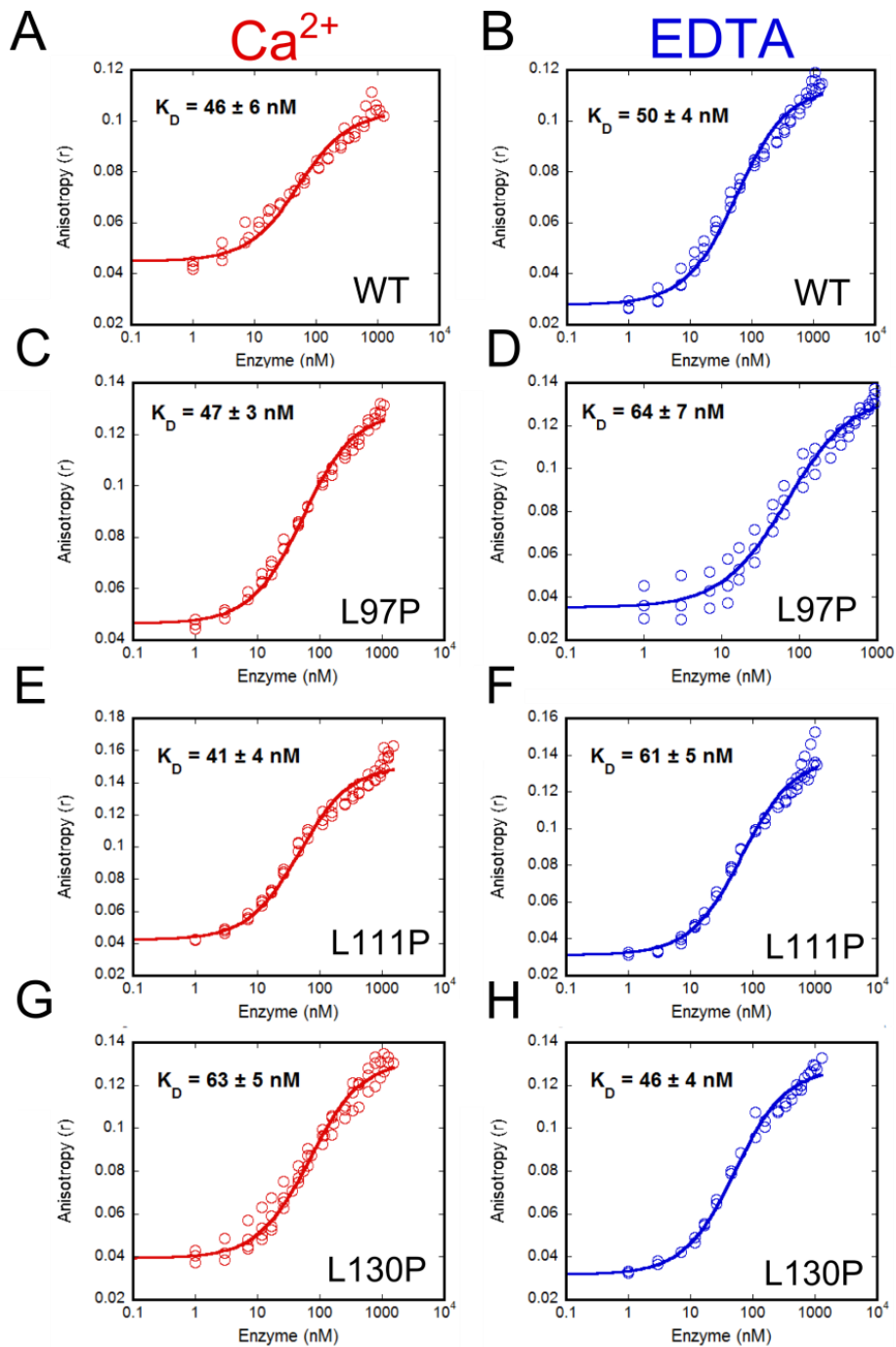
Figure 5.6: DF-5 binding curves for (A) hFEN1 / 10 mM CaCl₂, (B) hFEN1 / 2 mM EDTA, (C) L97P / 10 mM CaCl₂, (D) L97P / 2 mM EDTA, (E) L111P / 10 mM CaCl₂, (F) L111P / 2 mM EDTA, (G) L130P / 10 mM CaCl₂, (H) L130P / 2 mM EDTA. *Apart from the specified conditions, assays were performed in 110 mM KCl, 55 mM HEPESpH=7.5, 0.1 µg/µl BSA, 1 mM DTT and at 37 ˚C. The data points were fit to equation 3.3. The K_D value for each condition is shown in the insert of the particular graph, each K_D is representative of at least three repeats, and errors shown are standard error.*

5.4 Helical Arch variants can still accommodate 5'-flaps

The data in *section 5.3* implies that correct accommodation of the 5'-flap is an absolute requirement for the unpairing mechanism. Therefore, prior to exciton-coupling 2-AP experiments with the altered proteins, the impact of proline mutations on 5'-flap accommodation was investigated.

To distinguish between trapped and blocked FEN complexes the EMSA assay previously used to visualise hFEN1:DNA complexes (*section 3.5.1*) was adapted, described below. This method was chosen, as opposed to a kinetic approach that examines reaction (Patel et al., 2012), because the rates of reactions of the proline variants were drastically slowed. A 10 nt 5'-biotinylated flap substrate (DF-10B) was created for this assay. This construct had slightly longer downstream and upstream regions as the substrate used in FA and ST experiments and the length of the 5'-flap was increased to minimize any potential interactions between SA and the ES complex. In addition, a competitor oligonucleotide (DF-10) was created for these experiments of identical sequence to the bimolecular substrate (DF-B10) but lacking the biotin moiety.

The method to produce trapped and blocked assemblies was analogous to that described for ECCD experiments (*section 5.3*), schemes are repeated below for convenience (*figure 5.7*). As before, two types of FEN1-substrate-SA complexes were formed, either conjugation of SA to DF-10B prior to the addition of enzyme (50 nM) in Ca^{2+} or assembly of the FEN:DF-10B(Ca^{2+}) complex followed by SA conjugation; in both cases a 5-fold molar ratio of SA was used. After 10 mins incubation at room temperature the complexes were challenged with a 5-fold excess of unlabelled competitor (with respect to the final enzyme concentration) and incubated, now at 37 ˚C, for a further 10 mins.

Complexes were detected using an SA-HRP conjugate and enhanced chemiluminescence reagents. However, it was believed that this method would be incapable of detecting biotin labelled species already conjugated with SA used to form the trapped and blocked complexes. Therefore, initially an SA-HRP antibody was used to visualise SA-bound species. However this method gave a high background signal and only excess SA in the samples could be detected (*figure 5.8*). In addition, only pre-SA-conjugated species (DF-10B:SA, blocked and trapped complexes) can be detected using this methodology therefore the location of unbound DF-10B and the FEN:DF-10B species would not be visible. Surprisingly, when additional blots were probed with SA-HRP all the bands could be assigned to the relevant species. Despite the saturation of the biotin by SA, the SA-HRP could still bind. There are two possible explanations for this. Firstly, it seems possible that the cross-linking procedure results in denaturation of the

SA bound to the biotinylated oligonucleotides in our electrophoresed samples. The cross-linking procedure involved exposure of the membrane to UV light (312 nm) immediately after the transfer from the gel, which induces covalent bonds between the DNA and free amine of the nylon membrane. Alternatively monomer units of the SA used in the initial formation of the complexes may be able to exchange with monomer units of the SA-HRP conjugate.

It can been seen from the EMSA blots in *figure 5.7* that the unbound substrate migrated fastest followed by the enzyme-substrate complex and then the substrate-SA complex. The slowest migrating species were the blocked and trapped complexes (DF-10B-SA-E) which had similar electrophoretic mobility. For WThFEN1 and all the mutated FENs tested the trapped complex was more stable than the equivalent blocked and unmodified enzyme-substrate complexes. Thus, in each case the amount of blocked complexes substantially decreased, by incubation with the competitor DNA indicated by an increase in the amount of unbound DF-10B-SA (compare lanes 5 and 7). However, in all cases trapped complex persisted after competitor incubation (compare lanes 6 and 8).

It is interesting to note that the stability of the trapped complex varied across the hFEN1 proteins tested. WT, K93A, R100A, and Y40A trapped complexes appeared to be more stable than L97P, L111P, L130P and K93AR100A trapped complexes. One possibility is that the apparent stability of the trapped complex reflects the portion of the substrate that is threaded through the helical arch when SA is added. Any substrate that is not threaded through the helical arch effectively becomes blocked when SA is added. This portion of blocked substrate may be readily displaced leaving the true trapped complex. Nevertheless, all the mutated proteins tested could produce a trapped complex, suggesting that the 5'-flap portion of the substrate could be accommodated in the same position as WT protein. Furthermore, the experiments with helix breaking Pro residues show that a unique helical arch conformation is not required for 5'-flap accommodation.
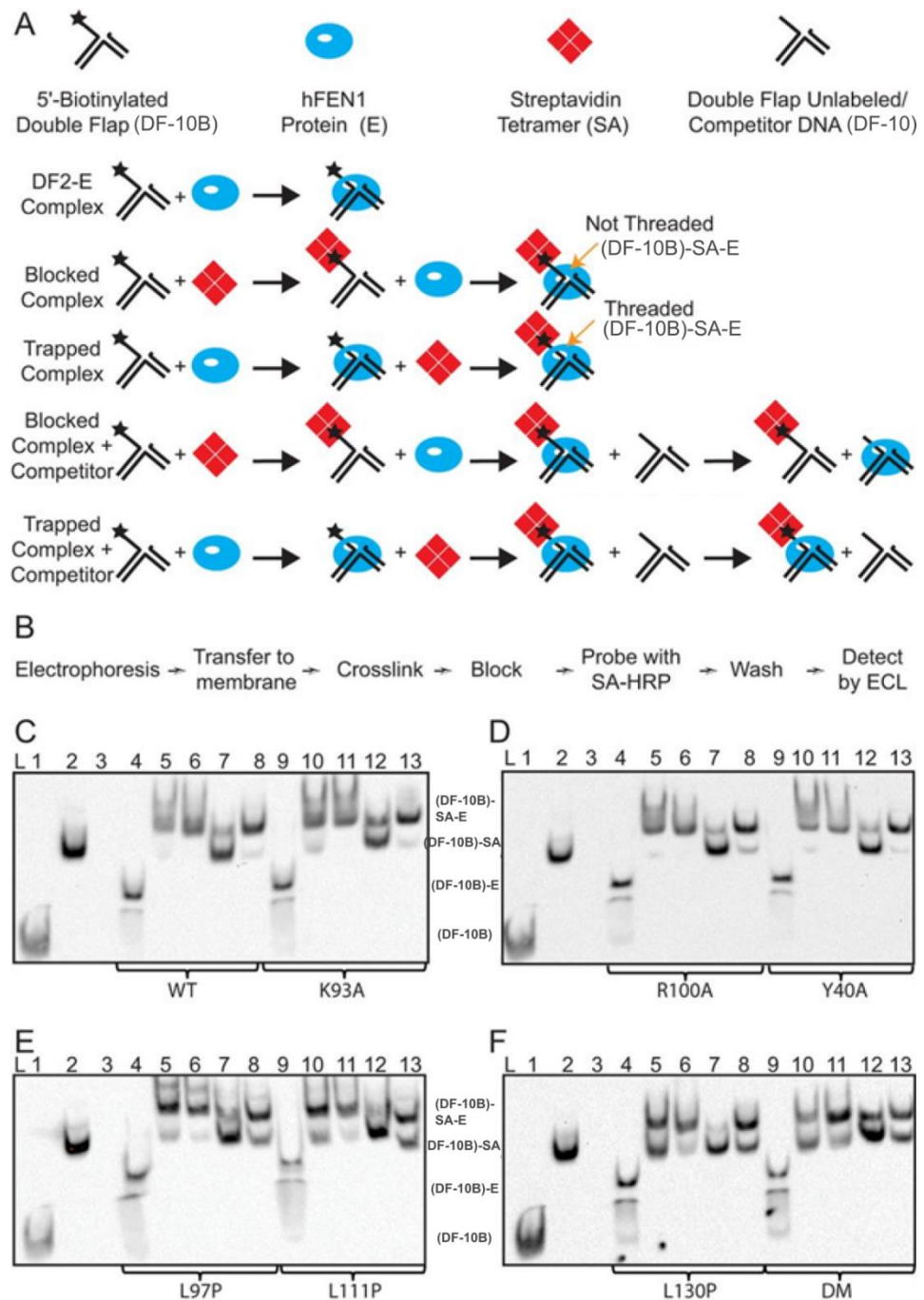
Figure 5.7: 5'-Flap accommodation assessment by EMSA. (A) schematic of procedure for the formation of complexes. Trapped complexes were assembled at 20 °C by the addition of SA to an E-DF-10B complex, whereas blocked complexes involved the addition of E to SA-DF-10B. Competition experiments were carried out with 5-fold excess nonbiotinylated DF-10 at 37 °C for 10 min. (**B**) workflow diagram for detection of biotinylated oligonucleotides using SA-HRP conjugate. (**C–F**) flap accommodation assessed by EMSA that was detected using the SA-HRP conjugate. Lane 1, DF-10B; lane 2, DF-10B-SA; lane 3, DF-10 with SA (control, no signal observed without biotinylated DNA); lanes 4 and 9, DF-10B-E; lanes 5 and 10, blocked DF-10B-SA-E; lanes 6 and 11, trapped DF-10B-E-SA; lanes 7 and 12, blocked DF-10B-SA-E competed with DF-10; lanes 8 and 13, trapped DF-10B-E-SA competed with DF-10. (**C**) WT hFEN1 (lanes 4–8) and K93A (lanes 9–13). (**D**) R100A (lanes 4–8) and Y40A (lanes 9–13). (**E**) L97P (lanes 4–8) and L111P (lanes 9–13). (**F**) L130P (lanes 4–8) and K93AR100A (lanes 9–13). Taken form Patel, N. et al. JBC (2013) 288, 34239-34248.

167
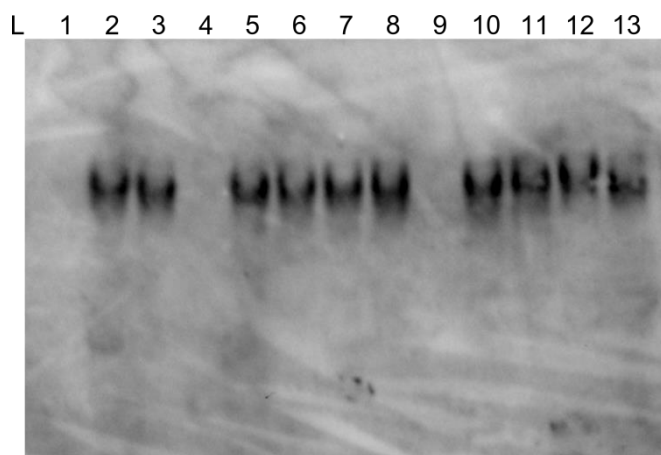
Figure 5.8: Detection of SA-conjugated species using SA-HRP antibody, only excess SA can be observed. *Lane 1, DF-10B; lane 2, DF-10B-SA; lane 3, DF-10 with SA (control, SA was added to this lane); lanes 4 and 9, DF-10B-E; lanes 5 and 10, blocked DF-10B-SA-E; lanes 6 and 11, trapped DF-10B-E-SA; lanes 7 and 12, blocked DF-10B-SA-E competed with DF-10; lanes 8 and 13, trapped DF-10B-E-SA competed with DF-10. (**C**) WT hFEN1 (lanes 4–8) and K93A (lanes 9–13).*

5.5 Integrity of helical cap is crucial for substrate unpairing

The characterisation of L97P, L111P and L130P, so far suggests that the inability to perform hydrolysis efficiently is not a consequence of decreased substrate affinity or the incorrect accommodation of the 5ʹ-flap. Therefore, the low-energy CD spectra of the S$_{-1-2}$ assemblies were recorded to determine whether these mutations affected the substrate unpairing dynamics. As observed previously with the active site variants (K93A, R100A and Y40A) only WThFEN1 was capable of altering the conformation of the 2-APs in the absence of Ca^{2+} ions (presence of EDTA); L97P, L111P and L130P do not show any change in the equilibrium distribution of the 2-AP dimer compared to unbound substrate in buffer containing EDTA (*figure 5.9b*). The addition of L97P to S$_{-1-2}$ in the presence of calcium ions produced a spectrum that was essentially identical to that of WThFEN1 (*figure 5.9a*). This implies that the adjacent -1 and -2 2-APs are in the same position when in a complex with L97P as WT protein. However, surprisingly the CD intensity at 326 nm for L111P and L130P substrate complexes remained similar to that of the unbound substrate despite the presence of divalent metal ions (*figure 5.9a*). This indicated that the helical cap mutations (L111P and L130P), opposed to the helical gateway mutation (L97P), alter the ability of hFEN1 to unpair at the ds-ss junction.
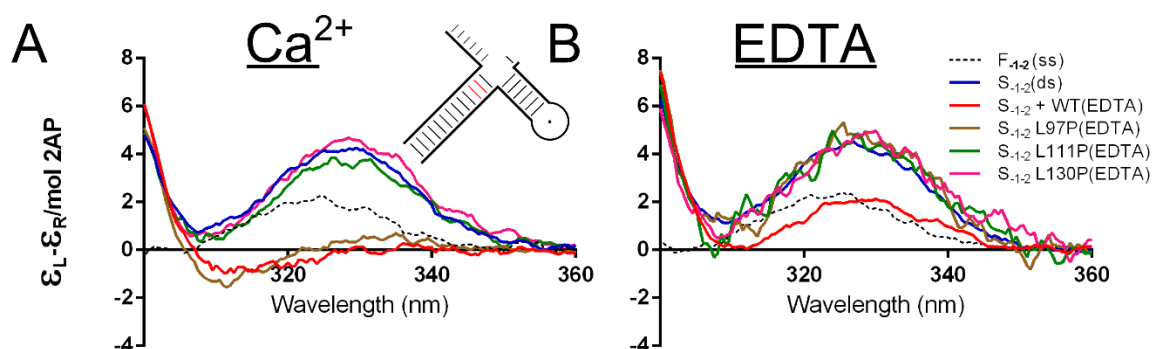
Figure 5.9 Low-energy CD spectra of 2AP-containing DNA substrate construct and hFEN1 WT- and variant-DNA substrate complexes. (A) CD spectra recorded in buffer containing 10 mM Ca^{2+} ions, *unbound S$_{-1-2}$ (blue), the corresponding single strand (F$_{-1-2}$, gray), S$_{-1-2}$ bound to WThFEN1 (red), L97P (gold), L111P (green), and L130P (magenta).***(D)** CD spectra in buffer containing no available divalent metal ions (10 mM Ca^{2+} + 25 mM EDTA) same colouring as in C. Samples were buffered at pH =7.5, and measurements performed at 20 ˚C; specific conditions given in section 2.7.1. Each measurement was independently repeated and gave equivalent results, spectra were normalized as described in experimental section 2.7.1. Substrate construct is illustrated schematically (red 2AP) with sequence shown in figure 3.5.*

The effect of the gateway L97P mutation upon unpairing is essentially null. Conceptually, this would be expected for a mutation of this type at this position. The insertion of proline at the base of α4 is expected to alter the distribution of conformations the helix is able to adopt compared to those adopted by WThFEN1. As a result, it is assumed that this mutation will considerably change the likelihood that K93 and R100 will be able to access the same positions (or at least for the same amount of time) compared with the WT protein. As discussed above, K93 and R100 are not required for conformational rearrangement at the -1 and -2 positions. Therefore, the result obtained for the L97P enzyme is in agreement with previous observations. In addition, it reveals that the structural integrity of the base of α4 is not required for unpairing.

The finding that Leu to Pro mutations within the helical cap (L111P and L130P) can abolish unpairing was surprising and unexpected. Firstly, a number of active site mutations including K93A, R100A, D181A, R100AK93A had already been tested and each of these was capable of bringing about local substrate conformational change in the presence of Ca^{2+} ions. Secondly, this result was particularly unanticipated given the distance of these altered residues from the active site and from the DNA itself. Moreover, the fact that alteration of a remote part of the protein can prevent unpairing suggests the protein participates in this DNA conformational change and it is not the result of a passive fraying of the bound duplex DNA.

5.6 Discussion and conclusions

In FEN1-DNA substrate and product structures α5 of the helical cap packs against the α2-α3 loop. The upper part of α2 along with α3, and the loop connecting the two, forms the upstream dsDNA and 3'-flap pocket binding region. However, in DNA-free structures of FEN1 both the helical cap and the α2 – α3 loop are disordered. The dynamics of the α2 – α3 loop appear to be sensitive to the presence of the substrate 3'-flap and this interaction is thought to have a domino effect on the structural rigidity of other regions of the protein. It is thought that binding of upstream DNA promotes the formation of a well-ordered 3'-flap binding pocket which is accompanied by increased ordering of the α4 - α5 helical arch (Ivanov et al., 2006). Thus the binding of the 3'-flap is communicated to the helical arch by packing interactions between the α2 – α3 loop and α5.

These structural observations led *Tustakawa et al* to suggest a model for FEN1 action. Upon binding the downstream dsDNA of a double flap substrate the helical gateway, which is ordered, directs the 5'-flap under the disordered helical cap. This step is argued to select for ss or so-called gapped 5'-flaps only as a structured helical gateway is unable to accommodate DNA containing secondary structure. Once the 3'-flap has been selected conformational changes in the α2 – α3 loop are transmitted to the helical arch triggering ordering of the helical cap which is suggested to push down on the 5'-flap. At this point the residues involved in unpairing are in a position to capture the unpaired state.

The major provisos of this model stipulate that:

- The 5'-flap has already threaded through a disordered helical arch before unpairing
- The structured helical gateway enforces the need for unpairing
- The role of the helical cap is to select for 5'-termini and is not involved in unpairing
- Unpairing only occurs after helical arch structuring placing Lys-93 and Arg-100 in a position to facilitate unpairing

Some of the results presented in this chapter are in accord with this model. For example, the fact that the 5'-portion of substrates needs to be properly accommodated before the DNA can unpair is demonstrated by lack of change in $(2\text{-AP})_2$ ECCD signal on addition of FEN1(Ca^{2+}). This suggests that the flap must be correctly positioned before local DNA conformational change can occur. However, earlier results from *chapters 3* and *4* obtained with K93A, R100A and Y40A suggest that it is not necessary for these gateway residues to be correctly positioned for DNA unpairing, but that their presence is required for efficient catalysis. In particular it was found that the variant L97P(Ca^{2+}), containing a helix breaking Pro residues, could still unpair DNA. Thus

the α2 – α4 gateway, where the FEN1 conserved amino acid residues reside, need not be structured to produce unpairing as suggested in the previous model.

Most importantly the data presented here suggests unpairing requires proper ordering of the helical cap region of the helical arch but not the base of α4. Thus, the helical cap has a more important role in the hFEN1 mechanism than previously thought and its role does not appear confined to selecting for 5'-termini. One possibility is that helical cap could promote unpairing by restraining/positioning one or both of the unpaired nucleobases, although there is no obvious structural rationale for this scenario. The spatial relationship and packing of the helical cap with the α2 – α3 loop that fashions the 3'-flap and duplex binding pocket is intriguing. Both of these regions of the protein pack and interact with α2 on which base paired DNA resides. It is possible that the helical cap, together with the 3'-flap sensing α2 – α3 loop, participate in protein conformational changes that subtly alter the positioning of α2 to afford the unpaired state. In the future, it would be intriguing to look at the importance of the helical cap in the mechanism of other FEN family members; for those which have very similar caps like EXO1 and particularly for those which do not (GEN1 and XPG).

Chapter 6: Investigating the inhibition of hFEN1 by cyclic *N*-hydroxy urea compounds

6.1 Identification of hFEN1 inhibitors

As discussed in the introduction *section 1.7.1*, human flap endonuclease-1 is an attractive oncology target due to its key roles in DNA replication and repair. The most potent hFEN1 inhibitor published to date was discovered by a fragment-based approach starting with a series of *N*-hydroxy urea compounds (Tumey et al., 2005). In addition, as part of the same study the authors reported selectivity for hFEN1 over XPG; however, the origins of this bias was not investigated. Furthermore, a separate more recent study also identified inhibitors of FEN1 and also included inhibitors containing *N*-hydroxy urea core (van Pel et al., 2013). Therefore, in collaboration with Astra Zeneca, the mechanism of inhibition by *N*-hydroxy urea compounds was investigated with a view to aid the design of next-generation inhibitors.

The inhibitor selected as the subject for investigation was reported by Tumey, N. L., *et al* as one of the most promising lead compounds, with a measured IC_{50} of 11 nM from the initial screen (*figure 6.1*). This compound is subsequently referred to as inhibitor **1**. In the earlier study no direct evidence for binding to the protein was reported for any compound in the series. However, based on inhibitors with a related pharmacophore they were suspected to coordinate to divalent metal ions located in the active site via their *N*-hydroxy urea 'war-head' (*figure 6.1*).
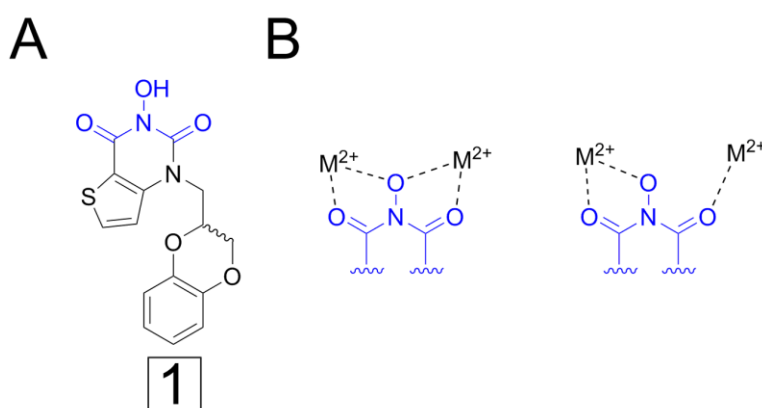
Figure 6.1: N-hydroxy urea hFEN1 inhibitor (*A*) Inhibitor 1 / 1-((2,3-dihydrobenzo[b][1,4]dioxin-2-yl)methyl)-3-hydroxythieno[3,2-d]pyrimidine-2,4(1H,3H)-dione (compound 17 form Tumey et al (2005)), divalent metal-binding 'war-head' shown in blue. (*B*) Two proposed binding interactions of nuclease inhibitors through their 'war-head' moiety, which is likely to be ionised at physiological pH. Adapted from Tumey, N. L., et al. (2005) Bioorganic & Medicinal Chemistry Letters 15, 277-281.

As part of the investigation crystallisation screens were performed by Judit Debreczeni at Astra Zeneca in an attempt to image the inhibitor **1** bound in the active site of hFEN1Δ336 in the presence of Mg^{2+} ions. Human FEN1 (hFEN1) is a 380 amino acid protein. The first 336 amino acid residues comprise the nuclease core domain, whereas amino acid residues 337 to 380 (called the extended C-terminus) are known to be important for hFEN1 regulation. The extended C-terminus (C-terminal tail) contains the PCNA interaction motif as well as interaction motifs for other proteins and is the site of post-translation modifications (Guo et al., 2012; Sakurai et al., 2005). Furthermore, the C-terminal tail has been reported to facilitate substrate binding (*section 6.3.3*) (Stucki et al., 2001). However, the C-terminal tail is not essential for nuclease activity or specificity (*section 6.3.3*) (Tsutakawa et al., 2011).

The ternary complex, hFEN1Δ336-$(Mg^{2+})_2$-inhibitor **1**, was successfully crystallised with a 3.4 Å resolution structure. As predicted, inhibitor **1** binds to hFEN1Δ336 through the war-head moiety which directly coordinates the two active site Mg^{2+} ions (*figure 6.2*). hFEN1 and inhibitor **1** bound with a 1:1 stoichiometry. However, and somewhat disappointingly, with the exception of a methionine to the sulphur of the thiophene group interaction, there are no other specific interactions between inhibitor **1** and residues within the active site. There is free rotation about the nitrogen and carbon bond of the pyrimidine and the sidechain aromatic group. It was thought that the aromatic group might stack against aromatic residues within the active site, driven by favourable π-π stacking interactions. However, in the structure this portion of the inhibitor sits within an open cleft, and raising the possibility that this compound may be capable of inhibiting related, and more importantly, unrelated two metal-ion dependent nucleases. On the other hand, this result was encouraging because it was foundation for rational based design, and an opportunity to form favourable, and more specific, interactions with residues within the cleft.
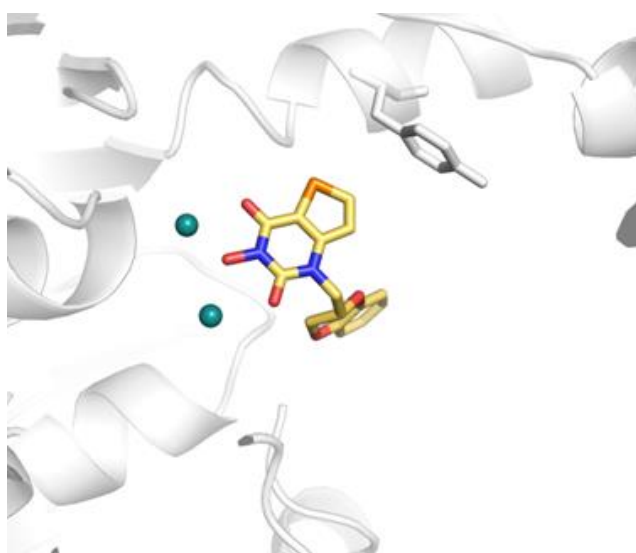
Figure 6.2: Crystal structure of inhibitor 1 bound in the active site of hFEN1Δ336 in the presence of magnesium ions. *Protein residues (grey) are shown as cartoon, with the two magnesium ions (aqua) and tyrosine-40, a potential stacking partner, shown as sticks. Inhibitor **1** colouring: carbon-yellow, oxygen-red, nitrogen-blue and sulphur-orange.*

6.2 Measurement of binding constants for *N*-hydroxy urea inhibitors using Isothermal Titration Calorimetry

Ligand binding to macromolecules can be determined calorimetrically (Ababou and Ladbury, 2006; Perozzo et al., 2004). The independent variables of thermodynamic interest used to characterise interactions between a biological macromolecule M and a small ligand X are the dissociation constant, K_D, heat of binding, $\Delta H°$, and the number of binding sites, n. Isothermal titration calorimetry (ITC) is a direct method used to determine these thermodynamic parameters. A heat signal is a nearly universal property of binding interactions, so it is applicable to most ligand macromolecule-interactions (Wiseman et al., 1989). Isothermic calorimeters have two compartments, sample and reference, which are treated equally, but ligand is only introduced to the sample compartment (*figure 6.3*). For the duration of the titration both the reference and sample cell are kept at the same temperature. Therefore, the heat required to compensate the endo- and exothermic binding events is registered.
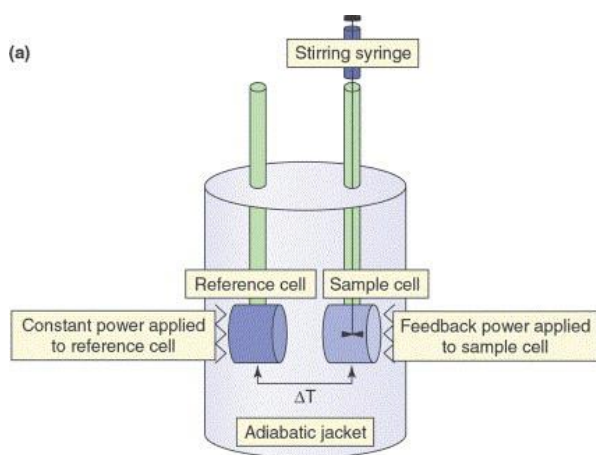
Figure 6.3: Schematic representation of a power compensation isothermal titration calorimeter. *A constant power is applied to the reference cell, activating a feedback circuit, which applies a variable power to the sample cell. This maintains a very small and monitored temperature difference between the cells. The feedback power is the baseline in the ITC experiment. Exothermic reactions generate heat, thus triggering a decrease in the feedback power, whereas endothermic reactions increase the feedback power. Taken from Holdgate, G. A., and Ward, W. H. J., (2005). Drug Discovery Today 10, 1545-1550.*

Binding of a ligand to a macromolecule with one binding site is described by,

$$M + X \rightleftharpoons MX \quad \text{equation 6.1}$$

The heat change of the system, Q, after the ith injection is dependent on the amount of ligand added $(dQ/d[X]_t)$ given by,

$$Q_i = V_0 \Delta H° \Theta_i \qquad \text{equation 6.2}$$

, where V_0 is the cell volume and $\Delta H°$ is the molar heat of ligand binding. The quantity Θ_i is fractional saturation of the macromolecule after the ith injection. The individual heat change per injection (q_i) or the cumulative heat change (Q) must be expressed in terms of total ligand and macromolecular concentrations as these are the known experimental quantities. For a simple 1:1 binding interaction, and more complicated cases, *equation 6.2* has been solved for Θ_i in terms of total ligand $([L]_t)$ and total macromolecular $([M]_t)$ concentration (Freire et al., 1990). A nonlinear fit base on *equation 6.2* of Q vs $([L]_t)$ to a hyperbolic saturation curve yields the parameters K_D, $\Delta H°$ and *n* from a single experiment.

6.2.1 *N*-hydroxy urea inhibitors bind to hFEN1 with nanomolar affinity

ITC was employed to quantify the strength of the association between inhibitor **1** and hFEN1(Mg^{2+}). In addition, the binding affinity of two previously uncharacterised compounds was investigated. The *N*-hydroxy urea warhead moiety is common to all three compounds but inhibitor **2** and inhibitor **3** differ in ring system A and B as shown in *figure 6.4*. Ring system A for inhibitor **1** includes a fused thiophene, this is replaced by a fused unfunctionalised fused phenyl group in inhibitor **2** and **3**. The major difference in the structure of all three inhibitors was in ring system B. In place of the large aromatic side chain, inhibitor **2** had a cyclopropyl, commonly used in structure-activity-relationships (SARs) as it is one of the smallest metabolically stable ring systems (Taylor et al., 2014). Inhibitor **3** lacked any substitution at the pyrimidine nitrogen (*figure 6.4*), therefore this compound is essentially just the active site metal binding "war-head". Here an assessment between ring system B size and binding affinity was made using ITC. The measured K_D, $\Delta H°$ and *n* values and the derived entropy ($\Delta S°$) of binding to hFEN1 are compiled for the three inhibitors in *table 6.1*.
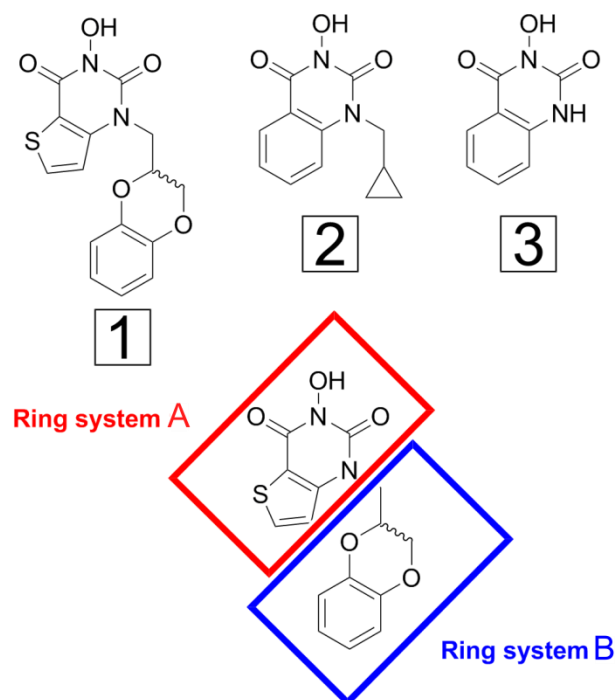
Figure 6.4: Structure of inhibitors 1, 2, and 3. For clarity the compounds can be divided into two ring systems A and B.

Table 6.1: Binding parameters of inhibitors 1,2, and 3 bound to hFEN1 in the presence of 8 mM MgCl₂ from global analysis of triplicate ITC experiments: *n-stoichiometry, K_D-dissociation constant (95% confidence interval), $\Delta H°$-enthalpic change, $\Delta S°$-entropic change and χ^2 critical chi-square value for 95% limits on global analysis. Accompanying graphs are shown in figure 6.5.*

Inhibitor	Protein	K_D, nM (95% C.I.)	n (S.E.)	$H°$, kcal mol⁻¹ (S.E.)	$S°$, cal mol⁻¹ deg⁻¹	Reduced critical χ^2
1	hFEN1Δ336	182 (72 – 395)	1.10 (± 0.02)	1.60 (± 0.02)	36	31
2	hFEN1Δ336	194 (62 – 492)	1.05 (± 0.03)	1.51 (± 0.03)	36	27
3	hFEN1Δ336	1577 (725 – 3333)	0.75 (± 0.03)	1.90 (± 0.03)	33	35
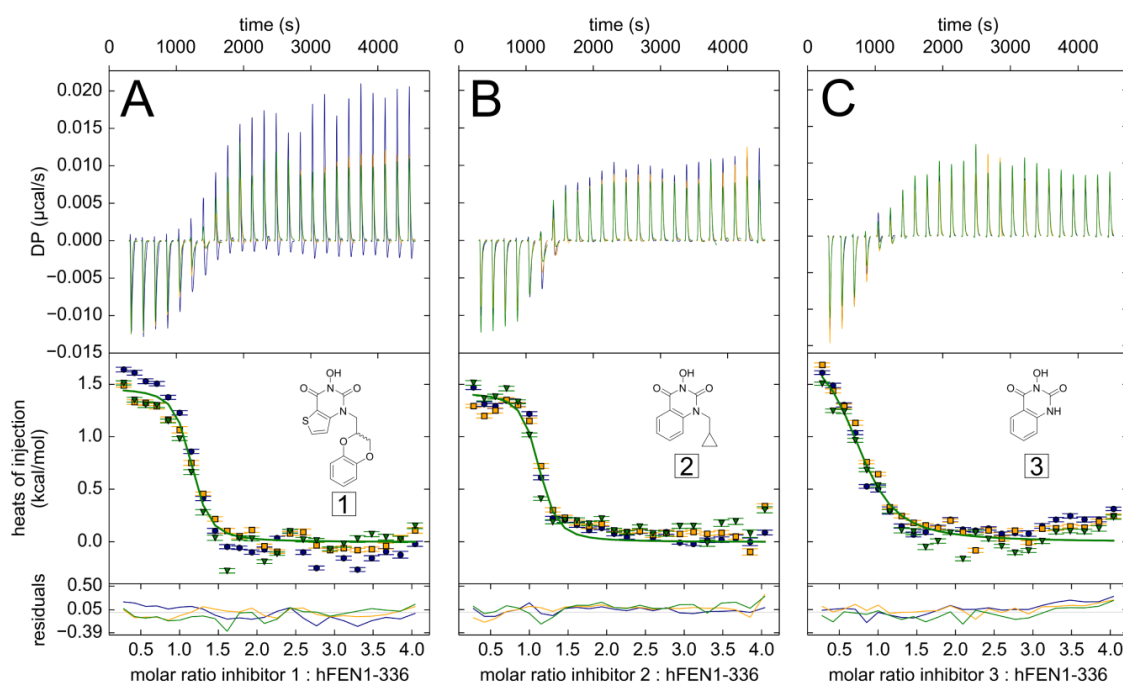
Figure 6.5: Inhibitor binding to hFEN1. *Affinity of (**A**) inhibitor **1**, (**B**) inhibitor **2**, and (**C**) inhibitor **3**, for hFEN1Δ336 assessed by ITC. The top, middle and bottom panels show the raw data, processed data and residuals, respectively. The orange, blue, and green circles represent data from the three respective replicates used in this analysis. The residual lines in the lower graph are coloured according to their respective replicate. Each measurement was conducted at pH 7.5 and 25 ˚C in the presence of 8 mM Mg²⁺. Parameters from global analysis of at least triplicate experiments are shown in table 6.1.*

ITC experiments were performed with the crystallographic construct hFEN1Δ336. The C-terminal tail does not form part of the active site therefore it was not expected to interfere with the inhibitors interaction with the Mg²⁺ bound in the active site. However, the C-terminal tail does contain a large proportion of positively charged residues. Therefore, the reasoning at the time was that interactions of the war-head and this region would possibly result in a false dissociation

constant and binding stoichiometry. In addition, as discussed above *in vivo* the C-terminal tail is proposed to be involved in protein-protein interactions, thus, it is likely that this region of the protein will not be accessible for possible inhibitor interactions *in vivo*. Titrations were performed in buffer containing Mg^{2+} ions. This is the catalytically relevant cofactor. The crystallographic data shows that the war-head moiety is able to chelate Mg^{2+} bound in the active site and in the original study inhibitor **1** was confirmed to bind to hFEN1 in Mg^{2+} containing buffer.

To determine the true heat of binding separate experiments measure the heat of dilution of ligand and also of the macromolecule which can be subtracted from the measured heat of the binding titration. A rigorous correction, as shown in *equation 6.3*, would also include the heat change associated with injection of buffer from the syringe into the same buffer in the sample cell, which is theoretically zero and any heat signal is instrument specific.

$$Q_{corrected} = Q_{measured} - Q_{dil.ligand} - Q_{dil.macromolecule} - Q_{blank} \quad equation\ 6.3$$

Representative calorimetric titrations of the heat changes associated with injection of inhibitor **1** into buffer, buffer into hFEN1, and buffer into buffer are shown in the *figure 6.6*. The data shows that the heat of injection associated with all three control experiments is essentially identical, 1.2 µcal per injection. Theoretically the instrument blank, buffer into buffer titration, should have much smaller heats of injection than either dilution of the ligand or macromolecule, which can have identical heat changes but normally do not. This result can be rationalised as the final buffer composition contained 1 % dimethyl sulfoxide (DMSO) (v/v) to aid inhibitor solubility. All the inhibitors tested were stored in pure DMSO and dilutions to the final appropriate concentration and 1 % DMSO (v/v) was made using a pipette. Because 1 % DMSO is a large difference in buffer composition, hFEN1 and buffer solutions were also made 1 % DMSO (v/v) using a pipette. Therefore, it was believed that a mis-match in DMSO concentration, linked to pipetting error, was the major contribution to the heat signal. Therefore, only the ligand dilution experiment was used to correct the measured heat of binding; a common protocol implemented in ITC studies (Holdgate and Ward, 2005). In addition, the buffer into buffer experiment described here is not a true instrument blank experiment and this experiment was not performed. Finally, the heat change associated with the first injection was ignored for all the titrations performed in this thesis. This is a common practice when performing ITC experiments as there is an endemic error associated with currently available isothermal titration calorimeters and deliverance of the first injection (Mizoue and Tellinghuisen, 2004).
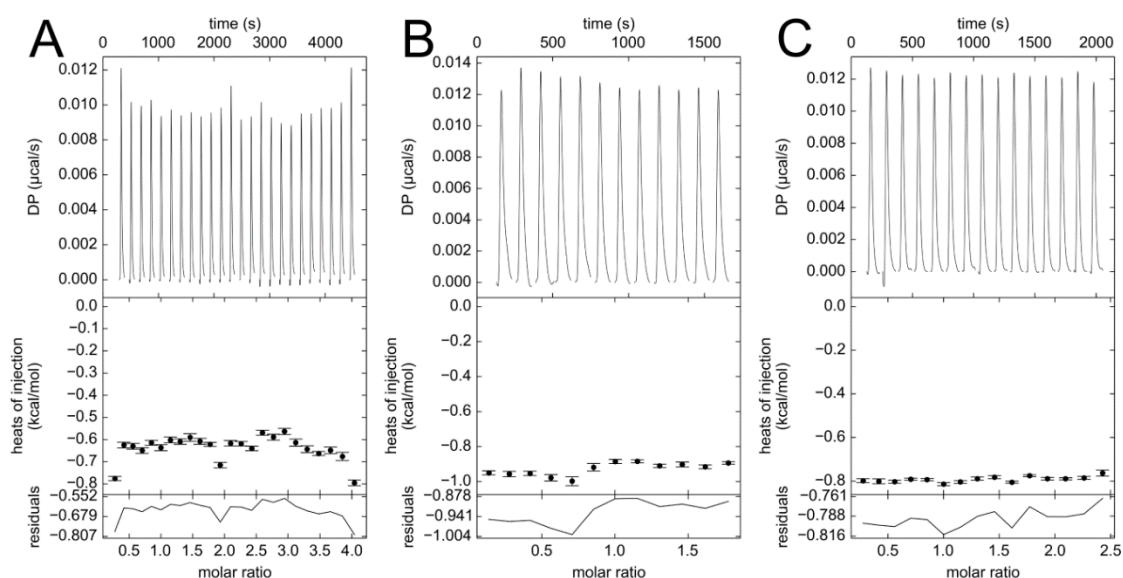
Figure 6.6: Isothermal titration calorimetry control experiments. (A) *inhibitor **1** into buffer. **(B)** buffer into hFEN1. **(C)** buffer into buffer. Titrations were performed at 25 °C in buffer containing 100 mM KCl, 50 mM HEPES-NaOH pH = 7.5, 8 mM MgCl₂ and 1 % DMSO (v/v). To allow for subtraction from the binding experiment injection volumes and injections intervals were matched to the titration of inhibitor **1** into hFEN1. The top, middle and bottom panels show the raw, processed data and reciprocals, respectively.*

Figure 6.5 shows the results of titrating inhibitors **1**,**2** and **3** into hFEN1Δ336 in the presence of

8 mM MgCl₂. In agreement with crystallographic data, all three inhibitors bound to hFEN1Δ336

with a stoichiometry close to one (*table 6.1*), presumably with inhibitors **2** and **3** bound in the

active site as was previously observed for inhibitor **1** (*figure 6.2*). Analysis of the results shows

that inhibitor **1** binds hFEN1Δ336(Mg^{2+}) with a dissociation constant of 184 nM (K_D). Inhibitor

2 had a K_D of 192 nM similar to that of inhibitor **1** (*table 6.1*). In contrast, the binding affinity of

inhibitor **3** to hFEN1Δ336(Mg^{2+}) was weaker than either **1** or **2**. This suggests that despite the

lack of interactions in the crystal structure of hFEN1(Mg^{2+}):inhibitor**1** the presence of a ring

system B is contributing to the binding affinity. However, if the confidence intervals on the K_D

values are taken into consideration the difference between the affinities is little as 2-fold (*table

6.1*), implying that the major contribution to binding affinity is provided by the interaction

between the war-head and Mg^{2+} ions. The error values on the K_D values may seem large. One

possible explanation for this is that the heat of binding was measured as 1.5 μcal per injection

ideally this would be closer to 10 μcal per injection. Therefore, errors associated with DDT

oxidation and baseline drift become more apparent. In addition, because the heat of dilution

becomes comparable to the heat of interaction (both ~ 1.5 ucal/inj although opposite signs),

then further errors may be introduced on subtracting background processes from the data. The

heat signal of binding could not be increased for reasons outlined below.

The power-versus time plots (top panel *figure 6.5*) show that binding off all three inhibitors is endothermic; less power is supplied to the reference cell on injection of ligand. As the concentration of unoccupied binding sites begins to decrease, the endothermic heat changes decreases correspondingly as ligand is added, until at saturating concentrations of ligands the heat changes are equal to the exothermic heat of dilution of the ligand (*figure 6.5*). The endothermic binding enthalpy of inhibitor **1** and **2** is less than inhibitor **3**. This suggests that binding of inhibitor **1** and **2** is more spontaneous as less heat is taken in on binding. These differences may be due to subtle contributions from ring system B. ITC titrations were performed at 25 °C, at 37 °C the binding of all three inhibitors should be more favourable. The endothermic binding suggests that this process in entropically driven. Although the fitting program (SEDPHAT) does not report errors for the binding entropies, inhibitor **1** and **2** displayed greater entropies of binding than inhibitor **3.** Overall inhibitors **1** and **2** possibly displaced more water molecules or ions upon binding. In summary, the binding enthalpies and entropies for all three inhibitors were similar suggesting that there was no drastic difference in the way the inhibitors interact with the hFEN1(Mg^{2+}) *complex*.

Wiseman *et al*. noted that for a simple non-interacting one site model (Wiseman et al., 1989), the shape of such a binding isotherm was related to the ratio of the receptor concentration and the dissociation constant. This led to the untiless Wiseman 'c' parameter (*equation 6.4*), which has been noted that curve fitting is optimal when 10 ≤ c ≥ 500. Recently, it has been suggested that ITC experiments are not limited to this so-called experimental window and fitting can be achieved with c ≥ 0.1 (Turnbull and Daranas, 2003). However, this was determined for a theoretical system with very small errors, and working with c values of 10-500 allows for saturation to be achieved by adding as little as 2 equivalents of ligand. For inhibitors **1** and **2** the c value is ~100, but is ~11 for inhibitor **3**. Therefore, one way to improve the experiment would be to increase this value by increasing the protein concentration. However, the [hFEN1] is limited by the solubility of the inhibitor. All three inhibitors were deemed to be soluble up to 200 μM in buffer containing 1 % DMSO (v/v) (Astra Zeneca personal communication). This concentration was discovered to yield detectable heat signals when added at 2 μL increments. Therefore, the concentration of hFEN1 cannot be increased substantially as this would mean that to achieve saturation the inhibitor concentration would also have to increase. For the same reason, a greater heat signal, which would presumably reduce the errors associated with fitting, cannot be simply achieved by increasing both [hFEN1] and [inhibitor].

$$c = K_A[M]_t \qquad equation\ 6.4$$

6.3 Kinetic Studies demonstrate hFEN1 inhibitors bind to DNA-free and DNA-bound hFEN1

6.3.1 Michaelis-Menten kinetics

As an initial simplification enzyme catalysis of a mono-substrate enzyme-catalysed reaction is assumed to consist of two distinct processes depicted by the reaction scheme below (*figure 6.7*),

$$[E] + [S] \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} [ES] \xrightarrow{k_2} [E] + [P]$$
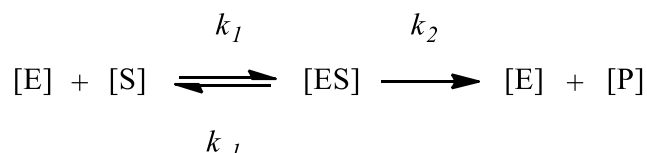
Figure 6.7: Michaelis-Menten mechanism. *[E] – enzyme, [S] – substrate, and [P] – product.*

An equation describing this mechanism, known as the Michaelis-Menten equation allows enzymes to be evaluated and compared using the afforded catalytic parameters, k_{cat} and K_M. In a simple Michaelis-Menten mechanism the formation of the ES complex is in rapid equilibrium, so that k_{cat} is equal to k_2 and K_M is the equilibrium dissociation constant of the ES complex. Although as discussed previously in this thesis the FEN1 reaction mechanism is proposed to be more complex (*section 3.4*) than the original Michaelis-Menten reaction scheme, fitting to *equation 6.5* still yields catalytic parameters, although these are more complex composites of individual rate constants.

Michaelis-Menten equation (*equation 6.5*),

$$\frac{v}{[E]_0} = \frac{k_{cat}[S]}{K_M + [S]} \qquad equation\ 6.5$$

6.3.2 Reversible Inhibition

Unless there is evidence to suggest otherwise, enzyme inhibition is assumed to be reversible and different inhibition models exist depending on the inhibitor interaction with the enzyme. The enzyme:substrate:inhibitor (ESI) complex may be inactive or active and the reversible inhibition can be classified as partial or dead-end, accordingly. Data presented below suggests that, in this instance, ESI is completely inactive therefore can be designated as a dead-end complex. The general mechanism for (dead-end) reversible enzyme inhibition is given by the following reaction scheme (*figure 6.8*),
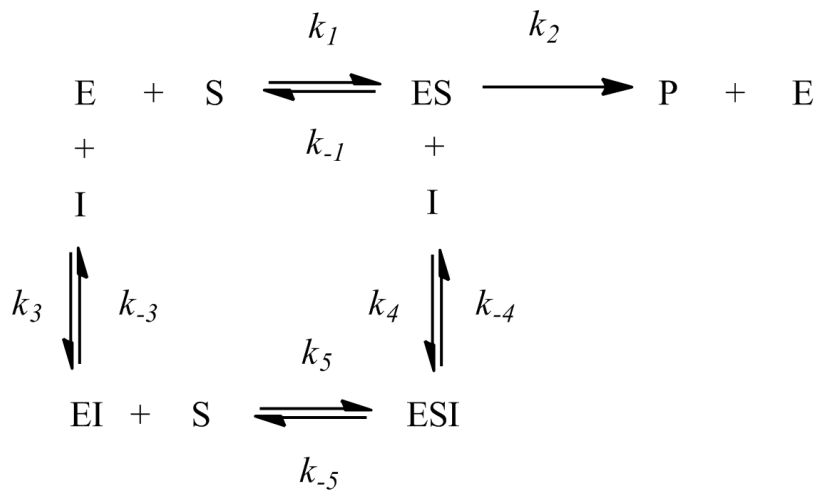
Figure 6.8: General reaction mechanism for reversible enzyme inhibition. *[E] – enzyme, [S] – substrate, [I] – inhibitor, and [P] – product The rate constants of the separate processes are described in the main text.*

The reaction mechanism in *figure 6.8* contains one rate constant, k_2 for the uninhibited reaction, and two dissociation constants each for the substrate S and for the inhibitor I:

$$K_S = \frac{k_{-1}}{k_1} = \frac{[E][S]}{[ES]} \; ; \; K_{Si} = \frac{k_{-5}}{k_5} = \frac{[EI][S]}{[ESI]}$$

$$K_{ic} = \frac{k_{-3}}{k_3} = \frac{[E][I]}{[EI]} \; ; \; K_{iu} = \frac{k_{-4}}{k_4} = \frac{[ES][I]}{[ESI]}$$

The simplest, and most common, types of reversible inhibition and their effects on a monosubstrate enzyme catalysed reaction have been depicted in *figure 6.9*. Studies on the effect of an inhibitor on the steady-state catalytic parameters (K_M and k_{cat}) can be used to quantify the change in equilibrium position of all substrate free and enzyme bound species,

affording an insight into the mode of inhibition (Bisswanger, 1999). Briefly, an inhibitor (I) defined as purely competitive binds to the same site on the enzyme (E) as the substrate (S), resulting in an unproductive enzyme-inhibitor (EI) complex. Because the inhibitor and substrate directly compete for the same site only ES or EI complexes, not ESI complexes, are possible. The binding constant for the competitive inhibitor, the competitive inhibition constant, K_{ic}, defines the affinity of the inhibitor for the free enzyme. As a consequence of competitive inhibition only parameters containing K_M are altered as excess substrate will displace the inhibitor from the binding site leaving k_{cat} equal to the uninhibited turnover rate. In contrast, an uncompetitive inhibitor binds exclusively to the ES complex quantified by the uncompetitive inhibition constant, K_{iu}. Since for uncompetitive inhibition the inhibition does not bind to the free enzyme the K_M is unaltered and only the k_{cat} is affected. This type of inhibition is rare and normally only found for bi-substrate reactions were binding of the first substrate allows binding of the inhibitor, preventing binding of the second substrate.

For a monosubstrate reaction, a non-competitive inhibitor binds to a distinct site separate to that of the substrate; therefore the substrate can still bind to the EI complex but the resulting complex is inactive. For pure non-competitive inhibition K_{ic} and K_{iu} substrate binding does not affect the interaction of the inhibitor with the enzyme and importantly the presence of the inhibitor will not influence the affinity of the substrate so the apparent Michaelis constant is unaffected.

Mixed inhibition contains characteristics of competitive and noncompetitive inhibition, the inhibitor binds at the same site as the substrate, but formation of an EI complex does not abolish substrate binding. The term mixed is considered by some to be insufficient to describe this type of inhibition, as it can either by mixed *with* competitive ($K_{ic} < K_{iu}$) or mixed *with* uncompetitive ($K_{ic} > K_{iu}$). The more common scenario is mixed with competitive as an already bound substrate will impede the binding of the inhibitor and thus the inhibitor will have a higher affinity for the free enzyme than the enzyme substrate complex. Throughout these analyses only mixed with competitive has been considered and termed simply mixed inhibition. The modified Michaelis-Menten equations describing these four types of inhibition are shown below (*equations 6.6 -6.9*).
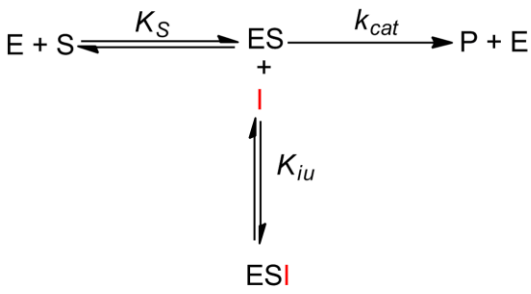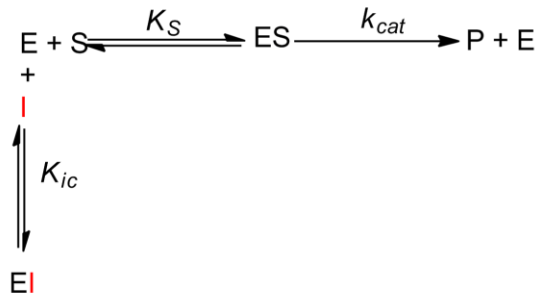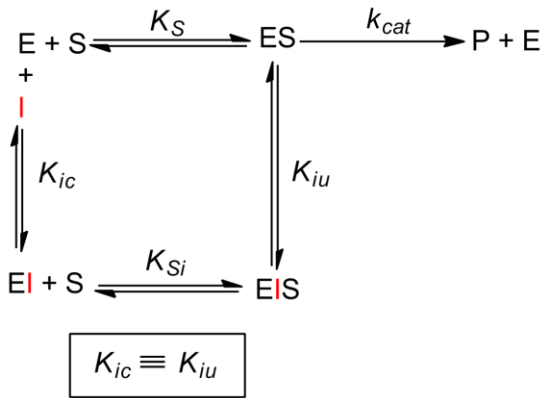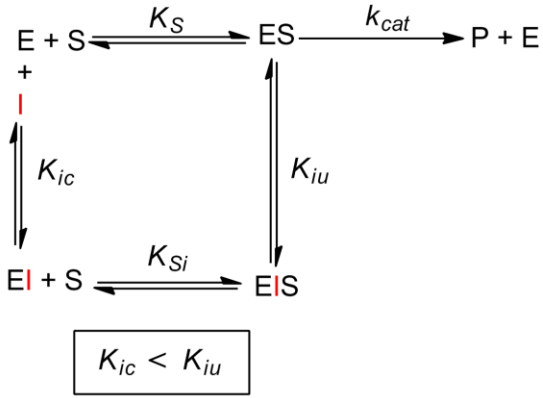
Figure 6.9: The four simplest kinetic models for a reversible inhibitor obeying linear inhibition. E - free enzyme, S - substrate, I - inhibitor (red), P - product, ES - enzyme:substrate complex, EI - enzyme:inhibitor complex, EIS - enzyme:inhibitor:substrate complex. K_S - affinity of S for E, K_{Si} - affinity of S for EI, k_{cat} -maximum multiple turnover rate, K_{ic} -affinity of I for E, and K_{iu} -affinity of I for ES.

Uncompetitive:

$$\frac{v}{[E]_0} = \frac{k_{cat}[S]}{K_M + \left(1 + \frac{[I]}{K_{iu}}\right)[S]}$$

equation 6.6

Competitive:

$$\frac{v}{[E]_0} = \frac{k_{cat}[S]}{K_M\left(1 + \frac{[I]}{K_{ic}}\right) + [S]}$$

equation 6.7

Non − competitive:

$$\frac{v}{[E]_0} = \frac{k_{cat}[S]}{K_M\left(1 + \frac{[I]}{K_{ic}}\right) + \left(1 + \frac{[I]}{K_{iu}}\right)[S]}$$

equation 6.8

Mixed:

$$\frac{v}{[E]_0} = \frac{k_{cat}[S]}{K_M\left(1 + \frac{[I]}{K_{ic}}\right) + \left(1 + \frac{[I]}{K_{iu}}\right)[S]}$$

equation 6.9

6.3.3 Kinetic characterisation of hFEN1 and hFEN1Δ336

Prior to kinetic studies with inhibitors the steady-state catalytic parameters were determined for full-length hFEN1 and the crystallographic construct hFEN1Δ336. The activity of WThFEN1 and hFEN1Δ336 was determined upon a static double flap substrate. This substrate had a 5nt 5'-flap and a 1nt 3'-flap but the sequence was different to the DF-5 substrate previously described. It was discovered that the template strand of DF-5 could form relatively stable homodimers (*figure 6.10*), therefore, the sequence was altered to avoid their formation. Since the newly designed substrate had the same dimensions as the original substrate (*figure 6.10*) it was denoted as DF-5*. Unsurprisingly, for a structure-specific nuclease the change in sequence did not affect the cleavage site specificity. Therefore, endonucleoytic cleavage of this DF-5* was analysed in an analogous manner to that described in *section 3.4* for determination of the single-turnover rate on DF-5. The difference in catalytic and binding parameters of DF-5* and DF-5 are statistically insignificant. Nevertheless, it was felt that to have the utmost confidence in the effect the inhibitor has upon the FEN catalysed reaction a substrate which would not form aberrant structures should be used.
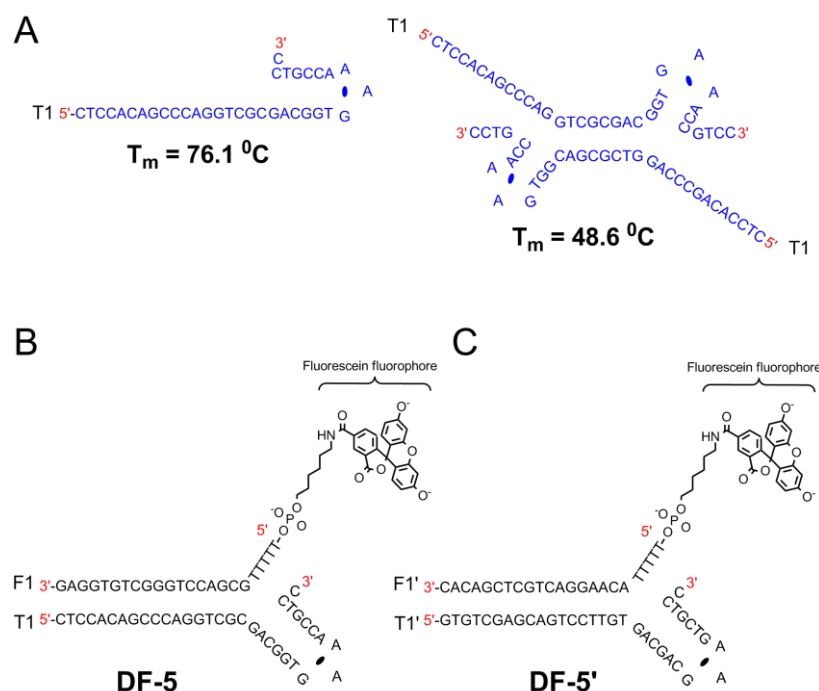
***Figure 6.10: Comparison of DF-5 and DF-5* oligonucleotide structures. (A)** Schematic representation of possible homodimer structures formed by the template strand (T1 / blue) of DF-5. Left, expected homodimer structure, this forms due to presence of hairpin, the flap strand (F1 / black) is still capable of binding and forming the expected heteroduplex (DF-5). Right, unexpected homodimer structure, folding into this structure does not allow the heteroduplex to form. The melting temperature (Tm) shown below each structure is predicted using 'DINAmelt' software. **(B)** Schematic representation of DF-5 substrate and comparison with **(C)** DF-5* substrate, heteroduplex formed from flap strand (F1*) and template strand (T1*). The newly designed template strand (T1*) is not capable of forming a homodimer structure of significant stability.*

Experimentally, the Michaelis-Menten relationship is valid when $[E] \ll [S]$. Therefore, the final enzyme concentration was kept to less than 0.01 % of the total substrate concentration. In addition, to ensure the rate being measured, v (nM min^{-1}), is the initial rate of the product formation, reaction was only ever taken to 10 % product or less. This is particularly relevant for hFEN1 to avoid seeing effects due to product accumulation and resulting competitive inhibition.

Figures 6.11a and *6.11b* show the graphs resulting from fitting of the normalized initial rates of reaction ($v_0/[E]_0$ (min^{-1})) vs. substrate concentration for WThFEN1 and hFEN1Δ336, respectively. The normalized initial rate was measured in at least triplicate at twelve substrate concentrations that span four orders of magnitude. The concentration range selected ensures that data are collected above and below the suspected K_M. In addition, the assay was designed in such a way that the salt concentration was constant (110 mM KCl) over the substrate range assayed (*section 2.9*). Examination of the results show that FEN1 follows classic Michaelis-Menten behaviour. At low *[S]*, $v_0/[E]_0$ increases linearly with substrate concentration, then the normalized initial rate follows saturation kinetics with respect to substrate concentration. At sufficiently high *[S]*, $v_0/[E]_0$ tends towards a limiting value, which is k_{cat}. At very low substrate concentrations, where $K_M \gg [S]$, the slope of the $v_0/[E]_0$ against *[S]* plot is equivalent to k_{cat}/K_M, representing the apparent second-order rate constant for the conversion of free E and S to free E and P.

Inspection of the residuals from the non-linear regression plot of substrate concentration vs. normalized initial rate for hydrolysis of DF-5 by both WThFEN1 and hFEN1Δ336 (*figure 6.11c and 6.11d*) shows that the variance increases with increasing substrate concentration; this is a consequence of how the percentage product is calculated (*equation 3.4*). Application of 1/Y^2 weighting (*equation 6.11*), produces residuals that show a constant coefficient of variation (*figure 6.11g and 6.11h*). Although the actual values of the k_{cat} and K_M have changed relatively little, the error estimates are far more accurate. As a consequence 1/Y^2 weighting was applied to all steady-state rates measured over a range of substrate concentrations including studies with inhibitors (*section 6.3.5*).
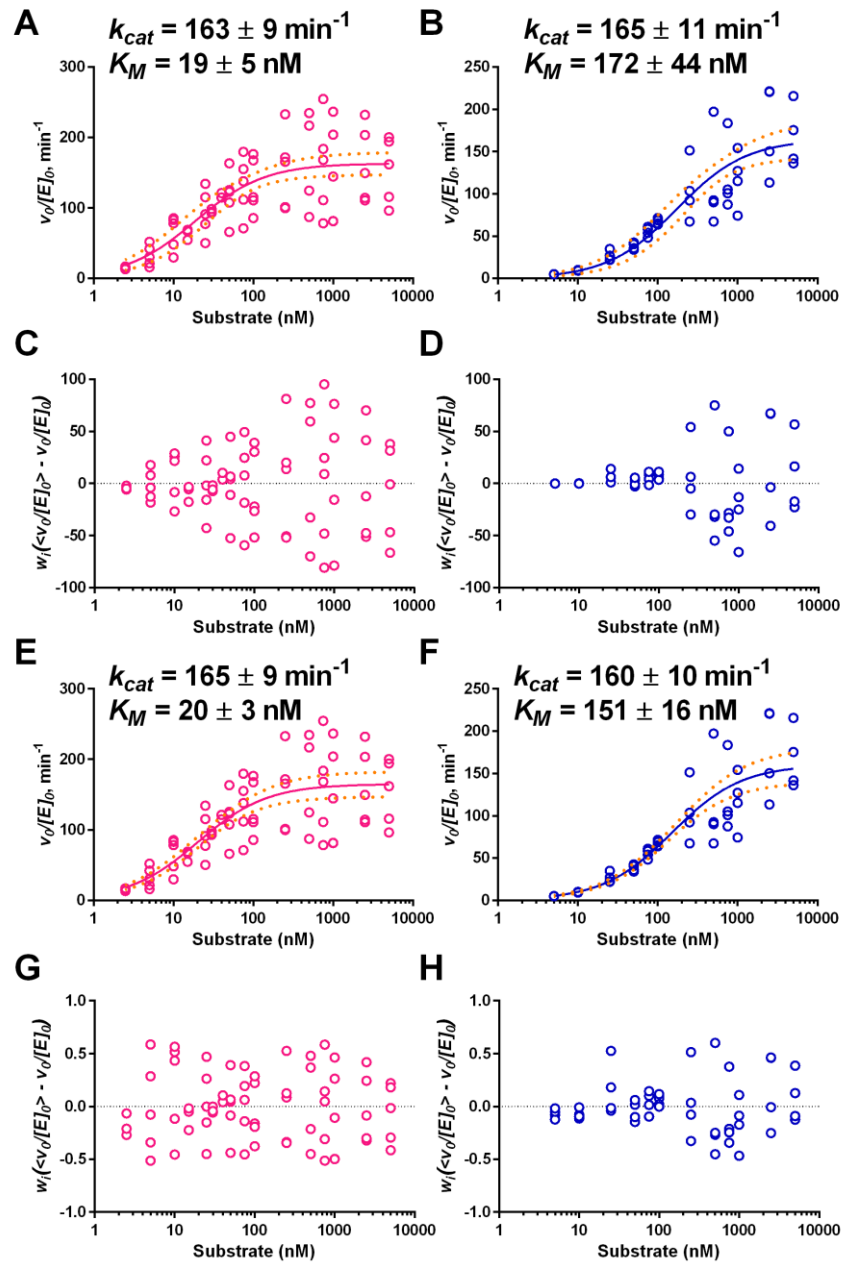
Figure 6.11: **Comparison of Michaelis-Menten plots for endonucleoytic cleavage of DF-5* by WT *hFEN1 and hFEN1Δ336*.** *For WThFEN1 catalysis, normalized initial rate for each specific substrate concentration shown as open circle (magenta), and the 95% confidence limit is shown as a dashed line (orange). (**A**) Result of fitting data using non-linear regression without applying weights, insert displays the derived multiple turnover kinetic parameters from this fit, corresponding residual plot shown in (**C**). (**E**) Result of non-linear regression when variance is weighted $1/Y^2$, insert displays the derived multiple turnover kinetic parameters from this fit, corresponding residual plot shown in (**G**). For hFEN1Δ336 normalized initial rate for each specific substrate concentration shown as open circle (blue), and the 95% confidence limit is shown as a dashed line (orange). (**B**) Result of fitting data using non-linear regression without applying weights, insert displays the derived multiple turnover kinetic parameters from this fit, corresponding residual plot shown in (**D**). (**F**) Result of non-linear regression when variance is weighted $1/Y^2$, insert displays the derived multiple turnover kinetic parameters from this fit, corresponding residual plot shown in (**H**).*

Table 6.2: Comparison of catalytic and binding parameters of DF-5* with WThFEN1 and hFEN1Δ336. *The turnover number (k_{cat}) and Michaelis constant (K_M) were determined under steady state conditions, accompanying graphs shown in figure 6.11. The values of k_{cat}/K_M were determined independently from the slope of the Michaelis-Menten plot at low substrate concentrations. The binding constant (K_D) was measure by FA, accompanying graphs shown in figure 6.12. The single turnover rate (k_{ST}) was determined under enzyme saturating conditions ($10 \times K_D$), accompanying graphs shown in figure 6.13. Errors represent standard errors of experiments performed in at least triplicate. Reactions were performed in 110 mM KCl, 55 mM HEPES-NaOH pH=7.5, 0.1 mg/ml BSA, 1 mM DTT and 8 mM MgCl$_2$ when measuring catalytic parameters or 10 mM CaCl$_2$ for binding equilibria.*

Protein	k_{cat} (min^{-1})	K_M (nM)	k_{cat}/K_M (nM^{-1}min^{-1})	K_D (nM)	k_{ST} (min^{-1})	$t_{1/2}$ (s)
WThFEN1	165 ± 9	20 ± 3	9.17 ± 1.10	77 ± 7	918 ± 50	4.5×10^{-2}
hFEN1Δ336	160 ± 10	151 ± 16	0.70 ± 0.04	150 ± 13	775 ± 35	5.4×10^{-2}

Comparison of the catalytic parameters for endonucleoytic cleavage of DF-5* by WThFEN1 and hFEN1Δ336 show that at saturating *[S]* there is no statistical difference in the turnover rate, 165 ± 9 and 160 ± 10 min^{-1}, respectively (*table 6.2*). To complement steady state parameters the single-turnover rates of DF-5* hydrolysis were measured. The single-turnover rate for WThFEN1 and hFEN1Δ336 was significantly faster than their corresponding multiple-turnover rates (*table 6.2*). Thus, as observed in previous FEN studies, hFEN1-catalysed reactions performed under multiple turnover conditions in this study are rate-limited by product release . The data suggests that there are small differences in the stability of all enzyme bound species, (i.e., forms of ES and EP), with WThFEN1 compared to hFEN1Δ336. In support of this the K_D of WThFEN1 compared to hFEN1Δ336 was 2-fold lower (*table 6.2*).

The single turnover rate was measured at [E] 10-fold higher than the respective K_D value of WThFEN1 and hFEN1Δ336. The k_{ST} value for WThFEN1 was also determined at a final enzyme concentration 10-fold higher than the K_M value. As a consequence of the difference in K_M and K_D value the rate was measured at a final [E] of 200 nM and 1000 nM. However, there was no difference in the observed rate (*Appendices, figure A4*). Therefore, this suggests the k_{ST} rate of 918 ± 50 min^{-1} is the maximal single turnover rate (k_{STmax}). In summary, loss of the C-terminal tail results in a modest decrease in affinity for the substrate, but once the enzyme:susbtrate complex has formed has no effect of the molecular mechanism. This analysis was performed because steady-state reactions with FEN1 inhibitors were performed with hFEN1Δ336. Therefore, results acquired with hFEN1Δ336 and inhibitors would be applicable to WThFEN1 (*section 6.3.5*).
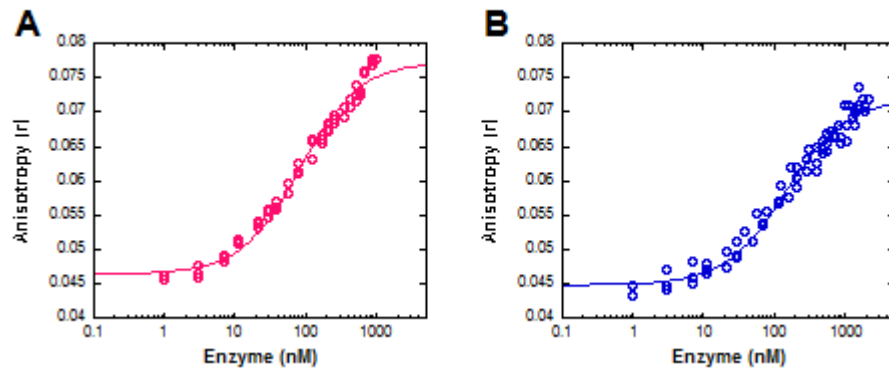
Figure 6.12: DF-5* binding curves for (A) WThFEN1 and (B) hFEN1Δ336 measured by FA. WThFEN1, K_D = 77 ± 7 nM, data points-open circles (magenta), line of best fit- solid (magenta). hFEN1Δ336, K_D = 150 ± 13 nM, data points-open circles (blue), line of best fit- solid (blue). *Assays were performed in 10 mM CaCl$_2$, 110 mM KCl, 55 mM HEPESpH=7.5, 0.1 µg/µl BSA, 1 mM DTT and at 37 ˚C. The data points were fit to equation 3.3.*
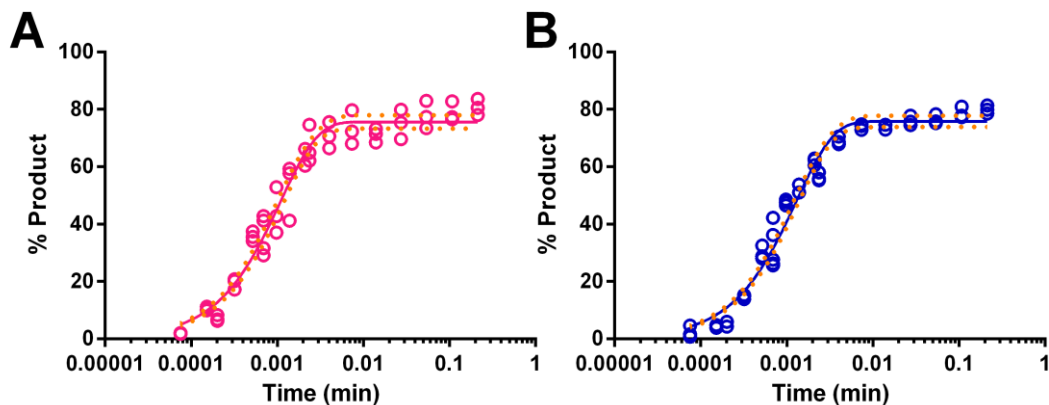
Figure 6.13: The single turnover rate profiles for DF-5* cleavage by hFEN1 and truncated hFEN1 proteins. WThFEN1, k_{ST} = 918 ± 50 min^{-1}, $t_{1/2}$ = 5.6X10^{-2} s, data points-open circles (magenta), line of best fit- solid (magenta) and the 95% confidence limit is shown as a dashed line (orange). hFEN1Δ336, k_{ST} = 775 ± 35 min^{-1}, $t_{1/2}$ = 5.6X10^{-2} s, data points-open circles (blue), line of best fit- solid (blue) and the 95% confidence limit is shown as a dashed line (orange). The substrate concentration was 2.5 nM for both enzymes. The final enzyme concentrations were 1000 nM for WThFEN1 and 1500 nM for hFEN1Δ336. Reactions were performed in 8mM MgCl$_2$, 110mM KCl, 55mM HEPES-NaOH pH = 7.5, 1mM DTT and 0.1µg/µl BSA and at 37 ˚C. Data were fitted to equation 3.5, and each single turnover rate constant is representative of at least three replicates.

6.3.4 Estimation of K_i for hFEN1 inhibitors

For initial evaluation of inhibitor potency, the reaction velocity was measured over a range of inhibitor concentrations at a single substrate concentration. The substrate concentration was fixed at the K_M value (150 nM) of hFEN1Δ336 so that both non-competitive and competitive contributions of the potential inhibitor interaction would be measured. The data were fit to a simple competitive inhibition model (*equation 6.12*) to determine an estimate of the K_i value for each inhibitor (*figure 6.14*).

$$\frac{k_{obs}}{k_0} = \frac{K_i}{K_i + [I]} \qquad equation\ 6.12$$

Where k_{obs} is the observed normalized initial rate $v_0/[E]_0$ at a given concentration of inhibitor and k_0 is $v_0/[E]_0$ in the absence of inhibitor.
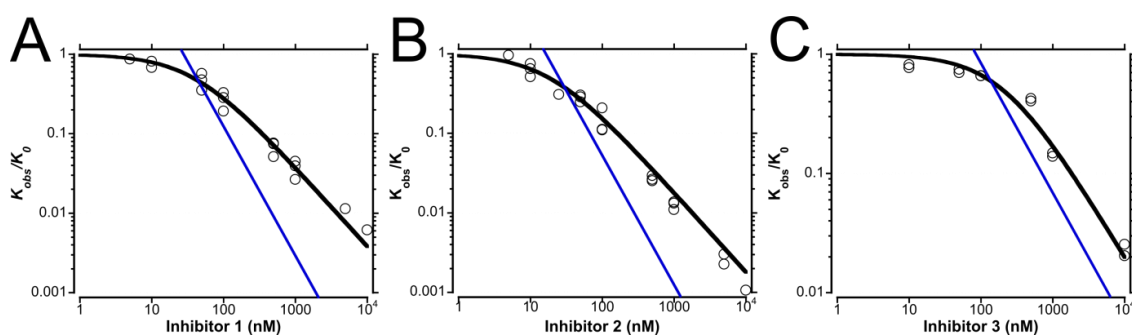
Figure 6.14: The inhibitory effects of N-hydroxy urea compounds on the hFEN1Δ336 catalysed hydrolysis of DF-5*. *Experiments were conducted at 150 nM [DF-5*] with varying amounts of inhibitor. Reactions were performed in 8mM MgCl₂, 110mM KCl, 55mM HEPES-NaOH pH = 7.5, 1mM DTT and 0.1µg/µl BSA and at 37 ˚C. (**A**) Variation in k_{obs}/k_0 as a function of inhibitor **1** concentration. (**B**) Variation in k_{obs}/k_0 as a function of inhibitor **2** concentration. (**C**) Variation in k_{obs}/k_0 as a function of inhibitor **3** concentration. Data were fitted to equation 6.12 to yield respective values of apparent K_i of 38 ± 3 nM, 18 ± 2 nM, and 203 ± 37 nM. The black line shows the best fit to a simple competitive inhibition scheme where one inhibitor displaces one substrate. The blue line is a slope of -1.*

From this analysis the apparent K_i value (the concentration of inhibitor where $k_{obs}/k_0 = 0.5$) of inhibitor **1** was 38 ± 3 nM for hFEN1Δ336. This initial value of the inhibition constant was in good agreement with the previously published IC_{50} value of inhibitor **1** (11 nM). In addition, inhibitors **2** and **3** were confirmed as active inhibitors of hFEN1, displaying apparent K_i values of 18 ± 2 and 203 ± 37, respectively. The results resemble the binding constants of the inhibitors determined by ITC for the free protein (*section 6.2.1*).

To ensure the probe compound remained soluble, assays were designed in such a way that the DMSO concentration was 1 % of the final reaction solution. The apparent second-order rate constant (k_{cat}/K_M) of FEN-catalyzed hydrolysis of the DF-5* substrate reactions were determined at 5 % DMSO. The normalized initial rates obtained under these conditions were identical within standard error, and statistically there was no change in reaction rate compared to 0 % DMSO (data not shown). In addition, the hFEN1 reaction has been shown to be able to tolerate up to 10 % DMSO (McWhirter et al., 2013). Furthermore, as would be expected, DMSO did not alter the cleavage site specificity of hFEN1.

6.3.5 Steady state evaluation of hFEN1 inhibition demonstrate inhibitor 1 and inhibitor 2 binds to DNA-free and DNA-bound hFEN1

Kinetic studies are routinely employed to determine the type of inhibition and the inhibition constants. Variation of the substrate and inhibitor concentrations allows both of these parameters to be obtained simultaneously. At low substrate concentrations any competitive element of inhibition will be more noticeable; conversely any uncompetitive element of the inhibition will manifest itself at high substrate concentrations. The estimation of K_i (*section 6.3.4*) was performed at a range of inhibitor concentrations at a constant amount of substrate. Seven further series were performed in the same way, changing only the substrate concentration. It was decided at this point it was not worth pursuing the full kinetic characterisation of inhibitor **3** as initial results (ITC and estimation of K_i) were more promising for inhibitors **1** and **2**.

Substrate and inhibitor concentrations were varied around their respective Michaelis and inhibition constants. Specifically, the lowest substrate concentration was set to < $0.2K_M$ and the highest substrate concentration was $50K_M$. The concentration of inhibitors **1** and **2** used were extended from approximately $0.2K_i$ to approximately $20K_i$, the upper limit was determined by the rate that could be measured accurately on a suitable timescale. The initial rates of reaction of DF-5* hydrolysis by hFEN1 in the presence of either inhibitor **1** or inhibitor **2** were globally fit using non-linear least-squares regression to the four inhibition model described above, competitive, uncompetitive, non-competitive and mixed inhibition models (*equations 6.6-6.9*).

Inhibitor 1 results

As expected an attempt to fit the data for inhibitor **1** to an uncompetitive inhibition model where the inhibitor is able to bind reversibly to the enzyme-substrate (ES) complex but does not bind

to the free enzyme did not produce a good fit (*figure 6.15a*). The competitive model, where the binding of inhibitor and substrate are mutually exclusive, does not fit the data at high concentrations of inhibitor **1** (*figure 6.15b*). Although an uncompetitive and competitive inhibition model could be ruled out based on visual evaluation of the respective fits, a choice between a purely non-competitive (*figure 6.15c*) and mixed (*figure 6.15d*) inhibition model could not be made with any confidence. Both models produced reasonable fits to the data and both assume that the inhibitor can bind to both the free enzyme and the enzyme-substrate complex. However, it is apparent at the highest substrate concentrations that the non-competitive model fits poorly at high concentrations of inhibitor **1**, and the mixed inhibition model is imperfect at low inhibitor **1** concentrations. The kinetic parameters derived from each of these fits are displayed in *table 6.3*. These parameters can also be used to assess the goodness of fit by comparing them to previously determined values for hFEN1Δ336 in the absence of inhibitor (*table 6.2*). Michaelis–Menten parameters, k_{cat} and K_M, produced by the mixed inhibition model are in better agreement than those produced by the non-competitive inhibition model (*table 6.3*).

Table 6.3: Variation of the steady-state catalytic parameters of hFEN1Δ336 as a function of inhibitor 1 concentration. *The data have been fit to equations 6.6, 6.7, 6.8 and 6.9 to obtain the turnover number (k_{cat}), Michaelis constant (K_M), competitive inhibition constant (K_{ic}), and uncompetitive inhibition constant (K_{iu}) for the respective inhibition models, accompanying graphs shown in figure 6.15. Reactions were performed in 8mM MgCl₂, 110mM KCl, 55mM HEPES-NaOH pH = 7.5, 1mM DTT and 0.1µg/µl BSA and at 37 ˚C. All measurements represent the result of three independent experiments with standard errors shown. Parameters could not be determined (n.d) when the model fit the data poorly, and were a particular parameter did not enter into the model this was not applicable (n.a).*

Inhibition model	k_{cat}, min⁻¹ (± S.E.)	K_M, nM (± S.E.)	K_{ic}, nM (± S.E.)	K_{iu}, nM (± S.E.)
Uncompetitive	n.d	n.d	n.d	n.d
Competitive	108 ± 7	200 ± 26	31 ± 4	n.a
Non-competitive	173 ± 10	433 ± 33	68 ± 5	68 ± 5
Mixed	140 ± 9	297 ± 31	48 ± 5	117 ± 27
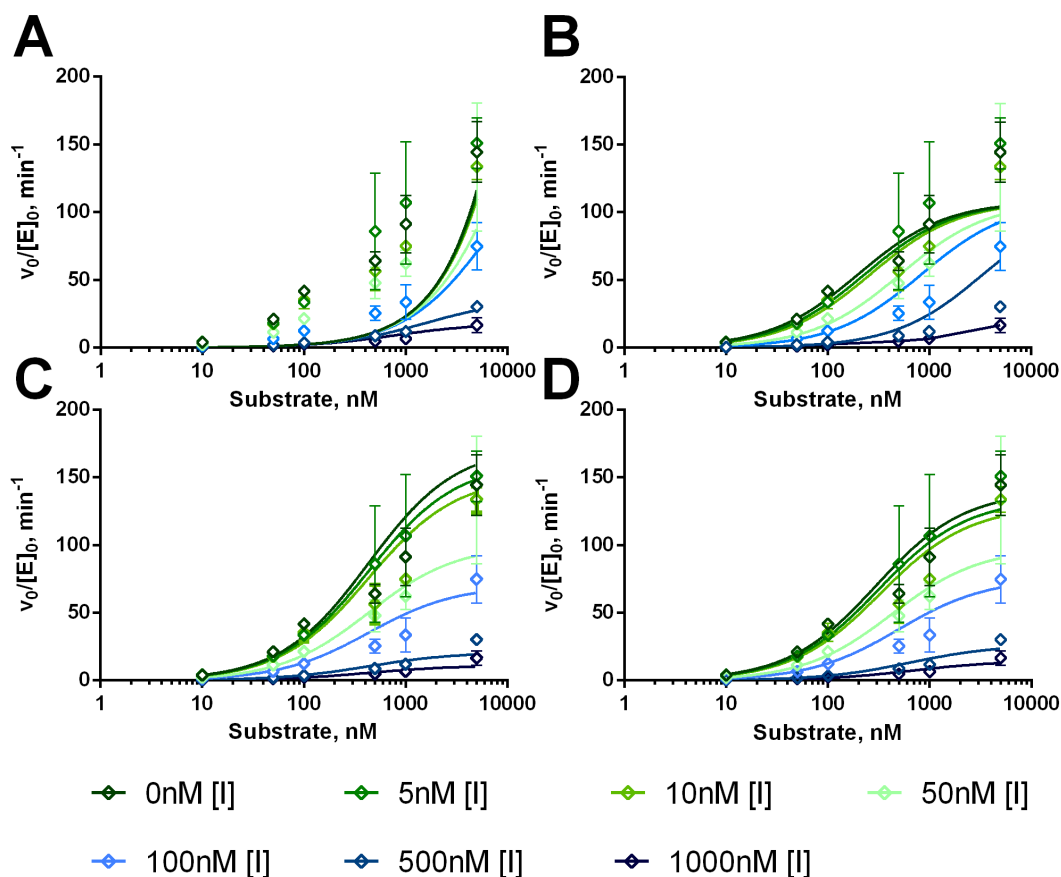
***Figure 6.15: Nonlinear regression plots of normalized initial rate vs. substrate concentration for substrate (DF-5*) cleavage at varying inhibitor 1 concentrations. (A)** Uncompetitive, data was fit to equation 6.6. **(B)** Competitive, data was fit to equation 6.7. **(C)** Non-competitive, data was fit to equation 6.8. **(D)** Mixed, data was fit to equation 6.9. Reactions were performed in 8mM MgCl₂, 110mM KCl, 55mM HEPES-NaOH pH = 7.5, 1mM DTT and 0.1µg/µl BSA and at 37 ˚C. The data points (open diamonds) are the average of three replicates and standard errors are shown.*

Reciprocal plots can be used to determine K_{ic} and K_{iu} values but due to problems associated with error propagation non-linear regression is the preferred method. However, such plots can still be useful to guide model selection. A Dixon plot, in which the reciprocal initial velocity $(1/v_0)$ is plotted against the inhibitor concentration [I], at different concentrations of substrate, is limited in this purpose because competitive and mixed inhibitions models cannot be discerned visually. However, a modified version of the Dixon plot in which S/v_0, instead of $1/v_0$ is plotted against the inhibitor concentration at the same values of [S] distinguishes mixed and competitive inhibition models allowing the inhibition model to be determined graphically (Cornish-Bowden,A., 1974). For reference characteristic Dixon and modified Dixon plots for the common

types of inhibition are compared alongside each other in *Appendices figure A5*. Plots of S/v_0 against [I] for reaction rates obtained with inhibitor **1** were generated for mixed, non-competitive and competitive inhibition models (*figure 6.16*). For each inhibition model S/v_0 values were calculated using the catalytic parameters obtained by non-linear regression and plotted using the relevant equation in the reciprocal form. These values were then used to construct a reciprocal plot, raw S/v_0 values were then plotted on the same graph to examine were they fell on the best fit lines generated using the calculated S/v_0 values. This was a complementary approach to the non-linear regression fits that allowed the goodness of fit to be assessed visually. The reciprocal data more clearly shows how poor the competitive fit was a high concentrations of enzyme (*figure 6.16c*). In addition, the raw reaction rates lie closer to the straight line generated for the highest substrate concentration (5000 nM) for the mixed inhibition model (*figure 6.16c*) compared with the non-competitive inhibition model. The choice between a mixed inhibition model and non-competitive inhibition model was not so clear based on comparisons at the other substrate concentrations. However, it can be concluded from the analyses that inhibitor **1** is capable of binding to the free enzyme and the enzyme:substrate complex.
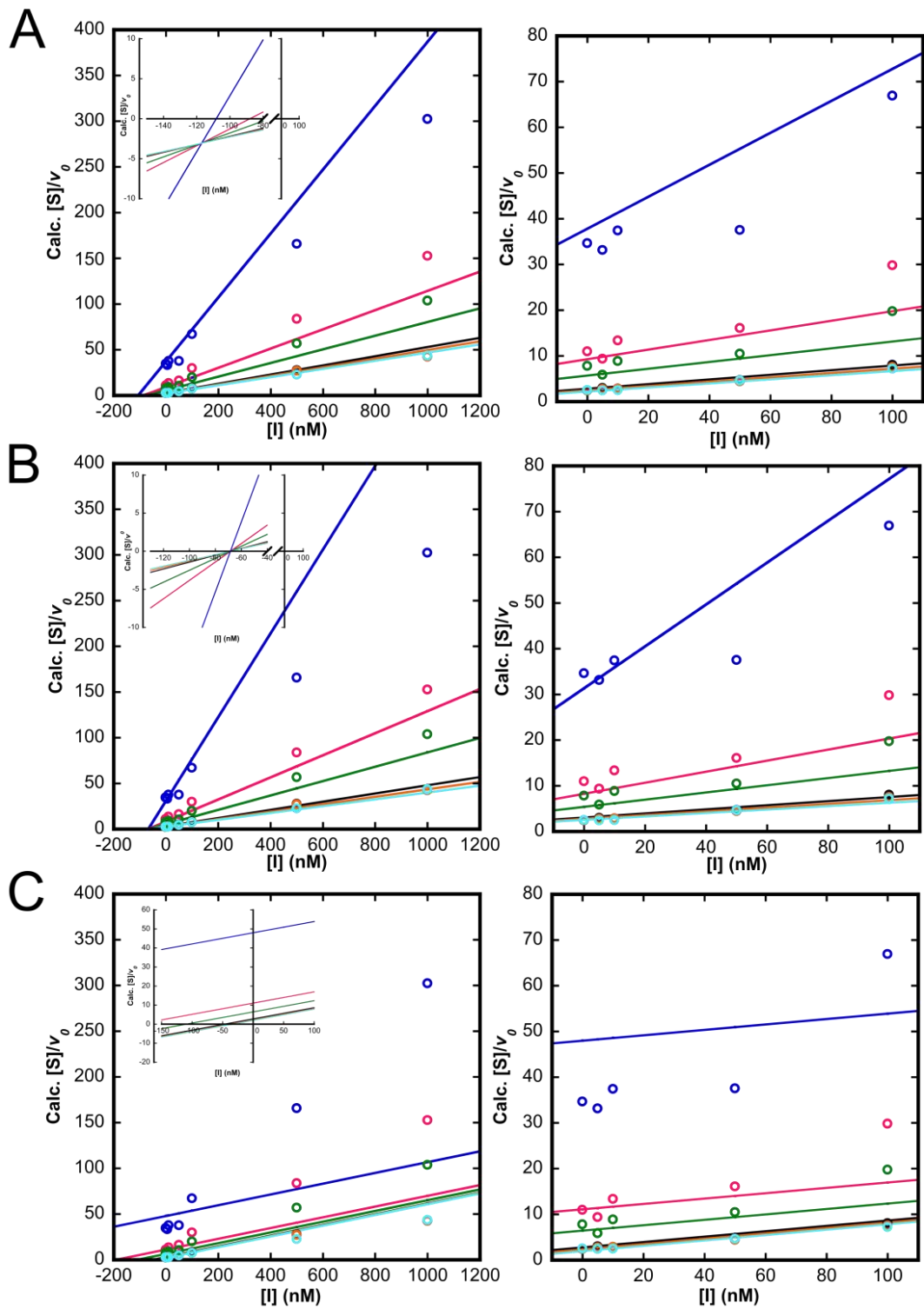
Figure 6.16: Plots of $[S]/v_0$ ***against inhibitor 1 concentration,*** *for (**A**) mixed, (**B**) non-competitive, (**C**) competitive inhibition models. 5000 nM [S] blue, 1000 nM [S] magenta, 500 nM [S] green, 100 nM [S] black, 50 nM [S] orange, 10 nM [S] cyan. In every case the intersection point in the plot provides a measure of* K_{ic}.

A more complicated model will always fit the data better. Consequently, a comparison of the sum-of-squares is unsufficient to aid model selection. Therefore, models must be compared statistically. Two such statistical tests are the classical F test, and Akaike's Information criterion (AIC). Unlike the F test, AIC does not require a null hypothesis to be stated and a statistical significance to be set. Rather, this method determines which model is correct and quantifies how much more likely. The theoretical basis of how this is achieved is explained in *Model Selection and Multimodel Selection* (Anderson and Burnham, 2002), far beyond the scope of this thesis. An AIC value is calculated for each model, and the difference in AIC ($\Delta \text{AIC}_\text{C}$) reports on the likelihood that the more complicated model is correct. A difference in AIC of 6 is considered to represent a 99.5 % probability that the more complicated model is correct. In addition, the advantage of AIC over the F test is that it allows comparison of three or more models as there is no relation back to an arbitrary P value. As such AIC was used to assist in model selection.

Application of the AIC test in Prism® indicated with greater than 99.9% confidence that the mixed inhibition model is the preferred model for the data. Comparison of the mixed inhibition model with the competitive and non-competitive model yielded $\Delta \text{AIC}_\text{C}$ values of 76 and 25, respectively. This shows that more complicated model, the mixed inhibition model, is overwhelmingly preferred. This model assumes that the inhibitor can bind to substrate-free and substrate-bound forms of the enzyme. The affinity of inhibitor **1** for the free enzyme (K_{ic}) and enzyme:substrate complex (K_{iu}) were determined to be 48 and 117 nM, respectively (*table 6.3*). This indicated that inhibitor **1** had a higher affinity for the enzyme:substrate complex than the free enzyme. The kinetically determined value of K_{ic} is in good agreement with dissociation constant determined by ITC for inhibitor **1** and the free enzyme in the presence of Mg^{2+} ions (*table 6.1*). As mention above, it is unlikely that inhibitor **1** binding will increase the affinity of the substrate for the enzyme, thus, this data also suggests that the inhibition is partly as consequence of preventing substrate binding.

Unsurprisingly, due to the similarity between the mixed and non-competitive inhibition models comparison between non-competitive and competitive models for inhibitor **1** predicted with greater than 99.9% confidence that the non-competitive inhibition model is correct; these two models have same number of parameters. Interestingly, the F test does not reject either model.

Finally, as an additional check that we have selected the correct model, the residuals from both the non-competitive and mixed inhibitions models were inspected (*Appendices, figure A6*). From

this, one can see that the residuals are better distributed around the mean in the mixed inhibition model than the non-competitive model.

Inhibitor 2 results

The data from the inhibitor **2** assay were analysed analogously to that above. Comparison of inhibition models indicated that like inhibitor **1**, inhibitor **2** was also a mixed inhibitor. In this instance, visual discrimination can rule out not only an uncompetitive inhibition model but also a non-competitive one (*figure 6.18a and c, respectively*). In contrast, competitive and mixed produced acceptable fits (*figure 6.1b and d, respectively*) and the kinetic parameters derived from each of these fits are displayed in *table 6.4*. However, inspection of either non-linear regression (*figure 6.18*) or reciprocal plots (*figure 6.19*) cannot aid in selection between mixed and competitive inhibition models. In addition, the Michaelis-Menten k_{cat} and K_M parameters were very similar for both models (*table 6.4*). Therefore, AIC was again employed to discern between these two models. Like inhibitor **1**, the mixed inhibition model was again the preferred model for inhibitor **2**; however, the ΔAIC$_c$ was calculated to be 10 (c.f. 76 for inhibitor **1**), suggesting that the confidence that the correct model was chosen is less than compared with inhibitor **1**, but still overwhelmingly in favour of the mixed inhibition model for this particular experimental design. In addition, the residuals from both the competitive and mixed inhibitions models were inspected and suggested the mixed inhibition model fitted the data better (*Appendices, figure A7*). Finally, as an additional check, the F test was utilised to determine the statistically favoured model. This test rejects the competitive model with a P value of < 0.0005, well below the traditional value of 0.05, suggesting the simpler model can be rejected. Unlike inhibitor **1**, however, the kinetically derived dissociation constants (K_{ic} and K_{iu}) vary by an order of magnitude (*table 6.4*). Taken together these suggest a much more competitive element to the inhibition mode of inhibitor **2** compared with inhibitor **1**.

Table 6.4: Variation of the steady-state catalytic parameters of hFEN1Δ336 as a function of inhibitor 2 concentration. *The data have been fit to equations 6.6, 6.7, 6.8 and 6.9 to obtain the turnover number (k_{cat}), Michaelis constant (K_M), competitive inhibition constant (K_{ic}), and uncompetitive inhibition constant (K_{iu}) for the respective inhibition models, accompanying graphs shown in figure 6.18. Reactions were performed in 8mM MgCl$_2$, 110mM KCl, 55mM HEPES-NaOH pH = 7.5, 1mM DTT and 0.1μg/μl BSA and at 37 ˚C. All measurements represent the result of three independent experiments with standard errors shown. Parameters could not be determined (n.d) when the model fit the data poorly, and were a particular parameter did not enter into the model this was not applicable (n.a).*

Inhibition model	k_{cat}, min^{-1} (± S.E.)	K_M, nM (± S.E.)	K_{ic}, nM (± S.E.)	K_{iu}, nM (± S.E.)
Uncompetitive	n.d	n.d	n.d	n.d
Competitive	169 ± 11	374 ± 43	16 ± 2	n.a
Non-competitive	n.d	n.d	n.d	n.d
Mixed	182 ± 13	422 ± 50	17 ± 2	306 ± 125
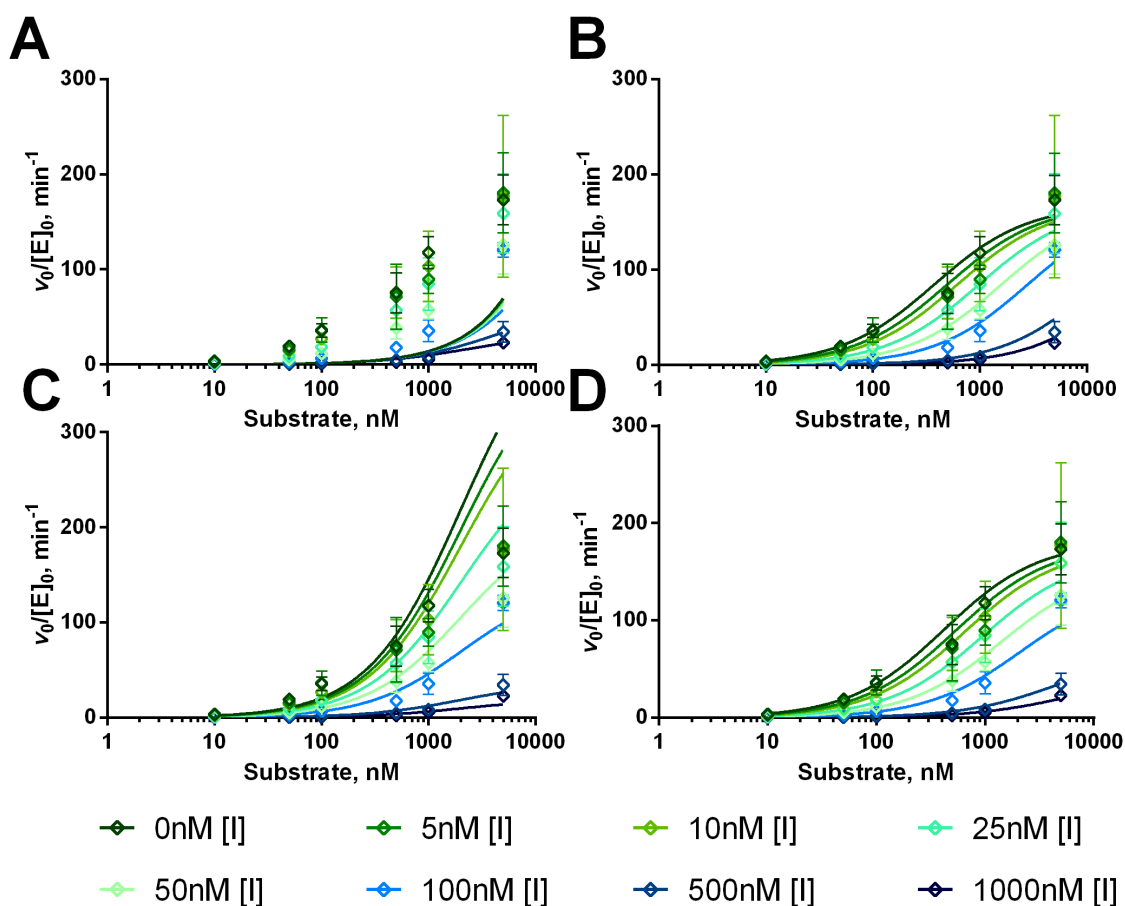
Figure 6.18: Nonlinear regression plots of normalized initial rate vs. substrate concentration for substrate (DF-5*) cleavage at varying inhibitor 2 concentrations. *(A) Uncompetitive, data was fit to equation 6.6. (B) Competitive, data was fit to equation 6.7. (C) Non-competitive, data was fit to equation 6.8. (D) Mixed, data was fit to equation 6.9. Reactions were performed in 8mM MgCl$_2$, 110mM KCl, 55mM HEPES-NaOH pH = 7.5, 1mM DTT and 0.1μg/μl BSA and at 37 ˚C. The data points (open diamonds) are the average of three replicates and standard errors are shown.*
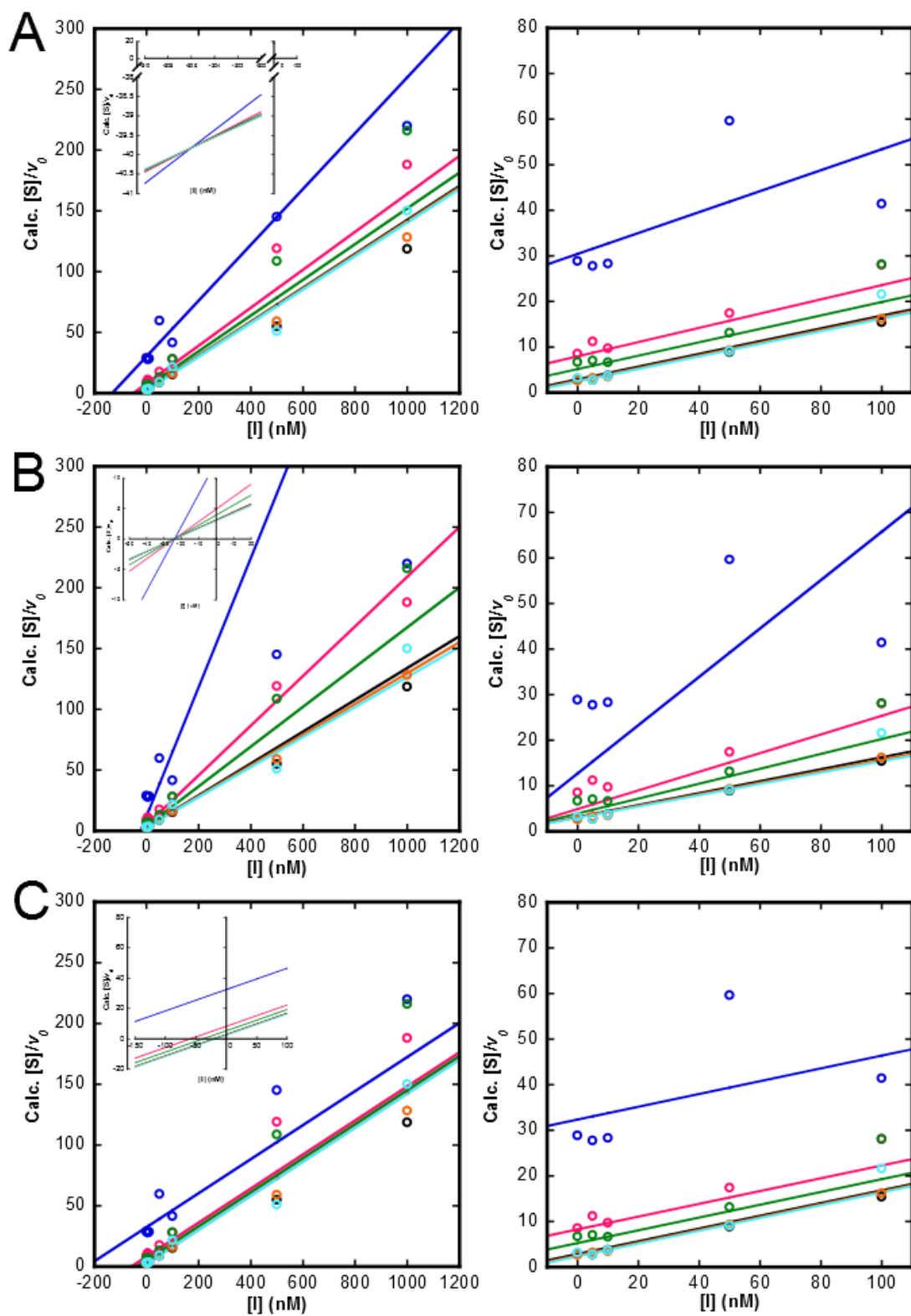
Figure 6.19: Plots of $[S]/v_0$ ***against inhibitor 1 concentration,*** *for (**A**) mixed, (**B**) non-competitive, (**C**) competitive inhibition models. 5000 nM [S] blue, 1000 nM [S] magenta, 500 nM [S] green, 100 nM [S] black, 50 nM [S] orange, 10 nM [S] cyan. In every case the intersection point in the plot provides a measure of* K_{ic}.

6.4 Investigation of an an hFEN1–Mg^{2+}–Inhibitor–DNA complex using biophysical approaches

6.4.1 Ca^{2+} ions in the active site do not support inhibitor binding

Consistent with the metal chelating potential of inhibitor **1**, binding is not observed in the presence of EDTA (data not shown). As stated previously Ca^{2+} does not support hFEN1 catalysis. Moreover, Ca^{2+} ions have been shown to be a competitive inhibitor of 5'-nuclease reactions with respect to Mg^{2+} ions (Syson et al., 2008; Tomlinson et al., 2011), implying that Ca^{2+} and Mg^{2+} occupy similar sites on the protein. Therefore, it is commonly used as a replacement divalent ion in substrate binding studies. As we were interested to determine whether the compound interacts with hFEN1 in the presence of substrate, we sought to determine whether inhibitor **1** can bind the free enzyme in the presence of 10 mM Ca^{2+}. Surprisingly, the data show that Ca^{2+} does not support binding of inhibitor **1** (*Figure 6.20*). The affinity of inhibitor **2 f**or hFEN1(Ca^{2+}) was not determined as it has a similar "war-head" group and it therefore also unlikely to bind to calcium ion in the active site. Although the reason Ca^{2+} does not support inhibitor binding is not known, it is possible that the Ca^{2+} ions are not positioned correctly in the active site to support inhibitor binding. Thus, interaction of these types of compound with hFEN1 was specific to the nuclease core domain and dependent on the presence of Mg^{2+} ions, again emphasizing that a significant contribution to hFEN1–inhibitor interaction was via the active site magnesium ions.
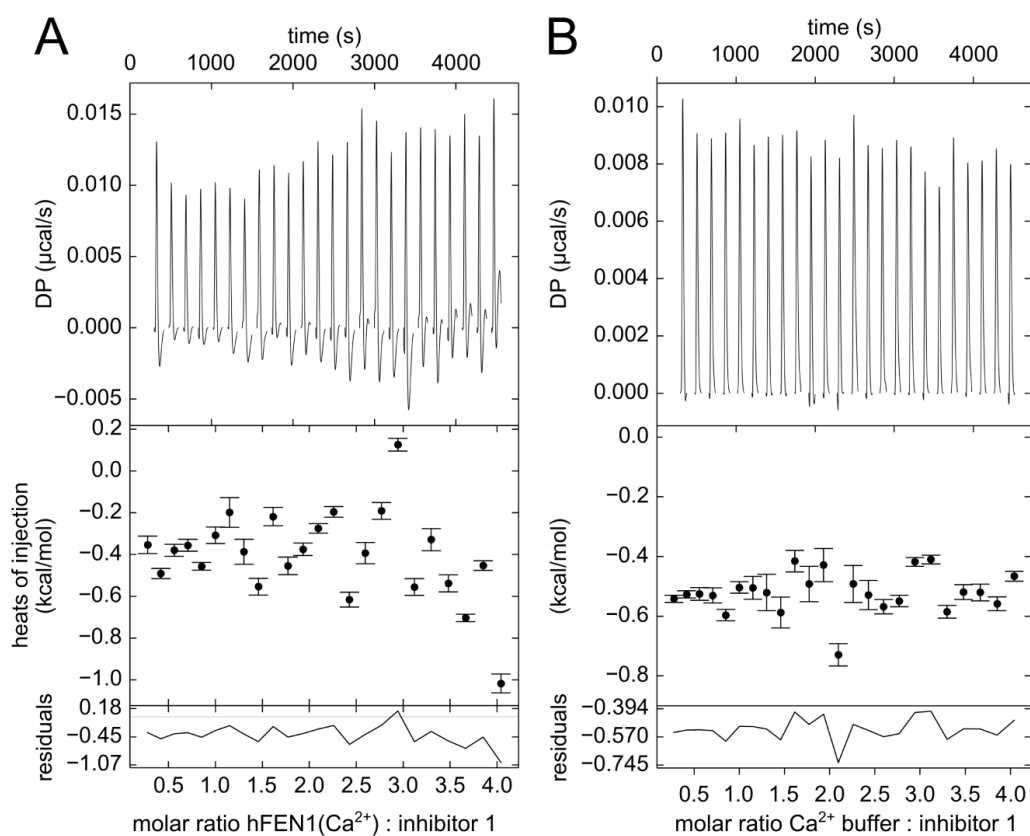
Figure 6.20: Inhibitor binding to hFEN1 in the presence of Ca²⁺ ions. (A) *Affinity of inhibitor **1** for hFEN1Δ336 assessed by ITC at pH 7.5 and 25 ˚C in the presence of 10 mM Ca²⁺. **(B)** Titration of inhibitor **1** in buffer containing 10 mM Ca²⁺. The top, middle and bottom panels show the raw, processed data and reciprocals, respectively.*

6.4.2 Full-length R100A is suitable for biophysical assays with inhibitors.

To verify that a quaternary complex of enzyme(Mg²⁺)–inhibitor–DNA (E(Mg²⁺)–I–DNA) was formed, the ability of E(Mg²⁺)–I to form complexes with DNA was tested. Biophysical studies were performed using the full length WThFEN1 construct opposed to hFEN1Δ336. The full-length protein has a lower K_D than the C-terminal truncated form (*table 6.2*). Therefore, it was perceived that any changes in the binding affinity of the substrate for WThFEN1 in the presence and absence of inhibitor would be more easily detected. Accordingly, binding of inhibitor **1** was also evaluated with full-length protein in the presence of Mg²⁺ using ITC and the result was similar to the parameters for hFEN1Δ336 (*table 6.21*). The stoichiometry of binding was also determined to be approximately 1 suggesting that the inhibitor was binding in the active site of the full-length protein and not to the positively charge C-terminal tail.

Knowing that the inhibitor cannot bind to the free enzyme in the presence of Ca^{2+} ions, FA experiments including inhibitors would have to be performed in the presence of Mg^{2+} ions. Initially, inhibitors were used at large excess (100 µM) with WThFEN1 to see whether this could prevent reaction on a time scale appropriate for binding experiments. To this end reaction rates were measured under single-turnover conditions (2.5 nM [S], 1000 nM [E]) in the presence of 8 mM Mg^{2+} and 100 µM inhibitor **1** and **2**. Data were fit to a single-exponential (*equation 3.5*) and yielded single-turnover rates of 0.48 and 1.52 min⁻¹, for WThFEN1 in the presence of inhibitor **1** and **2**, respectively; between 600-2000-fold lower than WThFEN1 in the absence of any inhibitor (*table 6.3*). However, these rates corresponded to half-life values of approximately 1.5 and 0.5 mins, which is not sufficiently long enough to prevent reaction during FA and FRET experiments that typically take 40 minutes.

Therefore, to prevent reaction, catalytically deficient variants, K93A and R100A, where chosen as potential candidates to model substrate binding to the enzyme-inhibitor complex. These particular variants shown to have comparable K_D values to WT for DF-5 (*section 3.3*). As a check that the substrate (DF-5*) remained intact for the duration of the FA titration, using the combination of inhibitor and mutation, reactions were again performed under single-turnover (ST) conditions (2.5 nM [S], 1000 nM [E]) in the presence of 8 mM Mg^{2+} and 100 µM inhibitor **1**.

Single turnover experiments in the presence of the inhibitor **1** revealed that R100A, but not K93A, was a suitable candidate for FA experiments. The half-life of K93A in the presence of 100 µM inhibitor 1 was 23 minutes (*table 6.3*) suggesting that reaction would likely be at completion at the end of typical binding assays. In contrast, the combination of the R100A mutation in the presence of 100 µM inhibitor **1** meant that only ~7% of product formed after 6hrs (*figure 6.22*). The low levels of product are not a consequence of enzyme-degradation or denaturation because reaction still occurs up to 47 hours after initiation. ITC experiments with hFEN1R100A (full-length), performed analogously to experiments described in *section 6.2,* gave a very similar K_D (~117 nM) for inhibitor **1** to the WT enzyme in the presence of 8 mM Mg^{2+} (*table 6.2*). Therefore, this particular mutation does not affect the ability of the compound to bind presumably because it does not alter metal ion sequestration. Because we planned to assess the effect of ring system B (*figure 6.4*) on the ability to form a quaternary complex, the single turnover experiments were repeated with inhibitor **2** and **3** in combination with R100A. Although inhibitors **2** and **3** were less effective at preventing reaction (*figure 6.22*), the data shows that the predicted half-life values are greater than 200 mins, thus these inhibitors are also suitable for appropriately planned biophysical assays. Inhibitor **2** and **3** are proposed to have a very similar K_D for WT as inhibitor **1** the same assumption is made here for the R100A protein

in the presence of Mg^{2+} ions. Because hFEN1R100A has essentially the same K_D as DF-5 for WThFEN1(Ca^{2+})$_2$ (*section 3.3*), this means that any disruption in the binding affinity of the substrate can be attributed to the inhibitors and not the mutation itself. Therefore, in this instance, hFEN1R100A can act as a suitable model for the WThFEN1 behaviour.

Table 6.2: Thermodynamic binding parameters of inhibitor 1 bound to hFENΔ336 1 in the presence of 8 mM MgCl₂ from global analysis of triplicate experiments: n-stoichiometry, K_D-dissociation constant (95% confidence limits), ΔH-enthalpic change, ΔS-entropic change and $χ^2$ critical chi-square value for 95% limits on global analysis. Accompanying graphs shown in figure 6.21.

Inhibitor	Protein	K_D, nM (95% C.I.)	n (S.E.)	ΔH°, kcal mol^{-1}	ΔS°, cal mol^{-1} deg^{-1}	Reduced critical $χ^2$
1	hFEN1Δ336	182 (72 – 395)	1.10 ± 0.02	1.60 ± 0.01	36	31
1	WT	260 (80 - 687)	1.12 ± 0.09	1.40 ± 0.01	35	17
1	R100A	117 (22 – 387)	1.17 ± 0.05	1.55 ± 0.02	37	70
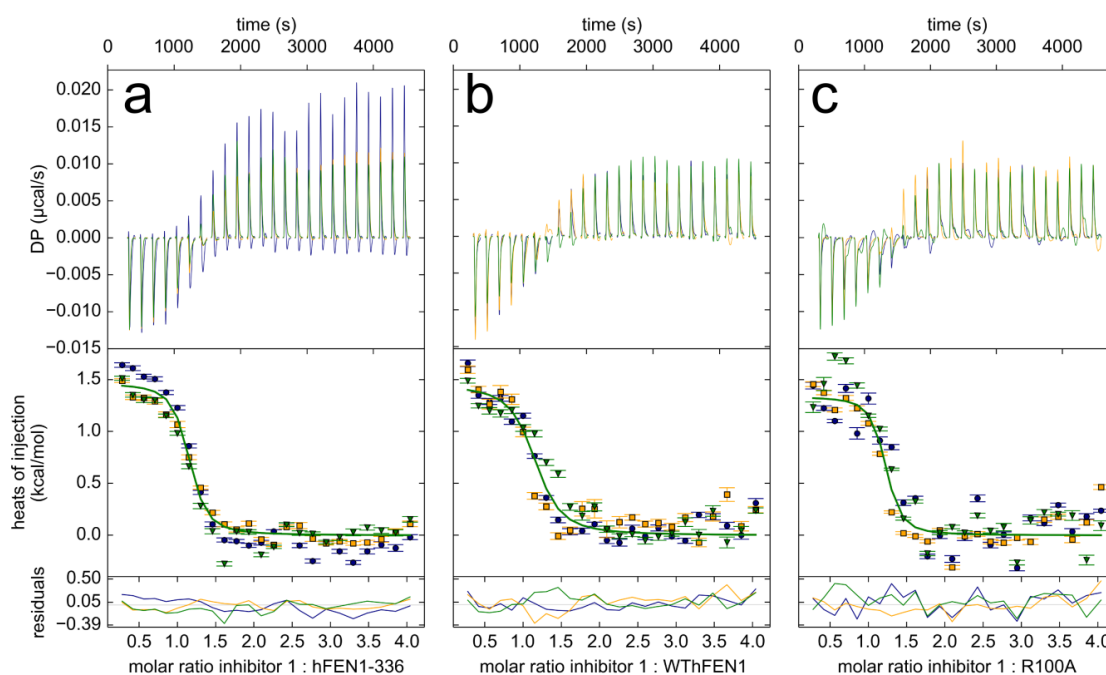
Figure 6.21: Comparison of Inhibitor 1 binding to (A) hFEN1Δ336, (B) WThFEN1 and (C) hFEN1R100A assessed by isothermal titration calorimetry (ITC). The top, middle and bottom panels show the raw data, processed data and residuals, respectively. The orange, blue, and green circles represent data from the three respective replicates used in this analysis. The residual lines in the lower graph are coloured according to their respective replicate. Each measurement was conducted at pH 7.5 and 25 ˚C in the presence of 8 mM Mg^{2+}. Parameters from global analysis of at least triplicate experiments are shown in table 6.2.

Table 6.3: Single turnover rate profiles for DF-5* cleavage by WThFEN1 and variants in the absence and presence of 100μM inhibitor. *Apart from the specified conditions, assays were performed in 110 mM KCl, 55 mM HEPES-NaOH, pH=7.5, 0.1 mg/ml BSA, 1 mM DTT and at 37 ^{0}C. The data points were fit to equation 3.5. Accompanying graphs are shown in figure 6.22.*

Enzyme	Inhibitor	k_{ST}, min^{-1}	$t_{1/2}$, min
hFEN1	none	916 ± 49	8×10^{-4}
hFEN1	1	0.48 ± 0.04	1.50
hFEN1	2	1.52 ± 0.09	0.50
K93A	none	0.36 ± 0.02	2
K93A	1	0.03 ± 8×10^{-4}	23
R100A	none	0.09 ± 3×10^{-3}	8
R100A	1	~4×10^{-4}	≤1750
R100A	2	~2×10^{-3}	≤360
R100A	3	~3×10^{-3}	≤230
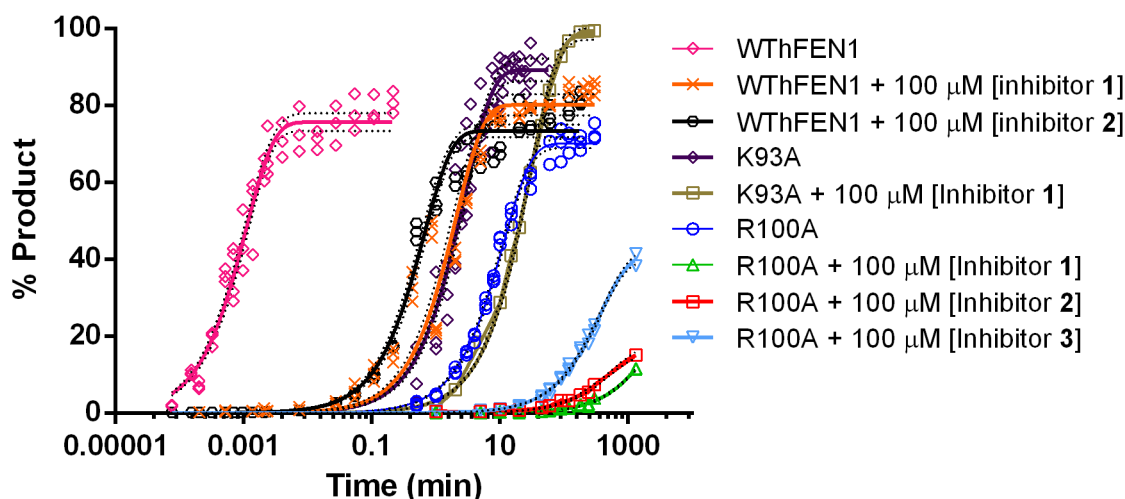
Figure 6.22: Hydrolysis of substrate DF-5* by FEN1 proteins under single turnover conditions, in the presence and absence of inhibitors. (**a**) Maximal single turnover rate profiles (% conversion *vs.* time) measured at 37 °C at pH 7.5 in the presence of Mg^{2+} (note log scale of x-axis). Magenta plot (open diamonds) full length hFEN1 no inhibitor; cyan plot (inverted open triangles); orange plot (crosses) hFEN1 plus 100 μM inhibitor **1**; black plot (open circles) hFEN1 plus 100 μM inhibitor **2**; purple plot (open diamonds) hFEN1K93A; gold plot (open squares) hFEN1K93A plus 100 μM inhibitor **1**; blue plot (open circles) hFEN1R100A; green plot (open triangles) hFEN1R100A plus 100 μM inhibitor **1**; red plot (open squares) hFEN1R100A plus 100 μM inhibitor **2**; cyan plot (open inverted triangles) hFEN1R100A plus 100 μM inhibitor **3**. In each case, dashed black lines represent 95% confidence interval of the fit.

6.4.3 Fluorescent anisotropy experiments confirm the formation of a quaternary complex.

FA was employed to monitor DNA binding to R100A(Mg^{2+})–I as described earlier for WThFEN1 and mutated FENs (*section 3.2*). FA results suggest that the substrate is still able to bind to the enzyme, albeit with a weakened affinity, when *N*-hydroxy urea inhibitors are bound in the active site (*table 6.4*). For all three inhibitors tested the *r* value increased upon addition of R100A(Mg^{2+})–I, suggesting the successful formation of R100A(Mg^{2+})–I–DNA complexes (*figure 6.23*). The data displayed hyperbolic behaviour, saturating at high concentrations of R100A–Mg^{2+}–I, and thus could be fit to a simple binding isotherm.

The dissociation constant of DF-5* for R100A(Mg^{2+})–inhibitor **1** was approximately threefold higher than for the R100A(Ca^{2+}) (*table 6.4*). This indicted that inhibitor **1** may weaken, but does not prevent, substrate binding to hFEN1R100A supporting kinetic data predicting a mixed inhibition model (*section 6.3.5*). Furthermore, the K_D of DF-5* measured in EDTA was identical in the presence and absence of inhibitor **1** as would be expected under conditions where the compound cannot bind to the enzyme (absence of Mg^{2+}) (*table 6.4*). Bearing in mind the issues encountered during FA experiments, it was decided to determine the maximal single turnover rate for hFEN1R100A using FA. The reaction was performed in the same buffer conditions using the same substrate and enzyme concentration to that used in the bench top The k_{ST} measured by FA is very similar to that measured using a discontinuous assay (*Appendices, figure A8*). This result was important because it showed how much of a dramatic effect binding of the inhibitor had on hFEN1 reaction mechanism.

Table 6.4: K_D *parameters for 10 nM DF-5* bound to free R100A, and R100A-inhibitor complexes determined by fluorescence anisotropy. All measurements were made at 37 ˚C in 110 mM KCl, 55 mM HEPES-NaOH pH = 7.5, 1 mM DTT, 0.1 mg/ml BSA and 10 mM Ca^{2+} or 10 mM Mg^{2+} or 2 mM EDTA as indicated. The dissociation constants derived from equation 3.3 are reported with the standard errors in parentheses. The minimum and maximum anisotropy values,* r$_{min}$ *and* r$_{max}$, *respectively, for triplicate experiments are reported with the standard errors in parentheses. Accompanying graphs are shown in figure 6.23. ^{a} determined using DF-5.*

Buffer Conditions	R100A K_D, nM	r$_{min}$	r$_{max}$
^{a}Ca^{2+}	92 (± 10)	0.04 ± (0.001)	0.12 ± (0.001)
Mg^{2+} + Excess inhibitor **1**	311 (± 10)	0.03 ± (0.001)	0.14 ± (0.001)
Mg^{2+} + Excess inhibitor **2**	1033 ± (31)	0.02 ± (0.001)	0.15 ± (0.004)
Mg^{2+} + Excess inhibitor **3**	341 ± (23)	0.03 ± (0.001)	0.14 ± (0.002)
^{a}EDTA	111 ± (8)	0.03 ± (0.001)	0.13 ± (0.001)
EDTA + Excess inhibitor **1**	111 ± (6)	0.02 ± (0.001)	0.13 ± (0.001)
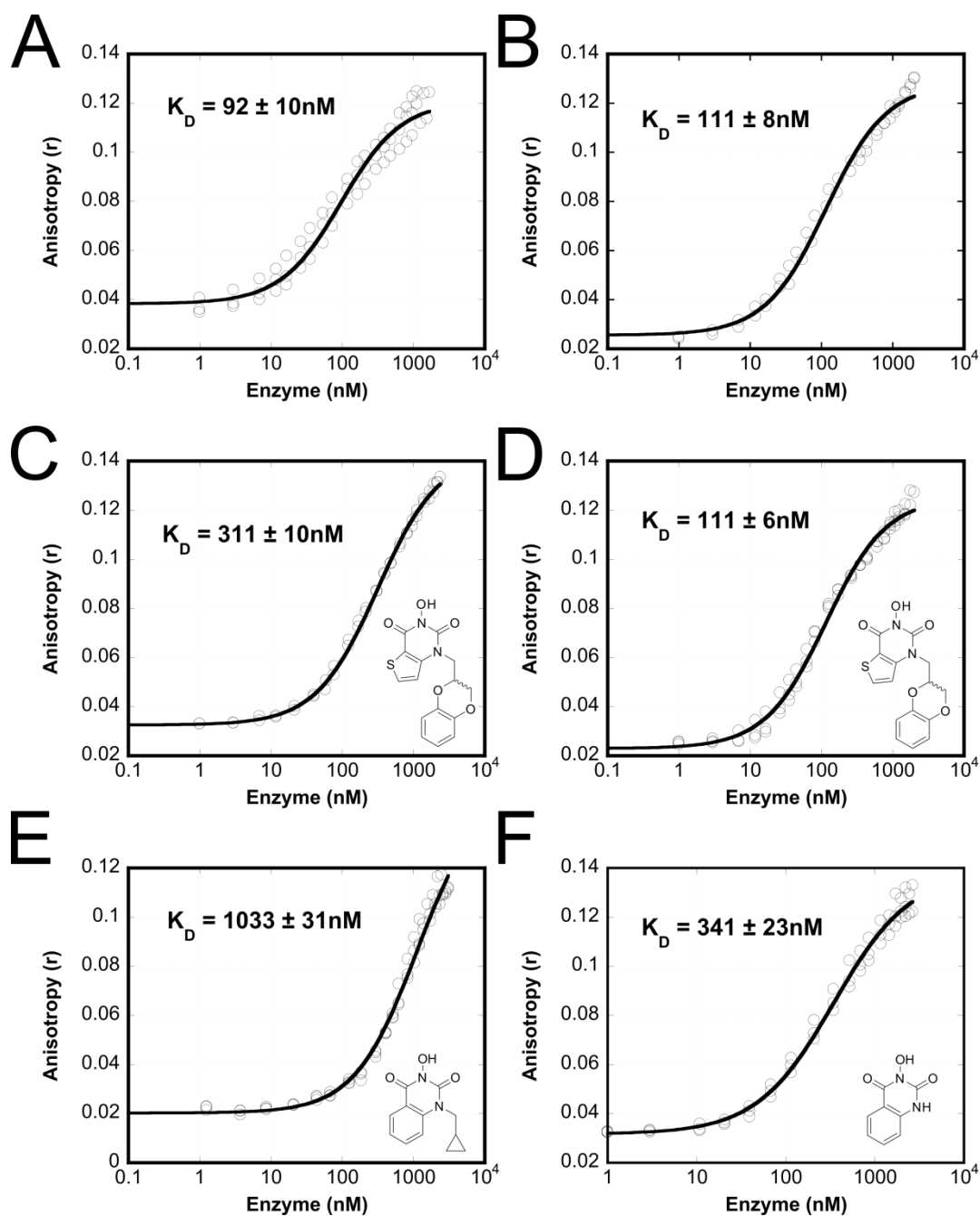
Figure 6.23: DF-5* binding curves for hFEN1R100A (A) *10 mM Ca²⁺, (**B**) 2 mM EDTA, (**C**) 8 mM Mg²⁺ and 100 μM inhibitor **1**, (**D**) 2 mM EDTA and 100 μM inhibitor **1**, (**E**) 8mM Mg²⁺ and 100 μM inhibitor **2**, and (**F**) 8mM Mg²⁺ and 100 μM inhibitor **3**. Binding assays were performed at 37 ⁰C in 110 mM KCl, 55 mM HEPES-NaOH, pH=7.5, 1 mg/ml BSA, and 1 mM DTT buffering conditions. The data points are fit to equation 3.3. The K_D value for each condition is shown in the insert of the particular graph, each K_D is representative of at least two repeats, and errors shown are standard error.*

Although the formation of putative quaternary complexes detected by FA appeared plausible, one problem with FA is light scattering that results from formation of precipitates. As scattered light is polarised, this can potentially be misinterpreted as formation of an enzyme substrate complex. To rule this out competition experiments were used to confirm the formation of the proposed quaternary complex. An unlabelled version of the substrate (UDF-5*) which has the same sequence as DF-5* (*table 2.5*) but no 5'-FAM was used to compete away the labelled substrate from the R100A-(Mg^{2+})$_2$–DF5–inhibitor **1** complex (at approximately 80% saturation of S by E). The data shows (*figure 6.24*) that the *r* value decreases upon addition of the unlabelled substrate and eventually returns to the starting value (r_{min}), confirming that all the labelled substrate has been successfully competed from the complex. This data implies that the inhibitor is bound solely at a sub-binding site and does not function as a competitive inhibitor by preventing overall substrate binding.
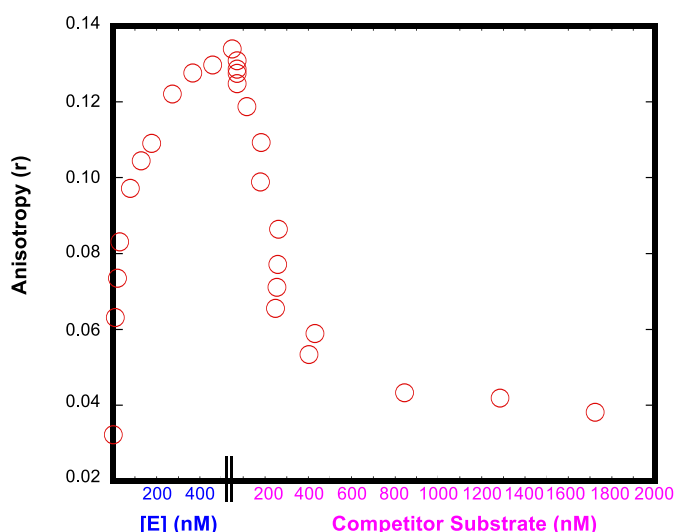
Figure 6.24: The effect of addition of unlabelled competitor substrate on the outcome of the proposed R100A–(Mg^{2+})$_2$–DF5–inhibitor 1 complex. *Enzyme was added until saturation of the substrate reached ~80%. At this point unlabelled substrate (UDF-5*) was added to the same cuvette. The competition assay was performed at 37 ^{0}C in 100 µM inhibitor 1, 8 mM MgCl$_2$, 110 mM KCl, 55 mM HEPES-NaOH, pH=7.5, 0.1 mg/ml BSA, and 1 mM DTT. The X-axis has been spilt, blue axis labels representing the enzyme concentration and the added UDF-5* concentration is represented by pink axis labels.*

The ability of DF-5* to bind to R100A(Mg^{2+})–inhibitor **2** and R100A(Mg^{2+})–inhibitor **3** were also tested. The dissociation constant measured for substrate binding to R100A(Mg^{2+}) in the presence of inhibitor **2** was threefold higher than the equivalent inhibitor **1** complex and tenfold higher when there was no inhibitor present R100A(Ca^{2+}) (*table 6.4*). Counter-intuitively, this implies that replacement of the larger aromatic side chain with the cyclopropyl group was more

disruptive to substrate binding. However, this is in good agreement with kinetic data for **2** that shows that although mixed inhibition can occur that characteristics of this are more competitive in nature than non-competitive (*section 6.3.5*). Interestingly, inhibitor **3** which had no functionalization at the same position did not improve substrate binding by removing potential steric blocking effects of ring system B. In the presence of both inhibitor **1** and **3** R100A(Mg^{2+})–I displayed an equivalent dissociation constant for the substrate (table 6.4). This analysis assumes that any difference in binding was not the consequence of the thiophene group, which is replaced by a benzene group in the two related inhibitors (ring system A, *figure 6.4*).

To ensure that reported differences were not a consequence of product production, an aliquot of the sample was taken after each titration. The dHPLC trace (*figure 6.25*) shows only a small amount of cleavage occurred (<5%), in the presence of inhibitor **1** and **2**. As expected from single-turnover data slightly more product was formed (<10%), in the presence of inhibitor **3** however it is believed that this is not a significant enough to affect the overall outcome.
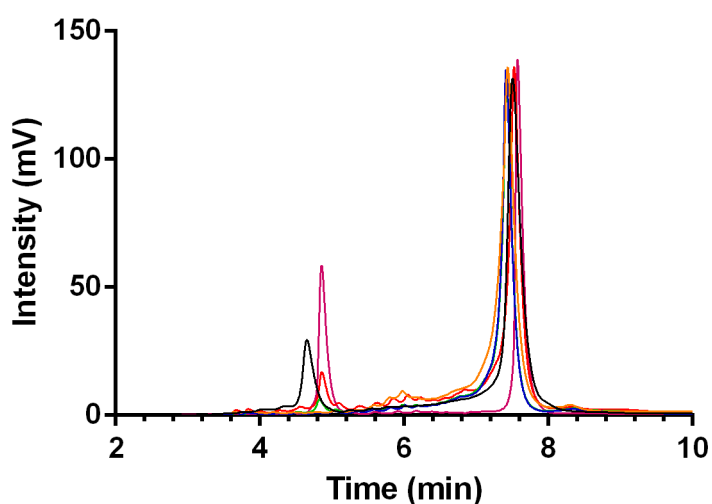
Figure 6.25: dHPLC trace to show the amount of product produced after the FA experiment to determine the K_D of DF-5* for hFEN1R100A in the presence of 8 mM MgCl$_2$ and 100 μM inhibitor 1, 2, and 3. *The orange trace is unreacted DF-5*, and the magenta trace is DF-5* reacted with WThFEN1 in the presence of Mg^{2+} to show where the product is expected to elute. The blue, green, red, and black traces are samples taken after FA of R100A in the presence of 10 mM Ca^{2+}, 8 mM Mg^{2+} and inhibitor 1, 8 mM Mg^{2+} and inhibitor 2, and 8 mM Mg^{2+} and inhibitor 3. Prior to analysis, an aliquot of each sample was diluted in 8 M Urea and 80 mM EDTA to a final concentration of 100 nM. Denaturing HPLC retention times for the product and the uncleaved flap are approximately 4.8 min and 7.5 min, respectively. Denaturing HPLC was performed on a WAVE® system, Transgenomic, UK equipped with UV detector using tetrabutyl ammonium bromide containing buffers and a linear gradient of acetonitrile as described (section 2.4.1).*

6.4.4 The substrate is in the same conformation when bound to R100A(Mg²⁺)–I and R100A(Ca²⁺) complexes

In addition to anisotropy, the change in FRET signal of the DF-4 substrate (*section 3.5.2*) was used to validate the formation of the E(Mg²⁺)–I–S complex. As discussed, this assay also had the advantage of being able to report on the conformation of the substrate when bound to FEN. An increase in FRET efficiency was observed upon titration of both inhibitor complexes (*figure 6.26*). This suggested that the downstream and upstream ds regions of the double flap substrate were capable of binding to the appropriate FEN1 motifs enforcing a kink in the substrate. The difference in dissociation constants between R100A(Ca²⁺) and R100A(Mg²⁺)–inhibitor **1** resembled changes in affinity determine by FA. The dissociation constant of the substrate for R100A(Mg²⁺)–inhibitor **1** was 436 ± 74 nM threefold higher than for R100A(Ca²⁺) (c.f. 135 ± 22 nM). The FRET$_{eff}$-values at saturation were similar for R100A(Ca²⁺) and R100A(Mg²⁺)–inhibitor **1** complexes suggesting that the substrate is bound in the same conformation irrespective of the inhibitor bound in the active site (data not shown). Although an increase in FRET efficiency was detected upon titration of R100A(Mg²⁺)–inhibitor **2** into DF-4, the FRET$_{eff}$-value had not stabilised by the end of the titration. Suggesting that even at the highest protein concentration (3.5 μM) not all of the substrate in solution is bound, and bent. Overall these results reiterate that although in the presence of inhibitor **2** a quaternary complex is formed the stability of the complex is much less than the equivalent complex in the presence of inhibitor **1**.
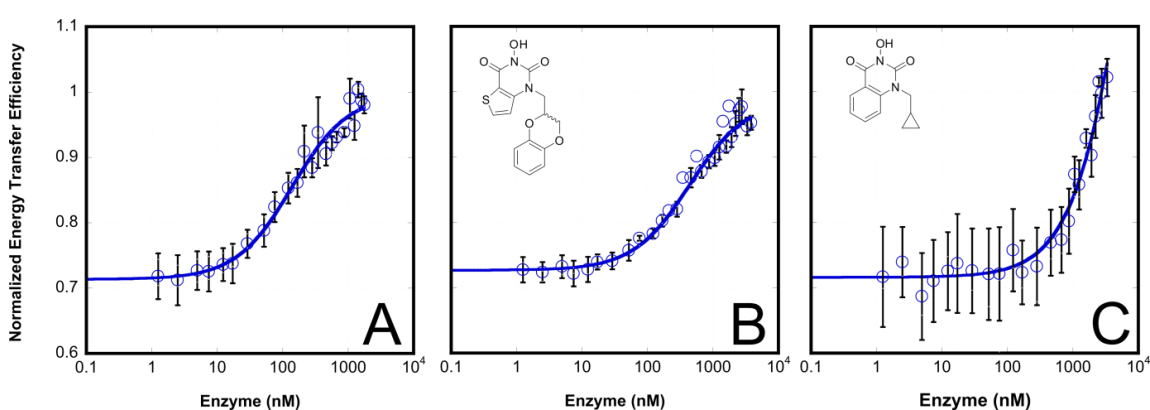
Figure 6.26: DNA is bent in both R100A–Ca²⁺–DNA and R100A–Mg–I–DNA complexes. *hFEN1R100A binding to 10 nM donor and acceptor-labeled double-flap substrate (DF-4) causes a change in FRET efficiency indicating a decrease in the end-to-end distance of the double-labeled DNA consistent with kinking of double-flap DNA in buffer containing (**A**) 10 mM Ca²⁺: $K_{D-bending}$ = 135 ± 22 nM, (**B**) 8 mM Mg²⁺ and 100 μM inhibitor **1**: $K_{D-bending}$ = 436 ± 74 nM and (**C**) 8 mM Mg²⁺ and 100 μM inhibitor **2**: $K_{D-bending}$ < 1000 nM. Apart from the specified conditions, assays were performed in 110 mM KCl, 55 mM HEPESpH=7.5, 1 mM DTT and at 37 ⁰C.*

6.4.5 Active Site Metal-Binding N-Hydroxyurea Inhibitors of hFEN1 Block DNA Unpairing

As established in *chapter 4*, an exciton-coupling signal can be used to monitor the substrate conformational changes brought about by FEN1-Ca^{2+}. As the inhibitors interact with active-site Mg^{2+} ions, it is likely that they could prevent this DNA conformational change. Therefore the metal-dependent unpairing of FEN1 substrate DNA (*section 4.5.1*) was investigated in the presence of inhibitor **1** and **2**. R100A was used in combination with the S$_{-1-2}$ construct (*figure 4.5*) as this displayed a large metal-dependent substrate unpairing signal (*section 4.6.1*); duplicated in *figure 6.27* for convenience. Experiments were conducted at 10 μM substrate and 12.5 μM protein, under these conditions the substrate is proposed to be fully bound to both the R100A(Mg^{2+})–inhibitor **1** and R100A(Mg^{2+})–inhibitor **2** complex. However, no decrease in the CD intensity was observed when the substrate was bound to R100A(Mg^{2+}) in the presence of either inhibitor **1** or **2** (*figure 6.27 b and c, respectively*). This suggested that when inhibitor **1** or **2** is bound in the active site the -1 nucleotide, which is reported to undergo a large conformational change in the FEN1 mechanism remains well stacked in dsDNA and strongly coupled with the -2 nt. Therefore, the data implies that because the flanking nts must undergo a conformational selection step prior to hydrolysis these inhibitors prevent reaction by preventing association of the catalytic metal ions with the scissile phosphate.

The HTV traces (*Appendices, figure A9*), which report on the amount of light passing through the sample at each wavelength displayed an slight signal variation for both inhibitors, but this deviation is accounted for in the correction process (*section 2.7.1*), and is not suspected to impose on the upairing signal at 326 nm. Removing the metal ions from the active site, removes the ability of the inhibitor to bind but also the substrate to unpair. Therefore, experiments in EDTA can be seen as a negative control. The ECCD traces are very similar for R100A(EDTA) in the presence and absence of inhibitors (*figure 6.27*).
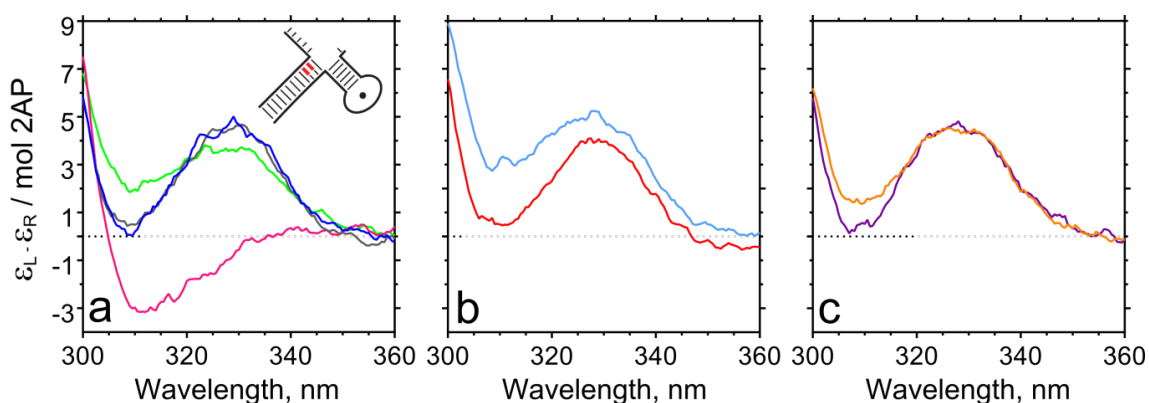
Figure 6.27: Inhibitor prevents reaction by preventing unpairing. *Low-energy CD spectra of hFEN1R100A of (a) Unbound S_{-1-2} in the presence of 10 mM Ca^{2+} (blue) and 25 mM EDTA (grey), S_{-1-2} bound to hFEN1R100A(Ca^{2+}) (magenta) and S_{-1-2} bound to hFEN1R100A(EDTA) (green). (b) S_{-1-2} bound to hFEN1R100A(Mg^{2+})+100 μM inhibitor **1** (cyan), S_{-1-2} bound to hFEN1R100A(EDTA)+100 μM inhibitor **1** (red) (c) S_{-1-2} bound to hFEN1R100A(Mg^{2+})+100 μM inhibitor **2** (orange) and S_{-1-2} bound to hFEN1R100A(EDTA)+100 μM inhibitor **2** (purple). Samples were buffered at pH =7.5, and measurements performed at 20 ˚C; specific conditions give in section 2.7.1. Each measurement was independently repeated and gave equivalent results, spectra were normalized as describe in experimental section 2.7.1. Substrate construct is illustrated schematically (red 2AP) with sequence shown in figure 4.5. The associated HTV traces are shown in Appendices, figure A9.*

6.5 Discussion:

The data presented here demonstrates that the related *N*-hydroxy urea inhibitors **1,2** and **3** bind to the active-site Mg^{2+} ions of FEN1 with nanomolar affinity. Furthermore, functionalization of the 'war-head' nitrogen increases the affinity for the free enzyme, and is important for the interaction with bound DNA. A mixed inhibition model described the kinetic data best for inhibitor **1** and **2**. In agreement with the ITC data the kinetic data suggested that inhibitors **1** and **2** bind to the free enzyme with similar affinity. The term mixed inhibition involves both competitive and non-competitive elements. Based on statistical analysis inhibitor **1** may be more mixed with non-competitive whereas inhibitor **2** might be mixed with competitive. Thus, the small cyclopropyl group of inhibitor **2** was unable to bind as tightly to the FEN:DNA complex as the larger aromatic group of inhibitor **1**.

Normally, a mixed inhibition model also assumes that binding of the inhibitor disrupts binding of the substrate. Binding assays under conditions that prevented hydrolysis of the substrate validated the formation of a quaternary complex, E(Mg^{2+})–I–DNA. The data implies that there was no obvious steric hindrance between larger inhibitor side chains and the substrate. When inhibitor **3** that lacks a side chain was bound to FEN1(Mg^{2+}) the affinity of the substrate for the

resultant complex was identical to that of the equivalent inhibitor **1** complex that has a large sidechain.. However, inhibitor **2**, bound in the active site dramatically reduced the affinity of DNA for hFEN1, despite its smaller size with the small cyclopropyl sidechain. Thus the larger aromatic group is tolerated when DNA is bound but it is more difficult to accommodate the smaller hydrophobic cyclopropyl sidechain. If it is assumed that the position of the side chains are equivalent in inhibitor **1**- and **2**-DNA complexes with DNA then this suggests that the benzyl group is capable of interacting with the DNA. Intercalation or grove binding of the benzyl group with the bound DNA would provide an explanation for the increased affinity of the substrate for the enzyme inhibitor complex of **1** compared to **2**. These possible binding modes merit further investigation but are encouraging to observe because disruption of whole DNA repair pathways is a likely consequence if the inhibitors prevent reaction but not DNA binding capabilities.

The ability of FEN1 to bind and recognize bent DF structures has been reported to be important for the FEN1 mechanism. The FRET measurements imply that binding of the inhibitors does not alter the conformation of the substrate when bound to FEN1, and it is the proposed selective unpairing step which is prevented by the steric blocking effects of the inhibitors. Nucleotides flanking the scissile phosphate cannot move into the required position for catalysis to take place because of the association between the inhibitor and active-site magnesium ions.

Transfer of DNA bound at an initial binding site to the active site through an affirming step, whether this is unpairing or fraying, is also suspected to be an important mechanistic step for λ exonuclease, Mus81 and Mre11 nucleases (*section 1.4*) acting on different nucleic acid structures. Therefore, the unpairing process can potentially be exploited as a new strategy to generate inhibitors of DNA replication and repair enzymes. However, in any general strategy there is a challenge of specificity. Potentially any nuclease could be inhibited by *N*-hydroxy urea inhibitors if they rely on a two metal ion mechanism for catalysis. However, selectivity has been reported for hFEN1 over XPG (Tumey et al., 2005). In light of this surprising result the ability of *N*-hydroxy urea compounds to inhibit T5FEN and XRN1, two closely related family members were tested.

6.6 Closely related family members are not inhibited by *N*-hydroxy Urea compounds

6.6.1 T5FEN inhibition studies

T5FEN is an early evolution bacteriophage flap endonuclease that plays a similar role to hFEN1 processing Okazaki fragments during replication (Mitsunobu et al., 2014). T5FEN reactions were studied using a unimolecular pseudo Y (pY) substrate in the presence and absence of inhibitor

1. It has been documented that the preferred substrates of bacteriophage FENs *in vitro* is a pY structure (Williams et al., 2007a), and so this structure was used for assay of T5 activity (*figure 6.28*). Reactions were performed under multiple turnover conditions, with the substrate concentration set at the previously determined K_M value of T5FEN with this pY-1 substrate (Patel 2012). In order to saturate the enzyme or any resulting ES complex, the probe compound concentration was set at 100 μM (100000000 fold excess, inhibitor **1** to free T5FEN).

The cleavage product was observed using denaturing HPLC albeit with a different gradient to that used to separate DF-5* substrate and product structures (*figure 6.28*) (*section 2.13*). Due to the lack of a 3'-flap binding pocket, T5FEN demonstrates less reaction site specificity. Although the major cleavage site is still one nucleotide into the dsDNA of the flap strand additional products are also observed (*figure 6.28*). The normalized initial rate for endonucleoytic phosphate diester hydrolysis was calculated analogously to those of hFEN1. It came as somewhat of a surprise when the normalized initial rates were compared and it became apparent that inhibitor **1** did not significantly perturb the reaction rate of T5FEN even at high concentration (*table 6.5*).
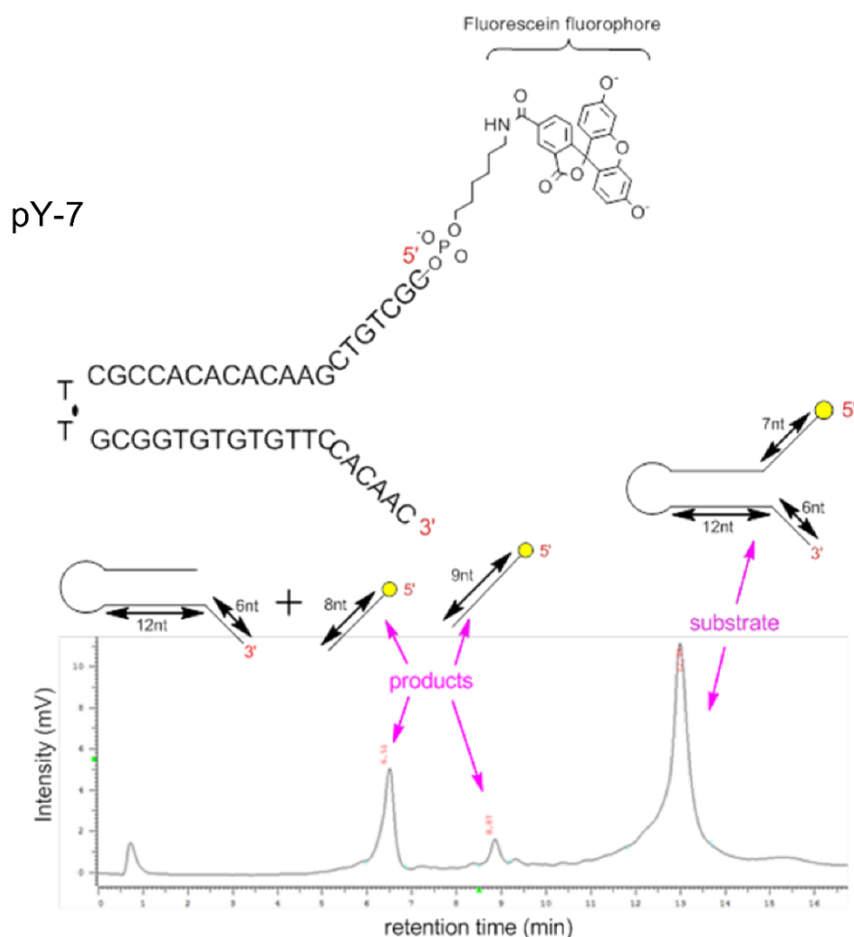
Figure 6.28: Structure of the pseudo-Y (pY) oligonucleotide substrate pY-7, unimolecular substrate with a six nucleotide overhang region (upstream region), twelve nucleotide downstream and seven nucleotide 5'-flap, possessing a 5'-FAM. Separation of products by dHPLC, indicating a major eight nucleotide product, minor nine nucleotide product and thirty-nine nucleotide substrate, quantitative detection by fluorescently labelled substrate, with the yellow ball representing fluorescein.

The T5FEN pH-rate profile shows that the activity of T5FEN upon pY structures is optimal at pH=9.3 and therefore this pH was used in the first experiment. Therefore, in order to rule out the possibility that the ionisation state of the war-head affects association with the active site metals, the reaction was repeated under the same conditions except now at pH=7.5, the same pH that the hFEN1Δ336 inhibition assays were performed at. The hydroxyl functional group of the war-head has an estimated pK_a~5, due to stabilization through delocalization into the empty nitrogen sp2 orbital, and as such is expected to be largely deprotonated at pH=7.5 and moving to pH=9.3 would only increase the proportion of ionized inhibitor **1**. Therefore, it did not come as a surprise when results demonstrated the inhibitor **1** was still unable to inhibit the T5FEN catalysed reaction at pH=9.3 (*table 6.5*). It is also worth mentioning that inhibitor **1** was just as

effective at inhibiting the hFEN1336 catalysed reaction of DF-5* at pH=9.3 as it was at pH=7.5 (*table 6.5*).

Table 6.5: Normalized initial rates for 35 nM pY-7 and 150 nM DF-5* cleavage by T5FEN and hFEN1Δ336, respectively. Reactions were carried out in appropriate buffers, *section 2.13.1*. Errors represent standard error. T5FEN used in these experiments was purified by Nikesh Patel.

Enzyme	Inhibitor 1 (μM)	pH	$v_0/[E]_0$ (min^{-1})
T5FEN	0	9.3	272 ± 6
	100	9.3	240 ± 8
	0	7.5	112
	100	7.5	97 ± 3
hFEN1Δ336	0	9.3	87 ± 4
	100	9.3	1.6 ± 0.7
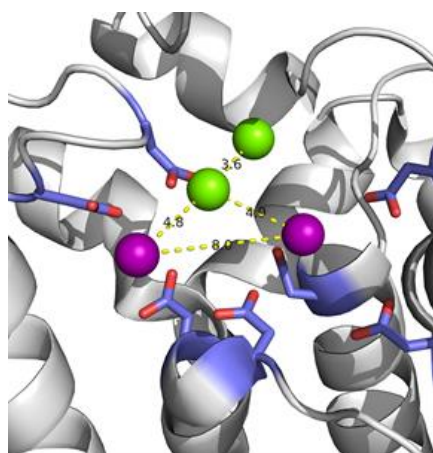
Figure 6.29: The active site of T5FEN superimposed with the structure of hFEN1 to show the relative position of crystalographically observed divalent metal ions. Seven acidic residues of T5FEN (1UT5) shown as sticks (carbon-blue, oxygen-red) coordinate at least two Mg^{2+} (purple), divalent metal ions (green) of hFEN1 (1UL1).

The active sites of T5FEN and hFEN1 are very similar. Each have seven positionally conserved carboxylates, superfamily conserved lysine and arginine residues and a stacking partner for unpaired DNAs (Tyr in hFEN1, Phe in T5). Moreover, the overall protein folds are very similar. Although several structural features are conserved between hFEN1 and T5FEN, the position and number of active site divalent metal ions observed in substrate-free structures is not (*figure*

215

6.29) (Grasby et al., 2012). Crystallographic data shows only two metal ions bound in the T5FEN active site, but these are at a distance greater than the 4 Å required to perform 2-metal ion catalysis. However, recent biochemical data suggest the participation of at least three metal ions for T5FEN catalysed reaction (Syson et al., 2008). It was therefore posited, that a third Mg^{2+} binds after the DNA has bound, at a site close enough to either of the crystalographically observed metal ions to participate in the two metal ion mediated hydrolysis of the phosphate diester backbone of DNA. This idea of 'positive cooperativity' whereby binding of the highly negatively charged DNA causes additional metals to bind that then participate in catalysis is not unheard of and there is evidence for other unrelated enzymes (Lai et al., 2003). It is possible that the lack of inhibition observed with the T5 enzyme arises because unlike human FEN1 the substrate-free enzyme does not have two close divalent metal ions within the active site.

6.6.2 XRN1 inhibition studies

Recently XRN1 has been identified as a potential member of the FEN-superfamily (Grasby et al., 2012). The processive exoribonuclease activity of XRN1 is involved in the degradation of messenger RNA (mRNA) in the cytoplasm (Jones et al., 2012). XRN1 is triggered into action by the presence of a 5'-monophosphate, generated by 5'-7-methylguanosine (m^7GpppN) cap removal (Coller and Parker, 2004) and endonucleolytic mRNA cleavage (Gatfield and Izaurralde, 2004).

The full-length *Kluyveromyces lactis* XRN (1-1453) shows weak sequence conservation with human XRN1 and *S. cerevisiae* XRN up to residue 1180 and 1240, respectively (Chang et al., 2011). This region includes all the superfamily conserved active site features of the FEN superfamily. Removal of the unique C-terminal tail (1245-1453) does not affect phosphodiesterase activity and facilitated crystallisation (Chang et al., 2011). Sequencing confirmed the subcloning of residues 1-1245 of *Kluyveromyces lactis* XRN1 into a pET26b plasmid background suitable for expression in a BL21 (DE3) *E.coli* system (*figure Appendices, figure A10*). The recombination protein was purified using a very similar procedure to that employed to isolate hFEN1 (*section 2.14*). This construct had a molecular weight of approximately 145 kDa, and purity was confirmed by SDS-page (*figure 6.30*).
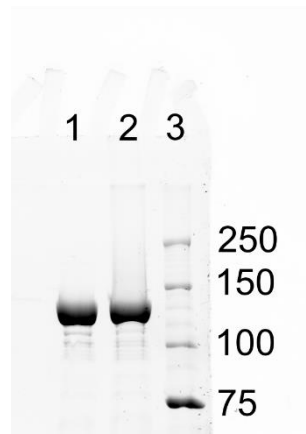
Figure 6.30: XRN1 purification results. *Lane 1: 2 µg/µl XRN1, lane 2: 4 µg/µl XRN1, and lane 3: protein marker, with molecular weight bands shown in kDa.*

In accord with XRN1s *in vivo* function, the rate of exonuclease activity *in vitro* is dependent on the nature of the 5'-end of RNA (Sinturel et al., 2009). The rate of cleavage is significantly reduced with 5'-triphosphate or an m^{7}GpppN cap compared to a 5'-monophosphate ssRNA substrate (Poole and Stevens, 1997). As XRN1 had not been previously studied in Sheffield, experiments began with a biochemical characterisation of substrate preference. Reactions of both 5'-phosphorylated and non-phosphorylated ssRNA went to completion, (50nM E, 500nM S / multiple turnover) but the time to complete reactions was markedly reduced for the substrate bearing a 5'-monophosphate (*figure 6.31*). The 5'-monophosphate and non-phosphorylated ssRNA strands are dennoted, RNA1 and RNA1-OH, respectively (*figure 6.31*). In agreement with previous studies, the activity appeared to be processive, as intermediates could not be detected in the reaction (*figure 6.31*). In an attempt to estimate the size of the final product, a ladder was produced by partial alkaline hydrolysis of the RNA strand (*figure 6.31a, lane L*). Co-electrophoresed XRN1 reaction and alkaline hydrolysis samples suggested that XRN1 hydrolyses the complete strand and potentially even the phosphate between the 3'-FAM and final nucleotide.

XRN1 can cleave 5'-flapped duplex RNA structures, although it is not an endonuclease. *In vitro* exoribonuclease activity on RNA:DNA hybrids has been confirmed (Sinturel et al., 2009), and it has been shown that increasing 5'-flap length promotes activity (Jinek et al., 2011). Here, using a RNA:DNA hybrid with a 13 nt 5'-overhang (*RNA1:DNA1, figure 6.32*), *K.lactus* XRN1 was shown to have processive activity (*figure 6.32*). As seen in ssRNA no intermediates were detected suggesting XRN1 possess solely exonuclease activity. In addition, no products were detected as a consequence of DNA strand digestion (*figure 6.32*), presumably aided by the 5'-fluorphore

preventing access to the active site. The RNA strands in duplexes containing a 3'-overhang or a blunt-end were not as efficiently degraded as in the duplex containing the 5'-overhang (*figure 6.33*). To observe reaction on these non-optimal substrates, the starting enzyme and substrate concentrations were equimolar. The reaction was exclusively on the RNA strand in both these constructs (data not shown). The mechanism of XRN1 has been proposed to be an extension of the ratchet mechanism of λ exonuclease (*section 1.4.4*) and involves duplex melting (unpairing) at the ss-ds junction which is facilitated by the presence of a 5'-flap.

A RNA1-OH
5' HO-ACUCACUCACUCACCAAAAAAAAAAAAAAACC-FAM 3'

B RNA1
5' p-ACUCACUCACUCACCAAAAAAAAAAAAAAACC-FAM 3'
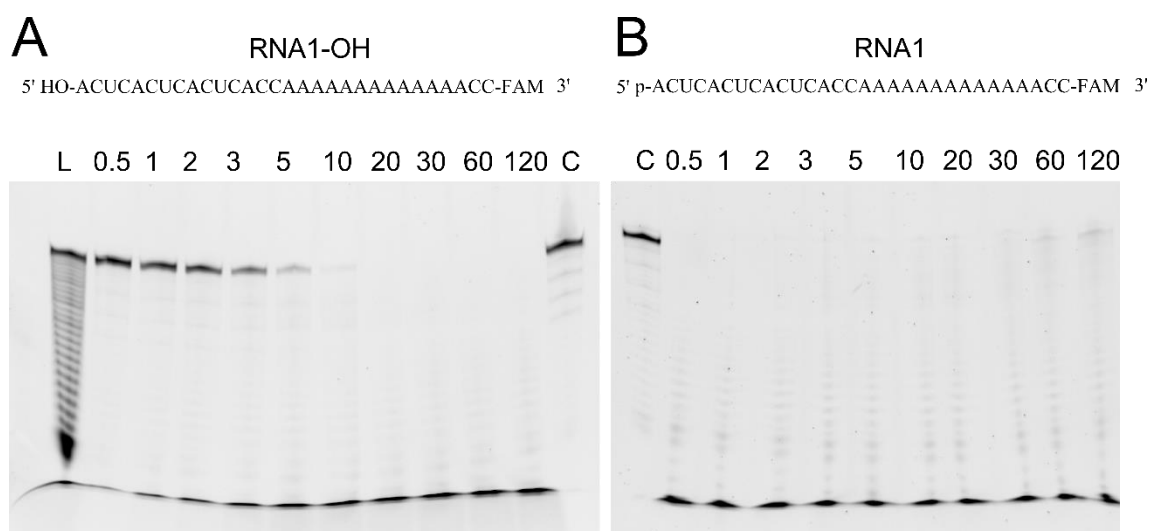
Figure 6.31: The activity of K.lactus XRN1 on ssRNA is affected by the nature of the 5'-terminus.
XRN1 was incubated with ssRNA bearing either (A) 5'-hydroxyl or (B) 5'-monophosphate, the oligonucleotides had the same sequence as shown above each gel. Reactions were performed at 37 °C with 500 nM substrate and 50 nM enzyme in appropriate buffers, see section 2.16. A time-course of the reaction is shown with the numbers above each lane representing minutes after initiation with XRN1. Before reaction was initialled an aliquot of the reaction mixture was removed as a control and placed under the same conditions, shown in the lane denoted C. The products of partial alkaline hydrolysis of RNA1 is shown in the lane denoted L. Reactions were analysed on a 20% denaturing gel and RNA1 products imaged by detecting FAM.

RNA1   5' p-ACUCACUCACUCACCAAAAAAAAAAAAACC-FAM 3'
DNA1   3' HO-GGTTTTTTTTTTTTTGG-TAMRA 5'
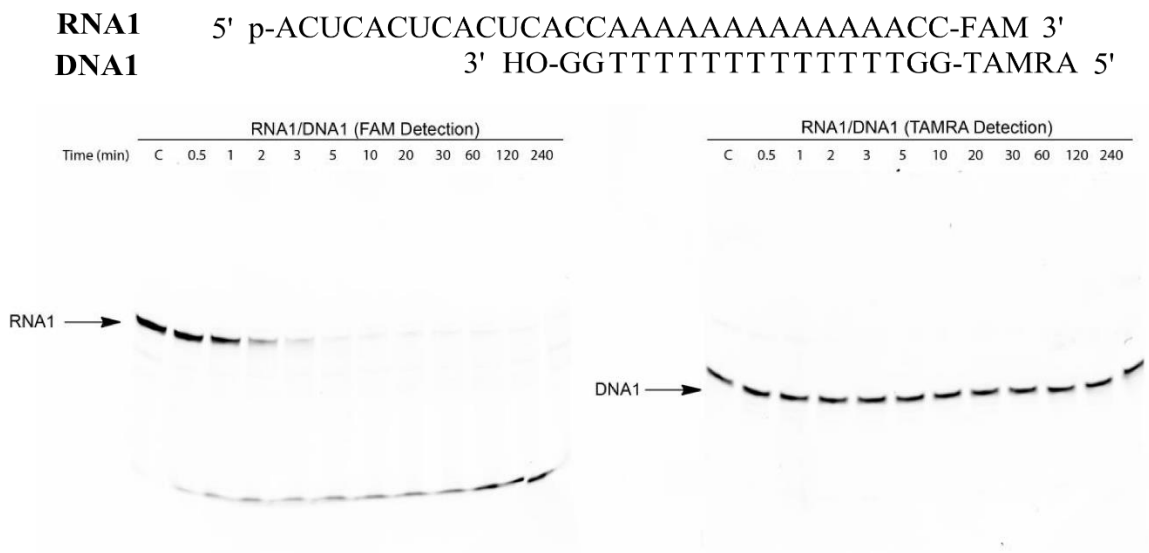
Figure 6.32: The activity of XRN1 on a RNA:DNA duplex with a 5'-overhang. *The sequences of the RNA and DNA strands forming the duplex are shown above the gel, the RNA strand has a 5'-monophosphate. Reactions were performed at 37 °C with 500 nM substrate and 50 nM enzyme in appropriate buffers, see section 2.16. A time-course of the reaction is shown with the numbers above each lane representing minutes after initiation with XRN1. Before reaction was initialled an aliquot of the reaction mixture was removed as a control and placed under the same conditions, shown in the lane denoted C. Reactions were analysed on a 20% denaturing gel, RNA1 products (left) and DNA1 (right) products imaged by detecting FAM and TAMRA, respectively.*
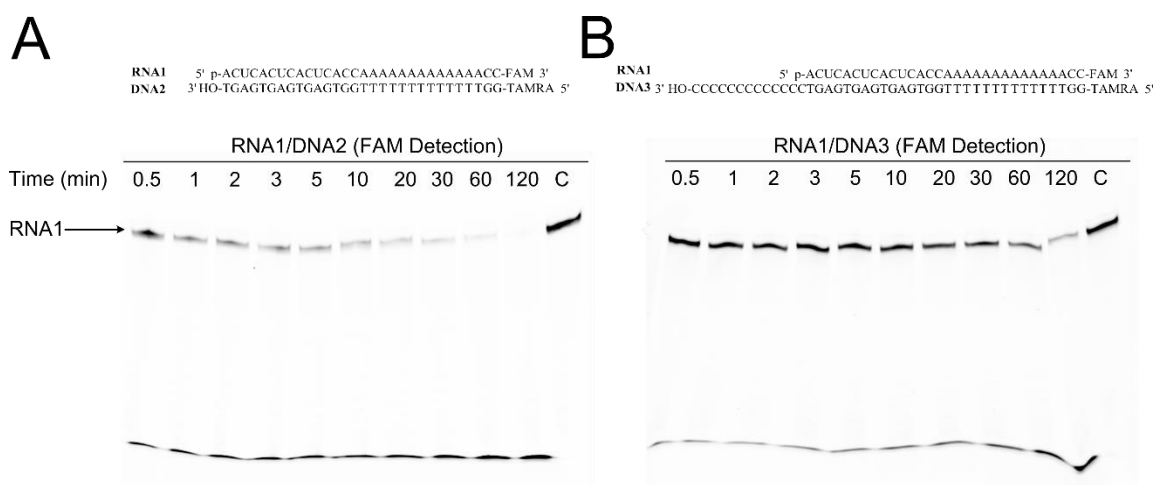
Figure 6.33: The activity of XRN1 on a RNA:DNA duplex with (A) a blunt-end, and (B) a 3'-overhang. *The sequences of the RNA and DNA strands forming the duplex are shown above the gel, the RNA strand has a 5'-monophosphate. Reactions were performed at 37 °C with 500 nM substrate and 500 nM enzyme in appropriate buffers, see section 2.16. A time-course of the reaction is shown with the numbers above each lane representing minutes after initiation with XRN1. Before reaction was initialled an aliquot of the reaction mixture was removed as a control and placed under the same conditions, shown in the lane denoted C. Reactions were analysed on a 20% denaturing gel, RNA1 and DNA2/3 products imaged by detecting FAM and TAMRA, respectively.*

To examine the ability of inhibitor **1** to prevent hydrolysis by XRN1, the activity of XRN1 toward the RNA:DNA duplex with a 5'-overhang was assayed in the presence of a large excess of inhibitor **1 (**100 μM). This substrate was chosen because substrate depletion and product accumulation was very clear over a reasonable timescale. The result showed that there was no decrease in exoribonuclease activity in the presence of inhibitor **1** (*figure 6.34*). This was surprising as XRN1 and hFEN1 possess similar active-site architecture. The binding of inhibitor **1** to nucleic acid free XRN1 and T5FEN was not investigated in this study. Therefore, the ability of *N*-hydroxy urea compounds to bind to the free enzymes needs to be determined to further elucidate the reasons these enzymes are not inhibited by compound **1**. Interestingly, very recent experiments preformed by Steven Shaw and David Finger show that another superfamily member human exonuclease-1 is only inhibited by compound **1** at very high concentrations **(**10-100 μM) suggesting that achieving specificity with active site inhibitors is a realistic possibility.
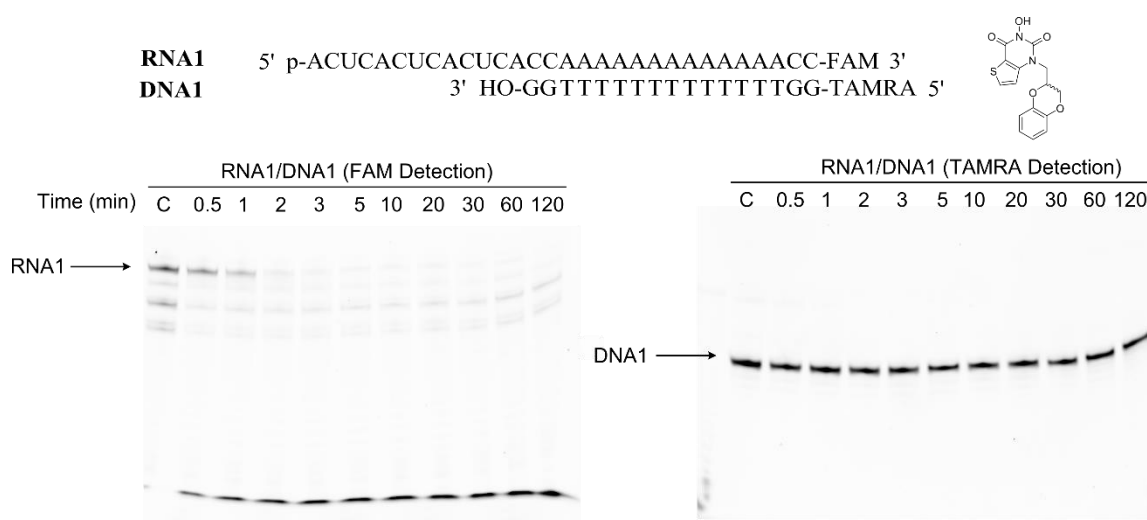
Figure 6.34: XRN1 activity on a RNA:DNA duplex with a 5'-overhang is not inhibited by N-hydroxy urea compound. *The sequences of the RNA and DNA strands forming the duplex are shown above the gel, the RNA strand has a 5'-monophosphate. Reactions were performed at 37 °C with 500 nM substrate, 50 nM enzyme, and 100 μM inhibitor **1** in appropriate buffers, see section 2.16. Pre-incubation of XRN1 and inhibitor **1** was carried out on ice. A time-course of the reaction is shown with the numbers above each lane representing minutes after initiation with XRN1. Before reaction was initialled an aliquot of the reaction mixture was removed as a control and placed under the same conditions, shown in the lane denoted C. Reactions were analysed on a 20% denaturing gel, RNA1 and DNA1 products imaged by detecting FAM and TAMRA, respectively.*

6.7 Summary and Conclusions

In this chapter the proposed metal-dependent binding of *N*-hydroxy urea compounds to human FEN1 was confirmed by Isothermal titration calorimetry. These experiments revealed that the common 'war-head' of inhibitor **1**, **2,** and **3** was only capable of coordinating to magnesium ions bound in the active-site as calcium ions did not support binding, presumably dues to changes in geometry. Kinetic assays established that inhibitor **1** and **2** could bind to both the free enzyme and enzyme substrate complex. Having established that the mode of inhibition was mixed, binding equilibrium assays independently confirmed the formation of an enzyme(Mg^{2+}):DNA:inhibitor complex. Significantly, DNA could be competed away from the enzyme:inhibitor complex. Taken together this data suggested that these inhibitors bind at a site that allows the DNA to still associate with the protein surface. This proposal is supported by data suggesting that these inhibitors are effective at preventing the FEN1 reaction because they prevent the DNU mechanism, which transfers the DNA from an initial binding site to the active site. It was surprising to observe that inhibitor **1** did not perturb the catalysis by the closely related family members XRN1 and T5FEN. While this deserves further investigation, from a selectivity aspect it is promising that they cannot inhibit closely related family members because they are unlikely to have off target effects, a criticism levelled at active site metal-binding inhibitors.

Chapter 7: Conclusions

The work described by this thesis has used several different approaches to study the mechanism and inhibition of human flap endonuclease 1. Based predominantly on structural data a double nucleotide unpairing mechanism has been proposed, in which, the nucleotides flanking the scissile phosphate must change conformation during the reaction course to place the scissile phosphate in the catalytic radius of active-site divalent metal ions (Syson et al., 2008; Tsutakawa et al., 2011). With regards to the inhibition of hFEN1, independent studies have identified *N*-hydroxyurea compounds as small molecule inhibitors of hFEN1 but no work has been conducted to characterise their mode of inhibition (Abdel-Fatah et al., 2014; Tumey et al., 2005; van Pel et al., 2013).

Previous biochemical studies have identified residues proposed to be directly involved in hFEN1 mechanistic steps. The residues studied in this thesis, Tyr-40, Lys-93, Arg-100 were chosen because they were proposed to be important for the double nucleotide unpairing mechanism. In addition, the residue Asp-181 coordinates one of the catalytic magnesium ions, therefore the impact of perturbing metal ion sequestration on unpairing was also studied. Creation of alanine variants have been studied previous under various different enzymatic conditions and all mutations were found to be detrimental towards catalysis (Greene et al., 1999; Storici et al., 2002; Tsutakawa et al., 2011). However, the full impact of their removal upon catalytic activity was not known. Therefore, in chapter 3 these single-point mutations were re-evaluated under single-turnover conditions. The first order rate constant was reduced by ~100 − 10,000 fold compared to wild-type hFEN1. Specifically, the decrease in activity compared to wild-type hFEN1 varied in the order WT<Y40A < K93A < R100A < D181A. This demonstrated that these residues were important for catalysis, which was known, but the range of rates suggest that they play very different roles in the FEN mechanism. A double mutation was also introduced were by both basic residues were mutated to alanine. Surprisingly, the variant, K93AR100A was only as detrimental towards catalysis as the single R100A mutation. This suggests that the Lys-93 and Arg-100 act in a concerted manner in the same step of the mechanism.

In chapter 3 the decrease in activity due to a decreased affinity for the substrate was ruled out by measuring the dissociation constant for wild-type and the variant proteins. Various techniques were employed to measure the binding affinity. Subtle differences between the wild-type and variant proteins were observed using FA. However, the validity of these results was brought into question by complications associated with non-specific binding, a phenomenon confirmed by EMSA. Therefore, another florescence based equilibrium technique, FRET, was

used to measure binding. It has been previously established that a change in FRET is only observed upon FEN binding in the 'productive' mode (Chapados et al., 2004; Craggs et al., 2014; Sobhy et al., 2013). Therefore, this ruled out the prospect of non-specific binding affecting affinities. Results with the FRET assay also suggested that WT hFEN1 and the particular variants studied have different affinities for double-flap substrates. In addition, the affinity in the presence and absence of divalent metal ions has always thought to be very similar. The FRET assay confirmed this but also suggested that the substrate may not be in equivalent conformations. Based on these results FRET would be the most appropriate method to determine binding affinities of hFEN1 in the future.

The dynamics at the various positions adjacent to the ss-ds junction of a double flap substrate were investigated by monitoring the exciton-coupling signal of two tandem 2-AP nucleotides (Johnson et al., 2004). In chapter 4, the data collected with unbound constructs agreed with previous studies of similar structures using this technique, indicating stacking of nucleotides upon duplex formation. Specific to double flap structures the nucleotides at the base of the flap appear destabilised by the presence of the flap. It was established using a product construct mimic that ECCD is an appropriate technique to monitor unpairing, as spectroscopic changes were dependent on active-site metal ions and a $5'$-monophosphate; results which can be rationalised by previous crystal structures. All of the variants studied were unable to reproduce the WThFEN1(Ca^{2+})-product signal suggesting that these residues are involved in the unpairing mechanism. The residues Tyr-40, Lys-93 and Arg-100 are believed to directly aid unpairing by contacting and therefore stabilising extra-helical nucleotides were as Asp-181 is likely indirectly affecting unpairng by weakening FEN's affinity for active-site divalent metal ions. Changes in the unpairing equilibrium positions of the substrate construct were also dependent on the presence of active-site divalent metal ions. However, nucleotide conformational changes of a double flap *substrate* could be measured when bound to variants although the exact position of the nucleotides in R100, K93AR100A and Y40-complexes differed compared to WT. Unexpectedly, K93A produced a very similar exciton-coupling signal to that of WThFEN1. However, the double mutant K93AR100A was more similar to R100A. Coupled with the knowledge that the single turnover rate of K93AR100A is as slow as R100A this suggests that the Arg-100 residues is involved in the mechanism before Lys-93 and Arg-100 is required to position the nucleotide allowing the involvement of Lys-93.

Movement of the (2-AP)$_2$ spectral probe within the double flap substrate allowed the nature of the conformational to be investigated. Preliminary results indicate that the equilibrium is shifted to a position were both nucleotides undergo a conformational change simultaneously. The time

resolution of the ECCD is unable to decipher the precise order of unpairing, that is to say whether it is a concentred or multistep process, therefore this merits further work. In addition, replacement of the nucleotide on the reacting strand at the immediate base of the 5ʹ-flap creating a C:C mismatch resulted in a change in equilibrium position of the adjacent nucleotide when in complex with WThFEN1(Ca^{2+}) compared to a completely Watson-Crick base paired substrate.

In chapter 5, the role of the helical arch in the unpairing mechanism was explored. Experiments to trap and block access of the 5ʹ-flap through the helical arch imply that the 5ʹ-flap must be threaded for unpairing to occur. This was the first direct evidence for involvement of the helical arch in the unpairing mechanism. It was previously hypothesised that an ordered helical arch promoted unpairing (Tsutakawa et al., 2011). However, disruption of the α-helical arch through insertion of the proline residues contradicted this proposal. It was discovered that proper ordering of the top of the helical arch (helical cap) is required for unpairing but not the base of the helical arch. Therefore, the initial hypothesis must be re-evaluated and allow for these results. An ordered helical cap might be required to push down on the 5ʹ-flap to position the substrate for proper involvement of other hFEN1 residues.

An *N*-hydroxyurea based series of inhibitors showing variable selectivity for FEN1 over XPG has previously been reported (Tumey et al., 2005). Their FEN1 inhibitory activity was thought to derive from a metal-chelating 'warhead', preventing association between the scissile phosphate and active-site metal ions. However, in the initial investigation no direct evidence was reported. In chapter 6, isothermal titration calorimetry confirmed active-site magnesium ion dependent nanomolar affinity of these inhibitors. Additionally, ITC showed that calcium ions within the active site were unable to support inhibitor binding. Together these results indicate that *N*-hydroxyurea compounds are sensitive to the active-site metal ions probably due to a change in ionic radius and/or coordination geometry. In chapter 6, a kinetic approach was used to determine the inhibition model of *N*-hydroxyurea compounds upon hFEN1 catalysed phosphate diester hydrolysis. Statistical discrimination favoured a mixed inhibition model for both inhibitor **1** and inhibitor **2**. However, and counterintuitively, inhibitor **2** with the smaller cyclopropyl group compared to inhibitor **1,** which bore a large aromatic side chain, had a much greater competitive element to its mixed inhibition. From this it can be concluded that the moiety in ring system B seems to be important for the interaction with bound DNA. Preferentially a FEN1 inhibitor would interact favourably with any DNA substrate as this would mean it could potentially inhibit multiple DNA repair pathways if hFEN1 inhibition occurred when bound to DNA. The formation of an enzyme-substrate-inhibitor (ESI) complex was independently confirmed by fluorescence

anisotropy and FRET experiments. The same pattern seen in the kinetic analyses was also observed here, inhibitor **2** weakened the affinity of FEN1 for DNA by approximately 3-fold compared to inhibitor **1**, but there was still evidence for an enzyme-substrate-inhibitor **2** complex. Exciton-coupling with the double flap substrate which produced the largest signal change upon protein binding was used in conjunction with the inhibitor **1** and **2** to assess their effect on unpairing. These set of experiments revealed that these inhibitors prevent reaction by sterically blocking the DNU step. FRET experiments suggest that the bound DNA is capable of bending in the presence of the inhibitor. DNA unpairing is dependent on a threaded substrate, and it is more than likely that bending is also a prerequisite for unpairing, therefore in the future it would be interesting to investigate if the inhibitor blocked unpairing by preventing threading. Finally, it was intriguing to observe that these potent FEN1 inhibitors were unable to prevent reaction of closely related family members T5FEN and XRN1 at saturating concentrations. *N*-hydroxyurea compounds did however inhibit EXO1, the most structurally similar homologue, but not to the same extent. Appreciation of differing inhibition levels may aid in the understanding of the FEN1 mechanism and differences in processivity between family members.

References

Ababou, A., and Ladbury, J.E. (2006). Survey of the year 2004: literature on applications of isothermal titration calorimetry. Journal of Molecular Recognition *19*, 79-89.

Abdel-Fatah, T., Arora, A., Gorguc, I., Abbotts, R., Beebeejaun, S., Storr, S., Mohan, V., Hawkes, C., Soomro, I., Lobo, D.N.*, et al.* (2013). Are DNA Repair Factors Promising Biomarkers for Personalized Therapy in Gastric Cancer? Antioxidants & Redox Signaling *18*, 2392-2398.

Abdel-Fatah, T.M.A., Russell, R., Albarakati, N., Maloney, D.J., Dorjsuren, D., Rueda, O.M., Moseley, P., Mohand, V., Sun, H., Abbotts, R.*, et al.* (2014). Genomic and protein expression analysis reveals flap endonuclease 1 (FEN1) as a key biomarker in breast and ovarian cancer. Molecular Oncology *8*, 1326-1338.

Aboussekhra, A., Biggerstaff, M., Shivji, M.K.K., Vilpo, J.A., Moncollin, V., Podust, V.N., Protic, M., Hubscher, U., Egly, J.M., and Wood, R.D. (1995). Mammalian DNA nucleotide excision-repair reconstituted with purified protein-components. Cell *80*, 859-868.

Alian, A., Griner, S.L., Chiang, V., Tsiang, M., Jones, G., Birkus, G., Geleziunas, R., Leavitt, A.D., and Stroud, R.M. (2009). Catalytically-active complex of HIV-1 integrase with a viral DNA substrate binds anti-integrase drugs. Proceedings of the National Academy of Sciences of the United States of America *106*, 8192-8197.

Anderson, D.R., and Burnham, K.P. (2002). Avoiding pitfalls when using information-theoretic methods. Journal of Wildlife Management *66*, 912-918.

Anderson, S., and Depamphilis, M.L. (1979). Metabolism of Okazaki fragments during simian virus-40 DNA-replication. Journal of Biological Chemistry *254*, 1495-1504.

Anker, M., and Corales, R.B. (2008). Raltegravir (MK-0518): a novel integrase inhibitor for the treatment of HIV infection. Expert Opinion on Investigational Drugs *17*, 97-103.

Arnedos, M., Soria, J.-C., Andre, F., and Tursz, T. (2014). Personalized treatments of cancer patients: A reality in daily practice, a costly dream or a shared vision of the future from the oncology community? Cancer Treatment Reviews *40*, 1192-1198.

Asante-Appiah, E., and Skalka, A.M. (1999). HIV-1 integrase: Structural organization, conformational changes, and catalysis. Advances in Virus Research, Vol 52 *52*, 351-369.

Azumi, T., and McGlynn, S.P. (1962). Polarization of luminescence of phenanthrene. Journal of Chemical Physics *37*, 2413-&.

Bambara, R.A., Murante, R.S., and Henricksen, L.A. (1997). Enzymes and reactions at the eukaryotic DNA replication fork. Journal of Biological Chemistry *272*, 4647-4650.

Bapat, A., Glass, L.S., Luo, M., Fishel, M.L., Long, E.C., Georgiadis, M.M., and Kelley, M.R. (2010). Novel Small-Molecule Inhibitor of Apurinic/Apyrimidinic Endonuclease 1 Blocks Proliferation and Reduces Viability of Glioblastoma Cells. Journal of Pharmacology and Experimental Therapeutics *334*, 988-998.

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. Journal of Biological Chemistry *271*, 29624-29631.

Bastin-Shanower, S.A., Fricke, W.M., Mullen, J.R., and Brill, S.J. (2003). The mechanism of Mus81-Mms4 cleavage site selection distinguishes it from the homologous endonuclease Rad1-Rad10. Molecular and Cellular Biology *23*, 3487-3496.

Beddows, A., Patel, N., Finger, L.D., Atack, J.M., Williams, D.M., and Grasby, J.A. (2012). Interstrand disulfide crosslinking of DNA bases supports a double nucleotide unpairing mechanism for flap endonucleases. Chemical Communications *48*, 8895-8897.

Beernink, P.T., Segelke, B.W., Hadi, M.Z., Erzberger, J.P., Wilson, D.M., and Rupp, B. (2001). Two divalent metal ions in the active site of a new crystal form of human apurinic/apyrimidinic endonuclease, Ape1: Implications for the catalytic mechanism. Journal of Molecular Biology *307*, 1023-1034.

Beese, L.S., Derbyshire, V., and Steitz, T.A. (1993). Structure of DNA-Polymerase-I Klenow fragment bound to duplex DNA. Science *260*, 352-355.

Beese, L.S., and Steitz, T.A. (1991). Structural basis for the 3'-5' Exonuclease activity of Escherichia-coli DNA-Polymerase-I a 2 metal-ion mechanism. Embo Journal *10*, 25-33.

Bhagwat, M., Meara, D., and Nossal, N.G. (1997). Identification of residues of T4 RNase H required for catalysis and DNA binding. Journal of Biological Chemistry *272*, 28531-28538.

Bisswanger, H. (1999). Enzyme Kinetics. Enzyme Kinetics: Principles and Methods, 51-160.

Boiteux, S., and Guillet, M. (2004). Abasic sites in DNA: repair and biological consequences in Saccharomyces cerevisiae. DNA Repair *3*, 1-12.

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.

Brautigam, C.A., Sun, S., Piccirilli, J.A., and Steitz, T.A. (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*, 696-704.

.

Burgers, P.M.J. (2009). Polymerase Dynamics at the Eukaryotic DNA Replication Fork. Journal of Biological Chemistry *284*, 4041-4045.

Caldecott, K.W. (2003). XRCC1 and DNA strand break repair. DNA Repair *2*, 955-969.

Cantor, C.R., and Schimmel, P.R. Biophysical Chemistry, Part II: Techniques for the Study of Biological Structure and Function, 1980. WH Freeman, NY *539*, 191.

Carter, D.M., and Radding, C.M. (1971). Role of exonuclease and beta protein of phage lambda in genetic recombination .2. Substrate specificity and mode of action of lambda exonuclease. Journal of Biological Chemistry *246*, 2502-&.

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.

Chang, J.H., Xiang, S., Xiang, K., Manley, J.L., and Tong, L. (2011). Structural and biochemical studies of the 5 '-> 3 ' exoribonuclease Xrn1. Nature Structural & Molecular Biology *18*, 270-U249.

Chapados, B.R., Hosfield, D.J., Han, S., Qiu, J.Z., Yelent, B., Shen, B.H., 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.

Ciccia, A., McDonald, N., and West, S.C. (2008). Structural and functional relationships of the XPF/MUS81 family of proteins. Annual Review of Biochemistry *77*, 259-287.

Cisneros, G.A., Perera, L., Schaaper, R.M., Pedersen, L.C., London, R.E., Pedersen, L.G., and Darden, T.A. (2009). Reaction Mechanism of the epsilon Subunit of E-coli DNA Polymerase III: Insights into Active Site Metal Coordination and Catalytically Significant Residues. Journal of the American Chemical Society *131*, 1550-1556.

Clegg, R.M. (1992). Fluorescence resonance energy-transfer and nucleic-acids. Methods in Enzymology *211*, 353-388.

Clegg, R.M., Murchie, A.I.H., and Lilley, D.M.J. (1994). The solution structure of the 4-way DNA junction at low-salt conditions - a fluorescence resonance energy-transfer analysis. Biophysical Journal *66*, 99-109.

Clegg, R.M., Murchie, A.I.H., Zechel, A., Carlberg, C., Diekmann, S., and Lilley, D.M.J. (1992). Fluorescence resonance energy transfer analysis of the structure of the four-way DNA junction. Biochemistry *31*, 4846-4856.

Coller, J., and Parker, R. (2004). Eukaryotic mRNA decapping. Annual Review of Biochemistry *73*, 861-890.

Connolly, B.A., Eckstein, F., and Pingoud, A. (1984). The stereochemical course of the restriction endonuclease ECORI-catalyzed reaction. Journal of Biological Chemistry *259*, 760-763.

Cornish-Bowden, A., (1974). Simple graphical method for determining inhibition constants of mixed, uncompetitive and non-competitive inhibitors. Biochemical Journal *137*, 143-144.

Craggs, T.D., Hutton, R.D., Brenlla, A., White, M.F., and Penedo, J.C. (2014). Single-molecule characterization of Fen1 and Fen1/PCNA complexes acting on flap substrates. Nucleic Acids Research *42*, 1857-1872.

Dale, R.E., Eisinger, J., and Blumberg, W.E. (1979). The orientational freedom of molecular probes. The orientation factor in intramolecular energy transfer. Biophysical journal *26*, 161-193.

Datta, K., Johnson, N.P., LiCata, V.J., and von Hippel, P.H. (2009). Local Conformations and Competitive Binding Affinities of Single- and Double-stranded Primer-Template DNA at the Polymerization and Editing Active Sites of DNA Polymerases. Journal of Biological Chemistry *284*, 17180-17193.

Datta, K., Johnson, N.P., and von Hippel, P.H. (2006). Mapping the conformation of the nucleic acid framework of the T7 RNA polymerase elongation complex in solution using low-energy CD and fluorescence spectroscopy. Journal of Molecular Biology *360*, 800-813.

Datta, K., Johnson, N.P., and von Hippel, P.H. (2010). DNA conformational changes at the primer-template junction regulate the fidelity of replication by DNA polymerase. Proceedings of the National Academy of Sciences of the United States of America *107*, 17980-17985.

Datta, K., and von Hippel, P.H. (2008). Direct spectroscopic study of reconstituted transcription complexes reveals that intrinsic termination is driven primarily by thermodynamic destabilization of the nucleic acid framework. Journal of Biological Chemistry *283*, 3537-3549.

de Laat, W.L., Appeldoorn, E., Jaspers, N.G.J., and Hoeijmakers, J.H.J. (1998). DNA structural elements required for ERCC1-XPF endonuclease activity. Journal of Biological Chemistry *273*, 7835-7842.

Debrauwere, H., Loeillet, S., Lin, W., Lopes, J., and Nicolas, A. (2001). Links between replication and recombination in Saccharomyces cerevisiae: A hypersensitive requirement for homologous recombination in the absence of Rad27 activity. Proceedings of the National Academy of Sciences of the United States of America *98*, 8263-8269.

Delelis, O., Carayon, K., Saib, A., Deprez, E., and Mouscadet, J.-F. (2008). Integrase and integration: biochemical activities of HIV-1 integrase. Retrovirology *5*.

Derbyshire, V., Freemont, P.S., Sanderson, M.R., Beese, L., Friedman, J.M., Joyce, C.M., and Steitz, T.A. (1988). Genetic and crystallographic studies of the 3',5'-exonucleolytic site of DNA-Polymerase-I. Science *240*, 199-201.

Derbyshire, V., Grindley, N.D.F., and Joyce, C.M. (1991). The 3'-5' exonuclease of DNA Polymerase-I of Escherichia-coli - contribution of each amino-acid at the active-site to the reaction. Embo Journal *10*, 17-24.

Dervan, J.J., Feng, M., Patel, D., Grasby, J.A., Artymiuk, P.J., Ceska, T.A., and Sayers, J.R. (2002). Interactions of mutant and wild-type flap endonucleases with oligonucleotide substrates suggest an alternative model of DNA binding. Proceedings of the National Academy of Sciences of the United States of America *99*, 8542-8547.

Deshpande, R.A., and Shankar, V. (2002). Ribonucleases from T2 family. Critical Reviews in Microbiology *28*, 79-122.

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

Dianov, G.L., and Huebscher, U. (2013). Mammalian Base Excision Repair: the Forgotten Archangel. Nucleic Acids Research *41*, 3483-3490.

Dianov, G.L., Sleeth, K.M., Dianova, II, and Allinson, S.L. (2003). Repair of abasic sites in DNA. Mutation Research-Fundamental and Molecular Mechanisms of Mutagenesis *531*, 157-163.

Doan, L., Handa, B., Roberts, N.A., and Klumpp, K. (1999). Metal ion catalysis of RNA cleavage by the influenza virus endonuclease. Biochemistry *38*, 5612-5619.

Doublie, S., Tabor, S., Long, A.M., Richardson, C.C., and Ellenberger, T. (1998). Crystal structure of a bacteriophage T7 DNA replication complex at 2.2 angstrom resolution. Nature *391*, 251-258.

Dupre, A., Boyer-Chatenet, L., Sattler, R.M., Modi, A.P., Lee, J.-H., Nicolette, M.L., Kopelovich, L., Jasin, M., Baer, R., Paull, T.T.*, et al.* (2008). A forward chemical genetic screen reveals an inhibitor of the Mre11-Rad50-Nbs1 complex. Nature Chemical Biology *4*, 119-125.

Durant, S.T. (2012). Telomerase-Independent Paths to Immortality in Predictable Cancer Subtypes. Journal of Cancer *3*, 67-82.

Eckstein, F. (1985). Nucleoside phosphorothioates. Annual Review of Biochemistry *54*, 367-402.

El-Khamisy, S.F., Masutani, M., Suzuki, H., and Caldecott, K.W. (2003). A requirement for PARP-1 for the assembly or stability of XRCC1 nuclear foci at sites of oxidative DNA damage. Nucleic Acids Research *31*, 5526-5533.

Eritja, R., Kaplan, B.E., Mhaskar, D., Sowers, L.C., Petruska, J., and Goodman, M.F. (1986). Synthesis and properties of defined DNA oligomers containing base mispairs involving 2-aminopurine. Nucleic Acids Research *14*, 5869-5884.

Espeseth, A.S., Felock, P., Wolfe, A., Witmer, M., Grobler, J., Anthony, N., Egbertson, M., Melamed, J.Y., Young, S., Hamill, T.*, et al.* (2000). HIV-1 integrase inhibitors that compete with the target DNA substrate define a unique strand transfer conformation for integrase. Proceedings of the National Academy of Sciences of the United States of America *97*, 11244-11249.

Evans, R.J., Davies, D.R., Bullard, J.M., Christensen, J., Green, L.S., Guiles, J.W., Pata, J.D., Ribble, W.K., Janjic, N., and Jarvis, T.C. (2008). Structure of PolC reveals unique DNA binding and fidelity determinants. Proceedings of the National Academy of Sciences of the United States of America *105*, 20695-20700.

Fairclough, R.H., and Cantor, C.R. (1978). The use of singlet-singlet energy transfer to study macromolecular assemblies. Methods in enzymology *48*, 347-379.

Farmer, H., McCabe, N., Lord, C.J., Tutt, A.N.J., Johnson, D.A., Richardson, T.B., Santarosa, M., Dillon, K.J., Hickson, I., Knights, C.*, et al.* (2005). Targeting the DNA repair defect in BRCA mutant cells as a therapeutic strategy. Nature *434*, 917-921.

Feng, M., Patel, D., Dervan, J.J., Ceska, T., Suck, D., Haq, I., and Sayers, J.R. (2004). Roles of divalent metal ions in flap endonuclease-substrate interactions. Nature Structural & Molecular Biology *11*, 450-456.

Fidalgo da Silva, E., Mandal, S.S., and Reha-Krantz, L.J. (2002). Using 2-aminopurine fluorescence to measure incorporation of incorrect nucleotides by wild type and mutant bacteriophage T4 DNA polymerases. The Journal of biological chemistry *277*, 40640-40649.

Finger, L.D., Atack, J.M., Tsutakawa, S., Classen, S., Tainer, J., Grasby, J., and Shen, B. (2012). The wonders of flap endonucleases: structure, function, mechanism and regulation. Sub-cellular biochemistry *62*, 301-326.

Finger, L.D., Patel, N., Beddows, A., Ma, L., Exell, J.C., Jardine, E., Jones, A.C., and Grasby, J.A. (2013). Observation of unpaired substrate DNA in the flap endonuclease-1 active site. Nucleic Acids Research *41*, 9839-9847.

Fogg, J.M., Kvaratskhelia, M., White, M.F., and Lilley, D.M.J. (2001). Distortion of DNA junctions imposed by the binding of resolving enzymes: A fluorescence study. Journal of Molecular Biology *313*, 751-764.

Forster, T. (1948). *Zwischenmolekulare energiewanderung und fluoreszenz. Annalen Der Physik *2*, 55-75.

Freemont, P.S., Friedman, J.M., Beese, L.S., Sanderson, M.R., and Steitz, T.A. (1988). Cocrystal structure of an editing complex of Klenow fragment with DNA. Proceedings of the National Academy of Sciences of the United States of America *85*, 8924-8928.

Freire, E., Mayorga, O.L., and Straume, M. (1990). Isothermal titration calorimetry. Analytical Chemistry *62*, A950-A959.

Frey, M.W., Sowers, L.C., Millar, D.P., and Benkovic, S.J. (1995). The nucleotide analog 2-aminopurine as a spectroscopic probe of nucleotide incorporation by the Klenow fragment of Escherichia-coli Ploymerase-I and bacteriophage-T4 DNA-Polymerase. Biochemistry *34*, 9185-9192.

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 Research *32*, 2520-2528.

Fung, H., and Demple, B. (2005). A vital role for Ape1/Ref1 protein in repairing spontaneous DNA damage in human cells. Molecular Cell *17*, 463-470.

Furgason, J.M., and Bahassi, E.M. (2013). Targeting DNA repair mechanisms in cancer. Pharmacology & Therapeutics *137*, 298-308.

Garg, P., and Burgers, P.M.J. (2005). DNA polymerases that propagate the eukaryotic DNA replication fork. Critical Reviews in Biochemistry and Molecular Biology *40*, 115-128.

Gatfield, D., and Izaurralde, E. (2004). Nonsense-mediated messenger RNA decay is initiated by endonucleolytic cleavage in Drosophila. Nature *429*, 575-578.

Gerlinger, M., Rowan, A.J., Horswell, S., Larkin, J., Endesfelder, D., Gronroos, E., Martinez, P., Matthews, N., Stewart, A., Tarpey, P.*, et al.* (2012). Intratumor Heterogeneity and Branched Evolution Revealed by Multiregion Sequencing. New England Journal of Medicine *366*, 883-892.

Gerlt, J.A., Coderre, J.A., and Mehdi, S. (1983). Oxygen chiral phospahte-esters. Advances in Enzymology and Related Areas of Molecular Biology *55*, 291-380.

Gloor, J.W., Balakrishnan, L., and Bambara, R.A. (2010). Flap Endonuclease 1 Mechanism Analysis Indicates Flap Base Binding Prior to Threading. Journal of Biological Chemistry *285*, 34922-34931.

Goodman, M.F., Hopkins, R., and Gore, W.C. (1977). 2-aminopurine-induced mutagenesis in T4-bacteriophage - model relating mutation frequency to 2-aminopurine incorporation in DNA. Proceedings of the National Academy of Sciences of the United States of America *74*, 4806-4810.

Gorman, M.A., Morera, S., Rothwell, D.G., delaFortelle, E., Mol, C.D., Tainer, J.A., Hickson, I.D., and Freemont, P.S. (1997). The crystal structure of the human DNA repair endonuclease HAP1 suggests the recognition of extra-helical deoxyribose at DNA abasic sites. Embo Journal *16*, 6548-6558.

Gottlin, E.B., Rudolph, A.E., Zhao, Y., Matthews, H.R., and Dixon, J.E. (1998). Catalytic mechanism of the phospholipase D superfamily proceeds via a covalent phosphohistidine intermediate. Proceedings of the National Academy of Sciences of the United States of America *95*, 9202-9207.

Grasby, J.A., and Connolly, B.A. (1992). Stereochemicaal outcome of the hydrolysis reaction catalyzed by the ECORV restriction endonuclease. Biochemistry *31*, 7855-7861.

Grasby, J.A., Finger, L.D., Tsutakawa, S.E., Atack, J.M., and Tainer, J.A. (2012). Unpairing and gating: sequence-independent substrate recognition by FEN superfamily nucleases. Trends in Biochemical Sciences *37*, 74-84.

Greene, A.L., Snipe, J.R., Gordenin, D.A., and Resnick, M.A. (1999). Functional analysis of human FEN1 in Saccharomyces cerevisiae and its role in genome stability. Human molecular genetics *8*, 2263-2273.

Guest, C.R., Hochstrasser, R.A., Sowers, L.C., and Millar, D.P. (1991). Dynamics of mismatched base-pairs in DNA. Biochemistry *30*, 3271-3279.

Guo, Z., Kanjanapangka, J., Liu, N., Liu, S., Liu, C., Wu, Z., Wang, Y., Loh, T., Kowolik, C., Jamsen, J.*, et al.* (2012). Sequential Posttranslational Modifications Program FEN1 Degradation during Cell-Cycle Progression. Molecular Cell *47*, 444-456.

Gupta, A.P., and Benkovic, S.J. (1984). Stereochemical course of the 3'- 5'-exonuclease activity of DNA-Polymerase-I. Biochemistry *23*, 5874-5881.

Gwon, G.H., Jo, A., Baek, K., Jin, K.S., Fu, Y., Lee, J.-B., Kim, Y., and Cho, Y. (2014). Crystal structures of the structure-selective nuclease Mus81-Eme1 bound to flap DNA substrates. Embo Journal *33*, 1061-1072.

Habraken, Y., Sung, P., Prakash, L., and Prakash, S. (1993). Yeast excision-repair gene RAD2 encodes a single-stranded-DNA endonuclease. Nature *366*, 365-368.

Hardeland, U., Kunz, C., Focke, F., Szadkowski, M., and Schaer, P. (2007). Cell cycle regulation as a mechanism for functional separation of the apparently redundant uracil DNA glycosylases TDG and UNG2. Nucleic Acids Research *35*, 3859-3867.

Hariharan, C., and Reha-Krantz, L.J. (2005). Using 2-aminopurine fluorescence to detect bacteriophage T4 DNA polymerase-DNA complexes that are important for primer extension and proofreading reactions. Biochemistry *44*, 15674-15684.

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

Harrington, J.J., and Lieber, M.R. (1995). DNA structural elements required for FEN-1 binding. Journal of Biological Chemistry *270*, 4503-4508.

Hazuda, D.J., Felock, P., Witmer, M., Wolfe, A., Stillmock, K., Grobler, J.A., Espeseth, A., Gabryelski, L., Schleif, W., Blau, C.*, et al.* (2000). Inhibitors of strand transfer that prevent integration and inhibit HIV-1 replication in cells. Science *287*, 646-650.

Hengge, A.C. (2005). Mechanistic studies on enzyme-catalyzed phosphoryl transfer. Advances in Physical Organic Chemistry, Vol 40 *40*, 49-108.

Hohl, M., Dunand-Sauthier, I., Staresinic, L., Jaquier-Gubler, P., Thorel, F., Modesti, M., Clarkson, S.G., and Schaerer, O.D. (2007). Domain swapping between FEN-1 and XPG defines regions in XPG that mediate nucleotide excision repair activity and substrate specificity. Nucleic Acids Research *35*, 3053-3063.

Holdgate, G.A., and Ward, W.H.J. (2005). Measurements of binding thermodynamics in drug discovery. Drug Discovery Today *10*, 1543-1550.

Holmen, A., Norden, B., and Albinsson, B. (1997). Electronic transition moments of 2-aminopurine. Journal of the American Chemical Society *119*, 3114-3121.

Hopfner, K.P., Karcher, A., Craig, L., Woo, T.T., Carney, J.P., and Tainer, J.A. (2001). Structural biochemistry and interaction architecture of the DNA double-strand break repair Mre11 nuclease and Rad50-ATPase. Cell *105*, 473-485.

Horton, N.C. (2008). DNA Nucleases. Protein-Nucleic Acid Interactions: Structural Biology, 333-366.

Hosfield, D.J., Frank, G., Weng, Y.H., Tainer, J.A., and Shen, B. (1998a). Newly discovered archaebacterial flap endonucleases show a structure-specific mechanism for DNA substrate binding and catalysis resembling human flap endonuclease-1. Journal of Biological Chemistry *273*, 27154-27161.

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

Houtman, J.C.D., Brown, P.H., Bowden, B., Yamaguchi, H., Appella, E., Samelson, L.E., and Schuck, P. (2007). Studying multisite binary and ternary protein interactions by global analysis of isothermal titration calorimetry data in SEDPHAT: Application to adaptor protein complexes in cell signaling. Protein Science *16*, 30-42.

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

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

Hwang, K.Y., Baek, K., Kim, H.Y., and Cho, Y. (1998). The crystal structure of flap endonuclease-1 from Methanococcus jannaschii. Nature Structural Biology *5*, 707-713.

Illuzzi, J.L., and Wilson, D.M., III (2012). Base Excision Repair: Contribution to Tumorigenesis and Target in Anticancer Treatment Paradigms. Current Medicinal Chemistry *19*, 3922-3936.

Ishikawa, G., Kanai, Y., Takata, K., Takeuchi, R., Shimanouchi, K., Ruike, T., Furukawa, T., Kimura, S., and Sakaguchi, K. (2004). DmGEN, a novel RAD2 family endo-exonuclease from Drosophila melanogaster. Nucleic Acids Research *32*, 6251-6259.

Ivanov, I., Chapados, B.R., McCammon, J.A., and Tainer, J.A. (2006). Proliferating cell nuclear antigen loaded onto double-stranded DNA: dynamics, minor groove interactions and functional implications. Nucleic Acids Research *34*, 6023-6033.

Jean, J.M., and Hall, K.B. (2001). 2-Aminopurine fluorescence quenching and lifetimes: Role of base stacking. Proceedings of the National Academy of Sciences of the United States of America *98*, 37-41.

Jencks, W.P. (1972). General acid-base catalysis of complex reactions in water. Chemical Reviews *72*, 705-&.

Jinek, M., Coyle, S.M., and Doudna, J.A. (2011). Coupled 5 ' Nucleotide Recognition and Processivity in Xrn1-Mediated mRNA Decay. Molecular Cell *41*, 600-608.

Johnson, D.A. (2005). C-terminus of a long α-neurotoxin is highly mobile when bound to the nicotinic acetylcholine receptor: a time-resolved fluorescence anisotropy approach. Biophysical chemistry *116*, 213-218.

Johnson, N.P., Baase, W.A., and von Hippel, P.H. (2004). Low-energy circular dichroism of 2-aminopurine dinucleotide as a probe of local conformation of DNA and RNA. Proceedings of the National Academy of Sciences of the United States of America *101*, 3426-3431.

Johnson, N.P., Baase, W.A., and von Hippel, P.H. (2005a). Investigating local conformations of double-stranded DNA by low-energy circular dichroism of pyrrolo-cytosine. Proceedings of the National Academy of Sciences of the United States of America *102*, 7169-7173.

Johnson, N.P., Baase, W.A., and von Hippel, P.H. (2005b). Low energy CD of RNA hairpin unveils a loop conformation required for lambda N antitermination activity. Journal of Biological Chemistry *280*, 32177-32183.

Johnson, W.C., and Tinoco, I. (1969). Circular dichrosim of polynucleotides. A generaal method applied to dimers. Biopolymers *8*, 715-&.

Jones, C.I., Zabolotskaya, M.V., and Newbury, S.F. (2012). The 5'-3' exoribonuclease XRN1/Pacman and its functions in cellular processes and development. Wiley Interdisciplinary Reviews-Rna *3*, 455-468.

Jose, D., Datta, K., Johnson, N.P., and von Hippel, P.H. (2009). Spectroscopic studies of position-specific DNA "breathing'' fluctuations at replication forks and primer-template junctions. Proceedings of the National Academy of Sciences of the United States of America *106*, 4231-4236.

Jose, D., Weitzel, S.E., and von Hippel, P.H. (2012). Breathing fluctuations in position-specific DNA base pairs are involved in regulating helicase movement into the replication fork. Proceedings of the National Academy of Sciences of the United States of America *109*, 14428-14433.

Kaelin, W.G. (2005). The concept of synthetic lethality in the context of anticancer therapy. Nature Reviews Cancer *5*, 689-698.

Kahn, J.D., Yun, E., and Crothers, D.M. (1994). Detection of localized DNA flexibility. Nature *368*, 163-166.

Kaiser, M.W., Lyamicheva, N., Ma, W.P., Miller, C., Neri, B., Fors, L., and Lyamichev, V.I. (1999). A comparison of eubacterial and archaeal structure-specific 5 '-exonucleases. Journal of Biological Chemistry *274*, 21387-21394.

Kakolyris, S., Kaklamanis, L., Engels, K., Fox, S.B., Taylor, M., Hickson, I.D., Gatter, K.C., and Harris, A.L. (1998). Human AP endonuclease 1 (HAP1) protein expression in breast cancer correlates with lymph node status and angiogenesis. British Journal of Cancer *77*, 1169-1173.

Kanazhevskaya, L.Y., Koval, V.V., Lomzov, A.A., and Fedorova, O.S. (2014). The role of Asn-212 in the catalytic mechanism of human endonuclease APE1: Stopped-flow kinetic study of incision activity on a natural AP site and a tetrahydrofuran analogue. DNA Repair *21*, 43-54.

Kao, H.-I., Henricksen, L.A., Liu, Y., and Bambara, R.A. (2002). Cleavage specificity of Saccharomyces cerevisiae flap endonuclease 1 suggests a double-flap structure as the cellular substrate. Journal of Biological Chemistry *277*, 14379-14389.

Keller, S., Vargas, C., Zhao, H., Piszczek, G., Brautigam, C.A., and Schuck, P. (2012). High-Precision Isothermal Titration Calorimetry with Automated Peak-Shape Analysis. Analytical Chemistry *84*, 5066-5073.

Kim, C.Y., Park, M.S., and Dyer, R.B. (2001). Human flap endonuclease-1: Conformational change upon binding to the flap DNA substrate and location of the Mg2+ binding site. Biochemistry *40*, 3208-3214.

Kim, C.Y., Shen, B.H., Park, M.S., and Olah, G.A. (1999). Structural changes measured by x-ray scattering from human flap endonuclease-1 complexed with Mg2+ and flap DNA substrate. Journal of Biological Chemistry *274*, 1233-1239.

Kim, J.M., Sohn, H.Y., Yoon, S.Y., Oh, J.H., Yang, J.O., Kim, J.H., Song, K.S., Rho, S.M., Yoo, H.S., Kim, Y.S., *et al.* (2005). Identification of gastric cancer-related genes using a cDNA microarray containing novel expressed sequence tags expressed in gastric cancer cells. Clinical Cancer Research *11*, 473-482.

Kim, K., Biade, S., and Matsumoto, Y. (1998). Involvement of flap endonuclease 1 in base excision DNA repair. Journal of Biological Chemistry *273*, 8842-8848.

Kim, Y., Eom, S.H., Wang, J.M., Lee, D.S., Suh, S.W., and Steitz, T.A. (1995). Crystal-structure of thermus-aquaticus DNA-polymerase. Nature *376*, 612-616.

Kirby, A.J., and Jencks, W.P. (1965). Reactivity of nucleophilic reagents toward P-nitrophenyl phosphate dianion. Journal of the American Chemical Society *87*, 3209-&.

Kirby, A.J., and Varvogli.Ag (1967). Reactivity of phospahte esters. Monoester hydrolysis. Journal of the American Chemical Society *89*, 415-&.

Klenow, H., and Overgaar.K (1970). Proteolytic cleavage of DNA polymerase from Escherichia-coli B into an exonuclease unit and a polymerase unit. Febs Letters *6*, 25-&.

Klostermeier, D., and Millar, D.P. (2002). Energetics of hydrogen bond networks in RNA: Hydrogen bonds surrounding G+1 and U42 are the major determinants for the tertiary structure stability of the hairpin ribozyme. Biochemistry *41*, 14095-14102.

Klungland, A., and Lindahl, T. (1997). Second pathway for completion of human DNA base excision-repair: Reconstitution with purified proteins and requirement for DNase IV (FEN1). Embo Journal *16*, 3341-3348.

Kostrewa, D., and Winkler, F.K. (1995). Mg^{2+} binding to the active-site of ECORV endonuclease - a crystallographic study of complexes with substrate and product DNA at 2-angstrom resolution. Biochemistry *34*, 683-696.

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

Krause, A., Combaret, V., Iacono, I., Lacroix, B., Compagnon, C., Bergeron, C., Valsesia-Wittmann, S., Leissner, P., Mougin, B., and Puisieux, A. (2005). Genome-wide analysis of gene expression in neuroblastomas detected by mass screening. Cancer Letters *225*, 111-120.

Krishnan, L., Li, X., Naraharisetty, H.L., Hare, S., Cherepanov, P., and Engelman, A. (2010). Structure-based modeling of the functional HIV-1 intasome and its inhibition. Proceedings of the National Academy of Sciences of the United States of America *107*, 15910-15915.

Kubota, Y., Nash, R.A., Klungland, A., Schar, P., Barnes, D.E., and Lindahl, T. (1996). Reconstitution of DNA base excision-repair with purified human proteins: Interaction between DNA polymerase beta and the XRCC1 protein. Embo Journal *15*, 6662-6670.

Kunkel, T.A. (2004). DNA replication fidelity. Journal of Biological Chemistry *279*, 16895-16898.

Kuzminov, A. (2001). Single-strand interruptions in replicating chromosomes cause double-strand breaks. Proceedings of the National Academy of Sciences of the United States of America *98*, 8241-8246.

Lafemina, R.L., Schneider, C.L., Robbins, H.L., Callahan, P.L., Legrow, K., Roth, E., Schleif, W.A., and Emini, E.A. (1992). Requirement of active human-immunodeficiency-virus type-1 integrase enzyme for productive infection of human T-lymphoid cells. Journal of Virology *66*, 7414-7419.

Lagomarsino, S., Iotti, S., Farruggia, G., Trapani, V., Cedola, A., Fratini, M., Bukreeva, I., Mastrototaro, L., Notargiacomo, A., McNulty, I.*, et al.* (2010). Combined X-ray Microfluorescence and Atomic Force Microscopy Studies of Mg Distribution in Whole Cells. Paper presented at: 10th International Conference on X-ray Microscopy (Univ Chicago, Chicago, IL).

Lai, B., Li, Y., Cao, A., and Lai, L.H. (2003). Metal ion binding and enzymatic mechanism of Methanococcus jannaschii RNase HII. Biochemistry *42*, 785-791.

Lakowicz, J.R. (1999). Energy transfer. In Principles of fluorescence spectroscopy (Springer), pp. 367-394.

Lam, J.S., Seligson, D.B., Yu, H., Li, A., Eeva, M., Pantuck, A.J., Zeng, G., Horvath, S., and Belldegrun, A.S. (2006). Flap endonuclease 1 is overexpressed in prostate cancer and is associated with a high Gleason score. Bju International *98*, 445-451.

Lamichhane, R., Berezhna, S.Y., Gill, J.P., Van der Schans, E., and Millar, D.P. (2013). Dynamics of Site Switching in DNA Polymerase. Journal of the American Chemical Society *135*, 4735-4742.

Lau, J.P., Weatherdon, K.L., Skalski, V., and Hedley, D.W. (2004). Effects of gemcitabine on APE/ref-1 endonuclease activity in pancreatic cancer cells, and the therapeutic potential of antisense oligonucleotides. British Journal of Cancer *91*, 1166-1173.

Lee, J.-B., Cho, W.-K., Park, J., Jeon, Y., Kim, D., Lee, S.H., and Fishel, R. (2014). Single-molecule views of MutS on mismatched DNA. DNA Repair *20*, 82-93.

Li, Y., Korolev, S., and Waksman, G. (1998). Crystal structures of open and closed forms of binary and ternary complexes of the large fragment of Thermus aquaticus DNA polymerase I: structural basis for nucleotide incorporation. Embo Journal *17*, 7514-7525.

Liao, X.M., Anjaneyulu, P.S.R., Curley, J.F., Hsu, M., Boehringer, M., Caruthers, M.H., and Piccirilli, J.A. (2001). The Tetrahymena ribozyme cleaves a 5 '-methylene phosphonate monoester similar to 10(2)-fold faster than a normal phosphate diester: Implications for enzyme catalysis of phosphoryl transfer reactions. Biochemistry *40*, 10911-10926.

Lin, S., Horning, D.P., Szostak, J.W., and Chaput, J.C. (2009). Conformational Analysis of DNA Repair Intermediates by Time-Resolved Fluorescence Spectroscopy. Journal of Physical Chemistry A *113*, 9585-9587.

Liu, L., and Gerson, S.L. (2004). Therapeutic impact of methoxyamine: blocking repair of abasic sites in the base excision repair pathway. Current opinion in investigational drugs (London, England : 2000) *5*, 623-627.

Liu, L., Taverna, P., Whitacre, C.M., Chatterjee, S., and Gerson, S.L. (1999). Pharmacologic disruption of base excision repair sensitizes mismatch repair-deficient and -proficient colon cancer cells to methylating agents. Clinical Cancer Research *5*, 2908-2917.

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

Liu, Y., Beard, W.A., Shock, D.D., Prasad, R., Hou, E.W., and Wilson, S.H. (2005). DNA polymerase beta and flap endonuclease 1 enzymatic specificities sustain DNA synthesis for long patch base excision repair. Journal of Biological Chemistry *280*, 3665-3674.

Liu, Y., Kao, H.I., and Bambara, R.A. (2004). Flap endonuclease 1: A central component of DNA metabolism. Annual Review of Biochemistry *73*, 589-615.

Longley, D.B., and Johnston, P.G. (2005). Molecular mechanisms of drug resistance. Journal of Pathology *205*, 275-292.

Longley, M.J., Prasad, R., Srivastava, D.K., Wilson, S.H., and Copeland, W.C. (1998). Identification of 5'-deoxyribose phosphate lyase activity in human DNA polymerase γ and its role in mitochondrial base excision repair in vitro. Proceedings of the National Academy of Sciences *95*, 12244-12248.

Lundblad, J.R., Laurance, M., and Goodman, R.H. (1996). Fluorescence polarization analysis of protein-DNA and protein-protein interactions. Molecular Endocrinology *10*, 607-612.

Lyamichev, V., Brow, M.A.D., and Dahlberg, J.E. (1993). Structure-specific endonucleoytic cleavage of nucleic-acids by Eubacterial DNA-polymerases. Science *260*, 778-783.

Madhusudan, S., Smart, F., Shrimpton, P., Parsons, J.L., Gardiner, L., Houlbrook, S., Talbot, D.C., Hammonds, T., Freemont, P.A., Sternberg, M.J.E.*, et al.* (2005). Isolation of a small molecule inhibitor of DNA base excision repair. Nucleic Acids Research *33*, 4711-4724.

Mandal, S.S., da Silva, E.F., and Reha-Krantz, L.J. (2002). Using 2-aminopurine fluorescence to detect base unstacking in the template strand during nucleotide incorporation by the bacteriophage T4 DNA polymerase. Biochemistry *41*, 4399-4406.

Manvilla, B.A., Pozharski, E., Toth, E.A., and Drohat, A.C. (2013). Structure of human apurinic/apyrimidinic endonuclease 1 with the essential Mg2+ cofactor. Acta Crystallographica Section D-Biological Crystallography *69*, 2555-2562.

Marchand, C., Johnson, A.A., Karki, R.G., Pais, G.C.G., Zhang, X.C., Cowansage, K., Patel, T.A., Nicklaus, M.C., Burke, T.R., and Pommier, Y. (2003). Metal-dependent inhibition of HIV-1 integrase by beta-diketo acids and resistance of the soluble double-mutant (F185K/C280S). Molecular Pharmacology *64*, 600-609.

Matsui, E., Musti, K.V., Abe, J., Yamasaki, K., Matsui, I., and Harata, K. (2002). Molecular structure and novel DNA binding sites located in loops of flap endonuclease-1 from Pyrococcus horikoshii. Journal of Biological Chemistry *277*, 37840-37847.

McCullough, A.K., Dodson, M.L., and Lloyd, R.S. (1999). Initiation of base excision repair: Glycosylase mechanisms and structures. Annual Review of Biochemistry *68*, 255-285.

McManus, K.J., Barrett, I.J., Nouhi, Y., and Hieter, P. (2009). Specific synthetic lethal killing of RAD54B-deficient human colorectal cancer cells by FEN1 silencing. Proceedings of the National Academy of Sciences of the United States of America *106*, 3276-3281.

McWhirter, C., Lund, E.A., Tanifum, E.A., Feng, G., Sheikh, Q.I., Hengge, A.C., and Williams, N.H. (2008). Mechanistic study of protein phosphatase-1 (PP1), a catalytically promiscuous enzyme. Journal of the American Chemical Society *130*, 13673-13682.

McWhirter, C., Tonge, M., Plant, H., Hardern, I., Nissink, W., and Durant, S.T. (2013). Development of a High-Throughput Fluorescence Polarization DNA Cleavage Assay for the Identification of FEN1 Inhibitors. Journal of biomolecular screening *18*, 567-575.

Miki, Y., Swensen, J., Shattuckeidens, D., Futreal, P.A., Harshman, K., Tavtigian, S., Liu, Q.Y., Cochran, C., Bennett, L.M., Ding, W.*, et al.* (1994). A strong candidate for the breast and ovarian-cancer susceptibility gene BRCA1. Science *266*, 66-71.

Mitsunobu, H., Zhu, B., Lee, S.-J., Tabor, S., and Richardson, C.C. (2014). Flap Endonuclease Activity of Gene 6 Exonuclease of Bacteriophage T7. Journal of Biological Chemistry *289*, 5860-5875.

Mizoue, L.S., and Tellinghuisen, J. (2004). The role of backlash in the "first injection anomaly" in isothermal titration calorimetry. Analytical Biochemistry *326*, 125-127.

Mol, C.D., Hosfield, D.J., and Tainer, J.A. (2000a). Abasic site recognition by two apurinic/apyrimidinic endonuclease families in DNA base excision repair: the 3 ' ends justify the means. Mutation Research-DNA Repair *460*, 211-229.

Mol, C.D., Izumi, T., Mitra, S., and Tainer, J.A. (2000b). DNA-bound structures and mutants reveal abasic DNA binding by APE1 DNA repair and coordination. Nature *403*, 451-456.

Moore, D.H., Michael, H., Tritt, R., Parsons, S.H., and Kelley, M.R. (2000). Alterations in the expression of the DNA repair/redox enzyme APE/ref-1 in epithelial ovarian cancers. Clinical Cancer Research *6*, 602-609.

Moynahan, M.E., Chiu, J.W., Koller, B.H., and Jasin, M. (1999). Brca1 controls homology-directed DNA repair. Molecular Cell *4*, 511-518.

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.

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. Journal of Biological Chemistry *270*, 30377-30383.

Navarro, M.S., Bi, L., and Bailis, A.M. (2007). A mutant allele of the transcription factor IIH helicase gene, RAD3, promotes loss of heterozygosity in response to a DNA replication defect in Saccharomyces cerevisiae. Genetics *176*, 1391-1402.

Neely, R.K., Tamulaitis, G., Chen, K., Kubala, M., Siksnys, V., and Jones, A.C. (2009). Time-resolved fluorescence studies of nucleotide flipping by restriction enzymes. Nucleic Acids Research *37*, 6859-6870.

Newman, M., Murray-Rust, J., Lally, J., Rudolf, J., Fadden, A., Knowles, P.P., White, M.F., and McDonald, N.Q. (2005). Structure of an XPF endonuclease with and without DNA suggests a model for substrate recognition. Embo Journal *24*, 895-905.

Nikolova, T., Christmann, M., and Kaina, B. (2009). FEN1 is Overexpressed in Testis, Lung and Brain Tumors. Anticancer Research *29*, 2453-2459.

Nimonkar, A.V., Ozsoy, A.Z., Genschel, J., Modrich, P., and Kowalczykowski, S.C. (2008). Human exonuclease 1 and BLM helicase interact to resect DNA and initiate DNA repair. Proceedings of the National Academy of Sciences of the United States of America *105*, 16906-16911.

Nishino, T., Komori, K., Ishino, Y., and Morikawa, K. (2005). Structural and functional analyses of an archaeal XPF/Rad1/Mus81 nuclease: Asymmetric DNA binding and cleavage mechanisms. Structure *13*, 1183-1192.

Nolan, J.P., Shen, B.H., Park, M.S., and Sklar, L.A. (1996). Kinetic analysis of human flap endonuclease-1 by flow cytometry. Biochemistry *35*, 11668-11676.

Nowotny, M., Gaidamakov, S.A., Crouch, R.J., and Yang, W. (2005). Crystal structures of RNase H bound to an RNA/DNA hybrid: Substrate specificity and metal-dependent catalysis. Cell *121*, 1005-1016.

Odonovan, A., Davies, A.A., Moggs, J.G., West, S.C., and Wood, R.D. (1994). XPG endonuclease makes the 3' incision in human DNA nucleotide excision-repair. Nature *371*, 432-435.

Oferrall, R.A. (1970). Relationships between E2 and E1CB mechanism of Beta-elimination. Journal of the Chemical Society B-Physical Organic, 274-&.

Ogawa, T., and Okazaki, T. (1980). Discontinuous DNA-replication. Annual Review of Biochemistry *49*, 421-457.

Ollis, D.L., Brick, P., Hamlin, R., Xuong, N.G., and Steitz, T.A. (1985). Structure of large fragment of Escherichia-coli DNA-Polymerase-I complexed with DTMP. Nature *313*, 762-766.

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.

Papeo, G., Casale, E., Montagnoli, A., and Cirla, A. (2013). PARP inhibitors in cancer therapy: an update. Expert Opinion on Therapeutic Patents *23*, 503-514.

Parkes, K.E.B., Ermert, P., Fassler, J., Ives, J., Martin, J.A., Merrett, J.H., Obrecht, D., Williams, G., and Klumpp, K. (2003). Use of a pharmacophore model to discover a new class of influenza endonuclease inhibitors. Journal of Medicinal Chemistry *46*, 1153-1164.

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

Patel, N., Exell, J.C., Jardine, E., Ombler, B., Finger, L.D., Ciani, B., and Grasby, J.A. (2013). Proline Scanning Mutagenesis Reveals a Role for the Flap Endonuclease-1 Helical Cap in Substrate Unpairing. Journal of Biological Chemistry *288*, 34239-34248.

Paull, T.T., and Gellert, M. (1998). The 3 ' to 5 ' exonuclease activity of Mre11 facilitates repair of DNA double-strand breaks. Molecular Cell *1*, 969-979.

Perozzo, R., Folkers, G., and Scapozza, L. (2004). Thermodynamics of protein-ligand interactions: History, presence, and future aspects. Journal of Receptors and Signal Transduction *24*, 1-52.

Pickering, T.J., Garforth, S., Sayers, J.R., and Grasby, J.A. (1999). Variation in the steady state kinetic parameters of wild type and mutant T5 5 '-3 '-exonuclease with pH - Protonation of Lys-83 is critical for DNA binding. Journal of Biological Chemistry *274*, 17711-17717.

Poole, T.L., and Stevens, A. (1997). Structural Modifications of RNA Influence the 5' Exoribonucleolytic Hydrolysis by XRN1 and HKE1 ofSaccharomyces cerevisiae. Biochemical and biophysical research communications *235*, 799-805.

Potter, B.V.L., Connolly, B.A., and Eckstein, F. (1983). Synthesis and configuration analysis of a dinucleoside phospahte isotopically chiral at phosphorus - stereochemical course of penicillium-citrum nuclease P1 reaction. Biochemistry *22*, 1369-1377.

Purohit, V., Grindley, N.D.F., and Joyce, C.M. (2003). Use of 2-aminopurine fluorescence to examine conformational changes during nucleotide incorporation by DNA polymerase I (Klenow fragment). Biochemistry *42*, 10200-10211.

Qiu, J.Z., Bimston, D.N., Partikian, A., and Shen, B.H. (2002). Arginine residues 47 and 70 of human flap endonuclease-1 are involved in DNA substrate interactions and cleavage site determination. Journal of Biological Chemistry *277*, 24659-24666.

Qiu, J.Z., Liu, R., Chapados, B.R., Sherman, M., Tainer, J.A., and Shen, B.H. (2004). Interaction interface of human flap endonuclease-1 with its DNA substrates. Journal of Biological Chemistry *279*, 24394-24402.

Ramstein, J., and Lavery, R. (1988). Energetic coupling between DNA bending and base pair opening. Proceedings of the National Academy of Sciences of the United States of America *85*, 7231-7235.

Rass, U., Compton, S.A., Matos, J., Singleton, M.R., Ip, S.C.Y., Blanco, M.G., Griffith, J.D., and West, S.C. (2010). Mechanism of Holliday junction resolution by the human GEN1 protein. Genes & Development *24*, 1559-1569.

Richards, J.D., Johnson, K.A., Liu, H., McRobbie, A.-M., McMahon, S., Oke, M., Carter, L., Naismith, J.H., and White, M.F. (2008). Structure of the DNA repair helicase Hel308 reveals DNA binding and autoinhibitory domains. Journal of Biological Chemistry *283*, 5118-5126.

Richardson, T.T., Wu, X., Keith, B.J., Heslop, P., Jones, A.C., and Connolly, B.A. (2013). Unwinding of primer-templates by archaeal family-B DNA polymerases in response to template-strand uracil. Nucleic Acids Research *41*, 2466-2478.

Roberts, J.A., and White, M.F. (2005). DNA end-directed and processive nuclease activities of the archaeal XPF enzyme. Nucleic Acids Research *33*, 6662-6670.

Robertson, A.B., Klungland, A., Rognes, T., and Leiros, I. (2009). DNA Repair in Mammalian Cells. Cellular and Molecular Life Sciences *66*, 981-993.

Robins, P., Pappin, D.J.C., Wood, R.D., and Lindahl, T. (1994). Structural and functional homology between mammalian DNAse-IV and the 5ˈ-nuclease domain of Escherichia-coli DNA-Polymerase-I. Journal of Biological Chemistry *269*, 28535-28538.

Rosta, E., Yang, W., and Hummer, G. (2014). Calcium Inhibition of Ribonuclease H1 Two-Metal Ion Catalysis. Journal of the American Chemical Society *136*, 3137-3144.

Rothwell, P.J., Berger, S., Kensch, O., Felekyan, S., Antonik, M., Wöhrl, B.M., Restle, T., Goody, R.S., and Seidel, C.A.M. (2003). Multiparameter single-molecule fluorescence spectroscopy reveals heterogeneity of HIV-1 reverse transcriptase: primer/template complexes. Proceedings of the National Academy of Sciences *100*, 1655-1660.

Sabir, T., Schroeder, G.F., Toulmin, A., McGlynn, P., and Magennis, S.W. (2011). Global Structure of Forked DNA in Solution Revealed by High-Resolution Single-Molecule FRET. Journal of the American Chemical Society *133*, 1188-1191.

Saharia, A., and Stewart, S.A. (2009). FEN1 contributes to telomere stability in ALT-positive tumor cells. Oncogene *28*, 1162-1167.

Sajesh, B.V., Guppy, B.J., and McManus, K.J. (2013). Synthetic genetic targeting of genome instability in cancer. Cancers *5*, 739-761.

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 Journal *24*, 683-693.

Samson, L., Derfler, B., Boosalis, M., and Call, K. (1991). Cloning and characterisation of a 3-methyladenine DNA glycosylase cDNA from human-cells whose gene maps to chromosome-16. Proceedings of the National Academy of Sciences of the United States of America *88*, 9127-9131.

Sancar, A., Lindsey-Boltz, L.A., Unsal-Kacmaz, K., and Linn, S. (2004). Molecular mechanisms of mammalian DNA repair and the DNA damage checkpoints. Annual Review of Biochemistry *73*, 39-85.

Sato, M., Girard, L., Sekine, I., Sunaga, N., Ramirez, R.D., Kamibayashi, C., and Minna, J.D. (2003). Increased expression and no mutation of the Flap endonuclease (FEN1) gene in human lung cancer. Oncogene *22*, 7243-7246.

Sauer, U.H., Sun, D.P., and Matthews, B.W. (1992). Tolerance of T4-lysozyme to proline substitutions within the long interdomain alpha-helix illustrates the adaptability of proteins to potentially destabilizing lesions. Journal of Biological Chemistry *267*, 2393-2399.

Sayers, J.R., and Artymiuk, P.J. (1998). Flexible loops and helical arches. Nature Structural Biology *5*, 668-670.

Sayers, J.R., and Eckstein, F. (1990). Properties of overexpressed Phage-T5 D15 exonuclease - similarities with Escherichia-coli DNA-Polymerase-I 5'-3' exonuclease. Journal of Biological Chemistry *265*, 18311-18317.

Schaerer, O.D. (2008). XPG: Its products and biological roles. Molecular Mechanisms of Xeroderma Pigmentosum *637*, 83-92.

Schlacher, K., Christ, N., Siaud, N., Egashira, A., Wu, H., and Jasin, M. (2011). Double-Strand Break Repair-Independent Role for BRCA2 in Blocking Stalled Replication Fork Degradation by MRE11. Cell *145*, 529-542.

Schwartz, E.K., and Heyer, W.-D. (2011). Processing of joint molecule intermediates by structure-selective endonucleases during homologous recombination in eukaryotes. Chromosoma *120*, 109-127.

Seiple, L.A., Cardellina, J.H., Akee, R., II, and Stivers, J.T. (2008). Potent inhibition of human apurinic/apyrimidinic endonuclease 1 by arylstibonic acids. Molecular Pharmacology *73*, 669-677.

Sengerova, B., Tomlinson, C., Atack, J.M., Williams, R., Sayers, J.R., Williams, N.H., and Grasby, J.A. (2010). Bronsted Analysis and Rate-Limiting Steps for the T5 Flap Endonuclease Catalyzed Hydrolysis of Exonucleolytic Substrates. Biochemistry *49*, 8085-8093.

Shah, S., Dunten, P., Stiteler, A., Park, C.K., and Horton, N.C. (2015). Structure and specificity of FEN-1 from Methanopyrus kandleri. Proteins-Structure Function and Bioinformatics *83*, 188-194.

Shamoo, Y., and Steitz, T.A. (1999). Building a replisome from interacting pieces: Sliding clamp complexed to a peptide from DNA polymerase and a polymerase editing complex. Cell *99*, 155-166.

Shen, B.H., Nolan, J.P., Sklar, L.A., and Park, M.S. (1997). Functional analysis of point mutations in human flap endonuclease-1 active site. Nucleic Acids Research *25*, 3332-3338.

Shen, B.H., Qiu, J.Z., Hosfield, D., and Tainer, J.A. (1998). Flap endonuclease homologs in archaebacteria exist as independent proteins. Trends in Biochemical Sciences *23*, 171-173.

Shibata, A., Moiani, D., Arvai, A.S., Perry, J., Harding, S.M., Genois, M.-M., Maity, R., van Rossum-Fikkert, S., Kertokalio, A., Romoli, F.*, et al.* (2014). DNA Double-Strand Break Repair Pathway Choice Is Directed by Distinct MRE11 Nuclease Activities. Molecular Cell *53*, 7-18.

Siegal, G., Turchi, J.J., Myers, T.W., and Bambara, R.A. (1992). A 5' to 3' exonuclease functionally interacts with calf DNA polymerase-epsilon. Proceedings of the National Academy of Sciences of the United States of America *89*, 9377-9381.

Silber, J.R., Bobola, M.S., Blank, A., Schoeler, K.D., Haroldson, P.D., Huynh, M.B., and Kolstoe, D.D. (2002). The apurinic/apyrimidinic endonuclease activity of Apel/Ref-1 contributes to human glioma cell resistance to alkylating agents and is elevated by oxidative stress. Clinical Cancer Research *8*, 3008-3018.

Simeonov, A., Kulkarni, A., Dorjsuren, D., Jadhav, A., Shen, M., McNeill, D.R., Austin, C.P., and Wilson, D.M., III (2009). Identification and Characterization of Inhibitors of Human Apurinic/apyrimidinic Endonuclease APE1. Plos One *4*.

Singh, P., Yang, M., Dai, H., Yu, D., Huang, Q., Tan, W., Kernstine, K.H., Lin, D., and Shen, B. (2008). Overexpression and Hypomethylation of Flap Endonuclease 1 Gene in Breast and Other Cancers. Molecular Cancer Research *6*, 1710-1717.

Singh, P., Zheng, L., Chavez, V., Qiu, J., and Shen, B. (2007). Concerted action of exonuclease and gap-dependent endonuclease activities of FEN-1 contributes to the resolution of triplet repeat sequences (CTG)(n)- and (GAA)(n)-derived secondary structures formed during maturation of Okazaki fragments. Journal of Biological Chemistry *282*, 3465-3477.

Sinturel, F., Pellegrini, O., Xiang, S., Tong, L., Condon, C., and Benard, L. (2009). Real-time fluorescence detection of exoribonucleases. Rna-a Publication of the Rna Society *15*, 2057-2062.

Sobhy, M.A., Joudeh, L.I., Huang, X., Takahashi, M., and Hamdan, S.M. (2013). Sequential and Multistep Substrate Interrogation Provides the Scaffold for Specificity in Human Flap Endonuclease 1. Cell Reports *3*, 1785-1794.

Sowers, L.C., Fazakerley, G.V., Eritja, R., Kaplan, B.E., and Goodman, M.F. (1986). Base-pairing and mutagenesis - observation of a protonated base pair between 2-aminopurine and cytosine in an oligonucleotide by proton NMR. Proceedings of the National Academy of Sciences of the United States of America *83*, 5434-5438.

Steinmetzer, K., Behlke, J., Brantl, S., and Lorenz, M. (2002). CopR binds and bends its target DNA: a footprinting and fluorescence resonance energy transfer study. Nucleic Acids Research *30*, 2052-2060.

Steitz, T.A., and Steitz, J.A. (1993). A general 2-metal-ion mechanism for catalytic RNA. Proceedings of the National Academy of Sciences of the United States of America *90*, 6498-6502.

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

Stryer, L., and Haugland, R.P. (1967). Energy-transfer - a spectroscopic ruler. Proceedings of the National Academy of Sciences of the United States of America *58*, 719-&.

Stucki, M., Jonsson, Z.O., and Hubscher, U. (2001). In eukaryotic flap endonuclease 1, the C terminus is essential for substrate binding. Journal of Biological Chemistry *276*, 7843-7849.

Subramanian, K., Rutvisuttinunt, W., Scott, W., and Myers, R.S. (2003). The enzymatic basis of processivity in lambda exonuclease. Nucleic Acids Research *31*, 1585-1596.

Subuddhi, U., Hogg, M., and Reha-Krantz, L.J. (2008). Use of 2-aminopurine fluorescence to study the role of the beta hairpin in the proofreading pathway catalyzed by the phage T4 and RB69 DNA polymerases. Biochemistry *47*, 6130-6137.

Summa, V., Petrocchi, A., Bonelli, F., Crescenzi, B., Donghi, M., Ferrara, M., Fiore, F., Gardelli, C., Paz, O.G., Hazuda, D.J.*, et al.* (2008). Discovery of Raltegravir, a potent, selective orally bioavailable HIV-integrase inhibitor for the treatment of HIV-AIDS infection. Journal of Medicinal Chemistry *51*, 5843-5855.

Swan, M.K., Johnson, R.E., Prakash, L., Prakash, S., and Aggarwal, A.K. (2009). Structural basis of high-fidelity DNA synthesis by yeast DNA polymerase delta. Nature Structural & Molecular Biology *16*, 979-U107.

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. Journal of Biological Chemistry *283*, 28741-28746.

Taylor, R.D., MacCoss, M., and Lawson, A.D.G. (2014). Rings in Drugs. Journal of Medicinal Chemistry *57*, 5845-5859.

Tentori, L., Leonetti, C., Scarsella, M., d'Amati, G., Portarena, I., Zupi, G., Bonmassar, E., and Graziani, G. (2002). Combined treatment with temozolomide and poly(ADP-ribose) polymerase inhibitor enhances survival of mice bearing hematologic malignancy at the central nervous system site. Blood *99*, 2241-2244.

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 Journal *22*, 995-1004.

Tomkinson, A.E., Vijayakumar, S., Pascal, J.M., and Ellenberger, T. (2006). DNA ligases: Structure, reaction mechanism, and function. Chemical Reviews *106*, 687-699.

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. Biochemical Society Transactions *38*, 433-437.

Tomlinson, C.G., Syson, K., Sengerova, B., Atack, J.M., Sayers, J.R., Swanson, L., Tainer, J.A., Williams, N.H., and Grasby, J.A. (2011). Neutralizing Mutations of Carboxylates That Bind Metal 2 in T5 Flap Endonuclease Result in an Enzyme That Still Requires Two Metal Ions. Journal of Biological Chemistry *286*, 30878-30887.

Trujillo, K.M., and Sung, P. (2001). DNA structure-specific nuclease activities in the Saccharomyces cerevisiae Rad50 center dot Mre11 complex. Journal of Biological Chemistry *276*, 35458-35464.

Tsoi, P.Y., Zhang, X.M., Sui, S.F., and Yang, M.S. (2003). Effects of DNA mismatches on binding affinity and kinetics of polymerase-DNA complexes as revealed by surface plasmon resonance biosensor. Analyst *128*, 1169-1174.

Tsutakawa, S.E., Classen, S., Chapados, B.R., Arvai, A.S., Finger, L.D., Guenther, G., Tomlinson, C.G., Thompson, P., Sarker, A.H., Shen, B.*, et al.* (2011). Human Flap Endonuclease Structures, DNA Double-Base Flipping, and a Unified Understanding of the FEN1 Superfamily. Cell *145*, 198-211.

Tsutakawa, S.E., Shin, D.S., Mol, C.D., Izumi, T., Arvai, A.S., Mantha, A.K., Szczesny, B., Ivanov, I.N., Hosfield, D.J., Maiti, B., *et al.* (2013). Conserved Structural Chemistry for Incision Activity in Structurally Non-homologous Apurinic/Apyrimidinic Endonuclease APE1 and Endonuclease IV DNA Repair Enzymes. Journal of Biological Chemistry *288*, 8445-8455.

Tumey, L.N., Bom, D., Huck, B., Gleason, E., Wang, J.M., 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. Bioorganic & Medicinal Chemistry Letters *15*, 277-281.

Turchi, J.J., Huang, L., Murante, R.S., Kim, Y., and Bambara, R.A. (1994). Enzymatic completion of mammalian lagging-strand DNA-replication. Proceedings of the National Academy of Sciences of the United States of America *91*, 9803-9807.

Turnbull, W.B., and Daranas, A.H. (2003). On the value of c: Can low affinity systems be studied by isothermal titration calorimetry? Journal of the American Chemical Society *125*, 14859-14866.

van Oijen, A.M., Blainey, P.C., Crampton, D.J., Richardson, C.C., Ellenberger, T., and Xie, X.S. (2003). Single-molecule kinetics of lambda exonuclease reveal base dependence and dynamic disorder. Science *301*, 1235-1238.

van Pel, D.M., Barrett, I.J., Shimizu, Y., Sajesh, B.V., Guppy, B.J., Pfeifer, T., McManus, K.J., and Hieter, P. (2013). An Evolutionarily Conserved Synthetic Lethal Interaction Network Identifies FEN1 as a Broad-Spectrum Target for Anticancer Therapeutic Development. Plos Genetics *9*.

Walker, L.J., Craig, R.B., Harris, A.L., and Hickson, I.D. (1994). A role for the human DNA-repair enzyme HAP1 in cellular-protection against DNA-damaging agents and hypoxic stress. Nucleic Acids Research *22*, 4884-4889.

Wang, M.L.J., Stellwagen, R.H., and Goodman, M.F. (1981). Evidence for the absence of DNA proofreading in Hela-cell nuclei. Journal of Biological Chemistry *256*, 7097-7100.

Wedge, S.R., Porteous, J.K., May, B.L., and Newlands, E.S. (1996a). Potentiation of temozolomide and BCNU cytotoxicity by O-6-benzylguanine: A comparative study in vitro. British Journal of Cancer *73*, 482-490.

Wedge, S.R., Porteous, J.K., and Newlands, E.S. (1996b). 3-aminobenzamide and/or O-6-benzylguanine evaluated as an adjuvant to temozolomide or BCNU treatment in cell lines of variable mismatch repair status and O-6-alkylguanine-DNA alkyltransferase activity. British Journal of Cancer *74*, 1030-1036.

Weinberg, R.L., Freund, S.M.V., Veprintsev, D.B., Bycroft, M., and Fersht, A.R. (2004). Regulation of DNA binding of p53 by its C-terminal domain. Journal of molecular biology *342*, 801-811.

Widom, J.R., Johnson, N.P., von Hippel, P.H., and Marcus, A.H. (2013). Solution conformation of 2-aminopurine dinucleotide determined by ultraviolet two-dimensional fluorescence spectroscopy. New Journal of Physics *15*.

Williams, A. (1992). Effective charge and transition-state structure in solution. Advances in Physical Organic Chemistry, Vol 27 *27*, 1-55.

Williams, N.H., Cheung, W., and Chin, J. (1998). Reactivity of phosphate diesters doubly coordinated to a dinuclear cobalt(III) complex: Dependence of the reactivity on the basicity of the leaving group. Journal of the American Chemical Society *120*, 8079-8087.

Williams, R., Sengerowa, B., Osborne, S., Syson, K., Ault, S., Kilgour, A., Chapados, B.R., Tainer, J.A., Sayers, J.R., and Grasby, J.A. (2007a). Comparison of the catalytic parameters and reaction specificities of a phage and an archaeal flap endonuclease. Journal of Molecular Biology *371*, 34-48.

Williams, R.S., Moncalian, G., Williams, J.S., Yamada, Y., Limbo, O., Shin, D.S., Groocock, L.M., Cahill, D., Hitomi, C., Guenther, G., *et al.* (2008). Mre11 dimers coordinate DNA end bridging and nuclease processing in double-strand-break repair. Cell *135*, 97-109.

Williams, R.S., Williams, J.S., and Tainer, J.A. (2007b). Mre11-Rad50-Nbs1 is a keystone complex connecting DNA repair machinery, double-strand break signaling, and the chromatin template. Biochemistry and Cell Biology-Biochimie Et Biologie Cellulaire *85*, 509-520.

Wiseman, T., Williston, S., Brandts, J.F., and Lin, L.N. (1989). Rapid measurement of binding constants and heats of binding using a new titration calorimeter. Analytical Biochemistry *179*, 131-137.

Wright, S., and Dobzhansky, T. (1946). Genetics of natural populations .12. experimental reproducion of some of the changes caused by natural selection in certain populations of Drosophila-pseudoobscura. Genetics *31*, 125-156.

Wu, X.T., Li, J., Li, X.Y., Hsieh, C.L., Burgers, P.M.J., and Lieber, M.R. (1996). Processing of branched DNA intermediates by a complex of human FEN-1 and PCNA. Nucleic Acids Research *24*, 2036-2043.

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

Yang, M.S., and Millar, D.P. (1997). Fluorescence resonance energy transfer as a probe of DNA structure and function. Fluorescence Spectroscopy *278*, 417-444.

Yang, W. (2008). An equivalent metal ion in one- and two-metal-ion catalysis. Nature Structural & Molecular Biology *15*, 1228-1231.

Yang, W. (2011). Nucleases: diversity of structure, function and mechanism. Quarterly Reviews of Biophysics *44*, 1-93.

Yang, W., Lee, J.Y., and Nowotny, M. (2006). Making and breaking nucleic acids: Two-Mg2+-ion catalysis and substrate specificity. Molecular Cell *22*, 5-13.

Zalatan, J.G., and Herschlag, D. (2006). Alkaline phosphatase mono- and diesterase reactions: Comparative transition state analysis. Journal of the American Chemical Society *128*, 1293-1303.

Zawahir, Z., Dayam, R., Deng, J., Pereira, C., and Neamati, N. (2009). Pharmacophore Guided Discovery of Small-Molecule Human Apurinic/Apyrimidinic Endonuclease 1 Inhibitors. Journal of Medicinal Chemistry *52*, 20-32.

Zhang, J., McCabe, K.A., and Bell, C.E. (2011). Crystal structures of lambda exonuclease in complex with DNA suggest an electrostatic ratchet mechanism for processivity. Proceedings of the National Academy of Sciences of the United States of America *108*, 11872-11877.

Zhang, J., Stevens, M.F.G., and Bradshaw, T.D. (2012). Temozolomide: mechanisms of action, repair and resistance. Current molecular pharmacology *5*, 102-114.

Zhang, J., Xing, X., Herr, A.B., and Bell, C.E. (2009). Crystal Structure of E-coli RecE Protein Reveals a Toroidal Tetramer for Processing Double-Stranded DNA Breaks. Structure *17*, 690-702.

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

Appendices

```
sequence 1: K93AR100A forward
sequence 2: hFEN1 mRNA

Alignment of Sequence_1:  [038_df_28b-DM_T7F.seq.xdna] with  Sequence_2: [hFEN1_mRNA.xdna]

Similarity : 894/1161 (77.00 %)

          H  N  F  F  N  F  K  K  E  N  T  M  G  I  Q  G  L  A  K  L
Seq_1  1  CCATAATTTTTTTAACTTTAAGAAGGAGAATACCATGGGAATTCAAGGCCTGGCCAAACT   60
          ################################||||||||||||||||||||||||||||
Seq_2  1  --------------------------------atgggaattcaaggcctggccaaact   26
                                           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  Y  F  G  R
Seq_1 61  AATTGCTGATGTGGCCCCCAGTGCCATCCGGGAGAATGACATCAAGAGCTACTTTGGCCG  120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Seq_2 27  aattgctgatgtggcccccagtgccatccgggagaatgacatcaagagctactttggccg   86
          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  V  R  Q  G
Seq_1 121 TAAGGTGGCCATTGATGCCTCTATGAGCATTTATCAGTTCCTGATTGCTGTTCGCCAGGG  180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Seq_2 87  taaggtggccattgatgcctctatgagcatttatcagttcctgattgctgttcgccaggg  146
          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  G  M  F  Y
Seq_1 181 TGGGGATGTGCTGCAGAATGAGGAGGGTGAGACCACCAGCCACCTGATGGGCATGTTCTA  240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Seq_2 147 tggggatgtgctgcagaatgaggagggtgagaccaccagccacctgatgggcatgttcta  206
          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  D  G  K  P
Seq_1 241 CCGCACCATTCGCATGATGGAGAACGGCATCAAGCCCGTGTATGTCTTTGATGGCAAGCC  300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Seq_2 207 ccgcaccattcgcatgatggagaacggcatcaagcccgtgtatgtctttgatggcaagcc  266
          R  T  I  R  M  M  E  N  G  I  K  P  V  Y  V  F  D  G  K  P


          P  Q  L  A  S  G  E  L  A  K  A  S  E  R  R  A  E  A  E  K
Seq_1 301 GCCACAGCTCGCGTCAGGCGAGCTGGCCAAAGCCAGTGAGCGGCGGGCTGAGGCAGAGAA  360
          ||||||||||##||||||||||||||||||##||||||||||||||||||||||||||||
Seq_2 267 gccacagctcaagtcaggcgagctggccaaacgcagtgagcggcgggctgaggcagagaa  326
          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  F  T  K  R
Seq_1 361 GCAGCTGCAGCAGGCTCAGGCTGCTGGGGCCGAGCAGGAGGTGGAAAAATTCACTAAGCG  420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Seq_2 327 gcagctgcagcaggctcaggctgctggggccgagcaggaggtggaaaaattcactaagcg  386
          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  S  L  M  G
Seq_1 421 GCTGGTGAAGGTCACTAAGCAGCACAATGATGAGTGCAAACATCTGCTGAGCCTCATGGG  480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Seq_2 387 gctggtgaaggtcactaagcagcacaatgatgagtgcaaacatctgctgagcctcatggg  446
          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  L  V  K  A
Seq_1 481 CATCCCTTATCTTGATGCACCCAGTGAGGCAGAGGCCAGCTGTGCTGCCCTGGTGAAGGC  540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Seq_2 447 catcccttatcttgatgcacccagtgaggcagaggccagctgtgctgccctggtgaaggc  506
          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  G  S  P  V
Seq_1 541 TGGCAAAGTCTATGCTGCGGCTACCGAGGACATGGACTGCCTCACCTTCGGCAGCCCTGT  600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Seq_2 507 tggcaaagtctatgctgcggctaccgaggacatggactgcctcaccttcggcagccctgt  566
          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  V  K  P  K  S  C  Q  S  R  N  S  T  *
Seq_1 601 GCTAATGCGACACCTGACTGC-AGTGAAGCCAAAAAGCTGCCAATCCAGGAATTCCACCT  659
          |||||||||||||||||||||#||||||||||||||||||||||||||||||||||||||
Seq_2 567 gctaatgcgacacctgactgccagtgaagccaaaaagctgccaatccaggaattccacct  626
```

```
                    L  M  R  H  L  T  A  S  E  A  K  K  L  P  I  Q  E  F  H  L


                    A  G  F  C  R  S  W  A  *  T  R  N  S  L  W  I  C  A  S  C
Seq_1  660  GAGCCGGATTCTGCAGGAGCTGGGCCTGAACCAGGAACAGTTTGTGGATCTGTGCATCCT  719
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Seq_2  627  gagccggattctgcaggagctgggcctgaaccaggaacagtttgtggatctgtgcatcct  686
                    S  R  I  L  Q  E  L  G  L  N  Q  E  Q  F  V  D  L  C  I  L


                    *  A  V  T  T  V  R  V  S  G  V  L  G  P  S  G  L  W  T
Seq_1  720  GCTAGGCAGTGACTACTGTGAGAGTATCCGGGGTATTGGGCCCAAGCGGGCTGTGGAC-T  778
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||#|
Seq_2  687  gctaggcagtgactactgtgagagtatccggggtattgggcccaagcgggctgtggacct  746
                    L  G  S  D  Y  C  E  S  I  R  G  I  G  P  K  R  A  V  D  L


                    H  P  E  A  Q  E  H  R  G   I  V  R  R  L  D  P  T    S  T  L
Seq_1  779  CATCCAGAAGCACAAGAGCATCGAGGA-ATCGTGCGGCGACTTGACCCCA-CAAGTACCC  836
            |||||||||||||||||||||||||#|||||||||||||||||||||||#|||||||||
Seq_2  747  catccagaagcacaagagcatcgaggagatcgtgcggcgacttgaccccaacaagtaccc  806
                    I  Q  K  H  K  S  I  E  E  I  V  R  R  L  D  P  N  K  Y  P


                    C  Q  K  I  G  S  T   G  G  S  P  A  L  L  G
Seq_1  837  TGTGCCAGAAAATTGGCTCCACA-GGAGGCTCACCAGCTCTTCTTGG-------------  882
            ||||||||||||||||||||||||#||||||||||||||||||||||#############
Seq_2  807  tgtgccagaaaattggctccacaaggaggctcaccagctcttcttggaacctgaggtgct  866
                    V  P  E  N  W  L  H  K  E  A  H  Q  L  F  L  E  P  E  V  L


                          *  G  W  T  I  W  A  E  W  R  A  N  E  K  L  I  K  F
Seq_1  883  -----CT-GAGGGTGGACCATCTGGGCTGAATGGCGAGCCAATGAAAAACTGATCAAGTT  936
            #####|#########|####|###|########||#|####|#||#|#||||||||||||
Seq_2  867  ggacccagagtctgtggagctgaagtggagcgagccaaatgaagaagagctgatcaagtt  926
                    D  P  E  S  V  E  L  K  W  S  E  P  N  E  E  E  L  I  K  F


                    C  V  G  K  A  L  R  S  Q  G  H  D  N  P  A  I   H  V  Y  A
Seq_1  937  CTGTGTGGGAAAAGCTCTGCGATCCCAGGGGCATGATAACCCCGCGAT-TCACGTCTACG  995
            |######||##||###|#|###||##|||#||####|###|##|#|###|||#|#####|
Seq_2  927  catgtgtggtgaaaagcagttctctgaggagcgaa-tccgca-gtggggtcaagaggctg  984
                    M  C  G  E  K  Q  F  S  E  E  R  I  R  S  G  V  K  R  L


                    K  G  G  X
Seq_1  996  CCAAGGGGG-GTA-----------------------------------------------  1007
            ###|#|#|##|###|##############################################
Seq_2  985  agtaagagccgccaaggcagcacccagggccgcctggatgatttcttcaaggtgaccggc  1044
                    S  K  S  R  Q  G  S  T  Q  G  R  L  D  D  F  F  K  V  T  G


Seq_1  1008 ------------------------------------------------------------  1007
            ############################################################
Seq_2  1045 tcactctcttcagctaagcgcaaggagccagaacccaagggatccactaagaagaaggca  1104
                    S  L  S  S  A  K  R  K  E  P  E  P  K  G  S  T  K  K  K  A


                                                                         X
Seq_1  1008 -----------------------------------------------------  1007
            #####################################################
Seq_2  1105 aagactggggcagcagggaagtttaaaagggggaaaaCATCATCATCATCATCACTAA  1161
                    K  T  G  A  A  G  K  F  K  R  G  K  H  H  H  H  H  H  *
```

243

sequence 1: K93AR100A reverse
sequence 2: hFEN1 mRNA

Alignment of Sequence_1: [046_df_28b-DM_T7R.seq.xdna] with Sequence_2: [hFEN1_mRNA.xdna]

Similarity : 888/1161 (76.49 %)


```
Seq_1  1009  ------------------------------------------------------------  1010
             ############################################################
Seq_2  1     atgggaattcaaggcctggccaaactaattgctgatgtggcccccagtgccatccgggag  60
             M  G  I  Q  G  L  A  K  L  I  A  D  V  A  P  S  A  I  R  E


Seq_1  1009  ------------------------------------------------------------  1010
             ############################################################
Seq_2  61    aatgacatcaagagctactttggccgtaaggtggccattgatgcctctatgagcatttat  120
             N  D  I  K  S  Y  F  G  R  K  V  A  I  D  A  S  M  S  I  Y


Seq_1  1009  ------------------------------------------------------------  1010
             ############################################################
Seq_2  121   cagttcctgattgctgttcgccagggtggggatgtgctgcagaatgaggagggtgagacc  180
             Q  F  L  I  A  V  R  Q  G  G  D  V  L  Q  N  E  E  G  E  T


                                      H  P  H  *  L  R  V  G  V  Q  D  Q  A
Seq_1  1009  ----------------------CATCCCCATTAGCTTCGTGTGGGGGTGCAGGACCAGG  973
             ####################|#|#||#######|########||#####||######
Seq_2  181   accagccacctgatgggcatgttc-taccgcaccattcgcatgatggagaacggcatcaa  239
             T  S  H  L  M  G  M  F  Y  R  T  I  R  M  M  E  N  G  I  K


             P  C  M  R  A  H  V  S  *  A  A  A  A  Q  A  S  C  N
Seq_1  972   CTCCCTGCATGAGAGCTCACGTATCTTGAGCAGCAGCTGCTCAGGCGAGCTGCAA-----  918
             ##||#||#|||||#########|#|####|########|###|||||||||||########
Seq_2  240   gcccgtgtatgtctttgatggcaagccgccacagctcaagtcaggcgagctggccaaacg  299
             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  E  A  E  K  Q  L  Q  Q  A  Q  L  L  G  P  T
Seq_1  917   CAGTGAGCGGCGGGCTGAGGCAGAGAAGCAGCTGCAGCAGGCTCAG-CTGCTGGGGCCGA  859
             |||||||||||||||||||||||||||||||||||||||||#||||||||||||||||
Seq_2  300   cagtgagcggcgggctgaggcagagaagcagctgcagcaggctcaggctgctggggccga  359
             S  E  R  R  A  E  A  E  K  Q  L  Q  Q  A  Q  A  A  G  A  E


             G  G  G  K  I  H  K  R  L  V  K  V  I  S  S  T  M  M  S
Seq_1  858   -CAGGAGGTGGAAAAATTCA-TAAGCGGCTGGTGAAGGTCA-TAAGCAGCACAATGATGA  802
             #||||||||||||||||||#|||||||||||||||||||#|||||||||||||||||
Seq_2  360   gcaggaggtggaaaaattcactaagcggctggtgaaggtcactaagcagcacaatgatga  419
             Q  E  V  E  K  F  T  K  R  L  V  K  V  T  K  Q  H  N  D  E


             A  N  I  A  E  L  M  G  I  P  Y  L  D  A  P  S  E  A  E
Seq_1  801   GTGCAAACAT-TGCTGAG-CTCATGGGCATCCCTTATCTTGATGCACCCAGTGAGGCAGA  744
             ||||||||#||||||||#||||||||||||||||||||||||||||||||||||
Seq_2  420   gtgcaaacatctgctgagcctcatgggcatcccttatcttgatgcacccagtgaggcaga  479
             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  L  V  K  A  G  K  V  Y  A  A  A  T  E  D  M
Seq_1  743   GGCCAGCTGTGCTGCCCTGGTGAAGGCTGGCAAAGTCTATGCTGCGGCTACCGAGGACAT  684
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Seq_2  480   ggccagctgtgctgccctggtgaaggctggcaaagtctatgctgcggctaccgaggacat  539
             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  G  S  P  V  L  M  R  H  L  T  A  S  E  A  K
Seq_1  683   GGACTGCCTCACCTTCGGCAGCCCTGTGCTAATGCGACACCTGACTGCCAGTGAAGCCAA  624
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Seq_2  540   ggactgcctcaccttcggcagccctgtgctaatgcgacacctgactgccagtgaagccaa  599
             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  I  L  Q  E  L  G  L  N  Q
Seq_1  623   AAAGCTGCCAATCCAGGAATTCCACCTGAGCCGGATTCTGCAGGAGCTGGGCCTGAACCA  564
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Seq_2  600   aaagctgccaatccaggaattccacctgagccggattctgcaggagctgggcctgaacca  659
```

244

```
                    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   L   C   I   L   L   G   S   D   Y   C   E   S   I   R   G
Seq_1   563   GGAACAGTTTGTGGATCTGTGCATCCTGCTAGGCAGTGACTACTGTGAGAGTATCCGGGG   504
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Seq_2   660   ggaacagtttgtggatctgtgcatcctgctaggcagtgactactgtgagagtatccgggg   719
                    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   K   H   K   S   I   E   E   I   V
Seq_1   503   TATTGGGCCCAAGCGGGCTGTGGACCTCATCCAGAAGCACAAGAGCATCGAGGAGATCGT   444
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Seq_2   720   tattgggcccaagcgggctgtggacctcatccagaagcacaagagcatcgaggagatcgt   779
                    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   E   N   W   L   H   K   E   A   H
Seq_1   443   GCGGCGACTTGACCCCAACAAGTACCCTGTGCCAGAAAATTGGCTCCACAAGGAGGCTCA   384
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Seq_2   780   gcggcgacttgaccccaacaagtaccctgtgccagaaaattggctccacaaggaggctca   839
                    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   E   S   V   E   L   K   W   S   E
Seq_1   383   CCAGCTCTTCTTGGAACCTGAGGTGCTGGACCCAGAGTCTGTGGAGCTGAAGTGGAGCGA   324
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Seq_2   840   ccagctcttcttggaacctgaggtgctggacccagagtctgtggagctgaagtggagcga   899
                    Q   L   F   L   E   P   E   V   L   D   P   E   S   V   E   L   K   W   S   E


                    P   N   E   E   L   I   K   F   M   C   G   E   K   Q   F   S   E   E   R
Seq_1   323   GCCAAATGAAGAAGAGCTGATCAAGTTCATGTGTGGTGAAAAGCAGTTCTCTGAGGAGCG   264
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Seq_2   900   gccaaatgaagaagagctgatcaagttcatgtgtggtgaaaagcagttctctgaggagcg   959
                    P   N   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   R   Q   G   S   T   Q   G   R   L
Seq_1   263   AATCCGCAGTGGGGTCAAGAGGCTGAGTAAGAGCCGCCAAGGCAGCACCCAGGGCCGCCT   204
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Seq_2   960   aatccgcagtggggtcaagaggctgagtaagagccgccaaggcagcacccagggccgcct   1019
                    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   S   A   K   R   K   E   P   E   P
Seq_1   203   GGATGATTTCTTCAAGGTGACCGGCTCACTCTCTTCAGCTAAGCGCAAGGAGCCAGAACC   144
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Seq_2   1020  ggatgatttcttcaaggtgaccggctcactctcttcagctaagcgcaaggagccagaacc   1079
                    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   A   A   G   K   F   K   R   G   K
Seq_1   143   CAAGGGATCCACTAAGAAGAAGGCAAAGACTGGGGCAGCAGGGAAGTTTAAAAGGGGAAA   84
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Seq_2   1080  caagggatccactaagaagaaggcaaagactggggcagcagggaagtttaaaaggggaaa   1139
                    K   G   S   T   K   K   K   A   K   T   G   A   A   G   K   F   K   R   G   K


                    H   H   H   H   H   H   *   A   C   G   R   T   R   A   P   P   P   P   P   R
Seq_1   83    ACATCATCATCATCATCACTAAGCTTGCGGCCGCACTCGAGCACCACCACCACCACCACG   24
              |||||||||||||||||||||||############################
Seq_2   1140  aCATCATCATCATCATCACTAA---------------------------------------   1161
                    H   H   H   H   H   H   *


                    D   R   R   *   Q   S   P   X
Seq_1   23    AGACCGGCGCTAACAAAGCCCGA   1
              #######################
Seq_2   1162  ---------------------   1161
                                X   X
```

Figure A1: Sequence alignment of hFEN1K93AR100A and hFEN1 mRNA: *Both T7 forward and reverse sequencing reactions were performed by the University of Sheffield Medical Scholl and subsequent alignment of the ORF of recombinant plasmid was done using Serial Cloner 2.1 (Serial Basics, Freeware). Mutation sites are highlighted (yellow) and nucleotides that could not be properly determined are denoted with a *.The histidine tag used in affinity chromatography is highlighted (red).*
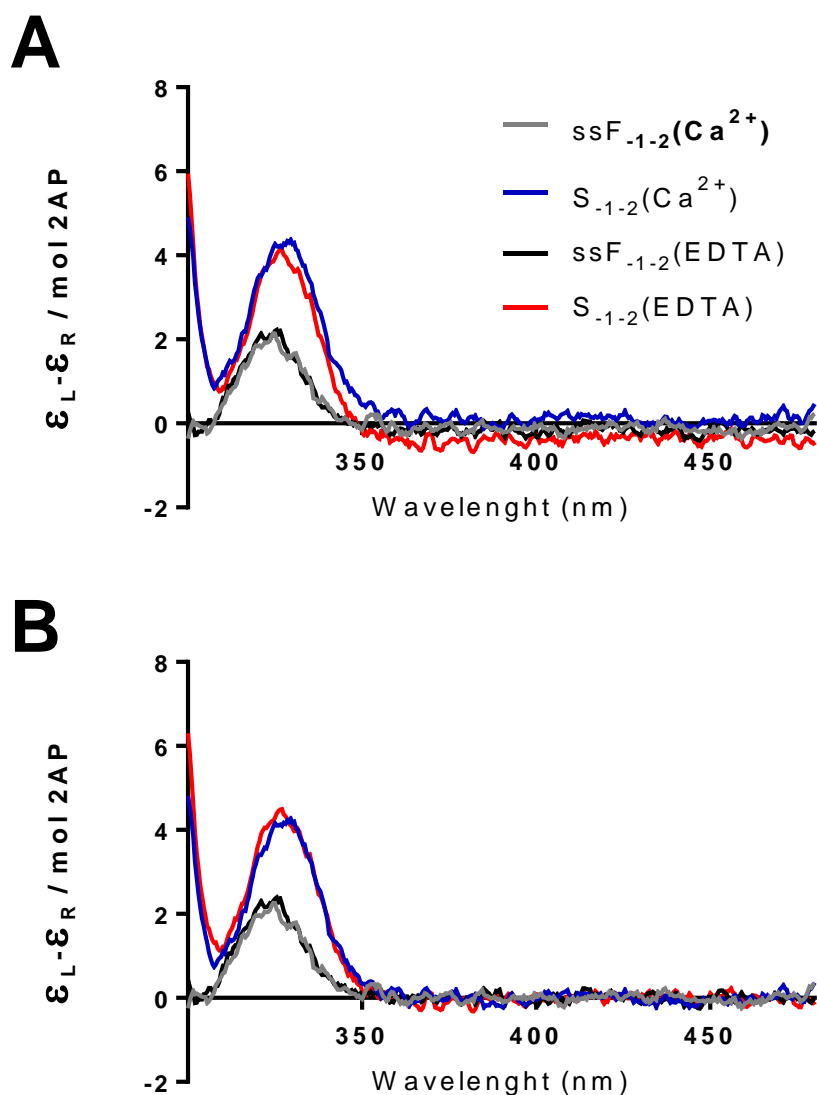
Figure A2: Low-energy CD spectra of unbound ss and ds 2-AP containing oligonucleotides demonstrating that the spectroscopic behaviour is identical in the presence and absence of Ca²⁺ ions and there is no absorption by 2-AP above 360 nm. (A) *Raw molar ellipticity changes from 300 – 480 nm for unbound S₋₁₋₂ (blue), the corresponding single strand (F₋₁₋₂, gray) in buffer containing 10 mM Ca²⁺ and unbound S₋₁₋₂ (red), the corresponding single strand (F₋₁₋₂, black) in buffer containing 2 mM EDTA .***(B) Demonstrating the requirement to normalize the signal due to base-line shift**. *The same spectra as in (A) but have been normalized by averaging the signal between 400 -480 nm and subtracting the average value from the raw data.*
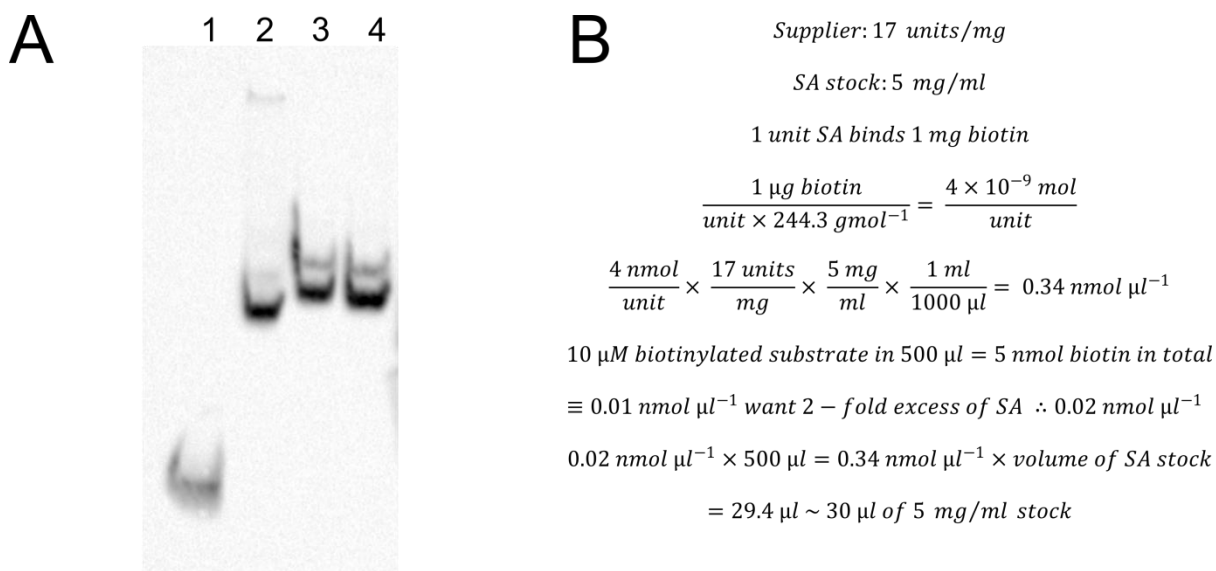
Figure A3: Streptavidin gel-shift and calculation (**A**) EMSA of 10 nM DF-B10 (lane 1) and, 10 nM DF-B10 with 5- (lane 2), 2- (lane 3), and 1-molar equivalents of SA (lane 4). The free and conjugated substrate were visualised using chemiluminescence as described in section 2.5. (B) Example calculation of SA stock (5 mg/ml) required for 2-molar excess in a 500 µl, 10 µM biotinylated substrate sample.

The calculation in panel B:

$$Supplier: 17\ units/mg$$

$$SA\ stock: 5\ mg/ml$$

$$1\ unit\ SA\ binds\ 1\ mg\ biotin$$

$$\frac{1\ \mu g\ biotin}{unit \times 244.3\ gmol^{-1}} = \frac{4 \times 10^{-9}\ mol}{unit}$$

$$\frac{4\ nmol}{unit} \times \frac{17\ units}{mg} \times \frac{5\ mg}{ml} \times \frac{1\ ml}{1000\ \mu l} = 0.34\ nmol\ \mu l^{-1}$$

$$10\ \mu M\ biotinylated\ substrate\ in\ 500\ \mu l = 5\ nmol\ biotin\ in\ total$$

$$\equiv 0.01\ nmol\ \mu l^{-1}\ want\ 2-fold\ excess\ of\ SA \therefore 0.02\ nmol\ \mu l^{-1}$$

$$0.02\ nmol\ \mu l^{-1} \times 500\ \mu l = 0.34\ nmol\ \mu l^{-1} \times volume\ of\ SA\ stock$$

$$= 29.4\ \mu l \sim 30\ \mu l\ of\ 5\ mg/ml\ stock$$
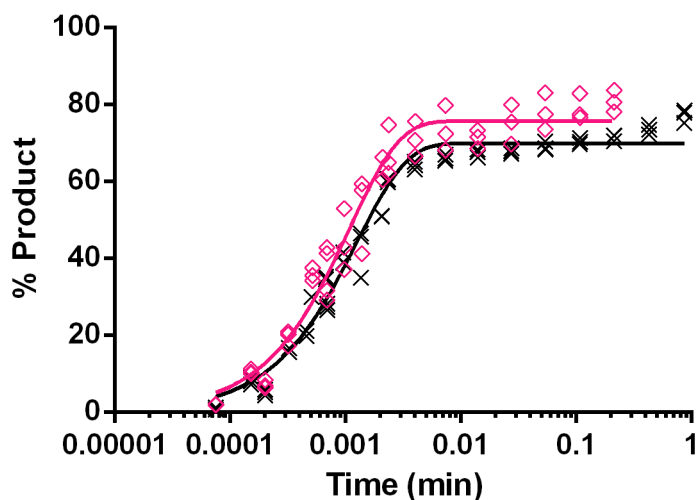
Figure A4: The single turnover rate profiles for DF-5* cleavage by hFEN1 at 10 times K_D and 10 times K_M. Plot of percentage product at 1000 nM [E] (open diamonds magenta) and 200 nM [E] (crosses black) versus time. The black and magenta solid line represent lines of best fit to a single exponential equation and give k_{ST} values of $918 \pm 50\ min^{-1}$ and $761 \pm 40\ min^{-1}$, respectively. The final substrate concentration was 2.5 nM in both experiments. Reactions were performed in 8 mM MgCl$_2$, 110 mM KCl, 55 mM HEPES-NaOH pH = 7.5, 1 mM DTT and 0.1 µg/µl BSA and at 37 ˚C.

247
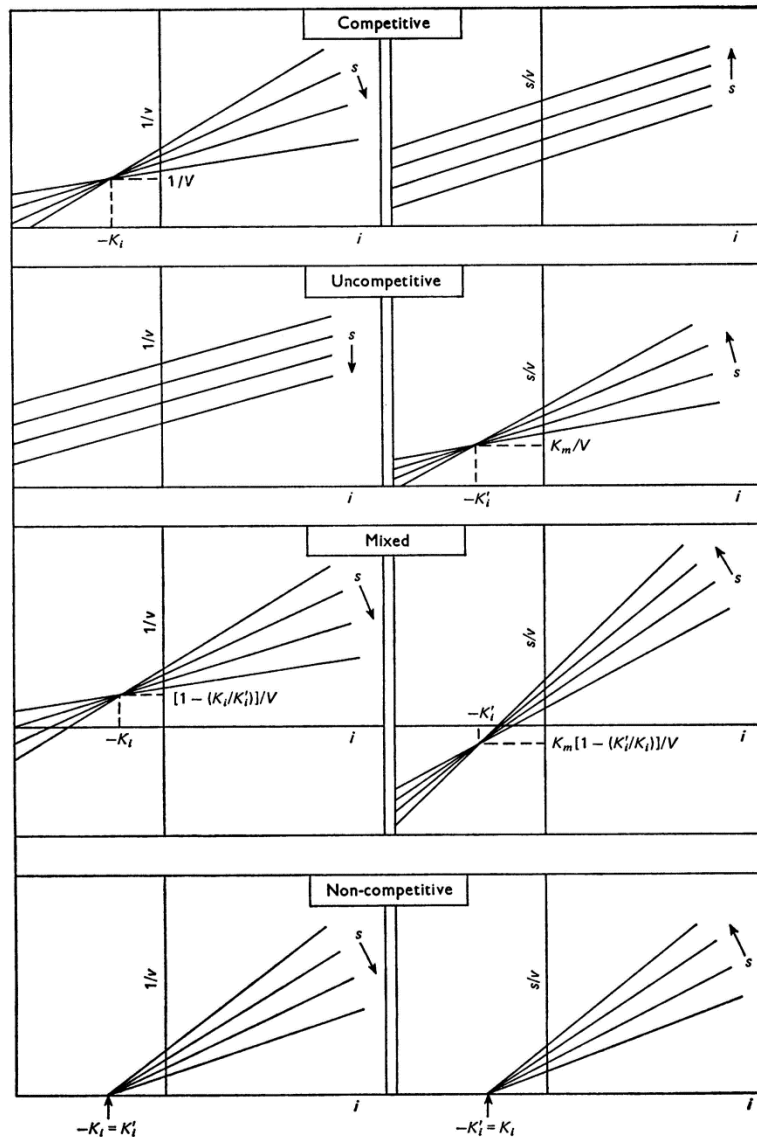
Figure A5: Representative plots of the most common types of linear inhibition. *(left) Plots of 1/v against i, intersect K_{iu} (Dixon plots) and (right) s/v against i, intersect K_{ic}, for the inhibition models indicated. Taken from Cornish-Boden, A. Biochem J. (1974) 137, 143–144.*

F

G
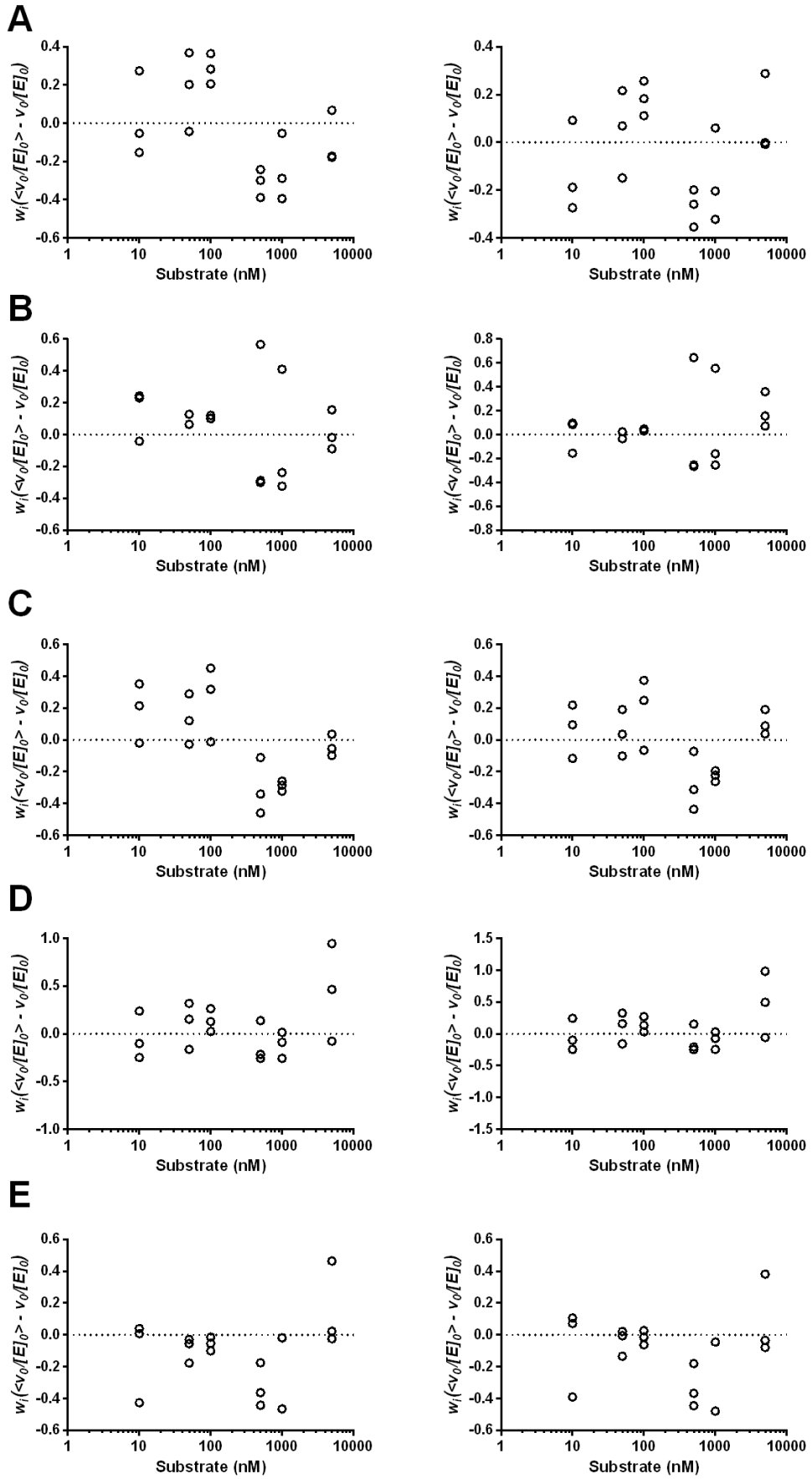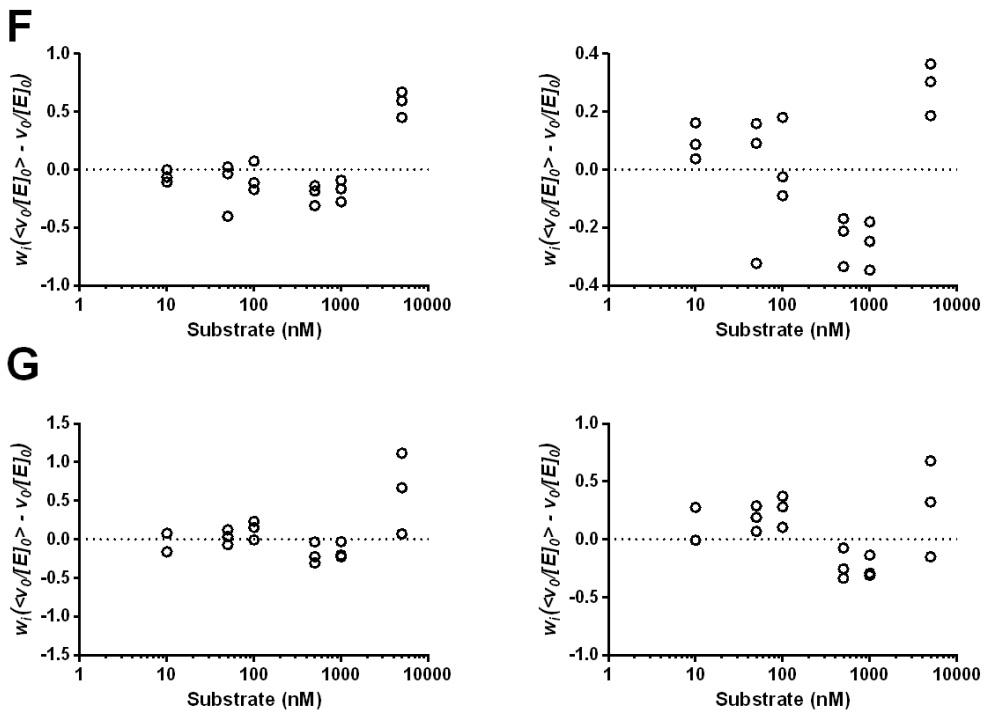
Figure A6: Residuals for inhibitor 1 kinetic data fit to non-competitive inhibition model (left) and mixed inhibition model (right), at (A) 5 nM, (B) 10 nM, (C) 50 nM, (D) 100 nM, (E) 500 nM, (F) 1000 nM, and (G) 5000 nM [inhibitor **1**].

A

B

C
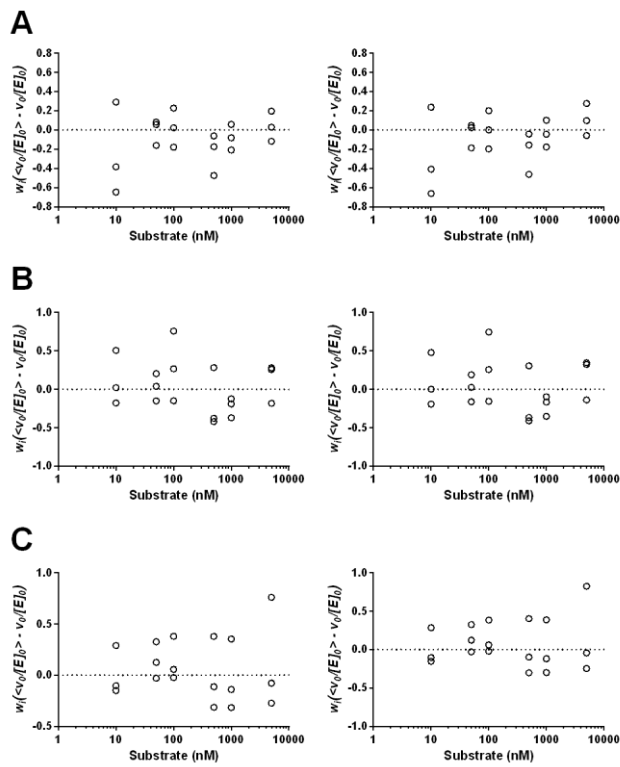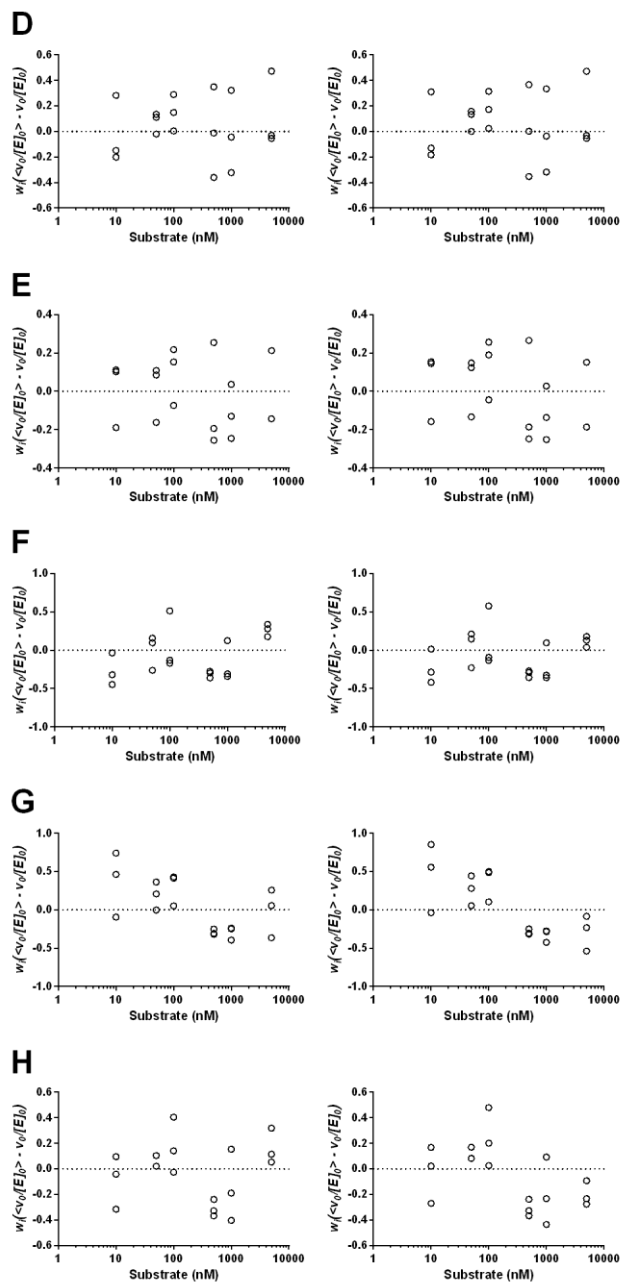
Figure A7: Residuals for inhibitor 2 kinetic data fit to competitive inhibition model (left) and mixed inhibition model (right), *at (A) 5 nM, (B) 10 nM, (C) 25 nM, (D) 50 nM, (E) 100 nM, (F) 500 nM, (G) 1000 nM, and (H) 5000 nM [inhibitor **2**].*

A) $K_{STmax} = 0.087 \pm 3\text{X}10^{-3}$ min^{-1}

B) $K_{STmax} = 0.11 \pm 3\text{X}10^{-3}$ min^{-1}
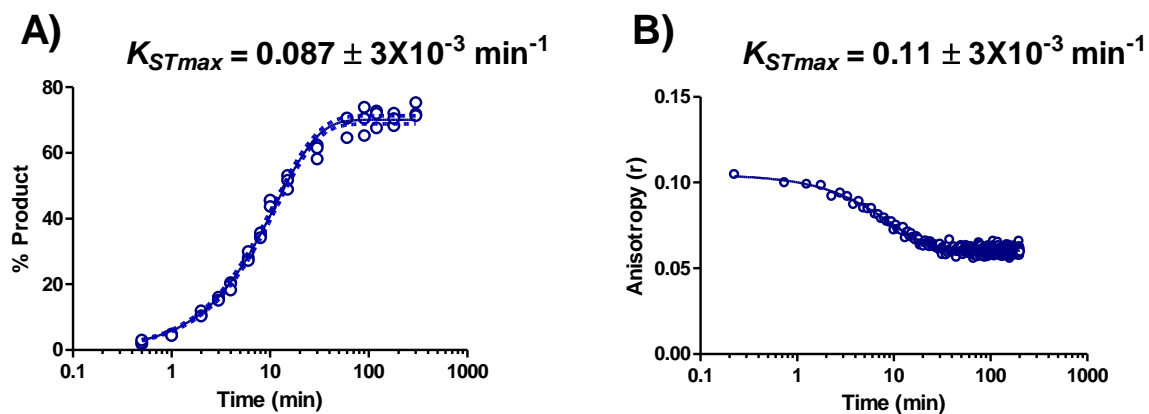
Figure A8: The single turnover rate profiles for DF-5 cleavage by hFEN1R100A determined by (A) discontinuous assay, (B) fluorescence anisotropy. Both were fit to single-phase exponential to yield single-turnover rate constants shown above the respective graph. Reactions were performed in 8mM MgCl$_2$, 110mM KCl, 55mM HEPES$^{pH=7.5}$, 1mM DTT and 0.1µg/µl BSA and at 37 ^{0}C. Final R100A/DF-5 concentrations were 1000/5 nM, for the discontinuous assay and 1000/10 nM when measured by fluorescence anisotropy, respectively.
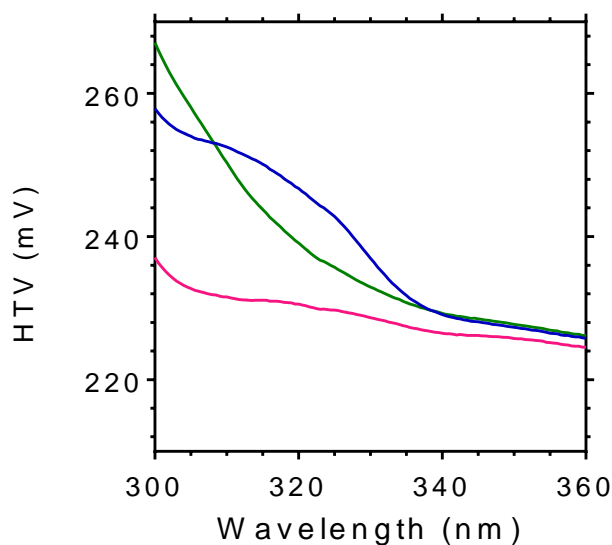
Figure A9: High-tension voltage (HTV) trace of ECCD samples in absence and presence of inhibitors. The total absorbance of the sample (protein and buffer) can be monitored by High Tension Voltage (HTV) trace. To avoid excessive noise and demonstrate reliable absorbance data, the HTV trace should be less than 700 mV. Samples contained 12.5 µM R100A, 110 mM KCl, 55 mM HEPES-NaOH pH=7.5, 1 mM DTT and 10 mM CaCl$_2$ (magenta), 10 mM MgCl$_2$ + 100 µM inhibitor 1 (green), and 10 mM MgCl$_2$ + 100 µM inhibitor 2 (blue). These samples were used as blanks for the appropriate samples containing 10 µM DNA.

```
                                                                  M  G  I  P  K  F
Seq_1    1    -----------------------------------------------atgggtattccaaaat   16
                                                             ||||||||||||||||
Seq_2    1    ttCCcTcTanaataTTTtGTTtaCTTtAaGAaGGAGATATACATATGGGTATTCCAAAAT   60
              P   S   X   I   F   C   L   L   *   E   G   D   I  H  M  G  I  P  K  F


              F  H  F  I  S  E  R  W  P  Q  I  S  Q  L  I  D  G  S  Q  I
Seq_1   17    tctttcatttcatctctgagagatggcctcaaatttctcaattgatcgatggatcacaga   76
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Seq_2   61    TCTTTCATTTCATCTCTGAGAGATGGCCTCAAATTTCTCAATTGATCGATGGATCACAGA   120
              F  H  F  I  S  E  R  W  P  Q  I  S  Q  L  I  D  G  S  Q  I


              P  E  F  D  N  L  Y  L  D  M  N  S  I  L  H  N  C  T  H  G
Seq_1   77    ttccagagttcgacaatttgtacttggatatgaattctattttgcataattgtacgcatg   136
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Seq_2  121    TTCCAGAGTTCGACAATTTGTACTTGGATATGAATTCTATTTTGCATAATTGTACGCATg   180
              P  E  F  D  N  L  Y  L  D  M  N  S  I  L  H  N  C  T  H  G


              D  G  S  E  V  N  S  R  L  S  E  E  E  V  Y  S  K  I  F  S
Seq_1  137    gtgatggtagcgaggtgaattcaagactatcagaggaagaagtttattccaaaattttca   196
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Seq_2  181    GTGATGGTAGCGAGgtGAATTCAAGACTATCAGAGGAAGAAGTTTATTCCAAAATTTTCA   240
              D  G  S  E  V  N  S  R  L  S  E  E  E  V  Y  S  K  I  F  S


              Y  I  D  H  L  F  H  T  I  K  P  K  Q  T  F  Y  M  A  I  D
Seq_1  197    gttatattgatcatcttttccatactattaaaccaaaacagacattttatatggcaatcg   256
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Seq_2  241    .GTTATATTGATCATCTTTTCCATACTATTAAACCAAAACAGACATTTTATATGGCAATCG   300
              Y  I  D  H  L  F  H  T  I  K  P  K  Q  T  F  Y  M  A  I  D


              G  V  A  P  R  A  K  M  N  Q  Q  R  A  R  R  F  R  T  A  M
Seq_1  257    atggtgtggccccaagagcaaagatgaaccaacaaagagctcgtagattcagaactgcca   316
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Seq_2  301    ATGGTGTGGCCCCAAGAGCAAAGATGAACCAACAAAGAGCTCGTAGATTCAGAACTGCCA   360
              G  V  A  P  R  A  K  M  N  Q  Q  R  A  R  R  F  R  T  A  M


              D  A  E  K  A  L  Q  K  A  I  E  N  G  D  E  L  P  K  G  E
Seq_1  317    tggatgccgaaaaggctttgcagaaggccattgaaaatggtgacgagttgcctaagggag   376
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Seq_2  361    TGGATGCCGAAAAGGCTTTGCAGAAGGCCATTGAAAATGGTGACGAGTTGCCTAAGGGAG   420
              D  A  E  K  A  L  Q  K  A  I  E  N  G  D  E  L  P  K  G  E


              P  F  D  S  N  A  I  T  P  G  T  E  F  M  A  K  L  T  E  N
Seq_1  377    agccatttgattccaacgctattaccccaggaacagaatttatggcaaaattaaccgaga   436
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Seq_2  421    AGCCATTTGATTCCAACGCTATTACCCCAGGAACAGAATTTATGGCAAAATTAACCGAGA   480
              P  F  D  S  N  A  I  T  P  G  T  E  F  M  A  K  L  T  E  N


              L  K  Y  F  I  H  D  K  I  T  N  D  T  R  W  Q  N  V  K  V
Seq_1  437    atttgaaatatttcatccatgacaagatcaccaacgataccagatggcagaacgtgaagg   496
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Seq_2  481    ATTTGAAATATTTCATCCATGACAAGATCACCAACGATACCAGATGGCAGAACGTGAAGG   540
              L  K  Y  F  I  H  D  K  I  T  N  D  T  R  W  Q  N  V  K  V


              I  F  S  G  H  E  V  P  G  E  G  E  H  K  I  M  D  Y  I  R
Seq_1  497    ttattttctccgggcatgaggttcctggtgaaggtgaacataagatcatggattacatca   556
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Seq_2  541    TTATTTTCTCCGGGCATGAGGTTCCTGGTGAAGGTGAACATAAGATCATGGATTACATCA   600
              I  F  S  G  H  E  V  P  G  E  G  E  H  K  I  M  D  Y  I  R


              A  I  R  A  Q  E  D  Y  N  P  N  T  R  H  C  I  Y  G  L  D
Seq_1  557    gagcaattagagcacaagaggattacaatccaaatacaagacattgtatctacgggttag   616
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Seq_2  601    GAGCAATTAGAGCACAAGAGGATTACAATCCAAATACAAGACATTGTATCTACGGGTTAG   660
              A  I  R  A  Q  E  D  Y  N  P  N  T  R  H  C  I  Y  G  L  D


              A  D  L  I  I  L  G  L  S  T  H  D  H  H  F  C  L  L  R  E
Seq_1  617    atgctgatttgatcatcctaggtttatccacccatgatcaccactttgtttattaagag   676
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Seq_2  661    ATGCTGATTTGATCATCCTAGGTTTATCCACCCATGATCACCACTTTTGTTTATTAAGAG   720
```

253

```
          E   V   T   F   G   K   R   S   S   S   V   K   T   L   E   T   Q   N   F   F
Seq_1 677 aagaagttacttttggtaaacgttcgtcttctgtgaaaactctagaaacacagaacttct  736
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Seq_2 721 AAGAAGTTACTTTTGGTAAACGTTCGTCTTCTGTGAAAACTCTAGAAACACAGAACTTCT  780
          E   V   T   F   G   K   R   S   S   S   V   K   T   L   E   T   Q   N   F   F


          L   L   H   L   S   I   L   R   E   Y   L   A   L   E   F   E   E   I   T   D
Seq_1 737 tcttgttgcatttgtctatcttgagagaatatttggcattagagttcgaagaaataacag  796
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Seq_2 781 TCTTGTTGCATTTGTCTATCTTGAGAGAATATTTGGCATTAGAGTTCGAAGAAATAACAG  840
          L   L   H   L   S   I   L   R   E   Y   L   A   L   E   F   E   E   I   T   D


          S   V   Q   F   E   Y   D   F   E   R   V   L   D   D   F   I   F   V   L   F
Seq_1 797 attctgtgcagtttgaatacgactttgaaagagtattagatgatttcatctttgtattat  856
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Seq_2 841 ATTCTGTGCAGTTTGAATACGACTTTGAAAGAGTATTAGATGATTTCATCTTTGTATTAT  900
          S   V   Q   F   E   Y   D   F   E   R   V   L   D   D   F   I   F   V   L   F


          F   I   F   V   L   F   T   I   G   N   D   F   L   P   N   L   P   D   L   H
Seq_1 841 ttcatctttgtattatttaccatcggtaatgatttcttaccaaatttgcccgatttgcat  900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Seq_2 176 TTCATCTTtgTATTATTTACCATCGGTAATGATTTCTTACCAAATTTGCCCGATTTGCAT  235
          F   I   F   V   L   F   T   I   G   N   D   F   L   P   N   L   P   D   L   H


          L   K   K   G   A   F   P   V   L   L   Q   T   F   K   E   A   L   Q   H   M
Seq_1 901 ttgaaaaaaggtgcattccctgtgctattacaaacttttaaagaagctctccaacatatg  960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Seq_2 236 TTGAAAAAAGGTGCATTCCCTgtGCTATTACAAACTTTTAAAGAAGCTCTCCAACATATG  295
          L   K   K   G   A   F   P   V   L   L   Q   T   F   K   E   A   L   Q   H   M


          D   G   Y   I   N   E   Q   G   K   I   N   L   A   R   F   S   I   W   L   K
Seq_1 961 gatggttacattaatgaacaaggtaagataaatttggcaagattttccatttggttgaag  1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Seq_2 296 GATGGTTACATTAATGAACAAGGTAAGATAAATTTGGCAAGATTTTCCATTTGGTTGAAG  355
          D   G   Y   I   N   E   Q   G   K   I   N   L   A   R   F   S   I   W   L   K


          Y   L   S   D   F   E   Y   L   N   F   E   K   K   D   I   D   V   E   W   F
Seq_1 1021 tacttgtccgattttgaataccttaactttgagaagaaagatattgacgttgaatggttc  1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Seq_2 356 TACTTGTCCGATTTTGAATACCTTAACTTTGAGAAGAAAGATATTGACGTTGAATGGTTC  415
          Y   L   S   D   F   E   Y   L   N   F   E   K   K   D   I   D   V   E   W   F


          N   Q   Q   L   E   N   I   S   L   E   G   E   R   K   R   T   R   M   G   K
Seq_1 1081 aatcaacaacttgaaaatatttccttggaaggtgagcgtaaacgtactaggatgggtaaa  1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Seq_2 416 AATCAACAACTTGAAAATATTTCCTTGGAAGGTGAGCGTAAACGTACTAGGATGGGTAAA  475
          N   Q   Q   L   E   N   I   S   L   E   G   E   R   K   R   T   R   M   G   K


          K   L   L   M   K   Q   Q   K   K   L   I   G   A   V   K   P   W   L   L   K
Seq_1 1141 aagttgttgatgaaacaacaaaagaaattgattggcgccgtaaaaccatggttattgaaa  1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Seq_2 476 AAGTTGTTGATGAAACAACAAAAGAAATTGATTGGCGCCGTAAAAACCATGGTTATTGAAA  535
          K   L   L   M   K   Q   Q   K   K   L   I   G   A   V   K   P   W   L   L   K


          T   V   Q   R   K   V   T   S   E   L   Q   D   A   D   F   E   I   F   P   L
Seq_1 1201 accgttcaacggaaggtcacttctgaattacaggatgccgatttcgaaattttccctctt  1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Seq_2 536 ACCGTTCAACGGAAGGTCACTTCTGAATTACAGGATGCCGATTTCGAAATTTTCCCTCTT  595
          T   V   Q   R   K   V   T   S   E   L   Q   D   A   D   F   E   I   F   P   L


          E   D   K   E   L   V   R   A   N   L   D   F   L   K   E   F   A   F   D   L
Seq_1 1261 gaggataaagaattggttcgggccaacctggatttcttgaaggaattcgcatttgatttg  1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Seq_2 596 GAGGATAAAGAATTGGTTCGGGCCAACCTGGATTTCTTGAAGGAATTCGCATTTGATTTG  655
          E   D   K   E   L   V   R   A   N   L   D   F   L   K   E   F   A   F   D   L


          G   L   I   L   A   H   S   K   S   K   D   L   Y   Y   F   K   L   D   L   D
Seq_1 1321 ggtttaattcttgctcattctaaatcaaaagatttgtactacttcaagttggatttagac  1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Seq_2 656 GGTTTAATTCTTGCTCATTCTAAATCAAAAGATTTGTACTACTTCAAGTTGGATTTAGAC  715
          G   L   I   L   A   H   S   K   S   K   D   L   Y   Y   F   K   L   D   L   D
```

254

```
              S  I  N  V  Q  E  T  D  E  E  H  E  A  R  I  H  E  T  R  R
Seq_1  1381   tccatcaatgtgcaagaaacagatgaagagcatgaagcccgtattcacgaaacaagacgt   1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Seq_2   716   TCCATCAATGTGCAAGAAACAGATGAAGAGCATGAAGCCCGTATTCACGAAACAAGACGT   775
              S  I  N  V  Q  E  T  D  E  E  H  E  A  R  I  H  E  T  R  R


              S  I  K  K  Y  E  Q  G  I  I  I  A  S  E  E  E  L  E  E  E
Seq_1  1441   tctattaagaaatatgaacaaggtattattattgcatctgaagaagaattagaagaagag   1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Seq_2   776   TCTATTAAGAAATATGAACAAGGTATTATTATTGCATCTGAAGAAGAATTAGAAGAAGAG   835
              S  I  K  K  Y  E  Q  G  I  I  I  A  S  E  E  E  L  E  E  E


              R  E  I  Y  S  E  R  F  V  E  W  K  D  Q  Y  Y  K  D  K  L
Seq_1  1501   cgtgaaatttatagcgagagattcgttgagtggaaagatcaatactataaagataaatta   1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Seq_2   836   CGTGAAATTTATAGCGAGAGATTCGTTGAGTGGAAAGATCAATACTATAAAGATAAATTA   895
              R  E  I  Y  S  E  R  F  V  E  W  K  D  Q  Y  Y  K  D  K  L


              D  F  S  I  N  D  T  D  S  L  K  E  M  T  E  N  Y  V  G  G
Seq_1  1561   gattttttccatcaatgatactgatagtctaaaagaaatgacagaaaactatgtcggaggt   1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Seq_2   896   GATTTTTTCCATCAATGATACTGATAGTCTAAAAGAAATGACAGAAAACTATGTCGGaggT   955
              D  F  S  I  N  D  T  D  S  L  K  E  M  T  E  N  Y  V  G  G


              L  Q  W  V  L  Y  Y  Y  Y  R  G  C  P  S  W  S  W  Y  Y  R
Seq_1  1621   ttacaatgggttttatactactattatcgtggatgtccatcttggtcctggtattacaga   1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Seq_2   956   TTACAATGGGTTTTATACTACTATTATCGTGGATGTCCATCTTGGTCCTGGTaTTACAGA   1015
              L  Q  W  V  L  Y  Y  Y  Y  R  G  C  P  S  W  S  W  Y  Y  R


              Y  H  Y  A  P  R  I  S  D  V  I  K  G  I  D  Q  N  I  E  F
Seq_1  1681   tatcattacgctcctcgtatatcagatgtgatcaaaggtatagatcagaacatagaattt   1740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Seq_2  1016   TATCATTACGCTCCTCGTATATCAGATGTGATCAAaGGTATAGATCAGAACATAGAATTt   1075
              Y  H  Y  A  P  R  I  S  D  V  I  K  G  I  D  Q  N  I  E  F


              Y  H  Y  A  P  R  I  S  D  V  I  K  G  I  D  Q  N  I  E  F
Seq_1  1681   tatcattacgctcctcgtatatcagatgtgatcaaaggtatagatcagaacatagaattt   1740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Seq_2   395   TATCATTACGCTCCTCGTATATCAGATGTGATCAAAGGTATAGATCAGAACATAGAATTT   454
              Y  H  Y  A  P  R  I  S  D  V  I  K  G  I  D  Q  N  I  E  F


              H  K  G  Q  P  F  K  P  F  Q  Q  L  M  A  V  L  P  E  R  S
Seq_1  1741   cacaagggacagcctttcaagcccttccaacaattgatggccgtcttgcctgagagatca   1800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Seq_2   455   CACAAGGGACAGCCTTTCAAGCCCTTCCAACAATTGATGGCCGTCTTGCCTGAGAGATCA   514
              H  K  G  Q  P  F  K  P  F  Q  Q  L  M  A  V  L  P  E  R  S


              K  N  L  I  P  V  V  Y  R  P  L  M  Y  D  E  H  S  P  I  L
Seq_1  1801   aagaatttgattccagttgtgtacagaccactcatgtatgatgaacactctcctatcttg   1860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Seq_2   515   AAGAATTTGATTCCAGTTGTGTACAGACCACTCATGTATGATGAACACTCTCCTATCTTG   574
              K  N  L  I  P  V  V  Y  R  P  L  M  Y  D  E  H  S  P  I  L


              D  F  Y  P  N  E  V  E  L  D  L  N  G  K  T  A  D  W  E  A
Seq_1  1861   gacttctatcctaacgaagtagagttggatctaaatggcaaaacagcagactgggaagct   1920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Seq_2   575   GACTTCTATCCTAACGAAGTAGAGTTGGATCTAAATGGCAAAACAGCAGACTGGGAAGCT   634
              D  F  Y  P  N  E  V  E  L  D  L  N  G  K  T  A  D  W  E  A


              V  V  K  I  S  F  V  D  Q  K  R  L  V  E  A  M  A  P  Y  D
Seq_1  1921   gtcgtcaaaatttcattcgtggatcaaaagcgtttagtagaagcaatggctccttatgat   1980
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Seq_2   635   -GTCGTCAAAATTTCATTCGTGGATCAAAAGCGTTTAGTAGAAGCAATGGCTCCTTATGAT   694
              V  V  K  I  S  F  V  D  Q  K  R  L  V  E  A  M  A  P  Y  D


              A  K  L  S  P  D  E  K  K  R  N  S  F  G  T  D  L  I  F  I
Seq_1  1981   gctaagctttctccggatgaaaagaaaagaaattcgttcggaactgatctgattttcata   2040
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Seq_2   695   GCTAAGCTTTCTCCGGATGAAAAGAAAAGAAATTCGTTCGGAACTGATCTGATTTTCATA   754
              A  K  L  S  P  D  E  K  K  R  N  S  F  G  T  D  L  I  F  I
```

255

```
          F   N   P   Q   V   D   T   V   Y   K   T   P   L   A   G   L   F   N   D   I
Seq_1 2041 ttcaatcctcaggtagacacagtttacaaaacaccactagcggggttgtttaatgatatt 2100
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Seq_2  755 TTCAATCCTCAGGTAGACACAGTTTACAAAACACCACTAGCGGGGTTGTTTAATGATATT 814
          F   N   P   Q   V   D   T   V   Y   K   T   P   L   A   G   L   F   N   D   I


          E   H   N   H   C   I   E   R   E   F   I   P   E   S   M   E   N   V   K   F
Seq_1 2101 gaacacaatcattgtatcgaaagggaattcattccagaatcgatggaaaacgttaagttc 2160
           |||||||||||||||||||||||||||||||||||||||||||||||:||||||||||||
Seq_2  815 GAACACAATCATTGTATCGAAAGGGAATTCATTCCAGAATCGATGGAnAACGTTAAGTTC 874
          E   H   N   H   C   I   E   R   E   F   I   P   E   S   M   X   N   V   K   F


          L   F   G   L   P   K   G   A   K   L   G   A   S   S   L   A   G   F   P   S
Seq_1 2161 ttatttggtttgccaaagggcgctaaactcggagctagctctttggccggtttcccatcc 2220
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Seq_2  875 TTATTTGGTTTGCCAAAGGGCGCTAAACTCGGAGCTAGCTCTTTGGCCGGTTTCcCATCC 934
          L   F   G   L   P   K   G   A   K   L   G   A   S   S   L   A   G   F   P   S


          L   K   T   L   P   L   T   A   E   L   A   Y   N   S   S   V   V   F   N   F
Seq_1 2221 ctaaagacactaccactaactgcagagcttgcttataattcttctgttgtcttcaatttc 2280
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Seq_2  935 CTAAAGACACTACCACTAACTGCAGAGCTTGCTTATAATTCTTCTgttGTCTTCAATTTC 994
          L   K   T   L   P   L   T   A   E   L   A   Y   N   S   S   V   V   F   N   F


          P   S   K   Q   Q   S   M   V   L   H   I   Q   D   L   Y   K   E   N   G   I
Seq_1 2281 ccttctaaacaacaatctatggtgttgcatattcaggacctttacaaggaaaatggcatc 2340
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Seq_2   80 CCTTCtAAACAACAATCTATGGTGTTGCATATTCAGGACCTTTACAAGGAAAATGGCATC 139
          P   S   K   Q   Q   S   M   V   L   H   I   Q   D   L   Y   K   E   N   G   I


          S   L   S   D   L   A   K   R   H   M   G   K   I   V   Y   S   R   W   P   F
Seq_1 2341 tctctgtcagatctagcaaaaagacatatgggtaagattgtttattcaagatggccgttt 2400
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Seq_2  140 TCTCTGTCAGATCTAGCAAAAAGACATATGGGTAAGATTGTTTATTCAAGATGGCCGTTT 199
          S   L   S   D   L   A   K   R   H   M   G   K   I   V   Y   S   R   W   P   F


          L   R   E   S   K   L   L   S   L   I   T   E   E   T   V   Y   E   G   V   K
Seq_1 2401 ctaagagaatctaaacttttgtcgttgattacagaggaaactgtgtatgaaggagttaag 2460
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Seq_2  200 CTAAGAGAATCTAAACTTTTGTCGTTGATTACAGAGGAAACTGTGTATGAAGGAGTTAAG 259
          L   R   E   S   K   L   L   S   L   I   T   E   E   T   V   Y   E   G   V   K


          S   G   K   L   T   K   V   I   E   R   K   P   Q   D   F   E   R   K   E   F
Seq_1 2461 tcaggcaaattaacaaaggtcattgaaagaaagcctcaggattttgaaaggaaagaattt 2520
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Seq_2  260 TCAGGCAAATTAACAAAGGTCATTGAAAGAAAGCCTCAGGATTTTGAAAGGAAAGAATTT 319
          S   G   K   L   T   K   V   I   E   R   K   P   Q   D   F   E   R   K   E   F


          R   E   L   K   M   T   L   K   S   N   Y   Q   R   T   K   A   I   L   L   D
Seq_1 2521 agagagttgaagatgactctcaaatcgaattatcaaaggacaaaggccattcttttggat 2580
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Seq_2  320 AGAGAGTTGAAGATGACTCTCAAATCGAATTATCAAAGGACAAAGGCCATTCTTTTGGAT 379
          R   E   L   K   M   T   L   K   S   N   Y   Q   R   T   K   A   I   L   L   D


          D   I   S   A   L   A   K   V   V   P   V   N   G   L   V   R   N   S   D   G
Seq_1 2581 gacatttctgctttggctaaagtggttccagtaaatggattggtgagaaactctgatggt 2640
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Seq_2  380 GACATTTCTGCTTTGGCTAAAGTGGTTCCAGTAAATGGATTGGTGAGAAACTCTGATGGT 439
          D   I   S   A   L   A   K   V   V   P   V   N   G   L   V   R   N   S   D   G


          S   Y   S   K   S   F   N   E   T   I   E   Y   Y   P   L   Q   L   I   V   E
Seq_1 2641 tcctattctaagtctttcaatgaaactatcgaatactatcctttacaattaatcgtagag 2700
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Seq_2  440 TCCTATTCTAAGTCTTTCAATGAAACTATCGAATACTATCCTTTACAATTAATCGTAGAG 499
          S   Y   S   K   S   F   N   E   T   I   E   Y   Y   P   L   Q   L   I   V   E


          D   V   K   N   K   D   E   R   Y   I   E   K   E   P   L   P   I   N   K   E
Seq_1 2701 gacgtaaaaaataaggacgaaagatatattgaaaaagagccgttaccaattaataaagaa 2760
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Seq_2  500 GACGTAAAAAATAAGGACGAAAGATATATTGAAAAAGAGCCGTTACCAATTAATAAAGAA 559
          D   V   K   N   K   D   E   R   Y   I   E   K   E   P   L   P   I   N   K   E
```

256

```
                 F   P   K   G   S   K   V   V   F   L   G   D   Y   A   Y   G   G   E   A   T
Seq_1  2761  tttccaaaaggttcaaaagttgtgttttaggtgattatgcgtatggtggggaggcgacc   2820
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Seq_2  560   TTTCCAAAAGGTTCAAAAGTTGTGTTTTTAGGTGATTATGCGTATGGTGGGGAGGCGACC   619
                 F   P   K   G   S   K   V   V   F   L   G   D   Y   A   Y   G   G   E   A   T


                 V   D   G   Y   N   S   E   T   R   L   K   L   T   V   K   K   G   S   L   R
Seq_1  2821  gttgatggttataatagtgagactagattaaaacttacagtcaaaaagggttctctcaga   2880
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Seq_2  620   GTTGATGGTTATAATAGTGAGACTAGATTAAAACTTACAGTCAAAAAGGGTTCTCTCAGA   679
                 V   D   G   Y   N   S   E   T   R   L   K   L   T   V   K   K   G   S   L   R


                 A   E   P   N   I   G   K   V   R   A   K   L   D   S   Q   A   L   R   F   Y
Seq_1  2881  gcagagcctaacatcggaaaagttagagcgaaattggattctcaagccttgagattctac   2940
             |||||||||||||||||||||:||||||||||||||||||||||||||||||||||||||
Seq_2  680   GCAGAGCCTAACATCGGAnAAGTTAGAGCGAAATTGGATTCTCAAGCCTTGAGATTCTAC   739
                 A   E   P   N   I   G   X   V   R   A   K   L   D   S   Q   A   L   R   F   Y


                 P   T   Q   V   F   S   K   I   A   R   V   H   P   L   F   L   S   K   I   T
Seq_1  2941  ccaacacaagtgttttcaaaaatagctcgtgtccaccctctcttcttgtcaaaaattact   3000
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Seq_2  740   CCAACACAAGTGTTTTCAAAAATAGCTCGTGTCCACCCTCTCTTCTTGTCAAAAATTACT   799
                 P   T   Q   V   F   S   K   I   A   R   V   H   P   L   F   L   S   K   I   T


                 S   R   Y   L   V   N   D   S   K   K   K   S   H   N   V   G   L   M   I   K
Seq_1  3001  tcaagatatttggtcaatgattctaaaaagaaaagccataacgtcggtttgatgatcaaa   3060
             |||||||||||||||||||||||||||||:||||||||||||||||||||||||||||||
Seq_2  800   TCAAGATATTTGGTCAATGATTCTAAAAAGAAnAGCCATAACGTCGGTTTGATGATCAAA   859
                 S   R   Y   L   V   N   D   S   K   K   X   S   H   N   V   G   L   M   I   K


                 F   K   A   R   N   Q   K   V   L   G   Y   A   R   C   S   S   N   K   W   E
Seq_1  3061  ttcaaagcaagaaatcaaaaagttctcggttatgccagatgcagctcaaacaaatgggaa   3120
             |||||||||||:||||||||||||||||||||||||||||||||||||||:|||||||||
Seq_2  -860  TTCAAAGCAAGAnATCAAAAAGTTCTCGGTTATGCCAGATGCAGCTCAnACAAATGGGAA   919
                 F   K   A   R   X   Q   K   V   L   G   Y   A   R   C   S   S   X   K   W   E


                 Y   S   D   V   A   L   G   L   L   E   Q   F   R   S   T   F   P   E   F   F
Seq_1  3121  tactctgacgtcgctcttggtttgttagagcagttcagatctacattccctgagttcttt   3180
             .||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Seq_2  920   TACTCTGACGTCGCTCTTGGTTTGTTAGAGCAGTTCAGATCTACATTCCCTGAGTTCTTT   979
                 Y   S   D   V   A   L   G   L   L   E   Q   F   R   S   T   F   P   E   F   F


                 A   K   L   S   N   S   K   E   Q   A   I   P   S   I   T   D   L   F   P   N
Seq_1  3181  gcgaaactttctaactcgaaggaacaagcaattccatcgatcactgatctcttccctaac   3240
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Seq_2  980   GCGAAACTTTCTAACTCGAAGGAACAAGCAATTCCATCGATCACTGATCTCTTCCCTAAC   1039
                 A   K   L   S   N   S   K   E   Q   A   I   P   S   I   T   D   L   F   P   N


                 K   S   S   A   E   A   D   S   I   L   K   T   V   A   D   W   L   S   E   A
Seq_1  3239  acaaatctagcgcggaagctgattccattttgaaaacagtggctgattggctctcagaag   3298
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Seq_2  543   ACAAATCTAGCGCGGAAGCTGATTCCATTTTGAAAACAGTGGCTGATTGGCTCTCAGAAG   484
                 K   S   S   A   E   A   D   S   I   L   K   T   V   A   D   W   L   S   E   A


                 R   K   P   F   V   V   V   S   L   E   S   D   S   L   T   K   A   S   M   A
Seq_1  3299  caagaaaaccattcgtggtggtgtctttggaaagtgattcgctaaccaaggcttcgatgg   3358
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Seq_2  483   CAAGAAAACCATTCGTGGTGGTGTCTTTGGAAAGTGATTCGCTAACCAAGGCTTCGATGG   424
                 R   K   P   F   V   V   V   S   L   E   S   D   S   L   T   K   A   S   M   A


                 A   V   E   S   E   I   I   K   Y   V   S   L   P   D   S   S   E   Q   K   K
Seq_1  3359  cagctgttgaatctgaaatcataaaatacgtttctttaccagattcaagcgagcagaaga   3418
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Seq_2  423   CAGCTGTTGAATCTGAAATCATAAAATACGTTTCTTTACCAGATTCAAGCGAGCAGAAGA   364
                 A   V   E   S   E   I   I   K   Y   V   S   L   P   D   S   S   E   Q   K   K


                 L   A   K   V   P   R   E   A   I   L   N   A   E   S   S   Y   V   L   L   R
Seq_1  3419  aattagctaaggtcccacgtgaggcaatcttaaatgcggaatcgtcatatgttctattgc   3478
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Seq_2  363   AATTAGCTAAGGTCCCACGTGAGGCAATCTTAAATGCGGAATCGTCATATGTTCTATTGC   304
                 L   A   K   V   P   R   E   A   I   L   N   A   E   S   S   Y   V   L   L   R
```

257

```
                 S  Q  R  F  H  L  G  D  R  V  M  Y  I  Q  D  S  G  K  V  P
Seq_1   3479  gctcccaaaggttccacttaggtgatagggtaatgtacattcaagattcaggcaaggttc   3538
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Seq_2   303   GCTCCCAAAGGTTCCACTTAGGTGATAGGGTAATGTACATTCAAGATTCAGGCAAGGTTC   244
                 S  Q  R  F  H  L  G  D  R  V  M  Y  I  Q  D  S  G  K  V  P


                 L  H  S  K  G  T  V  V  G  Y  T  S  I  G  K  N  V  S  I  Q
Seq_1   3539  cacttcacagcaaaggtactgtcgttggctacacttcaattggcaagaacgtctcaatcc   3598
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Seq_2   243   CACTTCACAGCAAAGGTACTGTCGTTGGCTACACTTCAATTGGCAAGAACGTCTCAATCC   184
                 L  H  S  K  G  T  V  V  G  Y  T  S  I  G  K  N  V  S  I  Q



Seq_1   3599  aagttctatttgacaatgaaataattgcaggaaacaactttggtggtaggttgcagacca   3658
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Seq_2   183   AAGTTCTATTTGACAATGAAATAATTGCAGGAAACAACTTTGGTGGTAGGTTGCAGACcA   124
                 V  L  F  D  N  E  I  I  A  G  N  N  F  G  G  R  L  Q  T  R


                 R  G  L  G  L  D  S  S  F  L  L  N  L  S  D  R  Q  L  V  Y
Seq_1   3659  gacgtggtttgggattggactcttccttcttattaaacttgtctgatagacaattggtat   3718
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Seq_2   123   GACGTGGTTTGGGATTGGACTCTTCCTTCTTATTAAACTTGTCTGATAGACAATTGGTAT   64
                 R  G  L  G  L  D  S  S  F  L  L  N  L  S  D  R  Q  L  V  Y


                 H                                                          X
Seq_1   3719  atcat-------------------------------------------------------   3723
              |||||
Seq_2   63    ATCATTCAAAGGCATCGCTCGAGCACCACCACCACCACCACTGAGATCCGGctgctaAnc   4
                 H  S  K  A  S  L  E  H  H  H  H  H  H  *  D  P  A  A  X  X


                 X
Seq_1   3724  ---  3723

Seq_2   3    nAa  1
                 X
```

Figure A10: Sequence alignment of residues **1-1245** *of Kluyveromyces lactis* **XRN1 with**
Kluyveromyces lactis mRNA: *Both T7 forward and reverse sequencing reactions were performed*
by the GATC sequencing® and subsequent alignment of the ORF of recombinant plasmid was
done using Serial Cloner 2.1 (Serial Basics, Freeware). Nucleotides that could not be properly
determined are denoted with a *.

Observation of unpaired substrate DNA in the flap endonuclease-1 active site

L. David Finger[1], Nikesh Patel[1], Amanda Beddows[1], Long Ma[2], Jack C. Exell[1], Emma Jardine[1], Anita C. Jones[2,*] and Jane A. Grasby[1,*]

[1]Department of Chemistry, Centre for Chemical Biology, Krebs Institute, University of Sheffield, Sheffield, S3 7HF, UK and [2]EaStCHEM School of Chemistry and Collaborative Optical Spectroscopy, Micromanipulation and Imaging Centre, The University of Edinburgh, West Mains Road, Edinburgh EH9 3JJ, UK

ABSTRACT

The structure- and strand-specific phosphodiesterase flap endonuclease-1 (FEN1), the prototypical 5'-nuclease, catalyzes the essential removal of 5'-single-stranded flaps during replication and repair. FEN1 achieves this by selectively catalyzing hydrolysis one nucleotide into the duplex region of substrates, always targeting the 5'-strand. This specificity is proposed to arise by unpairing the 5'-end of duplex to permit the scissile phosphate diester to contact catalytic divalent metal ions. Providing the first direct evidence for this, we detected changes induced by human FEN1 (hFEN1) in the low-energy CD spectra and fluorescence lifetimes of 2-aminopurine in substrates and products that were indicative of unpairing. Divalent metal ions were essential for unpairing. However, although 5'-nuclease superfamily-conserved active-site residues K93 and R100 were required to produce unpaired product, they were not necessary to unpair substrates. Nevertheless, a unique arrangement of protein residues around the unpaired DNA was detected only with wild-type protein, suggesting a cooperative assembly of active-site residues that may be triggered by unpaired DNA. The general principles of FEN1 strand and reaction-site selection, which depend on the ability of juxtaposed divalent metal ions to unpair the end of duplex DNA, may also apply more widely to other structure- and strand-specific nucleases.

INTRODUCTION

Structure-specific phosphodiesterases play essential cellular roles by recognizing and acting on aberrant nucleic acid structures. Nucleic acid structures that require processing in this way include bubbles, flaps, nicks, gaps and four-way DNA junctions, which occur as intermediates during DNA replication, repair and recombination. To return to the duplex state and thereby restore the genome, the ends of intact duplexes contained within these more complex structures must undergo site-selective strand-specific phosphate diester hydrolyses. In line with this key role in maintaining genome integrity, defects in structure-sensing nucleases lead to a range of diseases, including cancer (1,2).

Exemplary strand-specific duplex-targeting phosphodiesterases are the flap endonuclease (FEN)-family nucleases, also known as the 5'-nuclease superfamily, whose activities involve processing a range of aberrant nucleic acid substrates (3). FEN1 is an essential component of the lagging-strand DNA replication, long-patch base excision repair and ribonucleotide excision repair pathways. FEN1 removes 5'-single-stranded protrusions, known as flaps, that are formed during DNA polymerase-catalyzed strand displacement synthesis. Other protein-sequence-related 5'-nucleases include EXO1, the 5'-nuclease that catalyzes resection of duplex, nicked and gapped DNAs during mismatch and double-strand break repair, and XPG, the 5'-nuclease of nucleotide excision repair that targets DNA bubbles. Another family member, GEN1, processes DNA four-way (Holliday) junctions. Although members of the 5'-nuclease family act on a diverse range of substrates, one feature of their

---

*To whom correspondence should be addressed. Tel: +44 1142 229478; Fax: +44 1142 229346; Email: j.a.grasby@sheffield.ac.uk
Correspondence may also be addressed to Anita C. Jones. Tel: +44 1316 506449; Fax: +44 1316 504743; Email: a.c.jones@ed.ac.uk
Present addresses:
Nikesh Patel, School of Molecular and Cellular Biology, LC Miall Building, University of Leeds, LS2 9JT, UK.
Long Ma, Biomolecular Sciences Research Complex, EaStCHEM School of Chemistry, University of St Andrews, Fife KY16 9ST, UK.

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

reactions is universal: all FEN family nucleases catalyze hydrolysis one nucleotide into a double-stranded region of their more complex target DNAs. Moreover, this reaction always occurs on the 5′-strand of the duplex abutting the nucleic acid junction.

This precise reaction site selection is critical to 5′-nuclease action. For example, FEN1 is tasked to create ligatable nicked-DNA products so as to prevent erroneous genome-endangering incision that would require the subsequent action of additional repair pathways. The reaction-site-specificity of superfamily proteins has been explained by a novel double nucleotide unpairing mechanism that only allows a specific phosphate diester at the 5′-end of the duplex to contact the catalytic active-site divalent metal ions (Figure 1A) (3–7). This mechanism was initially inferred from biochemical studies and X-ray structures of a FEN-family protein with substrate, where the scissile phosphate to undergo reaction was seen bound in front of, not within, the active site (4,5,8).

More recently, structures of human FEN1 (hFEN1) and human EXO1 (hEXO1) bound to product DNAs provided support for this mechanism (6,7). In product complexes, hydrolyzed 5′-phosphate monoester was directly coordinated to two metal ions within the active site; this conformation required the terminal nucleotide of product to be extra-helical. Recently, we showed that interstrand cross-linking of the nucleobases proposed to unpair prevented hFEN1 reaction within the duplex DNA, lending further support for the double nucleotide unpairing mechanism (9). However, although a double nucleotide unpaired conformation of substrate is implied by biochemical and structural data, it has yet to be observed. Here, we study the 5′-nuclease unpairing mechanism by monitoring local DNA conformational changes using low-energy (>300 nm) CD and fluorescence decays of substrate and product containing the nucleobase analogue 2-aminopurine (2AP), which can replace adenine in DNA without significant structural perturbation (10–15). These data provide the first observations of unpaired conformations in DNA substrates bound to hFEN1 in the presence of catalytically inert active-site metal ions and define the role of amino acid residues and divalent metal ion cofactors in this process.

MATERIALS AND METHODS

DNA constructs

DNA oligonucleotides including those containing site-specific 2AP substitutions were purchased with purification by high performance liquid chromatography from DNA Technology A/S. DNA concentrations were determined by ultraviolet absorbance at 260 nm (20°C). Heteroduplexes were formed by heating the appropriate 2AP ssDNA with the complementary template in a 1:1.1 ratio at 80°C for 5 min in 50 mM Tris–HCl (pH 7.5) and 100 mM KCl with subsequent cooling to room temperature. Sequences are given in Supplementary Figure S1.
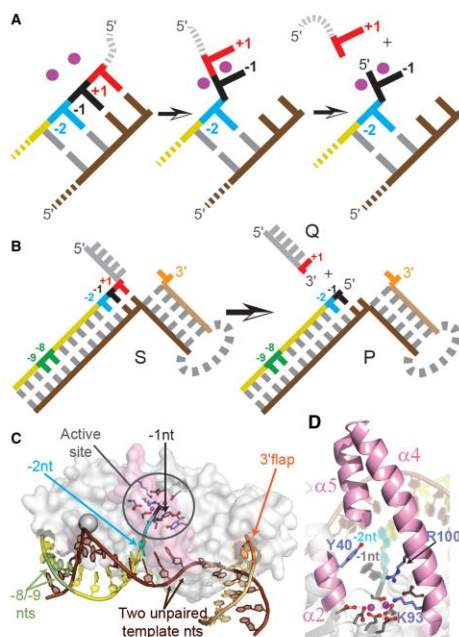

Figure 1. Schematic of proposed 5′-nuclease double nucleotide unpairing mechanism, reaction and supporting hFEN1-product structures. (**A**) Proposed unpairing of duplex ends that places the scissile phosphate diester in contact with active-site divalent ions (magenta). (**B**) The reaction of a static double-flap substrate (S) catalyzed by hFEN1 generates a 5′-phosphorylated product (P) and a single-stranded product (Q). Hydrolysis occurs between the +1 and −1 nucleotides (nts) as shown. (**C**) Structure of the hFEN1-product complex (3Q8K) highlighting two unpaired nucleotides of the template strand (brown), a single unpaired nt (−1 nt, black) of the CF strand (yellow) positioned on active-site metal ions (magenta), the 3′-flap nt (orange), the −2 nt (cyan), and −8 and −9 nts (green). (Product Q was not observed.) Protein is shown transparent with α4-α5 in pink. (**D**) The active site of hFEN1-P (colors as C). Side chains of superfamily conserved carboxylate residues (gray), K93, R100 and semi-conserved Y40 (light blue) are illustrated as sticks.

Enzymes

hFEN1 and mutants were over-expressed and purified as described (6).

CD spectroscopy

Samples containing 10 μM of the indicated DNA construct, 50 mM Tris–HCl (pH 7.5), 100 mM KCl, 1 mM dithiothreitol (DTT) and, where appropriate, 11–15 μM protein and either 10 mM CaCl$_2$ or 10 mM CaCl$_2$ + 25 mM EDTA were prepared with subsequent acquisition of CD spectra (300–480 nm) at 20°C using a JASCO J-810 CD spectrophotometer. CD spectra are an

average of five scans recorded in 0.5 nm steps (0.5 s response time) that were baseline corrected using spectra recorded on samples containing the same components, but lacking DNA. The baseline-subtracted spectra were then corrected by smoothing using the Means Movement option with a convolution width of five in the JASCO Spectra Analysis software (version 1.53.07). The CD spectra were plotted as $\Delta\varepsilon$ per mol 2AP residue versus wavelength. Each measurement was independently repeated (typically in triplicate) and gave identical results.

Time-resolved fluorescence spectroscopy

Samples contained 10 μM of the indicated DNA construct, 50 mM 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES) (pH 7.5), 100 mM KCl, 1 mM dithiothreitol (DTT) 10 mM $CaCl_2$ and, where appropriate, 10 μM protein and were monitored at 20°C. Fluorescence decays were measured using the technique of time-correlated single photon counting, following the same procedure reported previously (13). The excitation source was the third harmonic of the pulse-picked output of a Ti-Sapphire femtosecond laser system (Coherent, 10 W Verdi and Mira Ti-Sapphire), consisting of pulses of ~200 fs duration at a repetition rate of 4.75 MHz and a wavelength of 315 nm. Fluorescence decays were measured in an Edinburgh Instruments spectrometer equipped with TCC900 photon counting electronics. The instrument response of the system was ~80 ps full-width at half-maximum. Fluorescence decay curves were analyzed using a standard iterative reconvolution method, assuming a multi-exponential decay function. Decays were collected at three emission wavelengths (370, 380 and 390 nm) and analyzed globally using Edinburgh Instruments 'FAST' software (i.e. they were fitted simultaneously) with lifetimes, τ_i, as common parameters. The quality of fit was judged on the basis of the reduced chi-square statistic and the randomness of residuals. Results from repeat measurements (typically 3) were in good agreement, and the uncertainties in reported values of lifetimes and A factors are ≤10%.

RESULTS

Substrate design

The structure that is specifically recognized by eukaryotic FEN1 proteins is a two-way junction known as a 5′-3′-double flap that has a 5′-flap of any length (including none) and a single nucleotide 3′-flap (Figure 1B). When FEN1-catalyzed hydrolysis occurs one nucleotide into the 5′-duplex (between the +1 and −1 residues in the 5′-flap strand), the resulting product is nicked DNA that can be immediately joined by DNA ligase. Nicked DNA is the result of FEN1 action because both 5′-flaps and 3′-flaps are complementary to the template strand. However, such equilibrating 5′-3′-double flaps can adopt a variety of structures complicating analyses; ambiguity can be prevented by using non-complementary (i.e. static) double-flap constructs that undergo reactions with the same specificity as their equilibrating counterparts (16,17). To monitor conformational changes in hFEN1 substrates,

using low-energy CD and time-resolved fluorescence spectroscopies, we designed static, double-flap substrates (S) and the corresponding product (P) from two oligomers, where the flap (F) or cleaved-flap (CF) strand contained one or more 2AP residues (Figure 1B).

Spectroscopic analyses

Initially, we used time-resolved fluorescence of 2AP-containing DNA constructs to explore local conformational distortion induced by hFEN1 binding. Although these experiments proved informative, the interpretation of the 2AP decay parameters was complicated by the effect of interaction of 2AP with the Y40 residue of hFEN1 as well as with neighboring bases in the duplex. We, therefore, turned to low-energy CD, which reports only on the extent of stacking interaction between two adjacent 2APs strategically positioned within the duplex and gives a clear qualitative indication of local distortion. The low-energy CD results will be reported first, before considering the more-detailed information provided by time-resolved fluorescence.

Product unpairing monitored by low-energy CD

To establish a signal for altered nucleic acid conformations in hFEN1-DNA complexes, we began by investigating the low-energy CD spectra of free and bound product. In X-ray structures of hFEN1-product complexes, a single unpaired nucleotide in the −1 position is unstacked from its nearest neighbor, the −2 nucleotide (Figure 1A and C) (6). A product containing a 2AP dimer, $P_{−1−2}$, (numbering denotes the positioning of 2AP modifications), was created to monitor any changes in relative positions of the −1 and −2 nucleobases by low-energy CD (Figures 1B; Supplementary Figure S1). Adjacent 2APs dimers form an exciton-pair with two singly excited electronic transitions of unequal energies and oscillator strengths (absorption intensities). The energies and intensities of the electronic transitions are dependent on the respective orientation of the electronic transition dipoles of each 2AP in the dimer pair, and therefore, spectroscopic changes reflect the local DNA conformation. These changes are most easily observed in CD spectra where a bisignate signal is observed for exciton-coupled pairs. In our case, only the red-shifted portion of the bisignate signal is observed due to interference from the DNA absorption below 300 nm. Moreover, this red-shifted signal reaches is maximum at 320−330 nm, a region of the CD spectra where protein and unmodified DNA are transparent (10). Thus, 2AP-dimer containing DNAs can be used to study conformational changes, such as unpairing, in DNA and DNA-protein complexes (10).

The amplitude of the signal for the exciton-coupled 2APs contained within the single-stranded (ss) DNA ($CF_{−1−2}$) was considerably enhanced and red-shifted when the double-stranded (ds) $P_{−1−2}$ construct was formed (Figure 2A) (10). On formation of the hFEN1-$P_{−1−2}$ complex in the absence of divalent metal ions (presence of EDTA), only a small (25% *cf.* product in EDTA) decrease in the CD signal at 330 nm was
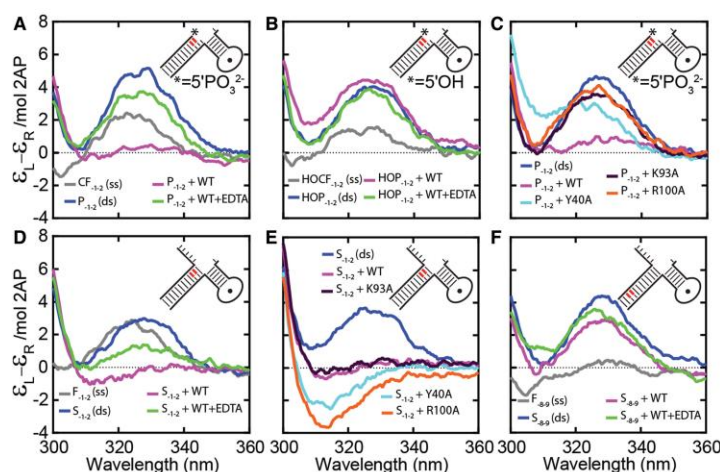
Figure 2. Low-energy CD spectra of 2AP-containing DNAs and hFEN1 WT- and mutant-DNA complexes. DNA constructs are illustrated schematically (red 2AP) with sequences shown in Supplementary Figure S1. (**A**) A dramatic divalent metal ion dependent reduction in 2AP exciton coupling signal occurred when product P_{-1-2} was bound to hFEN1, indicative of product unpairing. Unbound P_{-1-2} (blue), the corresponding single strand (CF_{-1-2}, gray) and P_{-1-2} bound to hFEN1 (magenta) all in Ca^{2+} containing buffer. P_{-1-2} bound to hFEN1 in buffer containing 25 mM EDTA (green). (**B**) When product is not 5'-phosphorylated, ($HO-P_{-1-2}$), there was no reduction in 2AP exciton coupling on addition of hFEN1- Ca^{2+}. Unbound $HO-P_{-1-2}$ (blue), the corresponding 2AP single strand ($HO-CF_{-1-2}$, gray) and $HO-P_{-1-2}$ bound to hFEN1 (magenta) all in Ca^{2+} containing buffer. $HO-P_{-1-2}$ bound to hFEN1 in buffer containing 25 mM EDTA (green). (**C**) The presence of 5'-nuclease superfamily conserved K93 and R100 were required to reduce exciton coupling of P_{-1-2}. Unbound P_{-1-2} (blue), P_{-1-2} bound to WT hFEN1 (magenta), Y40A (cyan), K93A (purple) and R100A (orange) in Ca^{2+} containing buffer. See Supplementary Figure S3A for spectra in EDTA buffer. (**D**) A similar dramatic divalent metal ion-dependent reduction in 2AP exciton coupling signal occurred when substrate S_{-1-2} was bound to hFEN1, indicative of substrate unpairing. Unbound S_{-1-2} (blue), the corresponding 2AP single strand (F_{-1-2}, gray), and S_{-1-2} bound to WT hFEN1 (magenta) in Ca^{2+} buffer. S_{-1-2} bound to hFEN1 in buffer containing 25 mM EDTA (green). (**E**) The presence of 5'-nuclease superfamily conserved K93, R100 and Y40 were not required to reduce exciton coupling of S_{-1-2}. Unbound S_{-1-2} (blue) and S_{-1-2} bound to WT hFEN1 (magenta), Y40A (cyan), K93A (purple), R100A (orange) all in Ca^{2+} containing buffer. (**F**) No changes in low energy CD signal occurred on complexation when 2APs were located away from the region of the substrate proposed to unpair. Unbound S_{-8-9} (blue), the corresponding single strand (F_{-8-9}, gray) and S_{-8-9} bound to hFEN1 (magenta) all in Ca^{2+} containing buffer. S_{-8-9} bound to hFEN1 in buffer containing 25 mM EDTA (green). Each measurement was independently repeated and gave equivalent results.

observed, despite electrophoretic mobility shift assay results showing that at least 95% of the product was bound (Supplementary Figure S2A). Thus, only a minor alteration of the respective orientations of the two 2AP nucleobases occurs on binding hFEN1 without active-site divalent metals. In contrast, addition of hFEN1 in the presence of catalytically inert Ca^{2+} ions dramatically reduced the amplitude of the signal at 330 nm (83% cf. product in Ca^{2+} buffer) to near zero, implying a significant change of the relative orientation of the two 2AP residues (Figure 2A). However, when product lacked its 5'-phosphate monoester (HO-P_{-1-2}), negligible changes in the CD spectra occurred on addition of protein, regardless of the presence of Ca^{2+} ions (Figure 2B). Notably, both P_{-1-2} and HO-P_{-1-2} in complex with hFEN1 produced the identical signals in EDTA. Thus, hFEN1-induced conformational change of the 3'-product (P) DNA requires both divalent metal ions and the 5'-phosphate monoester, consistent with the extra-helical conformation of the −1 nucleotide observed in hFEN1 structures, where phosphate monoester is directly bound to active-site metal ions (Figure 1C and D).

In hFEN1-unpaired product structures, the metal-bound 5'-phosphate monoester is contacted by superfamily conserved residues K93 and R100 and the unpaired −1 base is stacked with Y40 (Figure 1D) (6). When P_{-1-2} complexes were formed in the presence of Ca^{2+} ions with K93A and R100A mutated FENs, no evidence for conformational change about the −1 and −2 nt was observed, when compared with similar complexes formed in the presence of EDTA (Figures 2C; Supplementary FigureS3A). In contrast, Y40A complexes displayed a decrease in signal at 330 nm (51% cf. product in Ca^{2+} buffer) with Ca^{2+} ions (Figure 2C), but this was not as dramatic as that observed for wild-type (WT) protein complex (Figure 2A).

Substrate unpairing monitored by low-energy CD

To investigate substrate unpairing in hFEN1 complexes, CD spectra of the bound double-flap substrate S_{-1-2} were determined ± catalytically inert Ca^{2+} ions. Like the product complex, a Ca^{2+}-dependent reduction in signal (98% cf. substrate duplex in Ca^{2+} buffer) was observed in S_{-1-2} complexes (Figure 2D). Thus, both substrate and

262

product undergo conformational changes that have a similar effect on the relative positions of the -1 and -2 2APs in the presence of hFEN1-Ca^{2+}. This decrease in the magnitude of exciton-coupling (83% cf. substrate in Ca^{2+} buffer) was also observed in the presence of K93A-mutated hFEN1 and Ca^{2+} ions (Figure 2E), showing that, unlike product, substrate conformational change with respect to the -1 and -2 nucleotides is not dependent on this residue. Mutated hFEN1s, R100A and Y40A, also produced a dramatic change in the CD spectra, with a deep minimum present at 315 nm, in the presence of Ca^{2+} ions, implying that reorientation of the -1 and -2 nt also occurs in these complexes, but the precise juxtaposition or the partition between paired and unpaired forms differs from those in the WT and K93A complexes (Figure 2E).

In all cases, addition of EDTA to the Ca^{2+}-enzyme-substrate complexes, which returned the signal to that of a complex formed in EDTA alone and allowed subsequent analysis of the samples by high performance liquid chromatography, confirmed that the spectral changes were not the result of conversion of substrate to product (Supplementary Figures S2B and S4). Unlike complexes of S_{-1-2}, only minor spectral changes were observed for hFEN1-S_{-8-9} ± Ca^{2+}, where the adjacent 2APs are firmly embedded in the substrate duplex (Figure 2F). Furthermore, electrophoretic mobility shift assay showed that all the substrates were at least 95% bound by hFEN1 under the conditions used (Supplementary Figure S2B and C). Therefore, we have spectral evidence of protein- and metal-ion-dependent conformational change of the substrate, analogous to that of product complex and indicative of unpairing of the substrate.

Time-resolved fluorescence of 2AP single-strands, products and substrates

More detailed information about the nature of DNA conformational changes brought about by hFEN1-Ca^{2+} and mutants in the presence of metal ions was provided by time-resolved fluorescence spectroscopy. F or CF strands containing single 2AP substitutions were created. Where necessary, these were hybridized to template DNAs to form substrates S_{+1}, S_{-1}, S_{-9} and products, P_{-1} and Q_{+1} (numbering denotes the site of 2AP modification, Figure 1A and B, Supplementary Figure S1) for analysis by time-resolved fluorescence spectroscopy.

Initial fluorescence studies focused on the characterization of substrates and products in the absence of hFEN1. As seen previously, the fluorescence decays of the 2AP-containing DNAs were described by four lifetime components (τ_{1-4}), reflecting the existence of a variety of conformational states (Figure 3A, Supplementary Table S1A) in which 2AP experiences different quenching efficiencies (11–15). Quenching is used here to mean the reduction in fluorescence lifetime (increase in non-radiative decay rate) as a result of intermolecular interaction. Earlier studies indicate that the shortest lifetime component (ca. 50 ps) corresponds to 2AP that is well stacked in the duplex and subject to rapid, fluorescence quenching, primarily due to inter-base electron transfer. The longest component of ca. 10 ns results from extra-helical 2AP (free
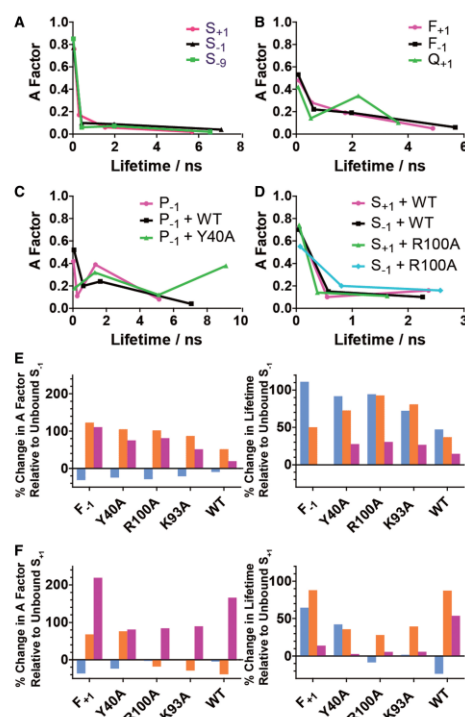

Figure 3. Graphical representation of the fluorescence decay parameters of 2AP-containing DNAs and hFEN1 WT- and mutant-DNA complexes. All spectra were recorded in Ca^{2+} containing buffers. (**A–D**) show plots of fractional amplitude (A factor) versus lifetime for unbound DNA constructs and selected complexes: (A) unbound S_{+1} (orange), S_{-1} (blue) and S_{-9} (green); (B) unbound F_{+1} (magenta), F_{-1} (black) and Q_{+1} (green); (C) unbound P_{-1} (magenta) and complexes of P_{-1} (black) and Y40A (green) with WT hFEN-1; (D) complexes of S_{+1} with WT (magenta) and R100A (green) and of S_{-1} with WT (black) and R100A (cyan). (**E**) and (**F**) show the percentage change in the fractional amplitudes (left) and corresponding fluorescence lifetime components (right) of each double-stranded substrate, S_{-1} (E) and S_{+1} (F), on conversion to the corresponding single strand or formation of the indicated protein complexes; τ_1 and A_1 cyan, τ_2 and A_2 orange, τ_3 and A_3 magenta Full sequences are shown in Supplementary Figure S1.

from quenching), whereas two intermediate lifetimes are attributed to partially stacked forms. Thus, the magnitude of the lifetime reveals the nature of microenvironment of the 2AP (extent of stacking). The fractional amplitudes (A factors) of the four components (A_{1-4}) indicate the fraction of the emitting population with a given lifetime and correspond to the relative populations of the various conformational states.

The fluorescence decay parameters of S_{+1}, S_{-1} and S_{-9} are summarized graphically in Figure 3A as a plot of A

factor versus lifetime, and the values are given in Supplementary Table S1A. They are all typical of well-formed duplex structures in which the vast majority of the 2AP population displays a short lifetime (τ_1), characteristic of a highly stacked conformation. Interestingly, although the population of this rapidly quenched state is greatest ($A_1 = 85\%$) when 2AP is embedded well within the duplex (S_{-9}), it remains almost as high (ca. 75%) in the double-flap substrates with 2AP at the duplex termini (S_{+1} and S_{-1}). The maintenance of the terminal bases of duplex in an optimally stacked state within the double flaps is presumably a consequence of coaxial stacking of the two-way junction. This contrasts with measurements of simple duplex termini or forked DNAs, which show much greater amounts of unpairing at duplex ends (10). By analogy, a more heterogeneous conformational population than seen in the substrates would be expected for the 5'-phosphorylated product P_{-1}, with a 5'-terminal 2AP, as here the 3'-flap is not complementary to template DNA, reducing coaxial stacking. Indeed, in P_{-1}, 2AP was less well stacked than in the intact substrate with a large fraction of the population (ca. 60%) in weakly stacked states (Figure 3C, Supplementary Table S1A).

The fluorescence decay of 2AP is not sensitive to hydrogen-bonding interactions *per se*; thus, 2AP is not a direct probe of base pairing. However, the perturbation of base-stacking interactions that occurs as a result of local unpairing of the duplex gives rise to characteristic changes in the fluorescence decay parameters of 2AP, as we have shown previously in a systematic comparison of single strands and duplexes (15). This is evident here in the decay parameters of ss F_{-1} and F_{+1} (Figure 3B, Supplementary Table S1A), when compared with those of the corresponding duplexes, as illustrated in Figure 3A and Supplementary Table S1A. Compared with the duplex, there was a sizeable transfer of population from the highly stacked state (decrease in A_1) to states with sub-optimal stacking (increase in A_2 and A_3). There was also a substantial lengthening of decay times τ_1 and τ_2, indicative of a weakening of stacking interactions in the single strands. The population of the longest lifetime component remains small (only a few percent) in the single strands, indicating that the propensity for complete destacking of 2AP from its neighboring bases (i.e. base flipping) is no greater than in the duplex. When 2AP was the terminal 3'-nucleotide in ss product Q_{+1} and less constrained than intrastrand 2AP in ss F_{+1} and F_{-1}, there was an even more marked increase in occupancy of poorly stacked states (Figure 3B, Supplementary Table S1A).

The changes in the 2AP decay parameters that accompany the transition from duplex to single strand (i.e. duplex unpairing) can be represented schematically by the percentage change in lifetimes and A factors of components 1 3 in the single strand relative to the values in the duplex, as illustrated in Figure 3E, for F_{-1} (ss) compared with S_{-1} (ds), and in Figure 3F, for F_{+1} (ss) compared with S_{+1} (ds). (Component 4 has been omitted for simplicity because, as noted earlier in the text, it

constitutes a small fraction of the emitting population in both single strands and duplexes).

Unpaired hFEN1-product complexes detected by time-resolved fluorescence

To study the fluorescence decay, a WT hFEN1-Ca^{2+}-P_{-1} complex was formed, analogous to the P_{-1-2} complex that provided evidence of unpairing by low energy CD. Surprisingly, a greater proportion of 2AP was rapidly quenched in hFEN1-Ca^{2+}-P_{-1} compared to the P_{-1} in the unbound state (increase in A_1 and A_2, decrease in A_3 and A_4) as shown in Figure 3C and Supplementary Table S1B. We rationalized that when bound to WT hFEN1, fluorescence of 2AP in P_{-1} is quenched by Y40, observed to stack on the unpaired -1 nucleotide in the crystal structure (Figure 1D). Previously, stacking of 2AP with tyrosine residues was shown to result in efficient charge-transfer quenching (13,14). To test this hypothesis, we prepared a Y40A-P_{-1}-Ca^{2+} complex and observed that quenching of 2AP in P_{-1} was dramatically reduced (Figure 3C, Supplementary Table S1B). In the absence of Y40, ca. 40% of 2AP decayed with a lifetime of 9 ns, characteristic of an unquenched extrahelical environment. Nevertheless, there is also evidence that 2AP occupies intrahelical states in Y40A-P_{-1}-Ca^{2+}; 18% of the 2AP still decays with a short lifetime (τ_1) implying that a portion of the population remains well stacked on its nearest neighbor, although this is substantially reduced compared with the unbound duplex. This is in accord with the conclusions of low energy CD measurements of Y40A-Ca^{2+}-P_{-1-2}.

Unpairing of substrate by mutated hFEN1s monitored by time-resolved fluorescence

When Y40A, K93A and R100A complexes were formed with substrate S_{-1} in the presence of Ca^{2+} ions, the fluorescence decays and associated parameters resembled those of the unbound ss flap F_{-1} (Supplementary Tables S1A and C). This is illustrated in Figure 3E, where the changes in the decay parameters on binding of S_{-1} by the mutated enzymes can be seen to closely resemble the changes that accompanied the transition from duplex to single strand. Thus, the 2AP in all these complexes sees a local environment that resembles a single strand, implying that both the -1 nucleotide and the adjacent $+1$ nucleotide are unpaired, in line with the conclusions of exciton coupling measurements. However, in enzyme-bound S_{-1}, the population of well-stacked states (A_1) is somewhat greater than in ssF_{-1} and the population of poorly stacked states (A_3) somewhat less, suggesting an equilibrium between unpaired and paired conformations. It is notable that the decay parameters of 2AP in S_{-1} are not sensitive to the presence of Y40 in the mutated complexes (the decay parameters of all three complexes are almost identical), implying that the -1 nucleotide does not stack with this tyrosine residue in the R93A or R100A substrate complexes.

As shown in Figure 3F (and Supplementary Tables S1A and C), the decay of S_{+1} also showed a clear shift toward that of an unpaired structure on binding to Y40A-Ca^{2+};

there was a significant decrease in A_1 (from 0.75 to 0.58) and corresponding increases in A_2 (from 0.17 to 0.3) and A_3 (from 0.06 to 0.10), together with increases in τ_1 (from 60 ps to 80 ps) and τ_2 (from 300 ps to 400 ps), in comparison with the unbound substrate. However, S_{+1} showed a different response when the substrate was bound to R100A or K93A, suggesting that 2AP in the +1 position may be affected by interaction with Y40. Indeed, in these complexes, there was enhanced quenching of 2AP compared with the Y40A complex (increase in A_1 and decrease in τ_1) reminiscent of that seen in the WT-P_{-1} product complex (see earlier in the text). In crystal structures of hFEN1 with substrates that lack a 5'-phosphate, the DNA is seen to remain base-paired, with the +1 nucleotide positioned to stack with Y40 (Supplementary Figure S5A and B) (6). Therefore, the enhanced quenching seen in R100A-Ca^{2+} and K93A-Ca^{2+} S_{+1} complexes can be attributed to interaction of the +1 nucleotide with Y40, which is again suggestive of an equilibrium between unpaired and paired conformations.

Unpaired hFEN1-substrate complexes detected by time-resolved fluorescence

When WT hFEN1 was bound to S_{-9} in the presence of Ca^{2+} ions, no alteration of the decay parameters with respect to the free substrate was observed (Supplementary Table S1C). In contrast, the decay parameters of the WT- Ca^{2+}-S_{+1} complex were altered with respect to free S_{+1} and resembled those of S_{+1} bound to K93A and R100A, with the quenching effect of Y40 clearly apparent (Figure 3D and F; Supplementary Table S3A). However, the values of A_3 and τ_3 were greater than in the mutated complexes, signifying greater disruption of interbase stacking; this increase may be the consequence of distortion required to stack the −1 base with Y40 (see later in the text). It is particularly significant that when S_{-1} was bound to WT hFEN1, its decay parameters were different from those of all three mutated complexes and closely resembled those of the WT hFEN1-Ca^{2+}-S_{+1} complex (Figure 3F, Supplementary Table S1C). This is emphasized in Figure 3D, where the mutated complexes are exemplified by R100A (the decay parameters of S_{-1} in the three mutated complexes are almost identical). In complex with the WT protein, 2AP in each position in the substrate, +1 and −1, sees a similar microenvironment; both residues are quenched by stacking with Y40, presumably as a consequence of a paired-unpaired equilibrium. It is only in the WT complex that the −1 nucleotide of substrate appears able to stack with Y40, implying the necessity of residues R100 and K93 to form this particular unpaired state in substantial quantities.

DISCUSSION

The necessity for double nucleotide unpairing of 5'-nuclease superfamily substrates has been inferred from biochemical studies and structural studies, but not yet observed (3 7). Our spectroscopic studies using 2AP-containing substrates provide the first direct evidence that the

prototypical 5'-nuclease, hFEN1, can facilitate the unpairing of nucleotides surrounding the scissile phosphate diester bond. Our studies also reveal the components of the 5'-nuclease active site required for unpairing DNAs. An essential prerequisite for unpairing of both substrates and products is the presence of active site bound divalent metal ions, whose absence prevents unpairing in complexes with both DNAs. It has been noted previously that two (or more) metal ion active sites are associated with nucleases where DNA distortion is linked to active-site metal ion positioning of the scissile phosphate diester to afford specificity (18). The 5'-nuclease mechanism is an example of this phenomenon.

Surprisingly, 5'-nuclease superfamily strictly conserved (K93, R100) and semi-conserved residues (Y40), whose mutations are severely detrimental to catalysis in a range of family members (6,7,19), are not required to effect unpairing of substrates per se. Instead, active-site divalent metal ions are sufficient, although there is clear local variation in the respective conformation of unpaired nucleotides observed by CD with mutated proteins Y40A and R100A. Interestingly, in hFEN1 complexes with active-site metals, but devoid of DNAs, the region of the protein bearing K93 and R100 is disordered (α4-α5), as is the top portion of the Y40 helix (α2) (3,6,20,21). Moreover, residues of α2 pack against α5 when these regions are structured. As DNA conformational change can occur without these conserved residues, the possibility that nucleotide unpairing promotes arrangement of these regions of the protein to catalyze reaction is attractive. Such a mechanism would associate assembly of constituents of the active site with the presence of correctly unpaired substrate, thereby linking catalysis to specificity.

In contrast to the two unpaired nucleotides of substrate, both the conserved basic residues K93 and R100 are required to stabilize the single nucleotide unpaired product complex. Although some unpaired product can be formed with Y40A-Ca^{2+}, the absence of Y40 appears to alter the partition between paired and unpaired forms of product or the respective orientation of the −1 and −2 nucleotides. As K93, R100 and Y40 are key to efficient catalysis, it seems likely that the requirements for unpaired product conformation most closely mimic those of double nucleotide unpaired substrate that is correctly positioned to react.

Time-resolved fluorescence identifies a unique equilibrium in WT hFEN1-substrate complexes, where interaction of Y40 with both the −1 and +1 nucleobases either side of the scissile bond can be detected. Although we cannot rule out an intermediate orientation of substrate that is unpaired, but positioned in such a way that allows Y40 to interact with the +1 base, the simplest explanation of this phenomenon is that paired substrate DNAs are present alongside their unpaired counterparts. The detected interaction of Y40 with the −1 nucleotide of the unpaired substrate, unique to the WT protein, requires the presence of the basic residues of the active site. This indicates that the characteristic 5'-nuclease superfamily active-site residues cooperate to produce what is likely to be close to the reactive conformation and would be

consistent with an unpaired substrate inducing assembly of key active-site residues.

FEN1 substrate conformational change is reminiscent of the unpairing observed when DNA polymerases transfer the 3'-terminus of substrates to the 3'-exonuclease domain in editing mode. This substrate unpairing has also been studied using low energy CD of 2AP dimer substrates and time-resolved fluorescence measurements involving both 2AP and dyes (22 25). In DNA polymerases, occupation of the divalent metal ion binding 3'-exonuclease site is increased when the duplex has 3'-mismatched ends, although this site is still occupied to a lesser extent by base-paired DNAs, including their 2AP-T analogue. With the static double-flap substrate alone, the target duplex is not in a state that is substantially frayed, and neither is this state induced when bound to protein without divalent metal ions, despite the protein bending the substrate by 100° and breaking the coaxial stacking of the double flap. In contrast to polymerase editing, the FEN1 superfamily is tasked with recognizing base-paired ends (that abut DNA junctions), presumably to identify and preserve perfect duplex during replication or restoration of the genome. Nevertheless, similar overall principles, despite subtle differences, apply to both 5'-nucleases and 3'-exonucleases; a duplex binding site separated from an active site requires substrates to unpair for hydrolysis to occur. In both cases, this affords reaction specificity for duplex ends that can be unpaired, whereas the particular strand of duplex that undergoes reaction is selected by the precise juxtaposition of duplex binding and active sites.

These principles may also apply more widely to other structure- and strand-specific nucleases. Unrelated 5'-nucleases, such as the toroidal lambda nuclease, appear likely to achieve strand selectivity by unpairing ends (26), and DNA co-crystal structures of the 3'-nuclease Mre11 also suggest substrate unpairing is necessary for contact between active-site metal ions and the target phosphate diester (27). Degradative exoribonucleases (28 30), including some that conserve the FEN active site (3), seem to use unpairing of ends to deal with regions of RNA secondary structure and preserve reaction polarity. Thus, duplex-end unpairing, used by FEN1 family members to locate their reaction site and ensure reaction is specific for the 5'-strand, appears likely to be a common feature of both DNA- and RNA-acting structure- and strand-specific nucleases. The combination of low-energy CD and time-resolved fluorescence of 2AP-containing substrates will be a valuable methodology to explore the possible generality of this phenomenon.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

The authors thank Barbara Ciani (Sheffield) for advice on CD measurements, Jochen Arlt (COSMIC, Edinburgh) for assistance with time-resolved fluorescence experiments and Mark Thompson (Sheffield) for critical reading of the manuscript.

REFERENCES

1. Kucherlapati,M., Yan,K., Kuraguchi,M., Zhao,J., Lia,M., Heyer,J., Kane,M.F., Fan,K.H., Russell,R., Brown,A.M.C. *et al.* (2002) Haploinsufficiency of flap endonuclease (FEN-1) leads to rapid tumour progression. *Proc. Natl Acad. Sci. USA,* **99**, 9924–9929.
2. Zheng,L., Dai,H., Zhou,M., Li,M., Singh,P., Qui,J., Tsark,W., Huang,Q., Kernstine,K., Zhang,X. *et al.* (2007) Fen1 mutations result in autoimmunity, chronic inflammation and cancers. *Nat. Med.,* **13**, 812–818.
3. Grasby,J.A., Finger,L.D., Tsutakawa,S.E., Atack,J.M. and Tainer,J.A. (2012) Unpairing and gating: sequence-independent substrate recognition by FEN superfamily nucleases. *TIBS,* **37**, 74–84.
4. Tomlinson,C.G., Atack,J.M., Chapados,B.R., Tainer,J.A. and Grasby,J.A. (2010) Substrate recognition and catalysis by flap endonucleases and related enzymes. *Biochem. Soc. Trans.,* **38**, 433–437.
5. 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 T5 flap endonuclease catalysed DNA hydrolysis. *J. Biol. Chem.,* **283**, 28741–28746.
6. Tsutakawa,S.E., Classen,S., Chapados,B.R., Arvai,A.S., Finger,L.D., Guenther,G., Tomlinson,C.G., Thompson,P., Sarker,A.H., Shen,B. *et al.* (2011) Human flap endonuclease structures, DNA double-base flipping, and a unified understanding of the FEN1 superfamily. *Cell,* **145**, 198–211.
7. 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.
8. 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.
9. Beddows,A., Patel,N., Finger,L.D., Atack,J.M., Williams,D.M. and Grasby,J.A. (2012) Interstrand disulfide crosslinking of DNA bases supports a double nucleotide unpairing mechanism for flap endonucleases. *Chem. Comm.,* **48**, 8895–8897.
10. Jose,D., Datta,K., Johnson,N.P. and von Hippel,P.H. (2009) Spectroscopic studies of position-specific DNA "breathing" fluctuations at replication forks and primer-template junctions. *Proc. Natl Acad. Sci. USA,* **106**, 4231–4236.
11. Guest,C.R., Hochstrasser,R.A., Sowers,L.C. and Millar,D.P. (1991) Dynamics of mismatched base-pairs in DNA. *Biochemistry,* **30**, 3271–3279.
12. Lenz,T., Bonnist,E.Y.M., Pljevaljcic,G., Neely,R.K., Dryden,D.T.F., Scheidig,A.J., Jones,A.C. and Weinhold,E. (2007) 2-Aminopurine flipped into the active site of the adenine-specific DNA methyltransferase M.TaqI: Crystal structures and time-resolved fluorescence. *J. Am. Chem. Soc.,* **129**, 6240–6248.
13. Neely,R.K., Daujotyte,D., Grazulis,S., Magennis,S.W., Dryden,D.T.F., Klimasauskas,S. and Jones,A.C. (2005) Time-

266

resolved fluorescence of 2-aminopurine as a probe of base flipping in M.HhaI–DNA complexes. *Nucleic Acids Res.*, **33**, 6953–6960.

14. Neely,R.K., Tamulaitis,G., Chen,K., Kubala,M., Siksnys,V. and Jones,A.C. (2009) Time-resolved fluorescence studies of nucleotide flipping by restriction enzymes. *Nucleic Acids Res.*, **37**, 6859–6870.

15. Sabir,T., Toulmin,A., Ma,L., Jones,A.C., McGlynn,P., Schroeder,G.F. and Magennis,S.W. (2012) Branchpoint expansion in a fully complementary three-way DNA junction. *J. Am. Chem. Soc.*, **134**, 6280–6285.

16. Finger,L.D., Atack,J.M., Tsutakawa,S., Classen,S., Tainer,J., Grasby,J. and Shen,B. (2012) The Wonders of Flap Endonucleases: Structure, Function, Mechanism and Regulation. In: MacNeill,S. (ed.), *The Eukaryotic Replisome: A guide to Protein Structure and Function*, Vol. 62. Springer, Netherlands, pp. 301–326.

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

18. Yang,W., Lee,J.Y. and Nowotny,M. (2006) Making and breaking nucleic acids: Two-Mg2+-ion catalysis and substrate specificity. *Mol. Cell*, **22**, 5–13.

19. Sengerová,B., Tomlinson,C.G., 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.

20. 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.

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

22. Datta,K., Johnson,N.P. and von Hippel,P.H. (2010) DNA conformational changes at the primer-template junction regulate the fidelity of replication by DNA polymerase. *Proc. Natl Acad. Sci. USA*, **107**, 17980–17985.

23. Hochstrasser,R.A., Carver,T.E., Sowers,L.C. and Millar,D.P. (1994) Melting of a DNA Helix Within the Active-Site of a DNA-Polymerase. *Biochemistry*, **33**, 11971–11979.

24. Lam,W.C., Van der Schans,E.J.C., Joyce,C.M. and Millar,D.P. (1998) Effects of mutations on the partitioning of DNA substrates between the polymerase and 3′-5′ exonuclease sites of DNA polymerase I (Klenow fragment). *Biochemistry*, **37**, 1513–1522.

25. Richardson,T., Wu,X., Keith,B., Heslop,P., Jones,A. and Connolly,B. (2013) Unwinding of primer-templates by archaeal family-B DNA polymerases in response to template-strand uracil. *Nucleic Acids. Res.*, **41**, 2466–2478.

26. Zhang,J., McCabe,K.A. and Bell,C.E. (2011) Crystal structures of lambda exonuclease in complex with DNA suggest an electrostatic ratchet mechanism for processivity. *Proc. Natl Acad. Sci. USA*, **108**, 11872–11877.

27. Williams,R.S., Moncalian,G., Williams,J.S., Yamada,Y., Limbo,O., Shin,D.S., Groocock,L.M., Cahill,D., Hitomi,C., Guenther,G. *et al.* (2008) Mre11 dimers coordinate DNA end bridging and nuclease processing in double-strand-break repair. *Cell*, **135**, 97–109.

28. Frazao,C., McVey,C.E., Amblar,M., Barbas,A., Vonrhein,C., Arraiano,C.M. and Carrondo,M.A. (2006) Unravelling the dynamics of RNA degradation by ribonuclease II and its RNA-bound complex. *Nature*, **443**, 110–114.

29. Lee,G., Bratkowski,M.A., Ding,F., Ke,A. and Ha,T. (2012) Elastic coupling between RNA degradation and unwinding by an exoribonuclease. *Science*, **336**, 1726–1729.

30. Lorentzen,E., Basquin,J., Tomecki,R., Dziembowski,A. and Conti,E. (2008) Structure of the active subunit of the yeast exosome core, Rrp44: Diverse modes of substrate recruitment in the RNase II nuclease family. *Mol. Cell*, **29**, 717–728.

267

Proline Scanning Mutagenesis Reveals a Role for the Flap Endonuclease-1 Helical Cap in Substrate Unpairing[*]

Nikesh Patel[1,2]**, Jack C. Exell**[1]**, Emma Jardine, Ben Ombler**[3]**, L. David Finger**[4]**, Barbara Ciani**[5]**, and Jane A. Grasby**[6]

From the Centre for Chemical Biology, Department of Chemistry, Krebs Institute, University of Sheffield, Sheffield S3 7HF, United Kingdom

Background: Human flap endonuclease-1, the prototypical 5′-nuclease, removes 5′-flaps by incising one nucleotide into duplex DNA using a double nucleotide unpairing mechanism.
Results: Alteration of the hFEN1 helical cap structure, but not removal of conserved basic residues, prevents substrate unpairing.
Conclusion: The hFEN1 helical cap is required for substrate rearrangement.
Significance: Mechanistic details of 5′-nuclease catalysis are crucial for understanding DNA replication and repair.

The prototypical 5′-nuclease, flap endonuclease-1 (FEN1), catalyzes the essential removal of single-stranded flaps during DNA replication and repair. FEN1 hydrolyzes a specific phosphodiester bond one nucleotide into double-stranded DNA. This specificity arises from double nucleotide unpairing that places the scissile phosphate diester on active site divalent metal ions. Also related to FEN1 specificity is the helical arch, through which 5′-flaps, but not continuous DNAs, can thread. The arch contains basic residues (Lys-93 and Arg-100 in human FEN1 (hFEN1)) that are conserved by all 5′-nucleases and a cap region only present in enzymes that process DNAs with 5′ termini. Proline mutations (L97P, L111P, L130P) were introduced into the hFEN1 helical arch. Each mutation was severely detrimental to reaction. However, all proteins were at least as stable as wild-type (WT) hFEN1 and bound substrate with comparable affinity. Moreover, all mutants produced complexes with 5′-biotinylated substrate that, when captured with streptavidin, were resistant to challenge with competitor DNA. Removal of both conserved basic residues (K93A/R100A) was no more detrimental to reaction than the single mutation R100A, but much less severe than L97P. The ability of protein-Ca^{2+} to rearrange 2-aminopurine-containing substrates was monitored by low energy CD. Although L97P and K93A/R100A retained the ability to unpair substrates, the cap mutants L111P and L130P did not. Taken together, these data challenge current assumptions related to 5′-nuclease family mechanism. Conserved basic amino acids are not required for double nucleotide unpairing and appear to act cooperatively, whereas the helical cap plays an unexpected role in hFEN1-substrate rearrangement.

Lagging strand DNA replication depends on the ability to process Okazaki fragments by removing single-stranded nucleic acid protrusions known as flaps that result from DNA polymerase-catalyzed strand displacement synthesis. In mammals, yeast, and some bacteria, FEN1 [7] elegantly catalyzes flap removal by recognizing a unique double-flap Okazaki fragment conformer with a 5′-flap of variable length, but only a single nucleotide 3′-flap (1–5). When subject to FEN1-catalyzed 5′-nucleolytic action, an incision is created one nucleotide into the double strand (see Fig. 1A). This results in a nicked DNA product that is immediately suitable for ligation, the last step of Okazaki fragment maturation. Similarly, DNA repair events that also depend upon DNA polymerase-catalyzed strand displacement synthesis, such as long patch base and ribonucleotide excision repair, have the same absolute requirement for the structure- and strand-specific nuclease FEN1.

In addition to flaps, other aberrant DNA structures also require 5′-nucleolytic action. Several of these aberrant structures are processed by alternative nucleases with sequence homology to FEN1; together, they form the FEN1 or 5′-nuclease superfamily (6). In humans, these include the DNA repair protein XPG that acts on bubble structures in nucleotide excision repair, EXO1 that acts on nicked or gapped DNAs in mismatch repair and recombination, and GEN1 that acts on Holliday junctions. An intriguing question raised by the superfamily is how such a diverse range of nucleic acid structures is processed with the exquisite substrate specificity necessary for each of the individual proteins using similar protein architecture.

Biochemical and structural studies of hFEN1 and other FEN1 family members have begun to reveal a mechanism for the action of 5′-nucleases (5–13). FEN1 family proteins all catalyze divalent metal ion-dependent phosphate diester hydrolysis one nucleotide into a duplex portion of their respective substrates. This reaction selectivity is achieved by double nucleotide unpairing of the duplex portion of substrate that abuts the DNA junction (see Fig. 1A). Unpairing allows the scissile phosphate

[1] These two authors should be regarded as joint first authors.
[2] Present address: School of Molecular and Cellular Biology, LC Miall Bldg., University of Leeds, LS2 9JT, United Kingdom.
[3] Present address: Reckitt Benckiser, Dansom La., Hull HU8 7DS, United Kingdom.
[4] To whom correspondence may be addressed. Tel.: 44-114-222-9478; E-mail: d.finger@sheffield.ac.uk.
[5] To whom correspondence may be addressed. Tel.: 44-114-222-9478; E-mail: b.ciani@sheffield.ac.uk.
[6] To whom correspondence may be addressed. Tel.: 44-114-222-9478; E-mail: j.a.grasby@sheffield.ac.uk.

[7] The abbreviations used are: FEN1, flap endonuclease-1; hFEN1, human FEN1; EXO1, exonuclease-1; GEN1, gap endonuclease-1; 2AP, 2-amino purine; DF, double-flap; SA, streptavidin; XPG, xeroderma pigmentosum complementation group G.

268

TABLE 1
Sequences of individual oligonucleotides used
FAM, 5′-fluorescein; **Biotin-TEG**, 5′-biotin connected through a tri(ethylene) glycol linker.

Oligo	Sequence	Construct
F₁	5′-**FAM** TTT TTG CGA CCT GGG CTG TGG AG-3′	DF1
T₁	5′-CTC CAC AGC CCA GGT CGC GAC GGT GAA ACC GTC C-3′	
F₂	5′-**Biotin-TEG**- TTT TTT TTT TAG TTA GGA CTC GTC ATC-3′	DF2
T₂	5′-GAT GAC GAG CAG TCC TAA CTG GAA ATC TAG CTC TGT GGA CGA AGT CCA CAG AGC TAG ATT TCC C-3′	
F₃	5′-TTT TTT TTT TAG TTA GGA CTG CTC GTC ATC-3′	DF3
T₂	5′-GAT GAC GAG CAG TCC TAA CTG GAA ATC TAG CTC TGT GGA CGA AGT CCA CAG AGC TAG ATT TCC C-3′	
F₄	5′-TTT TTG (**2AP**)(**2AP**)A GGC AGA GTG-3′	DF4
T₃	5′-CAC TCT GCC TTT CGA CAG CGA AGC TGT CC-3′	

diester bond to contact catalytic active site metal ions bound by a superfamily conserved carboxylate-rich active site (see Fig. 1, *B–D*).

FEN1 proteins have a structural feature known as the "helical arch" ($\alpha4$-$\alpha5$) that straddles the active site (see Fig. 1, *B–E*). The arch is partly disordered in substrate-free hFEN1 structures, as is the 3′-flap and duplex binding $\alpha2$-$\alpha3$ loop that packs against it when substrate is present (see Fig. 1*B*) (9, 14). The helical arch can be further subdivided into a region that contains two basic amino acids that are conserved in all 5′-nucleases (*i.e.* the base of $\alpha4$, see Fig. 1, *D* and *E*) and a part that is not conserved in sequence alignments in all superfamily members (*i.e.* the C-terminal region of $\alpha4$ and all of $\alpha5$; ~30 amino acids, see Fig. 1*E*). Here, we use the term helical arch to designate all of $\alpha4$-$\alpha5$, including the superfamily conserved region. We refer to the nonconserved 30 amino acids as the "helical cap" (6, 9, 15). When hFEN1 is bound to unpaired product, the nuclease domain of the protein is fully structured and the basic residues Lys-93 and Arg-100 contact the metal ion-bound 5′-phosphate monoester (see Fig. 1, *C* and *D*).

Possible roles for the helical arch and, in particular, its cap have received much attention (6, 8–10, 15–18). The helical cap is only present in 5′-nucleases hFEN1 and human EXO1 (hEXO1) that have a specificity for substrates that have 5′ termini-like flapped and gapped DNAs (6, 9, 15). The corresponding region in other family members is either much longer (XPG) or much shorter (GEN1) (see Fig. 1*E*). Biochemical data are consistent with a model where 5′-flaps are passed under the cap (15, 16). This would exclude processing of continuous DNA substrates such as bubbles and four-way junctions that have no 5′ termini, explaining the various substrate preferences within the superfamily (11, 15–19); however, this model is not universally accepted (10). It is also suggested that the helical cap may function to position superfamily conserved basic residues within the active site (6, 9, 16).

Here, to gain further insights into the role of the superfamily conserved $\alpha4$-basic residues and cap secondary structure in hFEN1 catalysis, we study the impact of mutations on substrate accommodation, unpairing, and hydrolysis. Mutation of superfamily conserved basic residues and introduction of helix breaking proline within the arch are severely detrimental to FEN1 function (3, 9). We demonstrate that these mutations

compromise neither protein stability nor the ability to bind substrate. Moreover, all the proteins can still accommodate 5′-flaps in a manner consistent with threading this portion of the substrate underneath the cap. Instead, we find that the structural integrity of the cap, but not the superfamily conserved portion of the helical arch, is necessary to unpair hFEN1 substrates.

EXPERIMENTAL PROCEDURES

Enzymes—Constructs for the preparation of mutant proteins were prepared using the pET28b-hFEN1 plasmid (1) and the protocol outlined in the QuikChange® site-directed mutagenesis kit (Stratagene/Agilent). hFEN1 proteins were expressed and purified as described previously (9) except for L130P and K93A/R100A, which were only expressed for 4 h at 18 °C.

DNA Constructs—Oligodeoxynucleotides including those synthesized with 5′-fluorescein-CE-phosphoramidite (for DF1), 5′-biotin connected through a triethylene glycol linker (biotin-TEG) phosphoramidite (for DF2), or containing site-specific 2-amino purine (2AP) substitutions (for DF4) were purchased with HPLC purification from DNA Technology A/S. DNA concentrations were determined by UV absorbance at 260 nm using extinction coefficients generated using OligoAnalyzer 3.1. Heteroduplex double-flap (DF) substrates were prepared by heating the appropriate flap strand with the complementary template in a 1:1.1 ratio at 80 °C for 5 min in 50 mM Tris-HCl, pH 7.5, and 100 mM KCl with subsequent cooling to room temperature (see Fig. 2A). Sequences are given in Table 1.

Maximal Single Turnover Rates—Maximal turnover rates of reaction were determined in triplicate using 2.5 nM DF1 and 1 μM hFEN1 or mutant protein in 55 mM HEPES, pH 7.5, 110 mM KCl, 8 mM MgCl₂, 1 mM DTT (dithiothreitol), and 0.1 mg/ml BSA using rapid quench apparatus or manual sampling where appropriate at 37 °C as described (1). Reactions were analyzed by denaturing HPLC equipped with a fluorescence detector (Wave® system, Transgenomic) as described (1, 4, 20, 21). The data for the formation of product over time (P_t) were modeled by nonlinear least squares regression in KaleidaGraph 4.0 using Equation 1 to determine the maximal first order rate constant (k) where P_∞ is the amount of product at end point.

$$P_t = P_\infty(1 - e^{-kt}) \tag{Eq. 1}$$

269

Fluorescence Anisotropy—Protein-DNA equilibrium dissociation constants were measured by fluorescence anisotropy using a Horiba Jobin Yvon FluoroMax-3® fluorometer with automatic polarizers. The excitation wavelength was 490 nm (excitation slit width 8 nm) with emission detected at 510 nm (10-nm-wide slit). Samples contained 2 mM EDTA or 10 mM CaCl$_2$, 10 nM DF1, 55 mM HEPES, pH 7.5, 110 mM KCl, 0.1 mg/ml BSA, and 1 mM DTT at 37 °C. The first measurement was taken prior to the addition of protein with subsequent readings taken on the cumulative addition of enzyme, with corrections made for dilution. Data were modeled by nonlinear least squares regression in KaleidaGraph 4.0 using Equation 2,

$$r = r_{min} + (r_{max} - r_{min})$$
$$\times \left(\frac{([S] + [E] + K_D) - \sqrt{([S] + [E] + K_D)^2 - 4[S][E]}}{2[S]} \right) \quad \text{(Eq. 2)}$$

where r is the measured anisotropy, $[E]$ is the total protein concentration, $[S]$ is the total DF1 concentration, r_{min} is the anisotropy of free DNA, r_{max} is the anisotropy of the DNA-protein complex, and K_D is the equilibrium dissociation constant. Each measurement was independently repeated in triplicate.

Protein Circular Dichroism Spectroscopy—CD measurements of hFEN1 and mutants thereof were performed using a 0.1-cm path length cuvette and samples containing 3 μM protein in 1 mM potassium phosphate, pH 7.5, 30.7 mM (NH$_4$)$_2$SO$_4$, 8 mM MgSO$_4$, 0.1 mM EDTA, and 0.25 mM tris(hydroxypropyl)phosphine. This buffer composition was used due to its low absorption in the near UV region (19, 22) and similar overall ionic strength to that used for determination of maximal single turnover rates of reaction. The CD spectra were collected between 190 and 260 nm in 1-nm steps with a response time and bandwidth of 8 s and 2 nm, respectively, at 20 °C using a JASCO J-810 CD spectrophotometer. Using the JASCO Spectra Analysis software (version 1.53.07), spectra were converted to mean residue molar ellipticity and smoothed by the means-movement method using a convolution width of five.

Fluorescence-detected Equilibrium Unfolding by Urea-induced Denaturation—Urea stock solution concentrations were determined by the method of Warren and Gordon (23). Prior to unfolding studies, proteins were subjected to size exclusion chromatography using 50 mM Tris, pH 8.5, 0.02% sodium azide, 1 mM DTT, and 100 mM KCl, and subsequently, concentration before determination of protein concentration by UV spectroscopy (A_{280}, $\epsilon = 22{,}920$ M^{-1} cm^{-1}). Urea-induced denaturation of hFEN1 proteins (1 μM) was conducted in 50 mM Tris, pH 8.5, 0.02% sodium azide, 1 mM DTT with either 100 mM KCl, 8 mM MgCl$_2$ or 120 mM KCl, 1 mM EDTA at 26 °C between 0 and 8 M urea in 0.5 M increments. Samples were prepared by removal of protein from the native protein stock and subsequent addition of an equal volume of denatured protein in buffer containing 8 M urea. At each urea concentration, the samples were allowed to attain equilibrium for 2 min, which was independently confirmed to be sufficient to reach equilibrium. Fluorescence measurements were performed on a Horiba Jobin Yvon FluoroMax-3® fluorometer using an excitation pulse at 295 nm and by collecting emission spectra using 10-nm slit widths

between 300 and 500 nm. The final emission spectra were an average of five emission spectra.

To obtain parameters for unfolding, the urea concentration dependence at the point of maximum change (362 nm) was fit to a two-state model appropriate for a monomer.

$$F = \frac{(B_N + b_N[D]) + (B_U + b_U[D])\exp\left[\frac{m([D] - [D]^{50\%})}{RT}\right]}{1 + \exp\left[\frac{m([D] - [D]^{50\%})}{RT}\right]}$$

(Eq. 3)

Here, B_N and B_U are the base-line CD values for the native (N) and denatured (D) states, b_N and b_U are the slopes of the pre- and post-transitions, $[D]^{50\%}$ is the denaturant concentration at which the protein is 50% unfolded (24), and the m-value is the slope of the unfolding transition. Data were modeled by nonlinear least squares regression using KaleidaGraph 3.6 and Equation 3 with only five parameters as free variables in the fit. Moreover, due to the lack of pre-transition base line, this slope parameter was excluded from the fit. Parameters for the unfolding data in Fig. 3B were not derived given the lack of pre- and post-transition base lines; thus, the fit to the data is merely cosmetic. $\Delta G_{D-N}^{H_2O}$ values were calculated from m-values and $[D]^{50\%}$ using Equation 4.

$$\Delta G_{D-N}^{H_2O} = m[D]^{50\%}$$

(Eq. 4)

Streptavidin Blocking/Trapping Experiments—All FEN1-substrate binding reactions were prepared with 1 nM of the biotinylated substrate DF2 in 50 mM HEPES-KOH, pH 7.5, 100 mM KCl, 8 mM CaCl$_2$, 0.25 mg/ml BSA, 2 mM DTT, and 5% glycerol. Where appropriate, the indicated protein (50 nM final concentration), streptavidin (Roche Diagnostics; 0.2 μg), and/or the nonbiotinylated competitor DNA (DF3; 250 nM final concentration) were added in specific orders to create blocked or trapped conditions. Blocked complexes refer to binding reactions that were prepared by incubating the substrate with streptavidin for 10 min at room temperature prior to the addition of the indicated hFEN1 protein. Trapped complexes refer to binding reactions that were prepared by incubating the substrate with the indicated hFEN1 protein for 10 min at room temperature before the addition of streptavidin to binding reactions. After the addition of the second component in both blocked and trapped binding reactions, the reactions were incubated for a further 10 min. For reactions where FEN1-substrate binding was challenged with DNA substrate lacking biotin, DF3 was added after the second incubation, and the mixture was allowed to incubate for 10 min at 37 °C.

Electrophoretic Mobility Shift Assay—The extent of binding in the various samples described above was assessed using a previously described electrophoretic mobility shift assay (EMSA) (16) in combination with the protocol of the chemiluminescent nucleic acid detection module kit (Thermo Scientific). Briefly, reaction mixtures (5 μl) were loaded onto native 6% polyacrylamide gels (29:1) containing 0.5× Tris-borate-EDTA while running at 200 V at room temperature and then further electrophoresed at 200 V until the bromphenol blue

270

indicator in a separate well reached the bottom of the gel. The gels were then electroblotted onto Biodyne B nylon membranes (Thermo Scientific) at 4 °C in 0.5× Tris-borate-EDTA at 380 mA for 30 min using a Mini Trans-Blot® module (Bio-Rad). The membranes were removed from the blot sandwich and were exposed face-down to a transilluminator equipped with 312-nm bulbs for 15 min to cross-link all complexes. The blots were then incubated in nucleic acid blocking buffer (Thermo Scientific) for 30 min, and subsequently, incubated in 1:300 dilution of stabilized streptavidin-horseradish peroxidase (SA-HRP) in nucleic acid blocking buffer for 30 min. The blots were washed five times in 20 ml of 1× wash buffer and once in 20 ml of substrate equilibration buffer (Thermo Scientific). Enhanced chemiluminescence (ECL) detection was accomplished using chemiluminescence substrate (Thermo Scientific) and a ChemiDoc XRS system (Bio-Rad).

2AP Exciton-coupled Circular Dichroism Spectroscopy (Low Energy CD)—Spectra were recorded on samples containing 10 μM DF4, 50 mM Tris-HCl, pH 7.5, 100 mM KCl, 1 mM DTT, either 10 mM CaCl$_2$ or 10 mM CaCl$_2$ + 25 mM EDTA and, where appropriate, 12.5 μM protein using a JASCO J-810 CD spectrophotometer (300–480 nm) at 20 °C as described (12). The CD spectra were plotted as $\Delta\epsilon$ per mol of 2AP residue *versus* wavelength. Each measurement was independently repeated (typically in triplicate) and gave similar results.

RESULTS

Arch Mutations Inhibit hFEN1-catalyzed Reactions—To investigate the role of conserved and nonconserved regions of the helical arch (α4-α5) in the hFEN1 catalytic pathway, we used leucine to proline mutations to disrupt or alter the structure. Earlier, these same point mutations were isolated from an *in vivo* screen of dominant-negative hFEN1 mutants that were toxic to yeast (3). The proline mutations were either at the base of α4 (L97P) between the two superfamily conserved basic residues (Lys-93 and Arg-100) or in the helical cap region at the top of α4 (L111P) and end of α5 (L130P) (Fig. 1, D and E). A failure to properly position the α4 and/or α5 helices could alter the presentation of superfamily conserved basic residues Lys-93 and Arg-100 into the active site. We, therefore, studied mutations of these residues to alanine (K93A, R100A) and the double mutant (K93A/R100A) alongside the proline-mutated hFEN1s.

The maximal rates of reaction of the mutated hFEN1s were characterized under single turnover conditions (*i.e.* enzyme excess), using an optimal, nonequilibrating (*i.e.* static) double-flap substrate, DF1 (Fig. 2, A and B, and Table 2). In line with their strict conservation throughout the 5'-nuclease superfamily, replacement of either Lys-93 or Arg-100 with alanine has previously been shown to be detrimental in several FEN1 family proteins (9, 10, 25). Here, under maximal single turnover conditions, removal of either of the basic side chains decreased the rate of reaction by factors of ~2000-fold (K93A) and 7000-fold (R100A), respectively (Fig. 2B). Interestingly, the magnitudes of the decreases in rate of reaction relative to WT hFEN1 were much greater under maximal single turnover conditions than when the catalytic efficiencies (k_{cat}/K_m) of the enzymes were measured (9). This likely arises because, at low concentrations

of substrate (i.e. k_{cat}/K_m conditions), the rate of the wild WT hFEN1 reaction is rate-limited by encounter of enzyme and substrate and not by the proceeding steps (25). The double mutation K93A/R100A also decreased the rate of reaction by 3 orders of magnitude, producing a similar rate constant to R100A.

Unsurprisingly, the mutation L97P, which is located between the two 5'-nuclease conserved basic residues, was highly detrimental to catalysis, drastically reducing the rate of reaction >100,000-fold. L97P is by far the most catastrophic point mutation of hFEN1 studied to date, and its impact cannot be accounted for solely by preventing participation of the key basic residues in catalysis. As L111P and L130P mutations are contained in the helical cap region of the protein that are not superfamily conserved, they were expected to have less of an impact on the rate of reaction. Nevertheless, L111P was also severely detrimental to the reaction, decreasing the maximal rate by a factor of ~15,000-fold. Similarly, L130P decreased the rate constant ~3,000-fold, further underscoring the importance of cap structural integrity to hFEN1-catalyzed reactions. Furthermore, the severities of rate reduction correlate with the relative severities of the reported dominant-negative phenotypes of these mutations upon overexpression in yeast (3).

Arch Mutations Do Not Alter Substrate Affinity—To examine the impact of the arch mutations on substrate binding, the dissociation constants of the respective proteins were measured by fluorescence anisotropy using DF1, which was synthesized with a 5'-fluorescein label (Fig. 2, A and C, and Table 1). In the presence of EDTA, the WT protein and all mutants had similar dissociation constants, showing only 2-fold variation. Calcium ions do not support hFEN1 catalysis and are a competitive inhibitor of FEN1 family reactions with respect to viable cofactor ions such as Mg^{2+}, implying that they occupy the same active site positions (7, 26). However, even with occupation of the carboxylate-rich active site by Ca^{2+} ions, the dissociation constants of the WT and mutated hFEN1 complexes were relatively unchanged.

Proline Mutations Do Not Destabilize hFEN1—To investigate the impact of proline mutations on protein structure, we measured the CD spectra of the proteins in the presence of Mg^{2+} ions (Fig. 3A) and monitored unfolding of the proteins in the presence of urea (Fig. 3, B and C). The CD spectra of each of the mutated proteins closely resembled that of WT hFEN1, thereby demonstrating that no gross changes to the secondary structure of the proteins occurred. Equilibrium unfolding studies of the mutants were monitored by tryptophan fluorescence using urea as the denaturant. Data were fitted assuming two-state unfolding (Equation 3). The Trp residues of hFEN1 are buried in the main saddle region of the protein on either side of the β-sheet (Fig. 1, A and B). In the presence of EDTA, the denaturation curves of WT and mutant enzymes lacked pre- and post-transition base lines, which is indicative of the absence of cooperativity during the unfolding transitions (Fig. 3B). The addition of magnesium ions induced cooperativity in both the WT and the mutant equilibrium unfolding, which is consistent with the higher melting temperatures observed for WT samples monitored by differential scanning fluorometry under similar conditions (data not shown). In the presence of Mg^{2+} ions, the
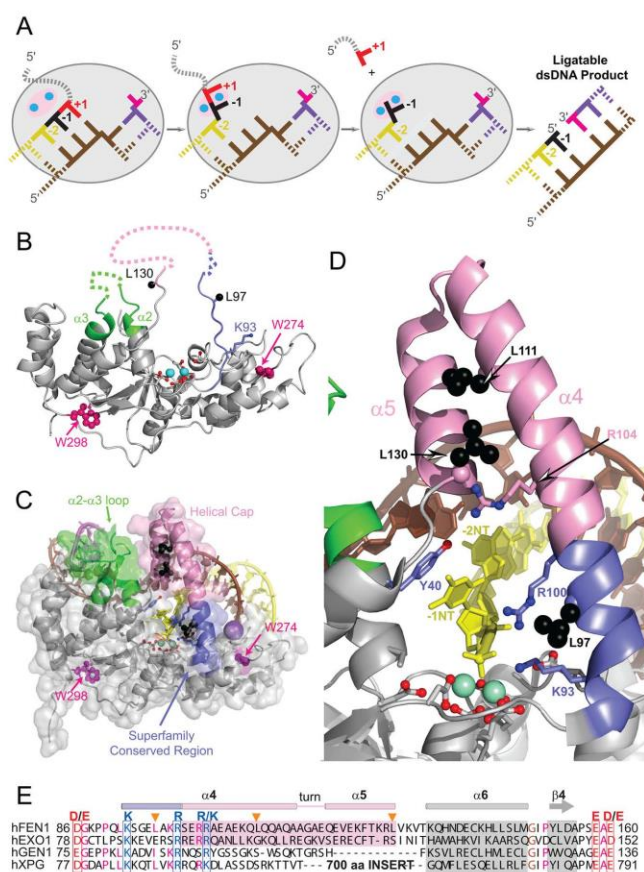
271

FIGURE 1. **hFEN1 structures, helical arch/cap structural transitions, and relationship to the other members of the 5′-nuclease superfamily.** *A*, schematic of an equilibrating double flap undergoing double nucleotide unpairing to allow hFEN1-catalyzed reaction, thereby generating a nicked DNA product. *B*, graphic representation of hFEN1 without substrate (1UL1:X) highlighting the disordered arch and α2-α3 loop. Superfamily conserved region of the arch (*blue*) and helical cap (*pink*) are shown with missing regions as *dotted lines*. Active site metal ions (*cyan*), carboxylates (*red*), and α2-α3 loop (*green*) are also highlighted. The side chains of the superfamily conserved basic residues Lys-93 and Arg-100 (*blue*), the two hFEN1 Trp residues (Trp-274, Trp-298; *dark pink*), and the positions of Leu residues that are mutated in this study (L97P, L111P, L130P, *black*) are shown (when defined in this structure). *C*, transparent surface representation of hFEN1 bound to product (3Q8K) with unpaired −1 nucleotide (*yellow*) with protein colored as in *B*. *D*, magnified view of the active site in 3Q8K and helical arch/cap with protein and residues colored as in *B* and *C*. hFEN1 conserved Tyr 40 (*blue*) is also shown. *E*, sequence alignment of hFEN1, human EXO1 (*hEXO1*), human XPG (*hXPG*), and human GEN1 (*hGEN1*) between Asp-86 and Asp-160 colored as in *B* highlighting conservation of basic amino acids and lack of conservation of the helical cap in hGEN1 and hXPG. The positions of the Leu residues mutated in this study are highlighted with an *orange triangle*.

midpoints of transition for urea denaturation ([D]$^{50\%}$, Table 2) were similar or higher than the WT and mutated proteins, but were surprisingly low for an ~40-kDa protein (27).

Arch Mutants Can Still Accommodate 5′-Flaps—Earlier work (15, 16) devised a methodology that allows the ability to properly accommodate the 5′-flap portion of substrates to be assessed. Here, "5′-flap accommodation" refers to positioning the 5′-flap in a manner that can readily proceed to reaction in the WT protein (15, 16). When streptavidin (SA) is added to 5′-biotinylated substrates in complex with WT hFEN1-Ca^{2+}, the majority of the DNA can be trapped on the protein (15),

indicating that the 5′-flap is trapped in a conformation that easily proceeds to reaction because the addition of Mg^{2+} ions results in reaction rates comparable with those of unmodified complexes. This is the case even when the trapped complex is challenged with excess unlabeled substrate before the addition of viable cofactor, demonstrating that trapped DNA cannot exchange. In contrast, when SA is conjugated to substrate before interaction with hFEN1-Ca^{2+}, termed a blocked complex, reaction is drastically slowed. Blocked complexes readily dissociate in the presence of excess substrate, as do unmodified complexes (15, 16). These observations are interpreted as indi-
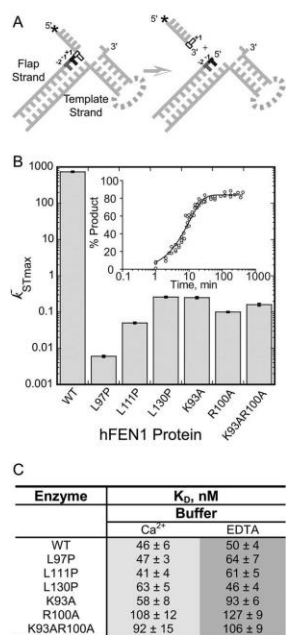
272

FIGURE 2. **Catalysis and substrate binding of arch hFEN-1 mutants.** *A*, schematic of the DF substrates used in this study illustrating the hFEN1-catalyzed reaction that occurs between the +1 and −1 nucleotides. Each static DF is composed of a flap and template strand, where 3′- and 5′-flaps are not complementary to template strand. Full sequence information is available in Table 1. In DF1, * = 5′-fluorescein; in DF2, * = 5′-biotin; in DF3, * = OH (unmodified); and in DF4, * = OH with both the −1 and −2 nucleobases as 2AP. *B*, maximal single turnover rates (k_{STmax}) of WT and mutant hFEN1s (note log scale) with standard errors (S.E.) shown, measured at 37 °C and pH 7.5 using DF1. The *inset* shows example data for hFEN1 R100A fitted to Equation 1, and further details are under "Experimental Procedures." *C*, substrate dissociation constants (± S.E.) in the presence of Ca^{2+} or EDTA for WT and mutant hFEN1s as determined by fluorescence anisotropy using DF1 at 37 °C and pH 7.5.

TABLE 2

Denaturation parameters at 25 °C for hFEN1 and proline mutants

Parameters were calculated fitting raw fluorescence data to a 2-state unfolding model in urea with sloping baselines (Equations 3 and 4).

Sample	$[D]^{50\%}$ ± S.E.[a]	*m*-value ± S.E.[a]	ΔG^{H_2O} ± S.E.[a]
	M	*kcal mol*$^{-1}$ *M*$^{-1}$	*kcal mol*$^{-1}$
Wild type	2.26 ± 0.12	0.66 ± 0.08	1.49 ± 0.20
L97P	2.80 ± 0.17	0.63 ± 0.09	1.76 ± 0.27
L111P	3.40 ± 0.27	0.50 ± 0.07	1.70 ± 0.27
L130P	2.27 ± 0.07	0.92 ± 0.09	2.10 ± 0.21

[a] S.E. is the standard error of the fit.

cating that a nonblocked 5′-flap is able to pass under the helical cap and is able to be trapped in this state by SA, a conclusion supported by recent single-molecule studies (15, 16, 19).

As the rates of reaction of the mutants studied here were very slow, we examined the ability of WT and mutated proteins to form blocked and trapped complexes with 5′-biotinylated substrate DF2 by EMSA (16). Thus, we prepared two types of enzyme-Ca^{2+}-5′-SA-DF2 complexes (Fig. 4*A*). A blocked com-
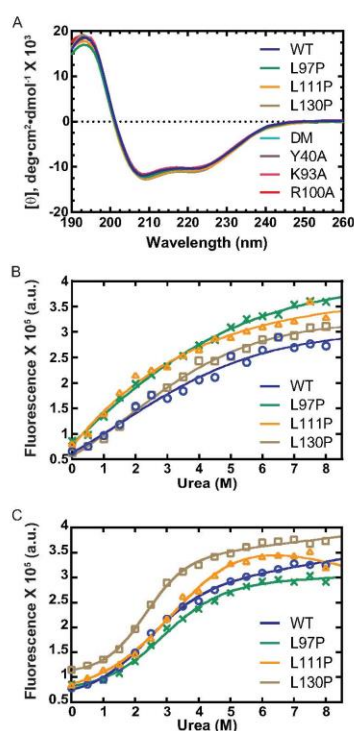

FIGURE 3. **Structure and stability of WT and mutated hFEN1s.** *A*, CD spectra of WT and mutated hFEN1 proteins (see *legend inset*) at pH 7.5 and 20 °C in Mg^{2+}-containing buffer. *B*, urea denaturation curves of WT hFEN1 and proline mutants thereof (see *legend inset*) in the absence of metal ions and monitored by Trp fluorescence emission at 362 nm at 26 °C, pH 8.5. For clarity, data are presented with a cosmetic fit to Equation 3 without including pre- and post-transition base lines. (*a. u.* = arbitrary units) *C*, urea denaturation of WT hFEN1 and proline mutants thereof (see *legend inset*) in the presence of Mg^{2+} and monitored by Trp fluorescence emission at 362 nm at 26 °C, pH 8.5. Data are fitted to Equation 3 without including the pre-transition base-line slope to afford the midpoints of transition, $[D]^{50\%}$, which are reported in Table 2. Full details of conditions are provided under "Experimental Procedures."

plex was formed by conjugation of SA to DF2 prior to the addition of the enzyme, preventing the passage of the 5′-flap underneath the cap. A second complex was formed by prior assembly of 5′-biotinylated hFEN1-Ca^{2+}-DF2 complex followed by the addition of SA to trap the DNA in this conformer. In the trapped complexes, the 5′-flap could still have passed under the cap. Complexes were detected using an SA-HRP conjugate and enhanced chemiluminescence reagents (Thermo Scientific), which together detect biotinylated DNAs (Fig. 4*B*). Before developing this method, we rationalized that because some of our samples already contained biotinylated oligonucleotide conjugated to SA, the SA-HRP conjugate would be unable to detect any complexes of DF2 to which SA had been added prior to electrophoresis. Thus, we initially detected biotinylated DF2 complexes indirectly using an anti-SA antibody that was con-
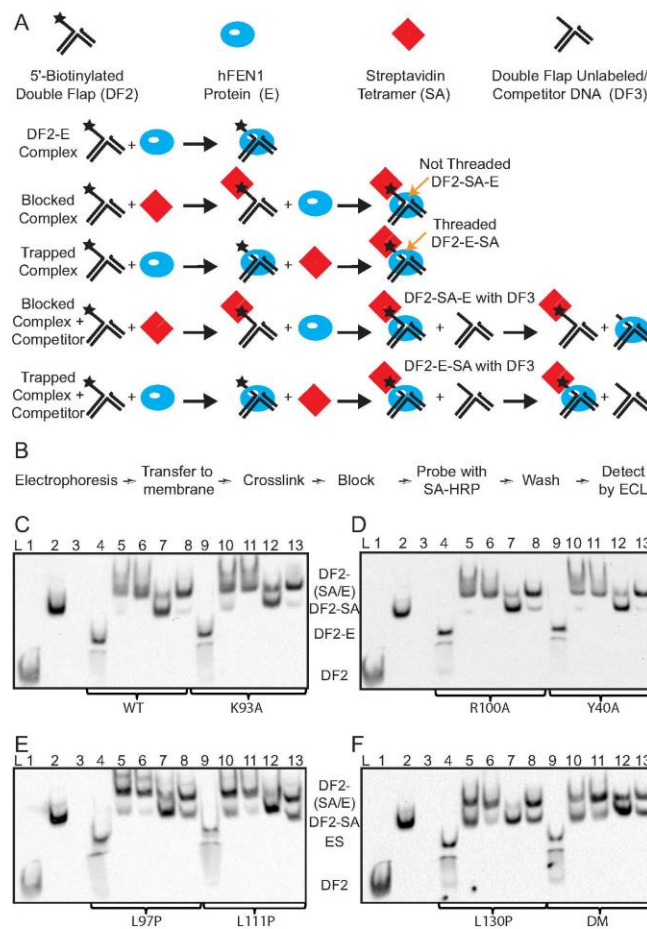
273

FIGURE 4. **5'-Flap accommodation assessment by EMSA.** *A*, schematic of procedure for the formation of complexes. Trapped complexes were assembled at 20 °C by the addition of SA to an E-DF2 complex, whereas blocked complexes involved the addition of E to SA-DF2. Competition experiments were carried out with 5-fold excess nonbiotinylated DF3 at 37 °C for 10 min. *B*, workflow diagram for detection of biotinylated oligonucleotides using SA-HRP conjugate. *C–F*, flap accommodation assessed by EMSA that was detected using the SA-HRP conjugate. *Lane 1*, DF2; *lane 2*, DF2-SA; *lane 3*, DF3 with SA (control, no signal observed without biotinylated DNA); *lanes 4* and *9*, DF2-E; *lanes 5* and *10*, blocked DF2-SA-E; *lanes 6* and *11*, trapped DF2-E-SA; *lanes 7* and *12*, blocked DF2-SA-E competed with DF3; *lanes 8* and *13*, trapped DF2-E-SA competed with DF3. *C*, WT hFEN1 (*lanes 4–8*) and K93A (*lanes 9–13*). *D*, R100A (*lanes 4–8*) and Y40A (*lanes 9–13*). *E*, WT L97P (*lanes 4–8*) and L111P (*lanes 9–13*). *F*, L130P (*lanes 4–8*) and K93A/R100A (*DM*) (*lanes 9–13*).

jugated to HRP (Abcam). Although this method gave a high background, we were able to detect the various SA complexes in addition to the excess SA in samples to which it was added (data not shown). However, when we probed additional blots of the same samples with the SA-HRP conjugate, we were surprisingly able to observe all types of DF2 complexes in the context of a low background (Fig. 4, *C* and *F*). Despite the saturation of biotinylated oligonucleotides with SA in specific samples, the SA-HRP conjugate still detected all the biotinylated species after cross-linking/blotting. It seems likely that the cross-link-

ing procedure results in denaturation of the SA bound to the biotinylated oligonucleotides in our electrophoresed samples.

Blocked (Fig. 4, *C–F*, *lanes 5* and *10*) and trapped (Fig. 4, *C–F*, *lanes 6* and *11*) SA complexes had similar electrophoretic mobility, but blocked and unmodified complexes were less stable under gel conditions. Blocked and trapped complexes were then challenged with excess unlabeled (*i.e.* no 5'-biotin moiety) competitor substrate, DF3 (Fig. 4, *C–F*, *lanes 7* and *12* (blocked) and *lanes 8* and *13* (trapped)) prior to electrophoresis. Regardless of hFEN1 protein, we found that blocked complexes were

readily competed away, whereas the majority of trapped complexes were not displaced, as was observed earlier for the WT enzyme (15, 16). The degree of competition of trapped complexes does differ for some of the mutants, namely L97P, L111P, L130P, and K93A/R100A. We suspect that this indicates a change in the equilibrium constant between threaded and unthreaded forms of the enzyme-substrate complex. Nevertheless, all the mutated proteins could produce a trapped complex upon the addition of SA, suggesting that the ability to properly accommodate the 5'-flap substrate under the helical arch is not prevented.

Helical Cap Proline Mutants Cannot Unpair Substrates—Because the proline mutations neither altered the ability to bind substrates nor impaired protein stability and each protein was able to accommodate a 5'-flap, we asked whether the mutant FENs could induce substrate conformational changes essential to the hFEN1 catalytic cycle (Fig. 1*A*). Previously, we monitored the ability of WT hFEN1-Ca^{2+} to unpair DNAs by reduction in the low energy CD signal resulting from the presence of a 2AP exciton pair in the substrate (12). Both protein and the natural nucleotides of the DNA have been shown to be transparent in this region of the CD spectrum. The adjacent 2APs were located at positions -1 and -2 in DF4. In the double nucleotide unpairing model, the -2 nucleotide will remain base-paired when complexed with hFEN1, whereas the -1 nucleotide will become extrahelical (Figs. 1*A* and 2*A* and Table 1). The addition of hFEN1 in the presence of Ca^{2+} ions, but not the presence of hFEN1 alone (EDTA), was necessary to effect a reduction in CD signal at 330 nM (Fig. 4, *B* and *C*). Extending these studies to the mutated proteins, we found that the addition of L97P-Ca^{2+} ions produced a decrease in the exciton pair signal, which indicated that the substrate had undergone an analogous conformational change to that when bound to WT hFEN1-Ca^{2+} (Fig. 5*A*). This reduction in signal was strictly dependent on the presence of Ca^{2+} as it is for the WT protein (Fig. 5*B*). Interestingly, the K93A/R100A double mutation also produced an altered CD signal in the presence of Ca^{2+} ions. However, a deep minima was observed, analogous to the signal detected previously with R100A and Y40A mutated hFEN1s (12). This implied that although the nucleotides positioned at -1 and -2 were reoriented in the K93A/R100A-Ca^{2+} complex, their juxtaposition at equilibrium differed from that afforded by the WT protein. In contrast, modest enhancements rather than reduction in low energy CD signals at 330 nM were observed with L111P-Ca^{2+} and L130P-Ca^{2+} when compared with the free substrate. Furthermore, the cap proline mutants afforded low energy CD signals that were very similar $\pm Ca^{2+}$ (Fig. 4, *B* and *C*). This indicated that even in the presence of divalent metal ions, L111P and L130P lacked the ability to unpair a detectable amount of the DNA substrate.

DISCUSSION

Although biochemical and structural studies suggest a role for the conserved basic residues of the helical arch in 5'-nuclease catalysis, their precise function has remained elusive. These residues have been hypothesized to play a key role in capturing the unpaired state of substrates in FEN1 family proteins (9). In addition, the role of the hFEN1 helical cap, absent in
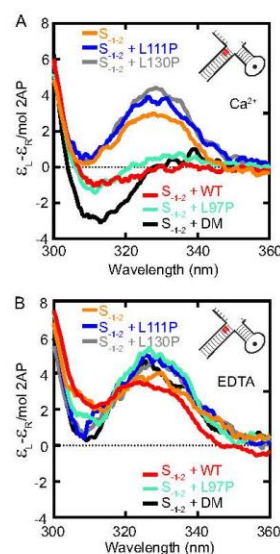
FIGURE 5. **Substrate unpairing by WT and mutated hFEN1 proteins.** *A* and *B*, low energy CD spectra of 2AP dimer containing DF4 and protein-DF4 complexes with divalent metal ions (10 mM $CaCl_2$) (*A*) or without available divalent ions (10 mM $CaCl_2$ + 25 mM EDTA) (*B*). Samples were buffered at pH 7.5, and measurements performed at 20 °C. DF4 is illustrated schematically with the position of 2APs (full sequence is shown in Table 1).

some superfamily members, has also remained enigmatic (6, 8–10, 15–18). It has been argued that the helical cap of the protein may serve to position the required basic superfamily residues, while ensuring specificity for substrates such as flaps with 5' termini (6, 9, 16). However, the results herein imply more complex roles for both superfamily conserved basic residues and the helical cap in the hFEN1 catalytic cycle.

In line with the observation that L97P, L111P, and L130P produce a dominant-negative phenotype in yeast, we observe negligible changes in the affinity of substrates for the mutated proteins (3). Similarly, we also demonstrate the ability to trap 5'-biotinylated substrate with streptavidin on all the mutated proteins. Moreover, each trapped complex also persisted in the presence of excess unlabeled substrate. This strongly suggests that the 5'-flap of substrate has been appropriately accommodated. Furthermore, it demonstrates that the ability to produce a trapped conformation does not require $\alpha4$-$\alpha5$ to adopt a defined structure. In fact, earlier experiments using 5'-flaps with duplex regions too large to pass through a structured arch demonstrate that threading the 5'-flap requires the arch to be flexible as it is in the substrate-free enzyme (15). In line with this expectation, proline mutations that grossly impact the secondary structure of $\alpha4$ and $\alpha5$ have negligible effects on the ability to accommodate 5'-flaps of substrates.

The substrate conformational change that hFEN1 facilitates to transfer the scissile phosphate diester to active site metal ions involves double nucleotide unpairing of duplex, a process that places the unpaired substrate within the helical arch (9, 10, 12,

19). As shown recently, Lys-93 and Arg-100 are not required for this substrate rearrangement when active site divalent metal ions are present (12). As we show here, even the introduction of proline in between these two basic residues (L97P), a mutation expected to considerably change the positioning of Lys-93 and Arg-100 side chains, does not prevent substrate unpairing. Even when both side chains are removed in K93A/R100A, DNA conformational rearrangement still occurs, although the orientation of the −1 and −2 nucleotides differs from that of DNA in WT complex. These observations rule out models where arch basic residues are required to stabilize an unpaired state. However, as these mutations to basic residues produce severely retarded reaction rates, either the substrate remains incorrectly positioned for reaction to proceed without conserved residues in place and/or the superfamily conserved basic residues are required to directly participate in catalysis.

The helical cap mutations, L111P and L130P, inhibit hFEN1 catalysis at a different catalytic step to L97P. These mutated hFEN1s are surprisingly unable to unpair any detectable amount of their DF substrates. This occurs despite evidence that divalent metal ions alter the urea-induced denaturation behavior of L111P and L130P, suggesting that the ability to bind the requisite divalent active site ions is not perturbed. In hFEN1-product structures where the hydrolyzed phosphate monoester contacts active site metal ions, the unpaired nucleobase is constrained within a cavity formed by Tyr-40 in $\alpha2$ and residues from the C-terminal end of $\alpha5$ (Fig. 1, C and D). However, although −1 nucleotide is stacked on Tyr-40, there are no direct helical cap-unpaired nucleobase interactions. Residue Leu-130 is at the C-terminal end of $\alpha5$; as such, a proline substitution here could have disrupted the structure in this region. However, it is less straightforward, but more intriguing, to explain why the L111P mutation prevents substrate unpairing. This could be related to a role played by Leu-111 in stabilizing the two-helix cap structure through the formation of a mini hydrophobic $\alpha4$-$\alpha5$ helix-helix core. Moreover, $\alpha5$ packs against $\alpha2$ and the $\alpha2$-$\alpha3$ loop when substrate is present. These same regions of the protein are partly disordered in structures without DNA (Fig. 1, B and C), but form interactions with substrate when present. An altered helical cap region could sterically exclude substrate unpairing by preventing accommodation of the 5′-flap through the arch; however, the ability to form trapped complexes with these mutants suggests that the 5′-flap of the substrate can pass under the helical cap, at least while duplex remains in a paired state. Alternatively, the inability to rearrange the DNA close to active site metal ions suggests a role for the helical cap in formation of unpaired DNA, either directly by the requirement to restrain one or both of the unpaired nucleobases and/or by participation in protein conformational changes necessary to afford the unpaired state.

Interestingly, although K93A and R100A independently have large impacts on the ability to catalyze the reaction, the double mutant K93A/R100A is no more detrimental to catalysis than R100A alone. This indicates cooperation between the basic residues of the active site because loss of Arg-100 interaction impacts on the ability of Lys-93 to participate in the reaction. A similar phenomenon has also been observed independently in time-resolved fluorescence studies of 2AP substrates. Interac-

tion of FEN1-conserved Tyr-40 of $\alpha2$ with the −1 nucleobase could not be detected in R100A and K93A complexes (12). Despite this evidence for the cooperative assembly of conserved residues around unpaired DNA, the impact of the secondary structure altering L97P is notably much greater than those of removal of the side chains of the basic conserved residues Lys-93 and Arg-100. This implies that a lack of these interactions is not the only consequence of L97P mutation. It would appear that conformational rigidity of proline locks the hFEN1-DNA complex into a conformation that is catalytically compromised in either a substrate-unpaired state (L97P) or a base-paired state (L111P and L130P).

The finding that perturbation of the helical cap can prevent double nucleotide unpairing of substrate raises questions about superfamily members XPG and GEN1, which lack this feature. These enzymes have specificities for reaction one nucleotide into the double strand of their respective substrates, bubbles, and four-way-junctions, suggesting that they also unpair DNA. It seems likely that these larger enzymes or their partner proteins make alternative interactions to position unpaired DNA within the active site. The addition of an archaeal helical cap together with removal of the 650-residue spacer domain of XPG did confer the ability to process substrates with free 5′ termini to the XPG-FEN1 chimera XPF2 (9). Unlike bubble-incising XPG, XFX2 is a reasonably proficient flap endonuclease, although it lacks the single scissile phosphate specificity of FEN1.

In conclusion, the helical cap is revealed to have a more complex role in the hFEN1 catalytic cycle in addition to its roles in positioning of essential residues or acting as a barrier to DNA substrates without 5′ termini. A role for substrate unpairing in threading flaps under the cap is ruled out by our studies; instead, 5′-flap accommodation, which does not require a unique helical arch conformation, must precede DNA conformational change. Unpairing does not require conserved basic residues, although they are essential to catalyze reactions on a biologically relevant time scale. It is interesting that despite a requirement to unpair DNAs, hFEN1 does not process substrates with non-Watson Crick base pairs at the end of reacting duplexes efficiently (13, 19). Indeed, hFEN1 appears to initially stabilize the base-paired state by stacking interactions with $\alpha2$, thereby selecting against substrates with frayed ends. There is an intriguing association between an intact helical cap and substrate conformational change, alongside an unusually low stability of the protein to denaturant and the ability of rigid proline mutations to inhibit reaction. Altogether, these data suggest that conformational dynamics of hFEN1-substrate complexes are likely to play a role in creating the unpaired state.

REFERENCES

1. 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
2. Lyamichev, V., Brow, M. A., Varvel, V. E., and Dahlberg, J. E. (1999) Com-

276

parison of the 5′ nuclease activities of *Taq* DNA polymerase and its isolated nuclease domain. *Proc. Natl. Acad. Sci. U.S.A.* **96**, 6143–6148

3. 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 replication and repair. *EMBO J.* **21**, 5930–5942

4. 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

5. Finger, L. D., Atack, J. M., Tsutakawa, S., Classen, S., Tainer, J., Grasby, J., and Shen, B. (2012) The wonders of flap endonucleases: structure, function, mechanism, and regulation. in *The Eukaryotic Replisome: A Guide to Protein Structure and Function* (MacNeill, S., ed), pp. 301–326, Springer Publishing, New York

6. Grasby, J. A., Finger, L. D., Tsutakawa, S. E., Atack, J. M., and Tainer, J. A. (2012) Unpairing and gating: sequence-independent substrate recognition by FEN superfamily nucleases. *Trends Biochem. Sci.* **37**, 74–84

7. 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

8. 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

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

10. 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

11. Devos, J. M., Tomanicek, S. J., Jones, C. E., Nossal, N. G., and Mueser, T. C. (2007) Crystal structure of bacteriophage T4 5′ nuclease in complex with a branched DNA reveals how FEN-1 family nucleases bind their substrates. *J. Biol. Chem.* **282**, 31713–31724

12. Finger, L. D., Patel, N., Beddows, A., Ma, L., Exell, J. C., Jardine, E., Jones, A. C., and Grasby, J. A. (2013) Observation of unpaired substrate DNA in the flap endonuclease-1 active site. *Nucleic Acids Res.* 10.1093/nar/gkt737

13. Beddows, A., Patel, N., Finger, L. D., Atack, J. M., Williams, D. M., and Grasby, J. A. (2012) Interstrand disulfide crosslinking of DNA bases supports a double nucleotide unpairing mechanism for flap endonucleases. *Chem. Comm.* **48**, 8895–8897

14. 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

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

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

17. 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

18. 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

19. Sobhy, M. A., Joudeh, L. I., Huang, X., Takahashi, M., and Hamdan, S. M. (2013) Sequential and multistep substrate interrogation provides the scaffold for specificity in human flap endonuclease 1. *Cell Reports* **3**, 1785–1794

20. 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

21. 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

22. Kelly, S. M., Jess, T. J., and Price, N. C. (2005) How to study proteins by circular dichroism. *Biochim. Biophys. Acta* **1751**, 119–139

23. Warren, J. R., and Gordon, J. A. (1966) On the refractive indices of aqueous solutions of urea. *J. Phys. Chem.* **70**, 297–300

24. Clarke, J., and Fersht, A. R. (1993) Engineered disulfide bonds as probes of the folding pathway of barnase: increasing the stability of proteins against the rate of denaturation. *Biochemistry* **32**, 4322–4329

25. 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

26. Tomlinson, C. G., Syson, K., Sengerová, B., Atack, J. M., Sayers, J. R., Swanson, L., Tainer, J. A., Williams, N. H., and Grasby, J. A. (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

27. Fersht, A. (1999) *Structure and Mechanism in Protein Science: Guide to Enzyme Catalysis and Protein Folding*, 2nd Ed., W.H. Freeman & Co Ltd.

277