*3332–3338* *Nucleic Acids Research, 1997, Vol. 25, No. 16* 

# Functional analysis of point mutations in human flap endonuclease-1 active site

**Binghui Shen[1], John P. Nolan[2,3], Larry A. Sklar[3,4] and Min S. Park[2,*]**

[1]Department of Cell and Tumor Biology, City of Hope National Medical Center, 1500 East Duarte Road, Duarte, CA 91010-0269, USA, [2]Life Sciences Division, M888 and [3]National Flow Cytometry Resource, Los Alamos National Laboratory, Los Alamos, NM 87545, USA and [4]Cytometry, School of Medicine, University of New Mexico, Albuquerque, NM 87131, USA

Received April 2, 1997; Revised and Accepted July 2, 1997

## ABSTRACT

**Human flap endonuclease-1 (hFEN-1) is highly homologous to human XPG, *Saccharomyces cerevisiae* RAD2 and *S.cerevisiae* RTH1 and shares structural and functional similarity with viral exonucleases such as T4 RNase H, T5 exonuclease and prokaryotic DNA polymerase 5′ nucleases. Sequence alignment of 18 structure-specific nucleases revealed two conserved nuclease domains with seven conserved carboxyl residues and one positively charged residue. In a previous report, we showed that removal of the side chain of each individual acidic residue results in complete loss of flap endonuclease activity. Here we report a detailed analysis of substrate cleavage and binding of these mutant enzymes as well as of an additional site-directed mutation of a conserved acidic residue (E160). We found that the active mutant (R103A) has substrate binding and cleavage activity indistinguishable from the wild type enzyme. Of the inactive mutants, one (D181A) has substrate binding properties comparable to the wild type, while three others (D34A, D86A and E160A) bind with lower apparent affinity (2-, 9- and 18-fold reduced, respectively). The other mutants (D158A, D179A and D233A) have no detectable binding activity. We interpret the structural implications of these findings using the crystal structures of related enzymes with the flap endonuclease activity and propose that there are two metal ions ($Mg^{2+}$ or $Mn^{2+}$) in hFEN enzyme. These two metal coordinated active sites are distinguishable but interrelated. One metal site is directly involved in nucleophile attack to the substrate phosphodiester bonds while the other may stabilize the structure for the DNA substrate binding. These two sites may be relatively close since some of carboxyl residues can serve as ligands for both sites.**

## INTRODUCTION

A structure-specific flap endonuclease-1 (FEN-1) activity in mammalian cells was first identified by Harrington and Lieber (1). Once the mouse FEN-1 cDNA, deduced amino acid sequence and specific biochemical activities were available, mammalian enzymes purified from murine cells (2), human cells (3–6) and calf thymus (7) have been re-identified as the same enzyme as FEN-1. The 42 kDa enzyme has been demonstrated to have two activities: (i) a flap endonuclease (1,8,9) and (ii) a nick-specific 5′→3′ exonuclease (1,4,10,11). The enzyme specifically recognizes a 5′ DNA flap structure, which has been proposed to exist in replication, repair and recombination. It slides through the 5′ free arm (12) and then removes the unannealed region as an intact segment via a single endonucleolytic cleavage. The calf FEN-1 5′→3′ exonuclease cooperatively functions with calf DNA polymerase to perform nick translation (7). FEN-1 has also been shown to be a necessary component in DNA replication. It removes the last base of RNA primer attached to an Okazaki fragment after RNase H action in the SV40 *in vitro* replication system (4).

Genetic studies showed that deleting FEN-1 homologues in the yeasts *Saccharomyces cerevisiae* and *Schizosaccharomyces pombe* resulted in a marked sensitivity to the alkylating reagent methylmethane sulfonate (MMS), modest UV sensitivity and chromosome instability (5,13–15). These data imply that the protein is involved in fundamental processes such as DNA replication, repair and recombination. The FEN-1 gene is required for the stability of the simple repetitive DNA in the yeast system. A FEN-1 null mutation was shown to increase the mutation rate of simple repetitive DNA by as much as 280 times and to increase the spontaneous mutation rate 30-fold (15). More recently, it has been shown in the yeast *S.cerevisiae* that the majority of the resulting mutations have a structure in which sequences ranging from 5 to 108 bp flanked by direct repeats of 3–12 bp are duplicated. Epistasis analysis indicates that FEN-1 does not play a major role in MSH2-dependent mismatch repair, but it can contribute to regional duplication in the genome, which in turn can lead to genomic instability (16).

Human FEN-1 is highly homologous to human XPG, *S.cerevisiae* RAD2 and *S.cerevisiae* RTH1 and shares structural and functional similarity with viral exonucleases such as T4 RNase H and T5 exonuclease and prokaryotic DNA polymerase 5′ nucleases. Previously, we identified several critical residues in flap substrate catalysis or binding (17). In this report, we further characterize site-directed mutants using a kinetic approach based on flow cytometry (18). We estimated the binding affinities of catalytically-deficient mutants (D34A, D86A, E160A and D181A) and confirmed that removal of conserved arginine 103 did not affect either substrate binding or catalysis. In addition, we have identified

*To whom correspondence should be addressed. Tel: +1 505 667 5701; Fax: +1 505 665 3024; Email: park@telomere.lanl.gov

and characterized another acidic amino acid residue (E160) involved in $Mg^{2+}$ coordination. We interpret the structural implications of these findings using the crystal structures of related enzymes with the flap endonuclease activity and propose that there are two metal ions ($Mg^{2+}$ or $Mn^{2+}$) in hFEN enzyme with distinguishable functions.

## MATERIALS AND METHODS

### Materials

Wild type recombinant human flap endonuclease was purified as described previously (18). All biological and chemical reagents were obtained from vendors listed previously (17).

### Sequence alignment and molecular superimposition

All sequence comparison and alignments were performed with DNAStar software (DNAStar, Inc., Madison, WI), and all necessary DNA sequences were retrieved from GenBank. Molecular modeling was carried out in Mark Sherman's laboratory at the City of Hope Cancer Center core facility using the T4 RNase H and T5 exonuclease-1 coordinates available in the Brookhaven data bank (Brookhaven file 1TFR; 19,20). The crystallographic coordinates were visualized on an Indigo High Impact workstation (Silicon Graphics) running commercially available software (Insight II, Biosym Technologies).

### Site-directed mutagenesis, protein expression and purification

We removed all seven residues listed in Table 1 by site-directed mutagenesis. In addition, R103 is another conserved amino acid residue in these 18 nucleases. It is replaced by a lysine in the lower organisms. Alanine has been chosen to replace all of the desired mutation because it eliminates the functional group while retaining the space-filling properties of the original residues. Mutant enzyme expression and purification were carried out according to the protocol described previously (17) unless otherwise indicated.

### Kinetic analysis

For active enzymes, nuclease cleavage kinetics was measured as described previously (18) using a Becton Dickenson FACSCalibur flow cytometer running CellQuest software and a stopped-flow flow cytometer built at the National Flow Cytometry Resource, Las Alamos National Laboratory (21). Flow cytometry data was processed using IDLYK software (created at Los Alamos National Laboratory). Single turnover $Mg^{2+}$ jump kinetics to measure cleavage kinetics and enzyme substrate binding were analyzed by least-square curve fitting in SigmaPlot (Jandel). For inactive enzymes, substrate binding was assessed by competition with wild type enzyme for substrate. The initial reaction velocities were plotted versus mutant protein concentration and the $K_d$ for binding estimated by fitting these data to an equation describing binding to a single class of sites.

## RESULTS

The critical amino acid residues were identified by sequence alignment with other nucleases in the family of structure-specific nucleases, superimposition of the hFEN-1 protein sequence onto the crystallographic structure of T4 RNase H (18), site-directed mutagenesis and nuclease activity screening.

### Sequence alignment and molecular superimposition

We have aligned the hFEN-1 sequence to 17 other structure-specific endo/exonucleases (Fig. 1). These nucleases include four subfamilies: flap endonucleases (17), XPG nucleases (22), 5′ nuclease domains of prokaryotic DNA polymerases (23) and viral exonucleases (23). Six carboxyl amino acid residues were identical among these 18 nucleases. R103 and E160 of hFEN-1 are absolutely conserved in mammalian enzymes. hFEN-1 R103 is replaced by a lysine while E160 is replaced by an aspartate in prokaryotic organisms. We therefore hypothesized that these conserved acidic residues are important for the nuclease activity of hFEN-1.

Based on the known three dimensional structure of T4 RNase H (19) and the primary sequence alignment between these two enzymes, we have superimposed the hFEN-1 sequence onto the 3-D structure of T4 RNase H. Figure 2 shows the region surrounding the enzyme cofactor $Mg^{2+}$s. The conserved acidic residues identified in the sequence alignment are all located very close to the metal ions in the enzyme. They interact directly or indirectly via water molecules with $Mg^{2+}$.

### Site-directed mutagenesis and the mutant protein expression and purification

All of the identified conserved amino acid residues including D34, D86, R103, E158, E160, D179, D181 and D233 have been converted to an alanine. The altered DNA sequences have been confirmed in a mutagenesis vector pBluescript and in the overexpression vector pET28b after subcloning. When the site-direct mutagenesis used to explore the role of a particular amino acid residue, a critical question is whether the observed effects are due to the removal of the active side chain or the distortion of the native conformation caused by mutation. This is particularly important when the three dimensional structure of the protein is not available. In our case the precaution has been taken to purify the protein from the soluble fraction under non-denaturing conditions. The soluble part is ~30% of total expressed protein for the most mutant proteins. Retention of the binding ability of some of the mutants is a good indication of the native conformation of inactive mutants. For one particular mutant (D233A), the solubility of the expressed protein is very low when the cell is grown and induced to express to the mutant protein in LB broth at 37°C (Fig. 3, lanes 2 and 3). In order to increase the solubility of the protein, a sorbitol medium has been used in this case (24). The cells were grown in the medium with a high concentration of sorbitol (1 M) at 37°C and then the temperature was shifted to 25°C after induction; then grown for another 3 h before harvest. Under these conditions the solubility of the expressed mutant protein was dramatically improved. The cells provided enough material in the soluble part to be purified (Fig. 3).

### Flap endonuclease activity screening

Initial screening of flap endonuclease activities of mutants has been reported previously (17). The removal of conserved acidic amino acid residues including the newly constructed E160A results in complete loss of the flap endonuclease activity except in the case of R103A, which retains the wild type-like enzyme activity. All of the inactive mutants have been subjected to a competition experiment to test their ability to bind to the flap substrate. Four out of seven mutants retained substrate binding ability. Therefore, all of the mutants produced in our studies can be classified into three categories based on their substrate binding and cleavage

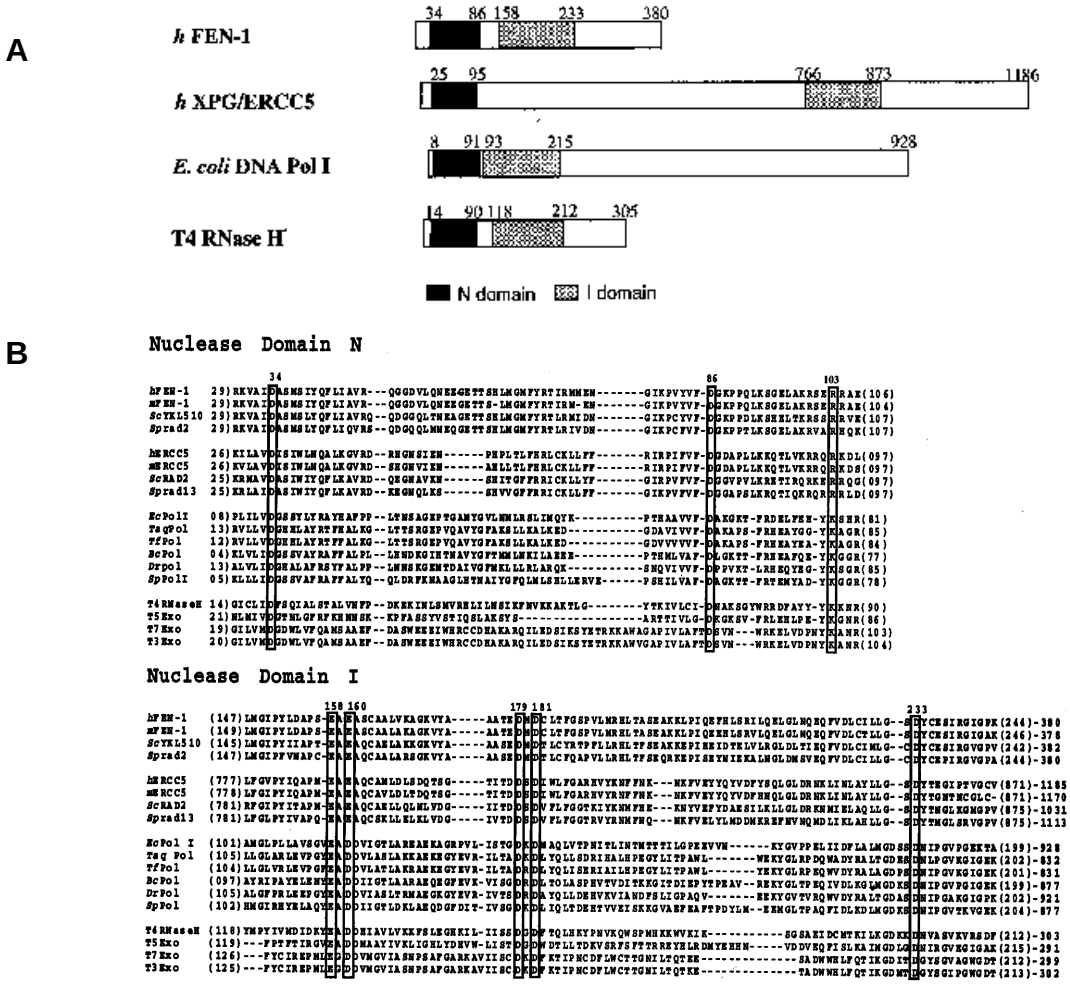


**Figure 1.** The conservative regions of four categories of nucleases (**A**) and amino acid sequence comparison of these regions (**B**). The boxed residues are conserved in all 18 enzymes and are possible ligands to $Mg^{2+}$. The site-directed mutagenesis has been performed on these sites. Keys: *h* or *m*FEN-1: human or mouse flap endonuclease; *Sc*YKL510: *S.cerevisiae* YKL510 gene product; *Sp*rad2: *S.pombe* rad2 gene product; *h* or *m*XPG: human or mouse protein responsible for *Xeroderma pigmentosum* group G; *Sc*RAD2: *S.cerevisiae* RAD2 gene product; *Sp*rad13: *S.pombe* rad13 gene product; *Ec*PolI: *E.coli* DNA polymerase I; TaqPol: *Thermus aquaticus* DNA polymerase; *Tf*Pol: *Thurmus flavus* DNA polymerase; *Bc*Pol: *Bacillus caldotenax* DNA polymerase; DrPol: *Deinococcus radiodurans* DNA polymerase; *Sp*Pol: *Streptococcus pneumoniae* DNA polymerase; T4RNaseH: phage T4 RNase H 5′–3′ DNA exonuclease; T5Exo: phage T5 D15 exonuclease; T7Exo: phage T7 gene 6 exonuclease; T3Exo: phage T3 gene 6 exonuclease. The sequence alignment has been done using DNA-star program (DNAstar, Madison, WI).

activities. The first category of mutants has completely lost the flap nuclease activity but retained the substrate binding ability. The mutants of this category are D34A, D86A, E160A and D181A. These residues possibly play a role in substrate catalysis. The second category includes the mutants which have lost both substrate binding and cleavage activity completely (E158A, D179A and D233A). Their activity loss might be due to the deficiency in substrate binding. The last type of mutant has no effect on the flap endonuclease activity (R103A). R103 is a conserved amino acid residue among the 18 nucleases. It has been suggested that the T5 exonuclease counterpart of hFEN-1 R103 (K83) is directly involved in substrate interaction as it is located in the inner side of the arch (20). Further questions to be answered are: (i) does the conversion of the conserved arginine 103 to alanine have any effect on the overall cleavage and binding kinetics compared to the kinetics of the wild type enzyme?; (ii) do the catalytically-deficient mutants have any deficiency in substrate binding?

## Overall kinetics of mutant R103A

As the first step to examine the *in vitro* functional alteration of the mutant enzyme R103A, we have studied the rate constant for the reaction as a function of the mutant enzyme concentration (Fig. 4). No significant difference between the mutant R103A and wild type enzymes has been observed. This rate constant is a product of the rate constant for each individual step in the reaction (enzyme binding and dissociation, $Mg^{2+}$ binding and dissociation, cleavage and product release). In order to examine the possible functional alteration of this mutant in the single steps, we have performed the following two experiments.

## Cleavage kinetics of the mutant R103A

The rate of cleavage increased with increasing enzyme concentration and appeared to approach a saturating rate. At the highest enzyme concentration shown, enzyme is in >1000-fold excess

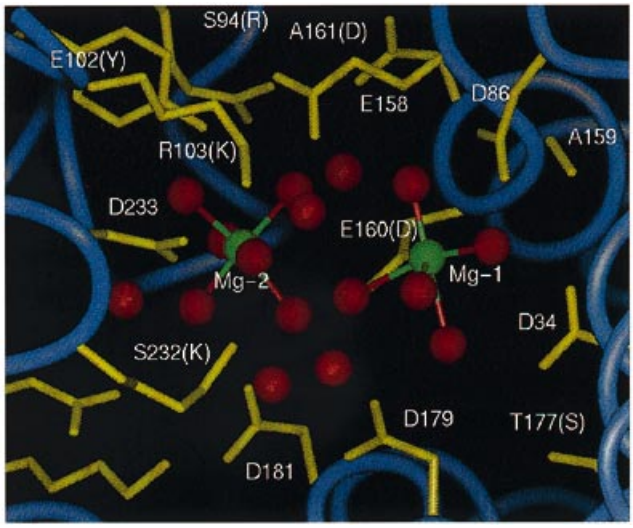


**Figure 2.** Proposed T4 RNase H active site (Brookhaven file 1TFR, 19) with superimposed human flap endonuclease sequence. For non-conserved residues, the identity of the side chain in T4 RNase H is indicated in parentheses. Shown are the two magnesium atoms (green) with coordinated water molecules (oxygen atoms, red), as well as additional water (red) and side chains (yellow) that participate in the active hydrogen-bonding network.

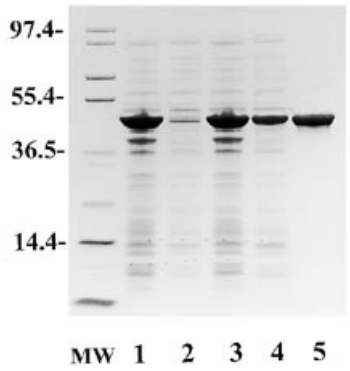


**Figure 3.** The solubilization of *E.coli* overexpressed human flap endonuclease. MW: molecular weight markers; lane 1: crude exact of *E.coli* strain BL21 (DE3) with a overexpression plasmid pET.FCH8 containing mutant D233A hFEN-1 grown and induced in LB medium; lane 2: the soluble part of lane 1; lane 3: crude exact of *E.coli* strain BL21 (DE3) with an overexpression plasmid pET.FCH8 containing mutant D233A hFEN-1 grown and induced in the sorbitol medium (25); lane 4: the soluble part of lane 3; lane 5: the purified mutant FEN-1 protein D233A. Approximately 70 µg protein has been loaded onto each lane.

over substrate, and cleavage kinetics likely represents single turnover kinetics, in which each enzyme is able to bind and cleave only a single substrate molecule before all substrate is consumed. When saturating concentrations of R103A FEN-1 enzyme are pre-incubated with the flap DNA substrate in the presence of 1 mM EDTA, addition of excess (10 mM) $MgCl_2$ induces DNA cleavage with first order kinetics and a half time of ~7 s, which is very close to that of wild type enzyme (Fig. 5A).

## Equilibrium binding of R103A

When FEN-1 is allowed to pre-bind the flap DNA substrate at saturating concentrations, the $Mg^{2+}$ jump kinetics represent a single turnover of FEN-1 flap DNA cleavage. We have used single

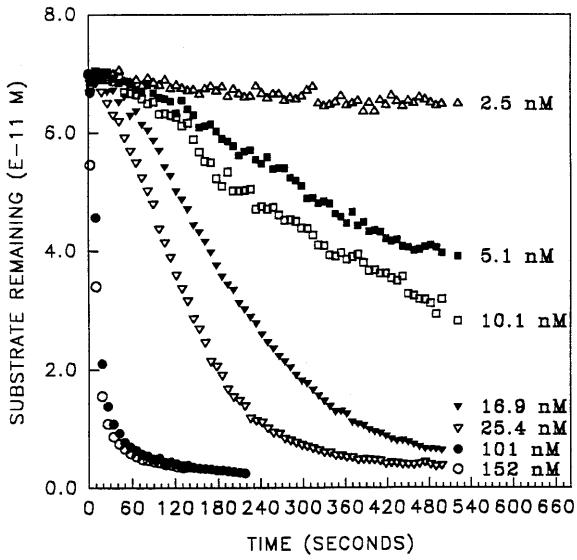


**Figure 4.** Concentration dependence of cleavage activity of FEN-1 site-directed mutant R103A. Enzyme was added to flap substrate beads in buffer TBM and cleavage kinetics measured by flow cytometry.

turnover kinetics to determine the equilibrium binding of FEN-1 to the flap DNA substrate. Because the burst of cleavage upon addition of $Mg^{2+}$ is rapid compared to subsequent enzyme turnover, the burst amplitude is a measure of the amount of enzyme–substrate complex formed during that pre-incubation. Presented in Figure 5B is the mutant R103A concentration dependence of the burst size, and thus mutant enzyme–flap DNA complex formation. The size of the cleavage burst increases with increasing R103A FEN-1 enzyme concentration and saturates >20 nM enzyme. A fit of these data to a hyperbolic binding equation reveals a $K_d$ for flap DNA binding of 8 nM for this mutant, compared to 7.5 nM for wild type enzyme (18).

## Competition between catalytically-deficient mutant and wild type enzymes

We evaluated the ability of each of the inactive mutant proteins to bind to the flap substrate by measuring the competition of these proteins with wild type enzyme for substrate binding sites. Mutants that are able to bind to the substrate compete with the wild type enzyme for binding sites and slow its cleavage kinetics. For this experiment, the inactive mutant enzymes were pre-incubated with the substrate at room temperature for 15 min to allow them to bind to the substrate followed by the addition of wild type enzyme and measurement of cleavage. We observed that mutants D34A, D86A, E160A and D181A can inhibit wild type activity while other inactive mutants did not demonstrate any effect in the same conditions. The concentration dependence of cleavage inhibition by different mutants reflects the binding ability of the catalytically inactive mutants (Fig. 6A–D). The initial rate of cleavage by wild type enzyme was plotted versus mutant protein concentration (Fig. 6E), and the concentration required for half-maximal inhibition is presented as a measure of the affinity of the mutants for substrate (Table 1). Also presented in Table 1 are the binding and cleavage parameters for the other mutants and for the wild type enzyme. The other three mutants (E158A, D179A and D233A) did not show any inhibition under the same

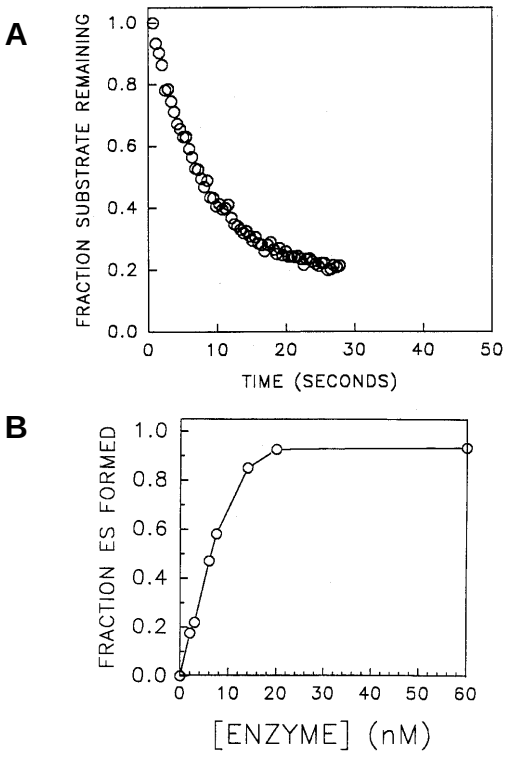


**Figure 5.** Binding and cleavage of the flap DNA by mutant R103A. (**A**) Estimated $K_d$ for binding of R103A to the flap DNA substrate. Enzyme at the indicated concentrations was incubated with flap DNA substrate for 5 min in buffer TBE. Cleavage was initiated by the addition of 10 mM $MgCl_2$, and the burst of cleavage was measured and plotted versus enzyme concentration. (**B**) $Mg^{2+}$ jump kinetics of flap DNA cleavage by mutant R103A. Enzyme (100 nM) was incubated with flap DNA substrate for 5 min in buffer TBE. Preformed enzyme–substrate complex was then mixed with an equal volume of buffer TB containing 20 mM $MgCl_2$ and the cleavage kinetics were measured using the Rapid Kinetic Flow Cytometry.

condition, indicating that they lost binding ability completely. An error on the D179A mutation was introduced during subcloning for overexpression of the mutant protein and reported in our previous paper (17). We have reconstructed the mutant and confirmed the presence of the mutation in the overexpression construct. The purified mutant protein has a complete loss of substrate binding and catalysis.

**Table 1.** Binding and cleavage parameters for FEN-1 and site-directed mutagenesis

| Protein | $K_d$ (nM)[a] | $k_{cleave}$ ($s^{-1}$)[b] |
|---|---|---|
| Wild type | 7.5[a] | 0.1[a] |
| D34A | 150[b] | na |
| D86A | 18[b] | na |
| R103A | 6[c] | 0.1[d] |
| E158A | nc | na |
| E160A | 72 | na |
| D179A | nc | na |
| D181A | 8[b] | na |
| D233A | nc | na |

[a]From Nolan *et al.* (18).
[b]Estimated as half-maximal inhibition of wild type activity in Figure 6E.
[c]Estimated from half-maximal burst size in Figure 5B.
[d]Estimated from $Mg^{2+}$ jump kinetics in Figure 5A.
na, not active; nc, did not compete.

## DISCUSSION

Among the 18 nucleases considered for sequence comparison, three of them have 3-D structures available: *Taq* DNA polymerase 5′ nuclease (25), T4 RNaseH (19) and T5 exonuclease-1 (20). All three enzymes have flap endonuclease activity. The 5′ nuclease domains of *Escherichia coli* DNA polymerase I and *Taq* DNA

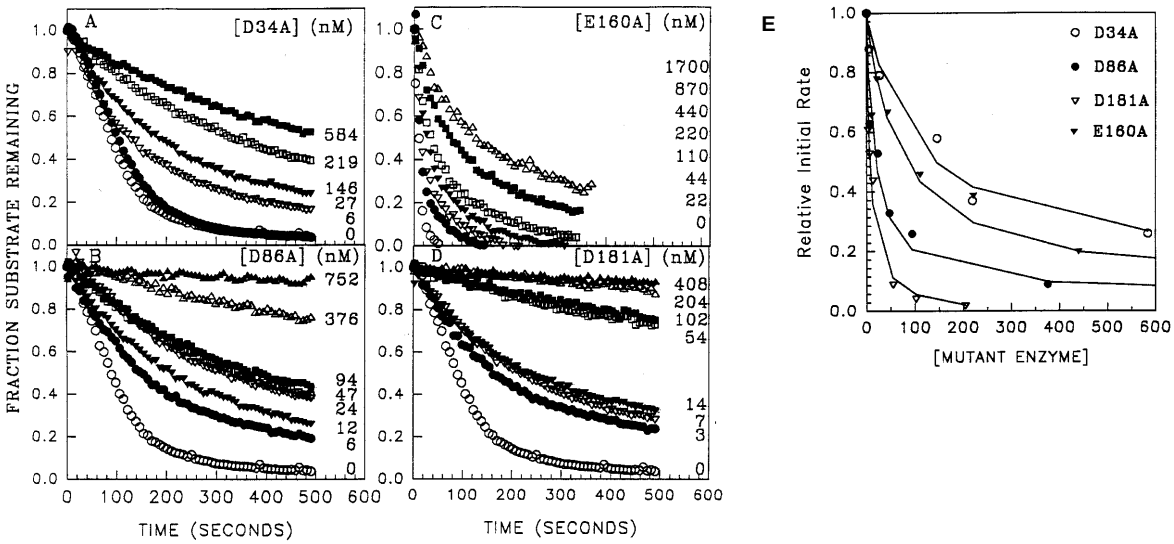


**Figure 6.** Competition of wild-type FEN-1 with inactive mutant proteins. (**A–D**) Concentration-dependent inhibition of the wild type activity by mutants. Mutant proteins were incubated with flap DNA beads in buffer TBM at the indicated concentration for 15 min. Wild type enzyme (50 nM) was added, and the cleavage kinetics measured. (**E**) Initial rate. The initial rate at each concentration of mutant protein was calculated relative to that of wild type enzyme alone, and plotted versus mutant protein concentrations.

polymerase are functional homologues of the human flap endonuclease-1 (26). The 5′ exo/endonuclease exists as an independent domain from the other two functional domains (3′ exonuclease and DNA polymerase). It is functional after proteolytic removal from the rest of the protein (27). The crystallographic structure of 5′ nuclease shows that it has a deep cleft that contains at its bottom the conserved carboxylates to ligate divalent metal ions (25). A central β-sheet lies at the heart of the domain and is flanked on both sides by assemblies of five and six α-helices which form the wall of the active-site cleft. Nine acidic amino acid residues are conserved with other nucleases as listed in Table 2. Seven of these residues (Asp18, Asp67, Glu117, Asp119, Asp120, Asp142 and Asp144) cluster within a sphere of 7 Å radius around the metal ions.

The other structure-specific endonuclease that has crystal structure available is T4 RNase H. It also has highly conserved cluster of acidic residues at the base of the structural cleft. The carboxylate oxygen atoms of conserved acidic residues Asp19, Asp71, Asp132 and Asp155 directly or indirectly coordinate essential metal ion $Mg^{2+}$[1] by forming hydrogen bonds to water molecules. Five other conserved carboxyl residues (E130, D133, D157, D197, D200) coordinate the $Mg^{2+}$[2] through hydrogen-bonding to six water molecule ligands (19). Another viral counterpart of the human FEN-1 is T5 exonuclease, which has the crystal structure available (20). The 3-D structure has revealed similar information in terms of the active site of the enzyme. The structure specifically revealed that conserved aspartic acid residues D26, D86, D128 and D131 coordinate Mn[1] while the other conserved residues D130, D153, D155, D201 and D204 coordinate Mn[2] (Table 2).

Table 2 lists conserved amino acid residues of the human flap endonuclease-1, their counterparts in the three nucleases with known 3-D structures and their possible roles in $Mg^{2+}$ ($Mn^{2+}$) coordinations. Our mutagenesis data have demonstrated that the removal of the conserved corresponding acidic residues individually in FEN-1 leads to complete loss of the flap endonuclease activity. However, it is still a difficult task to assign the role of individual conserved acidic acid residues in the $Mg^{2+}$ coordinations and their physical interacting relationship to an individual metal ion in FEN-1 protein without a 3-D structure. Even in the cases of two viral proteins, the ligands coordinating individual metal ions do not correspond exactly to their physical location in the primary sequences (Table 2). However, it seems to be true that two invariable aspartate from the N-terminal domain of the nucleases serve as ligands to the Mg[1] or Mn[1] even though the primary locations of the other two ligands vary. It is also interesting that these two residues (D34, D86) in human FEN-1 retained binding ability in spite of complete loss of activity with the other two residues (E160, D181). We hypothesize that these four amino acid residues may be involved in Mg[1] coordination and are important in DNA catalysis in hFEN-1. In this work, we have demonstrated that the wild type enzyme activity can be inhibited by these mutant proteins in the *in vitro* flap endonuclease assay. However, their inhibitory abilities or binding affinities are different. The concentration-dependent inhibition of these four mutants (Table 1) indicates that D181A has approximately the same binding ability as wild type, and demonstrates that conversion of aspartate 181 to alanine altered the enzymatic function only in the cleavage step. In the cases of D86A, E160A and D34A, their binding abilities have been decreased by 2-, 9- and 18-fold. Therefore, these three amino acid residues also play a role in substrate binding. Previously, it has been demonstrated that FEN-1 still binds to flap DNA without cutting in the absence of divalent cations (12,28).

FEN-1 mutant proteins E158A, D179A and D233A lost substrate binding and catalytic activities completely. We speculate that these amino acid residues are involved in $Mg^{2+}$[2] coordination and have roles in substrate binding. The coordination formed in this case may stabilize the structure for the DNA substrate binding. However, we cannot exclude the possibility that the mutations have caused dramatic conformational change of the enzyme, leading to loss of substrate binding ability. From the structures of homologous proteins, one can see that these residues are clustered around the $Mg^{2+}$ in the middle and less accessible to the surface of protein molecule (19). Removal of these residues would likely cause structural distortion. All three mutants were less soluble and went into an inclusion body when they were overexpressed in the *E.coli* cells compared with wild type and other mutant proteins. This may indicate a non-natural conformation of the mutant proteins. The soluble fraction of the mutant D233A is too small to be purified under standard conditions and required a special modification of the culture medium and conditions.

**Table 2.** Correspondence of conserved carboxyl residues coordinating the $Mg^{2+}$ ($Mn^{2+}$) for active sites in human flap endonuclease with ones in other three structure-specific nuclease of known structures

| *Taq* Pol I 5′ Exo (26) | | T4 RNase H (19) | | T5 Exonuclease-1 (20) | | *h*FEN-1 (17, this work) | |
|---|---|---|---|---|---|---|---|
| aa residue | metal site | aa residue | metal site | aa residue | metal site | aa residue | functions |
| D18 | MI | D19 | Mg[1] | D26 | Mn[1] | D34 | cleavage/binding |
| D67 | – | D71 | Mg[1] | D68 | Mn[1] | D86 | cleavage/binding |
| E117 | MIII | E130 | Mg[2] | D128 | Mn[1] | E158 | binding |
| D119 | MI/III | D132 | Mg[1]/[2] | D130 | – | E160 | cleavage/binding |
| D120 | MIII | D133 | Mg[2] | D131 | Mn[1] | absent | absent |
| D142 | MI/II | D155 | Mg[1] | D153 | Mn[1]/[2] | D179 | binding |
| D144 | MII | D157 | Mg[2] | D155 | Mn[2] | D181 | cleavage |
| D188 | – | D197 | Mg[2] | D201 | Mn[2] | absent | absent |
| D191 | – | D200 | Mg[2] | D204 | Mn[2] | D233 | binding |

– , No interaction between residues and metal site from the structural data.

An error on the D179A mutation was introduced during the subcloning for overexpression of the mutant protein and reported in our previous paper (17). We have reconstructed the mutant and confirmed the presence of the mutation in the overexpression construct. D179 corresponding to D142 in *Taq* polymerase and D155 in T4 RNase is more closely located with $Mg^{2+}$[1]. However, our mutagenesis result showed that it is defective in substrate binding. It is possible that the FEN-1 has different arrangement of the catalytic center dependent on where the additional ligands come from. Based on our sequence analysis, only seven conserved acidic amino acid residues in the FEN-1 nuclease were identified corresponding to the ones in *Taq* polymerase, T4 RNase H and T5 exonuclease (Fig. 1 and Table 2). There are at least two more residues that are not identified in FEN-1 enzyme if the $Mg^{2+}$s are coordinated by nine ligands as in the case of *Taq* polymerase, T4 RNase H and T5 exonuclease. Other conserved acidic amino acid residues conserved only in FEN-1 and XPG subfamilies could be used as additional $Mg^{2+}$ ligands. Moreover, some conserved amino acid residues such as D34, D86 and E160 could be involved in coordinating both magnesium ions. Therefore removal of these can cause deficiency in both substrate binding and catalysis. From molecular modeling, we have also observed there might be more non-carboxyl amino acid residues are involved in the $Mg^{2+}$ coordinations (Fig. 2). A good example is T177. It is conserved throughout the 18 enzymes compared (it is replaced by a serine in some organisms). This residue may interact with D34 according to the T4 RNase structure or directly provide -OH group in coordinating the $Mg^{2+}$ cofactor. The role of this conserved residue will be assessed in our future experiments.

All of the conserved acidic amino acid residues are clustered in the N-terminal (N) and intermediate (I) regions of the protein. This may indicate that they are the conserved nuclease domains. The recently published crystal structure of T5 exonuclease (20) demonstrates that this is the case. It also pointed out that the region between the two conserved domains (N and I) was missing in the structures of *Taq* polymerase and T4 RNase H. This region connects the two conserved domains and makes a molecular arch three dimensionally. There is a hole in the middle of the molecule, which is only large enough for single-stranded DNA substrate to thread through. This may explain the enzyme specificity to the flap substrate. There are two clusters of positively charged and aromatic amino acid residues in the inner side of this arch. They may directly interact with DNA substrate. One of these amino acid residues is conserved in hFEN-1 enzyme, R103. However, removal of R103 residue surprisingly turned out to have no effect on substrate binding. One possible reason is that R103 is adjacent to a unique R104 in the FEN-1/XPG subfamilies. The function of the R103 could be replaced by R104 in the mutant protein. We will further investigate the roles of this conserved positively charged amino acid residue by a double mutation R103A and R104A.

## ACKNOWLEDGEMENTS

We thank Dr Mark Sherman's help with molecular modeling and superimposition of the human FEN-1 sequence onto the known three dimensional structure of T4 RNase H, Ms Kieran G. Cloud for synthesis of all the oligos used in the experiments, Mr Robb C. Habbersett for customizing the IDLYK flow cytometry data analysis program to facilitate kinetic analysis, Miss Allison Jolly and Lina Somsouk for technical assistance, and Dr J. Qiu for preparation of some figures and critical reading of the manuscript. We also thank Drs Tim Mueser and Craig Hyde at NIH for sharing the 3-D structural information of the T4-RNase H prior to publication. This research was supported by the Office of Health and Environmental Research of the Department of Energy, Los Alamos National Laboratory Directed Research Funds, NIH grants RR01315 and CA71630, and a City of Hope National Medical Center institutional fund.

## REFERENCES

1 Harrington, J. J. and Lieber, M. R. (1994) *EMBO J.* **13**, 1235–1246.
2 Goulian, M., Richards, S. H., Heard, C. J. and Bigsby, B. M. (1990) *J. Biol. Chem.* **265**, 18461–18471.
3 Ishimi, Y., Claude, A., Bullock, P. and Hurwitz, J. (1988) *J. Biol. Chem.* **263**, 19723–19733.
4 Waga, S., Bauer, G. and Stillman, R. (1994) *J. Biol. Chem.* **269**, 10923–10934.
5 Murray, J. M., Tavassoli, M., Al-harithy, R., Sheldrick, K. S., Lehmann, A. R., Carr, A. M. and Watts, F. Z. (1994) *Mol. Cell. Biol.* **14**, 4878–4888.
6 Robins, P., Pappin, D. J. C., Wood, R. D. and Lindahl, T. (1994) *J. Biol. Chem.* **269**, 28535–28538.
7 Siegal, G., Truchi, J. J., Myers, T. W. and Bambara, R. A. (1992) *Proc. Natl. Acad. Sci. USA* **89**, 9377–9381.
8 Harrington J. J. and Lieber, M. R. (1994) *Genes Dev.* **8**, 1344–1355.
9 Murante, R. S., Huang, L., Turchi, J. J. and Bambara, R. (1994) *J. Biol. Chem.* **269**, 1191–1196.
10 Turchi, J. J. and Bambara, R. A. (1993) *J. Biol. Chem.* **268**, 15136–15141.
11 Turchi, J. J., Huang, L., Murante, R. S., Kim, Y. and Bambara, R. A. (1994) *Proc. Natl. Acad. Sci. USA* **91**, 9803–9807.
12 Murante, R. S., Rust, L. and Bambara, R. A. (1995) *J. Biol. Chem.* **270**, 30377–30383.
13 Reagan, M. S., Pittenger, C., Siede, W. and Friedberg, E. C. (1995) *J. Bacteriol.* **177**, 364–371.
14 Sommers, C. H., Miller, E. J., Dujon, B., Prakash, S. and Prakash, L. (1995) *J. Biol. Chem.* **370**, 4193–4196.
15 Johnson, R. E., Kovvali, G. K., Prakash, L. and Prakash, S. (1995) *Science* **269**, 238–240.
16 Tishkoff, D. X., Filosi, N., Gaida, G. M. and Kolodner, R. D. (1997) *Cell* **88**, 253–263.
17 Shen, B., Nolan, J. P., Sklar, L. A. and Park, M. S. (1996) *J. Biol. Chem.* **271**, 9173–9176.
18 Nolan, J. P., Shen, B., Park, M. S. and Sklar, L. A. (1996) *Biochemistry* **35**, 11668–11676.
19 Mueser, T. C., Nossal, N. G. and Hyde, C. C. (1996) *Cell* **85**, 1101–1112.
20 Ceska, T. A., Sayers, J. R., Stier, G. and Suck, D. (1996) *Nature* **382**, 90–93.
21 Nolan, J. P., Posner, R. G., Martin, J. C., Habbersett, R. C. and Sklar, L. A. (1995) *Cytometry* **21**, 223–229.
22 Carr, A. M., Sheldrick, K. S., Murray, J. M., Al-Harrithy, R., Watts, F. Z. and Lehmann, R. (1993) *Nucleic Acids Res.* **21**, 1345–1349.
23 Gutman, P. D. and Minton, K. W. (1993) *Nucleic Acids Res.* **21**, 4406–4407.
24 Blackwell, J. R. and Horgan, R. (1991) *FEBS Lett.* **295**, 10–12.
25 Kim, Y., Eom, S. H., Wang, J., Lee, D.-S., Suh, S. W. and Steitz, T. A. (1995) *Nature* **376**, 612–616.
26 Lyamichev, V., Brow, M. A. V. and Dahlberg, J. E. (1993) *Science* **260**, 778–783.
27 Lundquist, R. C. and Olivera, B. M. (1982) *Cell* **31**, 53–60.
28 Harrington, J. J. and Lieber, M. R. (1995) *J. Biol. Chem.* **270**, 4503–4508.