

## Interaction Interface of Human Flap Endonuclease-1 with Its DNA Substrates\*<sup>§</sup>

Received for publication, February 10, 2004, and in revised form, March 17, 2004  
Published, JBC Papers in Press, March 22, 2004, DOI 10.1074/jbc.M401464200

Junzhuan Qiu<sup>‡§</sup>, Ren Liu<sup>‡¶§</sup>, Brian R. Chapados<sup>||</sup>, Mark Sherman<sup>\*\*</sup>, John A. Tainer<sup>||</sup>,  
and Binghui Shen<sup>‡¶‡</sup>

From the Departments of <sup>‡</sup>Radiation Research and <sup>\*\*</sup>Bio-informatics and the <sup>¶</sup>Graduate Program in Biological Science, City of Hope National Medical Center and Beckman Research Institute, Duarte, California 91010 and the <sup>||</sup>Department of Molecular Biology and Skaggs Institute of Chemical Biology, the Scripps Research Institute, La Jolla, California 92122

Flap endonuclease-1 or FEN-1 is a structure-specific and multifunctional nuclease critical for DNA replication, repair, and recombination; however, its interaction with DNA substrates has not been fully understood. In the current study, we have defined the borders of the interaction between the FEN-1 protein and its DNA substrates and identified six clusters of conserved positively charged amino acid residues, which are in direct contact with DNA substrate. To map further the corresponding interactions between FEN-1 residues and DNA substrates, we performed biochemical assays employing a series of flap DNA substrates lacking some structural components and a series of binding-deficient point mutants of FEN-1. It was revealed that Arg<sup>47</sup>, Arg<sup>70</sup>, and Lys<sup>326</sup>-Arg<sup>327</sup> of FEN-1 interact with the upstream duplex of DNA substrates, whereas Lys<sup>244</sup>-Arg<sup>245</sup> interact with the downstream duplex. This result indicates the orientation of the FEN-1-DNA interaction. Moreover, Arg<sup>70</sup> and Arg<sup>47</sup> were determined to interact with the sites around the 2nd nucleotide (Arg<sup>70</sup>) or the 5th/6th nucleotide (Arg<sup>47</sup>) of the template strand in the upstream duplex portion counting from the nick point of the flap substrate. Together with previously published data and the crystallographic information from the FEN-1-DNA complex that we published recently (Chapados, B. R., Hosfield, D. J., Han, S., Qiu, J., Yelent, B., Shen, B., Tainer, J. A. (2004) *Cell* 116, 39–50) we are able to propose a reasonable model for how the human FEN-1 protein interacts with its DNA substrates.

Flap endonuclease-1 (FEN-1)<sup>1</sup> is a critical structure-specific nuclease for  $\alpha$ -segment processing during Okazaki fragment maturation and for DNA base excision repair (1–5). Deficiency of these pathways consequently results in genome instability, including significantly enhanced mutation frequency and microsatellite instability (5–11). The enzyme is also involved in

preventing illegitimate crossover activities such as short repeated sequence recombination and hence enhances genome stability (12). More recently, it has been demonstrated that FEN-1 has a novel function in promoting apoptotic DNA fragmentation (13).

Despite its multiple biochemical activities, protein-protein interaction partners and involvement in several DNA metabolic pathways, FEN-1 is a unique enzyme that solely recognizes abnormal DNA structures, typified by a flap DNA substrate. Initial motif analysis based on protein sequence comparison and biochemical assays identified two major conserved motifs, the N (N-terminal) and I (Intermediate) motifs, which are essential for the nuclease activities of FEN-1 proteins (4, 14). These two regions contain 7–8 conserved acidic amino acid residues that coordinate two magnesium ions and form an active center for catalysis (15). A third motif toward the C terminus is involved in the interaction between FEN-1 and proliferating cell nuclear antigen (16–18). This interaction is required for the recruitment of FEN-1 to proliferating cell nuclear antigen at sites of Okazaki fragment processing and DNA damage repair (4, 17, 19). In eukaryotic cells, FEN-1 also has a C-terminal motif with multiple clusters of positively charged amino acid residues, which are important for the localization of FEN-1 into the nucleus (20).

Crystal structures of FEN-1 enzymes have identified structural elements that may be important for interactions with DNA flap substrates. Structures of FEN-1 and its homologs, including three archaeal FEN-1 proteins, T4 RNase H, and T5 exonuclease (21–24) exhibit an arch element formed by a helix-loop-helix motif located over the active center. This “helical arch” structure contains a number of positively charged and bulky amino acid residues on its inner side, which could be in the direct contact with the single-stranded DNA flap of DNA flap substrates (23, 24). Corresponding residues in the 5′-nuclease of *Escherichia coli* DNA polymerase I have been shown to play a role in DNA binding (25). Furthermore, partial deletions of this region in *Methanococcus jannaschii* FEN-1 (mjFEN-1) abolish enzyme activity (24). This arch structure may allow FEN-1 to slide through the single-stranded flap of flap DNA substrate to perform structure-specific cleavages, a sliding-through mechanism proposed for FEN-1 by Barnes *et al.* (26). In addition, this structure was shown to be critical for catalysis in human FEN-1 (27). Recent evidence suggests that the helical arch region undergoes a conformational change upon binding to substrate DNA, resulting in increased helical content, which may aid in positioning the single-stranded DNA flap near the catalytic site (28, 29).

In addition to the helical arch region, analysis of the *Pyrococcus furiosus* FEN-1 (pfFEN-1) crystal structure revealed a positively charged groove containing the active center and a

\* This work was supported by NCI, National Institutes of Health Grants R01CA073764 (to B. S.) and R01CA081967 (to J. A. T.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

<sup>§</sup> The on-line version of this article (available at <http://www.jbc.org>) contains Table S1.

<sup>¶</sup> These authors contributed equally to this work.

<sup>‡‡</sup> To whom correspondence should be addressed: Radiation Research, City of Hope National Medical Center, 1500 East Duarte, Duarte, CA 91010. Tel.: 626-301-8879; Fax: 626-301-8280; E-mail: bshen@coh.org.

<sup>1</sup> The abbreviations used are: FEN-1, flap endonuclease-1; exoIII, exonuclease III.

H3TH motif, which were proposed to mediate binding of the double- or single-stranded regions of the flap DNA substrates to FEN-1 (23). The recent structure of *Archaeoglobus fulgidus* FEN-1 (affFEN-1) bound to DNA identified two additional helix-loop-helix motifs that contact the upstream portion of the DNA flap substrate (29). Consistent with this report, mutation of Arg<sup>47</sup> and Arg<sup>70</sup> in human FEN-1, which are both located in these regions, affects substrate binding specificity (30). Furthermore, biochemical and mutational analysis of *Pyrococcus horikoshii* FEN-1 (phFEN-1) (31) revealed a total of five loop regions important for protein-DNA interactions, which include the helical arch, H3TH, and other structurally identified DNA binding regions.

Although evidence for the DNA-protein interaction of FEN-1 is accumulating, it is still important to identify all of the elements directly involved in substrate binding to understand how the enzyme positions its substrates in a way that optimizes cleavage. Here we present mutational and biochemical analyses of human FEN-1 which define how FEN-1 interacts with DNA. We defined the borders of the interactions between the FEN-1 and its DNA substrates by exonuclease III (exoIII) footprinting. In addition, we identified 14 FEN-1 mutations that affect DNA binding. Using FEN-1 substrates designed to test the interactions of these mutants with known specific regions of the flap substrate, we determined the substrate binding orientation for human FEN-1 and implicated key structural elements in DNA binding. Moreover, Arg<sup>70</sup> and Arg<sup>47</sup>, which are known to be involved in FEN-1-DNA binding, were determined to interact with the sites around the 2nd nucleotide (Arg<sup>70</sup>) and the 5th/6th nucleotide (Arg<sup>47</sup>) of the template strand in the upstream duplex portion counting from the nick point of the flap substrate. Together, these results and the protein-DNA complex structure (29) allowed us to propose a reasonable model for how the human FEN-1 protein interacts with its DNA substrates.

#### EXPERIMENTAL PROCEDURES

**Site-directed Mutagenesis, Protein Overexpression, and Purification**—All human FEN-1 mutants used for this study were prepared using the QuikChange site-directed mutagenesis kit from Stratagene (La Jolla, CA). Mutagenic primers were synthesized at the City of Hope DNA/RNA/peptide synthesis core facility. Mutations and corresponding oligonucleotide sequences for primers are listed in Table S1 for clarity as supplementary material and are available upon request. Site-directed mutagenesis, overexpression, and purification of wild type and mutant FEN-1 enzymes were carried out based on our previously published procedures (30). Mutagenesis reactions were performed using pET-28-derived plasmids harboring the wild type human FEN-1 sequence as a template so that the isolated plasmids containing a mutation could be used directly for protein expression in *E. coli*.

**Protein Sequence Alignment and Three-dimensional Structure Modeling of Human FEN-1 and FEN-1-DNA Complexes**—The sequences of FEN-1 proteins were aligned using ClustalW 1.8 Multiple Sequence Alignment Algorithm at BCM Search Launcher, Baylor College of Medicine HGSC (searchlauncher.bcm.tmc.edu). The structure of human FEN-1 was modeled using the crystal structures of *M. jannaschii* and *P. furiosus* FEN-1 proteins as templates (PDB files 1A76 and 1B43). Using the homology module within the molecular modeling program Insight II (Accelrys, San Diego), the sequence of human FEN-1 was aligned with the archaeal FEN-1 sequences. Structural information was then used to modify the final alignment of the three FEN-1 sequences so that insertions and deletions fell between secondary structure elements. The model was built and refined using scripts provided with the program.

The hFEN-1-DNA model was constructed using initial DNA coordinates from the crystal structure of affFEN-1 bound to DNA (PDB code 1RXW). The upstream DNA substrate was initially positioned by aligning the hFEN-1 model with the affFEN-1 based on C $\alpha$  positions. Close contacts in the initial hFEN-1-DNA were eliminated by rigid body minimization based only on stereochemical parameters using crystallography NMR software (32). The downstream DNA and 5'-flap regions were then added to the model and positioned in a way that satisfied the greatest number of contacts to residues known to affect DNA binding.

The steric clashes between side chains and the DNA were resolved by minimizing only the side chain atoms of hFEN-1 using crystallography NMR software.

**DNA Substrate Preparation and FEN-1 Nuclease Activity Assays**—Protocols for DNA substrate preparation and nuclease activity assays were performed as published previously (30). Briefly, oligonucleotides as shown in the relevant figures were individually phosphorylated at the 5'-end. This was done by incubating 40 pmol of the oligonucleotide with 10  $\mu$ Ci of [ $\gamma$ -<sup>32</sup>P]ATP and 1  $\mu$ l (10 units/ $\mu$ l) of polynucleotide kinase at 37 °C for 60 min. Polynucleotide kinase was then inactivated by heating at 72 °C for 10 min. 80 pmol each of remaining oligonucleotides for individual substrates listed in the relevant figures were added to the labeled oligonucleotides, respectively. The samples were incubated at 70 °C for 5 min followed by slow cooling to 25 °C, thus allowing the oligonucleotides to anneal and form the flap and nick duplex substrates. Substrates were precipitated at -20 °C overnight after adding 20  $\mu$ l of 3 M NaOAc and 1 ml of 100% ethanol. Substrates were collected by centrifugation and washed once with 70% ethanol and resuspended in sterile water.

Reactions were carried out with the indicated amount of hFEN-1 and 80 nM flap or nick duplex substrate in reaction buffer containing 50 mM Tris (pH 8.0) and 10 mM MgCl<sub>2</sub>. Each reaction was then brought to a total volume of 10  $\mu$ l with water. All reactions were incubated at 30 °C for 15 min and terminated by adding an equal volume of stop solution (95% formamide, 20 mM EDTA, 0.05% bromophenol blue, 0.05% xylene cyanol). An aliquot of each reaction was then run on a 15% denaturing polyacrylamide gel at 1900 V for 1 h. The gel was dried at 70 °C for 50 min, and the bands were visualized by autoradiography.

**Kinetic Analysis**—FEN-1 cleavage kinetic assays were performed using various concentrations of DNA substrates (30–500 nM) and constant amounts of FEN-1 (92 nM) following the procedures described by Hosfield *et al.* (33). Briefly, reactions were initiated by combining standard reaction buffer, substrate, and enzyme in order. Samples were mixed and incubated for 2 min. The products and substrates were separated by denaturing gel electrophoresis. The initial velocity was calculated by measuring product and substrate intensity on the gel using the IPLab Gel program and the equation,

$$v = (I_1/I_0 + 0.5I_1)t \times [\text{substrate}] \quad (\text{Eq. 1})$$

where  $t$  = time in seconds,  $I_1$  = product intensity, and  $I_0$  = final substrate concentration. The substrate concentration was expressed in nM.  $V_{\text{max}}$  and  $K_m$  values were calculated by directly fitting the data into the Michaelis-Menten equation, and then  $K_{\text{cat}}$  was calculated by  $V_{\text{max}}/[\text{E}]_0$ .  $K_m$ ,  $K_{\text{cat}}$ , and  $K_{\text{cat}}/K_m$  values were finally used for plotting.

**Direct Enzyme/Substrate Binding Analysis**—This assay was conducted using a method modified from one established by Harrington and Lieber (34). In brief, the indicated amounts of FEN-1 were mixed with 80 nM labeled DNA in a final volume of 10  $\mu$ l containing 50 mM Tris (pH 8.0), 10 mM NaCl, 5 mM EDTA, 10% glycerol, and 50  $\mu$ g/ml bovine serum albumin. After a 15-min incubation at room temperature, each reaction was loaded onto a 5% polyacrylamide gel containing 0.5 $\times$  TBE. The reactions were then electrophoresed for 45 min at 100 V at 4 °C. The gel was dried and then exposed to Kodak x-ray film for imaging.

**ExoIII Footprinting**—For this experiment, we followed a procedure described by Hohl *et al.* (35). Human FEN-1 mutant R100A or D181A was used to bind 0.27-pmol flap substrates with 5'-<sup>32</sup>P end labeling of the flap strand or template strand. They were incubated in 50 mM Tris (pH 6.8), 20 mM NaCl, 15% glycerol, and 50  $\mu$ g/ml bovine serum albumin with 2 mM MgCl<sub>2</sub> and 100-fold excess cold double-stranded DNA competitor for 20 min. Subsequently, 20 units of exoIII were added to allow a 15-min digestion at 15 °C. The reaction was stopped by the addition of EDTA to a final concentration of 50 mM and formamide loading buffer. After heating for 10 min at 75 °C, the reactions were analyzed on a 15% sequencing gel and the digestion patterns were visualized by autoradiography.

**Circular dichroism (CD) Measurements**—CD measurements were performed using a J-600 spectropolarimeter (Japan Spectroscopic Co., Ltd., Tokyo) as described previously (36). Briefly, the far-UV (200–250 nm) CD spectrum was obtained at 25 °C using a solution containing 90  $\mu$ g/ml enzyme, 20 mM sodium phosphate buffer (pH 7.6) in a 2-mm path length cuvette. CD data were analyzed using the K2D program (37) for the calculation of the relative ratio of  $\alpha$ -helices,  $\beta$ -sheets, and random coils.

#### RESULTS

**Defining the Region of DNA Substrates Contacted by FEN-1**—To obtain an insight into how FEN-1 binds to DNA sub-



TABLE I  
Nuclease activity and binding activity alterations of human FEN-1 mutants

+++; wide type enzyme activity; ++; decreased activity; +; minimal activity; -, no activity; ND not defined; \*, with binding deficiency; Up-D, upstream duplex; Down-D, downstream duplex; ss-Flap, single-stranded DNA Flap.

Mutations	FEN-1 motifs	Nuclease activities		Binding capacity		Binding deficiency	Interactive DNA region
		Endo-activity	Exo-activity	Flap substrate	Nick substrate		
Wild type		+++	+++	+++	+++		
K29A/K80A	ND	+++	+++	+++	+++		
R47A	Small loop 1	++	+	++	++	*	Up-D
R70A	ND	++	+	++	+	*	Up-D
R73A	ND	+++	+++	+++	+++		
K93A	Large loop	-	-	++	+++	*	ss-Flap
K99A		+++	+++	+++	+++		
R100A		-	-	+++	+++		
R103A		+++	+++	+++	+++		
R103A/R104A		+	+	+	+	*	
K125A		++	++	++	+++	*	ss-Flap
K128A/R129A		+	+	+	+	*	
K132A		-	-	-	-	*	
R192A	Small loop 2	-	-	-	-	*	
K200A		+	+	+	+	*	Down-D
K201A		-	-	-	-	*	
R239A	Small loop 3	+++	+++	+++	+++		
K244A/R245A		++	+	++	++	*	Down-D
K252A/K254A	Small loop 4	+	-	-	-	*	Down-D
R261A/R262A		+++	+++	+++	+++		
K267A		+++	++	++	++	*	ND
R320A	ND	+++	+++	+++	+++		
K326A/R327A	ND	+++	++	++	++	*	Up-D

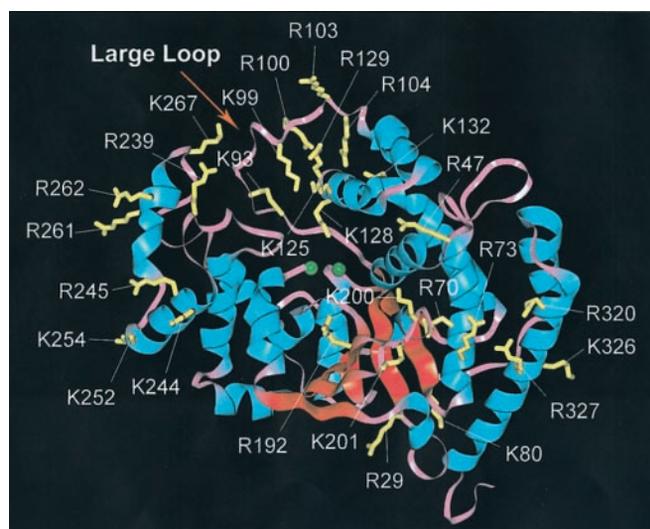


FIG. 2. Three-dimensional view of the human FEN-1 positively charged amino acid residues selected for site-directed mutagenesis. The human FEN-1 structure was modeled based on the crystal structures of *M. jannaschii* and *P. furiosus* FEN-1 proteins. The residues emphasized in this study are indicated in yellow.

both the flap and nick duplex DNA substrates. Mutants with no detectable enzyme activity (6 of 15) were excluded in this analysis because their kinetic parameters were not measurable. In this analysis, Michaelis-Menten kinetics was used to derive  $V_{max}$  and  $K_m$  values, and  $K_{cat}$  and  $K_{cat}/K_m$  values were then calculated (Fig. 5). Overall, the mutants R47A, R70A, R103A/R104A, K125A, K128A/R129A, K200A, K244A/K245A, and K326A/R327A have increased  $K_m$ , decreased  $K_{cat}$ , and decreased  $K_{cat}/K_m$  values (Fig. 5). The increased  $K_m$  values of the mutants suggest a substrate binding deficiency. In particular, the mutants R103A/R104A, K128A/R129A, K200A, and K244A/K245A had significantly higher  $K_m$  values compared with the wild type protein, whereas their  $K_{cat}$  and  $K_{cat}/K_m$  values were proportionally reduced. These results indicate that the decreased activity of these mutants is mainly the result of

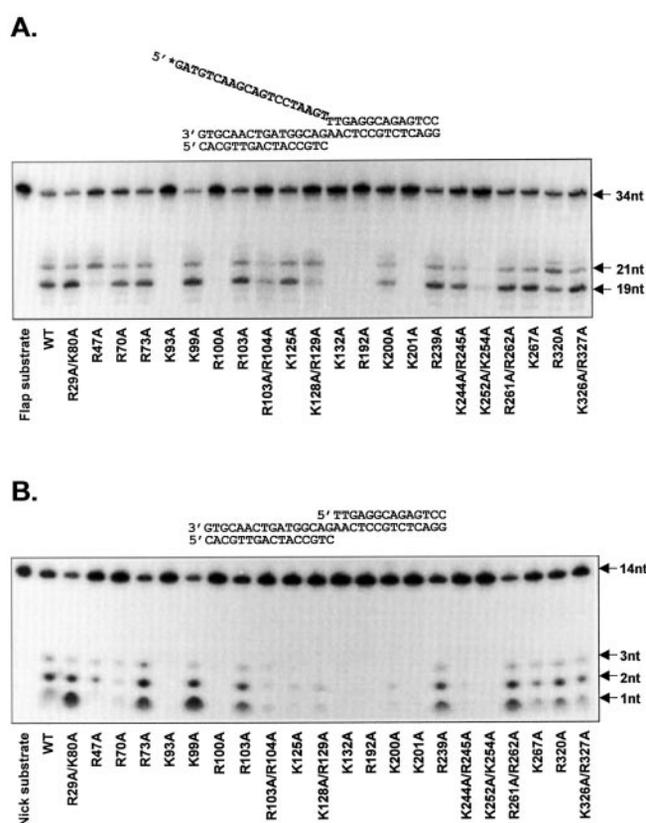


FIG. 3. Enzyme activities of wild type and mutant FEN-1 proteins. A, flap endonuclease activity of FEN-1 proteins. B, exonuclease activity of FEN-1 proteins. For both A and B, the top bands represent the uncut substrates, and the bottom bands are the cleavage products. The oligonucleotide size is indicated in nucleotides (nt). Reactions were carried out with 133 nM hFEN-1 protein and 80 nM of flap or nick duplex substrate in reaction buffer containing 50 mM Tris (pH 8.0) and 10 mM  $MgCl_2$  in a total volume of 10  $\mu$ l with water. Reactions were incubated at 30 °C for 15 min.

a substrate binding deficiency and therefore, the corresponding amino acid residues of these mutants are important for DNA substrate binding. These results are consistent with the gel

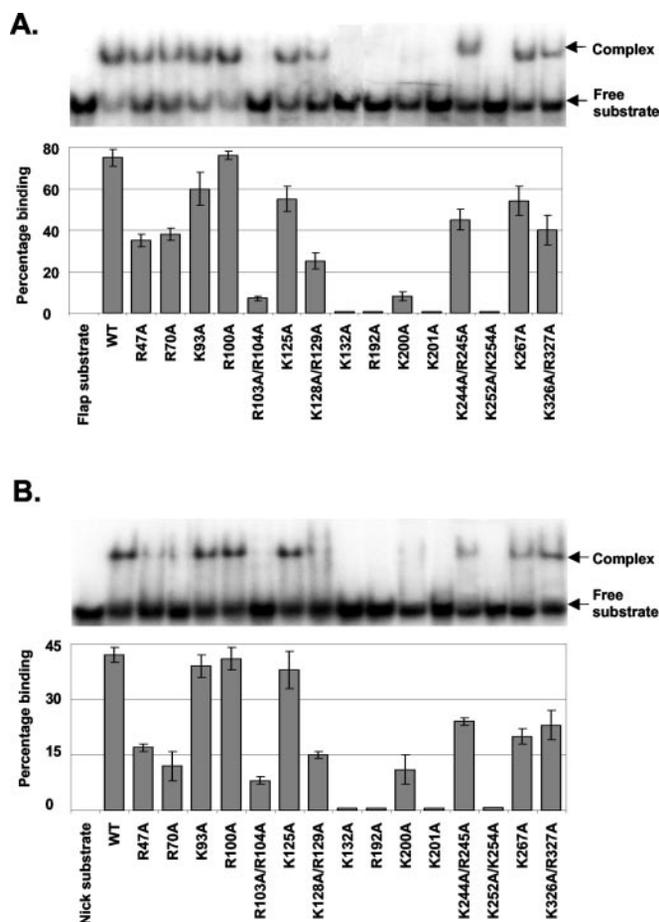


FIG. 4. **Substrate binding assay of hFEN-1 proteins.** *A*, binding assay based on flap substrate. *B*, binding assay based on nick duplex substrate. FEN-1 (666 nM), DNA substrate (80 nM), and binding buffer were mixed and incubated for 15 min at room temperature. The reactions were then loaded onto a 5% native polyacrylamide gel for electrophoresis and radioimaging. The substrates are the same as those shown in Fig. 3*A* (flap substrate) and *B* (nick duplex substrate).

shift assays (Fig. 4), supporting the role of these residues in DNA binding.

**Determination of the Binding Orientation between FEN-1 and Its 5'-Overhang Substrate**—A typical flap DNA substrate can be designated into three portions: upstream duplex, single-stranded flap, and downstream duplex (Fig. 1). To understand fully FEN-1-DNA interactions, it is critical to determine the identified protein structure elements (amino acid residues) that interact with each portion of the DNA substrate. Compared with the standard flap substrates, the 5'-overhang substrate has no upstream duplex portion, which allows us to determine which amino acid residues of FEN-1 interact with the upstream duplex. If the amino acid residues interact with this component, we expect that mutations of these residues will significantly affect FEN-1 flap cleavage activity using a standard flap substrate. However, when a 5'-overhang substrate that lacks the upstream duplex is used, the mutations should have no effect.

A comparison of the substrate cleavage patterns of wild type and mutant FEN-1 enzymes using the 5'-overhang substrate identifies residues that may contact the upstream duplex (Fig. 6). Notably, three mutants, R47A, R70A, and K326A/R327A, have similar or even higher enzyme activity than wild type FEN-1 at the normal cleavage site near the flap-duplex junction. The same mutants have weaker enzyme activities than wild type FEN-1 in the presence of normal flap or nick duplex substrates (Fig. 3). The sharp contrast in activities of the mu-

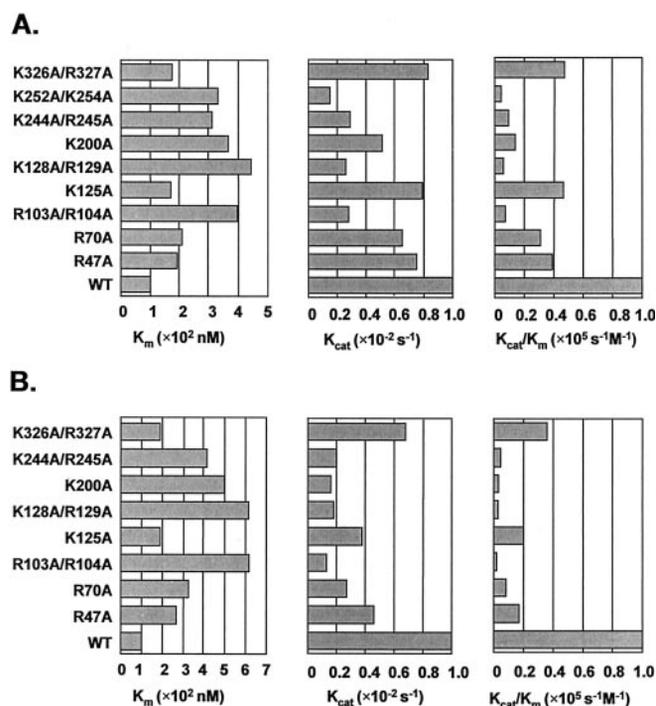


FIG. 5. **Kinetic assays of FEN-1 mutants.** Relative  $K_m$ ,  $K_{cat}$ , and  $K_{cat}/K_m$  values of FEN-1 mutants with partial endonuclease (*A*) and exonuclease (*B*) activity are shown. The assay conditions were described under "Experimental Procedures." The protein concentration for WT and mutants was 92 nM, and the substrate concentrations were 30–500 nM. The substrates are the same as those shown in Fig. 3*A* (flap substrate) and *B* (nick duplex substrate).

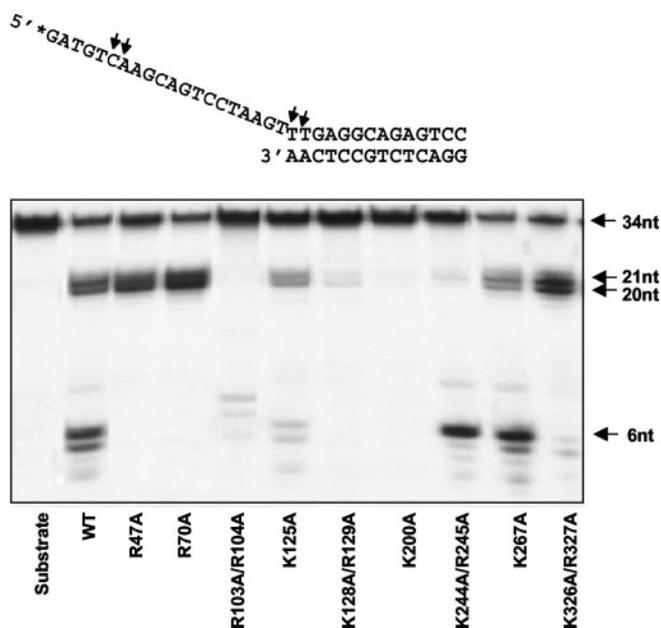


FIG. 6. **FEN-1 nuclease activities based on 5'-overhang substrate.** The top bands represent the uncut substrates, and the bottom bands are the cleavage products. The oligonucleotide size is indicated in nucleotides (nt). Reactions were carried out with 400 nM hFEN-1 protein and 80 nM flap or nick duplex substrate in reaction buffer containing 50 mM Tris (pH 8.0) and 10 mM MgCl<sub>2</sub> in a total volume of 10  $\mu$ l with water. Reactions were incubated at 30  $^{\circ}$ C for 15 min.

tants on different substrates suggests that Arg<sup>47</sup>, Arg<sup>70</sup>, and Lys<sup>326</sup>-Arg<sup>327</sup> interact with the upstream duplex portion of FEN-1 DNA substrates.

In addition, we noticed that R47A, R70A, and K326A/R327A had a cleavage pattern different from that of wild type FEN-1 on cleaving the 5'-overhang substrate. The wild type FEN-1

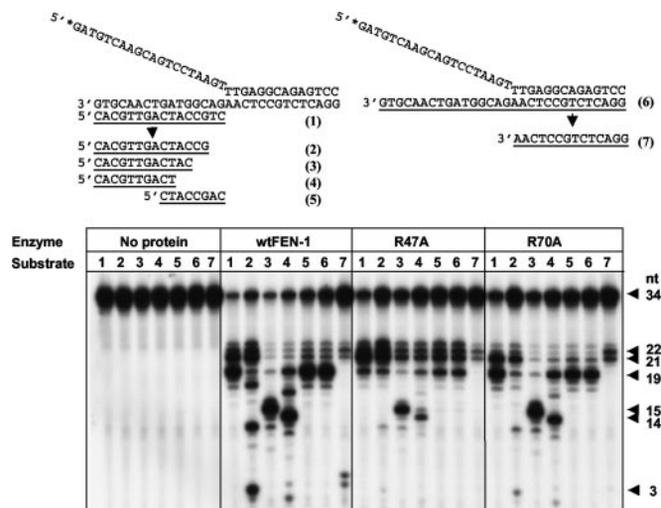
appeared to cleave the 5'-overhang substrates at two main sites: one near the flap-duplex junction and a second site close to the 5'-end of the single-stranded flap. Although the mutants R47A, R70A, and K326A/R327A cleaved this substrate near the flap-duplex junction with similar or greater efficiency than wild type FEN-1, the mutants did not cleave the substrate at the second site near the 5'-end of the flap. This result may indicate a "pulling-back" mechanism of the single-stranded flap: In the absence of the upstream duplex portion of the 5'-overhang substrate, the positively charged residues Arg<sup>47</sup>, Arg<sup>70</sup>, and Lys<sup>326</sup>-Arg<sup>327</sup> in FEN-1 become free and exposed, thus they are able to interact with the single-stranded flap. The interactions may pull the single-stranded flap to form a structure, which leads its 5'-end portion to be close to the active center of FEN-1 and generates the additional cleavage sites. However, mutation of any of the residues Arg<sup>47</sup>, Arg<sup>70</sup>, and Lys<sup>326</sup>-Arg<sup>327</sup> affects the stability of the unknown structure, which leads to the alteration of the cleavage sites close to the 5'-end of the flap strand.

In contrast to the R47A, R70A, and K326A/R327A mutants, K244A/R245A differs from the wild type FEN-1 in the opposite manner with regard to the two cleavage sites. The cleavage site of K244A/R245A mutant shifted from one near the flap-duplex junction to the one near the 5'-end of the single-stranded flap (Fig. 6). This result suggests that K244/R245 interacts with downstream duplex such that the residues are close to the cleavage sites near the flap-duplex junction but distant from the cleavage sites toward the 5'-end of the single-stranded flap. This explains why the mutation of K244A/R245A significantly affects cleavage near the flap-duplex junction but has little effect on the cleavage near the 5'-end of the single-stranded flap.

Although the cleavage patterns of the R103A/R104A and K244A/R245A mutants look similar, they differ in the relative effects on substrate cleavage at each site. At the sites near the flap-duplex junction, the R103A/R104A mutant only makes slight cleavages, whereas K244A/R245A makes much stronger cleavages. Because Arg<sup>103</sup> and Arg<sup>104</sup> are conserved residues in the large loop region, mutation of these residues should alter the interaction between the large loop and the single-stranded flap, which may affect cleavage at either site. Thus, this result is consistent with the previous proposals that the large loop region interacts with the single-stranded flap (21–24, 40). As discussed previously, Lys<sup>125</sup> and Lys<sup>128</sup>-Arg<sup>129</sup> might also interact with the single-stranded flap. Therefore, the cleavage patterns of the two mutants K125A and K128A/R129A could be similarly explained. In addition, the K200A mutation totally abolished the FEN-1 activity on the 5'-overhang substrate. Because Lys<sup>200</sup> is located near the active center (Fig. 2), its mutation could affect protein-DNA interactions near the active site and therefore prevent cleavage of the substrate.

**Mapping of Site-specific Interactions of FEN-1 with DNA**—Thus far, we have determined the orientation of DNA substrate binding to FEN-1, which shows that Arg<sup>47</sup> and Arg<sup>70</sup> of FEN-1 interact with the upstream duplex portion, Lys<sup>244</sup>-Arg<sup>245</sup> interacts with the downstream duplex portion, and the large loop region represented by Arg<sup>103</sup>-Arg<sup>104</sup> and Lys<sup>128</sup>-Arg<sup>129</sup> probably interacts with the single-stranded flap. In this experiment, we attempted to map the specific interaction sites of Arg<sup>47</sup> and Arg<sup>70</sup> with the upstream duplex. Because FEN-1 activity is only partially affected by the absence of an upstream duplex in the DNA substrates, it is reasonable to map the interaction sites using substrates with gradual changes in the length of the upstream duplex (Fig. 7).

The R47A and R70A mutants cleave both a flap substrate (Fig. 7, substrate 1) and a pseudo-Y substrate (Fig. 7, substrate

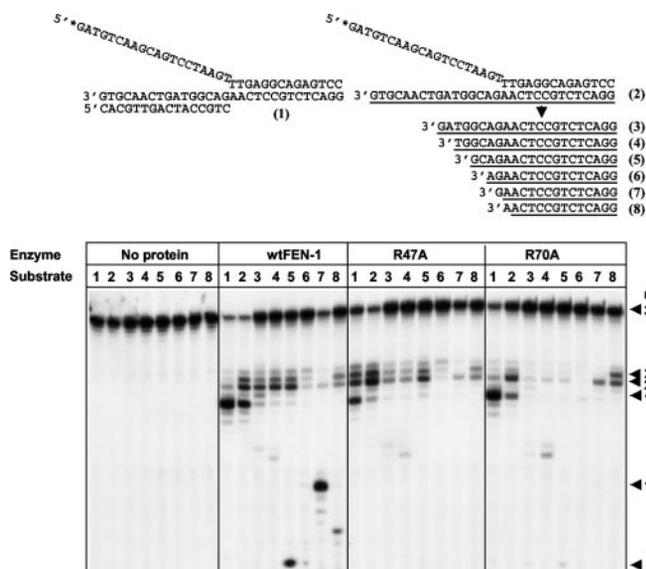


**FIG. 7. Enzyme activities and patterns of R47A, R70A, and wild type FEN-1 based on flap substrates with variant upstream gap lengths.** The top panel lists all the substrate used for the analysis. The arrow indicates the sequences of the changed upstream primers or template strand in the following substrate(s). For the bottom panel, the top band represents the uncut substrates, and the bottom bands are the cleavage products. Lane numbers correspond to the numbers of substrates used for the assay. The oligonucleotide size is indicated in nucleotides (nt). Reactions were carried out with 133 nM hFEN-1 protein and 80 nM flap or nick duplex substrate in reaction buffer containing 50 mM Tris (pH 8.0) and 10 mM MgCl<sub>2</sub> in a total volume of 10  $\mu$ l with water. Reactions were incubated at 30  $^{\circ}$ C for 15 min.

6), generating products that are similar to wild type FEN-1 for each substrate. Because the substrate cleavage patterns are not affected by the absence of the upstream primer, it is not likely that Arg<sup>47</sup> or Arg<sup>70</sup> interacts with the upstream primer. This interpretation is supported further by the cleavage of the mutants on other substrates with different gap lengths (Fig. 7, substrates 2–5), which show a proportional reduction in enzyme activities of the mutants compared with the wild type FEN-1.

Notably, the substrate cleavage pattern resulting from both the wild type and mutant FEN-1 proteins is altered by the introduction of gaps into the DNA substrates. It appears that the 3'-portion of the upstream primer is critical because the shortening of the primer at the 3'-end (substrates 2–4) reduces activity and alters the cleavage pattern. Although the exact reason for how the gap length alteration (especially from substrate 2 to 4) causes such a significant change in the cleavage pattern is unknown, we speculate that the single-stranded flap might go through a conformational change, which leads to cleavage site shifts. The conformational alteration could be triggered by the exposed amino acid residues caused by the introduction of gaps at the 3'-end of the upstream primer or the interaction of the nucleotide residue at 3'-end of the upstream primer with the single-stranded flap, or by both of the factors.

To identify the specific interaction sites of Arg<sup>47</sup> and Arg<sup>70</sup> on the complementary (template) stand of the upstream primer in the upstream duplex portion of flap substrates, we examined the substrate cleavage of the R47A and R70A mutants using pseudo-Y substrates with various 3'-overhang lengths. Shortening the length of the upstream template strand by up to 10 nucleotides from the end significantly reduced the flap endonuclease activity of R47A compared with wild type FEN-1 with respect to cleavage sites near the flap-duplex junction (Fig. 8, substrates 1–4). However, further reduction of the length of the upstream template strand does not affect R47A cleavage at the flap junction compared with wild type FEN-1 (Fig. 8, substrates 5–8). Because substrates 4 and 5 differ by only 2 nu-



**FIG. 8. Enzyme activities and cleavage patterns of R47A, R70A, and wild type FEN-1 on pseudo-Y substrates with variant 3'-overhang lengths.** The top panel lists all of the substrates used for the analysis. The arrow indicates the sequences of the changed template strands in the following substrates. For the bottom panel, the top band represents the uncut substrates, and the bottom bands are the cleavage products. Lane numbers correspond to the numbers of substrates used for the assay. The oligonucleotide size is indicated in nucleotides (nt). The reaction conditions are the same as in Fig. 6 except different substrates were used.

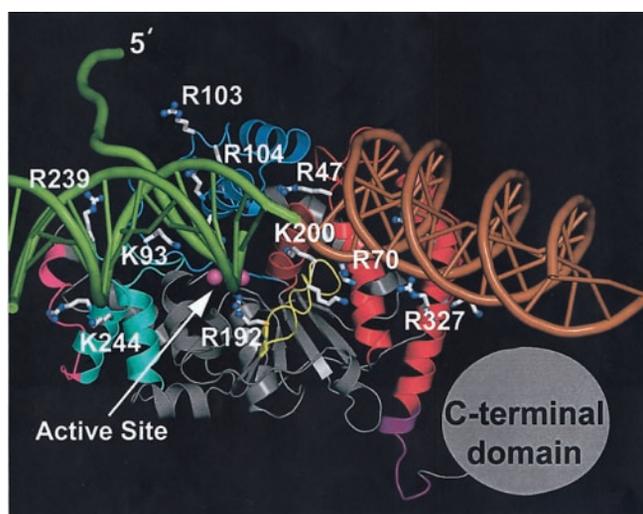
cleotides it is reasonable to propose that Arg<sup>47</sup> interacts with nucleotide 5 or 6 on the upstream template.

For the mutant R70A, shortening of the upstream template strand reduced substrate cleavage efficiency at the flap-duplex junction compared with wild type FEN-1 (Fig. 8, substrates 1–6). However, in the presence of very short upstream template strands (Fig. 8, substrates 7 and 8), R70A activity was comparable with wild type FEN-1. Because substrates 6 and 7 differ by 1 nucleotide (nucleotide 2 counting from the flap-duplex junction), Arg<sup>70</sup> likely interacts with the 2nd nucleotide after the flap junction. Surprisingly, pseudo-Y substrates with short upstream template strands were cleaved at additional sites by wild type FEN-1, generating products of 3, 6, or 10 nucleotides in length (Fig. 8, substrates 5, 7–8). We think the additional cleavage by wild type FEN-1 is caused by the pulling-back of the single-stranded flap as we discussed in the section of using the 5'-overhang substrates to determine the orientation of FEN-1 binding.

#### DISCUSSION

To function properly, FEN-1 nuclease requires the appropriate interaction between the protein and its DNA substrates. It is important, therefore, to understand how FEN-1 recognizes, binds, and cleaves its substrates. In the current study, we defined the region of substrate DNA protected by FEN-1 binding and identified a number of positively charged amino acids critical for FEN-1 binding to DNA.

The exoIII footprinting analysis indicates that FEN-1 protects 13 nucleotides of the downstream duplex and 20 nucleotides of the upstream duplex. The significantly larger length protected length by FEN-1 in the upstream duplex portion may be explained by the existence of a C-terminal portion in FEN-1, which contains a nuclear localization signal (20) and was proposed to play a role in substrate binding (42). Although the structure of this C-terminal portion remains unknown, it may interact with the upstream duplex as we proposed in Fig. 9. Based on a previous footprinting analysis (26), the border of the



**FIG. 9. A model showing the general orientation, borders, and some specific interaction sites in the human FEN-1-DNA substrate complex.** A structural model of human FEN-1 bound to a flap DNA substrate (ribbon representation) was created based on the crystal structure of affFEN-1 bound to DNA (PDB ID: 1RXW). Residues determined to affect DNA binding (white residues) interact with either the upstream DNA (20 bp, brown), the downstream DNA (13 bp, green), or the 5'-flap (10 nucleotides, green), defining the substrate binding orientation. Residues from two key  $\alpha$ -loop- $\alpha$  regions (red) interact with the upstream DNA duplex. Several different regions of FEN-1, including the H3TH motif (small loop 2, cyan), along with other loops (small loop 3, cyan; small loop 4, magenta) interact with the downstream duplex. Residues in the helical arch (large loop, blue) likely interact with the 5'-flap. The C-terminal domain (gray circle), which follows the proliferating cell nuclear antigen binding domain (purple), could interact with the upstream DNA and thus protect the full 20 bp.

single-stranded flap is located near the 25th nucleotide counting from the flap duplex junction. Taken together, we now have comprehensive information on the borders of a DNA substrate bound to FEN-1. In addition, the current study provides the information on the orientation of DNA substrate bound to FEN-1. Our results indicate that FEN-1 residues Arg<sup>47</sup> and Arg<sup>70</sup> contact the template strand within upstream duplex, residues Lys<sup>244</sup>-Arg<sup>245</sup> and Lys<sup>251</sup>-Lys<sup>254</sup> interact with the downstream duplex, and that residues located in the large loop region of FEN-1 bind to the single-stranded flap. Therefore, these data allow us to draw an overall picture of interaction of human FEN-1 with a DNA flap substrate (Fig. 9).

We further determined key elements of human FEN-1 in interaction with different portions of flap substrates. Although a number of archaeal FEN-1 crystal structures were solved at high resolution, for human and other eukaryotic FEN-1s, no crystal structure is available. It becomes a difficult task, hence, for researchers to establish an accurate model to address protein-DNA interaction of eukaryotic FEN-1 enzymes. In the current study, we performed activity, gel shift, and kinetic assays, which enable us to determine 14 mutants, R47A, R70A, K93A, R103A/R104A, K125A, K128A/R129A, K132A, R192A, K200A, K201A, K244A/R245A, K252A/K254A, K267A, and K326A/R327A with deficiency in the interaction with DNA substrates. These 14 mutants cover 18 positively charged surface and conserved amino acid residues that appear important for substrate binding. These amino acid residues can be assigned into different loop regions with reference to the archaeal FEN-1 (31) (Table I). Among these residues, Arg<sup>47</sup> is in small loop 1, which is a close neighbor to Arg<sup>70</sup>. Both of these residues were determined in our previous studies (30) to be important in substrate binding. Arg<sup>192</sup>, Lys<sup>200</sup>, and Lys<sup>201</sup> are located in small loop 2. This loop is extremely critical as mutation of any of the three residues Arg<sup>192</sup>, Lys<sup>200</sup>, and Lys<sup>201</sup> significantly

affects both DNA binding and activity. Residues Lys<sup>244</sup>, Arg<sup>245</sup>, Lys<sup>252</sup>, Lys<sup>254</sup>, and Lys<sup>267</sup> are in adjacent small loops 3 and 4, whereas Lys<sup>326</sup> and Arg<sup>327</sup> are located on the other side of FEN-1, closer to Arg<sup>70</sup> (Fig. 2). Finally, Lys<sup>93</sup>, Arg<sup>104</sup>, Lys<sup>125</sup>, Lys<sup>128</sup>, Arg<sup>129</sup>, and Lys<sup>132</sup> all belong to the large loop in the arch structure of FEN-1, which was proposed to be important in interaction with single-stranded DNA flap (21–24, 40).

For years, the large loop region has been recognized as the most critical element for DNA recognition and interaction. A thread-through or tracking-down model was proposed to account for the potential role of this region in recognizing the single-stranded flap (40, 43). The corresponding residues of the large loop region in the 5'-nuclease of *E. coli* DNA polymerase I and archaeal FEN-1 (24) were also shown to have roles in DNA binding (24, 25). However, in a recent study some amino acid residues in the large loop region were revealed to be critical for catalysis (27). The mutations of Lys<sup>93</sup> into arginine and Ser<sup>94</sup>, Arg<sup>100</sup>, Leu<sup>97</sup>, or Leu<sup>130</sup> into proline in this region significantly affected enzyme activities of human FEN-1 but had little effect on DNA binding (27). These results imply that the large loop region has dual roles in both DNA interaction and catalysis. To address this possibility, we mutated almost all the positively charged amino acid residues located in the large loop into alanine. Our results show that the large loop is indeed involved in DNA interactions because the mutation of the majority of these amino acids into alanine either completely abolishes (K132A) or significantly decreases (R104A, K125A, K128A/R129A) FEN-1 enzyme activity and substrate binding affinity. In addition, our study identified two mutants, K93A and R100A, which retained an intact or slightly diminished substrate binding capability but have no enzyme activity, suggesting that the large loop region could participate in catalysis.

It was proposed that the large loop region could be important for efficient cleavage by positioning the 5'-flap near the catalytic site (27). This hypothesis is supported by recent evidence that Arg<sup>94</sup> in *P. furiosus* FEN-1, which corresponds to Arg<sup>100</sup> in human FEN-1, makes ionic contact with the phosphate bonds near the cleavage site (44). However, according to crystal structures of archaeal FEN-1 enzymes (23, 24) and the structural model of human FEN-1 (Fig. 2), Arg<sup>100</sup> is located at the end of the large loop region, far away from the active site of FEN-1. Therefore, for Arg<sup>100</sup> to play a role in catalysis, the large loop region must undergo a significant conformational change, which would position the residue near the active site. A conformation change would also explain why mutating large loop residues Ser<sup>94</sup>, Leu<sup>97</sup>, or Leu<sup>130</sup> into proline, which would affect the flexibility of the large loop, significantly decreases human FEN-1 activity (27). Alternatively, the elements for catalysis in the large loop region of FEN-1, such as Arg<sup>100</sup> and Lys<sup>93</sup>, might participate in a mechanism for maintaining an extensive catalytic solvent network within the active site. Such a mechanism was described by Chevalier *et al.* (41) to address DNA catalysis by a homing endonuclease (I-CreI). In this nuclease, residues with basic side chains form a network surrounding the nucleophilic water molecules, extending around the scissile phosphate to the 3'-oxygen leaving group. The network supports a DNA cleavage mechanism in which the scissile phosphates contact two divalent cations that are extensively hydrated by several water molecules. The hydration is structured and polarized by interactions with a number of basic side chains. Disturbance of such a network is similar to mutations of metal ion ligands, which can abolish enzyme activities (15).

To establish a detailed model of FEN-1 interaction with DNA we mapped the specific interaction sites of residues Arg<sup>47</sup> and Arg<sup>70</sup> on the upstream duplex. Because the lack of an upstream duplex portion in flap DNA substrates does not completely

abolish the activity of wild type FEN-1 or the R47A and R70A mutants, it is possible to map the protein-DNA interaction sites in the upstream duplex using flap substrates with gradual change of the upstream duplex lengths. This method is more precise than footprinting because it uses substrates with well defined sequence changes. Based on this method, we determined that Arg<sup>70</sup> interacts with DNA near the 2nd nucleotide and Arg<sup>47</sup> interacts with either the 5th or 6th nucleotide of the template strand in the upstream duplex, counting from the nick point of the flap substrate (Fig. 8). Although we did not expect that Arg<sup>47</sup> could interact with the template strand in the upstream duplex portion, we predicted that Arg<sup>70</sup> interacts with the template strand in the upstream duplex portion based on previous results (29). Notably, Arg<sup>70</sup> in human FEN-1 appears to interact with DNA at the same site as Arg<sup>64</sup> in both pfFEN-1 and afFEN-1 (28, 44). This result indicates that human and archaeal FEN-1 enzymes may have a very similar protein-DNA interaction mechanism.

Compared with the model proposed for archaeal FEN-1 (44), our results experimentally confirmed proposed roles of a number of positively charged amino acid residues in substrate binding and identified several new residues such as Arg<sup>192</sup>, Lys<sup>200</sup>, and Lys<sup>201</sup> that are critical for substrate binding in human FEN-1. In addition, we determined the specific interaction sites of Arg<sup>47</sup> and Arg<sup>70</sup> with DNA. The new information supports the proposal that the DNA substrate might be kinked upon binding to FEN-1 (28). Upon binding, there could be a significant conformational change in FEN-1, especially in the large loop region, which positions amino acid residues, such as Arg<sup>100</sup>, closer to the active site to participate in catalysis. In addition, a conformational change in the region of small loop 1 may allow Arg<sup>47</sup> to shift its orientation from that shown in Fig. 9 toward favoring the interaction with nucleotide 5 or 6 on the upstream template as revealed by biochemical analysis.

**Acknowledgments**—We are grateful for the contributions or efforts made by David N. Bimston, Arthur Partikian, and Paul Luu in the initial stage of this work. Thanks also go to Steve Alas for critical reading of the manuscripts and valuable comments, and to Li Zheng and other members of the Shen laboratory for technical assistance and stimulating discussions.

#### REFERENCES

1. Turchi, J. J., Huang, L., Murante, R. S., Kim, Y., and Bambara, R. A. (1994) *Proc. Natl. Acad. Sci. U. S. A.* **91**, 9803–9807
2. Bambara, R. A., Murante, R. S., and Henricksen, L. A. (1997) *J. Biol. Chem.* **272**, 4647–4650
3. Hubscher, U., and Seo, Y. S. (2001) *Mol. Cell* **12**, 149–157
4. Klungland, A., and Lindahl, T. (1997) *EMBO J.* **16**, 3341–3348
5. Johnson, R. E., Kovvali, G. K., Prakash, L., and Prakash, S. (1995) *Science* **269**, 238–240
6. Gordenin, D. A., Kunkel, T. A., and Resnick, M. A. (1997) *Nat. Genet.* **16**, 116–118
7. Freudenreich, C. H., Kantrow, S. M., and Zakian, V. A. (1998) *Science* **279**, 853–856
8. Spiro, C., Pelletier, R., Rolfsmeier, M. L., Dixon, M. J., Lahue, R. S., Gupta, G., Park, M. S., Chen, X., Mariappan, S. V. S., and McMurray, C. T. (1999) *Mol. Cell* **4**, 1079–1085
9. Tishkoff, D. X., Filosi, N., Gaida, G. M., and Kolodner, R. D. (1997) *Cell* **88**, 253–263
10. Kunkel, T. A., Resnick, M. A., and Gordenin, D. A. (1997) *Cell* **88**, 155–158
11. Chen, C., and Kolodner, R. D. (1999) *Nat. Genet.* **23**, 81–85
12. Negritto, M. C., Qiu, J., Ratay, D. O., Shen, B., and Bailis, A. M. (2001) *Mol. Cell. Biol.* **21**, 2349–2358
13. Parrish, J. Z., and Xue, D. (2003) *Mol. Cell* **11**, 987–996
14. Harrington, J. J., and Lieber, M. R. (1994) *Genes Dev.* **8**, 1344–1355
15. Shen, B., Nolan, J. P., Sklar, L. A., and Park, M. S. (1996) *J. Biol. Chem.* **271**, 9173–9176
16. Shen, B., Qiu, J., Hosfield, D., and Tainer, J. A. (1998) *Trends Biochem. Sci.* **23**, 171–173
17. Warbrick, E., Lane, D. P., Glover, D. M., and Cox, L. S. (1997) *Oncogene* **14**, 2313–2321
18. Frank, G., Qiu, J., Zheng, L., and Shen, B. (2001) *J. Biol. Chem.* **276**, 36295–36302
19. Kim, K., Biade, S., and Matsumoto, Y. (1998) *J. Biol. Chem.* **273**, 8842–8848
20. Qiu, J., Li, X., Frank, G., and Shen, B. (2001) *J. Biol. Chem.* **276**, 4901–4908
21. Ceska, T. A., Sayers, J. R., Stier, G., and Suck, D. (1996) *Nature* **382**, 90–93
22. Mueser, T. C., Nossal, N. G., and Hyde, C. C. (1996) *Cell* **85**, 1101–1112

23. Hosfield, D., Mol, C. D., Shen, B., and Tainer, J. A. (1998) *Cell* **95**, 135–146
24. Hwang, K. Y., Baek, K., Kim, H.-Y., and Cho, Y. (1998) *Nat. Struct. Biol.* **5**, 707–713
25. Xu, Y., Potapova, O., Leschziner, A. E., Grindley, N. D. F., and Joyce C. M. (2001) *J. Biol. Chem.* **276**, 30167–30177
26. Barnes, C. J., Wahl, F., Shen, B., Park, M. S., and Bambara, R. A. (1996) *J. Biol. Chem.* **271**, 29624–29631
27. Storici, F., Henneke, G., Ferrari, E., Gordenin, D. A., Hubscher, U., and Resnick, M. A. (2002) *EMBO J.* **21**, 5930–5942
28. Kim, C. Y., Park, M. S., and Dyer, R. B. (2001) *Biochemistry* **40**, 3208–3214
29. Chapados, B. R., Hosfield, D. J., Han, S., Qiu, J., Yelent, B., Shen, B., and Tainer, J. A. (2004) *Cell* **116**, 39–50
30. Qiu, J., Bimston, D. N., Partikian, A., and Shen, B. (2002) *J. Biol. Chem.* **277**, 24659–24666
31. Matsui, E., Musti, K. V., Abe, J., Yamasaki, K., Matsui, I., and Harata, K. (2002) *J. Biol. Chem.* **277**, 37840–37847
32. Brunger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J.-S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) *Acta Crystallogr. D Biol. Crystallogr.* **54**, 905–921
33. Hosfield, D. J., Frank, G., Weng, Y., Tainer, J. A., and Shen, B. (1998) *J. Biol. Chem.* **273**, 27154–27161
34. Harrington, J. J., and Lieber, M. R. (1994) *EMBO J.* **13**, 1235–1246
35. Hohl, M., Thorel, F., Clarkson, S. G., and Scharer, O. D. (2003) *J. Biol. Chem.* **278**, 19500–19508
36. Chai, Q., Qiu, J., Chapados, B. R., and Shen, B. (2001) *Biochem. Biophys. Res. Commun.* **286**, 1073–1081
37. Yang, J. T., Wu, C. S., and Martinez, H. M. (1986) *Methods Enzymol.* **130**, 208–269
38. Lyamichev, V., Brow, M. A. D., Varvel, V. E., and Dahlberg, J. E. (1999) *Proc. Natl. Acad. Sci. U. S. A.* **96**, 6143–6148
39. Kao, H. I., Henricksen, L. A., Liu, Y., and Bambara, R. A. (2002) *J. Biol. Chem.* **277**, 14379–14389
40. Murante, R., Rust, L., and Bambara, R. A. (1995) *J. Biol. Chem.* **270**, 30377–30383
41. Chevalier, B. S., Monnat, R. J., Jr., and Stoddard, B. L. (2001) *Nature Struct. Biol.* **8**, 312–316
42. Stucki, M., Jonsson, Z. O., and Hubscher, U. (2001) *J. Biol. Chem.* **276**, 7843–7849
43. Bornarth, C. J., Ranalli, T. A., Henricksen, L. A., Wahl, A. F., and Bambara, R. A. (1999) *Biochemistry* **38**, 13347–13354
44. Allawi, H. T., Kaiser, M. W., Onufriev, A. V., Ma, W. P., Brogaard, A. E., Case, D. A., Neri, B. P., and Lyamichev, V. I. (2003) *J. Mol. Biol.* **328**, 537–554