Essential Amino Acids for Substrate Binding and Catalysis of Human Flap Endonuclease 1*

(Received for publication, January 16, 1996)

Binghiu Shent, John P. Nolan†, Larry A. Sklar‡, and Min S. Park§

From the§Life Sciences Division and §National Flow Cytometry Resource, Los Alamos National Laboratory, Los Alamos, New Mexico 87545 and the Department of Cytometry, School of Medicine, University of New Mexico, Albuquerque, New Mexico 87131

Human flap endonuclease 1 (FEN-1) is a member of the structure-specific endonuclease family and is involved in DNA repair. Eight restrictively conserved amino acids in FEN-1 have been converted individually to alanine to elucidate their roles in specific DNA substrate binding and catalysis. Flap endonuclease activity of the wild type and mutant enzymes was measured by kinetic flow cytometry. Mutants D34A, D86A, and D181A lost their cleavage activity completely but retained substrate binding ability, as measured by their ability to inhibit the wild type enzyme in a competition assay. This indicates that these amino acids contribute to integrity of the enzyme active site. Loss of both binding and cleavage competency for the flap substrate by mutants E156A, G231A, and D233A suggests that these amino acids are involved in substrate binding. Mutants R103A and D179A retained wild type-like enzyme activity.

Human flap endonuclease 1, a structure-specific endonuclease, recognizes a specific 5' flap DNA structure, the junction where the two strands of duplex DNA adjoin a single-stranded arm (Lymichev et al., 1993). FEN-1 binds to the arm with a free 5' end, slides down to the double-strand-single-strand junction of the substrate, and then cleaves the single-stranded arm (Murante et al., 1995). This in vitro activity suggests an in vivo function of the enzyme in DNA damage repair pathways (Harrington and Lieber, 1994b; Murray et al., 1994; Sommers et al., 1995; Reagan et al., 1995; Jhnson et al., 1995). Mouse FEN-1 and the yeast FEN-1 homologs, Saccharomyces cerevisiae YKL510 and Schizosaccharomyces pombe rad2 have been identified (I acquier et al., 1992; Harrington and Lieber, 1994a; Murray et al., 1994). Sequence comparison of FEN-1 with XPG family proteins revealed that two regions of the protein sequences are conserved (Carr et al., 1993; Madnnes et al., 1993; Scherly et al., 1993). The in vitro activity of the flap endonuclease has been found in the human XPG nuclease (Cloud et al., 1995). This indicates that members of the XPG family including human and mouse XPG and the yeast counterparts, S. cerevisiae RAD2 and S. pombe rad13, and the FEN-1 family are conserved not only in sequence but also in function.

Previously known 5' exo/endonuclease domains of several prokaryotic DNA polymerases such as Escherichia coli DNA polymerase I and Thermus aquaticus (Taq) DNA polymerase have now been shown to possess a structure-specific endonuclease activity that cleaves the 5' flap structure (Lymichev et al., 1993). When four bacteriophage exonuclease sequences have been compared with the above exo/endonuclease domains of the DNA polymerases, five conserved regions of these two groups of nucleases were identified (Gutman and Minton, 1993). Interestingly, the sequence of human flap endonuclease has some similarities to the conserved exo/endonuclease domain of E. coli DNA polymerase I (Robins et al., 1994). Taken together, this indicates that the structure and function of this 5' structure-specific nuclease activity are conserved among bacteria, yeasts, and mammals.

In this report, we compared sequences of the two conserved regions of 8 eukaryotic FEN-1/XPG family members, 6 exonuclease domains of prokaryotic DNA polymerases, and 4 bacteriophage exonucleases. We have identified eight amino acids that are restrictively conserved throughout these 18 exonuclease structures (Fig. 1). We predict that these amino acids might be essential for substrate binding or catalysis. To test this hypothesis, we have converted these amino acids individually to alanine, expressed the mutant genes in E. coli, and purified the mutant FEN-1 proteins to near homogeneity. The mutant enzymes were characterized for their ability to recognize and cleave the flap DNA structure by a recently developed assay employing kinetic flow cytometry. 3 This approach allows us to identify critical amino acids involved in DNA substrate binding and catalysis.

**EXPERIMENTAL PROCEDURES**

Materials—Restriction enzymes, EcoRI, BamHI, and NcoI, and 100 mg/ml bovine serum albumin solution were obtained from New England Biolabs. Sequenase Version 2.0 DNA sequencing kit was from United States Biochemical Corp. Expression vectors and the host E. coli strains were from Novagen (Madison, WI). All electrophoresis gels and reagents were obtained from Novex (San Diego, CA). Chelating Sepharose fast flow resin was from Pharmacia Biotech Inc. Protein assay kit was from Bio-Rad. All buffers and salts were molecular biology grade from Fisher Scientific. Streptavidin-coated polystyrene microspheres (6.2-μm diameter) were from Spherotech (Libertyville, IL). Fluorescein isothiocyanate-labeled standard microspheres were from Flow Cytometry Standards Corp (San Juan, PR). Sodium fluorescein was from Sigma. Oligonucleotide substrates were synthesized on an Applied Biosystems (ABI) DNA synthesizer, using nucleotide or BioTEG CPG support and nucleotide or fluorescein phosphoramidites (Glen Research, Sterling, VA).

Site-directed Mutagenesis—Table I shows the sequences of the oligonucleotides used to construct the eight site-directed mutants used in this study. The altered nucleotides are underlined. Nucleotide substitutions were incorporated at the desired locations in the donor FEN-1 pBluescript vector by using Chameleon Double-stranded Site-directed Mutagenesis Kit from Stratagene. The mutations were confirmed by double-stranded DNA sequencing using U. S. Biochemical Corp. Sequenase Version 2.0 kit.

Mutant Enzyme Expression and Purification—The pET-FCH plasmid has been made to overexpress the wild type human FEN-1 protein as described previously. 4 All of the mutated sequences described above

---

* This research has been supported by funds from the Los Alamos National Laboratory, Department of Energy, and National Institutes of Health. 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.

† To whom correspondence should be addressed: Los Alamos National Laboratory, Life Sciences Division, MS M888, Los Alamos, NM 87545. Tel: 505-667-5701; Fax: 505-665-3024; E-mail: park@elomere.lanl.gov.

‡ B. Shen, J. P. Nolan, L. A. Sklar, and M. S. Park, submitted for publication.

§ The abbreviations used are: FEN-1, human flap endonuclease 1; XPG, xeroderma pigmentosum G.
Substrate Binding and Catalysis of Human FEN-1

**Fig. 1.** Comparison of the protein sequences from human FEN-3, XPG and E. coli exonuclease domain of the DNA polymerase I, A, alignment of the amino acid sequences. The alignment was created by the DNA Star program using the dot matrix method using the weighted residue weight table. Conserved amino acids are indicated by a ; similar amino acids are indicated by a . The amino acids in the boxes are identical throughout all 18 structure-specific endo/exonuclease enzyme sequences compared. Top, E. coli exonuclease domain of DNA polymerase I representing 6 prokaryotic DNA polymerase exo/endonucleases and 4 bacteriophage exonucleases (Gutman and Minton, 1993). Middle, human FEN-3 (Murray et al., 1994) representing mouse FEN-1 (Harrington and Lieber, 1994a), S. cerevisiae, and S. pombe rad13 (Carr et al., 1993). Bottom, representation of the DNA polymerase I exonuclease domain (Lyamichev et al., 1989; Hollingsworth and Nossal, 1991), we extended this analysis to include 18 total amino acids for human and mouse FEN-1.

were subcloned into pET-FCH with Ncol and BamHI sites. The resultant plasmids were transformed into the E. coli host strain BL21(DE3) for protein expression. Enzyme expression procedures for most of the mutant proteins were essentially the same as for the wild type. For some mutant proteins with low solubility in LB medium, a sorbitol/betaine medium was used according to Blackwell and Horgan (1991), and the expression temperature was shifted to 25°C in order to enhance solubility. All of the purification steps were carried out either on ice or in a 4°C cold room. Bacterial lysates were made by sonicating in buffer S (50 mM Tris, pH 8.0, 50 mM sorbitol). The extract was then centrifuged at 30,000 x g for 1 h. The recombinant proteins were purified by passing the supernatant over a Ni2+-Sepharose column under nondenaturing conditions. The column was washed extensively with buffer A (20 mM Tris-HCl, pH 7.9, 0.5 mM NaCl, 5 mM imidazole). Buffer A containing 50 mM imidazole was used to remove weakly bound proteins. The FEN-1 mutant proteins were eluted with buffer A containing 1.0 mM imidazole. The eluted proteins were diluted 10-fold with phosphate-buffered saline (50 mM sodium phosphate, pH 7.0, 100 mM NaCl). The diluted proteins were dialyzed against 3 liters of phosphate-buffered saline for 12 h with 3 changes and then concentrated with polyethylene glycol powder (molecular mass = 12–20 kDa). The purity of the protein was analyzed on 10–20% gradient SDS-polyacrylamide gel electrophoresis. The protein was quantitated using a Bio-Rad detergent-compatible protein assay based on the Lowry assay. Enzyme concentrations were calculated based on the protein mass measurements and an estimated molecular mass of 42.5 kDa.

Flow Cytometric Nuclease Activity Assay—The flap substrate labeling, immobilization of the flap substrate to the microsphere, and calibration were carried out according to the protocol described recently. In a typical reaction, carried out at room temperature, 5–7 nM immobilized flap substrate was mixed in R buffer (50 mM Tris, pH 8.0, 10 mM MgCl2, 100 μg/ml bovine serum albumin) in total volume of 500 μl. The substrate alone was analyzed—5 s to establish a baseline. The sample tube was removed from the tube holder, approximately 50 nM enzyme was added, the tube was vortexed, and sample was reintroduced to the instrument. The first time point was acquired 8–9 s after mixing.

The internal dyes of the FACScan/IIr’s Macintosh acquisition computer was used to track the time parameter. The data obtained were analyzed using IDLYK, a flow cytometry data analysis program developed at Los Alamos National Laboratory. Time data were collected at 1024-channel resolution (500 ms/channel) and rebinned at 64-channel resolution (8 s/channel) for displaying. The time value for a given data point was the midpoint of the time window measured. For clarity, only every fifth data point is displayed for each time point.

**Results**

The amino acid sequences of eight proteins from the XPG/FEN-1 family in mammals and yeast were compared using computer program DNA Star. This analysis revealed the regions in the N-terminal (29–114 for hFEN-1) and in the middle part (129–260 for hFEN-1) of the protein amino acid sequence as described by Murray et al. (1994). Within these two regions (128 total amino acids for hFEN-1), most amino acids are identical. Moreover, exonuclease domains of prokaryotic DNA polymerases have been shown to have flap endonuclease activity (Lyamichev et al., 1993). When we extended this comparison to 6 exonuclease domains of the prokaryotic DNA polymerases, including E. coli DNA polymerase I and Taq DNA polymerase and 4 bacteriophage exonucleases (Lopez et al., 1989; Beck et al., 1989; Hollingsworth and Nossal, 1991), we found that 11 amino acids are similar and 8 of them are identical among these 18 enzymes (Fig. 1, A and B).

Each of these 8 charged or hydrophilic amino acids was converted to alanine by site-directed mutagenesis. The mutant proteins were overexpressed to a level similar to that of the wild type enzyme (Fig. 2A). The solubility of the mutant proteins differed among most of mutant proteins were comparable to the recombinant wild type protein. The most insoluble mutant proteins (G231A and D233A) were solubilized in vivo by employing the sorbitol/betaine medium according to Blackwell and Horgan (1991). The His-6-tagged FEN-1 mutant proteins were purified from the soluble part of the total crude E. coli cell extract using a nickel-Sepharose column as shown in Fig. 2B. The purified mutant proteins migrated as wild type protein in a 10–20% SDS-polyacrylamide gel electrophoresis

**Table I**

| Mutant | Mutation site | Oligonucleotides containing altered codons |
|--------|---------------|--------------------------------------------|
| D34    | 5′-GTGGGCCATTGGCTGGCT-3′ | |
| D86A   | 5′-GCTTTTGCTGGAACGCG-3′ | |
| R103A  | 5′-CGGATGGCGCCCGGCTG-3′ | |
| E158A  | 5′-GCCACCACGGGGGGAGAGGG-3′ | |
| D179A  | 5′-GGCTTACGGGAGCATGACGTG-3′ | |
| D183A  | 5′-CGAGGATGGGCTGCTTCAA-3′ | |
| G231A  | 5′-CTTGGTACGGTGGTACTGCT-3′ | |
| D233A  | 5′-GCAAAAGGGCTCCTGGGAGG-3′ | |
Flap endonuclease activity of the wild type and mutant proteins was screened by a flow cytometry-based assay system which has been described recently. Six of the eight mutant proteins (D34A, D86A, E158A, D181A, G231A, D233A) exhibited no flap endonuclease activity when the assay was carried out at protein concentrations of 50 nM (Fig. 3) or at 150 nM even after 30 min of incubation (data not shown). Conversion of an arginine at position 103 or aspartate 179 to an alanine did not affect the flap endonuclease activity compared to the wild type enzyme.

Mutants with no catalytic activity are expected to have a defect in substrate binding, cleavage, or both. To examine whether an individual inactive mutant was capable of substrate binding, we performed a competition assay with the wild type enzyme. For these experiments, the inactive mutant enzymes were preincubated 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. If an inactive mutant is capable of binding to the flap substrate, it will compete with and reduce the apparent activity of the wild type enzyme. Fig. 4 shows that of the six inactive mutants, three (D34A, D86A, E158A, D181A, G231A, D233A) partially inhibit the wild type FEN-1 activity at a concentration three times higher than wild type FEN-1. By increasing the concentration of mutant enzyme, the wild type activity can be completely inhibited (data not shown) although the concentration required for complete inhibition by these three mutants are different. On the other hand, the other three catalytically inactive mutants E158A, G231A, and D233A did not inhibit the wild type enzyme activity at all under the same experimental conditions. The results are summarized in Table II.

**DISCUSSION**

Isolation and characterization of homologous genes from different organisms provide valuable information on structurally conserved regions that are likely to be important for protein function or for macromolecular interactions. Flap endonuclease 1 genes from human, mouse, and two yeast species constitute a highly conserved family. The encoded proteins have homologous regions with the XPG family (Carr et al., 1993; MacInnes et al., 1993; Scherly et al., 1993; Murray et al., 1994). When the comparison of these protein sequences was extended to the prokaryotic and viral proteins that exhibit exonuclease/endonuclease activities, the number of absolutely conserved amino acids is reduced to a few. In order to evaluate the roles of these evolutionarily conserved amino acids in catalysis, we changed these amino acids to alanine. FEN-1 enzyme activity was assayed with a recently developed flow cytometer-based nuclease assay. Substrate binding by catalytically inactive mutants was measured by competition against the wild-type enzyme. Table II summarizes the ability of eight mutant enzymes to catalyze DNA cleavage and compete with the wild type enzyme. These data indicate that mutation of Asp-34, Asp-86, and...
Asp-181 lead the enzyme to lose catalytic activity completely but retain the binding ability to the flap substrate which is on the same order as wild type substrate binding. Therefore, we propose that these amino acids are required for the catalytic activity of the enzyme. Three other amino acids, Glu-158, Gly-231, and Asp-233, are important for binding of DNA substrate because their replacement by alanine leads not only to the loss of cleavage activity, but also to the loss of the ability to compete with the wild type enzyme. This indicates that the affinity of these mutants for the flap DNA substrate has been reduced several orders of magnitude relative to the wild type enzyme, and that the $K_D$ for substrate binding must be much greater than 150 nM, the concentration of mutant enzyme tested here. Mutation of Arg-103 and Asp-179 did not change activity of the enzyme. Interestingly, the amino acids constituting the active site of the enzyme (Asp-34, Asp-86, Asp-181) are all aspartate. The amino acids with a carboxyl group have been identified as components of an active site in several nucleases involved in DNA repair or recombination. We have found amino acids essential for substrate binding and catalysis. The information provided here should be useful for further studies on structure and function of this enzyme and other nucleases involved in DNA repair.

Acknowledgments—We thank Kieran G. Cloud for synthesis of all the oligonucleotides used in the experiments and Robb C. Habberset for customizing his IDLYK flow cytometry data analysis program to facilitate kinetic analysis. We also thank Dr. Ronald Gary and Paul Gasser for their stimulating discussions and critical reading of the manuscript.

REFERENCES

Barzilay, G., Walker, L. J., Robson, C. N., and Hickson, I. D. (1995) Nucleic Acids Res. 23, 1544–1550

Beck, P., Gonzalez, S., Ward, C., and Molineux, I. (1989) J. Mol. Biol. 210, 687–701

Breeze, L. S., and Steltz, T. A. (1991) EMBO J. 10, 25–33

Blackwell J. R., and Steitz, T. A. (1991) FEBS Lett. 295, 10–12

Carr, A. M., Sheldon, K. S., Murky, J. S., Al-Harithy, R., Watta, F. Z., and Lehman, A. R. (1993) Nucleic Acids Res. 21, 1345–1349

Cloud, K., Shen, B., Strniste, G. F., and Park, M. S. (1995) Mutat. Res. 347, 55–60

Derbyshire, V., Grindley, N. D. F., and Joyce, C. M. (1991) EMBO J. 10, 17–24

Gutman, P. D., and Minton, K. W. (1993) Nucleic Acids Res. 21, 4406–4407

Harrington, J. J., and Lieber, M. (1994a) Genes & Dev. 8, 1344–1355

Hollworth, H. C., and Nossal, N. G. (1991) J. Biol. Chem. 266, 1888–1897

Jacquier, A., Legrain, P., and Dujon, B. (1992) Yeast 8, 121–132

Johnson, R. E., Kovvali, G. K., Prakash, L., and Prakash, S. (1995) Science 269, 238–240

Lyamichev, V., Brow, M. A. V., and Dahlberg, J. E. (1993) Science 260, 778–783

Lopez, P., Martinez, S., Diaz, A., Espinao, M., and Lacks, S. A. (1989) J. Biol. Chem. 264, 4255–4263

Madrane, M. A., Dickson, J. A., Hernandez, R. R., Learmonth, D., Lin, G. Y., Mudgett, J. S., Park, M. S., Shauer, S., Reynolds, R. J., Strniste, G. F., and Yu, Y. J. (1993) Mol. Cell. Biol. 13, 6393–6402

Murante, R. S., Rust, L., and Bambara, R. A. (1995) J. Biol. Chem. 270, 30377–30383

Murray, J. M., Tavassoli, M., Al-Harithy, R., Sheldon, K. S., Lehman, A. R., Carr, A. M., and Watts, F. Z. (1994) Mol. & Cell Biol. 14, 4878–4888

Reagan, M. S., Pittenger, C., Siede, W., and Friedman, E. C. (1995) J. Bacteriol. 177, 364–371

Robins, R., Papin, D. J. C., Wood, R. D., and Lindahl, T. (1994) J. Biol. Chem. 269, 28535–28538

Saeng, W. (1991) Curr. Opin. Struct. Biol. 1, 130–138

Saito, A., Iwasaki, H., Ariyoshi, M., Murky, J. S., and Shinkawa, H. (1995) FEBS Lett. 372, 7470–7474

Scherly, D., Nouspikel, T., Corlet, J., Uda, C., Bairoch, A., and Clarkson, S. G. (1993) Nature (Lond.) 363, 182–185

Shomi, T., Harada, Y., Saito, T., Shomi, N., Okuno, Y., and Yamaizumi, M. (1994) Mutat. Res. 314, 167–175

Sommers, C. H., Miller, E. J., Dujon, B., Prakash, S., and Prakash, L. (1995) J. Biol. Chem. 270, 4193–4196

Steltz, T. A., and Steltz, J. A. (1993) Proc. Natl. Acad. Sci. U.S.A. 90, 7470–7474

Steltz, T. A., and Steltz, J. A. (1993) Proc. Natl. Acad. Sci. U.S.A. 90, 6498–6502

Suck, D., and Oefner, C. (1986) Nature (Lond.) 321, 620–625

Suck, D., Lahm, A., and Oefner, C. (1988) Nature (Lond.) 332, 464–468

Volbeda, A., Lahm, A., Sakiyama, F., and Suck, D. (1991) EMBO J. 10, 1607–1618

Weber, D. J., Meeker, A. K., and Mildvan, A. S. (1991) Biochemistry 30, 6103–6114

\footnote{B. Shen, J. P. Nolan, L. A. Sklar, and M. S. Park, unpublished data.}
