Characterization of a Putative Helix-Loop-Helix Motif in Nucleotide Excision Repair Endonuclease, XPG*

Min S. Park†, Joseph Valdez, Lawrence Gurley, and Chang-Yub Kim

From the Life Sciences Division, Los Alamos National Laboratory, Los Alamos, New Mexico 87545

Complementation group G of xeroderma pigmentosum (XPG) is one of the most rare and pathophysiological heterogeneously inherited disorders. XPG patients exhibit varying phenotypes, from having a very mild defect in DNA repair to being severely affected, and a few cases are also associated with the neurological degeneracy and growth retardation of Cockayne’s syndrome. The XPG gene encodes a 134-kDa nuclear protein that is essential for the incision steps of nucleotide excision repair. XPG protein contains a putative helix-loop-helix (HLH) motif in the region that is most conserved among the members of structure-specific endonuclease family. To establish the functional significance of the HLH motif, we used several approaches, including theoretical modeling, functional complementation assay, structure-specific endonuclease assay, and DNA binding assay. A secondary structure of the motif was predicted by energy minimization and the Monte Carlo simulation and empirically proven using the circular dichroism to contain a high content of α-helix. When an XPG mutant lacking the HLH was overexpressed in UV135 cells, which have defects in the hamster homolog of XPG, the mutant gene failed to confer to the hamster cells the resistance to UV light. A recombinant XPG protein lacking the HLH motif was purified from insect cells and tested for a structure-specific endonuclease activity. The mutant protein failed to cleave the flap strand. A recombinant peptide containing the HLH (amino acids 758–871) was expressed in and purified from bacteria, tested for DNA binding activity, and found to bind to a DNA substrate with the flap structure. These results suggest that the HLH motif is required for the catalytic and DNA binding activities of XPG.

The genome of living organisms is constantly exposed to both external and internal sources of physically and chemically damaging agents. The inherent stability of the DNA molecule provides the first level of protection from such attacks. Moreover, organisms have developed intricate machineries of DNA repair to correct various types of damages in the genome (1). These DNA repair machineries are well conserved among bacteria, yeasts, and human (2). During the last 2 decades, investigators have been identifying and characterizing different DNA repair pathways, including base excision, nucleotide ex-}

---

* This work was supported in part by National Institutes of Health Grant CA71630 (to M. S. P.), Los Alamos National Laboratory Directed Research Development Grant XL2E94205 (to M. S. P.), and Department of Energy Grant KP-0302000. 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 1754 solely to indicate this fact.

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

---

1 The abbreviations used are: NER, nucleotide excision repair; XP, xeroderma pigmentosum; XPG, complementation group G of xeroderma pigmentosum; CHO, Chinese hamster ovary; HLH, helix-loop-helix; bHLH, basic helix-loop-helix; aa, amino acids; HPLC, high performance liquid chromatography; TFE, trifluoroethanol; DMS, dimethyl sulfoxide; Tricine, N-[2-hydroxy-1,1-bis(hydroxymethyl)ethyl]glycine.

---

This paper is available on line at http://www.jbc.org

27823
The predicted primary structure of XPG presents three regions that are evolutionarily conserved among members of a newly emerging family, the structure-specific endonuclease (Fig. 1). One of the regions contains a motif that exhibits homology with the second amphipathic helix of a family of regulatory proteins known as helix-loop-helix (HLH) proteins, including MyoD (Fig. 1). The basic helix-loop-helix (bHLH) family of DNA-binding proteins regulate cell type-specific transcription (reviewed in Ref. 28). This family of proteins is distinguished by a conserved dimerization motif (the HLH domain) and an adjacent cluster of basic amino acid residues shown to be responsible for site-specific DNA binding (Refs. 29 and 30 and references therein). The HLH domain is composed of two amphipathic α-helices separated by a loop region of variable length, and this motif also facilitates homodimerization and heterodimerization between members of the HLH protein family (31, 32).

To determine potential functional domains such as catalytic and DNA binding domains of XPG protein, we performed a motif search, made a secondary structure prediction, and reconstructed biochemical and genetic experiments to assimilate the information obtained by theoretical studies. As a first step for this approach, we analyzed the evolutionary conserved domain of XPG and found that this region contains a putative HLH motif. Biophysical characterization was performed to validate the theoretical prediction of the motif, and biochemical and genetic experiments were carried out to establish the functional significance of the HLH motif of XPG both in vivo and in vitro. In this report we present evidences that support the potential role of the HLH domain as a part of catalytic and/or DNA binding domain of XPG. The implication of these observations is discussed in terms of catalytic and DNA binding activities of XPG in specific and structure-specific endonuclease in general.

EXPERIMENTAL PROCEDURES

Cell Culture, DNA Transfection, and UV Survival Experiments

The UV-sensitive CHO mutant UV135 cells were used and cultured as described previously (16). Cells were seeded at a density of 105 cells/cm2 in 6-well dish and transfected with 10 μg of wild-type or mutant DNAs by calcium phosphate method (33). Stable transformants were selected with G418 (250 μg/ml) and expanded for UV survival test. UV survival test was performed as described previously (16).

Plasmid Constructs

The pBSK-XPGC was made by removing 5′-untranslated regions and adding an adenine residue at −3 position to introduce the Kozak's consensus sequence (34). The original cDNA of XPG has a cytosine at this position (16). The modified cDNA was transferred to pBKCMV digested with SphI/XhoI for expression in UV135 cells. The mutant XPG (XPG-DHLH) was created from the pBSK-XPG by removing the region (aa 771–799) containing the amphipathic helix by using ExSite mutagenesis kit (Stratagene, CA). The mutagenic primers were: P1, 5′-CAG GTA CAT CTG TCC GGT GAC AGT 3′ and P2, 5′-CTG ACT GAT CAG ACA ACC 3′. Once the presence of deletion was confirmed by nucleotide sequencing, the insert was transferred to either pCMV-XPG- or pVL-XPG-DHLH.

To test DNA binding activity of the XPG regions that contain the HLH motif, a bacterial expression construct was made in pET-28 vector (Novagen). The inserts coding for the peptides (aa 747–804 and 758–871) were cloned into pET28 digested with NcoI and XhoI. These constructs were designated as pET-XPG(747–804) and pET-XPG(758–871).

Protein Expression and Purification

Expression and purification of wild-type XPG have been described previously (27). The mutant transfer construct pVL-XPGd(771–799) was transfected into Sf9 cells, and the recombinant virus was obtained using the procedures used for wild-type XPG recombinant virus. Protein was expressed and purified from insect cells as described (27). *Escherichia coli* BL21(DE3) (35) was transformed with pET-XPG(747–804) or pET-XPG(758–871) and cultured in 1 liter of LB medium. Isopropyl-β-D-thiogalactopyranoside was added to 0.1 mM final concentration to a culture of A600 = 0.6. Cultures were further incubated for another 3 h. Cells were harvested and washed twice with a wash buffer (100 mM NaCl, 1 mM EDTA, 50 mM Tris-HCl, pH 8.0), resuspended in 100 ml of a column binding buffer, and sonicated. The sample was centrifuged at 100,000 × g for 1 h, and the supernatant was filtered through 0.45-μm filter before loading onto a metal chelate column and purified as suggested by the manufacturer (Novagen).

Reverse Phase Column Chromatography

The purified recombinant peptide (aa 747–804) from the metal chelate column was injected on to a reverse phase high performance liquid chromatography (HPLC) column (Sem-Prep CN, 7.8 × 300 mm, Waters) and run at the flow rate of 4.0 ml/min. Solvent compositions were: A, 0.2% trifluoroacetic acid, 100% acetonitrile; B, 0.2% trifluoroacetic acid, 50% acetonitrile; and C, 0.2% trifluoroacetic acid, H2O. After equilibrating the column with a 100% solvent C isocratic run-in for 5 min, the sample was pumped onto the column, and the gradient was started immediately. A linear gradient was applied from 20 to 50% acetonitrile at a rate of 10%/h. Fractions containing either the HLH monomer or dimer peak were collected separately, lyophilized, and stored at −20 °C under argon gas.

Peptide Synthesis

Two peptides (17-mer, FGIPYIQAPMEAQAQ and 36-mer, ATVT-GQMFLESSL禄FLGFIPYIQAAPMEAQAQ) that contained the putative HLH motif were chemically synthesized on a Milligen system by a solid phase chemistry. The synthetic peptides were further purified by a reverse phase HPLC, and the purity was confirmed by mass spectrometry analysis.

Circular Dichroism, Mass Spectrometry, and Amino Acid Composition Analysis

The amount of helical content was determined by circular dichroism spectroscopy (36). The purified HLH peptide was dissolved in varying concentrations of triffluoroethanol (TFE) and subjected to CD analysis in a JASCO model J-710. Molecular weights of monomer and dimer fractions was determined by gel desorption/ionization time-of-flight mass spectrometry on a VESTEC Laser Research Tec. model equipped with the VESTEC 201 electrospay unit. The amino acid composition of the monomer and dimer peak of a recombinant peptide (aa 747–804) containing the HLH motif was determined by using the Pico-Tag amino acid analysis system (Waters) following the manufacturer's suggestion.

Chemical Cross-linking

The bifunctional cross-linking reagent DMS (dimethyl suberimidate, Pierce) was used for protein cross-linking studies. DMS was always prepared immediately before use by dissolving 30 mg of DMS in 100 μl of ice-cold triethanolamine-HCl (0.15 M, pH 8.2). The pH of the DMS solution was readjusted to 8.2 with 1 M NaOH. Cross-linking was initiated by adding DMS (10 mg/ml) to the XPG 36-mer peptide (10 μm) that was preincubated in 50 mM triethanolamine buffer, pH 8.2, that contained 10 mM MgCl2, 100 mM NaCl, and 1 mM dithiothreitol at 22 °C. The cross-linking reactions were stopped after intervals ranging from 2 to 20 min by adding an equal volume of 1 M glycine to aliquots of the reactions. The quenched samples were mixed with SDS sample buffer (125 mM Tris-HCl, pH 6.8, 4% SDS, 10% mercaptoethanol, 20% glycerol, and 0.002% bromphenol blue), heated at 90 °C for 2 min, and subjected to Tricine-SDS-polyacrylamide gel electrophoresis using 10–20% gradient gels. Proteins were visualized by Coomassie Brilliant Blue R-250 staining.

Assays

Flap-endonuclease assays were performed as described previously (27, 37). Protein or peptide samples were analyzed by Tricine-SDS-polyacrylamide gel electrophoresis (38). DNA binding assay was performed as described previously (39).

Prediction of Secondary Structures and Theoretical Modeling

The secondary structural elements and energy-minimized structures of the HLH motif of XPG protein were predicted by the following procedures.

**Step 1 (Prediction of Secondary Structures for the 36-Amino Acid-...**
long Peptide Containing the Putative Helix-Loop-Helix Motif—The secondary structural elements are predicted by computing the probability $S$ of a given residue $i$ in this sequence to adopt a $k$ type of conformation ($k$ = helix (h), $\beta$-sheet (b), coil (c), or turn (t)), where

$$S(k, i) = \sum_{j} P(k, i + 1) \frac{1}{1 + 1}$$

(Eq. 1)

(where $\gamma$ = size of the window chosen to account for the effect of the neighboring amino acid residues: $\gamma = 5$ for $h$; $\gamma = 3$ for $b$; and $\gamma = 4$ for $c$ or $t$). $P(k, i)$ = potential for the $k$ type of conformation of individual residue $i$ derived from the analysis of the single crystal structures of about 65 proteins. The highest $S(k, i)$ determines the conformation $k$ for the $i$ residue. This methodology is adopted from Delage and Roux (40).

Use of any existing algorithm for secondary structure prediction is only 60% accurate. To improve accuracy, we performed step 2.

Step 2 (Generation of an Energy-minimized Helix-Loop-Helix Model)—This step involves obtaining an energetically stable helix-loop-helix motif given the secondary structural states of the constituent amino acid residues as obtained after step 1. Appropriate ranges of ($\psi$, $\phi$) values are assigned to all amino acids. For example, $\psi$ = $-55^\circ$ $\pm$ $25^\circ$, $\phi$ = $-55^\circ$ $\pm$ $25^\circ$ for residues in a helix; $\psi$ = $-140^\circ$ $\pm$ $30^\circ$, $\phi$ = $140^\circ$ $\pm$ $30^\circ$ for residues in a $\beta$-strand; $\psi$ = $1 + 1 = -65^\circ$, $\phi$ = $1 + 1 = -50^\circ$, $\psi$ = $1 + 2 = -90^\circ$, $\phi$ = $1 + 2 = 0^\circ$ for residues in a type I turn; $\psi$ = $1 + 1 = 20^\circ$, $\phi$ = $1 + 1 = 150^\circ$ $\pm$ $20^\circ$, $\psi$ = $1 + 1 = 0^\circ$ for residues in a type II turn; ($\psi$, $\phi$) of residues in the coil state are set free to choose any point in the allowed space (for definitions of different secondary structures and corresponding ($\psi$, $\phi$) values; see Ref. 41). We simplified the sequence by assuming Ala for residues with side chains extending beyond the C terminus of the protein, except for the Pro residues and the Gly residues. Our rationale for doing this is that the allowed ($\psi$, $\phi$) space of residues with a side chain longer than Ala is only a subspace of that allowed for Ala. Finally, appropriate side chains are attached to generate an actual helix-loop-helix sequence, and the potential energy of the system is minimized in the ($\psi$, $\phi$, $\omega$, $\chi$) space using the force field of Scheraga and co-workers (42). Several energy-minimized structures of the helix-loop-helix sequence are obtained by choosing a number of starting structures within the specified ranges of ($\psi$, $\phi$) values predicted in step 1. If the secondary structural states of one or more residues, as predicted in step 1, are energetically unfavorable for the HLH motif, those states are altered in step 2. Wrongly predicted secondary structural states of a residue by step 1 are corrected in this step using the energy criteria.

Step 3 (Simulated Annealing to Obtain an Ensemble of Low Energy Structures of the Helix-Loop-Helix Motif)—The simulated annealing is performed in the following manner (43–45). First, a starting energy-minimized structure is chosen and Monte Carlo simulations are performed for 50,000 steps at 100 K in the ($\psi$, $\phi$, $\omega$, $\chi$) space; the last accepted configuration is stored to be subsequently used as a starting configuration in the next lower temperature cycle. Second, 50,000 Monte Carlo steps are repeated in several cycles of gradually decreasing temperature until a temperature of 100 K is reached. Third, the lowest energy configuration at 100 K is further energy-minimized to a low energy gradient (46). This is the “temperature quenching” step in which thermally excited single bond rotations around the equilibrium position are quenched. Finally, the first through third steps are repeated for 20 different starting configurations.

RESULTS

Secondary Structure Prediction and Theoretical Modeling Reveals a Putative HLH Motif in the Most Conserved Region among the Members of Structure-specific Nucleases—Using the BLAST search and alignment algorithms, we identified a block of XPG sequence that shares high homology with other members of the structure-specific endonuclease family (Fig. 1). Energy minimization approaches predicted a strong propensity for this region to form a HLH structure (Fig. 2A and see below). We confirmed this prediction by using the X-ray crystal structures of known HLH proteins, the first helix of XPG was similar with the second amphipathic helix II of HLH proteins, including MysD (30). The positions of hydrophobic residues were identical, and the amino acid residues were highly conserved in terms of hydrophobicity (Fig. 1). The sampled low energy structures of the helix-loop motif show the following conformational features, i.e.
the evolutionarily conserved region, we hypothesized that this region is important for either catalysis or DNA binding. Before we extended our study along this line, we first established the functional significance of this region both in vivo and in vitro.

To test the significance of the HLH motif in vivo, a mutant cDNA construct lacking the HLH motif (pCMV-XPG-DHLH) was introduced into CHO UV135 cells to assess its ability to complement genetic defects of UV135, which has mutation(s) in the hamster homolog of XPG. As a control we also transfected UV135 cells with a construct containing the wild-type XPG (pCMV-XPG). The UV135 cells transfected with the pCMV-XPG construct became resistant to 75% of the parental AA8 cells at a dose of 8 J/m². In contrast, the UV135 cells transfected with the mutant construct were still sensitive to UV lights (Fig. 4A).

The functional significance of this motif was further examined in vitro by measuring the structure-specific endonuclease activity of the mutant protein lacking the HLH motif from XPG protein, which was shown previously to be able to cleave a flap structure (27). The protein was overexpressed in insect cells, purified to apparent homogeneity (Fig. 4C), and tested for its ability to cleave the flap substrate. Since the level of expression and the chromatographic behavior of the mutant protein were similar as those of the wild-type XPG protein, we assumed that the mutant protein was folded properly. The fact that the mutant XPG proteins still retained DNA binding activity further supports the notion of proper folding of these XPG mutant proteins (Fig. 6B). The wild-type XPG cleaved the branch structure and released the 21-base-long fragments. However, the HLH mutant protein failed to cleave the flap structure (Fig. 4B).

Dimerization of HLH of XPG in Vitro—Based upon the amphipathic nature of the peptide, we predicted that the putative HLH motif can form a homodimeric structure through intermolecular interactions. To test this notion, the homodimeric interactions of a recombinant peptide containing the HLH motif was characterized in vitro using the reverse phase HPLC. Purified recombinant peptide (aa 747–804) that contains the HLH motif was fractionated as two distinctive peaks on a reverse phase column (Fig. 5A). The presence of monomeric and dimeric species were evident on a reverse phase HPLC column. Two peaks with equal stoichiometry were present on the purification profile. When these peaks were analyzed by amino acid composition analysis and mass spectrophotometry, it was evident that these peaks had identical amino acid composition, but the first peak was twice the size of the second peak's molecular mass (data not shown).

To remove the influence of other amino acid residues outside of the HLH motif on the homodimerization characteristics of the recombinant peptide (aa 747–804), we chemically synthesized peptides with varying lengths, 17-mer (FGIPYIQAPMEAEAQCA) and 36-mer (ATVTGQMFLESQELLRLFGIPYIQAPMEAEAQCAIL), and tested for their ability to form homodimers using a chemical cross-linker. The synthetic peptide (36-mer) containing the HLH motif formed a dimer in the presence of a chemical cross-linker, DMS. The peptide dimerized in an incubation time (Fig. 5B)- and peptide concentration-dependent manner (data not shown). Similar results were also obtained with the other synthetic peptide (17-mer) and the recombinant peptide (aa 747–804) (data not shown).

DNA Binding Activity of a Recombinant Peptide That Contains HLH Motif—Since the HLH motifs in some of the HLH proteins provide DNA binding activity besides the inter/intramolecular interactions, DNA binding activity of the HLH motif of XPG was tested using electrophoretic mobility shift assay. The XPG mutant protein, which lacks the HLH motif, was tested for its DNA binding ability to the flap substrate. Compared with the DNA binding ability of wild-type XPG
protein, XPG-ΔHLH mutant protein exhibited reduced DNA binding activity (Fig. 6B). To further demonstrate the HLH-mediated DNA binding activity, we tested DNA binding ability of the various peptides containing the HLH motif. These peptides were 1) a synthetic peptide, 17-mer (FGIPYIQAPAEMAQCA), 2) a synthetic peptide, 36-mer (ATVTGQMFLESQELLRLGIPYIQAPMEAEA-QCA), 3) a recombinant peptide (aa 747–804), and 4) a recombinant peptide (aa 758–871) containing the HLH motif (6 μg, lane 3). The synthetic 17- and 36-mer peptides, and the recombinant peptide (aa 747–804), did not exhibit the DNA binding activity to a flap substrate (data not shown), but the other recombinant peptide (aa 758–871) bound to the flap substrate (Fig. 6C).

**DISCUSSION**

Here we present evidence that an evolutionarily conserved region in XPG protein forms a putative HLH motif. We further demonstrate that this region is important for nucleotide excision repair in vivo and structure-specific DNA endonuclease activity, in vitro. This motif appears to be required for catalytic and DNA binding activities of XPG protein.

**Amphipathic Nature of the Putative HLH Motif**—Our secondary structure prediction was made by the energy minimization on the evolutionarily conserved region in eight different proteins that belong to the family of structure-specific endonuclease. Although they have a diversity in their primary amino acid sequences, their predicted secondary structures were similar in terms of the propensity for forming an HLH structure. Since the theoretical prediction by energy minimization analysis suggested the potential of this region to form a putative helix-loop-helix structure, we wanted to establish the propensity of this region to form an α-helix by measuring the helical content of the region. Our CD spectroscopy analysis clearly indicates that this region has a strong propensity to form a helical structure. More direct evidence for the presence of HLH motif will become available when the three-dimensional structures of XPG or the region containing the putative HLH become solved by either x-ray crystallography and/or multidimensional NMR spectroscopy.

The HLH motif is present in many sequence-specific transcription factors. A hallmark feature of the HLH motif is the amphipathic nature of helices. This amphipathy is the key for both protein-DNA and protein-protein interactions of HLH proteins. Although the HLH of XPG is not a typical HLH motif found in sequence-specific transcription factors, it exhibits a high degree of conservation in terms of amphipathy. Actually it shares conserved hydrophobic amino acid residues and space relationships that are almost identical with the second amphipathic helix of other known typical HLHs (30). The amphipathic nature of the XPG HLH was experimentally supported by the self-dimerization of the peptides containing the putative HLH motif (Fig. 5). It is well documented that the amphipathic helices provide homodimerization and heterodimerization of the HLH proteins (29–31). We originally hypothesized that this amphipathic helices could mediate the dimerization of XPG protein. However, so far we have not obtained any evidence for dimeric nature of XPG. Thus, the self-dimerization of the putative HLH peptide appears to be inherent to the characteristics of the peptide itself.

**Functional Significance of the HLH Motif**—Because of the high degree of sequence conservation and the apparent amphipathic nature of the putative HLH motif, we decided to evaluate the functional significance of the HLH motif of the XPG protein. When the HLH motif was removed from XPG protein, the mutant protein was not able to provide a UV resistance to mutant hamster cells, or it lost a flap structure-specific DNA endonuclease activity. These results indicate that this particular region is required for proper function of XPG protein both in vitro and in vivo. Based on these results, we postulate that this motif and/or surrounding region of XPG is important for catalytic activity of the XPG nuclease. This notion is supported by the fact that this motif and the surrounding region contain several amino acid residues that are conserved in the catalytic domains of many nucleases (Ref. 51 and references therein). Particularly, an acid amino acid residue that is equivalent to the glutamic acid 789 of XPG is present in the catalytic cores of various nucleases (51). The recently solved x-ray crystal structures of some of these nucleases indicate that this residue is one of the key amino acids that coordinate Mg$^{2+}$ ion(s), which is essential for cleavage activity (51). We have recently demonstrated that an amino acid substitution of the glutamic acid 181 of human FEN-1 nuclease, which is equivalent to Glu$^{789}$ of XPG, completely abolished catalytic activity of the enzyme (52). Moreover, one of the amino acid residues
FIG. 5. Homodimerization of a recombinant peptide (aa 747–804) that contains the putative HLH motif of XPG. A, chromatogram of reverse phase HPLC analysis of the purified peptide. Peaks: first peak, dimer; second peak, monomer. B, time-dependent chemical cross-linking of the purified peptide.

FIG. 6. Role of the putative HLH motif in DNA binding activity of XPG. A, DNA binding activity of the wild-type XPG protein in various concentration (lanes 1–4: 0, 50, 250, and 500 ng, respectively). B, DNA binding activity of the mutant XPG protein (XPG-dHLH) in various concentration (lanes 1–4: 0, 50, 250, and 500 ng, respectively). C, DNA binding activity of a recombinant peptide (aa 758–871 of XPG) that contains the putative HLH motif in various concentration (lanes 1–4: 0, 100, 250, and 500 ng, respectively). Flap DNA substrate was used for the DNA binding assay. The positions of free and complexed labeled substrates are shown with arrowheads.

(alanine 792) in this particular region was found to be mutated to valine in one of the XPG patients (XP125LO) (19).

DNA binding could be another potential function of the putative HLH motif of XPG. It is possible that the HLH motif is at the center of the active site where DNA substrates make a contact. This hypothesis gains supports by the DNA binding activity of the recombinant peptides (aa 758–871) and by the reduced DNA binding activity of a XPG mutant that lacks the HLH motif. The HLH itself was not sufficient for DNA binding activity, because the short peptides that contain a part or whole HLH did not exhibit DNA binding activity. This suggests that the HLH motif is required but not sufficient for proper DNA binding activity of XPG, and there might be other regions in XPG that are required to form a catalytically relevant DNA binding domain(s) in XPG.

Other indirect, but corroborating, evidence that supports the hypothesis of the involvement of HLH in DNA binding of XPG can be derived from our observation of the unusual DNA binding characteristics of a point mutant D181A of FEN-1. As mentioned earlier, the glutamic acid 181 of FEN-1 is equivalent to glutamic acid 789 of XPG. When the DNA binding activity of the D181A mutant was compared with that of the wild-type FEN-1, the protein binds to the flap substrate, but exhibits an abnormal mobility. Our DNA footprinting data show that the mutant D181A makes a larger footprint than the wild-type FEN-1, suggesting the involvement of the Asp181 in DNA binding.2 Taken together, the putative HLH motif appears to contribute to the DNA binding activity of XPG. A more definitive answer to this hypothesis should be possible with a more rigorous definition of the DNA binding domain of XPG using a comprehensive mutagenesis experiment.

In summary, using theoretical modeling and biochemical, biophysical, and genetic approaches, we established the functional significance of a putative HLH motif that is located in the most conserved region among members of the structure-specific endonucleases. The basic understanding of its role in catalytic and DNA binding activities of XPG will be useful to understand the potential functions of this motif in other members of this enzyme family.

Acknowledgments—We thank Dr. Mark MacInnes for the initial HLH motif search and Dr. Goutam Gupta for theoretical modeling of the putative HLH motif in XPG.

REFERENCES

1. Friedberg, E. C., Walker, G. C., and Siede, W. (1995) DNA Repair and Mutagenesis. American Society for Microbiology, Washington, D. C.
2. Aboussekhra, A., and Wood, R. D. (1994) Curr. Opin. Genet. Dev. 4, 212–220
3. Mu, D., Park, C.-H., Matsumaga, T., Hsu, D. S., Reardon, J. T., and Sancar, A. (1995) J. Biol. Chem. 270, 2415–2418
4. Aboussekhra, A., Biggerstaff, M., Shiivi, M. K. K., Vilpo, J. A., Moncollin, V., Podust, V. N., Proti, M. M., Hulsbecher, U., Egly, J.-M., and Wood, R. D. (1995) Cell 80, 859–868
5. Fischel, R., Lescoe, M. K., Rae, M. R. S., Copeland, N. G., Jenkins, N. A., Garber, J., Kane, M., and Kolodner, R. (1993) Cell 75, 1027–1038
6. Brunner, C. E., Baker, S. M., Morrison, P. T., Warren, G., Smith, L. G., Lescoe, M. K., Kane, M., Esarabino, C., Lipford, J., Lindholm, A., Tannergard, P., Bollard, R., Godwin, A. R., Ward, D. C., Nordenskiold, M., Fischel, R., Kolodner, R., and Liskay, R. (1994) Nature 368, 258–260
7. Wood, R. D. (1995) Philos. Trans. R. Soc. Lond. B. Biol. Sci. 347, 69–74
8. Sancar, A., and Tago, M. S. (1990) Photochem. Photobiol. 57, 905–921
9. Cleaver, J. E., and Kraemer, K. H. (1989) In The Metabolic Basis of Inherited Disease (Scriber, C. R., Beaudet, A. L., Sly, W. S., and Valle, D., eds) Sixth Ed., pp. 2849–2971, MaGraw-Hill, New York
10. Bootsma, D. (1995) Phìlos Trans R. Soc. Lond. B. Biol. Sci. 347, 75–81
11. Vermeulen, W., Osseweijer, P., de Jonge, A. J. R., and Hoeijmakers, J. H. J. (1986) Mutat. Res. 165, 199–206
12. Norris, P. G., Hawk, J. L., Avery, J. A., and Giannelli, F. (1987) Br. J. Dermatol. 116, 861–866
13. Vermeulen, W., Jaeken, J., Janpier, N. G., Bootsma, D., and Hoeijmakers, J. H. J. (1986) Am. J. Hum. Genet. 39, 145–152
14. Busch, D., Grenier, C. E., Lewis, K., Ford, R., Adair, G., and Thomson, L. H. (1989) Mutagenesis 4, 359–354
15. Mudgett, J. S., and MacInnes, M. A. (1990) Genomics 6, 623–633
16. MacInnes, M. A., Dickson, J. A., Hernandez, R. H., Learmount, D., Lin, G., Mudgett, J. S., Park, M. S., Schafer, S., Reynolds, R. J., Srinvest, G. F., and Yu, J. Y. (1993) Mol. Cell. Biol. 13, 6393–6402
17. Sancher, D., Naupskel, T., Curlett, U., Cada, C., Biure, A., and Clarkson, S. G. (1993) Nature 363, 182–185
18. O’Donovan, A., and Wood, R. D. (1993) Nature 363, 185–188
19. Naupskel, T., and Clarkson, S. G. (1994) Hum. Mol. Genet. 3, 963–967
20. Madura, R., and Prakash, S. (1986) J. Bacteriol. 166, 914–923
21. Shimizu, T., Harada, Y., Saito, T., Shiomi, N., Okuno, Y., and Yamaizumi, M. (1994) Mutat. Res. 314, 167–175
22. Harada, Y. N., Matsuoka, Y., Shiomi, N., and Shiomi, T. (1995) Genomics 28, 69–65
23. O’Donovan, A., Sancher, D., Clarkson, S. G., and Wood, R. D. (1994) J. Biol. Chem. 269, 15965–15968
24. O’Donovan, A., Davies, A. A., Moog, J. G., West, S. C., and Wood, R. D. (1994b) Nature 371, 432–435
25. Habraken, Y., Seng, P., Prakash, L., and Prakash, S. (1994) J. Biol. Chem. 269, 31542–31545

8 C. J., Barnes, R. A., Bambara, and M. S. Park, unpublished observation.
Helix-Loop-Helix Motif in XPG

27829

26. Habraken, Y., Sung, P., Prakash, L., and Prakash, S. (1994) *Nucleic Acids Res.* **22**, 3312–3316
27. Cloud, K. G., Shen, B., Strniste, G. F., and Park, M. S. (1995) *Mutat. Res. Lett.* 347, 55–60
28. Jones, N. (1990) *Cell* **61**, 9–11
29. Lassar, A. B., Buskin, J. N., Lockshon, D., Davis, R. L., Apone, S., Hauschka, S. D., and Weintraub, H. (1989) *Cell* **58**, 823–831
30. Murre, C., McCraw, P. S., and Baltimore, D. (1989) *Cell* **56**, 777–783
31. Murre, C., McCraw, P. S., Vaessin, H., Caudy, M., Jan, L. Y., Jan, Y. N., Cabrera, C. V., Buskin, J. N., Hauschka, S. D., Lassar, A., Weintraub, H., and Baltimore, D. (1989) *Cell* **58**, 537–544
32. Sun, X.-H., and Baltimore, D. (1991) *Cell* **64**, 459–470
33. Kozak, M. (1991) *J. Biol. Chem.* **266**, 19867–19870
34. Studier, F. W., Rosenberg, A. H., Dunn, J. J., and Dubendorff, J. W. (1990) *Methods Enzymol.* **185**, 60–89
35. Johnson, W. C. (1988) *Annu. Rev. Biophys. Biophys. Chem.* **17**, 145–166
36. Harrington, J. J., and Lieber, M. (1994) *EMBO J.* **13**, 1235–1246
37. Schagger, H., and Jagow, G. V. (1987) *Anal. Biochem.* **166**, 368–379
38. Harrington, J. J., and Lieber, M. R. (1995) *J. Biol. Chem.* **270**, 4503–4508
39. Delage, G. and Roux, B. (1989) *Prediction of Protein Structure and Principles of Protein Conformation*, pp.587–635, Plenum Press, New York
40. Ramachandran, G. N., and Sasisekharan, V. (1968) *Adv. Protein Chem.* **23**, 283–294
41. Silk, M. J., Nemethy, G., and Scheraga, H. A. (1984) *J. Phys. Chem.* **88**, 6231–6239
42. Gupta, G., Nemethy, G., and Scheraga, H. A. (1984) *J. Phys. Chem.* **88**, 6231–6239
43. G. Gupta and G. Myers (1990) *Cinquieme Colloque des Cent Gardes* **99**, 143–149
44. Gupta, G., and Myers, G. (1977) in *Peptides: Design, Synthesis, Biological Activity* (Basava, C., and Anantharamaiah, G. M., eds) pp. 324–354, Birkhauser, New York
45. Kirkpatrick, S., Gelatt, C. D., Jr., and Vecchi, M. P. (1983) *Science* **220**, 671–675
46. Fletcher, R. (1984) *Practical Methods of Optimization*, Vol. 1, John Wiley & Sons, New York
47. Carr, A. M, Sheldrick, K. S., Murray, J. M., Al-Harithy, R., Watts, F. Z., and Lehman, A. R. (1995) *Nucleic Acids Res.* **23**, 1345–1349
48. Jacquier, A., Legrain, P., and Dujon, B. (1992) *Yeast* **8**, 121–132
49. 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
50. Hiraoka, L. R., Harrington, J. J., Gerhard, D. S., Lieber, M. R., and Hsieh, C. (1995) *Genomics* **25**, 220–225
51. Muser, T. C, Nossal, N. G, and Hyde, C. C. (1996) *Cell* **85**, 1101–1112
52. Shen, B., Nolan, P. J., Sklar, L. A., and Park, M. S. (1996) *J. Biol. Chem.* **271**, 9173–9176