Domain Mapping of the DNA Binding, Endonuclease, and ERCC1 Binding Properties of the Human DNA Repair Protein XPF†

Sandra L. McCutchen-Maloney, Cindi A. Giannecchini, Mona H. Hwang, and Michael P. Thelen*

Molecular and Structural Biology Division, Lawrence Livermore National Laboratory, Livermore, California 94550

Received March 15, 1999; Revised Manuscript Received May 20, 1999

ABSTRACT: During nucleotide excision repair, one of the two incisions necessary for removal of a broad spectrum of DNA adducts is made by the human XPF/ERCC1 protein complex. To characterize the biochemical function of XPF, we have expressed and purified the independent 104 kDa recombinant XPF protein from *E. coli* and determined that it is an endonuclease and can bind DNA in the absence of the ERCC1 subunit. Endonuclease activity was also identified in a stable 70 kDa proteolysis fragment of XPF obtained during protein expression, indicating an N-terminal catalytic domain. Sequence homology and secondary structure predictions indicated a second functional domain at the C-terminus of XPF. To investigate the significance of the two predicted domains, a series of XPF deletion fragments spanning the entire protein were designed and examined for DNA binding, endonuclease activity, and ERCC1 subunit binding. Our results indicate that the N-terminal 378 amino acids of XPF are capable of binding and hydrolyzing DNA, while the C-terminal 214 residues are capable of binding specifically to ERCC1. We propose that the N-terminal domain of XPF contributes to the junction-specific endonuclease activity observed during DNA repair and recombination events. In addition, evidence presented here suggests that the C-terminal domain of XPF is responsible for XPF/ERCC1 complex formation. A working model for the XPF protein is presented illustrating the function of XPF in the nucleotide excision pathway and depicting the two functional domains interacting with DNA and ERCC1.

XPF is a member of a family of orthologous proteins with dual functions in nucleotide excision repair (NER)[1] and recombination repair processes. Like its counterpart Rad1 of *Saccharomyces cerevisiae*, XPF is the large subunit of a heterodimer that functions as a DNA structure-specific endonuclease (*1*−*4*). Although much is known about several multiprotein complexes that participate in the NER pathway functioning to remove a broad spectrum of DNA damage (reviewed in refs *5*, *6*), relatively little is understood about the biochemical properties of the individual proteins involved. What is known, however, is that mutations within *XPF* (formerly known as *ERCC4*; refs *7*, *8*), or in any of the five other genes (*XPA*−*XPD, XPG*) essential for this DNA repair pathway, can result in xeroderma pigmentosum, characterized by extreme sensitivity to sunlight, a high incidence of skin cancer (reviewed in refs *9*, *10*), and, in some cases, neurological disorders (*11*, *12*).

The major emphasis in NER research over the past several years has been the in vitro reconstitution of NER using purified proteins and the identification of the complex interactions between the proteins involved (*3*, *13*−*18*). However, a mechanistic understanding of NER and the underlying causes of xeroderma pigmentosum will require detailed biochemical characterization of the individual NER proteins themselves. To achieve this, we initially set out to characterize the XPF protein apart from its complex partner ERCC1 for the basic biochemical functions of DNA binding, endonuclease activity, and ERCC1 binding.

Studies of the *S. cerevisiae* incision endonuclease subunits Rad1 and Rad10 have revealed a close parallel with human XPF and ERCC1 proteins, and sequence homology between XPF, Rad1, and other orthologs is particularly striking. Rad1 and Rad10 proteins form a specific, constitutive complex that is necessary for NER (*19*−*21*), and the Rad1/Rad10 protein interaction has been mapped to C-terminal, hydrophobic regions of both polypeptides (*22*). These observations were the first evidence of a specific interaction between eukaryotic proteins required for the NER pathway. A major insight into the biological role of XPF and ERCC1 came with the demonstration that the Rad1/Rad10 complex is an endonuclease having substrate preference for single-stranded DNA (*23*, *24*). Although most evidence regarding this activity indicates the involvement of both proteins in the complex, Rad1 itself has been characterized as possessing structure-specific endonucleolytic activity (*25*). In conflict with this report but in keeping with the original notion that both proteins are necessary for nuclease action is prior work demonstrating the inability of Rad1 to act separately (*26*).

† We gratefully acknowledge financial support from the National Institutes of Health to M.P.T. (Grant GM52120) with additional funding provided to S.L.M.-M. by an NIH Postdoctoral Fellowship (GM18388) and by an appointment to the Alexander Hollaender Distinguished Postdoctoral Fellowship Program, sponsored by the U.S. Department of Energy, Office of Biological and Environmental Research, and administered by the Oak Ridge Institute for Science and Education. This research was performed under the auspices of the U.S. DOE by LLNL under Contract W-7405-ENG-48.

* Correspondence should be addressed to this author at the Lawrence Livermore National Laboratory, P.O. Box 808, L-452, Livermore, CA 94550. Telephone: (925) 422-6547. Fax: (925) 422-2282. E-mail: mthelen@llnl.gov.

[1] Abbreviations: NER, nucleotide excision repair; XP, xeroderma pigmentosum; ERCC, excision repair cross-complementing; mAb, monoclonal antibody; ORF, open reading frame; BSA, bovine serum albumin; sERCC1, S-tagged ERCC1 protein.

10.1021/bi990591+ CCC: $18.00
Published on Web 06/24/1999

The requirement for both Rad1 and Rad10 to cleave the 3′ single-strand extension at or near DNA duplex/single-strand junctions has also been demonstrated (*27*). Nevertheless, Rad1 is likely to provide catalytic function to the complex, as suggested by a study using monoclonal antibodies (mAbs) to identify N-terminal and central regions of Rad1 that are required for junction-specific cleavage (*28*).

Similar to Rad1 and Rad10, the XPF and ERCC1 proteins (104 and 33 kDa, respectively) have been isolated from mammalian cells as an endonuclease complex (*1*, *15*). We have previously shown that recombinant XPF and ERCC1 interact as a specific complex by coexpressing and purifying the XPF/ERCC1 complex from bacteria and demonstrating endonuclease activity (*29*). In addition, we have presented evidence that the XPF/ERCC1 heterodimer, coexpressed in insect cells, was capable of cleaving one strand of duplex DNA at the 5′ side of a junction with single-stranded DNA (*3*). Other reports including work from our lab have shown that the XPF/ERCC1 complex is necessary for reconstitution of the NER pathway and specifically contributes to the 5′ incision on a synthetic structural analogue of a damaged DNA intermediate (*3*, *14*). Similar to the Rad1/Rad10 complex, XPF/ERCC1 can cleave synthetic substrates that contain 3′, but not 5′, branched DNA structures (*4*). However, several biochemical features of the XPF/ERCC1 endonuclease complex that are relevant to repair and recombination have yet to be investigated. Enzymatic and structural contributions of the individual proteins in the XPF/ERCC1 complex are of particular interest, especially since mutations in *XPF* can lead to DNA repair deficiencies that result in cancer (*2*, *8*, *12*, *30*).

To investigate the characteristics of XPF that contribute to the incision endonuclease activity of NER, we have examined functions that are potentially inherent to the XPF protein itself. These include endonuclease activity, DNA binding, and the protein−protein interaction with ERCC1. Proteolysis within the C-terminus of XPF, observed during protein expression, resulted in a stable, endonucleolytic protein fragment which lead us to explore potential functional domains within the XPF protein. This observed proteolysis within XPF, along with sequence homology to other DNA repair proteins and secondary structure predictions, was used to design polypeptide fragments representing several specific N-terminal and C-terminal regions representing predicted functional domains. The full-length XPF protein and polypeptides were expressed, purified, and analyzed for biochemical function. We demonstrate that endonuclease activity is intrinsic to the XPF protein, and that its catalytic and DNA binding properties reside in an N-terminal domain. During the preparation of this manuscript, deLaat and co-workers reported that XPF and ERCC1 interact through short C-terminal regions of both proteins (*31*). Although this short region of XPF is capable of binding to ERCC1, we report here that the protein/protein interaction spans a larger region within the C-terminus of XPF and that a specific C-terminal domain directs the interaction of XPF with ERCC1.

EXPERIMENTAL PROCEDURES

Construction of Expression Plasmids. XPF and deletion fragments were produced using the polymerase chain reaction (PCR) and subsequently cloned into the bacterial expression vector pET28. Numbering for the primers and polypeptide products throughout refers to the amino acid residues derived from the full-length *XPF(ERCC4)* cDNA (*8*). To produce XPF and fragments N166, N378, N720, 584C, and 667−854, oligodeoxynucleotide primers used in PCR reactions each contained either *Nde*I or *Sal*I sites (underlined), and are listed as follows:12A: 5′ GAG CAT ATG GCG CCG CTG CTG GAG 3′ 166B: 5′ CTC GTC GAC TCA GTC TGT GAA AGC TTT AAT AAA 3′ 378B: 5′ CTC GTC GAC TCA GTT GCT TTC TAG GAC CAG 3′ 702B: 5′ CTC GTC GAC TCA CCG ACG ATG GAT CAG AGA 3′ 584A: 5′ CTC CAT ATG GAG CTA ACC TTT GTT CGG 3′ 667A: 5′ CTC CAT ATG ACT GAC ACT CGG AAA GCC 3′ 854B: 5′ CTC GTC GAC TCA GTT TTT GGC ATT CAC CCC 3′ 916B: 5′ CTC GTC GAC TCA CTT TTT CCC TTT TCC TTT 3′ For each PCR reaction, 10 ng of cDNA template (*XPF* clone cER4-40; *8*) and 10 μmol of each primer were used. The reaction mix contained 1.5 mM $MgCl_2$, 10 mM dNTPs, 0.1 volume of 10× buffer, and 2.5 units of AmpliTaq enzyme (Perkin-Elmer). Samples were overlayed with mineral oil and subjected to 30 cycles of 94 °C 1 min, 55 °C 1 min, and 72 °C 1 min. To amplify the regions of interest (see Figure 2), the following primer pairs were used: XPF (encoding amino acids 12−916), primers 12A and 916B; N166 (12−166), primers 12A and 166B; N378 (12−378), primers 12A and 378B; N702 (12−702), primers 12A and 702B; 584C (584−916), primers 584A and 916B; 667−854, primers 667A and 854B. PCR products were analyzed by gel electrophoresis, purified using a QIAquick column (Qiagen), and subcloned using the *Nde*I and *Sal*I sites of pET28 (Novagen). Each cloned product was sequenced to ensure the integrity of the amplified product. Sequencing was performed by Dr. Yiding Wang at Biotech Core (Palo Alto, CA). N260 and N310 were excised directly from the *XPF* cDNA (cER4-40) using *Nco*I and *Eco*RI (for N260) or *Bam*HI (for N310), and these were inserted into the corresponding sites of the pET expression vector. For the initial in situ activity assays described in Figure 1, XPF cDNA was excised directly from the *XPF* cDNA (cER4-40) using *Nco*I and *Not*I and subcloned into the corresponding sites of the pET32 expression vector. For immobilization of ERCC1 on S-protein agarose (*32*), an expression plasmid was prepared using pET29T (Novagen) that fused the S-peptide tag to the N-terminus of ERCC1. To amplify the *ERCC1* open reading frame (ORF), PCR conditions were the same as those described for *XPF*, using the *ERCC1* cDNA (*33*) as template and oligodeoxynucleotide primers as follows: SE1A: 5′ GAT GGA CCC TGG GAA GGA CAA AGA G 3′ SE1B: 5′ TCA AGA AGG GCT CGT GCA GGA CAT C 3′. The PCR product was processed and subcloned into pET29T directly, according to the manufacturer's instructions (Novagen), and sequenced to confirm that the junctions and insertion site were correct. For the purposes of cloning the ORF in-frame with the vector's N-terminal S-peptide and C-terminal hexahistidine tag sequences, the last two native codons of the *ERCC1* ORF were eliminated. The protein product of the resulting expression plasmid, p29TE1, is referred to throughout as sERCC1.

Expression of XPF and Polypeptides. Full-length XPF and all XPF fragments were expressed with an N-terminal hexahistidine-tag using the pET28 expression system in *E. coli* BL21(DE3) cells. Bacteria were cultured in 2 L of LB

medium containing 30 mg/L kanamycin, overnight at 25 °C with shaking at 80 rpm. When cells reached an OD_{600} of approximately 0.6, IPTG was added to a final concentration of 0.3 mM, and shaking was increased to 225 rpm. After 4 h, cells were harvested by centrifugation, and cell pellets were frozen at −80 °C. For lysis, cells were resuspended in 200 mL of buffer (listed below for each polypeptide) containing 0.5 mg/mL lysozyme and 0.1% Triton X-100. After standing on ice for 30 min, samples were sonicated, and the insoluble fraction was removed by centrifugation (15000*g* for 10 min at 4 °C). Prior to column chromatography, soluble fractions were passed through 0.45 μm filters.

Purification of XPF and Polypeptides. (A) General Methods. All chromatographic steps were carried out at 4 °C using the GradiFrac system (Pharmacia Biotech), run at a flow rate of 5 mL/min. Fractions were analyzed for proteins by SDS−PAGE, in which all overexpressed proteins were detected by Coomassie blue staining and confirmed by immunodetection using XPF fragment-specific mAbs, as described below.

(B) XPF and N702. Bacteria containing the XPF and N702 expression vectors were lysed in buffer H (20 mM Tris-HCl, pH 7.6, 0.5 mM DTT, 0.5 mM EDTA, 100 mM NaCl, 10% glycerol), in preparation for chromatography on heparin−Sepharose (5 mL HiTrap column, Pharmacia Biotech). Both XPF and N702 proteins eluted from separate columns at 600 mM NaCl, using 50 mL steps of 100 mM, 300 mM, 600 mM, and 1 M NaCl in buffer H. Peak fractions were combined and dialyzed against buffer Q (20 mM Tris-HCl, pH 8.5, 1 mM DTT, 100 mM NaCl, 10% glycerol). XPF and N702 samples were loaded onto separate Q-Sepharose columns (5 mL HiTrap column, Pharmacia), and each protein was eluted at approximately 500 mM NaCl in a 150 mL linear gradient from 100 mM to 1 M NaCl in buffer Q. Peak fractions were pooled, dialyzed overnight against two changes (1 L) of buffer D (20 mM Tris-HCl, pH 7.5, 100 mM NaCl, 1 mM DTT, 10% glycerol), and stored at −80 °C. Purified XPF was used in screening for mAbs, as well as in biochemical characterizations described here.

(C) N378, 584C, and 667−854. Bacteria containing the expression vectors for these proteins were lysed in buffer N (20 mM Tris-HCl, pH 7.5, 200 mM NaCl, 5 mM β-mercaptoethanol, 10% glycerol). His-tagged proteins were purified on separate Ni−NTA agarose columns (8 mL, Qiagen), using a 150 mL linear gradient from 0 to 500 mM imidazole in buffer N. Proteins eluting at approximately 250 mM imidazole were then pooled and dialyzed as above for storage at −80 °C.

(D) N166, N260, and N310. Bacteria containing the vectors for expressing these proteins were lysed in buffer N, but the overexpressed proteins were confined solely to the insoluble fraction. Proteins were solubilized in buffer N containing 8 M urea. The His-tagged proteins were then isolated by denaturing chromatography on Ni−NTA agarose using a step gradient of 40, 80, and 250 mM imidazole in buffer N containing 4 M urea. The 250 mM fractions containing purified protein were pooled and dialyzed against buffer D to remove urea and permit renaturation of the proteins. Little or no soluble protein remained after dialysis with these three polypeptides. (However, this resolubilization method proved successful with a fraction containing insoluble, full-length XPF, such that it was useful as antigen for the generation of antibodies in mice.)

sERCC1 Expression and Purification. S-Peptide-tagged ERCC1 (sERCC1) was expressed from p29TE1 in *E. coli* B834(DE3)pLysS. Expression was carried out as described for XPF, with modifications given as follows. Culture medium contained 35 mg/L chloramphenicol and 30 mg/L kanamycin, and expression was induced with 1 mM IPTG. Cells from a 2 L culture were lysed in 100 mL of cold TNE (20 mM Tris-HCl, pH 7.5, 0.5 mM EDTA, 0.5 mM DTT, 100 mM NaCl, 4% glycerol) on ice for 30 min. After filtering the soluble fraction, it was diluted with 100 mL of cold water and loaded onto heparin−Sepharose. Fractions were analyzed for sERCC1 by SDS−PAGE and by Western blot, in which peroxidase-conjugated S-protein was used for detection of sERCC1, as described by the manufacturer (Novagen). Peak fractions containing sERCC1 (eluting between 0.3 and 0.6 M NaCl) were pooled and dialyzed overnight against two changes, 1 L each, of buffer S (20 mM Tris-HCl, pH 7.5, 150 mM NaCl, 0.1% Triton X-100). After dialysis, 3 mL (bed volume) of S-protein agarose (Novagen) was mixed with the sERCC1 sample by gently rocking at room temperature for 30 min, and then packed into a column. The column was washed with 30 mL of buffer S, and sERCC1 was eluted with 2 M guanidine thiocyanate (Aldrich Chemical Co.) in buffer S. The first 3 column volumes containing purified sERCC1 were then dialyzed as described for XPF.

Endonuclease Activity Assay. A standard assay for cleavage of supercoiled plasmid DNA was used to test purified proteins for endonuclease activity. The reaction mix contained pUC19 DNA (1 μg) in 20 mM Tris-HCl, pH 8.0, 20 mM NaCl, 5 mM $MgCl_2$, 1 mM DTT, 5% glycerol, and 50 μg/mL acetylated bovine serum albumin (BSA). Reactions were initiated by the addition of XPF or polypeptide fragments in the range of 10−200 ng, as indicated, to a total volume of 20 μL. Following incubation for 2 h at 37 °C, SDS was added to a final concentration of 0.5% and incubated for an additional 30 min at 37 °C to disrupt protein−DNA complexes. To visualize digestion products, samples were subjected to electrophoresis on a 1% agarose gel, stained for several minutes in 2 μg/mL ethidium bromide, and destained in water. Mung bean nuclease (4 units) was used as a positive control, according to the specifications of the vendor (Gibco BRL). As a blank, buffer D was added in place of the protein sample.

To determine the protein responsible for endonuclease activity in the above assay for full-length XPF, an in situ endonuclease assay was performed as previously described (*34*). Briefly, XPF was separated by SDS gel electrophoresis on a 10% polyacrylamide gel containing 10 μg/mL DNA substrate (pUC19 or single-stranded M13) and 40 μg/mL acetylated BSA. Following electrophoresis, SDS was removed by washing to permit protein renaturation. The gel was then incubated for enzymatic activity for 24 h or more at 37 °C. The gel was stained for 30 min in 2 μg/mL ethidium bromide, destained in water for 1 h, and photographed on a 260 nm transilluminator. To visualize the protein by standard staining methods, a second gel lacking BSA was run in parallel, stained using Coomassie blue, and photographed using a Kodak DC120 digital camera. Bovine pancreatic DNase I (Gibco BRL) was used as a positive control for in situ nucleolytic activity under identical conditions as XPF, with 1 unit resulting in an intense band on a gel containing

pUC19 DNA. Fluorescence photographs for the standard and in situ endonuclease assays were scanned using a flatbed scanner at 600 pixel/inch resolution.

DNA Binding. Denatured DNA−cellulose (Pharmacia Biotech) was used to test the affinity of purified polypeptides for DNA in the absence of Mg^{2+}. DNA−cellulose was washed in 200 μL (bed volume) batches 3 times by centrifugation (10000*g* for 5 min at 4 °C), each with 400 μL of NETN buffer (50 mM Tris-HCl, pH 8.0, 0.25 M NaCl, 1 mM EDTA, 0.5% Nonidet P-40), and finally resuspended by the addition of 150 μL of NETN. Purified XPF or polypeptide fragments (1 μg of each) were added to 50 μL aliquots of washed DNA−cellulose, and the volume was adjusted to 500 μL with NETN. Samples were incubated with gentle rocking for 2 h at room temperature and washed 3 times with 200 μL of NETN. DNA−cellulose samples were mixed with 50 μL of 5× SDS sample buffer, heated at 37 °C for 5 min, and briefly centrifuged, and 25 μL of supernatant was loaded onto an SDS−polyacrylamide gel. Proteins binding to DNA were detected by Western blot, described below.

ERCC1 Binding. sERCC1 immobilized on S-protein agarose was used to test interactions with purified proteins. S-Protein agarose was washed in 200 μL (bed volume) batches 3 times by centrifugation (40*g* for 5 min at 4 °C), each with 800 μL of buffer D, and the agarose was resuspended with 200 μL of buffer D. Purified sERCC1 (1 μg) was added to the agarose to a total volume of 400 μL and incubated with rocking for 1 h at room temperature. The agarose was washed 3 times with NETN and resuspended with 200 μL of buffer D. Purified XPF or polypeptide fragments (1 μg of each) were added to 50 μL of the sERCC1−agarose suspension, and the volume was adjusted to 500 μL with buffer D. Samples were incubated with rocking for 1 h at room temperature, and then collected and washed 3 times with 200 μL of NETN each. Agarose samples were resuspended in 250 μL of SDS sample buffer, heated in a boiling water bath for 2 min, and briefly centrifuged, and 25 μL of supernatant was loaded onto an SDS−polyacrylamide gel. Proteins binding to sERCC1 were detected by Western blot, described below.

Immunoprecipitation. mAbs 27, 13, and 218 used in this study were raised against human XPF and affinity-mapped to regions 12−166, 702−854, and 854−916, respectively.[2] To probe ERCC1/XPF complex formation, protein-A-purified mAb (20 μg) was mixed with purified XPF (6 μg) in a total volume of 400 μL using NETN as diluent, and incubated with gentle rocking for 1 h at 4 °C. sERCC1 (5 μg) was then added, and incubation was continued for 1 h, followed by the addition of 25 μL (bed volume) of protein-A−Sepharose 6MB (Sigma) that had been previously washed in NETN. After 1 h, the suspension was centrifuged (500*g* for 5 min at room temperature), washed 3 times with 0.5 mL of NETN, and finally resuspended in 50 μL of 1× SDS sample buffer. The suspension was heated at 37 °C for 5 min and briefly centrifuged, and 30 μL of supernatant was loaded onto an SDS−polyacrylamide gel. Proteins were separated by electrophoresis and transferred to a nitrocellulose membrane filter. sERCC1 and S-peptide-tagged protein markers were detected using alkaline phosphatase-conjugated S-protein according to the manufacturer (Novagen).

[2] M. Hwang and M.P. Thelen, unpublished.

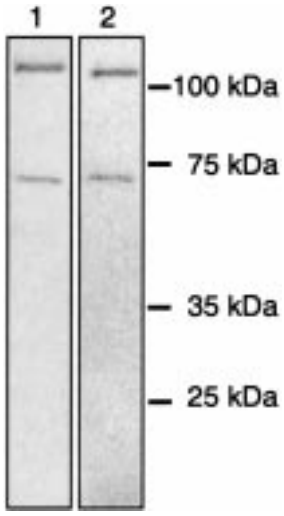

FIGURE 1: Identification of endonuclease activity in XPF and N-terminal 70 kDa proteolysis fragment. Purified XPF separated by SDS gel electrophoresis was subjected to the in situ endonuclease assay using a polyacrylamide gel containing plasmid DNA. Endonuclease activity was detected by darkened bands resulting from degradation of DNA within the background of ethidium bromide fluorescence. A gel run in parallel was stained separately. Lane 1, 1 μg of XPF, assayed for activity. Lane 2, 1 μg of XPF, directly stained for protein detection. Endonuclease activity was observed in XPF and the N-terminal 70 kDa proteolysis fragment of XPF.

Immunodetection. XPF polypeptides transferred to nitrocellulose filters were detected by enhanced chemiluminescence (ECL, Amersham) for Figures 4 and 5 or by indirect staining with alkaline phosphatase for Figure 6. For ECL, peroxidase-conjugated F(ab′) fragment (Amersham) was used at 1:10 000 dilution, and the immunoblot was developed by incubation with a Luminol reagent (Pierce) and then exposed to autoradiographic film. For alkaline phosphatase, filters were incubated with 0.125 μg/mL purified mAbs, washed, and further incubated with a secondary anti-mouse antibody. Alkaline phosphatase-conjugated IgG (BioRad) was used at 1:5000 dilution, and the immunoblot was stained with BCIP/NBT (5-bromo-4-chloro-3-indolyl phosphate *p*-toluidine/nitroblue tetrazolium chloride) as described by the vendor (Gibco BRL).

RESULTS

Domain Prediction and Experimental Design. In our initial characterization of the XPF protein, endonuclease activity was demonstrated in XPF; however, proteolysis of the 104 kDa protein observed during expression yielded what appeared to be a stable 70 kDa protein fragment. Immunodetection using the affinity-mapped XPF mAbs revealed that this fragment retained the N-terminus, indicating that proteolysis had occurred within the C-terminus. As shown in Figure 1, further characterization for endonuclease activity using an in situ nuclease assay revealed that both the 104 kDa protein and the 70 kDa protein fragment were nucleolytically active. This observation suggested a catalytic domain within the N-terminus of the XPF protein and was the starting point for the work presented here. To test our hypothesis that the N-terminus of XPF contained a functional domain,

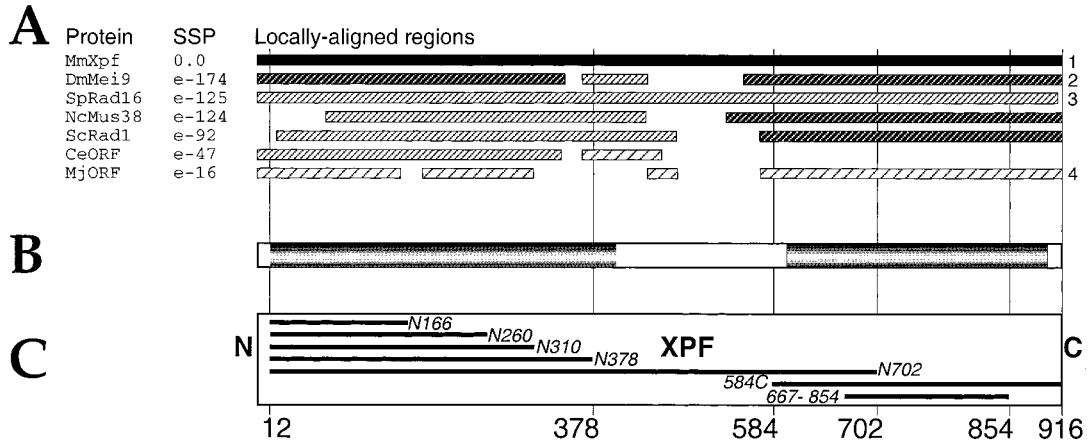

FIGURE 2: Prediction of domains in XPF by sequence homology and secondary structure. (A) Similarity between human XPF (916 residues) and homologues in *Mus musculus* (MmXPF), *Drosophila melanogaster* (DmMei9), *Schizosaccharomyces pombe* (SpRad16), *Neurospora crassa* (NcMus38), *Saccharomyces cerevisiae* (ScRad1), and open reading frames predicted from *Caenorhabditis elegans* (CeORF) and *Methanococcus jannaschii* (MjORF), as determined by Gapped-BLAST (37). The regions most highly conserved with XPF are indicated as shaded boxes, with the level of homology as indicated: 87% (bar 1); 61−69% (bar 2); 53−59% (bar 3); and 44−48% (bar 4). SSP, smallest sum probability. (B) Regions of XPF likely to represent structural domains as predicted by PSI−BLAST comparison of related sequences (37), and by secondary structure algorithms. Information was combined from Chou−Fasman/Robson−Garnier (using MacVector software, Oxford Molecular Group) and PHDsec (36). (C) Design of recombinant XPF polypeptides used in this study based on homology and structure predictions. Numbers at bottom indicate XPF amino acid positions.

Table 1: Summary of XPF and XPF Fragment Purification and Characterization[a]

XPF fragment (kDa)	purification scheme	yield (mg/L)	endonuclease activity	DNA binding	ERCC1 binding
full length (104)	H, Q	<2	+	+	+
N378 (42)	N	2	+	+	−
N702 (80)	H, Q	2	+	+	−
584C (37)	N	5	−	−	+
667−854 (21)	N	>50	−	−	+
N166, N260, N310	N (dn)	nd	nd	nd	nd

[a] All recombinant polypeptides constituted a major fraction of the total bacterial protein after overexpression. Purification was carried out using the chromatographic methods indicated: H, heparin−Sepharose; Q, Q-Sepharose; N, nickel-chelate; dn, denaturing. Yield is given in milligrams of soluble protein purified per liter of bacterial culture. Biochemical function was examined by assays for plasmid DNA cleavage, DNA binding, and ERCC1 affinity by the in vitro assays described. N, fragments beginning at the N-terminus of XPF; C, fragment ending at the C-terminus; 667−854, region near the C-terminus; (+), detected in assay; (−), not detected in assay; nd, not determined due to the lack of soluble proteins.

we examined sequence homology and secondary structure predictions and used this information to predict functional domains within XPF (Figure 2). Homology comparisons were extended beyond the original XPF alignments (*8*) to include the sequences encoded by the *Xpf* cDNA from *Mus musculus*,[3] the *Neurospora crassa* ortholog *mus-38* (*35*), and putative protein sequences from *Caenorhabditis elegans* and *Methanococcus jannaschii* that show significant conservation with this protein family (Figure 2A). From these alignments, it is clear that regions of XPF in the first 400 and the last 300 residues are highly conserved across evolution, whereas sequences in the midsection of these proteins are considerably divergent. Furthermore, the N-terminal 350 amino acid sequence of the XPF protein family was found to be similar to many RNA and DNA helicases, such as eukaryotic initiation factors and RecG-related proteins. Although this homology is not as significant as that found in the orthologous family, it indicates a conserved structural fold or protein domain in the N-terminal region with potential DNA binding properties. Moreover, analysis of the XPF sequence using secondary structure algorithms such as PHDsec (*36*) predicted α helical and *β* strand structures within the conserved regions of XPF, but less defined structure in the nonconserved midsection (Figure 2B). To test whether the conserved N-terminal region contains a functional domain with endonuclease activity, recombinant polypeptides N166, N260, N310, N378, and N702, representing N-terminal fragments of XPF that exclude the highly conserved C-terminal block, were designed (Figure 2C). From homology and structural information, we also targeted C-terminal fragments 584C and 667−854 to examine XPF for a second functional domain capable of binding to ERCC1.

Expression and Purification of XPF and Polypeptide Fragments. For biochemical characterization, the XPF protein, separate from its complex partner ERCC1, was overexpressed in *E. coli* using the commercially available pET28 vector. The seven polypeptide fragments of XPF designed for domain studies were similarly expressed and purified, as summarized in Table 1. Polypeptides N378, N702, 584C, and 667−854 were each overexpressed as a major fraction of the bacterial protein, and were readily purified by conventional ion exchange or nickel affinity chromatography methods (Table 1). Polypeptide 667−854 was the most highly expressed, easily purified, and stable over long-term storage, properties that are consistent with

[3] M.E. Shannon, J. Lamerdin, S.L. McCutchen-Maloney, L. Richardson, M. Hwang, M.A. Handel, L. Stubbs, and M.P. Thelen, manuscript submitted.

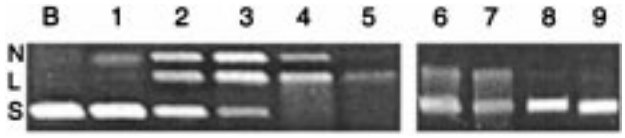

FIGURE 3: Endonuclease activity of XPF and N-terminal polypeptide fragments. Purified XPF (left panel) and polypeptides (right panel) were tested for cleavage of supercoiled plasmid DNA in the standard endonuclease assay. Supercoiled DNA substrate and products of digestion were separated by agarose gel electrophoresis and visualized by ethidium bromide fluorescence. 0.5 μg of DNA was loaded in each lane. B, blank sample containing only BSA. XPF was added at increasing concentrations in lanes 1−5: 10 nM, 25 nM, 50 nM, 100 nM, and 200 nM, respectively. Lanes 6−9 contain the different XPF polypeptides, 100 nM each: N378, N702, 584C, and 667−854, respectively. DNA species indicated at left are S, supercoiled; N, nicked; L, linear.

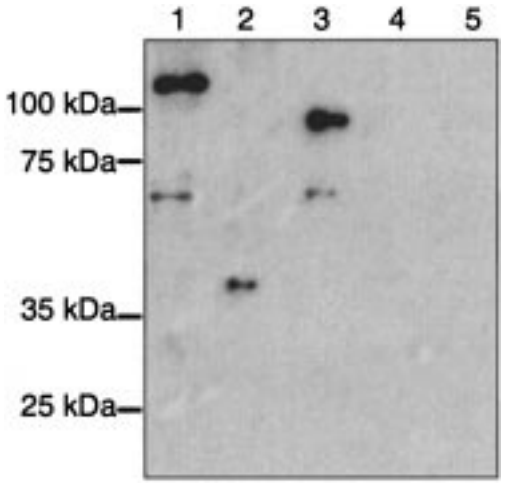

FIGURE 4: DNA binding by XPF and N-terminal polypeptide fragments. Purified proteins were examined for DNA binding using DNA−cellulose for affinity-capture. After washing to remove nonspecifically bound proteins, samples were processed for Western blotting. XPF polypeptides remaining bound to DNA were detected using a combination of mAbs 27 and 13 (see legend in Figure 6A). Lanes 1, XPF; 2, N378; 3, N702; 4, 584C; and 5, 667−854. XPF, N378, and N702 were detected in the bound fraction. Note that for both the XPF and N702 samples, the N-terminal 70 kDa proteolysis fragment of XPF was also detected.

that of an autonomously folding domain. In contrast, overexpression of N-terminal regions containing 310 residues or less resulted in proteins that remained insoluble, even after denaturing in urea and purification by affinity chromatography. The insolubility of N166, N260, and N310 may indicate incomplete structures that were unable to fold properly in solution. For XPF and the polypeptides that were purified by nondenaturing chromatography, biochemical functions were examined using assays for DNA binding, plasmid DNA cleavage, and ERCC1 binding.

N-Terminal Endonuclease and DNA Binding Domain. In the presence of supercoiled plasmid DNA, purified XPF was found to cause nicking and double-strand breaks, and at higher enzyme concentration catalyzed degradation (Figure 3, lanes 1−5). In these standard assays, the activity of XPF on single-stranded DNA appeared to be greater than that for duplex DNA, and the activity was dependent on a divalent cation such as Mg^{2+} (data not shown). A direct comparison of the purified polypeptide fragments in this assay indicated that N378 and N702 possess endonuclease activity, and that this activity was not detectable in 584C and 667−854 (Figure 3, lanes 6−9). Interestingly, N378 and N702 demonstrated a different digestion pattern from that of XPF, and the main product visualized on the gel was linearized DNA. This may reflect the lack of a control element in the C-terminus of XPF which confers part of the duplex/single-strand junction specificity of the XPF/ERCC1 complex.

An essential function that must accompany the incision event catalyzed by XPF during repair is DNA binding. Since it was not obvious that the nucleolytic domain of XPF was sufficient for DNA binding, affinity-capture of the proteins on DNA coupled to cellulose was used to investigate the region of XPF required for DNA binding (Figure 4). This was performed in the absence of Mg^{2+} and in the presence of EDTA in order to uncouple DNA binding and the nuclease activity found in the first 378 amino acids of XPF. After incubating with DNA−cellulose followed by extensive washing, XPF and the N-terminal 70 kDa proteolysis fragment remained specifically bound to DNA, as did fragments N378 and N702 (lanes 1−3). The C-terminal polypeptides 584C and 667−854 were not detected in the bound fraction in the same experiment (lanes 4 and 5), but instead were detected in the unbound fraction (data not shown). Hence, the N-terminal 378 amino acids of XPF are capable of binding to DNA. Results from this DNA binding study and the endonuclease assays indicate that a region within the first 378 amino acids of XPF is sufficient for catalytic activity, including DNA binding, and therefore may constitute a single domain within XPF. This represents the first evidence that XPF contributes to the endonuclease activity of the XPF/ERCC1 complex.

C-Terminal Protein Interaction Domain. The interaction between XPF and ERCC1 proteins was investigated by two different techniques utilizing the S-peptide affinity tag (*32*) fused to the N-terminus of ERCC1 (sERCC1). In the first of these tests, direct interaction of XPF polypeptides with ERCC1 was examined using sERCC1 immobilized on S-protein agarose beads (Figure 5). Since sERCC1 was bound to the S-protein through its N-terminus, presumably most of ERCC1 was still available for interaction with XPF. Following affinity-capture, XPF polypeptides that bound to sERCC1 were identified by immunodetection. Specific and stable binding to sERCC1 was observed with XPF, 584C, and 667−854 polypeptides (lanes 1, 4, and 5); however, N378, N702 (lanes 2 and 3), and the N-terminal 70 kDa fragment of XPF (lane 1) were not detected in the bound fraction. These results indicate that the interaction domain extends well beyond the C-terminal 91 amino acids that was reported previously (*31*). We find that XPF amino acids 667−854 can bind to ERCC1 while the previous report found that amino acids 825−916 could bind to ERCC1. This 29 amino acid overlap between the two reports may indicate the region of XPF possessing the highest affinity for ERCC1.

To further define the ERCC1 binding domain, a second approach to monitor XPF binding to ERCC1 was used in which binding was challenged with mAbs[2] directed against various regions of the XPF protein. The mAbs were selected by their epitope mapping to XPF (Figure 6A). mAb 27 is specific for the N-terminal 166 residues of XPF, mAb 13 is specific for the region between amino acids 702 and 854, and mAb 218 is specific for the C-terminal 62 amino acids of XPF. Each mAb was found to efficiently bind to XPF in this pull-down assay (Figure 6B). Only mAb 27 was able to

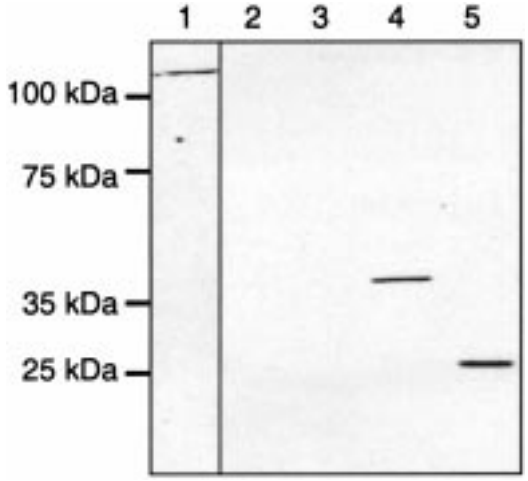

FIGURE 5: ERCC1 binding by XPF and C-terminal polypeptide fragments. Purified proteins were assessed for protein interaction using affinity-capture by Sepharose-bound sERCC1. After washing to remove unbound polypeptides, samples were processed for Western blotting, and XPF polypeptides were detected using a combination of mAbs 27 and 13 (see legend in Figure 6A). Lanes 1, XPF; 2, N378; 3, N702; 4, 584C; and 5, 667−854. Only XPF and the two C-terminal polypeptides were detected in the bound fraction.

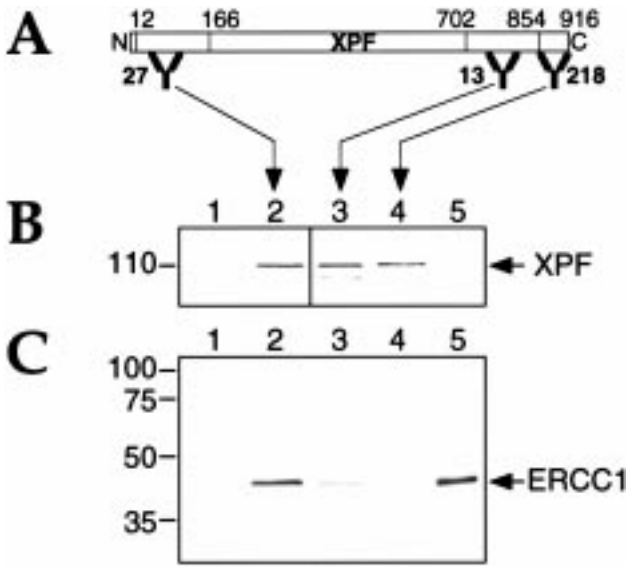

FIGURE 6: Identification of the C-terminal protein interaction region of XPF. (A) Schematic of XPF indicating binding sites of individual mAbs used in immunoprecipitation. mAbs binding to specific regions of XPF are numbered as follows: 27, affinity for amino acids 12−166; 13, affinity for amino acids 702−854; 218, affinity for amino acids 854−916. **(**B and C) XPF/ERCC1 protein complex formation was assessed by adding sERCC1 to the XPF/mAb mixture, and the resulting complex was captured with Protein-A Sepharose. Samples were then processed on separate Western blots for immunodetection (B, XPF) and S-tag detection (C, sERCC1). Lanes 1−4 were immunoprecipitated with the following: 1, an unrelated mAb (specific for the XRCC1 protein, used here as a negative control); 2, mAb 27; 3, mAb 13; 4, mAb 218. As a reference, sERCC1 without the other components was loaded in lane 5.

capture the intact XPF/sERCC1 complex (Figure 6C, lane 2). This result was anticipated in light of the fact that polypeptides 667−854 and 584C can bind to ERCC1 while N378 and N702 are not capable of interacting with ERCC1. Since mAb 27 binds to the N-terminal 166 amino acids, we reasoned that it should not affect the protein/protein interaction at the C-terminus. However, a 4-fold reduction in ERCC1 binding was observed by immunoprecipitation with mAb 13 (Figure 6C, lane 3), and complex formation was completely inhibited when the C-terminal 62 amino acids were bound by mAb 218 (Figure 6C, lane 4). This, together with the results from affinity-capture of C-terminal XPF polypeptides by sERCC1, provides evidence that a domain within residues 702 and 916 directs the interaction of XPF with ERCC1.

DISCUSSION

The focus of this study was to determine the biochemical contribution of XPF functional domains to the NER 5′ incision endonuclease complex. Given that XPF has not previously been characterized as an isolated protein separate from its association with ERCC1, we initially investigated endonuclease activity in the XPF protein. Subsequently, we identified potential domains within XPF based on our finding that a 70 kDa N-terminal proteolysis fragment of XPF generated during overexpression retains endonuclease activity (Figure 1). Homology of XPF to yeast, fly, and mouse NER proteins as well as secondary structural predictions for XPF were also useful in predicting functional domains (see Figure 2). Our approach to designing deletion fragments of XPF to test for functional domains was further directed by a previous study describing inhibition of *S. cerevisiae* Rad1/Rad10 endonuclease and NER activity by mAbs specific for N-terminal and central regions of Rad1 (*28*). We present here the expression, purification, and characterization of several XPF deletion fragments which were used to determine the existence of functional domains within XPF (Table 1).

Endonuclease activity, similar to that of other Mg^{2+}-dependent single-strand-specific nucleases, was identified as an intrinsic function of XPF. We found that XPF was capable of nicking supercoiled DNA presumably due to the partial single-stranded nature of supercoiled duplex DNA (Figure 3). Once nicked, the substrate was linearized and subsequently degraded at longer incubation times or at higher XPF concentrations. This activity is analogous to that previously reported for the Rad1/Rad10 complex; however, neither Rad1 nor Rad10 possessed activity alone (*39*). Our results represent the first evidence that XPF alone functions as an endonuclease in the absence of its protein partner ERCC1, although ERCC1 likely modulates the specificity required for incision during the repair processes as predicted by a model illustrating the function of XPF in NER (Figure 7).

The ability to make double-strand breaks in plasmid DNA was also observed in the N378 and N702 N-terminal fragments. Endonuclease activity was not detected in the C-terminal fragments of XPF, and this lack of activity provided a negative control for similar nuclease activity commonly seen as a contaminant from bacterial overexpression. Correspondingly, DNA binding was detected for XPF and the two N-terminal polypeptides, N378 and N702 (Figure 4). The C-terminal XPF fragments showed no affinity for DNA. Other XPF fragments that were truncated short of the highly conserved N-terminal region were highly insoluble and apparently incapable of folding properly after overexpression, suggesting an incomplete domain. Thus, straightforward biochemical characterization for endonuclease ac-

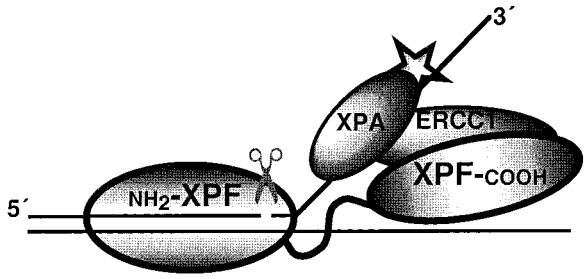

FIGURE 7: Model depicting the function of XPF in the NER pathway. The two functional domains of the 916 amino acid XPF protein are illustrated. The N-terminus of XPF is shown interacting with DNA. This interaction is indicated near the double-strand/single-strand junction that is located 5′ to the DNA damage (represented by the star). Incision near the junction is known to occur during NER and is denoted by scissors and a break in the damaged DNA strand. The C-terminus of XPF is shown binding to ERCC1 which in turn interacts with the XPA protein that has previously been reported to have affinity for ERCC1 and damaged DNA (40, 41). The open bubble structure of DNA in complex with XPF and other proteins involved in the junction-specific incision, including the single-strand binding protein RPA and the multiprotein complex TFIIH, are not shown for clarity purposes.

tivity and DNA binding was not feasible for these smaller N-terminal deletion fragments (Table 1). However, we clearly show that the N-terminus of XPF is capable of binding DNA and contains endonuclease activity. Future studies using defined DNA substrates similar to those we have previously used to characterize the XPF/ERCC1 complex (*3*, *29*) will provide information regarding the specificity of XPF for repair intermediate substrates; however, this work focused on identifying the domain responsible for cleaving DNA. In addition, through site-directed mutagenesis, it may be possible to map the residues directly involved in catalysis. Nonetheless, results from the endonuclease and DNA binding studies indicate that a single domain with the ability to bind DNA resides in the first 378 residues of XPF. This domain likely imparts the endonuclease function necessary for the XPF/ERCC1 complex to incise at duplex/single-strand junctions during NER and recombination events (*1*, *3*, *4*). As XPF homology to Rad1 would suggest, this nucleolytic activity is similar to that of the Rad1/Rad10 complex from *S. cerevisiae* (*23*, *24*, *39*), and may reflect activity that was observed in Rad1 apart from its complex with Rad10 (*25*). Our evidence indicates that the endonuclease function of the XPF subunit provides the catalytic activity required for the incision events during NER.

We have also determined that the XPF/ERCC1 interaction occurs within the C-terminus of XPF. This was concluded from the fact that ERCC1 has no affinity for the first 702 residues of XPF but binds specifically to the 584C and 667−854 C-terminal deletion fragments (Figure 5). In a previous report, the XPF/ERCC1 interaction appeared to be restricted to the C-terminal 91 residues of XPF (*31*); however, interaction within a larger domain of XPF could have been obstructed by the polyclonal antibodies used for immunoprecipitation in this study. In the approach presented here, ERCC1 was immobilized on S-protein agarose using an N-terminal S-tag. By immobilizing ERCC1 in this way, binding of full-length XPF was not affected. Since XPF and the polypeptides 584C and 667−854 were capable of binding to sERCC1, our work clearly shows that amino acids 667−854 are sufficient for ERCC1 binding. In light of the report that amino acids 825 through 916 are required for binding (*31*), two interpretations can be drawn from our results. One is that the 29 amino acids between residues 825 and 854 are sufficient for XPF to interact with ERCC1. The second interpretation is that, although this short region of XPF can bind to ERCC1, the entire region of XPF that interacts with ERCC1 encompasses a larger C-terminal domain. The second interpretation was corroborated by more specific probing of the XPF/ERCC1 interaction using epitope-mapped mAbs. Our results showed that mAbs directed against the C-terminus of XPF could block the XPF/ERCC1 protein interaction (Figure 6). Specifically, mAbs 13 and 218 disrupted complex formation. These results were not intuitive based on the binding sites of the individual monoclonal antibodies. mAb 13, which affinity-maps to amino acids 702 through 854, was expected to inhibit binding of XPF to ERCC1. We did not, however, expect mAb 218 to affect the interaction since it has affinity for amino acids 854 through 916. Our reasoning was due to the fact that since polypeptide 667−854 could interact with ERCC1, mAb 218 would not affect the interaction because the epitope recognized by mAb 218 falls outside of this region. In addition, since mAb 13 maps to a region which is involved in the protein interaction, we reasoned that mAb 13 would inhibit the association of XPF and ERCC1. The possibility that binding of mAbs 218 and 13 to XPF induces a conformational change that inhibits ERCC1 binding or that the mAbs obstruct access of ERCC1 to the binding site cannot be ruled out; however, antibody-targeted disruption of complex formation by mAbs spanning the C-terminus of XPF indicates a larger region of XPF is involved in the protein interaction with ERCC1. This further demonstrates that complex formation occurs between residues 702 and the C-terminus of XPF. Our interpretation of the combined data is that although shorter regions of XPF can indeed bind to ERCC1, the entire protein interaction spans a larger portion of the C-terminus of XPF. We propose that a domain within at least the C-terminal 214 amino acids of XPF governs the interaction of the XPF and ERCC1 proteins and that XPF amino acids 825 through 854 are absolutely required for the interaction.

Further importance of this highly conserved C-terminal region of XPF is underscored by the location of biologically relevant mutations within the homologous domains of the *S. cerevisiae* and *S. pombe* Rad1 and Rad16 proteins, respectively. The *S. cerevisiae rad1−20* mutation impairs Rad1/Rad10 complex formation (*21*) through substitution of Asp for Gly, conserved as Gly 714 in XPF. Data from *S. pombe* alleles in the analogous region of Rad16 indicate a negative effect on the Rad16/Swi10 interaction (*38*), including truncation of the entire C-terminal region in *rad16−1*, and substitution of a Lys for Glu in *swi-9(rad16)*, conserved as Glu 690 in XPF. More importantly, two mutations linked to human disease in XPF (*2*) are also located in this proposed interaction domain at highly conserved residues, justifying speculation that alterations in this region could directly impact the stability of the XPF/ERCC1 complex.

The data presented here provide the first biochemical characterization of XPF as an isolated protein. Based on our experimental findings, we introduce a model for the function of XPF in the NER pathway (Figure 7). This model illustrates the 2 functional domains within the 916 amino acid protein.

As depicted, the N-terminal domain which spans the first 378 amino acids imparts the DNA binding and endonuclease activity, while the C-terminal domain which encompasses the last 214 amino acids is involved in the protein interaction with ERCC1. The specific affinity of the C-terminal domain of XPF for ERCC1, which is known to interact with XPA (*40*, *41*), is likely to regulate the general endonucleolytic activity provided by the N-terminal domain of XPF. As such, both domains of XPF are essential in the controlled incision events that take place during repair of damaged DNA and perhaps in the resolution of recombination intermediates. Further characterization of the XPF endonuclease activity and function of the catalytic domain through use of synthetic DNA substrates containing site-specific damage or mimicking NER or recombination intermediates will help elucidate the precise role that XPF plays in conferring specificity to the endonuclease activity. Understanding the crucial nature of complex formation with ERCC1 will also contribute to our knowledge of DNA excision repair and the role of XPF in human health.

ADDED IN PROOF

We acknowledge two recent publications that report homology of N-terminal and C-terminal domains of XPF to ancient DNA repair protein sequences, as we also describe here in the text referring to Figure 2: the first is by Aravind et al. (1999) (*42*) and the second by Sgouros et al. (1999) (*43*).

ACKNOWLEDGMENT

We are grateful to Dr. K. Fidelis for help with secondary structure predictions, and to Drs. I. McConnell, J. Felton, and J. George for critical comments on the manuscript.

REFERENCES

1. Park, C.-H., Bessho, T., Matsunaga, T., and Sancar, A. (1995) *J. Biol. Chem. 270*, 22657−22660.
2. Sijbers, A. M., de Laat, W. L., Ariza, R. R., Biggerstaff, M., Wei, Y. F., Moggs, J. G., Carter, K. C., Shell, B. K., Evans, E., de Jong, M. C., Jaspers, N. G., Wood, R. D., Bootsma, D., and Hoeijmakers, J. H. (1996) *Cell 86*, 811−822.
3. Bessho, T., Sancar, A., Thompson, L. H., and Thelen, M. P. (1997) *J. Biol. Chem. 272*, 3833−3837.
4. de Laat, W. L., Appeldoorn, E., Jaspers, N. G. J., and Hoeijmakers, J. H. (1998) *J. Biol. Chem. 273*, 7835−7842.
5. Sancar, A. (1996) *Annu. Rev. Biochem. 65*, 43−81.
6. Wood, R. D. (1997) *J. Biol. Chem. 272*, 23465−23468.
7. Thompson, L. H., Brookman, K. W., Weber, C. A., Salazar, E. P., Reardon, J. T., Sancar, A., and Siciliano, M. J. (1994) *Proc. Natl. Acad. Sci. U.S.A. 91*, 6855−6859.
8. Brookman, K. W., Lamerdin, J. E., Thelen, M. P., Hwang, M., Reardon, J. T., Sancar, A., Zhou, Z. Q., Walter, C. A., Parris, C. N., and Thompson, L. H. (1996) *Mol. Cell. Biol. 16*, 6553−6562.
9. Bootsma, D., Kraemer, K. H., Cleaver, J. E., and Hoeijmakers, J. H. J. (1998) in *The Genetic Basis of Human Cancer* (Vogelstein, B., and Kinzler, K. W., Eds.) pp 245−274, McGraw-Hill, New York.
10. Thompson, L. H. (1998) in *DNA Damage and Repair Vol. 2: DNA Repair in Higher Eukaryotes* (Nickoloff, J. A., and Hoekstra, M., Eds.) pp 335−393, Humana Press, Totowa, NJ.
11. Reardon J. T., Bessho, T., Kung, H. C., Bolton, P. H., and Sancar, A. (1997) *Proc. Natl. Acad. Sci. U.S.A. 94*, 9463−9468.
12. Sijbers, A. M., van Voorst Vader, P. C., Snoek, J. W., Raams, A., Jaspers, N. G., and Kleijer, W. J. (1998) *J. Invest. Dermatol. 110*, 832−836.
13. Park, C.-H., and Sancar, A. (1994) *Proc. Natl. Acad. Sci. U.S.A. 91*, 5017−5021.
14. Mu, D., Park, C. H., Matsunaga, T., Hsu, D. S., Reardon, J. T., and Sancar, A. (1995) *J. Biol. Chem. 270*, 2415−2418.
15. Aboussekhra, A., Biggerstaff, M., Shivji, M. K., Vilpo, J. A., Moncollin, V., Podust, V. N., Protic, M., Hübscher, U., Egly, J. M., and Wood R. D. (1995) *Cell 80*, 859−868.
16. Matsunaga, T., Park, C.-H., Bessho, T., Mu, D., and Sancar, A. (1996) *J. Biol. Chem. 271*, 11047−11050.
17. Wakasugi, M., and Sancar, A. (1998) *Proc. Natl. Acad. Sci. U.S.A. 95*, 6669−6674.
18. de Laat, W. L., Appeldoorn, E., Sugasawa, K., Weterings, E., Jaspers, N. G. J., and Hoeijmakers, J. H. (1998) *Genes Dev. 12*, 2598−2609.
19. Bardwell, L., Cooper, A. J., and Friedberg, E. C. **(**1992) *Mol. Cell. Biol. 12*, 3041−3049.
20. Bailly, V., Sommers, C. H., Sung, P., Prakash, L., and Prakash S. (1992) *Proc. Natl. Acad. Sci. U.S.A. 89*, 8273−8277.
21. Siede, W., Friedberg, A. S., and Friedberg, E. C. (1993) *J. Bacteriol. 175*, 6345−6347.
22. Bardwell, A. J., Bardwell, L., Johnson, D. K., and Friedberg, E. C. (1993) *Mol. Microbiol. 8*, 1177−1188.
23. Tomkinson, A. E., Bardwell, A. J., Bardwell, L., Tappe, N. J., and Friedberg, E. C. (1993) *Nature 362*, 860−862.
24. Sung, P., Reynolds, P., Prakash, L., and Prakash, S. (1993) *J. Biol. Chem. 268*, 26391−26399.
25. Habraken, Y., Sung, P., Prakash, L., and Prakash, S. (1994) *Nature 371*, 531−534.
26. Davies, A. A., Friedberg, E. C., Tomkinson, A. E., Wood, R. D., and West, S. C. (1995) *J. Biol. Chem. 270*, 24638−24641.
27. Bardwell, A. J., Bardwell, L., Tomkinson, A. E., and Friedberg, E. C. (1994) *Science 265*, 2082−2085.
28. Rodriguez, K., Wang, Z., Friedberg, E. C., and Tomkinson, A. E. (1996) *J. Biol. Chem. 271*, 20551−20558.
29. McCutchen-Maloney, S. L., Hwang, M., Bessho, T., Sancar, A., and Thelen, M. P. (1997) *FASEB J. 11*, A13723.
30. Matsumura, Y., Nishigori, C., Yagi, T., Imamura, S., and Takebe, H. (1998) *Hum. Mol. Genet. 7*, 969−974.
31. de Laat, W. L., Sijbers, A. M., Odijk, H., Jaspers, N. G. J., and Hoeijmakers, J. H. (1998) *Nucleic Acids Res. 26*, 4146−4152.
32. Kim, J. S., and Raines, R. T. (1993) *Protein Sci. 2*, 348−356.
33. Westerveld, A., Hoeijmakers, J. H., van Duin, M., de Wit, J., Odijk, H., Pastink, A., Wood, R. D., and Bootsma, D. (1984) *Nature 310*, 425−429.
34. Blank, A., Sugiyama, R. H., and Dekker, C. A. (1982) *Anal. Biochem. 120*, 267−275.
35. Hatakeyama, S., Ito, Y., Shimane, A., Ishii, C., and Inoue, H. (1998) *Curr. Genet. 33*, 276−283.
36. Rost, B., and Sander, C. (1993) *J. Mol. Biol. 232*, 584−599.
37. Altschul, S. F., Madden, T. L., Schäffer, A. A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D. J. (1997) *Nucleic Acids Res. 25*, 3389−3402.
38. Carr, A. M., Schmidt, H., Kirchhoff, S., Muriel, W. J., Sheldrick, K. S., Griffiths, D. J., Basmacioglu, C. N., Subramani, S., Clegg, M., Nasim, A., and Lehman, A. (1994) *Mol. Cell. Biol. 14*, 2029−2040.
39. Tomkinson, A. E., Bardwell, A. J., Tappe, N., Ramos, W., and Friedberg, E. C. (1994) *Biochemisty 33*, 5305−5311.
40. Li, L., Elledge, S. J., Peterson, C. A., Bales, E. S., and Legerski, R. J. (1994) *Proc. Natl. Acad. Sci. U.S.A. 91*, 5012−5016.
41. Saijo, M., Kuraoka, I., Masutani, C., Hanaoka, F., and Tanaka, K. (1996) *Nucleic Acids Res. 24*, 4719−4724.
42. Aravind, L., Walker, D. R., and Koonin, E. V. (1999) *Nucleic Acids Res. 27*, 1223−1242.
43. Sgouros, J., Gaillard, P. H. L., and Wood, R. D. (1999) *Trends Biochem. Sci. 24*, 95−97.

BI990591+