Structural Determinants for Substrate Binding and Catalysis by the Structure-specific Endonuclease XPG*

Marcel Hohli, Fabrizio Thorel, Stuart G. Clarkson, and Orlando D. Schärer†‡¶

From the ¶Institute of Molecular Cancer Research, University of Zürich, August Forel Strasse 7, CH-8008 Zürich, Switzerland and the §Department of Genetics and Microbiology, University Medical Centre, 1 rue Michel-Server, 1211 Geneva 4, Switzerland

Received for publication, December 23, 2002, and in revised form, March 12, 2003
Published, JBC Papers in Press, March 18, 2003, DOI 10.1074/jbc.M213155200

XPG belongs to the Fen1 family of structure-specific nucleases and is responsible for the 3’ endonucleolytic incision during mammalian nucleotide excision repair. In addition, it has ill-defined roles in the transcription-coupled repair of oxidative DNA damage and likely also in transcription that are independent of its nuclease activity. We have used DNA binding and footprinting assays with various substrates to gain insight into how XPG interacts with DNA. Ethylation interference footprinting revealed that XPG binds to its substrates through interaction with the phosphate backbone on one face of the helix, mainly to the double-stranded DNA. By comparing DNA binding and cleavage activity using single-/double-stranded DNA junction substrates differing in the length of the single-stranded regions, we have found that the 3’ but not the 5’ single-stranded arm was necessary for DNA binding and incision activity. Furthermore, we show that although a 5’ overhang is not required for XPG activity, an overhang containing double-stranded DNA near the junction inhibits the nuclease but not substrate binding activity. Apparently, junction accessibility or flexibility is important for catalysis but not binding of XPG. These results show that XPG has distinct requirements for binding and cleaving DNA substrates.

Nucleases with specificity for DNA structural features rather than sequence are important for many aspects of DNA metabolism, including replication, recombination, and repair (1, 2). One such structure-specific endonuclease, human XPG, is involved in nucleotide excision repair (NER)† in a poorly defined pathway of transcription-coupled base excision repair of oxidative lesions and likely also in efficient RNA polymerase II-mediated transcription (3–7). Two inherited human syndromes are associated with deficiencies in XPG: xeroderma pigmentosum (XP) and Cockayne syndrome (CS). Although the XP phenotype arises from a defect in NER and is characterized by acute sensitivity to UV light and an extremely high incidence of skin cancer (8), defects in transcription-coupled base excision repair and transcription are likely responsible for certain manifestations of CS such as skeletal, neurological, and developmental abnormalities (9). Whether a defect in XPG causes XP or combined XP/CS is dependent on the position and type of mutation in the XPG gene. Point mutations resulting in a defect in the nuclease activity generally lead to an XP phenotype, whereas mutations resulting in truncation of the protein cause an XP/CS phenotype (10). These observations have led to the model that the nuclease activity of XPG is important for NER, whereas the involvement of XPG in transcription-coupled base excision repair and transcription requires binding of XPG to specific DNA structures but not its nuclease activity.

At the biochemical level, only the involvement of XPG in NER is understood in any detail. In mammals, NER is the major DNA repair pathway responsible for the removal of adducts formed by UV light and environmental mutagens from DNA (11). Two partially overlapping NER subpathways have been discerned: one process called global genome repair, which is able to eliminate NER lesions from all locations in the genome, and transcription-coupled repair, which is responsible for the accelerated removal of lesions from actively transcribed DNA strands (12). Global genome repair has been elucidated in some detail and can be divided in three major stages: (i) damage recognition and DNA opening to form a bubble-shaped structure around the lesion by the XPC-HR23B, TFIIH, XPA, RPA, and XPG factors, (ii) endonucleolytic incision 5’ and 3’ of the lesion by ERCC1-XPF and XPG, respectively, and release of an oligonucleotide of 24–32 residues, and (iii) gap-filling repair synthesis and ligation to restore the original DNA sequence (13–15). Within NER, XPG also fulfills roles that are independent of its nuclease activity in the opening of the DNA helix around the lesion and the formation of a stable damage recognition complex (16, 17). The physical presence but not the catalytic activity of XPG is also required for the subsequent 5’ incision by ERCC1-XPF (18). Conversely, the nuclease function of XPG may be regulated through interaction with other NER factors, notably TFIIH, RPA, and XPA (19–21). In transcription-coupled repair, the damage is thought to be recognized by the damage-stalled RNA polymerase. Although the detailed mechanism of transcription-coupled repair has not yet been clarified, it is thought to proceed similarly to global genome repair in steps subsequent to damage recognition.

The enzymatic activities of XPG have been studied in vitro, where it shows remarkable structure-specific properties, cleaving artificial DNA structures that contain ss/dsDNA junctions including bubbles, splayed arms, stem-loops, and flap substrates with the same polarity as in the NER reaction (22–24). The XPG protein (1186 amino acids) is a member of the Fen1 (flap endonuclease 1) family of structure-specific nucleases (2). Members of this family have a conserved nu-

* This work was supported by the Swiss National Science Foundation. 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.
† START Fellow of the Swiss National Science Foundation, EMBO Young Investigator. To whom correspondence should be addressed. Tel.: 41-1-634-8919; Fax: 41-1-634-8904; E-mail: scharer@imr.unizh.ch.
‡ The abbreviations used are: NER, nucleotide excision repair; CS, Cockayne syndrome; ss/dsDNA, single-/double-stranded DNA; XP, xeroderma pigmentosum; WT, wild type; nt, nucleotide(s); PMSF, phenylmethylsulfonyl fluoride; EMSA, electrophoretic mobility shift assay; ExoIII, exonuclease III; TBE, Tris borate/EDTA electrophoresis buffer.

© 2003 by The American Society for Biochemistry and Molecular Biology, Inc. Printed in U.S.A.
interaction of XPG endonuclease with its DNA substrates

Although the Fen1 family members Fen1, XPG, and Exo1 (exonuclease 1) share a conserved nucleotide domain and bind to ss/dsDNA junctions, there are differences in the exact substrate specificity of these three proteins. Although XPG efficiently processes bubble or stem-loop substrates, Fen1 and Exo1 only act on substrates that have free 5′ ssDNA ends (31, 32). In the absence of high resolution structures of Fen1 family members bound to DNA, indications of how these proteins might interact with their DNA substrates are based on DNA footprinting experiments of Fen1, Exo1, and the 5′ exonuclease domain of DNA polymerase I and on modeling studies based on the structures of the free proteins (27, 33–37). Here, we have analyzed the requirements of XPG for binding to ss/dsDNA junctions using band shift and footprinting methods. Furthermore, we have studied the structural elements required for DNA binding and cleavage activities of XPG by using various substrates. Unexpectedly, XPG displays distinct requirements for binding and cleaving its DNA substrates. These observations might help to elucidate the catalytic and structural roles of XPG in various repair processes.

EXPERIMENTAL PROCEDURES

All of the enzymes were purchased from New England Biolabs unless indicated otherwise. γ-32P-ATP and α-32P-dATP were obtained from Amersham Biosciences. All of the other chemicals and reagents were from Sigma or Fluka. Routine molecular biological procedures were performed using standard protocols (38).

Oligonucleotides Used in This Study—Oligonucleotides used were either synthesized on an Expedite 8909 nucleic acid synthesis system or purchased from Metabion (Germany) and purified by denaturing PAGE. 5′ end labeling was carried out using T4 polynucleotide kinase and [γ-32P]-ATP. For 3′ labeling, the oligonucleotides were incubated with terminal deoxynucleotidyl transferase and [α-32P]dATP using the protocol supplied by the manufacturer (Promega). The labeled oligonucleotide was annealed with the complementary strand in annealing buffer (50 mM NaCl, 10 mM Tris-Cl, pH 8.0) by heating for 5 min at 95°C and cooling to room temperature. Substrates consisting of two oligonucleotides (substrates 1, 2, and 9) were annealed with a 3-fold excess of the unlabeled cold complementary strand, whereas for substrates with more than two oligonucleotides (substrates 3–8), a 2-fold excess of the two unlabeled oligonucleotides was used. Annealing efficiencies were tested for every substrate by native PAGE and were found to be higher than 95%. The sequences of the oligonucleotides used were indicated otherwise.

Interaction of XPG Endonuclease with its DNA Substrates

### Table I

| substrate | length | oligo name |
|-----------|--------|------------|
| 1         | n = 20 | A20, B19   |
| 2         | n = 2  | A2, "     |
| 3         | n = 3  | B3, "     |
| 4         | n = 2  | B2, "     |
| 5         | n = 1  | B1, "     |
| 6         | n = 0  | B0, "     |
| 7         | n = 19 | A20, B19, C19 |
| 8         | n = 17 | A20, B19, D17 |
| 9         | n = 39 | A20, B19, E39a |
| 10        | n = 30 | A0, F30    |

The lengths of the individual regions of the oligonucleotides are indicated (in nucleotides) for fixed lengths in the drawing; variable lengths are expressed as n, and the lengths are indicated in the list. Asterisks indicate the 5′-32P-labeled end of the substrates. In figure legends, we use the substrate labels 1–9 and show the length of n in subscript. Oligonucleotide sequences are given under "Experimental Procedures."

Construction of Recombinant Baculoviruses for XPG Expression in S9 Insect Cells—A 246-bp BglII/Asp718 fragment containing the 3′ terminus of XPG cDNA was excised from the pBS-XPG plasmid (39). Oligonucleotides sc-404 (5′-GATCTGTGTTTGGGAAGAAAGAAGGAAACATAGGGCGGAGAACGTGAGGGAGAGAAGAAGGAAACTAAGAGCTGCGAGGGGAAGAAAAAGGAAAACTCACCATGACCAGCTTGGTACCTTAGTTGTGGTGA-3′) and sc-405 (5′-GTACCATTTGATTGTGGTGA-TGGTGAATTCGTTCCTTTTCATTCTTCCCTCAGCGTCTTTAGTTCTTTCCCTCAAAACAC-3′) were annealed to generate a BglII/Asp718 fragment encoding the 3′ end of XPG followed by 6 histidine residues. This fragment was then cloned between the BglII and Asp718 sites of pBS-XPG to form pBS-XPG-His6. pBS-XPG791A-His6 was generated by replacing the 1843-bp NdeI-BglII fragment from pBS-XPG-His6 with the corresponding fragment from pGEM-XPG791A encoding the E791A substitution (26). pFastBac1-XPG-His6 and pFastBac1-XPG791A-His6 were generated by cloning the 3646-bp Nof/Isp718 fragment from pBS-XPG-His6 and pBS-XPG791A-His6, respectively, into the NolI and Asp718 sites of the vector pFastBac1 (Invitrogen). Recombinant XPG baculoviruses were prepared according to the manufacturer’s instructions (Invitrogen).

Expression and Purification of WT and Mutant XPG Proteins—Se-
run-free conditioned SF9 insect cells were grown in suspension at 27 °C under extensive stirring (according to instructions from Invitrogen). For protein expression, 300 ml of SF9 cells (1.5 × 10^6 cells/ml) were infected with the appropriate baculovirus at an multiplicity of infection of 5 and were grown in an Erlenmeyer flask with air-permeable filter caps at 27 °C and 200 rpm. The cells were harvested 48–54 h after infection by low speed centrifugation and resuspended in 40 ml of buffer A (50 mM potassium phosphate, pH 7.6, 10% glycerol, 5 mM β-mercaptoethanol, 1 mM PMSF, containing one EDTA-free protease inhibitor mixture tablet from Roche Applied Science). After incubation on ice for 20 min, NaCl was added to a final concentration of 800 mM, and the lysate was sonicated on ice (Bandelin Sonicator, 60% power at 30% interval) by three 1-min intervals, with at least 2 min of cooling on ice between each interval. The lysate was clarified by centrifugation at 40,000 × g for 45 min at 4 °C. The supernatant was supplemented with 1 mM PMSF and incubated overnight with 4 ml of nickel-nitriolactic acid-agarose beads (Qiagen) in presence of 1 mM imidazole. The beads were collected by low speed centrifugation and resuspended in 10 ml of buffer A containing 800 mM NaCl and 1 mM imidazole and packed in column (1 × 5 cm). The column was washed stepwise with 20 ml of buffer A containing 1 mM imidazole with 500 ml of NaCl, 20 ml of 10 mM imidazole with 300 mM NaCl, or 10 ml of 50 mM imidazole with 300 mM NaCl. XPG was eluted with 6 ml of buffer A containing 100 mM imidazole with 300 mM NaCl. The eluate was then applied on a 1-ml HiTrap SP column (Amersham Biosciences) after a 1:2 dilution in buffer B (25 mM HEPES, pH 8.0, 10% glycerol, 1 mM EDTA, 10 mM β-mercaptoethanol, 1 mM PMSF) at a flow rate of 0.5 ml/min. The protein was eluted with a 25-ml gradient from 100 mM to 1 M NaCl. XPG eluted between 450 and 700 mM NaCl. Fractions containing XPG were pooled and diluted to 150 mM NaCl with buffer C (25 mM HEPES, pH 7.5, 10 mM β-mercaptoethanol, 1 mM PMSF) and loaded on a 1-ml HiTrap Q column (Amersham Biosciences). XPG eluted in 2.5 ml between 400 and 500 mM NaCl with a 30-ml gradient from 150 mM to 1 M NaCl. The protein concentration was estimated from absorption at 280 nm. Purified protein was flash frozen in aliquots and stored at −80 °C. Typically, 0.3–1.2 mg of pure XPG protein at a concentration of 0.2–0.5 mg/ml was obtained.

**Endonuclease assay**—2.5 μM of 5′-32P end-labeled DNA substrate was incubated with 1.25–5 μM XPG in nuclease buffer (25 mM Tris-Cl, pH 8.0, 25 mM KCl, 15 mM β-mercaptoethanol, 0.5 mM bovine serum albumin, 30 mM KCl) in the presence of 2 mM MgCl2 or 0.5 mM MnCl2 in a reaction volume of 20 μl for 1–2 h. The reactions were stopped by the addition of an equal volume of formamide loading buffer (90% (w/v) formamide, 10 mM β-mercaptoethanol, 5 mM β-mercaptoethanol, 1 mM PMSF, containing one EDTA-free protease inhibitor mixture tablet from Roche Applied Science). After incubation on ice for 20 min, NaCl was added to a final concentration of 800 mM, and the lysate was sonicated on ice (Bandelin Sonicator, 60% power at 30% interval) by three 1-min intervals, with at least 2 min of cooling on ice between each interval. Following quantification of the samples in a scintillation counter, the samples were ethanol-precipitated and resuspended in gel loading buffer (25 μl) and annealed in annealing buffer (see above) with a 5-fold excess of the complementary strand by heating for 5 min at 90 °C and cooling to room temperature.

For the DNA binding reaction, 25 μl of substrate 1 (A50-B50) were incubated under EMSA conditions described above in presence of a 100-fold excess of cold dsDNA competitor (substrate 9) with 150 nM of WT XPG in a volume of 20 μl for 5–10 min at room temperature, resulting in about 40–50% complex formation. The mixture was run on a 5% native polyacrylamide gel for 90 min at 90 V. The bands of bound and unbound fractions were visualized by autoradiography, excised, and electroeluted (1× TBE buffer) into 10 μl NH4OAc buffer for 4 h at 90 V. The volume of the electroeluted fractions was reduced on Microcon YM-10 filters (Millipore) and washed twice with TE buffer before precipitation with 100% ethanol in presence of 10 mg/ml tRNA followed by washing twice with 80% ethanol. DNA strand cleavage at ethylated phosphate sites was achieved by incubation in 100 μl of 140 mM NaOH at 90 °C for 30 min. After neutralization with 15 μl of 1x AcOH, the samples were ethanol-precipitated and resuspended in gel loading buffer. Following quantification of the samples in a scintillation counter, the samples were loaded onto a 15% 7M urea denaturing gel along with T, C, and G/A sequencing ladders prepared from the same target DNA (40, 41). Briefly, the T-specific modification was prepared by incubation of 0.5 pmol of denatured labeled oligonucleotide DNA in 1.27 mM potassium permanganate for 20 min, for C in 3 M hydroxylamine hydrochloride, pH 6.0, and for G/A in 80% formic acid in a volume of 20 μl. Modified oligonucleotides were ethanol-precipitated and cleaved at modified sites by incubation for 15 min at 90 °C in 1.1 M pyridoline and evaporated to dryness. The gels were visualized and quantified by PhosphorImager.

**RESULTS**

**Activity of Purified XPG Protein on ds/dsDNA Junctions**—Wild type and E791A mutant human histidine-tagged XPG endonuclease was purified from SF9 insect cells as described under “Experimental Procedures” (Fig. 1A). We compared the...
structure-specific nuclease activity of XPG on splayed arm, bubble, and stem-loop DNA substrates (Fig. 1B). Incubation of XPG in the presence of either MgCl$_2$ (lanes 2, 5, and 8) or MnCl$_2$ (lanes 3, 6, and 9) resulted in efficient DNA cleavage of the splayed arm (22 and 91%, respectively) and bubble substrates (70 and 96%), whereas only weak activity was found on the stem-loop substrate (3 and 4%). The cleavage positions were mapped 1 nt into the duplex region of the junction for the splayed arm substrate, whereas the bubble and stem-loop structures were cleaved 1 nt into the ssDNA just 5’ of the ss/dsDNA junction (data not shown). Although a bubble-shaped structure might formally resemble a reaction intermediate of NER more closely than a splayed arm structure, we decided to perform our further analysis of substrate binding and cleavage by XPG with the splayed arm substrate, because binding of the protein at two equivalent positions of a bubble would complicate such an analysis. We designate a 5’ ssDNA overhang of a splayed arm substrate as an unpaired overhang with a 5’ end, whereas a 3’ ssDNA overhang refers to the unpaired overhang with a 3’ ssDNA end.

**Structure-specific Binding of XPG to ss/dsDNA Junctions—** We used EMSA to define conditions that allow specific binding of the XPG endonuclease to ss/dsDNA junctions. Incubation of XPG with a splayed arm substrate (substrate 1 in Table I) in the absence of metal cofactor resulted in the formation of a band indicative of a protein-DNA complex (Fig. 2A, lane 2). Optimal binding was observed over a broad range of pH (6–8.5) in the presence of bovine serum albumin and glycerol and at a KCl concentration of <25 mM. Stable DNA binding was observed if the concentration of substrate was ≥1 nM and that of MnCl$_2$ (2 mM) was incubated with XPG (20 nM) with or without divalent metal ions. Lanes 2 and 7, no metal. Lanes 3 and 8, 2 mM MgCl$_2$. Lanes 4 and 9, 0.5 mM MnCl$_2$. Lanes 5 and 10, 2 mM CaCl$_2$. The positions of XPG-DNA complex (c), free substrate (fs), unannealed oligonucleotide A$_{ss}$ (as), and cleaved substrate (cs) are indicated. B, influence of oligonucleotide competitors on substrate binding by XPG. 5’-$^{32}$P-labeled substrate 1$_{ss}$ (1.25 nM) was incubated for 20 min at room temperature with XPG (9.75 nM) in the presence of unlabeled competitor oligonucleotides. The following competitors were used: lanes 3–5, A$_{ss}$; lanes 6–8, 9$_{ss}$; lanes 9–11, 3$_{ss}$, and lanes 12–14, 1$_{ss}$. The concentrations of competitor were: lanes 3, 6, 9, and 12, 97.5 nM (10-fold excess); lanes 4, 7, 10, and 13, 195 nM (20-fold excess); and lanes 5, 8, 11, and 14, 390 nM (40-fold excess). XPG was ≥6 nM. At high XPG concentrations (≥10 nM), we observed the appearance of a second band, which may be due to the formation of a dimeric XPG-DNA complex.

**Fig. 2. Junction-specific DNA binding by the XPG endonuclease.** A, influence of divalent metal ions on substrate binding by wild type and E791A mutant XPG proteins. Substrate 1$_{ss}$ (2.5 nM) was incubated with XPG (20 nM) with or without divalent metal ions. Lanes 2 and 7, no metal. Lanes 3 and 8, 2 mM MgCl$_2$. Lanes 4 and 9, 0.5 mM MnCl$_2$. Lanes 5 and 10, 2 mM CaCl$_2$. The positions of XPG-DNA complex (c), free substrate (fs), unannealed oligonucleotide A$_{ss}$ (as), and cleaved substrate (cs) are indicated. B, influence of oligonucleotide competitors on substrate binding by XPG. 5’-$^{32}$P-labeled substrate 1$_{ss}$ (1.25 nM) was incubated for 20 min at room temperature with XPG (9.75 nM) in the presence of unlabeled competitor oligonucleotides. The following competitors were used: lanes 3–5, A$_{ss}$; lanes 6–8, 9$_{ss}$; lanes 9–11, 3$_{ss}$, and lanes 12–14, 1$_{ss}$. The concentrations of competitor were: lanes 3, 6, 9, and 12, 97.5 nM (10-fold excess); lanes 4, 7, 10, and 13, 195 nM (20-fold excess); and lanes 5, 8, 11, and 14, 390 nM (40-fold excess).

**Fig. 3. ExoIII protection footprinting reveals the 3’ borders of XPG junction-specific DNA binding.** The incised or complementary strands of the splayed arm substrate 1$_{ss}$ (2.5 nM) were 5’-$^{32}$P-end labeled substrate and incubated with no (lanes 5 and 11), 12 nM (lanes 6 and 12), or 30 nM (lanes 7 and 13) XPG (E791A) in the presence of 2 mM MgCl$_2$ and a 100-fold excess of double strand cold competitor (substrate 9) and digested with ExoIII. Sequencing ladders are shown for T (lanes 1 and 8), C (lanes 2 and 9), and G/A (lanes 3 and 10). The position of incision by WT XPG is shown in lane 4. Regions of single- and double-stranded DNA are indicated by bars on the right sides of the panels. Bands indicating the 3’ border of XPG substrate binding and their distance from the junction are shown. Through small variations in pH or enzyme concentration, cleavage could be prevented entirely without compromising DNA binding (data not shown). The presence of Ca$^{2+}$, which does not support nuclease activity in most enzymes, resulted in the formation of stable protein-DNA complex with both WT and the E791A mutant (Fig. 2A, lanes 8 and 9). To determine whether the interaction of XPG with the substrate was specific for a ss/dsDNA junction, we tested the effect of various DNA competitors on the binding reaction. Oligonucleotides containing only the ssDNA (Fig. 2B, lanes 3–5) or dsDNA (lanes 6–8) part of the splayed arm substrate had no effect on the binding reaction (a >500-fold excess was needed to detect weak competition), whereas the addition of a 40-fold excess of specific competitors, either flab substrate 3 (lanes 9–11) or splayed arm substrate 1 (lanes 12–14), abolished the binding reactions under the same conditions. These observations demonstrate that XPG shows a strong binding preference for DNA molecules containing ss/dsDNA junctions.
protein is bound to the DNA, ExoIII stalls on the 3′ side of the protein-binding site, leading to preferential incision at that position. Because the exonuclease requires the presence of Mg2+ for activity, footprinting was performed with the XPG active site mutant E791A under conditions allowing only ≤5% of incision to occur (data not shown). To detect the 3′ borders of the junction binding, XPG was incubated with the 5′ end-labeled substrate and the unprotected DNA degraded by ExoIII following equilibration (Fig. 3, lanes 6, 7, 12, and 13).

On the incised strand, where XPG WT protein cleaves 1 nt into the duplex region (lane 4), strong bands could be detected 5–7 bases in the duplex region away from the junction. Bands of higher mobility, 3–5 nt in the single-stranded 5′ overhang, were detectable. The meaning of these bands is unclear, because our further footprinting experiments (see below) did not reveal any DNA interactions upstream of this region. Furthermore, very weak bands were observed 14–15 nt in the duplex region, which were more pronounced using shorter digestion times or lower ExoIII concentrations (data not shown). However, these milder digestion conditions failed to reveal the prominent boundary of 5–7 nt into the duplex region. Apparently, ExoIII is able to displace XPG from the ss/dsDNA junction under certain conditions. On the complementary strand, a clear band was obvious on the first nt on either side of the junction, indicating that the binding site of XPG starts around the junction in the dsDNA region. The faster migrating fragments are most likely products of nonspecific degradation, because the control without XPG (lane 11) shows the same pattern. A second XPG-specific band was mapped close to the end of the 3′ overhang. Again, our studies were not able to explain the nature of this specific band. The pattern of ExoIII footprinting suggests that XPG contacts the incision as well as the complementary strand of its substrate, mainly in the dsDNA region.

**Phosphate Ethylation Interference DNA Footprinting**—To gain more detailed insight of how XPG binds to junctions, we performed ethylation interference footprinting (42), which directly detects contacts of the protein with the phosphates in the substrate DNA. We treated either the incised or complementary 5′ end-labeled substrate and the unprotected DNA degraded by ExoIII following equilibration (Fig. 3, lanes 2 and 8), C (lanes 3 and 9), and G/A (lanes 4 and 10) are shown. The black squares in lanes 3 and 10 mark the first paired nt at the ss/dsDNA junction. The cleavage profile of ethylN-isourea-treated DNA are shown. Lanes 5 and 11, hydrolysis of ethylated 1Δ prior to the binding reaction. Unbound and bound fractions are shown in lanes 6 and 12 and lanes 7 and 13, respectively. The strongest phosphate contacts are marked by filled circles, and their positions are indicated by numbers; weaker contacts are represented by open circles. Positive numbers correspond to the double-stranded regions counted from the junction, and negative numbers refer to the single-stranded region counted from the junction.

**Fig. 4.** Mapping of XPG DNA backbone phosphate contacts by ethylation interference footprinting. Ethylated substrate 1Δ (25 nM) was incubated for 15 min at room temperature with 250–300 nM XPG in the absence of metal cofactor. DNA-protein complexes were separated from free DNA on a native 5% polyacrylamide gel, electroeluted, and cleaved at the modified sites by alkaline hydrolysis, as described under “Experimental Procedures.” A, ethylN-isourea footprint of WT XPG on the incised and complementary strand. The position of incision by XPG is shown in lane 1, and the sequencing ladders for T (lanes 2 and 8), C (lanes 3 and 9), and G/A (lanes 4 and 10) are shown. The black squares in lanes 3 and 10 mark the first paired nt at the ss/dsDNA junction. The cleavage profile of ethylN-isourea-treated DNA are shown. Lanes 5 and 11, hydrolysis of ethylated 1Δ prior to the binding reaction. Unbound and bound fractions are shown in lanes 6 and 12 and lanes 7 and 13, respectively. The strongest phosphate contacts are marked by filled circles, and their positions are indicated by numbers; weaker contacts are represented by open circles. Positive numbers correspond to the double-stranded regions counted from the junction, and negative numbers refer to the single-stranded region counted from the junction. B, quantification of differences in bound/unbound fractions of three independent experiments on the incised and complementary strand. For each lane, the radioactivity present in each band was quantified using PhosphorImager and Image Quant software and was normalized to the total input of radioactivity/lane. The normalized value for each band of the bound fraction was then divided by the normalized value of the corresponding band of the unbound fraction and expressed as a percentage (relative intensity). No ethylation interference refers to a value of 100%, inhibition of substrate binding by ethylation is indicated by a value of <100%, and increased binding to ethylated DNA results a value of >100%. In the diagram the values are indicated as a deviation from 100%, and the standard deviations are given in error bars.
XPG endonuclease makes numerous phosphate contacts on both strands around the junction, mainly in the dsDNA region.

**Minimal Size of the Single-stranded Overhang Required for Substrate Cleavage and Binding**—Because our EMSA and footprinting data revealed specific binding to the double-stranded region at the ss/dsDNA junction but not extensive DNA contacts to the single-stranded regions of the substrate, we decided to determine the minimal lengths of the 5' and 3' single-stranded overhangs needed for DNA binding and cleavage activity. We therefore tested the incision activity of XPG on splayed arm substrates with various lengths of the ssDNA region of the incised strand, the complementary strand, or both. Shortening the length of the 5' overhang from 20 to 2, 1, or 0 nt resulted in a 1.5-fold (2 nt) or 2-fold (1 and 0 nt) reduction of incision activity (Fig. 5, A and C). The position of the incision remained unchanged, yielding 5' end-labeled cleavage products of 21, 3, 2, and 1 nt in length, respectively, and a product of 20 nt in length in all cases if the incised strand was labeled at the 3' end (data not shown). Truncation of the 3' ssDNA overhang from 19 to 3, 2, 1, and 0 nt had a more pronounced influence leading to 2- and 4-fold reduction (3- and 2-nt overhangs) and a trace or no activity (1 and 0 nt overhangs), respectively (Fig. 5, B and C). The efficiency of cleavage of substrates with truncations of the complementary strand was partially restored by substituting Mn$^{2+}$ for Mg$^{2+}$, although this led to the formation of products with different incision sites in some cases (Fig. 5B, lanes 15–20). Apparently, the two catalytically divalent metal ions can influence the conformation of the DNA substrates, which can affect protein activity under certain conditions. We then examined the effect of simultaneous truncation of both the 5' and 3' ssDNA overhangs of the splayed arm substrate (Fig. 5D). We used substrates with 2- and 0-nt 5' overhangs together with a 3' overhang of 20, 3, 2, or 0 nt (Fig. 5C). We observed a combined decrease in incision efficiency by shortening both ssDNA regions. The smallest substrate that was cleaved, albeit at very low efficiency, consisted of a 5' overhang of 0 nt and a 3' overhang of 2 nt (Fig. 5C, lane 29).

We wondered whether the reduction of cleavage efficiency upon truncation of the 3' ssDNA overhang was caused by a reduction in the DNA binding or phosphodiester bond hydrolysis step. Therefore, we performed EMSA assays to test DNA binding activity on substrates with truncated 3' overhangs. Indeed, DNA binding decreased monotonically with shortening of the 3' ssDNA overhang from 20 to 3, 2, 1, and 0 nt (Fig. 6), indicating that the 3' ssDNA overhang is required for binding of XPG to ss/dsDNA junctions. In contrast, truncation of the 5' overhang did not show a significant decrease in DNA binding activity (data not shown).
Interaction of XPG Endonuclease with its DNA Substrates

Fig. 7. Determination of substrate requirements of XPG for DNA binding and cleavage activity. The various substrates used are indicated above of the figures with numbers corresponding to those in Table I. A, EMSA with 1 nM substrate in presence of a 150-fold excess (150 nM) of unlabeled dsDNA. Substrates 1–5 (lanes 1–5), 3 (lanes 6–10), 4 (lanes 11–15), 5 (lanes 16–20), 6 (lanes 21–25), 7 (lanes 26–30), and 8 (lanes 31–35) were incubated with XPG at the following concentrations: 0 nM (lanes 1, 6, 11, 16, 21, 26, and 31), 0.8 nM (lanes 2, 7, 12, 17, 22, 27, and 32), 1.6 nM (lanes 3, 8, 13, 18, 23, 28, and 33), 3.2 nM (lanes 4, 9, 14, 19, 24, 29, and 34), 6.4 nM (lanes 5, 10, 15, 20, 25, 30, and 35). B, incision assay with 2.5 nM substrates 1–2 (lanes 1 and 2), 3 (lane 3), 4 (lane 4), 5 (lane 5), 6 (lane 6), 7 (lane 7), and 8 (lane 8) and 0 nM (lanes 1) or 5 nM (lanes 2–8) XPG protein in the presence of 2 mM MgCl2.

Separation of DNA Binding and Cleavage Activity Using Different Junction Structures—Based on our results demonstrating that the minimal requirement for processing a ss/dsDNA junction is a 2 nt 3’ overhang, we became interested in determining how replacing the ssDNA overhangs partially or entirely with dsDNA would affect XPG activity (Fig. 7). We first investigated flap substrates, in which either the 5’ or 3’ ssDNA overhang was replaced with dsDNA. Substrates in which the 3’ overhang on the nonincised strand was double-stranded (substrate 3) were bound and cleaved with similar efficiency as the splayed arm substrate 2 (Fig. 7, compare A, lanes 6–10 and 1–5, with B, lanes 3 and 2). Substrates containing dsDNA on the 5’ overhang (substrate 4) were bound efficiently (Fig. 7A, lanes 11–15) but were not processed under the same conditions (Fig. 7B, lane 4). Apparently, a 5’ dsDNA overhang interferes with incision but not with binding by XPG. A related substrate in which the dsDNA 5’ overhang was recessed by 3 nt on the side of the junction (substrate 5) yielding 3 nt of ssDNA adjacent to the junction was bound (Fig. 7A, lanes 16–20) and processed by XPG (Fig. 7B, lane 5), indicating that ssDNA character of the 5’ overhang was only required near the junction but not at the 5’ end. We then investigated substrates in which both the 5’ and 3’ overhangs were made of dsDNA (substrate 6), and three-way junctions with (substrate 8) or without mismatches (substrate 7) at the site of the junction. These substrates were bound preferentially compared with the splayed arm, flap, or duplex 5’ overhang structures (Fig. 7A, compare lanes 21–35 and 1–20), but displayed minimal (gap structure 6; Fig. 7B, lane 6) or no incision (three-way junctions with or without mismatches; Fig. 7B, lanes 7 and 8). Increasing XPG concentration by a factor of 30 and labeling of the complementary strands did not result in cleavage (data not shown). Apparently, the nature of the overhangs in the junctions is not crucial for DNA binding, whereas the nuclease activity of XPG requires the presence of a single-stranded or no 5’ overhang and is blocked by dsDNA immediately adjacent to the junction.

DISCUSSION

As part of our studies on the structure-specific endonucleases involved in NER (43), we were interested in characterizing the determinants of XPG interaction with its DNA substrates using binding, cleavage, and footprinting techniques. XPG is a member of the Fen1 family of structure-specific endonucleases, which also includes Fen1 and Exo1. Consistent with their contribution to different repair pathways, these two proteins and XPG show distinct but overlapping substrate specificities and cleave ss/dsDNA junctions in the dsDNA region adjacent to a 5’ ssDNA overhang (2, 31, 32).

Junction-specific DNA Contacts of XPG—We established conditions that allowed the stable binding of XPG to its substrates by preventing incision either by omitting the metal cofactor or by using the E791A mutant of XPG, which displays greatly reduced nuclease activity (26). Identical interference patterns were observed in our footprinting studies using WT XPG in the absence (Fig. 4) or the E791A mutant in the presence of the metal cofactor (data not shown), indicating that the metal ion does not significantly influence substrate binding prior to incision. We used several protecting footprinting techniques with XPG-DNA complexes, including exonuclease III,
DNase I, hydroxyl radical, or copper phenanthroline footprinting. In these techniques, the region of the DNA bound by the protein is protected from degradation, and a fundamentally kinetic parameter, the effect of the protein bound to DNA on the rate of attack of the footprinting reagent, is observed. Except for the studies with exonuclease III (Fig. 3), we failed to observe a clear footprint with any of these methods (data not shown), which might be due to fast on and off rates of XPG binding to DNA. We also used ethylation interference footprinting, a technique that measures the effect of ethylation of individual phosphodiester bonds in terms of DNA binding. In contrast to protection footprinting techniques, with interference methods the effect of DNA modifications on the equilibrium binding strength of a protein is observed. Interference footprinting is thus a thermodynamically rather than kinetically controlled experiment (44).

The interface of XPG binding to DNA revealed by these studies is summarized in Fig. 8. XPG makes phosphate contacts mainly in the double-stranded region of the DNA on both strands spanning about 12 nucleotides. The strongest phosphate contacts mapped to one face of the B-DNA helix, suggesting that XPG interacts preferentially with the splayed arm substrate on the side where the 5′ single-stranded overhang protrudes from the junction.

XPG makes three additional phosphate contacts in the 5′ ssDNA overhang adjacent to the junction on the same face as the contacts to the dsDNA, which have however only a minor influence on binding affinity. Footprinting studies of exonuclease I and DNA polymerase I revealed similar modes of binding, with main contacts being made on one face of the dsDNA and less extensive contacts on the ssDNA overhangs. The orientation of binding of exonuclease I and DNA polymerase I appears to be different, because it makes contacts on the 3′ ssDNA overhang and that face of the helix (34, 35). Overall it appears that XPG and the other Fen1 family nucleases interact in a similar way with the double-stranded portion of a junction containing substrate through multiple nonspecific phosphodiester backbone contacts.

An intriguing observation from the footprinting studies is that although XPG is highly specific for binding to ss/dsDNA junctions, not many contacts to the ssDNA overhangs were observed. These observations were further emphasized when we investigated how far the ssDNA overhangs could be truncated without losing cleavage and binding activity of XPG. Although deletion of the 5′ overhang only reduced XPG incision activity by 50%, at least two nucleotides of the 3′ overhang were required to retain substantial XPG binding and cleavage activity. Similar observations were made for Fen1 and the exonuclease domain of DNA polymerase I, where shortening of the 5′ overhang only had a minor effect on the activity of the two enzymes (35, 45). Because we failed to observe phosphate contacts to the 3′ overhang in our footprinting assays, it will be of interest to determine what contacts are responsible for the specific binding to the ss/dsDNA junctions at the molecular level. XPG might recognize distortion or bending of the DNA and the 3′ protruding overhang, a technique that measures the effect of ethylation of individual phosphodiester bonds in terms of DNA binding. The nature of the 3′ overhang is more critical; 5′ dsDNA overhangs are more critical; 5′ dsDNA overhangs strongly interfere with XPG cleavage activity, regardless of the nature of the 3′ overhang. The inhibition of cleavage of a 5′ dsDNA overhang by XPG is in agreement with a study of cleavage activity of XPG on model substrates of transcription-induced immunoglobulin switch regions, in which XPG cleaved substrates with a ssDNA overhang, but not dsDNA 5′ overhangs (46). Substrates in which the 5′ dsDNA overhang is recessed to yield 3 nt of ssDNA adjacent to the junction are efficiently processed by XPG, indicating that the substrates containing 5′ dsDNA overhangs likely inhibit XPG activity by restricting accessibility and/or flexibility of the substrate at the junction. Because 5′ dsDNA overhangs do not interfere with substrate binding by XPG, this accessibility/flexibility appears to be only important for catalysis but not binding. Particularly instructive in this respect are three-way junction substrates, which are bound severalfold more efficiently than splayed arm substrates by XPG but are completely resistant to cleavage. It was shown that cationic surfaces of proteins can induce substantial DNA bending by neutralizing phosphates on one face of the helix (47). Because we have found that also XPG binds to the phosphodiester backbone on one face of the helix, it is possible that substrates like three-way junctions that already contain a bent backbone favor the XPG-substrate interaction. XPG and other structure-specific nucleases might recognize ss/dsDNA junctions by sensing the bending of the junction. The flexibility/accessibility of the junction could be a necessary second step for rendering the enzyme catalytically competent. DNA binding of XPG to specific DNA structures may be important for the additional roles of the protein in transcription-coupled base excision repair and transcription that do not require its endonuclease activity (4–7). Therefore, the binding of XPG to three-way junctions raises the question about the potential physiological significance of this activity. Three-way junctions are possible intermediates in homologous recombination and replication (48). Although XPG has not yet been directly implicated in these processes, it is tempting to speculate that the severe phenotype of certain XPG patients and XPG knockout mice might partially stem from a special role of XPG in binding distorted DNA structures that are intermediates in transcription, replication, or recombination (49).

Further studies are required to determine in detail how XPG and other Fen1 family members recognize ss/dsDNA junctions. A mechanistically particularly intriguing question is why certain substrates are bound but not cleaved by XPG. Understanding this difference will be crucial for elucidating the structural and catalytic roles of XPG.

Acknowledgments—We are grateful to Rolf Buff for providing the bubble substrate, to Ludovic Gillet, Jacqueline Enzlin, and Rachel Jossen for thoughtful discussions, and to Oliver Fleck, Pavel Janscak, Josef Jiricny, and Primo Schar for helpful comments on the manuscript.

REFERENCES
1. Suck, D. (1997) Biopolymers 44, 405–421
2. Lieber, M. R. (1997) Bioessays 19, 233–240
3. O’Donovan, A., Davies, A. A., Maggs, J. G., West, S. C., and Wood, R. D. (1994) Nature 371, 432–435
4. Klungland, A., Hess, M., Gunz, D., Constantinoiu, A., Clarksson, S. G., Doetsch, P. W., Bolton, P. H., Wood, R. D., and Lindahl, T. (1999) Mol. Cell 3, 33–42
5. Le Page, F., Kwoh, E. E., Avrutskaya, A., Gentil, A., Leandon, S. A., Sarasin, A., and Cooper, P. K. (2005) Cell 121, 159–171
6. Shiomi, N., Hayashi, E., Sasanuma, S., Mita, K., and Shiomi, T. (2001) Mutat. Res. 487, 127–135
7. Lee, S. K., Yu, S. L., Prakash, L., and Prakash, S. (2002) Cell 109, 823–834
8. Stary, A., and Sarasin, A. (2002) Biochimie (Paris) 84, 40–49
9. van Gool, A. J., van der Horst, G. T., Citterio, E., and Hoeijmakers, J. H. (1997) EMBO J. 16, 4155–4162
10. Nosuppel, T., Lalle, P., Leandon, S. A., Cooper, P. K., and Clarksson, S. G. (1997) Proc. Natl. Acad. Sci. U. S. A. 94, 3116–3121
11. Friedberg, E. C., Walker, G. C., and Siede, W. (1995) DNA Repair and Mutagenesis, American Society for Microbiology, Washington D.C.
12. Svejstrup, J. Q. (2002) Proc. Natl. Acad. Sci. U. S. A. 99, 3116–3121
13. de Laat, W. L., Jaspers, N. G., and Hoeijmakers, J. H. (1999) Genes Dev. 13, 768–785
14. Friedberg, E. C. (2001) Nat. Rev. Cancer 1, 22–33
15. Volker, M., Mone, M. J., Karmakar, P., van Hofven, A., Schul, W., Vermeulen, W., Hoeijmakers, J. H., van Driel, R., van Zeeland, A. A., and Mullenders,
Structural Determinants for Substrate Binding and Catalysis by the Structure-specific Endonuclease XPG
Marcel Hohl, Fabrizio Thorel, Stuart G. Clarkson and Orlando D. Schärer

J. Biol. Chem. 2003, 278:19500-19508.
doi: 10.1074/jbc.M213155200 originally published online March 18, 2003

Access the most updated version of this article at doi: 10.1074/jbc.M213155200

Alerts:
• When this article is cited
• When a correction for this article is posted

Click here to choose from all of JBC’s e-mail alerts

This article cites 46 references, 26 of which can be accessed free at http://www.jbc.org/content/278/21/19500.full.html#ref-list-1