Effects of DNA Adduct Structure and Sequence Context on Strand Opening of Repair Intermediates and Incision by UvrABC Nuclease†

Yue Zou,*,‡ Steven M. Shell,‡ Christopher D. Utzat,§ Charlie Luo,‡ Zhengguan Yang,‡ Nicholas E. Geacintov,|| and Ashis K. Basu§

Department of Biochemistry and Molecular Biology, James H. Quillen College of Medicine, East Tennessee State University, Johnson City, Tennessee 37614, Department of Chemistry, University of Connecticut, Storrs, Connecticut 06269, and Department of Chemistry, New York University, New York, New York 10003

Received March 19, 2003; Revised Manuscript Received July 31, 2003

ABSTRACT: DNA damage recognition of nucleotide excision repair (NER) in *Escherichia coli* is achieved by at least two steps. In the first step, a helical distortion is recognized, which leads to a strand opening at the lesion site. The second step involves the recognition of the type of chemical modification in the single-stranded region of DNA during the processing of the lesions by UvrABC. In the current work, by comparing the efficiencies of UvrABC incision of several types of different DNA adducts, we show that the size and position of the strand opening are dependent on the type of DNA adducts. Optimal incision efficiency for the C8-guanine adducts of 2-aminofluorene (AF) and *N*-acetyl-2-aminofluorene (AAF) was observed in a bubble of three mismatched nucleotides, whereas the same for C8-guanine adduct of 1-nitropyrene and N^2-guanine adducts of benzo[*a*]pyrene diol epoxide (BPDE) was noted in a bubble of six mismatched nucleotides. This suggests that the size of the aromatic ring system of the adduct might influence the extent and number of bases associated with the opened strand region catalyzed by UvrABC. We also showed that the incision efficiency of the AF or AAF adduct was affected by the neighboring DNA sequence context, which, in turn, was the result of differential binding of UvrA to the substrates. The sequence context effect on both incision and binding disappeared when a bubble structure of three bases was introduced at the adduct site. We therefore propose that these effects relate to the initial step of damage recognition of DNA structural distortion. The structure−function relationships in the recognition of the DNA lesions, based on our results, have been discussed.

Nucleotide excision repair (NER),[1] as one of the primary DNA repair pathways in cells, is capable of removing an extensive variety of bulky lesions, induced by chemicals and radiation, with varying efficiencies (*1*−*3*). Its ability to recognize and excise such a broad repertoire of substrates has been the subject of intensive research, and it is generally believed that the success of NER depends on efficient damage recognition. The UvrABC nuclease system, which initiates the NER in *Escherichia coli*, represents a paradigm for understanding the general mechanism of DNA damage recognition and incision (*1*). This model system has been widely used to study the interactions of NER with various types of DNA damages. Briefly, the enzymatic action of the UvrABC system begins with the dimerization of UvrA. The $UvrA_2$ dimer interacts with UvrB to form an $UvrA_2B$ complex, which binds to the damaged DNA site (*4*). Upon binding of $UvrA_2B$ to the damage, the $UvrA_2$ protein dissociates from the DNA, leading to the formation of a stable UvrB−DNA complex (*5*). Recent studies suggested that two UvrB molecules participated in the formation of the UvrAB complex (*6*, *7*). After dissociation of $UvrA_2$, UvrC is recruited by the interaction with UvrB (*8*), resulting in the formation of the UvrBC−DNA pre-incision intermediate. Formation of this structure-specific intermediate is believed to trigger 3′ incision followed immediately by 5′ incision by the endonuclease activity of UvrC (*9*) resulting in cleavage of the phosphodiester bonds four to seven phosphates 3′ and eight phosphates 5′ to the damaged residue, respectively (*1*, *10*).

Our recent studies indicated that the DNA damage recognition in *E. coli* NER occurs through a sequential two-step mechanism in which $UvrA_2$ recognizes the adduct-induced disruption of the Watson−Crick DNA structure at the initial step, while, following a strand opening, UvrB recognizes the type of nucleotide modification (*12*). The strand opening is driven by the helicase activity of $UvrA_2B$ (*11*), which plays a key role in this process by facilitating the transition from the first to the second step of recognition. Our results further suggest that this strand opening leads to the dissociation of $UvrA_2$ and the formation of a stable

† This study was supported by NCI Grant CA86927 (Y.Z.) and NIEHS Grants ES09127 and ES00318 (A.K.B.).

* To whom correspondence should be addressed. Phone: (423) 439-2124. Fax: (423) 439-2030. E-mail: zouy@etsu.edu.

‡ East Tennessee State University.

§ University of Connecticut.

|| New York University.

[1] Abbreviations: NER, nucleotide excision repair; BPDE, benzo[*a*]pyrene diol epoxide or 7,8-dihydroxy-9,10-epoxy-7,8,9,10-tetrahydrobenzo[*a*]pyrene; 1-NP, 1-nitropyrene; C8-AP-dG, *N*-(deoxyguanosin-8-yl)-1-aminopyrene; AF, 2-aminofluorene; C8-AF-dG, *N*-(deoxyguanosin-8-yl)-2-aminofluorene; AAF, *N*-acetyl-2-aminofluorene; C8-AAF-dG, *N*-(deoxyguanosin-8-yl)-*N*-acetyl-2-aminofluorene; DTT, dithiothreitol; EDTA, ethylenediaminetetraacetic acid.

10.1021/bi034446e CCC: $25.00
Published on Web 10/08/2003

UvrB−DNA complex. The opening is also necessary for the 5′-incision by UvrABC (*11*). However, it is unclear if the dimension of this strand opening driven by $UvrA_2B$ depends on the type of DNA lesions. The effects of DNA sequence context on the damage recognition and incision by UvrABC are also not known. A major objective of the current work is to address these questions.

Recent studies of solution structures of BPDE-, AF-, AAF-, and AP-DNA adducts by NMR spectroscopy indicate that the C8-AAF-dG, C8-AP-dG, and N^2-(+)-*cis*-BPDE-dG adducts in DNA duplexes adopt a conformation in which the AF, AP, and BPDE rings, respectively, are intercalated into the DNA helix and stack between the neighboring base pairs. The modified guanine of the C8-AAF-dG and C8-AP-dG is displaced into the major groove of DNA (*13*−*17*), while the modified guanine of the (+)-*cis*-BPDE-dG adduct and the cytosine residue on the complementary strand are displaced into the minor and major grooves, respectively (*18*). In contrast, the C8-AF-dG adduct equilibrates between AF-intercalated and AF-external (in the major groove) structures, although the external one exists as the predominant form (*19*−*22*). Similarly, the N^2-(+)-*trans*-BPDE-dG adduct lies in the minor groove with the benzo[*a*]pyrene moiety oriented in the 5′ direction of the adducted strand (*23*). The similarities and differences in chemical structure and induced conformation of these adducts provide us with a tool to study the potential effects of adduct on the DNA strand opening during damage recognition by UvrABC.

In this paper, we demonstrate that the number of unpaired bases and positioning of the strand-opened region relative to the site of the adduct depends on the type of DNA adduct. To our surprise, we found that the type of DNA distortion induced by the adduct had little influence in generating the optimal opened region catalyzed by UvrABC. Instead, the size and structure of the aromatic ring systems of the adduct molecules appear to play a role in determining the extent of the DNA strand opening during NER. We also found that the neighboring sequences influence the incision efficiency of the DNA adduct, presumably due to their effect on the efficiency of initial damage recognition by UvrA.

EXPERIMENTAL PROCEDURES

Chemicals. Tris base, boric acid, ethylenediaminetetraacetic acid (EDTA), and $CaCl_2$ were purchased from Sigma. Acrylamide, ammonium persulfate, *N*,*N*′-methylenebisacrylamide, and urea were obtained from National Diagnostics and Fisher Scientific. The [γ-^{32}P]ATP was bought from either New England Nuclear Inc. (DuPont) or Amersham Biosciences. All other chemicals were purchased from Fisher Scientific.

Purification of UvrA, UvrB, and UvrC Proteins. UvrA was purified from *E. coli* strain MH1 ΔUvrA containing the overproducing plasmid, pSST10 (graciously supplied by L. Grossman, Johns Hopkins University), in which the *uvrA* gene is under the control of the heat-inducible PL promoter. UvrB was purified in one step through a chitin column from *E. coli* strain XL-1 Blue transformed with the overexpressing plasmid pUTG97 containing the *uvrB* gene under the control of the IPTG-induced P_{tac} promoter as described previously (*11*). UvrC was overproduced from *E. coli* C41(DE3) cells (*25*) harboring plasmid pUTG98 containing the PCR-amplified *uvrC* gene, which was subcloned via *Nde* I and *Kpn* I restriction sites into the vector pTYB1 (IMPACT T7 system, New England Biolabs). The UvrC protein was also purified on a chitin column in one step following the same procedures as described previously for UvrB (*11*) except that 500 mM NaCl rather than 100 mM was used in the cleavage and elution buffers.

DNA Substrate Construction. The 50-bp oligonucleotides containing a single AF, AAF, AP, or BPDE adduct were prepared as described previously (*10*−*12*, *24*). Briefly, 30 pmol of phosphorylated 11mer d(CCATCG*CTACC) modified at G* with the C8 guanine adducts of AF and AAF and of AP or N2 guanine adducts of BPDE (*26*−*28*) were ligated with stoichiometric quantities of 20mer and phosphorylated 19mer (or none). This process used T4 DNA ligase in the presence of a 28mer bottom strand containing the middle part of the complementary sequence to top strands in a 30 μL solution containing 50 mM Tris-HCl, pH 7.8, 10 mM $MgCl_2$, 10 mM DTT, 1 mM ATP, and 50 μg/mL of BSA. The ligation was carried out at 16 °C for 12 h. After ligation, the products were purified and reannealed with various 50mer bottom strands to make appropriate substrates as shown in Figure 1B. The annealed substrates were purified on a nondenaturing 8% polyacrylamide gel. The double-stranded character and homogeneity of the 50-bp substrates were examined by a restriction assay (*10*) and analyzed on a 12% polyacrylamide sequencing gel under denaturing conditions with TBE as the running buffer (50 mM Tris-HCl, 50 mM boric acid, 1 mM EDTA, pH 8.0).

Gel Mobility Shift Assays. Binding of the UvrA protein to the DNA substrates was determined by gel mobility shift assays. Typically, the substrate (2 nM) was incubated with UvrA with varying concentrations as indicated at 37 °C for 15 min in 20 μL of UvrABC buffer (50 mM Tris-HCl, pH 7.5, 50 mM KCl, 10 mM $MgCl_2$, 5 mM DTT) in the presence of 1 mM ATP. After incubation, 2 μL of 80% (v/v) glycerol was added, and the mixture was immediately loaded onto a 3.5% native polyacrylamide gel in TBE running buffer and electrophoresed at room temperature. For an estimation of the dissociation constant for the UvrA interaction with a DNA adduct, radioactivity of the DNA bands on gel was quantified with a Fiji FLA-5000 image scanner. The binding isotherm was then generated, and the K_d was estimated from the UvrA titration concentration at which half of the DNA substrate molecules had been bound.

Incision Assays. The 5′-terminally labeled DNA substrates (2 nM) were incised by UvrABC (UvrA, 15 nM; UvrB, 250 nM; and UvrC, 100 nM) or UvrBC in the absence of UvrA in the UvrABC buffer (1 mM ATP) at 37 °C for a given period. The Uvr subunits were diluted and premixed into storage buffer before mixing with DNA. The reactions were terminated by adding EDTA (20 mM) or heating to 90 °C for 3 min. The samples were denatured with formamide and heated to 90 °C for 5 min and then quick-chilled on ice. The digested products were analyzed by electrophoresis on a 12% polyacrylamide sequencing gel under denaturing conditions with TBE buffer.

RESULTS

Construction of Adducted DNA Substrates. The structures of DNA adducts used in this study are presented in Figure

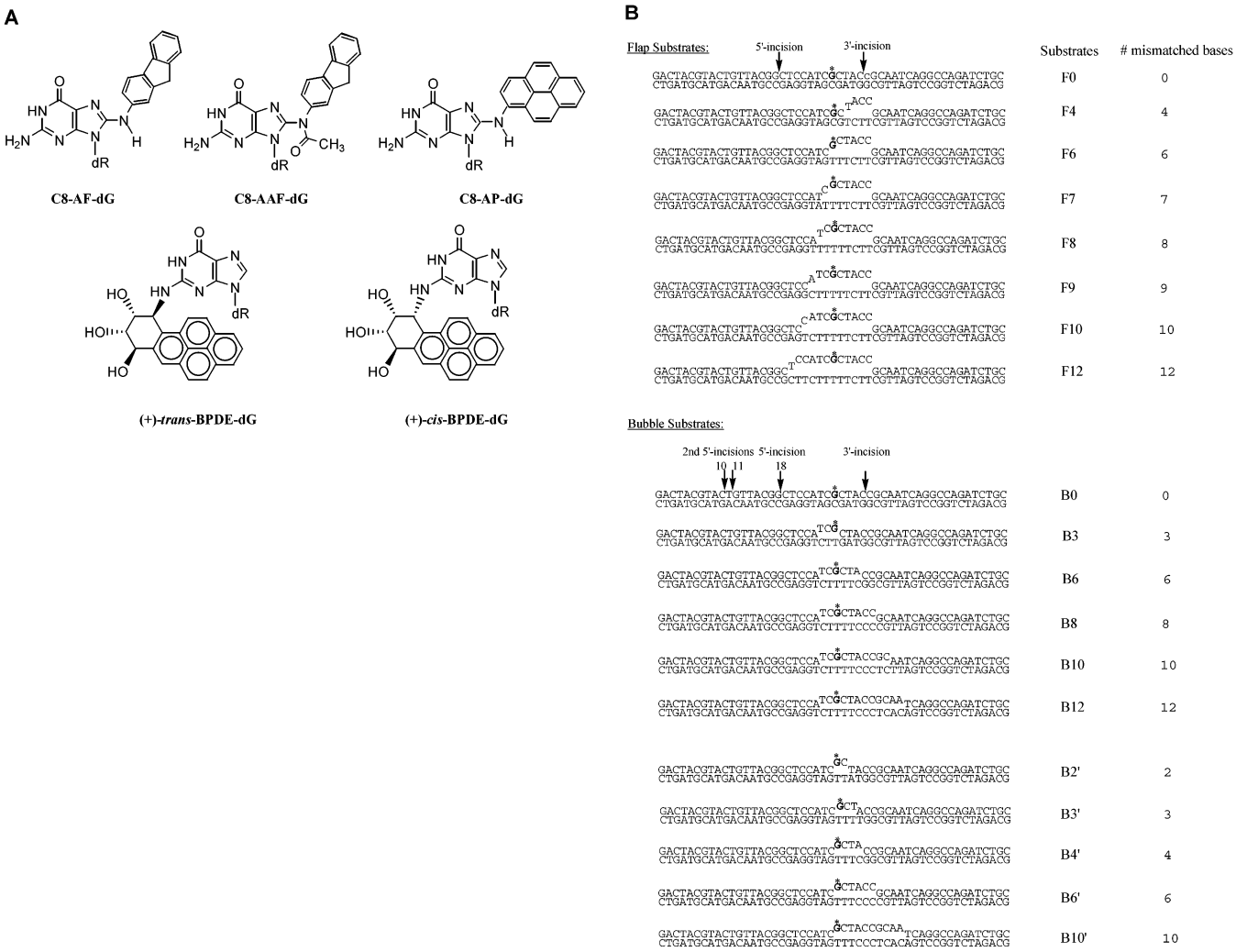

FIGURE 1: Chemical structure of DNA adducts (A) and the list of substrates (B) used in the present study. The bold G with an asterisk in the sequence (B) represents the adducted base, the arrows indicate the major incision sites, the lowercase base (C) in the top strand of the flap substrate F0 indicates there is a nick between the bases (CG), and the c represents a free 3′ end. Only the flap substrates contain an internal nick or free 3′ end, and the bubble substrates do not. The flap size increases in the direction of 3′ to 5′, while the bubble expands from 5′ to 3′.

1A. Each of these guanine adducts contains a covalent link to either the N^2 or C8 position of the base with an aromatic ring system of a different size and/or conformation. Each substrate was composed of a 50-bp oligodeoxynucleotide with a defined sequence containing the adduct in the middle of the top strand (Figure 1B) as reported earlier (*10*−*12*, *24*). There were two groups of substrates, namely, the flap and bubble substrates. The flap substrates have been used to examine the dimension of the strand opening formed during NER on the 5′-side of the adduct, while the purpose of the bubble substrates was to determine the extent of the opening on the 3′-side. The flap substrates contained a nick between the fifth and sixth bases 3′ to the adducted dG, and each one contained a string of mismatched bases 3′ to the adduct with a free 3′-end. For many substrates (F6−F12), additional mismatches were introduced at 5′ to the adduct site as shown in Figure 1B, representing the different degree of strand opening in the 5′ direction of the damage. The 3′-free end was made to mimic the 3′-cleavage (a prerequisite for 5′-cleavage to occur) made by UvrABC as these flap substrates were designed to characterize the strand opening on the 5′-side of the adduct, which may greatly influence the 5′-incision of the adduct.

Incision of Flap Substrates by UvrABC Nuclease. As shown in Figure 2, the UvrABC incisions of the flap substrates containing a specific DNA adduct were performed. The substrates were labeled at the 5′-end of the top strand. Two types of 5′-incision were observed: the normal 5′-incision between the seventh and eighth phosphodiester linkage and the second 5′-incisions between the 14th and 15th (or 15th and 16th) phosphodiester bond 5′ to the adduct. It has been previously reported that the second 5′-incision is coupled with the first 5′-incision (*12*), so that the second incisions can occur only after the primary incision. It is believed that this second 5′-incision might be the result of a damage-independent endonuclease activity of UvrABC or UvrBC that makes incisions at the seventh and eighth phosphodiester bonds 5′ to a nicked site of nondamaged DNA (just like the product of the first 5′-incision) (*39*, *40*). Since the top strand of flap substrates had been 5′-radiolabeled with ^{32}P, the second incision made the prior incision invisible in the gel. As a result, calculation of the overall products of normal 5′-incision required inclusion of the products from both types of 5′-incisions (*12*). It was clear that the incision efficiency varied with the size of the opened strand region for each adducted DNA. In the case of the

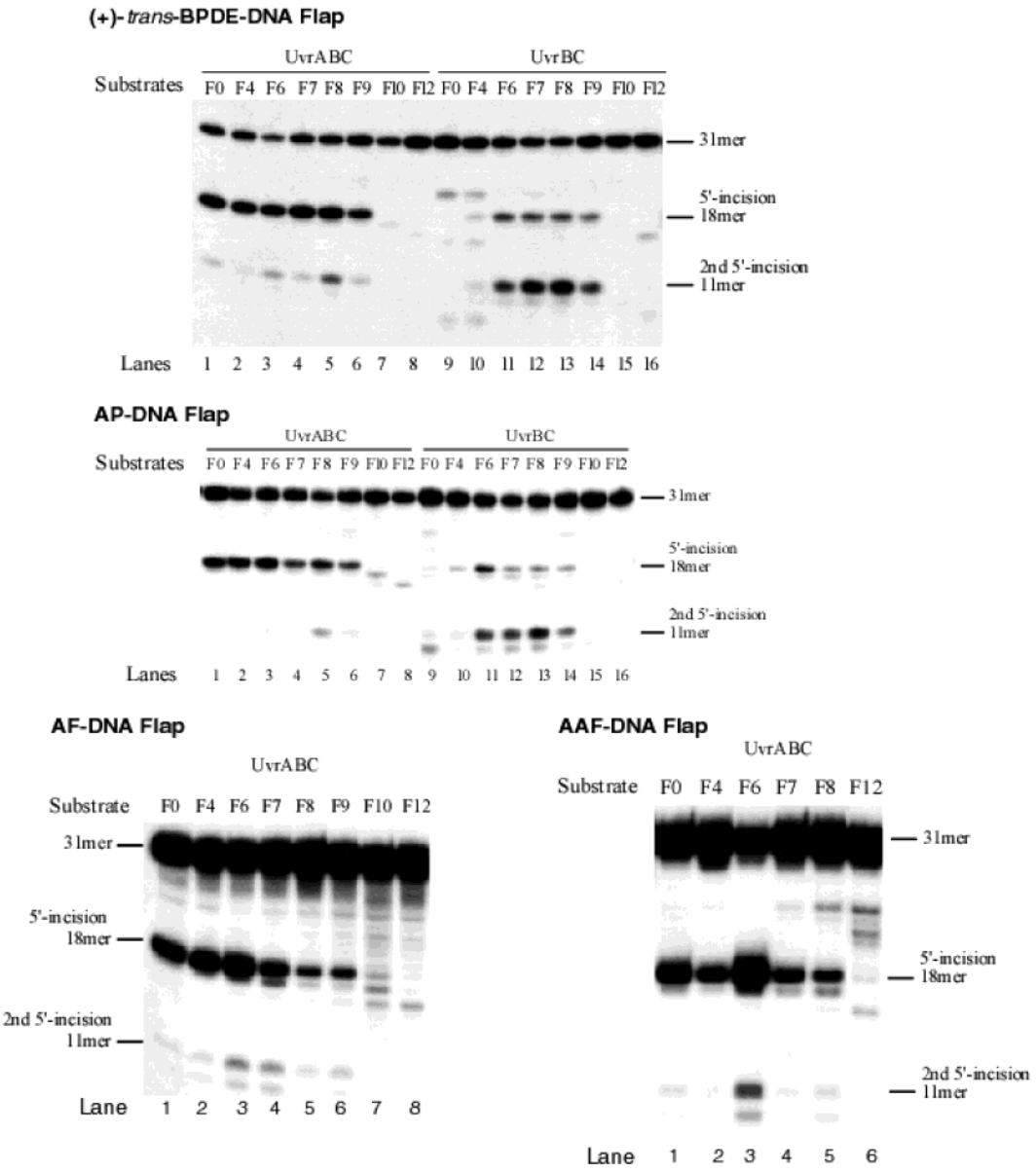

FIGURE 2: Incisions of DNA adducts of flap substrate by UvrABC and UvrBC nucleases. The 5′-terminally labeled flap substrates were incubated with UvrABC (UvrA, 15 nM; UvrB, 250 nM; and UvrC, 100 nM) or UvrBC in the UvrABC buffer at 37 °C for 30 min. The incision products were analyzed on a 12% polyacrylamide sequencing gel. The 31mer represents the intact top strand DNA (labeled), while the 18mer and 10−11mer are the products of normal 5′-incision and second 5′-incisions, respectively. The second 5′-incisions are coupled with and dependent on the normal 5′-incision so that both types of incision products should be combined for determination of incision efficiency. The four different types of adducts, (+)-*trans*-BPDE-, AP-, AF-, and AAF-DNA were assayed under the same conditions.

(+)-*trans*-BPDE adduct, the incision efficiency did not increase for the substrates of F0−F8 as the double-/single-stranded DNA (ds-ssDNA) junction moved away from the 3′-nick up to a flap size of eight bases (F8). But efficiency of the incision decreased significantly as the size of the strand opening increased to nine or more bases (F9−F12). Interestingly, there was a one or two nucleotide shift of the 5′-incision site further away from the adduct for F10 or larger flaps (Figure 2). Little or no incision was observed for the flap substrate with 12 unpaired bases (F12). The same experiments were also conducted for other DNA adducts including AP-, AF-, and AAF-DNA adducts (Figure 2). Since the $UvrA_2B$ protein complex contains DNA helicase activity (*11*), as expected, the substrates with small flap structures were incised efficiently, and a change of the number of unpaired bases for substrates F0−F8 (for (+)-*trans*-BPDE) resulted in relatively little effect on incision efficiency (*11*). We conclude that the $UvrA_2B$ opens up the DNA helix 5′ to the adduct to an extent similar to one of the defined structures in the substrates. However, the incision became very inefficient as the strand opening became larger than the intermediate structure required for UvrABC incision or larger than the opening induced by $UvrA_2B$ during damage recognition of NER.

In contrast to the incision by UvrABC, the incision of the flap substrates by UvrBC in the absence of UvrA showed different patterns. The main difference was noted for the substrates with small flap structures that were not incised, or at best, incised with low efficiencies. This is consistent with the observation that UvrB contains a helicase activity that can only be activated in the presence of UvrA as in the form of $UvrA_2B$ (*11*, *36*−*38*). Because of the lack of helicase

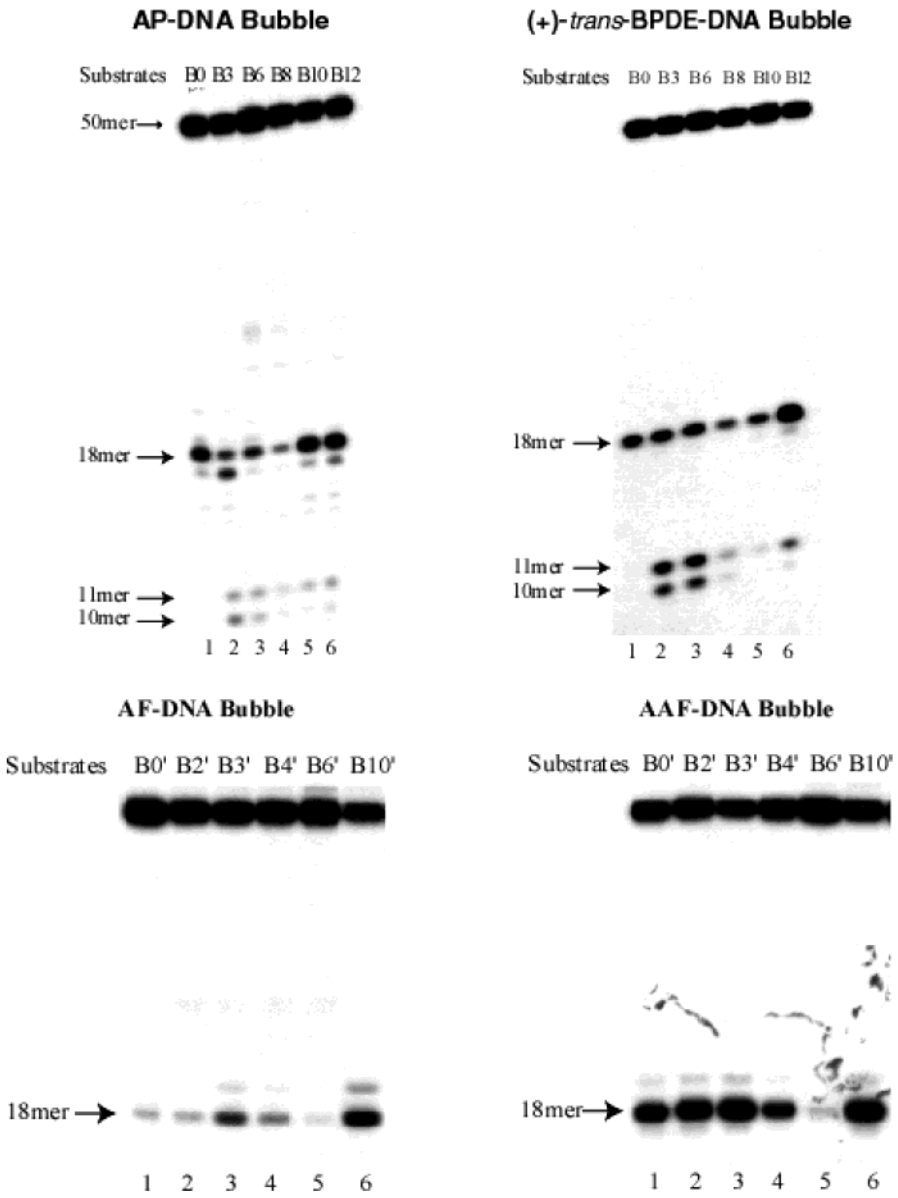

FIGURE 3: Incisions of DNA adducts of bubble substrate by UvrABC nucleases. The 5′-terminally labeled DNA substrates with varying bubble sizes were incubated with UvrABC (UvrA, 15 nM; UvrB, 250 nM; and UvrC, 100 nM) in the UvrABC buffer at 37 °C for 30 min. The incision products were analyzed on a 12% polyacrylamide sequencing gel. The 50mer represents the intact DNA substrate. The 18mer and 11mer stand for the normal 5′- and the second 5′-incision products, respectively. The second 5′-incisions are coupled with and dependent on the normal 5′-incision so that both types of incision products should be combined for determination of incision efficiency. The four different types of adducts, (+)-*trans*-BPDE-, AP-, AF-, and AAF-DNA were assayed under the same conditions.

activity in the absence of UvrA, the small flap substrates could not be opened up to the structure required for efficient incisions by UvrBC. As shown in Figure 2, maximum incision efficiency was achieved for the substrate F8 with the (+)-*trans*-BPDE adduct that represents the ideal flap-substrates for the UvrBC incision. The data presented in Figure 2 suggest the preferred structure of strand opening for the optimal 5′-incision, and this structure varied with the types of DNA adducts.

Incision of Bubble Substrates by UvrABC Nuclease. The experiments with the flap DNA substrates defined the intermediate DNA structure at the 5′-side of the adduct formed in the pre-incision complex. To obtain a comprehensive picture of the complex, we have examined the formation of strand opening at the 3′-side of the adduct using a set of DNA bubble substrates with varying sizes (Figure 1B). These DNA bubble substrates were constructed on the basis of information gained from the experiments with flap substrates. Specifically, for (+)-*trans*-BPDE- and AP-DNA adducts, substrates were built with bubble starting from the second base 5′ to the adduct, while for AF and AAF adducts, the substrates were constructed with the duplex opened from the adducted guanine. The size of the bubbles varied from zero to 12 unpaired bases on the 3′-side of these lesions.

Similar to the case with flap substrates, efficient incisions occurred when the bubble size varied from zero to six (B0−B6) for (+)-*trans*-BPDE- and AP-DNA adducts and from zero to three (B0′−B3′) for AF and AAF adducts (Figure 3). Again, the total incision products include both from the normal 5′-incision (18mer) and the second 5′-incision (11mer and 10mer), although little or no second 5′-incision was observed for the AF and AAF adducts. As shown in Figure 3, the incision efficiency dropped dramatically as the substrates were further opened to a bubble size of eight

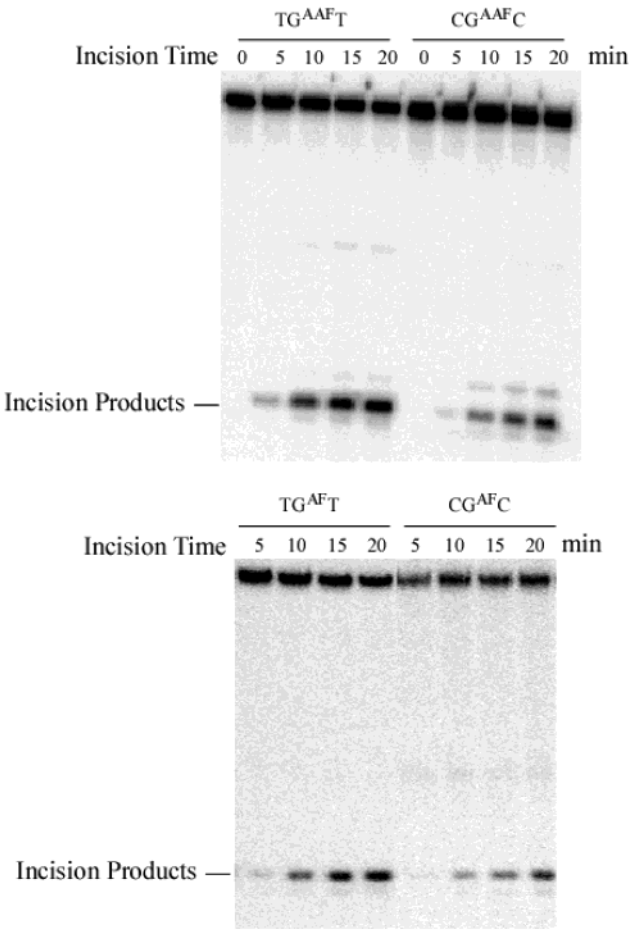

FIGURE 4: Effect of DNA sequence context at adducts on the incision efficiency of UvrABC. The 5′-terminally labeled DNA substrates adducted with AF and AAF in TG*T and CG*C sequences were incubated with UvrABC (UvrA, 15 nM; UvrB, 250 nM; and UvrC, 100 nM) in the UvrABC buffer at 37 °C for the indicated periods. The incision products were analyzed on a 12% polyacrylamide urea denatureing gel. Incision efficiency or rate was determined for each of the substrates as the slope of a linear regression line of the kinetic data.

unpaired bases, B8 (for (+)-*trans*-BPDE- and AP), and of four to six unpaired bases, B4′−B6′ (for AF and AAF). The drop in incision activity implies that the size of the artificial bubble was larger than that of the bubble structure generated during the normal processing of the lesion by UvrABC. It is noteworthy, however, that a further larger bubble structure, such as B10−B12 or B10′, led to restoration of the incision efficiency. These unusual incisions have been noted in other studies (*11*) and are believed to be due to uncoupled 5′-incision without a prior 3′-incision triggered by the extensive strand opening 3′ to the adduct. The incisions of these bubble substrates by UvrBC in the absence of UvrA were also conducted, and the results were similar (data not shown).

UvrABC Incision of Adducts with Different Neighboring DNA Sequences. To address if the sequence context changes incision efficiency of an adduct, two types of substrates (50 bp) with different neighboring sequences of TG*T and CG*C were constructed, where G* is the modified nucleotide with AF or AAF (Figure 1). The sequences of CG*C and TG*T substrates are identical except for the immediate 5′ and 3′ neighbor of the adducted guanine. The UvrABC incisions on these substrates were investigated. As shown in Figure 4 and Table 1, the substrate with either AF or AAF adduct in the TG*T sequence was incised 70% more efficiently by UvrABC than in the CG*C sequence. These results suggest that the incision efficiency also depends on the neighboring DNA sequence of the lesion. Similar results were observed with BPDE adducts (*30*).

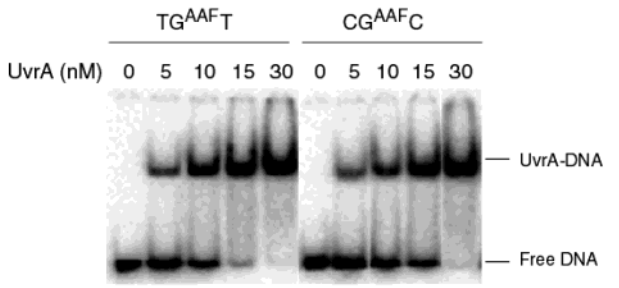

FIGURE 5: Binding of UvrA protein to the AAF-DNA substrates in TG*T and CG*C sequences. The 5′-terminally labeled DNA substrates were incubated with varying concentrations of UvrA at 37 °C for 15 min in the UvrABC buffer without ATP. The binding products were analyzed on a 3.5% native polyacrylamide gel. Free DNA represents the unbound DNA substrates, and the UvrA−DNA represents the protein−DNA complex formation.

Table 1: UvrABC Incision Rates of AF- and AAF-DNA Adducts in TG*T and CG*C Sequences

sequence context	% DNA incised per min	SD
TG(AF)T	1.21	0.12
CG(AF)C	0.71	0.03
TG(AAF)T	1.76	0.19
CG(AAF)C	1.02	0.03

Binding of UvrA to Adducts with Different Neighboring DNA Sequences. Since the sequential two-step mechanism dominates the process of damage recognition in *E. coli* NER (*12*), we explored how the local sequence of an adduct influences the recognition process that ultimately changes the efficiency of the incision. We have examined the interactions of UvrA with the AAF adduct in the TG*T and CG*C sequences using a gel mobility shift assay (Figure 5). The slower migrated bands in the assay represented the formation of the $UvrA_2$−DNA complex, while the faster migrating bands represent free DNA. Similar to the incision results, the binding of UvrA to the adducted TG*T substrate was better than to the CG*C substrate. For the AAF-DNA adduct, the estimated dissociation constant, K_d, was about 12.5 and 8 nM for the CG*C and TG*T substrates, respectively, exhibiting 50% more efficient binding to the TG*T substrate. The data are in agreement with the results from incision experiments, suggesting that the DNA sequence context, which in turn influences the helix distortion by the adduct, was recognized by $UvrA_2$ or $UvrA_2B$ at the very first step of recognition.

Incision of AAF-DNA Bubble Substrates with TGT and CGC Neighboring Sequences. To confirm the results presented previously, the 50-bp substrates containing an AAF-DNA adduct in the sequences of TG*T and CG*C with a bubble structure were constructed, which was subjected to the UvrABC incision and the binding by the UvrA protein. On the basis of the results from the first part of this study (Figure 3), the bubble structure contained a three-base strand opening from the modified nucleotide to the second base 3′ to the adduct (B3′ in Figure 1B). The data presented in the

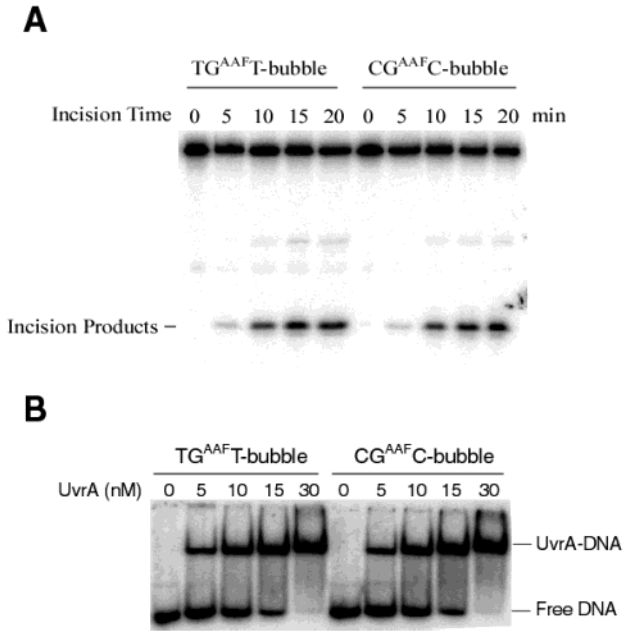

FIGURE 6: Incision and binding of AAF-DNA bubble substrates. (A) The AAF-DNA bubble substrates (B3′, Figure 1B) with the adduct sequences TG*T and CG*C were incised by UvrABC nuclease. The 5′-terminally labeled substrates were incubated with UvrABC (UvrA, 15 nM; UvrB, 250 nM; and UvrC, 100 nM) in the UvrABC buffer at 37 °C for various periods. The incision products were analyzed on a 12% polyacrylamide sequencing gel. (B) Binding of UvrA to the same bubble substrates. The 5′-terminally labeled DNA substrates were incubated with varying concentrations of UvrA at 37 °C for 15 min in the UvrABC buffer without ATP. The binding products were analyzed on a 3.5% native polyacrylamide gel.

Figure 6A demonstrated that, different from the case of the nonbubble substrates, the bubble DNA substrates with the TG*T and CG*C sequences were incised equally well by the UvrABC. Similar results were also obtained from the binding assay in which the UvrA also bound equally well to the bubble substrates with the two different sequences (Figure 6B). It is evident, therefore, that the neighboring DNA sequence exerted its effects on NER by directly influencing the extent of DNA structural alteration or distortion induced by the adduct (Figures 5 and 6). We conclude that the alteration was recognized by UvrA at the initial step of damage recognition (*12*). By contrast, interaction of the neighboring base with the repair proteins may play little or no role in these effects.

DISCUSSSION

In *E. coli* nucleotide excision repair, DNA strand opening or unwinding carried out by $UvrA_2B$ is an indispensable step of damage recognition, which is followed by recognition of the chemical modifications at the second step (*11*, *12*). The strand opening overrides the local duplex distortions induced by the adduct and allows for the direct access of the repair protein to the adducted molecule (*11*, *12*). Unlike other DNA repair pathways, the strand opening for recognition is a unique characteristic of NER. In the present study, we have systematically determined the effects of DNA adducts on strand opening mediated by $UvrA_2B$ using a series of structure-specific DNA substrates to establish the relationships between the types of adducts and the dimensions of strand opening. The influence of the DNA sequence context of the adduct on damage recognition and incision was also examined.

Our results indicated that both the size and the position of the strand opening produced by $UvrA_2B$ around a lesion varied with the type of DNA adduct. Specifically, the action of $UvrA_2B$ or UvrABC on (+)-*trans*-BPDE- and AP-DNA adducts shows optimal recognition in a DNA bubble structure of six bases. This structure contained two unpaired bases 5′ and three unpaired bases 3′ to the adduct. The same structure was also optimal for recognition of the (+)-*cis*-BPDE-DNA adduct (*11*). By contrast, a three-base strand opening was preferred for the AF- and AAF-DNA adducts. This three-base strand separation was positioned from the modified nucleotide through the second base 3′ to the adduct. The result for AAF is generally consistent with that of a previous study (*29*). Taken together, our results demonstrated the dependence of strand opening on the type of DNA adduct during DNA damage recognition.

Although the AF- and AAF-DNA adducts are structurally distinct as the former resides in the major groove while the latter intercalates in DNA, and thus recognized and incised by UvrABC with significantly different efficiencies (*24*), the dimension of the strand opening produced by $UvrA_2B$ for both adducts was the same. This was also true for the stereoisomeric adducts of (+)-*trans*-BPDE- and (+)-*cis*-BPDE-DNA, which may be classified as the minor groove conformer and intercalator, respectively (*11*, *12*). The same dimension of strand opening was observed for both adducts, even though the incision of the intercalating stereoisomer was better than that of the minor groove conformer. These observations suggested that the strand opening was not directly related to the DNA distortion induced by the adducts and that there was no direct correlation between the dimension of the opening and incision efficiency. This is consistent with the assumption that the adduct-induced DNA distortion is recognized at the initial step prior to strand opening. Further evidence for this mechanism was obtained from the result that the (+)-*trans*-BPDE and AP adducts exhibited the same strand opening during NER, despite many differences in their structural and conformational features. All of these results indicated that although the incision strongly depends on the adduct structures, the dependence of strand opening on the structure of adducts occurred in a different way and seemed to be only related to the second step of the damage recognition. Why do adducts containing a pyrene ring system require six-base strand opening as compared to three-base opening for adducts with a fluorene ring? We hypothesize that efficient damage recognition at the second step requires larger strand opening as the size of the aromatic ring system increases to allow access of UvrB (or UvrBC) to the chemical modification. At this time, there are no additional data supporting this model, and testing of this hypothesis would require appropriate structural studies and/or examination of many adducts containing different aromatic ring systems, which are beyond the scope of this work.

DNA sequence context is another important aspect with respect to damage recognition by NER (*31*−*35*). Our preliminary investigation showed that the local DNA sequence exhibited an effect both on the recognition and on the incision. We used AF- and AAF-adducts in this study because their effects on DNA are different, but they are

similar with respect to the strand opening and the second step of damage recognition (*24*). Interestingly, the two adducts in the sequence TG*T were always incised more efficiently than in the sequence CG*C, and the difference appeared to be greater with AAF than with AF. Similar results were obtained with the BPDE adducts by Ruan et al. (*30*). Further study using a gel mobility shift assay showed that the UvrA protein had a higher affinity for the DNA in which the AAF adduct was located in a TG*T relative to the CG*C sequence. Since TG*T/ACA should have a lower melting temperature than CG*C/GCG, as shown in the case of (+)-*trans*-BPDE adducts (*35*), it is conceivable that more DNA distortion may be induced by the AAF adduct in the TG*T sequence than in the CG*C sequence. We believe that the sequence dependent effects are eliminated at the second step of damage recognition due to the bubble structure formation, which was demonstrated in our study with the AAF-DNA bubble substrates containing different sequence contexts.

In summary, DNA strand opening, a crucial step in DNA damage recognition and incision of NER, varies in size and dimension depending on the chemical structures of the DNA adducts, and the size of the aromatic ring systems seems very important. The DNA sequence context also influences the incision efficiency of UvrABC. This effect is related to the extent of the DNA helical alteration induced by the adduct and thus depends on the first step of damage recognition.

REFERENCES

1. Van Houten, B. (1990) *Microbiol. Rev. 54*, 18−51.
2. Sancar, A. (1996) *Annu. Rev. Biochem 65*, 43−81.
3. Lindahl, T., and Wood, R. D. (1999) *Science 286*, 1897−1905.
4. Orren, D. K., and Sancar, A. (1990) *J. Biol. Chem. 265*, 15796−15803.
5. Shi, Q., Thresher, R., Sancar, A., and Griffith, J. (1992) *J. Mol. Biol. 226*, 425−432.
6. Hildebrand, E. L., and Grossman, L. (1999) *J. Biol. Chem. 274*, 27885−27890.
7. Verhoeven, E. E., Wyman, C., Moolenaar, G. F., and Goosen, N. (2002) *EMBO J. 21*, 4196−4205.
8. Lin, J.-J., Phillips, A. M., Hearst, J. E., and Sancar, A. (1992) *J. Biol. Chem. 267*, 17693−17700.
9. Verhoeven, E. E. A., Kesteren, M. V., Moolenaar, G. F., Visse, R., and Goosen, N. (2000) *J. Biol. Chem. 275*, 5120−5123.
10. Zou, Y., Liu, T. M., Geacintov, N. E., and Van Houten, B. (1995) *Biochemistry 34*, 13582−13593.
11. Zou, Y., and Van Houten, B. (1999) *EMBO J. 18*, 4889−4901.
12. Zou, Y., Luo, C., and Geacintov, N. E. (2001) *Biochemistry 40*, 2923−2931.
13. O'Handley, S. F., Sanford, D. G., Xu, R., Lester, C. C., Hingerty, B. E., Broyde, S., and Krugh, T. R. (1993) *Biochemistry 32*, 2481−2497.
14. Milhe, C., Dhalluin, C., Fuchs, R. P., and Lefevre, J. F. (1994) *Nucleic Acid Res. 22*, 4646−4652.
15. Milhe, C., Fuchs, R. P., and Lefevre, J. F. (1996) *Eur. J. Biochem. 235*, 120−127.
16. Cho, B. P., and Zhou, L. (1999) *Biochemistry 38*, 7572−7583.
17. Mao, B., Vyas, R. R., Hingerty, B. E., Broyde, S., Basu, A. K., and Patel D. J. (1996) *Biochemistry 35*, 12659−12670.
18. Cosman, M., de los Santos, C., Fiala, R., Hingerty, B. E., Ibanez, V., Luna, E., Harvey, R., Geacintov, N. E., Broyde, S., and Patel, D. J. (1993) *Biochemistry 32*, 4145−4155.
19. Eckel, L. M., and Krugh, T. R. (1994) *Nature Struct. Biol. 1*, 89−94.
20. Cho, B. P., Beland, F. A., and Marques, M. M. (1994) *Biochemistry 33*, 1373−1384.
21. Mao, B., Hingerty, B. E., Broyde, S., and Patel, D. J. (1998) *Biochemistry 37*, 81−94.
22. Mao, B., Hingerty, B. E., Broyde, S., and Patel, D. J. (1998) *Biochemistry 37*, 95−106.
23. Cosman, M., de los Santos, C., Fiala, R., Hingerty, B. E., Singh, S. B., Ibanez, V., Margulis, L. A., Live, D., Geacintov, N. E., Broyde, S., and Patel, D. J. (1992) *Proc. Natl. Acad. Sci. U.S.A. 89*, 1914−1918.
24. Luo, C., Krishnasamy, R., Basu, A. K., and Zou, Y. (2000) *Nucleic Acid Res. 28*, 3719−3724.
25. Miroux, B., and Walker, J. E. (1996) *J. Mol. Biol. 260*, 289−298.
26. O'Handley, S. F., Sanford, D. G., Xu, R., Lester, C. C., Hingerty, B. E., Broyde, S., and Krugh, T. R. (1993) *Biochemistry 32*, 2481−2497.
27. Cho, B. P., Beland, F. A., and Marques, M. M. (1994) *Biochemistry 33*, 1373−1384.
28. Vyas, R. R., Nolan, S. J., and Basu, A. K. (1993) *Tetrahedron Lett. 34*, 2247−2250.
29. Gordienko, I., and Rupp, W. D. (1998) A specific 3′ exonuclease activity of UvrABC, *EMBO J. 17*, 626−633.
30. Ruan, Q., Kolbanovskiy, A., Skorvaga, M., Zou, Y., Lader, J., Malkani, B., Van Houten, B., and Geacintov, N. E. (2003), manuscript in preparation.
31. Xu, R., Mao, B., Amin, S., and Geacintov, N. E. (1998) *Biochemistry 37*, 769−778.
32. Fountain, M. A., and Krugh, T. R. (1995) *Biochemistry 34*, 3152−3161.
33. Lui, T., Xu, J., Tsao, H., Li, B., Xu, R., Yang, C., Amin, S., Moriya, M., and Geacintov, N. E. (1996) *Chem. Res. Toxicol. 9*, 255−261.
34. Tsao, H., Mao, B., Zhuang, P., Xu, R., Amin, S., and Geacintov, N. E. (1998) *Biochemistry 37*, 4993−5000.
35. Ruan, Q., Zhuang, P., Li, S., Perlow, R., Srinivasan, A. R., Lu, X. J., Broyde, S., Olson, W. K., and Geacintov, N. E. (2001) *Biochemistry 40*, 10458−10472.
36. Theis, K., Chen, P. J., Skorvaga, M., Van Houten, B., and Kisker, C. (1999) *EMBO J. 18*, 6899−6907.
37. Machius, M., Henry, L., Palnitkar, M., and Deisenhofer, J. (1999) *Proc. Natl. Acad. Sci. U.S.A. 96*, 11717−11722.
38. Nakagawa, N., Sugahara, M., Masui, R., Kato, R., Fukuyama, K., and Kuramitsu, S. (1999) *J. Biochem. 126*, 986−990.
39. Moolenaar, G. F., Bazuine, M., van Knippenberg, I. C., Visse, R., and Goosen, N. (1998) *J. Biol. Chem. 273*, 34896−34903.
40. Gordienko, I., and Rupp, W. D. (1998) *EMBO J. 17*, 626−633.

BI034446E