In Vitro Replication of Primer-Templates Containing Benzo[*a*]pyrene Adducts by Exonuclease-Deficient *Escherichia coli* DNA Polymerase I (Klenow Fragment): Effect of Sequence Context on Lesion Bypass†

Yuriy O. Alekseyev and Louis J. Romano*

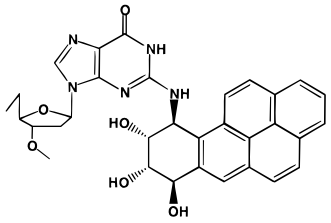
Department of Chemistry, Wayne State University, Detroit, Michigan 48202

Received March 10, 2000; Revised Manuscript Received June 26, 2000

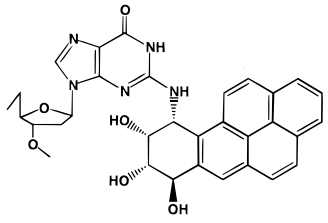
ABSTRACT: The presence of benzo[*a*]pyrene diol epoxide (B[*a*]PDE) adducts in DNA is known to interfere with DNA replication. Kinetic studies of nucleotide insertion by exonuclease-deficient *E. coli* DNA polymerase I (Klenow fragment) across from either the (+)-*trans*- or the (+)-*cis*-B[*a*]P−N^2-dG adduct in the 5′-CGT-3′ sequence context indicated that the rate of nucleotide incorporation followed the order: dAMP > dGMP > dTMP > dCMP, which did not correlate with the mutational spectrum observed for these adducts in this sequence in *E. coli* (mostly G→A transitions). Interestingly, a kinetic analysis of extension past the adduct showed that, unlike other sequences studied, the primer-template was extended best when dT was positioned at the 3′-terminus of the primer across from either a (+)-*trans*- or a (+)-*cis*-B[*a*]P−N^2-dG adduct. In contrast, when the (+)-*trans*-B[*a*]P−N^2-dG adduct was positioned in the 5′-TGC-3′ sequence context, which gives predominantly G→T mutations in *E. coli,* extension was detectable only when dA was positioned across from the adduct. These data provide the first in vitro evidence that may explain why G→A transitions, rather than the G→T transversions found in other sequences, are preferred in the 5′-CGT-3′ sequence in vivo.

Benzo[*a*]pyrene (B[*a*]P)[1] is an important environmental pollutant that is among the best studied of the family of carcinogens called polycyclic aromatic hydrocarbons. Upon metabolic activation to highly reactive diol epoxides, these compounds become capable of covalent binding to DNA and inducing mutations, some of which correlate with tumor initiation and progression (*1, 2*). In most cases, the mutation spectra resulting from DNA modifications by these compounds are quite complex and depend on the stereochemistry of the adduct, the type of polymerase involved in their replication, and the DNA sequence context within which they lie [reviewed in (*3, 4*)]. However, despite extensive studies, the molecular mechanism by which a specific DNA adduct induces a specific mutation remains illusive.

(±)-*anti*-Benzo[*a*]pyrene diol epoxide [(±)-*anti*-B[*a*]PDE] is thought to be the most relevant product of B[*a*]P metabolic activation (*5*). It has been shown to react predominantly with the exocyclic amine of guanine in duplex DNA via the cis or trans opening of the epoxy ring at the C-10 position to form four major adducts, two of which are shown in Figure 1 (*1*). These stereoisomers adopt strikingly different conformations in double-stranded DNA and at the primer-template junctions [reviewed in (*6*)]. These adducts have been shown to be capable of inducing base substitutions, frameshifts, insertions, and deletions in *E. coli* (*3*). (±)-*anti*-B[*a*]PDE-induced mutagenesis studies have shown that the major base substitution mutation is a G→T transversion in both bacteria (*7, 8*) and mammalian cells (*9*−*11*). More recent studies with site-specifically-modified DNA have shown that all three base substitution mutations (G→T, A, and C) can occur (*12, 13*), but the relative level of each is determined by the



(+)-*trans*



(+)-*cis*

FIGURE 1: Structures of (+)-*anti*-B[*a*]P−N^2-dG adducts.

† This investigation was supported by U.S. Public Health Service Grant CA40605 awarded by the Department of Health and Human Services.

* To whom correspondence should be addressed. Tel.: 313-577-2584; Fax: 313-577-8822; E-mail: LJR@chem.wayne.edu.

[1] Abbreviations: B[*a*]P, benzo[*a*]pyrene; *anti*-B[*a*]PDE, benzo[*a*]-pyrene diol epoxide or r7,t8-dihydroxy-t9,10-epoxy-7,8,9,10-tetrahydro-benzo[*a*]pyrene; KF, Klenow fragment; (+)-*anti*-B[*a*]PDE, 7(*R*),8(*S*),-9(*S*),10(*R*) absolute configuration; (−)-*anti*-B[*a*]PDE, 7(*S*),8(*R*),9(*R*),10(*S*) absolute configuration; B[*a*]P−N^2-dG, benzo[*a*]pyrene adduct at N-2 position of guanine; PAGE, polyacrylamide gel electrophoresis; K_M, Michaelis constant; V_{max}, maximum rate of reaction; F_{ins}, frequency of insertion; F_{ext}, frequency of extension.

10.1021/bi000552u CCC: $19.00
Published on Web 08/03/2000

sequence context surrounding the lesion. These types of studies with the major (+)-*trans*-B[*a*]P−N^2-dG adduct showed that the adduct caused mostly G→T transversions in most cases (*14*, *15*) [including in 5′-CGG-3′ and 5′-TGC-3′ sequence contexts (*13*, *16*)] but in at least two other sequence contexts [5′-CGT-3′ (*17*) and 5′-AGA-3′ (*18*)] G→A transitions predominated.

In an attempt to understand the mutational mechanisms that generate these mutation spectra, in vitro replication systems have been developed that allow the quantification of the mutation frequencies opposite DNA lesions (*19*, *20*). This type of system has been used to show that the rates of incorporation by DNA polymerase I (Klenow fragment) (KF) across from each of the four sterioisomers in both a 5′-CGC-3′ and a 5′-TGC-3′ sequence follow the order: dAMP > dGMP > dTMP > dCMP (*21*, *22*). Consistent with these results, thermodynamic studies have shown that the relative destabilizing effect of the (±)-*anti*-B[*a*]PDE adducts in these two sequence contexts for all stereoisomers is least when a B[*a*]P−N^2-dG adduct is positioned across from an adenine and most when the adduct is positioned across from a cytosine (*23*). These results suggest that KF preferentially incorporates the nucleotide that causes the least local distortion of the DNA double helix. Prior studies have shown that the purines stack better with the B[*a*]P moiety than pyrimidines (*24*), providing a related explanation for why dAMP and dGMP are incorporated significantly faster than dTMP and dCMP. Other in vitro studies using T7 DNA polymerase (Sequenase 2) and DNA polymerase I (Klenow fragment) and its exonuclease-deficient version (*25*, *26*) have recently confirmed and extended these results by showing that purines (dAMP > dGMP) are more readily incorporated compared with pyrimidines across from B[*a*]P−N^2-dG adducts when lesion bypass takes place. Finally, extension past the adduct was shown to occur in the 5′-CGC-3′ (*21*) or 5′-CGG-3′ (*25*) sequence context only when dAMP had been incorporated opposite the adduct, further accounting for the G→T mutations that occur in these sequences.

In the present study, we report the results of a kinetic analysis of nucleotide insertion across from the (+)-*trans*- and (+)-*cis*-B[*a*]P−N^2-dG adducts in a 5′-CGT-3′ sequence context, one of the sequence contexts that leads to G→A mutations in vivo (*17*). Prior studies have shown that the thermodynamic stability of (±)-*anti*-B[*a*]PDE adducts in this sequence gave the same trends that were determined for the 5′-TGC-3′ and 5′-CGC-3′ sequences (*23*, *27*). Thus, it is not surprising that we find in this study that for both the (+)-*trans*- and (+)-*cis*-B[*a*]P−N^2-dG adducts in the 5′-CGT-3′ sequence the rates of incorporation also followed the same order: dAMP > dGMP > dTMP > dCMP. These results do not correlate with the mutation spectrum observed for these lesions in the same sequence in vivo. However, when extension past these adducts in this sequence was analyzed, it was found that only when a dT was positioned opposite either adduct could extension be measured. Thus, even though dTMP is incorporated more slowly in this sequence, it allows the most efficient bypass of the lesion, thus possibly accounting for the G→A mutations that are seen in vivo (*17*).

MATERIALS AND METHODS

Materials. The Klenow fragment of *E. coli* DNA polymerase I was purchased from Amersham Pharmacia Biotech. The protein had been overproduced and purified from a strain carrying the double mutation D355A, E357A, which results in about 10^5-fold reduction of endogenous 3′-5′ exonuclease activity (*28*). T4 polynucleotide kinase and T4 DNA ligase were also purchased from Amersham Pharmacia Biotech.

Oligonucleotides were obtained from Midland Certified Inc. dNTPs were purchased from Promega. [γ-^{32}P]ATP was from ICN Biomedicals. The racemic (±)-*anti*-B[*a*]PDE was purchased from the National Cancer Institute Chemical Reference Standard Repository (Kansas City, MO). All other general reagents and chemicals were obtained from Fisher and VWR.

Synthesis and Purification of the B[a]PDE-Modified Oligonucleotides. The sequences of oligonucleotides that were used in this study are shown in Figure 2A. All oligonucleotides (approximately 1 OD unit) were purified by 20% denaturing polyacrylamide gel electrophoresis (PAGE) in the presence of 8 M urea according to the procedures described in (*29*). The oligonucleotides were detected under UV light, and the corresponding area was cut out, eluted overnight, and desalted with Centricon-3 microconcentrators (Amicon) according to the manufacturer's protocol. The 11-mers were modified by racemic (±)-*anti*-B[*a*]PDE, and the four stereoisomers were separated by reverse-phase HPLC, gel-purified, and characterized as described (*23*). The 11-mers with (+)-*trans*- and (+)-*cis*-B[*a*]P−N^2-dG adducts, as well as unmodified 11-mers, were ligated to the phosphorylated 17-mer, using fully complementary 22-mers as scaffolds (Figure 2B) according to the manufacturer's protocol (Amersham Pharmacia Biotech). Briefly, a mixture containing the modified or unmodified 11-mer (1 nmol) and the 17-mers (1.5 nmol) and the 22-mer (1.5 nmol) in 1.5 mL volume containing 50 mM Tris-HCl, pH 7.5, 10 mM $MgCl_2$, 10 mM DTT, 1 mM ATP, and 25 mg/mL bovine serum albumin was heated to 90 °C and then slowly cooled over 60 min. T4 DNA ligase (40 units) was added, and the samples were incubated at 16 °C for 16 h. After desalting of the ligation mixture with Centricon-3 microconcentrators (Amicon), the ligated 28-mers were separated from the ligation mixture using 20% denaturing PAGE and desalted as described above. Aliquots (0.5 pmol) were 5′-end-labeled with T4 polynucleotide kinase and [γ-^{32}P]ATP and analyzed using 20% PAGE (0.4 mm) in the presence of 8 M urea. The analysis of the purity of the products is shown in Figure 2C.

Kinetics of Nucleotide Insertion Across from the Adduct and Chain Extension Past the Adduct. The kinetic analyses were performed under conditions similar to those described (*21*). The purified 22-mer and 23-mers (10 pmol) were 5′-end-labeled with T4 polynucleotide kinase and [γ-^{32}P]ATP, purified using 20% PAGE in the presence of 8 M urea (the corresponding bands were detected by Molecular Dynamics PhosphorImager SF, cut out, and eluted overnight as described above), and desalted with Centricon-3 microconcentrators (Amicon). The modified templates (28-mers) were annealed to the ^{32}P-labeled primers (22-mer or one of the 23-mers) in a 4:1 molar ratio by heating to 95 °C and slow-cooling to room temperature in a 50 μL volume containing 50 mM Tris-HCl, pH 7.5, 10 mM $MgCl_2$, 1 mM dithiothreitol, and 0.05 mg/mL bovine serum albumin. The extent of duplex formation was estimated by analysis of an aliquot taken from the annealing mixtures using a 7% native PAGE

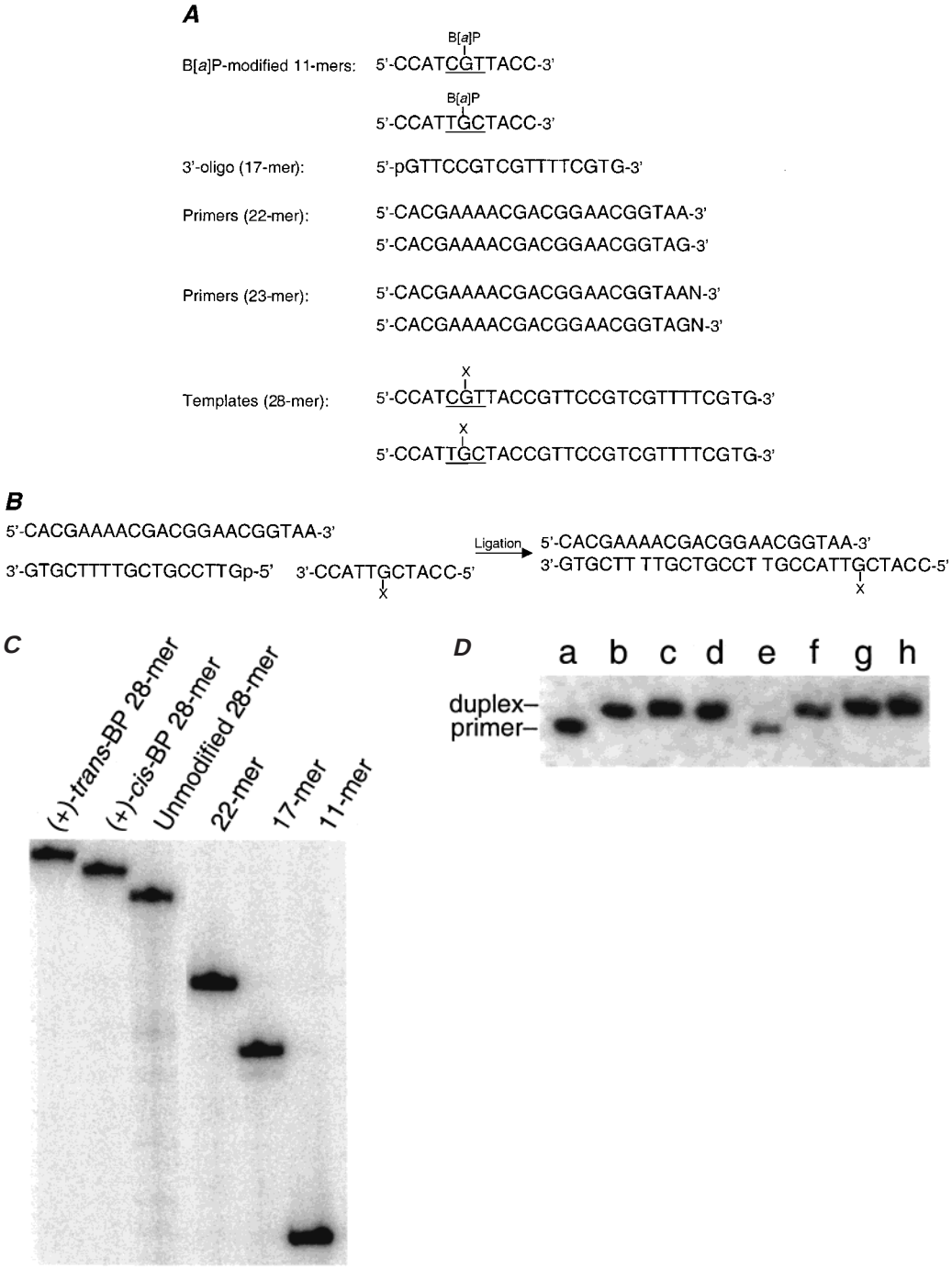


FIGURE 2: (A) Sequences of the oligonucleotides used in this study. The 22-mer and the 23-mers (N stands for one of the four possible nucleotides) were used as primers in the primer extension studies with the 28-mer templates (X stands for either H or (+)-*cis*- or (+)-*trans*-B[*a*]P−N^2-dG). (B) To obtain the 28-mer templates, the modified or unmodified 11-mer was ligated with the phosphorylated 17-mer using complementary 22-mer as a scaffold as described under Materials and Methods. (C) All oligonucleotides and ligation products were gel-purified and desalted as described under Materials and Methods. To confirm the purity, the ^{32}P-labeled aliquots were analyzed by 20% PAGE under denaturing conditions as described under Materials and Methods. (D) The templates were annealed to the ^{32}P-labeled primers according to the procedure described under Materials and Methods. The extent of duplex formation was estimated by 7% native PAGE. Lane a, 23-mer primer with N = A; lanes b, c, d, duplexes formed by this primer with unmodified, (+)-*trans*-, or (+)-*cis*-B[*a*]PDE-modified templates, respectively; lane e, 23-mer primer with N = C; lanes f, g, h, duplexes formed by this primer with unmodified, (+)-*trans*-, or (+)-*cis*-B[*a*]PDE-modified templates, respectively. Effective annealing was achieved in all cases. The results for the other primer-templates used in the study were identical (data not shown).

analysis at 4 °C according to the procedure described in (*29*) (Figure 2D). After addition of KF (final concentration 0.01 unit/μL) and 200 or 400 μM samples of one of the four dNTPs (dGTP or dATP for chain extension past the adduct), 8 μL aliquots were removed from the reaction mixtures at the indicated times. The samples were then heated for 3 min in the presence of formamide and analyzed using 15% denaturing PAGE (0.4 mm). Band intensity was determined using a Molecular Dynamics PhosphorImager SF. To derive the kinetic parameters, the same DNA duplexes were incubated in the presence of KF and increasing amounts of one of the four nucleotides for times that provide no more

than 20% primer extension based on the results of the time courses. The velocity of the reaction, expressed as percent per unit time, was measured as $v = 100[I_1/(I_0 + 0.5I_1)t]$ (*30*), where I_0 is the intensity of the original primer, I_1 is the intensity of the primer extended by one nucleotide, and t is the reaction time. The K_M and V_{max} values were obtained from Hanes−Woolf plots of [dNTP]/v vs [dNTP] based on the results of three independent experiments where a standard deviation did not exceed 10% (data not shown). Insertion (F_{ins}) and extension (F_{ext}) frequencies were determined relative to the correct unmodified base pair, according to equations derived by Mendelman et al. (*31, 32*): $F = (V_{max}/K_m)$[wrong pair]/(V_{max}/K_m)[right pair = G·C], where the wrong pair is any base pair containing a mismatch or B[*a*]P−N^2-dG adduct.

RESULTS

Kinetics of Nucleotide Insertion Across from (+)-*anti-B*[*a*]*-PDE Adducts.* As a first step in the determination of the nucleotide incorporation parameters across from the (+)-*trans*- and (+)-*cis*-B[*a*]P−N^2-dG adducts in a 5′-CGT-3′ sequence, the times to achieve 20% primer extension were determined. These levels are needed so that the remaining experiments can be carried out under conditions that satisfy single-hit kinetic requirements (*31, 32*). As shown in Figure 3, modified 28-mers were used as templates and a complementary ^{32}P-labeled 22-mer was used as a primer so that the 3′-terminus was positioned one nucleotide before the adduct. Each time course was conducted in the presence of a single nucleotide. It is evident from these results that incorporation across from a (+)-*trans*-B[*a*]P−N^2-dG adduct by KF was more rapid than for the (+)-*cis*-B[*a*]P−N^2-dG adduct. For the modified templates, no extension above the primer + 1 band was observed. From this time course study, it is also possible to conclude that the purines were incorporated much faster than pyrimidines.

Next the V_{max} and K_M were determined using steady-state conditions as described previously (*21*). For these experiments, the primer-templates shown in Figure 3A were incubated with KF and increasing levels of each dNTP, and the F_{ins} was determined as described under Materials and Methods. The data obtained for unmodified templates and templates containing (+)-*trans*- and (+)-*cis*-B[*a*]P−N^2-dG adducts are presented in Table 1. It was found that trends for the rates of the nucleotide incorporation for the 5′-CGT-3′ sequence correlated with those reported by Shibutani et al. (*21*), Mekhovich (*22*), and Hanrahan et al. (*25*) for 5′-CGC-3′, 5′-TGC-3′, and 5′-CGG-3′ contexts: for both stereoisomers, the insertion frequences followed the order: dAMP > dGMP > dTMP > dCMP. The unmodified templates displayed the opposite order of the F_{ins}. For both stereoisomers in the sequence studied here, the overall bypass was less than that reported by Shibutani for the 5′-CGC-3′ sequence (*21*), which may be explained by the fact that there is a less stable A·T base pair preceding the lesion. Finally, unlike the latter sequence, for the 5′-CGT-3′ context the F_{ins} values were much higher for any of the four nucleotides across from (+)-*trans*-B[*a*]P−N^2-dG than those for the same nucleotide across from (+)-*cis*-B[*a*]P−N^2-dG (Table 1).

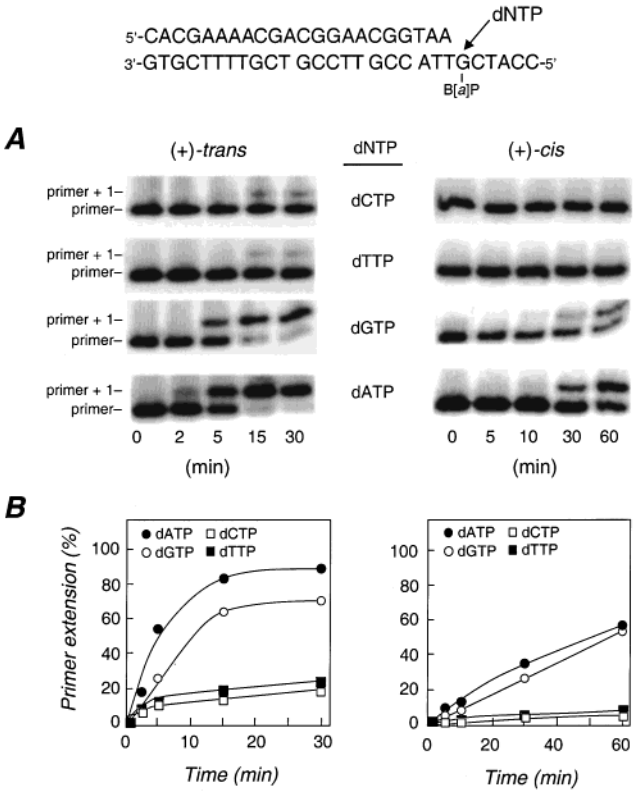


FIGURE 3: Time course of dNTP insertion across from the (+)-*trans*- and (+)-*cis*-B[*a*]P−N^2-dG adducts in 5′-CGT-3′ sequence contexts by KF. The ^{32}P-labeled 22-mer primers were annealed to 28-mer templates and incubated with KF in the presence of one of the four dNTPs under conditions described under Materials and Methods. (A) Aliquots were taken out of the reaction mixture at the indicated time points, heated in the presence of formamide, and analyzed by denaturing PAGE (15%) as described under Materials in Methods. (B) Primer extension (%), calculated as the ratio of the band intensity of the extension product (primer + 1) to that of the total DNA, was plotted as a function of time.

Table 1: Insertion Frequencies of Nucleotide Incorporation Across from the Adduct[a]

	dNMP			
template	dCMP	dTMP	dGMP	dAMP
unmodified	1	8.2×10^{-4}	4.6×10^{-4}	2.8×10^{-4}
(+)-*trans*	1.5×10^{-7}	1.5×10^{-6}	1.1×10^{-5}	1.8×10^{-5}
(+)-*cis*	7.1×10^{-9}	3.2×10^{-8}	1.3×10^{-7}	2.9×10^{-6}

[a] Frequencies of nucleotide insertion were estimated by the equation: $F_{ins} = (V_{max}/K_m)$[wrong pair]/(V_{max}/K_m)[right pair = dG·dC], where the kinetic parameters V_{max} and K_m were obtained from analyzing the primer extension at increasing dNTP concentration as described under Materials and Methods.

Chain Extension Past the Adduct. The analysis of the time course for chain extension past the adducts is shown in Figure 4. In these experiments, the same (+)-*anti*-B[*a*]PDE-modified 28-mers were used as templates, but four different 23-mers were used as primers, so that the 3′ end of the primer terminated with either dC, dT, dG, or dA opposite the adduct on the template. The ability of KF to incorporate dGMP, the next correct nucleotide, using each of these primer-templates is shown in Figure 4A. For the modified templates, no extension above the primer + 1 band was observed as in the case of the 22-mer primer (see above). In contrast to the 5′-CGC-3′ and 5′-CGG-3′ sequences analyzed previously (*21, 25*), where only the primers with dA at the 3′-terminus across

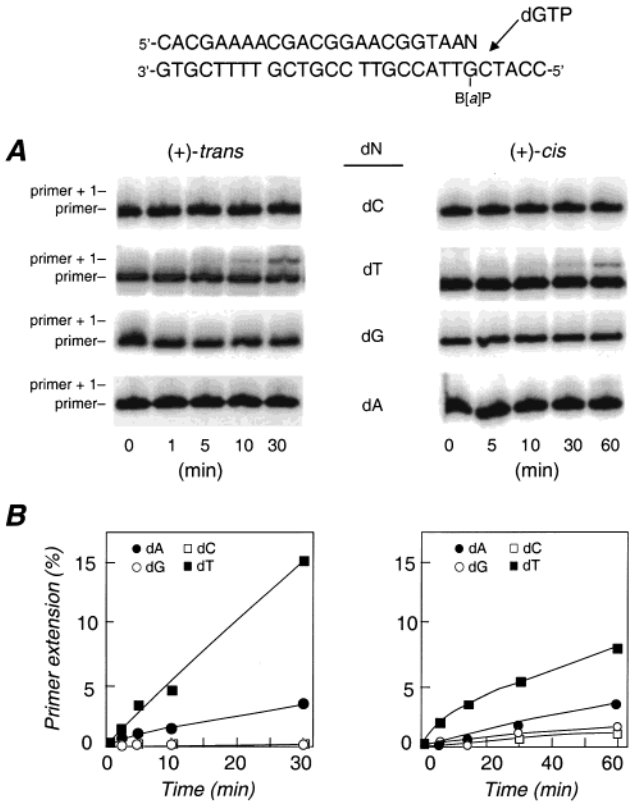


FIGURE 4: Time course of primer extension past the (+)-*trans*- and (+)-*cis*-B[*a*]P−N^2-dG adducts in a 5′-CGT-3′ sequence context by KF. The ^{32}P-labeled 23-mer primers with one of the four nucleotides on the 3′-terminus (dN) were annealed to 28-mer templates and incubated with KF in the presence of the next correct nucleotide dGTP under conditions described under Materials and Methods. (A) Aliquots were taken out of the reaction mixture at the indicated time points, heated in the presence of formamide, and analyzed by denaturing PAGE (15%) as described under Materials in Methods. (B) Primer extension (%), calculated as the ratio of the band intensity of the extension product (primer + 1) to that of the total DNA, was plotted as a function of time.

Table 2: Extension Frequencies of Nucleotide Incorporation Past the Adduct[a]

	N[b]			
template	C	T	G	A
unmodified	1	4.2×10^{-3}	7.4×10^{-4}	2.6×10^{-4}
(+)-*trans*	nd[c]	8.4×10^{-7}	nd	nd
(+)-*cis*	nd	2.1×10^{-8}	nd	nd

[a] Extension frequencies were estimated by the equation: $F_{ext} = (V_{max}/K_m)$[wrong pair]/(V_{max}/K_m)[right pair = dG·dC], where the kinetic parameters V_{max} and K_m were obtained from analyzing the primer extension at increasing dGTP concentration as described under Materials and Methods. [b] Nucleotide in primer across from the adduct as indicated in Figure 4, top. [c] nd, not detectable.

from the B[*a*]P−N^2-dG were extended, in the case of the 5′-CGT-3′ sequence the extension of the primer with dA at the 3′-terminus was hardly detectable (Figure 4B). No extension was observed when either dG or dC was positioned across from the adducts. However, when dT was positioned across from either of the stereoisomers, the extension was significantly higher, allowing the kinetic parameters V_{max} and K_M to be determined. The extension frequencies derived from these parameters for the unmodified and (+)-*anti*-B[*a*]PDE-modified templates are summarized in Table 2. Although extension of the dA-terminated primer could just barely be

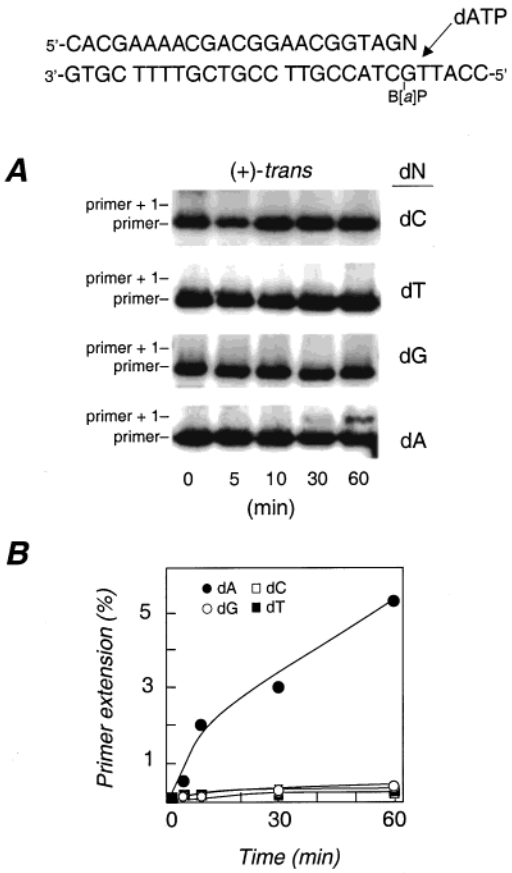


FIGURE 5: Time course of primer extension past the (+)-*trans*-B[*a*]P−N^2-dG adduct in the 5′-TGC-3′ sequence context by KF. The ^{32}P-labeled 23-mer primers with one of the four nucleotides on the 3′-terminus (dN) were annealed to 28-mer templates and incubated with KF in the presence of the next correct nucleotide dATP under conditions described under Materials and Methods. (A) Aliquots were taken out of the reaction mixture at the indicated time points, heated in the presence of formamide, and analyzed by denaturing PAGE (15%) as described under Materials in Methods. (B) Primer extension (%), calculated as a ratio of the band intensity of the extension product (primer + 1) to that of the total DNA, was plotted as a function of time.

observed in Figure 4, the level was not high enough to obtain the V_{max} and K_M values. The frequency of lesion bypass can be calculated as the product of F_{ins} and F_{ext} and was found to be equal to 1.3×10^{-12} and 6.7×10^{-16} when dTMP was incorporated across from (+)-*trans*- and (+)-*cis*-B[*a*]P−N^2-dG adducts, respectively. Interestingly, for the unmodified primer-templates, the extension of a G·T mismatch was significantly faster than either a G·G or a G·A mismatch (Table 2).

To ensure that the trend observed for the 5′-CGT-3′ sequence was not an artifact related to the experimental design, the same experiment was performed using a template containing a (+)-*trans*-B[*a*]P−N^2-dG adduct in a 5′-TGC-3′ sequence context in an otherwise identical template and using the analogous four primers as described in Figure 4. The time course of extension shows that the incorporation of the next correct nucleotide, dAMP, was detectable only when the primer was terminated with dA instead of dT (Figure 5) as had been determined for the 5′-CGT-3′ in Figure 4 and Table 2. Thus, sequence context appears to be a major factor in determining bypass efficiency.

DISCUSSION

The fact that in some cases there is an apparent lack of correlation between in vitro and in vivo replication of damaged DNA suggests that caution is needed when attempting to make generalizations about the cellular processing of DNA lesions based on data obtained in vitro (*4, 33*). For example, most of the in vitro studies have shown that the incorporation of the purines across from the B[*a*]P−N^2-dG adducts is significantly faster than that of the correct cytosine (*21, 22, 25, 26*) even though B[*a*]PDE-induced mutagenesis studies find that the majority of the bypass products in the cell exhibit the insertion of the correct nucleotide across the B[*a*]PDE-modified guanine. This is presumably because of DNA repair prior to replication, the 3′-exonuclease proofreading activity of the DNA polymerase, and the possible involvement of a more accurate multicomponent replicative machinery (*4, 14, 25, 34, 35*). Moreover, it has also been shown that the mutation spectrum strongly depends on the cell type and consequently on the enzyme involved in replication (*14, 25*). Based on the observed discrepancies between in vitro and in vivo replication of the damaged DNA some authors have concluded that their in vitro assay was not predictive of the mutagenic spectrum observed in vivo (*33*).

One of the strongest arguments against the legitimacy of predictions based on in vitro studies is that replicative machinery in the cell is fundamentally different from that in vitro. In numerous in vitro studies, including those described here, KF is used as a model enzyme because of its ability to bypass bulky lesions and the availability of mutants that lack proofreading activity, even though its role in in vivo replication is quite restricted. However, even the use of *E. coli* DNA polymerase III, which, in fact, is completely blocked by B[*a*]PDE lesions in vitro (*21, 25*), does not adequately model the in vivo system (*33*) since it is not clear which protein assemblies are involved in replicative bypass of the lesions in bacterial cells. One model suggests that a different subassembly of Pol III is involved (*33*), while other studies propose that Pol I (*36*) or Pol II, which can become processive in the presence of Pol III accessory factors (*37*), might substitute for Pol III in replicating the exogenous DNA. The recent discovery of Pol IV (*38*) and Pol V (*39, 40*), which appear to participate in lesion bypass, make this process even more uncertain.

Another argument against the significance of in vitro studies is that the conformation of the adduct during DNA replication in the cell might be different from that present during in vitro DNA synthesis. The proponents of this hypothesis have suggested that a single sterioisomer of (+)-*trans*-B[*a*]P−N^2-dG is capable of numerous sequence-dependent conformations in duplex DNA (*41*) and that different mutations result when the polymerase encounters these different structures (*35*). NMR studies have shown that B[*a*]PDE adducts adopt very different conformations when they are present at the primer-template junction compared with their structures in single-stranded or duplex DNA (*6*) and therefore it is not clear how appropriate it is to correlate adduct structures with those that occur inside the active site of a polymerase during replication.

Despite all of the arguments listed above, there have been several revealing in vitro studies involving different lesions that seem to provide some insight into the mechanism of mutagenesis (*20, 21, 25, 42*−*45*), and it is tempting to assume that the bases for these correlations are measurable physical properties that are independent of the replication machinery and which can be studied in an in vitro model system. In addition, a second factor that must contribute to the mechanism of mutagenesis is the fact that the identity of the nucleotide incorporated across from a lesion may be unrelated to the nucleotide that is needed to allow full bypass. If the major nucleotide incorporated does not allow further extension while a minor product does, then the mutation observed would result from this latter process. For example, in the above-mentioned study (*33*), which did not reveal any correlation between in vivo and in vitro replication of damaged DNA, only incorporation of the nucleotides across from the adducts was studied, and extension past the adduct was not analyzed. It is likely that studies that do not take into account this important component of in vitro DNA synthesis may be missing a significant factor in mutation induction.

In the present study, it was found that the in vitro kinetics of nucleotide incorporation across from either a (+)-*trans*- or a (+)-*cis*-B[*a*]P−N^2-dG adduct in a 5′-CGT-3′ sequence did not correlate with the mutation caused by these lesions in this sequence in vivo (*17*). The predominant (82%) mutation that is observed in *E. coli* for the (+)-*trans*-B[*a*]P−N^2-dG adduct in this sequence involves the incorporation of dTMP across from the adduct leading to a G→A transition (*17*), while we found that dAMP has an insertion frequency 12- or 90-fold higher than dTMP opposite the (+)-*trans*-B[*a*]P−N^2-dG or (+)-*cis*-B[*a*]P−N^2-dG adducts in this sequence, respectively (Table 1). However, when the kinetics of extension were measured using primers that positioned each of the four nucleotides opposite either adduct, the presence of a dT across from either adduct allowed faster extension than if any of the other nucleotides was positioned there. This is consistent with the preference for G→A mutations in this sequence in vivo (see below). A similar result has been obtained for the bypass of B[*a*]P−N^2-dG adducts by T7 RNA polymerase where it was found that a majority of the transcripts terminated opposite the adduct and contained a purine on their 3′-terminus, while full transcripts had incoporated only pyrimidines across from the adduct (*46*).

The rationale for performing these studies in the 5′-CGT-3′ sequence context is due to the fact that, as mentioned above, in vivo studies have shown that when a B[*a*]PDE-modified guanine is positioned in this context it gives rise predominantly to G→A transitions, while most other contexts give predominantly G→T transversions. Thus, while a (+)-*trans*-B[*a*]P−N^2-dG adduct in a 5′-CGT-3′ sequence gave rise to 82% G→A transitions in *E. coli* (*17*), (±)-anti-B[*a*]-PDE-modified guanine in 5′-TGC-3′ or 5′-CGG-3′ produced 97% and 57% G→T transversions, respectively (*12, 16*). Moreover, prior studies on the in vitro replication of primer-templates containing a (+)-*trans*-B[*a*]P−N^2-dG adduct in a 5′-CGG-3′(*25*), and (+)-*trans*-, (−)-*trans*-, (+)-*cis*-, or (−)-*cis*-B[*a*]P−N^2-dG adducts in 5′-CGC-3′ or 5′-TGC-3′ (*21, 22*) have shown that in each case the insertion frequencies followed the order: dAMP > dGMP > dTMP > dCMP. The extension past these adducts in the first two sequences was reported to be significant only when the primer was

terminated with dA across from the adduct on the template (*21*, *25*). We have extended these studies here by showing that extension past a (+)-*trans*-B[*a*]P−N^2-dG adduct in the 5′-TGC-3′ sequence also takes place only from the primer terminated with dA across from the adduct (Figure 5). All of these in vitro findings are consistent with the observed in vivo mutation spectrum of predominantly G→T in 5′-CGG-3′ or 5′-TGC-3′and G→A in 5′-CGT-3′ sequences.

Two questions arise from the results presented here. First, what factors contribute to the selection of the nucleotide to be incorporated opposite the B[*a*]PDE-modified guanosine? Second, what factors allow a primer-template that terminates opposite the lesion to be extended? As stated above, we believe that there are measurable physical parameters that contribute to both of these processes, which may allow the prediction of the in vivo mutational spectrum.

Prior studies have shown that B[*a*]PDE adducts are significant blocks to DNA replication in vitro for all tested DNA polymerases. In most cases, the polymerase becomes stalled before the lesion, although some polymerases, including KF, are able to incorporate a nucleotide across from the adduct (*21*, *22*, *25*, *26*, *47*). In general, most studied DNA polymerases tend to incorporate purines across from BPDE-modified guanine in a process that is independent of DNA sequence and adduct steriochemistry (*21*, *22*, *25*, *26*). This is true even for such distantly related enzymes as T7 RNA polymerases (*46*).

The usual explanation provided as the basis for these results is that the structure of the DNA helix induced by B[*a*]PDE modification is interpreted by the polymerase as noninstructional, so that incorporation takes place according to the "A-rule" [discussed in (*21*, *25*)]. However, there are several studies that suggest that there may be a physical basis for this preferred incorporation. One of these may be the stability of the base pair that forms, since it has been shown that the relative thermodynamic stability of each B[*a*]PDE adduct isomer in both a 5′-CGC-3′ and a 5′-TGC-3′ sequence exactly correlates with the trends of nucleotide incorporation observed in vitro (*23*). Comparison of the insertion frequencies obtained here (Table 1) with the thermodynamic data from a previous study (*27*) shows that the same correlation exists also for the 5′-CGT-3′ sequence. Additional insight into this question has been obtained using a spectroscopic analysis (*24*), which showed that a B[*a*]P moiety is able to stack much better with the purine bases. It was suggested that the incoming purine might form a weak van der Waals complex with the B[*a*]P moiety prior to base-pairing and phospodiester bond formation. Other spectroscopy studies and several NMR solution structures of a B[*a*]P−N^2-dG adduct in 5′-CGC-3′ (*48*−*51*) and 5′-TGC-3′ (*52*, *53*) sequence contexts reveal that the nucleotides flanking the modified guanine have an influence on the structure of the adduct in duplex DNA. In addition, computational analyses of (+)-*trans*-B[*a*]P-dG in 5′-TGC-3′ and 5′-CGT-3′ sequence contexts in duplex DNA suggest that for this adduct there are numerous possible conformations of relatively similar stability and that in different sequences the predominant conformations are not the same (*41*). The authors (*41*) propose that an intercalative conformation in which the modified base is displaced into the major groove pointing in the 5′ direction leads to a G→T mutation while an intercalative structure where the modified base is in the minor groove pointing in the 3′ direction leads to a G→A mutation (*35*).

All of the above analyses deal with the fully duplex structures and thus do not take into account the effect of having a lesion at the primer terminus. Moreover, these studies cannot address the effect of having the adduct present in the active site of the DNA polymerase. Although there is not enough published physical data to estimate this effect completely, it has been shown that the impact of bulky lesions on polymerase/DNA complex formation is significant (*54*, *55*). In addition, only a single structure has been determined of a B[*a*]PDE-modified primer-template [(+)-*trans*-B[*a*]P−N^2-dG adduct located in a 5′-CGC-3′ sequence context] in which cytosine was positioned across from the adduct (*56*). However, without data showing the effect of different sequence contexts, it is difficult to make predictions on how these structures relate to the mutation spectra. A better understanding of these processes awaits NMR structures of these adducts in primer-templates in different sequence contexts and crystal structures of polymerases bound to modified templates.

ACKNOWLEDGMENT

We thank Dr. Mohsen B. Arghavani for kindly providing the (±)-*anti*-B[*a*]PDE-modified 11-mer oligonucleotides.

REFERENCES

1. Harvey, R. G. (1991) *Polycyclic Aromatic Hydrocarbons: Chemistry and Carcinogenicity*, University Press, Cambridge.
2. Denissenko, M. F., Pao, A., Tang, M., and Pfeifer, G. P. (1996) *Science 274*, 430−432.
3. Loechler, E. L. (1995) *Mol. Carcinog. 13*, 213−219.
4. Loechler, E. L. (1996) *Carcinogenesis 17*, 895−902.
5. Newbold, R. F., and Brookes, P. (1976) *Nature 261*, 52−54.
6. Geacintov, N. E., Cosman, M., Hingerty, B. E., Amin, S., Broyde, S., and Patel, D. J. (1997) *Chem. Res. Toxicol. 10*, 111−146.
7. Bernelot-Moens, C., Glickman, B. W., and Gordon, A. J. (1990) *Carcinogenesis 11*, 781−785.
8. Eisenstadt, E., Warren, A. J., Porter, J., Atkins, D., and Miller, J. H. (1982) *Proc. Natl. Acad. Sci. U.S.A. 79*, 1945−1949.
9. Mazur, M., and Glickman, B. W. (1988) *Somat. Cell. Mol. Genet. 14*, 393−400.
10. Carothers, A. M., and Grunberger, D. (1990) *Carcinogenesis 11*, 189−192.
11. Keohavong, P., and Thilly, W. G. (1992) *Environ. Health Perspect. 98*, 215−219.
12. Rodriguez, H., and Loechler, E. L. (1993) *Carcinogenesis 14*, 373−383.
13. Jelinsky, S. A., Liu, T., Geacintov, N. E., and Loechler, E. L. (1995) *Biochemistry 34*, 13545−13553.
14. Moriya, M., Spiegel, S., Fernandes, A., Amin, S., Liu, T., Geacintov, N., and Grollman, A. P. (1996) *Biochemistry 35*, 16646−16651.
15. Page, J. E., Zajc, B., Oh-hara, T., Lakshman, M. K., Sayer, J. M., Jerina, D. M., and Dipple, A. (1998) *Biochemistry 37*, 9127−9137.
16. Mackay, W., Benasutti, M., Drouin, E., and Loechler, E. L. (1992) *Carcinogenesis 13*, 1415−1425.
17. Shukla, R., Liu, T., Geacintov, N. E., and Loechler, E. L. (1997) *Biochemistry 36*, 10256−10261.
18. Shukla, R., Geacintov, N. E., and Loechler, E. L. (1999) *Carcinogenesis 20*, 261−268.
19. Shibutani, S., Takeshita, M., and Grollman, A. P. (1991) *Nature 349*, 431−434.
20. Shibutani, S., and Grollman, A. P. (1993) *J. Biol. Chem. 268*, 11703−11710.

21. Shibutani, S., Margulis, L. A., Geacintov, N. E., and Grollman, A. P. (1993) *Biochemistry 32*, 7531−7541.
22. Mekhovich, O. (1997) Ph.D. Dissertation, *Enzymatic Analysis of Aminofluorene and Benzo[a]pyrene DNA Structures*, Wayne State University.
23. Arghavani, M. B., SantaLucia, J., Jr., and Romano, L. J. (1998) *Biochemistry 37*, 8575−8583.
24. Geacintov, N. E., Zhao, R., Kuzmin, V. A., Kim, S. K., and Pecora, L. J. (1993) *Photochem. Photobiol. 58*, 185−194.
25. Hanrahan, C. J., Bacolod, M. D., Vyas, R. R., Liu, T., Geacintov, N. E., Loechler, E. L., and Basu, A. K. (1997) *Chem. Res. Toxicol. 10*, 369−377.
26. Lipinski, L. J., Ross, H. L., Zajc, B., Sayer, J. M., Jerina, D. M., and Dipple, A. (1998) *Int. J. Oncol. 13*, 269−273.
27. Arghavani, M. (1998) Ph.D. Dissertation, *Study of Structural and Biological Effects of Carcinogen-Modified Oligonucloetides: Thermodynamics, Incorporation, and Repair*, Wayne State University.
28. Derbyshire, V., Freemont, P. S., Sanderson, M. R., Beese, L., Friedman, J. M., Joyce, C. M., and Steitz, T. A. (1988) *Science 240*, 199−201.
29. Maniatis, T., Fritsch, E. F., and Sambrook, J. (1982) *Molecular Cloning. (A Laboratory Manual)*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
30. Petruska, J., Goodman, M. F., Boosalis, M. S., Sowers, L. C., Cheong, C., and Tinoco, I., Jr. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 6252−6256.
31. Mendelman, L. V., Boosalis, M. S., Petruska, J., and Goodman, M. F. (1989) *J. Biol. Chem. 264*, 14415−14423.
32. Mendelman, L. V., Petruska, J., and Goodman, M. F. (1990) *J. Biol. Chem. 265*, 2338−2346.
33. McNees, A. G., O'Donnell, M., Horton, P. H., Kim, H. Y., Kim, S. J., Harris, C. M., Harris, T. M., and Lloyd, R. S. (1997) *J. Biol. Chem. 272*, 33211−33219.
34. Fernandes, A., Liu, T., Amin, S., Geacintov, N. E., Grollman, A. P., and Moriya, M. (1998) *Biochemistry 37*, 10164−10172.
35. Kozack, R. E., Shukla, R., and Loechler, E. L. (1999) *Carcinogenesis 20*, 95−102.
36. Maki, H., Bryan, S. K., Horiuchi, T., and Moses, R. E. (1989) *J. Bacteriol. 171*, 3139−3143.
37. Bonner, C. A., Stukenberg, P. T., Rajagopalan, M., Eritja, R., O'Donnell, M., McEntee, K., Echols, H., and Goodman, M. F. (1992) *J. Biol. Chem. 267*, 11431−11438.
38. Wagner, J., Gruz, P., Kim, S. R., Yamada, M., Matsui, K., Fuchs, R. P., and Nohmi, T. (1999) *Mol. Cell. 4*, 281−286.
39. Tang, M., Shen, X., Frank, E. G., O'Donnell, M., Woodgate, R., and Goodman, M. F. (1999) *Proc. Natl. Acad. Sci. U.S.A. 96*, 8919−8924.
40. Maor-Shoshani, A., Reuven, N. B., Tomer, G., and Livneh, Z. (2000) *Proc. Natl. Acad. Sci. U.S.A. 97*, 565−570.
41. Kozack, R. E., and Loechler, E. L. (1999) *Carcinogenesis 20*, 85−94.
42. Wang, C. I., and Taylor, J. S. (1992) *Biochemistry 31*, 3671−3681.
43. Reardon, D. B., Bigger, C. A., and Dipple, A. (1990) *Carcinogenesis 11*, 165−168.
44. Lindsley, J. E., and Fuchs, R. P. (1994) *Biochemistry 33*, 764−772.
45. Belguise-Valladier, P., Maki, H., Sekiguchi, M., and Fuchs, R. P. (1994) *J. Mol. Biol. 236*, 151−164.
46. Choi, D. J., Roth, R. B., Liu, T., Geacintov, N. E., and Scicchitano, D. A. (1996) *J. Mol. Biol. 264*, 213−219.
47. Keohavong, P., Shukla, R., Melacrinos, A., Day, B. W., and Reha-Krantz, L. (1998) *DNA Cell Biol. 17*, 541−549.
48. 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., et al. (1992) *Proc. Natl. Acad. Sci. U.S.A. 89*, 1914−1918.
49. de los Santos, C., Cosman, M., Hingerty, B. E., Ibanez, V., Margulis, L. A., Geacintov, N. E., Broyde, S., and Patel, D. J. (1992) *Biochemistry 31*, 5245−5252.
50. 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.
51. Cosman, M., Hingerty, B. E., Luneva, N., Amin, S., Geacintov, N. E., Broyde, S., and Patel, D. J. (1996) *Biochemistry 35*, 9850−9863.
52. Fountain, M. A., and Krugh, T. R. (1995) *Biochemistry 34*, 3152−3161.
53. Pradhan, P., Jernstrom, B., Seidel, A., Norden, B., and Graslund, A. (1998) *Biochemistry 37*, 4664−4673.
54. Dzantiev, L., and Romano, L. J. (1999) *J. Biol. Chem. 274*, 3279−3284.
55. Rechkoblit, O., Amin, S., Geacintov, N. E. (1999) *Biochemistry 38*, 11834−11843.
56. Feng, B., Gorin, A., Hingerty, B. E., Geacintov, N. E., Broyde, S., and Patel, D. J. (1997) *Biochemistry 36*, 13769−13779.

BI000552U