Sequence Dependence and Characteristics of Bends Induced by Site-Specific Polynuclear Aromatic Carcinogen−Deoxyguanosine Lesions in Oligonucleotides†

Hong Tsao,‡ Bing Mao,‡,§ Ping Zhuang,‡ Rong Xu,‡,|| Shantu Amin,⊥ and Nicholas E. Geacintov*,‡

Chemistry Department, New York University, New York, New York 10003-5180, and American Health Foundation, Valhalla, New York 10595

Received February 4, 1998

ABSTRACT: The tumorigenic metabolite of benzo[*a*]pyrene, the (+)-7*R*,8*S*,9*S*,10*R* enantiomer, and the nontumorigenic mirror-image isomer, (−)-7*S*,8*R*,9*R*,10*S*, of *r*7,*t*8-dihydroxy-*t*9,10-epoxy-7,8,9,10-tetrahydrobenzo[*a*]pyrene (*anti*-BPDE) bind covalently to the exocyclic amino group of deoxyguanosine (N^2-dG) in native DNA. These adducts can cause structural perturbations such as DNA bends, which in turn may influence the cellular processing of these lesions. The characteristics of bends in site-specifically modified oligodeoxyribonucleotide duplexes induced by single (+)- and (−)-*anti*-[BP]-N^2-dG lesions were examined by self-ligation and gel electrophoresis techniques. The modified residues (dG*) were centrally positioned in the 11-mer oligonucleotide d(CACAXG*XACAC) complexed with the natural complementary strands, with X = T or C, or in oligonucleotides 16 or 22 base pairs long with the same centrally positioned 11-mer. Among the four stereochemically distinct lesions, the 10*S* (+)-*trans*-*anti*-[BP]-N^2-dG adducts were significantly more bent than any of the other three stereoisomeric adducts and were selected for detailed studies. In the TG*T sequence context (X = T), the retardation factor R_L (apparent length of multimer/sequence length) is approximately independent of the phasing (distance, in base pairs, between the lesions) of the adducts with respect to the helical repeat (10.5 base pairs/helix turn). In contrast, in the CG*C sequence context (X = C), R_L is markedly lower in the case of ligated 16-mers than in the case of ligated 11-mer duplexes. The dependence of R_L on the phasing of the bends as a function of the helical repeat, indicate that the bends associated with (+)-*trans*-*anti*-[BP]-N^2-dG lesions are relatively rigid in the d(...CG*C...)·d(...GCG...) sequences, and flexible in the d(...TG*T...)·d(...ACA...) sequence context. These differences are attributed to the orientations of the pyrenyl residues on the 5′-side of the modified deoxyguanosine residues in the minor groove and to the intrinsic roll and tilt characteristics of DNA dinucleotide steps CG, GC, TG, and GT. The influence of flanking bases on the extent and character of DNA bending suggest that base sequence effects may be important in the cellular processing of (+)-*trans*-*anti*-[BP]-N^2-dG lesions.

Structural distortions such as bends and flexible hinge joints induced by the binding of chemical carcinogens to DNA may affect the interactions of the modified DNA with repair, replicative, and other cellular proteins. The processing of carcinogen-modified DNA sequences in vivo, if not repaired, may result in mutations and, ultimately, in cancer (*1*). Examples of covalent mutagen− or carcinogen−DNA adducts that cause bending include O^4-alkylthymine (*2*) and 1,N^2-propano-2′-deoxyguanosine lesions (*3*). Adducts derived from the binding to DNA of the aromatic amines 2-aminofluorene (AF) and *N*-acetyl-2-aminofluorene (AAF) give rise to flexible hinge joints (*4*). Metabolites of benzo[*a*]pyrene that can bind covalently to DNA also cause bends or flexible hinge joints at the sites of the lesions (*5*−*7*), and the extent of bending is markedly dependent on the stereochemical characteristics of these adducts (*8*).

Benzo[*a*]pyrene, a ubiquitous environmental carcinogen (*9*), is itself not biologically active, but is metabolized by microsomal mixed-function oxidases to isomeric benzo[*a*]pyrene diol epoxide derivatives of varying tumorigenic and mutagenic activities (*10*). Of particular interest for structure−function studies are the (+)-7*R*,8S,*9S*,*10R* and the (−)-7*S*,8*R*,9*R*,10*S* enantiomers of 7*r*,8*t*-dihydroxy-*t*9,10-epoxy-7,8,9,10-tetrahydrobenzo[*a*]pyrene [(+)- and (−)-*anti*-BPDE].[1] The (+)-enantiomer is highly tumorigenic in rodent

† This research was supported by a grant from the Office of Health and Environmental Research, Department of Energy (DE-FGO2-86ER60405), and in part by the NIH, National Cancer Institute (Grant CA 20851). At the American Health Foundation, the work was supported by NIH Grants CA-17613 (S.A.) and NCI-CB-77022−75. The benzo[*a*]pyrene diol epoxides were supplied by the National Cancer Institute Carcinogen Reference Repository.

‡ New York University.

§ Present address: Memorial Sloan-Kettering Cancer Center, New York, NY 10021.

|| Present address: Rockefeller University, Box 163, New York, NY 10021.

⊥ American Health Foundation.

[1] Abbreviations: (+)-*anti*-BPDE, (+)-7*R*,8*S*-dihydroxy-9*S*,10*R*-epoxy-7,8,9,10-tetrahydrobenzo[*a*]pyrene; (−)-*anti*-BPDE, (−)-7*S*,8*R*-dihydroxy-9*R*,10*S*-epoxy-7,8,9,10-tetrahydrobenzo[*a*]pyrene; (+)-*trans*- or (+)-*cis*-*anti*-[BP]-N^2-dG adducts, covalent products derived from the binding of (+)-*anti*-BPDE at its C10 position to the N^2-2′-deoxyguanosine residues in DNA by trans or cis addition, while adduct designations preceded by (−) arise from a similar reaction with the (−)-*anti*-BPDE stereoisomer.

S0006-2960(98)00291-8 CCC: $15.00
Published on Web 03/20/1998

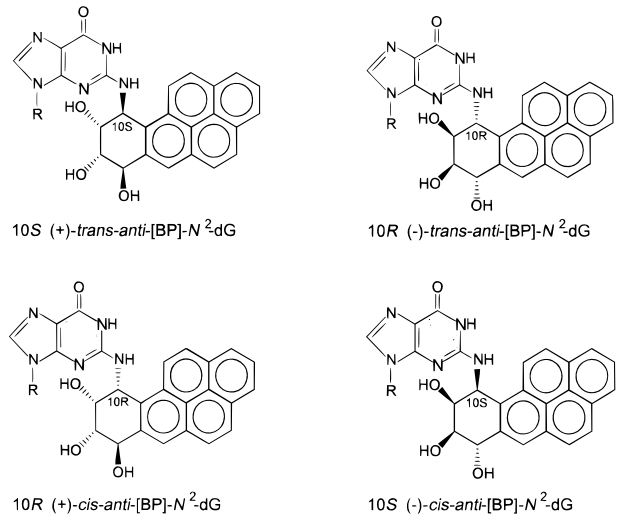

FIGURE 1: Structures of stereoisomeric *anti*-[BP]-N^2-dG adducts.

model systems, while the mirror-image (−)-enantiomer is not (*11*, *12*); furthermore, (+)-*anti*-BPDE is more mutagenic in mammalian cells than (−)-*anti*-BPDE (*13*, *14*). Both enantiomers form DNA adducts via cis or trans addition of N^2-dG in native DNA to the C10 position of *anti*-BPDE (*15*, *16*), giving rise to four stereoisomeric adducts (Figure 1). The mutagenic potentials of some of these lesions in vivo in different site-specific sequence contexts are beginning to be explored (*17*−*19*).

The solution structures of the four *anti*-[BP]-N^2-dG lesions (G*) in the sequence context 5′−d(...CG*C...)•d(...GCG...) have been investigated by high-resolution NMR techniques and are characterized by remarkably different stereochemistry-dependent conformations (*20*−*24*). In the two (+)- and (−)-*trans* adducts, the pyrenyl residues are positioned in the minor groove, with opposite orientations relative to the modified dG residues; in contrast, the (+)- and (−)-*cis* adducts are characterized by base-displaced intercalative conformations, again with opposite orientations relative to the major and minor grooves of the duplexes. Fountain and Krugh (*25*) found that the (+)-*trans*-*anti*-[BP]-N^2-dG lesions in the sequence context 5′−d(...TG*C...)•(...GCA...) are characterized by a predominant minor groove structure of the type observed by Cosman et al. (*20*) and a minor, presumably carcinogen-base stacked conformer. These two conformers appear to be in conformational equilibrium with one another (*25*), which may be associated with a local, sequence-dependent flexibility (*24*). Thus, a 5′-flanking dT residue introduces a conformational heterogeneity that is not observed in the presence of a dC residue flanking dG* on the 5′-side (*20*).

In this work, using gel electrophoresis techniques and the phasing methods developed for studying bending in $(dA)_n$•$(dT)_n$ tracts (*26*−*28*), the characteristics of bends associated with the (+)-*trans*-*anti*-[BP]-N^2-dG lesions flanked either by two dC or by two dT residues in otherwise similar sequence contexts have been examined. It is shown that dT residues flanking the (+)-*trans*-*anti*-[BP]-N^2-dG lesions give rise to flexible bends, while flanking dC residues give rise to predominantly rigid bends. These observations suggest that sequence effects can significantly affect the local bending or flexibility induced by (+)-*trans*-*anti*-[BP]-N^2-dG lesions and may thus facilitate the interconversion from one adduct conformation to another. In turn, these effects may play a role in the expression of the mutagenic potentials of these adducts (*25*).

Table 1: BPDE-Modified DNA Duplex Sequences Investigated

designation	sequence
TG*T-11	**5'-TGTAC ATGTGG CACATG*TACAC-5'**
TG*T-16	**5'-AGTGTAC ATGTGAGAG CTCACATG*TACACTCT**
TG*T-22	**5'-GAGAGTGTAC ATGTGAGAGTGG TCTCTCACATG*TACACTCTCAC-5'**
CG*C-11	**5'-TGTGC GTGTGG CACACG*CACAC-5'**
CG*C-16	**5'-AGTGTGC GTGTGAGAG CTCACACG*CACACTCT-5'**

MATERIALS AND METHODS

The oligodeoxyribonucleotide sequences studied in this work are shown in Table 1. By use of the twist angles provided by Kabsch et al. (*29*), the helical repeat of the TGT-11 and CGC-11 sequences can be estimated and are ∼10.4 and 10.5 base pairs/per helical turn, respectively. Thus, the sequences 11, 16, and 22 base pairs long are approximately 1, 1.5, and 2 helical turns in size, respectively.

All oligonucleotides were purchased from Midland Certified Reagents (Midland, TX), repurified by reverse-phase HPLC methods (*30*), and desalted using a Sephadex G25 column. The BPDE-modified oligonucleotides were generated by a direct synthesis method (*31*) using racemic *anti*-BPDE obtained from the National Cancer Institute Carcinogen Reference Repository. The procedures used and the methods of characterization and verification of adduct stereochemistry have already been described in detail in the case of the TG*T-11 (*32*), the TG*T-16 (*33*), and the TG*T-22 sequences (*34*).

About 2 μg of each unmodified and BPDE-modified oligonucleotide were 5′-end-labeled with [γ-^{32}P]ATP (New England Nuclear) using 10 units of T4 polynucleotide kinase (Gibco Biological Research Laboratories). After labeling with [γ-^{32}P]ATP, 2 μL of 0.1 M cold ATP (Pharmacia) and 10 units of T4 polynucleotide kinase were added, and the reaction was continued at 37 °C for an additional hour. The labeled unmodified and BPDE-modified single-stranded oligonucleotides were repurified using denaturing 20% polyacrylamide gel electrophoresis (7 M urea). The complementary strands were end-labeled with cold ATP using 3 μL of a 0.1 M ATP solution and 20 units of kinase. The end-labeled and purified oligonucleotides were mixed with complementary strands designed to produce cohesive single-stranded ends (Table 1), heated at 70 °C for 10 min, and allowed to cool slowly to 8 °C overnight, thus forming duplexes. The basic procedure of Koo et al. (*28*), with minor modifications, was used for ligating the different oligonucleotide duplexes (complementary:modified strand ratio 1.3:1). About 0.5 μg of these annealed double-stranded oligonucleotides were incubated in 45 μL of ligation buffer solution (50 mM Tris-HCl, 10 mM $MgCl_2$, 0.1 mM ATP, and 0.5 mM dithiothreitol, pH 7.6) and 9 units of T4 ligase (Gibco Biological Research Laboratories) at 12 °C for at least

24 h. The ligated multimers were subjected to electrophoresis on nondenaturing 8% polyacrylamide gels in 0.089 M Tris−borate (pH 8.3) and 2 mM EDTA (TBE buffer) at 4 °C. The relative intensities of each of the bands were analyzed using a Bio-Rad GS-525 imaging system (Bio-Rad, Hercules, CA).

RESULTS

In this work we have focused on a set of CG*C and TG*T sequences in which the (+)-*trans-anti*-[BP]-N^2-dG lesions are flanked on both sides either by dTs or by dCs; in turn, the central ...TG*T... or ...CG*C... residues are flanked by identical CACA or ACAC sequences on the 5′- and 3′-sides, respectively. The hypothesis was tested that dA·dT or dG·dC base pairs flanking the lesions influence the character of the BPDE adduct-induced bending and thus the conformational characteristics of the *anti*-[BP]-N^2-dG lesions. The pairs of sequences, TG*T-11 and TG*T-16 and CG*C-11 and CG*C-16, are ideal for this purpose because they are identical to one another except for the two bases flanking the BPDE-modified deoxyguanosyl residue dG*. Identical base sequences flanking the CGC and TGT base triplets were selected because sequence contexts involving elements greater than nearest-neighbor bases can also influence the overall bending properties of DNA *(35)*.

We have selected the (+)-*trans-anti*-[BP]-N^2-dG lesions for detailed studies of base sequence effects because they cause significantly more bending than the stereoisomeric (−)-*trans*-, (+)-*cis*-, and (−)-*cis-anti*-[BP]-N^2-dG lesions (Figure 2 and Figure S2, Supporting Information).

*CG*C Sequence Context.* An autoradiograph of a nondenaturing electrophoresis gel of ligation products of the unmodified and BPDE-modified sequences CG*C-11 is shown in Figure 2. The single band at the bottom of the gel in lane 1 is due to the unligated and unmodified 11-mer duplex. Lane 2 shows the results obtained upon ligating the unmodified 11-mer, while lanes 3−6 represent the ligation ladders derived from the *anti*-BPDE-modified CG*C-11 oligonucleotide duplexes with different adduct stereochemical properties: G* = 10*S* (+)-*trans*-, 10*R* (−)-*trans*-, 10*R* (+)-*cis*-, and 10*S* (−)-*cis-anti*-[BP]-N^2-dG, respectively. The BPDE-modified 11-mer oligomers with adducts characterized by 10*R* stereochemistry (lanes 4 and 5) are not as effectively ligated as the adducts with 10*S* absolute configuration (lanes 3 and 6). In lane 3, showing the mobilities of ligation products containing (+)-*trans* lesions, three dark bands (arrows) are observed near the top of the gel; two-dimensional gel experiments (*36*) indicate that these dark bands are due to circularized ligation products (data not shown). The densitometer traces of each of the lanes in Figure 2 are shown in Figure S1 (Supporting Information); the three bands attributed to the circular ligation products are clearly distinguishable from the background and the other, linear ligation products in that figure. As in the case of the BPDE-modified TG*T-11 oligonucleotides (*8*), the CG*C-11 oligonucleotides with (+)-*trans-anti*-[BP]-N^2-dG lesions exhibit the slowest electrophoretic mobilities (Figure 2).

An autoradiograph of a gel with CG*C-16 ligation products is shown in Figure 3. In contrast to the CG*C-11 oligomers (Figure 2), all four stereochemically distinct BPDE-modified 16-mers appear to be ligated efficiently. The

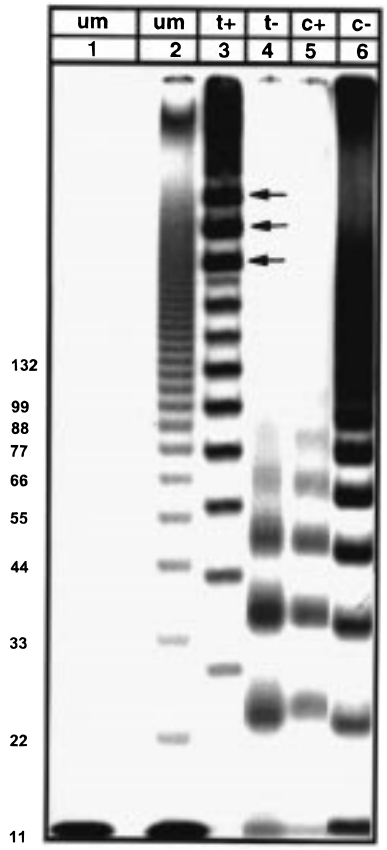

FIGURE 2: Ligation ladders of unmodified CGC-11 duplexes (lane 2) and BPDE-modified CG*C-11 duplexes with 10*S* (+)-*trans*- (lane 3), 10*R* (−)-*trans*- (lane 4), 10*R* (+)-*cis*- (lane 5), and 10*S* (−)-*cis* adduct stereochemistry (lane 6). The single band in lane 1 is due to an unligated standard 11-mer duplex. The three dark bands near the top of lane 3 (arrows) indicate the formation of circular ligation products. The ladder (numbers on the left margin of the figure) refer to the positions of the *unmodified* CGC-11 multimers (lane 2) of the designated sizes expressed in numbers of base pairs.

10*R* or 10*S* adduct stereochemistry appears to exert less influence on the efficiencies of ligation in the case of the 16-mers than in the case of the 11-mers. This difference is attributed to the greater distance between the cohesive ends of the molecules, the sites of ligation, and the BPDE-modified deoxyguanosine residues in the 16-mers than in the 11-mers.

While all of the BPDE-modified oligonucleotide fragments migrate more slowly than the unmodified oligonucleotides of the same sequence lengths, the DNA molecules with (+)-*trans-anti*-[BP]-N^2-dG lesions within the CG*C-16 ligation products again exhibit the slowest electrophoretic mobilities (Figure 3, lane 4).

*TG*T Sequence Context.* The autoradiograph of the ligation products of the TG*T-16 sequences are shown in Figure 4. The remarkably lower electrophoretic mobilities of the different ligated TG*T-16 sequences with (+)-*trans-anti*-[BP]-N^2-dG lesions are clearly maintained even in the case of the higher multimers. The asterisks in lanes 1, 2, and 3 in Figure 4 denote bands with the same sequence lengths of 160 base pairs. There are only minor differences in the positions of the bands of the unmodified TGT-16 and the (−)-*trans-anti*-BPDE-modified TG*T-16 10-mer sequences, but the distance migrated by (+)-*trans-anti*-BPDE-modified TG*T-16 multimer, 160 bp long is significantly

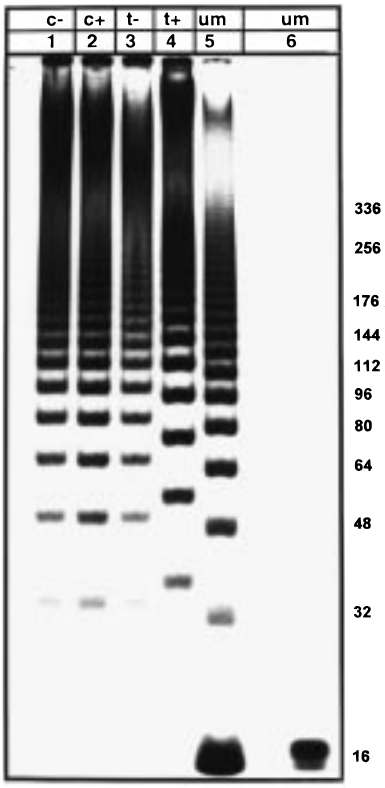

FIGURE 3: Ligation ladders of CG*C-16 duplexes with 10*S* (−)-*cis*- (lane 1), 10*R* (+)-*cis*- (lane 2), 10*R* (−)-*trans*- (lane 3), and 10*S* (+)-*trans* adduct stereochemistry (lane 4). (Lane 5) Ligation ladders for unmodified CGC-16 duplexes. The single band in lane 6 is due to an unligated CGC-16 duplex. The ladder (numbers on the right margin of the figure) refer to the positions of the unmodified CGC-16 multimers (lane 5) of the designated sizes, expressed in numbers of base pairs.

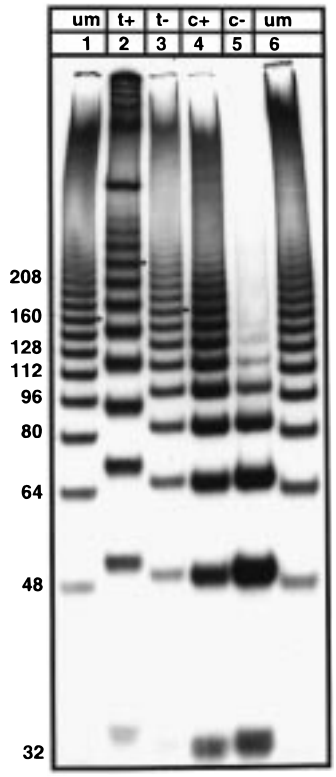

FIGURE 4: Ligation ladders of unmodified TGT-16 duplexes (lanes 1 and 6) and BPDE-modified TG*T-16 duplexes with 10*S* (+)-*trans*- (lane 2), 10*R* (−)-*trans*- (lane 3), 10*R* (+)-*cis*- (lane 4), and 10*S* (−)-*cis* adduct stereochemistry (lane 5). The asterisks in lanes 1, 2, and 3 indicate the ligated 10-mers of the TGT-16 or TG*T-16 sequences. The ladder (numbers on the left margin of the figure) refer to the positions of the unmodified TGT-16 multimers (lanes 1 and 6) of the designated sizes expressed in numbers of base pairs.

smaller (lane 2). The mobilities of oligonucleotide sequences with (+)-*cis*- and (−)-*cis*-*anti*-[BP]-N^2-dG lesions are close to those of the unmodified TGT-16 sequences of the same lengths (lanes 4, 5, and 6).

The ligation products of TG*T-11 (*8*) and TG*T-22 sequences with G* = (+)-*trans*-*anti*-[BP]-N^2-dG lesions (Figure S2, Supporting Information) exhibit gel electrophoretic patterns that are similar to those shown for ligated TG*T-16 sequences (Figure 4).

Retardation Factor. This factor is defined as the ratio of the apparent length of the oligonucleotide (determined by gel electrophoresis) to its true sequence length (*26*, *27*):

$$R_L = \frac{\text{(apparent length of multimer)}}{\text{(sequence length of multimer)}} \qquad (1)$$

The apparent length of a BPDE-modified multimer of a given sequence length (expressed in terms of the number of base pairs per molecule), is estimated from the distances migrated by this molecule and the sequence length of an *unmodified multimer* with the same migration distance. The BPDE-modified multimers are bent and migrate more slowly than the unmodified multimers of the same sequence length, and thus $R_L > 1.0$.

Plots of R_L vs the number of base pairs per molecule for TG*T-11, TG*T-16, and TG*T-22 oligonucleotides with (+)-*trans*-*anti*-[BP]-N^2-dG lesions are shown in Figure 5A, while those for the ligated CG*C-11 and CG*C-16 sequences

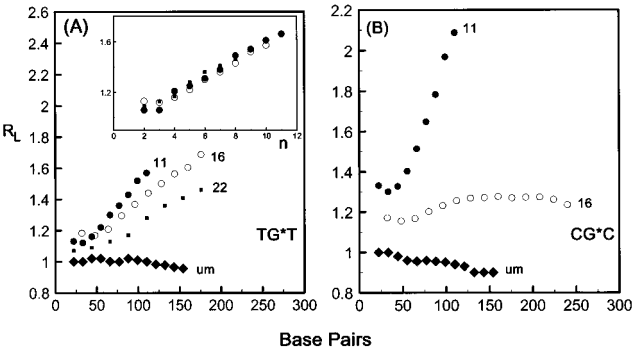

FIGURE 5: Plots of the R_L-ratio for duplexes with 10*S* (+)-*trans*-*anti*-[BP]-N^2-dG lesions as a function of the number of base pairs per multimer; the inset in (A) is a plot of R_L as a function of *n*, the number of ligated monomer units per multimer (or, the number of bends per molecule, see text). (A) -TGT- sequences (●: TG*T-11 (from ref 8), ○: TG*T-16, and ■: TG*T-22). (B) -CGC- sequences (●: CG*C-11, and ○: CG*C-16). The R_L values for the unmodified TC*T-11 (A) and CG*C-11 (B) ligation products (◆) relative to ligated *Bam*HI 10-mer linkers are also shown. The apparent lengths of the TG*T-11 and CG*C-11 multimers were deduced from the mobility patterns of the *Bam*HI linker ligation ladders in the same gels (Figure S4, Supporting Information), as described in more detail in the Supporting Information.

are shown in Figure 5B. In the case of the TG*T sequences (Figure 5A), the values of the R_L factors decrease with increasing size of the modified oligonucleotide sequence. This effect is associated with the decreasing mean number of lesions per helical turn (*37*), which is ∼1.0, 0.66, and 0.5 in the case of TG*T-11, TG*T-16, and TG*T-22 oligomers,

respectively. When the R_L values are replotted as a function of *n*, the number of ligated monomer units (or bending loci) per molecule, the data points for the sequences with ~1, ~1.5, and ~2 helical turns/monomer unit are superimposed on one another within experimental error (inset, Figure 5A). Thus, R_L is approximately independent of the number of adducts/helical turn, or the phasing of the bends relative to the helical repeat, but does increase as the number of adducts per multimer molecule increases.

While the anomalously slow electrophoretic mobility is maintained in the case of the TG*T-16 sequences, approximately 1.5 helical turns in size, there is a pronounced phasing effect in the case of the CG*C-11 and CG*C-16 sequences containing the same (+)-*trans-anti*-[BP]-N^2-dG lesions: the maximum values of R_L decrease from ~2.2 in the case of the ligated 11-mers to ~1.2 in the case of the ligated 16-mers (Figure 5B). Inspection of this figure indicates that the R_L values for the CG*C-11 and CG*C-16 multimers would still be quite different from one another even if R_L were plotted as a function of *n*.

Mobilities of Unmodified TGT-11 and CGC-11 Ligation Products Relative to BamHI Linker Size Standards. We examined the possibility that the TGT and CGC sequences used here are characterized by unusual mobilities, even in the absence of *anti*-[BP]-N^2-dG lesions. Xu et al. (*38*) showed that the electrophoretic mobilities of the unmodified TGT-11 ligation products are comparable to those of commercially available DNA size marker ladders (*38*). To further explore this question, we compared the mobilities of multimers derived from the ligation of unmodified TGT-11 and CGC-11 with those derived from the ligation of the *Bam*HI 10-mer linker 5′-d(CGGGATCCCG). The *Bam*HI ladders are often used as DNA size markers [see, for example, ref *28*]. These studies are described in detail in the Supporting Information. To summarize, the mobilities of the TGT-11 and CGC-11 multimers differ by at most 10% from those of the *Bam*HI ligation products of similar size. The R_L values of the unmodified multimers [(apparent length of the CGC-11 or TGT-11 multimers)/(sequence length)] are also shown in Figure 5; the apparent lengths of these multimers were determined from the sequence lengths of *Bam*HI multimers with the same electrophoretic mobilities. The unmodified TGT-11 or CGC-11 multimers either have the same or somewhat higher electrophoretic mobilities than the *Bam*HI linkers since $0.9 \geq R_L \leq 1.0$ (Figure 5). Since the mobilities of the unmodified TGT-11 and CGC-11 are not significantly different from those of the *Bam*HI linkers, the unmodified TGT-11, TGT-16, CGC-11, and CGC-16 multimers have been used as size markers in Figures 2-4. The R_L values of all of the BPDE-modified multimers have been evaluated relative to the mobilities of the *unmodified* multimers of the same base composition and sequence in each case.

DISCUSSION

Phasing Relationships and Carcinogen-Induced Bending. The bending of an oligonucleotide duplex is associated with a shorter end-to-end distance and thus to a lower electrophoretic mobility (*39, 40*). Upon ligation, the effects of static bends add coherently when the distances between the loci of the bends are in phase with the helical repeat, to produce multimers with the largest macroscopic curvature and the lowest electrophoretic mobilities (*26*−*28*). However, when the bends are out of phase with the helical repeat, the individual bends no longer add coherently to produce a macroscopic curvature, and both the average end-to-end distances of the ligated multimers and their electrophoretic mobilities increase, thus producing smaller values of R_L. However, if the bends are isotropically flexible (flexible hinges), the anomalous electrophoretic mobilities are observed regardless of the phasing of the bends with respect to the helical repeat (*27, 28, 41*). Thus, the behavior of the retardation factor R_L for ligated sequences with bends that are either in or out of phase with the helical repeat, can be used to distinguish between directed bends and flexible hinge joints (*4, 27, 28, 41, 42*). For a succinct discussion of flexible and static bends, the reader is referred to the paper by Hagerman (*27*).

*Differences in the Characteristics of the Bends in CG*C and TG*T Sequence Contexts.* When normalized to the same number of adducts per multimer molecule, *n*, in the TG*T sequence context, the anomalous R_L factors are similar in value in the case of ligated monomer units 1.0, 1.5, and 2.0 helical turns in size (inset, Figure 5A). Thus, when the (+)-*trans*- adducts are flanked by thymidine residues on both sides, the local structural distortions are due to *flexible hinge joints* rather than to static bends.

In contrast to the TG*T sequences, the CG*C sequences exhibit a marked effect of phasing of the bends with the helical repeat (Figure 5B). The anomalous electrophoretic mobility, expressed in terms of the factor $R_L - 1$ (*37*), is ~5 times greater for the CG*C-11 than for the CG*C-16 sequence. These effects suggest that the anomalous electrophoretic mobilities associated with the (+)-*trans-anti*-[BP]-N^2-dG lesions flanked by dC residues on both sides are due predominantly to *static bending* rather than to flexible hinge joints. However, in the ligated CG*C-16 sequences, corresponding to roughly 1.5 turns/bend, the anomalous electrophoretic mobility effect is not completely eliminated ($R_L \approx 1.25$); this suggest that there is also some flexibility associated with the lesions in the ...CG*C... sequence context.

The Dinucleotide Step Model of Bending. The base sequence effects and the differences in the characteristics of the bends associated with (+)- and (−)-*trans-anti*-[BP]-N^2-dG lesions flanked either by T or C on either side can be examined in terms of the well-known dinucleotide step models of bending of unmodified DNA (*43*−*55*). In these nearest-neighbor approximations, the effects of longer-range interactions are neglected, and the roll, tilt, and twist of the base pairs are assumed to be the dominant conformational parameters that govern DNA bending. Both static and dynamic wedge models and the cumulative effects of individual dinucleotide steps provide good descriptions of the bending of a variety of DNA sequences in gel electrophoresis experiments (*46, 51, 53*) and crystal structures (*54, 55*).

*Bending Induced by (+)- and (−)-trans Adducts in the CG*C Sequence Context.* As initially predicted by Ulyanov and Zhurkin (*47*), the CG steps are found to be most frequently bent toward the major groove (opening of the minor groove), while GC steps are frequently bent toward the minor groove (*55*). This characteristic sequence-dependent tendency for bending toward the major or minor grooves

suggests a mechanism for the observed differences in the bending of the CG*C-11 sequences with (+)- and (−)-*trans-anti*-[BP]-N^2-dG lesions. In the case of the (+)-*trans* adduct in the CG*C-11 sequence, the bulky pyrenyl residue on the 5′-side of G* in the minor groove (*20*), may reinforce the already natural tendency for bending into the major groove at CG dinucleotide steps. In contrast, unmodified GC steps are characterized by anisotropic opening of the major groove and a compression of the minor groove (*55*). Therefore, in the case of the (−)-*trans-anti*-[BP]-N^2-dG adducts, the pyrenyl residues positioned on the 3′-side of G* in the minor groove (*23*) may hinder the natural bending into the minor groove associated with this G*C dinucleotide step. Therefore, less bending is expected in the case of the (−)-*trans*- than in the case of the (+)-*anti*-[BP]-N^2-dG lesions in the CG*C-11 sequences, as observed experimentally.

*Flexibility Associated with trans-[BP]-N^2-dG Lesions in the TG*T Sequence Context.* The flexibility associated with ...TG*T... sequences and the relative rigidity of the ...CG*C... sequences may be associated with the properties of the TG* and CG* dinucleotide steps flanking the BPDE-modified G* residues. For example, the dinucleotide TG/CA step greatly enhances the flexibility of a DNA sequence when it replaces a CG/CG step (*56*): the TG/CA dinucleotide sequence increases the flexibility of a 21-mer oligonucleotide containing the O_R3 recognition site for the Cro protein of λ phage and facilitates the circularization of these sequences in the absence and presence of protein binding. Nagaich et al. (*57*) have shown that oligonucleotides with a TG/CA step at the 5′-end of oligo(A) tracts exhibit considerably slower electrophoretic mobilities than those with GT/AC and other dinucleotide steps and attributed this behavior to a kink. The latter are defined as abrupt deflections of the local helical axis and are associated with a loss of the normal stacking interactions between neighboring bases; Nagaich et al. (*57*) concluded that if the kink is sufficiently pronounced, it becomes a flexible hinge joint.

Other indications of the greater flexibilities associated with the TG/CA and the greater rigidities the CG/CG steps, are the NMR solution structures of oligonucleotide sequences with site-specific (+)-*trans-anti*-[BP]-N^2-dG lesions in ...CG*C... and ...TG*C... sequence contexts. While a unique structure was defined with the pyrenyl residues positoned in the minor groove on the 5′-side of G* in the case of the (+)-*trans*-adducts in a CG*C sequence context (*20*), in a TG*C sequence context, two conformations in equilibrium with one another were found (*25*). In the case of the TG*T-11 sequence, the loss of imino proton connectivities at the TG* step and other indications (*38*), are consistent with a considerable structural perturbation within the 5′-(..ATG*..)• (..CAT..) portion of the TG*T-11duplex, and to the coexistence of several different conformations. Taken together (*25, 38*), these NMR results suggest that the local flexibility is greater in the TG*T than in the CG*C case, in agreement with our gel electrophoresis results. The base pair stacking interactions are likely to be reduced in the presence of the bulky (+)-*trans*-adduct spanning the 5′-TG* step in the TG*T-11 oligonucleotides (*38*), thus enhancing the already intrinsic flexibility associated with TG/CA dinucleotide steps (*48, 56*−*58*).

In contrast to the (+)-*trans-anti*-[BP]-N^2-dG lesions, the (−)-*trans-anti*-[BP]-N^2-dG lesions in the TG*T-11 multimer sequences are characterized by almost normal electrophoretic mobilities (*8*). The NMR solution structures of these sequences have not been determined, but stereochemical considerations suggest that a 3′-orientation, along the G*T dinucleotide step in the minor groove, is a likely conformation for these (−)-*trans-anti*-[BP]-N^2-dG adducts (*24*). The intrinsic flexibilities of pur-pyr are lower than those of pyr-pur dinucleotide steps (*48*), the roll angles are smaller for GT/AC than for TG/CA steps (*45, 52*), and the thermal stability of GT/AC steps seems to be significantly higher than those of TG/CA steps (*57, 59*). All of these factors suggest a higher resistance to the unstacking of bases at the GT step than at the TG step, and thus to a lower degree of bending associated with (−)-*trans*- than (+)-*trans-anti*-[BP]-N^2-dG lesions in the TG*T-11 modified oligonucleotides.

Smaller Extents of Bending Associated with Base-Displaced Intercalative (+)-cis- and (−)-cis Lesions. In the CG*C-11 oligonucleotides, the (+)- and (−)-*cis-anti*-[BP]-N^2-dG moieties exert significantly smaller effects on the electrophoretic mobilities of the modified CG*C-11 sequences than the (+)-*trans-anti*-[BP]-N^2-dG lesions (Figure 2) since R_L does not exceed values of ∼1.2 for the two *cis* adducts (Figure S2, Supporting Information). The same order of relative electrophoretic mobilities as a function of adduct stereochemistry was observed by Suh et al. (*60*) for the same set of stereoisomeric lesions in the 11-mer duplex d(CCATCG*CTACC)•d(GGTAGCGATGG). In the latter sequence, detailed NMR studies have shown that the pyrenyl residues are predominantly intercalated between the neighboring C•G base pairs with the modified deoxyguanosyl residues displaced into the minor groove in the case of the (+)-*cis*-[BP]-modified duplexes and into the major groove in the case of the (−)-*cis*-[BP]-modified duplexes (*21, 22*). Assuming that the adduct conformations are similar in our CG*C-11 sequences, these observations suggest that base-displaced intercalative *cis-anti*-[BP]-N^2-adduct conformations do not strongly affect the electrophoretic mobilities, and thus bending in DNA. Relatively small values of R_L (≤ 1.1) were also observed in the case of the ligated AAF-modified oligonucleotides (AAF−C8-dG lesions) in 15- and 20-mer oligonucleotides (*4*) in which the dominant adduct conformations (∼70%) also involve a base-displaced intercalation (*61*).

Bending in Oligonucleotides Induced by Other Covalently Bound Molecules. There are other examples of flexible or rigid bends associated with covalently bound drugs or carcinogens, but the effects of base sequence have not been systematically investigated. Oligonucleotides containing single O^4-alkylthymine lesions in the CT*TG or CTT*G sequence context also give rise to flexible hinge joints, but the R_L values are quite small (≤1.1) (*2*). The extent of bending is thus much smaller than observed by us with (+)-*trans-anti*-[BP]-N^2-dG lesions in the TG*T-11 sequence. The antitumor drug *cis*-diamminedichloroplatinum(II), which forms intrastrand cross-links between the N^7 atoms of adjacent deoxyguanosine (dG) residues, gives rise to a ∼40° bend (*41*) and an increase in the local flexibility of the double helix (*62*). Mitomycin C forms intrastrand cross-links via binding to the exocyclic amino groups of adjacent guanine residues; the adducts give rise to rigid bends of ∼15° since the R_L factor depends strongly on the phasing of the adducts as a function of the helical repeat (*63*). Minor groove N^2-dG adducts derived from the binding of the antitumor drugs

anthramycin and maymycin (*64*), and the antitumor antibiotic (+)-CC-1065 that binds to N^3 of deoxyadenosine residues (*65*) also cause DNA bending.

Biological Implications. The pronounced sequence effects reported here suggest that base sequence context may strongly influence the dynamics and overall shapes of carcinogen-modified DNA sequences and thus determine the processing of these lesions by cellular enzymes. As discussed in more detail in one of our other, related publications (*38*), the sequence-dependent flexibility and bending of carcinogen-modified DNA sequences could be important in the interactions of DNA with repair proteins (*66*, *67*), transcription factors (*68*, *69*), and replicative enzymes (*70*, *71*).

ACKNOWLEDGMENT

One of us (N.E.G.) would like to thank Professor W. Olson for some very helpful discussions.

SUPPORTING INFORMATION AVAILABLE

Six figures, showing the gel autoradiograph for the TGT-22 and TG*T-22 sequences (Figure S1), the density profiles of each of the lanes in Figure 2 (Figure S2), plots of the retardation factor R_L for the CG*C-11 sequence with (+)-*trans*-, (+)-*cis*-, (−)-*trans*-, and (−)-*cis*-*anti*-[BP]-N^2-dG lesions (Figure S3), the electrophoretic mobility patterns of unmodified and (+)-*anti*-*trans*-[BP]-N^2-dG-modified TGT-11 and CGC-11 multimers compared to those associated with ligated *Bam*HI linker size markers (Figure S4), a densitometer tracing of the gel patterns (Figure S5), and a comparison of the mobilities of the unmodified TGT-11, CGC-11, and *Bam*HI multimers (Figure S6) (9 pages). Ordering information is given on any current masthead page.

REFERENCES

1. Singer, B., and Grunberger, D. (1983) *Molecular Biology of Mutagens and Carcinogens*, Plenum Press, New York.
2. Georgiadis, P., Xu, Y.-Z., and Swann, P. F. (1991) *Biochemistry 30*, 11725−11732.
3. Moe, J. G., Reddy, R., Marnett, L. J., and Stone, M. P. (1994) *Chem. Res. Toxicol. 7*, 319−328.
4. Schwartz, A., Marrot, L., and Leng, M. (1989) *J. Mol. Biol. 207*, 445−450.
5. Hogan, M. E., Dattagupta, N. and Whitlock, J. P., Jr. (1981) *J. Biol. Chem. 256*, 4504−4513.
6. Eriksson, M., Nordén, Jernström, B., and Gräslund, A. (1988) *Biochemistry 27*, 1213−1221.
7. Roche, C. J., Geacintov, N. E., Ibanez, V., and Harvey, R. G. (1989) *Biophys. Chem. 33*, 277−288.
8. Xu, R., Mao, B., Xu, J., Li, B., Birke, S., Swenberg, C. E., and Geacintov, N. E. (1995) *Nucleic Acids Res. 23*, 2314−2319.
9. Phillips, D. H. (1983) *Nature 303*, 468−472.
10. Conney, A. H. (1982) *Cancer Res. 42*, 4875−4917.
11. Buening, M. K., Wislocki, P. G., Levin, W., Yagi, H., Thakker, D. R., Akagi, H., Koreeda, M., Jerina, D. M., and Conney, A. H. (1978) *Proc. Natl. Acad. Sci. U.S.A. 75*, 5358−5361.
12. Slaga, T. J., Bracken, W. J., Gleason, G., Levin, W., Yagi, H., Jerina, D. M., and Conney, A. H. (1979) *Cancer Res. 39*, 67−71.
13. Brookes, P., and Osborne, M. R. (1982) *Carcinogenesis 3*, 1223−1226.
14. Wei, S.-J. C., Chang, R. L., Hennig, E., Cui, X. X., Merkler, K. A., Wong, C.-Q., Yagi, H., and Conney, A. H. (1994) *Carcinogenesis 15*, 1729−1735.
15. Meehan, T., and Straub, K. (1979) *Nature 277*, 410−412.
16. Cheng, S. C., Hilton, B. D., Roman, J. M., and Dipple, A. (1989) *Chem. Res. Toxicol. 2*, 334−340.
17. Mackay, W., Benasutti, M., Drouin, E., and Loechler, E. L. (1992) *Carcinogenesis 13*, 1415−1425.
18. Jelinsky, S. A., Liu, T., Geacintov, N. E., and Loechler, E. L. (1995) *Biochemistry 34*, 13545−13553.
19. Moriya, M., Spiegel, S., Fernandes, A., Amin, S., Liu, T.-M., Geacintov, N. E., and Grollman, A. P. (1995) *Biochemistry 35*, 16646−16651.
20. Cosman, M., de los Santos, C., Fiala, R., Hingerty, B. E., Ibanez, V., Margulis, L. A., Live, D., Geacintov, N. E., Broyde, S., & Patel, D. J. (1992) *Proc. Natl. Acad. Sci. U.S.A. 89*, 1914−1918.
21. Cosman, M., de los Santos, C., Fiala, R., Hingerty, B. E., Ibanez, V., Luna, E., Harvey, R. G., Geacintov, N. E., Broyde, S., and Patel, D. (1993) *Biochemistry 32*, 4145−4155.
22. Cosman, M., Hingerty, B. E., Luneva, N., Amin, S., Geacintov, N. E., Broyde, S., and Patel, D. J. (1996) *Biochemistry 35*, 9850−9863.
23. 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.
24. Geacintov, N. E., Cosman, M., Hingerty, B. E., Amin, S., Broyde, S., and Patel, D. J. (1997) *Chem. Res. Toxicol. 10*, 111−146.
25. Fountain, M. A., and Krugh, T. R. (1995) *Biochemistry 34*, 3152−3161.
26. Wu, H. M., and Crothers, D. M. (1984) *Nature 308*, 509−513.
27. Hagerman, P. (1985) *Biochemistry 24*, 7033−7037.
28. Koo, H.-S., Wu, H.-M., and Crothers, D. M. (1986) *Nature 320*, 501−506.
29. Kabsch, W., Sander, C., & Trifonov, E. N. (1982) *Nucleic Acids Res. 10*, 1097−1104.
30. McLaughlin, L. W., & Piel, N. (1984) in *Oligonucleotide Synthesis: A Practical Approach* (Gait, M. J., Ed.) pp 117−133, IRL Press, Oxford, England.
31. Cosman, M., Ibanez, V., Geacintov, N. E., and Harvey, R. G. (1990) *Carcinogenesis 11*, 1667−1672.
32. Geacintov, N. E., Cosman, M., Mao, B., Alfano, A., Ibanez, V., and Harvey, R. G. (1991) *Carcinogenesis 12*, 2099−2108.
33. Mao, B., Li, B., Amin, S., Cosman, M., and Geacintov, N. E. (1993) *Biochemistry 32*, 11785−11793.
34. Mao, B. (1993) Ph.D. Dissertation, New York University.
35. Dlakic, M. and Harrington, R. E. (1996) *Proc. Natl. Acad. Sci. U.S.A. 93*, 3847−3852.
36. Ulanovsky, L., Bodner, M., Trifonov, E. N., and Choder, M. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 862−866.
37. Koo, H.-S., and Crothers, D. M. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 1763−1767.
38. Xu, R., Mao, B., Amin, S., and Geacintov, N. E. (1998) *Biochemistry 37*, 769−778.
39. Lerman, L. S., and Frisch, H. L. (1982) *Biopolymers 21*, 995−997.
40. Lumpkin, O. J., and Zimm, B. H. (1982) *Biopolymers 21*, 2315−2316.
41. Rice, J. A., Crothers, D. M., Pinto, A. L., & Lippard, S. J. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 4158−4161.
42. Leng, M. (1990) *Biophys. Chem. 35*, 155−163.
43. Ulanovsky, L. E., and Trifonov, E. N. (1987) *Nature 326*, 720−722.
44. Calladine, C. R., Drew, H. R., and McCall, M. J. (1988) *J. Mol. Biol. 201*, 127−137.
45. De Santis, P., Palleschi, A., Savino, M., and Scipioni, A. (1990) *Biochemistry 29*, 9269−9273.
46. Bolshoy, A., McNamara, Harrington, R. E., and Trifonov, E. N. (1991) *Proc. Natl. Acad. Sci. U.S.A. 88*, 2312−2316.
47. Ulyanov, N. B., and Zhurkin, V. B. (1984) *J. Biomol. Struct. Dyn. 2*, 361−385.
48. Sarai, A., Mazur, J., Nussinov, R., and Jernigan, R. L. (1989) *Biochemistry 28*, 7842−7849.

49. Zhurkin, V. B., Ulyanov, N. B., Gorin, A. A., and Jernigan, R. L. (1991) *Proc. Natl. Acad. Sci. U.S.A. 88*, 7046−7050.
50. Olson, W. K., Marky, N. L., Jernigan, R. L., and Zhurkin, V. B. (1993) *J. Mol. Biol. 232*, 530−554.
51. Olson, W. K., Babcock, M. S., Gorin, A., Liu, G., Marky, N. L., Martino, J. A., Pedersen, S. C., Srinivasan, A. R., Tobias, I., Westcott, T. P., and Zhang, P. (1995) *Biophys. Chem. 55*, 7−29.
52. Olson, W. K. and Zhurkin, V. B. (1996) *Biological Structure and Dynamics* (Sarma, R. H., and Sarma, M. H., Eds.) pp 341−370, Adenine Press, Guilderland, NY.
53. DeSantis, P., Fuà, Palleschi, A., and Savino, M. (1995) *Biophys. Chem. 55*, 261−271.
54. Berman, H. M, Olson, W. K., Beveridge, D. L., Westbrook, J., Gelbin, A., Demeny, T., Hseih, S. H., Srinivasan, A. R., and Schneider, B. (1992) *Biophys. J. 63*, 751−759.
55. Young, M. A., Ravishanker, G., Beveridge, D. L., and Berman, H. M. (1995) *Biophys. J. 68*, 2454−2468.
56. Lyubchenko, Y. L., Schlyakhtenko, L. S., Appella, E., and Harrington, R. E. (1993) *Biochemistry 32*, 4121−4127.
57. Nagaich, A. K., Bhattacharyya, D., Brahmachari, S. K., and Bansal, M. (1994) *J. Biol. Chem. 269*, 7824−7833.
58. McNamara, P. T., Bolshoy, A., Trifonov, E. N., and Harrington, R. E. (1990) *J. Biomol. Struct. Dyn. 8*, 529−538.
59. SantaLucia, J., Jr., Allawi, H. T., and Seneviratne, P. A. (1996) *Biochemistry 35*, 3555−3562.
60. Suh, M., Ariese, F., Small, G. J., Jankowiak, R., Liu, T.-M., and Geacintov, N. E. (1995) *Biophys. Chem. 56*, 281−296.
61. 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.
62. Brabec, V., Sip, M., and Leng, M. (1993) *Biochemistry 32*, 11676−11681.
63. Rink, S. M., Lipman, R., Alley, S. C., Hopkins, P. B., and Tomasz, M. (1996) *Chem. Res. Toxicol. 9*, 382−389.
64. Kizu, R., Draves, P. H., and Hurley, L. H. (1993) *Biochemistry 32*, 8712−8722.
65. Lee, C.-S., Sun, D., Kizu, R., and Hurley, L. H. (1991) *Chem. Res. Toxicol. 4*, 203-213.
66. Zou, Y., Liu, T. M., Geacintov, N. E., and Van Houten, B. (1995) *Biochemistry 34*, 13582−13593.
67. Hess, T. M., Gunz, D., Luneva, N., Geacintov, N. E., and Naegeli, H. (1997) *Mol. Cell. Biol. 17*, 7069−7076.
68. MacLeod, M. C., Powell, K. L., and Tran, N. (1995) *Carcinogenesis 16*, 975−983.
69. MacLeod, M. C., Powell, K. L., Kuzmin, V. A., Kolbanovsky, A., and Geacintov, N. E. (1996) *Mol. Carcinogen. 15*, 44−52.
70. Moriya, M., Spiegel, S., Fernandes, A., Amin, S., Liu, T., Geacintov, N., and Grollman, A. (1996) *Biochemistry 35*, 16646−16651.
71. Shukla, R., Liu, T., Geacintov, N. E., and Loechler, E. L. (1997) *Biochemistry 36*, 10256−10261.

BI980291C