Published in final edited form as:

J Mol Biol. 2010 June 11; 399(3): 397–409. doi:10.1016/j.jmb.2010.04.004.

Distant neighbor base sequence context effects in human nucleotide excision repair of a benzo[a]pyrene-derived DNA lesion

Yueqin Cai^a, Konstantin Kropachev^b, Rong Xu^{b,†}, Yijin Tang^b, Marina Kolbanovskii^b, Alexander Kolbanovskii^b, Shantu Amin^c, Dinshaw J. Patel^d, Suse Broyde^a, and Nicholas E. Geacintov^{b,*}

^a Department of Biology, New York University, New York, N.Y., 10003

^b Department of Chemistry, New York University, New York, N.Y., 10003

^c Department of Pharmacology, Penn State College of Medicine, Hershey, P.A., 17033

^d Structural Biology Program, Memorial Sloan-Kettering Cancer Center, New York, N.Y., 10021

Summary

The effects of non-nearest base sequences, beyond the nucleotides flanking a DNA lesion on either side, on nucleotide excision repair (NER) in extracts from human cells were investigated. We constructed two duplexes containing the same minor groove-aligned 10S (+)-*trans-anti*-B[a]P-*N*²-dG (G*) DNA adduct, derived from the environmental carcinogen benzo[a]pyrene (B[a]P): 5'-C-C-A-T-C-G*-C-T-A-C-C-3' (CG*C-I), and 5'-C-A-C3-A4-C5-G*-C-A-C-A-C-3' (CG*C-II). We utilized gel electrophoresis to compare the extent of DNA bending, and molecular dynamics (MD) simulations to analyze the structural characteristics of these two DNA duplexes. The NER efficiencies are 1.6 ± 0.2 times greater in the case of the CG*C-II than the CG*C-I sequence context in 135-mer duplexes. Gel electrophoresis and self-ligation circularization experiments revealed that the CG*C-II duplex is more bent than the CG*C-I duplex, while MD simulations showed that the unique -C3-A4-C5- segment in the CG*C-II duplex plays a key role. The presence of a minor groove-positioned guanine amino group, namely, the Watson-Crick partner to C3, acts as a wedge; facilitated by a highly deformable local -C3-A4- base step, this amino group allows the B[a]P ring system to produce a more enlarged minor groove in CG*C-II than in CG*C-I, as well as a local untwisting and enlarged and flexible Roll only in the CG*C-II sequence. These structural properties fit well with our prior findings that in the case of the family of minor groove 10S (+)-*trans-anti*-B[a]P-*N*²-dG lesions, flexible bends and enlarged minor groove widths (Cai *et al.* (2009) *J. Mol. Biol.*, **385**: 30–44) constitute NER recognition signals, and extend our understanding of sequence context effects on NER to the neighbors that are distant to the lesion.

Keywords

benzo[a]pyrenyl-guanine lesion; nucleotide excision repair; sequence-dependence; DNA bending; flexibility

*Corresponding Author: Nicholas E. Geacintov, (212) 998–8407, Fax (212) 998–8421, ng1@nyu.edu.

†Present address: New York Structural Biology Center, New York, N.Y., 10027

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Introduction

The nucleotide excision repair (NER) machinery for removing bulky DNA lesions plays a central role in preserving the genome of mammalian cells and excises a 24 – 32 oligonucleotide stretch containing the lesion from the damaged strand.¹ Different bulky DNA lesions are excised with relative efficiencies that vary over several orders of magnitude.^{1,2} Nearest-neighbor base sequence context, i.e., the bases flanking the lesions, constitutes a significant element that can modulate NER efficiencies.^{3–6} The role of more distant neighboring bases remains, however, a relatively unexplored frontier.

Here we investigated the impact on NER excision efficiency of base sequence context beyond the bases flanking a benzo[*a*]pyrene-derived lesion. Benzo[*a*]pyrene (B[*a*]P) is a well-known environmental carcinogen^{7,8} that is metabolized in mammalian cells to the diol epoxide (+)-(7*R*,8*S*,9*S*,10*R*)-7,8-dihydroxy-9,10-epoxy-7,8,9,10-tetrahydrobenzo[*a*]pyrene ((+)-*anti*-B[*a*]PDE).⁹ This intermediate reacts predominantly with the exocyclic amino group of guanine in DNA to form the 10*S* (+)-*trans-anti*-B[*a*]P-*N*²-dG (**G***) adduct, a dominant reaction product (Figure 1a). The structure of this adduct has been characterized by high resolution NMR methods in several sequence contexts; the bulky aromatic B[*a*]P residue is positioned in the B-DNA minor groove, with the B[*a*]P ring system oriented in the 5'-direction of the damaged base **G*** (Figure 1b).^{10,11} Our earlier studies of sequence context effects on mammalian NER focused on the impact of the nearest-neighbor bases flanking **G*** on both sides;^{3–5,11} we found that the relative NER susceptibility could vary by as much as four-fold, depending on the bases flanking **G***. It was shown that perturbed Watson-Crick hydrogen bonding, flexible bends, diminished stacking and enlarged minor grooves are structurally destabilizing characteristics that are detected by the NER machinery. Impaired base stacking as an important repair-recognition factor has been highlighted by Yang.^{12,13} The potential impact of conformational flexibility associated with DNA lesions on the initial recognition by NER factors was reviewed earlier,¹⁴ and examined by us in specific base sequence contexts containing the lesion **G***.^{3–6}

In this work we examined the effects of the more distant base sequence on relative NER efficiencies in the following two sequence contexts:

5'-d(CCATCG*CTACC) (CG*C-I)
d(GGTAGC GATGG)-5'

5'-d(CACACG*CACAC) (CG*C-II)
d(GTGTGC GTGTG)-5'

These two sequences were selected for several reasons. The physical characteristics of the 10*S* (+)-*trans-anti*-B[*a*]P-*N*²-dG adduct positioned at an external minor groove binding site of the CG*C-I duplex have been extensively studied.^{10,15,16} The 10*S* (+)-*trans-anti*-B[*a*]P-*N*²-dG adduct in the CG*C-II sequence context is also positioned at an external binding site,¹⁷ residing in the minor groove according to molecular dynamics (MD) simulations with free energy calculations.³ Gel electrophoresis experiments have shown that this adduct introduces bends into both the CG*C-I¹⁶ and CG*C-II duplexes,^{18,19} but the differences in the extent of deformation of the CG*C-I and CG*C-II duplexes induced by the same 10*S* (+)-*trans-anti*-B[*a*]P-*N*²-dG adduct have not been determined. Here, we explored the hypothesis that the extent of bending is more pronounced in CG*C-II than in CG*C-I, and that this difference might result in differences in NER dual incision efficiencies. A slower electrophoretic mobility of CG*C-II is expected because we introduced bendable 5'-TG/5'-

CA dinucleotide steps flanking the central CG**C*/GCG motifs on both sides. The 5'-TG/5'-CA dinucleotide step is known to be highly deformable²⁰ and to strongly modulate the curvature of DNA in solution.²¹ Furthermore, 5'-pyrimidine-purine steps are generally known to represent flexible hinge-joints that favor DNA-protein complex formation,²² and are more numerous in the CG**C*-II than the CG**C*-I duplex. The effects of the more distant base pairs on the sequence-dependent structural differences between duplexes CG**C*-I and CG**C*-II were investigated by molecular dynamic (MD) simulation methods that provide a basis for interpreting the observed sequence effects.

Our experimental results indicate that the extent of bending is indeed greater in CG**C*-II than in CG**C*-I duplexes, and that this difference is correlated with a reproducible 1.6 ± 0.2 -fold greater NER dual incision efficiency for the CG**C*-II duplex. We propose that these differences stem directly from the sequence differences in the two duplexes beyond the central CG**C*/GCG trimer segment, which is identical in both cases (Figure 1c). A greater enlargement of the minor groove in the CG**C*-II than in the CG**C*-I duplex due to the bulky B[a]P residue positioned in the minor groove, produces an enhanced local untwisting as well as an enlarged and flexible Roll, and thus greater extent of bending in CG**C*-II. The new results presented here extend our understanding of sequence context effects on NER to non-nearest base sequences that are distant from the bulky 10*S* (+)-*trans-anti*-B[a]P-*N*²-dG lesion. With recent findings that polymorphisms in human NER repair genes impact lesion-repair and cancer-susceptibility, understanding base-sequence context effects becomes particularly significant.²³

Results

The CG**C*-II duplex is incised more efficiently

The extent of NER dual incisions of the CG**C*-I and CG**C*-II double-stranded sequences, embedded in 135-mer duplexes in human HeLa cell extracts as a function of incubation time, are depicted in Figure 2. A typical gel is shown in Figure 2a. Densitometric analysis of the percent dual incision rates in each experiment (each lane) is defined as the ratio of radioactivity in the 24–32 oligonucleotide range divided by the total radioactivity in the entire lane. These results are compared to the dual incision rates obtained with the 10*R* (+)-*cis-anti*-B[a]P-*N*²-dG adduct in the CG**C*-I sequence context used as a positive control of NER activity.⁵ This (+)-*cis* adduct has a base-displaced/intercalated conformation²⁴ and is an excellent NER substrate.²⁵ The dual incision NER efficiencies of the (+)-*cis* adduct are similar to those of the (6-4)-pyrimidine-pyrimidone T$\leftrightarrow$T photodimer⁵ generated by irradiating DNA with UVC light. In each set of experiments with a particular cell extract, the percent incision at 40 min obtained with the 10*R* (+)-*cis-anti*-B[a]P-*N*²-dG adduct was assigned the value of 100 percent (a typical gel is shown in Figure S1, Supporting Information), and all other results are expressed relative to this standard value, thus correcting for the variations in NER activities of different cell extracts. Results from multiple experiments with different cell extracts and substrate samples are summarized in Figure 2b, together with the standard deviations of each set of time-dependent data points. The straight lines represent least square fits to the data points and the slopes correspond to the initial rates of dual incisions. Our results show that the same 10*S* (+)-*trans-anti*-B[a]P-*N*²-dG adduct is repaired with 1.6 ± 0.2 -fold higher efficiency (Figure 2c) in CG**C*-II than in CG**C*-I, revealing that the non-nearest base pairs have a small but significant effect on the dual incision efficiencies. Repeated experiments with cell extracts prepared at different times consistently yielded similar results within the experimental error bars shown in Figure 2b.

To place these distant neighbor base sequence effects in context with nearest neighbor effects, we compare the NER obtained with the CG**C*-II sequence with a duplex in which

the guanines flanking the lesion were replaced by thymines, with all other more distant bases being identical:

The TG*T-II sequence context is of special interest because the two T:A pairs adjacent to G* introduce enhanced local flexibility into this duplex.^{18,26} Furthermore, the absence of guanine amino groups adjacent to G*, which constrain and compete for space with the B[a]P rings in the minor groove, also contribute to the flexibility of the TG*T-II duplex.³ The NER repair efficiencies of duplexes containing the identical modified guanine G* in different sequence contexts are significantly different. The NER efficiency of the TG*T-II duplex relative to CG*C-II is 1.5 ± 0.2 and relative to CG*C-I, it is 2.4 ± 0.2 (Figure 2c and Supporting Information, Figure S2).

The 10S (+)-*trans-anti*-B[a]P-*N*²-dG lesion induces a more pronounced bend in CG*C-II than in the CG*C-I duplex

A bent DNA duplex has a shorter end-to-end distance than a linear duplex, and thus exhibits a lower electrophoretic mobility in native polyacrylamide gels.^{27,28} The extent and characteristics of the bends in A-tracts have been extensively studied by ligating monomeric oligonucleotides to multimers of different lengths and examining the electrophoretic mobilities in polyacrylamide gels.^{29,30} The extent of bending of different oligonucleotides can be revealed by ligating double-stranded oligonucleotide sequences that are ~ one helical turn in length (11-mers) to form multimers. The bending induced by B[a]P-*N*²-dG adducts can be evaluated by comparing the gel electrophoretic mobilities of the B[a]PDE-modified multimers to those of the unmodified oligonucleotide multimers of the same sequence and length in native gels.^{18,19} The extent of bending of each multimer is revealed by the retardation factor $R_L = (\text{apparent length of B[a]PDE-modified multimer})/(\text{sequence length})$. The R_L factor is equal to 1.0 if there are no bends, and > 1.0 for B[a]PDE-modified duplexes; the greater R_L values are due to the local distortions of the DNA caused by the lesions that induce a bend, and thus a slower electrophoretic mobility.¹⁹

The gel electrophoresis patterns of the ligation ladders for the unmodified CGC-I/CGC-II and modified CG*C-I/CG*C-II 11-mer duplexes are compared in Figure 3. The properties of the ligated CG*C-II duplexes are discussed in detail elsewhere.¹⁸ The ligation patterns depend in a complex manner on the experimental conditions such as ligase concentration, starting DNA concentration, etc.³¹ Both linear and circularized ligation products are observable upon the ligation of B[a]PDE-modified DNA duplexes, as detailed in the following section.

Linear ligation products—The ligation experiments depicted in Figure 3a in the case of CG*C-I were conducted at 20 μM , and also at 4 μM starting oligonucleotide concentrations as shown in Figure S3, Supporting Information. A densitometric plot of the CG*C-I ligation ladder (lane 1) is shown in Figure 4a. Plots of the R_L factor for the linear multimers are shown in Figure 4b, and are practically superimposable for the experiments conducted at two different initial 11-mer concentrations of CG*C-I duplexes. We note that the R_L value is unusually high in the case of the CG*C-I (20 μM); as the length of the ligated molecules increases, the R_L value first diminishes and then increases. Since the overhangs in the shortest sequences (see Materials and methods section) can affect the electrophoretic mobilities, the longer ligation products (> 44 base pairs) are more reflective of the impact of the G* lesion on the electrophoretic characteristics of the lesion-containing duplexes. The R_L values of $\sim 1.3 - 2.1$ for the oligonucleotide adducts are within the range of values

exhibited by oligonucleotides with $(dA)_n \cdot (dT)_n$ tracts.^{29,30} When the bends are in phase with the helical repeat (~ 11 base pairs), the R_L values increase with increasing multimer size.^{30,32,33} The values of R_L for both types of CG*C-I and CG*C-II duplexes also follow this pattern for multimers longer than ~ 44 nucleotides (Figure 4b). The larger values of R_L for CG*C-II multimers indicate that the overall extent of bending of this sequence is greater than in the case of the CG*C-I sequence context. These results clearly indicate that the non-flanking, non-nearest neighbor DNA bases have a significant effect on the overall bending and retardation of electrophoretic mobilities induced by the 10S (+)-*trans-anti*-B[a]P- N^2 -dG lesion.

Circular ligation products—The self-ligation of oligonucleotides can also generate covalently closed, circular products.^{31,34} Covalent circularization requires the correct torsional alignment during the noncovalent hybridization of the “sticky ends”. Covalently closed circular DNA molecules are formed when ligase seals the resulting nicks. The circular products $c - i$ in Figure 4d were identified by two-dimensional native gel electrophoresis methods.³⁵ Typical results are shown for unmodified (Figure 3d) and B[a]PDE-modified oligonucleotide multimer ligation ladders (Figure 4d). The bands closer to the diagonal are due to the linear ligation products and no circular ligation products are evident in the case of the unmodified multimers (Figure 3d) since there are no off-diagonal spots.³⁶ However, such spots are clearly evident in the 2-D gel experiments carried out with CG*C-I multimers (Figure 4d). The nicked and covalently closed circular DNA molecules appear as spots above the diagonal.³⁶ The sizes of the CG*C-I circles in spots c, d, e, f, g, and h were found to consist of 77, 88, 99, 100, and 110 base pairs (the details are provided in Figure S4, Supporting Information). In contrast, in the CG*C-I case described here, there is a notable lack of circular ligation products when the starting CG*C-I 11-mer concentration is 4 μ M (Figure S3, Supporting Information), although such products are readily observed at a 20 μ M starting concentration (Figure 4b and d).

Overall, the formation of circular ligation products suggest that there is considerable torsional flexibility associated with the 10S (+)-*trans-anti*-B[a]P- N^2 -dG adducts in the two sequence contexts. The larger R_L factor in the case of CG*C-II is fully consistent with a greater extent of bending than in the case of the CG*C-I sequence context. This difference is attributed to the presence of the bendable and flexible T-G/C-A dinucleotide steps that are distant from the site of the lesion, but still influence the susceptibility of these modified DNA sequences to recognition by the NER machinery.

MD simulations show unique structural properties of CG*C-II

In order to gain molecular and structural insights into these experimentally observed distant neighbor base sequence context effects, we utilized 10.0 ns MD simulations to analyze the structural and dynamic conformational features of the CG*C-I4 and CG*C-II duplexes.³ An ensemble of 7000 structures collected in the final 7.0 ns of the MD simulation was analyzed to evaluate the structural dynamics of these duplexes. We find that, compared to CG*C-I, the CG*C-II duplex features a more enlarged minor groove (Figure 5), more impaired van der Waals base pair stacking interactions, enlarged and more flexible Roll and diminished Twist (Figure 6), and a greater Slide (relative motion of two stacked base pairs along their long axes)³⁷ (Figure S6, Supporting Information). These differences can account for the dependence of the observed differences in NER efficiencies on base sequence context beyond the bases flanking the 10S (+)-*trans-anti*-B[a]P- N^2 -dG lesion.

Minor groove widths—Minor groove widths are the smallest phosphorus-phosphorus atom separations across the groove. These widths are conventionally measured as the phosphorus atom center-to-center distances minus 5.8 Å which accounts for the van der

Waals radii of the two phosphate groups^{38,39} (Figure 5 and also Movie S1, Supporting Information). The minor grooves near the lesion site are enlarged to a greater extent in the CG*C-II duplex than in the CG*C-I case (Figure 5a and Supporting Information, Figure S5a, and Table S1). The greatest minor groove enlargement induced by the lesion occurs at the lesion site P7-P20 for the CG*C-II duplex, where the duplex is enlarged by ~ 5.1 Å compared to its corresponding unmodified case (Table S1, Supporting Information). By contrast, the greatest minor groove enlargement induced by the lesion for the CG*C-I case is only ~ 3.9 Å at the lesion site P8-P19 (Table S1, Supporting Information). This greater minor groove enlargement in CG*C-II compared to CG*C-I is a manifestation of the greater bend in CG*C-II, observed in our gel electrophoresis studies. The analyses in the next section describe further molecular manifestations of the greater bend in the CG*C-II duplex.

Roll, Twist and Slide—The distant neighbor base sequence effects are also manifested in several helicoidal parameters. Notably, the CG*C-II case features a more enlarged Roll with greater flexibility (Figure 6a), and greater and more dynamically flexible untwisting (Figure 6b) as compared to the CG*C-I case. The Roll is significantly enlarged at the C3-A4 Step 3 (See Figures 5c and 5d for definition of step number) for CG*C-II, with an ensemble average value of $12.7 \pm 6.6^\circ$ (Figure 6a and Table S2, Supporting Information). However, the ensemble average value for the analogous A3-T4 Step 3 in the CG*C-I case is only $0.3 \pm 5.6^\circ$, very similar to the unmodified duplex at this step (Figure 6a and Table S2, Supporting Information). Positive Roll bends the DNA into the major groove side.³⁷ Thus the CG*C-II duplex features a more prominent bend towards the major groove. Furthermore, the broader distribution of populations for this parameter and the greater standard deviation for the Roll at Step 3 in the CG*C-II duplex also suggest that this Roll is dynamically more flexible in CG*C-II than in CG*C-I. It is well-established that duplex bending occurs primarily via Roll⁴⁰ and the helicoidal parameters Twist and Roll are correlated, with greater Roll having lower Twist.^{22,41} The origin of this phenomenon is that enlarged Roll would lead to steric clashes in both grooves in the absence of simultaneous changes in some of the other helicoidal parameters adapted to relieve the steric collisions associated with Roll alone.⁴¹ Prominent among these parameters is Twist, with untwisting providing relief from the steric close contacts associated with increasing Roll.⁴¹ As expected, the CG*C-II duplex features an untwisting at the A4-C5 Step 4, with an ensemble average value of $24.7 \pm 5.3^\circ$, i.e., $\sim 4.1^\circ$ lower than that of CG*C-I duplex at the analogous T4-C5 Step 4 (Figure 6b and Table S2, Supporting Information). Furthermore, as seen from the greater standard deviation from the mean value, and the broader distribution of populations (Figures 5a and 5b), both Roll and untwisting are more flexible in CG*C-II than in CG*C-I. The Slide parameter in CG*C-II at the C5-G6* Step 5 is also greater, with an ensemble average value of 0.4 ± 0.5 Å, while the values for CG*C-I (and the unmodified duplex) at this step is only ~ 0.1 Å (Table S2, Supporting Information). We note that the difference in Slide at Step 5 results from the sequence difference at Step 4 in CG*C-I and CG*C-II. In summary, CG*C-II features an enlarged and more flexible Roll, greater and dynamically more flexible untwisting and larger Slide, as compared to the CG*C-I sequence. These altered features all occur on the 5'-side of the lesion modification site, because the B[a]P rings are oriented in this direction (Figure 1). We also observe some modest differences in minor groove widths and Roll on the 3'-side of G* (Tables S1 and S2, Supporting Information). We attribute these to intrinsic sequence-dependent structural differences on that side, combined with the impact of the lesion propagated from the 5'-side of G*.

The correlated enlarged, flexible Roll and flexible untwisting^{22,40,41} are molecular manifestations of correlated flexible bending and torsional flexibility. These are expected to lead to more efficient cyclization for the CG*C-II sequence. Experimentally, efficient cyclization was observed at the lower CG*C-II 11-mer starting concentration of 4 μ M in the self-ligation experiments of Tsao *et al*¹⁸ (Figure 4b), while in the case of the CG*C-I

ligation experiments, no circularization was evident at similar starting concentrations (Supporting Information), although circular ligation products were abundant at the 20 μM starting concentration (Figure 4b). The absence of cyclization at 4 μM suggests a lower torsional flexibility for CG*C-I (Figure 4b) than in the case of CG*C-II. This is consistent with the more dynamically flexible untwisting (Figure 6b) in the MD simulations in CG*C-II. Notably, the observed enlarged Roll and untwisting as well as the enlarged Slide are known to correlate with minor groove enlargement;⁴² together, these are reflections of the greater bend in CG*C-II duplexes.

Stacking energies—The G* lesion induces a greater local destabilization in CG*C-II because of lower van der Waals base stacking interactions than in CG*C-I (Table S3, Supporting Information). For the total van der Waals interaction energy summed over the central 9-mer base pairs, the same 10S (+)-*trans-anti*-B[a]P-*N*²-dG adduct destabilizes the CG*C-II duplex by $\sim 2.1 \text{ kcal}\cdot\text{mol}^{-1}$, while it induces almost no change in the van der Waals interaction energies in the CG*C-I case (Table S3, Supporting Information). Since stacking interactions are important in defining the Twist between adjacent base pairs,⁴³ the untwisting of the lesion-containing duplex in the CG*C-II case is consistent with the diminished local stacking energy.

B[a]P accommodation—The B[a]P moiety is better accommodated in the CG*C-II duplex; this is manifested by the total interaction energy between the B[a]P ring system and the DNA, which is $1.6 \text{ kcal}\cdot\text{mol}^{-1}$ more favorable in CG*C-II than CG*C-I (Table S4, Supporting Information). The better interactions between the B[a]P moiety and the DNA reduces the exposure of the hydrophobic surface area of the B[a]P to solvent in the CG*C-II case. This is shown by a lower solvent accessible surface area (SASA) of the B[a]P aromatic rings in CG*C-II than in CG*C-I (Figure 7a): the population distributions for the SASA show that 36 percent of the population has a SASA greater than $130 \text{ \AA}^2$ for CG*C-II, whereas 60 percent of the population has a SASA greater than this value for CG*C-I. The computed lower solvent exposure of the aromatic B[a]P residue in CG*C-II than in CG*C-I duplexes is experimentally supported by the fluorescence quenching experiments shown in Figure 7b. The relative solvent accessibility of the polycyclic aromatic pyrenyl residues can be measured by a standard fluorescence quenching method:¹⁵ the lower the solvent accessibility, the lower the ability of water-soluble fluorescence quenchers to access and quench the fluorescence of the B[a]P residue. The lower fluorescence quenching efficiency with the 10S (+)-*trans-anti*-B[a]P-*N*²-dG adduct in CG*C-II by the nucleoside dU (Figure 7b) is completely consistent with the computed SASA value.

Discussion

Intrinsic structural properties of the CG*C-I and CG*C-II DNA sequences determine the impact of the lesion

We consider the origin of the differences in the structural and dynamic properties of the CG*C-I and CG*C-II duplexes that can account for the greater susceptibility of the latter to mammalian NER. Specifically, we focus on the effects of the C3:G20 base pair in the CG*C-II duplex relative to the impact of the A3:T20 base pair at the analogous site of CG*C-I. It is well-known that in unmodified DNA, guanine amino groups in the minor groove of B-DNA cause an enlargement of this groove.^{41,44} This phenomenon is evident in our MD simulations for the unmodified DNA. However, the question arises: why is the minor groove more enlarged in the CG*C-II than in the CG*C-I duplex, although they are similar in the unmodified duplexes? The unique C3:G20 base pair in the CG*C-II sequence, combined with the unique C3-A4 step, play critical roles: the opening of the groove by the G20 guanine amino group acts as a wedge that takes advantage of the intrinsic higher

flexibility of the C3-A4 step.^{22,45–48} In the CG*C-II case, the C3-A4 pyrimidine-purine step is about fourfold as flexible as the analogous A3-T4 purine-pyrimidine step in the CG*C-I case.⁴⁷ This added flexibility, facilitated by the steric crowding in the minor groove caused by the amino group of G20, allows the 5'-directed B[a]P rings to assume a position with more favorable interactions with the DNA, with a more widened groove and increased Roll and untwisting. Consequently, the steric effect of the G20 guanine amino group in the minor groove is enhanced by the bulky lesion in the same groove and propagates to the lesion site; this causes the local duplex to unwind, as manifested by a flexible enlarged Roll and flexible untwisting. These disturbances are facilitated by the intrinsic thermodynamic stability and flexibility of the adjacent dinucleotide steps. In summary, the presence of a guanine amino group at C3:G20 only in the CG*C-II case, facilitated by the high flexibility and low thermodynamic stability of the C3-A4 dinucleotide step, produces the minor groove enlargement, and the flexible and higher Roll, flexible untwisting, and impaired van der Waals stacking interactions in the CG*C-II case. These disturbances, in turn, allow the B[a]P moiety to more optimally position itself in the minor groove with better interaction energies than in the CG*C-I case. Thus, the B[a]P rings are better accommodated at the price of causing a greater disturbance of the DNA duplex in the CG*C-II case.

Differences in base sequence contexts cause different structural distortions, and produce different NER dual incision efficiencies

The ~ 1.6-fold greater dual incision efficiency of the CG*C-II duplex relative to the CG*C-I duplex is consistent with the experimentally observed greater extent of bending and the enhanced flexibility of the CG*C-II duplex inferred from the MD simulation studies. The different base sequence contexts beyond the bases flanking G* and the high deformability and flexibility of the adjacent C-A dinucleotide base step in the CG*C-II duplex, produce these differences in the helical properties of the two duplexes. Steric hindrance effects associated with the bulky guanine amino groups that locally widen the minor groove, play a key role in establishing these differences.

In previous studies of flanking base sequence effects we found that base pair disruption in the CG*G sequence was a better NER recognition signal than a highly flexible bend observed in the sequence isomer GG*C that is characterized by intact hydrogen bonding.^{4,5} Flexibility is nevertheless among the contributing NER recognition factor,³ since the efficiencies of dual incision are higher in the case of the thermodynamically less stable⁶ and highly flexible TG*T-II²⁶ duplex than in the more rigid¹⁸ CG*C-II sequences in eukaryotic (Figure 2b) and in prokaryotic⁶ NER experiments in vitro. Furthermore, we now also show that flexibility enhances NER dual incisions even when the flexible base pairs are distant from the minor groove-positioned lesion G*. These observations are consistent with the concept that the extent of *local* thermodynamic destabilization^{49–51} governs the degree of recognition of duplex sequences with (+)-*trans-anti*-B[a]P-*N*²-dG lesions by XPC/HR23B. Regardless of the nature of the specific distortions, the degree of *local* destabilization induced may provide a unifying framework for understanding the differences in repairability of different lesions^{1,2} and in different sequence contexts.^{3–5}

Materials and methods

DNA substrates

The 11-mer CG*C-I and CG*C-II sequences, each containing the 10S (+)-*trans-anti*-B[a]P-*N*²-dG lesions at designated positions (Figure 1c) were prepared by routine direct synthesis methods involving the reactions of (±)-*anti*-B[a]PDE with the different oligonucleotides; the reaction products were purified and characterized as described in detail elsewhere.^{52,53} The purified oligonucleotides containing the single 10S (+)-*trans-anti*-B[a]P-*N*²-dG adducts

were ^{32}P -5'-end-labeled and incorporated into 135-mer oligonucleotides by standard ligation methods as described earlier.²⁵ These internally labeled sequences were then annealed with their complementary strands by heating the solutions to $\sim 90^\circ\text{C}$ and allowing the mixtures to cool slowly to 4°C , thus forming the 135-mer duplexes for NER assays.

Human NER assays in cell-free extracts

The cell-free extracts were prepared from standard HeLa cells (American Type Culture Collection GM00637, Coriell Institute for Medical research, Camden, N.J.). The cell extracts were prepared by standard protocols by cell lysis and ammonium sulfate precipitation^{25,54} as described in more detail in our earlier publication⁵ and in Supporting Information. After incubation of the modified or unmodified control oligonucleotide 135-mer duplexes for different time intervals, the oligonucleotide excision products and intact DNA were desalted by precipitation with an aqueous 80% methanol solution, and subjected to denaturing 12% acrylamide gel electrophoresis. The gels were analyzed by autoradiography methods using a Storm 840 phosphorimager.

Ligation and gel electrophoresis studies

The modified 11-mer CG**C*-I sequences were hybridized with the unmodified complementary 11-mer d(GGGTAGCGATA) strand in order to provide a one-base overhang for the self-ligation of the duplexes:

Forward reaction 5'-end labeling of both unmodified and modified oligonucleotides with $[\gamma\text{-}^{32}\text{P}]\text{ATP}$ (New England Nuclear) and cold ATP (Pharmacia) were employed as described previously;⁵⁵ the complementary oligonucleotides were phosphorylated at the 5'-ends with ATP. The $[\gamma\text{-}^{32}\text{P}]$ labeled oligonucleotides were purified by denaturing gel electrophoresis in 20% polyacrylamide gels. All samples were desalted using chromatography cartridge (Biorad) before further use. A two-fold excess of complementary strands over the radioactively labeled B[a]P-modified strands was used in the ligation reaction in order to maximize the fraction of DNA molecules in the duplex form and to enhance the yield of ligation products. The ligation reaction was conducted at two different DNA duplex concentrations of $4\ \mu\text{M}$ or $20\ \mu\text{M}$, with 1 unit of T4 ligase (GIBCO BRL) added for every 5 pmol or 25 pmol of double-stranded DNA in 1X ligation buffer (GIBCO BRL). The ligated DNA multimer mixtures were subjected to one-dimensional (1-D) and two-dimensional (2-D) non-denaturing (29:1 acrylamide to bisacrylamide) gel electrophoresis at 4°C . In the 2-D gel electrophoresis experiments, the polyacrylamide concentration was 6–8% for the first, and 8 to 12% for the second dimension; in the latter case, chloroquine ($1\ \mu\text{g}/\text{ml}$) was also present in the running buffer.³⁵

MD simulations

MD simulation studies of 10.0 ns with explicit water and counterions were carried out in the CG**C*-I4 and CG**C*-II3 sequences employing the AMBER 8 simulation package.⁵⁶ The PTRAJ module of AMBER 8⁵⁶ and the CARNAL module of the AMBER 7 package⁵⁷ were employed for structural analyses. Frames were selected at 1 ps intervals from the last 7.0 ns of simulations. DNA duplex groove dimensions were analyzed using MD Toolchest^{58,59} and $5.8\ \text{\AA}$ was subtracted from the pairwise phosphorous-phosphorous distances to account for the van der Waals radius of the P atoms.^{38,60} Stacking interactions were estimated by computing the van der Waals interaction energies between all adjacent base pairs, utilizing the ANAL module of the AMBER 8 package.⁵⁶ In these analyses, one

base pair at each end was excluded due to possible end effects. Total interaction energies between the B[a]P moiety (all atoms up to and including C10) and the DNA were obtained via the ANAL module and were composed of bond, angle, dihedral, van der Waals and electrostatic terms. Visualization and model building were performed in the INSIGHT II 2005 program (Accelrys Software, Inc.). PyMOL (Delano Scientific, LLC.)⁶¹ was employed to make molecular images and movies. Computations were carried out on our own cluster of Silicon Graphic Origin and Altix high-performance computers.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The experimental portion of this work was supported by NIH Grant CA-099194 (N.E.G.), and the computational aspects were supported by Grant CA-28038 (S.B.). Partial support for computational infrastructure and systems management was also provided by CA75449 (S.B.). Support for this work to D.J.P. was provided by CA-046533. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Cancer Institute or the National Institutes of Health.

Abbreviations

B[a]P	benzo[a]pyrene
B[a]PDE	benzo[a]pyrene diol epoxide
(+)-anti-B[a]PDE	(+)-(7R,8S,9S,10R)-7,8-dihydroxy-9,10-epoxy-7,8,9,10-tetrahydrobenzo[a]pyrene
NER	nucleotide excision repair
MD	molecular dynamics
SASA	solvent accessible surface area

References

1. Gillet LC, Schärer OD. Molecular mechanisms of mammalian global genome nucleotide excision repair. *Chem Rev.* 2006; 106:253–76. [PubMed: 16464005]
2. Wood RD. DNA damage recognition during nucleotide excision repair in mammalian cells. *Biochimie.* 1999; 81:39–44. [PubMed: 10214908]
3. Cai Y, Patel DJ, Geacintov NE, Broyde S. Dynamics of a benzo[a]pyrene-derived guanine DNA lesion in TGT and CGC sequence contexts: enhanced mobility in TGT explains conformational heterogeneity, flexible bending, and greater susceptibility to nucleotide excision repair. *J Mol Biol.* 2007; 374:292–305. [PubMed: 17942115]
4. Cai Y, Patel DJ, Geacintov NE, Broyde S. Differential nucleotide excision repair susceptibility of bulky DNA adducts in different sequence contexts: hierarchies of recognition signals. *J Mol Biol.* 2009; 385:30–44. [PubMed: 18948114]
5. Kropachev K, Kolbanovskii M, Cai Y, Rodriguez F, Kolbanovskii A, Liu Y, Zhang L, Amin S, Patel D, Broyde S, Geacintov NE. The sequence dependence of human nucleotide excision repair efficiencies of benzo[a]pyrene-derived DNA lesions: insights into the structural factors that favor dual incisions. *J Mol Biol.* 2009; 386:1193–203. [PubMed: 19162041]
6. Ruan Q, Liu T, Kolbanovskiy A, Liu Y, Ren J, Skovvaga M, Zou Y, Lader J, Malkani B, Amin S, Van Houten B, Geacintov NE. Sequence context- and temperature-dependent nucleotide excision repair of a benzo[a]pyrene diol epoxide-guanine DNA adduct catalyzed by thermophilic UvrABC proteins. *Biochemistry.* 2007; 46:7006–15. [PubMed: 17506530]

7. Luch A. Nature and nurture – Lessons from chemical carcinogenesis. *Nat Rev Cancer*. 2005; 5:113–25. [PubMed: 15660110]
8. Phillips DH. Fifty years of benzo(a)pyrene. *Nature*. 1983; 303:468–72. [PubMed: 6304528]
9. Conney AH. Induction of microsomal enzymes by foreign chemicals and carcinogenesis by polycyclic aromatic hydrocarbons: G. H. A. Clowes Memorial Lecture. *Cancer Res*. 1982; 42:4875–917. [PubMed: 6814745]
10. Cosman M, de los Santos C, Fiala R, Hingerty BE, Singh SB, Ibanez V, Margulis LA, Live D, Geacintov NE, Broyde S, Patel DJ. Solution conformation of the major adduct between the carcinogen (+)-*anti*-benzo[a]pyrene diol epoxide and DNA. *Proc Natl Acad Sci U S A*. 1992; 89:1914–8. [PubMed: 1311854]
11. Rodriguez FA, Cai Y, Lin C, Tang Y, Kolbanovskiy A, Amin S, Patel DJ, Broyde S, Geacintov NE. Exocyclic amino groups of flanking guanines govern sequence-dependent adduct conformations and local structural distortions for minor groove-aligned benzo[a]pyrenyl-guanine lesions in a GG mutation hotspot context. *Nucleic Acids Res*. 2007; 35:1555–68. [PubMed: 17287290]
12. Yang W. Poor base stacking at DNA lesions may initiate recognition by many repair proteins. *DNA Repair (Amst)*. 2006; 5:654–66. [PubMed: 16574501]
13. Yang W. Structure and mechanism for DNA lesion recognition. *Cell Res*. 2008; 18:184–197. [PubMed: 18157156]
14. Isaacs RJ, Spielmann HP. A model for initial DNA lesion recognition by NER and MMR based on local conformational flexibility. *DNA Repair (Amst)*. 2004; 3:455–64. [PubMed: 15162792]
15. Huang W, Amin S, Geacintov NE. Fluorescence characteristics of site-specific and stereochemically distinct benzo[a]pyrene diol epoxide-DNA adducts as probes of adduct conformation. *Chem Res Toxicol*. 2002; 15:118–26. [PubMed: 11849037]
16. Suh M, Ariese F, Small GJ, Jankowiak R, Liu TM, Geacintov NE. Conformational studies of the (+)-*trans*, (–)-*trans*, (+)-*cis*, and (–)-*cis* adducts of anti-benzo[a]pyrene diolepoxide to *N*²-dG in duplex oligonucleotides using polyacrylamide gel electrophoresis and low-temperature fluorescence spectroscopy. *Biophys Chem*. 1995; 56:281–96. [PubMed: 7578905]
17. Tang, Y. Ph.D thesis. New York University, Chemistry Department; 2008. Optical spectroscopic and NMR studies of covalent polycyclic aromatic hydrocarbons–DNA adducts: Influence of base sequence context and carcinogen topology.
18. Tsao H, Mao B, Zhuang P, Xu R, Amin S, Geacintov NE. Sequence dependence and characteristics of bends induced by site-specific polynuclear aromatic carcinogen-deoxyguanosine lesions in oligonucleotides. *Biochemistry*. 1998; 37:4993–5000. [PubMed: 9538018]
19. Xu R, Mao B, Xu J, Li B, Birke S, Swenberg CE, Geacintov NE. Stereochemistry-dependent bending in oligonucleotide duplexes induced by site-specific covalent benzo[a]pyrene diol epoxide-guanine lesions. *Nucleic Acids Res*. 1995; 23:2314–9. [PubMed: 7610061]
20. Lyubchenko YL, Shlyakhtenko LS, Appella E, Harrington RE. CA runs increase DNA flexibility in the complex of lambda Cro protein with the OR3 site. *Biochemistry*. 1993; 32:4121–7. [PubMed: 8471619]
21. Nagaich AK, Bhattacharyya D, Brahmachari SK, Bansal M. CA/TG sequence at the 5' end of oligo(A)-tracts strongly modulates DNA curvature. *J Biol Chem*. 1994; 269:7824–33. [PubMed: 8126009]
22. Olson WK, Gorin AA, Lu XJ, Hock LM, Zhurkin VB. DNA sequence-dependent deformability deduced from protein-DNA crystal complexes. *Proc Natl Acad Sci U S A*. 1998; 95:11163–8. [PubMed: 9736707]
23. Li C, Wang LE, Wei Q. DNA repair phenotype and cancer susceptibility--a mini review. *Int J Cancer*. 2009; 124:999–1007. [PubMed: 19065660]
24. Cosman M, de los Santos C, Fiala R, Hingerty BE, Ibanez V, Luna E, Harvey R, Geacintov NE, Broyde S, Patel DJ. Solution conformation of the (+)-*cis-anti*-[BP]dG adduct in a DNA duplex: intercalation of the covalently attached benzo[a]pyrenyl ring into the helix and displacement of the modified deoxyguanosine. *Biochemistry*. 1993; 32:4145–55. [PubMed: 8476845]

25. Hess MT, Gunz D, Luneva N, Geacintov NE, Naegeli H. Base pair conformation-dependent excision of benzo[a]pyrene diol epoxide-guanine adducts by human nucleotide excision repair enzymes. *Mol Cell Biol*. 1997; 17:7069–76. [PubMed: 9372938]
26. Xu R, Mao B, Amin S, Geacintov NE. Bending and circularization of site-specific and stereoisomeric carcinogen-DNA adducts. *Biochemistry*. 1998; 37:769–78. [PubMed: 9425101]
27. Lerman LS, Frisch HL. Why does the electrophoretic mobility of DNA in gels vary with the length of the molecule? *Biopolymers*. 1982; 21:995–7. [PubMed: 7082773]
28. Lumpkin OJ. Mobility of DNA in gel electrophoresis. *Biopolymers*. 1982; 21:2315–6. [PubMed: 7171739]
29. Koo HS, Crothers DM. Calibration of DNA curvature and a unified description of sequence-directed bending. *Proc Natl Acad Sci U S A*. 1988; 85:1763–7. [PubMed: 3162306]
30. Hagerman PJ. Sequence dependence of the curvature of DNA: a test of the phasing hypothesis. *Biochemistry*. 1985; 24:7033–7. [PubMed: 4084556]
31. Livshits MA, Lyubchenko YL. Kinetics of oligonucleotide cyclization and ligation. *Molecular Biology (Moscow) English Transl*. 1994; 28:687–690.
32. Koo HS, Wu HM, Crothers DM. DNA bending at adenine thymine tracts. *Nature*. 1986; 320:501–6. [PubMed: 3960133]
33. Wu HM, Crothers DM. The locus of sequence-directed and protein-induced DNA bending. *Nature*. 1984; 308:509–13. [PubMed: 6323997]
34. Shore D, Baldwin RL. Energetics of DNA twisting. I Relation between twist and cyclization probability. *J Mol Biol*. 1983; 170:957–81. [PubMed: 6315955]
35. Ulanovsky L, Bodner M, Trifonov EN, Choder M. Curved DNA: design, synthesis, and circularization. *Proc Natl Acad Sci U S A*. 1986; 83:862–6. [PubMed: 3456570]
36. Zahn K, Blattner FR. Direct evidence for DNA bending at the lambda replication origin. *Science*. 1987; 236:416–22. [PubMed: 2951850]
37. Dickerson RE. Definitions and nomenclature of nucleic acid structure components. *Nucleic Acids Res*. 1989; 17:1797–803. [PubMed: 2928107]
38. Fratini AV, Kopka ML, Drew HR, Dickerson RE. Reversible bending and helix geometry in a B-DNA dodecamer: CGCGAATTBrCGCG. *J Biol Chem*. 1982; 257:14686–707. [PubMed: 7174662]
39. Saenger, W. *Principles of Nucleic Acid Structure*. Springer-Verlag; New York: 1983. Polymorphism of DNA; p. 226
40. Dickerson RE. DNA bending: the prevalence of kinkiness and the virtues of normality. *Nucleic Acids Res*. 1998; 26:1906–26. [PubMed: 9518483]
41. Gorin AA, Zhurkin VB, Olson WK. B-DNA twisting correlates with base-pair morphology. *J Mol Biol*. 1995; 247:34–48. [PubMed: 7897660]
42. Johnson, RC.; Stella, S.; Heiss, JK. Bending and Compaction of DNA by Proteins. In: Rice, PAaCCC., editor. *Protein-Nucleic Acid Interactions*. The Royal Society of Chemistry; Cambridge: 2008. p. 397
43. Cooper VR, Thonhauser T, Puzder A, Schroder E, Lundqvist BI, Langreth DC. Stacking interactions and the twist of DNA. *J Am Chem Soc*. 2008; 130:1304–8. [PubMed: 18163624]
44. Calladine CR. Mechanics of sequence-dependent stacking of bases in B-DNA. *J Mol Biol*. 1982; 161:343–52. [PubMed: 7154084]
45. el Hassan MA, Calladine CR. Two distinct modes of protein-induced bending in DNA. *J Mol Biol*. 1998; 282:331–43. [PubMed: 9735291]
46. el Hassan MA, Calladine GR. Conformational characteristics of DNA: empirical classifications and a hypothesis for the conformational behaviour of dinucleotide steps. *Phil Trans ser A*. 1997:43–100.
47. Okonogi TM, Alley SC, Reese AW, Hopkins PB, Robinson BH. Sequence-dependent dynamics of duplex DNA: the applicability of a dinucleotide model. *Biophys J*. 2002; 83:3446–59. [PubMed: 12496111]
48. Packer MJ, Dauncey MP, Hunter CA. Sequence-dependent DNA structure: dinucleotide conformational maps. *J Mol Biol*. 2000; 295:71–83. [PubMed: 10623509]

49. Geacintov NE, Broyde S, Buterin T, Naegeli H, Wu M, Yan S, Patel DJ. Thermodynamic and structural factors in the removal of bulky DNA adducts by the nucleotide excision repair machinery. *Biopolymers*. 2002; 65:202–10. [PubMed: 12228925]
50. Gunz D, Hess MT, Naegeli H. Recognition of DNA adducts by human nucleotide excision repair. Evidence for a thermodynamic probing mechanism. *J Biol Chem*. 1996; 271:25089–98. [PubMed: 8810263]
51. Min JH, Pavletich NP. Recognition of DNA damage by the Rad4 nucleotide excision repair protein. *Nature*. 2007; 449:570–5. [PubMed: 17882165]
52. Mao B, Xu J, Li B, Margulis LA, Smirnov S, Ya NQ, Courtney SH, Geacintov NE. Synthesis and characterization of covalent adducts derived from the binding of benzo[*a*]pyrene diol epoxide to a -GGG- sequence in a deoxyoligonucleotide. *Carcinogenesis*. 1995; 16:357–65. [PubMed: 7859369]
53. Pirogov N, Shafirovich V, Kolbanovskiy A, Solntsev K, Courtney SA, Amin S, Geacintov NE. Role of hydrophobic effects in the reaction of a polynuclear aromatic diol epoxide with oligodeoxynucleotides in aqueous solutions. *Chem Res Toxicol*. 1998; 11:381–8. [PubMed: 9548810]
54. Shivji MK, Moggs JG, Kuraoka I, Wood RD. Dual-incision assays for nucleotide excision repair using DNA with a lesion at a specific site. *Methods Mol Biol*. 1999; 113:373–92. [PubMed: 10443435]
55. Xu R, Birke S, Carberry SE, Geacintov NE, Swenberg CE, Harvey RG. Differences in unwinding of supercoiled DNA induced by the two enantiomers of anti-benzo[*a*]pyrene diol epoxide. *Nucleic Acids Res*. 1992; 20:6167–76. [PubMed: 1475180]
56. Case, DA.; Darden, TA.; Cheatham, TE., III; Simmerling, CL.; Wang, J.; Duke, RE.; Luo, R.; Merz, KM.; Wang, B.; Pearlman, DA.; Crowley, M.; Brozell, S.; Tsui, V.; Gohlke, H.; Mongan, J.; Hornak, V.; Cui, G.; Beroza, P.; Schafmeister, C.; Kollman, PA. AMBER 8. University of California; San Francisco, CA: 2004.
57. Case, DA.; Pearlman, DA.; Caldwell, JW.; Cheatham, TE., III; Wang, J.; Ross, WS.; Simmerling, CL.; Darden, TA.; Merz, KM.; Stanton, RV.; Cheng, AL.; Vincent, JJ.; Crowley, M.; Tsui, V.; Gohlke, H.; Radmer, RJ.; Duan, Y.; Pitera, J.; Massova, I.; Seibel, GL.; Singh, UC.; Weiner, PK.; Kollman, PA. AMBER 7. University of California; San Francisco, CA: 2002.
58. Ravishanker, G.; Beveridge, DL. MD Toolchest 2.0. Wesleyan University; Middletown, CT 06459: 1993.
59. Ravishanker G, Swaminathan S, Beveridge DL, Lavery R, Sklenar H. Conformational and helicoidal analysis of 30 ps of molecular dynamics on the d(CGCGAATTCGCG) double helix: “curves”, dials and windows. *J Biomol Struct Dyn*. 1989; 6:669–99. [PubMed: 2619934]
60. Saenger, W. Principles of Nucleic Acid Structure. Springer-Verlag New York, Inc; New York, NY: 1984. p. 226
61. DeLano, WL. The PyMOL Molecular Graphics System. DeLano Scientific; Palo Alto, CA, USA: 2002.
62. Simmerling, C.; Elber, R.; Zhang, J. Modelling of biomolecular structure and mechanisms. Netherlands: Kluwer; 1995. p. 241-265.
63. Bloomfield, VA.; Crothers, DM.; Tinoco, JI. Nucleic Acids: Structures, Properties and Functions. University science books; Sausalito, C.A: 2000. Nucleic acid structures from diffraction methods; p. 90-96.

Figure 1.

(a) Chemical structure of the 10S (+)-trans-anti-B[a]P-N²-dG adduct. (b) Representative minor groove conformation of the 10S (+)-trans-anti-B[a]P-N²-dG adduct. The structure shown is the best representative structure⁶² from the 10.0 ns MD simulation⁴ for the CG*C-I sequence context. Only the central 5-mer is shown. (c) Cartoon representation of the lesion-containing duplexes, in which G6* designates the lesion-modified base. Note that CG*C-I and CG*C-II differ beginning with the next nearest neighbor to G6* and beyond. The locations of the key guanine amino group "1" and highly deformable C-A step "2" only in the CG*C-II case are marked in red balloons.

Figure 2.

(a) Autoradiograph of a denaturing gel showing the results of a typical dual incision experiment as a function of incubation time in extracts from human HeLa cells of the 10S (+)-*trans-anti*-B[a]P-*N*²-dG adduct (G*) in 135-mer duplexes; the G* is at position 67 counted from the 5'-end in the CG*C-I and CG*C-II sequence contexts. The exact sequence of the 135-mer is shown in the Supporting Information section of Kropachev et al.⁵ The incubation times (min) are designated for each lane, while lane M shows the mobilities of a single-stranded 135-mer (top) and oligonucleotide markers 32, 30, 28, 26, 24, etc. nucleotides in length used as size standards (lower half). (b) Time course of dual incision product formation elicited by the 10S (+)-*trans-anti*-B[a]P-*N*²-dG adduct in the CG*C-I, CG*C-II, and TG*T-II sequence contexts, and the 10R (+)-*cis-anti*-B[a]P-*N*²-dG adduct standard in the CG*C-I sequence context ((+)-*cis*-CG*C-I). The experimental points were obtained by summing the intensities of the pixels in the region of the 24–30 oligonucleotide fragments and dividing by the sum of intensities of all the pixels in the same lane, thus adjusting for any variations in DNA loading. The points shown are averages and standard deviations obtained in seven independent gel experiments with four different cell extracts. The experiments with the (+)-*cis*-CG*C-I duplex at 40 min were arbitrarily assigned a value of 100 to take into account variations in NER activities of different cell extracts (see the text); the 10 and 20 min time points shown are averages of four experiments obtained with four different cell extracts. A typical gel showing the kinetics of dual incision product formation derived from the (+)-*cis*-CG*C-I is shown in Figure S1, while a gel depicting a comparison of the CG*C-II and TG*T-II sequences is shown in Figure S2. The straight lines are the least square fits to the data points. (c) Comparisons of initial rates of dual incision efficiencies of the CG*C-I, CG*C-II, TG*T-II and the (+)-*cis*-CG*C-I adduct standard based on data points shown in panel (b).

Figure 3.

(a) Autoradiograph of 8% polyacrylamide gels (29:1 acrylamide: bis-acrylamide) showing the relative mobilities of ligation products of CG*C-I 11-mers containing single site-specific G* lesions (Lane 1), and of the CGC-I unmodified control (Lane 2). The starting concentrations in the ligation reactions of undecamer duplexes was 20 μ M. (b) Ligation ladder of CG*C-II 11-mers (3 μ M) containing single site-specific G* lesions (Lane 1), and of the CGC-II unmodified control (Lane 2). Adapted with permission from Tsao et al., *Biochemistry*. 1998, 37:4993–5000, Figure 2, Copyright © 1998, American Chemical Society.

Figure 4.

(a) Densitometric plot of the **CG*C-I** ligation ladder (Lane 1 in Figure 3(a)). The numbers of nucleotides in each multimer are indicated, while the letters c, d, e, f, g, h and i denote circular ligation products (see the text). (b) Plots of the R_L -values of the ligation products as a function of sequence length, expressed in terms of the number of base pairs per DNA molecule. The data for the **CG*C-I** sequence is from panel (a), Figure 3, while the data for **CG*C-II** is from panel (b), Figure 3, Tsao *et al.*¹⁸ We note that there was a notable lack of circular ligation products for the **CG*C-I** 11-mer when the starting concentration was 4 μM (See text and Figure S3, Supporting Information) and for the **CG*C-II** 11-mer when the starting concentration was 3 μM (See also Figure 3b). (c) and (d), two-dimensional gel electrophoresis of ligation products. (c), ligation products derived from the unmodified CGC-I undecamer. First dimension (horizontal): 6% polyacrylamide gel; second dimension: 7% polyacrylamide gel containing 50 $\mu\text{g/ml}$ chloroquine; no cyclic products are observed. (d), ligation products (from the gel shown in panel (a), Figure 3), obtained from **CG*C-I** undecamers. First dimension (horizontal): 6% polyacrylamide gel; second dimension: 8% polyacrylamide gel containing 50 $\mu\text{g/ml}$ chloroquine. Spots labeled “a” through “i” are cyclic ligation products (see text and Figure S4, Supporting Information).

Figure 5.

(a) Population distributions for minor groove widths. The red arrows indicate where the greatest differences between CG*C-I and CG*C-II occur. Comparisons with the unmodified sequences are given in Figure S5 (a), Supporting Information. Means and standard deviations are given in Table S1, Supporting Information. (b) Minor groove widths are the smallest phosphorus-phosphorus separations across the groove, less 5.8 Å which accounts for the van der Waals radii of the two phosphate groups³⁸ (See also Movie S1, Supporting Information). (c) Schematic illustration of minor groove widths for CG*C-I and CG*C-II, respectively.

Figure 6.

Population distributions for (a) Roll and (b) Twist. The red arrows indicate where the greatest differences between CG* C-I and CG* C-II occur. The definitions of the helicoidal parameters Roll and Twist are also shown in the cartoons; these cartoons are adapted with permission from Lu et al, *Nucleic Acids Res.* **31** (17): 5108–5121, Figure 1, Copyright © 2003, Oxford University Press. Roll is the angle between two successive base pair planes, while Twist is the angle by which one base pair has to rotate around the helical axis to match with the next base pair. Average values for Twist and Roll for B-DNA are 36 ° and 0 °, respectively, but these quantities vary as a function of base sequence.⁶³ Comparisons with the unmodified sequences are given in Figure S5 (b and c), Supporting Information. Means and standard deviations are given in Table S2, Supporting Information. (c) and (d), the numbering schemes for base steps for CG* C-I and CG* C-II.

Figure 7.

(a) Population distributions of the solvent accessible surface areas for the B[a]P aromatic rings in CG*C-I and CG*C-II. Ensemble average values and standard deviations (in parentheses) are given. The vertical magenta line shows that 36 percent of the population has a SASA greater than 130 Å^2 for CG*C-II, whereas 60 percent of the population has a SASA greater than this value for CG*C-I. (b) Solvent accessibility of B[a]P aromatic residues measured by fluorescence quenching methods as described by Huang et al.¹⁵ Typical Stern-Volmer fluorescence quenching curves for the pyrenyl-like residues of the 10S (+)-*trans-anti*-B[a]P-*N*²-dG adduct in CG*C-I and CG*C-II duplexes. Plots of the quenching factors $F_0/F(dU)$ as a function of quencher concentration ($dU = 2\text{-deoxyuridine}$), where F_0 is the fluorescence intensity in the absence of the quencher, and $F(dU)$ is the fluorescence intensity determined in the presence of dU . The aromatic B[a]P residue is more accessible to the quencher in CG*C-I than CG*C-II. Excitation wavelength: 325 nm. Fluorescence viewing wavelength: 360–400nm.