Solution Structures of Aminofluorene [AF]-Stacked Conformers of the *syn* [AF]−C^8-dG Adduct Positioned Opposite dC or dA at a Template-Primer Junction†

Zhengtian Gu,‡ Andrey Gorin,‡ Brian E. Hingerty,§ Suse Broyde,|| and Dinshaw J. Patel*,†

Cellular Biochemistry and Biophysics Program, Memorial Sloan-Kettering Cancer Center, New York, New York 10021, Life Sciences Division, Oak Ridge National Laboratory, Oak Ridge, Tennessee 37831, and Biology Department, New York University, New York, New York 10003

Received June 2, 1999

ABSTRACT: A solution structural study has been undertaken on the aminofluorene−C^8-dG ([AF]dG) adduct located at a single-strand−double-strand d(A1-A2-C3-[AF]G4-C5-T6-A7-C8-C9-A10-T11-C12-C13)·d(G14-G15-A16-T17-G18-G19-T20-A21-G22-N23) 13/10-mer junction (N = C or A) using proton−proton distance restraints derived from NMR data in combination with intensity-based relaxation matrix refinement computations. This single-strand−double-strand junction models one arm of a replication fork composed of a 13-mer template strand which contains the [AF]dG modification site and a 10-mer primer strand which has been elongated up to the modified guanine with either its complementary dC partner or a dA mismatch. The solution structures establish that the duplex segment retains a minimally perturbed B-DNA conformation with Watson−Crick hydrogen-bonding retained up to the dC5·dG22 base pair. The guanine ring of the [AF]dG4 adduct adopts a *syn* glycosidic torsion angle and is displaced into the major groove when positioned opposite dC or dA residues. This base displacement of the modified guanine is accompanied by stacking of one face of the aminofluorene ring of [AF]dG4 with the dC5·dG22 base pair, while the other face of the aminofluorene ring is stacked with the purine ring of the nonadjacent dA2 residue. By contrast, the dC and dA residues opposite the junctional [AF]dG4 adduct site adopt distinctly different alignments. The dC23 residue positioned opposite the adduct site is looped out into the minor groove by the aminofluorene ring. The *syn* displaced orientation of the modified dG with stacking of the aminofluorene and the looped out position of the partner dC could be envisioned to cause polymerase stalling associated with subsequent misalignment leading to frameshift mutations in appropriate sequences. The dA23 residue positioned opposite the adduct site is positioned in the major groove with its purine ring aligned face down over the van der Waals surface of the major groove and its amino group directed toward the T6·A21 base pair. The Hoogsteen edge of the modified guanine of [AF]dG4 and the Watson−Crick edge of dA23 positioned opposite it are approximately coplanar and directed toward each other but are separated by twice the hydrogen-bonding distance required for pairing. This structure of [AF]dG opposite dA at a model template-primer junctional site can be compared with a previous structure of [AF]dG opposite dA within a fully paired duplex [Norman, D., Abuaf, P., Hingerty, B. E., Live, D., Grunberger, D., Broyde, S., and Patel, D. J. (1989) *Biochemistry 28*, 7462−7476]. The alignment of the Hoogsteen edge of [AF]dG (*syn*) positioned opposite the Watson−Crick edge of dA (*anti*) has been observed for both systems with the separation greater in the case of the junctional alignment in the model template-primer system. However, the aminofluorene ring is positioned in the minor groove in the fully paired duplex while it stacks over the junctional base pair in the template-primer system. This suggests that the *syn* [AF]dG opposite dA junctional alignment can be readily incorporated within a duplex by a translation of this entity toward the minor groove.

The aromatic amines are a class of substances that are known carcinogens (*1*), present in the environment in tobacco smoke, automobile exhaust, broiled meats, and fish, and as byproducts of various industrial processes (*2*−*6*). 2-Aminofluorene (AF) and 2-acetylaminofluorene (AAF) are aromatic amines originally synthesized as insecticides but never used because they proved to be powerful mammalian carcinogens (*7*). Their highly potent carcinogenicity became the focus of a very large research effort to uncover the molecular origin of their tumorigenic effect.

Mutagenicity is now known to be an underlying cause of tumorigenesis, whose origin resides in structural damage to DNA in the case of chemical carcinogens such as aromatic amines and hydrocarbons (*8, 9*). Mutations may activate

† This research is supported by NIH Grant CA-49982 to D.J.P., by NIH Grants CA-75449, CA-28038, and RR-06458 and DOE Grant DE-FG02-90ER60931 to S.B., and by DOE Contract DE-AC05-960R22464 with Lockheed Martin Energy Research to B.E.H. Computations were carried out at the Department of Energy's National Energy Research Supercomputer Center and the National Science Foundation's San Diego Supercomputer Center.

* Corresponding author. Phone: 212-639-7207. FAX: 212-717-3066. E-mail: pateld@mskcc.org.

‡ Memorial Sloan-Kettering Cancer Center.

§ Oak Ridge National Laboratory.

|| New York University.

10.1021/bi991266p CCC: $18.00
Published on Web 07/24/1999

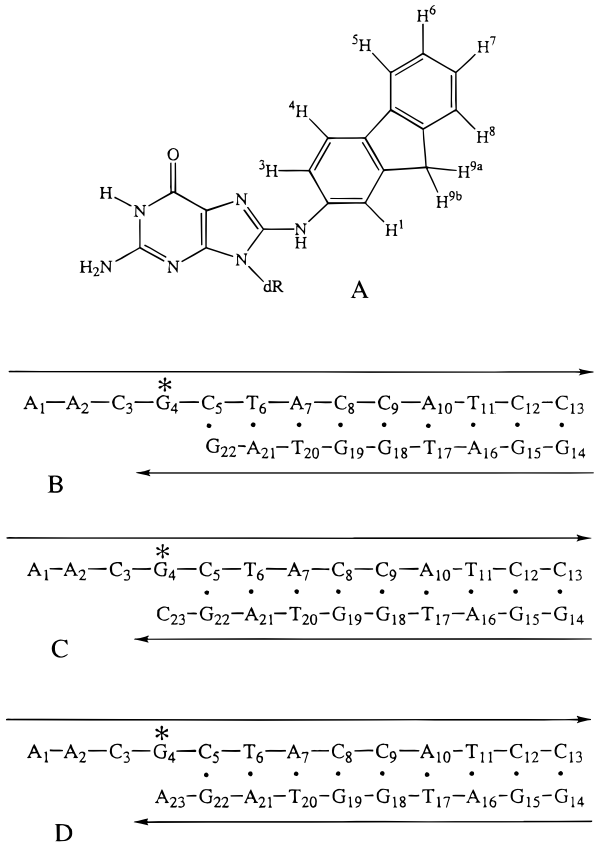

FIGURE 1: (A) Schematic of the [AF]dG adduct. Schematic and numbering system of (B) the [AF]dG 13/9-mer, (C) the [AF]dG·dC 13/10-mer, and (D) the [AF]dG·dA 13/10-mer.

oncogenes (*10, 11*) or inactivate tumor suppressors such as the p53 gene which has been found mutated in about half of all cancer cases (*12, 13*). Understanding the nature of this damage in atomic resolution detail is a necessary first step to understanding why these substances cause mutations that can lead to cancer.

It was determined in early work that AF and AAF are metabolically activated to reactive derivatives that form major in vivo adducts to C^8 of guanine. Moreover, the AAF−C^8-dG adduct can be enzymatically deacetylated in vivo to the AF−C^8-dG adduct (Figure 1A). Therefore, the most important adduct observed in many systems is the AF one, which is the focus of our interest here.

The extensive literature on biological effects of AF and AAF adducts has been reviewed by Heflich and Neft (*14*). Extensive mutagenicity studies both in vivo and in vitro have determined that the AF−C^8-dG adduct can have miscoding consequences. These are predominantly mismatches of the G·A type (*15*), although other types of mismatches and frameshifts have also been observed in certain systems. Moreover, the AF adduct also readily permits normal incorporation of dC opposite the lesion, and polymerases can generally replicate through the lesion, although specific base sequence contexts can foster blockage (*16*−*18*).

In an effort to elucidate the structural underpinnings associated with the treatment of the AF−C^8-dG adduct by polymerases and to understand the structural origins to the preference for the G·A mismatches, we have undertaken an investigation of the solution structures of the adducts at single-strand−double-strand junctions that mimic a primer-template complex at an arm of a replication fork by high-resolution NMR combined with molecular mechanics and subsequent relaxation matrix refinement computations. In previous work, we have studied the structure of the [AF]dG adduct opposite dA in a normal duplex (*19*) by a combined high-resolution NMR and molecular mechanics approach. It was found that the AF is sandwiched between the walls of the B-DNA minor groove, spanning the groove so that its distal ring is in proximity to the sugar phosphate backbone of the partner strand. The modified guanine adopts the *syn* conformation in the structure but remains stacked into the helix. The mismatched adenine is in the normal *anti* conformation and is also stacked intrahelically. A similar structure has also been observed when the [AF]dG is mismatched with dG or dI (*20*); we have also studied such a structure in the sequence context of Figure 1B, in which the modified guanine has no partner, modeling the situation in which replication has proceeded up to but not including the damaged guanine (*21*). The structure in this case contained a *syn* modified guanine displaced from its normal stacked position with insertion of the AF in its place.

In this paper on the [AF]dG adduct at single-strand−double-strand junctions, we investigated the same sequence context as previously reported for no base opposite the lesion site (Figure 1B) (*21*), except that either dC (Figure 1C) or dA (Figure 1D) was positioned opposite the lesion. This affords the opportunity for envisioning the situation following polymerase insertion of correct or mismatched base positioned opposite the lesion site. We find that the dA is situated in close proximity to the *syn* [AF]dG with structural features that are related to both those of the mismatched full duplexes and those of the single-strand−double-strand junction with no partner or dC opposite the [AF]dG. The structures provide insights into how a mismatched dA could be potentially stably incorporated opposite the [AF]dG, while the dC is looped out in this sequence context in solution.

MATERIALS AND METHODS

Materials. *N*-Acetoxy-2-acetylaminofluorene (AAAF) was purchased from Chemsyn Science Laboratories. The deoxyoligonucleotides d(A-A-C-G-C-T-A-C-C-A-T-C-C), d(G-G-A-T-G-G-T-A-G-C), and d(G-G-A-T-G-G-T-A-G-A) were synthesized on an Applied Biosystems Model 392 DNA synthesizer and purified by reverse-phase HPLC.

Preparation of Adduct. The AF-modified d(A-A-C-G-C-T-A-C-C-A-T-C-C) adducted sequence was prepared and purified using methods described previously (*19*). Briefly, the oligonucleotide was first reacted with 8 molar equiv of *N*-acetoxy-2-AAF in 2 mM sodium citrate (pH 7.0) containing 30% ethanol at 37 °C for 3 h to produce the [AAF]dG-modified oligomer. The purified [AAF]dG-modified oligomer was then converted to [AF]dG-modified oligomer by reacting with 1 M NaOH containing 0.3% 2-mercaptoethanol at a concentration of 2 mg/mL for 45 min at room temperature. The [AF]dG-modified oligomer was further purified by reverse-phase HPLC, desalted, and converted to sodium form.

The [AF]dG 13-mer strand was annealed either with the complementary d(G-G-A-T-G-G-T-A-G-C) 10-mer strand or with the complementary d(G-G-A-T-G-G-T-A-G-A) 10-mer strand at 70 °C, and the stoichiometry was followed by monitoring single proton resonances in both strands.

NMR Experiments. All NMR spectra were recorded on Varian Unity plus 600 and 500 MHz NMR spectrometers with a relaxation delay of 2.0 s unless otherwise indicated. A combination of through-space nuclear Overhauser effect (NOESY) and through-bond correlated (COSY and TOCSY) two-dimensional spectra were recorded in the States-TPPI mode (*22*) on approximately 6 mg of either the [AF]dG·dC 13/10-mer or the [AF]dG·dA 13/10-mer duplexes in 0.6 mL of 100 mM NaCl, 10 mM phosphate, aqueous buffer (pH 7.0 and pH 5.0) at 25 and 5 °C, and analyzed to assign the aminofluorene and nucleic acid protons. The NOESY spectra (150 ms mixing time) of the adduct 13/10-mers in H_2O buffer at 5 °C were collected using a jump−return pulse for solvent suppression. In addition, one-dimensional NOE difference experiments were recorded in H_2O buffer at 5 °C following saturation (1−3 s) of broad imino protons in the spectrum. NOESY spectra (50, 100, 150, and 200 ms mixing times) were collected to provide NOE buildup data in D_2O buffer at 25 °C with a relaxation delay of 2.5 s. The through-bond TOCSY data sets on the adduct in D_2O buffer were recorded at spin lock times of 40 and 80 ms at 25 °C.

The indirect proton−phosphorus correlation spectrum was recorded on the [AF]dG·dC 13/10-mer and [AF]dG·dA 13/10-mer duplexes in D_2O at 25 °C using the pulse sequence described previously (*23*). The phosphorus spectra were referenced relative to external 10% trimethyl phosphate (TMP). The $^1H-^{13}C$ HMQC correlation spectrum on the adduct 13/10-mers in D_2O buffer was recorded at 25 °C. The carbon spectra were referenced relative to external 3-(trimethylsiyl)propionate (TSP) using the method described previously (*24*).

The calculation of interproton distance bounds using volume buildup of NOE cross-peaks was based on the two-spin approximation using the dT(NH3)−dA(H2) distance of 2.92 Å for the NOESY data sets in H_2O and the dC(H5)−dC(H6) fixed distance of 2.45 Å for the NOESY data sets in D_2O solution. The upper and low bound ranges on the estimated interproton distance for nonexchangeable protons were determined based on the resolution of the cross-peaks in the two-dimensional contour plots and the quality of the NOE buildup plots.

The base proton to sugar H1′ NOE cross-peaks in the shortest mixing time NOESY data set in D_2O were evaluated to qualitatively differentiate *syn* and *anti* glycosidic torsion angles (*25*). The proton−proton vicinal coupling constants among sugar protons were analyzed from phase-sensitive COSY data to qualitatively distinguish between the C3′-endo and C2′-endo family of sugar puckers (*26*).

Molecular Mechanics Computations. Minimized potential energy calculations were carried out with DUPLEX, a molecular mechanics program for nucleic acids that performs potential energy minimizations in the reduced variable domain of torsion angle space (*27*). The vast diminution in the number of variables that must be simultaneously optimized, compared to Cartesian space, permits much larger movements from a given starting geometry, and also ensures correct internal geometry and chirality. DUPLEX uses a potential set similar to the one developed by Olson and coworkers for nucleic acids (*28*). Geometry and force field parameters, including partial charges, for the [AF]dG adduct were the same as those employed previously (*29, 30*). A hydrogen bond penalty function (*27*) was employed in all first-stage minimizations to aid the minimizer in locating the Watson−Crick hydrogen-bonded structures indicated by the NMR data. To locate minimum energy conformations with interproton distances available from the experimental NMR data, pseudo-potentials (permitting upper and lower bound restraints) were added to the energy, as described previously (*19, 31*). Briefly, the following functions were used:

$$F_N = W_N \sum_1^n (d - d_N)^2 \quad (1)$$

$$F_{NN} = W_{NN} \sum_1^n (d - d_{NN})^2 \quad (2)$$

The *W*'s are adjustable weights [in the range of 10−30 kcal/(mol·A^2)]. *d* is the current value of the interproton distance, d_N is a target upper bound, and d_{NN} a target lower bound. Equation 1 is implemented when *d* is greater than d_N, and eq 2 is implemented when *d* is less than d_{NN}. The functions are summed over all *n* target distances. F_N and F_{NN} can also be employed as relative indices of the goodness-of-fit to the NMR data. Here the *d* values are the achieved distances in a given model, and the *W*'s are the weights employed in the search. F_N and F_{NN} are composites, reflecting the overall fit of all the achieved distances to their targets. They both adopt values of zero when all model distances are within the upper and lower NMR distance bounds. Small deviations from the NMR targets, within the uncertainty of the data, are accepted in computed models, and these therefore have nonzero F_N and F_{NN} values.

Our search strategy employs a set of key NMR distance restraints, together with information from the NMR data on the hydrogen-bonding pattern between bases for the hydrogen bond penalty function. NMR information about the glycosidic torsion angle and the deoxyribose sugar pseudorotation parameter is also included. The resulting structures are ranked according to energy and goodness-of-fit (eqs 1 and 2). At this point, the first set of NMR-assigned distance bounds are evaluated in relation to achieved distances and energies in the ensemble of 16 structures. Cross-checks involving distances not employed as restraints and satisfaction of other criteria such as chemical shifts are also made. Especially in the case of conformational heterogeneity, this affords feedback to the NMR analysis and offers the opportunity for reassessment of the distance bounds. This procedure can be carried out iteratively until structures of lowest energy and best goodness-of-fit are located. This part of the computations was carried out at the Department of Energy's National Energy Research Supercomputer Center and the National Science Foundation's San Diego Supercomputer Center.

Relaxation Matrix Refinement. The final unrestrained energy-minimized structure obtained during the first-stage DUPLEX calculation was used as the starting structure for the second-stage refinement using *X-PLOR* (*32*). During the second stage, we performed molecular dynamics/simulated annealing calculations guided by the combination of the experimental NOESY intensities and NOE distances. The pseudo-energy function included two types of restraints: (1) intensity restraints for nonexchangeable protons were imposed as square-well potentials with an exponent of 2 in the penalty function, an isotropic correlation time of 5 ns, and

anisotropic bounds estimates of 10%; (2) distance restraints for nonexchangeable protons and exchangeable protons at the lesion site were retained through our protocol as square-well potentials with uniform 20% estimation of errors and 30 kcal/(mol·Å^2) force constant. A 4.5 Å cutoff was imposed for computing relaxation pathways, and the dynamics were carried out with a tolerance of 0.03 Å.

The relaxation matrix was set up for the nonexchangeable protons with the exchangeable imino, amino, and hydroxyl protons exchanged for deuterons. A total of 218 nonexchangeable distance restraints from the NOESY data sets in D_2O (872 intensities at 4 mixing times of 50 ms, 100 ms, 150 ms, and 200 ms) along with 3 exchangeable distance restraints at the lesion site from the NOESY data in H_2O were included in the calculations for the [AF]dG·dC 13/10-mer, and a similar set was used for the [AF]dG·dA 13/10-mer. Dihedral angle restraints (corresponding to B-DNA) were included with a very low weight of 5 kcal·rad^{-2} and restricted to residues which are two pairs below the modified deoxyguanosine [AF]dG4.

Six trials of the intensity refinement were performed. During each trial the starting structure was heated to 1000 K through the assignment of an arbitrary Maxwell−Boltzmann velocity distribution corresponding to a temperature of 1000 K. Then after 2.4 ps dynamics evolution at that temperature, the system was gradually cooled to 300 K during 7.2 ps with the "heat bath" method, and equilibrated at 300 K for 2.4 ps. After equilibration, the coordinates were subjected to energy minimization to a gradient of 0.1 kcal mol^{-1} Å^{-1}.

RESULTS

We first present the results for the [AF]dG·dC 13/10-mer sequence (Figure 1C) followed by their counterparts for the [AF]dG·dA 13/10-mer sequence (Figure 1D).

Exchangeable Nucleic Acid Protons in [AF]dG·dC 13/10-mer. The exchangeable proton NMR spectrum (10−14.5 ppm) of the [AF]dG·dC 13/10-mer in H_2O buffer solution, pH 7.0 at 5 °C, is plotted in Figure 2A. We observe eight well-resolved imino proton resonances between 12.5 and 14.0 ppm together with upfield-shifted imino resonances at 12.1 ppm and broadened resonances centered at 10.8 ppm. These imino protons have been assigned following analysis of the 150 ms mixing time H_2O NOESY spectrum based on established assignment procedures (*26, 33*).

The NOE connectivities between imino protons on adjacent base pairs in the [AF]dG·dC 13/10-mer can be traced from the dC5·dG22 pair adjacent to the single-strand−duplex junction to the dC13·dG14 pair located toward the opposite end of the helix in the expanded NOESY (150 ms mixing time) contour plot of the symmetrical 11.8−14.0 ppm imino proton region in Figure 3C. An expanded NOESY (150 ms mixing time) contour plot of the adduct 13/10-mer correlating NOE connectivities between the imino protons (11.8−14.0 ppm) and the base and amino protons (5.0−8.5 ppm) is plotted in Figure 3B. The observed NOE patterns establish Watson−Crick base pairing at all dA·dT pairs (thymine imino to adenine H2 protons) and at all dG·dC pairs (guanine imino to cytosine amino and H5 protons), including the dC5·dG22 base pair (peaks A, A′, and B, Figure 3B).

The imino proton of dG22, located at the junction site, is broadened and upfield-shifted (12.01 ppm) upon adduct formation relative to the control 13/10-mer (*34*). We were only able to observe few very weak intermolecular NOEs between the imino proton of dG22 and protons on the aminofluorenyl ring, and these are labeled numerically in Figure 3A,B.

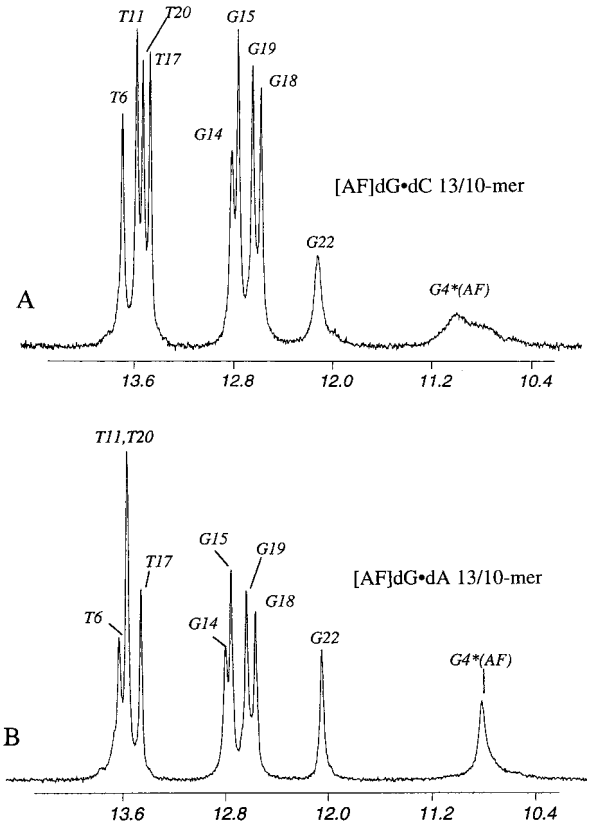

FIGURE 2: Imino proton spectra (10.4−14.0 ppm) of (A) the [AF]-dG·dC 13/10-mer and (B) the [AF]dG·dA 13/10-mer in H_2O buffer at 1 °C. The imino proton assignments are shown over the resonances in the spectra.

The imino proton of [AF]dG4 most likely resonates within the upfield-shifted imino proton envelope at 10.8 ppm. One-dimensional NOE experiments on this broad envelope failed to identify NOEs to other exchangeable and nonexchangeable protons in the [AF]dG·dC 13/10-mer for data recorded at pH 7.0 and 5.0. Its upfield chemical shift and exchange cross-peak with H_2O suggests that the guanine imino proton of [AF]dG4 is most likely looped out of the helix in the adduct 13/10-mer. The exchangeable imino and amino proton chemical shifts are listed for the entire [AF]dG·dC 13/10-mer in supplementary Table S1 (Supporting Information).

Nonexchangeable Nucleic Acid Protons in [AF]dG·dC 13/10-mer. The base and sugar H1′ nonexchangeable protons (5.1−8.5 ppm) of the [AF]dG·dC 13/10-mer in D_2O buffer, pH 7.0 at 25 °C, are shown in supplementary Figure S1A. We observe narrow and well-resolved resonances for both the single-strand and double-strand segments of the adduct 13/10-mer. The nonexchangeable proton assignments are based on an analysis of through-space NOESY (50 and 300 ms mixing times at 25 °C and 150 ms mixing time at 5 °C) data sets and through-bond COSY and TOCSY (40 and 80 ms spin lock times) data sets at 25 °C using established assignment procedures (*26, 33*).

The expanded NOESY (300 ms mixing time) contour plot establishing sequential connectivities between the base

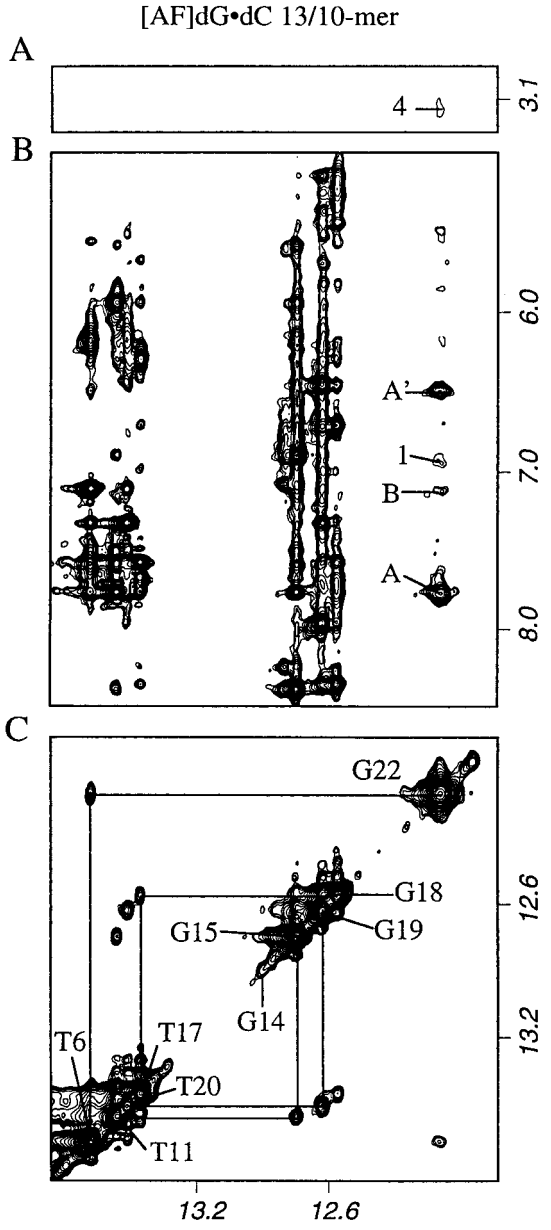

FIGURE 3: Expanded NOESY (150 ms mixing time) contour plots of the [AF]dG·dC 13/10-mer in H_2O buffer at 1 °C. (A) NOE connectivities between the imino protons (11.8−14.0 ppm) and the AF methylene protons (3.0−3.2 ppm). The carcinogen−DNA cross-peak 4 is assigned as follows: 4, G22(NH1)−AF(H9a,b). (B) NOE connectivities between the imino protons (11.8−14.0 ppm) and the base and amino proton regions (5.0 to 8.5 ppm). The NOE cross-peaks involving the imino protons are labeled in the figure as follows: A, A′, G22(NH1)−C5(NH_2-4b,e); B, G22(NH1)−A21-(H2). The carcinogen−DNA NOE cross-peak 1 is assigned as follows: 1, G22(NH1)−AF(H3,H4). (C) NOE connectivities in the symmetrical (11.8−14.0 ppm) region. The imino proton assignments are labeled along the diagonal. The lines trace the NOE connectivities between adjacent base pairs starting at dG22 toward one end of the helix and proceeding to dG14 toward the other end of the helix.

protons (6.4−8.4 ppm) and the sugar H1′ and cytosine H5 protons (5.15−6.35 ppm) of the [AF]dG·dC 13/10-mer in D_2O buffer, pH 7.0 at 25 °C, is plotted in duplicate in Figure 4. The base to sugar H1′ proton NOE connectivities are traced from dA1 to dA7 along the modified strand (solid line, Figure 4A) and from dT20 to dC23 along the complementary strand (dashed line, Figure 4B) in the contour plot of the adduct 13/10-mer. The observed break in the sequential NOE connectivity at the dC3−[AF]dG4 step on the modified strand is due to the absence of a purine H8 proton following AF modification at the C^8 position of dG4 in the adduct duplex. We observe weak interresidue NOE cross-peaks for the single-stranded d(A1-A2-C3) segment (Figure 4A) on the modified strand of the adduct 13/10-mer. These base and sugar H1′ proton assignments have been confirmed by cross-checks in other regions of the NOESY plot (Figure 5A), as well as from COSY (supplementary Figure S2A) and TOCSY plots of the adduct 13/10-mer. A complete set of nonexchangeable nucleic acid base and sugar proton chemical shifts for the entire [AF]dG·dC 13/10-mer is listed in Table S1.

We also observe an inversion of the H2′ and H2″ sugar proton chemical shifts of [AF]dG4 at the lesion site in the adduct 13/10-mer. The H2′ proton of [AF]dG4 undergoes a dramatical downfield chemical shift to 3.56 ppm as can be seen in both NOESY (Figure 5A) and COSY (Figure S2A) plots of the [AF]dG·dC 13/10-mer. By contrast, the H2″ proton of [AF]dG4 has an unperturbed chemical shift of 2.38 ppm.

Aminofluorene Protons in [AF]dG·dC 13/10-mer. The nonexchangeable aminofluorene protons have been assigned based on the analysis of the through-space NOE patterns and through-bond coupling connectivities in the [AF]dG·dC 13/10-mer. These nonexchangeable aminofluorene proton chemical shifts in the [AF]dG·dC 13/10-mer are compared in Figure S3 to the corresponding values of the AF-stacked conformer in the [AF]dG 13/9-mer (*21*) and those in the AF-intercalated (*35*) and AF-external (*36*) conformers of the [AF]-dG·dC 12-mer duplex published previously. The aromatic aminofluorene ring protons resonate between 6.5 and 7.3 ppm in the [AF]dG·dC 13/10-mer (closed squares, Figure S3), very similar to what was reported previously for the [AF]-dG 13/9-mer (open squares, Figure S3) in which the AF stacks on one face with the dG5·dC22 base pair and the other face with the purine ring of dA2 (*21*). By contrast, downfield shifts are observed for the AF protons in the AF-external conformer (*36*) of the [AF]dG·dC 12-mer duplex (open circles, Figure S3). The chemical shifts of the aminofluorene protons in the [AF]dG·dC 13/10-mer are listed in the caption to Figure S3.

We have not been able to locate the resonance from the NH proton involved in the covalent linkage between the aminofluorene and the G4 residue in spectra of the [AF]-dG·dC 13/10-mer recorded at both pH 7.0 and pH 5.0.

Intermolecular NOEs in [AF]dG·dC 13/10-mer. The intermolecular AF−DNA NOE cross-peaks which have been identified in the NOESY spectra of the [AF]dG·dC 13/10-mer are listed in Table 1. Some of these intermolecular AF−DNA NOE cross-peaks are labeled by numbers in the expanded NOESY plots of the exchangeable protons in H_2O solution (Figure 3) and nonexchangeable protons in D_2O solution (Figures 4 and 5A), and their assignments are listed in the figure captions.

The observed intermolecular NOE patterns (Table 1) are consistent with a single alignment of the aminofluorene at the junction site in the [AF]dG·dC 13/10-mer. The aminofluorene AF(H1) proton exhibits intermolecular NOEs to dA2(H1′), dA2(H8), dC5(H1′), and dC5(H5) protons, while the AF(H3) proton on the opposite side of the aminofluorene ring exhibits intermolecular NOEs to the minor groove sugar

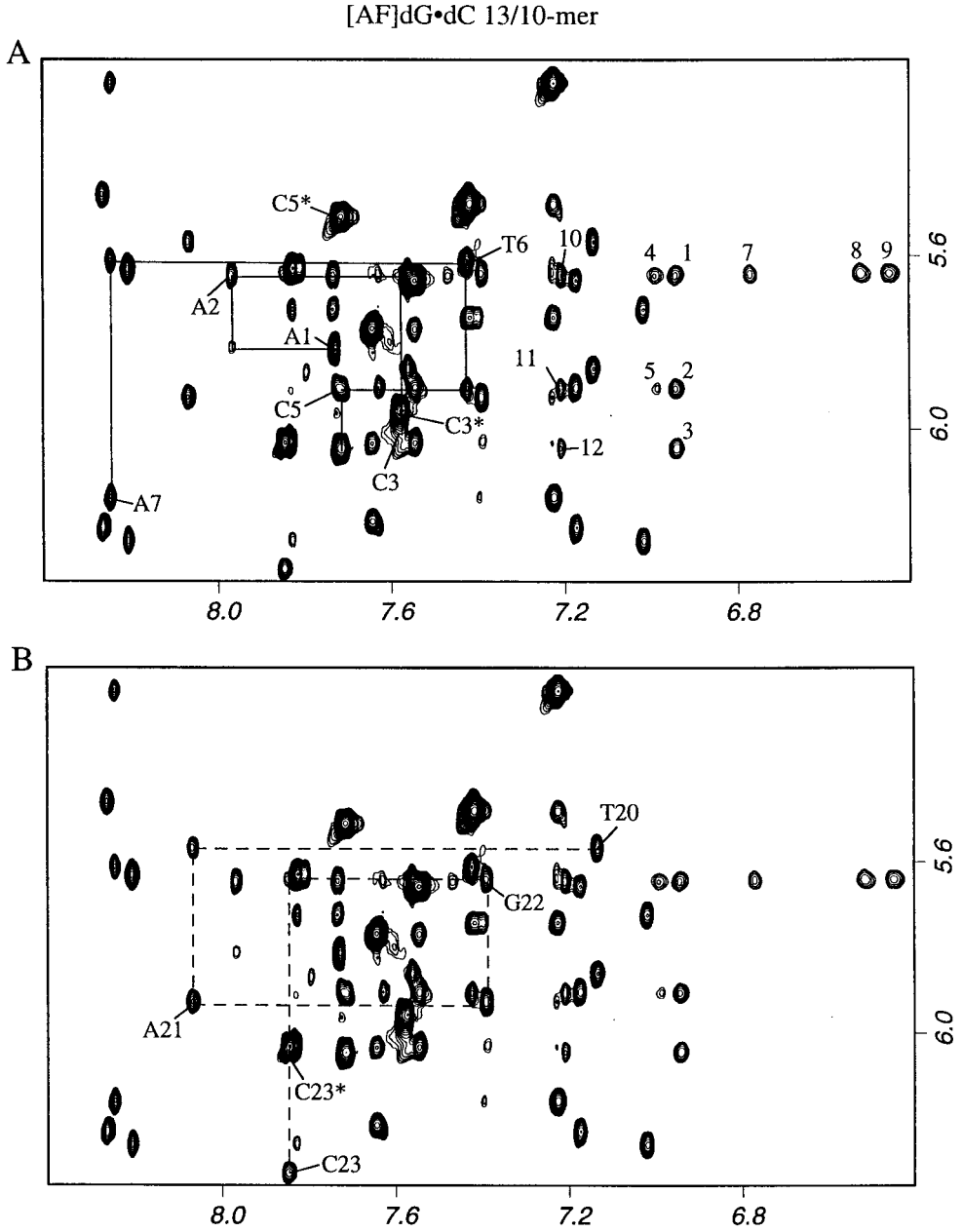

FIGURE 4: Duplicate expanded NOESY (300 ms mixing time) contour plots of the [AF]dG·dC 13/10-mer in D_2O buffer at 25 °C establishing distance connectivities between the base (purine H8 and pyrimidine H6) protons (6.4−8.4 ppm) and the sugar H1′ and deoxycytidine H5 protons (5.2−6.4 ppm). The NOE connectivities between the base and their own and 5′-flanking sugar H1′ protons are traced in (A) from dA1 to dA7 on the modified strand (solid lines) and in (B) from dT20 to dC23 on the unmodified partner strand (dashed lines). The assignments label the base to their own sugar H1′ NOEs, while the deoxycytidine H6−H5 NOEs are designated by asterisks. Note that the NOE cross-peak at the dC3-[AF]dG4 step is missing because of the absence of an H8 proton for [AF]dG4. The carcinogen−DNA NOE cross-peaks are assigned as follows: 1, AF(H3)−A2(H1′); 2, AF(H3)−C5(H1′); 3, AF(H3)−[AF]G4(H1′); 4, AF(H4)−A2(H1′); 5, AF(H4)−C5(H1′); 6, AF(H8)−G22(H1′); 7, AF(H6)−G22(H1′); 8, AF(H7)−G22(H1′); 9, AF(H1)−A2(H1′); 10, AF(H1)−C5(H1′); 11, AF(H1)−[AF]G4(H1′).

protons of dA2(H1′), [AF]dG4(H1′), and dC5(H1′) in the adduct 13/10-mer (Table 1). The AF(H4) proton only shows relatively weak NOEs to the minor groove sugar protons of dA2(H1′) and dC5(H1′), while the AF(H6) and AF(H7) protons exhibit NOEs to dG22(H1′), dG22(H2′), and dG22-(H2″) (Table 1). However, none of the aminofluorene protons exhibit any NOE with dC23 which is positioned opposite to the modified dG4 in the [AF]dG·dC 13/10-mer. These intermolecular NOE data help position the [AF]dG4 and dC23 bases relative to the dC5·dG22 base pair and the dA2 base in the adduct 13/10-mer.

Carbon Spectra of [AF]dG·dC 13/10-mer. The expanded contour plot of a natural-abundance proton−carbon HMQC experiment that correlates the H1′ and C1′ chemical shifts of individual residues of the [AF]dG·dC 13/10-mer in D_2O buffer at 25 °C is plotted in Figure S4A. The carbon C1′ resonances are assigned on the basis of the known sugar H1′ proton assignments in the adduct 13/10-mer. The C1′ chemical shift assignments for residues in the d(A2-C3-[AF]-G4-C5-T6)·d(A21-G22-C23) segment of the adduct duplex are labeled in Figure S4A. It is noted that the ^{13}C chemical shift of the C1′ carbon of [AF]dG4 resonates at 86.7 ppm,

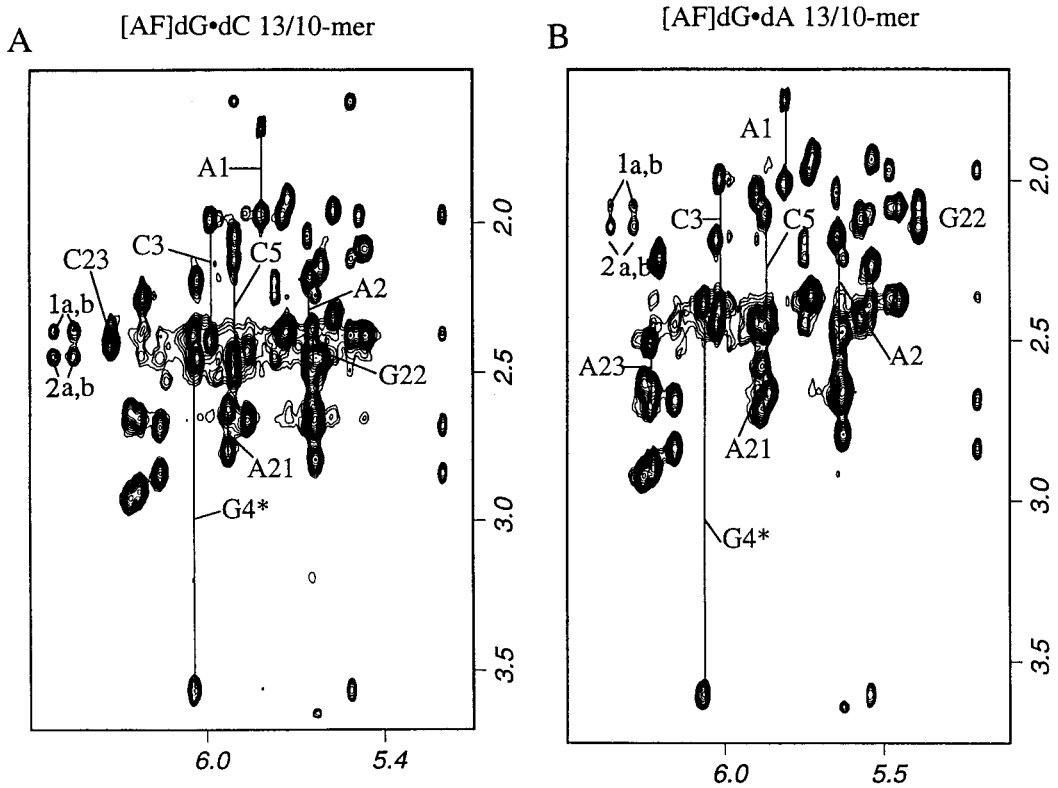

FIGURE 5: Expanded NOESY (300 ms mixing time) contour plots of (A) the [AF]dG·dC 13/10-mer and (B) the [AF]dG·dA 13/10-mer in D_2O buffer at 25 °C showing NOEs between the sugar H1′ protons (5.1−6.5 ppm) and the H2′, H2″ protons (1.6−3.7 ppm). The H2′ and H2″ protons of dA1, dA2, dC3, [AF]dG4, dC5, dA21, dG22, and dA23/dC23 are connected by lines and labeled. The H2′ protons resonate upfield of the H2″ protons for the majority of these residues except for [AF]dG4 and dA2 where the H2″ protons resonate upfield of the H2′ proton. The intermolecular AF−DNA NOE cross-peaks are assigned in both (A) and (B) as follows: 1a,b, AF(H6,H7)−G22(H2″); 2a,b, AF(H6,H7)−G22(H2′).

Table 1: Intermolecular NOEs between Aminofluorene Protons and DNA Protons in the [AF]dG·dC 13/10-mer[a]

AF protons	DNA protons with NOEs to AF protons
AF(H1)	A2(H8),*vw*; A2(H1′),*m*; A2(H2′), *m*; A2(H2″),*m*; G4(H1′),*m*; C5(H1′),*m*
AF(H3)	A2(H1′),*m*; A2(H2′),*vw*; A2(H2″),*vw*; G4(H1′),*m*; C5(H1′),*m*; G22(NH1),*vw*
AF(H4)	A2(H1′),*w*; G4(H1′),*vw*; C5(H1′),*w*; G22(NH1),*vw*
AF(H6)	G22(H8),*vw*; G22(H1′),*m*; G22(H2′),*m*; G22(H2″),*m*
AF(H7)	G22(H8),*w*; G22(H1′),*m*; G22(H2′),*m*; G22(H2″),*m*
AF(H8)	G22(H1′),*w*; G22(H2′),*w*; G22(H2″),*w*
AF(H9)	G22(NH1),*vw*

[a] The symbols *m*, *w*, and *vw* stand for medium, weak, and very weak NOEs.

which is downfield-shifted relative to other assignable dG residues in this segment of the adduct duplex (Figure S4A).

Phosphorus Spectra of [AF]dG·dC 13/10-mer. The proton-decoupled phosphorus spectrum of the [AF]dG·dC 13/10-mer has been recorded in D_2O buffer at 25 °C. The phosphorus resonances are dispersed over a 1.5 ppm range with three resonances shifted to lower field relative to the −3.8 to −4.5 ppm spectral region. The phosphorus resonances have been assigned from an analysis of the proton-detected phosphorus−proton heteronuclear correlation experiment with the expanded contour plot shown in Figure S4B. The three- and four-bond proton−phosphorus vicinal couplings permit the correlation of the phosphorus resonances with the 5′-linked H3′ proton and the 3′-linked H4′ and H5′, 5″ protons. The phosphorus chemical shifts for the dC3-[AF]dG4 (−3.61 ppm), dG22-dC23 (−3.56 ppm), and [AF]-dG4-dC5 (−3.24 ppm) steps are shifted to lower field relative to the −3.8 to −4.5 ppm unperturbed phosphodiester backbone chemical shift range.

Molecular Mechanics Computations on [AF]dG·dC 13/10-mer. The search strategy employed began with a B-DNA (*37*) central d(A1-A2-C3-[AF]G4-C5-T6)·d(A21-G22-C23) base pair segment of the [AF]dG·dC 13/10-mer with the computations guided by the intermolecular AF−DNA restraints listed in Table 1. The AF−DNA orientation space was searched with 16 energy minimization trials in which the linkage torsion angles α′ ([AF]dG4(N^9)-[AF]dG4(C^8)-[AF](N)-[AF](C^2)) and β′ ([AF]d4(C^8)-[AF](N)-[AF](C^2)-[AF](C^1)) were each started at 0°, 90°, 180°, and 270° in all combinations. Searching the orientation space at 90° intervals of α′ and β′ is a robust procedure for locating all the important potential energy wells because our minimization protocol permits torsion angle variations of up to 100° in each minimization step (*27*). Consequently, energy minima in each quadrant of α′ and β′ are accessible, and the reduced variable domain of torsion angle space greatly enhances the likelihood of finding the important structures. In these trials, the DUPLEX hydrogen bond penalty function (*27*) for Watson−Crick base pairing was utilized at all base pairs except at the [AF]dG4·dC23 site for which there was no indication of hydrogen bonding in the NMR data of the adduct 13/10-mer.

Of the 16 structures computed, 3 structures exhibited good fit to the NMR data and had low energies (Table S2, Supporting Information). These three structures for the [AF]-

Table 2: NMR Refinement Statistics for the [AF]dG·dC 13/10-mer

NMR restraints	
NOE distance restraints for entire 13/10-mer	218
NOE distance restraints in central 6/3-mer region [a]	117
Carcinogen−DNA distance restraints	27
entire 13/10-mer intensities	872 (total for 4 mixing times)
central 6/3-mer region[a] intensities	468 (total for 4 mixing times)
structure statistics	
NMR *R*-factor ($R_{1/6}$)	2.8%
rmsd of NOE violations	0.087 ± 0.004
no. of NOE violations >0.2 Å	14 ± 1
no. of NOE violations >0.2 Å in center 6/3-mer region[a]	7 ± 1
deviations from ideal geometry	
bond length (Å)	0.011 ± 0.001
bond angle (deg)	3.0 ± 0.1
impropers (deg)	0.4 ± 0.1
pairwise rmsd (Å) among six intensity refined structures[b]	
entire 13/10-mer	1.52 ± 0.26
central 6/3-mer region[a]	1.37 ± 0.20
6/3-mer region[a] without backbone	1.17 ± 0.21

[a] The d(A1-A2-C3-[AF]G4-C5-T6)·d(A21-G22-C23) segment. [b] Heavy atoms only.

dG·dC 6/3-mer segment viewed normal to the helix axis are superpositioned in Figure S5. The structure with the lowest value of the goodness-of-fit [$F_N = 5.7$, $F_{NN} = 0.1$, with $W = 30$ kcal/(mol·Å^2)] and the lowest energy (−192.7 kcal/mol) was embedded into an energy minimized B-form 13/10-mer and reminimized with all restraints. Subsequently, the hydrogen bond penalty function and the distance restraints were released with energy minimization in one step, yielding the unrestrained DUPLEX-based structure of the [AF]dG·dC 13/10-mer.

Relaxation Matrix Refinement on [AF]dG·dC 13/10-mer. The DUPLEX-based unrestrained structure of the [AF]dG·dC 13/10-mer served as the starting structure for intensity-based relaxation matrix refinements using the XPLOR-based computational protocol (*32*) outlined under Materials and Methods. An ensemble of six structures obtained during intensity refinement demonstrated an improved correspondence with experimental intensity and distance restraint data sets. The number of NOE distances violated by more than 0.2 Å decreased from 69 to 14 (with 7 violations in the central d(A1-A2-C3-[AF]G4-C5-T6)·d(A21-G22-C23) 6/3-mer segment with respect to different structures in the ensemble), and the NMR *R*-factor ($R_{1/6}$) improved from an initial value of 8.2% to 2.8% following intensity refinement. The refined structures exhibit good stereochemistry with rmsd values for bond length violations of 0.011 ± 0.001 Å, and bond angle and improper dihedral angle violations of 3.0° ± 0.1° and 0.4° ± 0.1°, respectively.

The rmsd of all heavy atoms between the six intensity refined structures and the initial DUPLEX structure is 2.3 ± 0.2 Å for all heavy atoms and 2.5 ± 0.3 Å for the 6/3-mer d(A1-A2-C3-[AF]G4-C5-T6)·d(A21-G22-C23). At the same time, good convergence was observed among the intensity refined structures. The pairwise rmsd in the set is 1.52 ± 0.26 Å for all heavy atoms, 1.37 ± 0.20 Å for the heavy atoms of the d(A1-A2-C3-[AF]G4-C5-T6)·d(A21-G22-C23) 6/3-mer segment, and 1.17 ± 0.21 Å for the heavy atoms of the 6/3-mer excluding backbone atoms (see Table 2).

Table 3: Intermolecular NOEs between Aminofluorene Protons and DNA Protons in the [AF]dG·dA 13/10-mer[a]

AF protons	DNA protons with NOEs to AF protons
AF(H1)	A2(H8),*w*; A2(H1′),*m*; A2(H2′),*m*; A2(H2″),*m*; G4(H1′),*m*; C5(H5),*w*; A23(H2),*w*
AF(H3)	A2(H8),*vw*; A2(H1′),*m*; A2(H2′),*vw*; A2(H2″),*w*; G4(H1′),*m*; G4(H2″),*vw*; C5(H1′),*m*; G22(NH1),*w*
AF(H4)	A2(H1′),*w*; A2(H2′),*vw*; A2(H2″),*vw*; A2(H8),*vw*; G4(H1′),*vw*; C5(H1′),*w*; G22(NH1),*w*; A23(H2),*vw*
AF(H6)	G22(NH1),*w*; G22(H1′),*m*; G22(H2′),*m*; G22(H2″),*m*
AF(H7)	G22(NH1),*w*; G22(H1′),*m*; G22(H2′),*m*; G22(H2″),*m*
AF(H8)	G22(H2″),*w*; A23(H2),*vw*
AF(H9)	G22(NH1),*w*

[a] The strengths of the NOE cross-peaks are designated by *m* (medium), *w* (weak), and *vw* (very weak).

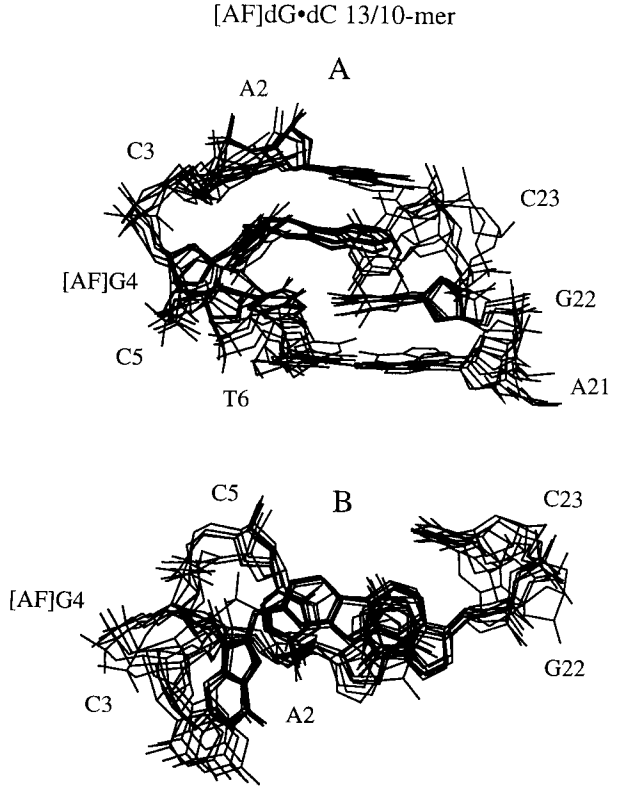

FIGURE 6: Superposition of six intensity refined structures of the [AF]dG·dC 13/10-mer. (A) View looking into the major groove and normal to the helix axis of the d(A2-C3-[AF]G4-C5-T6)·d(A21-G22-C23) segment. (B) View looking down the helix axis of the d(A2-C3-[AF]G4-C5)·d(G22-C23) segment.

The superpositioned view of six intensity refined structures of the d(A2-C3-[AF]G4-C5-T6)·d(A21-G22-C23) 5/3-mer segment of the [AF]dG·dC 13/10-mer is plotted in Figure 6A. The corresponding view looking down the helix axis of the central 4/2 segment is plotted in Figure 6B.

Solution Structures of [AF]dG·dC 13/10-mer. A view normal to the helix axis and looking into the major groove of the central d(A2-C3-[AF]G4-C5-T6)·d(A21-G22-C23) 5/3-mer segment of a representative structure from the ensemble of six intensity refined structures of the [AF]dG·dC 13/10-mer is shown Figure 7A. The corresponding structure for the entire adduct 13/10-mer is shown in stereo in Figure S6. The covalently linked aminofluorene ring stacks over the dC5·dG22 base pair and also stacks with the purine ring of the unpaired dA2 residue (Figure 7A). The guanine

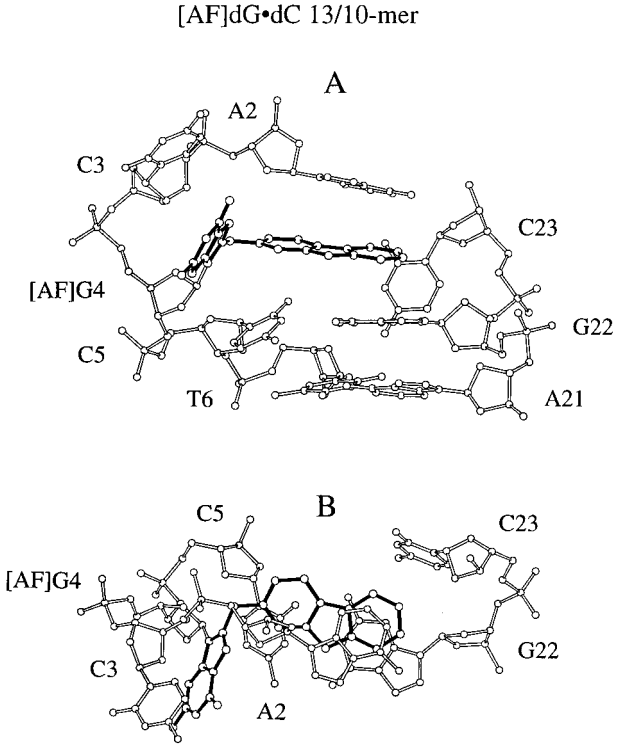

FIGURE 7: (A) View looking into the major groove and normal to the helix axis of the central d(A2-C3-[AF]G4-C5-T6-A7-C8)·d(A21-G22-C23) segment of a representative intensity refined structure of the [AF]dG·dC 13/10-mer. The AF ring system is shown in darkened bonds and is stacked over the dC5·dG22 base pair. The modified dG4 base is displaced toward the major groove. (B) View looking down the helix axis for the d(A2-C3-[AF]G4-C5)·d(G22-C23) segment of a representative intensity refined structure of the [AF]dG·dC 13/10-mer. Figures were prepared using Molscript VI.1 (*52*).

ring of the syn [AF]dG4 is displaced toward the major groove side while dC23 positioned opposite it is looped out toward the minor groove side and positioned outside the helix (Figure 7A).

A view looking down the helix axis of the central d(A2-C3-[AF]G4-C5)·d(G22-dC23) 4/2-mer segment of the [AF]-dG·dC 13/10-mer is shown in Figure 7B. This view emphasizes the overlapping geometry between the aminofluorene ring system and the flanking dC5·dG22 base pair and the dA2 residue. The long axis of the aminofluorene ring of AF is approximately parallel to the long axis of the dC5·dG22 base pair (Figures 7B). The carcinogen−base linkage site for the [AF]dG4 residue is defined by the torsion angles α' ([AF]dG4(N^9)-[AF]dG4(C^8)-[AF](N)-[AF](C^2)) and β' ([AF]dG4(C^8)-[AF](N)-[AF](C^2-[AF] (C^1)). These α', β' angles adopt values of 229 ± 10° and 10 ± 4°, respectively. The $\beta' = 10°$ orientation directs the C^9-containing edge of the aminofluorene toward the major groove side (Figure 7B) for the [AF]dG·dC 13/10-mer. The pseudorotation angles and glycosidic torsion angles for the d(A1-A2-C3-[AF]G4-C5-T6)·d(A21-G22-dC23) segment of the [AF]dG·dC 13/10-mer are listed in Table S3. The χ-(O4′−C1′−N9−C4) glycosidic torsion angle of the [AF]dG4 residue is *syn*, with a value of $\chi = 104 \pm 8°$. Torsion angles and pseudorotation parameters of the [AF]dG·dC 13/10-mer are within or near ranges observed in B-DNA crystals (*38*), although some values are in less common domains as indicated in Table S3. These values are essentially confined to the single-stranded region and its immediate vicinity.

Table 4: NMR Refinement Statistics for the [AF]dG·dA 13/10-mer Duplex

NMR restraints	
NOE distance restraints for entire 13/10-mer	216
NOE distances in the central 6/3-mer region[a]	128
carcinogen−DNA distance restraints	33
entire 13/10-mer intensities	864 (4 mixing times)
central 6/3-mer region[a] intensities	512 (4 mixing times)
structure statistics	
NMR *R*-factor ($R_{1/6}$)	2.5%
rmsd of NOE violations	0.060 ± 0.006
no. of NOE violations >0.2 Å	6 ± 2
no. of NOE violations >0.2 Å in center 6/3-mer region[a]	4 ± 1
deviations from ideal geometry	
bond length (Å)	0.009 ± 0.001
bond angle (deg)	2.8 ± 0.1
impropers (deg)	0.4 ± 0.1
pairwise rmsd (Å) among six intensity refined structures[b]	
entire 13/10-mer	1.86 ± 0.42
central 6/3-mer region[a]	1.36 ± 0.36
6/3-mer region[a] without backbone	1.04 ± 0.37

[a] The d(A1-A2-C3-[AF]G4-C5-T6)·d(A21-G22-A23) segment. [b] Heavy atoms only.

Comparison of Junctional [AF]dG·dC 13/10-mer and [AF]-dG 13/9-mer Structures. We previously reported on the solution structure of the [AF]dG 13/9-mer (sequence in Figure 1B), where there is no base opposite the [AF]dG adduct at the junctional site (*21*). This structure spanning the d(A2-C3-[AF]G4-C5)·d(G22) junctional segment is shown in Figure 8A. It can be compared with the corresponding solution structure of the [AF]dG·dC 13/10-mer (sequence in Figure 1C), whose d(A2-C3-[AF]G4-C5)·d(G22-C23) junctional segment is shown in Figure 8B.

Exchangeable Nucleic Acid Protons in [AF]dG·dA 13/10-mer. The exchangeable proton NMR spectrum (10.0−14.5 ppm) of the [AF]dG·dA 13/10-mer in H_2O buffer solution, pH 7.0 at 5 °C, is plotted in Figure 2B. Seven well-resolved imino proton resonances are observed between 12.5 and 14.0 ppm together with a narrow imino proton resonance at 12.1 ppm and a somewhat broadened imino proton resonance at 10.8 ppm. The connectivities between imino protons on adjacent base pairs in the [AF]dG·dA 13/10-mer are traced in the symmetrical 11.8−14.0 ppm imino proton region of the NOESY (150 ms mixing time) spectrum in Figure 9C. The NOE connectivities between the imino protons (11.8−14.0 ppm) and the base and amino protons (5.0−8.5 ppm) are plotted in Figure 9B. The observed NOE patterns establish Watson−Crick base pairing at all dA·dT and at all dG·dC pairs including the dC5·dG22 base pair (peaks A, A′, and B, Figure 9B). The somewhat broadened 12.01 ppm imino proton of dG22 of the dC5·dG22 base pair exhibits several intermolecular NOEs to aminofluorene protons, and these are labeled 1−4 in Figure 9A,B.

The upfield-shifted imino proton at 10.8 ppm is assigned to [AF]dG4 based on NOEs to the NH_2 and H2′, H2″ protons of dC5 identified in a one-dimensional difference NOE experiment on the adduct 13/10-mer in H_2O buffer at 1 °C (Figure S7). The imino proton of [AF]dG4 also shows a weak NOE to the H2 proton of dA23 in this difference spectrum (Figure S7). The exchangeable imino and amino proton chemical shifts for the entire [AF]dG·dA 13/10-mer at 5 °C are listed in Table S5.

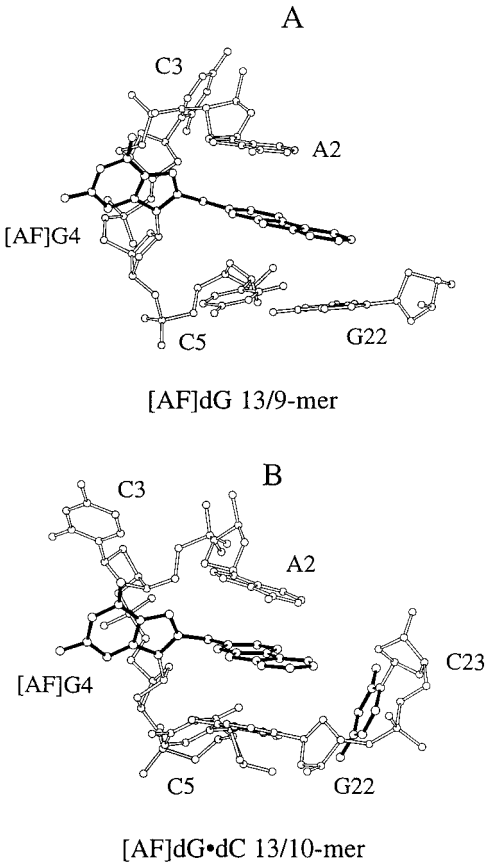

FIGURE 8: Comparison of the solution structures of (A) the [AF]-dG 13/9-mer showing the d(A2-C3-[AF]G4-C5)·d(G22) junctional segment (*21*) and (B) the [AF]dG·dC 13/10-mer showing the d(A2-C3-[AF]G4-C5)·d(G22-C23) junctional segment. The [AF]dG4 and dC23 residues are shown in darkened bonds. Figures were prepared using Molscript VI.1 (*52*).

Nonexchangeable Nucleic Acid Protons in [AF]dG·dA 13/10-mer. The base and sugar H1′ nonexchangeable proton spectrum region (5.0−8.5 ppm) of the [AF]dG·dA 13/10-mer in D_2O buffer, pH 7.0 at 25 °C, is shown in Figure S1B. The expanded NOESY (300 ms mixing time) contour plot establishing sequential connectivities between the base protons (6.75−8.3 ppm) and the sugar H1′ and cytosine H5 protons (5.1−6.4 ppm) of the [AF]dG·dA 13/10-mer in D_2O buffer, pH 7.0 at 25 °C, is plotted in duplicate in Figure 10. The base to sugar H1′ proton connectivities can be traced from dA1 to dA7 along the modified strand (solid line, Figure 10A) and from dT20 to dA23 along the complementary strand (dashed line, Figure 10B). The break in the tracing at the dC3-[AF]dG4 step on the modified strand is due to the absence of a purine H8 proton following AF modification at the C^8 position of dG4 in the adduct duplex. We also observe weaker interresidue cross-peaks for the single-stranded d(A1-A2-C3) segment on the modified strand of the adduct 13/10-mer (Figure 10A). These base and sugar H1′ proton assignments have been confirmed by cross-checks in other regions of the NOESY plot (Figure 5B), as well as from COSY (Figure S2B) and TOCSY plots of the adduct 13/10-mer. The H2 protons of A2 (7.52 ppm) and A23 (7.80 ppm) exhibit weak intermolecular NOEs to several aminoflourene protons in the [AF]dG 13/10-mer as shown in Figure S8.

The sugar H2′ proton chemical shift of [AF]dG4 (3.60 ppm) is dramatically downfield shifted as can be observed

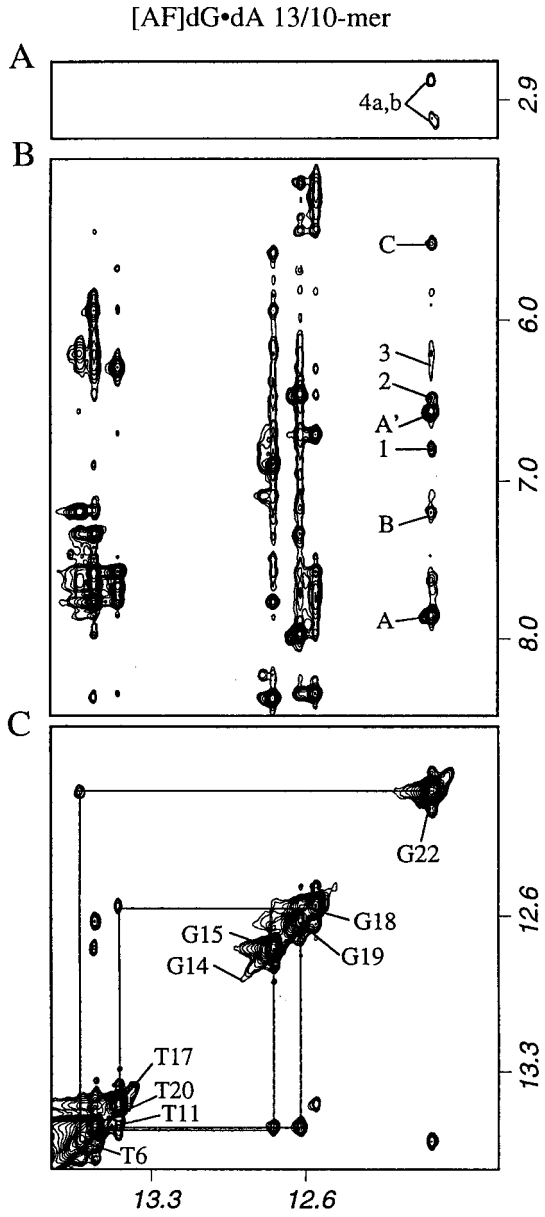

FIGURE 9: Expanded NOESY (150 ms mixing time) contour plots of the [AF]dG·dA 13/10-mer in H_2O buffer at 1 °C. (A) NOE connectivities between the imino protons (11.8−14.0 ppm) and the AF methylene protons (3.0−3.2 ppm). The intermolecular AF−DNA cross-peak 4 is assigned as follows: 4a,b, G22(NH1)−AF-(H9a,b). (B) NOE connectivities between the imino protons (11.8−14.0 ppm) and the base and amino proton regions (5.0−8.5 ppm). The NOE cross-peaks involving the imino protons are labeled in the figure as follows: A, A′, G22(NH1)−C5(NH_2-4b,e); B, G22-(NH1)−A21(H2); C, G22(NH1)−C5(H5). The intermolecular AF−DNA NOE cross-peaks 1−3 are assigned as follows: 1, G22-(NH1)−AF(H3,4); 2, G22(NH1)−AF(H5); 3, G22(NH1)−AF(H6,7). (C) NOE connectivities in the symmetrical imino proton (11.8−14.0 ppm) region. The imino proton assignments are labeled along the diagonal. The lines trace the NOE connectivities between adjacent base pairs starting at dG22 toward one end of the helix and proceeding to dG14 toward the other end of the helix.

in both NOESY (Figure 5B) and COSY (Figure S2B) plots of the [AF]dG·dA 13/10-mer. The nonexchangeable nucleic acid base and sugar proton chemical shifts for the entire [AF]-dG·dA 13/10-mer are listed in Table S5.

Aminofluorene Protons in [AF]dG·dA 13/10-mer. The aromatic aminofluorene ring protons resonate between 6.2 and 7.2 ppm in the [AF]dG·dA 13/10-mer (closed circles,

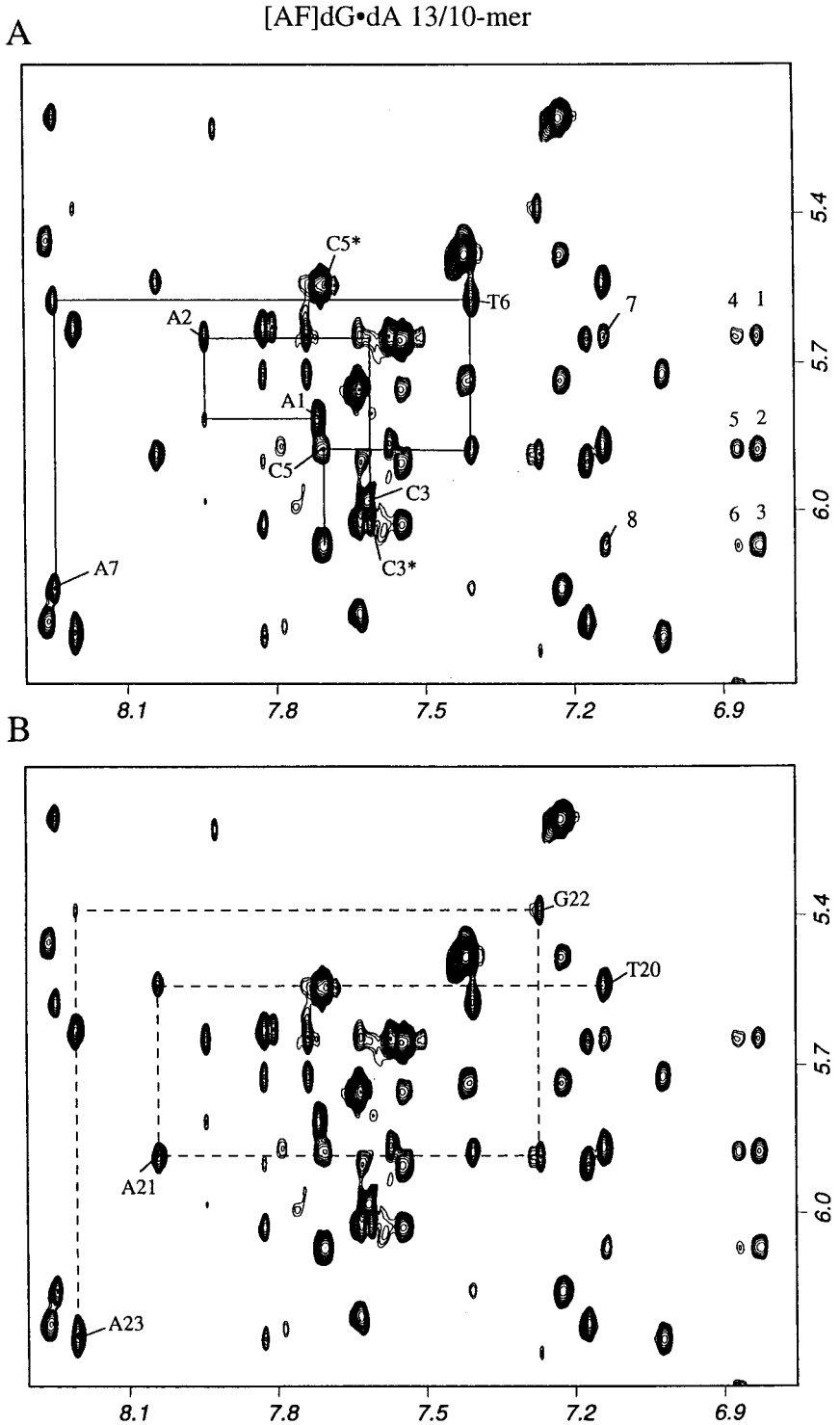

FIGURE 10: Duplicate expanded NOESY (200 ms mixing time) contour plots of the [AF]dG 13/9-mer in D_2O buffer at 25 °C establishing distance connectivities between the base (purine H8 and pyrimidine H6) protons (6.7−8.4 ppm) and the sugar H1′ and deoxycytidine H5 protons (5.1−6.4 ppm). The NOE connectivities between the base and their own and 5′-flanking sugar H1′ protons are traced in (A) from dA1 to dA7 on the modified strand (solid lines) and in (B) from dT20 to dA23 on the unmodified partner strand (dashed lines). The assignments label the base to their own sugar H1′ NOEs, while the deoxycytidine H6−H5 NOEs are designated by asterisks. Note that the NOE cross-peak at the dC3-[AF]dG4 step is missing because of the absence of an H8 proton for [AF]dG4. The intermolecular AF−DNA NOE cross-peaks are assigned as follows: 1, AF(H3)−A2(H1′); 2, AF(H3)−C5(H1′); 3, AF(H3)−[AF]G4(H1′); 4, AF(H4)−A2(H1′); 5, AF(H4)−C5(H1′); 6, AF(H1)−[AF]G4(H1′); 7, AF(H1)−A2(H1′); 8, AF(H1)−[AF]G4(H1′). The cross-peak between AF(H1) and C5-(H1′) is overlapped with T20(H6)−G19(H1′) and is not indicated.

Figure S9) with values very similar to those observed in the related [AF]dG·dC 13/10-mer (open circles, Figure S9) and [AF]dG 13/9-mer (open squares, Figure S9) (*21*) junctional sequences. The aminofluorene proton chemical shifts in the [AF]dG·dA 13/10-mer are listed in the caption to Figure S9. We have been unable to locate the linker NH proton resonance of [AF]dG4 in the NMR spectra of the [AF]dG·dA 13/10-mer.

Intermolecular NOEs in [AF]dG·dA 13/10-mer. The intermolecular AF−DNA NOE cross-peaks which have been identified in the NOESY spectra of the [AF]dG·dA 13/10-mer are listed in Table 3. Some of these intermolecular AF−DNA NOE cross-peaks are labeled by numbers in the expanded NOESY plots of the exchangeable protons (Figure 9) and nonexchangeable protons (Figures 10A, 10B, and 5B) of the [AF]dG·dA 13/10-mer, and their assignments are listed in the figure captions. The majority of the intermolecular NOEs are between the aminofluorene ring protons and the base and sugar protons of dA2 and the dC5·dG22 base pair in the adduct 13/10-mer (Table 3). In addition, the H2 proton of dA23 exhibits weak NOEs to the aminofluorene H1 proton and very weak NOEs to the aminofluorene H4 and H8 protons in the adduct 13/10-mer (Table 3).

Carbon and Phosphorus Spectra of [AF]dG·dA 13/10-mer. The expanded contour plot of a natural-abundance proton−carbon HMQC correlation experiment that correlates the H1′ and C1′ chemical shifts of individual residues in the [AF]dG·dA 13/10-mer in D_2O buffer at 25 °C is plotted in Figure S10A. The C1′ chemical shift of [AF]dG4 at 86.7 ppm resonates to low field of other assignable dG residues in this segment of the adduct duplex (Figure S10A).

The proton-decoupled phosphorus spectrum of the [AF]-dG·dA 13/10-mer has been recorded in D_2O buffer at 25 °C (Figure S10B). The phosphorus chemical shifts for the dC3-[AF]dG4 (−3.64 ppm), dG22-dA23 (−3.68 ppm), and [AF]-dG4-dC5 (−3.32 ppm) steps are shifted to the low field of the −3.8 to −4.5 ppm region characteristic of the unperturbed phosphodiester backbone chemical shift range.

Molecular Mechanics and Relaxation Matrix Refinement Computations on [AF]dG·dA 13/10-mer. The search strategy employed began with a B-DNA (*12*) central d(A1-A2-C3-[AF]G4-C5-T6)·d(A21-G22-A23) base pair segment of the [AF]dG·dA 13/10-mer with the molecular mechanics DUPLEX computations guided by the intermolecular AF−DNA restraints listed in Table 3. Of the 16 computed structures, 3 exhibited good fit to the NMR data and had low energies (Table S6). These three structures are shown in Figure S11, and their energies and goodness-of-fit indices are listed in Table S6. The structure with the lowest value of the goodness-of-fit [$F_N = 9.8$, $F_{NN} = 5.2$, with $W = 30$ kcal/(mol·Å^2)] was embedded into an energy-minimized B-form 13/10-mer and reminimized with all restraints.

The DUPLEX-based unrestrained structure of the [AF]dG·dA 13/10-mer served as the starting structure for intensity-based relaxation matrix refinements using the XPLOR-based computational protocol (*32*). The ensemble of six intensity-refined structures exhibited an improved correspondence with experimental intensity and distance restraint data sets relative to the DUPLEX-based structures. The number of NOE distances violated by more than 0.2 Å decreased from 29 to 6 (with only 4 violations in the central 6/3-mer d(A1-A2-C3-[AF]G4-C5-T6)·d(A21-G22-A23) segment with respect to different structures in the ensemble), and the NMR *R*-factor ($R_{1/6}$) improved from an initial value of 8.2% to 2.5%.

The rmsd of all heavy atoms between the six intensity refined structures and the initial DUPLEX structure is 3.3 ± 0.5 Å for all heavy atoms and 2.6 ± 0.4 Å for the d(A1-A2-C3-[AF]G4-C5-T6)·d(A21-G22-A23) 6/3-mer segment. The quality of convergence within the intensity refined

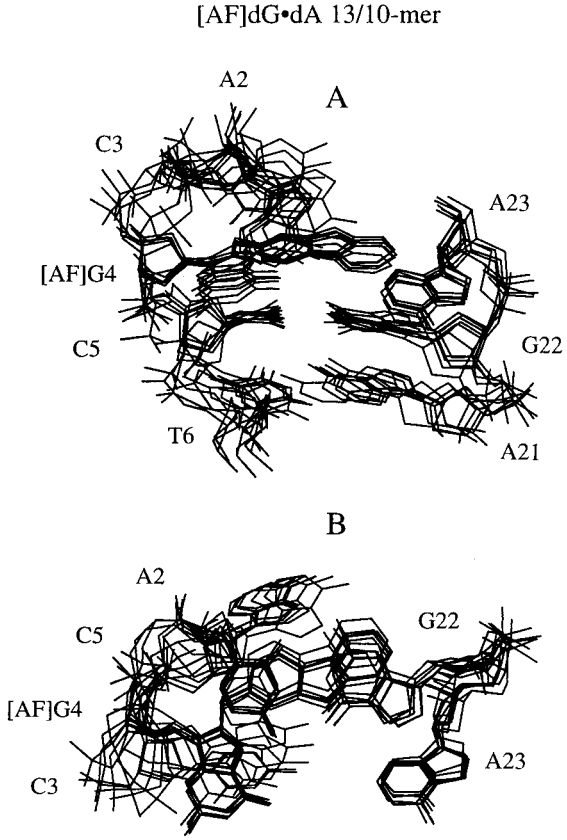

FIGURE 11: Superposition of six intensity refined structures of the [AF]dG·dA 13/10-mer. (A) View looking into the major groove and normal to the helix axis of the d(A2-C3-[AF]G4-C5-T6)·d(A21-G22-A23) segment. (B) View looking down the helix axis of the d(A2-C3-[AF]G4-C5)·d(G22-A23) segment.

structures is reflected in the pairwise rmsd of 1.86 ± 0.42 Å for all heavy atoms, 1.36 ± 0.36 Å for the heavy atoms of the d(A1-A2-C3-[AF]G4-C5-T6)·d(A21-G22-A23) 6/3-mer segment, and 1.04 ± 0.37 Å for the heavy atoms of the 6/3-mer segment excluding backbone atoms (see Table 4).

A superpositioned view of the six intensity refined structures of the d(A2-C3-[AF]G4-C5-T6)·d(A21-G22-A23) 5/3-mer segment of the [AF]dG·dA 13/10-mer is plotted in Figure 11A. The corresponding view looking down the helix axis of the central 4/2 segment is plotted in Figure 11B.

Solution Structures in [AF]dG·dA 13/10-mer. A view normal to the helix axis and looking into the major groove of the central d(A2-C3-[AF]G4-C5-T6)·d(A21-G22-A23) 5/3-mer segment of a representative structure from the ensemble of six intensity refined structures of the [AF]dG·dA 13/10-mer is shown Figure 12A. The corresponding structure for the entire adduct 13/10-mer is shown in stereo in Figure S12. The guanine ring of *syn* [AF]dG4 is displaced toward the major groove along with the dA23 positioned opposite it at the junctional site (Figure 12A).

A view looking down the helix axis of the central d(A2-C3-[AF]G4-C5)·d(G22-dA23) 4/2-mer segment of the [AF]-dG·dA 13/10-mer is shown in Figure 12B. The covalently linked aminofluorene ring stacks with the dC5·dG22 base pair in one direction and stacks with the purine ring of the unpaired dA2 residue in the other direction (Figure 12B). The long axis of the aminofluorene ring of AF is approximately parallel to the long axis of the dC5·dG22 base

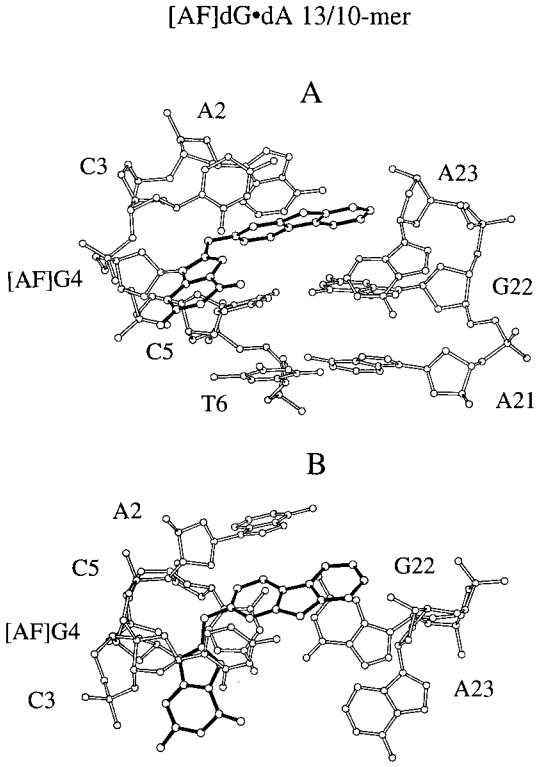

FIGURE 12: (A) View looking into the major groove and normal to the helix axis of the central d(A2-C3-[AF]G4-C5-T6)·d(A21-G22-A23) segment of a representative intensity refined structure of the [AF]dG·dA 13/10-mer. The AF ring system is shown in darkened bonds and is stacked over the dC5·dG22 base pair. The modified dG4 base and dA23 positioned opposite it are displaced toward the major groove. (B) View looking down the helix axis for the d(A2-C3-[AF]G4-C5)·d(G22-A23) segment of the [AF]dG·dA 13/10-mer. Figures were prepared using Molscript VI.1 (Kraulis, 1991).

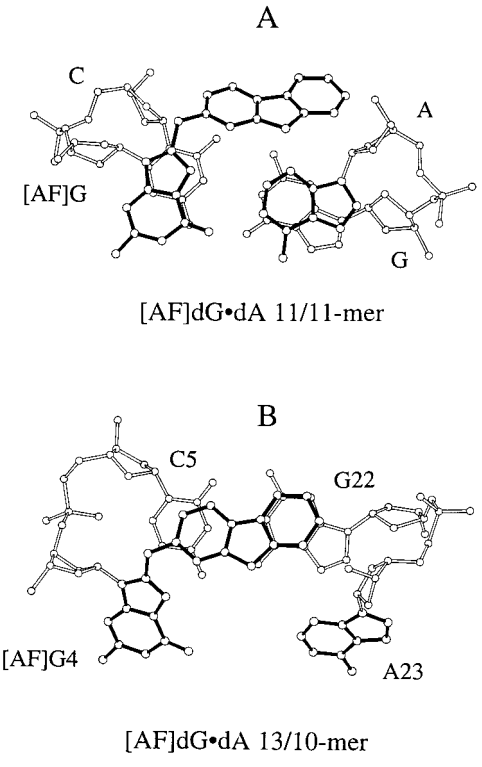

FIGURE 13: Overlap geometries between the [AF]dG opposite dA alignment and the flanking dC·dG base pair in (A) the [AF]dG·dA 11/11-mer duplex showing the d([AF]G-C)·d(G-A) segment (*19*) and (B) the [AF]dG·dA 13/10-mer showing the d([AF]G4-C5)·d(G22-A23) segment. The [AF]dG and dA positioned opposite it are shown in darkened bonds.

pair (Figures 12B). The position of the major groove displaced dA23 base is well-defined among the intensity refined structures of the [AF]dG·dA 13/10-mer (Figure 11B).

The carcinogen−base linkage site for the [AF]dG4 residue is defined by the torsion angles α' ([AF]dG4(N^9)-[AF]dG4-(C^8)-[AF](N)-[AF](C^2)) and β' ([AF]dG4(C^8)-[AF](N)-[AF]-(C^2)-[AF](C^1)) which adopt values of $\alpha' = 210 \pm 8°$ and $\beta' = 307 \pm 9°$. The pseudorotation angles and glycosidic torsion angles for the d(A1-A2-C3-[AF]G4-C5-T6)·d(A21-G22-dA23) 6/3-mer segment of the [AF]dG·dA 13/10-mer are listed in Table S7. The χ(O4′−C1′−N9−C4) glycosidic torsion angle of the [AF]dG4 residue is *syn*, with a value of $\chi = 98 \pm 11°$.

Comparison of [AF]dG·dA Alignments in 13/10-mer Junctional and 11/11-mer Duplex Structures. We have previously reported on the solution structure of [AF]dG·dA alignment within a duplex context (*19*). This alignment of the [AF]dG positioned opposite dA for the d([AF]G-C)·d(G-A) segment in the [AF]dG·dA 11/11-mer duplex is shown in Figure 13A (*19*). It can be compared with the corresponding solution structure of the [AF]dG·dA 13/10-mer (sequence in Figure 1D), whose d([AF]G4-C5)·d(G22-A23) junctional segment is shown in Figure 13B.

DISCUSSION

Spectral Quality. We observe well-resolved exchangeable (Figure 2A) and nonexchangeable (Figure S1A) proton resonances for the [AF]dG·dC 13/10-mer of a quality comparable to those observed previously for the control 13/10-mer (*34*) and the [AF]dG 13/9-mer (*21*). The imino proton of [AF]dG4 resonates within an envelope of broad peaks centered about 11.0 ppm. We do not know the origin of the other components within the 11.0 ppm broad envelope at this time. In addition, well-resolved NOE connectivities are observed involving exchangeable (Figure 3) and nonexchangeable (Figure 4) protons for the single-strand and double-strand segments in the NOESY data sets of the [AF]dG·dC 13/10-mer. The assignments of these proton resonances and the identification of the intermolecular AF−DNA NOEs provide sufficient restraints (Table 1) to define the solution structure at the junction site of the [AF]dG·dC 13/10-mer.

The one-dimensional (Figures 2B and S1B) and two-dimensional (Figures 9 and 10) spectra of the [AF]dG·dA 13/10-mer are also of unusually high quality. The imino proton of [AF]dG4 resonates as a single somewhat broadened resonance at 10.82 ppm in the [AF]dG·dA 13/10-mer (Figure 2B). This contrasts with a much broader resonance within an envelope of peaks for the imino proton of [AF]dG4 in the [AF]dG·dC 13/10-mer (Figure 2A). The exchangeable and nonexchangeable protons have been assigned in the spectra (Table S5) and have yielded sufficient intermolecular AF−DNA restraints (Table 3) to define the solution structure at the junction site of the [AF]dG·dA 13/10-mer.

[AF]dG Opposite dC at the Junction Site. The key structural feature of the [AF]dG·dC 13/10-mer solution structure is the displacement of the modified *syn* guanine of [AF]dG4 into the major groove while dC23 is looped out of

the helix (Figure 7A). Such an alignment results in the absence of pairing between junction site residues [AF]dG4 and dC23 which are positioned opposite each other in the sequence (Figure 7B).

The assignment of the *syn* glycosidic torsion angle at the [AF]dG4 lesion site has been based on the downfield shifts observed for the H2′ proton (3.56 ppm) (Figure 5A) and the C1′ carbon (86.7 ppm) (Figure S4A) resonances in the spectra of the [AF]dG 13/10-mer. The downfield shift of the H2′ proton of [AF]dG4 probably reflects in-plane ring current contributions from the modified guanine of the adduct in a *syn* alignment. Similar large downfield H2′ proton shifts have been reported for the benzo[*a*]pyrene−N^2-guanine adduct positioned at a duplex−single-strand junction (*39*) and specific guanines in G-quadruplexes (*40*). The *syn* alignment at guanines in these systems was independently established from the strong NOE between the base H8 and its own H1′ proton in short mixing time NOESY spectra (*25*).

The downfield shift of the C1′ carbon (86.7 ppm) of the modified guanine residue can be compared with the corresponding chemical shifts of adjacent dG22 (84.2 ppm) and dG19 (85.1 ppm) guanine residues in the [AF]dG·dC 13/10-mer (Figure S4A). It has been previously established that sugar C1′ carbon chemical shifts of the DNA residues adopting *syn* glycosidic torsion angles can shift downfield by up to 5 ppm provided the DNA residues retain C2′-*endo* sugar pucker geometries (*41, 42*). We have observed a coupling cross-peak between the H1′ (6.05 ppm) and H2′ (3.56 ppm) protons of [AF]dG4 (Figure S2A) consistent with placement of this sugar within the C2′-*endo* range.

We do not detect an NOE connectivity between the upfield-shifted imino proton of [AF]dG4 and other amino and imino protons from both one-dimensional NOE difference and two-dimensional NOESY data sets. This suggests that the imino proton of [AF]dG4 is exposed to solvent and not hydrogen-bonded to dC23 in the solution structure of the [AF]dG 13/10-mer. The observed NOE cross-peaks between the upfield-shifted dG22 imino proton and the amino and H5 protons of dC5 (Figure 3B) establish that the dC5·dG22 base pair maintains a Waston−Crick pairing alignment. The imino proton resonance of dG22 broadens and exchanges more rapidly with solvent following AF modification at the C^8 of the adjacent dG4 residue, indicating that the dC5·dG22 base pair is destabilized on going from the control 13/10-mer (*34*) to the [AF]dG·dC 13/10-mer.

Comparison of [AF]dG with and without Opposing dC at the Junction Site. There are similarities between the NMR spectra of the [AF]dG 13/9-mer (*21*) and the [AF]dG·dC 13/10-mer, and in their solution structures (Figure 8A,B). Both adducts show essentially the same NOE patterns in the NOESY spectra, and the chemical shifts of the aromatic protons of aminofluorene are also very similar (Figure S3). In both structures, the *syn*-modified guanine is displaced into the major groove and is replaced with the aminofluorene. The AF ring is stacked on one face with the dC5·dG22 base pair and on the other face with the purine ring of dA2 in the solution structures of the [AF]dG 13/9-mer (Figure 8A) (*21*) and the [AF]dG·dC 13/10-mer (Figure 8B). The chemical shifts of the imino protons of dG22 and the amino protons of dC5 are dramatically upfield-shifted in both cases when compared to their corresponding values in the control 13/10-mer (*34*) or 13/9-mer (*39*). This observation is in agreement with the stacking interaction between the aminofluorene ring and the dC5·dG22 base pair in the solution structure of the [AF]dG 13/9-mer (*21*) and [AF]dG·dC 13/10-mer. The purine ring of the dA2 residue is stacked with the AF ring in both cases.

[AF]dG Opposite dA at the Junction Site. The key structural feature of the [AF]dG·dA 13/10-mer solution structure is the displacement of the modified *syn* guanine of [AF]dG4 and *anti* dA23 positioned opposite each other into the major groove (Figure 12A). The Hoogsteen edge of the modified guanine of [AF]dG4 and the Watson−Crick edge of dA23 are approximately coplanar and directed toward each other but remain unpaired (Figure 12B).

The observed *syn* glycosidic torsion angle for [AF]dG4 is supported by the observation of a large downfield shift of the H2′ proton (3.60 ppm) (Figures 5B and S2B) and the downfield shift of the C1′ carbon (86.7 ppm) (Figure S10A) of the modified guanine residue for [AF]dG·dA 13/10-mer.

Similar Alignments of [AF]dG with No Partner and Opposite dC and dA at Junctional Sites. The key common structural features of the [AF]dG·dC 13/10-mer, the [AF]dG·dA 13/10-mer, and the [AF]dG 13/9-mer (*21*) solution structures are (1) the aminofluorene ring of [AF]dG4 stacks between the nonadjacent dA2 base and the dC5·dG22 base pair located at the junction site, (2) the modified guanine adopts a *syn* alignment, and (3) the duplex segment excluding the [AF]dG adduct remains in an unperturbed B-DNA conformation with all Waston−Crick alignments retained, including the dC5·dG22 base pair.

The observed stacking of the fluorene ring with the dC5·dG22 base pair results in displacement of the modified *syn* guanine toward the major groove in the solution structures of the [AF]dG·dC 13/10-mer (Figure 7), the [AF]dG·dA 13/10-mer (Figure 12), and the [AF]dG 13/9-mer (*21*). The linkage torsion angles α' ([AF]dG4(N^9)-[AF]dG4(C^8)-[AF]-(N)-[AF](C^2)) and β' ([AF]dG4(C^8)-[AF](N)-[AF](C^2)-[AF]-(C^1)) adopt similar values of $\alpha' = 229 \pm 10°$ and $\beta' = 10 \pm 4°$ for the [AF]dG·dC 13/10-mer, $\alpha' = 210 \pm 8°$ and $\beta' = 307 \pm 9°$ for the [AF]dG·dA 13/10-mer, and $\alpha' = 231 \pm 28°$ and $\beta' = 11 \pm 3°$ for the [AF]dG 13/9-mer (*21*). The similar stacking of the fluorene ring between the dA2 base and the dC5·dG22 base pair in the [AF]dG·dC 13/10-mer (Figure 7), the [AF]dG·dA 13/10-mer (Figure 12), and the [AF]dG 13/9-mer (*21*) is reflected in the similar chemical shifts of the fluorenyl protons for these three systems (Figure S9).

Different Alignments of dC and dA Positioned Opposite [AF]dG at the Junctional Site. There is an important distinction between the structures of the [AF]dG·dC 13/10-mer (Figure 7) and the [AF]dG·dA 13/10-mer (Figure 12) as related to the alignment of the base positioned opposite the [AF]dG4 adduct junctional site. The dC23 residue opposite [AF]dG4 did not show any detectable intermolecular NOEs and was looped out into the minor groove in the structures of the [AF]dG·dC 13/10-mer (Figure 7). By contrast, the dA23 residue opposite [AF]dG4 showed several intermolecular NOEs (Table 3 and Figure S8), which positions it in the major groove in an in-plane alignment with the modified guanine positioned opposite it in the structure of the [AF]dG·dA 13/10-mer (Figure 12). This orientation of the dA23 positions one face of its purine ring over the van der Waals surface of the major groove with its

amino group directed toward the dT6·dA21 base pair. The Hoogsteen edge of [AF]dG4 and the Watson−Crick edge of dA23 are directed toward each other but are separated by twice the normal hydrogen-bonding distances in the intensity refined structures of the [AF]dG·dA 13/10-mer (Figure 12B). The observed differences between dC23, which is looped out into the minor groove in the [AF]dG·dC 13/10-mer (Figure 7), and dA23, which adopts a defined orientation in the major groove in the [AF]dG·dA 13/10-mer (Figure 12), may reflect the relative abilities of the narrow minor and wide major grooves to accommodate pyrimidine versus purine residues.

The imino protons of both dG22 and [AF]dG4 are narrower in the imino proton spectrum of the [AF]dG·dA 13/10-mer (Figure 2B) compared to the [AF]dG·dC 13/10-mer (Figure 2A). This may reflect the greater participation of dA23 with the junctional site based on its positioning in the major groove (Figure 12) relative to dC23 which is looped out into the minor groove (Figure 7).

Comparison between Structures of [AF]dG·dA 13/10-mer and [AF]dG·dA 11/11-mer. We have previously characterized the solution structure of the [AF]dG adduct positioned opposite dA in the d(C-[AF]G-C)·d(G-A-G) sequence context at the 11/11-mer duplex level (*19*). The alignment of the AF ring in this complex is such that both its faces are buried within the walls of the minor groove with only its edges exposed, tilted slightly in the 3′-direction of the modified strand, and with its distal ring directed toward the sugar−phosphate backbone of the partner strand (*19*). In addition, the *syn* [AF]dG and its mismatched *anti* dA are stacked into the helix. There are no hydrogen bonds between the [AF]dG and the opposite dA at neutral pH, but at acidic pH the NMR data indicate a slight shift in the position of the AF ring such that a single hydrogen bond could form between the O^6 of the [AF]dG and the N^1 of the protonated dA ring. The neutral pH structure is stabilized by hydrophobic interactions between the faces of the AF moiety and the walls of the minor groove and by the maintenance of a stacked-in position of the [AF]dG and the mismatched dA, while the acidic pH structure adds the opportunity for one hydrogen bond to these stabilizing forces (*19*).

The relative orientations of the [AF]dG opposite dA pair and the flanking dC·dG pair looking down the helix axis are shown for the fully paired [AF]dG·dA 11/11-mer duplex in Figure 13A (*19*) and for the junctional [AF]dG·dA 13/10-mer in Figure 13B. Thus, the modified *syn* guanine and its opposing dA stack with the flanking dC·dG pair while the AF ring is positioned in the minor groove in the fully paired [AF]dG·dA 11/11-mer duplex (Figure 13A) (*19*). By contrast, the AF ring is stacked with the flanking dC·dG pair while the modified *syn* guanine and its opposing dA are positioned in the major groove in the junctional [AF]dG·dA 13/10-mer (Figure 13B). It therefore appears that [AF]dG opposite dA alignment is similar in both systems except that this alignment is shifted toward the minor groove in the fully paired [AF]dG·dA 11/11-mer duplex (Figure 13A) (*19*) while it is shifted toward the major groove in the junctional [AF]-dG·dA 13/10-mer (Figure 13B) (this study). In addition, the separation between the Hoogsteen edge of [AF]dG and the Watson−Crick edge of its opposing dA is smaller for the fully paired [AF]dG·dA 11/11-mer duplex (Figure 13A) (*19*) relative to the junctional [AF]dG·dA 13/10-mer (Figure 13B).

Biological Implications. DNA polymerases can, in many instances, readily synthesize past the [AF]dG adduct, and translesion synthesis can be highly accurate (reviewed in *14*). On the other hand, the [AF]dG adduct can cause mutations: mismatches (primarily G·A, and others), frame shifts, as well as polymerase blockage in certain systems. Both the base sequence and the nature of the polymerase are key determinants governing the biological response to this lesion (*14, 16, 17*). NMR-based studies have established that the [AF]-dG adduct is capable of adopting both *syn* and *anti* glycosidic torsion angles depending on base sequence context (*35, 36, 43*−*46*), and that this is likely to play a key role governing the biological response to the lesion. Presumably when accurate translesion synthesis takes place, the [AF]dG is in a normal *anti* conformation, as it is in part of the conformational mix in certain sequence contexts (*35, 36, 43, 44*). However, in the solution structures of [AF]dG at model primer-template junctions, the [AF]dG adopts the *syn* conformation, both in the [AF]dG 13/9-mer lacking a partner opposite the adduct site (*21*) and in the present [AF]dG·dC 13/10-mer where dC is positioned opposite the adduct site. The similarity between these two structures (Figure 8) is striking, except for the looped out position of the additional dC opposite [AF]dG in the [AF]dG·dC 13/10-mer. The position of the looped out dC may suggest how an incoming dCTP might fail to position itself properly for faithful extension opposite an [AF]dG containing a *syn* dG residue. A structure of this type could also possibly be responsible for polymerase blockage with the opportunity for slippage mispairing that can introduce frameshift mutations in a suitable sequence context (*47*−*51*).

The tendency for G to T transversions at sites of [AF]dG adduct formation may reflect the stabilizing interactions associated with the [AF]dG alignment opposite dA at a junctional site as determined in the present structural study. This stabilization could reflect contributions from the stacking of the AF ring of [AF]dG4 with the flanking dC5·dG22 base pair in the [AF]dG·dA 13/10-mer and the van der Waals contacts between the faces of the modified guanine and the opposing dA23 with the floor of the major groove. The preference for adenine may reflect both the larger ring size associated with a purine and the hydrogen-bonding potential of its exocyclic amino group which is directed toward the duplex segment within the major groove.

We can also draw correlations for [AF]dG aligned opposite dA at junctional sites and within fully paired duplexes. We note that the α' and β' linkage torsion angles are similar for the *syn* [AF]dG residue in the fully paired [AF]dG·dA 11/11-mer duplex ($\alpha' = 208°$ and $\beta' = 317°$) (*19*) and the junctional [AF]dG·dA 13/10-mer ($\alpha' = 210°$ and $\beta' = 307°$). Further, the Hoogsteen edge of [AF]dG is directed toward the Watson−Crick edge of the opposing dA in both the fully paired [AF]dG·dA 11/11-mer duplex (Figure 13A) (*19*) and the junctional [AF]dG·dA 13/10-mer (Figure 13B). This suggests that the [AF]dG opposite dA junctional alignment (Figure 13B) could be readily incorporated within a duplex (Figure 13A) by a translation of this entity toward the minor groove. Thus, there would be a switch from a stabilizing stacking interaction involving the AF ring and flanking bases/base pairs associated with the junctional alignment (Figure 13B) to a stabilizing stacking interaction involving the modified guanine and opposing adenine and flanking base

pairs associated with the fully paired duplex (Figure 13A). Similarly, there would be a switch from a stabilizing van der Waals interaction involving the modified guanine and opposing adenine aligned face down within the walls of the major groove in the junctional alignment to a stabilizing van der Waals interaction involving the aminofluorene ring sandwiched between the walls of the minor groove in the fully paired duplex (*19*). Such an analysis reasons that dA can be misincorporated because the [AF]dG can adopt the abnormal *syn* conformation, and when this happens a dA can be stably accommodated opposite the lesion at both the template-primer junction site and the full duplex level (*19*).

Coordinates Deposition. The coordinates of the [AF]dG·dC 13/10-mer (file name 1c0z) and [AF]dG·dA 13/10-mer (file name 1c0y) have been deposited in the Brookhaven Data Base.

SUPPORTING INFORMATION AVAILABLE

Eight tables listing the complete exchangeable and nonexchangeable proton chemical shifts, the energy and goodness-of-fit parameters for the three low-energy DUPLEX-based structures, the range of P and χ values for the central 7/4-mer segment among the intensity refined structures, and comparison of experimental and computed carcinogen−DNA distance restraints for the [AF]dG·dC 13/10-mer and [AF]-dG·dA 13/10-mer; and twelve figures showing one- and two-dimensional NMR spectra, intermolecular NOEs, plots of [AF]dG proton chemical shifts, ^{13}C and ^{31}P two-dimensional data, the superposition of the 5/3-mer segment of the three DUPLEX-based structures, and a representative full-length intensity refined structure of the [AF]dG·dC 13/10-mer and the [AF]dG·dA 13/10-mer (28 pages). This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Ashby, J., and Tennant, R. W. (1991) *Mutat. Res. 257*, 229−306.
2. Weissberger, J. H. (1988) in *Carcinogenic and Mutagenic Responses to Aromatic Amines and Nitroarenes* (King, C. M., Romano, L. J., and Schueltze, D., Eds.) pp 3−19, Elsevier, New York.
3. Beland, F. A., and Kadlubar, F. F. (1990) in *Hand Book of Experimental Pharmacology* (Cooper, C. S., and Grover, P. L., Eds.) pp 267−325, Springer-Verlag, Heidelberg.
4. Sugimura, T. (1992) *Science 258*, 603−607.
5. Vineis, P. (1994) *Environ. Health Perspect. 102*, 7−10.
6. Layton, D. W., Bogen, K. T., Knize, M. G., Hatch, F. T., Johnson, V. M., and Felton, J. S. (1995) *Carcinogenesis 16*, 39−52.
7. Wilson, R. H., DeEds, F., and Cox, A. J. J. (1941) *Cancer Res. 1*, 595−608.
8. Bishop, J. M. (1995) *Genes Dev. 9*, 1309−1315.
9. Weinberg, R. A. (1996) *Sci. Am. 275*, 62−70.
10. Bos, J. L. (1989) *Cancer Res. 49*, 4682−4689.
11. Conti, C. J. (1992) *Comp. Mol. Carcinog. 376*, 357−378.
12. Harris, C. C. (1993) *Science 262*, 1980−1981.
13. Husgafvel-Pursiainen, K., and Kannio, A. (1996) *Environ. Health Perspect. 104*, 553−556.
14. Heflich, R. H., and Neft, R. E. (1994) *Mutat. Res. 318*, 73−174.
15. Carothers, A. M., Urlaub, G., Mucha, J., Yuan, W., Chasin, L. A., and Grunberger, D. (1993) *Carcinogenesis 14*, 2181−2184.
16. Belguise-Valladier, P., and Fuchs, R. P. P. (1995) *J. Mol. Biol. 249*, 903−913.
17. Doisy, R., and Tang, M. S. (1995) *Biochemistry 34*, 4358−4368.
18. Hoffman, G. R., and Fuchs, R. P. (1997) *Chem. Res. Toxicol. 10*, 347−359.
19. Norman, D., Abuaf, P., Hingerty, B. E., Live, D., Grunberger, D., Broyde, S., and Patel, D. J. (1989) *Biochemistry 28*, 7462−7476.
20. Abuaf, P., Hingerty, B. E., Broyde, S., and Grunberger, D. (1995) *Chem. Res. Toxicol. 8*, 369−378.
21. Mao, B., Gu, Z., Gorin, A., Hingerty, B. E., Broyde, S., and Patel, D. J. (1997) *Biochemistry 36*, 14491−14501.
22. Marion, D., Ikura, M., Tschudin, R., and Bax, A. (1989) *J. Magn. Reson. 85*, 393−399.
23. Sklenar, V., Miyashiro, H., Zon, G., Miles, H. T., and Bax, A. (1986) *FEBS Lett. 208*, 94−98.
24. Bax, A., and Subramanian, J. (1986) *J. Magn. Reson. 67*, 565−570.
25. Patel, D. J., Kozlowski, S. A., Nordheim, A., and Rich, A. (1982) *Proc. Natl. Acad. Sci. U.S.A. 79*, 1413−1417.
26. van de Ven, F. J., and Hilbers, C. W. (1988) *Eur. J. Biochem. 178*, 1−38.
27. Hingerty, B. E., Figueroa, S., Hayden, T., and Broyde, S. (1989) *Biopolymers 28*, 1195−1222.
28. Taylor, E. R., and Olson, W. K. (1983) *Biopolymers 22*, 2667−2702.
29. Broyde, S., and Hingerty, B. E. (1983) *Biopolymers 22*, 2423−2441.
30. Hingerty, B. E., and Broyde, S. (1986) *J. Biomol. Struct. Dyn. 4*, 365−371.
31. 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.
32. Brunger, A. T. (1992) *X-Plor: A system for X-ray crystallography and NMR*, Yale University Press, New Haven and London.
33. Patel, D. J., Shapiro, L., and Hare, D. (1987) *Annu. Rev. Biophys. Biophys. Chem. 16*, 423−454.
34. Feng, B., Gorin, A., Hingerty, B. E., Geacintov, N. E., Broyde, S., and Patel, D. J. (1997) *Biochemistry 36*, 13769−13779.
35. Mao, B., Hingerty, B. E., Broyde, S., and Patel, D. J. (1998) *Biochemistry 37*, 81−94.
36. Mao, B., Hingerty, B. E., Broyde, S., and Patel, D. J. (1998) *Biochemistry 37*, 95−106.
37. Arnott, S., Bond, P. J., Selsing, E., and Smith, P. J. (1976) *Nucleic Acids Res. 2*, 2459−2470.
38. Berman, H. M., Olson, W. K., Beveridge, D. L., Westbrook, J., Gelbin, A., Demeny, T., Hsieh, S. H., Srinivasan, A. R., and Schneider, B. (1992) *Biophys. J. 63*, 751−759.
39. Cosman, M., Hingerty, B. E., Geacintov, N. E., Broyde, S., and Patel, D. (1995) *Biochemistry 34*, 15334−15350.
40. Wang, Y., and Patel, D. J. (1993) *Structure 1*, 263−282.
41. Ghose, R., Marino, J. P., Wiberg, K. B., and Prestegard, J. H. (1994) *J. Am. Chem. Soc. 116*, 8827−8828.
42. Greene, K. L., Wang, Y., and Live, D. (1995) *J. Biomol. NMR 5*, 333−338.
43. Eckel, L. M., and Krugh, T. R. (1994) *Nat. Struct. Biol. 1*, 89−94.
44. Cho, B. P., Beland, F. A., and Marques, M. M. (1994) *Biochemistry 33*, 1373−1384.
45. Mao, B., Cosman, M., Hingerty, B. E., Broyde, S., and Patel, D. J. (1995) *Biochemistry 34*, 6226−6238.
46. Mao, B., Hingerty, B. E., Broyde, S., and Patel, D. J. (1995) *Biochemistry 34*, 16641−16653.
47. Kunkel, T. A. (1990) *Biochemistry 29*, 8003−8011.
48. Kunkel, T. A. (1992) *J. Biol. Chem. 267*, 18251−18254.
49. Schaaper, R. M., Koffel-Schwartz, N., and Fuchs, R. P. P. (1990) *Carcinogenesis 11*, 1087−1095.
50. Shibutani, S., and Grollman, A. P. (1993) *J. Biol. Chem. 268*, 11703−11710.
51. Lambert, I. B., Napolitano, R. L., and Fuchs, R. P. P. (1992) *Proc. Natl. Acad. Sci. U.S.A. 89*, 1310−1314.
52. Kraulis, P. J. (1991) *J. Appl. Crystallogr. 24*, 946−950.

BI991266P