Solution Structure of the Aminofluorene-Intercalated Conformer of the *syn* [AF]-C[8]-dG Adduct Opposite a −2 Deletion Site in the *Nar*I Hot Spot Sequence Context[†,‡]

Bing Mao,[§] Andrey Gorin,[§] Zhengtian Gu,[§] 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 September 5, 1997[⊗]

ABSTRACT: This paper addresses structural issues related to the capacity of aminofluorene [AF] for frameshift mutations of the −2 type on C[8] covalent adduct formation at the G[3] site in the d(C-G[1]-G[2]-C-G[3]-C-C) *Nar*I hot spot sequence. This problem has been approached from a combined NMR and relaxation matrix analysis computational structural study of the [AF]dG adduct in the d(C-G-G-C-[AF]G-C-C)·d(G-G-C-C-G) sequence context at the 12/10-mer adduct level (designated [AF]dG·del(−2) 12/10-mer). The proton spectra of this system are of exceptional quality and are consistent with the formation of an AF-intercalated conformer with the modified guanine in a *syn* alignment displaced along with the 5′-flanking cytosine residue into the major groove. The solution structure has been determined by initially incorporating intramolecular and intermolecular proton−proton distances defined by lower and upper bound deduced from NOESY spectra as restraints in molecular mechanics computations in torsion angle space and subsequently refined through restrainted molecular dynamics calculations based on a NOE distance and intensity refinement protocol. Strikingly, the [AF]dG·del(−2) 12/10-mer duplex adopts only one of two potential AF-intercalation alignments for the [AF]dG adduct opposite the −2 deletion site in the *Nar*I sequence context with the extrusion of the dC-[AF]dG step favored completely over extrusion of the [AF]dG-dC step at the lesion site. This polarity establishes that the structural perturbation extends 5′ rather than 3′ to the [AF]dG lesion site in the adduct duplex. This structure of the [AF]dG adduct opposite a −2 deletion site shows distinct differences with conclusions reported on the alignment of the related acetylaminofluorene [AAF]dG adduct opposite a −2 deletion site in the identical *Nar*I sequence context [Milhe, C., Fuchs, R. P. P., and Lefevre, J. F. (1996) *Eur. J. Biochem. 235*, 120−127]. In that study, qualitative NMR data without computational analysis were employed to conclude that the extrusion at the lesion site occurs at the [AAF]dG-dC step for the AAF-intercalated conformer of the adduct duplex. The structure of the [AF]dG adduct opposite a −2 deletion site determined in our group provides molecular insights into the architecture of extended slipped mutagenic intermediates involving aromatic amine intercalation and base-displaced *syn* modified guanines in AF and, by analogy, AAF-induced mutagenesis in the *Nar*I hot spot sequence context.

The aromatic amine carcinogens 2-acetylaminofluorene [AAF] and 2-aminofluorene [AF] form covalent adducts following metabolic or chemical activation at the C[8] position of guanine on DNA (reviewed in ref 1). The more bulky [AAF]dG adduct differs only in the presence of an acetyl group linked to the N[2] fluorenyl moiety in the place of the hydrogen in the [AF]dG adduct **1** (Figure 1A). This substitution is responsible for the observed mutagenic differences produced by these two adducts, which in turn has provoked the decades long interest in the structural basis for their distinct mutagenic spectrum of activities.

The application of NMR methods has provided considerable insights into the structure of [AF]dG adducts as a function of base or absence of it on the partner strand opposite the lesion site and as a function of sequence context flanking the lesion site. Initial studies on [AF]dG adducts positioned opposite dA (*2*) and dG (*3*) established that the modified guanine in a *syn* alignment stacked into the helix with the aminofluorene positioned within the walls of the minor groove and directed toward the partner strand. Structural studies by Eckel and Krugh (*4*, *5*) established a slow interconversion between AF-intercalated and AF-external conformers for the [AF]dG adduct positioned opposite dC at the duplex level. These authors established that the [AF]dG adduct adopts an *anti* alignment for the AF-external conformer and also favored an *anti* alignment for the AF-intercalated conformer in the equilibrium. This conformational equilibrium between AF-intercalated and AF-external conformers in slow exchange has also been observed in other sequence contexts (*6*−*8*). However, a fundamental

[†] This research is supported by NIH Grant CA-49982 to D.J.P., by NIH Grants CA-28038 and CA-75449, NIH Grant RR-06458, and DOE Grant DE-FG02-90ER60931 to S.B., and by DOE Contract DE-AC05-960R22464 with Lockheed Martin Energy Research and DOE OHER Field Work Proposal ERKP931 to B.E.H.

[‡] The coordinates of the [AF]dG·del(−2) 12/10-mer have being deposited in the Brookhaven Protein Data Base (file name 1AX6).

[§] MemorialSloan-Kettering Cancer Center.

[∥] Oak Ridge National Laboratory.

[⊥] New York University.

[⊗] Abstract published in *Advance ACS Abstracts,* November 15, 1997.

S0006-2960(97)02205-8 CCC: $14.00

disagreement between the contributions from our laboratory (*7*, *8*) and that of Eckel and Krugh (*4*, *5*) relates to our demonstration that the slow conformational equilibrium involves interconversion between a *syn* [AF]dG alignment in the AF-intercalated conformer (*7*) and an *anti* [AF]dG alignment in the AF-external conformer (*8*). In addition, AF-intercalated conformers have also been observed for [AF]-dG adducts in a *syn* alignment when positioned opposite −1 deletion (*9*) and −2 deletion (*10*) sites within DNA duplexes.

The *Escherchia coli Nar*I hotspot sequence for frameshift mutations by the carcinogens AAF and AF has been a source of great interest and curiosity since it was first reported in the literature (*11*). This sequence, d(C-G^1-G^2-C-G^3-C-C), has the remarkable property of inducing −2 frameshift mutations leading to the sequence d(C-G^1-G^2-C-C) only when G^3 is modified by aromatic amines in the *Nar*I sequence context (*12*). Moreover, AAF has a much greater propensity than AF to induce frameshift mutations at this site (*12*−*16*), which highlights the issue of the relationships between the potential structural differences of these two carcinogen-bound DNAs and their functional distinctions.

This paper addresses structural issues related to the intriguing observation of the low but very interesting capacity of AF for frameshift mutations of the −2 type on covalent adduct formation at the G^3 site in the *Nar*I sequence context. Our combined NMR−computational structural studies have focused on a 12/10-mer duplex (alternate alignments **2** and **3** shown in Figure 1B,C) where the [AF]dG adduct has been located opposite a −2 deletion site in this sequence context. The research highlights formation of only one of two potential AF-intercalation alignments of [AF]dG opposite the −2 deletion site in the *Nar*I sequence context with structure determination greatly facilitated by the excellent quality of the NMR spectra.

MATERIALS AND METHODS

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

Preparation of Adduct. The preparation, separation, purification, and characterization of the d(C1-T2-C3-G4-G5-C6-[AF]G7-C8-C9-A10-T11-C12) adducted sequence was undetaken as follows: The d(C-T-C-G-G-C-G-C-C-A-T-C) sequence was reacted with excess AAAF in 2 mM sodium citrate buffer as reported previously (*17*). The potential modification sites are located at guanines G4, G5, and G7 in this 12-mer sequence. The three [AF]dG 12-mer adducts were separated by reversed-phase HPLC on a C18 ODS-Hypersil semipreparative column (Keystone Scientific Inc.). The solvent program was an isochratic step at 15% methanol in 20 mM phosphate buffer with a flow rate of 3 mL/min, pH 7.0, for 10 min, followed by a linear gradient to 40% methanol in 60 min. The monoadducted 12-mers that were subsequently characterized came off the column in the sequential order with modifications at [AAF]dG4 followed by [AAF]dG7 followed by [AAF]dG5. These monoadducts were followed by the elution of 12-mers containing bis-[AAF]dG adducts, which in turn were followed by 12-mers containing tris-[AAF]dG adducts at longer retention times. The three purifed monoadducted [AAF]dG 12-mers were converted into their corresponding monoadducted [AF]dG 12-mers by dissolving in 1 M NaOH containing 0.3% (v/v) 2-mercaptoethanol for 45 min at room temperature following purification by reverse-phase HPLC as outlined above. The pure monoadducted [AF]dG 12-mers were desalted on Sephadex G-25 and converted to their sodium forms on a Dowex 50 × 8 cation exchange resin column. Each of the three pure monoadducted [AF]dG 12-mers was annealed with the complementary d(G-A-T-G-G-C-G-C-C-G-A-G) 12-mer strand. These duplexes were then characterized by NMR to identify the position of the lesion site since the [AF]dG adduct lacks an H8 proton and results in a break in the sequential connectivities.

The [AF]dG·del (−2) 12/10-mer was prepared by annealing the [AF]dG monoadducted at the dG7 position in the 12-mer strand with the complementary d(G-A-T-G-G-C-C-G-A-G) 10-mer strand at 70 °C, and the stoichiometry was followed by monitoring single proton resonances in both strands.

NMR Experiments. A combination of through-space nuclear Overhauser effect (NOESY) and through-bond correlated (COSY, TOCSY) two-dimensional spectra were recorded on approximately 6 mg of the [AF]dG·del(−2) 12/10-mer duplex in 0.6 mL of aqueous buffer (0.1 M NaCl, 10 mM phosphate, pH 7.0). These two dimensional data sets were analyzed to assign the AF and nucleic acid protons in the [AF]dG·del(−2) 12/10-mer. All NMR experiments on the adduct duplex were carried out on Varian Unity plus 500 or 600 MHz spectrometers using the States-TPPI mode (*18*) with a 2.0-s relaxation delay between scans. The temperature of the sample was calibrated with an external methanol sample. The NOESY spectrum (150 ms mixing time) of the adduct duplex in H_2O buffer at 1 °C was collected using a jump-return pulse for solvent suppression. The corresponding NOESY spectra (50, 90, and 130 ms mixing times) and TOCSY (40 and 80 ms spin lock times) data sets on the adduct duplex were collected in D_2O buffer at 25 °C.

Heteronuclear $^1H-^{13}C$ and $^1H-^{31}P$ spectra on the adduct duplex were also collected as reported in the Methods section of the previous papers from our laboratory (*9*, *19*). Briefly, the indirect proton−phosphorus correlation spectrum was recorded on the [AF]dG·del(−2) 12/10-mer duplex in D_2O at 25 °C using the pulse sequence described previously in the literature (*20*). The phosphorus spectra were referenced relative to external 10% trimethylphosphate (TMP). The $^1H-^{13}C$ HMQC correlation spectra on the [AF]dG·del(−2) 12/10-mer duplex in D_2O buffer were also recorded at 25 °C. The carbon spectra were referenced relative to external 3-(trimethylsiyl)propionate (TSP) using the method described previously in the literature (*21*).

The calculation of interproton distance bounds using volume buildup of NOE cross-peaks for first stage DUPLEX computations was based on the two spin approximation using dT(NH3)-dA(H2) fixed 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 (*22*). The proton−proton vicinal coupling constants among sugar protons were analyzed from phase-sensitive COSY data to qualitatively distinquish between the C3′-endo and C2′-endo family of sugar puckers (*23*).

Molecular Mechanics Computations. Molecular mechanics calculations in torsion angle space were carried out with the DUPLEX program (*24*) as described in the Methods section of previously published paperes (*9*, *19*). 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.

Relaxation Matrix Refinement. The unrestrained energy minimized structure obtained following first stage DUPLEX calculations was used as the starting point for second stage XPLOR (*25*) based molecular dynamics/simulated annealing calculations guided by the combination of the experimental NOESY intensity and NOE distance restraints. 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, anisotropic bounds estimates of 10%; (2) distance restraints for nonexchangeable protons were retained during our protocol as square-well potentials with uniform 20% estimation of errors and force constant of 30 kcal/mol·Å^2 . A 4.5-Å cutoff was imposed for computing relaxation pathways, and the dynamics was 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 501 nonexchangeable intensity restraints from the NOESY data sets at three mixing times in D_2O (167 intensities per mixing time) and 167 nonexchangeable distance restraints were included in the calculations. Diheral angle restraints (corresponding to B-DNA) were included with a very low weight of 5 kcal/rad^2 and restricted to residues that are two pairs away in either direction from the [AF]dG lesion site.

Six intensity refinement trials were performed with the DUPLEX-based starting structure heated in each trial 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

Potential Pairing Alignments for [AF]dG positioned opposite a −2 Deletion Site in the NarI Hot Spot Sequence. There exist two potential pairing alignments for the stem segments flanking the [AF]dG7 adduct **1** (Figure 1A) positioned opposite a −2 deletion site in the [AF]dG·del(−2) 12/10-mer duplex. The dG17 base can potentially pair with dC8 in one alignment resulting in a two-base bulge consisting of the dC6-[AF]dG7 segment as shown schematically in **2** (Figure 1B). Alternately, the dG17 base can potentially pair with dC6 in a second alignment resulting in a two-base bulge consisting of the [AF]dG7-dC8 segment as shown schematically in **3** (Figure 1C). The potential exists that either one of these alignments or an interconverting

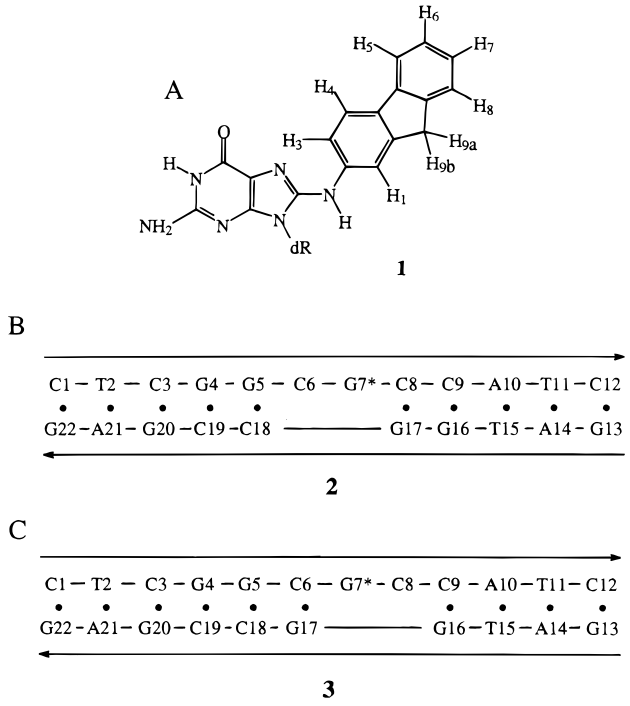

FIGURE 1: (A) Schematic of [AF]dG adduct **1**. (B) Schematic of the [AF]dG·del(−2) 12/10-mer **2** with a bulge at the dC6-[AF]-dG7 step. (C) Schematic of the [AF]dG·del(−2) 12/10-mer **3** with a bulge at the [AF]dG7-dC8 step.

mixture of **2** and **3** represent the solution structure of the [AF]dG·del(−2) 12/10-mer. The answer to this question can be approached by NMR methods that can also potentially define the solution structure of the adduct duplex.

Exchangeable Nucleic Acid Protons. The imino proton spectrum (10.0−14.0 ppm) of the [AF]dG·del(−2) 12/10-mer in H_2O buffer at 1 °C is plotted in Figure 2A. We observe 11 well-resolved imino protons that must originate in the [AF]dG adduct and the 10 base pairs in the adduct duplex. This establishes that the [AF]dG·del(−2) 12/10-mer adopts a single conformer in solution since we do not detect doubling of imino proton resonances in Figure 2A. Furthermore, the imino protons are both narrow and well resolved, which in turn should help in the analysis of NMR data on this adduct duplex.

The imino protons have been assigned following analysis of NOESY (150 mixing time) spectra of the [AF]dG·del(−2) 12/10-mer in H_2O buffer at 1 °C. We can trace the NOE connectivities between imino protons on adjacent base pairs starting from dG22 at one end and proceeding to dG5 in the middle of the helix and from dG17 in the middle to dG13 at the other end of the helix in an expanded NOESY contour plot of the symmetrical imino proton region of the adduct duplex as shown in Figure 3C. These data permit a differentiation between alignment **2** shown in Figure 1B in which the two-base bulge separates the imino protons of dG5 and dG17 from alignment **3** shown in Figure 1C in which the two-base bulge separates the imino protons of dG16 and dG17 in the adduct duplex. We do not observe an NOE between the imino protons of dG5 and dG17 (boxed region, Figure 3C) but do observe the NOE between the imino protons of dG16 and dG17 (Figure 3C), establishing that the [AF]dG·del(−2) 12/10-mer must align as shown in **2** (Figure 1B) with the two-base bulge formed by the dC6-[AF]dG7 segment.

The expanded NOESY contour plot correlating NOEs between the imino protons (10.5 and 14.5 ppm) and the base

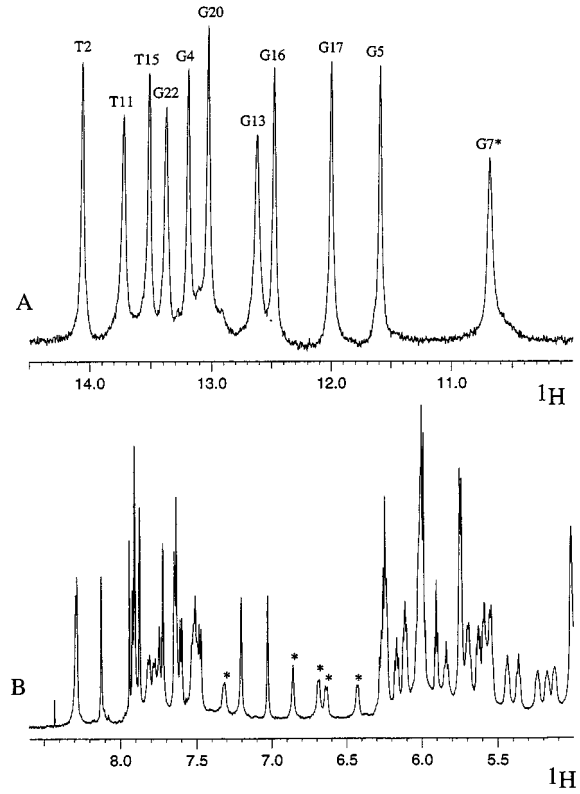

FIGURE 2: (A) Imino proton spectra (10.0−14.5 ppm) of the [AF]-dG·del(−2) 12/10-mer in H_2O buffer at 1 °C and (B) nonexchangeable proton spectra (5.0−8.6 ppm) of the [AF]dG·del(−2) 12/10-mer duplex in D_2O buffer at 25 °C. The buffer was 0.1 M NaCl, 10 mM phosphate, and 0.1 mM EDTA, at pH 7.0. The imino proton assignments are shown over the resonances in the spectrum in panel A. The AF carcinogen protons are designated by asterisks over resonances in the spectrum in panel B.

and amino protons (5.0−9.0 ppm) in the adduct duplex is plotted in Figure 3B. We detect NOEs between the guanine imino protons to their partner cytosine amino and H5 protons across the Watson−Crick dG·dC base pairs including dG5·dC18 (peaks B, B′, and D, Figure 3B) and dC8·dG17 (peaks C, C′, and E, Figure 3B) pairs that flank the dC6-[AF]dG7 bulge segment in adduct duplex **2** shown in Figure 1B.

Such an analysis yields the imino and amino protons in the adduct duplex that are listed for the d(G5-C6-[AF]G7-C8)·d(G17-C18) segment in Table 1 and for the entire adduct duplex in Table S1 (Supporting information). It should be noted that the imino proton of [AF]dG7 (10.70 ppm) at the lesion site together with the imino protons of dG5 (11.60 ppm) and dG17 (12.01 ppm) resonate to the high field of the remaining imino protons. This pattern is indicative of the base displacement−intercalation model (*26*, 27) for [AF]-dG positioned opposite a −2 deletion site for adduct duplex **2** (Figure 1B).

We also detect intermolecular NOEs between the protons on the fluorenyl ring and the imino protons of dG5 (peaks 1−3, Figure 3B) and dG17 (peaks 4 and 5, Figure 3A,B) indicative of intercalation of the fluorenyl ring of [AF]dG7 between the dG5·dC18 and dC8·dG17 Watson−Crick base pairs in adduct duplex **2** shown in Figure 1B.

Nonexchangeable Nucleic Acid Protons. The nonexchangeable base and sugar H1′ proton spectrum (5.0−8.5 ppm) of the [AF]dG·del(−2) 12/10-mer duplex in D_2O buffer at 25 °C is plotted in Figure 2B. We observe sharp resonances for nucleic acid base and sugar protons, which

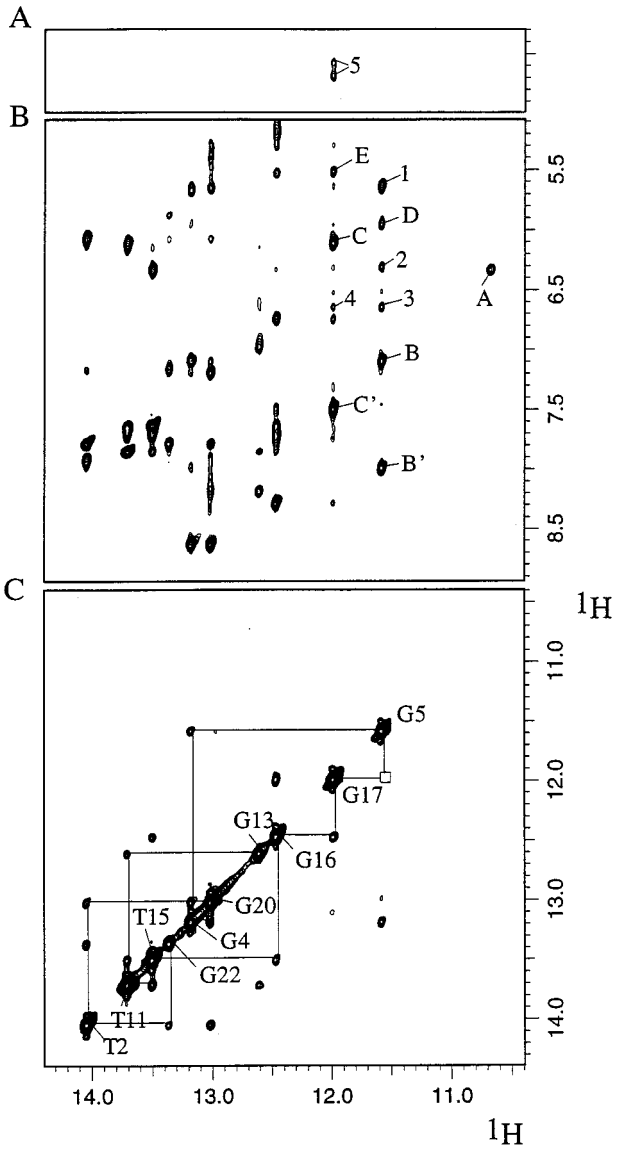

FIGURE 3: Expanded NOESY (150 ms mixing time) contour plots of the [AF]dG·del(−2) 12/10-mer in H_2O buffer at 1 °C. (A) NOE connectivities between the imino protons (10.4−14.4 ppm) and the AF methylene protons (2.8−3.5 ppm). The carcinogen−DNA cross-peaks labeled 5, are assigned as follows: 5: G17(NH1)-AF(H9a) and G17(NH1)-AF(H9b). (B) NOE connectivities between the imino protons (10.4−14.4 ppm) and the base and amino protons regions (5.1−8.9 ppm). The NOE cross-peaks involving the imino protons are labeled in the figure as follows: A, [AF]G7(NH1)-[AF]G7(NH_2-2); B′,B, G5(NH1)-C18(NH_2-4b,e); C′,C, G17(NH1)-C8(NH_2-4b,e); D, G5(NH1)-C18(H5); E, G17(NH1)-C8(H5). The intermolecular carcinogen−DNA NOE cross-peaks 1−4 are assigned as follows: 1, G5(NH1)-AF(H6,H7); 2, G5(NH1)-AF(H5); 3, G5(NH1)-AF(H8); 4, G17(NH1)-AF(H8). (C) NOE connectivities in the symmetrical (10.4−14.4 ppm) region. The imino 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 dG13 toward the other end of the helix. The connectivity between the imino protons of dG5 and that of dG17 is missing and marked by a box.

have been assigned following analysis of NOESY, COSY, and TOCSY two-dimensional data sets on the adduct duplex. An expanded NOESY (300 ms mixing time) data set outlining NOE connectivities between the base protons (6.8−8.4 ppm) and the sugar H1′ protons (5.1−6.3 ppm) is shown in Figure 4. The NOE connectivities between base protons to their own and 5′-flanking sugar H1′ protons are traced from dG4 to dA10 on the modified strand (solid lines, Figure 4) and from dT15 to dG19 on the unmodified strand (dashed

Table 1. Proton Chemical Shifts of the d(G5-C6-[AF]G7-C8)·d(G17-C18) Segment of the [AF]dG·del(−2) 12/10-mer in Aqueous Buffer

	Exchangeable Proton Chemical Shifts, ppm, 1 °C		
	G(NH1)	G(NH_2-2)	C(NH_2-4)
dG5·dC18	11.60		8.03,[a] 7.13[b]
dC6·del(−2)			
[AF]dG7·del(−2)	10.70	6.33	
dC8·dG17	12.01		7.54,[a] 6.13[b]

	Nonexchangeable Proton Chemical Shifts (ppm) 25 °C						
	H8/H6	H5	H1′	H2′,H2″	H3′	H4′	^{31}P
dG5	7.90		5.17	2.62,2.26	4.43	4.18	−3.65
dC6	7.63	6.23	6.01	2.04,2.43	4.70	4.29	−3.59
[AF]dG7			6.10	3.95,2.27	4.87	4.52	−2.79
dC8	7.82	5.55	6.01	2.53,2.53	4.88	4.36	−4.01
dG17	7.51		5.83	2.28,2.28	4.97	4.41	−4.26
dC18	7.78	6.03	6.00	2.30,2.30	4.91	4.21	−3.84

[a] Hydrogen-bonded amino proton. [b] Exposed amino proton. [c] ^{31}P chemical shift corresponds to residue (*n*) for the (*n*)-^{31}P-(*n* + 1) step.

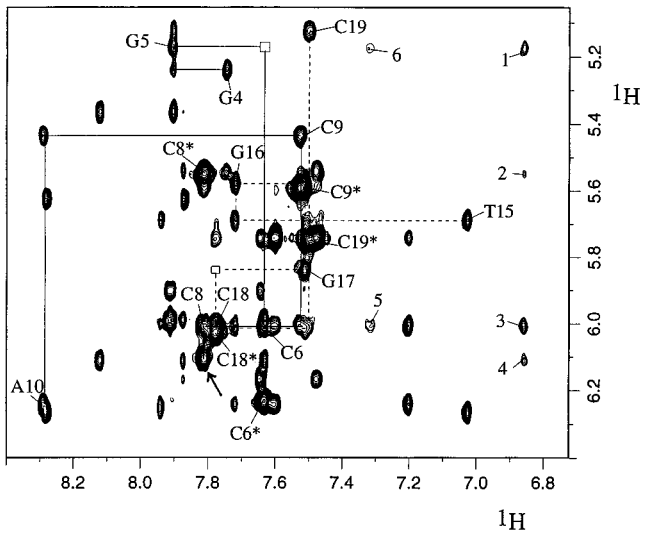

FIGURE 4: Expanded NOESY (300 ms mixing time) contour plot of the [AF]dG·del(−2) 12/10-mer in D_2O buffer at 25 °C establishing distance connectivities between the base (purine H8 and pyrimidine H6) protons (6.7−8.6 ppm) and the sugar H1′ and cytosine H5 protons (5.05−6.4 ppm). The NOE connectivities between the base and their own and the 5′-flanking sugar H1′ protons from dG4 to dA10 on the modified strand are shown by solid lines and from dT15 to dC19 on the unmodified partner strand are shown by dashed lines. The assignments label the base to their own sugar H1′ NOEs, while the cytosine H6-H5 NOEs are designated by asterisks. Note that the NOE cross-peak at the dC6-[AF]dG7 is missing because of the absence of an H8 proton for [AF]dG7. The intermolecular carcinogen−DNA NOE cross-peaks are assigned as follows: 1, AF(H1)-G5(H1′); 2, AF(H1)-C8(H5); 3, AF(H1)-C8(H1′); 4, AF(H1)-[AF]G7(H1′); 5, AF(H3)-C8(H1′); 6, AF(H3)-G5(H1′).

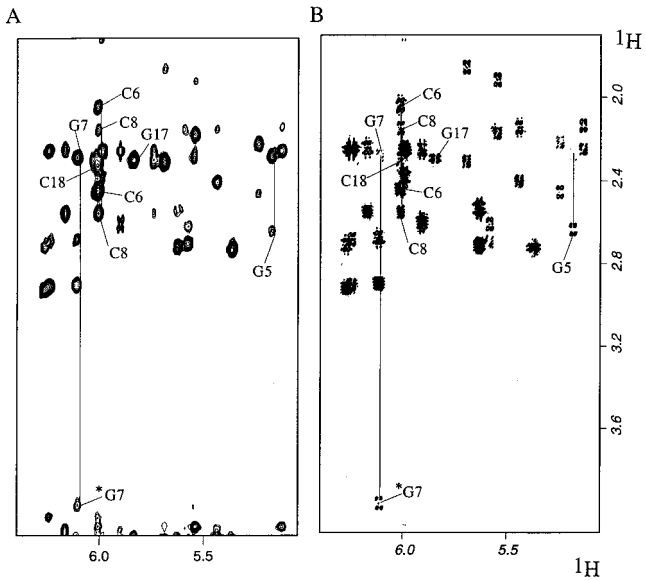

FIGURE 5: (A) Expanded NOESY (90 ms mixing time) contour plot of the [AF]dG·del(−2) 12/10-mer in D_2O buffer at 25 °C showing NOEs between the sugar H1′ protons (5.0−6.4 ppm) and H2′, H2″ protons (1.7−3.9 ppm). (B) Expanded phase sensitive COSY contour plot of the [AF]dG·del(−2) 12-mer duplex in D_2O buffer at 25 °C establishing coupling connectivities between the H1′ protons (5.0−6.6 ppm) and H2′, H2″ protons (1.7−4.1 ppm). In both panels A and B, the H2′ and H2″ protons of dG5, dC6, [AF]dG7, dC8, dG17 and dC18 are connected by lines and labeled. The H2′ protons resonate upfield of the H2" protons for the majority of these residues except for [AF]dG7 and dG5, where the H2" protons resonate upfield of the H2′ proton. Note the large downfield shift of the H2′ proton of [AF]dG7 at 3.95 ppm.

lines, Figure 4) in the adduct duplex. The [AF]dG7 adduct lacks an H8 proton due to covalent modification at this position, and hence there is a break in connectivity at this position in the dC6-[AF]dG7 segment of the modified strand (Figure 4). In addition, there are also breaks in NOE connectivities at the dG5-dC6 step on the modified strand and at the dG17-dC18 step on the unmodified strand (see boxed regions in Figure 4) in the expanded NOESY contour plot of the adduct duplex. The structural perturbations are restricted to the segment centered about the adduct site in the [AF]dG·del(−2) 12/10-mer duplex **2** shown in Figure 1B. It should be noted that the H5 (6.23 ppm) proton of dC6 (peak labeled C6* in Figure 4) resonates to the low field of the other cytosines in the adduct duplex. Such an analysis yields the base and sugar H1′ protons in the adduct duplex, which together with the remaining sugar proton assignments are listed for the d(G5-C6-[AF]G7-C8)·d(G17-C18) segment in Table 1 and for the entire adduct duplex in Table S2 (Supporting Information).

Expanded NOESY (90 ms mixing time) and COSY contour plots identifying connectivities between the sugar H1′ protons (5.0−6.5 ppm) and H2′,H2″ protons in the [AF]-dG·del(−2) 12/10-mer are plotted in Figure 5, panels A and B, respectively. We note that the sugar H2′ proton of [AF]-dG7 has shifted dramatically downfield to 3.95 ppm in the adduct duplex. This marker has been shown previously to be characteristic of a *syn* guanine orientation (*2, 7, 9, 10, 19, 28*) and is consistent with intercalation of the fluorenyl ring into the helix accompanied by displacement of the modified guanine into one of the grooves of the helix.

Aminofluorene Protons. The fluorenyl ring protons of the [AF]dG7 adduct resonate between 5.6 and 7.4 ppm in the nonexchangeable proton spectrum of the [AF]dG·del(−2)

Table 2: Comparison of Experimental Intermolecular Distance Restraints with Those Observed for Intensity-Refined NMR Structures of the [AF]dG·del(−2) 12/10-mer

	interproton distances (Å)	
	exptl bounds	observed
Exchangeable Protons (Carcinogen−DNA)		
AF(NH)-[AF]G7(H1′)	2.2−4.0	2.2 ± 0.1
AF(NH)-[AF]G7(H2′)	2.8−5.0	4.8 ± 0.1
AF(NH)-[AF]G7(H2″)	2.6−4.5	4.2 ± 0.2
AF(NH)-C8(H6)	2.9−5.0	3.5 ± 0.3
AF(NH)-C8(H5)	2.8−5.0	3.6 ± 0.3
AF(H5)-G5(NH1)	3.1−5.0	3.3 ± 0.2
AF(H6,H7)-G5(NH1)	2.8−5.0	3.9 ± 0.2
AF(H6,H7)-C18(NH4b)	3.0−6.0	5.0 ± 0.2
AF(H8)-G5(NH1)	3.2−5.0	5.1 ± 0.2
AF(H8)-G17(NH1)	3.2−5.0	5.0 ± 0.3
AF(H9a,b)-G17(NH1)	3.3−6.0	4.0 ± 0.3
Nonexchangeable Protons (Carcinogen−DNA)		
AF(H1)-G5(H1′)	3.0−4.6	4.3 ± 0.1
AF(H1)-G5(H2′)	3.0−6.0	4.4 ± 0.1
AF(H1)-[AF]G7(H1′)	3.2−4.8	4.3 ± 0.1
AF(H1)-C8(H1′)	2.9−4.3	3.8 ± 0.1
AF(H3)-G5(H8)	3.3−5.0	4.2 ± 0.1
AF(H3)-G5(H2′)	3.0−4.4	3.5 ± 0.1
AF(H3)-G5(H2″)	2.7−4.0	3.9 ± 0.1
AF(H3)-G5(H1′)	3.9−5.9	5.2 ± 0.1
Other Selected Nonexchangeable Proton Pairs		
G5(H1′)-C6(H6)	>5.0	6.0 ± 0.3
G5(H1′)-C6(H5)	>5.0	6.5 ± 0.5
G5(H1′)-C6(H1′)	> .0	7.7 ± 0.5
[AF]G7(H1′)-C8(H6)	2.3−3.3	2.8 ± 0.1

12/10-mer in D_2O buffer at 25 °C as shown in Figure 2B. The resolved fluorenyl ring protons (designated by asterisks in Figure 2B) are only slightly broader than their nucleic acid counterparts in the spectrum of the adduct duplex. Individual fluorenyl ring protons have been assigned from analysis of through space and through bond connectivities in nonexchangeable proton spectra of the adduct duplex. We have also been able to assign the NH proton (8.61 ppm) involved in the covalent linkage between the carcinogen and the modified guanine of [AF]dG7 in the adduct duplex. The fluorenyl ring proton chemical shifts in the [AF]dG·del(−2) 12/10-mer **2** shown in Figure 1B in the *Nar*I sequence context are compared with related data from a previous study of the [AF]dG adduct positioned opposite a −2 deletion site in another sequence context (*10*) in Figure S1 (Supporting Information).

Intermolecular AF-DNA NOEs and Restraints. We observe a large number of intermolecular NOEs involving both exchangeable (see numbered NOE peaks in Figure 3B) and nonexchangeable (see numbered peaks in Figure 4) protons for the [AF]dG·del(−2) 12/10-mer. A total of 20 intermolecular NOEs have been identified and assigned in the adduct duplex and are primarily between the AF protons and the protons of the dG5·dC18 and dC8·dG17 base pairs. For example, the imino proton of dG5 exhibits intermolecular NOEs to the H5, H6, H7, and H8 fluorene ring protons while the imino proton of G18 exhibits intermolecular NOEs to the H8 and H9a,b fluorene ring protons in the [AF]dG·del(−2) 12/10-mer. The corresponding experimental intermolecular distances (defined by lower and upper bounds) estimated (based on a two-spin approximation) from the volume integrals of the NOE cross-peaks in the adduct duplex are listed in Table 2.

Carbon Spectra. The expanded contour plot of a natural abundance proton−carbon HMQC correlation experiment that correlates the H1′ and C1′ chemical shifts of individual residues of the [AF]dG·del(−2) 12/10-mer in D_2O buffer at 25 °C is plotted in Figure 6A. The carbon resonances are assigned on the basis of the known H1′ proton assignments in the adduct duplex. The C1′ chemical shift assignments for residues in the d(G4-G5-C6-[AF]G7-C8-C9)·d(G16-G17-C18-C19) segment are labeled in Figure 6A. We note that the C1′ chemical shift of [AF]dG7 (86.92 ppm) is downfield of other assignable dG residues, which are observed to resonate between 84.2 and 85.2 ppm in the adduct duplex (Figure 6A). It has been previously observed that downfield sugar C1′ carbon chemical shifts (that can range up to 5 ppm) are observed for DNA residues, which adopt *syn* glycosidic torsion angles provided that they retain C2′-*endo* sugar pucker geometries (*29*, *30*). We detect a coupling cross-peak between the H1′ proton (6.10 ppm) and the H2′ proton (3.95 ppm) of [AF]dG7 (Figure 5B) placing this sugar within the C2′-*endo* range.

Phosphorus Spectra. The proton decoupled phosphorus spectrum of the [AF]dG·del(−2) 12/10-mer duplex has been recorded in D_2O buffer at 25 °C. The phosphorus resonances are dispersed over a 1.6 ppm range with a few resonances shifted to the low field of the −3.8 to −4.4 ppm range. The phosphorus resonances have been assigned from an analysis of the proton-detected phosphorus−proton heteronuclear correlation experiment on the adduct duplex with the expanded contour plot shown in Figure 6B. The phosphorus resonances for the individual steps in the d(G4-G5-C6-[AF]-G7-C8-C9)·d(G16-G17-C18-C19) segment of the adduct duplex are labeled in Figure 6B with downfield phosphorus chemical shifts observed for the dC6-[AF]dG7 (−3.59 ppm) and [AF]dG7-dC8 (−2.79 ppm) steps on the modified strand and the dC19-dG20 (−3.30 ppm) step on the unmodified strand in the adduct duplex.

Molecular Mechanics Computations. The search strategy employed began with the central d(G4-G5-C6-[AF]G7-C8-C9)·d(G16-G17-C18-C19) 6/4-mer segment of the [AF]-dG·del(−2) 12/10-mer with the computations guided by the intermolecular AF-DNA restraints listed in Table 2. The DNA starting conformation was the B-form (*31*) except for a *syn* ($\chi = 60°$) glycosidic torsion angle for the [AF]dG7 residue, which was not restricted to this value during the subsequent molecular mechanics calculations. The AF-DNA orientation space was searched with 16 energy minimization trials in which the linkage torsion angles α' ([AF]dG7(N^9)-[AF]dG7(C^8)-[AF](N)-[AF](C^2)) and β' ([AF]dG7(C^8)-[AF]-(N)-[AF](C^2)-[AF](C^1)) were each started at 0°, 90°, 180°, and 270° in all combinations. These 16 trials represent arbitrary unbiased high energy starting orientations of the carcinogen that equally survey the potential energy surface of the adduct. In these trials, the DUPLEX hydrogen-bond penalty function (*24*) for Watson−Crick base pairing was utilized at all paired bases, since the NMR data (Figure 3b) indicated that these hydrogen bonds were present.

Five of the 16 computed structures generated by DUPLEX exhibited good fit to the NMR data and had low energies. These five structures for the central 6/4-mer segment viewed normal to the helix axis are superpositioned in Figure S2A (Supporting Information), and for the central 4/2-mer segment viewed down the helix axis are superpositioned in Figure S2B (Supporting Information). The structure with the lowest energy was embedded into an energy-minimized B-form 12/10-mer and reminimized with all restraints. Subsequently, the hydrogen-bond penalty function and the

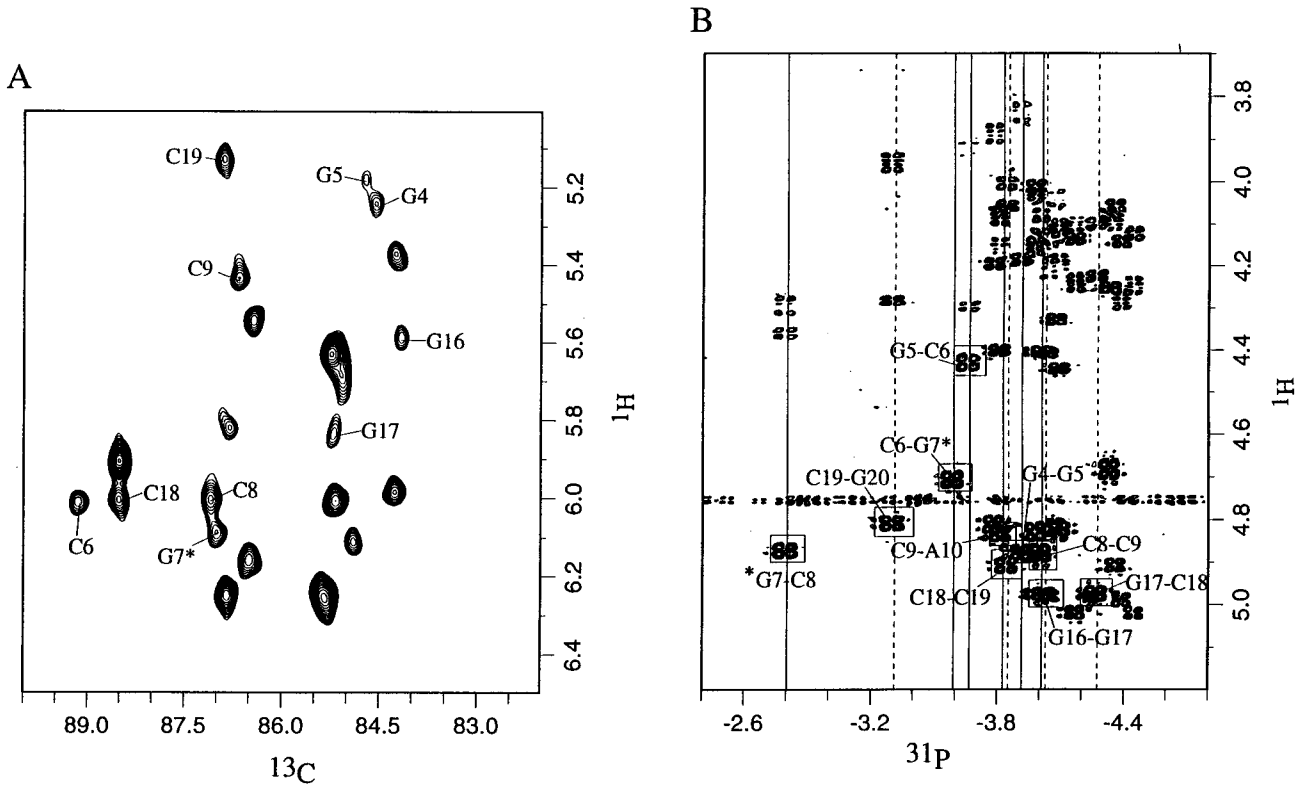

FIGURE 6: (A) Expanded contour plot of $^{1}H-^{13}C$ heteronuclear multiple-quantum coherence (HMQC) experiment on the [AF]dG·del(−2) 12/10-mer in D_2O buffer at 25 °C. The C1′ assignments are marked for residues in the d(G4-G5-C6-[AF]G7-C8-C9)·d(G16-G17-C18-C19) segment. (B) Expanded contour plot of the proton-detected phosphorus-proton heteronuclear correlation experiment on the [AF]dG·del(−2) 12/10-mer in D_2O buffer at 25 °C. The phosphorus assignments are listed for steps centered about the lesion site. The correlation cross-peaks between the phosphorus and its 5′-flanking sugar H3′ protons are boxed.

Table 3. NMR refinement statics for the [AF]dG·del(−2) 12/10-mer

NMR distance restraints	76
total no. of distance restraints	167
central segment distance restraints[a]	76
NMR intensity restraints	
total intensity restraints	167 per mixing time
central segment intensity restraints[a]	76 per mixing time
structure statistics	
NMR *R*-factor ($R_{1/6}$)	0.026 ± 0.006
rmsd of NOE violations	0.079 ± 0.022
no. of NOE violations >0.2 Å in entire adduct duplex	7.2 ± 2.7
no. of NOE violations >0.2 Å in central 6/4-mer region[a]	2.2 ± 0.1
deviations from the ideal geometry	
bond length (Å)	0.010 ± 0.0005
bond angle (deg)	2.64 ± 0.10
impropers (deg)	0.44 ± 0.30
pairwise rmsd (Å) among the six refined structures (heavy atoms only)	
entire 12/10-mer adduct duplex	2.61 ± 0.50
central 6/4-mer region[a]	1.66 ± 0.29
central 6/4-mer region[a] without backbone	1.25 ± 0.28

[a] The d(G4-G5-C6-[AF]G7-C8-C9)·d(G16-G17-C18-C19) segment.

distance restraints were released with energy minimization in one step, yielding the unrestrained DUPLEX-based structure of the [AF]dG·del(−2) 12/10-mer.

Relaxation Matrix Refinement. The DUPLEX-based unrestrained structure of the [AF]dG·del(−2) 12/10-mer served as the starting structure for intensity-based relaxation matrix refinements using the XPLOR-based computational protocol outlined in the Methods section. An ensemble of six structures obtained following intensity refinement exhibited an improved correspondence with experimental intensity and distance restraint data sets. The number of distance bounds violated by more than 0.2 Å decreased from 22 after first stage DUPLEX refinement to 7 after second stage intensity refinement [with only two to three violations in the central d(G4-G5-C6-[AF]G7-C8-C9)·d(G16-G17-C18-C19) 6/4-mer segment with respect to different structures in the ensamble]. Furthermore, the NMR *R*-factor ($R_{1/6}$) improved from an initial value of 8.0% to 2.6% on proceeding from the DUPLEX-based to intensity refinement-based structures (Table 3). The intensity-refined structures exhibit good stereochemistry with rmsd values for bond length , bond angle, and improper dihedral angle violations of 0.010 ± 0.001 Å, 2.64° ± 0.10°, and 0.44 ± 0.30°, respectively (Table 3).

The rmsd of all heavy atoms between the six intensity-refined structures and the initial DUPLEX structure is 2.3 ± 0.4 Å for all heavy atoms and 1.7 ± 0.4 Å for the d(G4-G5-C6-[AF]G7-C8-C9)·d(G16-G17-C18-C19) 6/4-mer seg-

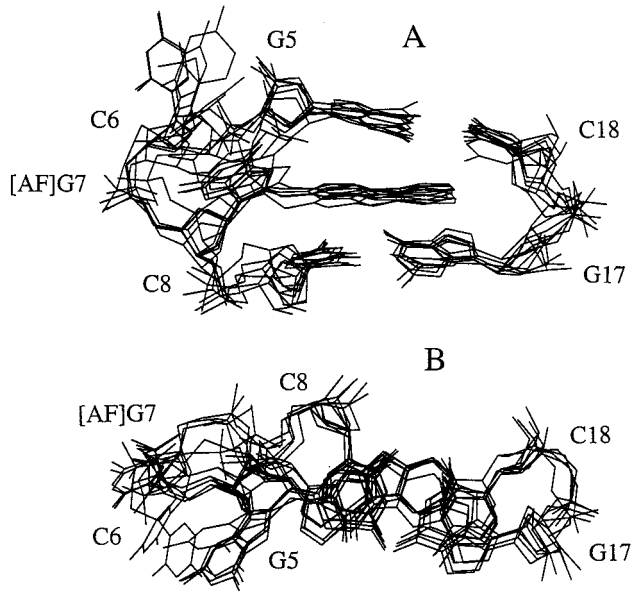

FIGURE 7: (A) Superposition of the d(G4-G5-C6-[AF]G7-C8-C9)·d(G16-G17-C18-C19) segments of the six intensity-refined structures of the [AF]dG·del(−2) 12/10-mer. View looking into the major groove and normal to the helix axis. (B) Superposition of the d(G5-C6-[AF]G7-C8)·d(G17-C18) segments of the six intensity-refined structures of the [AF]dG·del(−2) 12/10-mer duplex. View looking down the helix axis.

ment. The intensity-refined structures exhibit reasonable convergence with pairwise rmsd values in the set being 2.61 ± 0.50 Å for all heavy atoms, 1.66 ± 0.29 Å for the heavy atoms of the d(G4-G5-C6-[AF]G7-C8-C9)·d(G16-G17-C18-C19) 6/4-mer segment, and 1.25 ± 0.28Å for the heavy atoms of the 6/4-mer excluding backbone atoms (see Table 3)

The superpositioned view of six intensity-refined structures of the d(G4-G5-C6-[AF]G7-C8-C9)·d(G16-G17-C18-C19) segment of the [AF]dG·del(−2) 12/10-mer is plotted in Figure 7A. A view looking down the helix axis of the central 4/2 segment is plotted in Figure 7B.

Solution Structures. A view normal to the helix axis and looking into the major groove of the central d(G4-G5-C6-[AF]G7-C8-C9)·d(G16-G17-C18-C19) 6/4-mer segment of the representative structure from the ensemble of six intensity-refined structures of the [AF]dG·del(−2) 12/10-mer is shown in Figure 8A. A stereoview of the structure of the entire adduct duplex is shown in Figure 9. The covalently linked fluorene ring intercalates between Watson−Crick dG5·dC18 and dC8·dG17 base pairs by displacing the modified guanine ring of the *syn* [AF]dG7 residue and the adjacent dC6 residue into the major groove (Figure 8A). The guanine base plane of the *syn* [AF]dG7 adduct stacks over the minor groove sugar face of dC6 and is tilted somewhat toward the 5′-end of the modified strand in the structure of the adduct duplex (Figure 8A).

A view looking down the helix axis of the central d(G5-C6-[AF]G7-C8)·d(G17-C18) 4/2-mer segment of the structure of the [AF]dG·del(−2) 12/10-mer is shown in Figure 8B. This view emphasizes the overlap geometry between the fluorene ring system and the flanking Watson−Crick dG5·dC18 and dC8·dG17 base pairs with their long axis approximately parallel to each other (Figure 8B). The carcinogen-base linkage site for the [AF]dG7 residue is defined by the torsion angles α′ ([AF]dG7(N^9)-[AF]dG7-(C^8)-[AF](N)-[AF](C^2)) = 209.7 ± 7.5 and β′ ([AF]dG7-

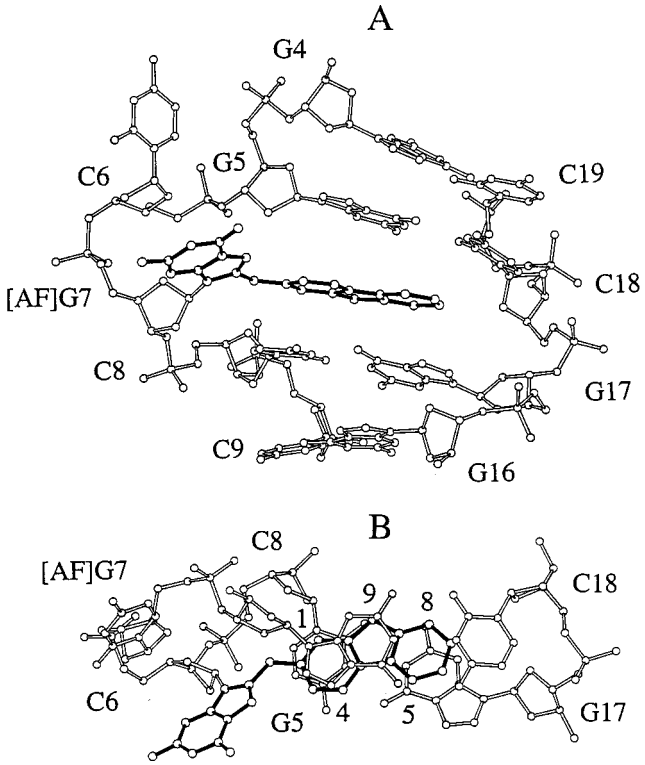

FIGURE 8: Representative structure of the [AF]dG·del(−2) 12/10-mer from the ensemble of six intensity-refined structures. (A) View looking into the major groove and normal to the helix axis of the central d(G4-G5-C6-[AF]G7-C8-C9)·d(G16-G17-C18-C19) segment. The AF ring system is shown in darkened bonds and is intercalated between the dG5·dC18 and dC8·dG17 base pairs. The modified dG7 base is displaced into the major groove. The unpaired dC6 residue is also displaced into the major groove. (B) View looking down the helix axis for the d(G5-C6-[AF]G7-C8)·d(G17-C18) segment. Figure were prepared using Molscript VI.1 (*51*).

(C^8)-[AF](N)-[AF](C^2)-[AF](C^1)) = 187.6 ± 14.4°. The glycosidic torsion angles and sugar pseudorotation **P** angles for the d(G4-G5-C6-[AF]G7-C8-C9)·d(G16-G17-C18-C19) 6/4-mer segment of the six intensity-refined structures of the [AF]dG·del(−2) 12/10-mer duplex are listed in Table S3 (Supporting Information). The χ(O4′-C1′-N9-C4) glycosidic torsion angle of the *syn* [AF]dG7 residue spans the range χ = 68 ± 16 °. The ensamble of six structures obtained from the intensity-based refinement was B-like, except that the sugar pucker pseudorotation **P** value for dG17 is 14 ± 20°, which places this sugar in the C3′-*endo* domain. The sugar pucker pseudorotation **P** value for dC6 is not well defined among the intensity-refined structures.

DISCUSSION

Spectral Quality. The exchangeable (Figure 2A) and nonexchangeable (Figure 2B) proton spectra of the [AF]-dG·del(−2) 12/10-mer are of exceptional quality with both narrow and resolved resonances observed for the nucleic acid and aromatic amine protons. The quality of these spectra have permitted a detailed structural analysis that has yielded the structure of the adduct duplex in solution.

Conformation at the Lesion Site. The two possible alignments of the [AF]dG·del(−2) 12/10-mer are represented schematically by **2** and **3** in Figure 1. The experimental data are consistent with formation of only one of these pairing alignments with the adduct duplex adopting exclusively alignment **2** shown in Figure 1B. The dC6 and [AF]dG7 residues lack pairing partners on the complementary strand

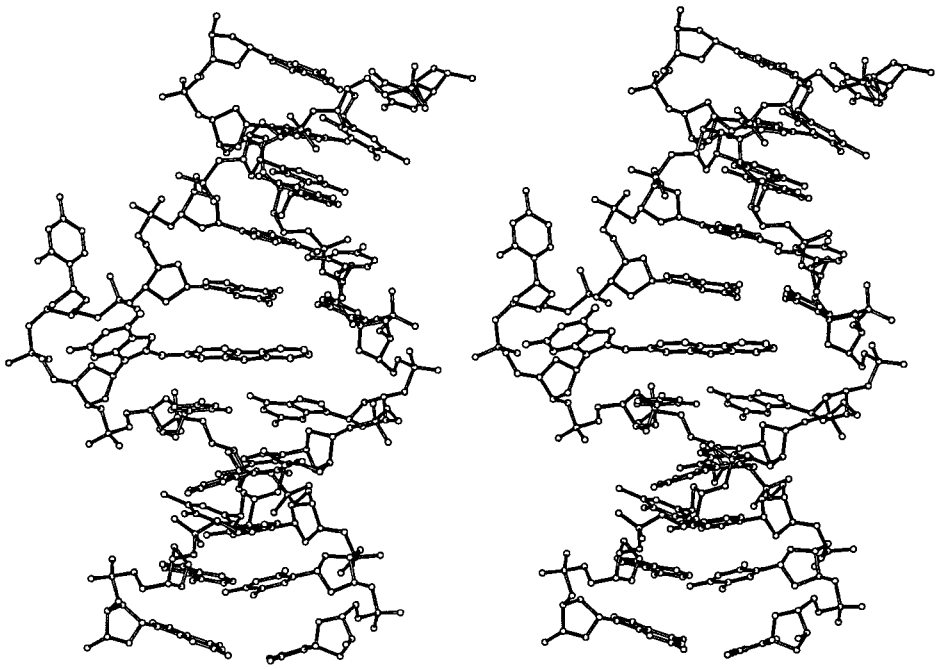

FIGURE 9: Stereo view normal to the helix axis of the representative intensity-refined structure of the entire [AF]dG·del(−2) 12/10-mer. The figure was prepared using Molscript V1.1 (*51*).

in alignment **2** shown in Figure 1B of the [AF]dG·del(−2) 12/10-mer duplex with the structure characterized by intercalation of the fluorenyl ring into the helix (Figure 8A). We discuss below some of the structural features centered about the lesion site in this AF-intercalated conformer of the adduct duplex.

Conformation of the Modified Guanine. The glycosidic torsion angle of the [AF]dG7 residue has been assigned a *syn* conformation in the AF-intercalated structure of the [AF]-dG·del(−2) 12/10-mer. This conclusion was reached primarily on the basis of the downfield shift of the H2′ proton of [AF]dG7 (3.95 ppm, Figure 5), which has been shown previously to be a characteristic marker of *syn* guanines in modified adduct duplexes (*2*, *7*, *9*, *10*, *19*, *28*). The magnitude of this downfield H2′ proton shift of the adducted guanine in the [AF]dG·del(−2) 12/10-mer is larger than what was observed previously (3.11 ppm) for a related study of a *syn* [AF]dG adduct positioned opposite a −2 deletion site in a different sequence context (*10*). Additional support for a *syn* alignment at the modified guanine is reflected in the downfield shift of the sugar C1′ of [AF]dG7 (86.92 ppm) relative to other guanines in the adduct duplex (Figure 6A), which is comparable to that observed previously (86.96 ppm) in the related study of a *syn* [AF]dG adduct positioned opposite a −2 deletion site in a different sequence context (*10*).

The modified guanine of [AF]dG7 is displaced into the major groove following intercalation of the fluorenyl ring into the helix in the solution structure of the [AF]dG·del-(−2) 12/10-mer (Figure 8A). The plane of the modified guanine is tilted slightly toward the 5′-direction relative to the helix axis and stacks over the sugar ring of dC6 in the structure of the adduct duplex (Figures 8A,B). We observe a strong NOE between the H6 proton of dC8 and the H1′ proton of [AF]dG7 for the [AF]dG7-dC8 step (see arrow, Figure 4) consistent with a separation between these protons of 2.8 ± 0.1 Å as measured in the intensity-refined structures of the adduct duplex (Table S3).

The imino and amino protons of the [AF]dG7 adduct are exposed to solvent in the solution structure of the [AF]-dG·del(−2) 12/10-mer (Figure 8A,B). This is consistent with the observed [AF]dG7 imino proton chemical shift of 10.70 ppm, which is to the high field of imino protons of dG·dC base pairs, and the observed averaged [AF]dG7 amino proton chemical shift of 6.33 ppm, which is to the high field of guanine amino protons participating in hydrogen-bonding alignments.

Conformation of Unpaired dC6. The dC6 base is looped out of the helix into the major groove in the solution structure of the [AF]dG·del(−2) 12/10-mer (Figure 8A). The displacement of the dC6 residue out of the helix results in a break in the base proton to its 5′-flanking sugar H1′ proton sequential NOE connectivities at the dG5-dC6 step (boxed region, Figure 4) in the adduct duplex. Indeed, this distance spans the range 6.0 ± 0.3 Å in the intensity-refined structures of the adduct duplex (Table 2), which readily explains the absence of a sequential NOE between this pair of protons. In addition, the dG5(H8) base proton to dC6(H5) base proton distance spans the range 6.5 ± 0.7 Å among the intensity-refined solution structure of the adduct duplex consistent with the absence of an NOE cross-peak between these pair of protons (Figure 4).

We have been unable to observe the amino protons of dC6 in the spectrum of the [AF]dG·del(−2) 12/10-mer. This is not surprising since dC6 is bulged out of the helix and not paired with any residue in the adduct duplex. The amino protons of dC6 are probably broadened out due to intermediate rotation about the C^4-N^4 exocyclic amino linkage.

Aminofluorene Intercalation into the Helix. Several lines of evidence support intercalation of the fluorenyl ring between dG5·dC18 and dC8·dG17 base pairs in the [AF]-dG·del(−2) 12/10-mer. These include upfield shifts in the imino protons of dG5 (11.60 ppm) and dG17 (12.01 ppm) as a result of stacking with the intercalated fluorenyl ring system (Figure 2A and Table 1) and upfield shifts of the fluorenyl ring protons as a result of stacking with flanking dG·dC base pairs (Figure S1) in the solution structure of the adduct duplex. The fluorenyl ring stacks primarily with the dG5 and dC8 on the modified strand of the adduct duplex (Figure 8B). By contrast, there is less overlap between the

fluorenyl ring and dG17 and none at all with dC18 on the unmodified strand in the adduct duplex (Figure 8B).

The intermolecular NOEs observed in the [AF]dG·del(−2) 12/10-mer (Table 2) are consistent with the H5−H8 proton-containing edge of the AF ring being directed toward the dG17-dC18 segment on the unmodified strand and the H1−H4 proton-containing edge of the AF ring being directed toward the dG5−dC8 segment on the modified strand (Figure 8A,B). We observe only one rotamer alignment about the β' ([AF]dG7(C^8)-[AF](N)-[AF](C^2)-[AF](C^1)) linkage bond for which the flourenyl H9a,b protons are directed toward the minor groove (Figure 8B).

We did not use coupling constant data as restraints during the computations. A list of sugar pseudorotation pucker (**P**) parameters and glycosidic torsion (χ) angles for residues at and adjacent to the lesion site in the intensity-refined structures of the [AF]dG·del(−2) 12/10-mer are listed in Table S3 (Supporting Information). The **P** value for dG17 is distinct and spans the range 14 ± 20°. Experimentally, we do not observe the H1′-H2′ coupling cross-peak for dG17 consistent with this residue adopting a sugar pucker in the C3′-*endo* range.

The observed downfield phosphorus chemical shifts for the [AF]dG7-dC8 step on the modified strand and, to a lesser extent, for the dC19-dG20 step on the partner strand of the [AF]dG·del(−2) 12/10-mer must reflect the conformational perturbations associated with accommodation of the [AF]-dG adduct opposite the −2 deletion site in the *Nar*I sequence context. We are reluctant to correlate these chemical shift changes with specific changes in the backbone torsion angles as has been proposed in the literature (*32*) since we have no direct measure of these torsion angles and there is a sufficient spread in values at these backbone torsion angles amongst the intensity-refined structures of the [AF]dG·del(−2) 12/10-mer.

Sequence Context of [AF]dG Opposite −2 Deletion Site. Our initial efforts to structurally characterize the [AF]dG adduct opposite −1 and −2 deletion sites were restricted to sequence contexts where there was only one possible pairing alignment for base pairs that flanked the deletion site (*9*, *10*). We studied both [AF]dG-dA and dA-[AF]dG steps opposite the −2 deletion site since it was not apparent which sequence context would provide NMR spectra suitable for the planned structural characterization of the system. We were successful with the d(C-[AF]G-A-C)·d(G-G) sequence context where a single conformation was observed for this case where the extra adenine is 3′ to the adduct site (*10*). The structure was of the AF-intercalation type with the *syn* modified guanine displaced into the major groove and partially stacked with the extra adenine in the 3′ direction, which was also displaced into the major groove. A parallel study was unsuccessfully attempted on an [AF]dG adduct opposite a −2 deletion site in the related d(C-A-[AF]G-C)·d-(G-G) sequence context where the extra adenine is 5′ to the adduct site. We were unable to structurally characterize this system due to the increased line widths of the residues at and adjacent to the lesion site suggestive of conformational equilibrium associated with this sequence context.

We have already established above that the [AF]dG·del-(−2) 12/10-mer in the *Nar*I sequence context exclusively adopts alignment **2** shown in Figure 1B containing a dC6-[AF]dG7 step with no evidence for the formation of alignment **3** shown in Figure 1C containing a [AF]dG7-dC8 step opposite the −2 deletion site. Thus, the extra base is 5′ to the [AF]dG adduct site in the current study on the −2 deletion site in the *Nar*I sequence context in contrast to our previous study where the extra base was 3′ to the [AF]dG adduct site (*10*). Both structures are of the AF-intercalation type with the modified guanine in a *syn* alignment and displaced into the major groove of the helix. The fluorenyl protons are shifted to high field as a consequence of intercalation in both structures of the [AF]dG adduct opposite −2 deletion sites (Figure S1). However, there are several differences between the two structures that are summarized below. The looped out modified *syn* guanine does not stack with its 5′ looped out dC neighbor in this study (Figure 8A) while the looped out modified *syn* guanine stacked with its 3′ looped out dA neighbor in the previous study (*10*). The DNA helix is less bent in the present study of the [AF]dG adduct opposite a −2 deletion site in the *Nar*I sequence context (Figure 9), with an average bend angle of about 21 ± 11.6° (using the program of ref 33), in contrast to a bend of about ≈45° centered about the lesion site in the previous study (the two structures are compared in Figure S3, Supporting Information) (*10*) . We compute an average helical twist of 39.5 ± 5° (using the program of ref 33) between dG·dC base pairs flanking the −2 deletion site in the present study (Figure 8B). Together, the two studies on the [AF]dG adduct opposite a −2 deletion site considerably extend our knowledge base of potential conformational alignments and their sequence dependence for these systems that are key to our understanding of the molecular basis of frame shift mutations.

Comparison of First Stage DUPLEX and Second Stage X-PLOR Intensity-Refined Structures. The key conformational features centered about the [AF]dG adduct site are shared by both the first stage DUPLEX and subsequent X-PLOR intensity-refined structures of the [AF]dG·del(−2) 12/10-mer. This includes the intercalation of the AF ring between the Watson-Crick dG5·dC18 and dC8·dG17 base pairs with C^9-containing edge of the aminofluorene ring facing the minor groove. Furthermore, the carcinogen-base linkage site torsion angles α' ([AF]dG7(N^9)-[AF]dG7(C^8)-[AF](N)-[AF](C^2)) and β' ([AF]dG7(C^8)-[AF](N)-[AF](C^2)-[AF](C^1)) are 227° and 186°, respectively, for the unrestrained DUPLEX-based structure and are 210 ± 8° and 188 ± 14°, respectively, for the X-PLOR intensity-refined structures.

Comparison of Structural Features of [AAF]dG and [AF]-dG Adducts Opposite −2 Deletion Sites in the NarI Sequence Context. NMR studies have been reported for DNA oligomer duplexes containing [AAF]dG (*34*) and [AF]dG (this study) substitutions at the G^3 position in the d(C-G^1-G^2-C-G^3-C-C)·d(G-G-C-C-G) *Nar*I sequence context with the adducts positioned opposite −2 deletion sites. The NMR data reported previously on the [AAF]dG adduct opposite a −2 deletion site in the *Nar*I sequence context could not be unambiguously analyzed by Milhe et al. (*34*) due to a heterogeneous mixture of conformers in equilibrium, broadening of resonances and lack of sequential connectivities within the d(G^2-C-[AAF]G^3-C-C) segment and the total absence of intermolecular NOEs between the AAF and the DNA. These limitations have precluded a structural characterization of the conformation of the lesion site containing the [AAF]dG adduct opposite the −2 deletion site in the *Nar*I sequence context (*34*). Nevertheless, sufficient information was available for the authors to reach certain qualitative structural conclusions as outlined below. The

upfield shift of the fluorenyl protons and the upfield shift of the imino and amino protons of the [AAF]dG adduct are consistent with intercalation of the AAF into the helix with base displacement of the modified guanine out of the helix for the predominant (≈80%) conformer in solution (*34*). The authors suggest that the modified guanine adopts a *syn* conformation based on its imino and amino proton chemical shifts, which in our opinion are not markers for distinguishing between *syn* and *anti* alignments for the glycosidic torsion angle in nucleic acids. The sugar protons of [AAF]dG were unassigned in the adduct duplex (*34*), and hence the authors could not use the chemical shift of the sugar H2′ proton to distinguish between the *anti* and *syn* glycosidic torsion angle for the adduct site. They also suggest that the two-base bulge involves the modified [AAF]dG7 residue and its 3′-flanking dC8 residue (alignment **3** as shown in Figure 1C) based on the narrowness of the proton resonances that were assigned to the 3′-flanking cytosine dC8 in the adduct duplex. They came to this conclusion without a definitive basis for differentiating between the protons of cytosines dC6 and dC8 that flank the [AAF]dG7 in the adduct duplex. It is our opinion that the published NMR data presented to date on the [AAF]dG adduct opposite a −2 deletion site in the *Nar*I sequence context (*34*) do not permit a definitive differentiation between pairing alignments **2** and **3** shown in Figure 1 for the predominant conformer in solution.

The availability of excellent NMR spectra and a large number of intermolecular NOEs corresponding to a single conformer for the [AF]dG adduct opposite a −2 deletion site in the *Nar*I sequence context (this study) has allowed us to undertake a combined NMR−computational characterization of the structure at and adjacent to the lesion site. Specifically, we have been able to define the details of the AF-intercalated conformer ranging from the stacking alignment of the intercalated fluorenyl ring with the flanking dG·dC pairs (Figure 8B) to the positioning of the base-displaced modified *syn* guanine and its 5′-linked cytosine in the major groove corresponding to the alignment **2** shown in Figure 1B for the adduct duplex (Figures 8 and 9). The acetyl group can be accommodated in the solution structure of the [AF]dG·del(−2) 12/10-mer corresponding to the alignment **2** shown in Figure 1B with minor perturbations to generate the corresponding plausible model of the [AAF]-dG·del(−2) 12/10-mer (Figure S4, Supporting Information) in the *Nar*I sequence context.

Related Structures of Adducts Opposite −2 Deletion Sites. There exist only a few examples in the literature on DNA lesions opposite −2 deletion sites where the conclusions are justified based on the quality of the NMR spectra and data sets. These in turn have permitted NMR−computational structural characterization of the adduct duplexes yielding novel structural and functional insights into the origin of frame shift mutagenesis. The first such studies were undertaken on a propano-dG adduct positioned opposite a −2 deletion site in a frameshift hotspot of *Salmonella typhimurium his*D3052 (*35*, *36*). Our contribution on [AF]-dG adducts opposite −2 deletion sites in different sequence contexts (*10*; this study) have defined the range of conformational alignments and yielded the first definitive insights toward a molecular understanding of covalent aromatic amine-guanine adduct-induced frameshifts in the *Nar*I hot spot site.

Biological Implications. The concept of a slipped mutagenic intermediate has a long intellectual history as a hypothesis to explain frameshift mutations in unmodified DNA (*37*−*39*) and in carcinogen-modified DNA (*34*, *40*−*43*). A potential mechanism involves stalling of the polymerase in the vicinity of the damaged base, which permits formation of a misaligned intermediate during replication. The d(C-G^1-G^2-C-G^3-C-C) *Nar*I hotspot sequence facilitates formation of a two-base bulge following adduct formation at G^3 but not at G^1 or G^2 positions in the sequence.

The present structural study defines the alignment of the [AF]dG adduct opposite a −2 deletion site in the *Nar*I sequence context. The lesion site adopts a single conformation with pairing as shown in **2** (Figure 1B) resulting in the extrusion of the dC6-[AF]dG7 step to the total exclusion of pairing as shown in **3** (Figure 1C), which requires the extrusion of the [AF]dG7-dC8 step. This polarity establishes that the structural perturbation extends 5′ rather than 3′ to the [AF]dG lesion site in the adduct duplex. The repair machinery could act on the lesion site through recognition of the helical distortions associated with intercalation of the AF ring into the helix and/or through recognition of the displaced modified guanine and 5′-cytosine that are positioned in the major groove.

DNA bending, under- or overwinding, disruption of base stacking interactions and increased deformability are features of DNA structure believed to promote recognition of lesions by DNA repair enzymes, especially the *Uvr*ABC excision repair system of *E. coli* (*44*−*46*). The present structural study of the [AF]dG adduct in the −2 deletion site in the *Nar*I sequence context manifests elements of all these features, with bend angles in the 21.9 ± 11.6° range surrounding the −2 deletion site, helical twist angles in the 39.5 ± 5° range flanking the −2 deletion site, looped out unstacked bases, and the increased deformability associated with intercalation of the aminofluorene chromophore into the helix and base displacement of the modified guanine into the major groove.

There is no NMR-based solution structure available of the [AAF]dG adduct positioned opposite a −2 deletion site in the *Nar*I sequence context. The available published NMR data of this adduct duplex (*34*) permits only qualitative structural analysis of the experimental data. Clearly, a direct comparison of structural features between [AF]dG and [AAF]-dG adducts opposite the −2 deletion site in the *Nar*I sequence context requires a solution structure of the latter adduct duplex.

Our structural studies to date on AF-intercalated conformers where the [AF]dG adduct is positioned opposite dC (*7*), opposite a −1 deletion site (*9*), and opposite a −2 deletion site (*10*; this study) exhibit a *syn* alignment of the modified guanine in the adduct duplexes. It is plausible that a *syn* alignment would also be favored for [AAF]dG adducts in AAF-intercalated conformers at the *Nar*I hotspot in a −2 deletion context, given that this alignment is strongly prefered based on previous computational studies on this bulkier (relative to AF) adduct (*47*−*49*). Indeed, an AAF-intercalated conformer with a *syn*-modified guanine has been characterized for [AAF]dG opposite dC in a DNA duplex (*50*), and more examples should emerge over time. Finally, AF's lesser propensity to induce −2 deletion mutations at the *Nar*I G^3 hotspot (*14*) might also be related to [AF]dG's ability to adopt both *anti* and *syn* alignments, while [AAF]-dG strongly favors the *syn* domain.

ACKNOWLEDGMENT

We thank Peter Wei for his technical assistance on this project.

SUPPORTING INFORMATION AVAILABLE

Three tables listing the complete exchangeable and nonexchangeable proton chemical shifts and the sugar pseudorotation and glycosidic torsion angles in the refined structures of the [AF]dG·del(−2) 12/10-mer and four figures comparing aminofluorene proton chemical shifts for [AF]-dG adducts positioned opposite −2 deletion sites, the superposition of three structures that best fit the NMR data obtained by molecular mechanics calculation using DU-PLEX, a comparision of the solution structure of the entire [AF]dG opposite −2 deletion sites reported in this study and that of Mao et al. (*10*), and a model of the [AAF]dG·del-(−2) 12/10-mer based on the solution structure of the [AF]-dG·del(−2) 12/10-mer (9 pages). Ordering information is given on any current masterhead page.

REFERENCES

1. Beland, F. A., and Kadlubar, F. F. (1990) in *Handbook of Experimental Pharmacology, Vol. 94/I: Chemical Carcinogenesis and Mutagenesis* (Cooper, C. S., and Grover, P. L., Eds.) pp 267−325. Springer-Verlag, Heidelberg.
2. Norman, D., Abuaf, P., Hingerty, B. E., Live, D., Grunberger, D., Broyde, S., and Patel, D. J. (1989) *Biochemistry 28*, 7462−7476.
3. Abuaf, P., Hingerty, B. E., Broyde, S., and Grunberger, D. (1995) *Chem. Res. Toxcol. 8*, 369−378.
4. Eckel, L. M., and Krugh, T. R. (1994) *Nat. Struct. Biol. 1*, 89−94.
5. Eckel, L. M., and Krugh, T. R. (1994) *Biochemistry 33*, 13611−13624.
6. Cho, B. P., Beland, F. A., and Marques, M. M. (1994) *Biochemistry 33*, 1373−1384.
7. Mao, B., Cosman, M., Hingerty, B. E., Broyde, S., and Patel, D. J. (1997) *Biochemistry* (in press).
8. Mao, B., Cosman, M., Hingerty, B. E., Broyde, S., and Patel, D. J. (1997) *Biochemistry* (in press).
9. Mao, B., Cosman, M., Hingerty, B. E., Broyde, S., and Patel, D. J. (1995) *Biochemistry 34*, 6226−6238.
10. Mao, B., Cosman, M., Hingerty, B. E., Broyde, S., and Patel, D. J. (1995) *Biochemistry 34*, 16641−16653.
11. Fuchs, R. P. P., Schwartz, N., and Daune, M. P. (1981) *Nature 294*, 657−659.
12. Burnouf, D., Koehl, P., and Fuchs, R. P. P. (1989) *Proc. Natl. Acad, Sci. U.S.A. 86*, 4147−4151.
13. Fuchs, R. P. P. (1983) *J. Mol. Biol. 177*, 173−180.
14. Bichara, M., and Fuchs, R. P. P. (1985) *J. Mol. Biol. 183*, 341−351.
15. Melchior, W. B., Jr., Marques, M. M., and Beland, F. A. (1994) *Carcinogenesis 15*, 889−899.
16. Tebbs, R.-S., and Romano, L. J. (1994) *Biochemistry 33*, 8998−9006.
17. Koehl, P., Valladier, P., Lefevre, J. F., and Fuchs, R. P. P. (1989) *Nucleic Acids Res. 17*, 9531−9541.
18. Marion, D., Ikura, M., Tschudin, R., and Bax, A. (1989) *J. Magn. Reson. 85*, 393−399.
19. Mao, B., Vyas, R. R., Hingerty, B. E., Broyde, S., Basu, A. K., and Patel, D. J. (1996) *Biochemistry 35*, 12659−12670.
20. Sklenar, V., Miyashiro, H., Zon, G., Miles, H. T., and Bax, A. (1986) *FEBS Lett. 208*, 94−98.
21. Bax, A., and Subramanian, J. (1986) *J. Magn. Reson. 67*, 565−570.
22. Patel, D. J., Kozlowski, S. A., Nordheim, A., and Rich, A. (1982) *Proc. Natl. Acad. Sci. U.S.A. 79*, 1413−1417.
23. van de Ven, F. J., and Hilbers, C. W. (1988) *Eur. J. Biochem. 178*, 1−38.
24. Hingerty, B. E., Figueroa, S., Hayden, T., and Broyde, S. (1989) *Biopolymers 28*, 1195−1222.
25. Brunger, A. T. (1992) *X-Plor: A system for X-ray Crystallography and NMR*, Yale University Press, New Haven and London.
26. Grunberger, D., Nelson, J. H., Cantor, C. R., and Weinstein, I. B. (1970) *Proc. Natl. Acad. Sci. U.S.A. 66*, 488−494.
27. Fuchs, R. P. P., and Daune, M. (1971) *FEBS Lett. 34*, 295−298.
28. Cosman, M., Hingerty, B. E., Geacintov, N. E., Broyde, S., and Patel, D. J. (1995) *Biochemistry 34*, 15334−15350.
29. Ghose, R., Marino, J. P., Wiberg, K. B., and Prestegard, J. H. (1994) *J. Am. Chem. Soc. 116*, 8827−8828.
30. Greene, K. L., Wang, Y., and Live, D. (1995) *J. Biomol. NMR 5*, 333−338.
31. Arnott, S., Bond, P. J., Selsing, E., and Smith, P. J. (1976) *Nucleic Acids Res. 2*, 2459−2470.
32. Nikonowicz, E. P., and Gorenstein, D. G. (1990) *Biochemistry 29*, 8845−8858.
33. Gorin, A. A., Zhurkin, V. B., and Olson, W. K. (1995). *J. Mol. Biol. 247*, 34−38.
34. Milhé, C., Fuchs, R. P. P., and Lefevre, J.-F. (1996) *Eur. J. Biochem. 235*, 120−127.
35. Moe, J., Reddy, G. R., Marnet, L. J., and Stone, M. P. (1994) *Chem. Res. Toxicol. 7*, 319−328.
36. Weisenseel, J. P., Moe, J. G., Reddy, G. R., Marnett, L. J., and Stone, M. (1995) *Biochemistry 34*, 50−64.
37. Streisinger, G., Okada, Y., Emrich, J., Newton, J., Tsugita, A., Terzaghi, E. and Inouye, M. (1966) *Cold Spring Harbor Symp. Quant. Biol. 31*, 77−84.
38. Drake, J. W., and Baltz, R. W. (1976) *Annu. Rev. Biochem. 45*, 11−37.
39. Kunkel, T. A. (1990) *Biochemistry 29*, 8003−8011.
40. Schaaper, R. M., Koffel-Schwartz, N., and Fuchs, R. P. P. (1990) *Carcinogenesis 11*, 1087−1095.
41. Lambert, I. B., Napolitano, R. L., and Fuchs, R. P. P. (1992) *Proc. Natl. Acad. Sci. U.S.A. 89*, 1310−1314.
42. Shibutani, S., and Grollman, A. P. (1993) *J. Biol. Chem. 268*, 11703−11710.
43. Napolitano, R. L., Lambert, I. B., and Fuchs, R. P. P. (1994) *Biochemistry 33*, 1311−1315.
44. Williams, L. D., and Gao, Q. *Biochemistry 31*, 4315−4324.
45. van Houten, B., and Snowden, A. (1993) *Bioessays 15*, 51−59.
46. Zou, Y., Liu, T. M., Geacintov, N. E., and van Houten, B. (1995) *Biochemistry 34*, 13582−13593.
47. Lipkowitz, K. B., Chevalier, T., Widdifield, M., and Beland, F. A. (1982) *Chem.-Biol. Interact. 40*, 57−76.
48. Hingerty, B. E., and Broyde, S. (1982) *Biochemistry 21*, 3243−3252.
49. Shapiro, R., Sidawi, D., Miao, Y. S., Hingerty, B. E., Schmidt, K. E., Moskowitz, J., and Broyde, S. (1994) *Chem. Res. Toxcol. 7*, 239−253.
50. O'Handley, F. F., Sanford, D. G., Xu, R., Lester, C. C., Hingerty, B. E., Broyde, S., and Krugh, T. R. (1993) *Biochemistry 32*, 2481−2497.
51. Kraulis, P. J. (1991) *J. Appl. Crystallogr. 24*, 946−950.

BI972205Z