NMR Solution Structure of an Oligodeoxynucleotide Duplex Containing the Exocyclic Lesion 3,N^4-Etheno-2′-deoxycytidine Opposite Thymidine: Comparison with the Duplex Containing Deoxyadenosine Opposite the Adduct†

David Cullinan,‡ Alexandre Korobka,‡ Arthur P. Grollman,‡ Dinshaw J. Patel,§ Moisés Eisenberg,‡ and Carlos de los Santos*,‡

Department of Pharmacological Sciences, State University of New York at Stony Brook, Stony Brook, New York 11794-8651, and Cellular Biochemistry and Biophysics Program, Memorial Sloan-Kettering Cancer Center, New York, New York 10021

Received March 7, 1996; Revised Manuscript Received June 24, 1996⊗

ABSTRACT: The exocyclic 3,N^4-etheno-2′-deoxycytidine adduct was incorporated at the center of the oligodeoxynucleotide duplex d(C-G-T-A-C-ϵC-C-A-T-G-C)·d(G-C-A-T-G-T-G-T-A-C-G), and its solution structure was analyzed using high-resolution proton NMR spectroscopy and molecular dynamics simulations. The experimental data indicate that the oligodeoxynucleotide duplex adopts a right-handed helical structure with sugar puckers in the C2′-endo/C3′-exo range and Watson−Crick hydrogen bond alignments for all base pairs. NOE connectivities established a *syn* orientation for the glycosidic torsion angle of the exocyclic adduct. Restrained molecular dynamics simulations, using the full relaxation matrix approach, produced a three-dimensional model in agreement with the experimental data. The structure shows only minor perturbations in the sugar-phosphate backbone and a 27° bend of the helical axis at the lesion site. On the refined model a well-formed hydrogen bond between T(N3H) and ϵC(N4) stabilizes the ϵC(*syn*)·T(*anti*) base pair alignment, reflecting the preference of the adduct for the *syn* orientation. Furthermore, the ϵC(*syn*)·T(*anti*) base pair stacks with flanking base pairs. We discuss a correlation between the mutagenic properties of the adduct and the three-dimensional structure of the ϵC·dA and ϵC·T duplexes.

Soon after the double helical nature of DNA was elucidated, it became evident that the fidelity of genetic information is mediated by the formation of specific hydrogen bonds between the bases. Watson−Crick hydrogen bonds define base complementarity and are believed to play a decisive role in the incorporation of correct bases during DNA replication. Genomic DNA is continuously exposed to damaging agents, including xenobiotics, UV radiation, and activated oxygen species, that may produce changes in its chemical structure. In turn, these modifications can affect base pair alignment and impair the fidelity of genetic replication. DNA lesions can direct preferential incorporation of specific deoxynucleotides by mechanisms generally associated with the formation of hydrogen bonds. The O^6-methyl-2′-deoxyguanosine adduct, a product of the reaction of alkylating agents with DNA, is a well-studied prototype of a miscoding lesion that induces G→A transitions (Snow *et al.*, 1984; Loechler *et al.*, 1984). This preference is likely to originate in the formation of a wobble G·T base pair stabilized by hydrogen bonds during DNA replication. Alternatively, DNA damage can impair the formation of hydrogen bonds originating a noncoding lesion. Abasic sites are an example of noncoding lesions: they partially block DNA replication and cause preferential incorporation of deoxyadenosine opposite the lesion (Kunkel, 1984; Rabkin & Strauss, 1984). This mutational specificity, known as "the adenine rule", appears to be a general property of noncoding lesions, but the exact mechanism has not been established.

3,N^4-Etheno-2′-deoxycytidine (ϵC)[1] adducts are produced when the human carcinogen vinyl chloride reacts with DNA (Barbin & Bartsch, 1986; Singer & Spengler 1986). These lesions are also generated by the reaction of lipid peroxidation products with DNA (el Ghissassi *et al.*, 1995) and have been detected as endogenous adducts in human and rat liver DNA (Nath *et al.*, 1994; Nair *et al.*, 1995). The ϵC adduct induces mainly ϵC→T transitions and ϵC→A transversions *in vitro* (Simha *et al.*, 1991; Zhang *et al.*, 1995a) and *in vivo* (Palejwala *et al.*, 1993; Basu *et al.*, 1993; Moriya *et al.*, 1994). The mutational frequency of ϵC is different in bacterial and mammalian cells, reflecting alternative processing of the lesion during DNA replication (Moriya *et al.*, 1994).

In the present study we assessed the structural properties of the ϵC·T base pair, which leads to the mutagenic ϵC→A transversion, by solving the solution structure of the d(C-G-T-A-C-ϵC-C-A-T-G-C)·d(G-C-A-T-G-T-G-T-A-C-G) oligodeoxynucleotide duplex (referred to in this paper as the

† This research was supported by Grant CA47995 (to A.P.G.) from the National Institutes of Health.

* To whom correspondence should be addressed. Phone: (516) 444-3649. Fax (516) 444-3218. E-mail: carlos@pharm.sunysb.edu.

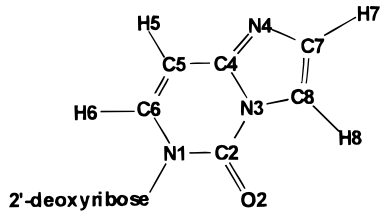
‡ State University of New York at Stony Brook.

§ Memorial Sloan-Kettering Cancer Center.

⊗ Abstract published in *Advance ACS Abstracts,* October 1, 1996.

[1] Abbreviations: NMR, nuclear magnetic resonance; EDTA, disodium ethylenediaminetetracetate; TSP, sodium (2,2,3,3-d_4)-3-(trimethylsilyl)propionate; NOESY, nuclear Overhauser effect spectroscopy; COSY, correlation spectroscopy; DQF-COSY, double quantum filtered correlation spectroscopy; HOHAHA, homonuclear Hartmann−Hahn spectroscopy; NOE, nuclear Overhauser effect; HPLC, high-performance liquid chromatography; ppm, parts per million; ϵC, 3,N^4-etheno-2′-deoxycytidine; dC, 2′-deoxycytidine; dA, 2′-deoxyadenosine; dG, 2′-deoxyguanosine; T, thymidine; dAMP, 2′-deoxyadenosine monophosphate; TMP, thymidine monophosphate; dNMP, 2′-deoxynucleotide monophosphate.

S0006-2960(96)00570-3 CCC: $12.00



3,N⁴-etheno-2'-deoxycytidine

(εC)

C_1 G_2 T_3 A_4 C_5 εC_6 C_7 A_8 T_9 G_{10} C_{11}
G_{22} C_{21} A_{20} T_{19} G_{18} T_{17} G_{16} T_{15} A_{14} C_{13} G_{12}

εC•T duplex

FIGURE 1: (A) Chemical structure of 3,N^4-etheno-2′-deoxycytidine (ϵC). (B) Sequence and numbering scheme of the ϵC·T duplex.

ϵC·T duplex), using NMR spectroscopy and restrained molecular dynamics simulations. We also compare the structural features of ϵC·dA [reported in the companion paper by Korobka *et al.* (1996)] and ϵC·T duplexes and correlate these structures with the biological properties of the lesion. The chemical structure of the 3,N^4-etheno-2′-deoxycytidine adduct and the numbering scheme of the duplex sequence employed in this study are shown in Figure 1.

MATERIALS AND METHODS

Synthesis and Purification of the Oligonucleotide Duplexes. 3,N^4-Etheno-2′-deoxycytidine was synthesized as previously described (Zhang *et al.*, 1995b) and incorporated into oligodeoxynucleotides using standard phosphoramidite chemistry procedures. Isolation and purification procedures were reported elsewhere [Korobka *et al.*, 1996 (companion paper)].

Duplex Formation and Sample Preparation. A 1:1 stoichiometry of the ϵC·T duplex was achieved by monitoring the intensity of the thymidine H6 and the sugar H1′ NMR signals during addition of the unmodified strand to the strand containing the ϵC adduct, at 55 °C. The sample used for NMR analysis consisted of 320 OD_{260} of duplex dissolved in 0.4 mL of 10 mM phosphate buffer, pH 6.8, containing 50 mM NaCl and 1 mM EDTA in either 99.96% D_2O or 90% H_2O−10% D_2O (v/v); this corresponds to a concentration of approximately 8 mM. The pH was adjusted by adding dilute HCl or NaOH solutions, and the samples were degassed under nitrogen before data collection. Values in D_2O buffer were uncorrected pH-meter readings.

NMR Experiments. One- and two-dimensional NMR spectra were recorded on a Bruker AM series spectrometer operating at a field strength of 500 MHz. NOE buildup experiments were collected on a Bruker AMX 600 MHz spectrometer. Proton chemical shifts were referenced relative to external TSP at 0.0 ppm. Phase-sensitive (States *et al.*, 1982) NOESY (150 ms mixing time) spectra in H_2O were recorded using a jump and return reading pulse (Plateau & Gueron, 1982). Phase-sensitive NOESY spectra in D_2O (50, 90, 130, 170, and 210 ms mixing times) and phase-sensitive COSY and DQF-COSY spectra were recorded with a repetition delay of 1.3 s during which the residual water signal was suppressed by presaturation. Time domain data sets normally consisted of 1024 complex data points in the t_2 dimension and 256 complex points in the t_1 dimension. NMR data were processed and analyzed using the program Felix (Biosym Technologies).

Computer Simulations. Restrained molecular dynamic simulations were performed on a Silicon Graphics Indigo 2 workstation using X-PLOR 3.1 (Brunger, 1993) with an all-atom force field set. Structures were visualized using Midas Plus (UCSF, Computer Graphics Laboratory) and structural parameters obtained with Curves (Lavery & Sklenar, 1988, 1989). Conditions used for the molecular dynamics simulations were described previously [Korobka *et al.*, 1996 (companion paper)]. Two different starting models were constructed from canonical A- and B-form DNA with the glycosidic angle of the ϵC residue set to 30° and 59°, respectively. The initial structures were energy minimized by 3000 steps of conjugate gradient prior to 37.5 ps of molecular dynamics simulations. Coordinates of the final 1 ps were averaged and subjected to 2500 steps of energy minimization. Distance-refined structures were further optimized by 10 ps of molecular dynamics restrained by the difference between experimental and back-calculated NOE peak intensities (Yip & Case, 1989; Nilges *et al.*, 1991). 899 experimental volumes measured from the NOESY (50, 90, 130, 170 and 210 ms mixing times) spectra were input as restraints with a 30% error range. A grid search revealed that an isotropic correlation time of 3.75 ns produced the closest initial fit and this value was used during the molecular dynamics simulations. Coordinates of the last picosecond were averaged and energy minimized, yielding the final NMR-restrained structures.

RESULTS

Assignment of the Nonexchangeable Protons. The one-dimensional proton spectrum in D_2O buffer, pH 6.8, 25 °C, is shown in Figure 1S, Supporting Information. The spectrum shows excellent resolution and is characterized by the presence of sharp lines as well as the absence of minor resonances. Assignment of the nonexchangeable base and H1′ sugar protons was accomplished by analysis of NOESY spectra recorded in D_2O buffer, pH 6.8, 25 °C, following standard procedures (Hare *et al.*, 1983; Wüthrich, K., 1986; van de Ven & Hilbers, 1988).

Figure 2 shows a region of the NOESY (300 ms mixing time) spectrum (in duplicate) that establishes distance interactions between the base (purine H8/pyrimidine H6) (6.7−8.5 ppm) and the H1′ (5.1−6.4 ppm) sugar proton regions. Sequential connectivities can be observed without interruptions from C1 to C11 on the modified strand and from G12 to G22 in the complementary strand (Figure 2, A and B, respectively). NOE cross peaks are observed between the base (purine H8/pyrimidine H6) and deoxycytidine H5 protons in the G12-C13, A20-C21, A4-C5, ϵC6-C7, and G10-C11 steps of the duplex (peaks A and B, Figure 4B, and peaks A−C, Figure 4A, respectively). The deoxyadenosine H2 protons are identified readily by their weak inter- and intra-strand connectivities with H1′ sugar protons, as is observed in the modified and unmodified strands of the ϵC·T duplex (peaks D−I, Figure 2A, and peaks C−J, Figure 2B, respectively). The exocyclic H7 and H8 protons were identified by the strong NOE peak between them and confirmed by the presence of a COSY cross peak (data not shown). Their specific assignment was based on structural considerations and the NOE peaks detected at 300 ms mixing

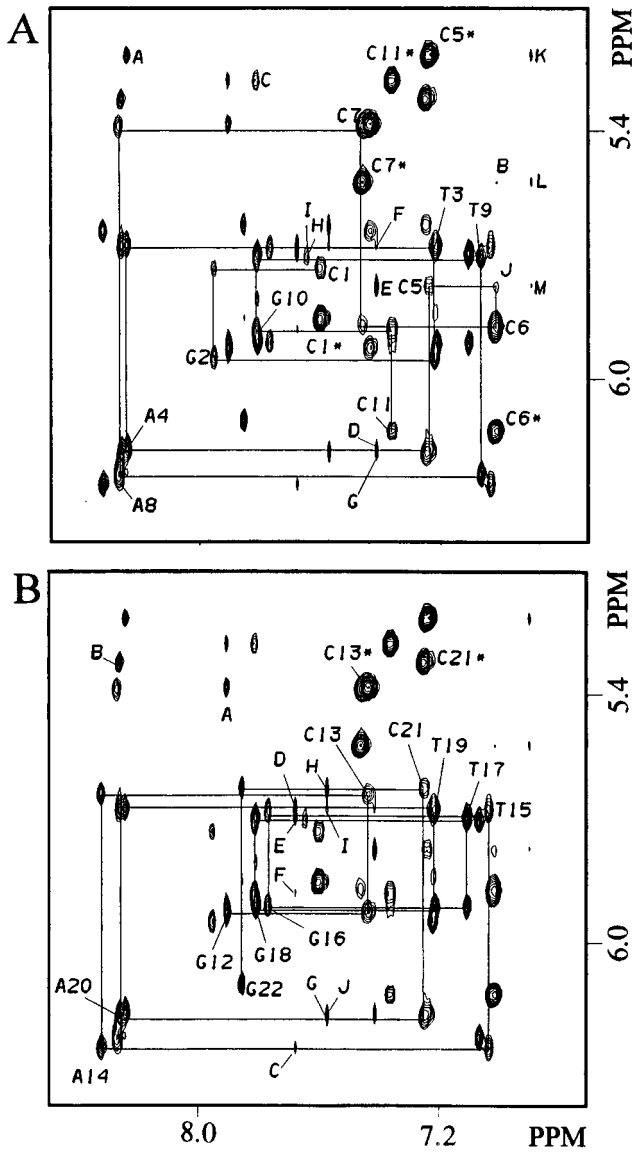


FIGURE 2: Duplicate counterplots of a portion of the NOESY (300 ms mixing time) spectrum recorded in D_2O buffer, pH 6.8, 25 °C, showing distance connectivities between base and H1′ sugar protons. Solid lines connect each base (purine H8/pyrimidine H6) proton to its own and 5′-flanking H1′ sugar proton in the (A) modified and (B) unmodified strands. Labeled peaks are assigned as follows: (A) A, A4(H8)-C5(H5); B, ϵC6(H6)-C7(H5); C, G10(H8)-C11(H5); D, A4(H2)-A4(H1′); E, A4(H2)-C5(H1′); F, A4(H2)-T19(H1′); G, A4(H2)-A20(H1′); H, A8(H2)-T9(H1′); I, A8(H2)-T15(H1′); J, C5(H1′)-ϵC6(H6); K, ϵC6(H8)-C5(H5); L, ϵC6(H8)-C7(H5), and M, ϵC6(H8)-C5(H1′). (B) A, G12(H8)-C13(H5); B, A20(H8)-C21(H5); C, A14(H2)-A14(H1′); D, A14(H2)-T15(H1′); E, A14(H2)-T9(H1′); F, A14(H2)-A10(H1′); G, A20(H2)-A20(H1′); H, A20(H2)-A21(H1′); I, A20(H2)-T3(H1′), and J, A20(H2)-A4(H1′). Asterisks indicate the cytosine H5-H6 cross peak.

time. When ϵC is *syn*, as established by the short mixing time NOESY experiment (see below), its exocyclic H8 proton is located in the major grove of the helix, where it showed NOE connectivities to the C5(H5) and C7(H5) major groove protons of flanking residues (Figure 2A, peaks K and L, respectively).

The intensities of NOE cross peaks originating from the ϵC6(H6) resonance are clearly unusual. The ϵC6(H6)-C5(H1′) and ϵC6(H6)-C7(H5) cross peaks are weaker than analogous connectivities observed on the sequence (Figure 2A, peaks J and B, respectively); at the same time, intensity

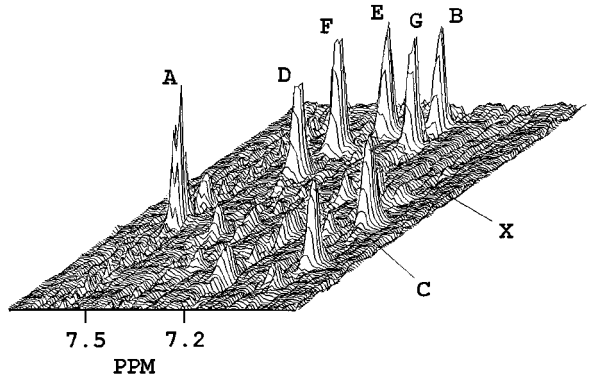


FIGURE 3: Stacked plot portion of a NOESY (50 ms mixing time) spectrum recorded in D_2O buffer, pH 6.8, 25 °C, showing distance interactions between base and H1′/cytidine H5 proton regions. Labeled peaks are assigned as follows: A, C1(H6)-C1(H5); B, C5-(H6)-C5(H5); C, ϵC6(H6)-ϵC6(H5); D, C7(H6)-C7(H5); E, C11(H6)-C11(H5); F, C13(H6)-C13(H5); G, C21(H6)-C21(H5), and X, ϵC6-(H6)-ϵC6(H1′).

of the ϵC6(H6)-ϵC6(H1′) NOE is particularly strong (Figure 2A, peak labeled C6). Comparison of NOE peak intensities is more reliable at short mixing times when spin diffusion effects are minimized. In the 50 ms mixing time NOESY spectrum, the intensity of the ϵC6(H6)-ϵC6(H1′) interaction is comparable to the deoxycytidine H5/H6 cross peaks (Figure 3, peaks X and A−G, respectively), establishing a distance shorter than 3 Å between the H6 and H1′ protons of the adduct.

Assignment of the H2′, H2″, H3′ and H4′ sugar protons was based on the NOESY (50 and 300 ms mixing time), COSY, DQF-COSY, and HOHAHA spectra following standard procedures (Hare *et al.*, 1983; Wüthrich, K., 1986; van de Ven & Hilbers, 1988). Chemical shift values of the nonexchangeable protons detected in D_2O buffer, pH 6.8, 25 °C, are listed in Table 1.

Assignment of the Exchangeable Protons. The imino and base/amino regions (6.0−14.5 ppm) of the one-dimensional proton spectrum in H_2O buffer, pH 6.8, 5 °C, is shown in Figure 2S, Supporting Information. Nine partially resolved signals are observed between 12.0 and 14.0 ppm, with no resonance detected between 9.0 and 12.0 ppm.

Assignment of the exchangeable protons followed the analysis of a NOESY (150 ms mixing time) spectrum in H_2O buffer, pH 6.8, 5 °C. An expanded contour plot showing distance connectivities between the imino (11.8−14 ppm) and base/amino (6.0−8.5 ppm) proton regions is shown in Figure 4A. A·T base pair formation was monitored readily by the presence of strong NOE cross peaks between thymidine imino and deoxyadenosine H2 protons. The independent assignment of all deoxyadenosine H2 protons from the analysis of NOESY spectra recorded in D_2O allowed the identification of these NOE peaks for the T3·A20, T9·A14, T15·A8, and T19·A4 base pairs of the duplex (Figure 4A, peaks A−D, respectively). Similarly, G·C base pairs were identified by the presence of distance connectivities between deoxyguanosine imino proton and the hydrogen-bonded and non-hydrogen-bonded deoxycytidine amino protons. We detected such interactions across the G2·C21, G10·C13, G12·C11, G16·C7, and G18·C5 base pairs of the sequence (Figure 4A, peaks E, E′; F, F′; G, G′; H, H′; and I, I′, respectively).

We have observed NOE interactions between G2(N1H)-A20(H2), T3(N3H)-A4(H2), G18(N1H)-A4(H2), G16(N1H)-

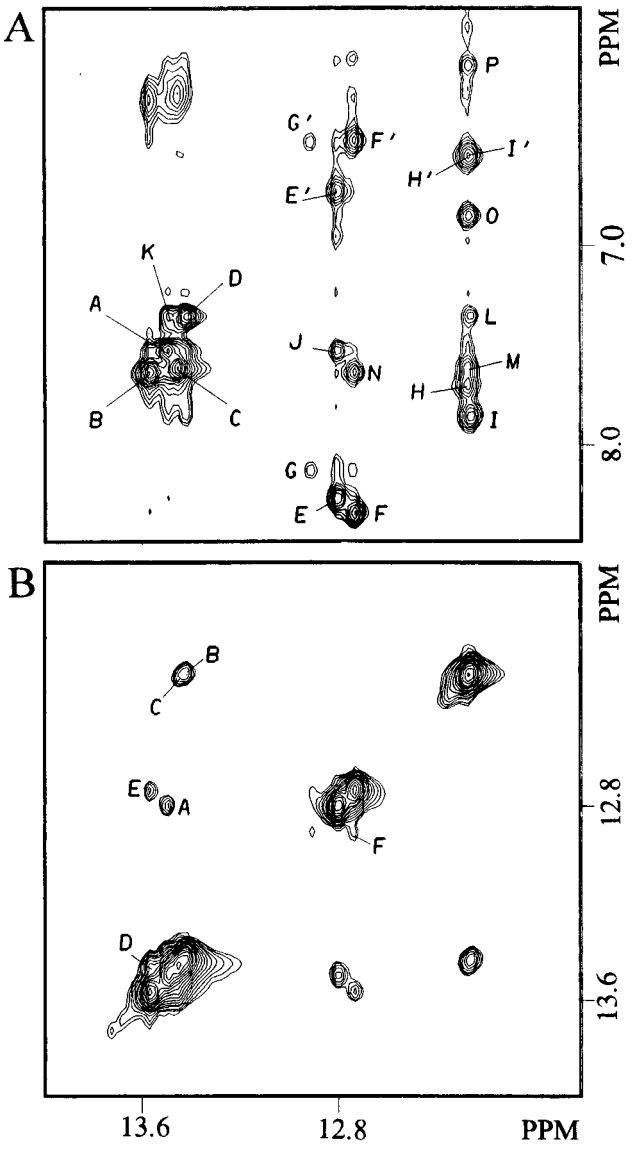


FIGURE 4: Counterplot portions of a NOESY (150 ms mixing time) spectrum recorded in H_2O buffer, pH 6.8, 5 °C, depicting distance interactions between (A) imino and base/amino proton regions and (B) symmetrical imino proton region. Labeled peaks are assigned as follows: (A) A, T3(N3H)-A20(H2); B, T9(N3H)-A14(H2); C, T15(N3H)-A8(H2); D, T19(N3H)-A4(H2); E, G2(N1H)-C21(N4H)$_{hb}$; E′, G2(N1H)-C21(N4H)$_{nhb}$; F, G10(N1H)-C13(N4H)$_{hb}$; F′, G10(N1H)-C13(N4H)$_{nhb}$; G, G12(N1H)-C11(N4H)$_{hb}$; G′, G12(N1H)-C11(N4H)$_{nhb}$; H, G16(N1H)-C7(N4H)$_{hb}$; H′, G16(N1H)-C7(N4H)$_{nhb}$; I, G18(N1H)-C5(N4H)$_{hb}$; I′, G18(N1H)-C5(N4H)$_{nhb}$; J, G2(N1H)-A20(H2); K, T3(N3H)-A4(H2); L, G18(N1H)-A4(H2); M, G16(N1H)-A8(H2); N, G10(N1H)-A14(H2); O, G16(N1H) and/or G18(N1H)-ϵC6(H7) and/or ϵC6(H8); P, G16(N1H) and/or G18(N1H)-ϵC6(H5). (B) A, G2(N1H)-T3(N3H); B, T19(N3H)-G18(N1H); C, G16(N1H)-T15(N3H); D, T15(N3H)-T9(N3H); E, T9(N3H)-G10(N1H), and F, G12(N1H)-G10(N1H). (N4H)$_{hb}$ and (N4H)$_{nhb}$ refer to the hydrogen-bonded and non-hydrogen-bonded cytidine amino protons, respectively.

A8(H2), and G10(N1H)-A14(H2) protons of neighboring base pairs (Figure 4A, peaks J−N, respectively). Base pair stacking was also detected between the G2(N1H)-T3(N3H), G18(N1H)-T19(N3H), G16(N1H)-T15(N3H), T15(N3H)-T9(N3H), T9(N3H)-G10(N1H), and G10(N1H)-G12(N1H) imino protons of the sequence (Figure 4B, peaks A−F, respectively).

Unlike the case of the ϵC·dA duplex [Korobka *et al.*, 1996 (companion paper)], the aromatic exocyclic protons exhibited

Table 1: Nonexchangeable Proton Chemical Shifts[a] in D_2O[b] at 25 °C

	H8/H6	H5/CH_3/H2	H1′	H2′	H2″	H3′	H4′
C1	7.61	5.86	5.74	1.99	2.38	4.68	4.05
G2	7.95		5.96	2.66	2.76	4.95	4.35
T3	7.22	1.50	5.68	2.04	2.39	4.87	4.17
A4	8.24	7.42	6.18	2.67	2.85	5.01	4.40
C5	7.24	5.22	5.78	2.00	2.48	4.73	4.19
ϵC6[c]	7.02	6.13	5.88	2.42	2.47	4.81	4.20
C7	7.47	5.53	5.39	2.12	2.36	4.82	4.22
A8	8.27	7.65	6.24	2.68	2.90	4.99	4.39
T9	7.06	1.41	5.71	1.92	2.31	4.82	4.12
G10	7.81		5.89	2.55	2.65	4.94	4.34
C11	7.36	5.28	6.13	2.17	2.17	4.45	4.02
G12	7.91		5.93	2.59	2.75	4.81	4.22
C13	7.44	5.38	5.65	2.13	2.45	4.87	4.19
A14	8.32	7.68	6.26	2.70	2.92	5.02	4.37
T15	7.03	1.42	5.68	1.85	2.27	4.82	4.25
G16	7.78		5.91	2.53	2.66	4.93	4.34
T17	7.11	1.49	5.70	2.03	2.30	4.78	4.34
G18	7.81		5.92	2.59	2.74	4.88	4.34
T19	7.22	1.35	5.68	2.07	2.42	4.86	4.34
A20	8.26	7.58	6.19	2.67	2.82	5.01	4.39
C21	7.25	5.32	5.63	1.83	2.25	4.76	4.13
G22	7.86		6.11	2.56	2.34	4.63	4.14

[a] Values in parts per million relative to TSP at 0.0 ppm. [b] 10 mM buffer phosphate, 50 mM NaCl, pH 6.8. [c] ϵC6(H7) and ϵC6(H8) protons resonate at 6.80 and 6.90 ppm, respectively.

distance connectivities with exchangeable protons of flanking G·C base pairs. NOE peaks were detected between G18-(N1H) and/or G16(N1H) and the ϵC6(H5), as well as the ϵC6(H7) and/or ϵC6(H8) exocyclic protons (Figure 4A, peaks P and O, respectively). ϵC6(H7) and ϵC6(H8) also exhibited distance connectivities to the hydrogen-bonded and non-hydrogen-bonded amino protons of C5 and C7 (peaks A−D respectively, Figure 3S, Supporting Information).

The imino proton of T17, which is opposite the exocyclic lesion, could not be accounted for in the one-dimensional spectrum in H_2O buffer, pH 6.8, 5 °C (Figure 2S, Supporting Information), nor could a cross peak be assigned to this proton in the NOE (150 ms mixing time) experiment; consequently, it was not identified. A plausible explanation is that T17(N3H) is too broad due to its exchange with water protons at pH 6.8. To reduce the water exchange rate we performed several NMR experiments in H_2O buffer under acidic conditions. A new broad resonance centered at 11.9 ppm and weak broad signal at 11.2 ppm were observed in the one-dimensional spectrum recorded in H_2O buffer, pH 5.0, 5 °C (Figure 4S, Supporting Information). NOESY experiments in H_2O buffer, pH 5.0, 2 °C, failed to detect any cross peaks originated from these new signals precluding their assignment. Further reduction of the pH did not sharpen the new resonances and produced a generalized broadening of other imino proton signals (data not shown). Chemical shifts values of the exchangeable protons detected in H_2O buffer, pH 6.8, 5 °C, are listed in Table 2.

Chemical Shifts. Assignment of the base and sugar protons on the ϵC·T and ϵC·dA [Korobka *et al.*, 1996 (companion paper)] duplexes, which differ only on the residue opposite the adduct, allowed us to identify characteristic variations in the proton chemical shifts. Differences in the chemical shift of base (purine H8/pyrimidine H6) and sugar H2′ major groove protons and of H1′, H2″ minor groove sugar protons observed between the duplexes are depicted graphically in Figure 5, A and B, respectively. Insignificant chemical shift

Table 2: Exchangeable Proton Chemical Shifts[a] in H_2O[b] at 5°C

	T(N3H)	G(N1H)	$C(N4H)_{hb}$[c]	$C(N4H)_{nhb}$[c]	A(H2)
C1·G22		nd[d]	8.09	6.97	
G2·C21		12.82	8.28	6.74	
T3·A20	13.50				7.54
A4·T19	13.42				7.36
C5·G18		12.27	7.87	6.56	
ϵC6·T17	na[e]				
C7·G16		12.29	7.72	6.54	
A8·T15	13.46				7.62
T9·A14	13.58				7.64
G10·C13		12.74	8.34	6.48	
C11·G12		12.91	8.13	6.50	

[a] Values in parts per million relative to TSP at 0.0 ppm. [b] 10 mM buffer phosphate, pH 6.8, 50 mM NaCl. [c] $C(N4H)_{hb}$ and $C(N4H)_{nhb}$ are the hydrogen-bonded and non-hydrogen-bonded cytidine amino protons, respectively. [d] nd, not detected. [e] Not assigned.

differences were detected for the base and sugar protons of the last three base pairs of the duplex. Closer to the lesion, we observed major differences for the exocyclic H7 and H8 protons (−0.59 and −0.41 ppm, respectively) as well as the H2′ and H1′ sugar protons of ϵC (Figure 5, panels A and B, respectively). Relatively minor shifts were detected for the ϵC6(H2″), C7(H6), C7(H1′), and C5(H1′) protons of residues at or near the lesion (Figure 5). On the complementary strand, significant shifts were limited to the G16(H1′), T15-(H2′) and T15(H2″) sugar protons, while smaller variations were seen on the G18(H8) and G18(H1′) protons. Among exchangeable protons, the only significant difference was observed for the hydrogen-bonded amino proton of C7, which moved 0.43 ppm upfield on the ϵC·T duplex.

Molecular Dynamics Simulations. The distance-refined structures starting from A- and B-form DNA models are shown in Figure 5S, Supporting Information. The root mean square (rms) deviation in the position of the atoms was reduced from 6.2 Å on the starting models to 1.5 Å between the refined structures, while the experimental restraints were fairly satisfied without major violations of the covalent geometry. Since assignment of the ϵC(H7,H8) protons was based on structural considerations, we interchanged their distance restraints and refined the initial models again. As a result, the rms deviations between inter-proton distances and NOE distance bounds on the central three-base pair segment increased 40%, supporting the original assignment of the exocyclic protons. Table 3 lists the violations of experimental restraints present on the initial and refined models as well as deviations from idealized nucleotide geometry.

Figure 6 is a stereoview of the three-dimensional model obtained after 47.5 ps of NMR-restrained molecular dynamics, starting from the B-form initial model, and Table 4 lists structural parameters. The duplex belongs to the B-form family, has all bases stacked into the helix and the sugar puckers are in the south (C2′-endo/C3′-exo) conformation. It is a regular right-handed helix with a slight perturbation at the C1·G22 end. The helical axis is bent toward the major groove by 27° at the lesion site and no major perturbation is observed on the sugar-phosphate backbone (Figure 6).

Figure 7 is a stereoview of the three central base pairs of the sequence seen from the major groove. The glycosidic torsion angle of the ϵC residue is 73°, which is within the *syn* range and close to the value of the initial model. The ϵC6(*syn*)·T17(*anti*) base pair is propeller-twisted by 19° without significant changes of inter-base pair distances. Flanking the lesion, Watson−Crick alignments are observed

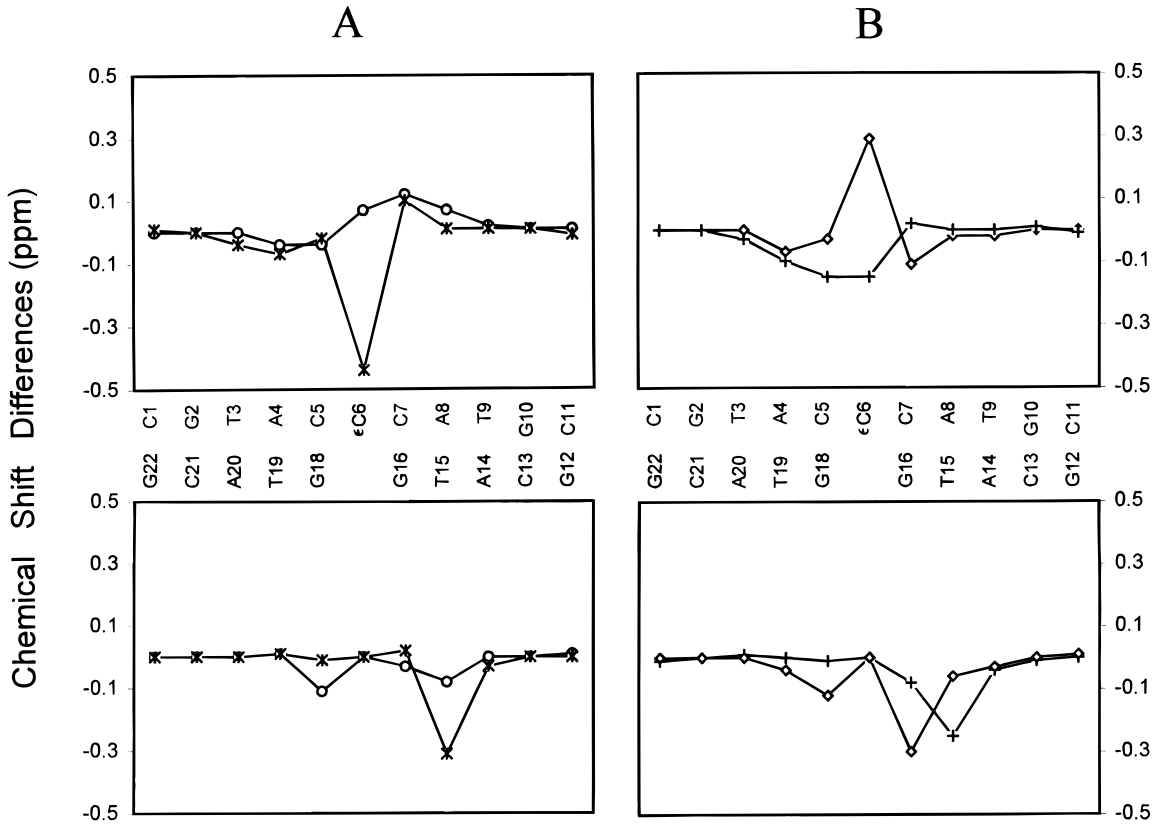


FIGURE 5: Differences of proton chemical shifts observed between the ϵC·dA [Korobka *et al.*, 1996 (companion paper)] and ϵC·T duplexes in D_2O buffer, pH 6.8, 25 °C. (A) major groove base (○) (purine H8/pyrimidine H6) and sugar H2′ (*) protons; (B) minor groove H1′ (◇) and H2″ (+) protons. Values are calculated as ϵC·dA minus ϵC·T chemical shifts. Differences between A17 and T17 protons are not considered.

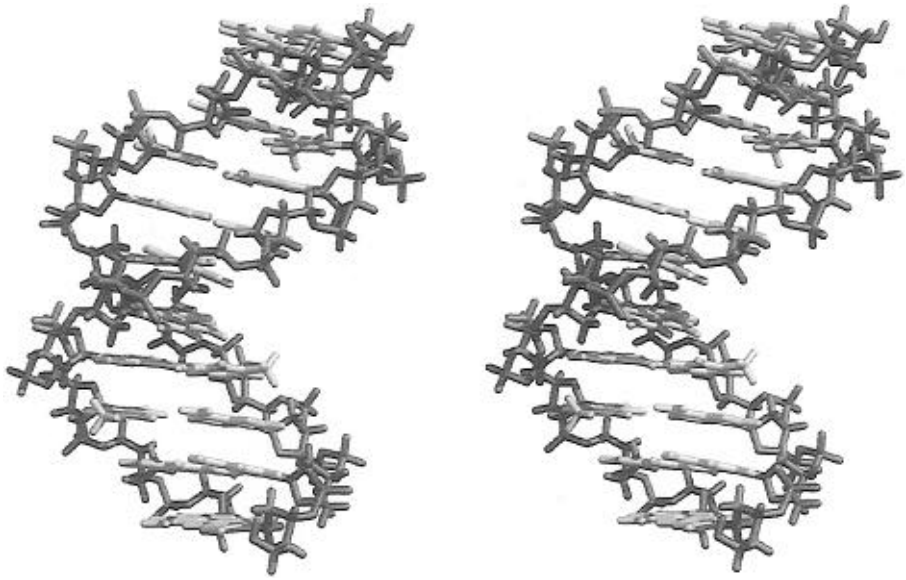


FIGURE 6: Stereoview of the three-dimensional structure after 47.5 ps of NMR-restrained molecular dynamics, starting from B-form DNA initial model. The figure depicts a slight bend of the helical axis.

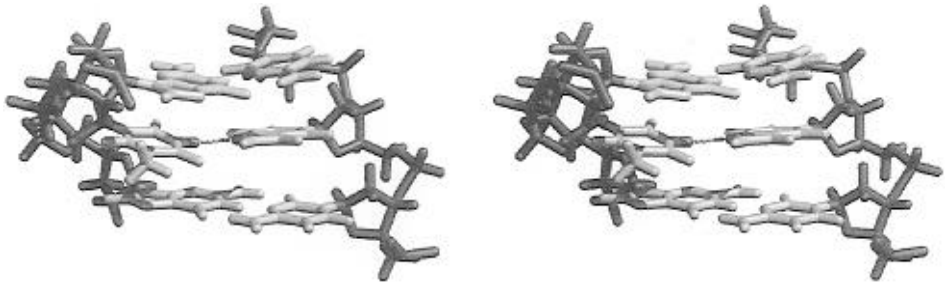


FIGURE 7: Stereoview of the central base pair segment of the refined duplex, seen from the major groove. The picture shows Watson−Crick alignment on flanking C5·G18 and C7·G16 base pairs. The label points to the strong hydrogen bond formed between ϵC6(N4) and T17(N3H) across the ϵC6·T17 base pair.

Table 3: Violations of Experimental Restraints and Idealized Covalent Geometry

RMS-Deviations from Experimental Inter-Proton Distances[a]				
	initial models		refined structures[b]	
	A-form	B-form	A-initial	B-initial
11-mer duplex[c]	0.65	0.36	0.26	0.23
central three-base pairs[d]	0.68	0.46	0.42	0.39
	R-factors[e]			
11-mer duplex (962 NOE-intensities)	0.680	0.459	0.069	0.069

RMS-Deviations from Idealized Covalent Geometry		
	refined structures[b]	
	A-initial	B-initial
bonds (Å)	0.006	0.006
bond angles (deg)	3.42	3.40
improper angles (deg)	0.18	0.17

[a] Values are given in angstroms. [b] A-initial and B-initial indicate that the refinement started from A- and B-form DNA initial structures, respectively. [c] Total of 216 distances. [d] Total of 59 distances. [e] 1/6 *R*-factor as defined by X-PLOR.

Table 4: Structural Parameters[a,b]

helical bend (deg)	27
ϵC6·T17 Alignment	
ϵC6 glycosidic angle (deg)	73
T17 glycosidic angle (deg)	−104
propeller twist (deg)	19
base pair stagger (Å)	0.79
base pair stretch (Å)	0.55
Base Pair Buckle	
C5·G18 (deg)	<1
C7·G16 (deg)	12
ϵC6·T17 Hydrogen Bonding	
ϵC6(N4)-T17(N3H) distance (Å)	2.05
ϵC6(N4)-T17(N3H)-T17(N3) angle (deg)	167
hydrogen bond energy (kcal/mol)[c]	-3.0

[a] Averaged values after refinement starting from A- and B-form DNA. [b] Parameters as defined by Curves. Values observed on canonical B-form DNA are: helical bend 0°; glycosidic torsion angle (pyrimidine) −118°; propeller twist, 3.7°; base pair stagger, 0.0 Å; base pair stretch, 0.0 Å; base pair buckle, 0° (Lavery & Sklenar 1989). [c] Calculated using the explicit hydrogen-bond energy term of X-PLOR without acceptor antecedent dependency. The energies of G16(N1H)-C7(N3) and G18(N1H)-C5(N3) hydrogen bonds are −2.7 and −3.2 kcal/mol, respectively.

on C5·G18 and C7·G16 base pairs; the latter exhibited a 12° buckle.

The distance between ϵC6(N4) and T17(N3) is 3.05 Å and the aromatic rings of these residues are coplanar, suggesting the formation of a hydrogen bond across the ϵC6(*syn*)·T17(*anti*) alignment (Figures 7 and 8). Stacking interactions are well-preserved between the ϵC6·T17 and C7·G16 base pairs and ϵC stacks partially with flanking C5·G18 base pair (Figure 8). Table 5 compares experimental distance bounds and inter-proton distances measured on the refined structure at the lesion site.

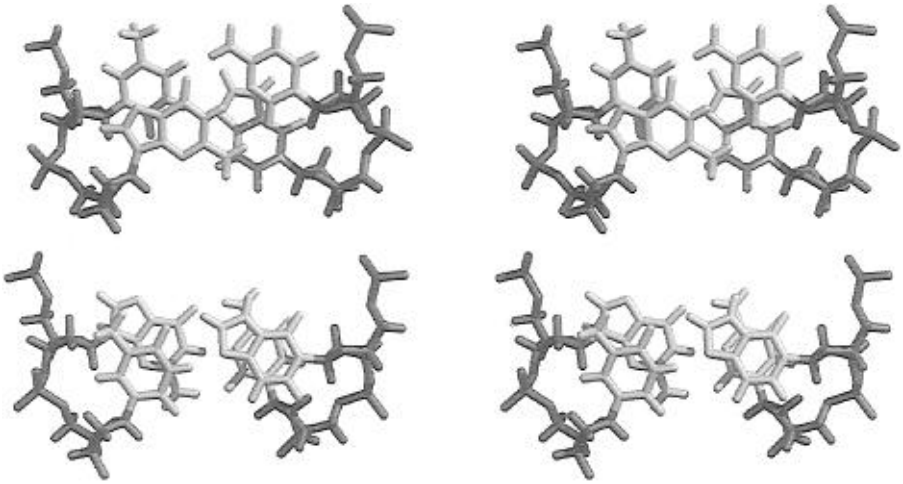


FIGURE 8: Stereoview depicting stacking interactions at the center of the helix between (top) C5·G18 and ϵC6·T17 and (bottom) ϵC6·T17 and C7·G16 base pairs, respectively.

Table 5: Comparison between Experimental and Refined Inter-Proton Distances (in Å)

	experimental distance bounds[a]	refined distances[b]
C5(H1′)-ϵC6(H6)	4.2−5.2	5.2
ϵC6(H1′)-C7(H6)	3.5−4.5	4.2
C5(H2′)-ϵC6(H8)	3.5−4.5	4.4
C5(H2″)-ϵC6(H8)	3.9−4.9	4.0
C5(H6)-ϵC6(H8)	4.5−5.5	4.5
ϵC6(H8)-C7(H6)	>5[c]	5.9
C5(H5)-ϵC6(H8)	4.3−5.3	4.6
ϵC6(H8)-C7(H5)	4.5−5.5	4.5
ϵC6(H6)-C7(H6)	4.6−5.6	4.8
C5(H5)-ϵC6(H7)	4.6−5.6	5.1
G16(N1H)-ϵC6(H7)	3.0−4.5	3.2
G16(N1H)-ϵC6(H8)	3.5−5.0	4.4
$C5(N4H)_{hb}$-ϵC6(H7)	3.0−4.5	3.6
$C5(N4H)_{nhb}$-ϵC6(H8)	3.5−5.0	5.2
G18(N1H)-ϵC6(H7)	nr	3.8
G18(N1H)-ϵC6(H8)	nr	4.7
$C5(N4H)_{hb}$-ϵC6(H8)	nr	4.9
$C5(N4H)_{nhb}$-ϵC6(H7)	nr	4.4
$C7(N4H)_{hb}$-ϵC6(H8)	nr	3.0
$C7(N4H)_{hb}$-ϵC6(H7)	nr	3.4
$C7(N4H)_{nhb}$-ϵC6(H8)	nr	3.2
$C7(N4H)_{nhb}$-ϵC6(H7)	nr	4.8
G16(N1H)-ϵC6(H5)	nr	4.7
G18(N1H)-ϵC6(H5)	nr	4.5

[a] Derived from NOE buildup curves. [b] Averaged values measured on the refined models. [c] NOE peaks observed only at long mixing time. nr (not restrained). Distance interactions detected in the NOESY spectrum recorded in H_2O buffer (Figures 4A and 3S, Supporting Information) but not restrained during refinement of the structures. $(N4H)_{hb}$ and $(N4H)_{nhb}$ refer to the hydrogen-bonded and non-hydrogen-bonded cytidine amino protons, respectively.

DISCUSSION

NMR Spectra. The complete assignment of the NMR resonances to a single set of base and sugar protons and the sharp lines observed for the exchangeable and nonexchangeable protons imply the existence of a single solution structure for the ϵC·T duplex in the NMR time scale. The sequential distance connectivities detected between base (purine H8/pyrimidine H6) and H1′, H2′, H2″, and H3′ sugar and thymine methyl protons, as well as deoxyadenosine H2 and H1′ sugar protons, establish a right-handed helix with all sugar puckers in the south (C2′-endo/C3′-exo) conformation and, with the exception of ϵC, glycosidic torsion angles oriented in the *anti* range.

The distance connectivities observed between thymidine imino and deoxyadenosine H2 protons and deoxyguanosine imino and hydrogen-bonded and non-hydrogen-bonded deoxycytidine amino protons (Figure 4A), confirm Watson−Crick base pair alignment for all A·T and G·C base pairs of the ϵC·T duplex. Furthermore, the sequential NOE interactions between imino protons and between imino and deoxyadenosine H2 protons (Figure 4, A and B, respectively) indicate base pair stacking throughout the duplex.

Alignment of the ϵC6·T17 Base Pair. An unusually strong NOE peak was detected between the ϵC6(H6) and its own H1′ sugar proton. On the shortest mixing time NOESY spectrum, for which spin diffusion effects are minimal, the intensity of this cross peak was similar to that of deoxycytosine H5-H6 NOE peaks, which represent a fixed 2.45 Å inter-proton distance (Figure 3). This short distance between these protons of the adduct establish a *syn* conformation adopted by the ϵC residue around the glycosidic torsion angle (van de Ven & Hilbers, 1988). In contrast with the ϵC·dA duplex, no stearic clashes would be expected if both the ϵC6 and T17 are coplanar and have their glycosidic torsion angles in *anti* conformation. Thus, the preference of the adduct for the *syn* orientation must be related to other energetic factors, including a gain in stability of the duplex due to hydrogen bonding. The imino proton of T17 opposite the exocyclic adduct was not observed in H_2O buffer, pH 6.8. Although a new broad signal, centered at 11.9 ppm, appeared under acidic conditions it could not be assigned due to water exchange (peak A, Figure 4S, Supporting Information). As a result, there is no spectroscopic evidence that T17(N3H) is involved in hydrogen bonding. Clearly, if a hydrogen bond is formed across the ϵC6·T17 base pair, it does not prevent exchange of the imino proton of T17 with water.

NMR-Refined Solution Structure. After restrained molecular dynamics the initial A- and B-form DNA models converged to structures within the B-form family that have an rms deviation of 1.5 Å in the atom position (Figure 5S, Supporting Information). Both models are in agreement with the structural features established by NMR, suggesting that the rms deviation between them reflects the precision of our data to define an averaged structure.

When the glycosidic torsion angle of the ϵC residue is *syn*, a ϵC6(*syn*)·T17(*anti*) base pair can be incorporated readily into a right-handed helix. It causes a small bend on

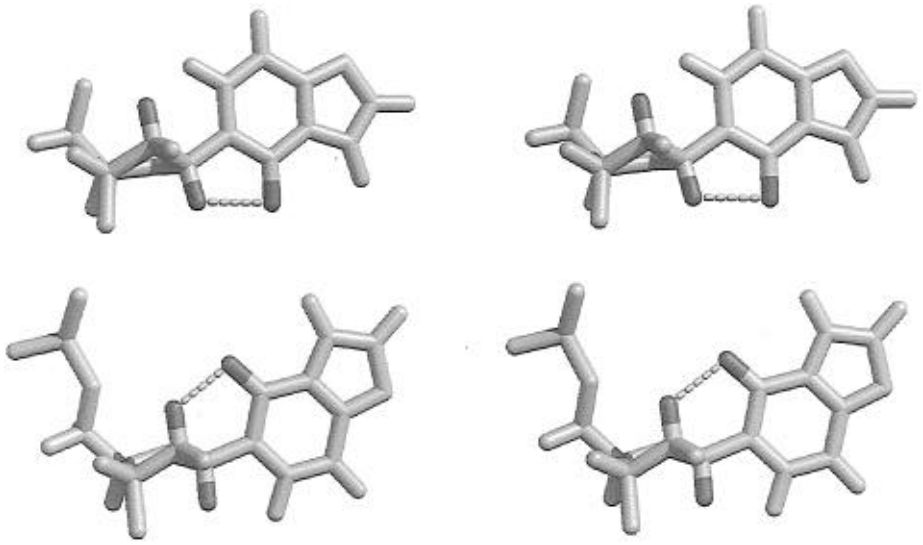


FIGURE 9: Stereoview showing the orientation adopted by the exocyclic adduct around the glycosidic torsion angle. The *anti* conformation observed in the ϵC·dA duplex [Korobka *et al.*, 1996 (companion paper)] (top); *syn* conformation present in the ϵC·T duplex (bottom). Dashed lines connect ϵC(O2) to the ϵC(H1′) (top) and the ϵC(H2′) (bottom) protons, respectively.

the helical axis with no significant perturbation of the sugar-phosphate backbone. The exocyclic adduct and its partner on the opposite strand are coplanar and stack with flanking base pairs (Figure 8).

An interesting feature of the structure is the presence of a well formed hydrogen bond between ϵC6(N4) and T17(N3H) stabilizing the ϵC6(*syn*)·T17(*anti*) alignment that may explain the conformational preference of the adduct for the *syn* orientation (Figures 7 and 8). The energy of this hydrogen bond, calculated by the explicit hydrogen bond term of the X-PLOR force field, is −3.0 kcal/mol. This value is similar to the energy of hydrogen bonds formed by the imino protons of flanking C5·G18 and C7·G16 base pairs (Table 4). The presence of this hydrogen bond in the refined structure seems to argue with the fact that NMR experiments failed to detect the T17(N3H) proton signal due to his fast exchange with water, while imino protons of flanking C·G base pairs were clearly observed. A plausible explanation is that formation of a single hydrogen bond in the ϵC6(*syn*)·T17(*anti*) base pair would not be sufficient to prevent base pair opening and subsequent water exchange. On flanking C·G base pairs, three hydrogen bonds are formed in the Watson−Crick alignment, reducing the opening rate of the base pairs and allowing NMR detection of the imino proton signals. The fact that inter-proton distances measured on the central segment of the model are within experimental bounds (Table 5) further supports the duplex structure and ϵC6(*syn*)·T17(*anti*) base pair alignment presented here.

The model also explains chemical shift differences observed between ϵC·T and ϵC·dA duplexes [Korobka *et al.*, 1996 (companion paper)]. The proton chemical shifts observed for the d(C1-G2-T3-A4)·d(T19-A20-C21-G22) and d(T9-G10-C11)·d(G12-C13-A14) terminal segments of the sequence are almost identical, limiting the structural differences between duplexes to the four central base pairs. On the modified strand, the largest chemical shift differences were observed for proton signals of the exocyclic lesion. Since the electronic distribution of the adduct is the same in both duplexes, these differences must originate from a distinct magnetic environment. The changes on the chemical shift of ϵC6(H1′) and ϵC6(H2′) sugar protons by 0.29 and −0.44 ppm, respectively (Figure 5), reflect the different glycosidic torsion angle adopted by the adduct. In the ϵC·dA duplex, where ϵC is *anti*, its H1′ proton is close to the ϵC6(O2) and almost coplanar with the aromatic moiety. In this orientation, the ϵC6(H1′) proton is deshielded by ring current effects resonating 0.29 ppm downfield. In the ϵC·T duplex, when ϵC adopts the *syn* conformation, the H2′ sugar proton is deshielded by the same ring currents. The stereoview in Figure 9 shows the relative orientation of the ϵC residue in the NMR-refined structures for both duplexes.

Deshielding of the exocyclic H7 and H8 protons in the ϵC·T duplex by 0.59 and 0.41 ppm, respectively, also correlates with the different glycosidic torsion angles of the adducts. In the *anti* orientation observed in the ϵC·dA duplex, the exocyclic H7 and H8 protons are positioned toward the center of the helix, intercalated between A17 and G18 and shielded by ring currents of these purines [Korobka *et al.*, 1996 (companion paper)]. When ϵC is *syn*, these protons are located on the major groove of the helix where shielding by adjacent pyrimidine residues is minimal. Minor changes are observed on the chemical shifts of the ϵC(H6) and ϵC(H5) resonances, which undergo similar translocation that the ϵC(H7) and ϵC(H8) protons. When ϵC is *syn*, ϵC(H6) proton is located in the minor groove and ϵC(H5) is directed toward the center of the helix but neither of them stacks with adjacent aromatic rings (Figure 8). When ϵC is *anti*, both protons are in the major groove of the helix showing limited overlap with flanking pyrimidine rings [Korobka *et al.*, 1996 (companion paper)]. As a result, similar ring current effects are experienced by these protons in both duplexes, explaining their close chemical shift values. Similarly, the downfield shift observed for the hydrogen-bonded C7(N4H) proton signal on the ϵC·T duplex may be explained as the result of increased shielding by the ϵC ring. In the ϵC·dA duplex, partial intercalation of A17 increases the separation between ϵC6 and C7, diminishing this effect.

On the complementary strand, electronic differences between T17 and A17 limit the possibility of correlating the chemical shift differences observed among protons of the G18 and G16 residues with structural features of the ϵC·T and ϵC·dA duplexes. We have no explanation for the

chemical shift differences observed in the T15(H2′,H2″) resonances.

Biological Implications. The hydrogen bonding potential of deoxycytidine is severely reduced when an etheno bridge is incorporated into the pyrimidine ring. Nevertheless, DNA polymerases permit translesional synthesis past ϵC *in vitro* and in cells. The duplex structures described in this paper and the companion study (Korobka *et al.*, 1996) represent possible intermediates in this process.

With ϵC in the *syn* conformation, a hydrogen bond forms between T(N3H) proton and ϵC(N4), maintaining a coplanar alignment and facilitating misincorporation of thymidine. The ϵC(*syn*)·T(*anti*) base pair can be accommodated with minimal distortion in a B-form DNA helix. Following the second round of replication, this intermediate would lead to ϵC→A transversions (Simha *et al.*, 1991; Palejwala *et al.*, 1993; Basu *et al.*, 1993; Moriya *et al.*, 1994).

When positioned opposite deoxyadenosine, ϵC assumes the *anti* conformation; steric factors preclude a coplanar alignment. Nevertheless, dA is incorporated preferentially opposite ϵC *in vitro* (Zhang *et al.*, 1995a; Shibutani, personal communication) and both dA and dG are introduced opposite this lesion in cells (Moriya *et al.*, 1994). Thus, a partially intercalated structure appears to promote translesional synthesis past the lesion.

One might expect that insertion of a dNMP opposite the lesion and/or chain extension from the 3′ primer terminus would be impeded by the absence of hydrogen bonding. Indeed, rates of these reactions for TMP and dAMP are several orders of magnitude lower than for the unmodified base (Zhang *et al.*, 1995a) but are significantly higher than those for other DNA bases. We conclude that stacking between an incoming nucleotide and a template base can support translesional synthesis past a damaged base; however, the process proceeds at the expense of fidelity afforded by Watson−Crick base pairs.

ACKNOWLEDGMENT

We thank Mr. Robert Rieger and Ms. Cecilia Torres for preparing modified oligonucleotides and Mr. Erich Bremer for assisting in the preparation of three-dimensional figures. Initial NMR data at 500 MHz were collected at the College of Physicians and Surgeons of Columbia University. 600 MHz experiments were recorded in the NMR Facility of the State University of New York at Stony Brook.

SUPPORTING INFORMATION AVAILABLE

Figures showing the one-dimensional spectra recorded in D_2O buffer, pH 6.8, 25 °C (Figure 1S), in H_2O buffer, pH 6.8, 5 °C (Figure 2S), and H_2O buffer, pH 5.0, 5 °C (Figure 4S); counterplot of the symmetrical base and amino proton region of a NOESY (150 ms mixing time) spectrum recorded in H_2O buffer, pH 6.8, 5 °C (Figure 3S); three-dimensional structures of the distance-refined models starting from A- and B-form DNA (Figure 5S) (6 pages). Ordering information is given in any current masthead page.

REFERENCES

Barbin, A., & Bartsch, H. (1986) in *The Role of Cyclic Nucleic Acid Adducts in Carcinogenesis and Mutagenesis*, IARC Scientific Publications 70, pp 345−358, International Agency for Research on Cancer, Lyon, France.

Basu, A. K., Wood, M. L., Niedernhofer, L. J., Ramos, L. A., & Essigmann, J. M. (1993) *Biochemistry 32*, 12793−12801.

Brunger, A. (1993) X-PLOR Version 3.1, A System for X-Ray Crystallography and NMR, Yale University Press, New Haven, CT.

el Ghissassi, F., Barbin, A., Nair, J., & Bartsch, H. (1995) *Chem. Res. Toxicol. 8*, 278−283.

Hare, D. R., Wemmer, D. E., Chou, S. H., Drobny, G & Reid, B. R. (1983) *J. Mol. Biol. 171*, 319−336.

Kunkel, T. A. (1984) *Proc. Natl. Acad. Sci. U.S.A. 81*, 1494−1498.

Lavery, R., & Sklenar, H. (1988) *J. Biomol. Struct. Dyn. 6*, 63−91.

Lavery, R., & Sklenar, H. (1989) *J. Biomol. Struct. Dyn. 6*, 655−667.

Loechler, E. L., Green, C. L., & Essigmann, J. M. (1984) *Proc. Natl. Acad. Sci. U.S.A. 81*, 6271−6275.

Moriya, M., Zhang, W., Johnson, F., & Grollman, A. P. (1994) *Proc. Natl. Acad. Sci. U.S.A. 91*, 11899−11903.

Nair, J., Barbin, A., Guichard, Y., & Bartsch, H. (1995) *Carcinogenesis 16*, 613−617.

Nath, R. G., & Chung, F.-L. (1994) *Proc. Natl. Acad. Sci. U.S.A. 91*, 7491−7495.

Nilges, M., Habazettl, J., Brunger, A. T., & Holak, T. A. (1991) *J. Mol. Biol. 219*, 499−510.

Palejwala, V. A, Rzepka, R. W., & Humayun, M. Z. (1993) *Biochemistry 32*, 4112−4120.

Plateau, P., & Gueron, M. (1982) *J. Am. Chem. Soc. 104*, 7310.

Rabkin, S. D., & Strauss, B. S. (1984) *J. Mol. Biol. 178*, 569−594.

Simha, D., Palejwala, V. A., & Humayun, M. Z. (1991) *Biochemistry 30*, 8727−8735.

Singer, B., & Spengler, S. J. (1986) in *The Role of Cyclic Nucleic Acid Adducts in Carcinogenesis and Mutagenesis*, IARC Scientific Publications 70, pp 345−358, International Agency for Research on Cancer, Lyon, France.

Snow, E. T., Foote, R. S., & Mitra, S. (1984) *J. Biol. Chem. 259*, 8095−8100.

States, D. J., Habekorn, R. A., & Ruben, D. J. (1982) *J. Magn. Reson. 48*, 286−292.

van de Ven, F. J. M & Hilbers, C. W. (1988) *Eur. J. Biochem. 178*, 1−38.

Wüthrich, K. (1986) in *NMR of Proteins and Nucleic Acids*, John Wiley, New York.

Yip, P., & Case, D. (1989) *J. Magn. Reson. 83*, 643−648.

Zhang, W., Johnson, F., Grollman, A. P., & Shibutani, S. (1995a) *Chem. Res. Toxicol. 8*, 157−163.

Zhang, W., Rieger, R., Iden, C., & Johnson, F. (1995b) *Chem. Res. Toxicol. 8*, 148−156.

BI9605705