Differential Base Stacking Interactions Induced by Trimethylene Interstrand DNA Cross-Links in the 5′-CpG-3′ and 5′-GpC-3′ Sequence Contexts

Hai Huang, Patricia A. Dooley,† Constance M. Harris, Thomas M. Harris, and Michael P. Stone*

Department of Chemistry, Center in Molecular Toxicology, and Center for Structural Biology, Vanderbilt University, Nashville, Tennessee 37235

Received July 4, 2009

Synthetically derived trimethylene interstrand DNA cross-links have been used as surrogates for the native cross-links that arise from the 1,N^2-deoxyguanosine adducts derived from α,β-unsaturated aldehydes. The native enal-mediated cross-linking occurs in the 5′-CpG-3′ sequence context but not in the 5′-GpC-3′ sequence context. The ability of the native enal-derived 1,N^2-dG adducts to induce interstrand DNA cross-links in the 5′-CpG-3′ sequence as opposed to the 5′-GpC-3′ sequence is attributed to the destabilization of the DNA duplex in the latter sequence context. Here, we report higher accuracy solution structures of the synthetically derived trimethylene cross-links, which are refined from NMR data with the AMBER force field. When the synthetic trimethylene cross-links are placed into either the 5′-CpG-3′ or the 5′-GpC-3′ sequence contexts, the DNA duplex maintains B-DNA geometry with structural perturbations confined to the cross-linked base pairs. Watson−Crick hydrogen bonding is conserved throughout the duplexes. Although different from canonical B-DNA stacking, the cross-linked and the neighbor base pairs stack in the 5′-CpG-3′ sequence. In contrast, the stacking at the cross-linked base pairs in the 5′-GpC-3′ sequence is greatly perturbed. The π-stacking interactions between the cross-linked and the neighbor base pairs are reduced. This is consistent with remarkable chemical shift perturbations of the C^5 H5 and H6 nucleobase protons that shifted downfield by 0.4−0.5 ppm. In contrast, these chemical shift perturbations in the 5′-CpG-3′ sequence are not remarkable, consistent with the stacked structure. The differential stacking of the base pairs at the cross-linking region probably explains the difference in stabilities of the trimethylene cross-links in the 5′-CpG-3′ and 5′-GpC-3′ sequence contexts and might, in turn, account for the sequence selectivity of the interstrand cross-link formation induced by the native enal-derived 1,N^2-dG adducts.

Introduction

Acrolein, crotonaldehyde, and 4-hydroxynonenal (4-HNE) are endogenous byproducts of lipid peroxidation, arising as a consequence of oxidative stress (*1–4*). Acrolein and crotonaldehyde exposures also occur from exogenous sources, for example, cigarette smoke (*5*) and automobile exhaust (*6*). These enals react with DNA nucleobases to give exocyclic adducts; they also react with proteins (*7*). The addition of these enals to dG involves Michael addition of the N^2-amine to give N^2-(3-oxopropyl)-dG adducts, followed by cyclization of N1 with the aldehyde, yielding the corresponding 1,N^2-dG products **1**. When placed opposite 2′-deoxycytosine in duplex DNA, these exocyclic 1,N^2-dG adducts typically undergo ring opening to the N^2-(3-oxopropyl)-dG aldehydes **2** [for a recent review, see Minko et al. (*8*)]. The latter induce N^2-dG:N^2-dG interstrand cross-links in the 5′-CpG-3′ sequence context but not in the 5′-GpC-3′ sequence context (*9–11*) (Scheme 1) [for a recent review, see Stone et al. (*12*)]. These cross-links exist as equilibrium mixtures of carbinolamines and imines in which the carbinolamines predominate (*12–14*).

The solution structure of the acrolein-derived (*R*)-γ-hydroxytrimethylene cross-link has been determined in the 5′-CpG-3′ sequence context (*15*). The carbinolamine was the major cross-link species, constituting 80−90% of the cross-link. The imino resonances of the cross-linked pairs were observed at 65 °C, indicating that this cross-link thermally stabilizes the DNA duplex with respect to melting. Additionally, for the 5′-neighbor G•C base pair, the guanine imino resonance remained sharp at 55 °C but broadened at 65 °C. In contrast, for the 3′-neighbor A•T base pair, the thymine imino resonance was severely broadened at 55 °C. Structural refinement indicated that the (*R*)-γ-hydroxytrimethylene linkage maintained the two cross-linked base pairs with minimal structural perturbations. The (*R*)-γ-hydroxytrimethylene linkage was located in the minor groove. The two guanine amino nitrogen atoms were in the gauche conformation with respect to the linkage, which maintained Watson−Crick hydrogen bonding of the cross-linked base pairs. The anti conformation of the carbinol hydroxyl group with respect to C^α of the tether minimized steric interactions and, more importantly, allowed the formation of a hydrogen bond between the carbinol hydroxyl and the cytosine O^2 atom located in the 5′-neighboring G•C base pair. It was proposed that this hydrogen bond may, in part, explain the thermal stability of this carbinolamine interstrand cross-link and the stereochemical preference for the (*R*)-configuration of the cross-link (*15*).

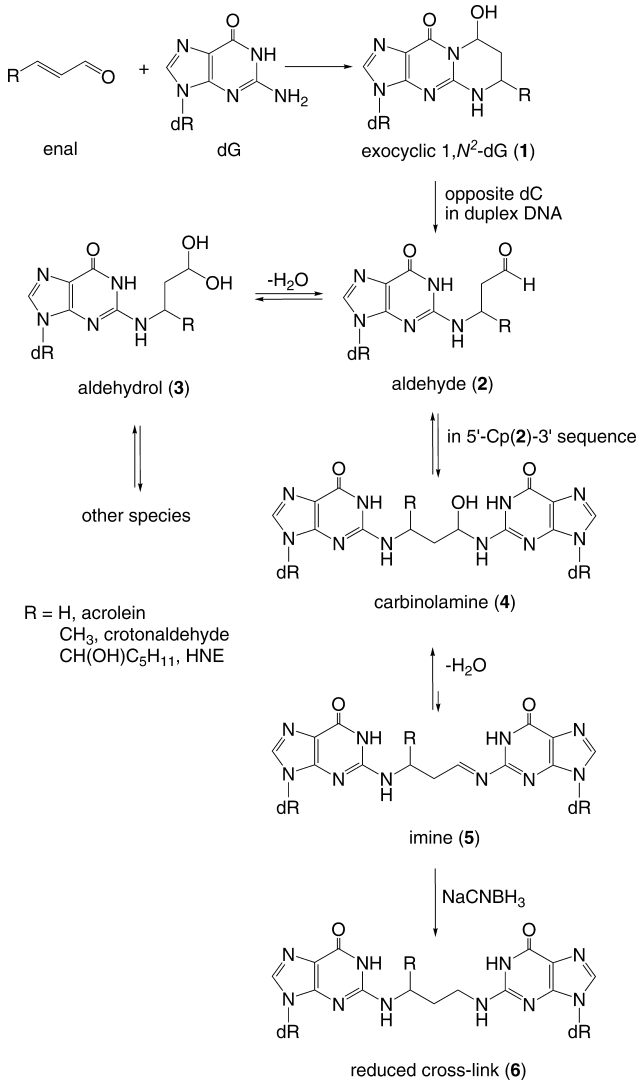
Previously, DNA duplexes containing the 5′-CpG-3′ or 5′-GpC-3′ sequence contexts were cross-linked with fully reduced trimethylene tethers (Scheme 2). This simulated the chemically

* To whom correspondence should be addressed. Tel: 615-322-2589. Fax: 615-322-7591. E-mail: michael.p.stone@vanderbilt.edu.

† Current address: Bard College at Simon's Rock, Great Barrington, MA 01230.

10.1021/tx900225c CCC: $40.75
Published on Web 10/09/2009

Scheme 1. Formation of the Native Carbinolamine N^2-dG:N^2-dG Interstrand Cross-Links by Enal-Derived Exocyclic 1,N^2-dG Adducts in the 5′-CpG-3′ Sequence in Duplex DNA[a]

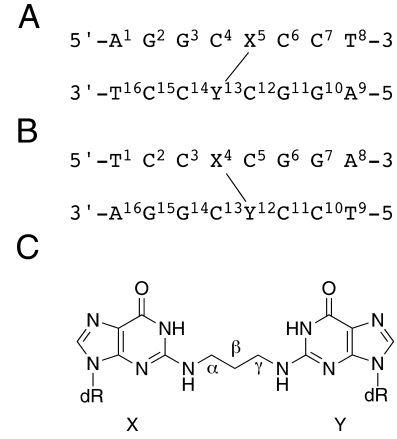


[a] The native cross-link is in equilibrium with the corresponding imine and can be reduced to the chemically stable trimethylene cross-link.

labile acrolein-derived (*R*)-carbinolamine cross-link and structurally similar cross-links arising from other enals (*16, 17*). The preferential formation of the interstrand cross-link in the 5′-CpG-3′ sequence as compared to the 5′-GpC-3′ sequence was attributed to the longer distance between the two N^2-dG atoms in the 3′-direction than in the 5′-direction, accompanied by greater destabilization of the DNA duplex by the cross-link in the 5′-GpC-3′ sequence (*16, 17*). However, the structural refinements of the two fully reduced cross-linked duplexes (*16, 17*) indicated greater perturbations of Watson−Crick hydrogen bonding of the cross-linked base pairs than was observed for the acrolein-derived (*R*)-γ-hydroxytrimethylene cross-link in the 5′-CpG-3′ sequence (*15*).

Because the reduced trimethylene interstrand DNA cross-links provide important and chemically stable surrogates for the labile native carbinolamine cross-links that arise from the 1,N^2-deoxyguanosine adducts, we have re-evaluated the earlier structures of the reduced cross-links (*16, 17*) using a more robust rMD protocol. The new protocol models solvation and includes additional restraints for Watson−Crick hydrogen-bonding interactions. The resulting more accurate structures of the fully

Scheme 2. Numbering Schemes of the Synthetically Derived Trimethylene Interstrand Cross-Links[a]



[a] Both sequences are self-complementary. X and Y represent the N^2-cross-linked 2′-deoxyguanosines. (A) The 5′-CpG-3′ sequence, (B) the 5′-GpC-3′ sequence, and (C) the trimethylene tether.

reduced interstrand cross-links in both the 5′-CpG-3′ and the 5′-GpC-3′ sequence contexts show decreased perturbations in Watson−Crick hydrogen bonding as compared to the earlier structures (*16, 17*) and agree well with the refined structure of the (*R*)-carbinolamine cross-link (*15*), suggesting that the fully reduced trimethylene cross-links do serve as reliable structural models for the labile carbinolamine cross-links. The differential stabilities of the cross-links in the 5′-CpG-3′ and in the 5′-GpC-3′ sequence contexts are attributed to differential stacking interactions at the cross-linked region.

Materials and Methods

Oligodeoxynucleotides. The cross-linked oligodeoxynucleotides were prepared by the reactions of the halopurine-containing 8-mers 5′-d(AGGCZCCT)-3′ or 5′-d(TCCZCGGA)-3′; Z = 2-fluoro-O^6-(trimethylsilylethyl)-2′-deoxyinosine, with diaminopropane to form the monoadducted 8-mers. Reaction with additional halopurine-containing 8-mers yielded cross-links (*16, 17*).

NMR Data. The oligodeoxynucleotides containing the interstrand trimethylene cross-links were dissolved in 10 mM sodium phosphate buffer containing 0.1 M NaCl and 50 μM Na_2EDTA (pH 7.1). For observation of nonexchangeable resonances, samples were dissolved in 0.6 mL of D_2O, giving a 2 mM concentration of the duplex. For assignments of water-exchangeable protons, the samples were dissolved in a 9:1 H_2O/D_2O buffer of the same composition as that above. Spectra were referenced to an internal standard of TSP. Spectra were recorded at a ^1H frequency of 500.13 MHz. The NMR data were previously reported (*16, 17*).

Restraints. The NOE distance restraints and torsion angles for backbone angles δ and ε for the trimethylene cross-links in the 5′-CpG-3′ and 5′-GpC-3′ sequence contexts were obtained from the earlier NMR data (*16, 17*). Additional empirical backbone torsion angle restraints derived from the canonical structure of B-DNA (*18*) were used for backbone torsion angles α, β, γ, and ϕ, except that these empirical torsion angle restraints were not used at the cross-linked and the terminal base pairs. Empirical distance restraints were also used to maintain Watson−Crick hydrogen bonding and to prevent excessive propeller twisting of the base pairs, again with the exception of the cross-linking site and at the terminal base pairs.

rMD Calculations. The trimethylene cross-links were constructed using Insight II. Starting structures were generated from both A- and B-DNA. Furthermore, for both A- and B-DNA starting structures, two starting structures were utilized for the trimethylene tether. Substituting the hydrogen-bonded amino protons of the cross-

linked guanines with the trimethylene tether created the first, while the second set was created by replacing the nonhydrogen-bonded amino protons with the trimethylene tether. The partial charges on the trimethylene tether were obtained from density function theory (DFT) calculations using a neutral total charge, utilizing the B3LYP/6-31G* basis set and the program GAUSSIAN (*19*). To obtain the A-form and B-form starting structures that were used for subsequent restrained molecular dynamics (rMD) calculations, these A-form or B-form modified duplexes were energy minimized using 200 iterations with the conjugate gradients algorithm, in the absence of experimental restraints.

Twenty sets of randomly seeded rMD calculations (10 from each A type and 10 from each B type DNA starting structure) were conducted using the program AMBER (v 9.0) (*20*) and the parm99 force field. The Hawkins, Cramer, Truhlar pairwise generalized Born (GB) model (*21, 22*) was used to simulate implicit waters. The parameters developed by Tsui and Case (*23*) were used. The cutoff radius for nonbonding interactions was 18 Å. The restraint energy function contained terms describing distance and torsion angle restraints, both in the form of square well potentials. Bond lengths involving hydrogens were fixed with the SHAKE algorithm (*24*). A 1000-step energy minimization was performed with an integrator time of 1 fs without experimental restraints, followed by a 100000-iteration simulated annealing protocol with an integrator time step of 1 fs. The system was heated to 600 K in 5000 iterations and kept at 600 K for 5000 iterations and then cooled to 100 K with a time constant of 4.0 ps over 80000 iterations. A final cooling was applied to relax the system to 0 K with a time constant of 1.0 ps over 10000 iterations. The force constants of the restraints were scaled from 3.2 to 32 kcal mol^{-1} Å^{-2} during the first 10 ps and were maintained at 32 kcal mol^{-1} Å^{-2} for the remainder of the calculations. Convergence was assessed for structures having the lowest number of deviations from the experimental distance and dihedral restraints, lowest van der Vaals energies, and the lowest overall energies. Finally, the 20 refined structures were energy minimized for 250 iterations without restraints and then averaged to obtain the final structures. Helicoidal analyses were carried out with the program 3DNA (*25*).

Results

Structure of the Trimethylene Cross-Link in the 5′-CpG-3′ Sequence. Figure 1A shows an expanded view of the average structure at the cross-linked and neighbor base pairs from the minor groove. The trimethylene cross-link maintained Watson−Crick hydrogen bonding throughout the duplex. The individual structures emergent from the rMD calculations were similar and independent of A- vs B-DNA starting structures, as suggested by pairwise root-mean-square deviation (rmsd) values less than 1.0 Å. The rMD-generated structures exhibited greater similarity to canonical B-DNA. As compared to the starting A- and B-DNA structures, the average rmsd values were 3.6 and 2.1 Å, respectively.

The rMD calculations generated two asymmetric conformations for the trimethylene tether. Figure 2A displays one of the conformations in which the C$^{\beta}$ orients toward the 5′-d(A^{9}G^{10}G^{11}C^{12}Y^{13}C^{14}C^{15}T^{16})-3′ strand. In the other conformation, C$^{\beta}$ orients toward the 5′-d(A^{1}G^{2}G^{2}C^{3}X^{4}C^{5}C^{6}T^8)-3′ strand. Because the DNA sequence is self-complementary, these two conformations are identical by rotating the DNA duplex 180°. Thus, the calculated average structure exhibited pseudodyad symmetry, consistent with the NMR spectroscopy of this self-complementary duplex (*16*). Figure 2B shows the Newman projection viewed along the C$^{\alpha}$−C$^{\beta}$ bond. The conformation was similar to the symmetric U-conformation proposed by Dooley et al. (*16*). In that conformation, the cross-linked X^5 N^2 overlaps with C$^{\gamma}$. In contrast, our present study indicates that X^5 N^2 is in the gauche conformation with expect to C$^{\gamma}$, which reduces steric hindrance in the 5′-CpG-3′ cross-link.

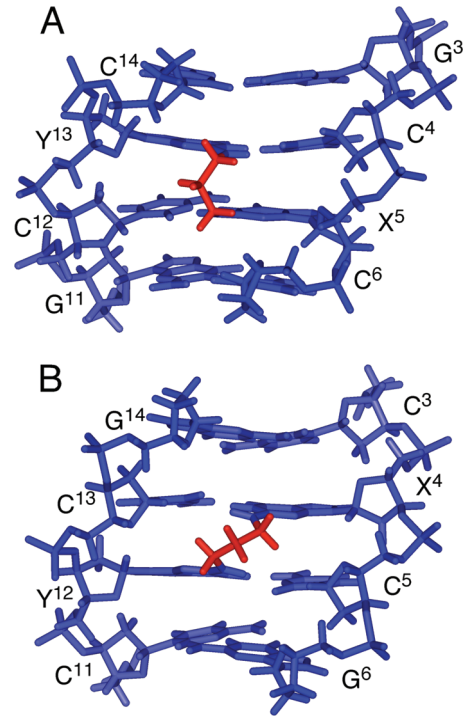


Figure 1. Expanded views of the trimethylene cross-links from the minor groove: (A) in the 5′-CpG-3′ sequence context and (B) in the 5′-GpC-3′ sequence context. Blue and red sticks represent the nucleotide and the tether, respectively.

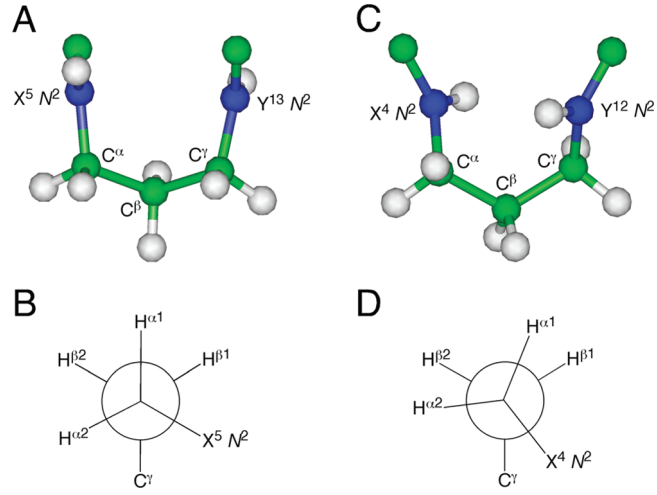


Figure 2. Conformations of the trimethylene tether: (A) in the 5′-CpG-3′ sequence context and (B) Newman projection of the tether viewed along the C$^{\alpha}$−C$^{\beta}$ bond in the 5′-CpG-3′ sequence context. The cross-linked X^5 N^2 is in gauche conformation with respect to C$^{\gamma}$. (C) In the 5′-GpC-3′ sequence context and (D) Newman projection of the tether viewed along the C$^{\alpha}$−C$^{\beta}$ bond in the 5′-GpC-3′ sequence context. The cross-linked X^4 N^2 is in half-gauche orientation with respect to C$^{\gamma}$; steric hindrance exists.

The experimentally determined δ and ε angles (*16*) were compared with the backbone torsion angles of the average structure emergent from the rMD calculations. They agreed with the average structure. Except at the cross-linked and the terminal base pairs, these torsion angles were consistent with B-DNA geometry. The χ glycosyl torsion angles of all nucleotides adopted the *anti* conformation. With regard to deoxyribose pseudorotation, the sugar puckers in the average structure were either C$_{1'}$-*exo*, C$_{2'}$-*endo*, C$_{3'}$-*exo*, or O$_{4'}$-*endo*.

Figure 3A,B displays the base stacking at the cross-linked base pairs. The cross-linked C^4•Y^{13} and X^5•C^{12} base pairing was normal, whereas they stacked with the neighbor G^3•C^{14} or C^6•G^{11}

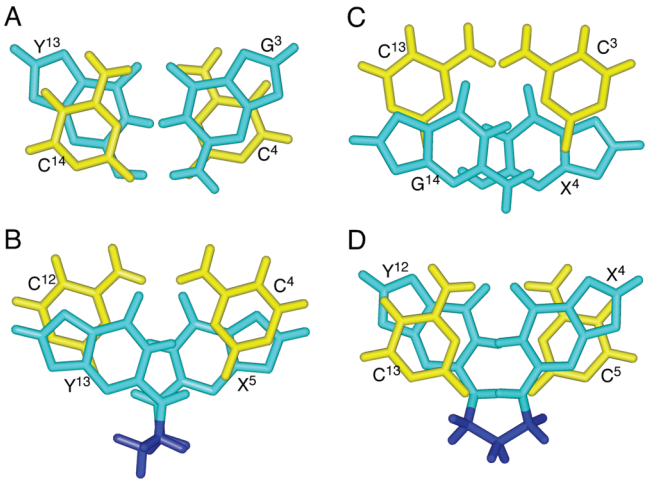


Figure 3. Base stacking of the cross-linked and the neighbor base pairs: (A and B) in the 5′-CpG-3′ cross-link and (C and D) in the 5′-GpC-3′ cross-link. Blue sticks represent the trimethylene tether.

base pairs in a pattern different from canonical B-DNA. Helicoidal analysis revealed that with the exception of roll, twist, inclination, and H-twist of the cross-linked base pairs, all other base pair parameters agreed with those of canonical B-DNA (*26*). The roll, twist, inclination, and H-twist of C^4•Y^{13} and X^5•C^{12} base pairs were 10.4, 8.8, 49.8, and 13.6°, respectively. As compared with those of B-DNA, they were remarkably shifted 9.8, −27.2, 47.7, and −22.9°, respectively.

Structure of the Trimethylene Cross-Link in the 5′-GpC-3′ Sequence. Figure 1B shows an expanded view of the average structure at the cross-linked and the neighbor base pairs from the minor groove. The 20 refined structures were similar, as suggested by pairwise rmsd values of less than 1.0 Å. The structures were compared with the starting A- and B-DNA model structures. The average rmsd values were 3.2 and 1.9 Å, respectively, suggesting that these rMD-generated structures were more similar to the canonical B-DNA structure. The trimethylene cross-link maintained Watson−Crick hydrogen bonding throughout the sequence.

Figure 2C,D displays the conformation of the trimethylene tether and the Newman projection viewed along the $C^{\alpha}-C^{\beta}$ bond. The tether adopted the symmetric "skew U" conformation (*16*). The X^4 N^2 atom was in the half-gauche conformation with respect to the C^{γ} of the trimethylene tether. The dihedral angles of $H^{\alpha 1}-C^{\alpha}-C^{\beta}-H^{\beta 1}$, $H^{\alpha 1}-C^{\alpha}-C^{\beta}-H^{\beta 2}$, and $H^{\alpha 2}-C^{\alpha}-C^{\beta}-H^{\beta 1}$ were 35, −80, and 151°, respectively. They agreed with the ^3J coupling constants of $H^{\alpha(\gamma)}-H^{\beta}$ (6, 0.5, and 8.5 Hz) (note that $H^{\alpha 1}$, $H^{\alpha 2}$, and $H^{\beta 1}$ are magnetically equal to $H^{\gamma 1}$, $H^{\gamma 2}$, and $H^{\beta 2}$, respectively) (*17*). The average structure exhibited pseudodyad symmetry, consistent with the NMR spectroscopy of the self-complementary duplex.

The experimental δ backbone torsion angles were compared with the average structure. These agreed with those of the average structure. Except for the cross-linked and the terminal base pairs, these torsion angles agreed well with the B-DNA geometry. All nucleotides adopted the *anti* conformation about the glycosyl torsion angle. The torsion angles of the deoxyriboses of the average structure were compared with experimental data of the 5′-GpC-3′ cross-link. The experimental torsion angles agreed with those of the average structure. Consistent with the B-DNA geometry, the sugar puckers were either $C_{1'}$-*exo*, $C_{2'}$-*endo*, or $O_{4'}$-*endo*.

Figure 3C,D displays the base stacking at the cross-linked base pairs. The cross-linked X^4•C^{13} and C^5•Y^{12} base pairs were normal, whereas the neighboring C^3•G^{14} and G^6•C^{11} base pairs exhibited minimal overlapping with the cross-linked X^4•C^{13} or C^5•Y^{12} base pairs. Except for the propeller twist of the cross-linked base pairs, all other helicoidal base pair parameters agreed with those of B-DNA (*26*). The propeller twisting of the X^4•C^{13} and C^5•Y^{12} base pairs was 11° and 10°, respectively, representing 23° and 22° shifts as compared with canonical B-DNA, indicating that the cross-linked X^4•C^{13} and C^5•Y^{12} base pairs were not parallel to the neighboring C^3•G^{14} and G^6•C^{11} base pairs. The interference of the base stacking at the cross-linked region was consistent with the remarkable chemical shift perturbations of the C^5 aromatic protons (*17*). As compared to the unmodified duplex, C^5 H5 and H6 shifted downfield by 0.4−0.5 ppm. The chemical shifts were comparable to those of deoxycytidine, suggesting that ring current effects arising from stacking of the neighboring bases was minimal. This was consistent with the observation that the X^4•C^{13} and C^5•Y^{12} base pairs exhibited minimal stacking with the neighboring C^3•G^{14} or G^6•C^{11} base pairs (Figure 3C).

Discussion

In duplex DNA, when placed opposite dC, the acrolein derived γ-OH-1,N^2-propano-2′-deoxyguanosine adduct **1** ring opens to N^2-dG aldehyde adduct **2** and induces the carbinolamine interstrand cross-link **4** in the 5′-CpG-3′ sequence, whereas it does not in the 5′-GpC-3′ sequence (*10*). Dooley et al. (*16, 17*) used the synthetically derived trimethylene cross-links **6** to simulate the labile carbinolamine cross-links. Whereas the 5′-CpG-3′ cross-link increased the T_m of the duplex by 30 °C, the 5′-GpC-3′ cross-link reduced the T_m of the duplex by 10 °C. The more accurate solution structures refined with the AMBER force field in this work now augment our understanding of why cross-linking prefers the 5′-CpG-3′ sequence context.

Trimethylene Cross-Link in the 5′-CpG-3′ Sequence. Dooley et al. (*16*) reported NMR data of the trimethylene cross-link in the 5′-CpG-3′ sequence. However, the Dooley et al. (*16*) structure suggested that the trimethylene tether of the cross-link adopted an unwound W-conformation that destroyed the B-DNA geometry by forming a bulge at the cross-linked base pairs. The two cross-linked guanines were in the same plane. The Watson−Crick proton of the guanine exocyclic amino group was replaced by the trimethylene tether, and moreover, hydrogen bonding and stacking of the base pairs were missing throughout the structure (*16*). This was inconsistent with the sharp imino resonances and their NOE cross-peaks arising from Watson−Crick hydrogen bonding and with the complete NOE connectivity of aromatic H6/H8 protons with deoxyribose H1′ protons (*16*), consistent with the B-DNA-like geometry of the refined structure presented here.

The present recalculation provides an improved depiction of the structure. The duplex maintains B-DNA geometry with minimal perturbations located at the cross-linked base pairs (Figure 1A). The present calculations show that the interstrand cross-link conserves Watson−Crick hydrogen bonding throughout the duplex, in agreement with the observation of X^5 N1H → X^5 N^2H, X^5 N1H → C^4 N^4H2, G^3 N1H → C^6 N^4H2, G^2 N1H → C^7 N^4H2, and G^3 N1H → X^5 N1H NOE correlations (*16*). Notably, the X^5 H8 → C^4 H1′ NOE correlation was weakened (*16*). Consistently, the distance between these two atoms in the present rMD structure is 5.3 Å. As compared with the unmodified duplex, large chemical shift perturbations were reported for the C^4 H2′ and H2″ protons (*16*), which are consistent with the structure of the cross-linked C^4•Y^{13} and X^5•C^{12} base pairs (Figure 1A). The trimethylene tether protons exhibited NOE

interactions with C^6 H1′ (*16*). Consistently, the corresponding atom distances in the present calculations are 2.6, 3.4, and 4.8 Å, respectively. In agreement with the corresponding atom distances in the refined structure, the trimethylene tether also exhibited strong NOEs with X^5 N^2H and medium NOEs with X^5 N1H (*16*). Collectively, these results suggest that the average structure provides an accurate depiction of the NMR data.

Trimethylene Cross-Link in the 5′-GpC-3′ Sequence. Dooley et al. (*17*) also reported NMR data of the synthetically derived trimethylene cross-link in the 5′-GpC-3′ sequence. In their structure, the trimethylene tether adopted the "Skew U" conformation. The B-DNA geometry was destroyed. The Watson−Crick proton of the guanine exocyclic amino group was replaced by the trimethylene tether. The hydrogen bonding and stacking of the base pairs were missing throughout the duplex. This structure was inconsistent with their observation of a complete NOE connectivity of aromatic H6/H8 protons with the deoxyribose H1′ protons, suggesting that the cross-link maintained B-DNA type geometry (*17*).

The recalculation provides a more accurate depiction of the structure. The duplex maintains B-DNA geometry with perturbations located at the cross-linked base pairs (Figure 1B). This is consistent with the complete NOE connectivities of the base H6/H8 protons with deoxyribose H1′ protons (*17*). The preservation of Watson−Crick hydrogen bonding agrees with the observation of sharp imino resonances and their NOE interactions (*17*). In contrast to the 5′-CpG-3′ sequence, in the 5′-GpC-3′ sequence, the C^5 H6 → X^4 H1′ NOE correlation was not weakened (*17*). Consistently, the distance of these two atoms in the new refined structure is 3.1 Å. As compared with the B-DNA duplex, the cross-linked X^4•C^{13} and C^5•Y^{12} base pairs are perturbed (Figure 1B). This explains the large chemical shift perturbations observed for the C^5 H5, H6, H1′, and H2′ resonances (*17*). Significantly, the interstrand cross-link conserves Watson−Crick hydrogen bonding in the average structure, also agreeing with the observation of the X^4 N1H → X^4 N^2H, X^4 N1H → C^5 N^4H2, G^6 N1H → C^3 N^4H2, G^7 N1H → C^2 N^4H2, G^6 N1H → X^4 N1H, and G^6 N1H → G^7 N1H NOE correlations. The distances between the trimethylene protons and X^4 N1H, N^2H in the average structure also agree with the NOEs of the trimethylene with X^4 N1H and N^2H. In addition, the torsion angle of the trimethylene protons is consistent with the experimental ^3J coupling constants. Collectively, these observations suggest that the average structure provides an accurate depiction of the NMR data.

Structure−Stability Relationship of the Trimethylene Cross-Links. Normally, interstrand cross-links increase the stability of the DNA duplex. The refined structures show that both cross-links maintain B-DNA geometry; however, the stacking at the cross-linked base pairs varies (Figure 3). The cross-linked C^4•Y^{13} and X^5•C^{12} base pairs incline toward each other in the 5′-CpG-3′ sequence context. Although different from the stacking pattern of canonical B-DNA, the cross-linked base pairs maintain π-stacking interactions with the neighbor base pairs (Figure 3A,B). In contrast, because the distance of two guanine N^2 atoms in the 5′-GpC-3′ sequence context is too long for the 3-carbon trimethylene tether, the cross-linking forces the cross-linked guanines to tilt toward each other. As compared with canonical B-DNA, the cross-link induces significant propeller twisting. Consequently, C^3•G^{14} and G^6•C^{11} base pairs are not parallel (Figure 1B) and barely overlap with the X^4•C^{13} or C^5•Y^{12} base pairs, respectively (Figure 3C). No π-stacking interaction exists between C^3•G^{14} and X^4•C^{13} or between C^5•Y^{12} and G^6•C^{11}. The chemical shift perturbations of both trimethylene interstrand cross-links as compared to the unmodified duplexes have been reported (*16, 17*). The NMR data are in agreement with the present structures showing differential base-stacking interactions for these two trimethylene interstrand cross-links. The chemical shifts of H5 and H6 of C^5 and C^{13} in the 5′-GpC-3′ sequence are comparable with those in the deoxycytidine, suggesting that the ring current effect of the neighboring bases is minimal due to the base stacking perturbation. In contrast, these chemical shift perturbations in the 5′-CpG-3′ sequence are not remarkable, consistent with the stacked structure. Base stacking rather than base pairing is the main stabilizing factor in the DNA duplex (*27–29*). Consistently, as compared with the unmodified DNA duplex, the trimethylene cross-link stabilizes the 5′-CpG-3′ duplex, whereas it destabilizes the 5′-GpC-3′ duplex. In addition, to form the cross-link, the trimethylene tether in the 5′-GpC-3′ sequence places X^4 N^2 and C^γ or Y^{12} N^2 and C^α in the half-gauche conformation and increases steric hindrance (Figure 2D), which may further destabilize the interstrand cross-link. Because the trimethylene cross-links in both 5′-CpG-3′ and 5′-GpC-3′ sequence contexts do not induce significant bending (*16, 17*), we conclude that the stability difference of these cross-links is mainly attributed to the differential stacking interactions at the cross-linked region. The ability of the enal-derived 1,N^2-dG adducts to induce interstrand cross-links in the 5′-CpG-3′ sequence and the inability to induce cross-links in the 5′-GpC-3′ sequence suggest that the formation of the cross-link in the 5′-GpC-3′ sequence is thermodynamically unfavorable.

Comparison with *R*-Carbinolamine Cross-Link. The structures of the acrolein-derived (*R*)-carbinolamine interstrand cross-link (*15*) and the trimethylene cross-link in the 5′-CpG-3′ sequence are similar. In both, the cross-linked guanine N^2 atoms are in the gauche orientation with respect to the tether. Importantly, the γ-hydroxyl group of the (*R*)-carbinolamine cross-link is critical for cross-link formation. It is positioned to form a hydrogen bond with the 5′-cytosine neighboring the cross-linked base pairs. The conformation of the trimethyl tether also allows the hydrogen bond if the corresponding H^γ is substituted by a hydroxyl group to produce the (*R*)-γ-hydroxytrimethylene tether. The conformational similarity of these two cross-links indicates that the trimethylene interstrand cross-links represent adequate models of the carbinolamine cross-links induced by the enal-derived 1,N^2-dG adducts.

Comparison with α-Methyltrimethylene Cross-Links. Cho et al. (*30*) reported structures of N^2-dG:N^2-dG interstrand cross-links bridged by stereoisomeric α-methyltrimethylene tethers in the 5′-CpG-3′ sequence context. The cross-links maintain B-DNA geometry with small perturbations located at the cross-linked region. The cross-linked base pairs are placed in proximity to each other. However, to minimize the steric hindrance of the α-methyl group, the orientation of the (*R*)-tether disfavors the formation of the hydrogen bond needed to stabilize the carbinolamine interstrand cross-link. The small structural difference explains the reduced ability of the crotonaldehyde-derived 1,N^2-dG adducts to produce interstrand cross-links in the 5′-CpG-3′ sequence as compared to acrolein adduct (*10*).

Comparison with Interstrand Cross-Links Containing 4-Carbon Tethers. Norman et al. (*31*) reported the structure of the N^2-dG:N^2-dG interstrand cross-link in the 5′-CpG-3′ duplex linked by mitomycin C, and Fagan et al. (*32*) reported the structure of the N^2-dG:N^2-dG interstrand cross-link linked by distamycin-pyrrole. The cross-linked base pairs of these cross-links maintain Watson−Crick hydrogen bonding. The DNA duplexes undergo small structural perturbations. The

distamycin-pyrrole cross-link induces a buckle in the opposite direction of the cross-linked dX•dC base pairs that results in the cross-linked guanines coming close to each other and decreases the helical twist at the lesion site. As compared with canonical B-DNA, the base pairs of the 3-carbon trimethylene cross-link are also located in proximity to each other. In contrast to these 4-carbon interstrand cross-links, the increased proximity of the cross-linked base pairs in the trimethylene cross-link is due to the inclination of the cross-linked base pairs. This may be the result of the 3-carbon trimethylene tether being still not long enough to bond two guanines in the 5′-direction; thus, a larger structural perturbation is required to form the cross-link.

Structures Deposited in RCSB Protein Data Bank. The cross-link in the 5′-CpG-3′ sequence is PDB ID 2KNK. The cross-link in the 5′-GpC-3′ sequence is PDB ID 2KNL.

Summary

The structures of interstrand cross-links bridged by trimethylene tethers were refined. In either 5′-CpG-3′ or 5′-GpC-3′ sequence contexts, the B-DNA duplex is minimally perturbed at the cross-linking region. Watson−Crick hydrogen bonding is conserved throughout the sequences. The stability difference of the cross-links in the 5′-CpG-3′ and 5′-GpC-3′ sequence contexts is attributed to differential stacking interactions at the cross-linked region, which might be responsible for the sequence-dependent formation of interstrand cross-links induced by the enal-derived γ-OH-1,N^2-propano-2′-deoxyguanosine adducts.

Acknowledgment. This work was supported by NIH Grant PO1 ES-05355 (T.M.H. and M.P.S.).

Supporting Information Available: Backbone torsion angles of the cross-link in the 5′-CpG-3′ sequence (Table S1), deoxyribose torsion angles of the cross-link in the 5′-CpG-3′ sequence (Table S2), backbone torsion angles of the cross-link in the 5′-GpC-3′ sequence (Table S3), deoxyribose torsion angles of the cross-link in the 5′-GpC-3′ sequence (Table S4), structural comparisons of the cross-link in the 5′-CpG-3′ sequence presented here with those previously reported (*16, 17*) (Figure S1), structural comparisons of the cross-link in the 5′-GpC-3′ sequence presented here with those previously reported (*16, 17*) (Figure S2), and base pairing parameters of the trimethylene cross-links (Figure S3). This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Burcham, P. C. (1998) Genotoxic lipid peroxidation products: Their DNA damaging properties and role in formation of endogenous DNA adducts. *Mutagenesis 13*, 287–305.

(2) Chung, F. L., Zhang, L., Ocando, J. E., and Nath, R. G. (1999) Role of 1,N^2-propanodeoxyguanosine adducts as endogenous DNA lesions in rodents and humans. *IARC Sci. Publ. 150*, 45–54.

(3) Chung, F. L., Nath, R. G., Nagao, M., Nishikawa, A., Zhou, G. D., and Randerath, K. (1999) Endogenous formation and significance of 1,N^2-propanodeoxyguanosine adducts. *Mutat. Res. 424*, 71–81.

(4) Nair, U., Bartsch, H., and Nair, J. (2007) Lipid peroxidation-induced DNA damage in cancer-prone inflammatory diseases: A review of published adduct types and levels in humans. *Free Radical Biol. Med. 43*, 1109–1120.

(5) Hecht, S. S., Upadhyaya, P., and Wang, M. (1999) Reactions of α-acetoxy-*N*-nitrosopyrrolidine and crotonaldehyde with DNA. *IARC Sci. Publ. 150*, 147–154.

(6) Nath, R. G., Ocando, J. E., Guttenplan, J. B., and Chung, F. L. (1998) 1,N^2-propanodeoxyguanosine adducts: Potential new biomarkers of smoking-induced DNA damage in human oral tissue. *Cancer Res. 58*, 581–584.

(7) Esterbauer, H., Schaur, R. J., and Zollner, H. (1991) Chemistry and biochemistry of 4-hydroxynonenal, malonaldehyde and related aldehydes. *Free Radical Biol. Med. 11*, 81–128.

(8) Minko, I. G., Kozekov, I. D., Harris, T. M., Rizzo, C. J., Lloyd, R. S., and Stone, M. P. (2009) Chemistry and biology of DNA containing 1,N^2-deoxyguanosine adducts of the α,β-unsaturated aldehydes acrolein, crotonaldehyde, and 4-hydroxynonenal. *Chem. Res. Toxicol. 22*, 759–778.

(9) Kozekov, I. D., Nechev, L. V., Sanchez, A., Harris, C. M., Lloyd, R. S., and Harris, T. M. (2001) Interchain cross-linking of DNA mediated by the principal adduct of acrolein. *Chem. Res. Toxicol. 14*, 1482–1485.

(10) Kozekov, I. D., Nechev, L. V., Moseley, M. S., Harris, C. M., Rizzo, C. J., Stone, M. P., and Harris, T. M. (2003) DNA interchain cross-links formed by acrolein and crotonaldehyde. *J. Am. Chem. Soc. 125*, 50–61.

(11) Wang, H., Kozekov, I. D., Harris, T. M., and Rizzo, C. J. (2003) Site-specific synthesis and reactivity of oligonucleotides containing stereochemically defined 1,N^2-deoxyguanosine adducts of the lipid peroxidation product *trans*-4-hydroxynonenal. *J. Am. Chem. Soc. 125*, 5687–5700.

(12) Stone, M. P., Cho, Y. J., Huang, H., Kim, H. Y., Kozekov, I. D., Kozekova, A., Wang, H., Minko, I. G., Lloyd, R. S., Harris, T. M., and Rizzo, C. J. (2008) Interstrand DNA cross-links induced by α,β-unsaturated aldehydes derived from lipid peroxidation and environmental sources. *Acc. Chem. Res. 41*, 793–804.

(13) Kim, H. Y., Voehler, M., Harris, T. M., and Stone, M. P. (2002) Detection of an interchain carbinolamine cross-link formed in a CpG sequence by the acrolein DNA adduct γ-OH-1,N^2-propano-2′-deoxyguanosine. *J. Am. Chem. Soc. 124*, 9324–9325.

(14) Cho, Y. J., Kim, H. Y., Huang, H., Slutsky, A., Minko, I. G., Wang, H., Nechev, L. V., Kozekov, I. D., Kozekova, A., Tamura, P., Jacob, J., Voehler, M., Harris, T. M., Lloyd, R. S., Rizzo, C. J., and Stone, M. P. (2005) Spectroscopic characterization of interstrand carbinolamine cross-links formed in the 5′-CpG-3′ sequence by the acrolein-derived γ-OH-1,N^2-propano-2′-deoxyguanosine DNA adduct. *J. Am. Chem. Soc. 127*, 17686–17696.

(15) Huang, H., Kim, H.-Y., Kozekov, I. D., Cho, Y.-J., Wang, H., Kozekova, A., Harris, T. M., Rizzo, C. J., and Stone, M. P. (2009) Stereospecific formation of the (*R*)-γ-hydroxytrimethylene interstrand N^2-dG:N^2-dG DNA cross-link arising from the γ-OH-1,N^2-propano-2′-deoxyguanosine adduct in the 5′-CpG-3′ sequence. *J. Am. Chem. Soc. 131*, 8416–8424.

(16) Dooley, P. A., Tsarouhtsis, D., Korbel, G. A., Nechev, L. V., Shearer, J., Zegar, I. S., Harris, C. M., Stone, M. P., and Harris, T. M. (2001) Structural studies of an oligodeoxynucleotide containing a trimethylene interstrand cross-link in a 5′-(CpG) motif: Model of a malondialdehyde cross-link. *J. Am. Chem. Soc. 123*, 1730–1739.

(17) Dooley, P. A., Zhang, M., Korbel, G. A., Nechev, L. V., Harris, C. M., Stone, M. P., and Harris, T. M. (2003) NMR determination of the conformation of a trimethylene interstrand cross-link in an oligodeoxynucleotide duplex containing a 5′-d(GpC) motif. *J. Am. Chem. Soc. 125*, 62–72.

(18) Arnott, S., and Hukins, D. W. L. (1972) Optimised parameters for A-DNA and B-DNA. *Biochem. Biophys. Res. Commun. 47*, 1504–1509.

(19) Frisch, M. J., Trucks, G. W., Schlegel, H. B., Scuseria, G. E., Robb, M. A., Cheeseman, J. R., Montgomery, J. A., Vreven, T., Kudin, K. N., Burant, J. C., Millam, J. M., Iyengar, S. S., Tomasi, J., Barone, V., Mennucci, B., Cossi, M., Scalmani, G., Rega, N., Petersson, G. A., Nakatsuji, H., Hada, M., Ehara, M., Toyota, K., Fukuda, R., Hasegawa, J., Ishida, M., Nakajima, T., Honda, Y., Kitao, O., Nakai, H., Klene, M., Li, X., Knox, J. E., Hratchian, H. P., Cross, J. B., Adamo, C., Jaramillo, J., Gomperts, R., Stratmann, R. E., Yazyev, O., Austin, A. J., Cammi, R., Pomelli, C., Pomelli, J., Ochterski, W., Ayala, P. Y., Morokuma, K., Voth, G. A., Salvador, P., Dannenberg, J. J., Zakrzewska, V. G., Daniels, A. D., Farkas, O., Rabuck, A. D., Raghavachari, K., and Ortiz, J. V. (2004) *GAUSSIAN 03*, Gaussian, Inc., Wallingford, CT.

(20) Case, D. A., Cheatham, T. E., 3rd, Darden, T., Gohlke, H., Luo, R., Merz, K. M., Jr., Onufriev, A., Simmerling, C., Wang, B., and Woods, R. J. (2005) The AMBER biomolecular simulation programs. *J. Comput. Chem. 26*, 1668–1688.

(21) Hawkins, G. D., Cramer, C. J., and Truhlar, D. G. (1995) Pairwise solute descreening of solute charges from a dielectric medium. *Chem. Phys. Lett. 246*, 122–129.

(22) Hawkins, G. D., Cramer, C. J., and Truhlar, D. G. (1996) Parametrized models of aqueous free energies of solvation based on pairwise descreening of solute atomic charges from a dielectric medium. *J. Phys. Chem. 100*, 19824–19839.

(23) Tsui, V., and Case, D. A. (2000) Theory and applications of the generalized Born solvation model in macromolecular simulations. *Biopolymers 56*, 275–291.

(24) Ryckaert, J.-P., Ciccotti, G., and Berendsen, H. J. C. (1977) Numerical integration of the cartesian equations of motion of a system with constraints: Molecular dynamics of n-alkanes. *J. Comp. Phys. 23*, 327–341.

(25) Lu, X. J., and Olson, W. K. (2003) 3DNA: A software package for the analysis, rebuilding and visualization of three-dimensional nucleic acid structures. *Nucleic Acids Res. 31*, 5108–5121.

(26) Olson, W. K., Bansal, M., Burley, S. K., Dickerson, R. E., Gerstein, M., Harvey, S. C., Heinemann, U., Lu, X. J., Neidle, S., Shakked, Z., Sklenar, H., Suzuki, M., Tung, C. S., Westhof, E., Wolberger, C., and Berman, H. M. (2001) A standard reference frame for the description of nucleic acid base-pair geometry. *J. Mol. Biol. 313*, 229–237.

(27) Petersheim, M., and Turner, D. G. (1983) Base-stacking and base-pairing contributions to helix stability: Thermodynamics of double-helix formation with CCGG, CCGGp, CCGGAp, ACCGGp, CCGGUp, and ACCGGUp. *Biochemistry 22*, 256–263.

(28) Turner, D. H., Petersheim, M., Albergo, D. D., Dewey, T. G., and Freier, S. M. (1995) Why do nucleic acids form helices? In *Biomolecular Stereodynamics* (Sarma, R. H., Ed.) pp 429−438, Adenine Press, New York.

(29) Yakovchuk, P., Protozanova, E., and Frank-Kamenetskii, M. D. (2006) Base-stacking and base-pairing contributions into thermal stability of the DNA double helix. *Nucleic Acids Res. 34*, 564–574.

(30) Cho, Y. J., Kozekov, I. D., Harris, T. M., Rizzo, C. J., and Stone, M. P. (2007) Stereochemistry modulates the stability of reduced interstrand cross-links arising from *R*- and *S*-α-CH_3-γ-OH-1,N^2-propano-2′-deoxyguanosine in the 5′-CpG-3′ DNA sequence. *Biochemistry 46*, 2608–2621.

(31) Norman, D., Live, D., Sastry, M., Lipman, R., Hingerty, B. E., Tomasz, M., Broyde, S., and Patel, D. J. (1990) NMR and computational characterization of mitomycin cross-linked to adjacent deoxyguanosines in the minor groove of the d(T-A-C-G-T-A)•d(T-A-C-G-T-A) dulplex. *Biochemistry 29*, 2861–2875.

(32) Fagan, P. A., Spielmann, H. P., Sigurdsson, S., Rink, S. M., Hopkins, P. B., and Wemmer, D. E. (1996) An NMR study of [d(CGCGAATTCGCG)]$_2$ containing an interstrand cross-link derived from a distamycin-pyrrole conjugate. *Nucleic Acids Res. 24*, 1566–1573.

TX900225C