J|A|C|S
ARTICLES
Published on Web 12/04/2002

DNA Interchain Cross-Links Formed by Acrolein and Crotonaldehyde

Ivan D. Kozekov, Lubomir V. Nechev,§ M. Scott Moseley, Constance M. Harris, Carmelo J. Rizzo, Michael P. Stone, and Thomas M. Harris*

Contribution from the Department of Chemistry and Center in Molecular Toxicology, Vanderbilt University, Nashville, Tennessee 37235

Received June 4, 2002

Abstract: Acrolein and higher α,β-unsaturated aldehydes are bifunctional genotoxins. The deoxyguanosine adduct of acrolein, 3-(2-deoxy-β-D-*erythro*-pentofuranosyl)-5,6,7,8-tetrahydro-8-hydroxypyrimido[1,2-*a*]purin-10(3*H*)-one (8-hydroxy-1,N^2-propanodeoxyguanosine, **2a**), is a major DNA adduct formed by acrolein. The potential for oligodeoxynucleotide duplexes containing **2a** to form interchain cross-links was evaluated by HPLC, CZE, MALDI-TOF, and melting phenomena. Interchain cross-links represent one of the most serious types of damage in DNA since they are absolute blocks to replication. In oligodeoxynucleotides containing the sequence 5′-dC-**2a**, cross-linking occurred in a slow, reversible manner to the extent of ∼50%. Enzymatic digestion to form 3-(2-deoxy-β-D-*erythro*-pentofuranosyl)-5,6,7,8-tetrahydro-8-(N^2-2′-deoxyguanosinyl)-pyrimido[1,2-*a*]purin-10(3*H*)one (**5a**) and reduction with $NaCNBH_3$ followed by enzymatic digestion to give 1,3-bis(2′-deoxyguanosin-N^2-yl)propane (**6a**) established that cross-linking had occurred with the exocyclic amino group of deoxyguanosine. It is concluded that the cross-link is a mixture of imine and carbinolamine structures. With oligodeoxynucleotide duplexes containing the sequence 5′-**2a**-dC, cross-links were not detected by the techniques enumerated above. In addition, $^{15}N-^{1}H$ HSQC and HSQC-filtered NOESY spectra carried out with a duplex having ^{15}N-labeling of the target amino group established unambiguously that a carbinolamine cross-link was not formed. The potential for interchain cross-link formation by the analogous crotonaldehyde adduct (**2b**) was evaluated in a 5′-dC-**2b** sequence. Cross-link formation was strongly dependent on the configuration of the methyl group at C6 of **2b**. The 6*R* diastereomer of **2b** formed a cross-link to the extent of 38%, whereas the 6*S* diastereomer cross-linked only 5%.

Introduction

The pervasiveness of unsaturated aldehydes in the environment and in vivo has raised concern about their toxicity, in particular, their role in carcinogenesis. Acrolein and crotonaldehyde are prepared in large quantities industrially for use as synthetic intermediates. Both compounds are also formed by combustion of plant material including tobacco.[1,2] In addition, acrolein, crotonaldehyde, and other unsaturated aldehydes are formed endogenously by oxidative degradation of unsaturated lipids.[3−7]

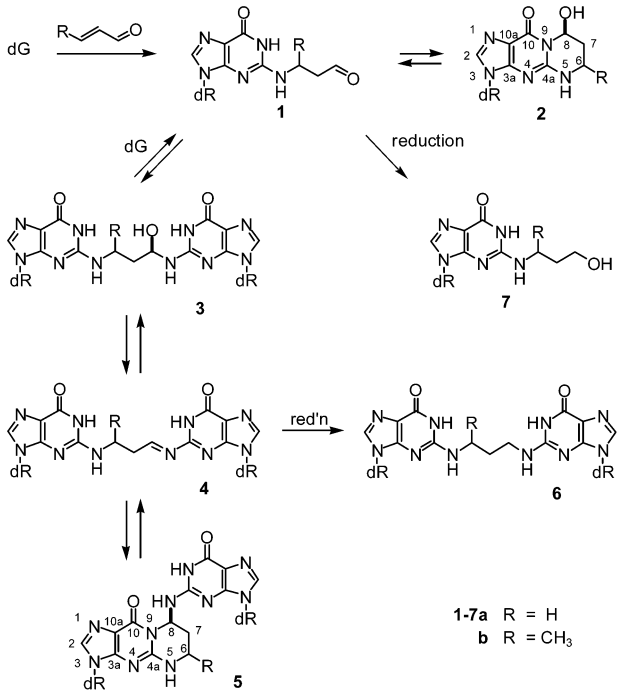
Unsaturated aldehydes are bis-electrophiles and have the capability of reacting with nucleosides via either the carbonyl group or the double bond. The simple nucleoside adducts are generally not observed, because they react a second time to fuse a ring on the nucleobase.[8−12] In Scheme 1 is shown the conjugate addition of the exocyclic amino group of deoxyguanosine to unsaturated aldehydes. Conjugate addition to acrolein followed by cyclization of the initially formed N^2-(3-oxopropyl) adduct **1a** gives the 8-hydroxypropano adduct **2a**, which is a major adduct of the reaction of acrolein with DNA.[4,6,8,13−15] A number of examples of 8-hydroxypropano adducts have been characterized resulting from reactions of other α,β-unsaturated aldehydes with nucleosides and DNA and also in exposed cells, laboratory animals, and humans. The unsaturated aldehydes that have received the most study are acrolein and crotonaldehyde; DNA adducts of both have been identified in living systems.[6,13,16−19]

* Corresponding author. Telephone: (615) 322-2649. Fax: (615) 322-7591. E-mail: harristm@toxicology.mc.vanderbilt.edu.
§ Current address: Transgenomic, Inc., Boulder, CO 80301.

(1) Treitman, R. D.; Burgess, W. A.; Gold, A. *Am. Ind. Hyg. Assoc. J.* **1980**, *41*, 796−802.
(2) Izard, C.; Valadaud-Barrieu, D.; Fayeulle, J. P.; Testa, A. *Mutat. Res.* **1980**, *77*, 341−344.
(3) Pan, J.; Chung, F. L. *Chem. Res. Toxicol.* **2002**, *15*, 367−372.
(4) Chung, F. L.; Nath, R. G.; Nagao, M.; Nishikawa, A.; Zhou, G. D.; Randerath, K. *Mutat. Res.* **1999**, *424*, 71−81.
(5) Wilson, V. L.; Foiles, P. G.; Chung, F. L.; Povey, A. C.; Frank, A. A.; Harris, C. C. *Carcinogenesis* **1991**, *12*, 1483−1490.
(6) Nath, R. G.; Ocando, J. E.; Chung, F. L. *Cancer Res.* **1996**, *56*, 452−456.
(7) Uchida, K.; Kanematsu, M.; Morimitsu, Y.; Osawa, T.; Noguchi, N.; Niki, E. *J. Biol. Chem.* **1998**, *273*, 16058−16066.
(8) Chung, F. L.; Young, R.; Hecht, S. S. *Cancer Res.* **1984**, *44*, 990−995.
(9) Chung, F. L.; Young, R.; Hecht, S. S. *Carcinogenesis* **1989**, *10*, 1291−1297.
(10) Smith, R. A.; Williamson, D. S.; Cerny, R. L.; Cohen, S. M. *Cancer Res.* **1990**, *30*, 3005−3012.
(11) Smith, R. A.; Williamson, D. S.; Cohen, S. M. *Chem. Res. Toxicol.* **1989**, *2*, 267−271.
(12) Sodum, R. S.; Shapiro, R. *Bioorganic Chem.* **1988**, *16*, 272−282.
(13) Nath, R. G.; Chung, F.-L. *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 7491−7495.
(14) Nath, R. G.; Chen, H.-J. C.; Nishikawa, A.; Young-Sciame, R.; Chung, F.-L. *Carcinogenesis* **1994**, *15*, 979−984.
(15) Galliani, G.; Pantarotto, C. *Tetrahedron Lett.* **1983**, *24*, 4491−4492.
(16) Eder, E.; Scheckenbach, S.; Deininger, C.; Hoffman, C. *Toxicol. Lett.* **1993**, *67*, 87−103.

10.1021/ja020778f CCC: $25.00

Scheme 1



Further reactions of the DNA adducts of unsaturated aldehydes are also possible. The reactions include formation of cross-links with proteins and nucleic acids. For example, 8-hydroxypropano adduct **2a** could react with nucleotides in DNA to form cross-linked species by an initial step involving reversion of **2a** to oxopropyl adduct **1a** which could then react with the exocyclic amino group of another guanine to form carbinolamine **3a**, which could dehydrate to imine **4a**. The latter might conceivably cyclize to pyrimidopurinone **5a** (Scheme 1). Adduct **1a** could also react with proteins by forming imines with lysine or N-terminal amino groups.

Many examples have been reported of DNA cross-link formation by bis-alkylating agents; prominent examples include the nitrogen mustards and mitomycin which are used in cancer chemotherapy.[20−22] These cross-linking reactions have received considerable attention because of the likelihood that the pharmacological effects are a function of their ability to form cross-links. Identification of the cross-linked species has been aided by the fact that the cross-linking reactions of alkylating agents are effectively irreversible. However, the evidence for unsaturated aldehydes forming cross-links in DNA has been sparse. It is likely that the reactions occur frequently, but the resulting carbinolamines and imines are difficult to detect due to hydrolytic instability.

Recently, Kawanishi et al. presented electrophoretic evidence for interchain and intrachain cross-linking of DNA by acrolein, although the cross-linked nucleosides were not isolated or characterized.[23] Wang et al. have reported formation of interchain cross-links by acetaldehyde.[24] The acetaldehyde reaction utilizes two equivalents of the aldehyde and involves the exocyclic amino groups of guanines. Wang isolated the pyrimidopurinone-type bis-nucleoside **5b** (along with crotonaldehyde adduct **2b**) by enzymatic digestion of acetaldehyde-treated calf thymus DNA. The cross-linking reaction is formally the equivalent of the reaction of crotonaldehyde, but they suggested that the primary route to this cross-link involves the Schiff base of acetaldehyde reacting with a second molecule of acetaldehyde to form aldehyde **1b** and the latter either cyclizing to form hydroxypropano adduct **2b** or undergoing condensation with another guanine to form imine-linked bis-nucleoside **4b** which in turn cyclizes to pyrimidopurinone **5b**.

We recently communicated preliminary results on the formation of acrolein-mediated interchain cross-links in duplex DNA. The study involved a 5′-d(CpG) sequence context.[25] Further characterization of the cross-linking reaction has revealed that cross-linking is strongly sequence-dependent. In a GpC sequence context, interchain cross-linking, if it occurs at all, is below the level of detection. The present studies, including those described in the accompanying paper[26] on the conformation of a duplex having an unfunctionalized trimethylene tether, provide an explanation for the sequence selectivity. The crotonaldehyde adduct of deoxyguanosine forms cross-links which are structurally identical to the acetaldehyde cross-links reported by Wang; cross-linking is observed in a CpG sequence context and is strongly dependent on the configuration of the stereogenic center derived from C3 of crotonaldehyde. The present study accounts for the stereoselectivity observed by Wang.

Experimental Section

Reference samples of nucleosides **2ab**, **5a**, and **6a** were prepared as previously described.[25,27,28] Modified oligonucleotides containing 1,N^2-hydroxypropano-2′-deoxyguanosine (**2a**) and the 6*R* and 6*S* methyl analogues **2b** were synthesized as described previously.[27,28] The ^{15}N-labeled samples of the complementary oligonucleotides in duplexes **B** and **C** (^{15}N label in the bold face guanine, Table 1) were prepared by treatment of analogous oligonucleotides containing 2-fluoro-O^6-(trimethylsilylethyl)-2′-deoxyguanosine with 6 M $^{15}NH_4OH$ (Isotec) for 24 h at 60 °C.[29] Unmodified oligonucleotides were purchased from Midland Certified Reagent Company, Midland, TX.

HPLC Separations. The purification of nucleosides and oligonucleotides and the analysis of reaction mixtures and mixtures of nucleosides obtained from enzymatic degradations were performed on a Beckman HPLC system (System Gold software, pump module 125) with a diode array UV detector (module 168) monitoring at 260 nm using YMC ODS-AQ columns (250 mm × 4.6 mm i.d., 1.5 mL/min for analysis and 250 mm × 10 mm i.d., 5 mL/min for purification) with H_2O and CH_3CN for nucleosides and 0.1 M aqueous ammonium formate and CH_3CN for oligonucleotides. *HPLC gradient*: 1−10% acetonitrile over 15 min, 10−20% acetonitrile over 5 min, hold for 5 min, 100% acetonitrile over 3 min, hold for 2 min, and then to 1% acetonitrile over 3 min. Cross-linked oligonucleotides eluted after 16−18 min (12−16% acetonitrile) followed by the uncross-linked species.

(17) Foiles, P. G.; Akerkar, S. A.; Miglietta, L. M.; Chung, F. L. *Carcinogenesis* **1990**, *11*, 2059−61.
(18) Foiles, P. G.; Akerkar, S. A.; Chung, F. L. *Carcinogenesis* **1989**, *10*, 87−90.
(19) Foiles, P. G.; Chung, F. L.; Hecht, S. S. *Cancer Res.* **1987**, *47*, 360−3.
(20) Gates, K. S. In *Comprehensive Natural Products Chemistry*; Barton, D., Nakanishi, K., Meth-Cohn, O., Eds.; Kool, E. T., Volume Ed.; Pergamon: Oxford, 1999; Vol. 7, Chapter 14, pp 491−552.
(21) Luce, R. A.; Hopkins, P. B. *Methods Enzymol.* **2001**, *340*, 396−412.
(22) Rajski, S. R.; Williams, R. M. *Chem. Rev.* **1998**, *98*, 2723−2796.
(23) Kawanishi, M.; Matsuda, T.; Nakayama, A.; Takebe, H.; Matsui, S.; Yagi, T. *Mutat. Res.* **1998**, *417*, 65−73.
(24) Wang, M.; McIntee, E. J.; Cheng, G.; Shi, Y.; Villalta, P. W.; Hecht, S. S. *Chem. Res. Toxicol.* **2000**, *13*, 1149−1157.
(25) Kozekov, I.; Nechev, L. V.; Sanchez, A.; Harris, C. M.; Lloyd, R. S.; Harris, T. M. *Chem. Res. Toxicol.* **2001**, *14*, 1482−1485.
(26) Dooley, P. A.; Zhang, M.; Korbel, G. A.; Nechev, L. V.; Harris, C. M.; Stone, M. P.; Harris, T. M. *J. Am. Chem. Soc.* **2002**, *124*, 62−72.
(27) Nechev, L. V.; Harris, C. M.; Harris, T. M. *Chem. Res. Toxicol.* **2000**, *13*, 421−429.
(28) Nechev, L. V.; Kozekov, I.; Harris, C. M.; Harris, T. M. *Chem. Res. Toxicol.* **2001**, *14*, 1506−1512.
(29) DeCorte, B. L.; Tsarouhtsis, D.; Kuchimanchi, S.; Cooper, M. D.; Horton, P.; Harris, C. M.; Harris, T. M. *Chem. Res. Toxicol.* **1996**, *9*, 630−637.

Table 1

Oligonucleotide Duplexes

A 5'-d(GCTAGC-**X**-AGTCC)
3'-d(CGATC**G**-C-TCAGG)

B 5'-d(CGTAC-**X**-CATGC)
3'-d(GCAT**G**-C-GTACG)

C 5'-d(CCTGA-**X**-CGATCG)
3'-d(GGACT-C-**G**CTAGC)

D 5'-d(GCTAGC-**X**-AGTCC)
3'-d(CGATCG-C-TCAGGATA)

E 5'-d(CCTGA-**X**-CGATCG)
3'-d(GGACT-C-GCTAGCATA)

Capillary Zone Electrophoresis. Electrophoretic analyses were carried out using a Beckman P/ACE instrument system 5500 series monitored at 260 nm on a 27 cm × 100 μm column packed with their 100-R gel (for ss-DNA) using the Tris-borate buffer system containing 7 M urea.

Mass Spectrometry. Low- and high-resolution FAB mass spectra were obtained at the Mass Spectrometry Facility at the University of Notre Dame, Notre Dame, IN. MALDI-TOF mass spectra (negative ion) of modified oligonucleotides were obtained on a Voyager Elite DE instrument (Perseptive Biosystems) at the Vanderbilt Mass Spectrometry Facility using a 3-hydroxypicolinic acid matrix containing ammonium hydrogen citrate (7 mg/mL) to suppress multiple sodium and potassium adducts. Unreduced cross-linked species were unstable in the MALDI matrix, rapidly reverting to the individual strands. To obtain satisfactory results, spectra had to be taken immediately after sample preparation. Even then, the signals for cross-linked species were weak relative to those of the individual strands.

NMR Spectroscopy. The spectra of adducted nucleosides were acquired on Bruker 300 and 400 MHz spectrometers. Spectra of ^{15}N-labeled oligonucleotide duplex **C** were acquired on a Bruker Avance 500 MHz spectrometer. For preparation of the sample, the two strands were combined in a 1:1.1 ratio of adducted and complement strands in 500 μL of buffer containing 0.1 M sodium chloride, 10 μm sodium phosphate, and 50 μM Na_2EDTA and then lyophilized. The oligonucleotide was then dissolved in 500 μL of 90:10 H_2O:D_2O and transferred to a standard 5 mm NMR tube (Wilmad, Buena, NJ). NMR experiments were carried out at 30 °C and 500.13 MHz proton resonance (^{15}N frequency 50.68 MHz). Decoupled $^1H-^{15}N$ dipole cross-relaxation (HSQC) spectra were measured using the standard inverse gradient pulse programs with Watergate suppression and States-TPPI mode. SW in the nitrogen dimension was set at 100 ppm with O3 set to 75 ppm. The pulse delay for $^1J_{1H-15N}$ was set at 90 Hz. All data were processed on Silicon Graphics (Mountain View, CA) Octane workstations using the Felix (v. 2000, Accelrys Inc., San Diego, CA) NMR processing software package.

Enzymatic Hydrolysis. Oligonucleotide (0.2−0.5 A_{260} units) was dissolved in 30 μL buffer (pH 7.0, 10 mM Tris-HCl, 10 mM $MgCl_2$) and incubated with DNase I, snake venom phosphodiesterase I, and alkaline phosphatase at 37 °C for 90 min. The mixture was analyzed by reverse phase HPLC. Adducted nucleosides were identified by comparison with authentic samples based on retention times, co-injection, and ultraviolet spectra.

2-(3-Hydroxypropyl)-2′-deoxyguanosine (7a). Compound **2a** (2.5 mg, 7.7 μmol) in 100 μL of 0.05 M, pH 7 phosphate buffer was treated with sodium borohydride (1.0 mg, 27.0 μmol). The mixture was stirred at room temperature for 10 min, then quenched with 0.5 mL of 5% acetic acid. HPLC purification gave 2.3 mg (93%) of **7a**. ^{1}H NMR (DMSO-d_6): δ (ppm) 10.15 (bs, 1H, N1H), 7.89 (s, 1H, H-8), 6.45 (m, 1H, N^{2}H), 6.15 (dd, 1H, H-1′, $J_1 = J_2 = 6.3$ Hz), 5.27 (d, 1H, 3′-OH, $J = 3.97$ Hz), 4.87 (t, 1H, 5′-OH, $J = 5.6$ Hz), 4.56 (bs, 1H, CH_2-OH), 4.35 (m, 1H, H-3′), 3.80 (m, 1H, H-4′), 3.55 (m, 1H, H-5′), 3.49 (m, 3H, H-5″ and CH_2-OH), 3.30 (2H, CH_2-NH buried under the water peak, identified by COSY), 2.61 (m, 1H, H-2′), 2.20 (m, 1H, H-2″), 1.68 (m, 2H, C-CH_2-C). HRMS (FAB+) *m*/*z* calcd for $C_{13}H_{20}N_5O_5$ $[M + H]^+$ 326.1464, found 326.1480.

3-(2-Deoxy-β-D-*erythro*-pentofuranosyl)-6(*R* and *S*)-methyl-5,6,7,8-tetrahydro-8-(*N*^{2}-deoxyguanosinyl)pyrimido[1,2-*a*]purin-10(3*H*)-one (5b). In two separate reactions, the 6*R* and 6*S* diastereomers of 3-(2-deoxy-β-D-*erythro*-pentofuranosyl)-8-hydroxy-6-methyl-5,6,7,8-tetrahydropyrimido[1,2-*a*]purin-10(3*H*)one[28] (**2b**, 2 mg, 5.9 μmol) and deoxyguanosine (3 mg, 11.0 μmol) were dissolved in DMSO (100 μL). The reaction mixtures were stirred at 100 °C for 8 days. The solvent was evaporated under vacuum. The residues were dissolved in 500 μL of pH 7.0, 0.05 M sodium phosphate buffer, and the products were isolated by HPLC. **6*R* diastereomer**: 1.0 mg (32%) ^{1}H NMR (DMSO-d_6): δ (ppm) 10.19 (bs, 1H, dG-N1H), 8.13 (s, 1H, NH-5), 7.97 (s, 1H, H-2), 7.93 (s, 1H, dG-H-8), 7.34 (m, 1H, dG-N^{2}H), 6.58 (m, 1H, H-8), 6.20 (t, 1H, H-1′, $J = 6.5$ Hz), 6.12 (t, 1H, H-1′, $J = 6.1$ Hz), 5.30 (d, 1H, 3′-OH, $J = 4.1$ Hz), 5.26 (d, 1H, 3′-OH, $J = 3.7$ Hz), 4.90 (m, 2H, 2 × 5′-OH), 4.35 (m, 2H, 2 × H-3′), 3.82 (m, 2H, 2 × H-4′), 3.65 (m, 1H, H-6), 3.53 (m, 4H, 2 × H-5′, 2 × H-5″), 2.68 (m, 1H, H-2′), 2.50 (1H, H-2′, buried under the DMSO peak, identified by COSY), 2.32 (m, 1H, H-7), 2.21 (m, 2H, 2 × H-2″), 1.62 (m, 1H, H-7), 1.24 (d, 3H, CH_3, $J = 6.4$ Hz). HRMS (FAB+) *m*/*z* calcd for $C_{24}N_{31}N_{10}O_8$ $[M + H]^+$ 587.2326, found 587.2308. **6*S* diastereomer**: 0.9 mg (29%) ^{1}H NMR (DMSO-d_6): δ (ppm) 10.19 (bs, 1H, dG-N1H), 8.12 (s, 1H, NH-5), 7.98 (s, 1H, H-2), 7.93 (s, 1H, dG-H-8), 7.33 (m, 1H, dG-N^{2}H), 6.57 (m, 1H, H-8), 6.21 (t, 1H, H-1′, $J = 6.7$ Hz), 6.12 (t, 1H, H-1′, $J = 6.3$ Hz), 5.30 (d, 1H, 3′-OH, , $J = 4.0$ Hz), 5.26 (d, 1H, 3′-OH, $J = 3.7$ Hz), 4.92 (m, 2H, 2 × 5′-OH), 4.36 (m, 2H, 2 × H-3′), 3.82 (m, 2H, 2 × H-4′), 3.65 (m, 1H, H-6), 3.55 (m, 4H, 2 × H-5′, 2 × H-5″), 2.67 (m, 1H, H-2′), 2.60 (m, 1H, H-2′), 2.32 (m, 2H, 1 × H-7 and H-2″), 2.18 (m, 1H, H-2″), 1.64 (m, 1H, H-7), 1.25 (d, 3H, CH_3, $J = 6.2$ Hz). HRMS (FAB+) *m*/*z* calcd for $C_{24}N_{31}N_{10}O_8$ $[M + H]^+$ 587.2326, found 587.2323.

2-(3-Hydroxy-1(*R* and *S*)-methylpropyl)-2′-deoxyguanosine (7b).[30] Each diastereomer of compound **2b** (1.8 mg, 5.3 μmol) in 100 μL of potassium phosphate buffer (0.05 M, pH 7) was treated with $NaBH_4$ (1.0 mg, 27.5 μmol). The mixture was stirred at room temperature for 10 min, quenched with 0.5 mL of 5% acetic acid, and purified by HPLC. **1*R* diastereomer**: 1.6 mg (88%) ^{1}H NMR (DMSO-d_6): δ (ppm) 10.33 (bs, 1H, N1H), 7.89 (s, 1H, H-8), 6.34 (d, 1H, NH, $J = 8.4$ Hz), 6.13 (dd, 1H, H-1′, $J_1 = J_2 = 7.5$ Hz), 5.28 (d, 1H, 3′-OH, $J = 4.1$ Hz), 4.86 (t, 1H, 5′-OH, $J = 5.6$), 4.52 (bs, 1H, CH_2-OH), 4.33 (m, 1H, H-3′), 4.02 (m, 1H, CH-CH_3), 3.81 (m, 1H, H-4′), 3.53 (m, 2H, H-5′, H-5″), 3.49 (m, 2H, CH_2-OH), 2.62 (m, 1H, H-2′), 2.20 (m, 1H, H-2″), 1.64 (q, 2H, CH-CH_2-C, $J = 6.5$ Hz), 1.16 (d, 3H, CH_3, $J = 6.5$ Hz). HRMS (FAB+) *m*/*z* calcd for $C_{14}H_{22}N_5O_5$ $[M + H]^+$ 340.1621, found 340.1638. **1*S* diastereomer**: 1.7 mg (94%) ^{1}H NMR (DMSO-d_6): δ (ppm) 10.31 (bs, 1H, N1H), 7.89 (s, 1H, H-8), 6.34 (d, 1H, NH, $J = 7.3$ Hz), 6.13 (dd, 1H, H-1′, $J_1 = J_2 = 6.5$ Hz), 5.28 (d, 1H, 3′-OH, $J = 3.8$ Hz), 4.87 (t, 1H, 5′-OH, $J = 5.2$), 4.53 (bs, 1H, CH_2-OH), 4.34 (m, 1H, H-3′), 4.02 (m, 1H, CH-CH_3), 3.81 (m, 1H, H-4′), 3.54 (m, 2H, H-5′, H-5″), 3.49 (m, 2H, CH_2-OH), 2.59 (m, 1H, H-2′), 2.20 (m, 1H, H-2″), 1.64 (q, 2H, CH-CH_2-C, $J = 6.5$ Hz), 1.16 (d, 3H, CH_3, J

(30) Gorelick, N. J.; Hutchins, D. A.; Tannenbaum, S. R.; Wogan, G. N. *Carcinogenesis* **1989**, *10*, 1579−1587.

= 6.5 Hz). HRMS (FAB+) *m*/*z* calcd for $C_{14}H_{22}N_5O_5$ $[M + H]^+$ 340.1621, found 340.1615.

2-(3-Amino-1(*R* and *S*)-methylpropyl)-2′-deoxyguanosine. Solutions of the 6*R* and 6*S* diastereomers of **2b** (1.8 mg, 5.3 μmol), 50 μL of ammonium hydroxide and 100 μL of methanol were stirred for 24 h at room temperature followed by treatment with $NaBH_4$ (1.0 mg, 27.5 μmol) for 10 min. The reactions were quenched with 0.5 mL of 5% acetic acid and the products isolated by HPLC. **1*R* diastereomer (acetate salt)**: 1.7 mg (80%) ^{1}H NMR (DMSO-d_6): δ (ppm) 7.87 (s, 1H, H-8), 6.13 (dd, 1H, H-1′, $J_1 = J_2 = 6.3$ Hz), 4.35 (m, 1H, H-3′), 4.00 (m, 1H, C<u>H</u>-CH$_3$), 3.80 (m, 1H, H-4′), 3.55 (dd, 1H, H-5′, $J_1 =$ 5.2 Hz, $J_2 =$ 11.5 Hz), 3.48 (dd, 1H, H-5″, $J_1 = 5.2$ Hz, $J_2 = 11.5$ Hz), 2.65 (m, 2H, C<u>H</u>$_2$-NH$_2$), 2.60 (m, 1H, H-2′), 2.20 (m, 1H, H-2″), 1.80 (s, 3H, CH$_3$COO), 1.61 (q, 2H, CH-C<u>H</u>$_2$, $J = 6.9$ Hz), 1.15 (d, 3H, CH$_3$, $J = 6.3$ Hz), N1H and OH signals not observed. HRMS (FAB+) *m*/*z* calcd for $C_{14}H_{23}N_6O_4$ $[M + H]^+$ 339.1781, found 339.1787. **1*S* diastereomer (acetate salt)**: 1.5 mg (70%) ^{1}H NMR (DMSO-d_6): δ (ppm) 7.87 (s, 1H, H-8), 6.13 (dd, 1H, H-1′, $J_1 = J_2 =$ 6.4 Hz), 4.35 (m, 1H, H-3′), 4.00 (m, 1H, C<u>H</u>-CH$_3$), 3.80 (m, 1H, H-4′), 3.55 (dd, 1H, H-5′, $J_1 = 5.2$ Hz, $J_2 = 11.5$ Hz), 3.48 (dd, 1H, H-5″, J_1 = 5.2 Hz, $J_2 = 11.7$ Hz), 2.65 (m, 2H, C<u>H</u>$_2$-NH$_2$), 2.60 (m, 1H, H-2′), 2.20 (m, 1H, H-2″), 1.80 (s, 3H, CH$_3$COO), 1.60 (q, 2H, CH-C<u>H</u>$_2$, J = 6.9 Hz), 1.15 (d, 3H, CH$_3$, $J = 6.1$ Hz), N1H and OH signals not observed. HRMS (FAB+) *m*/*z* calcd for $C_{14}H_{23}N_6O_4$ $[M + H]^+$ 339.1781, found 339.1756.

1-(*R* and *S*)-Methyl-1,3-bis(2′-deoxyguanosin-*N*^{2}-yl)propane (6b). The 1*R* and 1*S* diastereomers of 2-(3-amino-1-methylpropyl)-2′-deoxyguanosine were added to mixtures of *O*^{6}-trimethylsilylethyl-2-fluoro-2′-deoxyinosine (2.2 mg, 6.4 μmol), DMSO (100 μL), and diisopropylethylamine (50 μL). The mixtures were stirred at 55 °C for 12 h. The reactions were stopped, and the solvents were evaporated under vacuum. The residues were dissolved in 5% acetic acid (1 mL) and stirred at room temperature for 1 h followed by HPLC purification. **1*R* diastereomer**: 2.0 mg (80%) ^{1}H NMR (DMSO-d_6): δ (ppm) 10.30 (bs, 2H, 2 × N1H), 7.88 (bs, 2H, 2 × H-8), 6.40−6.47 (bd, 2H, N<u>H</u>-CH, N<u>H</u>-CH$_2$), 6.14 (m, 2H, 2 × H-1′), 5.27 (d, 2H, 2 × 3′-OH), 4.87 (t, 2H, 2 × 5′-OH), 4.34 (m, 2H, 2 × H-3′), 3.98 (m, 1H, C<u>H</u>-CH$_3$), 3.81 (m, 2H, 2 × H-4′), 3.54 (m, 4H, 2 × H-5′, 2 × H-5″), 3.30 (m, 2H, C<u>H</u>$_2$-NH buried under the water peak, identified by COSY), 2.60 (m, 2H, 2 x H-2′), 2.20 (m, 2H, 2 × H-2″), 1.77 (m, 2H, CH-C<u>H</u>$_2$-CH$_2$), 1.21 (d, 3H, CH$_3$, $J = 6.4$ Hz). HRMS (FAB+) *m*/*z* calcd for $C_{24}H_{33}N_{10}O_8$ $[M + H]^+$ 589.2483, found 589.2462. **1*S* diastereomer:** 1.2 mg (54%) ^{1}H NMR (DMSO-d_6): δ (ppm) 10.31 (bs, 2H, 2 × N1H), 7.87 (bs, 2H, 2 × H-8), 6.36−6.43 (bd, 2H, N<u>H</u>-CH, N<u>H</u>-CH$_2$), 6.11 (m, 2H, 2 × H-1′), 5.26 (d, 2H, 2 × 3′-OH), 4.86 (t, 2H, 2 × 5′-OH), 4.34 (m, 2H, 2 × H-3′), 3.99 (m, 1H, C<u>H</u>-CH$_3$), 3.80 (m, 2H, 2 × H-4′), 3.53 (m, 4H, 2 × H-5′, 2 × H-5″), 3.30 (m, 2H, C<u>H</u>$_2$-NH buried under the water peak, identified by COSY), 2.60 (m, 2H, 2 × H-2′), 2.21 (m, 2H, 2 × H-2″), 1.77 (m, 2H, CH-C<u>H</u>$_2$-CH$_2$), 1.21 (d, 3H, CH$_3$, $J = 6.7$ Hz).

The two diastereomers of compound **5b** (0.1 mg, 0.17 μmol) in 100 μL of methanol/water (1:1) were treated with $NaBH_4$ (0.1 mg, 2.7 μmol) for 20 h at room temperature. The mixtures were quenched with 0.1 mL of 5% acetic acid followed by HPLC purification. The products were identical with samples of the diastereomers of **6b** synthesized above.

Cross-Linking Reactions. In a typical reaction, a mixture of equimolar quantities (1−10 A_{260} units) of the adducted and complementary oligonucleotides in 20−100 μL of potassium phosphate buffer (0.05 M, pH 7.0) containing 0.1 or 1.0 M KCl was incubated at 37 °C. The reaction was analyzed periodically by HPLC or CZE or both. In the HPLC analysis the cross-linked species eluted ahead of the starting oligonucleotides, whereas in CZE the cross-linked species eluted later. Relative amounts of remaining starting materials and cross-linked product(s) were estimated from peak areas. Conversion of peak areas to relative molar amounts was made using molar absorptivities calculated on the basis of sequence using the method of Borer.[31] The molar absorptivity of nucleosides **2a** and **2b** were assumed to be identical to that of dG and those of bis-nucleosides **5a** and **5b** to be twice that of dG. The molar absorptivity contributions of the individual nucleosides were assumed to be identical for single-stranded and cross-linked oligonucleotides, that is, no allowance was made for the hypochromicity of cross-linked duplexes because of uncertainty as to the extent of denaturation under the HPLC and CZE conditions. The consequence of this assumption is that the extent of cross-linking may be slightly greater than the calculated values. In preparative experiments, the cross-linked oligonucleotide duplexes were collected, lyophilized, and then stored at −20 °C until used for reduction or analysis. Bis-nucleosides **5a** and **5b** were identified by HPLC comparison with authentic samples.

Reduction of the Cross-Linked Duplexes. Aliquots of cross-linked oligonucleotide (~1.0 A_{260} units) were treated with $NaCNBH_3$ (1.0 mg) in 30 μL of potassium phosphate buffer (pH 7.0, 0.05 M) for 24 h at room temperature. The reactions were quenched with 5% acetic acid (100 μL), and the reduced oligonucleotides were purified by HPLC. With oligonucleotide **A** (X = **2a**), enzymatic digestion of the reduced material gave 1,3-bis(2′-deoxyguanosin-*N*^{2}-yl)propane (**6a**) and 2-(3-hydroxypropyl)-2′-deoxyguanosine (**7a**), in addition to the four normal deoxynucleosides in the expected proportions. Compounds **6a** and **7a** were identified by coelution with authentic samples.[25] From oligonucleotides **A** (X = **2b**) containing the 6*R* and *6S* diastereomers of crotonaldehyde adduct **2b**, the corresponding diastereomers of bis-nucleoside **6b** were obtained in a similar fashion. Small amounts of reduction products **7b** were also observed. The two diastereomers of **6b** and **7b** were identified by comparison with authentic samples.

Melting Studies. Equal amounts of the modified oligonucleotides and the complementary strands (0.5 A_{260} units each) were dissolved in 1 mL of melting buffer (10 mM Na_2HPO_4/NaH_2PO_4, 1.0 M NaCl, 50 μM Na_2EDTA, pH 7.0). UV measurements (260 nm) were taken at 1-min intervals with a 1 °C/min temperature gradient. The temperature was cycled between 5 and 95 °C. The first derivative of the melting curve was used to establish melting transition (T_m) values.

Results

Cross-Linking by Acrolein Adduct 2a in a CpG Sequence Context. We carried out our cross-linking studies with oligonucleotides containing the acrolein and crotonaldehyde adducts **2a**−**b**, which had been prepared by published methodology.[27,28] A CpG sequence context seemed to be the most promising one with which to initiate the studies because several other types of cross-links have shown preference for formation in CpG sequences.[32−34] The 12-mer:12-mer duplex **A** (X = **2a**) (Table 1) in which the adducted strand contained the 8-hydroxypropano adduct in the sequence 5′-d(GCTAGC-**2a**-AGTCC) was incubated at 37 °C in a pH 7.0 potassium phosphate buffer containing 1 M KCl. The course of the reaction was monitored by reverse-phase HPLC with gradient elution using acetonitrile−0.1 M aqueous ammonium formate. The HPLC solvent mixture was sufficiently denaturing that signals for the individual strands were fully resolved from each other. Initially, the sample showed only the two signals for the starting oligonucleotides, but after 3 days a new, faster-moving component was observed (Figure 1, panel A). The reaction reached equilibrium within 7 days, at which point the new component represented ~50% of the mixture. Cross-link formation was also monitored by CZE using

(31) Borer, P. N. *Handbook of Biochemistry and Molecular Biology*, 1st ed.; CRC Press: Cleveland, 1975; Notes, p 359.
(32) Borowy-Borowski, H.; Lipman, R.; Tomasz, M. *Biochemistry* **1990**, *29*, 2999−3006.
(33) Hopkins, P. B. *Adv. DNA Sequence Specific Agents* **1996**, *2*, 217−239.
(34) Weidner, M. F.; Sigurdsson, S. T.; Hopkins, P. B. *Biochemistry* **1990**, *29*, 9225−9233.

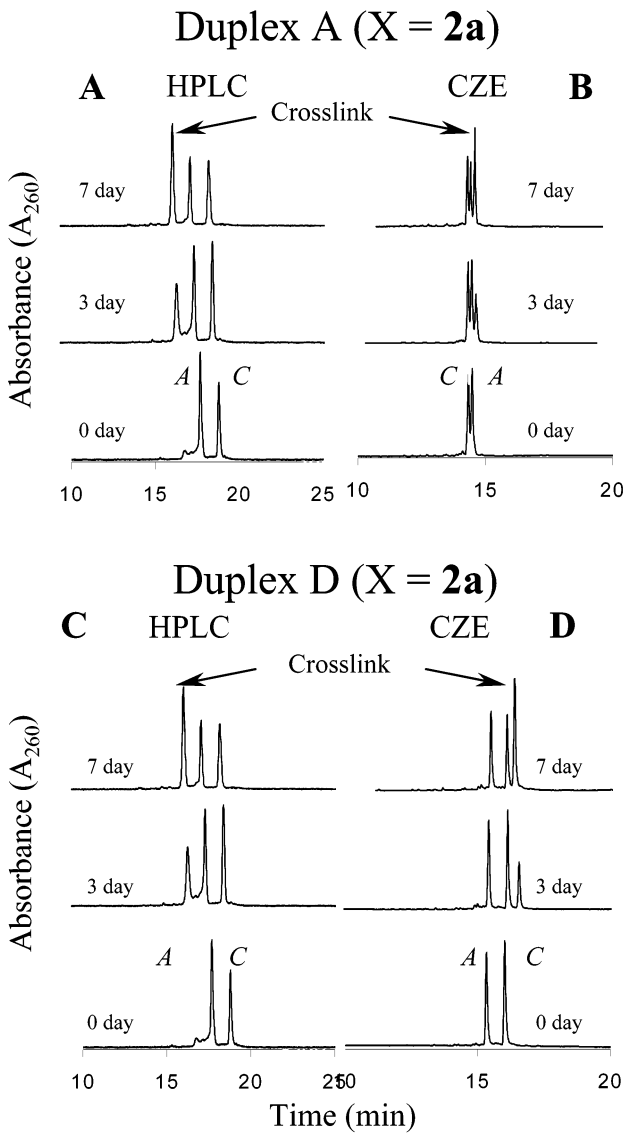


Figure 1. Cross-linking analysis on acrolein-adducted CpG duplexes **A** and **D**. (Panels A and B) HPLC and CZE results, respectively, with 12-mer:12-mer duplex **A**. (Panels C and D) HPLC and CZE results, respectively, with 12-mer:15-mer duplex **D**. The reactions were carried out at 37 °C as described in the Experimental Section. The bottom, middle and top tiles are 0, 3, and 7 days of reaction. The cross-linked duplex elutes first from the HPLC and last from the CZE. The peaks for the adducted and complementary strands are identified by the letters *A* and *C*, respectively.

a denaturing buffer to distinguish covalent links from Watson−Crick duplex formation. Again, the starting oligonucleotides were resolved from each other, and gradual formation of the new component was again observed; however, the temporal resolution of the three components by CZE was not as good as that by HPLC. (Figure 1, panel B).

The new component was isolated using HPLC. It was relatively stable in pH 7.0 buffer but rapidly (<1 h) reverted to starting materials in unbuffered water. MALDI-TOF mass spectrometry of the new species gave two ions, *m*/*z* 7345.2 and 7327.8, consistent with a cross-linked duplex containing **3a** ([M − H]$^-$ *m*/*z* 7345.3) plus **4a** and/or **5a** ([M − H]$^-$ *m*/*z* 7327.3). Rapid reversal of the cross-link in the MALDI matrix made the spectrum difficult to obtain and gave, at best, only weak signals for the cross-linked species along with intense signals for the starting oligonucleotides.

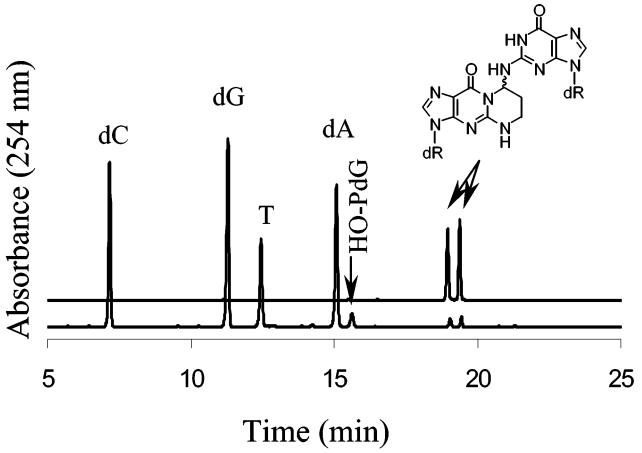


Figure 2. HPLC chromatogram of the products of enzymatic digestion of the cross-link formed by acrolein-adducted duplex **A**. Digestion procedure is provided in the Experimental Section. The diastereomers of **5a** elute at 19.1 and 19.5 min. Nucleoside **2a** elutes at 15.6 min. The chromatogram for the authentic sample of the diastereomers of **5a** is overlaid on the chromatogram.

Linkage of the acrolein moiety to N^2 of guanine was demonstrated by enzymatic degradation with DNase I, snake venom phosphodiesterase, and alkaline phosphatase to produce bis-nucleoside **5a** plus the starting mononucleoside **2a** (resulting from partial hydrolysis during the procedure) along with the normal four deoxynucleosides in the appropriate relative amounts (Figure 2). Bis-nucleoside **5a** chromatographed as two peaks of approximately equal area, reflecting the fact that it was a slowly equilibrating mixture of the two diastereomers at C8 of the pyrimidopurinone. Mononucleoside **2a**, although also a mixture of two diastereomers, appeared as only a single peak. The structure of bis-nucleoside **5a** was established by comparison with an authentic sample, prepared by the reaction of **2a** with deoxyguanosine in DMSO.[25] This reaction also gave a mixture of the two diastereomers. Condensation of the two nucleosides to form **5a** was much slower than the cross-linking reaction in the oligonucleotide duplex, requiring 6 days at 100 °C to reach equilibrium, where the yield was 65%. The structure of the bis-nucleoside was established spectroscopically. MS showed it was **4a** or **5a** rather than undehydrated carbinolamine **3a**. NMR studies showed it was pyrimidopurinone **5a** rather than imine **4a**. The chemical shift (6.57 ppm) of H8 on the pyrimidopurinone ring system was consistent with tetrahedral hybridization rather than trigonal. Furthermore, an HMBC spectrum established proximity of H8 to C10 (Figure 3). No signal was observed that would be consistent with the imine proton of **4a**. However, **4a** must be present in equilibrium with pyrimidopurinone **5a**, although the amount present is too small to be detected by NMR, because treatment of **5a** with $NaBH_4$ gave bis-nucleoside **6a**. The structure of **6a** was established by comparison with a sample prepared by reaction of 1,3-diaminopropane with the O^6-trimethylsilylethyl ether of 2-fluoro-2′-deoxyinosine.[35]

Cross-linked duplex **A** could be reduced in near quantitative yield with $NaCNBH_3$. The reaction was slow, requiring an extended reaction period (>16 h) to go to completion. MALDI-TOF MS of the reduced, cross-linked duplex showed an ion at

(35) Dooley, P. A.; Tsarouhtsis, D.; Korbel, G. A.; Nechev, L. V.; Shearer, J.; Zegar, I. S.; Harris, C. M.; Stone, M. P.; Harris, T. M. *J. Am. Chem. Soc.* **2001**, *123*, 1730−1739.

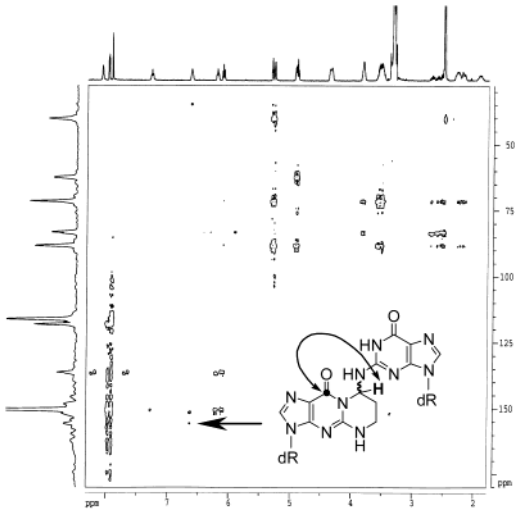


Figure 3. $^{13}C-^{1}H$ HMBC spectrum of mixed diastereomers of **5a**. The correlation between H8 and C10 of the pyrimidopurinone is noted on the spectrum.

7330.2, consistent with formation of **6a** in the duplex (calcd for $[M - H]^-$: *m*/*z* 7329.3). Enzymatic digestion gave bis-nucleoside **6a**, the structure of which was established by comparison with the authentic sample. Overall, the experiment established that adduct **2a** is capable of forming an interstrand cross-link with the exocyclic amino group of guanine in a CpG sequence context. The fact that the cross-link can be reduced with $NaCNBH_3$ means that it is, at least in part, in the imine form. However, the reduction is slow, requiring up to 24 h to reach completion. Our interpretation of this observation is that the cross-link is in the carbinolamine form and undergoes reduction only after eliminating water to form the imine.

Cross-Linking by the Acrolein Adduct 2a in a GpC Sequence Context. As a probe of sequence preference for formation of interchain cross-links, studies were made of cross-linking using duplex **B**, formed from the 11-mer 5′-d(CGTAC-**2a**-CATGC) with its complement. Duplex **B** has the potential to form a cross-link in either a CpG or a GpC context. In view of the fact that cross-linking is a reversible process, identification of the cross-linked product would establish the thermodynamically favored sequence context. To identify the site(s) of cross-linking, the complement strand, 3′-d(GCAT**G**CGTACG), was prepared with N^2 of the guanine bearing the ^{15}N label denoted in bold-faced type in Table 1. Again, immediately after preparation of the duplex, the adducted and unadducted strands eluted individually from the HPLC column. A signal for the cross-linked species arose over a period of ~5 days at 37 °C, with the new component eluting ahead of the individual strands. The cross-linking yield was similar to that obtained with duplex **A**. The cross-linked oligonucleotide derived from duplex **B** was isolated by preparative HPLC and analyzed by MALDI-TOF MS, giving weak ions at *m*/*z* 6726.7 and 6710.1, consistent with cross-linked duplexes containing **3a** plus **4a** or **5a** (or both), respectively. Enzymatic digestion gave bis-nucleoside **5a**, mononucleoside **2a** arising from partial hydrolysis of **5a,** plus the normal deoxynucleosides. Electrospray MS of **5a** showed the molecular ion ($M + H^+$) of the bis-nucleoside was *m*/*z* 574.2, consistent with **5a** containing ^{15}N (Figure 4, panel C). A small signal, no more than 2%, at 573.2 (^{14}N-**5a)** is ascribed to the residual ^{14}N in the ^{15}N-ammonia that had been used for synthesis

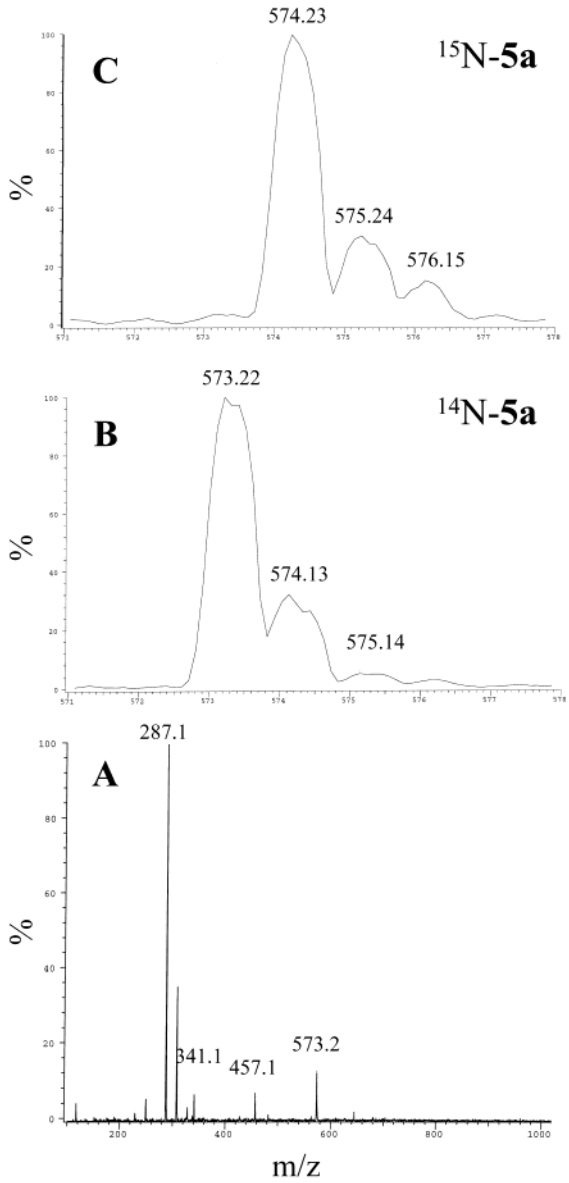


Figure 4. Electrospray mass spectra of bis-nucleoside **5a**. (Panel A) Full spectrum of **5a** with normal isotopic abundance. (Panel B) Expansion of the region of the parent ion in panel A. (Panel C) Region of the parent ion for **5a** derived from ^{15}N-substituted duplex **B**.

of the ^{15}N-labeled oligonucleotide. For comparison, panel A shows the spectrum of unlabeled **5a** with the region of the parent ion expanded in panel B. Consequently, we conclude that cross-linking occurs exclusively in a CpG context.

The next question to be addressed was whether cross-linking by adduct **2a** would occur in a GpC context if the more favorable CpG cross-linking opportunity were not present. A study was carried out using duplex **C,** which has the reversed sequence of oligonucleotide **A**. Duplex **C** was incubated and analyzed in the same manner as that used in the studies of oligonucleotides **A** and **B**. An HPLC analysis carried out as quickly as possible (<5 min) after the reaction mixture had been prepared showed the peaks for the two starting oligonucleotides plus a large new signal eluting earlier, similar to that of the cross-linked material observed in the CpG 12-mer experiment after 5 days incubation. A second HPLC analysis made after 40 min showed only marginal enhancement of the new signal and no further changes could be observed by HPLC over a period of 22 days. CZE

analysis was uninformative because the adducted and unadducted strands and any cross-linked species that might have been present failed to resolve from one another, giving only a single signal. The MALDI-TOF mass spectrum of the reaction mixture showed no signal for cross-linked species, only the signals for the individual strands. Enzymatic digestion provided no indication of bis-nucleoside **5a**; only 8-hydroxypropano adduct **2a** and the normal deoxynucleosides were observed. Reduction with $NaCNBH_3$ followed by enzymatic digestion failed to give bis-nucleoside **6a**; only N^2-(3-hydroxypropyl)deoxyguanosine (**7a**), the reduction product of starting nucleoside **2a**, was observed along with the normal nucleosides. Thus, HPLC results suggested a cross-link was being formed extremely rapidly in the GpC sequence context, but all other probes of cross-linking failed to provide confirmation.

The cross-linking reaction was reexamined with 12-mer:15-mer duplexes **D** and **E** (X = **2a**). With duplex **D** having the CpG sequence, the cross-linking reaction was monitored via HPLC analysis, giving results essentially identical to those obtained with duplex **A** (Figure 1, panel C). Likewise, CZE at the initiation of the cross-linking reaction showed individual signals for the two oligonucleotides. The signal for the cross-linked species gradually arose, reaching steady state within 5 days at 37 °C. The CZE results were identical to the observations made using duplex **A** except that the three signals were better resolved (Figure 1, panel D). It can be concluded that the conditions for the HPLC and CZE separations of the components present with CpG duplex **D** were adequately denaturing for the uncross-linked duplex formed from the two starting oligonucleotides to dissociate and for the two strands to elute individually. Furthermore, using either the HPLC or the CZE method, the cross-linked duplex was well-resolved from both of the starting materials.

The potential for cross-linking in the GpC sequence context was then reexamined using 12-mer:15-mer duplex **E** (X = **2a**). By HPLC, three signals were already present immediately after the sample was prepared; no changes occurred subsequently (Figure 5, panel A). The apparent cross-link was isolated; immediate reanalysis by HPLC gave the same ratio of species. Dilution of the sample did not alter this ratio. No imine could be trapped by reduction with $NaCNBH_3$. Whereas CZE of 12-mer:12-mer duplex **C** had been uninformative because the two starting oligonucleotides were not resolved from each other or from the putative cross-linked duplex, the 12- and 15-mers of duplex E were fully resolved from each other. Analysis of the reaction mixture by CZE immediately after sample preparation showed only the two signals for the individual oligonucleotides, and no changes could be detected thereafter (Figure 5, panel B).

From these results we conclude that imine is not being formed in the GpC sequence context, but we cannot exclude the possibility that a carbinolamine cross-link is being formed in a rapid but highly reversible manner in the GpC sequence context. The cross-link might stabilize the duplex sufficiently so that dissociation would be suppressed under HPLC conditions but insufficiently to protect the duplex from dissociation under the more completely denaturing conditions employed for CZE. In addition, the conditions present during mass spectrometry and enzymatic digestion might promote dissociation. Furthermore, the carbinolamine might be sterically incapable of eliminating

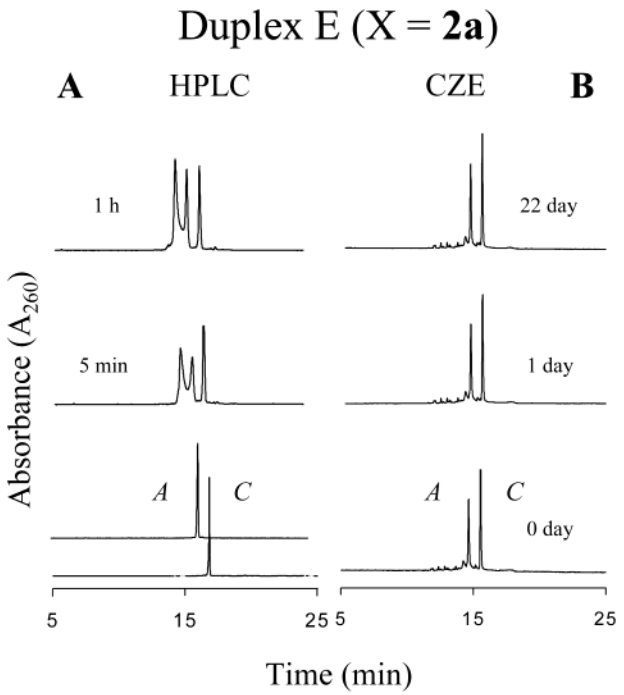


Figure 5. Cross-linking analysis on acrolein-adducted GpC duplex **E**. (Panel A) HPLC analysis. Bottom tile: Overlaid chromatograms of the individual strands. Middle tile: Results obtained immediately after mixing. Top tile: Results after 1 h. No changes were observed thereafter. (Panel B) CZE analysis. The bottom, middle, and top electropherograms are for 0, 1, and 22 days of reaction at 37 °C. See Experimental Section for details. The anomalous peak in the HPLC analysis appears first with a retention time of 15.8 min. The peaks for the adducted and complementary strands are identified by the letters *A* and *C*, respectively.

water to form the imine, leading to failure of the reduction with $NaCNBH_3$.

Melting Phenomena. As another probe of cross-link formation, duplex stability was examined as a function of temperature since it can be expected that the presence of a covalent cross-link would greatly enhance stability. The increase in absorbance at 260 nm that occurs as double-stranded DNA dissociates into single strands was monitored as the sample was heated from 5 to 90−95 °C. Reassociation was monitored as the samples were recooled to 5 °C. Concentrations were chosen to give total absorbances of ∼1 absorbance unit which produced hyperchromicity values of ∼0.2 absorbance units. The temperature was raised and lowered at 1.0 °C/min. The unadducted 12-mer:12-mer duplex **A** (X = dG) containing the CpG structural motif gave a T_m of 65 °C (Figure 6, panel A1); unadducted duplex **B** (X = dG) containing the GpC sequence melted at 67 °C (panel B1). In both cases the melting curves were highly reversible.

With adduct **2a** in place, 12-mer:12-mer CpG duplex **A** gave a melting transition of 55 °C immediately after preparation (panel A2). The measurement was repeated after allowing the sample to age. After 7 days at ambient temperature, a major melting transition at 55 °C was still observed, but a second transition could be seen above 90 °C. In two separate experiments, melting analysis after the samples were allowed to age for an additional 21 days at ambient temperature or for one week at 37 °C gave the 55 °C transition plus a second transition at 91 °C. The latter represented ∼30% of the total hyperchromicity (panel A3). In each case during the cooling cycle, the high-melting transition was almost completely absent. If after

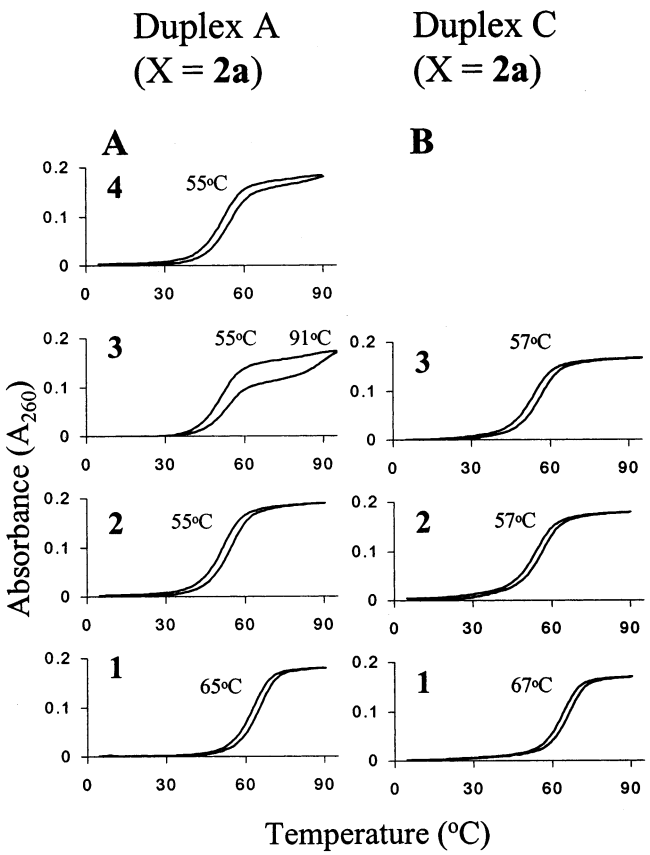


Figure 6. Melting analysis (T_m) of acrolein-adducted duplexes **A** and **C**. The incubations were carried out at 37 °C. Details of the incubations and analyses are provided in the Experimental Section. In each case the trace for the heating cycle lies below the trace for the cooling cycle. The T_m values are noted on the plots. (Panel A) CpG duplex **A**. Tile 1: Unadducted duplex **A** (X = dG). Tile 2: Adducted duplex **A** (X = **2a**) immediately after mixing. Tile 3: Same sample after 7 days at 37 °C. Tile 4: Repeat of T_m analysis immediately after the measurement shown in tile 3. (Panel b) GpC duplex **C**. Tile 1: Unadducted duplex **C** (X = dG). Tile 2: Adducted duplex **C** (X = **2a**) immediately after mixing. Tile 3: Same sample after 22-day incubation.

the analysis was complete the sample was immediately reheated from 5 to 95 °C, the second transition was largely gone (panel A4). Nevertheless, the 91 °C transition reappeared after the sample has been allowed to stand for several days at 37 °C, demonstrating the reversibility of the phenomena. It should be noted, though, that the sample is not infinitely reversible. As might be expected with an aldehyde, heating at 90 °C for extended periods (>1 h) caused gradual deterioration of the sample with CZE revealing the appearance of new, as yet, unidentified components.

Melting studies were repeated with CpG duplexes **B** and **D** where the results mimicked those with duplex **A**. Freshly prepared adducted duplexes **B** and **D** had T_m values of 49 and 56 °C, respectively. After aging, both showed two transitions, the second one appearing at 85 °C with **B** and at 91 °C with **D**. In both cases the second transition was largely gone during the cooling cycle.

Our interpretation of the melting experiments with adducted duplexes **A**, **B**, and **D** is that the cross-linked duplex in the aged samples has a T_m in the region of 90 °C but that thermally induced dissociation of the duplex causes rapid cleavage of the cross-link. Below the T_m , the cross-linking reaction is unimolecular which favors cross-link formation. Above the T_m, the cross-linking reaction is bimolecular and disfavored by low concentrations of reactants. Cross-link dissociation is unimolecular in both cases but much faster at elevated temperatures.

The GpC-type duplexes **C** and **E** were then studied. Adducted duplex **C** (X = **2a**) had a T_m of 57 °C immediately after preparation (Figure 6, panel B2). No change in melting behavior was observed as the sample aged. Panel B3 shows the result after 22 days at 37 °C (panel B3). The melting curve was similar to those of unadducted duplexes, showing only minimal hysteresis between the heating and cooling cycles. Repetition of these experiments with 12-mer:15-mer duplex **E** gave comparable results. These results provide additional evidence that cross-links are not being formed in the GpC sequence context.

NMR Studies with GpC Duplex C. Finally, unambiguous evidence was obtained that interchain carbinolamine-type cross-linking does not occur in the GpC sequence context using an NMR strategy recently reported by Kim et al.[36] They employed $^{15}N-^{1}H$ HSQC experiments on CpG duplex **A** in which the target guanine in the complementary strand (indicated in boldface in Table 1) bore ^{15}N substitution in the exocyclic amino group.[37] Uncross-linked duplexed species present in the sample gave dual correlations for the two protons on the $^{15}NH_2$ groups. For the cross-linked duplex, a pair of single correlations was observed between the ^{15}N and the N−H proton of the diastereomeric carbinolamines. An HSQC-filtered TOCSY experiment showed correlation between the ^{15}NH proton and the C−H proton of the carbinolamine as well as longer-range correlations of the amino proton with the both methylene groups.

In the present case, a similar strategy was employed with duplex **C** (X = **2a**) having ^{15}N substitution in the exocyclic amino group of guanine indicated in bold face in Table 1. The HSQC spectrum is shown in Figure 7. The experiment was carried out with ^{15}N decoupling to simplify the spectrum in the ^{1}H dimension. The HSQC spectrum showed that several species were present in the sample, the structures of which are presently under investigation. However, both protons are still present on the ^{15}N-labeled amino group in all of them, meaning the signals arise from duplexes that lack interchain cross-links to the labeled amino group. The pairs of protons are separated from each other by ~1.5 ppm, and their chemical shifts are consistent with their being amino groups involved in Watson−Crick base-pairing. The hydrogen-bonded protons (region A) are downfield of the non-hydrogen-bonded protons (region B). The signal in region C arises from the single-strand form of the ^{15}N-labeled oligonucleotide, which had been used in 10% excess. The signal appears as a singlet because of rapid exchange between the two amino protons. No signal consistent with a carbinolamine cross-link was observed in the HSQC spectrum.

Further confirmation of the absence of cross-linking was obtained from an HSQC-filtered NOESY spectra (not shown), which showed no correlation of any ^{15}NH proton signal with carbon-bound protons in the acrolein-derived tether. This result eliminates the possibility of rapidly reversible interchain carbinolamine-type cross-links being present in the sample. It

(36) Kim, H.-Y. H.; Voehler, M.; Harris, T. M.; Stone, M. P. *J. Am. Chem. Soc.* **2002**, *124*, 9324−9325.
(37) de los Santos et al. have also studied the structure of duplex **B**. Their study used unlabeled material and showed the presence of **1a** (and the hydrate of the aldehyde) in the duplex but did not reveal the presence of carbinolamine. de los Santos, C.; Zaliznyak, T.; Johnson, F. *J. Biol. Chem.* **2001**, *276*, 9077−9082.

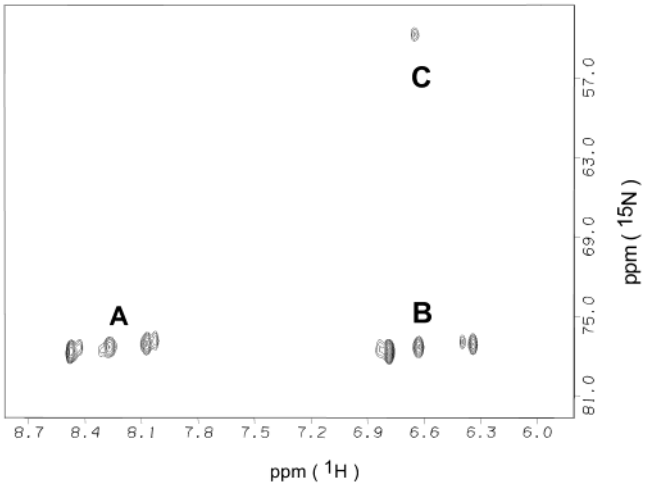


Figure 7. NMR evidence for the absence of an interstrand cross-link in the GpC sequence of duplex **C**. The decoupled ^{15}N HSQC spectrum observed at pH = 7.1. Peaks in regions A and B are the hydrogen-bonded and non-hydrogen-bonded N^2 protons, respectively, of the labeled deoxyguanosine in the complementary strand. Region C arises from excess complementary strand having equivalent, non-hydrogen-bonded protons. No other resonances were detected.

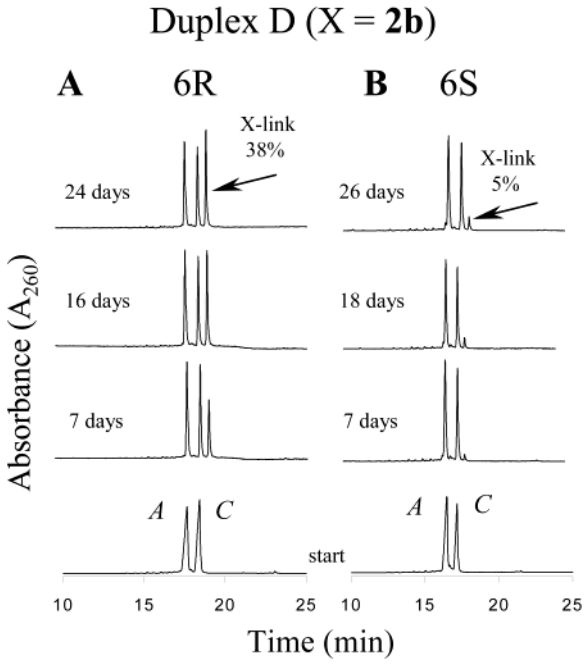


Figure 8. CZE analysis of cross-linking by crotonaldehyde-adducted duplex **D** (X = **2b**). (Panel A) Tiles 1−4: Duplex **D** (X = 6*R*-**2b**) after 0, 7, 16, and 24 days, respectively. (Panel B) Tiles 1−4: Duplex **D** (*X* = 6*S*-**2b**) after 0, 7, 18, and 26 days, respectively. Reactions carried out at 37 °C as described in the Experimental Section. The peaks for the adducted and complementary strands are identified by the letters *A* and *C*, respectively.

confirms our previous conclusion that the anomalous results obtained by HPLC with duplexes **C** and **E** arose from incomplete denaturation under the solvent conditions employed for elution.

Cross-Linking by Crotonaldehyde Adduct 2b. Cross-linking experiments were carried out using the two diastereomeric guanine adducts (**2b**) of crotonaldehyde[28] in CpG-type 12-mer:15-mer duplex **D** with monitoring by CZE. The cross-linking reaction was significantly slower than with the acrolein adduct, requiring 2−3 times longer to reach equilibrium. With the duplex containing the 6*R* diastereomer, a slower-moving peak for the cross-linked species appeared (Figure 8, panel A). After 21 days at 37 °C, the extent of its formation was ∼38%. Enzymatic digestion yielded bis-nucleoside **5b**. Only a single diastereomer was obtained. It was identified by HPLC comparison with an authentic sample of the 6*R* diastereomer of **5b**, prepared by reaction of deoxyguanosine with the 6*R* adduct of deoxyguanosine. The ^{1}H NMR spectrum compared well with that reported by Wang et al.[24] for the cross-link formed by the reaction of acetaldehyde with calf thymus DNA.

In sharp contrast with the results for the 6*R* diastereomer, the oligonucleotide containing the 6*S* diastereomer was only meagerly reactive with its complement. After 21 days at 37 °C, the reaction was at equilibrium, but cross-link formation was only ∼5% (Figure 8, panel B). Enzymatic digestion of the cross-linked component of the mixture gave bis-nucleoside **5b**, identical to **5b** prepared by reaction of the 6*S* diastereomer of **2b** with deoxyguanosine.

Melting studies carried out using freshly prepared 12-mer:15-mer duplexes **D** containing the *R* and *S* diastereomers of **2b** gave melting transitions of 51 °C, ∼5 °C lower than the analogous duplex containing the acrolein adduct. After standing 21 days at 37 °C, the duplex containing the 6*R* diastereomer exhibited two transitions, one at 49 °C and a second at 88 °C, the latter reflecting the presence of the cross-link (Figure 9, panel A2). The 6*S* diastereomer showed the transition at 50 °C and only a trace of the higher transition (panel B2).

Stability of Bis-nucleosides 5a−b. The pyrimidopurine-type bis-nucleosides **5a−b** are in equilibrium with imine-type bis-nucleosides **4a−b**. However, the equilibrium lies so far toward the pyrimidopurinones that the imines were below the level of detection in NMR samples of **5a−b**. Nevertheless, the presence of imine can be detected via chemical probes. Treatment of **5a** with $NaBH_4$ in methanol gives 1,3-bis(2′-deoxyguanosin-N^2-yl)propane (**6a**) in near quantitative yield. However, the reaction was slow, requiring ∼8 h at room temperature to reach completion.

With bis-nucleoside **5a**, equilibration could also be detected by monitoring its hydrolysis to form adduct **2a** and dG. The half-life for hydrolysis at ambient temperature (∼22 °C) was 64 h at pH 7. The reaction is base-catalyzed. Improved stability was observed in acidic solution; the half-life increased to 78.5 h at pH 6. The rate of hydrolysis is a composite function of the rate of ring opening and rate of imine hydrolysis. The rate of ring opening can be probed directly by observing by HPLC the epimerization of the diastereomers of **5a**, which occurs concurrently with the hydrolysis reaction. The epimerization is a pseudoracemization and eventually reaches an equal mixture of the two diastereomers. At pH 7.0, the half-life for the epimerization reaction is 90 min, revealing that the hydrolysis of imine is ∼100-fold slower than the ring-opening reaction. The epimerization reaction is catalyzed by base and is strongly retarded in acidic solution with half-lives increasing from 1.5 h at pH 7.0 to 14.8, 47.0, and 183.9 h at pH values of 6.0, 5.0, and 4.0, respectively. Ring-opening is base-catalyzed to a greater extent than the overall hydrolysis reaction.

The 6*R* and 6*S* diastereomers of bis-nucleoside **5b** exist as single, chromatographically distinguishable peaks that do not interconvert. The closure always occurs with the substituent at C8 trans to the methyl group at C6. The C6 stereogenic center

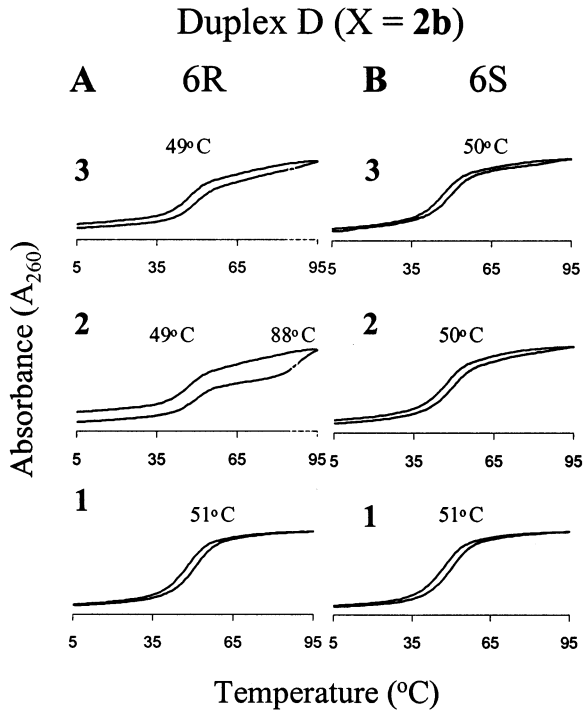


Figure 9. Melting analysis (T_m) of cross-linking by crotonaldehyde-adducted duplex **D** (X = **2b**). The incubations were carried out at 37 °C. Details of the incubations and analyses are provided in the Experimental Section. In each case the trace for the heating cycle lies below the trace for the cooling cycle. The T_m values are noted on the plots. (Panel A) Duplex **D** (*X* = 6*R*-**2b**). A1−3. After 0 and 24 days and then repeated immediately thereafter. (Panel B) Duplex **D** (*X* = 6*S*-**2b**). B1−3. After 0 and 26 days and then repeated immediately thereafter. Absorbance monitored at 260 nm.

cannot epimerize in **5b** or its cleavage products. As with **5a**, **5b** must be in equilibrium with vanishingly small amounts of imine **4b**, but epimerization is not observed in this case because the C6 stereogenic center is incapable of epimerization in **5b** or in its cleavage products. NMR studies of hydroxypropano nucleoside **2b** have shown that the 8-hydroxyl group has a strong preference for existing in a trans orientation relative to the 6-methyl group.[38,39] Bis-nucleoside **5b** also shows no evidence of epimerization at C8, and it is presumed that the deoxyguanosinyl substituent at C8 is oriented trans to the C6 methyl group. Reduction of **5b** to give **6b** can be achieved in good yield but is several-fold slower than the reduction of **5a**, a 20-h reaction period being required to bring the reaction to completion.

Discussion

Structure of the Acrolein Cross-Link in Duplex DNA. Three possible structures, **3**−**5**, for the guanine−guanine cross-link formed by the reaction of deoxyguanosine with deoxyguanosine adducts **2** are shown in Scheme 1. The products of reactions carried out with nucleosides are unambiguously pyrimidopurinones **5**, but the assignment of structure(s) of the cross-linked species in oligonucleotides is a more complex issue. The reduction reactions carried out on the cross-linked oligonucleotides with $NaCNBH_3$ demonstrate that the cross-link involves imine **4**, *at least in part*. However, **3**, **4**, and **5** are all in equilibrium with one another so that **5** might be the main species present in oligonucleotides but convert to imine during the course of the reduction. However, if **5a** were present, it would cause major distortion of the duplex structure, which is inconsistent with the high T_m values observed for the cross-linked oligonucleotides. The hydrolytic stability of the pyrimidopurinones also suggests that, if pyrimidopurinones had been present, the MALDI mass spectra would have given more intense signals for the cross-linked species and cross-link reversal would not have occurred so readily in unbuffered water.

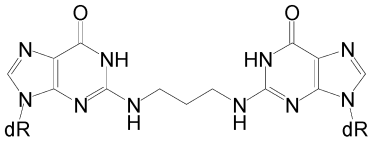
The recent NMR study by Kim et al. clearly established the presence of carbinolamine **3a**.[36] Consequently, we conclude that the acrolein and crotonaldehyde cross-links present in oligonucleotides are mixtures of the carbinolamine and imine forms. To maximize the reduction yield with $NaCNBH_3$, which only reduces the imine form, reaction times up to 20 h are required to permit time for the carbinolamine to fully convert to the imine. The long reaction period employed for the reduction may have led to the yield of the reduction product being greater than the amount of cross-link present in the equilibrium mixture because additional cross-link would have been formed as the reduction was taking place. However, we believe the effect of this is minor. Furthermore, the cross-linking yields presented in this paper are based on HPLC and CZE results rather the yields of reduced cross-links.

Sequence Selectivity for Cross-Link Formation. No evidence could be found for cross-linking in a GpC context either in competition with CpG or in an oligonucleotide in which the opportunity for CpG cross-linking was not present. There are precedents for interchain cross-links being formed preferentially in CpG sequence contexts.[32−34] However, the well-characterized examples involve reactions that are effectively irreversible, so that the preference reflects relative activation energies for formation of these cross-links as opposed to other types. In the present case the preference for the CpG sequence context is thermodynamic rather than kinetic.

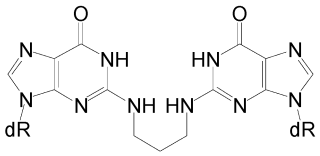
The conformations of DNA duplexes containing acrolein and crotonaldehyde cross-links are, at this juncture, unknown. Furthermore, because of the fact that a mixture of cross-linked and uncross-linked species is present in the samples, it may never be possible to obtain detailed structural solutions by NMR spectroscopy or X-ray crystallography. However, the structures of oligonucleotides containing reduced analogues of **3** and **4**, (i.e., with interchain trimethylene tethers between N^2 positions of guanines) have been established by NMR spectroscopy in both CpG and GpC sequence contexts.[26,35] Two classes of structures had to be considered to arrive at the structures, one in which the tether replaces the proton proximal to N1 leading to a W (i.e., zigzag) conformation for the tether and the other in which the distal proton is replaced, leading to a U conformation (Scheme 2). In both oligonucleotides the tether was found to be in a W conformation. One difference between the CpG and GpC structures is that the two guanines and the carbon atoms of the tether lie very nearly in a single plane in the CpG duplex, whereas the guanines in the GpC duplex lie in parallel planes with the tether assuming a stair-step between the purine rings such that no four atoms of the tether lie in a plane.

A second distinction between the two structures is that the GpC duplex is much more severely distorted than the CpG. This observation is reinforced by melting studies on the two duplexes

(38) Eder, E.; Hoffman, C. *Chem. Res. Toxicol.* **1992**, *5*, 802−808.
(39) Chung, F. L.; Hecht, S. S. *Cancer Res.* **1983**, *43*, 1230−1235.



W conformation



U conformation

where the T_m of the CpG duplex was raised by the trimethylene tether while the T_m of the GpC duplex was lowered. This leads us to conclude that the CpG duplex is stabilized by the trimethylene cross-link, while the GpC duplex is destabilized.

On the basis of these structures involving the trimethylene cross-link, some predictions can be made about imine- and carbinolamine-type cross-links. The imine tether differs from the trimethylene by the trigonal hybridization of the imine, forcing planarity of the four-atom sequence C−C═N−C, thus limiting conformational mobility in the duplex. The consequence is that a substitution of −CH═N− for $-CH_2-NH-$ can be accommodated in the CpG duplex but not in the GpC. To create the carbinolamine tether from the trimethylene, only minimal adjustments need to be made in duplex conformation to accommodate the hydroxyl substituent. However, for cross-links formed by unsaturated aldehydes, a key difference relative to the trimethylene cross-links is reversibility. The destabilizing effect of a three-carbon carbinolamine cross-link in a GpC sequence context will cause it to revert to the uncross-linked duplex.

Configurational Requirement for Cross-Link Formation by the Crotonaldehyde Adduct. For the reasons presented above for the acrolein cross-link, we think it probable that the crotonaldehyde cross-link is present in DNA duplexes in the imine and carbinolamine forms but not the pyrimidopurinone. The methyl substituent on the tether in the cross-linked species produces additional conformational constraints on the DNA duplex beyond those present with the acrolein cross-link. This is reflected in the observation that at equilibrium the yields of cross-links formed by the 6*R* and 6*S* diastereomers of **2b** are lower, 38 and 5%, respectively, in CpG duplex **A**, versus ∼50% for acrolein adduct **2a**.

The NMR structure of the duplex containing the trimethylene cross-link in a CpG sequence fails to provide a clear indication of the basis for the diastereomeric selectivity in formation of the crotonaldehyde cross-link. It is likely that the methyl group is accommodated by bending of the duplex and that the 6*R* diastereomer requires less deformation than the 6*S*. However, we cannot exclude the possibility that the presence of the methyl group forces the tether to assume a U conformation with the tether being linked to the distal (non-Watson−Crick) positions on the amino groups. NMR studies will be required to establish which scenario is correct.

Our cross-linking results with oligonucleotides containing **2b** explain the stereochemical preference observed by Wang et al.[24] in the formation of interchain cross-links by acetaldehyde where they obtained only one diastereomer of **6b**. We have established that the observed diastereomer has the 6*R* configuration.[28] In view of the fact that cross-linking is a reversible reaction, stereoselectivity is a thermodynamic phenomenon and independent of reaction mechanism. Cross-links having the *R* configuration have significant stability. Any duplex containing a cross-link having the *S* configuration will, if formed at all, revert almost completely to an uncross-linked duplex.

Biological Implications. The formation of interstrand cross-links in DNA creates a severe challenge to a cell. A cross-link is per se an absolute block to replication unless the cell has methodology at its disposal for repair of the cross-link. Even then, the repair process is likely to be error-prone because the damage is present in both strands. These questions have received much attention because of their importance with respect to the mechanisms of action of bifunctional xenobiotic carcinogens and chemotherapeutic agents.[40−46] The pathways for disposition of interchain cross-links achieve unique relevance with respect to acrolein and crotonaldehyde which have been shown by others to arise endogenously in mammalian cells and to form guanine adducts in the DNA of these cells.[3−8,13−15] In this work, we describe chemical experiments establishing that DNA interchain cross-links are formed by the acrolein and crotonaldehyde adducts. Consequently, it seems highly likely that the endogenously formed acrolein and crotonaldehyde form interchain DNA cross-links in living systems, as well. The availability of synthetic methodology for preparing DNA into which the adducted nucleosides have been placed in a site-specific manner now makes it possible to study the biological consequences of these adducts including the repair of cross-links derived from them.

Summary

Oligodeoxynucleotides containing acrolein adduct **2a**, which is a major form of DNA damage caused by acrolein, form interchain cross-links when the adducted base lies in a CpG sequence context but not when it is in a GpC. The cross-link exists as a mixture of carbinolamine **3a** and imine **4a** forms, although the linked nucleosides cyclize to give tetrahydropyrimidopurinone-type bis-nucleoside **5a** during enzymatic digestion with nucleases. The basis for the sequence selectivity has been shown to be thermodynamic rather than kinetic. The accompanying contribution demonstrates that a trimethylene tether between N^2 positions of guanines, which models the carbinolamine cross-link, severely disrupts duplex structure in a GpC sequence. An earlier study had shown that the trimethylene tether in a CpG sequence was strongly stabilizing.

(40) Kuraoka, I.; Kobertz, W. R.; Ariza, R. R.; Biggerstaff, M.; Essigmann, J. M.; Wood, R. D. *J. Biol. Chem.* **2000**, *275*, 26632−26636.
(41) Berardini, M.; Mackay, W.; Loechler, E. L. *Biochemistry* **1997**, *36*, 3506−3513.
(42) Mu, D.; Bessho, T.; Nechev, L. V.; Chen, D. J.; Harris, T. M.; Hearst, J. E.; Sancar, A. *Mol. Cell Biol.* **2000**, *20*, 2446−2454.
(43) Grossmann, K. F.; Ward, A. M.; Matkovic, M. E.; Folias, A. E.; Moses, R. E. *Mutat. Res.* **2001**, 73−83.
(44) Dronkert, M. L.; de Wit, J.; Boeve, M.; Vasconcelos, M. L.; van Steeg, H.; Tan, T. L.; Hoeijmakers, J. H.; Kanaar, R. *Mol. Cell. Biol.* **2000**, *20*, 4553−4561.
(45) Dronkert, M. L.; Kanaar, R. *Mutat. Res.* **2001**, *486*, 217−247.
(46) Dong, Q.; Johnson, S. P.; Colvin, O. M.; Bullock, N.; Kilborn, C.; Runyon, G.; Sullivan, D. M.; Easton, J.; Bigner, D. D.; Nahta, R.; Marks, J.; Modrich, P.; Friedman, H. S. *Cancer Chemother. Pharmacol.* **1999**, *4*, 73−79.

Oligodeoxynucleotides containing crotonaldehyde−deoxyguanosine adduct **2b** also form cross-links when the adducted base lies in a CpG sequence context. The effectiveness of cross-linking is dependent upon the configuration at C6 of **2b** with cross-linking being strongly favored by C6 having an *R* configuration. These results have important biological implications because interchain cross-links are a serious form of DNA damage and acrolein, crotonaldehyde, and their DNA adducts arise endogenously by oxidation of polyunsaturated fatty acids in mammalian cells.

Acknowledgment. We gratefully acknowledge financial support by the U.S. Public Health Service (NIEHS Grants ES00267 and ES05355).

Supporting Information Available: ^{1}H NMR spectra of compounds **5b** (6*R* and 6*S* diastereomers) and **6b** (1*R* and 1*S* diastereomers) and **7a** and **7b** (1*R* and 1*S* distereomers) (PDF). This material is available free of charge via the Internet at http://pubs.acs.org.

JA020778F