

Base-Displaced Intercalated Structure of the Food Mutagen 2-Amino-3-methylimidazo[4,5-f]quinoline in the Recognition Sequence of the *NarI* Restriction Enzyme, a Hotspot for -2 bp Deletions

Feng Wang, Nicholas E. DeMuro, C. Eric Elmquist, James S. Stover, Carmelo J. Rizzo,* and Michael P. Stone*

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

Received March 23, 2006; E-mail: michael.p.stone@vanderbilt.edu; carmelo.rizzo@vanderbilt.edu

Abstract: The solution structure of the oligodeoxynucleotide 5'-d(CTCGGCXCCATC)-3'-5'-d(GATGGCGC-CGAG)-3' containing the heterocyclic amine 8-[(3-methyl-3*H*-imidazo[4,5-f]quinolin-2-yl)amino]-2'-deoxyguanosine adduct (IQ) at the third guanine in the *NarI* restriction sequence, a hot spot for -2 bp frameshifts, is reported. Molecular dynamics calculations restrained by distances derived from 24 ^1H NOEs between IQ and DNA, and torsion angles derived from ^3J couplings, yielded ensembles of structures in which the adducted guanine was displaced into the major groove with its glycosyl torsion angle in the syn conformation. One proton of its exocyclic amine was approximately 2.8 Å from an oxygen of the 5' phosphodiester linkage, suggesting formation of a hydrogen bond. The carcinogen-guanine linkage was defined by torsion angles α' [N9-C8-N(IQ)-C2(IQ)] of $159 \pm 7^\circ$ and β' [C8-N(IQ)-C2(IQ)-N3(IQ)] of $-23 \pm 8^\circ$. The complementary cytosine was also displaced into the major groove. This allowed IQ to intercalate between the flanking C-G base pairs. The disruption of Watson-Crick hydrogen bonding was corroborated by chemical-shift perturbations for base aromatic protons in the complementary strand opposite to the modified guanine. Chemical-shift perturbations were also observed for ^{31}P resonances corresponding to phosphodiester linkages flanking the adduct. The results confirmed that IQ adopted a base-displaced intercalated conformation in this sequence context but did not corroborate the formation of a hydrogen bond between the IQ quinoline nitrogen and the complementary dC [Elmquist, C. E.; Stover, J. S.; Wang, Z.; Rizzo, C. J. *J. Am. Chem. Soc.* **2004**, *126*, 11189–11201].

Introduction

The browning of protein-rich foods imparts flavor during cooking. It leads to the formation of heterocyclic amines (HCA) such as 2-amino-3-methylimidazo[4,5-f]quinoline (IQ).^{1–4} Various HCAs, including IQ, have been identified in grilled foods at ppb levels.^{5,6} Daily human intakes of HCAs, estimated to be ~ 60 ng/day,⁷ are modest; however, exposure to these compounds, which have been isolated from human urine,⁸ is of concern with regard to human health.

Exposure to IQ is associated with carcinogenesis. Tumors in organs of rodents and in the livers of monkeys are induced by

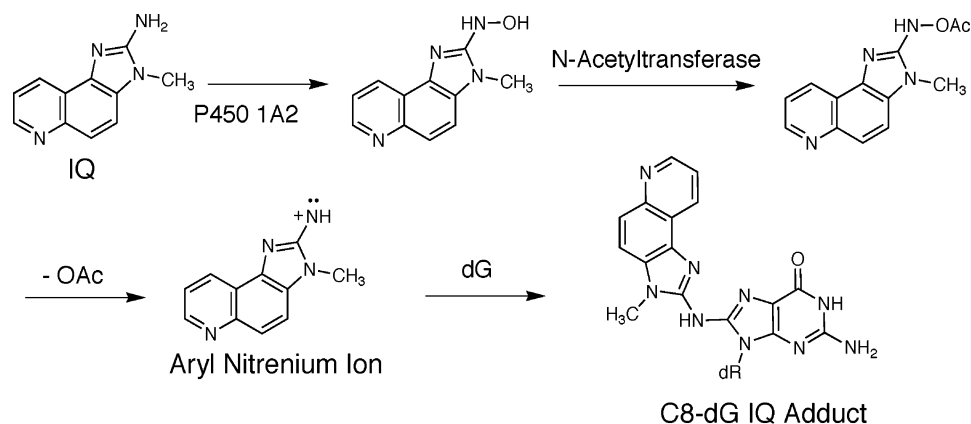
IQ.^{9–11,16} In mice, exposures lead to liver, forestomach, and lung tumors.¹² In rats, exposures lead to cancers in the liver, intestine, zymbal gland, clitoral gland, skin,¹³ mammary glands, liver, and ear ducts.¹⁴ TD50 values for in rats are 0.7 mg/kg/day, and in mice are 14.7 mg/kg/day.¹⁵ Human exposure to HCAs is associated with pancreatic,¹⁷ colon,¹⁸ prostate,¹⁹ and breast cancer.^{20,21}

In bacterial reversion assays,^{22–25} HCAs are active in point and frameshift tester strains.²⁶ IQ is one of the strongest chemical mutagens.²⁷ It is less prevalent than 2-amino-1-methyl-6-phenylimidazo[4,5-b]pyridine (PhIP)²⁸ but is 200-fold more

- (1) Wakabayashi, K.; Nagao, M.; Esumi, H.; Sugimura, T. *Cancer Res.* **1992**, *52*, 2092–2098.
- (2) Layton, D. W.; Bogen, K. T.; Knize, M. G.; Hatch, F. T.; Johnson, V. M.; Felton, J. S. *Carcinogenesis* **1995**, *16*, 39–52.
- (3) Sugimura, T. *Mutat. Res.* **1997**, *376*, 211–219.
- (4) Sugimura, T.; Wakabayashi, K.; Nakagama, H.; Nagao, M. *Cancer Sci.* **2004**, *95*, 290–299.
- (5) Kataoka, H.; Nishioka, S.; Kobayashi, M.; Hanaoka, T.; Tsugane, S. *Bull. Environ. Contam. Toxicol.* **2002**, *69*, 682–689.
- (6) Felton, J. S.; Knize, M. G.; Salmon, C. P.; Malfatti, M. A.; Kulp, K. S. *Environ. Mol. Mutagen.* **2002**, *39*, 112–118.
- (7) Kobayashi, M.; Hanaoka, T.; Nishioka, S.; Kataoka, H.; Tsugane, S. *Mutat. Res.* **2002**, *506–507*, 233–241.
- (8) Ushiyama, H.; Wakabayashi, K.; Hirose, M.; Itoh, H.; Sugimura, T.; Nagao, M. *Carcinogenesis* **1991**, *12*, 1417–1422.

- (9) Ohgaki, H.; Hasegawa, H.; Kato, T.; Suenaga, M.; Ubukata, M.; Sato, S.; Takayama, S.; Sugimura, T. *Environ. Health Perspect.* **1986**, *67*, 129–134.
- (10) Heflich, R. H.; Neft, R. E. *Mutat. Res.* **1994**, *318*, 73–114.
- (11) Thorgeirsson, U. P.; Snyderwine, E. G.; Gomez, D. E.; Adamson, R. H. *In Vivo* **1996**, *10*, 145–152.
- (12) Ohgaki, H.; Kusama, K.; Matsukura, N.; Morino, K.; Hasegawa, H.; Sato, S.; Takayama, S.; Sugimura, T. *Carcinogenesis* **1984**, *5*, 921–924.
- (13) Takayama, S.; Nakatsuru, Y.; Masuda, M.; Ohgaki, H.; Sato, S.; Sugimura, T. *Gann* **1984**, *75*, 467–470.
- (14) Tanaka, T.; Barnes, W. S.; Williams, G. M.; Weisburger, J. H. *Jpn. J. Cancer Res.* **1985**, *45*, 570–576.
- (15) Sugimura, T.; Nagao, M.; Wakabayashi, K. In *Complex Factors Pertinent to Human Hazard and Risk*; Wiley: New York, 2000; p 349–359.
- (16) Adamson, R. H.; Thorgeirsson, U. P.; Snyderwine, E. G.; Thorgeirsson, S. S.; Reeves, J.; Dalgard, D. W.; Takayama, S.; Sugimura, T. *Jpn. J. Cancer Res.* **1990**, *50*, 10–14.

Chart 1. Metabolic Activation of IQ



mutagenic than the latter in *Salmonella* reversion assays.³ IQ is an order of magnitude more mutagenic than is aflatoxin B₁. In bacteria, mutations occur primarily at G:C base pairs.^{29,30} It exhibits frameshift mutations in CG repeats. Similar levels of mutations are seen in mammalian *hprt*³¹ and *ef-2*³² gene assays. In mammalian cells, point mutations are observed.^{33–36} Sister chromatid exchanges are observed in rodent cells.^{37–39}

IQ is activated primarily by the enzyme CYP P450 1A2 to an *N*-hydroxyl oxidation product.^{40–43} Extra-hepatic CYP P450s oxidize HCAs with lower efficiencies.⁴⁴ The *N*-hydroxyl oxidation product is acetylated by cellular *N*-acetyl transferases, particularly NAT2.^{45–47} The resulting nitrenium ion is the ultimate reactive electrophile.^{36,44} The NAT2 fast acetylator polymorphism is associated with an increased risk of colorectal cancer in humans.^{48,49}

The C8-dG adducts of HCAs are observed both in rodents and primates, as measured by ³²P postlabeling.³⁵ The major

adduct formed by IQ occurs by substitution at C8-dG (Chart 1); a minor *N*²-dG adduct is also formed.⁵⁰ The structures of these adducts are established.^{51–53} The formation of the C8-dG adduct probably involves initial alkylation at N7-dG, followed by rearrangement.⁵⁴ High sensitivity LC/ESI-MS⁵⁵ has measured several adducts per 10⁷ nucleotides in animal tissues.^{19,56} The levels of C8 and *N*²-dG IQ adducts measured in tissues of rats and primates using mass spectrometry^{57,58} are in agreement with data obtained by ³²P postlabeling.

Heretofore, site-specific DNA adducts of HCAs have not been readily accessible. A synthesis of PhIP-adducted oligodeoxynucleotides involved reacting single-stranded DNA with the PhIP nitrenium ion.⁵⁹ The low yield, coupled with complexities of purification, limited the approach to oligodeoxynucleotides containing a single dG. In the COS-7 site-specific mutagenesis system,⁶⁰ if dC was at the 5'-flanking position to dG-C8 PhIP, incorporation of dC, the correct base, was observed. However, G → T transversions, and lesser amounts of G → A transitions and G → C transversions, were detected. If the dC 5'-flanking base was replaced by T, dA, or dG, the mutational spectra were similar, but greater mutational frequencies were observed with dC or dG than with dA 5' to the adduct. Single-base deletions were detected only when dG or T flanked the adduct. Thus, dG-C8 PhIP was mutagenic, generating primarily G → T transversions.⁶¹

A study of the C8-dG PhIP adduct in 5'-d(CCATCXCTACC)-3'·5'-d(GGTAGCGATGG)-3' represents the only conformational analysis of an HCA-adducted duplex.⁵⁹ This yielded a base-displaced intercalated structure, in which the adducted dG was in the syn conformation and situated in the major groove. The C6-phenyl and N3-methyl groups protruded into the minor groove, widening it and compressing the major groove, resulting in DNA bending.

- (17) Anderson, K. E.; Hammons, G. J.; Kadlubar, F. F.; Potter, J. D.; Kaderlik, K. R.; Ilett, K. F.; Minchin, R. F.; Teitel, C. H.; Chou, H. C.; Martin, M. V.; Guengerich, F. P.; Barone, G. W.; Lang, N. P.; Peterson, L. A. *Carcinogenesis* **1997**, *18*, 1085–1092.
- (18) Lang, N. P.; Butler, M. A.; Massengill, J.; Lawson, M.; Stotts, R. C.; Hauer-Jensen, M.; Kadlubar, F. F. *Cancer Epidemiol. Biomarkers Prev.* **1994**, *3*, 675–682.
- (19) Shirai, T.; Sano, M.; Tamano, S.; Takahashi, S.; Hirose, M.; Futakuchi, M.; Hasegawa, R.; Imaida, K.; Matsumoto, K.; Wakabayashi, K.; Sugimura, T.; Ito, N. *Cancer Res.* **1997**, *57*, 195–198.
- (20) Snyderwine, E. G. *Cancer* **1994**, *74*, 1070–1077.
- (21) Ronco, A.; De Stefani, E.; Mendilaharsu, M.; Deneo-Pellegrini, H. *Int. J. Cancer* **1996**, *65*, 328–331.
- (22) Josephy, P. D.; Evans, D. H.; Parikh, A.; Guengerich, F. P. *Chem. Res. Toxicol.* **1998**, *11*, 70–74.
- (23) Josephy, P. D.; Gruz, P.; Nohmi, T. *Mutat. Res.* **1997**, *386*, 1–23.
- (24) Oda, Y.; Yamazaki, H.; Watanabe, M.; Nohmi, T.; Shimada, T. *Mutat. Res.* **1995**, *334*, 145–156.
- (25) Oda, Y.; Aryal, P.; Terashita, T.; Gillam, E. M.; Guengerich, F. P.; Shimada, T. *Mutat. Res.* **2001**, *492*, 81–90.
- (26) Nagao, M. In *Food-Borne Carcinogens: Heterocyclic Amines*; Nagao, M., Sugimura, T., Eds.; Wiley: New York, 2000; pp 31–71.
- (27) Sugimura, T.; Sato, S. *Cancer Res.* **1983**, *43*, 2415s–2421s.
- (28) Hecht, S. S. *Environ. Mol. Mutagen.* **2002**, *39*, 119–126.
- (29) Kosakarn, P.; Halliday, J. A.; Glickman, B. W.; Josephy, P. D. *Carcinogenesis* **1993**, *14*, 511–517.
- (30) Watanabe, M.; Ohta, T. *Carcinogenesis* **1993**, *14*, 1149–1153.
- (31) Terada, M.; Nagao, M.; Nakayasu, M.; Sakamoto, H.; Nakasato, F.; Sugimura, T. *Environ. Health Perspect.* **1986**, *67*, 117–119.
- (32) Thompson, L. H.; Tucker, J. D.; Stewart, S. A.; Christensen, M. L.; Salazar, E. P.; Carrano, A. V.; Felton, J. S. *Mutagenesis* **1987**, *2*, 483–487.
- (33) Felton, J. S.; Fultz, E.; Dolbeare, F. A.; Knize, M. G. *Food Chem. Toxicol.* **1994**, *32*, 897–903.
- (34) Felton, J. S.; Knize, M. G.; Dolbeare, F. A.; Wu, R. W. *Environ. Health Perspect.* **1994**, *102*, 201–204.
- (35) Schut, H. A.; Snyderwine, E. G. *Carcinogenesis* **1999**, *20*, 353–368.
- (36) Turesky, R. J. *Drug Metab. Rev.* **2002**, *34*, 625–650.
- (37) Thompson, L. H.; Carrano, A. V.; Salazar, E.; Felton, J. S.; Hatch, F. T. *Mutat. Res.* **1983**, *117*, 243–257.
- (38) Aeschbacher, H. U.; Ruch, E. *Carcinogenesis* **1989**, *10*, 429–433.
- (39) Tohda, H.; Oikawa, A.; Kawachi, T.; Sugimura, T. *Mutat. Res.* **1980**, *77*, 65–69.

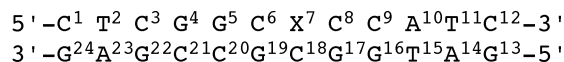
- (40) Yamazoe, Y.; Shimada, M.; Kamataki, T.; Kato, R. *Cancer Res.* **1983**, *43*, 5768–5774.
- (41) Shimada, T.; Hayes, C. L.; Yamazaki, H.; Amin, S.; Hecht, S. S.; Guengerich, F. P.; Sutter, T. R. *Cancer Res.* **1996**, *56*, 2979–2984.
- (42) Boobis, A. R.; Lynch, A. M.; Murray, S.; de la Torre, R.; Solans, A.; Farre, M.; Segura, J.; Gooderham, N. J.; Davies, D. S. *Cancer Res.* **1994**, *54*, 89–94.
- (43) Hammons, G. J.; Milton, D.; Stepps, K.; Guengerich, F. P.; Tukey, R. H.; Kadlubar, F. F. *Carcinogenesis* **1997**, *18*, 851–854.
- (44) Guengerich, F. P. *Drug Metab. Rev.* **2002**, *34*, 607–623.
- (45) Hickman, D.; Pope, J.; Patil, S. D.; Fakis, G.; Smelt, V.; Stanley, L. A.; Payton, M.; Unadkat, J. D.; Sim, E. *Gut* **1998**, *42*, 402–409.
- (46) Hein, D. W.; Doll, M. A.; Rustan, T. D.; Gray, K.; Feng, Y.; Ferguson, R. J.; Grant, D. M. *Carcinogenesis* **1993**, *14*, 1633–1638.
- (47) Minchin, R. F.; Reeves, P. T.; Teitel, C. H.; McManus, M. E.; Mojarrabi, B.; Ilett, K. F.; Kadlubar, F. F. *Biochem. Biophys. Res. Commun.* **1992**, *185*, 839–844.

An efficient strategy for synthesis of C8-dG arylamine adducts involving the Buchwald-Hartwig palladium-catalyzed *N*-arylation reaction of a protected 8-bromo-2'-dG derivative with arylamines facilitated preparation of the C8-IQ-adducted dG nucleoside, which was incorporated into oligodeoxynucleotides using phosphoramidite chemistry.^{62,63} A combination of thermal melting studies, and UV and circular dichroism spectroscopy, led to the proposal of a base-displaced intercalated conformation at the G³-position of the 5'-d(CG¹G²CG³CC)-3' recognition site of the *NarI* enzyme. It was proposed that this was stabilized by a hydrogen bond between the quinoline nitrogen of IQ and the complementary cytosine.⁶³ In contrast, molecular mechanics analysis of the C8-dG IQ-modified duplex 5'-d(G¹G²CX³CCA)-3'·5'-d(TGGCGCC)-3' suggested that the favored conformation featured the modified dG in the syn conformation with IQ in the minor groove and directed 3' with respect to the modified strand.⁶⁴ This suggested that the base-displaced intercalated conformation was ~10 kcal/mol higher in energy than the minor groove conformation.⁶⁴ A study of the C8-dG IQ adduct at the nucleoside level confirmed that the adducted dG was in the syn conformation about the glycosyl bond.⁵³

This work presents a study of the C8-dG IQ adduct in 5'-d(C¹T²C³G⁴G⁵C⁶X⁷C⁸C⁹A¹⁰T¹¹C¹²)-3'·5'-d(G¹³A¹⁴T¹⁵G¹⁶G¹⁷-C¹⁸G¹⁹C²⁰C²¹G²²A²³G²⁴)-3'; X = 8-[(3-methyl-3*H*-imidazo[4,5-*f*]quinolin-2-yl)amino]-2'-deoxyguanosine, named the *NarI*IQ3 sequence. It contains the 5'-d(CG¹G²CX³CC)-3' recognition site of the *NarI* restriction enzyme, in which the third guanine (G³ in the *NarI* sequence and X⁷ in this study) represents a hot spot for -2 bp frameshifts (Chart 2). The results reveal a base-displaced intercalated structure. The adducted dG adopts a syn conformation about the glycosyl bond and extrudes into the major groove, the IQ moiety intercalates into the DNA, and the complementary dC extrudes from the helix. A hydrogen bond between the IQ quinoline nitrogen and the complementary dC⁶³ is not observed.

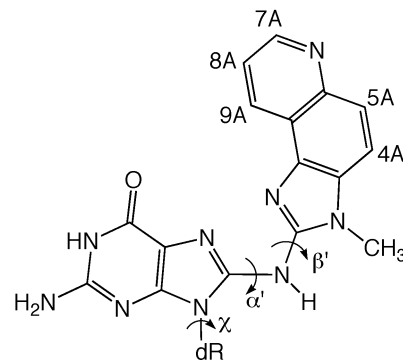
Chart 2. (A) *NarI*IQ3 Duplex and (B) C8-dG IQ Adduct^a

A



X = C8-dG IQ adduct

B



^a Torsion angles defining the IQ orientation in the duplex are γ , the glycosyl torsion angle (O4'-C1'-N9-C4), α' (N9-C8-N[IQ]-C2[IQ]), and β' (C8-N[IQ]-C2[IQ]-N3[IQ]).

Materials and Methods

Sample Preparation. The oligodeoxynucleotides 5'-d(CTCGGCGC-CATC)-3' and 5'-d(GATGGCGCGAG)-3' were obtained from the Midland Certified Reagent Company, purified by anion exchange chromatography. The oligodeoxynucleotide 5'-d(CTCGGCXCCATC)-3' was synthesized and purified as described.⁶³ All oligodeoxynucleotides were characterized by MALDI-TOF mass spectrometry and enzymatic digestion, and their purities were assessed by capillary zone electrophoresis (CZE). Oligodeoxynucleotide duplexes were annealed at 70 °C. Their stoichiometry was established by ¹H NMR. The duplexes were dissolved in 0.25 mL of buffer containing 0.1 M NaCl, 10 mM NaH₂PO₄, and 50 μ M Na₂EDTA (pH 7.0). The oligodeoxynucleotide concentrations were ~0.7 mM using an extinction coefficient of 1.10 $\times 10^5$ M⁻¹ cm⁻¹ at 260 nm.⁶⁵

NMR. ¹H NMR spectra were obtained at 500.13, 600.20, and 800.23 MHz. COSY spectra were collected at 15, 20, 25, 30, and 35 °C in 99.996% D₂O. ¹H NOESY experiments in D₂O were conducted at 15 °C. To obtain distance restraints, spectra were recorded at mixing times of 150, 200, and 250 ms at the ¹H NMR frequency of 800.23 MHz. The data were recorded with 1024 real data points in the t₁ dimension and 2048 real points in the t₂ dimension. The relaxation delay was 2 s. The data in the t₁ dimension were zero-filled to give a matrix of 2K \times 2K real points. NOESY spectra for the exchangeable protons were recorded at 5 °C, in 90:10 H₂O/D₂O, using the Watergate sequence⁶⁶ for water suppression and a 250-ms mixing time at a ¹H NMR frequency of 600.20 MHz. Chemical shifts of proton resonances were referenced to water. Double quantum-filtered ¹H correlation (DQF-COSY)^{67,68} and exclusive COSY (E-COSY)⁶⁹ spectra were collected at 25 °C at 500.13 MHz and zero-filled to give a matrix of 1024 \times 2048 real points. A skewed sine-bell square apodization function with a 90° phase shift and a skew factor of 1.0 was used in both dimensions. ¹H-³¹P HMBC spectra^{70,71} were obtained at 30 °C. The data matrix was 256 (t₁) \times

- (48) Le Marchand, L.; Hankin, J. H.; Pierce, L. M.; Sinha, R.; Nerurkar, P. V.; Franke, A. A.; Wilkens, L. R.; Kolonel, L. N.; Donlon, T.; Seifried, A.; Custer, L. J.; Lum-Jones, A.; Chang, W. *Mutat. Res.* **2002**, *506–507*, 205–214.
- (49) Ishibe, N.; Sinha, R.; Hein, D. W.; Kulldorff, M.; Strickland, P.; Fretland, A. J.; Chow, W. H.; Kadlubar, F. F.; Lang, N. P.; Rothman, N. *Pharmacogenetics* **2002**, *12*, 145–150.
- (50) Zhou, Y.; Chladek, S.; Romano, L. J. *J. Org. Chem.* **1994**, *59*, 556–563.
- (51) Snyderwine, E. G.; Roller, P. P.; Adamson, R. H.; Sato, S.; Thorgeirsson, S. S. *Carcinogenesis* **1988**, *9*, 1061–1065.
- (52) Nagaoka, H.; Wakabayashi, K.; Kim, S. B.; Kim, I. S.; Tanaka, Y.; Ochiai, M.; Tada, A.; Nukaya, H.; Sugimura, T.; Nagao, M. *Jpn. J. Cancer Res.* **1992**, *52*, 1025–1029.
- (53) Turesky, R. J.; Rossi, S. C.; Welti, D. H.; Lay, J. O., Jr.; Kadlubar, F. F. *Chem. Res. Toxicol.* **1992**, *5*, 479–490.
- (54) Guengerich, F. P.; Mundkowski, R. G.; Voehler, M.; Kadlubar, F. F. *Chem. Res. Toxicol.* **1999**, *12*, 906–916.
- (55) Gorlewska-Roberts, K.; Green, B.; Fares, M.; Ambrosone, C. B.; Kadlubar, F. F. *Environ. Mol. Mutagen.* **2002**, *39*, 184–192.
- (56) Takayama, K.; Yamashita, K.; Wakabayashi, K.; Sugimura, T.; Nagao, M. *Jpn. J. Cancer Res.* **1989**, *49*, 1145–1148.
- (57) Gangl, E. T.; Turesky, R. J.; Vouros, P. *Anal. Chem.* **2001**, *73*, 2397–2404.
- (58) Soglia, J. R.; Turesky, R. J.; Paehler, A.; Vouros, P. *Anal. Chem.* **2001**, *73*, 2819–2827.
- (59) Brown, K.; Hingerty, B. E.; Guenther, E. A.; Krishnan, V. V.; Broyde, S.; Turteltaub, K. W.; Cosman, M. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 8507–8512.
- (60) Moriya, M. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 1122–1126.
- (61) Shibutani, S.; Fernandes, A.; Suzuki, N.; Zhou, L.; Johnson, F.; Grollman, A. P. *J. Biol. Chem.* **1999**, *274*, 27433–27438.
- (62) Wang, Z.; Rizzo, C. J. *Org. Lett.* **2001**, *3*, 565–568.
- (63) Elmquist, C. E.; Stover, J. S.; Wang, Z.; Rizzo, C. J. *J. Am. Chem. Soc.* **2004**, *126*, 11189–11201.
- (64) Wu, X.; Shapiro, R.; Broyde, S. *Chem. Res. Toxicol.* **1999**, *12*, 895–905.

- (65) Borer, P. N. In *Handbook of Biochemistry and Molecular Biology*; CRC Press: Cleveland, OH, 1975.
- (66) Plietto, M.; Saudek, V.; Sklenar, V. *J. Biomol. NMR* **1992**, *6*, 661–665.
- (67) Piantini, U.; Sorensen, O. W.; Ernst, R. R. *J. Am. Chem. Soc.* **1982**, *104*, 6800–6801.
- (68) Rance, M.; Sorensen, O. W.; Hodenhausen, G.; Wagner, G.; Ernst, R. R.; Wuthrich, K. *Biochem. Biophys. Res. Commun.* **1983**, *117*, 479–485.
- (69) Griesinger, G.; Sorensen, O. W.; Ernst, R. R. *J. Am. Chem. Soc.* **1985**, *107*, 6394–6396.

2048 (t2) complex points. The data were Fourier transformed after zero filling in the t1 dimension, resulting in a matrix size of 512 (D1) \times 2048 (D2) real points. Trimethyl phosphate was used as an external standard. NMR data were processed using the program FELIX2000 (Accelrys, Inc., San Diego, CA) on Silicon Graphics (Mountain View, CA) Octane workstations.

Experimental Restraints. (a) Distance Restraints. Footprints were drawn around cross-peaks for the NOESY spectrum measured at a mixing time of 250 ms, using the program FELIX2000. Identical footprints were applied to the cross-peaks obtained at other mixing times. Cross-peak intensities were determined by volume integration. The intensities were combined with intensities generated from a complete relaxation matrix analysis of a starting DNA structure to generate a hybrid intensity matrix.⁷² The program MARDIGRAS (v. 5.2)^{73,74} was used to refine the hybrid matrix by iteration. The molecular motion was assumed to be isotropic. The noise level was set at half the intensity of the weakest cross-peak. Calculations were performed using DNA starting structures generated using the program INSIGHT II (Accelrys, Inc.), and NOE intensities derived from experiments at three mixing times, and with three τ_c values (2, 3, and 4 ns), yielding 18 sets of distances. Analysis of these data yielded experimental distance restraints and standard deviations used in restrained molecular dynamics calculations. For overlapped cross-peaks, the bounds on the distances were increased. The restraints were divided into four classes, reflecting the confidence level in the data.

(b) Torsion Angle Restraints. Deoxyribose pseudorotations were estimated by monitoring the $^3J_{\text{HH}}$ couplings of sugar protons.⁷⁵ The $J_{\text{H1}'\text{-H2}'}$ and $J_{\text{H1}'\text{-H2}''}$ couplings were measured from the E-COSY experiment,⁶⁹ whereas the intensities of $J_{\text{H2}'\text{-H3}'}$ and $J_{\text{H3}'\text{-H4}'}$ couplings were determined from the DQF-COSY experiment. The data were fit to curves relating the coupling constants to pseudorotation (P), sugar pucker amplitude (ϕ), and the percentage S-type conformation. The pseudorotation and amplitude ranges were converted to the five dihedral angles ν_0 to ν_4 .

Restrained Molecular Dynamics. Calculations were performed in vacuo using a simulated annealing protocol with the program X-PLOR.⁷⁶ The force field was derived from CHARMM⁷⁷ and adapted for nucleic acids. The empirical energy function treated hydrogens explicitly. The van der Waals energy term used the Lennard-Jones potential energy function. The electrostatic term used the Coulomb function, based on a full set of partial charges (-1 per residue) and a distance-dependent dielectric constant of $4r$. The nonbonded pair list was updated if any atom moved more than 0.5 \AA , and the cutoff radius for nonbonded interactions was 11 \AA . The effective energy function included terms describing distance and dihedral restraints, in the form of square-well potentials. Sets of rMD calculations for the unmodified and *Nar1IQ3* duplexes, and different starting structures of *Nar1IQ3* with IQ located in the minor groove (syn), major groove (anti), and intercalated position (syn), were considered. These were generated using INSIGHT II through modification at G⁷ C⁸, followed by energy minimization using X-PLOR. Partial charges and atom types for IQ used for X-PLOR calculations were those obtained by Wu et al.⁶⁴ Calculations were initiated by coupling to a heating bath, with a target temperature of 1100 K. The force constants were $25 \text{ kcal mol}^{-1} \text{ \AA}^{-2}$ for empirical hydrogen bonding, $10 \text{ kcal mol}^{-1} \text{ \AA}^{-2}$ for torsion angle restraints, and 50, 45, 40, and 35 $\text{kcal mol}^{-1} \text{ \AA}^{-2}$ for the four classes of NOE restraints. The target temperature was reached in 10 ps and

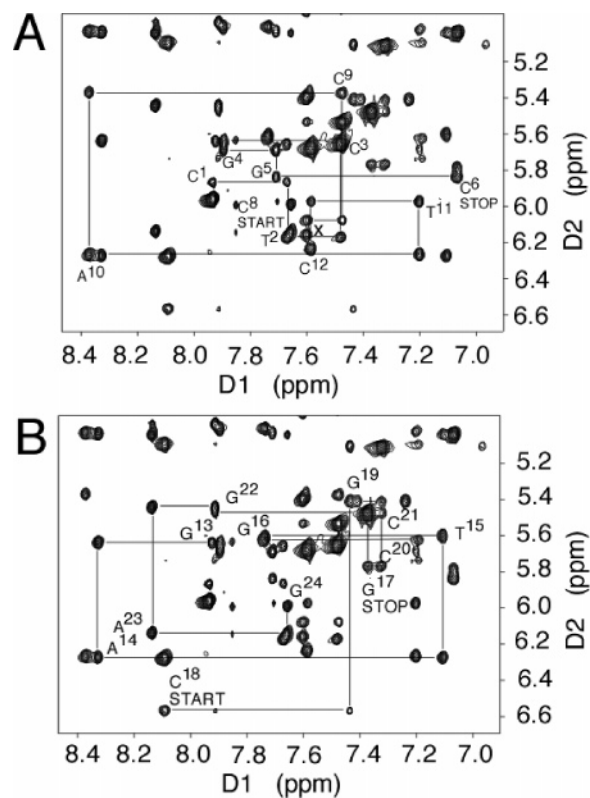


Figure 1. Aromatic-anomeric proton region of the 800.13 MHz NOESY spectrum for the *Nar1IQ3* duplex at 15 °C at 250 ms mixing time, showing sequential NOE connectivity. (A) Nucleotides C¹ \rightarrow C¹² of the modified strand. (B) Nucleotides G¹³ \rightarrow G²⁴ of the complementary strand.

was maintained for 25 ps. The system was cooled to 300 K over 10 ps and maintained at that temperature for 25 ps of equilibrium dynamics. The force constants for the four classes of NOE restraints were scaled during 10 ps of the heating period to 200, 180, 160, and 140 $\text{kcal mol}^{-1} \text{ \AA}^{-2}$ in the order of confidence factor. These weights were maintained during the remainder of the heating period and for the first 5 ps of equilibrium dynamics. They were then scaled to 100, 90, 80, and 70 $\text{kcal mol}^{-1} \text{ \AA}^{-2}$ in the order of confidence factor. The torsion angle and base pair distance force constants were scaled to 180 and 100 $\text{kcal mol}^{-1} \text{ \AA}^{-2}$ during the same period as for the NOE restraints. They were scaled to 70 and 45 $\text{kcal mol}^{-1} \text{ \AA}^{-2}$, also at the same time as the NOE restraints. Coordinate sets were archived every 0.1 ps, and 41 structures from the last 4.1 ps were averaged. These average rMD structures were subjected to 200 iterations of conjugate gradient energy minimization to obtain the final structures. Final structures were analyzed using X-PLOR to measure rmsd between the averaged and the converged structures. Back-calculation of NOE intensities from the emergent structures was performed using the program CORMA (v. 5.2).⁷² Helicoidal parameters were examined using the program 3DNA.⁷⁸

Results

NMR Spectroscopy. (a) DNA Nonexchangeable Protons. For the *Nar1IQ3* duplex, sequential NOE connectivities^{79,80} were interrupted (Figure 1). The absence of a purine imidazole proton in the C⁸-IQ-dG adduct X⁷ precluded observation of the C⁶ H1' \rightarrow X⁷ H8 and X⁷ H8 \rightarrow X⁷ H1' NOEs. The X⁷ H1' \rightarrow C⁸ H6 NOE was of normal intensity. In the complementary strand, the G¹⁷ H1' \rightarrow C¹⁸ H6 NOE was missing. The C¹⁸ H1' \rightarrow G¹⁹ H8 sequential NOE was weak. C¹⁸ is the nucleotide opposite

(70) Sklenar, V.; Bax, A.; Zon, G. *FEBS Lett.* **1986**, *208*, 94–98.

(71) Sklenar, V.; Miyashiro, H.; Zon, G.; Milles, H. T.; Bax, A. *FEBS Lett.* **1986**, *208*, 94–98.

(72) Keepers, J. W.; James, T. L. *J. Magn. Reson.* **1984**, *57*, 404–426.

(73) Borgias, B. A.; James, T. L. *J. Magn. Reson.* **1990**, *87*, 475–487.

(74) Liu, H.; Tonelli, M.; James, T. L. *J. Magn. Reson. B* **1996**, *111*, 85–89.

(75) Salazar, M.; Fedoroff, O. Y.; Miller, J. M.; Ribeiro, N. S.; Reid, B. R. *Biochemistry* **1993**, *32*, 4207–4215.

(76) Brunger, A. T. *X-Plor. Version 3.1. A System for X-ray Crystallography and NMR*; Yale University Press: New Haven, 1992.

(77) Nilsson, L.; Karplus, M. *J. Comput. Chem.* **1986**, *7*, 591–616.

(78) Lu, X. J.; Olson, W. K. *Nucleic Acids Res.* **2003**, *31*, 5108–5121.

(79) Reid, B. R. *Q. Rev. Biophys.* **1987**, *20*, 2–28.

(80) Patel, D. J.; Shapiro, L.; Hare, D. *Q. Rev. Biophys.* **1987**, *20*, 35–112.

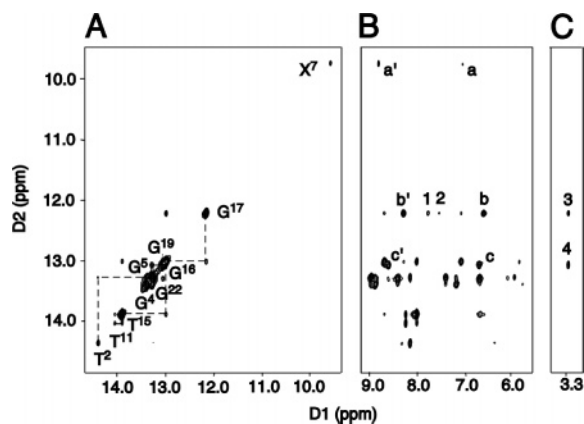


Figure 2. (A) Sequential NOE connectivity for the imino protons of base pairs T²·A²³ → T¹¹·A¹⁴ for the *Nar1IQ3* duplex at 5 °C. The labels represent the imino proton of the designated base. (B) NOE connectivity between the imino protons and the base and amino protons. The cross-peaks involving the imino protons are labeled as: a', a, X⁷ N1H → X⁷ NH₂-2b, e; b', b, G¹⁷ N1H → C⁸ NH₂-4b, e; c' and c, G¹⁹ N1H → C⁶ NH₂-4b, e; 1, G¹⁷ N1H → X⁷ H4A; 2, G¹⁷ N1H → X⁷ H5A. (C) NOE connectivity between the imino and the IQ methyl protons. The IQ-DNA cross-peaks labeled are as: 3, G¹⁷ N1H → X⁷ CH₃; and 4, G¹⁹ N1H → X⁷ CH₃.

to X⁷ in the complementary strand of the *Nar1IQ3* duplex. In the unmodified duplex, the G⁷ H8 → G⁷ H1' NOE was of normal intensity and all scalar cross-peaks between deoxyribose H1' and H2', H2'' protons were in the anticipated 1.6–2.8 ppm chemical-shift range. In contrast, for the *Nar1IQ3* duplex, the X⁷ H2' resonance shifted downfield to 3.61 ppm. This was characteristic of a *syn* dG orientation at X⁷. Complete sets of sequential NOEs were observed for both strands of the unmodified duplex (Figure S1 in the Supporting Information). The resonance assignments for the nonexchangeable protons of the *Nar1IQ3* and unmodified duplexes are found in Tables S1 and S2 of the Supporting Information.

(b) DNA Exchangeable Protons. The presence of the C8-dG IQ adduct resulted in chemical-shift dispersion of the imino proton resonances in the downfield region of the ¹H spectrum (Figure 2; Figure S2 in the Supporting Information shows an expanded contour plot of the imino proton resonances of the unmodified duplex). Whereas for the unmodified duplex the imino resonances arising from G⁴, G⁵, G⁷, G¹⁶, G¹⁷, and G²² were observed between 13 and 13.4 ppm, for the *Nar1IQ3* duplex, these imino protons resonated between 9.6 and 13.4 ppm. The imino protons of the *Nar1IQ3* duplex were assigned from NOEs between adjacent base pairs and NOEs to their corresponding base-paired amino protons.⁸¹ Interruptions in the NOEs between Watson–Crick hydrogen-bonded amino and imino protons of the *Nar1IQ3* duplex occurred between base pairs C⁶·G¹⁹ and X⁷·C¹⁸ and base pairs X⁷·C¹⁸ and C⁸·G¹⁷. All other sequential NOEs were observed. The X⁷ and G¹⁷ N1H resonances shifted upfield and were observed at 9.6 and 11.7 ppm, respectively. The X⁷ N1H proton exhibited NOEs to the X⁷ NH₂ protons. The amino protons resonated at 6.67 and 8.83 ppm, respectively. At the 5'-adjacent G⁶·C¹⁹ base pair, G¹⁹ N1H showed NOEs to the C⁶ NH₂ protons and to C⁶ H5. At the 3'-adjacent C⁸·G¹⁷ base pair, G¹⁷ N1H showed NOEs to the C⁸ NH₂ protons and to C⁸ H5. The T² N3H → A²³ H2 and T¹¹ N3H → A¹⁰ H2 NOEs were detected. The resonance assign-

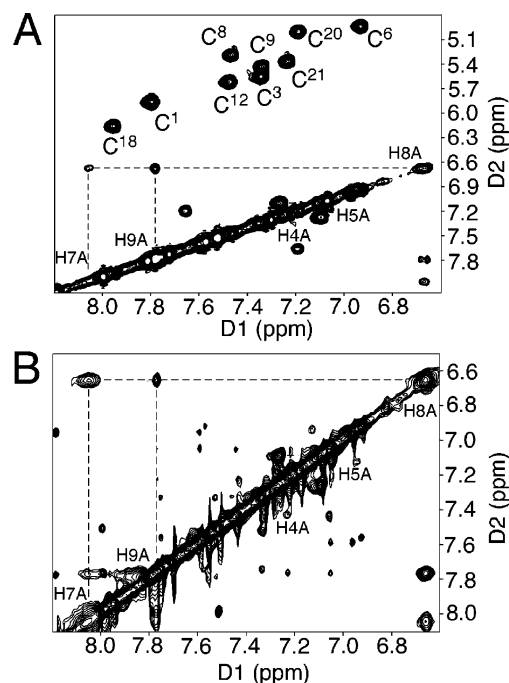


Figure 3. Expanded plots from the COSY spectrum at 30 °C and aromatic–aromatic region of the NOESY spectrum at 15 °C for the *Nar1IQ3* duplex, showing assignments for the IQ protons.

ments for the nonexchangeable protons of the *Nar1IQ3* and nonmodified duplexes are found in Table S3 of the Supporting Information.

(c) IQ Protons. The resonance assignments of the IQ protons were achieved using a combination of COSY and NOESY spectra, collected at 5 °C intervals between 15 and 45 °C. The COSY IQ H4A → H5A cross-peak was observed at all of these temperatures. However, the small scalar coupling between the IQ H8A and H7A protons was not observed below 25 °C, presumably due to line broadening at the lower temperatures. The COSY cross-peaks between the IQ protons also broadened above 35 °C. This could be due to thermal melting of the duplex as the temperature was increased and might also reflect conformational exchange at higher temperatures. Figure 3 compares the NOESY spectrum collected at 15 °C with a magnitude COSY spectrum collected at 30 °C. The IQ H4A proton was assigned at 7.28 ppm on the basis of a cross-peak to the IQ methyl protons in the NOESY spectrum. The IQ H5A proton resonance was assigned at 7.05 ppm on the basis of its scalar coupling to H4A. The IQ H7A, H8A, and H9A proton resonances were distinguished on the basis of comparison of scalar couplings and chemical shifts to those of pyridine. The resonances at 8.06 and 7.79 ppm were assigned to the H7A and H9A protons, and that at 6.70 ppm was assigned to the H8A proton. The H7A resonance exhibited broadening attributed to the pyridinyl nitrogen.

IQ-DNA NOEs. There were 24 NOEs observed between the IQ moiety and DNA protons (Figure 4, and Table S4 in the Supporting Information). The IQ H4A proton exhibited NOEs to G¹⁹ H1' and H5', and to G¹⁷ N1H. The IQ H5A proton exhibited NOEs to G¹⁹ H1', H5', and H5'', G¹⁷ N1H (weak), and C¹⁸ H3' and H4' (weak). The IQ H9A proton exhibited NOEs to G¹⁷ H1', H2', H2'', and H8, and to C¹⁸ H1', H3', H4', H5', and H5''. The IQ methyl protons exhibited NOEs to X⁷ H1', C⁸ H6, G¹⁷ N1H, G¹⁹ H8, and G¹⁹ N1H.

(81) Boelens, R.; Scheek, R. M.; Dijkstra, K.; Kaptein, R. *J. Magn. Reson.* **1985**, *62*, 378–386.

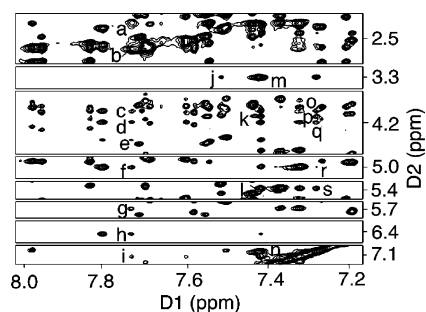


Figure 4. NOE cross-peaks between nonexchangeable protons of DNA and IQ protons in the *Nar1IQ3* duplex. (a–i) $G^{17} H2'$, $G^{17} H2''$, $C^{18} H5'$, $H4'$, $H5''$, $H3'$, $G^{17} H1'$, $C^{18} H1'$, and $G^{17} H8 \rightarrow IQ H9A$; (j) $C^8 H6 \rightarrow IQ CH_3$; (k–l) $G^{19} H5'$ and $G^{19} H1' \rightarrow IQ H4A$; (m–n) $G^{19} H8 \rightarrow IQ CH_3$ and $H5A$; (o–s) $G^{19} H5''$, $G^{19} H5'$, $C^{18} H4'$, $C^{18} H3'$, and $G^{19} H1' \rightarrow IQ H5A$, respectively.

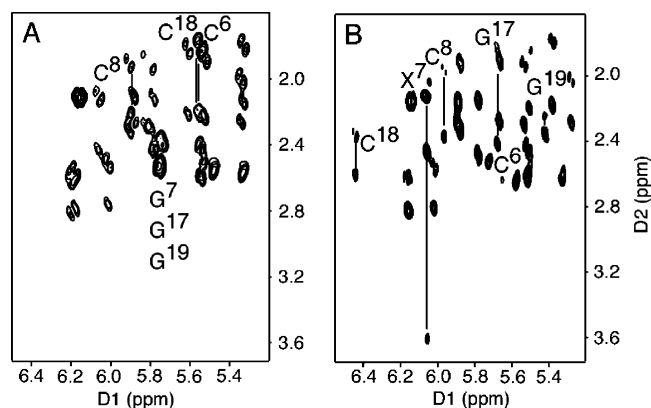


Figure 5. Expanded COSY spectra at 15 °C, establishing connectivity between the $H1'$ and $H2'$, $H2''$ protons. The $H2'$ and $H2''$ protons of the nucleotides adjacent to the lesion site are labeled. (A) Unmodified duplex. (B) *Nar1IQ3* duplex.

Torsion Angle Analysis. The glycosyl torsion angle, χ , was evaluated by inspection of chemical shift data at the deoxyribose $H2'$, $H2''$ protons.^{82,83} Expanded DQF–COSY plots identifying scalar couplings between deoxyribose $H1'$ (5.0–6.7 ppm) and $H2'$, $H2''$ protons in the unmodified and the *Nar1IQ3* duplexes are shown in Figure 5. For the *Nar1IQ3* duplex, the $X^7 H2'$ resonance shifted downfield to 3.61 ppm. This was characteristic of the syn dG orientation at X^7 . Analysis of DQF–COSY spectra suggested that all of the pyrimidine pseudorotation values were in the $C1'$ -exo range of $P = 126 \pm 18^\circ$, and all of the purines in the center 10 base pairs had pseudorotation values in the $C2'$ -endo range of $P = 162 \pm 18^\circ$. The sugar pucker of X^7 was in $C2'$ -endo region. The glycosyl torsion angles and the deoxyribose pseudorotations for the *Nar1IQ3* and unmodified duplexes are found in Table S5 of the Supporting Information.

Figure 6 shows the ^{31}P HMBC correlation spectrum for the *Nar1IQ3* duplex and its unmodified counterpart, and the assignments of P^6 , P^7 , P^{17} , and P^{18} , the phosphodiester linkages 5'- and 3' to the IQ adduct in the modified and complementary strands, respectively. The C^8 -dG IQ adduct dispersed these four ^{31}P resonances, with the most significant change occurring at P^6 , the phosphodiester 5' to X^7 in the modified strand. The downfield ^{31}P chemical shift at P^6 presumably reflects confor-

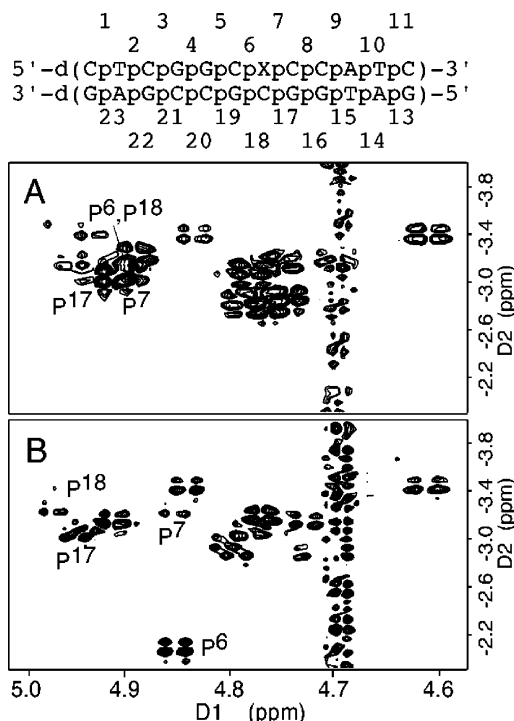


Figure 6. $H3'$ regions of nonselective excitation ^{31}P - 1H HMBC spectra of the unmodified (upper) and the *Nar1IQ3* duplexes (lower). Both spectra were collected at 15 °C.

mational perturbations associated with the P^6 phosphodiester.⁸⁴ The small differences observed for ^{31}P chemical shifts for P^{17} and P^{18} suggested that the phosphodiester linkages opposite to X^7 in the complementary strand were less perturbed. The carcinogen-base linkage site at X^7 residue is defined by the torsion angles α' ($N9-C8-N[IQ]-C2[IQ]$) and β' ($C8-N[IQ]-C2[IQ]-N3-[IQ]$). The absence of an NOE between the IQ NH and methyl protons suggested that the β' torsion angle must be in an eclipsed conformation, placing these protons far apart. Molecular modeling confirmed four stable syn conformations with α' and β' at 0 and 180° in all combinations.

Chemical-Shift Perturbations. The 1H NMR chemical shifts of the *Nar1IQ3* dodecamer were compared with those of the unmodified duplex (Figure 7). The largest perturbations were observed for the aromatic and anomeric protons of C^{18} in the complementary strand, opposite to the adduct. Smaller perturbations were also observed for the $G^{17} H8$, $G^{19} H8$ and $H1'$, $C^6 H6$ and $H1'$, $X^7 H1'$, and $C^8 H6$ and $H1'$ resonances.

Structural Refinement. For the *Nar1IQ3* duplex, a total of 488 NOE-based distance restraints were obtained, consisting of 148 inter- and 340 intra-nucleotide distances. They included 24 DNA–IQ distances. For the unmodified duplex, a total of 463 NOE-based distance restraints were obtained, consisting of 138 inter- and 325 intra-nucleotide distances. For the *Nar1IQ3* duplex, the pyrimidine pseudorotation values were restrained in the $C1'$ -exo range of $P = 126 \pm 18^\circ$, and the purines in the center 10 base pairs were restrained with pseudorotation values in the $C2'$ -endo range of $P = 162 \pm 18^\circ$. No backbone torsion angle restraints were used for the modified strand at the lesion site. Elsewhere, the backbone angles α , β , and ξ were restrained

(82) Fuchs, R. P.; Schwartz, N.; Daune, M. P. *Nature* **1981**, *294*, 657–659.

(83) Norman, D.; Abuaf, P.; Hingerty, B. E.; Live, D.; Grunberger, D.; Brody, S.; Patel, D. J. *Biochemistry* **1989**, *28*, 7462–7476.

(84) Gorenstein, D. G. Phosphorus-31 chemical shifts and spin–spin coupling constant principles and empirical observations. In *Phosphorus-31 NMR Principles and Application*; Gorenstein, D. G., Ed.; Academic Press: New York, 1984; pp 7–56.

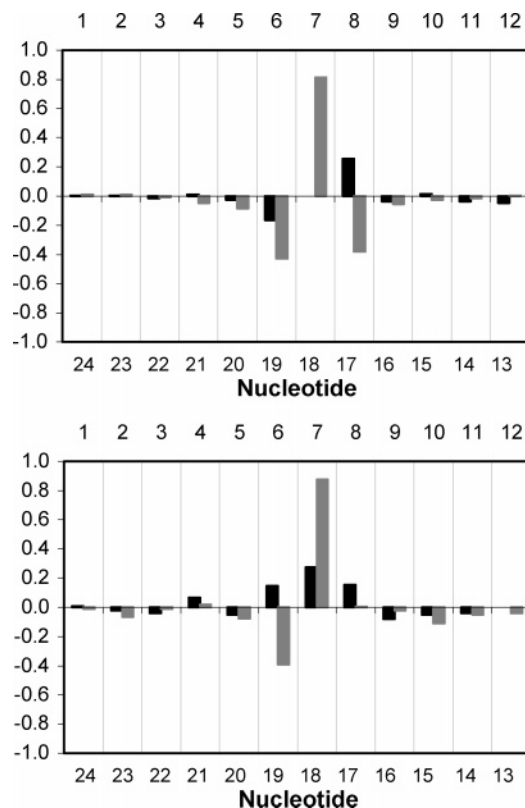


Figure 7. Chemical shift perturbations of (top) H6/H8 and (bottom) H1' protons of the *Nar1IQ3* duplex relative to the unmodified duplex. Dark bars represent the modified strand; light bars represent the complementary strand. $\Delta\delta = [\delta_{\text{modified oligodeoxynucleotide}} - \delta_{\text{unmodified oligodeoxynucleotide}}]$ (ppm).

to $-60 \pm 30^\circ$, $180 \pm 30^\circ$, and $-90 \pm 30^\circ$, respectively, to allow both A- and B-like geometry.⁸⁵ No empirical base-pairing restraints were used at the lesion site. Elsewhere, empirical base-pair planarity and Watson–Crick hydrogen-bonding restraints were used. These were consistent with crystallographic data.⁸⁶ Their inclusion was based on data that showed that DNA maintained Watson–Crick base pairing.

The restrained molecular dynamics calculation employed a simulated annealing protocol. This began with a searching strategy guided by intermolecular IQ–DNA restraints. The DNA starting conformation was B-like except for the syn glycosyl torsion angle at X⁷. The orientation space was searched with 16 energy minimization trials in which the IQ angles α' and β' (Chart 2) were started at 0, 90, 180, and 270° in all combinations. This generated four stable conformations with the glycosyl bond in the syn conformation and the angles α' and β' at 0 and 180°. For the conformations with $\alpha' \approx 0^\circ$ IQ oriented in minor groove, while for conformations with $\alpha' \approx 180^\circ$ IQ intercalated. These were checked with respect to NOE violations. One exhibited the best fit to the NOE data of the *Nar1IQ3* duplex.

Figure 8 shows a stereoview of an ensemble of 10 structures obtained from randomly seeded calculations. Their precision was determined by pairwise rmsd measurements. These exhibited a maximum pairwise rmsd of 0.68 Å, suggesting convergence. Figure S3 of the Supporting Information shows the corresponding results for the unmodified duplex.

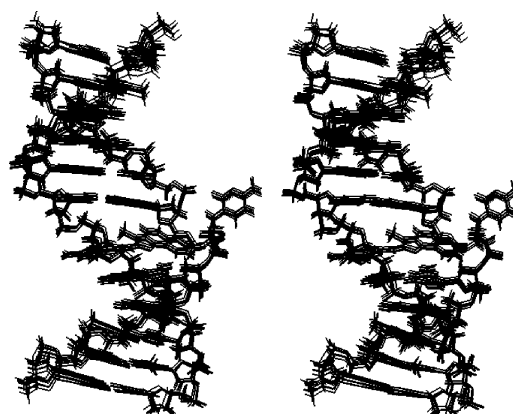


Figure 8. Stereoview of 10 superimposed structures emergent from randomly seeded rMD calculations on the *Nar1IQ3* duplex.

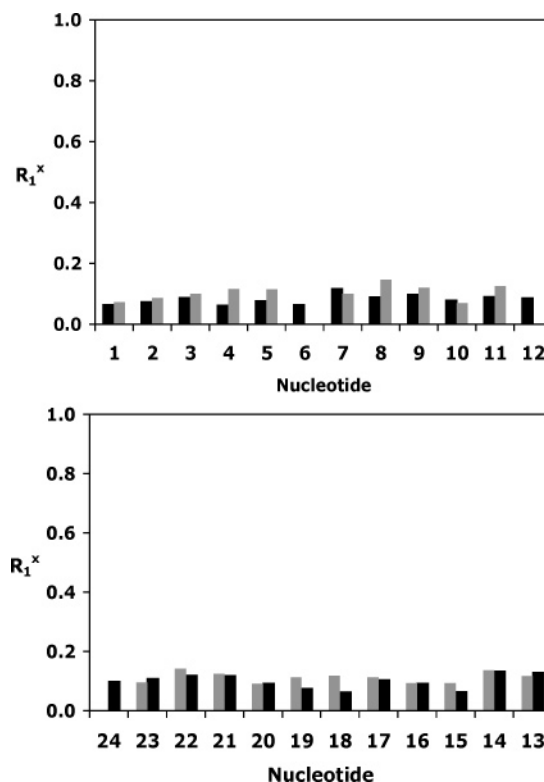


Figure 9. Sixth root residuals for NOE intensities of the *Nar1IQ3* duplex. (Top) Nucleotides C¹ → C¹². (Bottom) Nucleotides G¹³ → G²⁴. Dark bars represent intranucleotide values; light bars represent internucleotide values.

The accuracy of the refinement was determined by calculation of NOE intensities from the emergent structures using the program CORMA (v. 5.2)⁷² (Figure 9). The overall sixth root residual R_1^x for structures of the *Nar1IQ3* duplex was 8.1×10^{-2} . Figure S4 of the Supporting Information shows corresponding data for the unmodified duplex. For the *Nar1IQ3* duplex, inter- and intranucleotide R_1^x values were on the order of 10%. At the adduct site, the residuals were 6.2, 11, 5.2, and 6.1 ($\times 10^{-2}$) for C⁸, G¹⁷, C¹⁸, and G¹⁹, respectively. The structural statistics are found in Table 1.

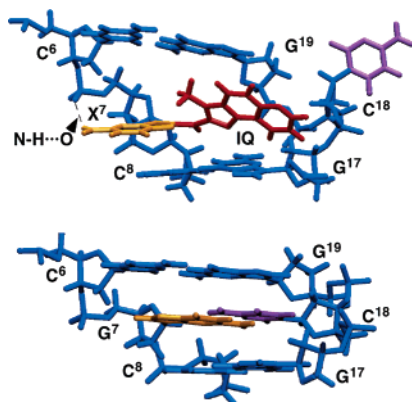
Solution Conformation. The IQ moiety was inserted into the helix with the modified guanine and its complement C¹⁸ displaced in the major groove. Views normal to the helix axis and looking into the major groove of the central 5'-d(C⁶X⁷C⁸)-3'-5'-d(G¹⁷C¹⁸G¹⁹)-3' segments of the unmodified and *Nar1IQ3* duplexes are shown in Figure 10. The IQ region inserted between

(85) Tjandra, N.; Tate, S.-I.; Ono, A.; Kainosho, M.; Bax, A. *J. Am. Chem. Soc.* **2000**, *122*, 6190–6200.

(86) Saenger, W. *Principles of Nucleic Acid Structure*; Springer: New York, 1984.

Table 1. Analysis of RMD-Generated Structures of the Unmodified and Nar1IQ3 Duplexes

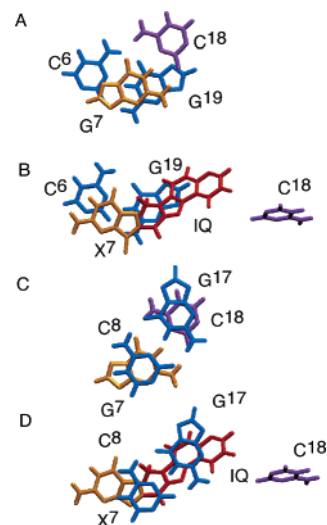
NMR restraints	Nar1	Nar1IQ3
total no. of distance restraints	463	488
interresidue distance restraints	138	148
intraresidue distance restraints	325	340
DNA-IQ distance restraints	0	24
IQ-IQ distance restraints	0	5
H-bonding restraints	33	30
dihedral planarity restraints	24	22
sugar pucker restraints	120	120
backbone torsion angle restraints	90	78
structural statistics		
NMR R-factor (R_1) ($\langle rMDR_i \rangle$)	0.0812 \pm 0.0003	0.0854 \pm 0.0005
rmsd of NOE violations (\AA)	0.00763 \pm 0.00001	0.00798 \pm 0.00002
no. of NOE violations > 0.2 \AA in the root-mean-square deviations from ideal geometry	0	0
bond length (\AA)	0.02402 \pm 0.00005	0.02783 \pm 0.00006
bond angle (deg)	2.613 \pm 0.007	2.672 \pm 0.006
improper angle (deg)	0.64 \pm 0.02	0.79 \pm 0.02
pairwise rmsd (\AA) over all atoms ($\langle rMDR_i \rangle$ vs $\langle rMD_{av} \rangle$)	0.65 \pm 0.01	0.68 \pm 0.02

**Figure 10.** Comparison of the average structures, looking into the major groove and normal to the helix axis of the central segment. (A) Nar1IQ3 duplex. The IQ ring is shown in red and is inserted between base pairs C⁶·G¹⁹ and C⁸·G¹⁷. (B) Unmodified duplex.

the C⁶·G¹⁹ and C⁸·G¹⁷ base pairs by displacing the modified guanine of the syn X⁷ nucleotide into the major groove. The glycosyl torsion angle χ (O4'–C1'–N9–C4) of the X⁷ residue was calculated as $85 \pm 10^\circ$.

The IQ methyl group faced into the helix. This placed the IQ H4A and H5A protons facing into the duplex, toward C¹⁸ and G¹⁹ in the complementary strand, whereas the IQ H7A, H8A, and H9A protons faced into the major groove in the vicinity of G¹⁷ and C¹⁸ in the complementary strand. One proton of the X⁷ exocyclic amino group was close to oxygen at the phosphodiester linkage P⁶ between C⁶ and X⁷. This yielded a N–H...O distance of 2.8 \AA , suggesting the formation of a hydrogen bond. Opposite to X⁷, the insertion of the IQ ring into the helix resulted in the displacement of C¹⁸ into the major groove.

Views looking down the helix axis of the 5'-d(C⁶X⁷C⁸)-3'-5'-d(G¹⁷C¹⁸G¹⁹)-3' segment are shown in Figure 11. The primary interaction involved base pair C⁸·G¹⁷, the 3'-neighboring base pair with respect to X⁷. The G¹⁷ imino proton was shielded by the IQ ring. However, the calculations suggested that the IQ ring tilted with respect to the DNA base pairs, presumably reducing stacking with base pairs C⁶·G¹⁹ and C⁸·G¹⁷. This might be attributed to steric hindrance from the IQ methyl group. This

**Figure 11.** Base stacking of the Nar1IQ3 and the unmodified duplexes. (A) Unmodified duplex. Stacking of C⁶·G¹⁹ and G⁷·C¹⁸. (B) Nar1IQ3 duplex. Stacking of C⁶·G¹⁹ and X⁷·C¹⁸. (C) Unmodified duplex. Stacking of G⁷·C¹⁸ and C⁸·G¹⁷. (D) Nar1IQ3 duplex. Stacking of X⁷·C¹⁸ and C⁸·G¹⁷.

tilt was defined by the IQ torsion angle β' which was measured from the refined structures as $-23 \pm 8^\circ$. The IQ torsion angle α' was calculated as $158 \pm 7^\circ$, resulting in the amine linkage of the IQ adduct being in plane with the C8-modified dG. The calculated glycosyl torsion angles and sugar pseudorotation P of the Nar1IQ3 duplex are found in Table S5 in the Supporting Information. The presence of the C8-dG IQ adduct opposite dC resulted in a bend at the adduct site of $23 \pm 5^\circ$, and helical twist angles of $-56 \pm 3^\circ$ and $-76 \pm 3^\circ$ for base pair steps C⁶ → X⁷ and X⁷ → C⁸, respectively.

Discussion

The synthesis of site-specific C8-dG arylamine oligodeoxy-nucleotide adducts^{62,63} enabled high-resolution conformational studies of a site-specific C8-dG IQ adduct in the Nar1IQ3 duplex. The conformation of the C8-dG IQ adduct in DNA has been of interest because this adduct represents one of the most mutagenic HCAs found in the human diet. Both base-displaced

insertion⁶³ and minor groove⁶⁴ conformations have been proposed for the C8-dG IQ DNA adduct, and indeed, the energetic differences between the two proposed conformations are likely to be modest and dependent upon sequence context.

Base-Displaced Insertion of the IQ Adduct. The present studies reveal that for the *NarI*IQ3 duplex, in which the C8-dG IQ adduct is located at position G³ of the 5'-d(CG¹G²CG³-CC)-3' recognition site of the *NarI* enzyme, the base-displaced insertion conformation is favored. The key evidence supporting the conclusion that the X⁷ glycosyl torsion angle χ was in the syn conformation was the downfield chemical shift for the X⁷ H2' resonance, observed at 3.61 ppm (Figure 5). This downfield shift of the H2' resonance is a characteristic marker of the syn conformation of dG in modified duplexes.^{82,83} This corroborated work showing that the dG-C8 IQ adduct was in the syn conformation at the nucleoside level.⁵³

Rotation of the glycosyl bond into the syn conformation at X⁷ placed the Watson-Crick hydrogen bonding edge of the modified dG into the major groove. The X⁷ imino and amino protons were exposed to solvent. Displacement of the modified dG into the major groove was consistent with the observed X⁷ imino proton chemical shift of 9.6 ppm, which was to the high field of the imino protons of other G•C base pairs. One proton of the amino group of the modified guanine was within 2.8 Å of the oxygen at the phosphodiester linkage P⁶ between C⁶ and X⁷. This suggested the potential for formation of a hydrogen bond that was consistent with the X⁷ amino proton chemical shifts of 6.67 and 8.83 ppm (Figure 2). This hydrogen bond might cause the local electrostatic potential at phosphodiester linkage P⁶ to be perturbed. This notion was supported by the significant ³¹P chemical shift perturbation observed at P⁶ (Figure 7). A strong X⁷ H1' → C⁸ H6 NOE for the X⁷ → C⁸ base step (Figure 1) was consistent with a separation between these protons of 3.0 ± 0.4 Å, as measured in the intensity-refined structures of the duplex.

Evidence supporting insertion of the IQ ring between C⁶•G¹⁹ and C⁸•G¹⁷ base pairs included the upfield shift of the G¹⁷ imino proton as a result of stacking with the intercalated IQ ring (Figure 2). The IQ ring stacked primarily with G¹⁷ and G¹⁹ (Figure 11). There was no stacking between IQ and C⁶ or C⁸. The observed NOEs (Table S4 in the Supporting Information) were consistent with the IQ H4A and H5A protons being directed toward C¹⁸ and G¹⁹ in the complementary strand, and the IQ H7A, H8A, and H9A protons being directed toward G¹⁷ and C¹⁸ in the complementary strand (Figure 10). The IQ methyl protons were closer to G¹⁹ than to C⁸ (Figure 10), also consistent with the observed NOEs (Table S4). The absence of an NOE between the IQ amine and methyl protons was attributed to rapid exchange of the amine proton with solvent.

On the basis of a decrease of the IQ absorption, Elmquist et al.⁶³ suggested a hydrogen bond between the IQ quinoline nitrogen and the exocyclic amine of complementary base C¹⁸. They proposed that this might stabilize the base-displaced insertion conformation with respect to a minor groove conformation. Spectroscopic evidence for this hydrogen bond was not observed. A series of rMD calculations that included this hydrogen bond as a restraint yielded structures that did not agree with the experimental NOEs. Instead, C¹⁸ was displaced into the major groove (Figure 10). This displacement of C¹⁸ resulted in a break in the ¹H sequential NOE connectivity at the G¹⁷ →

C¹⁸ step (Figure 1). This distance was predicted to be 7.9 ± 0.4 Å in the refined structures. The C¹⁸ amino proton resonances were not observed. This was consistent with the displacement of C¹⁸ into the major groove. These proton resonances were presumably broadened due to an intermediate rate of rotation about the C4-N⁴ bond and exchange with solvent.

The present results revealing a base-displaced insertion structure of the C8-dG IQ adduct in the *NarI*IQ3 duplex differ from the predictions of a molecular mechanics study on the duplex 5'-d(G¹G²CX³CCA)-3'•5'-d(TGGCGCC)-3'.⁶⁴ The latter study predicted small energetic differences between groove-bound and base-displaced intercalated conformations of the C8-dG IQ adduct. The favored conformation as predicted by molecular mechanics contained the modified dG in the syn conformation about the glycosyl bond with the IQ in the minor groove and directed in the 3' direction with respect to the modified strand. The calculations suggested that the base-displaced intercalated structure was ~10 kcal/mol higher in energy than the minor groove structure.

Sequence Dependence. The small energetic differences predicted by the molecular mechanics calculations⁶⁴ suggest that the conformations of C8-dG IQ adducts in duplex DNA are likely to be influenced by DNA sequence. Differing conformations may be responsible for increased genotoxicity in hotspot sequences. Sequence analysis of HCA-induced mutations in SOS-induced bacteria revealed the majority of mutations in the *lac Z* gene were clustered in hotspots involving iterated Gs.⁸⁷ The 5'-d(CG¹G²CX³CC)-3' recognition sequence of the *NarI* restriction enzyme examined herein represents a hot spot for -2 frameshift mutations⁸² at G³ when G³ is modified by aromatic amines.⁸⁸⁻⁹⁶ Elmquist et al.⁶³ examined the properties of the C8-dG IQ adduct located in the 5'-d(GGCAGXTGGTG)-3'•5'-d(CACCACCTGCC)-3' duplex, bearing codon 12 of the human *N-ras* protooncogene (underlined). Utilizing a combination of thermal UV melting studies, UV spectroscopy, and circular dichroism, they concluded that the C8-dG IQ adduct adopted a groove-bound conformation in the *ras12* sequence, similar to that predicted by Wu et al.⁶⁴ It will be of interest to examine the sequence dependence of the C8-dG IQ adduct in greater detail.

Comparison with the C8-dG PhIP Adduct. The solution structure of the C8-dG PhIP adduct was reported in 5'-d(CCATCXCTACC)-3'•5'-d(GGTAGCGATGG)-3'.⁵⁹ The PhIP-modified duplex with 5'-d(CXC)-3' sequence adopted a conformation similar to that of the C8-dG IQ adduct in the *NarI*IQ3 sequence. In the PhIP-modified duplex, the C8-dG PhIP adduct existed with the modified dG in the syn conformation and displaced into the major groove. The complementary dC was

- (87) Solomon, M. S.; Morgenthaler, P. M.; Turesky, R. J.; Essigmann, J. M. *J. Biol. Chem.* **1996**, *271*, 18368-18374.
 (88) Burnouf, D.; Koehl, P.; Fuchs, R. P. P. *Proc. Natl. Acad. Sci. U.S.A.* **1989**, *86*, 4147-4151.
 (89) Belguise-Valladier, P.; Fuchs, R. P. P. *Biochemistry* **1991**, *30*, 10091-10100.
 (90) Rodriguez, H.; Loechler, E. L. *Carcinogenesis* **1993**, *14*, 373-383.
 (91) Rodriguez, H.; Loechler, E. L. *Biochemistry* **1993**, *32*, 1759-1769.
 (92) Geacintov, N. E.; Cosman, M.; Hingerty, B. E.; Amin, S.; Broyde, S.; Patel, D. J. *Chem. Res. Toxicol.* **1997**, *10*, 111-146.
 (93) Shukla, R.; Jelinsky, S.; Liu, T.; Geacintov, N. E.; Loechler, E. L. *Biochemistry* **1997**, *36*, 13263-13269.
 (94) Shukla, R.; Liu, T.; Geacintov, N. E.; Loechler, E. L. *Biochemistry* **1997**, *36*, 10256-10261.
 (95) Page, J. E.; Zajc, B.; Oh-hara, T.; Lakshman, M. K.; Sayer, J. M.; Jerina, D. M.; Dipple, A. *Biochemistry* **1998**, *37*, 9127-9137.
 (96) Fernandes, A.; Liu, T.; Amin, S.; Geacintov, N. E.; Grollman, A. P.; Moriya, M. *Biochemistry* **1998**, *37*, 10164-10172.

displaced into the major groove. The IP ring inserted into the duplex, stacking with the flanking G¹⁸ purine and the C⁵ and G¹⁶ rings. However, the out-of-plane geometry of the phenyl ring with respect to the IP ring in the PhIP adduct contributed to a greater unwinding and twisting of the helix as compared to the C8-dG IQ adduct. The PhIP phenyl ring was inclined out-of-plane relative to the IP ring, rotating rapidly, precluding stacking with the flanking bases. Additionally, the PhIP methyl group was positioned toward the modified strand, directed toward the minor groove edge of the DNA, whereas in the NarIIQ3 duplex, the IQ methyl group was stacked between the flanking bases. Corresponding to the above differences, the PhIP-dG linkage site was defined by torsion angles α' and β' by 221.3 ± 3.0 and $132.5 \pm 8.0^\circ$.

Comparison with Aminofluorene and Acetyl-aminofluorene C8-dG Adducts. Adducts arising from the arylamines 2-aminofluorene (AF) and *N*-acetyl-2-aminofluorene (AAF) were extensively studied.¹⁰ The biological responses to the AF and AAF lesions differed,⁹⁷ although they were structurally similar. In bacterial mutagenesis assays, the AAF adduct gave -1 and -2 frameshift mutations,^{88,98–101} whereas base-pair substitutions were largely observed for AF. Both the C8-dG IQ (this work) and C8-dG AAF adducts¹⁰² exhibited single base-displaced inserted structures when placed opposite dC in the 5'-d(CXC)-3' context. NMR data for the C8-dG AF adduct in the NarIIQ3 sequence was equivocal due to a mixture of conformers.¹⁰³ An AF-intercalated conformer with the modified dG in the syn conformation and displaced with the 5'-flanking dC residue into the major groove was reported for the C8-dG AF adduct opposite -2 base deletion in the NarIIQ3 sequence.¹⁰⁴

Structure–Activity Relationships. The 5'-d(CG¹G²CG³CC)-3' NarI sequence represents the strongest known hotspot for frameshift mutagenesis.^{88,105} Within the NarI sequence, the propensity for frameshift mutagenesis is sequence-dependent. These mutations occur following adduct formation at the G³ but not the G¹ or G² positions.^{105,106} A single C8-dG acetyl-aminofluorene adduct located at position G³ induced -2 bp frameshifts more than 10⁷-fold over background mutagenesis in *Escherichia coli*.¹⁰⁷ In a study using site-specifically adducted oligonucleotides, it was found that AAF at the G³-position of the NarI sequence gave almost exclusively two-base deletion when replicated in bacteria, while giving largely base-pair substitution at the G¹ and G²-positions. In COS-7 cells, the C8-AAF adduct gave base-pair substitution at all three positions.¹⁰⁸ The -2 bp frameshift mutations induced at position G³ in the

NarI sequence by the aromatic amine AAF^{109,110} arose via AAF-induced stabilization of a transient strand slippage intermediate during trans-lesion replication,^{110–112} and it is thought that the -2 bp frameshifts induced by the PhIP C8-dG adduct arise via the same mechanism.⁶¹ Crystallographic analysis of the bypass polymerase Dpo4 from *Sulfolobus solfataricus* involving complexes with damaged DNA templates supports the notion that error-prone lesion bypass can involve the formation of transient slippage intermediates.^{113–115} Koffel-Schwartz and Fuchs demonstrated that the dinucleotide repeat GCGC was essential for the -2 bp frameshifts in the NarI sequence, whereas the flanking nucleotides N_aGCGCN_b, particularly N_b, modulated the relative mutagenic strength of the sequence.¹⁰⁷ In the case of AAF, it is thought that the 3'-neighboring base N_b forms favorable stacking interactions with the fluorene ring that stabilize the transient two-base strand slippage intermediate.¹¹⁶

The base-displaced intercalated structure may also play a role in modulating the repair of the C8-dG IQ adduct. Turesky et al.¹¹⁷ proposed that differences in the accumulation and rates of removal of C8-dG IQ and N2-dG IQ adducts in rodents and nonhuman primates may be attributable to differences in conformation about the glycosyl bond in the two classes of adducts. Adducts in the syn form are proposed to create greater distortions of the DNA duplex and, hence, be more easily recognized and excised. In contrast, adducts in the anti conformation are proposed to be more refractory toward repair. Turesky et al.¹¹⁷ observed a preferential removal of the C8-dG IQ adduct, whereas the N2-dG IQ adduct was more persistent. The latter existed in the anti conformation about the glycosyl bond at the nucleoside level.

Structural Coordinates Available. For the unmodified NarI duplex, the PDB ID code is 2HKB. For the IQ-modified NarIIQ3 duplex, the PDB ID code is 2HKC.

Acknowledgment. We thank Drs. Thomas M. and Constance M. Harris for helpful discussions and suggestions. We thank Ms. Pamela Tamura and Ms. Albenza Kozekova for assistance with the oligodeoxynucleotide synthesis and mass spectrometry and Mr. Markus Voehler for assistance with NMR experiments. This work was supported by NIH Grant CA-55678 (M.P.S.). The Vanderbilt Center for Molecular Toxicology is funded by a center grant from the National Institute of Environmental Health Sciences (ES-00267). J.S.S. was supported by NIH predoctoral traineeship ES-07028. Funding for the NMR spectrometers was supplied by NIH Grants RR-05805, ES-00267, and Vanderbilt University.

Supporting Information Available: Tables S1 and S2, nonexchangeable proton chemical shifts of the NarIIQ3 and

(97) Beland, F. A.; Kadlubar, F. F. *Environ. Health Perspect.* **1985**, *62*, 19–30.

(98) Fuchs, R. P. P. *J. Mol. Biol.* **1983**, *177*, 173–180.

(99) Bichara, M.; Fuchs, R. P. P. *J. Mol. Biol.* **1985**, *183*, 341–351.

(100) Melchior, W. B., Jr.; Marques, M. M.; Beland, F. A. *Carcinogenesis* **1994**, *15*, 889–899.

(101) Tebbs, R. S.; Romano, L. J. *Biochemistry* **1994**, *33*, 8998–9006.

(102) O'Handley, S. F.; Sanford, D. G.; Xu, R.; Lester, C. C.; Hingerty, B. E.; Broyde, S.; Krugh, T. R. *Biochemistry* **1993**, *32*, 2481–2497.

(103) Mao, B.; Hingerty, B. E.; Broyde, S.; Patel, D. J. *Biochemistry* **1998**, *37*, 95–106.

(104) Mao, B.; Gorin, A.; Gu, Z.; Hingerty, B. E.; Broyde, S.; Patel, D. J. *Biochemistry* **1997**, *36*, 14479–14490.

(105) Broschard, T. H.; Koffel-Schwartz, N.; Fuchs, R. P. P. *J. Mol. Biol.* **1999**, *288*, 191–199.

(106) Burnouf, D. Y.; Mitorski, R.; Fuchs, R. P. *Chem. Res. Toxicol.* **1999**, *12*, 144–150.

(107) Koffel-Schwartz, N.; Fuchs, R. P. P. *J. Mol. Biol.* **1995**, *252*, 507–513.

(108) Tan, X.; Suzuki, N.; Grollman, A. P.; Shibutani, S. *Biochemistry* **2002**, *41*, 14255–14262.

(109) Lambert, I. B.; Napolitano, R. L.; Fuchs, R. P. P. *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89*, 1310–1314.

(110) Garcia, A.; Lambert, I. B.; Fuchs, R. P. P. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 5989–5993.

(111) Milhe, C.; Dhalluin, C.; Fuchs, R. P.; Lefevre, J. F. *Nucleic Acids Res.* **1994**, *22*, 4646–4652.

(112) Milhe, C.; Fuchs, R. P.; Lefevre, J. F. *Eur. J. Biochem.* **1996**, *235*, 120–127.

(113) Ling, H.; Boudsocq, F.; Woodgate, R.; Yang, W. *Cell* **2001**, *107*, 91–102.

(114) Ling, H.; Boudsocq, F.; Woodgate, R.; Yang, W. *Mol. Cell* **2004**, *13*, 751–762.

(115) Zang, H.; Goodenough, A. K.; Choi, J. Y.; Irimia, A.; Loukachevitch, L. V.; Kozekov, I. D.; Angel, K. C.; Rizzo, C. J.; Egli, M.; Guengerich, F. P. *J. Biol. Chem.* **2005**, *280*, 29750–29764.

(116) Roy, D.; Hingerty, B. E.; Shapiro, R.; Broyde, S. *Chem. Res. Toxicol.* **1998**, *11*, 1301–1311.

(117) Turesky, R. J.; Box, R. M.; Markovic, J.; Gremaud, E.; Snyderwine, E. G. *Mutat. Res.* **1997**, *376*, 235–241.

unmodified duplexes; S3, exchangeable proton chemical shifts for the *Nar1IQ3* and unmodified duplexes; S4, comparison of experimental intermolecular distance restraints with those observed for intensity-refined structures of the *Nar1IQ3* duplex; and S5, pseudorotation and glycosyl torsion angles for the *Nar1IQ3* and unmodified duplexes. Figures S1, contour plots of the anomeric to aromatic region of the ^1H NOESY spectrum for the unmodified duplex; S2, contour plot of the imino protons

of the unmodified duplex; S3, stereoview of 10 superimposed structures emergent from rMD calculations on the unmodified duplex; and S4, R_1^x values for refined structures of the unmodified duplex calculated using the program CORMA (v. 5.2).⁷² This material is available free of charge via the Internet at <http://pubs.acs.org>.

JA062004V