involved.[10] The calculated bond lengths for HBCBH are appropriate for a linear molecule with two carbon–boron double bonds.

The ab initio calculated vibrational frequencies are also listed in Table I. The calculated $\nu_{BCB}$ frequencies for all isotopes are converted to the observed frequency values by a constant scaling factor of 0.930, which is expected for SCF calculations. The calculated $\nu_{BH}$ frequencies require a scale factor of 0.928; the failure to observe this mode in $CH_4$ experiments is probably due to parent molecule absorptions. The excellent prediction of isotopic frequencies by scaled SCF calculations confirms the identification of HB=C=BH.

The mechanism for reaction of B atoms with $CH_4$ will be considered in detail after all of the products are identified. Semiempirical calculations predict that B inserts into a C–H bond in $CH_4$ without activation energy.[11] Two successive B atom insertions into C–H bonds in methane, which must be highly exothermic reactions, followed by $H_2$ elimination are required for the formation of HBCBH.

None of the frequencies reported for carbon–boron double bonds are due to an isolated C=B subgroup.[2,3] The present SCF/DZP calculations predict antisymmetric and symmetric B=C=B stretching modes at 2012 and 1196 $cm^{-1}$, respectively. The average value scaled by 0.93 is 1492 $cm^{-1}$, which represents a prototype C=B subgroup stretching fundamental. It is perhaps noteworthy that SCF/DZP calculations for $H_2C$=BH predict the double-bond stretching mode at 1608 $cm^{-1}$, which scales (× 0.93) to 1495 $cm^{-1}$.

**Acknowledgment** is made for financial support from NSF Grant CHE 91-22556.

(1) Jeong, G. H.; Boucher, R.; Klabunde, K. J. *J. Am. Chem. Soc.* **1990**, *112*, 3332.
(2) Klusic, H.; Berndt, A. *Angew. Chem., Int. Ed. Engl.* **1983**, *22*, 877.
(3) Glaser, E.; Hanecher, E.; Noth, H.; Wager, E. *Chem. Ber.* **1987**, *120*, 659.
(4) Boese, R.; Paetzold, P.; Tapper, A. *Chem. Ber.* **1987**, *120*, 1069.
(5) Burkholder, T. R.; Andrews, L. *J. Chem. Phys.* **1991**, *95*, 8697.
(6) Hassanzadeh, P.; Andrews, L. *J. Phys. Chem.*, in press.
(7) Dupuis, M.; Watts, J. D.; Villar, H. D.; Hurst, G. J. B. *Comput. Phys. Commun.* **1989**, *52*, 415. Dupuis, M.; Rhys, J.; King, H. F. *J. Chem. Phys.* **1976**, *65*, 111.
(8) Cook, C. M.; Allen, L. C. *Organometallics* **1982**, *1*, 246.
(9) Frenking, G.; Schaefer, H. F., III *Chem. Phys. Lett.* **1984**, *109*, 521.
(10) Alberts, I. L.; Schafer, H. F., III *Chem. Phys. Lett.* **1990**, *165*, 250.
(11) Lebrilla, C. B.; Maier, W. F. *Chem. Phys. Lett.* **1984**, *105*, 183.

## Formaldehyde Preferentially Interstrand Cross-Links Duplex DNA through Deoxyadenosine Residues at the Sequence 5′-d(AT)

Huifang Huang, Marjorie S. Solomon, and Paul B. Hopkins*

*Department of Chemistry*
*University of Washington*
*Seattle, Washington 98195*

*Received July 17, 1992*

Formaldehyde is a toxic substance ubiquitous in living systems and the environment.[1] From dietary sources alone, the average adult human consumes up to an estimated 14 mg of formaldehyde per day.[1] Formaldehyde is a DNA-denaturing[2] and interstrand-cross-linking agent.[3] We report here that formaldehyde preferentially forms dA-to-dA cross-links as in **1** at the dinucleotide sequence 5′-d(AT) in certain AT-rich sequences of duplex DNA.

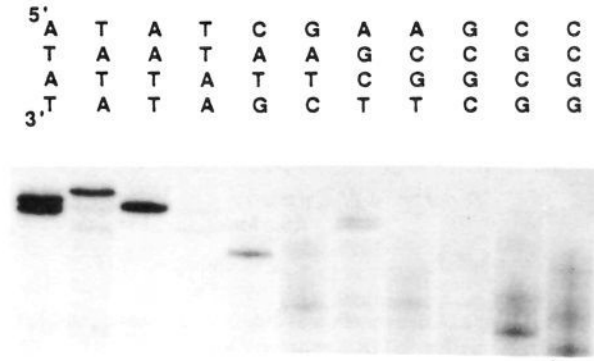



**Figure 1.** Autoradiogram of DPAGE analysis of interstrand-cross-linked products from formaldehyde-treated partially 5′-$^{32}$P-labeled [5′-d-(TACAACN$_4$GTTGT)]$_2$, $N_4$ as indicated.[4]

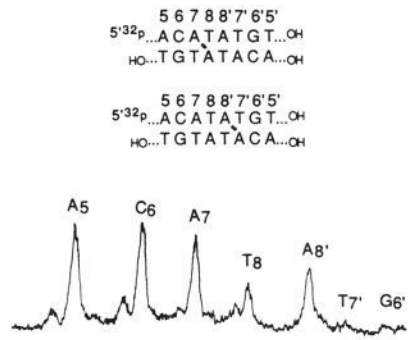
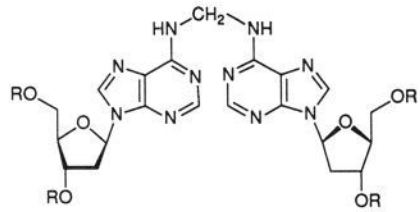



**Figure 2.** DPAGE analysis of partially 5′-$^{32}$P-labeled, formaldehyde-cross-linked [5′-d(TACAACATATGTTGT)]$_2$ following treatment with iron(II) EDTA/ascorbic acid/$H_2O_2$[5] reveals indicated sites of interstrand cross-linking.




**1** R= DNA chain

**2** R=H

A panel of self-complementary, 5′-$^{32}$P-radiolabeled duplexes 5′-d(TACAACN$_4$GTTGT) (Figure 1) was exposed to 25 mM formaldehyde (pH 6.0, 25 mM NaCl, 50 mM sodium phosphate buffer, 25 °C, 9 days). Each of the duplexes afforded several interstrand-cross-linked products (DPAGE), but the duplexes containing $N_4$ = ATAT and AATT preferentially afforded respectively two equiabundant products and one product, consistent with cross-linking centered on the shared 5′-d(AT) sequence.[4] A

(1) WHO Task Group on Environmental Health Criteria for Formaldehyde. Environmental Health Criteria 89, Formaldehyde; World Health Organization: Geneva, 1989.
(2) (a) McGhee, J. D.; von Hippel, P. H. *Biochemistry* **1975**, *14*, 1281, 1297. (b) McGhee, J. D.; von Hippel, P. H. *Biochemistry* **1977**, *16*, 3267, 3276.
(3) (a) Feldman, M. Ya. *Prog. Nucleic Acid Res. Mol. Biol.* **1973**, *13*, 1. (b) Freifelder, D.; Davison, P. F. *Biophys. J.* **1963**, *3*, 49. (c) van der Eb, A. J.; van Kestern, L. W.; van Brugger, E. F. J. *Biochim. Biophys. Acta* **1969**, *182*, 530. (d) Chaw, Y. F. M.; Crane, L. E.; Lange, P.; Shapiro, R. *Biochemistry* **1980**, *19*, 5525.

(4) Single strands and end-cross-linked products common to all DNAs are omitted in Figure 1 (see: Weidner, M. F.; Sigurdsson, S. Th.; Hopkins, P. B. *Biochemistry* **1990**, *29*, 9225). Because 5′-$^{32}$P-phosphorylated and excess hydroxyl-terminal strands were admixed, the symmetry of the two 5′-d(AT) sites in $N_4$ = ATAT is broken and two products are observed.[5b] Yields (%) of the major cross-linked product in each lane/total yield of cross-linked products exclusive of end-cross-linked products (phosphorimager) were as follows: ATAT 0.31 and 0.38/1.44; TATA 0.18/0.93; AATT 0.35/0.91; TTAA 0.04 and 0.04/0.39; CATG 0.16/0.91; GATC −/2.02; AGCT 0.08/1.19; ACGT −/1.45; GCGC −/1.14; CGCG 0.19/1.76; CCGG 0.18/2.03. These yields emphasize the distinction between *duplexes* which are efficiently cross-linked ($N_4$ = GATC and CCGG) and *specific sites within duplexes* [5′-d(AT)], which are preferentially cross-linked.

0002-7863/92/1514-9240$03.00/0 © 1992 American Chemical Society

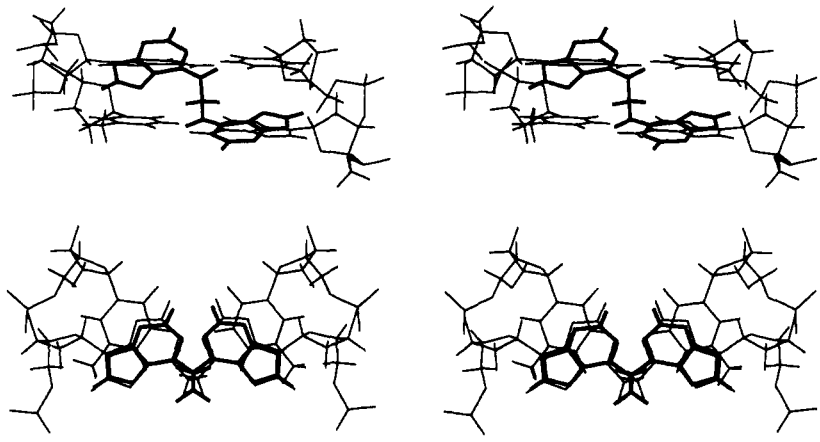


**Figure 3.** Overlay of B-DNA [5′-d(AT)]$_2$ superimposed upon the linked adenine groups (bold) of energy-minimized [5′-d(AT)]$_2$ cross-linked as in **1**.[7] Upper: View from the major groove. Lower: View down helix axis.

single interstrand-cross-linked product stood out in the duplexes $N_4$ = TATA and CATG, but these bands were comparable in intensity to those in the duplexes $N_4$ = CGCG and CCGG.[4] The nucleotide connectivity of cross-linked $N_4$ = ATAT was revealed by sequence-random fragmentation.[5*] The quantified distribution of products (Figure 2) (ca. 50% intensity of dT(8) and dA(8′); ca. 0% intensity of dT(7′); etc.) was consistent with an equibundant mixture of dA-to-dA cross-links at the pseudo-symmetry-related[4] 5′-d(AT) sequences.

Cross-linked $N_4$ = ATAT was digested with snake venom (type I) and spleen (type II) phosphodiesterases and calf intestinal alkaline phosphatase.[3d] Separation of the resulting mixture by RP-HPLC afforded substance **2**, based upon[3d] (a) *m/e* 515 (M + 1, electrospray ionization) consistent with bridging of two dA residues by a single methylene linkage, (b) 500-MHz $^1$H NMR spectrum containing nine[6] resonances indicative of $C_2$ symmetry, and (c) reduction[6] with aqueous sodium borohydride, which afforded a mixture of 2′-deoxyadenosine and $N^6$-methyl-2′-deoxyadenosine, defining $N^6$ as the site of alkylation on both strands. Assuming that $\epsilon_{260}$ for **2** is twice that of dA, 0.8 mol (expected: 1.0) of **2** was detected (RP-HPLC) per mole of cross-linked duplex.

The selectivity described herein was unexpected. Formaldehyde-treated DNA has previously afforded five pairings of dA, dC, and dG structurally analogous to and including **2**.[3d] From these could arise six distinguishable interstrand cross-links even if cross-linkable amino groups must reside in one groove and in adjacent base pairs of B-DNA. The origin of 5′-d(AT) selectivity is unknown, but may reflect preferential monoadduct formation in AT-rich regions[2b] and/or what we speculate may be rapid closure of monoadducts at 5′-d(AT) to cross-links due to proximity effects. Molecular mechanics energy minimization[7] of [5′-d(AT)]$_2$ cross-linked as in **1** (Figure 3) revealed propeller hypertwisting of the base pairs, but little torsional reorganization relative to B-DNA, suggesting the alternative explanation that thermodynamic stability may favor accumulation of this linkage.

**Acknowledgment.** We thank the NSF (CHE-9019129 and DIR-8219099) and NIH (AG00417). P.B.H. is a Cope Scholar; M.S.S. was an NIH predoctoral fellow (GM08268).

**Supplementary Material Available:** RP-HPLC trace of enzymatic digest of cross-linked $N_4$ = ATAT and $NaBH_4$-reduced **2**, MS and $^1$H NMR spectrum of **2**, and computed structure of **1** in [5′-d(AT)]$_2$ (4 pages). Ordering information is given on any current masthead page.

(5) (a) Weidner, M. F.; Millard, J. T.; Hopkins, P. B. *J. Am. Chem. Soc.* **1989**, *111*, 9270. (b) Millard, J. T.; Weidner, M. F.; Kirchner, J. J.; Ribeiro, S.; Hopkins, P. B. *Nucleic Acids Res.* **1991**, *19*, 1885.

(6) Resonances for $H^2$ and $H^8$ were coincident (500 MHz, $D_2O$). Samples of **2** for the NMR and $NaBH_4$ experiments were from enzymatic digests of $CH_2O$-treated 5′-d[(AT)$_{12}$]. **2** from both sources coeluted in RP-HPLC and gave comparable mass spectra.

(7) Performed using the AMBER force field as previously described: Kirchner, J. J.; Solomon, M. S.; Hopkins, P. B. In *Structure & Function Volume I: Nucleic Acids*; Sarma, R. H., Sarma, M. H., Eds.; Adenine Press: Albany, 1992; p 171.