The Mutagenic Damaged DNA Base, 5,6-Dihydrouracil (DHU), Incorporated into a 14-mer Duplex: NMR Evidence That DHU Is Intrahelical and Causes Minimal DNA Distortion

Julie M. Villanueva,[†] Jan Pohl,[§] Paul W. Doetsch,*[,‡,∥] and Luigi G. Marzilli*[,†]

Departments of Chemistry and Biochemistry
Emory Microchemical Facility
Division of Cancer Biology and
Department of Radiation Oncology
School of Medicine, Emory University
Atlanta, Georgia 30322

Received June 17, 1999

5,6-Dihydrouracil (DHU, Figure 1), formed in substantial amounts by ionizing radiation damage to cytosine under anoxic conditions,[1] is highly mutagenic at both replication and transcription levels.[2,3] DHU is a substrate for the *Escherichia coli* endonuclease III (Nth), a DNA base excision repair *N*-glycosylase found in several species including humans.[4−6] Nth enzymes, highly conserved structurally and functionally throughout nature, initiate repair of many types of toxic and mutagenic pyrimidine base damage.

No data describing the effect of DHU on the structure of a DNA duplex have been reported. DHU is sensitive to strongly basic conditions,[7,8] making deprotection of a phosphoramidite-synthesized oligonucleotide challenging. Procedures for obtaining DHU-containing oligonucleotides for enzymatic assays are impractical for obtaining quantities needed for NMR spectroscopy.[2,4−6] The self-complementary 14-mer, 5′-d($C_1G_2C_3G_4A_5C_6A_7T_8G_9T_{10}$-$DHU_{11}G_{12}C_{13}G_{14}$)-3′ (**DHU-14**) was synthesized in adequate quantities by a superior route using DHU, phenoxyacetyl-protected G and A, and acetyl-protected C phosphoramidites, and deprotected by using K_2CO_3 in anhydrous methanol at room temperature. These mild deprotection conditions afforded only one 14-mer, which could be altered under more basic conditions.[9] ^{1}H NMR and MALDI MS data confirm that unchanged DHU is present in **DHU-14** (m/z^+ Calcd: 4266.4. Found: 4267.1).

The human nucleotide excision repair protein XPG has been implicated as a cofactor for human Nth1 recognition and binding to both DHU- and thymine glycol (Tg)-containing substrates.[10] Tg,[11−13] produced by oxidizing agents and radiation,[1,6,10] is similar to DHU in that both are saturated, nonplanar bases. Thus, it was of particular interest to compare the structures of duplexes with DHU and Tg. The possibility of a common structural basis for the efficient recognition and excision of DHU and Tg by Nth proteins has not been directly investigated. XPG alone has been shown to cleave small non-hydrogen-bonded bulged structures in DNA.[14] A Tg-containing 11-mer has an extrahelical Tg, causing a significant distortion in the duplex.[15,16] On this basis, it was proposed that XPG might recognize extrahelical Tg, resembling a small bulged structure, and bind to it preferentially over normal, duplex B-DNA.[10] We found that *E. coli* Nth recognized and cleaved the DHU site in **DHU-14** (∼60% efficiency) but did not cleave the undamaged **C-14**, [d(CGCGACATGTCGCG)]$_2$. Thus, **DHU-14** represented an excellent model for further structural investigation.

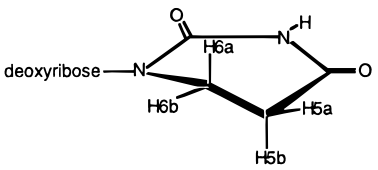

Figure 1. DHU_{11} puckering in **DHU-14**.

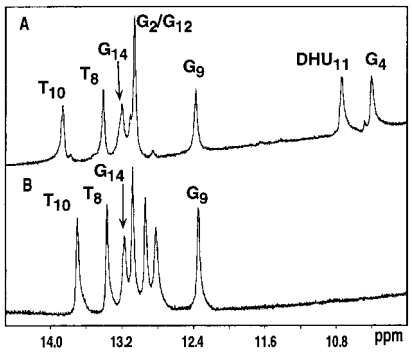

Figure 2. Imino proton spectrum (pH 6.5, 5 °C) of 2 mM (strands) **DHU-14** (A) and **C-14** (B). Imino signals were assigned as indicated by 1D NOE experiments. Smaller signals in the **DHU-14** spectrum indicate the presence of another conformer, probably a hairpin.

Downfield imino proton signals for the expected 6 (**DHU-14**) and 7 (**C-14**) Watson−Crick base pairs were observed in H_2O (Figure 2, two **DHU-14** signals overlap at ∼13.07 ppm). In addition, two upfield imino proton signals were found for **DHU-14**, indicating that these are G_4 and DHU_{11} NH signals; this is the first compelling evidence that DHU and its opposing G are intrahelical, since NH signals from extrahelical bases are broadened by exchange with water. More importantly, irradiation of either upfield signal produced a large NOE to the other upfield signal (Supporting Information). These data require that DHU_{11} and G_4 are in close proximity, thus occupying primarily intrahelical positions. Similar results were found for G−T or G−U wobble base pairs.[17,18] Only the terminal G_{14} imino signal had broadened and disappeared for both duplexes at 25 °C (not shown), indicating that both are relatively stable (UV T_m in 100 mM NaCl: 46 °C for **DHU-14**, 50 °C for **C-14**).

[†] Department of Chemistry.
[‡] Department of Biochemistry.
[§] Emory Microchemical Facility.
[∥] Division of Cancer Biology and Department of Radiation Oncology.

(1) Dizdaroglu, M.; Laval, J.; Boiteux, S. *Biochemistry* **1993**, *32*, 12105−12111.
(2) Liu, J.; Zhou, W.; Doetsch, P. W. *Mol. Cell. Biol.* **1995**, *15*, 6729−6735.
(3) Liu, J.; Doetsch, P. W. *Nucleic Acids Res.* **1998**, *26*, 1707−1712.
(4) Augeri, L.; Hamilton, K. K.; Martin, A. M.; Yohannes, P.; Doetsch, P. W. In *Methods of Enzymology*; Academic Press: 1994; Vol. 234; pp 102−115.
(5) Augeri, L.; Lee, Y.-M.; Barton, A. B.; Doetsch, P. W. *Biochemistry* **1997**, *36*, 721−729.
(6) You, H. J.; Swanson, R. L.; Doetsch, P. W. *Biochemistry* **1998**, *37*, 6033−6040.
(7) Sander, E. G. *J. Am. Chem. Soc.* **1969**, *91*, 3629−3634.
(8) Pojarlieff, I. G. *Tetrahedron* **1967**, *23*, 4307−4315.
(9) In initial studies, two methods of deprotection of 5′-CGCGAATT-(DHU)GCG-3′ (NH_4OH at room temperature for 24 h or anhydrous NH_3 gas by Glen Research) were evalutated. After ion exchange purification of both samples, we found two 12-mer products by electrophoresis.
(10) Klungland, A.; Hoss, M.; Gunz, D.; Constantinou, A.; Clarkson, S. G.; Doetsch, P. W.; Bolton, P. H.; Wood, R. D.; Lindahl, T. *Mol. Cells* **1999**, *3*, 33−42.
(11) Ide, H.; Kow, Y. W.; Wallace, S. S. *Nucleic Acids Res.* **1985**, *13*, 8035−8052.
(12) Rouet, P.; Essigmann, J. M. *Cancer Res.* **1985**, *45*, 6113−6118.
(13) Hayes, R. C.; Petrullo, L. A.; Huang, H.; Wallace, S. S.; LeClerc, J. E. *J. Mol. Biol.* **1988**, *201*, 239−246.
(14) Evans, E.; Fellows, J.; Coffer, A.; Wood, R. D. *EMBO J.* **1997**, *16*, 625−638.
(15) Kao, J. Y.; Goljer, I.; Phan, T. A.; Bolton, P. H. *J. Biol. Chem.* **1993**, *268*, 17787−17793.
(16) Kung, H. C.; Bolton, P. H. *J. Biol. Chem.* **1997**, *272*, 9227−9236.
(17) Hare, D.; Shapiro, L.; Patel, D. J. *Biochemistry* **1983**, *25*, 7445−7456.
(18) Carbonnaux, C.; Fazakerley, G. V.; Sowers, L. C. *Nucleic Acids Res.* **1990**, *18*, 4075−4081.

10.1021/ja9920516 CCC: $18.00
Published on Web 10/30/1999

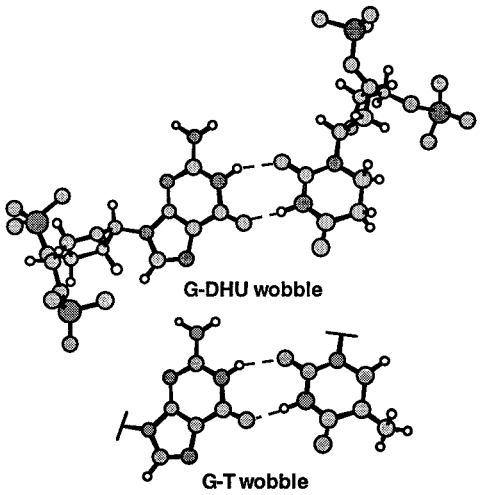

Figure 3. Schematic representation of the G_4-DHU_{11} wobble base pair in **DHU-14** and a G−T wobble base pair.

COSY and NOESY spectra (800 MHz, D_2O) of **DHU-14** exhibit cross-peaks among all four DHU_{11} base CH signals. The two signals with stronger NOE cross-peaks to DHU_{11}H1′ were assigned to H6′s. NOE cross-peaks from the $T_{10}CH_3$ signal to signals at 2.11 and 2.81 ppm assign these to H5a and H6a (Figure 1), respectively; the upfield shifts of these signals relative to the respective 2.70 ppm H5b and 3.44 ppm H6b signals (probably caused by shielding by the nearby T_{10} base) are consistent with a G_4-DHU_{11} wobble base pair (Figure 3). WATERGATE NOESY data (H_2O, Supporting Information) allow assignment of the upfield G_4N(1)H (10.39 ppm) and DHU_{11}N(3)H (10.73 ppm) signals through DHU_{11}N(3)H NOE cross-peaks to the $T_{10}CH_3$ and the four DHU base signals. The H5b signal had the strongest NOE cross-peak, followed by the H5a signal. These NOE data suggest the DHU ring pucker in Figure 1.

Base to H1′ (Supporting Information) and to H2′/2″ NOE cross-peaks were evident for all residues, suggesting a right-hand helix for **DHU-14**. The base proton/H1′ intraresidue NOE cross-peaks are weaker than the cytidine H5/H6 NOE cross-peaks, indicating that the residues are all *anti*.[19,20] Molecular modeling calculations performed without NMR-restraints predicted an N sugar for the DHU residue in a G-DHU base pair in a duplex.[21] However, the similar intensities of base to H3′ NOE cross-peaks (Supporting Information) and all H1′/H2′ COSY cross-peaks indicate an S sugar pucker for all residues.

What is the structural basis for substrate recognition by the Nth family of DNA repair enzymes? Our NMR data indicate that, in contrast to the extrahelical Tg,[15,16] no substantial structural departures from normal, B-form DNA were caused by the presence of two G-DHU wobble base pairs in **DHU-14**. Our findings suggest that significant disruptions in the duplex structure in the absence of bound repair enzyme are not likely key elements for substrate recognition by the Nth family of enzymes. The XPG protein improves binding of human Nth to damaged DNA and greatly promotes activity on Tg and DHU, possibly by bending DNA at damaged sites in order to expose the damaged base.[10] Thus, for modified bases that are extrahelical (Tg) or intrahelical (DHU), substrate recognition probably results from structural changes induced following interaction between damaged DNA and Nth and its homologues. Uracil-DNA glycosylase (UDG) recognizes and removes uracil, a frequently occurring mutagenic cytosine damage; the G-U base pair is also intrahelical.[18,22] However, the X-ray crystal structure of UDG bound to an oligonucleotide substrate showed that the cleaved uracil was in the UDG binding pocket, suggesting that UDG flipped uracil out from the major groove.[23] Nth may have a similar base-flipping mechanism and binding pocket for flipped-out damaged bases.

Acknowledgment. We thank Drs. John Glushka (CCRC, University of Georgia) and Sekar Chandrasekaran (Georgia State University) for assistance with 2D NMR experiments, Dr. Danny Jue at the CDC Biotechnology Facility for MS data, and Dr. Robert Somers (Glen Research) for helpful conversations. This work was supported by NIH grant CA 73041 (NCI).

Supporting Information Available: Details of the synthesis, purification, NMR spectroscopy, and MALDI MS experiments; figures of 800 MHz ^{1}H−^{1}H NOESY spectra, 600 MHz ^{1}H 1D NMR spectra, a UV-shadowed gel, and an intra- vs extrahelical base scheme (PDF). This material is available free of charge via the Internet at http://pubs.acs.org.

JA9920516

(19) Wüthrich, K. In *NMR of Proteins and Nucleic Acids*; John Wiley and Sons: New York, 1986; pp 233−251.
(20) Feigon, J.; Sklenar, V.; Wang, E.; Gilbert, D. E.; Macaya, R. F.; Schultze, P. In *Methods in Enzymology*; Lilley, D. M. J., Dahlberg, J. E., Eds.; Academic Press: New York, 1992; Vol. 211; pp 235−253.
(21) Cooney, M. G.; Doetsch, P. W. In *Radiation Damage in DNA: Structure/Function Relationships at Early Times*; Fuciarelli, A. F., Zimbrick, J. D., Eds.; Battelle Press: Columbus, OH, 1995; pp 299−302.
(22) Roongta, V. A.; Jones, C. R.; Gorenstein, D. G. *Biochemistry* **1990**, *29*, 5245−5258.
(23) Slupphaug, G.; Mol, C. D.; Kavli, B.; Arvai, A. S.; Krokan, H. E.; Tainer, J. A. *Nature* **1996**, *384*, 87−91.