Quinone Methide Alkylation of Deoxycytidine

Steven E. Rokita,*,† Jianhong Yang,‡ Praveen Pande,† and William A. Greenberg‡

Department of Chemistry and Biochemistry, University of Maryland, College Park, Maryland 20742, and Department of Chemistry, State University of New York at Stony Brook, Stony Brook, New York 11794

Received January 3, 1997

DNA alkylation is considered both a pathological and a therapeutic reaction since it is associated paradoxically with induction as well as treatment of cancer. Structural characterization of the resulting lesions is therefore central to both understanding the molecular basis of the disease and developing new anticancer antibiotics. Many of the pioneering studies in this field began by analyzing the fundamental reactions of DNA and simple alkylating agents.[1] More recently, modification of nucleobases by environmental toxins, natural and synthetic drugs, and related metabolites has received considerable attention.[2] Identifying the exact site of modification becomes particularly critical when designing new reagents for selective DNA cross-linking.[3]

The specificity of DNA alkylation depends greatly on the nature of the reaction pathway and the ability of the reactive intermediate to associate with particular nucleotide sequences or helical conformations. Some consistent trends have been observed in previous studies, but our predictive abilities remain limited. In general, reactions involving S_N2-like processes primarily modify nitrogen nucleophiles, most notably N7 of guanine and N1 of adenine.[4] In contrast, S_N1-like processes are less discriminating and often favor oxygen nucleophiles, most notably the phosphate and ribose oxygens as well as O^2 of cytosine and O^6 of guanine.[4] Conjugated and soft electrophiles exhibit yet another specificity that may in part be influenced by preferential binding and activation on the surface of DNA. For example, CC-1065 predominantly alkylates the N3 of adenine within A-T rich regions of duplex DNA.[5] Compounds that react through quinone methide and related intermediates couple to the *exo*-amino groups of guanine (N^2)[6] and adenine (N^6).[7] The relative reactivity of the cyclic nitrogen (N3) and exocylic nitrogen (N^4) of cytosine in the presence of such an intermediate is not well documented, and no consensus emerges from results of other alkylating agents. Dimethyl sulfate and methylnitrosourea principally alkylate N3, whereas ethylnitrosourea alkylates N^4 to a greater extent than N3.[4] Likewise, benzyl bromide modifies N3, *p*-methoxybenzyl bromide modifies N^4, and 7-(bromomethyl)benz[*a*]anthracene modifies both N3 and N^4.[8] A single report on a quinone methide-like intermediate formed by a pyrrolizidine derivative has suggested N3 as the possible site of alkylation,[9] but a relatively stable quinone methide resulting from oxidation of 2,6-di-*tert*-butyl-4-methylphenol may alkylate both N3 and N^4.[10] Structural assignments for the alkylated products have not been trivial and often relied solely on qualitative data such as ultraviolet absorption and NMR chemical shifts. To our knowledge, this report is the first to present NMR data providing direct connectivity for a quinone methide adduct of cytosine.

Results and Discussion

A major goal of our laboratory has been the development of quinone methide precursors that are capable of inducible and selective alkylation of DNA under alternative control of irradiation, reduction, fluoride or target binding.[11] Once the sites of modification are identified, a second generation of compounds may be constructed to optimize interstrand cross-linking of DNA.[3] Characterization began with the cytosine derivative, deoxycytidine (dC), since its reactivity was least predictable (Scheme 1). *O*-(*tert*-Butyldimethylsilyl)-2-(bromomethyl)-phenol was prepared by a literature procedure[11e] and combined (1.2 equiv) with dC in DMF. Reaction was initiated by addition of aqueous KF (1.5 equiv). After incubation at 20 °C for 24 h, the alkylation product of dC (**1**) was isolated in a 74% yield. When this reaction was repeated under a higher concentration of water, the product remained constant; only its yield decreased. In the absence of fluoride, no products were formed indicating that reaction in the presence of fluoride proceeded via quinone methide generation rather than direct displacement.[12] Similarly, a related methoxy (vs silyloxy) derivative was inert to DNA alkylation.[13]

Initial examination of **1** suggested that N3 rather than N^4 or O^2 had conjugated to the quinone methide. This product was stable under ambient conditions and was not subject to the rapid depyrimidization associated with O^2-alkylation of cytosine derivatives.[14] Also, alkylation resulted in an 8 nm bathochromic shift in the absorbance of the nucleotide (from 270 to 278 nm) in analogy to an equiv 10 nm shift detected upon alkylation of the N3 position by ethyl iodide.[15] In contrast, ethylation at N^4

† University of Maryland.
‡ State University of New York at Stony Brook.

(1) Singer, B.; Kuśmierek, J. T. *Annu. Rev. Biochem.* **1992**, *52*, 655.
(2) (a) Warpehoski, M. A.; Hurley, L. H. *Chem. Res. Toxicol.* **1988**, *1*, 315. (b) Basu, A. K.; Essigmann, J. M. *Chem. Res. Toxicol.* **1988**, *1*, 1. (c) Neidle, S.; Waring, M. *Molecular Aspects of Anticancer Drug-DNA Interactions*; CRC Press: Boca Raton, FL, 1993; Vol. 1.
(3) Tabone, J. C.; Stamm, M. R.; Gamper, H. B.; Meyer, R. B. *Biochemistry* **1994**, *33*, 375.
(4) Singer, B.; Grunberger, D. *Molecular Biology of Mutagens and Carcinogens*; Plenum Press: New York, 1983; Chapter 4.
(5) (a) Hurley, L. H.; Lee, C.-S.; McGovren, J. P.; Warpehoski, M. A.; Mitchell, M. A.; Kelly, R. C.; Aristoff, P. A. *Biochemistry* **1988**, *27*, 3886. (b) Lin, C. H.; Beale, J. M.; Hurley, L. H. *Biochemistry* **1991**, *30*, 3597. (c) Boger, D. L.; Munk, S. A.; Zarrinmayeh, H. *J. Am. Chem. Soc.* **1991**, *113*, 3980. (d) Warpehoski, M. A.; Harper, D. E. *J. Am. Chem. Soc.* **1995**, *117*, 2951.
(6) (a) Tomasz, M.; Chowdary, D.; Lipman, R.; Shimotakahara, S.; Veiro, D.; Walker, V.; Verdine, G. L. *Proc. Natl. Acad. Sci. U.S.A.* **1986**, *83*, 6702. (b) Egholm, M.; Koch, T. H. *J. Am. Chem. Soc.* **1989**, *111*, 8291.
(7) Angle, S. R.; Yang, W. *J. Org. Chem.* **1992**, *57*, 1092.
(8) (a) Dipple, A.; Brookes, P.; Mackintosh, D. S.; Rayman, M. P. *Biochemistry* **1971**, *10*, 4323. (b) Shapiro, R.; Shiuey, S.-J. *J. Org. Chem.* **1976**, *41*, 1597.
(9) Niwa, H.; Ogawa, T.; Okamoto, O.; Yamada, K. *Tetrahedron Lett.* **1991**, *32*, 927.
(10) Lewis, M. A.; Yoerg, D. G.; Bolton, J. L.; Thompson, J. A. *Chem. Res. Toxicol.* **1996**, *9*, 1368.
(11) (a) Chatterjee, M.; Rokita, S. E. *J. Am. Chem. Soc.* **1990**, *112*, 6397. (b) Chatterjee, M.; Rokita, S. E. *J. Am. Chem. Soc.* **1994**, *116*, 1690. (c) Kang, H.; Rokita, S. E. *Nucleic Acids Res.* **1996**, *24*, 3896. (d) Chatterjee, M.; Rokita, S. E. *J. Am. Chem. Soc.* **1991**, *113*, 5116. (e) Li, T.; Rokita, S. E. *J. Am. Chem. Soc.* **1991**, *113*, 7771. (f) Li, T.; Zeng, Q.; Rokita, S. E. *Bioconjugate Chem.* **1994**, *5*, 497.
(12) (a) Wakselman, M. *New J. Chem.* **1983**, *7*, 439. (b) Marino, J. P.; Dax, S. L. *J. Org. Chem.* **1984**, *49*, 3671.
(13) Zeng, Q.; Rokita, S. E. *J. Org. Chem.* **1996**, *61*, 9080.
(14) Singer, B.; Kröger, M.; Carrano, M. *Biochemistry* **1978**, *17*, 1246.
(15) Sun, L.; Singer, B. *Biochemistry* **1974**, *13*, 1095.

S0022-3263(97)00033-9 CCC: $14.00

Scheme 1

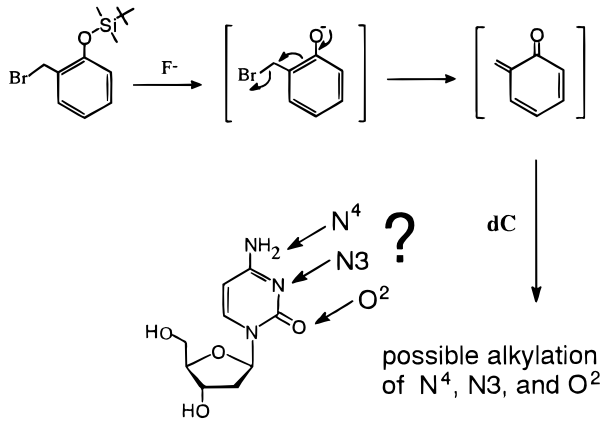

Table 1. ^{1}H and ^{13}C NMR Data (ppm) of 1 in DMSO-d_6 at 500 MHz[a]

position	δ_H	δ_C
2		149.3
4		158.9
5	6.07 d (7.9)	97.8
6	7.79 d (7.9)	137.7
$-CH_2-$	4.96 s	42.6
1′		156.3
2′		121.9
3′	7.20 d (7.5)	130.2
4′	6.74 t (7.5)	119.4
5′	7.12 t (7.4)	129.7
6′	6.77 d (8.0)	116.9

[a] Coupling constants (Hz) are provided in parentheses.

did not perturb the absorbance spectrum.[15] The N-substitution was further distinguished by a variety of NMR techniques. First, the ^{1}H and ^{13}C signals of **1** were assigned by HMQC[16] and DEPT[17] experiments using DMSO-d_6 as solvent (Table 1).[18] Although the exchangeable protons were not observed under these conditions, this solvent provided optimum solubility for the nucleoside derivative. The ^{13}C signals of the modified nucleobase are within 3.5 ppm of their counterparts in a derivative previously assigned as *N*3-methyl-dC.[19] Alkylation by either CH_3I or the quinone methide induced an upfield shift of 4−8 ppm for C2, C4, and C6 and a downfield shift of 2−5 ppm for C5.

While chemical shifts may suggest structure, direct connectivities are required for unambiguous assignment of structure. The identity of **1** was definitively established by long range ^{1}H−^{13}C interactions detected through a HMBC protocol.[20] The benzylic protons correlate with both the C2 and C4 of dC and the aromatic ring (C1′, C2′, and C3′) (Figure 1). This connectivity would only be satisfied by alkylation of N3 as illustrated. In contrast, modification at N^4 or O^2 would have generated single connectivities with the benzylic protons and either C4 or C2 of dC. The presence of the imine proton (δ 9.07) was confirmed after the hydroxyl and phenol oxygens of **1** were modified with *tert*-butyldimethylsilyl groups, and the resulting compound **2** was examined by NMR in a nonprotic solvent ($CDCl_3$).[18] In agreement with the the assignments above, a strong H5-imine proton NOE was

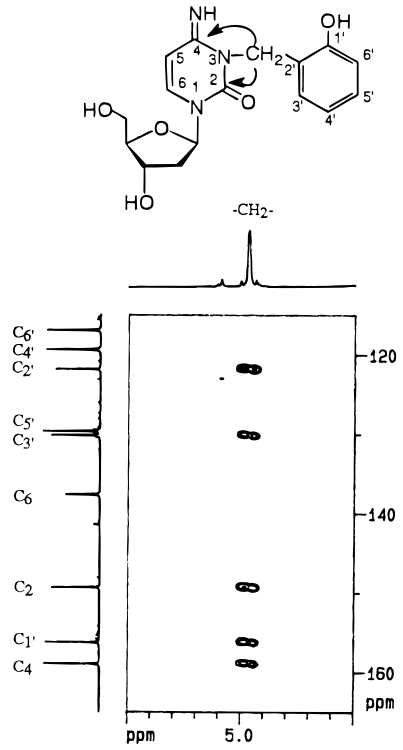

Figure 1. HMBC of **1** in DMSO-d_6 at 500 MHz.

detected and no cross-peaks were associated with the benzylic protons by a standard 2D-COSY experiment in $CDCl_3$.

The intrinsic selectivity of a simple quinone methide therefore modifies the cyclic nitrogen N3. This product is consistent with that suggested to form in the presence of another quinone methide-like precursor, dehydromonocrotaline,[9] but distinct from the *exo*-amine derivatives formed between purines and quinone methides.[6,7] Reaction of cytosine also contrasts with that of the purines since a single nucleophile of dC (N3) participates in both displacement and addition reactions. Although N3 is involved in the Watson−Crick hydrogen bonding that forms in duplex DNA, it remains a target of various alkylating agents in vivo[4] as well as a site for interstrand cross-linking by (haloethyl)nitrosourea dervatives.[21] Accordingly, this site may now be considered in the design of new cross-linking agents based on inducible formation of reactive quinone methide intermediates.

Experimental Section

General. ^{1}H NMR spectra were recorded at 250, 400, and 500 MHz, and chemical shifts are reported relative to the trace proton signals observed in the deuterated solvent. Coupling constants, *J*, are reported in hertz (Hz) and refer to apparent peak multiplicities and not true coupling constants. Decoupled ^{13}C NMR spectra were recorded at 100 MHz, and chemical shifts are reported relative to the ^{13}C signals of the solvent. Melting points are uncorrected. High-resolution mass spectra were obtained from the UC Riverside Mass Spectrometry Facility. Low-resolution mass spectra were obtained from the Mass Spectrometry Facility at Stony Brook. *O*-(*tert*-Butyldimethylsilyl)-2-(bromomethyl)phenol was prepared according to a literature procedure.[11e] All solvents and reagents were obtained from common suppliers and used without further purification.

(16) Bax, A.; Subramanian, S. *J. Magn. Reson.* **1986**, *67*, 565.
(17) Doddrell, D. M.; Pegg, D. T.; Bendall, M. R. *J. Magn. Reson.* **1982**, *48*, 323.
(18) Complete spectra are included in the Supporting Information.
(19) Box, H. C.; Lilga, K. T.; French, J. B.; Potiendo, G.; Alderfer, J. L. *J. Carbohydr. Nucleosides Nucleotides* **1981**, *8*, 189.
(20) Bax, A.; Summers, M. F. *J. Am. Chem. Soc.* **1986**, *108*, 2093.
(21) (a) MacFarland, J. G.; Kirk, M. C.; Ludlum, D. B. *Biochem. Pharmacol.* **1990**, *39*, 33. (b) Tong, W. P.; Kirk, M. C.; Ludlum, D. B. *Biochem. Pharmacol.* **1990**, *32*, 2011. (c) Bodell, W. J.; Pongracz, K. *Chem. Res. Toxicol.* **1993**, *6*, 434.

Aqueous solutions were made with water that had been purified by a standard filtration system to yield a resistivity of greater than 18 MΩ.

Alkylation of dC (1). Deoxycytidine (341 mg, 1.50 mmol) and *O*-(*tert*-butyldimethylsilyl)-2-(bromomethyl)phenol (542 mg, 1.800 mmol) were dissolved in 3.5 mL of DMF, and then an aqueous solution (1.5 mL) of $KF(H_2O)_2$ (212 mg, 2.25 mmol) was added to this mixture. This reaction mixture was stirred at room temperature (20 °C) for 24 h. DMF was removed under high vacuum, and the remaining residue was subjected to silica gel flash chromatography (CH_2Cl_2:MeOH = 5:1) to yield 369 mg (74%) of a white solid: mp 91−93 °C. ^{1}H NMR (DMSO-d_6): δ 2.09 (m, 1H), 2.15 (m, 1H), 3.55 (m, 2H), 3.80 (m, 1H), 4.22 (m, 1H), 4.96 (s, 2H), 6.07 (d, J = 7.9, 1H), 6.12 (t, J = 6.5, 1H), 6.74 (t, J = 7.5, 1H), 6.77 (d, J = 8.0, 1H), 7.12 (t, J = 7.4, 1H), 7.20 (d, J = 7.5, 1H), 7.79 (d, J = 7.9, 1H). ^{13}C NMR (DMSO-d_6): δ 40.0, 42.6, 61.3, 70.4, 86.2, 87.9, 97.8, 116.9, 119.4, 121.9, 129.7, 130.2, 137.7, 149.3, 156.3, 158.9. FAB HRMS (MEOH/SGLY): m/z 334.1404 (M + H$^+$). Calcd for $C_{16}H_{19}O_5N_3$ (M + H$^+$): 334.1403.

***tert*-Butyldimethylsilyl Derivatizations of 1.** *tert*-Butyldimethylsilyl chloride (866 mg, 5.74 mmol, 9 equiv) and **1** (213 mg, 0.638 mmol) were dissolved in 3 mL of anhydrous DMF. Imidazole (748 mg, 11.5 mmol, 18 equiv) was then added, and the reaction mixture was stirred at room temperature (20 °C) for 24 h. This was extracted with chloroform, concentrated under vacuum, and subjected to silica gel flash chromatography (1:6 ethyl acetate:hexane) yielding two compounds as colorless oils. One of the products exhibited only two *tert*-butyldimethyl groups by NMR, but the other product (166 mg, 38%) exhibited the three *tert*-butyldimethyl groups expected for complete protection of the hydroxyl groups. ^{1}H NMR ($CDCl_3$): δ 0.06 (s, 6H), 0.12 (s, 6H), 0.28 (s, 6H), 0.91 (s, 9H), 0.93 (s, 9H), 1.03 (s, 9H), 2.09 (m, 1H), 2.35 (m, 1H), 3.78 (m, 1H), 3.90 (m, 2H), 4.42 (m, 1H), 5.35 (s, 2H), 6.29 (t, 1H, J = 6.1), 6.41 (d, 1H, J = 8.1), 6.84 (m, 3H), 7.07 (t, 1H, J = 6.1), 7.91 (d, 1H, J = 8.1), 9.07 (s, 1H). ^{13}C NMR ($CDCl_3$): δ −5.53, −5.42, −4.90, −4.64, −4.11, 17.93, 18.31, 25.68, 25.85, 25.91, 42.11, 62.35, 71.10, 86.19, 87.89, 97.26, 118.5, 121.0, 125.3, 126.1, 127.7, 137.4, 149.5, 153.1, 158.7, 172.9. FAB HRMS ($CHCl_3$/NBA): m/z 676.4028 (M + H$^+$). Calcd for $C_{34}H_{61}O_5N_3Si_3$ (M + H$^+$): 676.3997.

Acknowledgment. This research was supported in part by the Center for Biotechnology, State University of New York at Stony Brook, in conjunction with the New York State Science and Technology Foundation. We thank the Mass Spectrometry Facility of UC Riverside for mass spectral analysis and our colleagues, Jeff Davis, Sampath Tirumala, and Yiu-Fai Lam, for help with the NMR determinations.

Supporting Information Available: HMQC, DEPT, and HMBC of **1** and ^{1}H NMR, ^{13}C NMR, and 2D-COSY of **2** (5 pages). This material is contained in libraries on microfiche, immediately follows this article in the microfilm version of the journal, and can be ordered from the ACS; see any current masthead page for ordering information.

JO9700336