Scheek, R. M., Boelens, R., Russo, N., van Boom, J. H., & Kaptein, R. (1984) *Biochemistry 23*, 1371-1376.
Singer, B., & Grunberger, D. (1983) in *Molecular Biology of Mutagens and Carcinogens*, Plenum, New York.
Sowers, L. C., Fazakerley, G. V., Kim, H., Dalton, L., & Goodman, M. F. (1986) *Biochemistry 25*, 3983-3988.
Weiss, M. A., Patel, D. J., Sauer, R. T., & Karplus, M. (1984) *Proc. Natl. Acad. Sci. U.S.A. 81*, 130-134.

Base Pair Mismatches and Carcinogen-Modified Bases in DNA: An NMR Study of G·T and G·O^{4}meT Pairing in Dodecanucleotide Duplexes†

Matthew W. Kalnik,‡ Michael Kouchakdjian,‡ Benjamin F. L. Li,§,‖ Peter F. Swann,§ and Dinshaw J. Patel*,‡

Department of Biochemistry and Molecular Biophysics, College of Physicians and Surgeons, Columbia University, New York, New York 10032, and Courtauld Institute of Biochemistry, Middlesex Hospital Medical School, London, England W1P 7PN

Received June 23, 1987; Revised Manuscript Received September 8, 1987

ABSTRACT: High-resolution two-dimensional NMR studies have been completed on the self-complementary d(C-G-C-*G*-A-G-C-T-*T*-G-C-G) duplex (designated G·T 12-mer) and the self-complementary d(C-G-C-*G*-A-G-C-T-*O^{4}meT*-G-C-G) duplex (designated G·O^{4}meT 12-mer) containing G·T and G·O^{4}meT pairs at identical positions four base pairs in from either end of the duplex. The exchangeable and nonexchangeable proton resonances have been assigned from an analysis of two-dimensional nuclear Overhauser enhancement (NOESY) spectra for the G·T 12-mer and G·O^{4}meT 12-mer duplexes in H_2O and D_2O solution. The guanosine and thymidine imino protons in the G·T mismatch resonate at 10.57 and 11.98 ppm, respectively, and exhibit a strong NOE between themselves and to imino protons of flanking base pairs in the G·T 12-mer duplex. These results are consistent with wobble pairing at the G·T mismatch site involving two imino proton–carbonyl hydrogen bonds as reported previously [Hare, D. R., Shapiro, L., & Patel, D. J. (1986) *Biochemistry 25*, 7445-7456]. In contrast, the guanosine imino proton in the G·O^{4}meT pair resonates at 8.67 ppm. The large upfield chemical shift of this proton relative to that of the imino proton resonance of G in the G·T mismatch or in G·C base pairs indicates that hydrogen bonding to O^{4}meT is either very weak or absent. This guanosine imino proton has an NOE to the OCH_3 group of O^{4}meT across the pair and NOEs to the imino protons of flanking base pairs. Taken together with data from the NMR of nonexchangeable protons, this shows that both G and O^{4}meT have anti-glycosidic torsion angles and are stacked into the duplex. Comparison of the intensity of the NOEs between the guanosine imino proton and the OCH_3 of O^{4}meT as well as other protons in its vicinity demonstrates that the OCH_3 group of O^{4}meT adopts the syn orientation with respect to N3 of the methylated thymidine. This rules out both the wobble pairing and the postulated structure in which the imino proton of G is hydrogen bonded to N3 of O^{4}meT for the G·O^{4}meT pair. We propose an alternate base pairing mode stabilized by one short hydrogen bond between the 2-amino group of guanosine and the 2-carbonyl group of O^{4}meT.

The base pairing properties of O^4-methylthymidine (O^{4}meT) are of considerable interest because of the importance of this alkylated base in the carcinogenic and mutagenic action of *N*-nitroso compounds (Swenberg et al., 1984; Singer, 1986). It is widely believed that *O*^{4}-alkylthymidine forms a base pair with guanosine in DNA because *O*^{4}-alkylthymidine is incorporated into DNA by DNA polymerase only in the presence of templates containing guanosine bases (Hall & Saffhill, 1983) and guanosine is incorporated when oligonucleotides containing O^{4}meT are used as templates for DNA synthesis (Singer et al., 1986). Further, the presence of O^{4}meT residues in the DNA of ϕX-174 phage results in A·T → G·C transitions at the original alkylation site (Preston et al., 1986). Hypothetical structures for this pair have been proposed (Brennan et al., 1986), but there is no structural information at the molecular level on the pairing properties of the alkylated base. This has led us to systematically study the NMR properties of the G·O^{4}meT pair in oligonucleotide duplexes and compare the spectral parameters with those for the G·T mismatch pair in otherwise identical duplexes. This paper focuses on a comparative study of self-complementary dodecanucleotides of the general sequence d(C-G-C-*X*-A-G-C-T-*Y*-G-C-G), which after annealing form a duplex containing internal G·T mismatches (X = G and Y = T) and G·O^{4}meT base pairs (X = G and Y = O^{4}meT).

The G·T mismatch, first studied with poly d(G-T) by NMR (Early et al., 1978), has been characterized in detail by one-dimensional NMR (Patel et al., 1982) and by two-dimensional NMR–distance geometry studies on solutions of oligodeoxydodecanucleotides (Hare et al., 1986), as well as by theoretical computation (Chuprina & Poltev, 1983; Keepers et al., 1984) and single-crystal X-ray diffraction (Kneale et al., 1985).

† The research was supported from Start-Up Funds provided by Columbia University, by NIH Grant CA-46533 to D.J.P., and by grants from the Cancer Research Campaign and the Medical Research Council to P.F.S. The NMR spectrometers were purchased from funds donated from the Robert Woods Johnson Jr. Trust and Matheson Trust toward setting up an NMR Center in the Basic Medical Sciences at Columbia University.

‡ Columbia University.
§ Middlesex Hospital Medical School.
‖ Present address: Institute of Molecular and Cellular Biology, National University of Singapore, Kent Ridge, Singapore 0511.

0006-2960/88/0427-0108$01.50/0

These studies unequivocally confirm the wobble G·T structure originally proposed by Crick (1966) in which hydrogen bonds form between O6 of G and N3 of T and between N1 of G and O2 of T. The present studies on the G·T mismatch are consistent with this conclusion, but the comparison between the G·T and G·O⁴meT 12-mer duplexes demonstrates that G·O⁴meT does not form an analogous pair nor does it form the structure in which the imino proton of G is hydrogen bonded to N3 of O⁴meT. Instead, we propose a structure in which the only short hydrogen bond is between the 2-NH_2 of G and O2 of T.

Experimental Procedures

The numbering system for the d(C-G-C-*G*-A-G-C-T-*T*-G-C-G) duplex (designated G·T 12-mer, I) and for the d(C-G-C-*G*-A-G-C-T-*O⁴meT*-G-C-G) duplex (designated G·O⁴meT 12-mer, II) is

C1-G2-C3-G4-A5-G6-C7-T8-T9-G10-C11-G12
G12-C11-G10-T9-T8-C7-G6-A5-G4-C3-G2-C1

I

C1-G2-C3-G4-A5-G6-C7-T8-T9-G10-C11-G12
G12-C11-G10-T9-T8-C7-G6-A5-G4-C3-G2-C1

II

The NMR studies were undertaken in 1.1 M NaCl to shift the equilibrium between the fully paired duplex and a minor form (probably a hairpin) toward the duplex state.

Sample Preparation. The G·T 12-mer and the G·O⁴meT 12-mer were synthesized in milligram quantities by the solution-phase phosphotriester approach. The synthesis, purification, and characterization of the G·O⁴meT 12-mer duplex has been reported recently (Li et al., 1987). The NMR spectra are recorded in aqueous-solution buffer containing 1.1 M NaCl, 10 mM phosphate, and 1 mM ethylenediaminetetraacetic acid (EDTA) (henceforth designated high-salt buffer). The G·T 12-mer and G·O⁴meT 12-mer samples were prepared to a concentration of 400 A_{260} units in 0.4 mL. The pH values quoted in D_2O solution are uncorrected pH meter readings.

NMR Experiments. The procedures for two-dimensional NOESY data collection in H_2O and D_2O solution and data processing were outlined in the preceding paper (Kalnik et al., 1988).

Results

We shall first present the NMR data for the control G·T 12-mer duplex (I) and then the corresponding NMR data for the G·O⁴meT 12-mer duplex (II).

G·T 12-mer

Exchangeable Proton Spectra. The exchangeable proton spectra of the G·T 12-mer duplex (7.0–15.0 ppm) in high-salt H_2O buffer, pH 6.20, recorded at 15, 40, and 55 °C are plotted in Figure 1. The exchangeable imino protons of the Watson–Crick A·T and G·C base pairs resonate between 12.6 to 14.0 ppm while the imino protons of the G·T pair resonate between 10.3 and 11.9 ppm.

The imino protons have been assigned by recording a two-dimensional magnitude nuclear Overhauser enhancement (NOESY) spectrum (mixing time 120 ms) of the G·T 12-mer duplex in high-salt H_2O buffer, pH 6.20, at 10 °C. A symmetrical contour plot covering the spectral range 14.5–6.0 ppm is plotted in Figure 2. The one-dimensional spectrum is on the diagonal while the off-diagonal cross-peaks monitor dipolar NOE interactions between pairs of adjacent protons. The magnitude of each NOE cross-peak is inversely proportional to the sixth power of the interproton distance and provides a useful indication of the distance between adjacent protons.

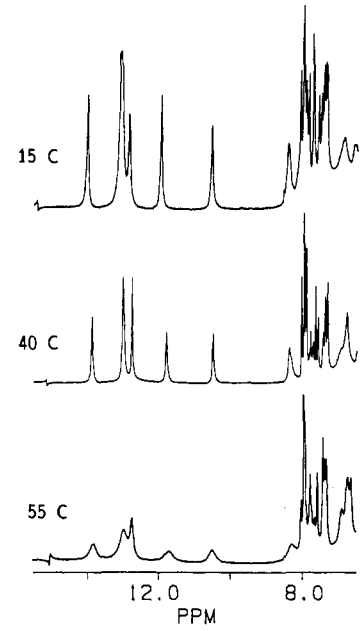


FIGURE 1: The 500-MHz proton NMR spectra (6–15 ppm) of the G·T 12-mer duplex in 1.1 M NaCl, 10 mM phosphate, and H_2O, pH 6.20, at 15, 40, and 55 °C.

We observe a strong NOE (cross-peak A, Figure 2) between the 11.98 ppm thymidine imino and the 10.57 ppm guanosine imino protons in the G·T pair, which establishes their close proximity at the mismatch site. Further, the imino proton of G4 in the G4·T9 pair exhibits NOEs to the imino proton of T8 (cross-peak B, Figure 2) and of G10 (cross-peak C, Figure 2) of the flanking A5·T8 and C3·G10 base pairs. Similarly, the imino proton of T9 at the wobble site also exhibits NOEs to the imino proton of T8 (cross-peak D, Figure 2) and of G10 (cross-peak E, Figure 2) of flanking base pairs. The expected NOE between the imino protons of flanking A5·T8 and G6·C7 base pairs (cross-peak F, Figure 2) is also observed.

Finally, additional NOEs are observed between the imino protons and the nonexchangeable base and amino protons within a base pair and between adjacent base pairs. Thus, NOEs are observed between the thymidine imino proton and the adenosine H2 proton within the A5·T8 Watson–Crick base pair (cross-peak G, Figure 2) and between the guanosine imino proton and the hydrogen-bonded cytidine amino proton within the G2·C11 (cross-peak J, Figure 2), C3·G10 (cross-peak K, Figure 2), and G6·C7 (cross-peak L, Figure 2) base pairs. In addition, NOEs between adjacent base pairs are observed between the H2 proton of A5 and the imino protons of G6 (cross-peak M, Figure 2), G4 (cross-peak H, Figure 2), and T9 (cross-peak I, Figure 2). The NOESY plots of the G·T 12-mer in high-salt H_2O buffer exhibit unusually well resolved cross-peaks (Figure 2), and the observed patterns rigorously established the alignment of G and T bases at the G·T mismatch site and demonstrate that the G·T mismatch pair is stacked into the duplex between the A5·T8 and C3·G10 Watson–Crick pairs.

The imino, hydrogen-bonded cytidine amino, and adenosine H2 proton assignments in the G·T 12-mer duplex are listed in Table I. We note that the imino protons of the G4·T9 wobble pair and the A5·T8 Watson–Crick pair broaden at approximately the same temperature and prior to the imino protons of the C3·G10 and G6·C7 base pairs (Figure 1).

Nonexchangeable Proton Spectra. The nonexchangeable base and sugar protons have been assigned from an analysis

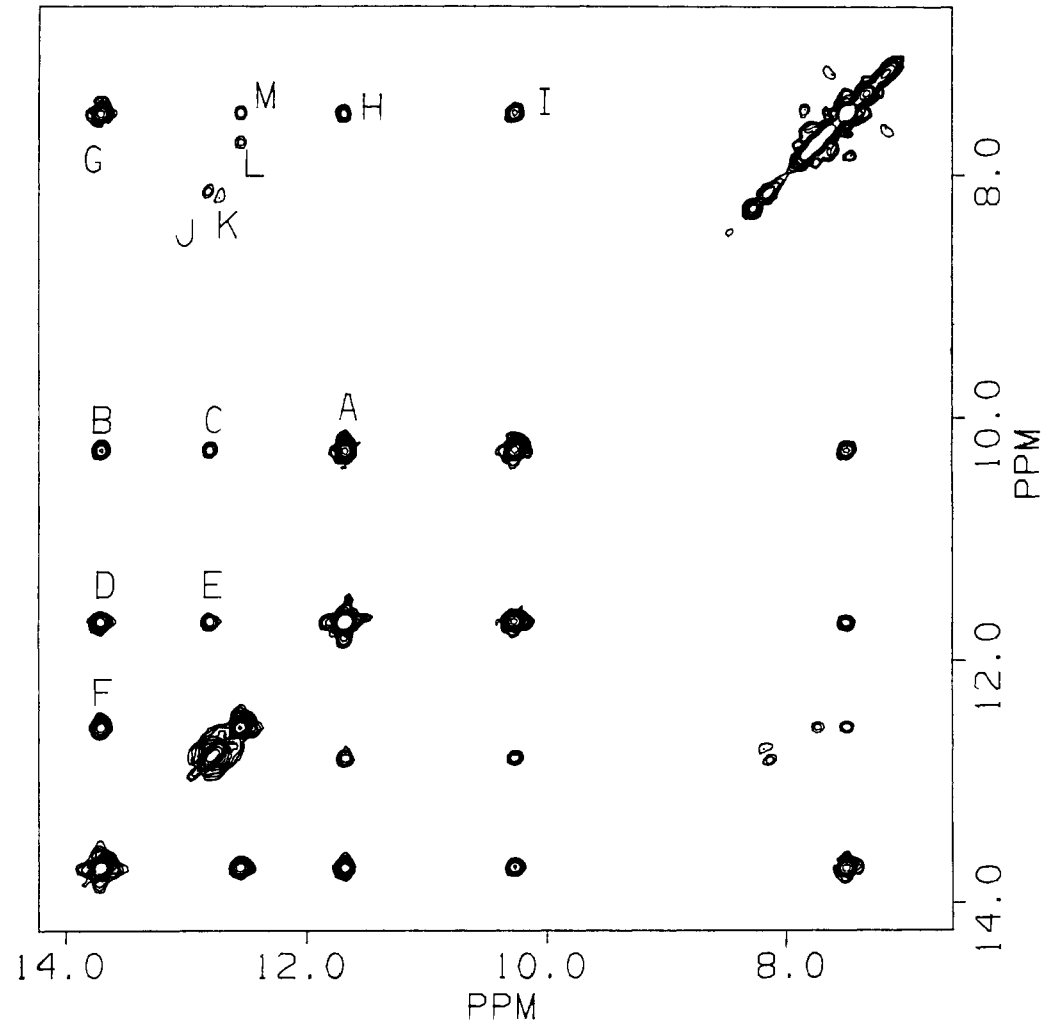


FIGURE 2: A magnitude NOESY spectrum (mixing time 120 ms) of the G·T 12-mer duplex in 1.1 M NaCl, 10 mM phosphate, and H_2O, pH 6.20, at 10 °C. The contour plot covers the symmetrical region between 6.5 and 14.0 ppm. The cross-peaks A–M are discussed in the text.

Table I: Exchangeable Proton Chemical Shifts in the G·T 12-mer Duplex at 10 °C[a]

	chemical shifts (ppm)			
base pair	T-H3	G-H1	C-H4[b]	A-H2
C1·G12		12.84		
G2·C11		13.02	8.47	
C3·G10		13.12	8.43	
G4·T9	11.98	10.57		
A5·T8	14.00			7.79
G6·C7		12.85	8.04	

[a] Conditions: 1.1 M NaCl, 10 mM phosphate, H_2O, and pH 6.20. [b] Hydrogen-bonded cytidine amino proton.

Table II: Nonexchangeable Proton Chemical Shifts in the G·T 12-mer Duplex at 25 °C[a]

	chemical shifts (ppm)							
base	H8	H2	H6	H5/CH_3	H1′	H2′	H2″	H3′
C1			7.61	5.82	5.74	1.94	2.43	
G2	8.01				5.96	2.75	2.83	5.06
C3			7.45	5.55	5.81	2.17	2.51	4.91
G4	7.86				5.71	2.63	2.86	5.07
A5	8.10	7.75			6.07	2.65	2.95	5.10
G6	7.72				5.79	2.58	2.68	5.04
C7			7.39	5.26	5.96	2.07	2.56	4.80
T8			7.46	1.62	6.18	2.14	2.65	4.94
T9			7.51	1.86	5.60	2.26	2.39	4.94
G10	8.02				5.99	2.76	2.80	5.08
C11			7.37	5.51	5.80	1.94	2.36	4.88
G12	7.94				6.15	2.46	2.69	

[a] Conditions: 1.1 M NaCl, 10 mM phosphate, D_2O, and pH 6.20.

of the phase-sensitive NOESY spectrum (250-ms mixing time) of the G·T 12-mer duplex in high-salt D_2O buffer, pH 6.20, at 25 °C and the assignments confirmed from an analysis of the through-bond two-dimensional correlated COSY spectrum of the G·T 12-mer duplex. The chemical shift values are listed in Table II.

Contour plots relating the base protons (7.2–8.2 ppm) with the sugar H1′ and cytidine H5 protons (5.2–6.2 ppm) are plotted in Figure 3A and with the thymidine CH_3 protons (1.5–2.1 ppm) are plotted in Figure 3B. We can readily trace the base protons (purine H8 or pyrimidine H6) with their own and 5′-linked sugar H1′ protons characteristic of right-handed helices (Hare et al., 1983).

Further support for a right-handed helix comes from the observed NOEs between adjacent base protons in purine H8(3′–5′)pyrimidine H5 G2-C3 (cross-peak C, Figure 3A), G6-C7 (cross-peak A, Figure 3A), and G10-C11 (cross-peak B, Figure 3A) steps and in pyrimidine H6(3′–5′)pyrimidine CH_3 C7-T8 (cross-peak K, Figure 3B) and T8-T9 (cross-peak J, Figure 3B) steps. In addition, the H2 proton of A5 exhibits NOEs to the H1′ proton of flanking G6 on the same strand (cross-peak E, Figure 3A) and the H1′ proton of flanking T9 on the partner strand (cross-peak D, Figure 3A), characteristic of right-handed helices.

Phosphorus Spectra. The proton-decoupled phosphorus NMR spectrum of the G·T 12-mer duplex in high-salt buffer,

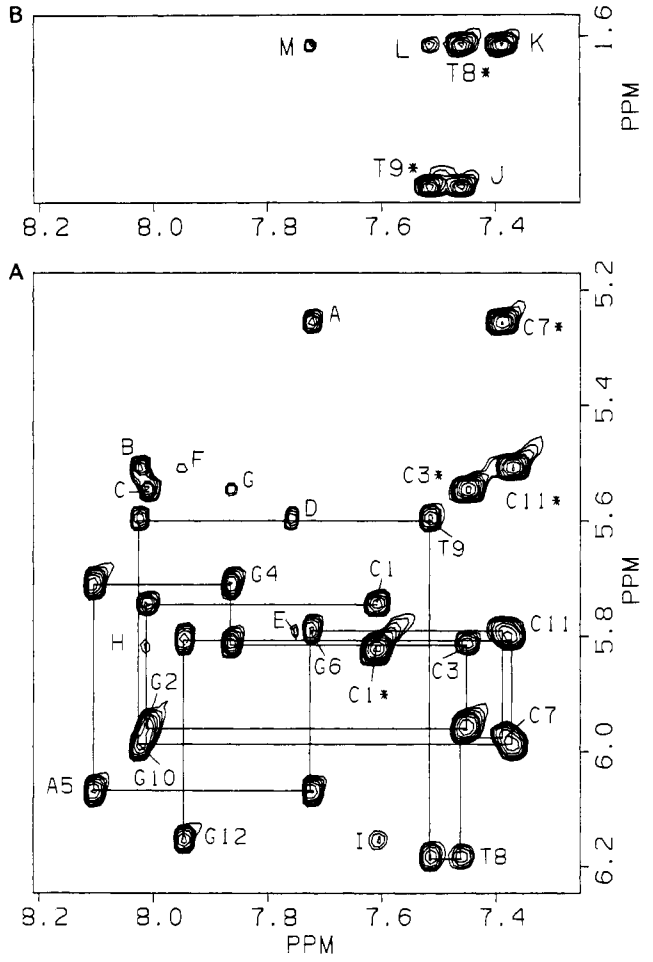


FIGURE 3: Expanded contour plots of the phase-sensitive NOESY spectrum (mixing time 250 ms) of the G·T 12-mer duplex in 1.1 M NaCl, 10 mM phosphate, and D_2O, pH 6.20, at 25 °C. (A) Distance connectivities between the base protons (7.2–8.2 ppm) with the sugar H1′ and cytidine H5 protons (5.2–6.2 ppm). The lines follow connectivities between adjacent base protons through their intervening sugar H1′ protons. (B) Distance connectivities between the base protons (7.2–8.2 ppm) and the thymidine CH_3 protons (1.5–1.9 ppm). The cytidine H6–H5 and thymidine H6–CH_3 cross-peaks are designated by asterisks. The cross-peaks A–M are discussed in the text.

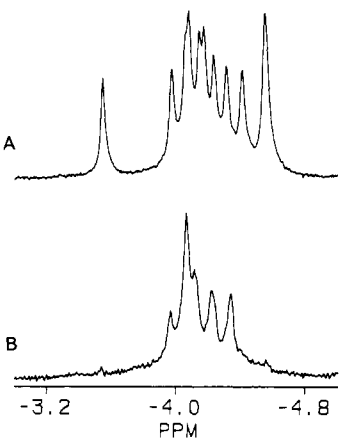


FIGURE 4: Proton Waltz-decoupled 121.5-MHz ^{31}P NMR spectra of (A) the G·T 12-mer duplex, pH 6.2, and (B) the G·O⁴meT 12-mer duplex, pH 6.35. The buffer was 1.1 M NaCl, 10 mM phosphate, and 1 mM EDTA, and data were collected at 25 °C.

pH 6.2, at 28 °C is plotted in Figure 4A. We detect phosphorus resonances outside the unperturbed 3.9–4.4 ppm range in which the phosphorus atoms of normal B-DNA resonate with one resonance downfield at 3.5 ppm and two superpositioned resonances at 4.55 ppm.

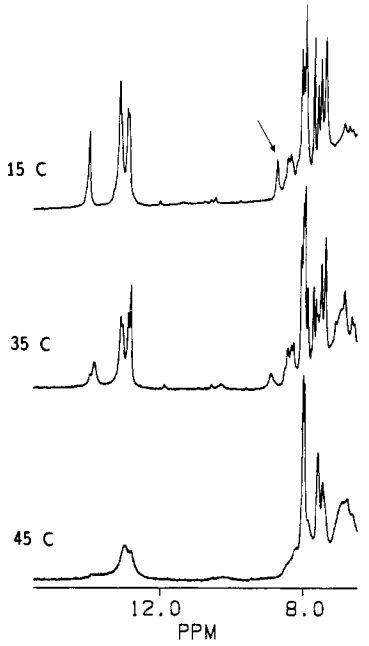


FIGURE 5: The 500-MHz proton NMR spectra (6–15 ppm) of the G·O⁴meT 12-mer duplex in 1.1 M NaCl, 10 mM phosphate, and H_2O, pH 6.05, at 15, 35, and 45 °C.

G·O⁴meT 12-mer Duplex

Exchangeable Proton Spectra. One-dimensional exchangeable proton spectra (6.5–15.0 ppm) of the G·O⁴meT 12-mer duplex in high-salt H_2O buffer, pH 6.05, recorded at 15, 35, and 45 °C are plotted in Figure 5. In addition to the normal exchangeable imino proton resonances between 12.7 and 14.0 ppm, one exchangeable imino proton resonance (designated by an arrow) was detected at 8.67 ppm, just downfield of the nonexchangeable base and hydrogen-bonded cytidine amino proton region (7.0–8.5 ppm).

A two-dimensional magnitude NOESY spectrum (mixing time 120 ms) of the G·O⁴meT 12-mer duplex in high-salt H_2O buffer, pH 6.05, was recorded at 10 °C. The symmetrical imino proton region (12.4–14.2 ppm) is plotted in Figure 6A, and the connectivities between the imino protons (12.4–14.2 ppm) and the base and amino protons (7.4–8.8 ppm) are plotted in Figure 6B. We observe the expected thymidine imino to adenosine H2 NOEs between the Watson–Crick A·T base pair (cross-peak C, Figure 6B) and between the guanosine imino and hydrogen-bonded cytidine amino protons for the three nonterminal Watson–Crick G·C base pairs (cross-peaks D–F, Figure 6B).

The imino proton of T8 in the A5·T8 base pair exhibits NOEs to the imino proton of G6 (cross-peak A, Figure 6A) and to the hydrogen-bonded amino proton of C7 (cross-peak H, Figure 6B) of the adjacent G6·C7 base pair in one direction and to the imino proton of G4 (cross-peak I, Figure 6B) of the G4·O⁴meT9 pair in the opposite direction. The imino proton of G10 in the C3·G10 base pair shows NOEs to the imino proton of G2 (cross-peak B, Figure 6A) and the hydrogen-bonded amino proton of C11 (cross-peak J, Figure 6B) of the adjacent G2-C11 base pair.

The observed NOEs to the 8.67 ppm imino proton resonance of G4 in the G4·O⁴meT9 pair are very weak in the NOESY plot partly due to the line width of this resonance and the truncation associated with apodization of magnitude two-dimensional data sets with an unshifted sine bell function. Therefore, we have also recorded one-dimensional NOE ex-

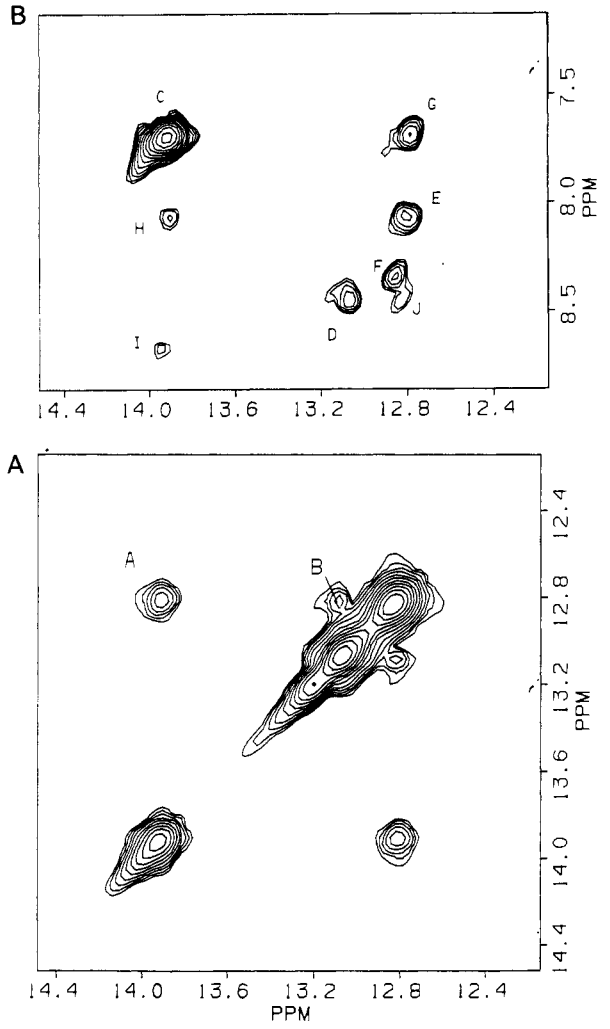


FIGURE 6: Expanded contour plots of the magnitude NOESY spectrum (mixing time 120 ms) of the G·O⁴meT 12-mer duplex in 1.1 M NaCl, 10 mM phosphate, and H_2O, pH 6.05, at 10 °C. (A) Distance connectivities between imino protons in the symmetrical 12.2–14.2 ppm spectral range. (B) Distance connectivities between the imino protons (12.2–14.2 ppm) and the base and amino protons (7.5–8.8 ppm). The cross-peaks A–J are discussed in the text.

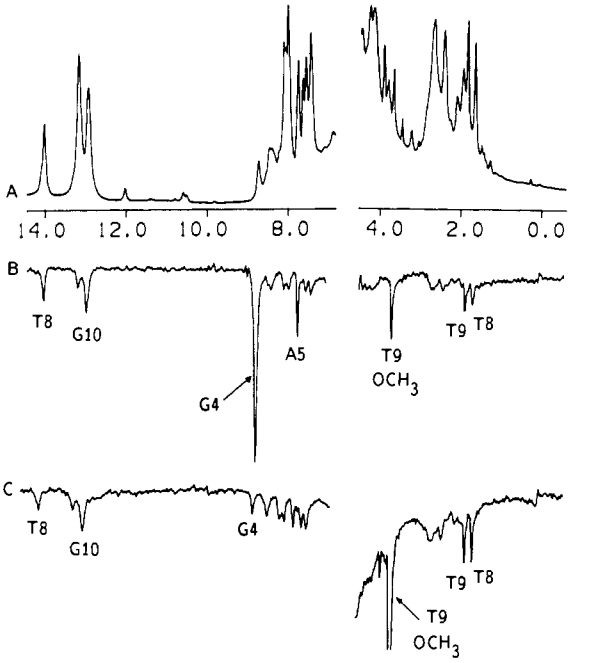


FIGURE 7: (A) The 500-MHz proton NMR spectrum (1–15 ppm) of the G·O⁴meT 12-mer duplex in 1.1 M NaCl, 10 mM phosphate, and H_2O, pH 6.05, at 10 °C. Difference spectra following 0.5-s saturation of (B) the 8.67 ppm imino proton of G4 and (C) the 3.63 ppm OCH_3 proton of O⁴meT9. The saturated resonance is designated by an arrow.

Table III: Exchangeable Proton Chemical Shifts in the G·O⁴meT 12-mer Duplex at 10 °C[a]

base pair	chemical shifts (ppm)			
	T-H3	G-H1	C-H4[b]	A-H2
C1·G12		13.18		
G2·C11		13.06	8.44	
C3·G10		12.84	8.35	
G4·O⁴meT9		8.67		
A5·T8	13.91			7.68
G6·C7		12.79	8.07	

[a] Conditions: 1.1 M NaCl, 10 mM phosphate, H_2O, and pH 6.05. [b] Hydrogen-bonded cytidine amino protons.

periments in the G·O⁴meT 12-mer duplex in 1.1 M NaCl buffer, pH 6.05, at 10 °C. The difference spectrum following partial saturation of the 8.67 ppm imino proton of G4 is plotted in Figure 7B, and the difference spectrum following partial saturation of the 3.85 ppm OCH_3 protons of O⁴meT9 is plotted in Figure 7C. The imino proton of G4 in the G4·O⁴meT9 pair shows NOEs to the H2 proton of A5 and to the imino proton of T8 of the flanking A5·T8 base pair in one direction and to the imino proton of G10 of the flanking C3·G10 base pair in the other direction (Figure 7B). Further, the imino proton of G4 shows an NOE to the OCH_3 of T9 across the G4·O⁴meT9 pair along with weaker NOEs of comparable magnitude to the CH_3 group of T9 within the same base pair and to the CH_3 group of T8 on the adjacent base pair (Figure 7B).

The OCH_3 protons of O⁴meT9 show NOEs to the imino protons of T8 and G10 of the flanking A5·T8 and C3·G10 base pairs (Figure 7C). Further, the NOEs from the OCH_3 of T9 to the CH_3 groups of T8 and T9 are of similar magnitude (Figure 7C). We also detect a large number of NOEs of comparable magnitude in the 7.5–9 ppm region one of which can be assigned to the imino proton of G4 within the G4·O⁴meT9 pair (Figure 7C). Further, weaker one-dimensional NOEs are detected in the difference spectra recorded in Figure 7B,C, which reflect the contributions of spin diffusion.

These results demonstrate that the G4·O⁴meT9 base pair is stacked between adjacent A5·T8 and C3·G10 base pairs and further establish the relative orientations of the OCH_3 group of T9 relative to the imino proton of G4 and the CH_3 group of T9 in the G4·O⁴meT9 pair.

The imino, hydrogen-bonded cytidine amino, and adenosine H2 proton chemical shifts for the G·O⁴meT 12-mer duplex at 10 °C are listed in Table III. As the temperature is increased we note that the imino proton of G4 in the G4·O⁴meT9 pair and the imino proton of T8 in the A5·T8 base pair broaden before the imino protons of the G2·C11, C3·G10, and G6·C7 base pairs in the G·O⁴meT 12-mer duplex (Figure 5).

Nonexchangeable Protons. We have assigned the nonexchangeable base and sugar protons following an analysis of the through-space NOESY and through-bond COSY connectivities in the G·O⁴meT 12-mer duplex in high-salt D_2O buffer, pH 6.15, at 25 °C (Table IV).

Expanded contour plots of the phase-sensitive NOESY spectrum (mixing time 250 ms) connecting the base protons (7.2–8.2 ppm) with the sugar H1′ and cytidine H5 protons (5.2–6.2 ppm) and with the CH_3 protons (1.6–2.0 ppm) are plotted in panels A and B of Figure 8, respectively. The duplex is a right-handed helix since we can readily trace the con-

Table IV: Nonexchangeable Proton Chemical Shifts of the G·O⁴meT 12-mer Duplex at 25 °C[a]

	chemical shifts (ppm)							
base	H8	H2	H6	H5/CH_3	H1′	H2′	H2″	H3′
C1			7.61	5.82	5.74	1.95	2.43	
G2	8.00				5.95	2.71	2.80	5.04
C3			7.38	5.46	5.80	2.04	2.39	4.86
G4	7.90				5.59	2.61	2.71	5.01
A5	8.06	7.71			6.06	2.69	2.87	5.08
G6	7.71				5.82	2.60	2.68	5.02
C7			7.40	5.26	5.98	2.09	2.53	4.87
T8			7.50	1.66	6.17	2.27	2.60	4.94
O⁴meT9			7.51	1.81	5.94	1.88	2.38	4.88
G10	7.95				5.85	2.70	2.71	4.86
C11			7.39	5.47	5.85	1.95	2.38	4.86
G12	7.95				6.16	2.46	2.69	

[a] Conditions: 1.1 M NaCl, 10 mM phosphate, D_2O, and pH 6.15.

nectivities from C1 to G12 without disruption through the G4·O⁴meT9 pair using base to sugar H1′ NOEs. In addition, detection of the expected cross-peaks between base protons on adjacent base pairs in the dinucleotide steps G2-C3 (cross-peak C, Figure 8A), G6-C7 (cross-peak A, Figure 8A), C7-T8 (cross-peak E, Figure 8B), and G10-C11 (cross-peak B, Figure 8A) and between the H2 proton of A5 and the H1′ proton of adjacent O⁴meT9 on partner strands (cross-peak D, Figure 8A) is consistent with formation of right-handed helices.

A one-dimensional slice taken through the OCH_3 protons of O⁴meT9 exhibits a stronger NOE to the CH_3 protons of T8 compared to the CH_3 protons of O⁴meT9, thus establishing the orientation of the OCH_3 group.

Phosphorus Spectra. The proton-decoupled phosphorus spectrum of the G·O⁴meT 12-mer duplex in high-salt D_2O buffer, pH 6.35, and 28 °C is plotted in Figure 4B. The resonances are dispersed between 3.9 and 4.4 ppm, which is the normal range of phosphorus resonances in unperturbed DNA. No phosphorus resonances are detected outside this unperturbed spectral range.

Discussion

The present two-dimensional NMR study has focused on the pairing properties of the G·O⁴meT lesion located four base pairs in from either end of the self-complementary duplex (II), and these results are compared with the corresponding G·T mismatch located at the same position in the sequence (I). The current studies on the G·T mismatch are in agreement with previous one-dimensional (Patel et al., 1982) and two-dimensional (Hare et al., 1986) NMR studies, which established wobble G·T base pair formation when the mismatch was located three base pairs in from either end of the self-complementary dodecanucleotide duplexes in aqueous solution.

We have not detected a change in conformation for the G·T 12-mer duplex and the G·O⁴meT 12-mer duplex on lowering the pH from neutrality to 5.0. This contrasts with results on similar duplexes containing A·C and A·O⁴meT pairs, which show a marked pH-dependent change in conformation (Kalnik et al., 1988). However, O⁴meT slowly demethylates to T at low pH (Singer et al., 1978), and as a result, the NMR spectra of the G·O⁴meT 12-mer duplex show that a small amount of the G·T 12-mer duplex had formed after the G·O⁴meT 12-mer had been kept in solution for extended periods at pH 6.

G·T Mismatch. The structure of the G·T mismatch can be characterized from an analysis of the NOESY spectra of exchangeable (Figure 2) and nonexchangeable (Figure 3) protons of the G·T 12-mer duplex. The 10.57 ppm G4 imino proton and the 11.98 ppm T9 imino proton in the G4·T9 mismatch are well resolved from each other and from other

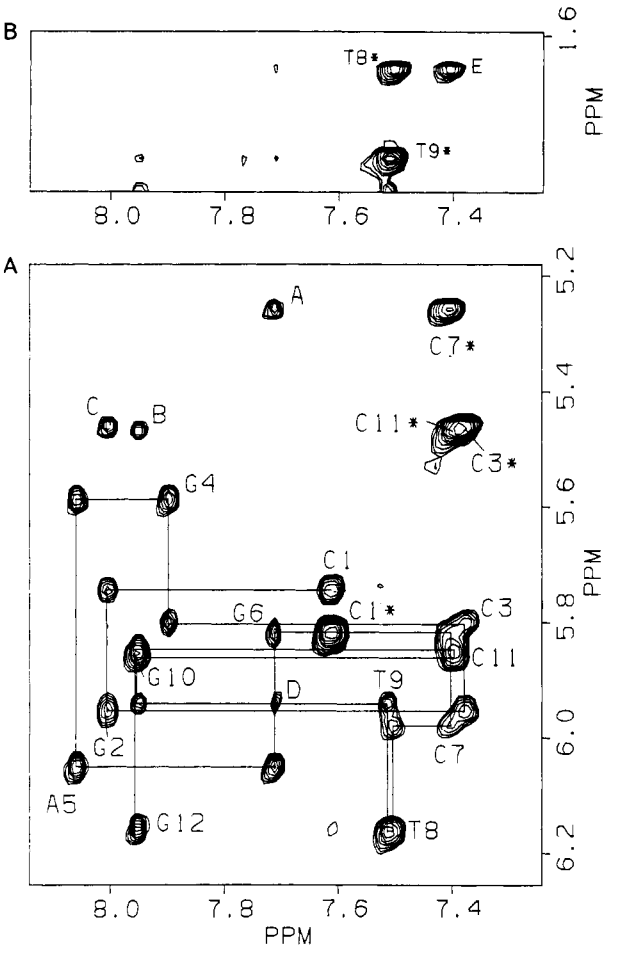
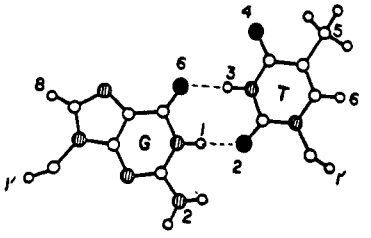


FIGURE 8: Expanded contour plots of the phase-sensitive NOESY spectrum (mixing time 250 ms) of the G·O⁴meT 12-mer duplex in 1.1 M NaCl, 10 mM phosphate, and D_2O, pH 6.15, at 25 °C. Distance connectivities between the base protons (7.2–8.2 ppm) with the sugar H1′ and cytidine H5 protons (5.2–6.2 ppm) and with the CH_3 protons (1.6–1.8 ppm) are plotted in (A) and (B), respectively.

exchangeable and nonexchangeable protons (Figure 1) such that NOEs can be detected within the G4·T9 pair and between this mismatch and flanking C3·G10 and A5·T8 Watson–Crick pairs in the G·T 12-mer duplex (Figure 2). These results establish that the G4 and T9 bases are stacked into the duplex and overlap with adjacent C3·G10 and A5·T8 base pairs. Further, the strong NOE between the imino protons of G4 and T9 (Figure 2) demonstrates wobble base pairing of the G4·T9 mismatch as shown in



III

The connectivities between the nonexchangeable base and sugar H1′ protons which is characteristic of right-handed helical DNA (Hare et al., 1983) can be readily traced through the C3-G4-A5 and T8-T9-G10 steps in the NOESY spectrum of 12-mer GT duplex (Figure 3), demonstrating that each of these segments forms part of a right-handed stacked helix centered about the mismatch site.

The structural conclusions reported in this study on G·T mismatches located four base pairs in from either end of the

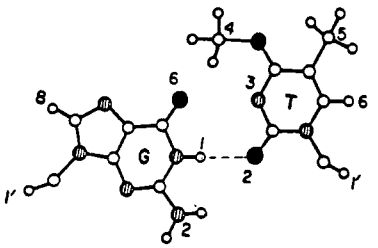
12-mer GT duplex (I) parallel earlier one-dimensional NMR (Patel et al., 1982) and two-dimensional NMR–distance geometry (Hare et al., 1986) studies on G·T mismatches located three base pairs in from either end of related self-complementary duplexes.

G·O⁴meT Lesion. The imino proton of G4 in the G·O⁴meT pair has an unusual upfield shift of 8.67 ppm (Figure 5). Such an upfield chemical shift indicates that the imino proton of G4 is not participating in a short intramolecular hydrogen bond and also that it is experiencing upfield ring current contributions through stacking with adjacent base pairs. The observed NOEs between the imino proton of G4 and the imino protons of flanking C3·G10 and A5·T8 base pairs (Figures 6 and 7) is consistent with the G4 base stacking into the helix.

It is important to determine whether the OCH_3 of O⁴meT9 is syn or anti with respect to N3 of the pyrimidine ring. If the OCH_3 is syn, and therefore projecting toward the guanosine to which the alkylated base is paired, it limits the possible orientation of the two bases. In the syn orientation of the OCH_3 group, the imino proton of G4 and the protons of the OCH_3 of O⁴meT9 are in close proximity, but in the anti orientation the OCH_3 group of O⁴meT9 is in close proximity to its own CH_3 group and directed away from the imino proton of G4. The observation of a NOE between the OCH_3 group of O⁴meT9 and the imino group of G4 and of a weak NOE between the OCH_3 and CH_3 groups of O⁴meT9 demonstrates that the OCH_3 group of O⁴meT9 is in the syn orientation with respect to the N3 of the O⁴meT.

The syn orientation of the OCH_3 and the observation that the imino proton of G4 is not strongly hydrogen bonded to the O⁴meT show that G·O⁴meT does not form the base-paired structure with one hydrogen bond between N1 of G and N3 of O⁴meT and another between the amino group of G and the O2 of O⁴meT, along with anti orientation of the OCH_3 group of O⁴meT.

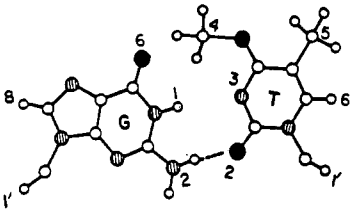
If G·O⁴meT adopted a wobble pair (IV) similar to the G·T



IV

wobble pair (III), we would have expected the imino proton of G4 to exhibit similar chemical shift in the G·O⁴meT and G·T 12-mer duplexes. However, the resonance frequencies of the imino protons of G4 in G·T 12-mer duplex (10.57 ppm) and G·O⁴meT 12-mer duplex (8.67 ppm) differ by approximately 2 ppm, therefore ruling out wobble pairing for the G·O⁴meT lesion site. Furthermore, one would expect a very weak NOE between the imino proton of G4 and the OCH_3 group of O⁴meT9 in the wobble G·O⁴meT pair, contrary to what was observed experimentally (Figure 7B).

We favor G4·O⁴meT9 pairing stabilized by one short intramolecular hydrogen bond as shown in structure V. The syn orientation of the OCH_3 group of O⁴meT9 increases the separation between O6 of G4 and O4 of O⁴meT9 in the major groove resulting in at best a very long intramolecular hydrogen bond between the imino proton of G4 and N3 ring nitrogen of O⁴meT9 (V). The proposed G·O⁴meT pairing outlined in structure V requires only small structural alterations in the helix, which would be consistent with the unperturbed tracing



V

of the C3-G4-A5 and T8-O⁴meT9-G10 segments in the expanded NOESY plots correlating the base and sugar H1′ protons (Figure 8). Finally, we do not detect strong intraresidue base to sugar H1′ NOE cross-peaks indicative of synglycosidic torsion angles at either the G4 or O⁴meT9 residue (Figure 8).

We also note that the phosphorus resonances of the G·O⁴meT 12-mer duplex resonate between 3.9 and 4.4 ppm, characteristic of an unperturbed phosphodiester backbone (Figure 4B), in contrast to phosphorus resonances in the G·T 12-mer duplex where resonances are shifted to low and high field of the normal spectral dispersion (Figure 4A). These results suggest a more pronounced structural perturbation in the backbone on incorporation of the G·T mismatch as compared with formation of the G·O⁴meT lesion.

The present studies show that the misreplication of oligodeoxynucleotides containing O⁴meT and the misincorporation of this base during DNA synthesis [reviewed by Singer (1986)] cannot be explained by formation of a base pair stabilized by two hydrogen bonds between O⁴meT and G and that some other explanation must be sought. To a large extent, the fidelity of DNA synthesis depends upon the enzymes involved (Loeb & Kunkel 1982; Loeb et al., 1986) rather than on the base pairing of the DNA. In this context the great lipophilicity of O⁴meT, much greater than of deoxyadenosine, the most lipophilic natural nucleotide, may be important because the DNA template is probably fixed in an hydrophobic cleft in the polymerase (Ollis et al., 1985).

These results have a wider implication for our conceptual approach to point mutations. Historically our understanding of DNA function has been dominated by the central role we have given to Watson–Crick base pairs, and the mutations following chemical modification of DNA bases have usually been rationalized by the postulation of abnormal base pairs. The present study raises the possibility that other structural features such as the lipophilicity of the base pair rather than solely the hydrogen-bonding character of the base are responsible for the pairings observed in vivo and in vitro for the G·O⁴meT base pair.

Registry No. G·T 12-mer, 111616-39-8; G·O⁴meT 12-mer, 106266-56-2; guanine, 73-40-5; thymine, 65-71-4; *O⁴*-methylthymine, 25902-89-0.

References

Brennan, R. G., Pyzalska, D., Blonski, W. J., Hruska, F. E., & Sundaralingam, M. (1986) *Biochemistry 25*, 1181–1185.

Chuprina, V. P., & Poltev, V. I. (1983) *Nucleic Acids Res. 11*, 5205–5222.

Crick, F. H. (1966) *J. Mol. Biol. 19*, 548–555.

Early, T. A., Olmsted, J., Kearns, D. R., & Lezius, A. G. (1978) *Nucleic Acids Res. 5*, 1955–1970.

Hall, J. A., & Saffhill, R. (1983) *Nucleic Acids Res. 11*, 4185–4193.

Hare, D. R., Wemmer, D. E., Chou, S. H., Drobny, G., & Reid, B. R. (1983) *J. Mol. Biol. 171*, 319–336.

Hare, D. R., Shapiro, L., & Patel, D. J. (1986) *Biochemistry 25*, 7445–7456.

Kalnik, M. W., Kouchakdjian, M., Li, B. F. L., Swann, P. F., & Patel, D. J. (1988) *Biochemistry* (preceding paper in this issue).
Keepers, J. W., Schmidt, P., James, T. L., & Kollman, P. A. (1984) *Biopolymers 23*, 2901–2929.
Kneale, G., Brown, T., Kennard, O., & Rabinovich, D. (1985) *J. Mol. Biol. 186*, 805–814.
Li, B. F., Resse, C. B., & Swann, P. F. (1987) *Biochemistry 26*, 1086–1093.
Loeb, L. A., & Kunkel, T. A. (1982) *Annu. Rev. Biochem. 51*, 429–457.
Loeb, L. A., Liu, P. K., & Fry, M. (1986) *Prog. Nucleic Acid Res. Mol. Biol. 33*, 57–110.
Ollis, D. L., Buch, P., Hamlin, R., Xuong, N. G., & Steitz, T. A. (1985) *Nature (London) 313*, 762–766.
Patel, D. J., Kozlowski, S. A., Marky, L. A., Rice, J. A., Broka, C., Dallas, J., Itakura, K., & Breslauer, K. J. (1982) *Biochemistry 21*, 437–444.
Preston, B. D., Singer, B., & Loeb, L. A. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 8501–8505.
Singer, B. (1986) *Cancer Res. 46*, 4879–4885.
Singer, B., Kroger, M., & Carrano, M. (1978) *Biochemistry 17*, 1246–1250.
Singer, B., Sagi, J., & Kusmierek, J. T. (1983) *Proc. Natl. Acad. Sci. U.S.A. 80*, 4884–4888.
Singer, B., Spengler, S. J., Fraenkel-Conrat, H., & Kusmierek, J. T. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 28–32.
Swenberg, J. A., Dryoff, M. C., Bedell, M. A., Popp, J. A., Huh, N., Kirstein, U., & Rajewsky, M. F. (1984) *Proc. Natl. Acad. Sci. U.S.A. 81*, 1692–1685.

Selective Reversible Deuteriation of Oligodeoxynucleotides: Simplification of Two-Dimensional Nuclear Overhauser Effect NMR Spectral Assignment of a Non-Self-Complementary Dodecamer Duplex†

Charles K. Brush,‡ Michael P. Stone,* and Thomas M. Harris*
Department of Chemistry and Center for Molecular Toxicology, Vanderbilt University, Nashville, Tennessee 37235
Received July 7, 1987; Revised Manuscript Received September 28, 1987

ABSTRACT: Oligodeoxynucleotides are reversibly deuteriated at the purine C8 and cytosine C5 positions with deuterioammonium bisulfite at pD 7.8. The exchange reaction is complete after 48 h at 65 °C. When an oligomer deuteriated under these conditions is analyzed by ^{1}H nuclear magnetic resonance (NMR) spectroscopy, the purine H8 and cytosine H5 proton signals are selectively removed from the spectrum. A non-self-complementary oligodeoxynucleotide that has been deuteriated in this manner may be annealed with its complement and the resulting heteroduplex analyzed by two-dimensional nuclear Overhauser enhancement (NOESY) spectroscopy. NOE cross-peaks arising from pyrimidine H6–deoxyribose H1′ dipolar interactions in both strands are observed, but purine H8–deoxyribose H1′ and purine H8–deoxyribose H2′,H2″ dipolar interactions are only observed for the nondeuteriated strand. The intense cytosine H5–H6 cross-peaks are also removed from the spectrum of the deuteriated strand, which further simplifies interpretation since these strong cross-peaks often interfere with less intense NOE cross-peaks arising from dipolar coupling between purine H8 or pyrimidine H6 and deoxyribose anomeric protons. The resulting spectral simplification allows unambiguous assignments to be made on NOEs that otherwise may be difficult to distinguish. The deuteriation procedure is demonstrated with the sequence d(CGTTATAATGCG)·d(CGCATTATAACG), which has previously been assigned by traditional NOESY methods [Wemmer, D. E., Chou, S.-H., Hare, D. R., & Reid, B. R. (1984) *Biochemistry 23*, 2262–2268]. Although the assignment of this dodecadeoxynucleotide may be completed without deuteriation, several NOEs must be assigned indirectly because of degeneracies in the chemical shift of the purine H8 protons. This methodology should have wide applicability to NMR spectral interpretation of oligodeoxynucleotides, particularly to oligonucleotides of 12 bases or longer.

Because of recent improvements in the chemical synthesis of oligodeoxynucleotides [Dorman et al., 1984; reviewed in the text edited by Gait (1984)], relatively large quantities of DNA fragments having defined sequences are accessible for study by physical methods. In particular, there is great interest in the use of high-field nuclear magnetic resonance spectroscopy (NMR)[1] to probe the solution structure of oligodeoxynucleotides and their interactions with various ligands, including drug and carcinogen molecules (Patel et al., 1981, 1986; Pardi et al., 1983; Chandrasekaran et al., 1984; Feigon

†This work was supported by funding from the National Institutes of Health, Grants ES-00267 and ES-03755. C.K.B. received support from a Harold Stirling Vanderbilt Scholarship and Dissertation Research Award.
* Address correspondence to this author.
‡Present address: Pharmacia P-L Biochemicals, Inc., 2202 North Bartlett Ave., Milwaukee, WI 53202.

[1] Abbreviations: NOE, nuclear Overhauser effect; HPLC, high-performance liquid chromatography; SAX, strong anion-exchange resin; UV, ultraviolet; EDTA, ethylenediaminetetraacetic acid; TPPI, time-proportional phase increments; 1D, one dimensional; 2D, two dimensional; tRNA, transfer RNA; DMSO, dimethyl sulfoxide; DSS, sodium 4,4-dimethyl-4-silapentanesulfonate. None of the oligodeoxynucleotides discussed in this paper have terminal phosphates. We abbreviate the notation for oligomers by leaving out the phosphodiester linkage. A, C, G, and T refer to mononucleotide units. A right superscript refers to position in the oligodeoxynucleotide sequence starting from the 5′-terminus of chain A and proceeding to the 3′-terminus of chain A and then from the 5′-terminus of chain B to the 3′-terminus of chain B. C2, C5, C6, C8, C1′, C2′, and C2″ represent specific carbon nuclei in nucleotides. H2, H5, H6, H8, H1′, H2′, H2″, etc., represent the protons attached to these carbons.

0006-2960/88/0427-0115$01.50/0