Supporting Information

Comparison of the structural and dynamic effects of 5-methylcytosine and 5-chlorocytosine in a CpG dinucleotide sequence

**Jacob A. Theruvathu¹, Whitney Yin^{1,2}, B. Montgomery Pettitt¹⁻³
and Lawrence C. Sowers^{1,2*}**

¹Department of Pharmacology & Toxicology, The University of Texas Medical Branch, 3.330 Basic Science Building, 301 University Blvd, Galveston, TX-77555-0617, Ph:409-772-9678, e-mail: lasowers@utmb.edu.

²Sealy Center for Structural Biology & Molecular Biophysics, The University of Texas Medical Branch, 301 University Blvd, Galveston, TX-77555

³Department of Biochemistry & Molecular Biology, The University of Texas Medical Branch, 301 University Blvd, Galveston, TX-77555

Table S1. Data collection and refinement statistics for the oligonucleotides ^mC3/C9 and ^{cl}C3/C9.

	^m C	^{cl} C
Wavelength (Å)	0.979	0.979
Resolution range (Å)	9.586 - 1.219 (1.263 - 1.219)	9.995 - 1.93 (1.998 - 1.929)
Space group	P 21 21 21	P 21 21 21
Unit cell	a=25.38 b=40.274 c=65.405 a= 90° β=90°, γ= 90°	a=24.629 b=39.918 c=65.364 a= 90° β=90°, γ= 90°
Total reflections	17420	4414
Unique reflections	17417 (798)	4411 (381)
Completeness (%)	84.42 (49.45)	85.25 (75.90)
Mean I/sigma(I)	26.29 (1.96)	13.63 (2.11)
Wilson B-factor	17.51	36.44
R-sym	0.06	0.056
R-factor	0.2492 (0.3488)	0.1957 (0.3047)
R-free	0.2782 (0.4093)	0.2629 (0.4219)
Number of atoms	874	795
Ligands	1	1
Water	104	38
Nucleoside residues	24	24
RMS(bonds)	0.013	0.010
RMS(angles)	1.93	1.98
Average B-factor	23.70	35.80
Macromolecules	22.10	35.50
Solvent	31.10	39.60

Statistics for the highest-resolution shell are shown in parentheses.

Table S2. Helicoid analyses of Dickerson-Drew dodecamers containing ^{13}C or ^{13}C at position 3 ($^{13}\text{C3/C9}$ and $^{13}\text{C3/C9}$) using the CURVES+ program (47).

	Helicoid parameters	^{13}C	^{13}C	C
BP-Axis	X-displacement bp3	0.4	0.33	0.08
	Avg	-0.02	-0.05	0.01
	Y-displacement bp3	-0.51	-0.84	-0.17
	Avg	-0.06	-0.07	0.03
	Inclination bp3	124.6	125.3	47.9
	Avg	7.2	10.3	4.7
	Tip bp3	-132.9	-154.2	-109.2
	Avg	-139.1	-136.9	-16.8
	Bend bp3	15.4	10.5	19.7
	Total bend	64.4	76.0	167.5
Intra-BP Translations	Shear bp3	13.00	14.39	14.06
	Avg	-7.23	-7.11	17.87
	Stretch bp3	18.02	17.66	-0.83
	Avg	-2.58	-2.74	-2.81
	Stagger bp3	-0.33	0.20	-0.17
	Avg	-0.74	-1.11	-0.02
	Buckle bp3	-145.2	-146.6	-145.5
	Avg	-106.1	-104.0	-126.8
	Propeller bp3	75.1	83.5	73.6
	Avg	139.6	140.6	133.
Rotation	Opening bp3	-22.2	-30.1	17.8
	Avg	-0.3	1.3	-8.7
Inter-BP Translations	Shift bp2-3	-0.32	-0.27	-0.30
	bp3-4	-0.25	-0.42	-0.24
	Avg	-0.21	-0.20	-0.14
	Slide bp2-3	-0.51	-0.53	-0.48
	Bp3-4	0.74	0.86	0.92
	Avg	0.04	0.09	0.07
	Rise bp2-3	0.17	0.20	0.15
	Bp3-4	-0.06	-0.17	-0.03
	Avg	0.11	0.14	0.14
	Tilt bp2-3	-1.9	-1.4	-2.1
Rotation	Bp3-4	169.5	166.8	169.8
	Avg	7.9	-5.4	8.0
	Roll bp2-3	1.0	-0.7	0.7
	Bp3-4	48.5	52.3	49.5
	Avg	-1.5	8.2	0.3
	Twist bp2-3	-36.4	-33.7	-36.0
	Bp3-4	14.0	18.4	9.6
	Avg	-3.0	2.6	7.2

Table S3: Chemical shifts of H6, H8, H1', CH3, H5 and H2 protons of the oligonucleotides modified at position 3 in the presence of 10% D₂O, 100 mM NaCl, 10 mM sodium phosphate and 0.2 mM EDTA at pH 7 and 300K. The chemical shifts were calibrated using 4,4-dimethyl-4-silapentane-1-sulphonate (DSS) as an internal standard reference

Position	H1'			H6/H8			CH3/H5/H2		
	C3/C9	^m C3/C9	^l C3/C9	C3/C9	^m C3/C9	^l C3/C9	C3/C9	^m C3/C9	^l C3/C9
C1	5.76	5.76	5.75	7.63	7.64	7.65	5.91	5.92	5.93
G2	5.88	5.97	5.96	7.94	7.95	7.98			
X3	5.60	5.51	5.51	7.25	7.04	7.54	5.36		
G4	5.42	5.40	5.40	7.83	7.82	7.85			
A5	5.99	5.99	5.98	8.09	8.10	8.11	7.23	7.23	7.23
A6	6.15	6.14	6.13	8.09	8.10	8.11	7.60	7.59	7.60
T7	5.91	5.90	5.88	7.09	7.09	7.10	1.26	1.26	1.26
T8	6.10	6.11	6.09	7.36	7.36	7.36	1.52	1.52	1.52
C9	5.65	5.65	5.63	7.45	7.45	7.46	5.62	5.64	5.65
G10	5.85	5.83	5.83	7.89	7.91	7.92			
C11	5.75	5.77	5.76	7.32	7.32	7.30	5.43	5.41	5.40
G12	6.16	6.14	6.13	7.93	7.93	7.93			

Table S4: Chemical shifts of H6, H8, H1', CH₃, H5 and H2 protons of the oligonucleotides modified at position 3 in the presence of 10% D₂O, 100 mM NaCl, 10 mM sodium phosphate and 0.2 mM EDTA at pH 7 and 300K. The chemical shifts were calibrated using 4,4-dimethyl-4-silapentane-1-sulphonate (DSS) as an internal standard reference.

Position	H1'			H6/H8			CH3/H5/H2		
	C3/ ^m C9	^m C3/ ^m C9	^l C3/ ^m C9	C3/ ^m C9	^m C3/ ^m C9	^l C3/ ^m C9	C3/ ^m C9	^m C3/ ^m C9	^l C3/ ^m C9
C1	5.76	5.76	5.79	7.63	7.64	7.65	5.91	5.91	5.93
G2	5.89	5.97	6.00	7.94	7.95	7.98			
X3	5.63	5.54	5.55	7.25	7.04	7.55	5.38		
G4	5.42	5.41	5.44	7.84	7.83	7.85			
A5	5.98	5.99	6.00	8.11	8.12	8.12	7.24	7.25	7.25
A6	6.16	6.16	6.17	8.11	8.12	8.12	7.59	7.58	7.58
T7	5.91	5.91	5.92	7.09	7.09	7.09	1.26	1.26	1.26
T8	6.10	6.11	6.12	7.39	7.40	7.39	1.53	1.54	1.54
^m C9	5.63	5.62	5.62	7.26	7.28	7.28			
G10	5.82	5.84	5.86	7.88	7.89	7.90			
C11	5.77	5.79	5.80	7.32	7.32	7.30	5.43	5.40	5.40
G12	6.15	6.15	6.16	7.93	7.93	7.93			

Table S5: Chemical shifts and linewidths of reference 2'-deoxynucleosides in 10%D₂O, 20 mM potassium phosphate pH 7, 278 K. The chemical shifts were calibrated using 4,4-dimethyl-4-silapentane-1-sulphonate (DSS) as an internal standard reference in D₂O.

Nucleoside	¹ pK _a	H6	Chemical shift		Linewidth (Hz)	
			H _a	H _b	H _a	H _b
^m dC	4.3	7.67	6.89	6.89	18.7	18.7
dC	4.2	7.83	6.76	7.24	33.4	23.0
^F dC	2.6	8.01	7.34	7.55	38.5	23.8
^{Cl} dC	2.7	8.13	7.36	7.24	29.2	13.1
^{Br} dC	2.7	8.23	7.41	7.06	23.5	11.7
^I dC	3.0	8.32	7.42	6.76	28.3	14.0

^IFrom reference 30.

Table S6: Chemical shifts of imino protons of the oligonucleotides studied here in the presence of 10% D₂O, 100 mM NaCl, 10 mM sodium phosphate and 0.2 mM EDTA at pH 7 and 300K. The chemical shifts were calibrated using 4,4-dimethyl-4-silapentane-1-sulphonate (DSS) as an internal standard reference.

Oligo	G2	G4	T7	T8	G10
^m C3/C9	13.04	12.65	13.61	13.73	12.95
^m C3/ ^m C9	13.04	12.71	13.62	13.73	12.94
C3/C9	13.06	12.69	13.61	13.76	12.90
C3/ ^m C9	13.05	12.72	13.62	13.74	12.86
^{Cl} C3/C9	13.06	12.68	13.61	13.75	12.54
^{Cl} C3/ ^m C9	13.05	12.73	13.61	13.75	12.52

Figure S1 : **(a)** GC/MS mass spectrum of 5-chlorocytosine (^{13}C) after acid hydrolysis and silylation of the ^{13}C 3/C9 oligonucleotide containing $[4-^{15}\text{N}]^{13}\text{C}$. The asterisk indicates the position of ^{15}N labeling. **(b)** MALDI spectrum of the $[4-^{15}\text{N}]^{13}\text{C}$ -containing ^{13}C 3/C9 oligonucleotide (solid line) with the theoretically simulated mass spectrum (dotted line). The expected monoisotopic ($\text{M}+\text{H}^+$) mass of the oligonucleotide is 3680.62 amu.

Figure S2: $1/T_m$ vs $\ln(Ct)$ plot for the ${}^mC3/C9$ oligonucleotide in the presence of 10 mM sodium phosphate, 100 mM sodium chloride, 0.1 mM EDTA at pH 7.0.

Figure S3 : ^1H NMR spectrum of the aromatic region of C3/C9 (top), $^m\text{C3/C9}$ (middle) and $^{\text{Cl}}\text{C3/C9}$ (bottom) oligonucleotides at pH 7 and 300K.

Figure S4 : 2D NOE of $[4-^{15}\text{N}]$ $^{\text{Cl}}\text{C3/C9}$ oligonucleotide at pH 7 (10% D_2O) and 300 K. The oligonucleotide was dissolved in a buffer containing 100 mM sodium chloride, 10 mM sodium phosphate and 0.2 mM EDTA.

Figure S5: Temperature-dependent ^1H NMR spectrum of the imino region of the $[4-^{15}\text{N}]$ C3/C9 oligonucleotide in the presence of 10% D_2O , 100 mM NaCl, 10 mM sodium phosphate and 0.2 mM EDTA at pH 7.

Figure S6: Magnetization transfer experiment from the nonWatson-Crick amino proton to the Watson-Crick amino proton of the $[4\text{-}^{15}\text{N}]^{\text{Cl}}\text{C3/C9}$ oligonucleotide at pH 7 and 300K. ^{15}N edited ^1H spectra were taken at various time delays after the inversion of the Watson-Crick amino proton with a selective rectangular pulse.

Figure S7: Relative intensity of W-C (▲) and non W-C (Δ) amino protons after the magnetization transfer experiments from the Watson-Crick to the nonWatson-Crick amino proton of the ^{13}C 3/C9 oligonucleotide at pH 7 and 300K. Rotational parameters are obtained from the non-linear curve fitting using Equation (1).

Figure S8: Temperature-dependent ^{15}N -edited amino proton (Watson-Crick ($\blacktriangle$) and nonWatson-Crick($\triangle$)) line widths of oligonucleotides at pH 7.