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

Jacob A. Theruvathu[1], **Whitney Yin**[1,2], **B. Montgomery Pettitt**[1,2,3], and **Lawrence C. Sowers**[1,2,*]

[1]Department of Pharmacology & Toxicology, The University of Texas Medical Branch, 3.330 Basic Science Building, 301 University Blvd, Galveston, TX-77555-0617.

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

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

Abstract

Inflammation-mediated reactive molecules can result in an array of oxidized and halogenated DNA damage products including 5-chlorocytosine (^{Cl}C). Previous studies have shown that ^{Cl}C can mimic 5-methylcytosine (^{m}C) and act as a fraudulent epigenetic signal, promoting the methylation of previously unmethylated DNA sequences. Although the 5-halouracils are good substrates for base excision repair, no repair activity has yet been identified for ^{Cl}C. Due to the apparent biochemical similarities of ^{m}C and ^{Cl}C, we have investigated the effects of ^{m}C and ^{Cl}C substitution on oligonucleotide structure and dynamics. In this study, we have constructed oligonucleotide duplexes containing C, ^{Cl}C and ^{m}C within a CpG dinucleotide. The thermal and thermodynamic stability of these duplexes are found to be experimentally indistinguishable. Crystallographic structures of duplex oligonucleotides containing ^{m}C and ^{Cl}C were determined to 1.2 and 1.9 Å, respectively. Both duplexes are B-form and are superimposable on a previously determined structure of a cytosine-containing duplex with RMSD of approximately 0.25 Å. NMR solution studies indicate that all duplexes containing cytosine or the cytosine analogs are normal B-form, and no structural perturbations are observed surrounding the site of each substitution. The

Dedication: "This manuscript is dedicated to the fond memory of Victor Fung, Ph.D., a former Program Officer at NCI and former Scientific Review Officer of the Cancer Etiology study section of CSR, NIH, for his wisdom, compassion, integrity, his love of sciences and the arts, his incredible culinary skills, and above all, his contributions to the career development of so many investigators during his own distinguished career."

Supporting Information Available: Crystallographic data collection and refinement statistics of the oligonucleotide ^{m}C3/C9 and ^{Cl}C3/C9 (Table S1); Helicoid analyses of Dickerson-Drew dodecamers containing ^{m}C or ^{Cl}C at position 3 (^{m}C3/C9 and ^{Cl}C3/C9) (Table S2); Chemical shifts of the selected non exchangeable protons of the oligonucleotides (Table S3 and Table S4); Chemical shifts and line width of amino protons of the nucleosides in 10% D$_2$O and 5 °C (Table S5); Chemical shift of imino protons of the oligonucleotides (Table S6); Characterization of the ^{Cl}C3/C9 oligonucleotide by GC-MS and MALDI ToF (Figure S1); $1/T_m$ *versus* ln(Ct) plot for the ^{m}C3/C9 oligonucleotide (Figure S2); 1D proton spectra of the aromatic region of the oligonucleotides (Figure S3); 2D NOE of ^{Cl}C3/C9 oligonucleotide with H6-H1' connectivity (Figure S4); Temperature dependent proton spectra in the imino region of ^{Cl}C3/C9 oligonucleotide (Figure S5); Magnetization transfer experiment of amino protons ^{Cl}C3/C9 oligo (Figure S6); Nonlinear curve fitting from magnetization transfer experiments of ^{Cl}C3/C9 oligo (Figure S7); Line widths of amino protons of C3/C9, ^{m}C3/C9 and ^{Cl}C3/C9 oligonucleotide as a function of temperature(Figure S8); This material is available free of charge via the internet at http://pubs.acs.org.

magnitude of the base-stacking induced upfield shifts for non-exchangeable base proton resonances are similar for each of the duplexes examined, indicating that neither ^{m}C nor ^{Cl}C significantly alter base stacking interactions. The ^{Cl}C analog is paired with G in an apparently normal geometry; however the G-imino proton of the ^{Cl}C-G base pair resonates to higher field relative to ^{m}C-G or C-G, indicating a weaker imino hydrogen bond. Using selective ^{15}N-enrichment and isotope-edited NMR, we observe that the amino group of ^{Cl}C rotates at roughly half the rate of the corresponding amino groups of the C-G or ^{m}C-G base pairs. The altered chemical shifts of hydrogen bonding proton resonances for the ^{Cl}C-G base pair, as well as the slower rotation of the ^{Cl}C amino group can be attributed to the electron-withdrawing inductive property of the 5-chloro substituent. The apparent similarity of duplexes containing ^{m}C and ^{Cl}C demonstrated here is in accord with results of previous biochemical studies and further suggests that ^{Cl}C is likely to be an unusually persistent form of DNA damage.

Keywords

5-chlorocytosine; 5-methylcytosine; amino group rotation; NMR; crystallography; duplex stability; epigenetics

The development of cancer is driven by both genetic mutations and epigenetic dysregulation (1-3). While an extensive body of work has established mechanisms by which DNA damage can induce genetic mutations, substantially less is known about the relationship between DNA damage and epigenetic changes. In higher eukaryotes, the primary epigenetic mark in DNA is 5-methylcytosine (^{m}C) which modulates the transcriptional activity of key genes primarily by altering DNA-protein interactions (4,5). In cancer cells, cytosine methylation patterns are corrupted, and frequently important tumor suppressor genes become silenced by aberrant promoter methylation (6-8). Currently, there are few clear mechanisms to explain how aberrant promoter hypermethylation occurs.

Emerging evidence indicates that inflammation-mediated reactive molecules can modify DNA, creating an additional type of endogenous DNA damage (9-11). Among the inflammation-mediated DNA adducts identified to date are the 5-halocytosines, 5-chlorocytosine (^{Cl}C) and 5-bromocytosine (^{Br}C), and both adducts have been identified in DNA from human tissues exposed to inflammation (12-17). While most types of DNA damage have been show to interfere with enzymatic DNA methylation and with DNA-binding proteins containing a methyl-binding domain, the 5-halocytosine analogs ^{Br}C and ^{Cl}C have been shown to mimic ^{m}C with respect to both DNA-binding proteins and methyltransferases (18-21). Therefore, ^{Br}C and ^{Cl}C could act as fraudulent epigenetic signals, in part explaining how inflammation-mediated DNA damage could drive aberrant promoter methylation. Studies in eukaryotic cells have established that ^{Cl}C in DNA can result in heritable changes in promoter methylation (22). While ^{Br}C paired with G can be repaired slowly by both MBD4 (18) and hTDG (23) glycosylases, no repair activities have yet been identified for ^{Cl}C, and therefore, ^{Cl}C could be a persistent form of DNA damage that would accumulate with increasing inflammation damage.

Previous studies have shown that the 5-halouracils can mimic thymine in DNA protein interactions as the sizes of the 5-Cl, Br and methyl substituents are similar (24, 25). However, as the inductive properties of the methyl substituent of thymine is electron donating and the 5-halo substituents are electron withdrawing, 5-halouracil analogs in DNA are rapidly repaired by multiple glycosylases (26, 27). We were therefore interested in understanding the impact of ^{Cl}C on oligonucleotide structure and dynamics in order to reveal how ^{Cl}C could serve as an excellent mimic for ^{m}C in several important biological interactions, as well as avoid surveillance and detection by DNA repair systems. In this

study, we have constructed oligonucleotide duplexes containing cytosine, ^{m}C and ^{Cl}C in the self-complementary Dickerson dodecamer sequence (28) within a CpG dinucleotide in one or both strands. We have studied these self-complementary duplexes in thermal melting studies, x-ray crystallography, and by ^{1}H-NMR spectroscopy in solution to determine how ^{m}C and ^{Cl}C might differentially influence duplex structure and dynamics.

Materials and methods

Synthesis and purification of oligonucleotides

Oligonucleotides used in this study were synthesized by standard phosphoramidite solid phase methods. To prepare oligonucleotides with 5-chlorocytosine, a phosphoramidite was prepared from O^{4}-ethyl, 5-chloro-2'-deoxyuridine as previously described (29). To prepare oligonucleotides containing ^{15}N-enriched 4-amino groups, phosphoramidites containing O^{4}-methyluracil (Glen research, Sterling, VA), O^{4}-methylthymine (Glen research) and O^{4}-ethyl-5-chlorouracil were coupled at the desired sequence position as shown in Figure 1. Oligonucleotides were then deprotected with ^{15}N-enriched aqueous ammonia (6 N, 98%, Cambridge Isotope Labs, Andover, MA) at room temperature for 36 h as previously describe for oligonucleotides containing 5-fluorocytosine (30).

Oligonucleotides containing a dimethoxytrityl (DMT) group were purified by reverse phase HPLC, using a PRP column (Hamilton, Reno, NV), detritylated in 80% acetic acid at room temperature for 30 min and then repurified by HPLC using a reverse phase column. Following purification, oligonucleotides were characterized by MALDI-ToF-MS (31) and the base composition of each was examined following formic acid hydrolysis and silylation by GC/MS (Supporting Information Figure S1).

Oligonucleotide UV Melting studies

The melting temperature (T_m) and thermodynamic values of the self-complementary oligonucleotides were measured as previously described (32, 33) using a Varian Bio 300 Cary UV Vis spectrophotometer (Palo Alto, CA). Oligonucleotides at selected total-strand concentrations from 2 μM to 60 μM were dissolved in a buffer containing 100 mM NaCl, 0.1 mM EDTA, and 10 mM sodium phosphate at pH 7. The T_m values are reported at 28 μM total strand concentration.

Crystallography

Dodecamer self-complimentary duplex oligos containing either ^{m}C or ^{Cl}C at position 3 ($^{m}C3/C9$ or $^{Cl}C3/C9$) were crystallized in solutions containing 50 mM HEPES, pH 7.0, 35-40-2-methyl-2,4-pentanediol (MPD), 5-10 mM spermine and 5-10 mM $MgCl_2$ at 17 °C. The crystals were directly frozen in liquid nitrogen. Diffraction data were collected at 100 Kelvin at the Advanced Photon Source, Argonne National Laboratory (Argonne, IL) using synchrotron radiation. Data were processed with the HKL3000 package (HKL Research, Inc., Charlottesville, VA), and structures were determined and refined using the PHENIX program (34).

Nuclear Magnetic Resonance Spectroscopy (NMR)

Solution NMR spectra were recorder with Bruker 500 or 300 MHz NMR systems (Billerica, MA). Proton NMR spectra were obtained in solution containing 10% D_2O/90% H_2O 100 mM NaCl, 10 mM sodium phosphate and 0.2 mM EDTA at pH 7.0. To ensure duplex formation, oligonucleotides (200 A_{260} OD units, 2.5 mM) were heated at 86 °C for 2 min and slowly cooled to room temperature prior to acquisition of NMR spectra. Each spectrum was calibrated using DSS (4,4-dimethyl-4-silapentane-1-sulphonate) as an internal standard

reference. Proton NMR spectra in aqueous solution were acquired with a water suppression double gradient echo WATERGATE W5 pulse program (35). The temperature of the sample was controlled by a variable temperature monitor (Eurotherm, BVT 3000). Proton resonances were assigned from 2D NOE spectra (36, 37). The ^{15}N edited proton NMR spectra were obtained using an HSQC pulse sequence (38).

Rotational dynamics

The ^{15}N edited proton NMR experiments were used to study the dynamics of cytosine amino group ($NH_{2)}$ rotation by observing the transfer of magnetization from one amino proton of the amino group to the other (39, 40). The resonance of each amino proton was selectively inverted by a rectangular soft pulse and the effect of magnetization transfer from the inverted amino proton to the other was studied. The signal intensity of both amino protons was monitored as a function of the delay time after the inversion pulse, ranging from 5 μsec to 4 sec. In the experiments described here, the magnetization of a particular amino proton at a particular time will be according to the equation 1,

$$M(t) = M^0 + C^* \, exp\,(\lambda_1 t) + F^* \, exp\,(\lambda_2 t) \quad (1)$$

where λ_1 and λ_2 are

$$\lambda_1 = -\,(R_{1A} + R_{1B})\,/\,(2 - \sigma) \quad (2)$$

and

$$\lambda_2 = -\,(R_{1A} + R_{1B})\,/\,(2 - 2k_r + \sigma) \quad (3)$$

M^0 is the equilibrium magnetization, C and F are constant coefficients depending on the relaxation rates of the two protons. R_{1A} and R_{1B} are the longitudinal relaxation rates, k_r is the rate of rotation, and σ is the cross relaxation rate. Values for λ_1 and λ_2 are obtained from the nonlinear fit of the plot of the intensity of the amino proton resonance *versus* delay time using equation 1. Rotation rates are obtained from λ_1 and λ_2 using equations 2 and 3. The cross relaxation rate (σ) is determined from the initial NOE buildup rate at lower mixing times (5-40 msec).

Quantum mechanical studies of 5-chlorocytosine

Quantum mechanical calculations (41,42) were performed with the Gaussian 03 program (Gaussian, Inc. Wallingford, CT) on a Pentium pro quad processor. Density functional theory (DFT) calculations were performed with the 6-31G basis set (B3LYP) within the Gaussian software. The template structure of 2'-deoxycytidine (dC) was used as the initial structure. The template structure of dC was edited to create the corresponding 5-methyl and 5-chloro-2'-deoxynucleosides. Bond lengths for each molecule were obtained from the energy-minimized structures.

Results

The synthetic approach presented here yields oligonucleotides containing selected ^{15}N-enriched cytosine amino groups. The ^{15}N-amino groups are introduced during the final deprotection of the oligonucleotides. Therefore this is an efficient method for inserting one or more labels site-specifically in either cytosine or cytosine analogs (Figure 1). Previously, ^{15}N-enriched cytosine amino groups have been introduced into oligonucleotides by synthesis of the ^{15}N-enriched phosphoramidite, or by introduction of a 4-triazole phosphoramidite (43-45). As precursors for the introduction of the 5-fluorocytosine or 5-chlorocytosine, we find the O^4-ethyl analogs to be suitably stable during phosphoramidite

and oligonucleotide synthesis and to be quantitatively displaced by ammonia during oligonucleotide deprotection with minimal hydrolysis to the corresponding uracil analogs. Base and isotopic composition of the oligonucleotides listed in Table 1 were verified by both MALDI-ToF MS and GC/MS (Supporting Information Figure S1).

As a first approach to identify potential perturbations in oligonucleotide structure and dynamics, the thermal and thermodynamic properties of a series of oligonucleotides containing C, ^{Cl}C or ^{m}C were investigated by examining duplex dissociation as a function of temperature (Figure 2). Thermodynamic parameters were obtained by fitting the experimental melting curves or from a plot of $1/T_m$ versus lnCt obtained at different oligonucleotide concentrations (Ct) as described previously (32, 33; Supporting Information Figure S2). The experimental values obtained in this study for the oligonucleotides shown in Figure 1 are listed in Table 1. The experimental T_m values obtained in this study are close to the value of 59 °C predicted for the cytosine-containing duplex based upon the nearest-neighbor model of SantaLucia et al. (46). In thermal melting studies, T_m's can be measured reproducibly to within 0.5 ° C. In the sequences studied here containing one modified cytosine analog and either cytosine or ^{m}C in the opposing strand of the CpG dinucleotide, observed melting temperatures for the ^{Cl}C and ^{m}C duplexes are experimentally indistinguishable (Figure 2). Therefore, in the sequences studied here, both ^{m}C and ^{Cl}C have no measurable impact on duplex thermal stability.

In order to determine the comparative impact of C, ^{m}C and ^{Cl}C on duplex structure, oligonucleotide duplexes were studied by x-ray crystallography and by NMR spectroscopy in aqueous solution. The crystal structure of a duplex containing ^{m}C was determined to 1.2 Å and refined to an R-factor of 0.249 (Rfree = 0.278) by the molecular replacement method using an unmodified Dickerson-Drew dodecamer (pdb accession number: 1FQ2). The structure of the ^{Cl}C-containing duplex was determined using the same method to 1.9 Å, and refined to an R-factor of 0.195 (Rfree=0.262) (Supporting Information Table S1). The ^{m}C and ^{Cl}C structures closely resemble each other, with an RMSD 0.25 Å. Both duplexes adopt a B-form DNA configuration, with RMSD 0.22 and 0.23 Å relative to the Dickerson-Drew duplex containing cytosine (Figure 3).

A more detailed comparative structural analysis was performed on each duplex examined here using the program CURVES+ (47) and the corresponding crystallographic PDB files as input. The CURVES+ program can reveal more subtle differences in the orientation of the base pair axis (BP-axis), base rotation and translation as well as other helical parameters. Data obtained with the CURVES+ program is tabulated in Supplementary Table S2 in which values for each modified base pair as well as average values for all base pairs in each duplex. Differences among the cytosine, ^{Cl}C and ^{m}C duplexes are modest, and usually the values for the ^{Cl}C and ^{m}C duplexes are more similar to one another that to the cytosine-containing oligonucleotide.

NMR studies in aqueous solution were performed to examine potential dynamic influences of each substitution. The aromatic base and H1' sugar residues for each duplex were observed and assigned by standard two-dimensional NMR methods (Supporting Information Figure S3). Chemical shifts are tabulated in Supporting Information Table S3 and Table S4. The observed NOE connectivities are consistent with a predominantly B-form helix in which all of the bases are intrahelical (Supporting Information Figure S4). Significant differences are noted for the chemical shifts of the H6 protons for C, ^{m}C and ^{Cl}C in each of the duplexes examined. In order to determine the effect of the 5-substituent on the chemical shift of the H6 proton, NMR spectra were also recorded for the corresponding free 2'-deoxynucleosides in D_2O (Supporting Information Table S5).

The proton NMR spectrum of the exchangeable proton region for each duplex at 300 K showed five imino proton resonances for these symmetric 12-base duplexes (Figure 4). The imino protons of the terminal base pairs are exchange-broadened under these conditions and are not observable. However, the terminal imino protons are observed at lower temperature (Supporting Information Figure S5). For the three duplexes containing C, ^{m}C or ^{Cl}C at position 3, the imino region is essentially the same, with one exception. The chemical shift of the imino proton of G10, paired with the C or substituted C at position 3 is different in each spectrum, ranging from 12.95 ppm when paired with ^{m}C to 12.54 ppm when paired with ^{Cl}C. A similar trend was observed in duplexes containing ^{m}C in the opposing strand (Supporting Information Table S6).

The amino protons of the cytosine residues are observed in a spectrally crowded region between 6 and 9.5 ppm (Figure 5). The amino proton hydrogen-bonded to the G residue is shifted downfield relative to the non-hydrogen bonded amino proton. In order to directly observe the amino protons of the cytosine residue at oligonucleotide position 3, the nitrogen of the amino groups was ^{15}N enriched. Selective enrichment permits direct observation of the *J*-coupled amino protons as shown in Figure 5a. The amino resonances observed in ^{15}N-decoupled spectra are shown in Figure 5b-d. The observed imino and amino proton resonances indicate that G pairs with ^{m}C and ^{Cl}C in a hydrogen-bonding pattern similar to the Watson-Crick C-G base pair.

In order to examine the dynamics of the target cytosine residue at position 3, the rotation rate of the amino group was measured by the saturation-transfer method described in the Experimental Section and shown in Supporting Information Figures S6 and S7. The rotational rates of C, ^{m}C and ^{Cl}C in the corresponding oligonucleotides are 32.9, 33.4 and 12.7 sec^{-1} respectively (Table 2). The placement of a ^{m}C residue in the opposing strand had minimal impact on the amino group rotation rate (Table 2), and in both sequence contexts, the amino group of the ^{Cl}C residue rotates at approximately half the rate of the corresponding ^{m}C residue. The linewidth of the G-imino proton of the X3-G10 bases pair was also measured as a function of temperature. For the three duplexes examined here, the temperature dependence of the G-imino proton resonance linewidth overlapped substantially as shown in Figure 6.

Discussion

A. Substitution with ^{Cl}C does not alter duplex stability as measured in thermal melting studies

In this study, a total of six symmetric oligonucleotides forming six duplexes were studied. Each oligonucleotide contained ^{m}C, C or ^{Cl}C at position X3 (Figure 1b). Three of the duplexes contained C at position Y9 (Figure 1b), while three contained ^{m}C at position Y9. In eukaryotic DNA, ^{m}C is usually found in symmetrically methylated ^{m}CpG dinucleotides, however, ^{m}C can also be found in hemimethylated CpG dinucleotides following DNA replication prior to maintenance methylation (20). Oligonucleotide melting temperatures are measured by observing optical absorbance as a function of temperature (Figure 2). Thermodynamic parameters are extracted from the melting studies either by curve-fitting or by measuring melting temperature as a function of oligonucleotide concentration (32,33) and differences obtained with each method were at or below experimental error. The thermal and thermodynamic parameters for duplex dissociation are given in Table 1. In the duplexes examined here, changes in duplex melting temperature resulting from the replacement of C by ^{m}C or ^{Cl}C were at or below the experimental error. Similarly, thermodynamic differences were at or below experimental error. Therefore, the replacement of ^{m}C by ^{Cl}C has little or no observable impact on the thermal or thermodynamic stability of an oligonucleotide duplex, at least within the sequence examined here. Previous studies have shown that

cytosine methylation generally increases oligonucleotide melting temperatures, however, the increase is modest and sequence-dependent (48, 49).

B. The crystal structures of the C, ^mC and ^{Cl}C duplexes are superimposable

The duplex structures of the ^{Cl}C3/C9 and ^mC3/C9 oligonucleotides were examined by crystallography and compared with the native Dickerson dodecamer (C3/C9). All the duplexes form typical B-form geometry. The ^{Cl}C or ^mC substitution in position 3 of the sequence had negligible effect upon the structures (Figure 3). Consistent with our findings, previous studies have shown that the 5-bromocytosine within the same sequence context is also B-form and is isomorphous to the native structure (28). Oligonucleotides containing 5-methylcytosine or 5-bromocytosine have been shown to promote alternative A or Z forms depending on the sequence context (50), however, these alternative forms require unusual DNA sequences.

C. Examination of non-exchangeable protons reveals the similarity of the ^{Cl}C and ^mC duplexes

Upon duplex formation, the solution ^{1}H-NMR chemical shifts of non-exchangeable protons are observed to move upfield relative to monomer chemical shifts due to base-base stacking interactions. In the duplex, the non-exchangeable resonances can be assigned by sequential NOE connectivities (36, 37) as shown and tabulated in Supporting Information (Table S3 and S4). The observed non-exchangeable base proton-H1' connectivities indicate that all duplexes studied here are predominantly B-form and that all bases are intrahelical.

The magnitudes of the upfield shifts of the non-exchangeable resonances are proportional to the degree of base-stacking interactions and are sensitive indicators of base geometry within the duplex. With one exception, the chemical shifts of the non-exchangeable aromatic proton resonances are invariant among the duplexes containing ^mC, C and ^{Cl}C at position X3 when Y9 is C or ^mC (Supporting Information Figure S3). The single exception in each case is the chemical shift of the H6 proton of the ^mC, C or ^{Cl}C residue at position X3 where these resonances are observed at 7.04, 7.25 and 7.54-7.55 ppm, respectively where Y9 is either C or ^mC. The observed chemical shift of the H6 resonance in a duplex is dependent upon both the intrinsic chemical shift of the proton resonance and the base-stacking change resulting from duplex formation. In pyrimidine analogs, a 5-substituent can significantly alter the chemical shift of H6, with electron donating substituents such as CH_3 moving the chemical shift to higher field, and electron withdrawing substituents shifting in the opposing direction. The chemical shifts of the H6 proton for the corresponding 2'-deoxynucleosides in water at 300K are 7.64, 7.81 and 8.11 ppm, for ^mdC, dC and ^{Cl}dC, respectively. The stacking-induced shifts in the sequence examined here are observed to be 0.60, 0.56 and 0.56 for ^mC, C and ^{Cl}C, respectively, and therefore the observed H6 chemical shift differences among the duplexes examined are dominated by the 5-substituent effect on the intrinsic chemical shift of the corresponding H6 proton. The similarity in the base-stacking induced change in the H6 chemical shift across the duplexes examined here is consistent with a negligible effect of ^{Cl}C substitution on duplex stability or conformation.

D. Examination of exchangeable protons resonances reveals differences between the duplexes

Upon duplex formation, the resonances of exchangeable amino and imino protons become narrower and shift downfield due to the formation of hydrogen bonds with bases in the complementary strand. The imino proton regions of duplexes containing ^mC, C and ^{Cl}C with C at position Y9 are shown in Figure 4. With the exception of the chemical shift of the G10 resonance of the X3-G10 base pair, the spectra are largely indistinguishable. The G10 imino

proton from the ^{Cl}C-G base pair is to high field of the other G imino protons, but well within the range expected for a normal Watson-Crick base pair.

A substituent in the 5-position of the cytosine ring can influence the basicity of the N^3 position which is the hydrogen bond acceptor for the G10 imino proton. The pK_a of the N^3 position of dC is 4.2, and an electron donating CH_3 substituent in the 5-position increases the pK_a to 4.3 (30). In contrast, an electron withdrawing Cl substituent decreases the pK_a of ^{Cl}dC to 2.7 (30). In Figure 7a, the G10 chemical shift is plotted versus the pK_a of the paired cytosine or cytosine analog, and a linear relationship is observed. The presence of a cytosine 5-substituent has an observable effect upon the G-C imino hydrogen bond which is proportional to the substituent effect on the corresponding N^3 pK_a value.

The dynamics of the base pair formed between X3 and G10 can also be investigated by examining the linewidth of the G10 imino resonance as a function of temperature. The G10 imino proton linewidths for duplexes containing ^{m}C, C and ^{Cl}C are shown in Figure 6. As the temperature is increased from 5°C, the linewidth for each duplex decreases and then increases due to enhanced base pair opening and proton exchange with solvent. However, the linewidths for the three duplexes are very similar and appear to converge around physiological temperature. These data suggest that although the presence of ^{Cl}C decreases the strength of the imino hydrogen bond with G10, the ^{Cl}C-G base pair is held intact by base-stacking interactions and does not open with greater frequency than the corresponding base pairs with C or ^{m}C, consistent with the thermodynamic results presented above.

The strength of the hydrogen bond can also be estimated by the downfield shift of the imino proton within the oligonucleotide relative to its monomer value. The intrinsic chemical shift of the imino proton of the 2'-deoxyguanosine nucleoside is 10.59 ppm (in DMSO) whereas the chemical shifts of guanine imino protons (G10) in C-G, ^{m}C-G and in ^{Cl}C-G base pairs are 12.90, 12.95 and 12.54 ppm respectively for the C3/C9, ^{m}C3/C9, ^{Cl}C3/C9 oligonucleotides. The difference in chemical shift between the oligo and corresponding nucleoside are 2.31, 2.36 and 1.95 ppm for the C3/C9, ^{m}C3/C9 and ^{Cl}C3/C9 oligonucleotides which indicates that the ^{m}C3-G10 base pair has the strongest hydrogen bond whereas the ^{Cl}C3-G10 base pair has the weakest. Similar relative magnitudes of base-pairing induced shifts were observed for oligonucleotides in which ^{m}C was placed in the opposing strand.

The dynamics of the base pairs formed between G10 and ^{m}C, C and ^{Cl}C can also be examined by observation of the amino protons of the paired cytosine or cytosine analog. The amino protons of cytosine analogs rotate at a relatively slow rate on the NMR time scale and can be observed as separate resonances. With monomeric 2'-deoxycytidine (dC) in solution, the hydrogen-bonding H_a or Watson-Crick (WC) proton is exchange-broadened and to high field of the non-hydrogen bonding, non-Watson Crick (nWC) proton, H_b (Figure 1c). With ^{m}dC, both amino protons are coincident, however, with ^{Cl}dC, the WC proton H_a is to low field of the nWC proton H_b (Supporting Information Table S5). Within the dC series, a 5-substituent has a much greater effect upon the chemical shift of the nWC proton. When comparing the linewidth of the H_a and H_b protons of dC and ^{Cl}dC at 5°C, the H_a proton is broader than H_b in all cases. The linewidths of the H_a protons for dC and ^{Cl}dC are 33.4 Hz versus 29.2 Hz, respectively, however, the linewidth of ^{Cl}dC H_b is only 13.1 Hz whereas H_b for dC is 23.0 Hz. The line widths for the ^{m}dC (monomer) amino protons couldn't be determined because of resonance overlap. The observed linewidth is a function of both chemical exchange with solvent and rotational line broadening. The data reported here suggest that in the monomers, dC and ^{Cl}dC, the rates of chemical exchange are similar (H_a). However, the narrower linewidth of H_b for ^{Cl}dC indicates that the amino group is rotating more slowly relative to dC.

The amino resonances of ^{15}N-isotope-enriched ^{m}C, C and ^{Cl}C within duplex oligonucleotides were edited from a crowded spectral region and assigned as shown in Figure 5. In each duplex, the H_a proton moves considerably downfield due to hydrogen bond formation with the O^6 position of G10 and becomes substantially narrower whereas the H_b proton moves slightly upfield due to base-stacking interactions. The magnitude of the hydrogen-bonding induced change in H_a chemical shift for ^{m}C, C and ^{Cl}C are 2.06, 1.89 and 1.95 ppm whereas the corresponding changes for the ^{m}C9 oligos are 2.11, 1.92 and 2.02 ppm. These data confirm the formation of a base pair between ^{Cl}C and G10 that approximates Watson-Crick geometry.

The rate of rotation of an amino group can be studied by using magnetization transfer as demonstrated previously by Russu and coworkers (39, 40), and values obtained with these methods are given in Table 2. The rotation rate of the ^{m}C amino group (X3 = ^{m}C) is approximately 33 sec^{-1}, when Y9 is C and 37 sec^{-1} when Y9 is ^{m}C. In the C-containing duplex (X3 = C) the amino group rotation rate is approximately 33 sec^{-1} when Y9 is either C or ^{m}C. In contrast, the apparent rate of the ^{Cl}C amino group (X3 = ^{Cl}C) is 12.7 sec^{-1} (Y9 = C) but increases to 14.1 sec^{-1} when ^{m}C is in the opposing strand (Y9 = ^{m}C). The temperature-dependent linewidth analysis shows that the labeled amino resonance of the ^{Cl}C3/C9 oligonucleotide broadens slower than those in the C3/C9 or ^{m}C3/C9 oligonucleotides which indicates that the ^{Cl}C amino group rotates more slowly (Supporting Information Figure S8).

The reduced rate of rotation of the amino group of ^{Cl}C relative to C or ^{m}C could be attributed to a reduced rate of base pair opening or to a slower intrinsic rotation rate resulting from the 5-chloro substituent. A previous study has shown that the C^4-N^4 bond of cytosine and its analogs have considerable double-bond character (41), consistent with its slow amino group rotation rate relative to adenine and guanine analogs. We investigated the geometry of the ^{m}dC, dC and ^{Cl}dC monomers through DFT calculations performed with the 6-31G basis set (B3LYP). The C^4-N^4 bond lengths of ^{m}dC, dC, ^{Cl}dC are found to be 1.363, 1.364 and 1.355 Å respectively. The reduced C^4-N^4 bond length of ^{Cl}dC is consistent with more double bond character, and would result in a reduced rate of amino group rotation, as observed in studies with the monomers discussed above.

Consistent with the computational results discussed above, the rate of rotation of the cytosine amino group in the duplexes examined here is related to the electronic inductive property of the 5-substituent. A linear relationship (Figure 7b) is obtained when the amino group rotation rate is plotted versus the Hammett parameter (σ_m) for ^{m}C, C and ^{Cl}C. The Hammett parameters for the selected substituents are obtained from Hansch et al. (51). The methyl group of ^{m}C is electron donating whereas the Cl substituent of ^{Cl}C is electron withdrawing. Therefore, the slower rotation rate of the ^{Cl}C amino group in the duplex can be attributed primarily to the electronic-inductive effect of the 5-Cl substituent.

E. Summary and conclusions

A series of oligonucleotide duplexes were constructed containing ^{m}C, C, and ^{Cl}C with either C or ^{m}C in the opposing strand within a CpG dinucleotide. Optical melting studies indicate that the replacement of C or ^{m}C with ^{Cl}C does not alter observably duplex thermal or thermodynamic stability. Results of crystallographic and ^{1}H-NMR studies establish that the overall geometries of the ^{Cl}C-containing duplexes are normal, B-form and that the base pair formed between ^{Cl}C and G approximates Watson-Crick geometry. A correlation is observed between the chemical shift of the imino proton of the G residue and the pK_a of the N^3 position of the paired cytosine analog for the series ^{m}C, C and ^{Cl}C, suggesting that the electron-withdrawing 5-chloro substituent, which reduces the pK_a of the N^3 position, proportionately decreases the strength of the G-^{Cl}C imino hydrogen bond. Despite an

apparently reduced imino hydrogen bond strength, the G imino proton of the G-^{Cl}C base pair is not unusually broadened with increasing temperature, suggesting that the local base-stacking, which is unchanged by ^{Cl}C substitution, is sufficient to maintain the base pair in a predominantly closed state. The H_a amino proton of ^{Cl}C forms a hydrogen bond with the O^6 position of the paired G residue, and the ^{Cl}C amino group rotates at approximately half the rate of the C or ^{m}C amino groups. The reduced rate of amino group rotation is attributed to the 5-chloro substituent increasing the double-bond character of the C^4-N^4 bond of ^{Cl}C as opposed to slowing the rate of base pair opening. These results are in accord with previous biochemical findings that ^{Cl}C is a very good mimic for ^{m}C and that it can act as a fraudulent epigenetic signal. Prior studies on DNA damage recognition have shown that glycosylases of the base-excision repair pathway can exploit structural and thermodynamic perturbations of damaged base pairs in distinguishing modified and normal bases in DNA (27, 52, 53). The apparent structural and dynamic similarities between ^{Cl}C and ^{m}C oligonucleotides shown here are in accord with the reported inability of glycosylases to recognize and remove ^{Cl}C from DNA. Therefore, the formation of ^{Cl}C in DNA poses a significant challenge for base excision repair and ^{Cl}C is likely to accumulate in DNA over time as a function of inflammation exposure and might in part explain the established relationship between inflammation and cancer development (54, 55).

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Abbreviations

mC	5-methylcystosine
ClC	5-chlorocytosine
BrC	5-bromocytosine
HOCl	hypochlorous acid
Tm	melting temperature
MALDI-ToF-MS	Matrix-Assisted Laser Desorption/Ionization -Time of Flight Mass Spectrometry
GC/MS	Gas chromatography-mass spectrometry
NOE	Nuclear Overhauser effect

References

1. Sandoval J, Esteller M. Cancer epigenomics: beyond genomics. Curr. Opin. Dev. 2012; 22:50–55.

2. You JS, Jones PA. Cancer genetics and epigenetics: two sides of the same coin. Cancer Cell. 2012; 22:9–20. [PubMed: 22789535]

3. Shen H, Laird PW. Interplay between the cancer genome and epigenome. Cell. 2013; 153:38–55. [PubMed: 23540689]

4. Bird AP, Wolffe AP. Methylation-induced repression—belts, braces, and chromatin. Cell. 1999; 99:451–454. [PubMed: 10589672]

5. Valinluck V, Tsai HH, Rogstad DK, Burdzy A, Bird A, Sowers LC. Oxidative damage to methyl-CpG sequences inhibits the binding of the methyl-CpG binding domain (MBD) of methyl-CpG binding protein 2 (MECP2). Nucleic Acids Res. 2004; 32:4100–4108. [PubMed: 15302911]

6. Jones PA, Baylin SB. The epigenomics of cancer. Cell. 2007; 128:683–692. [PubMed: 17320506]

7. McCabe MT, Brandes JC, Vertino PM. Cancer DNA methylation: molecular mechanisms and clinical implications. Clin. Cancer Res. 2009; 15:3927–3937. [PubMed: 19509173]

8. Kanwal R, Gupta S. Epigenetic modifications in cancer. Clin. Genet. 2012; 81:303–311. [PubMed: 22082348]

9. Lewis JG, Adams DO. Inflammation, oxidative DNA damage, and Carcinogenesis. Environ. Health Perspect. 1987; 76:19–27. [PubMed: 3129286]

10. Ames BN, Gold LS, Willett WC. The causes and prevention of cancer. Proc. Natl. Acad. Sci. U S A. 1995; 92:5258–5265. [PubMed: 7777494]

11. Lonkar P, Dedon PC. Reactive species and DNA damage in chronic inflammation: reconciling chemical mechanisms and biological fates. Int. J. Cancer. 2011; 128:1999–2009. [PubMed: 21387284]

12. Whiteman M, Jenner A, Halliwell B. Hypochlorous acid-induced base modifications in isolated calf thymus DNA. Chem. Res. Toxicol. 1997; 10:1240–1246. [PubMed: 9403176]

13. Winterbourn CC, Kettle AJ. Biomarkers of myeloperoxidase-derived hypochlorous acid. Free Radic. Biol. Med. 2000; 29:403–409. [PubMed: 11020661]

14. Badouard C, Masuda M, Nishino H, Cadet J, Favier A, Ravanat JL. Detection of chlorinated DNA and RNA nucleosides by HPLC coupled to tandem mass spectrometry as potential biomarkers of inflammation. J. Chromatogr B Analyt. Technol. Biomed. Life Sci. 2005; 827:26–31.

15. Kang JI, Sowers LC. Examination of hypochlorous acid-induced damage to cytosine residues in a CpG dinucleotide in DNA. Chem. Res. Toxicol. 2008; 21:1211–1218. [PubMed: 18826175]

16. Mangerich A, Knutson CG, Parry NM, Muthupalani S, Ye W, Prestwich E, Cui L, McFaline JL, Mobley M, Ge Z, Taghizadeh K, Wishnok JS, Wogan GN, Fox JG, Tannenbaum SR, Dedon PC. Infection-induced colitis in mice causes dynamic and tissue-specific changes in stress response and DNA damage leading to colon cancer. Proc. Natl. Acad. Sci. U S A. 2012; 109:E1820–1829. [PubMed: 22689960]

17. Seiberling KA, Chruch CA, Herring J, Sowers LC. Epigenetics of chronic rhinosinusitis and the role of the eosinophil. Int. Forum Allergy Rhinol. 2012; 2:80–84. [PubMed: 22311846]

18. Valinluck V, Liu P, Kang JI Jr. Burdzy A, Sowers LC. 5-halogenated pyrimidine lesions within a CpG sequence context mimic 5-methylcytosine by enhancing the binding of the methyl-CpG-binding domain of methyl-CpG-binding protein 2 (MeCP2). Nucleic Acids Res. 2005; 33:3057–3064. [PubMed: 15917437]

19. Lao VV, Darwanto A, Sowers LC. Impact of base analogues within a CpG dinucleotide on the binding of DNA by the methyl-binding domain of MeCP2 and methylation by DNMT1. Biochemistry. 2010; 49:10228–10236. [PubMed: 20979427]

20. Valinluck V, Sowers LC. Endogenous cytosine damage products alter the site selectivity of human DNA maintenance methyltransferase DNMT1. Cancer Res. 2007; 67:946–950. [PubMed: 17283125]

21. Valinluck V, Sowers LC. Inflammation-mediated cytosine damage: a mechanistic link between inflammation and the epigenetic alterations in human cancers. Cancer Res. 2007; 67:5583–5586. [PubMed: 17575120]

22. Lao VV, Herring JL, Kim CH, Darwanto A, Soto U, Sowers LC. Incorporation of 5-chlorocytosine into mammalian DNA results in heritable gene silencing and altered cytosine methylation patterns. Carcinogenesis. 2009; 30:886–893. [PubMed: 19279184]

23. Bennett MT, Rodgers MT, Hebert AS, Ruslander LE, Eisele L, Drohat AC. Specificity of human thymine DNA glycosylase depends on N-glycosidic bond stability. J. Am .Chem. Soc. 2006; 128:12510–12519. [PubMed: 16984202]

24. Brennan CA, Van Cleve MD, Gumport RI. The effects of base analog substitutions on the cleavage by the EcoR1 restriction endonuclease of octadeoxyribonucleotides containing modified Eco R1 recognition sequences. J. Biol. Chem. 1986; 261:7270–7278. [PubMed: 3011780]

25. Brennan CA, Van Cleve MD, Gumport RI. The effects of base analog substitutions on the methylation by the EcoR1 modification methylase of octadeoxyribonucleotides containing modified Eco R1 recognition sequences. J. Biol. Chem. 1986; 261:7279–7286. [PubMed: 3011781]

26. Liu P, Burdzy A, Sowers LC. Substrate Recognition by a Family of Uracil-DNA Glycosylases: UNG, MUG, and TDG. Chem. Res. Toxicol. 2002; 15:1001–1009. [PubMed: 12184783]

27. Darwanto A, Theruvathu JA, Sowers JL, Pascal T, Goddard W 3rd, Sowers LC. Mechanism of base selection by human single-stranded selective monofunctional uracil-DNA glycosylase. J. Biol. Chem. 2009; 284:15835–15846. [PubMed: 19324873]

28. Fratini AV, Kopka ML, Drew HR, Dickerson RE. Reversible bending and helix geometry in a B-DNA dodecamer: CGCGAATTBrCGCG. J. Biol. Chem. 1982; 257:14686–14707. [PubMed: 7174662]

29. Kang JI Jr. Burdzy A, Liu P, Sowers LC. Synthesis and characterization of oligonucleotides containing 5-chlorocytosine. Chem. Res. Toxicol. 2004; 17:1236–1244. [PubMed: 15377157]

30. Sowers LC. ^{15}N-enriched 5-fluorocytosine as a probe for examining unusual DNA structures. J. Biomol. Struct. Dyn. 2000; 17:713–723. [PubMed: 10698108]

31. Cui Z, Theruvathu JA, Farrel A, Burdzy A, Sowers LC. Characterization of synthetic oligonucleotides containing biologically important modified bases by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry. Anal. Biochem. 2008; 379:196–207. [PubMed: 18485883]

32. Theruvathu JA, Kim CH, Rogstad DK, Neidigh JW, Sowers LC. Base pairing configuration and stability of an oligonucleotide duplex containing a 5-chlorouracil-adenine base pair. Biochemistry. 2009; 48:7539–7546. [PubMed: 19618901]

33. Theruvathu JA, Kim CH, Darwanto A, Neidigh JW, Sowers LC. pH-Dependent configurations of a 5-chlorouracil-guanine base pair. Biochemistry. 2009; 48:11312–11318. [PubMed: 19863157]

34. Adams PD, Afonine PV, Bunkóczi GV, Chen B, Davis IW, Echols N, Headd JJ, Hung L-W, Kapral GJ, Grosse-Kunstleve RW, McCoy AJ, Moriarty NW, Oeffner RJ, Read, Richardson DD, Richardson JS, Terwilliger TC, Zwart PH. PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Cryst. 2010; D66:213–221.

35. Liu M, Mao X, Ye C, Huang H, Nicholson JK, Lindon JC. Improved WATERGATE Pulse Sequences for Solvent Suppression in NMR Spectroscopy. J. Magn. Reson. 1998; 132:125–129.

36. Hare DR, Wemmer DE, Chou SH, Drobny G, Reid BR. Assignment of the non-exchangeable proton resonances of d(C-G-C-G-A-A-T-T-C-G-C-G) using two-dimensional nuclear magnetic resonance methods. J. Mol. Biol. 1983; 171:319–336. [PubMed: 6317867]

37. Weiss MA, Patel DJ, Sauer RT, Karplus M. Two-dimensional 1H NMR study of the lambda operator site OL1: a sequential assignment strategy and its application. Proc. Natl. Acad. Sci. U S A. 1984; 81:130–134. [PubMed: 6229791]

38. Grzesiek S, Bax A. The importance of not saturating water in protein NMR. Application to sensitivity enhancement and NOE measurements. J. Am. Chem. Soc. 1993; 115:12593–12594.

39. Michalczyk R, Russu IM. Rotational dynamics of adenine amino groups in a DNA double helix. Biophys. J. 1999; 76:2679–2686. [PubMed: 10233082]

40. Jiang L, Russu IM. Internal dynamics in a DNA triple helix probed by (1)H-(15)N-NMR spectroscopy. Biophys. J. 2002; 82:3181–3185. [PubMed: 12023242]

41. Moser A, Guza R, Tretyakova N, York DM. Density Functional Study of the Influence of C5 Cytosine Substitution in Base Pairs with Guanine. Theor. Chem. Acc. 2009; 122:179–188. [PubMed: 19890472]

42. Frisch MJ, Trucks GW, Schlegel HB, Scuseria GE, Robb MA, Cheeseman JR, Montgomery JA Jr. Vreven T, Kudin KN, Burant JC, Millam JM, Iyengar SS, Tomasi J, Barone V, Mennucci B, Cossi M, Scalmani G, Rega N, Petersson GA, Nakatsuji H, Hada M, Ehara M, Toyota K, Fukuda R, Hasegawa J, Ishida M, Nakajima T, Honda Y, Kitao O, Nakai H, Klene M, Li X, Knox JE, Hratchian HP, Cross JB, Bakken V, Adamo C, Jaramillo J, Gomperts R, Stratmann RE, Yazyev O,

Austin AJ, Cammi R, Pomelli C, Ochterski JW, Ayala PY, Morokuma K, Voth GA, Salvador P, Dannenberg JJ, Zakrzewski VG, Dapprich S, Daniels AD, Strain MC, Farkas O, Malick DK, Rabuck AD, Raghavachari K, Foresman JB, Ortiz JV, Cui Q, Baboul AG, Clifford S, Cioslowski J, Stefanov BB, Liu G, Liashenko A, Piskorz P, Komaromi I, Martin RL, Fox DJ, Keith T, Al-Laham MA, Peng CY, Nanayakkara A, Challacombe M, Gill PMW, Johnson B, Chen W, Wong MW, Gonzalez C, Pople JA. Gaussian 03, revision A.1.

43. Kupferschmitt G, Schmidt J, Schmidt T, Fera B, Buck F, Ruterjans H. [15]N labeling of oligodeoxynucleotides for NMR studies of DNA-ligand interactions. Nucleic Acids Res. 1987; 15:6225–6241. [PubMed: 3498149]

44. Ramesh V, Xu YZ, Roberts GC. Site-specific ^{15}N-labelling of oligonucleotides for NMR: the trp operator and its interaction with the trp repressor. FEBS Lett. 1995; 363:61–64. [PubMed: 7729556]

45. Acedo M, Fabrega C, Avino A, Goodman M, Fagan P, Wemmer D, Eritja R. A simple method for N-15 labeling of exocyclic amino groups in synthetic oligodeoxynucleotides. Nucleic Acids Res. 1994; 22:2982–2989. [PubMed: 8065910]

46. SantaLucia J Jr. Allawi HT, Seneviratne PA. Improved nearest-neighbor parameters for predicting DNA duplex stability. Biochemistry. 1996; 35:3555–3562. [PubMed: 8639506]

47. Lavery R, Moakher M, Maddocks JH, Petkeviciute D, Zakrzewska K. Conformational analysis of nucleic acids revisited: Curves+. Nucleic Acids Research. 2009; 37:5917–5929. [PubMed: 19625494]

48. Sowers LC, Shaw BR, Sedwick WD. Base Stacking and Molecular Polarizability: Effect of a Methyl Group in the 5-Position of Pyrimidines. Biochem. Biophys. Research Commun. 1987; 148:790–794. [PubMed: 3689373]

49. Lefebvre A, Mauffret O, el Antri S, Monnot M, Lescot E, Fermandjian S. Sequence dependent effects of CpG cytosine methylation. A joint ^{1}H-NMR and ^{31}P-NMR study. Eur. J. Biochem. 1995; 229:445–454. [PubMed: 7744067]

50. Fujii S, Wang AH, van der Marel G, van Boom JH, Rich A. Molecular structure of (m5 dC-dG)3: the role of the methyl group on 5-methyl cytosine in stabilizing Z-DNA. Nucleic Acids Res. 1982; 10:7879–7892. [PubMed: 7155900]

51. Hansch C, Leo A, Taft RW. A Survey of Hammett substituent constants and resonance and field parameters. Chem. Rev. 1991; 91:165–195.

52. Petruska JP, Sowers LC, Goodman MF. Comparison of Nucleotide Interactions in Water, Proteins and Vacuum: Model for DNA polymerase Fidelity. Proc. Natl. Acad. Sci. U S A. 1986; 83:1559–1562. [PubMed: 3456600]

53. Zheng H, Cai Y, Ding S, Tang Y, Kropachev K, Zhou Y, Wang L, Wang S, Geacintov NE, Zhang Y, Broyde S. Base flipping free energy profiles for damaged and undamaged DNA. Chem. Res. Toxicol. 2010; 23:1868–1870. [PubMed: 21090780]

54. Balkwill F, Mantovani A. Inflammation and cancer: back to Virchow? Lancet. 2001; 357:539–545. [PubMed: 11229684]

55. Schmidt A, Weber OF. In memoriam of Rudolf Virchow: A historical retrospective including aspects of inflammation, infection and neoplasia. Contrib. Microbiol. 2006; 13:1–15. [PubMed: 16627956]

Figure 1.
(a) Reaction scheme for the synthesis of ^{15}N labeled oligonucleotide. **(b)** The 12mer sequence and the oligonucleotides synthesized in this study. **(c)** The chemical structure of the G-C base pair. H_a is the Watson-Crick hydrogen and H_b is non-Watson-Crick hydrogen.
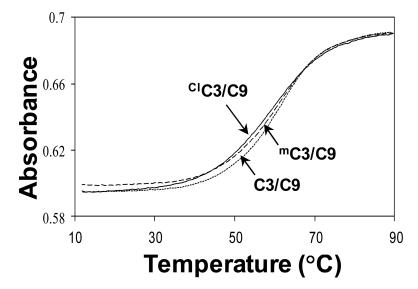
Figure 2.
Normalized melting profiles at 260 nm of 28 μM oligonucleotides in presence of 10 mM sodium phosphate, 100 mM sodium chloride, 0.1 mM EDTA at pH 7.0.
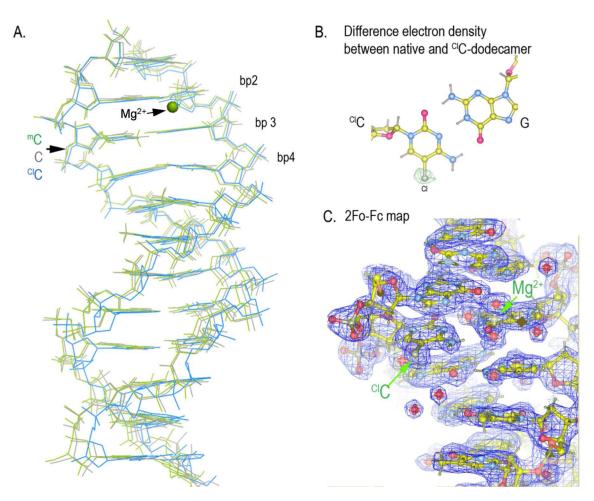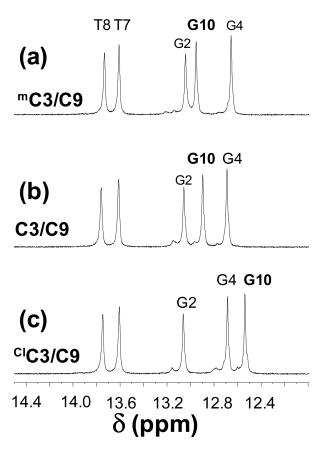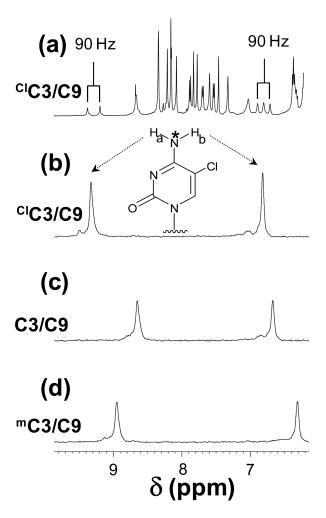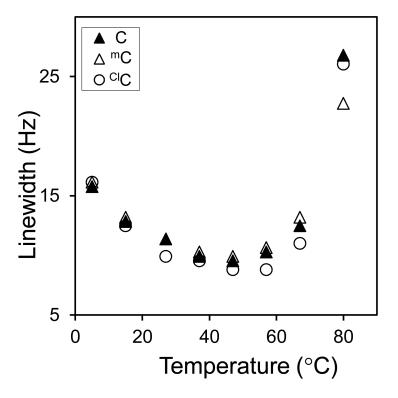
Figure 3.
Structure of the Dickerson-Drew dodecamer with a modified base at position 3. A) Superposition of ^{m}C (green) and ^{Cl}C (blue) containing duplex with native C dodecamer (gray). B) Difference map between the ^{Cl}C and C dodecamers (Fo^{ClC}-Fc^{C}, Φ^{ClC}) revealing electron density around Cl on ^{Cl}C at position 3. C) Electron density (2Fo-Fc) for the ^{Cl}C-containing dodecamer. Fo is the observed electron density map and Fc is the calculated electron density map. The structure factors and coordinates of the ^{Cl}C and ^{m}C-containing dodecamers are deposited to Protein Data Bank with accession number 4MGW and 4MKW, respectively.
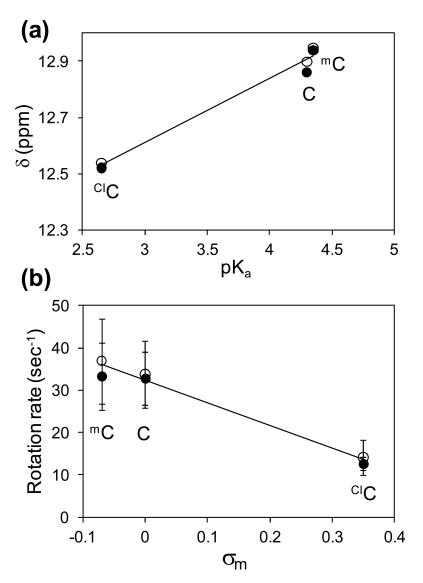
Figure 4.
^{1}H NMR spectra in the imino region of the [4-^{15}N] labeled (X3 position, Figure 1) oligonucleotides **(a)** ^{m}C3/C9, **(b)** C3/C9 and **(c)** ^{Cl}C3/C9 in the presence of 10% D_2O, 100 mM NaCl, 10 mM sodium phosphate and 0.2 mM EDTA at pH 7 and 300K.

Figure 5.
(a) ^{1}H NMR spectrum (unedited) of the [4-^{15}N] ^{Cl}C3/C9 oligo. (b-d) ^{15}N edited proton spectrum of ^{15}N labelled oligonucleotides in the presence of 10% D$_2$O, 100 mM NaCl, 10 mM sodium phosphate and 0.2 mM EDTA at pH 7 and 300K. ^{15}N label is indicated in the structure by the asterisk.

Figure 6.
Temperature dependent line width studies of G10 imino protons of C3/C9, ^{m}C3/C9
and ^{Cl}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 and 300K.

Figure 7.
(**a**) Proton chemical shifts of the G imino protons of the X3-G10 base pair of the oligonucleotide versus the *pK$_a$* of the corresponding nucleoside at pH 7, where X3 is ^mC, C or ^{Cl}C. (**b**) Amino group rotation rate as a function of the Hammett σ$_m$ parameter of the 5-substituent. Closed circles are for the X3/C9 oligo whereas open circles are for the X3/$^mC9$ oligo.

Table 1

T_m and thermodynamic parameters for 5'-CG**X**GAATT**Y**GCG-3' in 10 mM sodium phosphate, 100 mM sodium chloride and 0.1 mM EDTA at pH 7.0. T_m values are at 28 μM total strand concentration.

duplex	T_m (°C)	$\Delta H°$ (kcal mol^{-1})	$\Delta S°$ (cal mol^{-1}K^{-1})	$\Delta G°$ (25 °C) (kcal mol^{-1})
^{m}C3/C9	58.5 ± 0.5	−61.5 ± 5.8	−165 ± 18	−10.8 ± 0.5
^{m}C3/^{m}C9	58.7 ± 0.8	−61.9 ± 6.1	−166 ± 18	−10.8 ± 0.5
C3/C9	59.2 ± 0.1	−58.3 ± 9.0	−154 ± 27	−10.7 ± 0.8
C3/^{m}C9	59.9 ± 0.8	−59.3 ± 8.9	−157 ± 27	−10.9 ± 0.8
^{Cl}C3/C9	58.2 ± 0.2	−62.4 ± 10	−168 ± 31	−10.7 ± 0.7
^{Cl}C3/^{m}C9	58.9 ± 0.7	−60.5 ± 11	−162 ± 33	−10.6 ± 0.7

Table 2

Amino proton chemical shifts, rate of rotation of amino protons and line widths amino resonances of the 3rd base of the oligonucleotide (Figure 1) at 300 K and pH 7. Rate of rotation (k_r) determined by ^{15}N edited magnetization transfer experiments and line widths (half height width) using ^{15}N edited experiments. $\Delta\delta a$ = H_a-monomer H_a and $\Delta\delta b$ =H_b-monomer H_b.

duplex	k_r (sec^{-1})	Chemical shifts of C3 amino protons in oligos					Line width (Hz)	
		NH_2 H_a	NH_2 H_b	$\Delta\delta$ (H_a-H_b)	$\Delta\delta a$	$\Delta\delta b$	H_a	H_b
^{m}C3/C9	33.4	8.95	6.31	2.64	2.06	−0.58	31.9	26.8
^{m}C3/^{m}C9	36.9	9.00	6.23	2.77	2.11	−0.66	34.1	28.6
C3/C9	32.9	8.65	6.68	1.97	1.89	−0.56	33.0	25.7
C3/^{m}C9	33.8	8.68	6.59	2.09	1.92	−0.65	33.7	28.6
^{Cl}C3/C9	12.7	9.31	6.81	2.50	1.95	−0.43	26.8	21.6
^{Cl}C3/^{m}C9	14.1	9.38	6.74	2.64	2.02	−0.50	25.3	22.4