

Structural dynamics of double-stranded DNA with epigenome modification

Ayako Furukawa^{1,2}, Erik Walinda³, Kyohei Arita¹ and Kenji Sugase^{2,4,*}

¹Graduate School of Medical Life Science, Yokohama City University, 1-7-29 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan, ²Bioorganic Research Institute, Suntory Foundation for Life Sciences, 8-1-1 Seikadai, Seika, Soraku, Kyoto 619-0284, Japan, ³Department of Molecular and Cellular Physiology, Graduate School of Medicine, Kyoto University, Yoshida Konoe-cho, Sakyo-ku, Kyoto 606-8501, Japan and ⁴Department of Molecular Engineering, Graduate School of Engineering, Kyoto University, Kyoto-Daigaku Katsura, Nishikyo-Ku, Kyoto 615-8510, Japan

Received May 02, 2020; Revised November 25, 2020; Editorial Decision November 26, 2020; Accepted November 28, 2020

ABSTRACT

Modification of cytosine plays an important role in epigenetic regulation of gene expression and genome stability. Cytosine is converted to 5-methylcytosine (5mC) by DNA methyltransferase; in turn, 5mC may be oxidized to 5-hydroxymethylcytosine (5hmC) by ten-eleven translocation enzyme. The structural flexibility of DNA is known to affect the binding of proteins to methylated DNA. Here, we have carried out a semi-quantitative analysis of the dynamics of double-stranded DNA (dsDNA) containing various epigenetic modifications by combining data from imino ¹H exchange and imino ¹H $R_{1\rho}$ relaxation dispersion NMR experiments in a complementary way. Using this approach, we characterized the base-opening (k_{open}) and base-closing (k_{close}) rates, facilitating a comparison of the base-opening and -closing process of dsDNA containing cytosine in different states of epigenetic modification. A particularly striking result is the increase in the k_{open} rate of hemimethylated dsDNA 5mC/C relative to unmodified or fully methylated dsDNA, indicating that the Watson–Crick base pairs undergo selective destabilization in 5mC/C. Collectively, our findings imply that the epigenetic modulation of cytosine dynamics in dsDNA mediates destabilization of the GC Watson–Crick base pair to allow base-flipping in living cells.

INTRODUCTION

Cytosine modifications are known to play an important role in epigenetic regulation of gene expression and genome

stability (1,2). Cytosine is converted to 5-methylcytosine (5mC) by DNA methyltransferases (3,4); in turn, 5mC may be further oxidized to 5-hydroxymethylcytosine (5hmC) by ten-eleven translocation (TET) enzyme (5,6). Many proteins specifically recognize these DNA modifications. For example, UHRF1 (ubiquitin-like containing PHD and RING finger domains 1) specifically recognizes hemimethylated double-stranded DNA (dsDNA), in which one strand is methylated but the other is not. UHRF1 binds specifically to the ‘flipped-out’ 5mC structure (7–9), but does not interact with either unmodified dsDNA or full-methylated dsDNA, in which both strands are methylated (7). Although crystal structures of UHRF1 in the free and 5mC–DNA-bound forms have been reported (7–9), the mechanism by which the 5mC nucleotide flips out from the inside of the dsDNA for specific recognition by UHRF1 remains elusive.

One possibility is that DNA methylation increases the flexibility of dsDNA, thereby enabling the dsDNA at the position of 5mC to more frequently adopt a flipped-out structure or at least to adopt a more open structure that can then be recognized by UHRF1. Indeed, it has been shown that DNA methylation can alter DNA flexibility and affect the properties of proteins binding to the methylated DNA (10–12). Even without methylation, dsDNA is not completely rigid, but rather has a degree of flexibility. A particularly intriguing example of this flexibility is that the hydrogen bonds that form a Watson–Crick (WC) base pair in canonical dsDNA can be transiently broken, enabling the WC base pair to interconvert to a Hoogsteen (HG) base pair (13). The hydrogen bond break permits ¹H exchange (which can be experimentally observed) between water and the imino proton in thymine and guanine, even if the WC base pair is simply broken and does not fully interconvert into an HG base pair.

*To whom correspondence should be addressed. Tel: +81 75 383 7063; Fax: +81 75 383 2541; Email: sugase@moleng.kyoto-u.ac.jp
Present address: Kenji Sugase, Department of Molecular Engineering, Graduate School of Engineering, Kyoto University, Kyoto-Daigaku Katsura, Nishikyo-Ku, Kyoto 615-8510, Japan.

The ^1H exchange rate ($k_{1\text{H}}$) of the imino proton has been measured for various dsDNA constructs by using ^1H NMR. In this type of experiment, a catalyst, typically ammonium, is often used to enhance the $k_{1\text{H}}$ rate and to estimate the equilibrium constant K_{op} , which is the ratio of the base-opening rate (k_{open}) to the base-closing rate (k_{close}). However, ammonium can induce conformational changes in dsDNA and change the ^1H spin relaxation rate, thereby altering the derived $k_{1\text{H}}$ rate (14,15). In the case of 5hmC, for example, a hydrogen bond is formed between its hydroxyl group and the next base via water molecules, which may be disturbed by catalysts (16). Therefore, although it is difficult to estimate K_{op} without catalysts, their use should be avoided in the comparison of $k_{1\text{H}}$ rates among different dsDNA samples, such as those with cytosine bases in different states of epigenetic modification.

^1H exchange NMR experiments can indirectly characterize the base-opening and -closing process by permitting derivation of the equilibrium constant K_{op} . The base-opening and -closing process is usually a very fast process occurring on the microsecond timescale; therefore, it is very hard to probe under the conditions of physiological pH, temperature, and buffer, where the lifetime of the open state is very short-lived and its population is extremely low. By contrast, NMR relaxation dispersion experiments can provide the kinetic exchange rate (k_{ex}), which is the sum of k_{open} and k_{close} , even in the case where the population of the open state is too small to be directly observed by NMR (17,18). Indeed, the development of ^{13}C and ^{15}N $R_{1\rho}$ NMR relaxation dispersion has facilitated a direct characterization of the base-opening and -closing process, and even the above-mentioned WC to HG base pair exchange process (13,19–22).

In this study, we have carried out a semi-quantitative analysis of the base-opening and -closing process of unmodified and modified dsDNA samples by combining ^1H NMR $R_{1\rho}$ relaxation dispersion with ^1H exchange NMR to obtain insight into the mechanism by which the base-opening process may be exploited for the selective recognition of epigenetically modified dsDNA by proteins such as UHRF1. This combined NMR method can be applied without catalysts to study dsDNA dynamics even at neutral pH, provided that the base-closing rate is much faster than the intrinsic ^1H exchange rate, a condition that is commonly fulfilled for ^1H exchange NMR experiments of dsDNA (14–16). Using this method, we report a site-specific comparison of the base-opening and -closing process in dsDNA containing cytosine in different epigenetic modification states.

MATERIALS AND METHODS

Oligonucleotides

DNA oligonucleotides, synthesized and purified by high-performance liquid chromatography, were purchased from FASMAC (Figure 1 and Table 1). To prepare the experimental dsDNA samples, each oligonucleotide was suspended in 150 mM NaCl solution (pH 7.5 adjusted by NaOH), heated at 95°C for 5 min, and annealed by gradual cooling to 25°C.

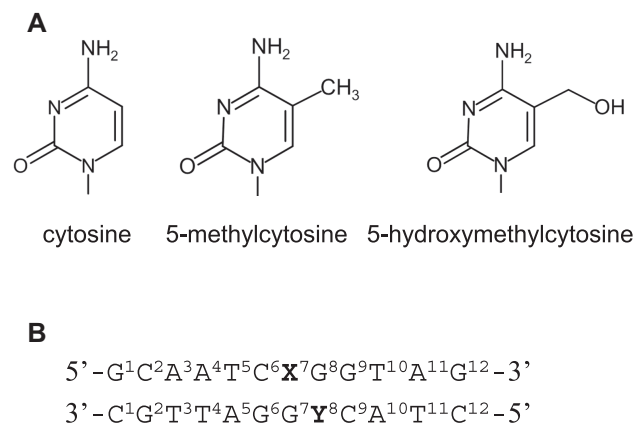


Figure 1. Structures of cytosine, the modified cytosine bases and the double-stranded DNA (dsDNA) sequence used. (A) Chemical structures of cytosine (C), 5-methylcytosine (5mC), and 5-hydroxymethylcytosine (5hmC). (B) The dsDNA sequence with the numbering used in this study. X and Y are C, 5mC or 5hmC.

Table 1. Melting temperature, T_m , of the six dsDNA samples

Sample	T_m (°C)
C/C	55.37 ± 0.09
5mC/C	55.43 ± 0.21
5mC/5mC	56.37 ± 0.28
C/5mC	56.42 ± 0.27
5hmC/C	55.04 ± 0.11
5hmC/5mC	56.00 ± 0.43

Thermal melting experiments

The melting temperature of 2 μM dsDNA samples was determined by using a V-650 UV-VIS spectrophotometer (JASCO). The samples were heated over a temperature range of 25–95°C at a rate of 1°C min⁻¹. The melting temperature (T_m) of each sample was derived from the first-order derivative of the melting curve.

NMR experiments

The dsDNA samples were dissolved at a final concentration of 300 μM in 150 mM NaCl solution (pH 7.5) containing 5% D₂O. All NMR experiments were carried out on an AVANCE DRX 600 spectrometer equipped with a TXI cryogenic probe and z-axis gradient or an AVANCE III HD 600 spectrometer equipped with a TCI cryogenic probe and z-axis gradient (Bruker). A 2D ^1H nuclear Overhauser spectroscopy (NOESY) spectrum was measured for each dsDNA at 15°C to obtain sequential NOE connectivities between imino ^1H resonances. The NMR spectra were processed by using NMRPipe (23), and imino ^1H resonances were assigned on the basis of NOE connectivities. The resultant assignments were transferred to those of the resonances at 30°C by measuring a series of one-dimensional (1D) NMR spectra at temperatures from 15 to 30°C. The imino ^1H resonances of G1, A11 and G12, which are located at or near either end of the dsDNA, remained unassigned because their flexibility provided no NOE connectivities at 15°C.

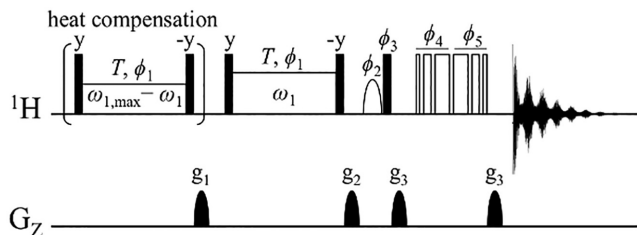


Figure 2. Pulse scheme for on-resonance imino proton $R_{1\rho}$ dispersion in dsDNA. All ^1H 90° radio-frequency (RF) pulses are shown as filled bars. The shaped 90° ^1H pulse shown as a half ellipse is the water-selective 1-ms sinc pulse. The heat compensation block is applied at 300 ppm downfield to maintain the sample at a constant temperature throughout the measurements. The RF power of the heat compensation pulse is varied as $\omega_{1,\max} - \omega_1$, where $\omega_{1,\max}$ is the maximum spin-lock power ω_1 used in a set of $R_{1\rho}$ dispersion measurements. After the g_1 gradient that eliminates initial unwanted transverse magnetization, the carrier frequency is jumped to an imino ^1H resonance of interest. The on-resonance ^1H magnetization is rotated by 90° and spin-locked at a spin-lock field strength ω_1 with a constant length T . The ^1H magnetization is flipped back to the z-axis, followed by a z-filter gradient g_2 . The 3–9–19 WATERGATE pulses are applied at the carrier frequency of 4.7 ppm to suppress the water signal, and lastly the 1D ^1H NMR spectrum is acquired. The phase cycle is given as follows: $\phi 1 = (x)$; $\phi 2 = (-x, x)$; $\phi 3 = (x, -x)$; $\phi 4 = (x, x, y, y, -x, -x, -y, -y)$; $\phi 5 = (-x, -x, -y, -y, x, x, y, y)$; rec = $(x, -x, -x, x)$.

Water–imino proton exchange rates, $k_{1\text{H}}$, were estimated from 1D CLEANEX-PM experiments at 30°C with mixing times τ_m of 5, 10, 15, 20, 50, 100 and 200 ms (24). Reference spectra were measured by the same pulse program with the CLEANEX-PM block omitted. NMR signals were deconvoluted if overlapping, and signal intensities were obtained by using TopSpin (Bruker). The obtained CLEANEX-PM intensities were fitted to the following equation by using the program GLOVE (25):

$$\frac{I(\tau_m)}{I_0} = \frac{k_{1\text{H}}}{R_A + k_{1\text{H}} - R_B} \times \{\exp(-R_B\tau_m) - \exp[-(R_A + k_{1\text{H}})\tau_m]\}, \quad (1)$$

where $I(\tau_m)$ and I_0 are the signal intensities measured at τ_m and measured as a reference, respectively; and R_A and R_B are the apparent relaxation rates of the imino proton and the water proton, respectively.

1D on-resonance imino ^1H $R_{1\rho}$ relaxation dispersion experiments were carried out at 30°C by using the pulse sequence shown in Figure 2. A similar pulse sequence in which the carrier frequency was set in the middle of the imino proton spectral region was recently reported (26). In our case, we set the carrier frequency on-resonance to each imino proton resonance. An advantage of the ^1H $R_{1\rho}$ relaxation dispersion experiment is that higher radio-frequency power can be used for ^1H than for other nuclei because of its higher gyromagnetic ratio, which means that ^1H $R_{1\rho}$ relaxation dispersion can detect faster conformational exchange as compared with its ^{13}C and ^{15}N counterparts. Furthermore, the ^1H $R_{1\rho}$ relaxation dispersion experiment does not require isotope-labeled samples. It is possible that the derived ^1H $R_{1\rho}$ relaxation rates may be influenced by ^1H – ^1H cross-relaxation. Therefore, we theoretically estimated the influence of this cross-relaxation on the ^1H $R_{1\rho}$ relaxation dispersion measurement (Supplementary Methods, Supple-

mentary Figures S1 and S2, Supplementary Tables S1 and S2), which showed that ^1H – ^1H cross-relaxation is negligible for the conditions (described below) under which the ^1H $R_{1\rho}$ relaxation dispersion profiles of the dsDNA samples were obtained.

^1H $R_{1\rho}$ relaxation spectra were measured at spin-lock powers, $\omega_1/(2\pi)$, ranging from 0.5 to 15 kHz with a constant relaxation period T of 20 ms. A reference spectrum was measured with the pulse sequence that omitted the spin-lock period. All $R_{1\rho}$ dispersion spectra were Fourier-transformed, baseline-corrected, and peak-integrated by using TopSpin. The effective $R_{1\rho}$ relaxation rate was calculated as:

$$R_{1\rho} = -\frac{1}{T} \ln \left(\frac{I_{\text{SL}}}{I_0} \right), \quad (2)$$

where I_{SL} and I_0 are the signal intensities of the imino proton resonances of interest measured in the presence and absence of the spin-lock, respectively. Errors in the effective $R_{1\rho}$ relaxation rates were derived from the amplitude of the baseline noise. The resulting $R_{1\rho}$ dispersion profiles were fitted to the following equation for a two-state exchange process ($A \leftrightarrow B$) by using GLOVE (25):

$$R_{1\rho} = R_2^0 + \frac{p_A p_B \Delta\omega^2 k_{\text{ex}}}{\frac{\omega_{\text{Ae}}^2 \omega_{\text{Be}}^2}{\omega_c^2} + k_{\text{ex}}^2 - p_A p_B \Delta\omega^2 \left(1 + \frac{2k_{\text{ex}}^2 (p_A \omega_{\text{Ae}}^2 + p_B \omega_{\text{Be}}^2)}{\omega_{\text{Ae}}^2 \omega_{\text{Be}}^2 + \omega_c^2 k_{\text{ex}}^2} \right)} \quad (3)$$

$$\omega_{\text{Ae}}^2 = \omega_1^2$$

$$\omega_{\text{Be}}^2 = \Delta\omega^2 + \omega_1^2$$

$$\omega_c^2 = \Omega_A^2 + \omega_1^2$$

$$\Omega_A = -\frac{p_B k_{\text{ex}}^2 \Delta\omega}{k_{\text{ex}}^2 + \Delta\omega^2},$$

where R_2^0 is the intrinsic transverse relaxation rate (18). p_A and p_B ($p_A + p_B = 1$) are the populations of the major and minor states, respectively. k_{ex} denotes the chemical exchange rate (i.e. $k_{\text{open}} + k_{\text{close}}$). $\Delta\omega$ is the chemical shift difference between the two states. Note that, for fast exchange on the chemical shift timescale ($k_{\text{ex}} \gg \Delta\omega$), it is impossible to separate p_A and p_B from $p_A p_B \Delta\omega^2$ by this method.

Molecular dynamics

All molecular dynamics simulations were performed in explicit water using GROMACS version 5.1 (27–31). The Amber99sb force field with the ParmBSC0 nucleic acid parameters was used (32). The starting structures were ideal DNA double helices obtained from the program 3DNA (33,34) using the sequence given in Table 1. Force field parameters for 5mC were obtained from ref. (35). Force field parameters for 5hmC were constructed as described previously (36). The respective DNA was solvated in a triclinic box with SPC/E water. The temperature and salt concentration in the simulations were chosen to mimic the experimental

conditions: that is, the system charge was neutralized by addition of sodium (Na^+) and chloride (Cl^-) ions to a total salt concentration of 150 mM. After solvation, energy minimization was carried out by the steepest-descent algorithm (maximum force: $1000 \text{ kJ mol}^{-1} \text{ nm}^{-1}$). The system was equilibrated as follows: first, the system was simulated under the canonical (NVT) ensemble for a time of 100 ps; and second, 100 ps of NPT simulation was performed. In both steps, position restraints were applied on all DNA atoms. Pressure coupling was achieved by the Parrinello–Rahman algorithm (37) (1 bar, 2.0-ps coupling constant). Temperature coupling was achieved by the modified Berendsen (29) thermostat (303 K, 0.1-ps coupling constant). The integration time step was set to 2 fs. Bonds were constrained by the LINCS algorithm (31). The parameters for fast particle mesh Ewald electrostatics (38) were as follows: Coulomb cut-off distance, 0.8 nm; fast Fourier transform grid spacing, 0.08 nm; Lennard–Jones cut-off distance, 0.8 nm. The interpolation order was cubic. Initial velocities were assigned from the Maxwell distribution at 303 K. Production molecular dynamics (MD) ran for a total duration of 1 μs . All calculations were performed on Linux workstations using CUDA-compatible NVIDIA graphics processing units ($\sim 50 \text{ ns/day}$). Processing of the trajectories was carried out by tools integrated in GROMACS and 3DNA. Distributions were analysed by numpy and scipy using matplotlib for visualization (39–41).

RESULTS

Specific design of epigenetically distinct dsDNA constructs

We designed six dsDNA sequences using the deoxyribonucleic acids carrying 5mC and 5hmC modifications based on the sequence recognized by UHRF1 (7). The samples were designated as unmodified dsDNA (C/C), hemi-methylated dsDNA at nucleotide position 7 (5mC/C), full-methylated dsDNA (5mC/5mC), hemi-methylated dsDNA at nucleotide position 8 (C/5mC), hemi-hydroxymethylated dsDNA (5hmC/C), and hydroxymethyl-methylated dsDNA (5hmC/5mC) (Figure 1). These dsDNA sequences were identical except for the state of cytosine modification at nucleotide positions 7 and 8. Because 5mC/C is recognized by UHRF1, it is the most physiologically relevant sequence, as highlighted by the crystal structure of the UHRF1–dsDNA complex (7–9). C/C and 5hmC/C were control samples, each differing from 5mC/C at a single (un)modified base, to analyse the effects of epigenetic cytosine modification on dsDNA dynamics. C/5mC, 5mC/5mC and 5hmC/5mC were also control samples designed to examine the effect of methylation at nucleotide position 8.

Epigenetic cytosine modification does not affect the overall thermal stability of dsDNA

To examine whether the overall thermal stability of dsDNA differs depending on the type of cytosine modification, we measured the DNA melting temperature (T_m) of the six dsDNA samples by using a UV/Vis spectrometer. Interestingly, the T_m values of all the dsDNA samples were almost identical (Table 1), indicating that site-specific cyto-

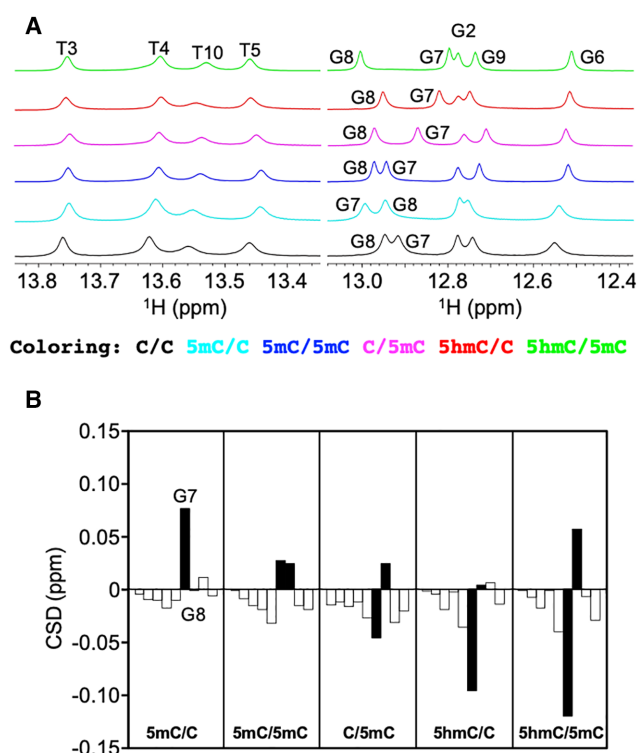


Figure 3. Dependence of imino proton chemical shifts on cytosine modification state. (A) Thymine and guanine imino proton regions in the 1D ^1H NMR spectra of C/C (black), 5mC/C (cyan), 5mC/5mC (blue), C/5mC (magenta), 5hmC/C (red) and 5hmC/5mC (green). (B) Imino proton chemical shift differences (CSD) of the bases from G2 to T10. To obtain CSD, the chemical shifts of the modified dsDNA samples were subtracted from those of the unmodified dsDNA sample. The CSD values of G7 and G8 are shown as filled bars.

sine modification does not significantly change the overall stability of dsDNA.

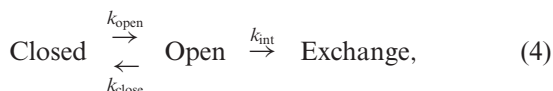
Cytosine modification alters the chemical environment of the guanine imino proton

The ^1H NMR spectra of the six dsDNA samples were measured, and the chemical shifts of the imino protons of the modified dsDNAs were compared with those of C/C to examine the effect of cytosine modification on dsDNA. Imino ^1H resonances are usually well dispersed and isolated from other resonances of dsDNA, thereby facilitating analysis. Although cytosine does not contain an imino proton, we expected that the chemical shift of the guanine imino proton that forms a WC base pair with the (modified) cytosine would sense the cytosine modification state. Indeed, we observed differences in the imino proton chemical shifts among the six dsDNA samples, and the most pronounced chemical shift changes were confined to bases adjacent to the modified cytosine (Figures 3A, B). Interestingly, the imino ^1H resonances of G7 in 5mC/C and 5mC/5mC, and G8 in 5mC/5mC, C/5mC and 5hmC/5mC were shifted downfield owing to cytosine methylation, whereas the imino ^1H resonances of G7 in 5hmC/C and 5hmC/5mC were shifted upfield owing to cytosine hydroxymethylation. It seems likely that the for-

mer downfield shift is due to the electron-donating nature of the methyl group, while the latter upfield shift is due to the electron-withdrawing nature of the hydroxymethyl group. Both the electron-donating (increased shielding) and electron-withdrawing (decreased shielding) effects should be sensed by the guanine imino proton owing to its proximity to the hydrogen-bonded cytosine nucleotide. However, the magnitude of the chemical shift differences cannot be explained solely by the modification-induced changes in electron density in the vicinity of the imino proton. Collectively, therefore, these results infer that cytosine modification alters the chemical environment of the imino proton, reflecting changes either in the local conformation of the dsDNA or in the structural fluctuations around that conformation (structural dynamics).

Cytosine modification affects the equilibrium of base-pair opening

Because cytosine modification is rather a small chemical change as compared with the size of the whole dsDNA molecule, we hypothesized that it might alter local structural flexibility. To test this hypothesis, we measured water-imino ^1H exchange rates ($k_{1\text{H}}$) by using the CLEANEX-PM experiment. If the stability of the hydrogen bond between the guanine imino proton and cytosine N3 is affected by the cytosine modification, the base-opening and closing rates (k_{open} and k_{close}) of the GC base pair will change, resulting in alteration of $k_{1\text{H}}$:



where k_{int} is the intrinsic ^1H exchange rate of the open state.

As shown by representative CLEANEX-PM profiles in Figure 4A, water-imino ^1H exchange was successfully detected. The intensity build-up was limited to values of $I(t)/I_0 < 0.1$ even at the longest ^1H exchange delay of 200 ms, indicating that the base-opening process is a very transient event (Figure 4A). Nevertheless, the $k_{1\text{H}}$ rates of thymine were faster than those of guanine (as expected from the number of hydrogen bonds in AT and GC base pairs), indicating that changes in local hydrogen bonding capability can indeed be discerned by CLEANEX-PM (Supplementary Figure S3A). Although the highest CLEANEX-PM build-up values of $I(t)/I_0$ for bases 6–9 around the modification sites were small, we observed differences in the $k_{1\text{H}}$ rates depending on the state of cytosine epigenetic modification (Figure 4B). Interestingly, this epigenetic effect on the $k_{1\text{H}}$ rates was not entirely local, but spread to both ends of the dsDNA, suggesting that a local base-opening process may be coupled to those of adjacent base pairs. Nevertheless, because both ends of dsDNA are intrinsically very flexible, a profound interpretation of the base-opening processes (i.e. the ^1H exchange rates) at base pairs distal from the modified cytosine was beyond the scope of the present study. Instead, we focused on nucleotide positions 6–9, which include the modification sites themselves and the flanking base pairs, in order to explore the short-range effects caused by cytosine modification.

Because no catalyst was added to accelerate k_{int} , the ^1H exchange process can be assumed to be governed by the EX2 regime, in which k_{close} is much faster than k_{int} (16):

$$k_{1\text{H}} = p_{\text{open}} k_{\text{int}}, \quad (5)$$

where p_{open} (i.e. $k_{\text{open}}/[k_{\text{open}}+k_{\text{close}}]$) denotes the population of the open state. p_{open} is usually further simplified as K_{op} (i.e. $k_{\text{open}}/k_{\text{close}}$) because k_{open} is much slower than k_{close} for ^1H exchange of the guanine imino proton in dsDNA (16):

$$k_{1\text{H}} = K_{\text{op}} k_{\text{int}} \quad (6)$$

^1H exchange of the guanine imino proton is base-catalysed by the cytosine residue that forms the GC base pair, and k_{int} is proportional to $10^{\text{pKa}(\text{N3};\text{C})-\text{pKa}(\text{N1};\text{G})}$, in which $\text{pKa}(\text{N3};\text{C})$ and $\text{pKa}(\text{N1};\text{G})$ represent the pKa of N3 in cytosine and N1 in guanine, respectively. Recently, Dai and co-workers determined the $\text{pKa}(\text{N3};\text{C})$ values of unmodified and epigenetically modified cytosine in dsDNA as 4.5 for cytosine, 4.4 for 5-methylcytosine, and 4.0 for 5-hydroxymethylcytosine (42). Using the $\text{pKa}(\text{N1};\text{G})$ value of 9.4 (43), $10^{\text{pKa}(\text{N3};\text{C})-\text{pKa}(\text{N1};\text{G})}$ is calculated to be 1.3×10^{-5} for C, 1.0×10^{-5} for 5mC, and 0.40×10^{-5} for 5hmC. Because of this difference in k_{int} among the unmodified and modified cytosine nucleotides, we normalized the obtained $k_{1\text{H}}$ rates of 5mC and 5hmC to that of C by multiplying by the factors 1.3 (i.e. $1.3 \times 10^{-5}/1.0 \times 10^{-5}$) and 3.3 (i.e. $1.3 \times 10^{-5}/0.40 \times 10^{-5}$), respectively. (Actually, ^1H exchange of the guanine imino proton base-paired with 5hmC was not observed.) This normalization enabled us to interpret the resultant ^1H exchange rate of each epigenetically modified GC base pair as the K_{op} value relative to that of the unmodified GC base pair (Equation 6; Figure 4C).

Strikingly, for nucleotide positions 6–9, the $k_{1\text{H}}$ rates of 5mC/C were the fastest among the dsDNA samples examined (Figure 4C). For G6 in particular, ^1H exchange was clearly detectable only for 5mC/C (and only slightly detectable for C/5mC). Moreover, the $k_{1\text{H}}$ rates of 5mC/C were >4-fold higher for G7 and G8 compared with those of C/C. For the more distal G9, the effect was less pronounced; nevertheless, a 2-fold faster ^1H exchange relative to unmodified dsDNA was still detected at this site. Similarly, hemi-methylation at position 8 (C/5mC) increased the $k_{1\text{H}}$ rate of G8 compared with C/C. As outlined above (Equations 5 and 6), these changes in normalized $k_{1\text{H}}$ value correspond to a relative increase in K_{op} or p_{open} . Therefore, methylation of cytosine specifically on one strand of dsDNA seems to stabilize a more open conformation of the GC base pair. Interestingly, the increase in the $k_{1\text{H}}$ rate of C/5mC was not as large as that of 5mC/C (Figure 4C), suggesting that the change in the $k_{1\text{H}}$ rate due to epigenetic modification depends on its position along the DNA sequence.

Surprisingly, the $k_{1\text{H}}$ rates at positions 6–9 of 5mC/5mC were comparable to those of C/C (Figure 4C). This indicated that, although methylation at position 7 (5mC/C) and position 8 (C/5mC) increased K_{op} at these particular base pairs (and in the vicinity in the case of hemi-methylation at

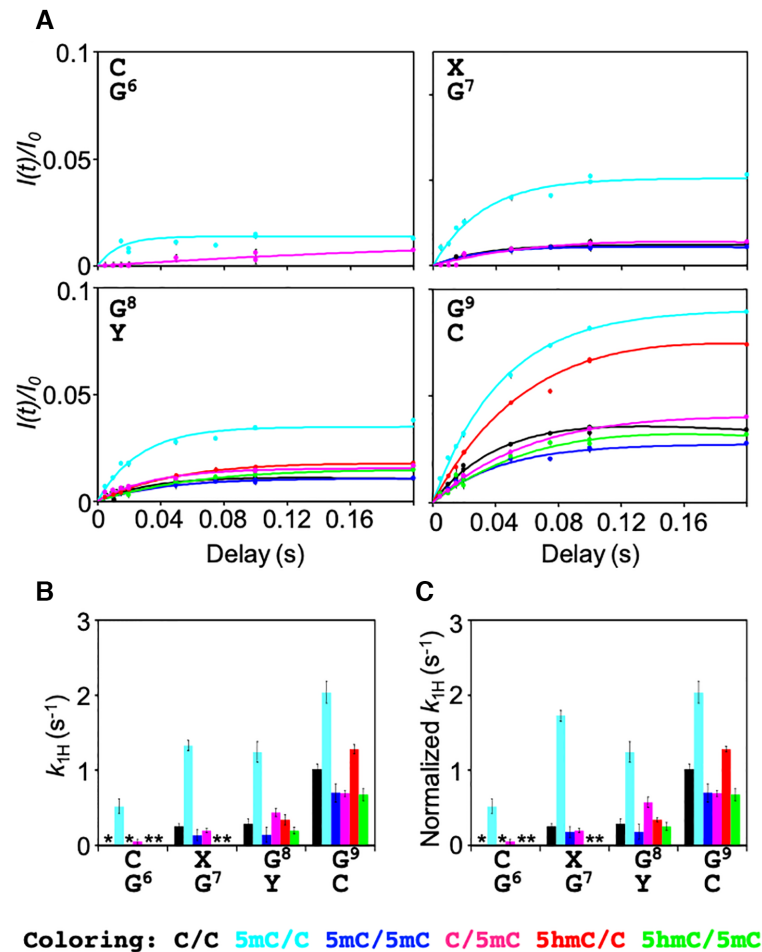


Figure 4. CLEANEX-PM experiments of imino protons in unmodified and cytosine-modified dsDNA. (A) CLEANEX-PM profiles for G6–G9 in C/C (black), 5mC/C (cyan), 5mC/5mC (blue), C/5mC (magenta), 5hmC/C (red), and 5hmC/5mC (green). Solid lines represent the best-fit lines to Equation 1. (B) ^1H exchange rates $k_{1\text{H}}$. Asterisks indicate the bases for which CLEANEX-PM signals were not detected. (C) Normalized $k_{1\text{H}}$ rates. The $k_{1\text{H}}$ rates of the guanine base-paired with 5mC were multiplied by 1.3.

position 7), this effect was reversed by further methylation at position 8 (5mC/5mC). Because hemi- and full methylation of cytosine at positions 7 and 8 had opposite effects on the $k_{1\text{H}}$ rates, the chemical nature of the methyl group cannot be the sole determinant of the change in the base-opening equilibrium. Instead, it is most likely that cytosine methylation changes the conformational properties of dsDNA in a more complex manner, which manifests as differences in the measured $k_{1\text{H}}$ rates.

In stark contrast, hydroxymethylation (5hmC/C) did not significantly alter $k_{1\text{H}}$ at any of the positions examined, indicating that K_{op} was not changed by the presence of an additional hydroxymethyl group at the modified cytosine (position 7; Figure 4C). Further methylation at position 8 (5hmC/5mC) again decreased $k_{1\text{H}}$, as observed for 5mC/5mC. Moreover, methylation at position 8 (5mC/5mC, C/5mC and 5hmC/5mC) decreased $k_{1\text{H}}$ of position 9 compared with those of C/C. Taken together, the data show that, among the cytosine modifications examined, hemi-methylation at position 7 has the most marked effect on the base-opening equilibrium.

Cytosine modification promotes the rate of local base-pair closing

The CLEANEX-PM experiments provided thermodynamic information on the base-opening and -closing process in the form of K_{op} , but did not provide any kinetic information on the process, namely, the k_{open} or k_{close} rates. To characterize the effect of epigenetic cytosine modifications on the kinetics of dsDNA, we used a similar strategy but conducted on-resonance ^1H $R_{1\rho}$ relaxation dispersion experiments on the base-paired (guanine and thymine) imino proton. This type of experiment can determine conformational exchange rates on the microsecond-to-millisecond timescale (see Figure 2 for detailed information about the NMR pulse sequence used).

The resulting imino proton $R_{1\rho}$ relaxation dispersion profiles indicated the presence of microsecond-to-millisecond conformational exchange in a manner dependent on the position of the respective base pairs in the DNA sequence (Figure 5A). Moreover, the contribution from conformational exchange (R_{ex}) to the effective $R_{1\rho}$ relaxation rates was larger for thymine than for guanine. This is in fine agree-

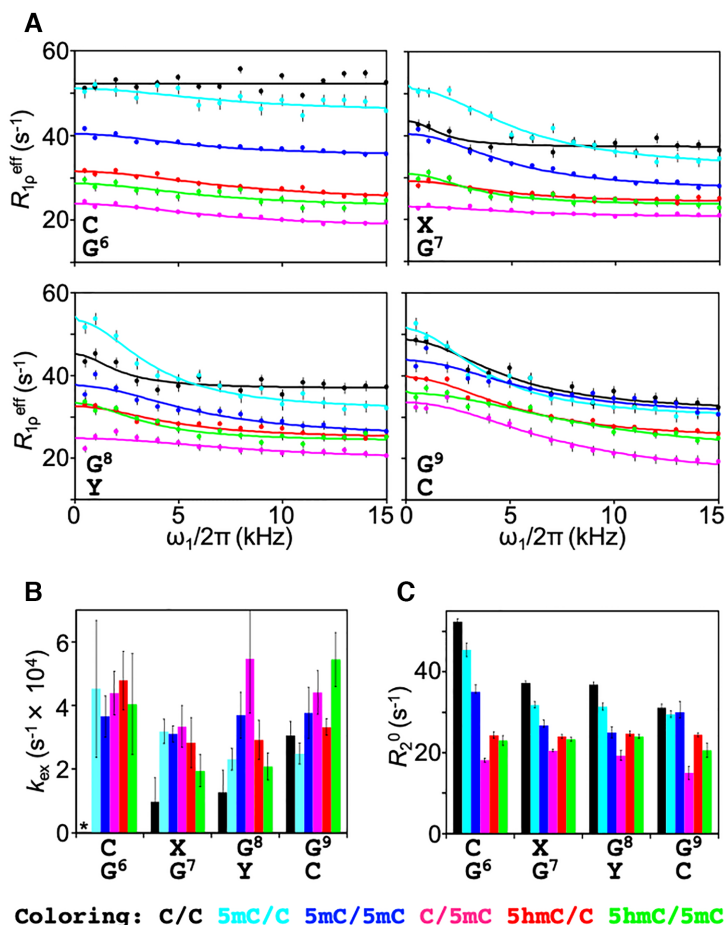


Figure 5. On-resonance R_{1p} dispersion experiments of imino protons in unmodified and cytosine-modified dsDNA. (A) On-resonance R_{1p} dispersion profiles for G6–G9 in C/C (black), 5mC/C (cyan), 5mC/5mC (blue), C/5mC (magenta), 5hmC/C (red) and 5hmC/5mC (green). Solid lines represent the best-fit lines to Equation 3. (B) The exchange rate, k_{ex} , for G6–G9. No R_{ex} was observed for G6 in C/C, as indicated by the asterisk. (C) Intrinsic transverse relaxation rate, R_2^0 , for G6–G9.

ment with the expectation that the AT base pair will be less stable than the GC base pair owing to the smaller number of hydrogen bonds and thereby a comparably higher population of broken hydrogen bonds for AT, leading to a larger R_{ex} (Supplementary Figure S3B).

By fitting the R_{1p} relaxation dispersion data to Equation 3, we obtained the exchange rate, k_{ex} (i.e. $k_{open} + k_{close}$), of the base-opening process (Figure 5B) (18). Because the resultant k_{ex} rates were very fast, on the order of 10^4 s⁻¹, the conformational exchange of the dsDNA samples was categorized as fast exchange on the chemical shift timescale, and thus, k_{ex} could not be separated into k_{open} and k_{close} under these conditions. Owing to the high tendency of dsDNA to form WC base pairs (i.e., the high thermodynamic stability of the closed double-stranded form of DNA), however, p_{open} of a given base pair should be very small at our experimental temperature of 303 K. In other words, k_{close} (i.e. $p_{close}k_{ex}$) would be much faster than k_{open} (i.e. $p_{open}k_{ex}$). In such a situation, k_{ex} can be assumed to be almost equal to k_{close} , meaning that a small change in k_{open} would not be reflected in k_{ex} .

As in the CLEANEX-PM experiments, we focused again on the central base positions 6–9. For G6 in C/C, no R_{1p}

relaxation dispersion ($R_{ex} \approx 0$) was observed, and thus k_{ex} could not be derived. However, the samples containing cytosine modifications did show relaxation dispersion (Figure 5A): for example, the k_{ex} ($\approx k_{close}$) rates for G6 among the different samples were identical to each other within experimental error (Figure 5B), although ¹H exchange for G6 had been observed only for 5mC/C and C/5mC (Figure 4). These results suggest that, although hemi-methylation at position 7 (5mC/C) and position 8 (C/5mC) increased k_{open} for G6 (observed by CLEANEX-PM), the tendency to return to the closed dsDNA structure was very similar among all dsDNA samples as reflected by the k_{ex} ($\approx k_{close}$) values (observed by relaxation dispersion).

For nucleotide positions 7–9, all k_{ex} rates were remarkably similar to each other with some exceptions (Figure 5B). First, G7 and G8 of unmodified dsDNA (C/C) showed lower k_{ex} values as compared with all modified forms of dsDNA examined. Second, G8 of C/5mC and G9 of 5hmC/5mC showed higher k_{ex} rates than the corresponding positions of the other dsDNA samples. Together, these results indicate that both methylation and hydroxymethylation can increase k_{close} ; in other words, these modifications can shorten the lifetime of the open state ($1/k_{close} \approx$

$1/k_{\text{ex}}$). By CLEANEX-PM, however, we also observed significantly increased $k_{1\text{H}}$ rates for G7 and G8 of 5mC/C and a slightly increased $k_{1\text{H}}$ rate for G8 of C/5mC owing to hemi-methylation at positions 7 and 8, respectively. These results suggest that hemi-methylation at position 7 (5mC/C) increases both k_{open} and k_{close} for the base pairs flanking the methylation site (G6, G7, G8) whereas hemi-methylation at position 8 (C/5mC) increases both k_{open} and k_{close} only for the base pair itself (G8).

Interestingly, R_2^0 rates were also affected by the state of cytosine modification (Figure 5C). There are a few possible reasons to account for the variation observed in R_2^0 rates. First, the difference in the number of ^1H nuclei in the dsDNA samples, which is related to the cross-relaxation rate, might explain variations in R_2^0 . However, this seems unlikely because C/C, which has fewer ^1H nuclei relative to the other samples, showed the fastest R_2^0 rates for all imino protons examined. Second, fast R_2^0 rates may originate from ^1H exchange, but C/C, which showed the fastest R_2^0 rates, showed slower $k_{1\text{H}}$ rates (Figure 4); thus, ^1H exchange is not a cogent explanation. In addition to these two possibilities, we note that the intrinsic transverse relaxation rate R_2^0 is determined by both the overall rotational correlation time and the local flexibility on the picosecond-to-nanosecond timescale. In particular, the overall rotational correlation time is determined by the tertiary structure of a biomolecule. Therefore, the observed variations in R_2^0 infer that the tertiary structure and/or local flexibility differed among the six dsDNA samples.

Cytosine modification alters local base-pair motion

Because neither tertiary structure nor local flexibility on the picosecond-to-nanosecond timescale can be captured by CLEANEX-PM or on-resonance imino ^1H $R_{1\rho}$ relaxation dispersion experiments, we carried out 1- μs MD simulations for each of the six dsDNA samples to examine their structure and flexibility (i.e. fluctuations around an average structure) and try to explain the variations in R_2^0 .

First, we examined the radius of gyration of the dsDNA samples because this parameter affects the overall rotational correlation time of a molecule. However, the radius of gyration did not display a significant dependency on the state of cytosine modification (Supplementary Figure S4), suggesting that the differences in R_2^0 are not caused by variations in overall rotational correlation time. Next, we compared the local flexibility of the dsDNA samples by analysing the MD trajectories using the program 3DNA (34), which can extract intra-base-pair, inter-base-pair, and base-pair-axis parameters. These parameters were extracted from all the dsDNA atomic coordinates sampled every 10 ps over a 1- μs MD run. For each parameter set, we calculated the averages and standard deviations (SD) to compare the distribution of parameters among the different dsDNA samples.

For most parameters, the derived average values were almost identical at all base pairs among all dsDNA samples, with the exception of the C6:G base pair (Figure 6 and Supplementary Figure S5). At this position, the average angles of roll, tilt, inclination, tip, and buckle were found to dif-

fer between unmodified and modified dsDNA except for C/5mC (Figure 6A), suggesting that these angles (at position 6) depend on the cytosine modification state of position 7 (among these samples only C/C and C/5mC are unmodified at position 7). The roll and tilt angles are defined in terms of local stacking configurations at successive base pairs, and are sensitive to, for example, the degree of DNA curvature. The inclination and tip angles reflect the orientation of the base-pair planes with respect to the helical axis, and the buckle angle measures deviations in planarity in a base pair. The average roll angles of C/C and C/5mC were larger than those of the other modified dsDNA samples, indicating that the C6:G base pairs in C/C and C/5mC are curved toward the major groove as compared with the other modified dsDNA samples. In addition, C/C and C/5mC showed negative values in the average tilt angle, whereas the other modified dsDNA samples displayed positive values, indicating that C/C and C/5mC are on average slightly curved at C6 in the opposite direction relative to the other modified dsDNA samples. Similar differences were observed for the average inclination and tip parameters, likewise depending on the modification state at position 7. Although the differences in the average angular parameters were only a few angular degrees, such changes in various angular parameters might have a significant effect on the local structure when taken together. Additional factors essentially untraceable by molecular dynamics, such as differences in tautomeric reactivity (44) occurring either classically over an activation energy-barrier or quantum-mechanically by tunnelling of hydrogen atoms along hydrogen bonds (45–47) that accompany changes in base-pair geometry may further contribute to the epigenetically distinct dynamics. Over longer timescales, these small differences may build up into larger distinctions between the different dsDNA samples that ultimately give rise to different dynamics, as experimentally observed by solution NMR.

In terms of the SD of the derived angular parameter distributions, which approximately represents fluctuations around the average local structure of the dsDNA over the course of the MD trajectories, the SD in the roll, tilt, inclination, tip, and buckle angles differed especially between the unmodified and modified dsDNA samples. Similar to the average values, the differences were pronounced for the C6:G base pair, where the unmodified dsDNA displayed larger SD values (i.e. broader distributions) as compared with the modified dsDNA (Figure 6B). In a comparison of the SD values of each parameter along the sequence, the SD values of tilt and roll angles from the C6:G to G9:C base pairs were smaller in the modified dsDNA than in the unmodified dsDNA (Figure 6C), indicating less flexibility at these base pairs in the modified dsDNA sample. In addition, all dsDNA samples containing methylation at position 8 (5mC/5mC, C/5mC and 5hmC/5mC) showed smaller SD values of tilt and roll angles of G8:C than the dsDNA samples unmethylated at position 8 (C/C, 5mC/C and 5hmC/C). Taken together, therefore, these results suggest that cytosine methylation and hydroxymethylation can decrease local flexibility in the dsDNA structure.

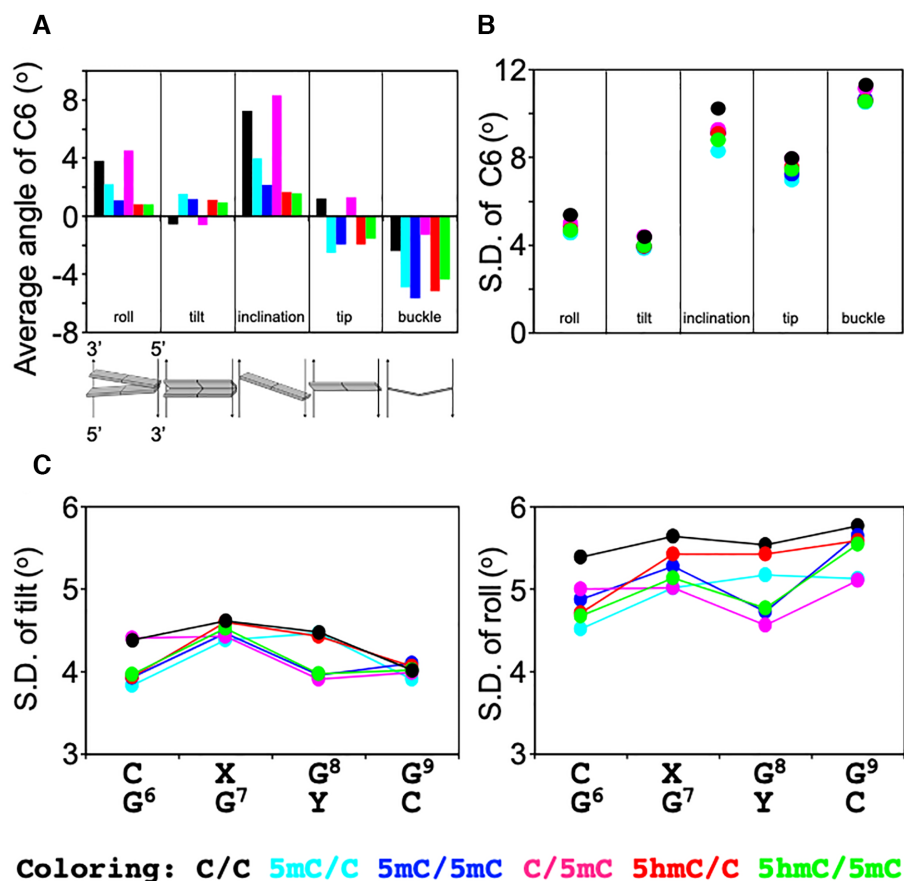


Figure 6. DNA helical parameters from MD simulations. (A) Average values of the roll, tilt, inclination, tip, and buckle angles for C6 in C/C (black), 5mC/C (cyan), 5mC/5mC (blue), C/5mC (magenta), 5hmC/C (red), and 5hmC/5mC (green) dsDNA. (B) Standard deviation of the tilt, roll, inclination, tip, and buckle angles for the C6:G base pair. (C) Standard deviation of the tilt and roll angles for C6-G9.

DISCUSSION

We have shown that site-specific cytosine modification alters local dynamics rather than the overall thermal stability of DNA. We have also demonstrated that the dsDNA dynamics altered by cytosine modification can be semi-quantitatively analysed by combining data from CLEANEX-PM and imino ^1H $R_{1\rho}$ relaxation dispersion experiments in a complementary way. Using this approach, we characterized the base-opening (k_{open}) and base-closing (k_{close}) processes and found that only hemi-methylation (5mC/C) at a single nucleotide site (position 7 of our sample) significantly increased both the k_{open} and k_{close} rates. By contrast, the other cytosine modifications examined (5mC/5mC, C/5mC, 5hmC/C and 5mC/5hmC) increased k_{close} but did not change k_{open} significantly. Because all the $\text{p}K_{\text{a}}(\text{N3};\text{C})$ values of modified cytosine N3 are lower than that of unmodified cytosine N3 (4.5 for cytosine, 4.4 for 5-methylcytosine, and 4.0 for 5-hydroxymethylcytosine), the N3 atoms of modified cytosine should be less protonated at neutral pH as compared with unmodified cytosine. In turn, this should be unfavourable for formation of the HG base pair because protonation of cytosine N3 is required to stabilize an HG GC base pair. Therefore, it seems likely that the increase in k_{close} represents either destabilization of an

HG base pair or a shorter lifetime of an HG base pair (proportional to $1/k_{\text{close}}$).

Previously, it was reported that the k_{ex} of WC to HG conformational exchange is on the order of 10^3 s^{-1} by ^{13}C and ^{15}N $R_{1\rho}$ relaxation dispersion (13,19–22). Most of those studies were conducted under low pH conditions because lower pH enhances the protonation of cytosine N3, which stabilizes the HG GC base pair. The k_{ex} rates obtained in the present study were on the order of 10^4 s^{-1} , which is faster by one order of magnitude than the previously determined k_{ex} rates. This difference arises because the dsDNA dynamics were analysed at neutral pH, at which the closing kinetics (k_{close}) should be much faster relative to lower pH. Moreover, the ^1H $R_{1\rho}$ relaxation dispersion experiment used in the present study can detect faster motions as compared with the previously used ^{13}C and ^{15}N $R_{1\rho}$ relaxation dispersion experiments. An alternative possibility is that the HG GC base pair, which has two inter-base hydrogen bonds, is not perfectly formed at neutral pH, and our ^1H $R_{1\rho}$ relaxation dispersion measurements detected conformational exchange between the WC base pair and such a state.

In the present study, the increase in the k_{open} rate of 5mC/C (Figure 4C) was particularly striking, and indicates that the WC base pairs of C6:G to G9:C are destabilized by cytosine methylation. Destabilization of the WC base pair

cannot be explained simply by the lower $pK_a(N3;C)$ value. The MD simulations indicated that the dynamics of local curvature and the flexibility of modified dsDNA differs from that of unmodified dsDNA. Therefore, local changes in structure and dynamics of dsDNA induced by cytosine modification may cause destabilization of the WC base pair. In addition, the destabilization of the WC base pair appears to also depend on the DNA sequence as we obtained the different k_{open} rates between 5mC/C and C/5mC. This is consistent with a previous study by Alvey *et al.* who also reported that dsDNA stability depends on the position of modification (21).

UHRF1 binds selectively to hemi-methylated dsDNA (5mC/C), but not to unmethylated (C/C) or fully methylated (5mC/5mC) dsDNA (7). As shown in the present study, the global stability of the dsDNA samples was identical irrespective of the cytosine modification state of dsDNA; however, local dynamics was altered by cytosine modification and especially by hemi-methylation of dsDNA (Figure 5). In particular, the k_{1H} rates of 5mC/C were markedly different from those of all the other dsDNA samples, including unmodified and fully methylated dsDNA (Figure 4). Thus, UHRF1 may exploit such differences in local dynamics to selectively bind to and thereby recognize hemi-methylated DNA.

During preparation of this manuscript, it was reported that UHRF1 binds to fully carboxyl-methylated dsDNA (5caC/5caC) with high affinity (48). Therefore, we additionally prepared this type of dsDNA and analysed its structural dynamics by the CLEANEX-PM and imino 1H $R_{1\rho}$ relaxation dispersion experiments. As a result, the imino 1H resonances of G7 and G8 in 5caC/5caC were shifted up-field, which is reasonable because the carboxyl group has the electron-withdrawing nature (Supplementary Figures S6B, C). Unexpectedly, 1H exchange and $R_{1\rho}$ relaxation dispersion of G7 and G8 were not observed (Supplementary Figure S6D), indicating that 5caC/5caC is more stable than C/C. In fine accordance with these NMR data, we observed that the melting temperature of 5caC/5caC was higher than that of C/C ($58.84 \pm 0.02^\circ C$ versus $55.37 \pm 0.09^\circ C$). Therefore, UHRF1 may recognize 5caC/5caC in a different manner from 5mC/C. Because structural data of a UHRF1–5caC/5caC complex are not yet available, it however remains elusive what the exact differences are that UHRF1 exploits to discriminate fully carboxyl-methylated from hemi-methylated DNA.

5-Methylcytosine and 5-hydroxymethylcytosine, which are found in most plants, animals, and fungi, have profound effects on genomic stability, gene expression, and development (1,2). Local changes in conformation and dynamics of dsDNA due to cytosine modification are therefore likely to play an important role in all these phenomena. In a physiological (nuclear) context, intriguingly it has been reported that HG base pairs are stabilized under conditions of molecular crowding, whereas WC base pairs are destabilized (49). This infers that, relative to *in vitro* experiments, dsDNA dynamics may be amplified in living cells, which are highly crowded environments. It seems that DNA-binding proteins including UHRF1 might exploit such dsDNA dynamics in their search for specific target DNA sequences.

For mice it has been shown that in addition to hemi-methylated dsDNA, hemi-hydroxymethylated/hemi-methylated dsDNA also exists at high levels in various tissues including kidney, heart, liver, and brain (cerebellum) (50). In this context, it is highly intriguing that we observed that 5mC/5mC and 5hmC/5mC exhibited different structural dynamics, in particular: k_{ex} of G7 and G8 (Figure 5B). Such differences in local dynamics might be exploited by proteins (by modulating the association or dissociation rates of binding to these specifically modified types of DNA), such as methyl-CpG-binding domain proteins or TET, to discern one epigenetic state from the other. Therefore, in addition to the specific case of UHRF1, it will be interesting to study structural dynamics of epigenetically (un)modified dsDNA in the presence of these epigenetic mark reading proteins in the future.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

This work was supported by funding from Suntory Holdings.

This research was partially supported by Platform Project for Supporting Drug Discovery and Life Science Research (Basis for Supporting Innovative Drug Discovery and Life Science Research (BINDS)) from AMED under Grant Number JP20am0101070 (support number 2555).

FUNDING

JSPS KAKENHI [15H01653 to A.F. and 18H04551 to K.S.]. Funding for open access charge: Ministry of Education, Culture, Sports, Science and Technology; Japan Society for the Promotion of Science.

Conflict of interest statement. None declared.

REFERENCES

- Law, J.A. and Jacobsen, S.E. (2010) Establishing, maintaining and modifying DNA methylation patterns in plants and animals. *Nat. Rev. Genet.*, **11**, 204–220.
- Wu, H. and Zhang, Y. (2014) Reversing DNA methylation: mechanisms, genomics, and biological functions. *Cell*, **156**, 45–68.
- Bestor, T., Laudano, A., Mattaliano, R. and Ingram, V. (1988) Cloning and sequencing of a cDNA encoding DNA methyltransferase of mouse cells. The carboxyl-terminal domain of the mammalian enzyme is related to bacterial restriction methyltransferases. *J. Mol. Biol.*, **203**, 971–983.
- Okano, M., Bell, D.W., Haber, D.A. and Li, E. (1999) DNA methyltransferases Dnmt3a and Dnmt3b are essential for *de novo* methylation and mammalian development. *Cell*, **99**, 247–257.
- Ito, S., Shen, L., Dai, Q., Wu, S.C., Collins, L.B., Swenberg, J.A., He, C. and Zhang, Y. (2011) Tet proteins can convert 5-methylcytosine to 5-formylcytosine and 5-carboxylcytosine. *Science*, **333**, 1300–1303.
- Hu, L., Lu, J., Cheng, J., Rao, Q., Li, Z., Hou, H., Lou, Z., Zhang, L., Li, W., Gong, W. *et al.* (2015) Structural insight into substrate preference for TET-mediated oxidation. *Nature*, **527**, 118–122.
- Arita, K., Ariyoshi, M., Tochio, H., Nakamura, Y. and Shirakawa, M. (2008) Recognition of hemi-methylated DNA by the SRA protein UHRF1 by a base-flipping mechanism. *Nature*, **455**, 818–821.
- Avvakumov, G.V., Walker, J.R., Xue, S., Li, Y., Duan, S., Bronner, C., Arrowsmith, C.H. and Dhe-Paganon, S. (2008) Structural basis for recognition of hemi-methylated DNA by the SRA domain of human UHRF1. *Nature*, **455**, 822–825.

9. Hashimoto, H., Horton, J.R., Zhang, X., Bostick, M., Jacobsen, S.E. and Cheng, X. (2008) The SRA domain of UHRF1 flips 5-methylcytosine out of the DNA helix. *Nature*, **455**, 826–829.
10. Derreumaux, S., Chaoui, M., Tevanian, G. and Fermandjian, S. (2001) Impact of CpG methylation on structure, dynamics and solvation of cAMP DNA responsive element. *Nucleic Acids Res.*, **29**, 2314–2326.
11. Nathan, D. and Crothers, D.M. (2002) Bending and flexibility of methylated and unmethylated EcoRI DNA. *J. Mol. Biol.*, **316**, 7–17.
12. Perez, A., Castellazzi, C.L., Battistini, F., Collinet, K., Flores, O., Deniz, O., Ruiz, M.L., Torrents, D., Eritja, R., Soler-Lopez, M. *et al.* (2012) Impact of methylation on the physical properties of DNA. *Biophys J.*, **102**, 2140–2148.
13. Nikolova, E.N., Kim, E., Wise, A.A., O'Brien, P.J., Andricioaei, I. and Al-Hashimi, H.M. (2011) Transient Hoogsteen base pairs in canonical duplex DNA. *Nature*, **470**, 498–502.
14. Gueron, M., Kochoyan, M. and Leroy, J.L. (1987) A single mode of DNA base-pair opening drives imino proton exchange. *Nature*, **328**, 89–92.
15. Folta-Stogniew, E. and Russu, I.M. (1996) Base-catalysis of imino proton exchange in DNA: effects of catalyst upon DNA structure and dynamics. *Biochemistry*, **35**, 8439–8449.
16. Szulik, M.W., Pallan, P.S., Nocek, B., Voehler, M., Banerjee, S., Brooks, S., Joachimiak, A., Egli, M., Eichman, B.F. and Stone, M.P. (2015) Differential stabilities and sequence-dependent base pair opening dynamics of Watson-Crick base pairs with 5-hydroxymethylcytosine, 5-formylcytosine, or 5-carboxylcytosine. *Biochemistry*, **54**, 1294–1305.
17. Korzhnev, D.M., Orekhov, V.Y. and Kay, L.E. (2005) Off-resonance $R_{1\rho}$ NMR studies of exchange dynamics in proteins with low spin-lock fields: an application to a Fyn SH3 domain. *J. Am. Chem. Soc.*, **127**, 713–721.
18. Palmer, A.G. 3rd and Massi, F. (2006) Characterization of the dynamics of biomacromolecules using rotating-frame spin relaxation NMR spectroscopy. *Chem. Rev.*, **106**, 1700–1719.
19. Nikolova, E.N., Gottardo, F.L. and Al-Hashimi, H.M. (2012) Probing transient Hoogsteen hydrogen bonds in canonical duplex DNA using NMR relaxation dispersion and single-atom substitution. *J. Am. Chem. Soc.*, **134**, 3667–3670.
20. Nikolova, E.N., Goh, G.B., Brooks, C.L. 3rd and Al-Hashimi, H.M. (2013) Characterizing the protonation state of cytosine in transient G.C Hoogsteen base pairs in duplex DNA. *J. Am. Chem. Soc.*, **135**, 6766–6769.
21. Alvey, H.S., Gottardo, F.L., Nikolova, E.N. and Al-Hashimi, H.M. (2014) Widespread transient Hoogsteen base pairs in canonical duplex DNA with variable energetics. *Nat. Commun.*, **5**, 4786.
22. Kimsey, I.J., Petzold, K., Sathyamoorthy, B., Stein, Z.W. and Al-Hashimi, H.M. (2015) Visualizing transient Watson-Crick-like mispairs in DNA and RNA duplexes. *Nature*, **519**, 315–320.
23. Delaglio, F., Grzesiek, S., Vuister, G.W., Zhu, G., Pfeifer, J. and Bax, A. (1995) NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR*, **6**, 277–293.
24. Hwang, T.L., van Zijl, P.C. and Mori, S. (1998) Accurate quantitation of water-amide proton exchange rates using the phase-modulated CLEAN chemical EXchange (CLEANEX-PM) approach with a Fast-HSQC (FHSQC) detection scheme. *J. Biomol. NMR*, **11**, 221–226.
25. Sugase, K., Konuma, T., Lansing, J.C. and Wright, P.E. (2013) Fast and accurate fitting of relaxation dispersion data using the flexible software package GLOVE. *J. Biomol. NMR*, **56**, 275–283.
26. Schlagnitweit, J., Steiner, E., Karlsson, H. and Petzold, K. (2018) Efficient detection of structure and dynamics in unlabeled RNAs: The SELOPE approach. *Chemistry*, **24**, 6067–6070.
27. Lindahl, E., Hess, B. and Van Der Spoel, D. (2001) GROMACS 3.0: a package for molecular simulation and trajectory analysis. *J. Mol. Modeling*, **7**, 306–317.
28. Berendsen, H.J., Van Der Spoel, D. and Van Drunen, R. (1995) GROMACS: a message-passing parallel molecular dynamics implementation. *Comput. Phys. Commun.*, **91**, 43–56.
29. Bussi, G., Donadio, D. and Parrinello, M. (2007) Canonical sampling through velocity rescaling. *J. Chem. Phys.*, **126**, 014101–014107.
30. Miyamoto, S. and Kollman, P.A. (1992) SETTLE: an analytical version of the SHAKE and RATTLE algorithm for rigid water models. *J. Comput. Chem.*, **13**, 952–962.
31. Hess, B. (1997) LINCS: a linear constraint solver for molecular simulations. *J. Comput. Chem.*, **18**, 1463–1472.
32. Hornak, V., Abel, R., Okur, A., Strockbine, B., Roitberg, A. and Simmerling, C. (2006) Comparison of multiple Amber force fields and development of improved protein backbone parameters. *Proteins*, **65**, 712–725.
33. Lu, X.-J. and Olson, W.K. (2003) 3DNA: a software package for the analysis, rebuilding and visualization of three-dimensional nucleic acid structures. *Nucleic Acids Res.*, **31**, 5108–5121.
34. Lu, X.-J. and Olson, W.K. (2008) 3DNA: a versatile, integrated software system for the analysis, rebuilding and visualization of three-dimensional nucleic-acid structures. *Nat. Protoc.*, **3**, 1213–1227.
35. Lankas, F., Cheatham, T.E. 3rd, Spacková, N., Hobza, P., Langowski, J. and Sponer, J. (2002) Critical effect of the N2 amino group on structure, dynamics, and elasticity of DNA polypurine tracts. *Biophys. J.*, **82**, 2592–2609.
36. Frauer, C., Hoffmann, T., Bultmann, S., Casa, V., Cardoso, M.C., Antes, I. and Leonhardt, H. (2011) Recognition of 5-hydroxymethylcytosine by the Uhrf1 SRA domain. *PLoS One*, **6**, e21306.
37. Parrinello, M. and Rahman, A. (1981) Polymorphic transitions in single crystals: a new molecular dynamics method. *J. Appl. Phys.*, **52**, 7182–7190.
38. Essmann, U., Perera, L. and Berkowitz, M.L. (1995) A smooth particle mesh Ewald method. *J. Chem. Phys.*, **103**, 8577–8593.
39. Hunter, J.D. (2007) Matplotlib: a 2D graphics environment. *Comput. Sci. Eng.*, **9**, 90–95.
40. Walt, S.V.d., Colbert, S.C. and Varoquaux, G. (2011) The NumPy array: a structure for efficient numerical computation. *Comput. Sci. Eng.*, **13**, 22–30.
41. Jones, E., Oliphant, T. and Peterson, P. (2014) SciPy: open source scientific tools for Python.
42. Dai, Q., Sanstead, P.J., Peng, C.S., Han, D., He, C. and Tokmakoff, A. (2016) Weakened N3 hydrogen bonding by 5-formylcytosine and 5-carboxylcytosine reduces their base-pairing stability. *ACS Chem. Biol.*, **11**, 470–477.
43. Nonin, S., Leroy, J.L. and Gueron, M. (1996) Acid-induced exchange of the imino proton in G.C pairs. *Nucleic Acids Res.*, **24**, 586–595.
44. Watson, J.D. and Crick, F.H. (1953) Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid. *Nature*, **171**, 737–738.
45. Brovarets, O.O., Oliynyk, T.A. and Hovorun, D.M. (2019) Novel tautomerisation mechanisms of the biologically important conformers of the reverse lowdin, Hoogsteen, and reverse Hoogsteen G*.C* DNA base pairs via proton transfer: a quantum-mechanical survey. *Front. Chem.*, **7**, 597.
46. Parker, B.R. and Everv, J.V. (1971) Quantum tunnelling in DNA. *Chem. Phys. Lett.*, **8**, 94–99.
47. Lowdin, P.O. (1963) Proton tunneling in DNA and its biological implications. *Rev. Mod. Phys.*, **35**, 724–732.
48. Schneider, M., Trummer, C., Stengl, A., Zhang, P., Szwagierczak, A., Cardoso, M.C., Leonhardt, H., Bauer, C. and Antes, I. (2020) Systematic analysis of the binding behaviour of UHRF1 towards different methyl- and carboxylcytosine modification patterns at CpG dyads. *PLoS One*, **15**, e0229144.
49. Miyoshi, D., Nakamura, K., Tateishi-Karimata, H., Ohmichi, T. and Sugimoto, N. (2009) Hydration of Watson-Crick base pairs and dehydration of Hoogsteen base pairs inducing structural polymorphism under molecular crowding conditions. *J. Am. Chem. Soc.*, **131**, 3522–3531.
50. Song, C.X., Diao, J., Brunger, A.T. and Quake, S.R. (2016) Simultaneous single-molecule epigenetic imaging of DNA methylation and hydroxymethylation. *Proc. Natl. Acad. Sci. U.S.A.*, **113**, 4338–4343.