Supporting Information

Probing Transient Hoogsteen Hydrogen Bonds in Canonical Duplex DNA Using NMR Relaxation Dispersion and Single-Atom Substitution

Evgenia N. Nikolova, Federico L. Gottardo, and Hashim M. Al-Hashimi*

Department of Chemistry and Biophysics, University of Michigan, Ann Arbor, MI, 48109

Supplementary Methods

Sample preparation

Gel-filtration grade unmodified DNA oligonucleotides were purchased from IDT, Inc., while c7A- and c7G-modified oligos were purchased from Midland Certified, Inc. DNA oligos were resuspended in the appropriate NMR buffer (15 mM Sodium phosphate, 25 mM NaCl, 0.1 mM EDTA, pH 6.8) at ~ 0.2 mM concentration and their exact concentration measured by UV absorbance at 260 nm using extinction coefficients provided by the manufacturer. Unlabeled DNA duplexes, A_6-DNA^{c7A16} and A_6-DNA^{c7G10}, were annealed by mixing an equimolar ratio of the complementary strands, heating for 2 min at 95°C and gradual cooling (~ 30 min) at room temperature. DNA preparations were further washed 3X in NMR buffer by microcentrifugation using an Amicon Ultra-4 centrifugal filter (3 kDa cutoff), concentrated to a volume of ~ 250 µl (~ 3 - 4 mM) for NMR studies and supplied with 10% D_2O. The sample pH was adjusted by titrating dilute HCl (or NaOH) directly into the sample, in necessary, and monitoring the pH with a pH meter (Thermo Scientific).

The semi-labeled $^{13}C/^{15}N$-labeled DNA duplex was prepared by annealing an unlabeled modified strand to a labeled unmodified strand. The fully $^{13}C/^{15}N$-labeled, unmodified DNA duplexes were prepared by annealing two labeled strands together. All labeled single strands were synthesized *in vitro* by the method of Zimmer *et al.*,[3] using a template hairpin DNA (IDT, Inc.) with the same hairpin sequence, Klenow fragment DNA polymerase and NEB2 buffer (NEB, Inc.), and uniformly $^{13}C/^{15}N$-labeled dNTPs (Isotec, Sigma-Aldrich). The single-stranded DNA product was purified by a 20% denaturing PAGE, isolated by passive elution from crushed gels and desalted on a C18 reverse-phase column (Sep-pak, Waters). The oligo was lyophilized and resuspended in NMR buffer. The semi-labeled DNA samples were prepared by titrating the

unlabeled strand directly into an NMR tube containing the $^{13}C/^{15}N$-labeled strand and monitoring the disappearance of single-stranded DNA peaks using quick HSQC experiments. The fully labeled samples were annealed in a similar fashion. Sample pH was adjusted as described above.

NMR spectroscopy

All NMR experiments were performed on a Bruker Avance 600 MHz NMR spectrometer equipped with a 5mm triple-resonance cryogenic probe. Unlabeled DNA duplexes were assigned using a conventional 2D $^{1}H,^{1}H$ NOESY experiment (175 ms mixing time, 16 scans, 768 indirect points) using excitation sculpting with gradients for water suppression, in 10% D_2O and at 26°C. 1D imino proton spectra (16 scans) were obtained at variable temperature ranging from 5 °C to 35 °C using a jump-return ^{1}H experiment. 2D $^{1}H,^{13}C$ HSQC and $^{1}H,^{15}N$ HSQC and correlation spectra were collected by appropriately modifying standard Bruker pulse sequences (32 scans, 256/160 indirect points for aromatic/ribose regions in unlabeled DNA; 8 scans, 300 indirect points for imino regions in labeled DNA).

Rotating frame ($R_{1\rho}$) relaxation dispersion profiles were measured at a single field (14.1 T) by using or adapting previously published ^{13}C and ^{15}N experiments that use a 1D acquisition scheme to inspect individual nuclei and heteronuclear Hartmann-Hahn polarization transfers for selective excitation.[1,4] The ^{15}N $R_{1\rho}$ relaxation dispersion pulse sequence adapted for nucleic acids is shown and described in Figure S1. ^{15}N spinlock power calibration was performed as described for ^{13}C experiments or by optimizing the power of a hard ^{15}N 90° pulse inserted immediately after the first INEPT in a conventional $^{1}H,^{15}N$ HSQC experiment. On- and off-resonance relaxation dispersion data were recorded at variable effective spin-lock power (ω_{eff} = 100 – 3500 Hz for ^{13}C; ω_{eff} 100 – 2000 Hz for ^{15}N) and spin-lock offset for $^{13}C/^{15}N$-labeled samples, listed in Table S2. Relaxation decays were collected with 16 or 32 scans and with time delays {0, 8, 16(2X), 24, 36, 48, 60(2X), 80 ms} for ^{15}N and {0, 4, 8, 12(2X), 16, 20, 24, 32(2X) ms} for ^{13}C data. $R_{1\rho}$ values were fitted to a model assuming a two-state equilibrium ($A \underset{k_B}{\overset{k_A}{\rightleftharpoons}} B$) with an asymmetric distribution ($p_A >> p_B$)[5] using Mathematica 7.0 (Wolfram Research):

$$R_{1\rho} = R_1 \cos^2\theta + R_2 \sin^2\theta + \sin^2\theta \frac{p_A p_B \Delta\omega_{AB}^2 k_{ex}}{\Omega_B^2 + \omega_1^2 + k_{ex}^2} \quad \text{(S1)}$$

R_1 and R_2 are the intrinsic longitudinal and transverse relaxation rates, respectively, ω_1 is the strength of the spinlock power, Ω (= Ω_A - Ω_{SL}) is the resonance offset from the spinlock carrier (Ω_{SL}), $\Delta\omega_{AB} = \Omega_B - \Omega_A$, p_A and p_B are the major state and minor state populations respectively, where $p_A + p_B = 1$, $k_{ex} = k_A + k_B$, where $k_A = k_{ex}\, p_B$ and $k_B = k_{ex}\, p_A$. F-statistics was used to assess the presence of chemical exchange (no exchange: $R_{1\rho} = R_1 \cos^2\theta + R_2 \sin^2\theta$) or individual versus global fits for two or more spins (data not shown).

Natural abundance ^{13}C on-resonance profiles were obtained for c7A16 C8/C1' and G10 C8 sites in A_6-DNA^{c7A16} (NMR buffer, pH 6.8 or pH 5.2), and c7G10 C8, C15 C6 and A16 C8 sites in A_6-DNA^{c7G10} (NMR buffer, pH 5.2). Relaxation decays were collected with 512 scans and three relaxation delays {0, 32(2X) ms} for C8/C6 and {0, 42(2X) ms} for C1'. Data were processed as described above. $R_{1\rho}$ on-resonance relaxation dispersion data was fit to a two-state equation that assumes fast chemical exchange ($k_{ex} >> \Delta\omega_{AB}$):[6]

$$R_{1\rho} = R_2 + R_{ex} = R_2 + \frac{\Phi_{ex} k_{ex}}{\omega_1^2 + k_{ex}^2}; \Phi_{ex} = p_A p_B \Delta\omega_{AB}^2 \qquad \text{(S2)}$$

This expression was used in place of Eq. S1 due to lack of off-resonance data to enable fitting of all parameters (specifically, p_A, p_B, and $\Delta\omega_{AB}$), and in order to distinguish between the presence and absence of relaxation dispersion in unlabeled DNA. For unmodified residues where exchange was observed (Figure 3 and Figure S6), the fits are reported for visual guidance only, while the values for chemical exchange rates obtained (data not reported) would not necessarily match values obtained from on/off-resonance data fitted to Eq. S1, especially in the case of intermediate chemical exchange ($k_{ex} \sim \Delta\omega_{AB}$).

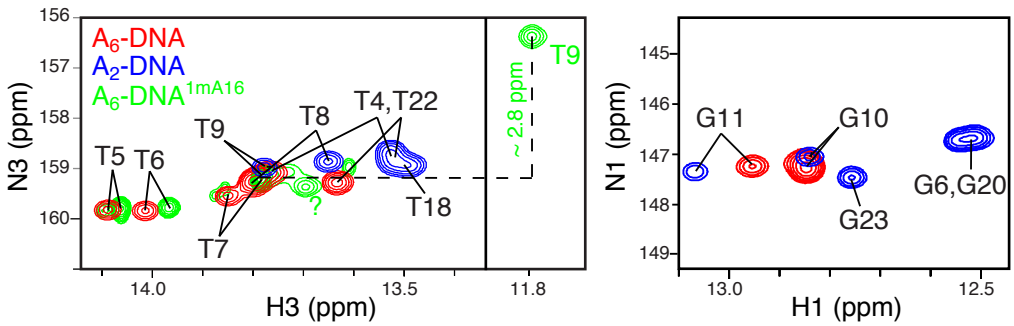
CPMG ^{13}C relaxation dispersion profiles were measured in 1D acquisition mode at 14.1 T and 26 ^{o}C by modifying a published pulse sequence for ^{15}N CPMG relaxation dispersion.[7] CPMG data was recorded at variable ν_{CPMG} (31 - 1000 Hz), frequency between 180^{o} pulses, and with two time delays {0, 32 ms} for C8. $R_{2,CPMG}$ relaxation rates were calculated by assuming a mono-exponential decay and errors were obtained from signal-to-noise estimation. CPMG profiles are included in Figure 3 and Figure S6.

$R_{1\rho}$ simulations

$R_{1\rho}$ simulations with variable p_B and $\Delta\omega_{AB}$ were carried out to test the reliability of $R_{1\rho}$ data fits when using Eq. S1, under conditions of low p_B and $\Delta\omega_{AB}$ values (observed here for ^{15}N relaxation dispersion). Sets of synthetic $R_{1\rho}$ relaxation data were generated with Mathematica 7.0 (Wolfram Research) using Eq. S1 by keeping R_1, R_2 and k_{ex} constant (best-fit values for T8 N3, pH 5.4 used: R_1 = 1.44 Hz, R_2 = 4.54 Hz, k_{ex} =1970 Hz) and varying p_B (= 0.1, 1.0 and 10 %) and $\Delta\omega_{AB}$ (= -50, -100, -200 and -500 Hz). The generated $R_{1\rho}$ data points corresponded to the same spinlock power/offset as in the experimental data. Gaussian noise with a standard deviation identical to the experimental error (for T8 N3) was added to the simulated $R_{1\rho}$ rates. 1000 simulations were performed for each set of $p_B/\Delta\omega_{AB}$ values by selecting random initial values for all parameters within appropriate bounds (R_1 = 0 - 2.5 Hz, R_2 = 0 - 15 Hz, k_{ex} = 0 - 10^6 Hz, p_B = 0 - 50 %, and $\Delta\omega_{AB}$ = -10^4 – 10^4 Hz) and fitting them to Eq. S1 using the same algorithm as in standard data fits. In general, combinations of low absolute p_B (< 1%) and $\Delta\omega_{AB}$ (< 100 Hz) values consistently failed to reproduce the starting values, while higher absolute values of p_B (> 1%) and $\Delta\omega_{AB}$ (> 200 Hz) reproduced the starting values with deviation of < 4 % in p_B and < 2.5 % in $\Delta\omega_{AB}$ (Figure S4), comparable to the experimental error. Therefore, the presence of low intrinsic p_B and $\Delta\omega_{AB}$ values in ^{15}N relaxation dispersion profiles here could explain the large deviations in p_B and $\Delta\omega_{AB}$ values we observe between individual ^{15}N data fits versus and ^{13}C data fits for nuclei within the same base pair (Table S3). Thus, performing a global analysis of multiple data sets where $p_B/\Delta\omega_{AB}/k_{ex}$ are better defined, i.e. ^{13}C and ^{15}N data together, and/or minimizing signal-to-noise by longer acquisition times and higher sample concentrations could reduce the inaccuracy in ^{15}N data fits in cases where p_B and $\Delta\omega_{AB}$ values are small.

DFT calculations

DNA conformations for WC and HG A•T and G•C base pairs were obtained as previously described.[2] The single-atom N7→C7H7 substitution was introduced into these conformations using the molecular modeling program UCSF Chimera.[8] DFT chemical shift calculations were conducted on a high-performance computing cluster with Gaussian 03[9] by using the GIAO method and the B3LYP/6-311+G(2d,p) basis set. ^{15}N and ^{13}C chemical shift predictions for unmodified and modified base pairs are reported in Table S1.



Supplementary Figure 1. $^{1}H,^{15}N$ HSQC spectra of A_6-DNA (red), A_2-DNA (blue), and A_6-DNA^{1mA16} (green) collected at 14.1 T, 26 °C and pH 5.4. Shown are T N3 and G N1 imino resonance assignments and the upfield shift of T9 N3 (highlighted in green) in the N1-methyladenine (1mA) Hoogsteen base pair relative to T N3 in the corresponding Watson-Crick base pairs.

Residue		$\Delta\omega_{(HG\text{-}WC)}$ **(ppm)**			$\Delta\omega_{(c7HG\text{-}c7WC)}$ **(ppm)**		$\Delta\omega_{(c7WC\text{-}WC)}$ **(ppm)**	
		N1/N3	C8	C1'	C8	C1'	C8	C1'
A/T	NMR	-2.8[a]	5.5[a]	3.6[a]	-	-	-18.0[b]	-0.1[b]
	DFT	-2.6 ± 2.0[c]	4.3 ± 1.0[c]	4.7 ± 1.7[c]	5.2[d]	5.6[d]	-24.1	-0.1
G	NMR	-	3.10[a]	3.70[a]	-	-	-18.0[b]	-0.2[b]
	DFT	-2.5 ± 1.2[c]	4.0 ± 2.0[c]	5.2 ± 1.4[c]	5.4	6.6	-26.5	0.0

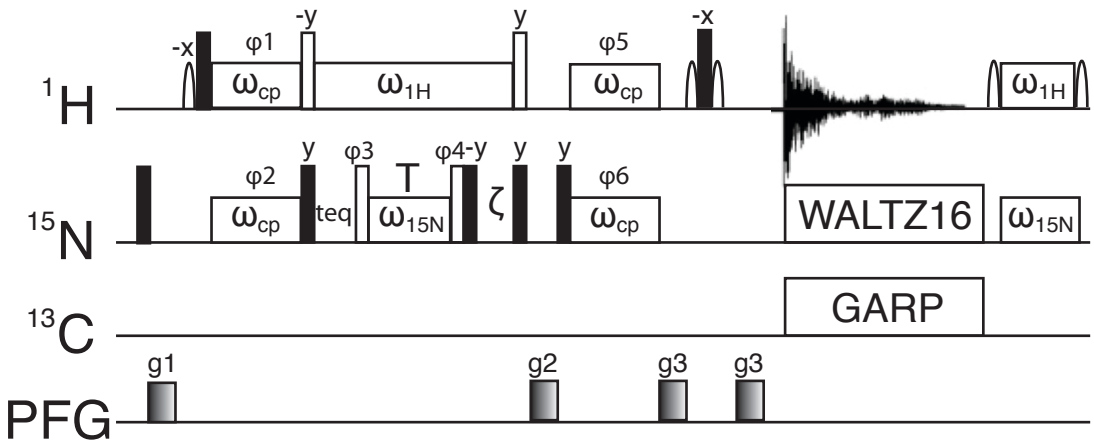
[a] chemical shift difference between WC and N1-methyladenine (1mA) modified Hoogsteen (HG) base pair

[b] chemical shift difference between WC and c7-modified base pair (c7WC)

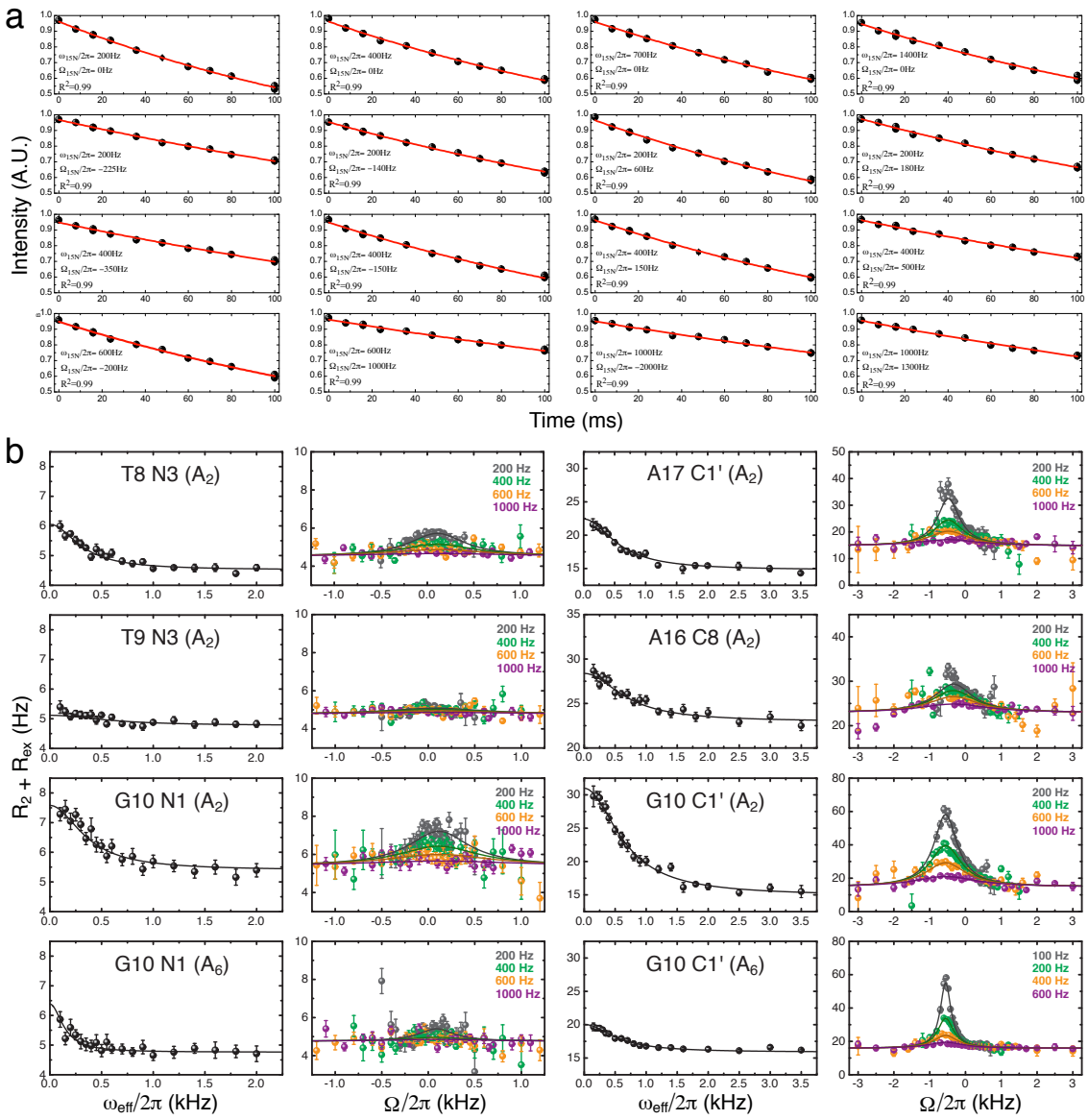
[c] reported as the average and standard deviation of DFT calculation for 12 different WC or HG structures, best matching DFT values reported in ref (2).

[d] single nucleotide instead of base pair was used for DFT due to steric clash between A H7 and T H3 in a hypothetical c7HG base pair.

Supplementary Table 1. Comparison of ^{13}C and ^{15}N chemical shift differences obtained from NMR and DFT calculations for various unmodified and modified Watson-Crick (WC) and Hoogsteen (HG) A•T and G•C base pairs. The WC to HG base pair transition causes significant shift in T N3 and G N1 (HG-WC), which is predicted to be similar in the c7WC to c7HG base pair transition (c7HG-c7WC). The c7-modification alone causes significant chemical shift perturbations in WC versus c7WC base pairs, as confirmed by DFT calculations (c7WC-WC). Carbon C8 and C1' values refer to A and G only.



Supplementary Figure 2. The selective 1D pulse scheme for measuring ^{15}N $R_{1\rho}$ relaxation dispersion is shown as previously described.[1] Black narrow bars represent hard 90° pulses, open narrow bars represent pulses applied with a tip angle $\theta = \text{arccot}(|\Omega_{15N}|/\omega_{15N})$ or $\text{arccot}(|\Omega_{1H}|/\omega_{1H})$, where Ω_{15N} is the offset from the spinlock and ω_{15N} is the ^{15}N spinlock field strength, Ω_{1H} is the offset from the frequency of water and ω_{1H} is the ^{1}H spinlock field strength. Open rectangles denote continuous wave (CW) irradiation periods on ^{1}H used to decouple ^{15}N-^{1}H DD/CSA cross-correlated relaxation and $^{1}J_{NH}$ evolution ($\omega_{1H} \approx$ 5000 Hz), the ^{15}N spinlock (ω_{15N} = 100 - 2000 Hz), the selective cross-polarization periods ($\omega_{cp} \approx$ 85 Hz) and the heat compensation element (directly after acquisition and prior to the first 90^{o} ^{15}N pulse). To normalize for heating effects between different time delays, an off-resonance heat compensation element is applied for T_{comp}= T_{max}-T, where T_{max} is the maximum and T is the relaxation time delay. t_{eq} that allows for equilibration of spins due to chemical exchange is set to 5 ms (~3/kex). The optional delay ζ is used to eliminate signals with different ^{1}H but similar ^{15}N frequencies, $\zeta=\pi/(2\Omega)$ where $\Omega/(2\pi)$ is the ^{15}N offset (in Hz) from the desired signal. Decoupling during acquisition is accomplished using WALTZ16 for ^{15}N and GARP for ^{13}C (optional). All pulses are applied with an *x* phase unless indicated otherwise. Additional purge elements after the ω_{1H} spinlock are included to improve water suppression. The phase cycle is as follows: ϕ_1= {8(*y*),8(-*y*)}, ϕ_2={4(*x*),4(-*x*)}; ϕ_3=-*y* and ϕ_4=*y* for spinlock offset $\Omega < 0$, ϕ_3=*y* and ϕ_4=-*y* for $\Omega > 0$; ϕ_5={4(x),4(-x)}, ϕ_6={2(x),2(-x)} and rec={x,-x,-x,x,-x,x,x,-x,-x,x,x,-x,x,-x,-x,x}. Pulse field gradients and carrier frequency changes were implemented as previously reported.[1]



Supplementary Figure 3. ^{15}N and ^{13}C $R_{1\rho}$ relaxation dispersion profiles. (a) Sample ^{15}N relaxation decays showing mono-exponential behavior (T8 N3, preliminary data, repeats at 16 and 100 ms). The symbol size reflects the experimental error derived from signal-to-noise. (b) Profiles for various residues in A_6-DNA (A_6) and A_2-DNA (A_2) measured at 14.1 T, 26 ^{o}C and pH 5.4. Shown are $R_2 + R_{ex}$ contributions as a function of spinlock power ($\omega_{eff}/2\pi$) and resonance offset ($\Omega/2\pi$) at four spinlock powers (see inset). Solid lines indicate best global fits to Eq. S1 with shared k_{ex}/p_B for sites in the same base pair. All error bars represent experimental uncertainty (one s.d.) estimated from mono-exponential fitting of duplicate sets of $R_{1\rho}$ data and analysis of signal-to-noise. Some data points at high offset are excluded for clarity.

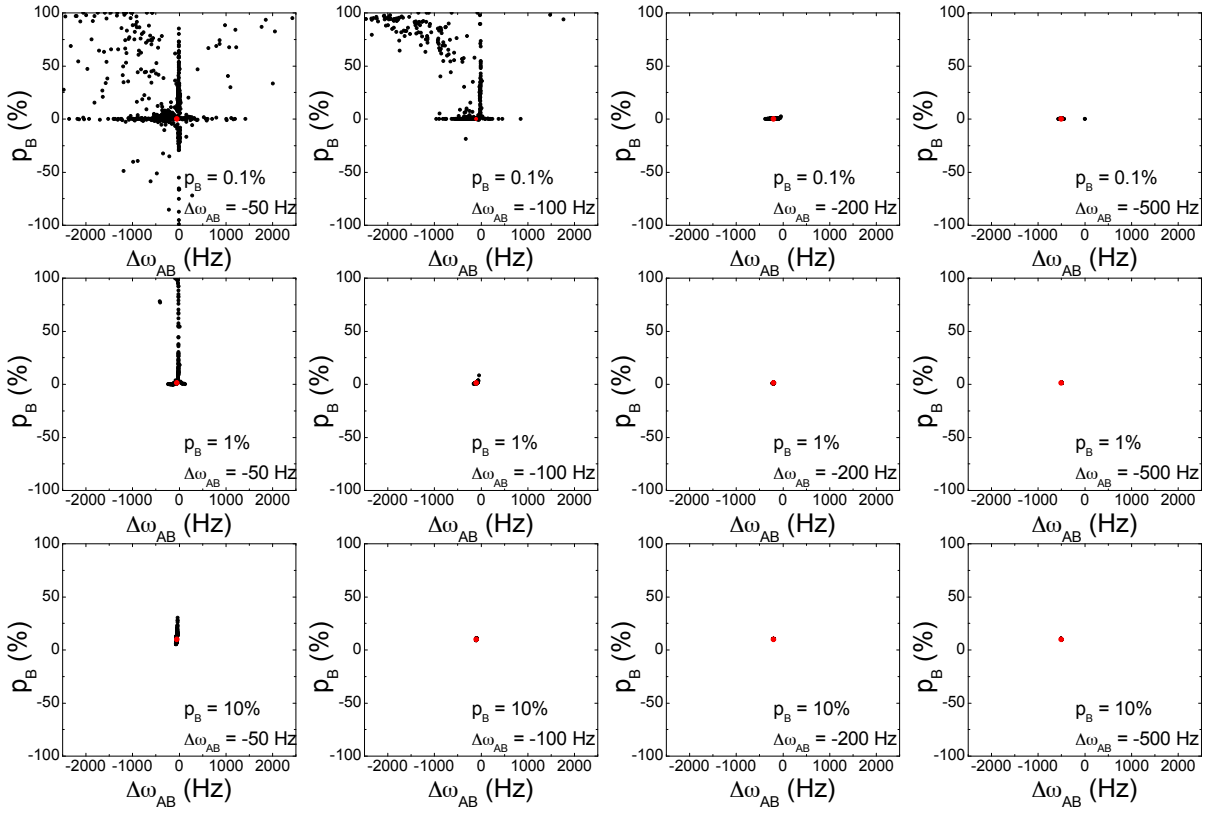
Residue (spin)	On-resonance spinlock power (ω)/ Off-resonance spinlock power (ω) & ±{offset (Ω)}
G10 (N1) T8 (N3) T9 (N3)	100, 150, 200, 250, 300, 350, 400, 450, 500, 550, 600, 700, 800, 900, 1000, 1200, 1400, 1600, 1800, 2000 Hz/ 200 Hz & ±{20, 40, 50, 60, 70, 80, 100, 120, 130, 140, 160, 170, 180, 200, 225, 250, 300, 350, 400, 500 Hz} 400 Hz & ±{30, 60, 90,110, 120, 140, 150, 170, 180, 210, 220, 250, 300, 350, 400, 500, 600, 700, 800, 1000 Hz} 600 Hz & ± {40, 80, 120, 150, 160, 200, 240, 250, 300, 350, 400, 450, 500, 600, 800, 1000, 1200, 1400, 1600, 2000 Hz} 1000 Hz & ± {100, 200, 250, 300, 400, 450, 500, 600, 650, 700, 900, 1100, 1300, 1500, 2000, 2500, 3000, 3500, 4000 Hz}
G10 (C1', A_6)	150, 200, 250, 300, 350, 400, 500, 600, 700, 800, 900, 1000, 1200, 1400, 1600, 1800, 2000, 2500, 3000, 3500 Hz/ 100 Hz & ±{30, 60, 90, 120, 150, 180, 210, 240, 270, 300, 360, 420, 480, 540, 600, 700, 800, 330 Hz} 200 Hz & ±{50, 100, 150, 200, 250, 300, 350, 400, 450, 500, 600, 700, 800, 900, 1000, 1200, 1500, 550 Hz} 400 Hz & ±{75, 150, 225, 300, 375, 450, 550, 650, 800, 1000, 1200, 1400, 1600, 2000, 2500, 3000, 3500, 4000, 5000 Hz} 600 Hz & ±{100, 200, 300, 400, 500, 600, 700, 900, 1100, 1300, 1500, 1700, 2000, 2500, 3000, 3500, 4000, 5000, 7500 Hz}
G10 (C1', A_2) A17 (C1') A16 (C8)	150, 200, 250, 300, 350, 400, 500, 600, 700, 800, 900, 1000, 1200, 1400, 1600, 1800, 2000, 2500, 3000, 3500 Hz/ 200 Hz & ±{30, 60, 90, 120, 150, 180, 210, 240, 270, 300, 360, 420, 480, 540, 600, 700, 800, 330 Hz} 400 Hz & ±{50, 100, 150, 200, 250, 300, 350, 400, 450, 500, 600, 700, 800, 900, 1000, 1200, 1500, 550 Hz} 600 Hz & ±{75, 150, 225, 300, 375, 450, 550, 650, 800, 1000, 1200, 1400, 1600, 2000, 2500, 3000, 3500, 4000, 5000 Hz} 1000 Hz & ±{100, 200, 300, 400, 500, 600, 700, 900, 1100, 1300, 1500, 1700, 2000, 2500, 3000, 3500, 4000, 5000, 7500 Hz}
C15 (C6,c7A16)	100, 150, 200, 250, 300, 350, 400, 450, 500, 550, 600, 700, 800, 900, 1000, 1200, 1400, 1600, 1800, 2000 Hz/ 200 Hz & ±{20, 40, 60, 80, 100, 120, 140, 160, 180, 200, 225, 250, 300, 350, 400, 500 Hz} 600 Hz & ±{30, 60, 90,120, 150, 180, 210, 250, 300, 350, 400, 500, 600, 700, 800, 1000 Hz}

Supplementary Table 2. On/off-resonance parameters used to collect $R_{1\rho}$ relaxation dispersion profiles at 14.1 T and 26.0 °C.

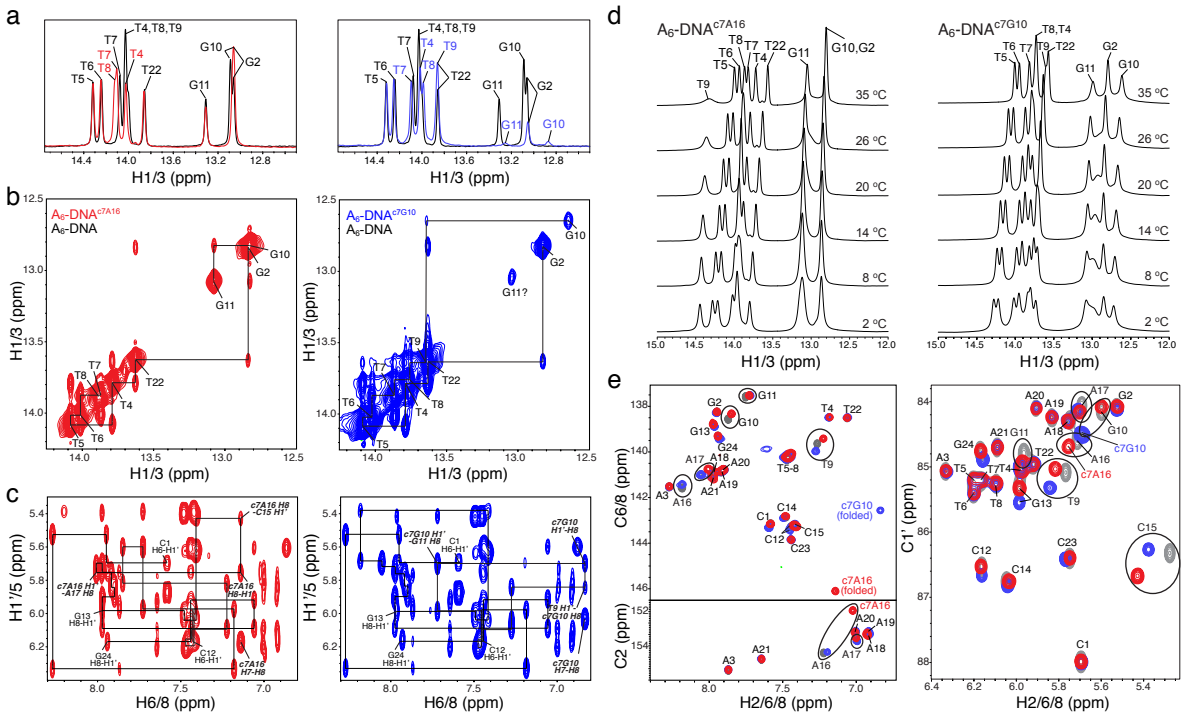
Residue	Parameter	*Individual Fit*	*Global Fit**
T8 (N3) A17 (C1') A_2-DNA	$p_{B,C1'}$ (%)	0.54 ± 0.02	0.55 ± 0.02
	$p_{B,N1}$ (%)	0.73 ± 0.27	
	$k_{ex,C1'}$ (s^{-1})	2097 ± 134	2024 ± 107
	$k_{ex,N1}$ (s^{-1})	1911 ± 185	
	$\Delta\omega_{AB,C1'}$ (ppm)	3.19 ± 0.10	3.18 ± 0.10
	$\Delta\omega_{AB,N1}$ (ppm)	-1.76 ± 0.34	-2.07 ± 0.07
	$R_{1,C1'}$ (Hz)	1.81 ± 0.05	1.81 ± 0.05
	$R_{1,N1}$ (Hz)	1.44 ± 0.02	1.44 ± 0.02
	$R_{2,C1'}$ (Hz)	14.7 ± 0.10	14.8 ± 0.10
	$R_{2,N1}$ (Hz)	4.54 ± 0.03	4.54 ± 0.03
T9 (N3) A16 (C8) A_2-DNA	$p_{B,C8}$ (%)	0.61 ± 0.09	0.60 ± 0.08
	$p_{B,N1}$ (%)	7.66 ± 7.05	
	$k_{ex,C8}$ (s^{-1})	3933 ± 287	3827 ± 269
	$k_{ex,N1}$ (s^{-1})	2111 ± 715	
	$\Delta\omega_{AB,C8}$ (ppm)	2.25 ± 0.17	3.24 ± 0.16
	$\Delta\omega_{AB,N1}$ (ppm)	-0.29 ± 1.22	-1.24 ± 0.13
	$R_{1,C8}$ (Hz)	2.21 ± 0.04	2.21 ± 0.04
	$R_{1,N1}$ (Hz)	1.70 ± 0.02	1.71 ± 0.02
	$R_{2,C8}$ (Hz)	22.9 ± 0.20	22.9 ± 0.20
	$R_{2,N1}$ (Hz)	4.83 ± 0.04	4.77 ± 0.04
G10 (N1/C1') A_2-DNA	$p_{B,C1'}$ (%)	1.02 ± 0.03	1.02 ± 0.02
	$p_{B,N1}$ (%)	10.9 ± 15.0	
	$k_{ex,C1'}$ (s^{-1})	2186 ± 87	2171 ± 78
	$k_{ex,N1}$ (s^{-1})	3104 ± 353	
	$\Delta\omega_{AB,C1'}$ (ppm)	3.73 ± 0.07	3.72 ± 0.06
	$\Delta\omega_{AB,N1}$ (ppm)	-0.89 ± 0.54	-1.90 ± 0.06
	$R_{1,C1'}$ (Hz)	1.73 ± 0.06	1.73 ± 0.06
	$R_{1,N1}$ (Hz)	1.63 ± 0.08	1.99 ± 0.03
	$R_{2,C1'}$ (Hz)	14.8 ± 0.20	14.9 ± 0.20
	$R_{2,N1}$ (Hz)	4.25 ± 0.17	5.39 ± 0.05
G10 (N1/C1') A_6-DNA	$p_{B,C1'}$ (%)	0.61 ± 0.03	0.61 ± 0.02
	$p_{B,N1}$ (%)	3.80 ± 0.19	
	$k_{ex,C1'}$ (s^{-1})	734 ± 42	736 ± 39
	$k_{ex,N1}$ (s^{-1})	739 ± 430	
	$\Delta\omega_{AB,C1'}$ (ppm)	3.82 ± 0.03	3.82 ± 0.03
	$\Delta\omega_{AB,N1}$ (ppm)	-1.84 ± 0.50	-1.46 ± 0.08
	$R_{1,C1'}$ (Hz)	1.86 ± 0.02	1.86 ± 0.02
	$R_{1,N1}$ (Hz)	1.59 ± 0.03	1.59 ± 0.03
	$R_{2,C1'}$ (Hz)	15.7 ± 0.10	15.7 ± 0.10
	$R_{2,N1}$ (Hz)	4.76 ± 0.04	4.76 ± 0.03

* $p_{B,}$ and k_{ex} were shared in global fits

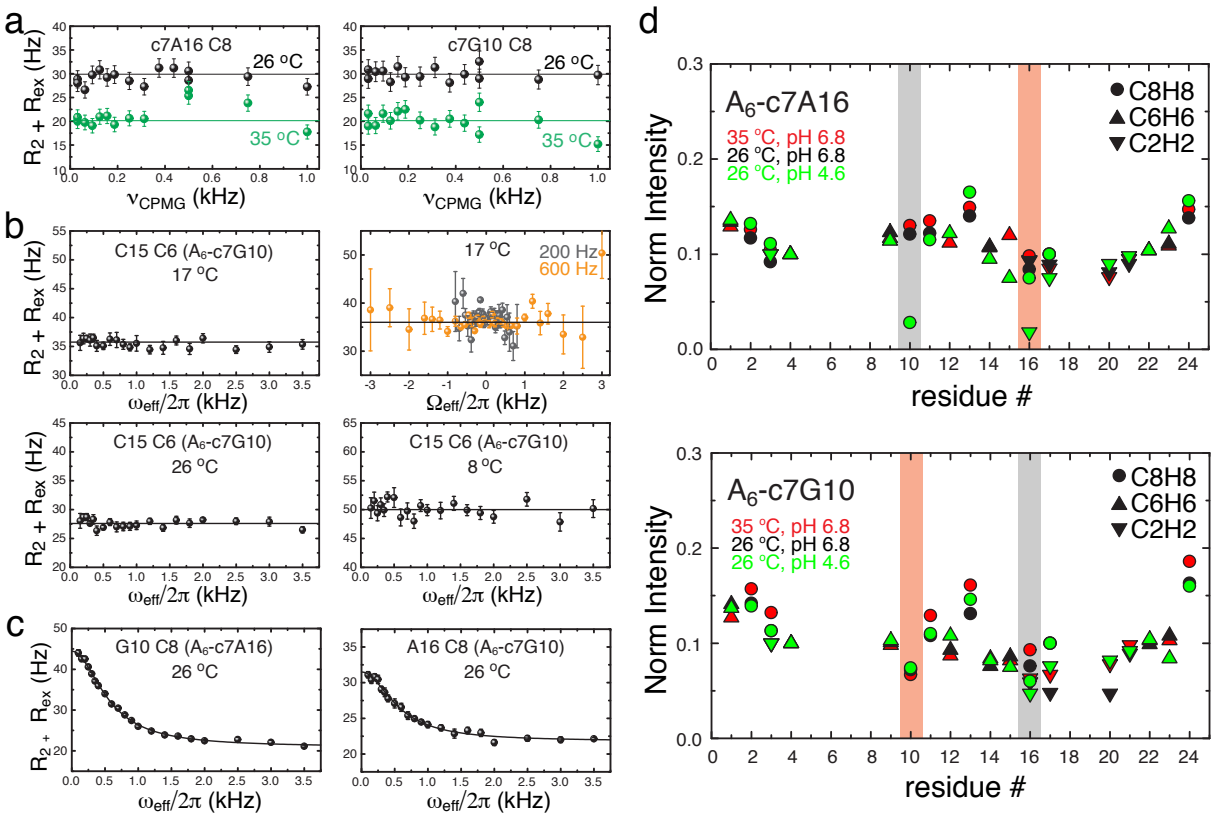
Supplementary Table 3: Best-fit parameters obtained from fitting individually or globally on- and off-resonance $R_{1\rho}$ relaxation dispersion data to Eq. S1.



Supplementary Figure 4. $R_{1\rho}$ simulations of noise-corrupted synthetic data with variable starting p_B and $\Delta\omega_{AB}$ values (inset) using Eq. S1. Black circles represent different simulation points, red circles represent starting parameters (see Methods for details).



Supplementary Figure 5. NMR spectra and resonance assignments for A_6-DNA^{c7A16} and A_6-DNA^{c7G10}. (a) 1D imino ^{1}H spectra and assignments (26 ^{o}C) showing overlay between unmodified A_6-DNA (black) and modified A_6-DNA^{c7A16} (red) and A_6-DNA^{c7G10} (blue) (excitation sculpting used for water suppression). Chemical shift perturbations and line broadening is observed for the modified (T9 or G10), neighboring and even remote base pairs. (b) Imino proton ^{1}H,^{1}H NOESY spectra (26 ^{o}C) showing an NOE walk between neighboring base pairs and significant broadening of T9 H3 - G10 H1/T8 H3 (left), and G10 H1 - G11 H1/T9 H3 (right) cross-peaks. (c) Corresponding ^{1}H,^{1}H NOESY spectra (26 ^{o}C) showing H6/H8 - H1' NOE walk typical for B-DNA. (d) Temperature dependence of 1D ^{1}H imino proton spectra (jump-return used for water suppression) showing significant line broadening at the modified sites, especially T9 H3. (e) 2D ^{1}H,^{13}C HSQC overlays and peak assignments showing large chemical shift perturbations at the modified site and neighboring base pairs (circled).



Supplementary Figure 6. Suppression of relaxation dispersion in c7-modified but not in neighboring base pairs. (a) CPMG profiles for c7A16 C8 (A_6-DNA^{c7A16}) and c7G10 C8 (A_6-DNA^{c7G10}) at 26 °C and 35 °C showing lack of chemical exchange. (b) On/off-resonance ^{13}C $R_{1\rho}$ relaxation dispersion profiles for $^{13}C/^{15}N$-enriched C15 C6 (A_6-DNA^{c7G10}) showing lack of chemical exchange in c7G10•C15 at a range of temperatures (8 to 26 °C). These data confirm that formation of transient HG base pairs is suppressed over a range of timescales. (c) On-resonance ^{13}C $R_{1\rho}$ relaxation dispersion profiles at G10 C8 (A_6-DNA^{c7A16}) and A16 C8 (A_6-DNA^{c7G10}) showing chemical exchange is retained relative to unmodified A_6-DNA.[2] Solid lines represent best fits to Eq. S2 (data not shown) because of insufficient data to use Eq. S1. (d) Resonance intensity plots for unlabeled A_6-DNA^{c7A16} and A_6-DNA^{c7G10} at various temperature (T) and pH (inset). C7-modified sites (highlighted in orange) that exhibit exchange in unmodified DNA show little to no sensitivity to T and pH, which correlates with lack of chemical exchange, while neighboring base pairs (highlighted in grey) that exhibit chemical exchange show pH dependent line broadening. The c7A16 C2 site (C2 does not show dispersion in unmodified DNA) is severely broadened, likely due to a different process.

Supporting Information References

(1) Korzhnev, D. M.; Orekhov, V. Y.; Kay, L. E. *J Am Chem Soc* 2005, *127*, 713.

(2) Nikolova, E. N.; Kim, E.; Wise, A. A.; O'Brien, P. J.; Andricioaei, I.; Al-Hashimi, H. M. *Nature* 2011, *470*, 498.

(3) Zimmer, D. P.; Crothers, D. M. *Proc Natl Acad Sci U S A* 1995, *92*, 3091.

(4) Hansen, A. L.; Nikolova, E. N.; Casiano-Negroni, A.; Al-Hashimi, H. M. *J Am Chem Soc* 2009, *131*, 3818.

(5) Trott, O.; Palmer, A. G., 3rd *J Magn Reson* 2002, *154*, 157.

(6) Davis, D. G.; Perlman, M. E.; London, R. E. *J Magn Reson Ser B* 1994, *104*, 266.

(7) Hansen, D. F.; Vallurupalli, P.; Kay, L. E. *J Phys Chem B* 2008, *112*, 5898.

(8) Pettersen, E. F.; Goddard, T. D.; Huang, C. C.; Couch, G. S.; Greenblatt, D. M.; Meng, E. C.; Ferrin, T. E. *J Comput Chem* 2004, *25*, 1605.

(9) M. J. Frisch, G. W. T., H. B. Schlegel, G. E. Scuseria, ; M. A. Robb, J. R. C., J. A. Montgomery, Jr., T. Vreven, ; K. N. Kudin, J. C. B., J. M. Millam, S. S. Iyengar, J. Tomasi, ; V. Barone, B. M., M. Cossi, G. Scalmani, N. Rega, ; G. A. Petersson, H. N., M. Hada, M. Ehara, K. Toyota, ; R. Fukuda, J. H., M. Ishida, T. Nakajima, Y. Honda, O. Kitao, ; H. Nakai, M. K., X. Li, J. E. Knox, H. P. Hratchian, J. B. Cross, ; C. Adamo, J. J., R. Gomperts, R. E. Stratmann, O. Yazyev, ; A. J. Austin, R. C., C. Pomelli, J. W. Ochterski, P. Y. Ayala, ; K. Morokuma, G. A. V., P. Salvador, J. J. Dannenberg, ; V. G. Zakrzewski, S. D., A. D. Daniels, M. C. Strain, ; O. Farkas, D. K. M., A. D. Rabuck, K. Raghavachari, ; J. B. Foresman, J. V. O., Q. Cui, A. G. Baboul, S. Clifford, ; J. Cioslowski, B. B. S., G. Liu, A. Liashenko, P. Piskorz, ; I. Komaromi, R. L. M., D. J. Fox, T. Keith, M. A. Al-Laham, ; C. Y. Peng, A. N., M. Challacombe, P. M. W. Gill, ; B. Johnson, W. C., M. W. Wong, C. Gonzalez, and J. A. Pople; Gaussian, Inc.: Wallingford CT, 2004.