PNAS

Supplementary Information for
Probing Watson-Crick and Hoogsteen base pairing in duplex DNA using dynamic nuclear polarization solid-state NMR spectroscopy

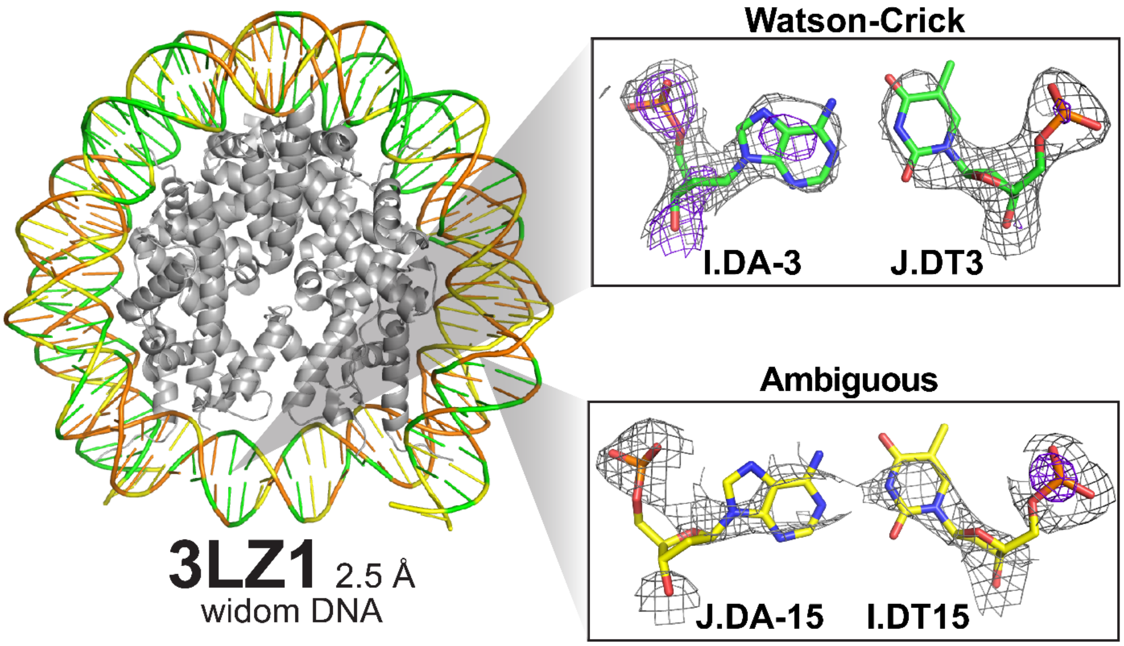
Daniel W. Conroy,[a] Yu Xu,[b] Honglue Shi,[b] Nicole Gonzalez Salguero,[a] Rudra N. Purusottam,[a] Matthew D. Shannon,[a] Hashim M. Al-Hashimi,[b,c,d,1] and Christopher P. Jaroniec[a,1]

[a]Department of Chemistry and Biochemistry, The Ohio State University, Columbus, OH 43210
[b]Department of Chemistry, Duke University, Durham, NC 27708
[c]Department of Biochemistry, Duke University Medical Center, Durham, NC 27710
[d]Department of Biochemistry and Molecular Biophysics, Columbia University, New York, NY 10032

[1]Hashim M. Al-Hashimi and Christopher P. Jaroniec
Email: ha2639@cumc.columbia.edu, jaroniec.1@osu.edu

This PDF file includes:

Figures S1 to S11
Tables S1 to S3
SI References



	Total	A-T	G-C
Watson-Crick	36.6%	40.0%	33.8%
Ambiguous	28.3%	35.4%	22.5%
Disorder	35.2%	24.6%	43.8%

Fig. S1. Analysis of the DNA electron density in the crystal structure (1) of a Widom 601 nucleosome core particle (PDB entry 3LZ1). DNA base pairs with clear Watson-Crick electron density, ambiguous electron density and located in two-fold disorder regions are highlighted in green, yellow and orange, respectively. This analysis shows that for approximately 28% of the base pairs the electron density is ambiguous in discriminating between Watson-Crick and Hoogsteen conformations. Insets show representative examples of base pairs with Watson-Crick and ambiguous electron density.

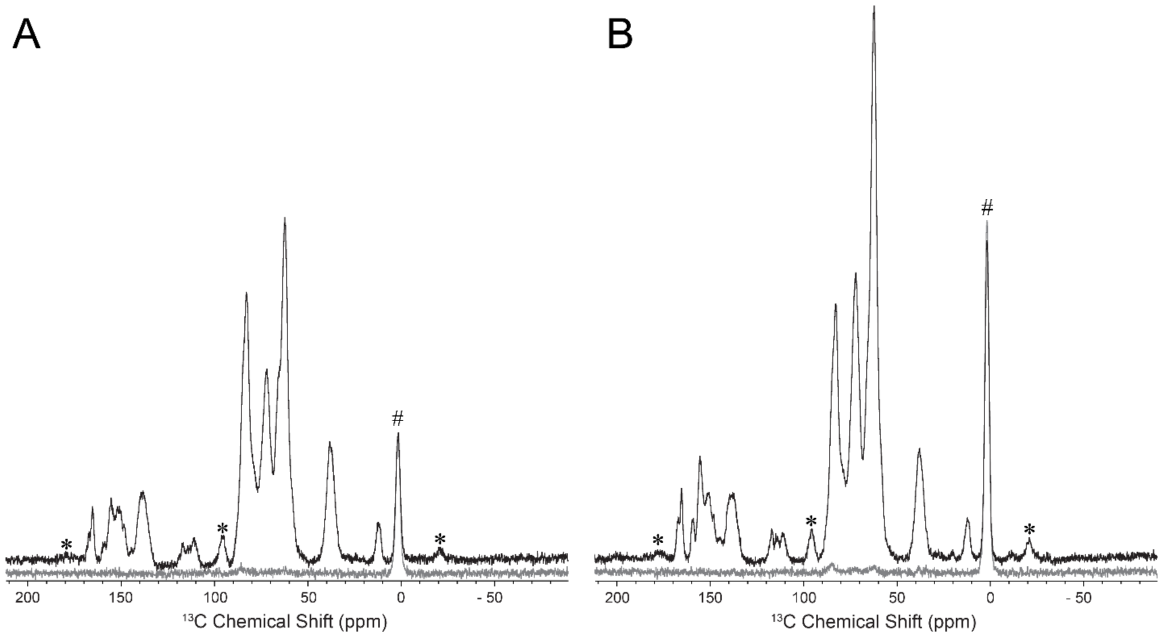


Fig. S2. ^{13}C CP-MAS solid-state NMR spectra of natural abundance 12-mer DNA duplex samples in (**A**) d_8,^{12}C-glycerol, D_2O and H_2O in 60:30:10 v/v/v ratio and (**B**) d_8,^{12}C-glycerol and H_2O in 60:40 v/v ratio, each containing 12 mM AMUPol. DNP-enhanced spectra recorded with microwaves turned on are shown in black, and the corresponding control spectra recorded with the microwaves turned off are shown in grey. Spinning sidebands and background signal arising from a silicone rotor insert are denoted by (*) and (#), respectively.

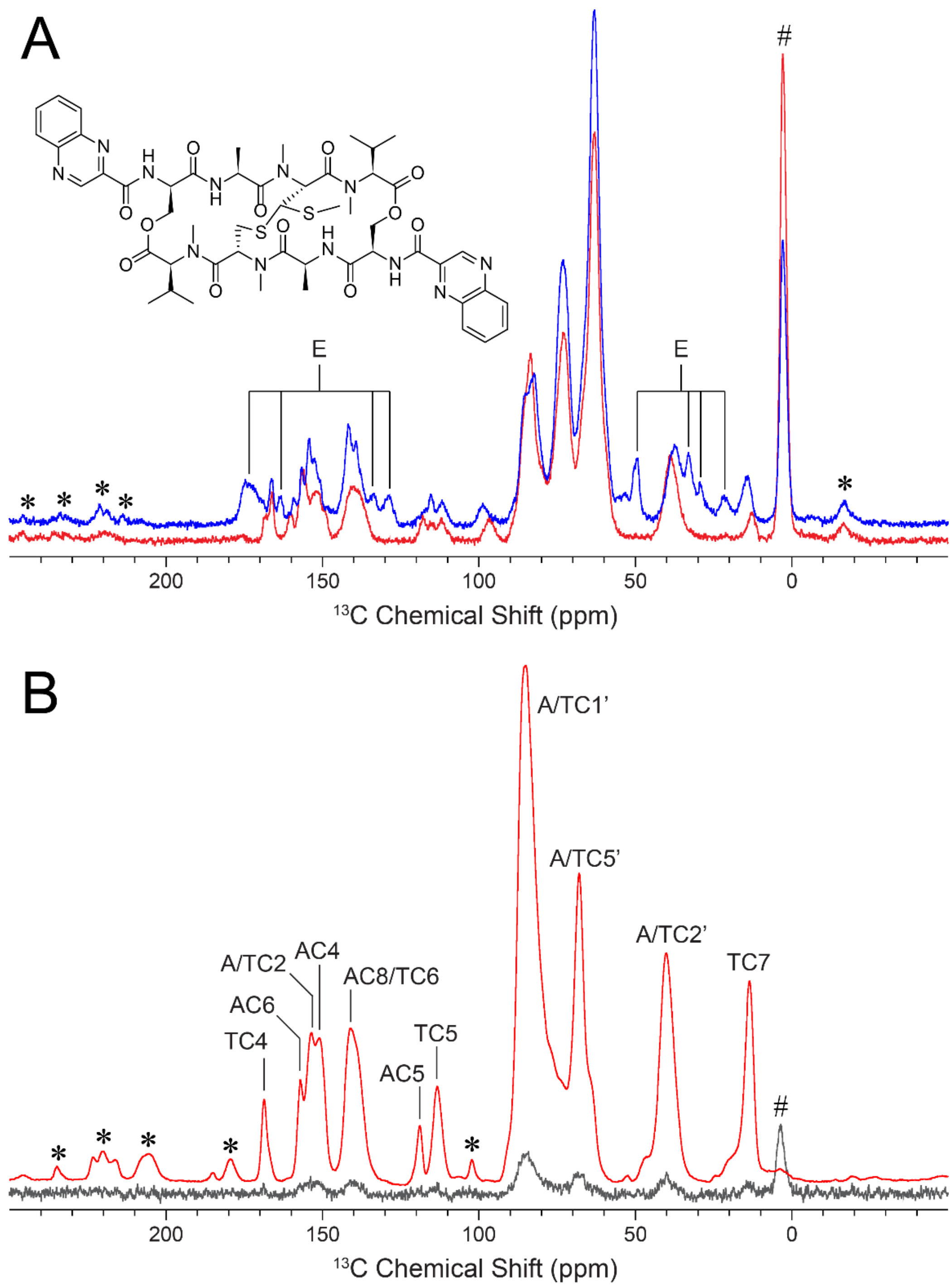


Fig. S3. (**A**) DNP-enhanced ^{13}C CP-MAS solid-state NMR spectra of 1.6 mM natural abundance 12-mer DNA duplex in d_8,^{12}C-glycerol and H_2O in 60:40 v/v ratio containing 12 mM AMUPol (red), and the corresponding sample of 1.1 mM DNA duplex in complex with echinomycin in a 1:2 molar ratio (blue). Resonances that can be uniquely attributed to the bound echinomycin are labeled with (E), while DNA resonances are not labeled in the spectra. (**B**) ^{13}C CP-MAS solid-state NMR spectra of the ^{13}C,^{15}N-TA DNA duplex with Watson-Crick base pairing recorded with microwaves on (red) and off (black; the spectrum is shown at 10x intensity), showing DNP NMR signal enhancement of 156. Spinning sidebands and background signal arising from a silicone rotor insert are denoted by (*) and (#), respectively.

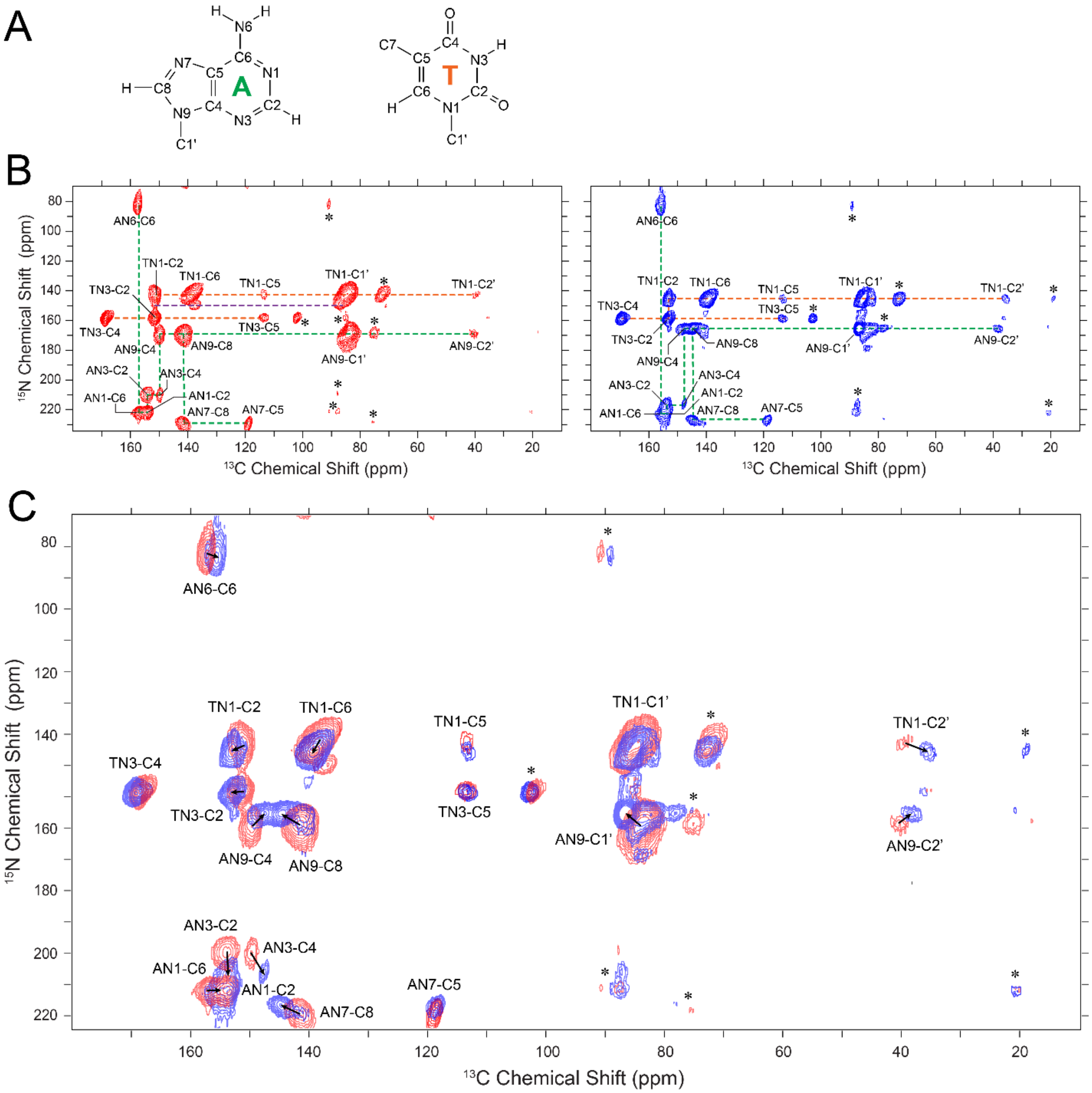


Fig. S4. (**A**) Structures of adenine and thymine nucleotides. (**B**) DNP enhanced ^{15}N-^{13}C ZF-TEDOR solid-state NMR spectra of the model Watson-Crick (red) and Hoogsteen (blue) ^{13}C,^{15}N-TA DNA duplex samples recorded with dipolar mixing time of 1.33 ms and total measurement time of 12 h per spectrum. The resonance assignments for adenine and thymine are traced with green and orange dashed lines, respectively, and the purple dashed line indicates a minor thymine hairpin state corresponding to monomer DNA (2–4). Spinning sidebands are denoted by (*). (**C**) Overlay of the spectra for model Watson-Crick (red) and Hoogsteen (blue) ^{13}C,^{15}N-TA DNA duplex samples shown in panel (B), with the most significant chemical shift differences indicated with black arrows.

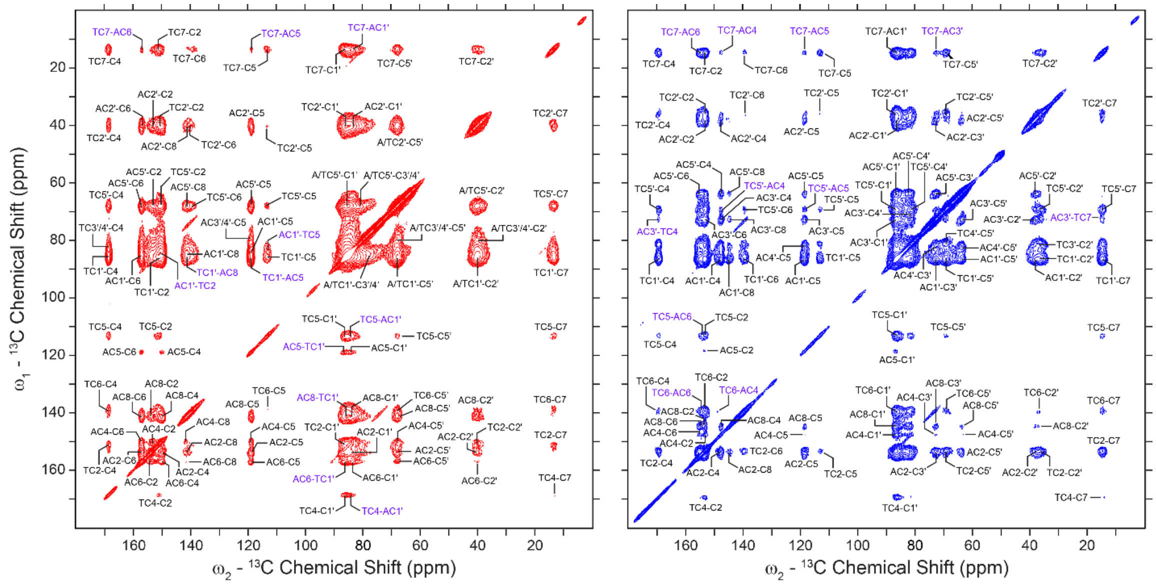


Fig. S5. DNP enhanced ^{13}C-^{13}C DARR solid-state NMR spectra of the model Watson-Crick (red) and Hoogsteen (blue) ^{13}C,^{15}N-TA DNA duplex samples recorded with dipolar mixing time of 500 ms. Intra- and inter-base correlations are labeled in black and purple font, respectively.

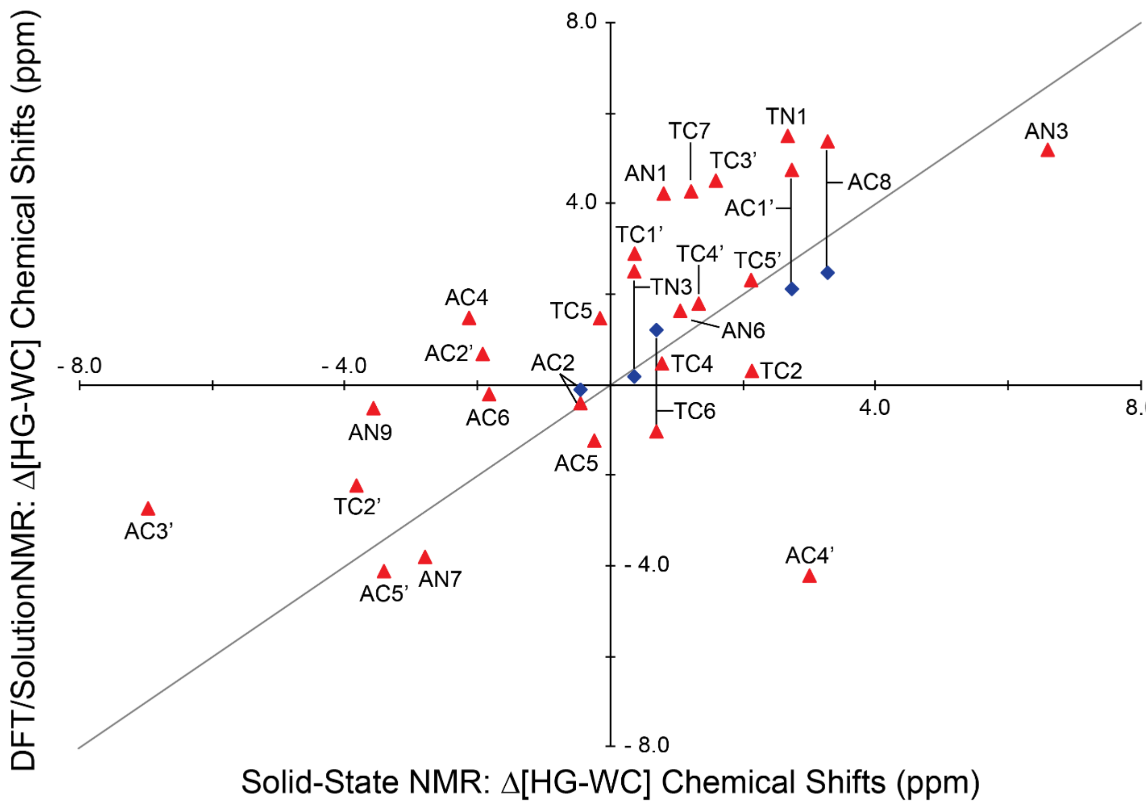


Fig. S6. Comparison of experimental ^{13}C and ^{15}N solid-state NMR chemical shift differences ($\Delta\delta$) for nucleotides 6T and A7 in the model Hoogsteen and Watson-Crick 12-mer DNA duplex samples with the corresponding experimental solution NMR $\Delta\delta$ values (blue diamonds) (2) and DFT-calculated $\Delta\delta$ values (red triangles) (c.f., SI Table S1).

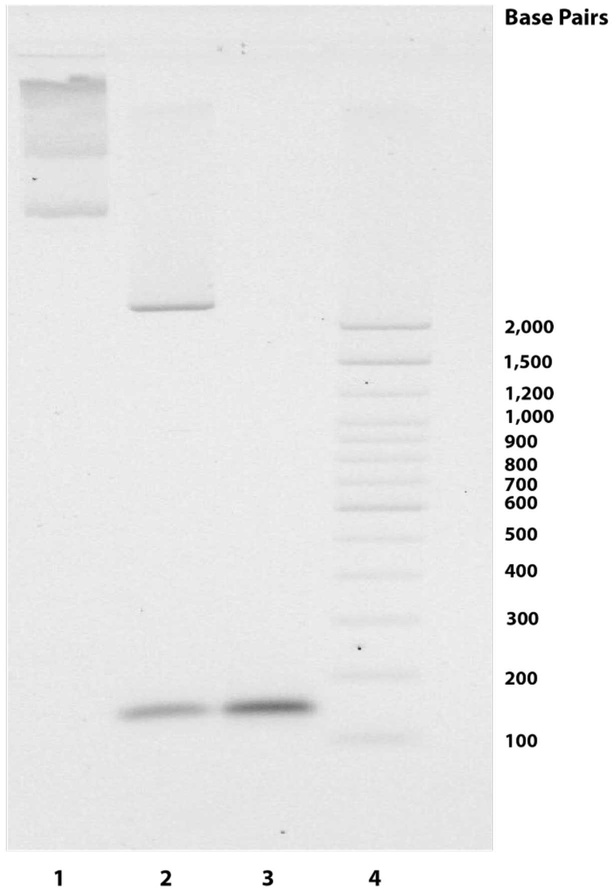


Fig. S7. Agarose gel (1%) analysis of Widom 601 DNA used for nucleosome core particle reconstitution following various purification steps. Lane 1: purified plasmid DNA; Lane 2: plasmid DNA after digestion with EcoRV; Lane 3: purified 147 bp Widom 601 DNA; Lane 4: DNA ladder.

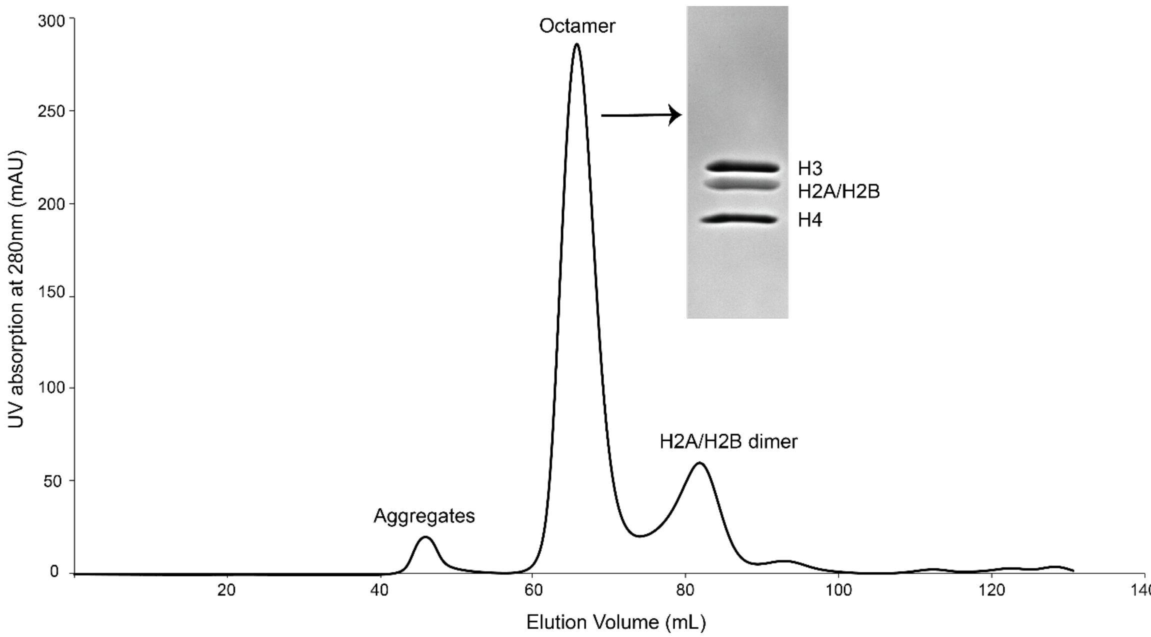


Fig. S8. Elution profile from size exclusion chromatography showing histone octamer assembly from purified histones H2A, H2B, H3 and H4 as described in the text. Protein content of the histone octamer was analyzed by SDS-PAGE (inset).

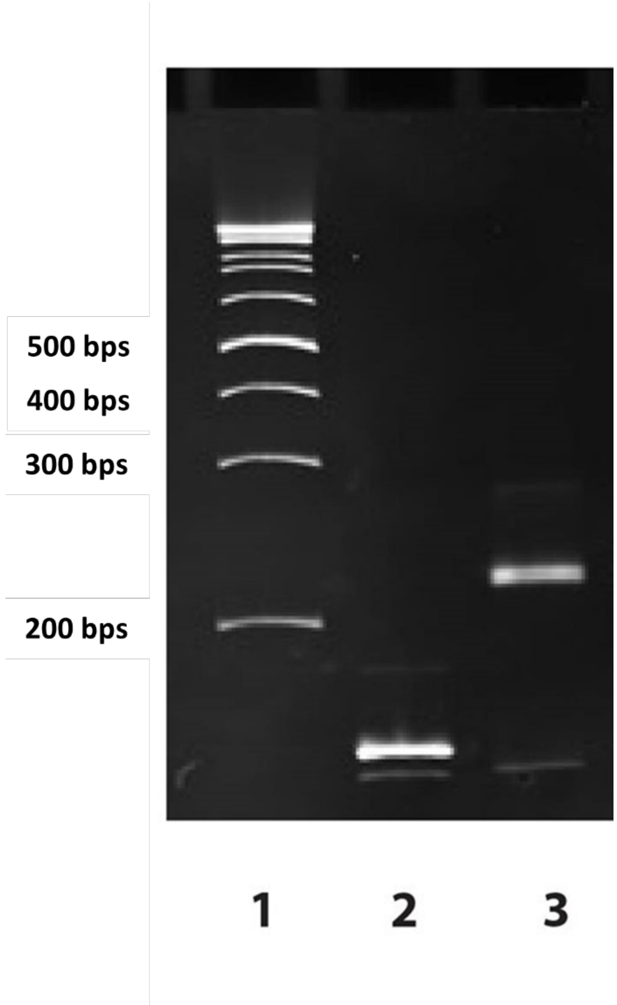


Fig. S9. Native acrylamide gel (5%) showing the shift in electrophoretic mobility between 147 bp Widom 601 DNA and nucleosomes. Lane 1: DNA ladder; Lane 2: 147 bp Widom 601 DNA; Lane 3: Sucrose gradient purified nucleosomes.

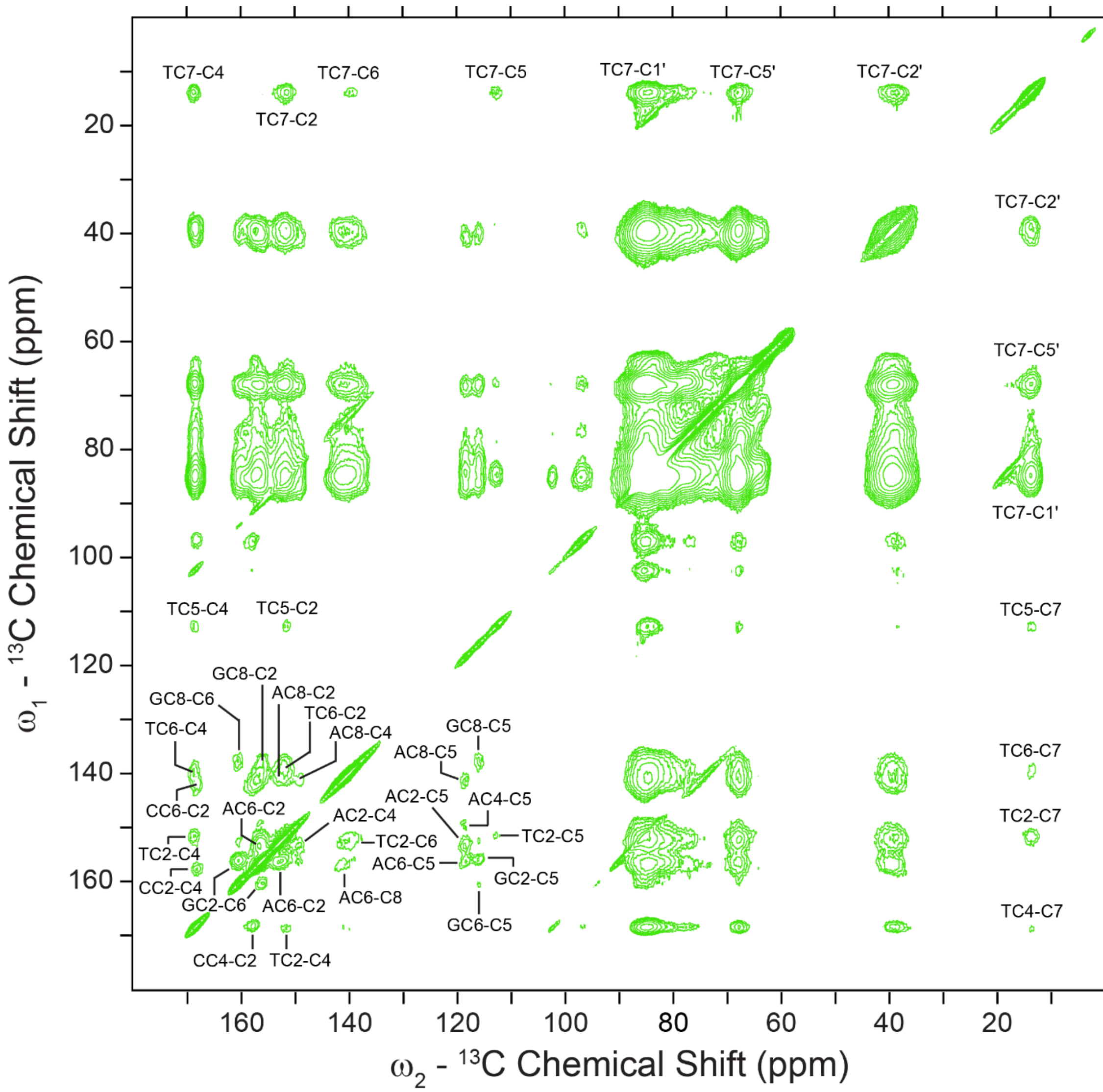


Fig. S10. DNP enhanced ^{13}C-^{13}C DARR solid-state NMR spectrum of NCPs reconstituted with ^{13}C,^{15}N-Widom 601 DNA recorded with dipolar mixing time of 500 ms. Assignments are indicated for the unambiguous correlations.

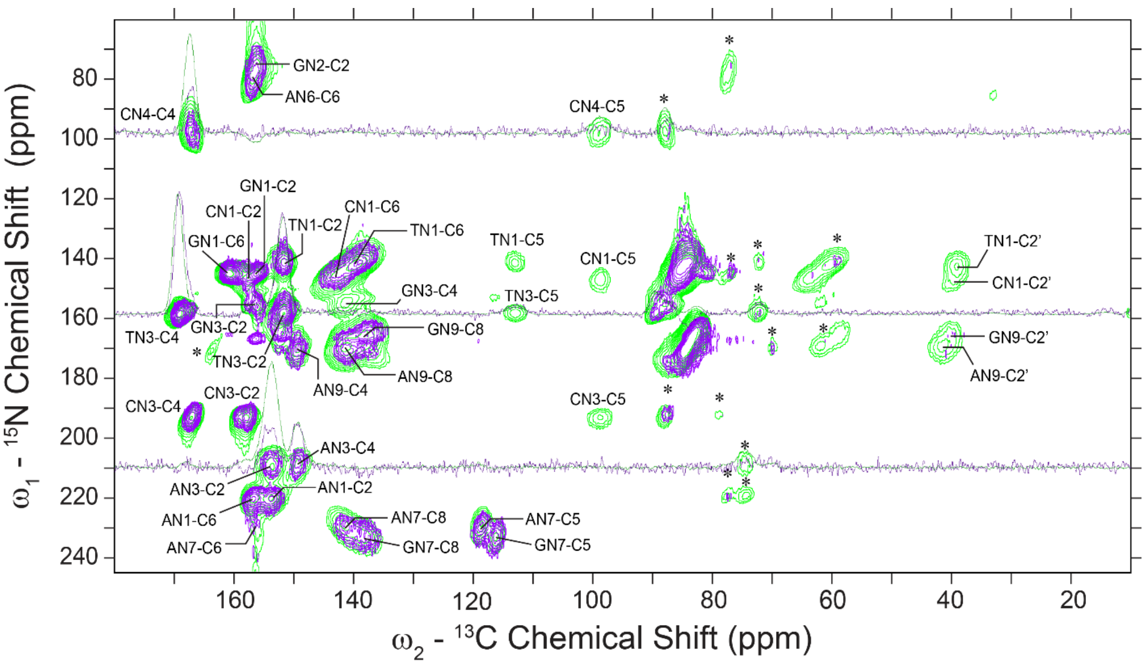


Fig. S11. DNP enhanced ^{15}N-^{13}C ZF-TEDOR solid-state NMR spectra of NCPs reconstituted with ^{13}C,^{15}N-Widom 601 DNA (green) and free ^{13}C,^{15}N-Widom 601 DNA (purple) recorded with dipolar mixing time of 1.33 ms. Spinning sidebands are denoted with an asterisk (*). For each spectrum the cross-peaks are drawn with the lowest contour level at 5 times the root-mean-square noise level, and representative 1D ^{13}C traces are shown corresponding to the CN4, TN3 and AN3 ^{15}N frequencies. The free Widom 601 DNA sample was prepared similarly to the NCP sample (see Materials and Methods) and consisted of ~150 ng of ^{13}C,^{15}N-DNA dissolved in a matrix of 60:40 d_8,^{12}C-glycerol and H_2O (v/v) with 12 mM AMUPol.

Table S1. ^{13}C and ^{15}N solid-state NMR chemical shifts and shift differences ($\Delta\delta$) for nucleotides T6 and A7 in model Watson-Crick and Hoogsteen 12-mer DNA duplexes, and corresponding experimental solution NMR and DFT-calculated $\Delta\delta$ values.

Base	Atom	Watson-Crick Chemical Shift§, δ_{WC} [FWHM]† (ppm)	Hoogsteen Chemical Shift§, δ_{HG} [FWHM]† (ppm)	$\Delta\delta$ (δ_{HG} - d_{WC}) Solid-State NMR‡ (ppm)	$\Delta\delta$ (δ_{HG} - δ_{WC}) DFT (ppm)	$\Delta\delta$ (δ_{HG} - δ_{WC}) Solution NMR (ppm)
A	C1'	84.03 ± 0.74 [5.21]	86.77 ± 0.25 [2.76]	2.74 ± 0.78	4.76	2.13
A	C2	153.94 ± 0.32 [2.85]	153.50 ± 0.16 [2.10]	-0.44 ± 0.36	-0.39	-0.09
A	C2'	40.22 ± 0.22 [4.07]	38.31 ± 0.21 [3.59]	-1.91 ± 0.30	0.70	-----
A	C3'	79.79 ± 0.52 [6.19]*	72.84 ± 0.16 [2.96]	-6.95 ± 0.55	-2.71	-----
A	C4	149.85 ± 0.33 [0.99]	147.73 ± 0.26 [1.50]	-2.12 ± 0.42	1.50	-----
A	C4'	79.79 ± 0.52 [6.19]*	82.80 ± 0.28 [4.07]	3.01 ± 0.59	-4.19	-----
A	C5	118.85 ± 0.05 [1.55]	118.62 ± 0.09 [2.01]	-0.23 ± 0.10	-1.21	-----
A	C5'	67.23 ± 0.69 [2.86]#	63.83 ± 0.17 [2.20]	-3.40 ± 0.71	-4.09	-----
A	C6	157.21 ± 0.21 [1.78]	155.40 ± 0.28 [1.94]	-1.81 ± 0.35	-0.19	-----
A	C8	141.45 ± 0.33 [3.73]	144.73 ± 0.16 [1.53]	3.28 ± 0.37	5.38	2.49
A	N1	221.78 ± 0.28 [4.99]	222.59 ± 0.28 [5.20]	0.81 ± 0.40	4.24	-----
A	N3	209.56 ± 0.43 [7.20]	216.16 ± 0.54 [5.94]	6.60 ± 0.69	5.20	-----
A	N6	82.63 ± 0.58 [9.01]	83.70 ± 0.42 [8.72]	1.07 ± 0.71	1.65	-----
A	N7	229.20 ± 0.43 [6.02]	226.42 ± 0.40 [4.06]	-2.78 ± 0.59	-3.77	-----
A	N9	169.03 ± 0.29 [6.23]	165.47 ± 0.21 [4.07]	-3.56 ± 0.35	-0.50	-----
T	C1'	85.80 ±0.40 [3.07]	86.18 ± 0.55 [2.35]	0.38 ± 0.68	2.91	-----
T	C2	151.09 ± 0.33 [2.94]	153.23 ± 0.18 [2.27]	2.14 ± 0.38	0.32	-----
T	C2'	39.95 ± 0.31 [3.36]	36.14 ± 0.37 [3.82]	-3.81 ± 0.48	-2.20	-----
T	C3'	79.79 ± 0.52 [6.19]*	81.39 ± 0.22 [2.60]	1.60 ± 0.57	4.52	-----
T	C4	168.66 ± 0.15 [1.53]	169.45 ± 0.18 [1.92]	0.79 ± 0.23	0.49	-----
T	C4'	79.79 ± 0.52 [6.19]*	81.13 ± 0.11 [2.23]	1.34 ± 0.53	1.81	-----
T	C5	113.37 ± 0.19 [2.11]	113.22 ± 0.13 [1.87]	-0.15 ± 0.23	1.49	-----
T	C5'	67.23 ± 0.69 [2.86]#	69.36 ± 0.19 [3.15]	2.13 ± 0.71	2.33	-----
T	C6	138.74 ± 0.23 [2.11]	139.45 ± 0.20 [2.48]	0.71 ± 0.30	-1.01	1.23
T	C7	13.45 ± 0.23 [2.65]	14.68 ± 0.17 [2.24]	1.23 ± 0.28	4.29	-----
T	N1	142.86 ± 0.31 [6.48]	145.54 ± 0.23 [4.61]	2.68 ± 0.38	5.51	-----
T	N3	158.45 ± 0.33 [4.14]	158.82 ± 0.30 [4.01]	0.37 ± 0.44	2.52	0.2

§ Errors in δ_{WC} and δ_{HG} values correspond to the standard deviations in the chemical shifts of every assigned resonance for each site in all recorded 2D ^{15}N-^{13}C and ^{13}C-^{13}C spectra.

*,# Shifts are not unique in the Watson-Crick DNA duplex due to resonance overlap.

† Full-width at half-maximum (FWHM) linewidth measurements correspond to the most intense, isolated correlations for each site in 2D ^{15}N-^{13}C or ^{13}C-^{13}C spectra.

Table S2. Inter- and intra-strand A-T distances for model Watson-Crick and Hoogsteen DNA duplexes corresponding to A-T ^{13}C-^{13}C correlations observed in DARR solid-state NMR spectra.

Atoms		Watson-Crick		Atoms		Hoogsteen	
		Inter-strand distance (Å)	Intra-strand distance (Å)			Inter-strand distance (Å)	Intra-strand distance (Å)
AC1'	TC2	8.5	5.1	AC3'	TC4	9.6	8.4
AC1'	TC4	8.8	6.9	AC3'	TC7	12.1	9.2
AC1'	TC5	10.2	7.0	AC4	TC5'	10.6	7.2
AC1'	TC7	11.3	6.3	AC4	TC6	7.8	4.1
AC5	TC1'	8.4	4.6	AC4	TC7	8.3	5.5
AC5	TC7	8.0	5.5	AC5	TC5'	9.7	7.7
AC6	TC1'	7.3	5.5	AC5	TC7	7.0	4.9
AC6	TC7	6.6	5.3	AC6	TC5	6.1	3.4
AC8	TC1'	10.5	3.9	AC6	TC6	7.3	3.7
				AC6	TC7	6.8	3.9

Distances for the Watson-Crick DNA duplex were determined from a structure generated using 3DNA assuming an idealized B-form geometry (5). Distances for the Hoogsteen DNA duplex were determined from the crystal structure (6).

Table S3. ^{13}C and ^{15}N chemical shift assignments for different nucleotide types in the nucleosome core particle.

Base	Atom	Chemical Shift (ppm)	Base	Atom	Chemical Shift (ppm)
A	C2	153.40	G	C2	156.29
A	C2'	41.32	G	C2'	40.08
A	C4	149.30	G	C4	141.12
A	C5	118.52	G	C5	116.02
A	C6	156.67	G	C6	160.54
A	C8	141.26	G	C8	137.85
A	N1	220.72	G	N1	145.15
A	N3	210.24	G	N2	76.83
A	N6	81.68	G	N3	155.11
A	N7	229.88	G	N7	233.54
A	N9	169.84	G	N9	167.77
C	C2	157.92	T	C2	151.73
C	C2'	39.16	T	C2'	38.72
C	C4	168.07	T	C4	168.80
C	C5	97.57	T	C5	112.83
C	C6	142.65	T	C6	139.46
C	N1	146.85	T	C7	13.75
C	N3	193.83	T	N1	142.21
C	N4	97.79	T	N3	158.48

SI References

1. D. Vasudevan, E. Y. D. Chua, C. A. Davey, Crystal structures of nucleosome core particles containing the “601” strong positioning sequence. *J. Mol. Biol.* 403, 1–10 (2010).
2. Y. Xu, J. McSally, I. Andricioaei, H. M. Al-Hashimi, Modulation of Hoogsteen dynamics on DNA recognition. *Nat. Commun.* 9, 1–10 (2018).
3. D. E. Gilbert, G. A. Van Der Marel, J. H. Van Boom, J. Feigon, Unstable Hoogsteen base pairs adjacent to echinomycin binding sites within a DNA duplex. *Proc. Natl. Acad. Sci. U.S.A.* 86, 3006–3010 (1989).
4. D. E. Gilbert, J. Feigon, The DNA Sequence at Echinomycin Binding Sites Determines the Structural Changes Induced by Drug Binding: NMR Studies of Echinomycin Binding to [d(ACGTACGT)]2 and [d(TCGATCGA)]2. *Biochemistry* 30, 2483–2494 (1991).
5. X. J. Lu, W. K. Olson, 3DNA: A software package for the analysis, rebuilding and visualization of three-dimensional nucleic acid structures. *Nucleic Acids Res.* 31, 5108–5121 (2003).
6. J. A. Cuesta-Seijo, G. M. Sheldrick, Structures of complexes between echinomycin and duplex DNA. *Acta Crystallogr. Sect. D Biol. Crystallogr.* 61, 442–448 (2005).