Supporting Information

NMR solution structure of an N^2-guanine DNA adduct derived from the potent tumorigen dibenzo[*a,l*]pyrene: Intercalation from the minor groove with ruptured Watson-Crick base pairing

Yijin Tang[§], Zhi Liu[§], Shuang Ding[†], Chin H. Lin[§], Yuqin Cai[†], Fabian A. Rodriguez[§], Jane M. Sayer[‡], Donald M. Jerina[⊥], Shantu Amin[||], Suse Broyde[†], Nicholas E. Geacintov[§,]*.

[§]Department of Chemistry and [†]Biology, New York University, New York, NY 10003

[‡] Laboratory of Chemical Physics, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, DHHS, Bethesda, MD 20892

[⊥]Deceased.

[||]Department of Pharmacology,Penn State College of Medicine, Hershey, PA 17033

*Corresponding author: (212) 998-8407, Fax (212) 998-8421, ng1@nyu.edu

Table of Contents

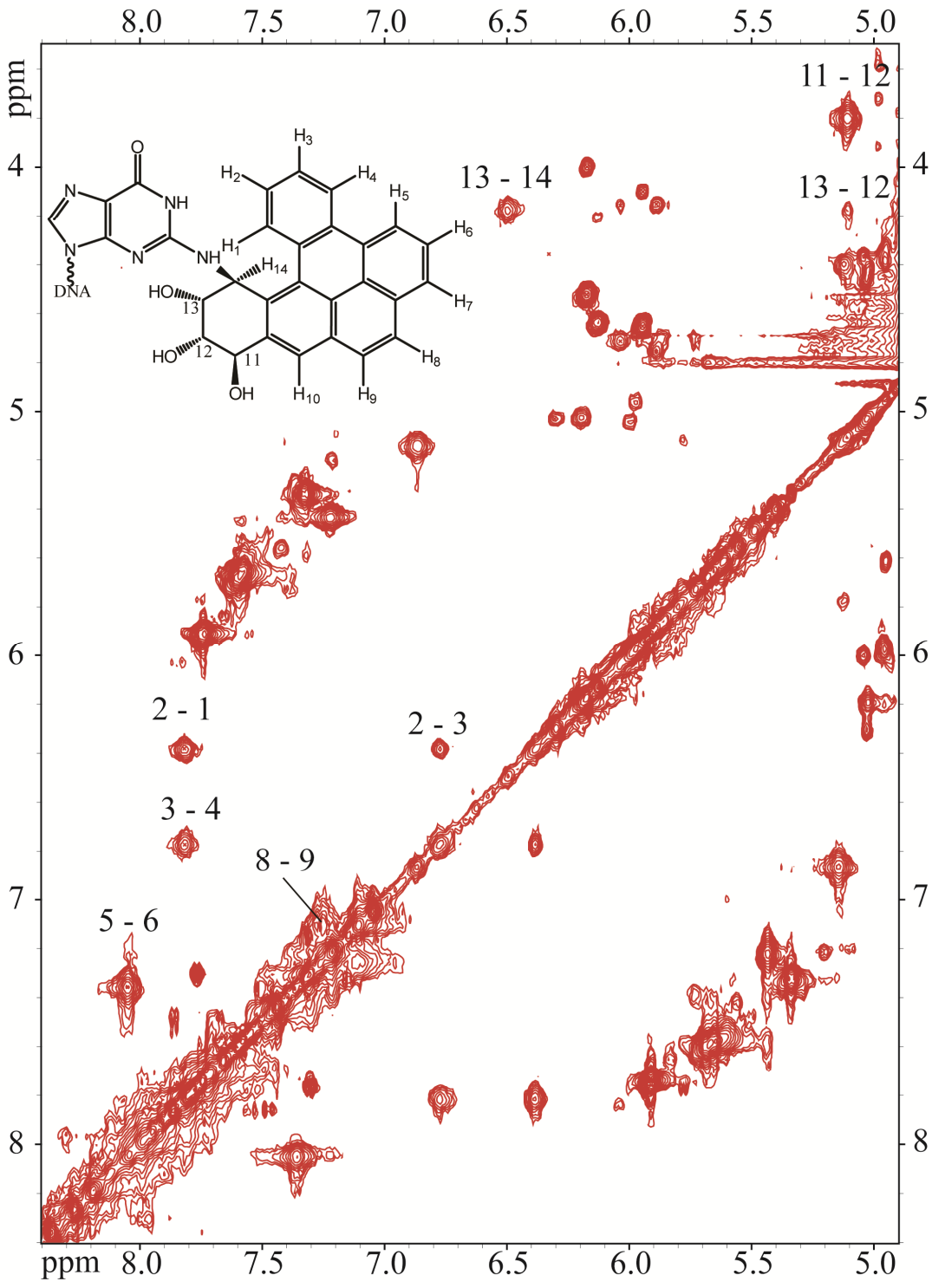

Figure S1. Expanded contour plot of a TOCSY spectrum (80 ms mixing time) of the DB[*a,l*]P-dG lesion in the same 11mer duplex in D_2O aqueous buffer solution using a 500MHz spectrometer with a cryoprobe, at 15°C showing through-bond connectives of protons in the aromatic DB[*a,l*]P residue.

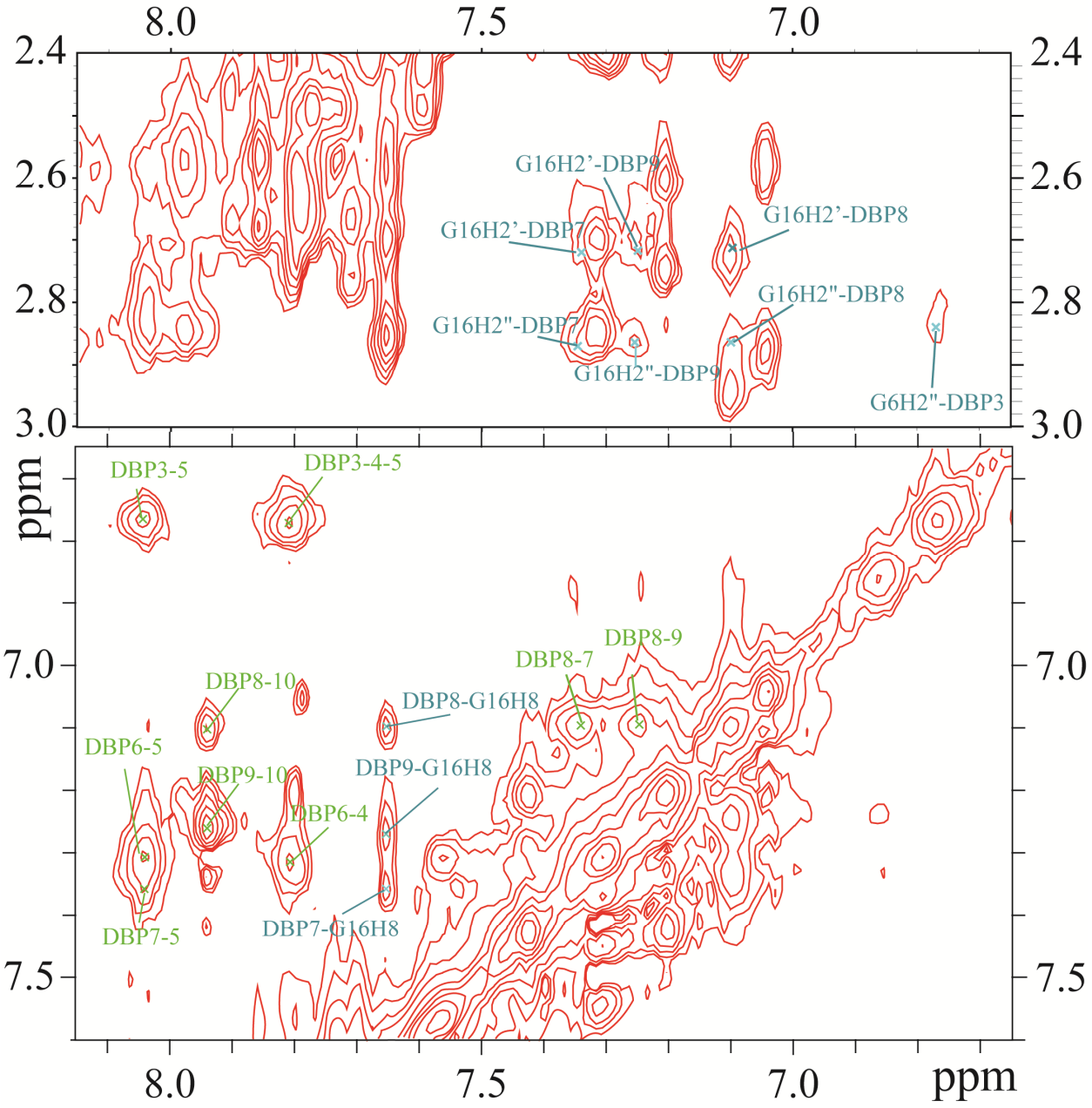

Figure S2. Expanded contour plot of a NOESY spectrum (300 ms mixing time) of the DB[*a,l*]P-dG lesion in the 11mer duplex in D_2O aqueous buffer solution using a 500MHz spectrometer with a cryoprobe, at 15°C showing additional examples of intermolecular NOEs. The DB[*a,l*]P moiety is denoted by DBP.

Table S3. Nucleic acid chemical shifts of the DB[*a,l*]P-dG•dC 11 mer duplex.

	H1'	H2'	H2''	H3'	H4'	H6/H8	H2/H5/Me	Imino& amino*
C1	5.93	2.07	2.49	4.63	4.09	7.73	5.91	N.A.
C2	5.39	2.18	2.42	N.A.	4.11	7.59	5.67	8.56_{b}/6.98_{nb}
A3	6.29	2.73	2.94	5.03	4.44	8.35	7.74	6.44_{nb}/ N.A.
T4	5.84	1.96	2.40	N.A.	N.A.	7.10	1.30	13.43
C5	5.63	1.57	2.21	N.A.	N.A.	7.21	5.43	8.20_{b}/6.64_{nb}
G6	6.09	N.A.	2.82	N.A.	N.A.	8.04	----	11.79/6.51_{nb}
C7	5.72	2.11	2.35	4.69	4.29	7.55	N.A.	7.30_{b}/6.43_{nb}
T8	5.48	2.07	2.40	N.A.	N.A.	7.30	1.49	13.36
A9	6.18	2.70	2.85	5.03	4.41	8.26	7.42	6.22_{nb}/N.A.
C10	5.88	2.02	2.38	4.74	4.14	7.31	5.33	8.21_{b}/6.74_{nb}
C11	6.16	N.A.	2.22	4.51	3.99	7.57	5.63	N.A.
G12	5.67	2.57	2.68	4.79	4.16	7.86	----	N.A.
G13	5.97	2.60	2.75	4.95	4.37	7.80	----	12.80
T14	5.54	2.02	2.30	N.A.	N.A.	7.21	1.36	13.56
A15	5.99	2.57	2.84	5.05	4.36	7.98	7.18*	6.00_{nb}/N.A.
G16	5.77	2.71	2.86	5.13	4.38	7.66	----	11.16
C17	4.25	1.46	1.85	N.A.	N.A.	6.86	5.14	N.A.
G18	5.55	2.67	2.76	N.A.	N.A.	7.71	----	12.36
A19	6.19	2.58	2.88	5.03	4.46	8.19	7.77	5.91_{nb}/N.A.
T20	5.62	1.78	2.17	4.80	4.08	7.04	1.38	13.68
G21	5.60	2.62	2.67	N.A.	4.31	7.79	----	12.90
G22	6.12	2.32	2.49	4.63	4.19	7.77	----	N.A.

*Assigned from water NOESY.

'b' and 'nb' define hydrogen-bonded and non-hydrogen-bonded amino protons. Imino proton: no label.

N.A.: Unable to assign.

Table S4. Nucleic acid chemical shifts of the 5'-CCATCGCTACC-3'• 5'-GGTAGCGATGG-3' 11 mer unmodified duplex (Unpublished data, kindly provided by Dr. Monique Cosman).

	H1'	H2'	H2"	H3'	H4'	H6/H8	H2/H5/Me	Imino& amino*
C1	5.95	2.07	2.50	4.64	4.08	7.72	5.91	$7.93_{b}/6.93_{nb}$
C2	5.44	2.18	2.44	4.85	4.12	7.6	5.70	$8.56_{b}/6.94_{nb}$
A3	6.32	2.75	2.97	5.04	4.45	8.38	7.78	N.A.
T4	5.87	2.01	2.43	4.83	4.17	7.15	1.41	13.58
C5	5.65	2.00	2.38	4.75	4.1	7.4	5.54	$8.38_{b}/6.80_{nb}$
G6	5.85	2.61	2.7	4.94	4.33	7.86		12.83
C7	5.83	2.42	2.63	4.69	4.15	7.32	5.27	$8.16_{nb}/6.55_{nb}$
T8	5.64	2.10	2.43	4.84	4.12	7.36	1.63	13.7
A9	6.19	2.70	2.85	5.00	4.4	8.28	7.49	N.A.
C10	5.91	2.04	2.41	4.81	4.18	7.32	5.38	$8.20_{b}/6.71_{nb}$
C11	6.18	2.22	2.53	4.51	4.00	7.59	5.67	$8.17_{nb}/7.0_{nb}$
G12	5.67	2.53	2.65	4.97	4.31	7.84	----	N.A.
G13	5.99	2.63	2.78	4.95	4.38	7.82	----	12.83
T14	5.63	2.10	2.43	4.86	4.19	7.26	1.43	13.58
A15	6.05	2.70	2.89	5.02	4.39	8.17	7.37	N.A.
G16	5.67	2.45	2.58	4.94	4.35	7.63	----	12.74
C17	5.58	1.83	2.28	4.79	4.08	7.18	5.17	$8.15_{b}/6.28_{nb}$
G18	5.61	2.58	2.79	4.93	4.35	7.85	----	12.67
A19	6.18	2.58	2.87	4.99	4.42	8.15	7.75	N.A.
T20	5.65	1.76	2.17	4.79	4.07	7.02	1.37	13.68
G21	5.62	2.51	2.61	4.98	4.34	7.85	----	12.89
G22	6.12	2.48	2.34	4.63	4.19	7.76	----	N.A.

*Assigned from water NOESY.

'b' and 'nb' define hydrogen-bonded and non-hydrogen-bonded amino protons. Imino proton: no label.

N.A.: Unable to assign.

MD simulation protocol

The duplex starting model was neutralized with 20 Na^+ ions and solvated with explicit water using the LEAP module of AMBER (*43*). A periodic rectangular box of TIP3P water with 10.0Å buffer was created around the DNA. The particle-mesh Ewald (PME) method with a 9.0 Å cutoff for the nonbonded interactions was used in the subsequent energy minimizations and dynamic simulations. Initially, 500 steps of steepest descent minimization followed by 500 steps of conjugate gradient minimization were conducted for the waters and sodium ions with 500 kcal/(mol·$Å^2$) restraints placed on the DNA duplex. Then, 1000steps of steepest descent minimization followed by 1500 steps of conjugate gradient minimization were carried out for the entire system without restraints. A 2fs time step and the SHAKE algorithm were applied in the MD simulations. The system was heated from 0 K to 300K over 20pswith a weak 10 kcal/(mol$Å^2$) restraint on the DNA at constant volume. The equilibration was continued with further 100psof simulation without restraint at constant atmospheric pressure. Finally, production was carried out for 3 ns at 300K with a 4.0 ps heat bath coupling parameter.

Table S1. AMBER atom type, connection type, and partial charge assignments for the DB[*a,l*]P-dG adduct.

Atom name	Atom type	Connection type	Partial Charge
P	P	M	1.221554
O1P	O2	E	-0.79326
O2P	O2	E	-0.79326
O5'	OS	M	-0.495
C5'	CT	M	-0.02793
H5'1	H1	E	0.083211
H5'2	H1	E	0.083211
C4'	CT	M	0.154291
H4'	H1	E	0.117104
O4'	OS	S	-0.36758
C1'	CT	B	0.080502
H1'	H2	E	0.147042
N9	N*	B	0.001865
C8	CK	S	0.140741
H8	H5	E	0.179994
C4	CB	S	0.04708
C5	CB	B	0.296566
N7	NB	E	-0.59984
C6	C	B	0.43052
O6	O	E	-0.53893
N1	NA	B	-0.27652
H1	H	E	0.29729
C2	CA	B	0.18175
N3	NC	E	-0.33615
N2	N2	B	-0.36203
H21	H	E	0.307353
C14(B)[a]	CT	B	0.040413
H14(B)	H1	E	0.07281
C13(B)	CT	3	0.101951
H13(B)	H1	E	0.094819
O13(B)	OH	S	-0.6675
HO13(B)	HO	E	0.43427
C12(B)	CT	3	0.065868
H12(B)	H1	E	0.116362
O12(B)	OH	S	-0.65771
HO12(B)	HO	E	0.450859
C11(B)	CT	3	0.156184
H11(B)	H1	E	0.099265
O11(B)	OH	S	-0.65249
HO11(B)	HO	E	0.417551
C16(B)	CA	S	0.029842
C10(B)	CA	B	-0.22052
H10(B)	HA	E	0.157593
C17(B)	CA	B	0.059497

C18(B)	CA	E	-0.01349
C9(B)	CA	B	-0.13114
H9(B)	HA	E	0.137418
C8(B)	CA	B	-0.25707
H8(B)	HA	E	0.159075
C20(B)	CA	B	0.129922
C21(B)	CA	E	0.080957
C7(B)	CA	B	-0.19611
H7(B)	HA	E	0.151182
C6(B)	CA	B	-0.18236
H6(B)	HA	E	0.157873
C5(B)	CA	B	-0.16249
H5(B)	HA	E	0.149984
C22(B)	CA	S	-0.02138
C23(B)	CA	S	0.058432
C4(B)	CA	B	-0.19154
H4(B)	HA	E	0.146901
C3(B)	CA	B	-0.13701
H3(B)	HA	E	0.149525
C2(B)	CA	B	-0.1179
H2(B)	HA	E	0.132053
C1(B)	CA	B	-0.20264
H1(B)	HA	E	0.170153
C24(B)	CA	S	0.058928
C19(B)	CA	S	-0.09264
C15(B)	CA	E	0.058864
C3'	CT	M	0.063028
H3'	H1	E	0.116117
C2'	CT	B	-0.10959
H2'1	HC	E	0.069536
H2'2	HC	E	0.069536
O3'	OS	M	-0.5228

[a] (B): DB[*a,l*]P moiety.

Table S2. Added force field parameters for the DB[*a,l*]P-dG adduct.

angle	K_θ(kcal/(mol radian^2))	θ_{eq} (deg)
OH-CT-CA	50	109.5
H1-CT-CA	50	109.5
N2-CT-CA	80	111.2

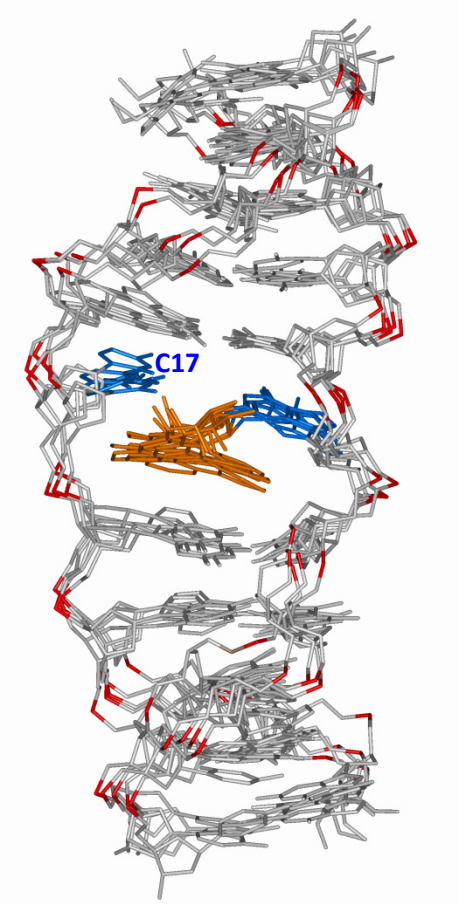

Figure S3. Superpositioned five structures of the DB[*a*,*l*]P-dG modified duplex that were extracted from the restrained MD simulation to best represent the NMR data; the view is from the minor groove. The DB[*a,l*]P moiety is in orange, the modified guanine and its partner cytosine are in blue, the phosphorus atoms are in red, and the rest of the DNA is in light gray. Hydrogen atoms in the DNA duplexes are not displayed for clarity.

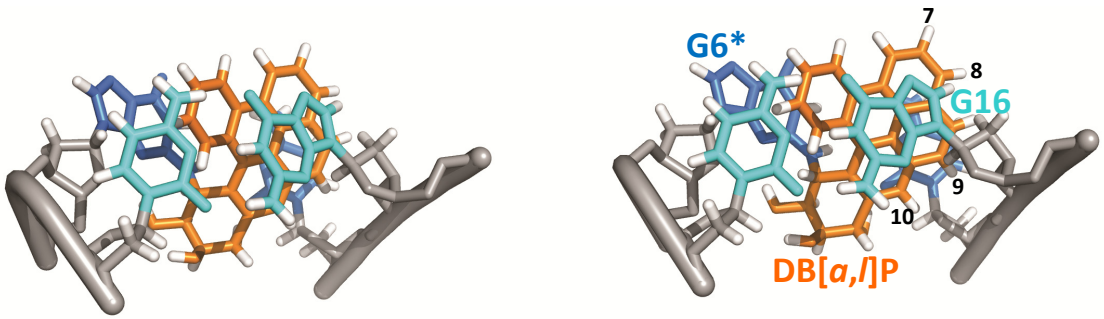

Figure S4. Stereo view of theDB[*a,l*]P moiety intercalated between dG6*:dC17 and dC7:dG16. The DB[*a,l*]P is in orange, the modified guanine and its partner cytosine are in blue, the 3'-side base pair C7:G16 are in cyan, displayed hydrogen atoms are in white, and the rest of the DNA is in light gray. DB[*a,l*]P protons are numbered.

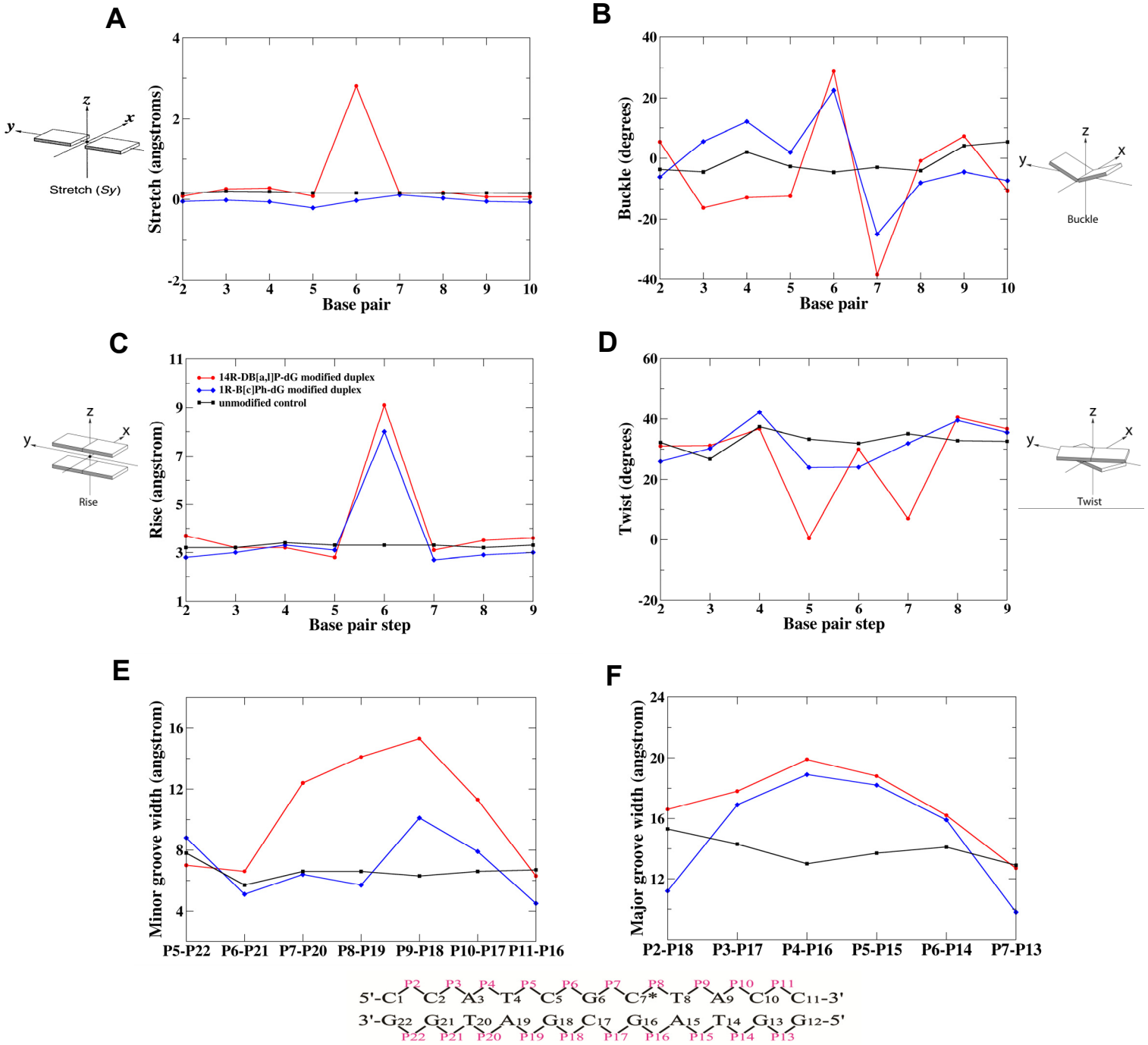

Figure S5. Average values of helicoidal parameters and groove widths for the five selected DB[*a,l*]P-dG structures (red). Unmodified control data (black) is from *(67)*, and 1*R*-B[*c*]Ph-N^2-dG (blue) is from the nine NMR solution structures deposited in the PDB (1HX4) (*38*). For Buckle and Propeller, C1·G22 is base pair 1, C2·G21 is base pair 2, etc.; For Rise and Twist, C1·G22 to C2·G21 is step 1, C2·G21 to A3·T20 is step 2, etc. The phosphate defining the groove dimensions are illustrated. The cartoons of helical parameters are from Lu X. and Olson W. K. *Nucleic Acids Res.*(2003) 31, 5108-5121.