Supporting Information

Nucleotide excision repair efficiencies of bulky carcinogen-DNA adducts are governed by a balance between stabilizing and destabilizing interactions

Yuqin Cai[a], *Nicholas E. Geacintov*[b], *and Suse Broyde*[a,*]

Department of [a]Biology and [b]Chemistry, New York University, New York, N.Y., 10003

***Corresponding Author:** Suse Broyde, broyde@nyu.edu Tel. (212) 998-8231 Fax. (212) 995-4015

Running Title: Lesion topology effects on NER

Supporting Methods

MD computation protocols. Our computational protocols follow recommendations of the AMBER development team (http://amber.scripps.edu/tutorials/basic/tutorial1/section5.htm). First, the starting models were subjected to 50 steps of conjugate gradient energy minimization, using a distance dependent dielectric function with dielectric constant 4.0. Then each system was neutralized with 20 Na^+ counterions and solvated with explicit water using the LEAP Module of the AMBER 9 simulation package.[1] A periodic rectangular box of TIP3P water[2] with 10.0 Å buffer was created around the DNA for each sequence. The Particle-Mesh Ewald[3, 4] method with 9.0 Å cutoff for the non-bonded interactions was used in subsequent energy minimizations and MD simulations. Subsequently, 500 steps of steepest descent minimization followed by 500 cycles of conjugate gradient minimization were conducted for the waters and counterions with 500 $kcal \cdot mol^{-1} \cdot Å^{-2}$ restraint on the DNA. Then, 500 steps of conjugate gradient minimization were carried out for the whole system without restraints. A 2.0 fs time step and the SHAKE algorithm[5] were applied in the MD simulations. All other parameters were default values in the AMBER 9 simulation package. The system was heated from 0 K to 300 K over 20 ps with the DNA fixed with a weak restraint of 10.0 $kcal \cdot mol^{-1} \cdot Å^{-2}$ at constant volume, using the Berendsen coupling algorithm[6] with a 1.0 ps coupling parameter. A constant pressure dynamics at 300K for 60 ps followed to further equilibrate the system. Finally, a 30.0 ns production was conducted at a temperature of 300 K and constant pressure of 1 Atm. Temperature and pressure coupling constants were both 1.0 ps.

Best representative structures. The best representative structure in the MD trajectory for each adduct was obtained by using the cluster analysis option in MOIL-View.[7] The best representative structure is a real frame from the ensemble; as determined by a cluster analysis, it is the most populated conformation or the conformation that is the closest to all other snapshots in an ensemble.

Hydrogen bond quality index analyses. We employed our hydrogen bond quality index (HBI),[8] to quantitatively assess the quality of Watson-Crick hydrogen bonding, in terms of the deviation from ideal Watson-Crick hydrogen bond distances and angles:

$$I_H^{Distance} = \sum_{D-H\ldots A}[(d_{DA} - d_{DA}^0)^2 + (1+\cos\gamma)^2]$$

$$I_H^{Angle} = \sum_{D-H\ldots A}(1+\cos\gamma)^2$$

where d_{DA} is the instantaneous donor-acceptor distance, d_{DA}^0 is an ideal donor-acceptor distance [9] [O6 (G) to N4 (C) is 2.91 Å, N1 (G) to N3 (C) is 2.95 Å, and N2 (G) to O2 (C) is 2.86 Å] and γ is the instantaneous donor-hydrogen···acceptor (D-H···A) hydrogen bond angle with an ideal value of 180°. The summation is over all the Watson-Crick hydrogen bonds in a base pair over the trajectory. The lower the value of I_H, the better the quality of the Watson-Crick hydrogen bonding. I_H is zero when the Watson-Crick hydrogen bonding is ideal during the dynamics. However, in real, even unmodified DNA,

sequence-dependent deviations from ideal Watson-Crick hydrogen bonding are normal.[10] Therefore, comparing the I_H value of a modified step to its analogue step in the unmodified control duplex provides an estimate of the perturbation of its Watson-Crick hydrogen bond upon modification.

Supporting Tables and Figures

Table S1A. Hydrogen bond occupancies (%)[a] for the central 3-mer [a]

Hydrogen bond	B[*c*]Ph-N^6-dA		B[*a*]P-N^6-dA		DB[*a,l*]P-N^6-dA		Unmodified
	R	***S***	***R***	***S***	***R***	***S***	
C5:G18							
N4-H41…O6	99	96	93	92	99	98	97
N2-H22…O2	100	100	99	97	100	100	100
N1-H1…N3	100	100	95	99	100	100	100
A6:T17							
N6-H6…O4	97	90	98	0	98	96	95
N3-H3…N1	99	95	84	0	98	95	99
C7:G16							
N4-H41…O6	96	98	92	94	97	98	96
N2-H21…O2	100	99	100	98	100	100	99
N1-H1…N3	99	100	98	100	100	100	100

[a] Criteria for hydrogen bonding: heavy atom-heavy atom distance < 3.3 Å and donor-hydrogen-acceptor angle > 140 °.

Table S1B. Hydrogen bond quality for the Watson-Crick base pairs adjacent to the lesion site and their corresponding unmodified control base pairs.

HBI-Angle

Base pair	B[*c*]Ph-N^6-dA		B[*a*]P-N^6-dA		DB[*a,l*]P-N^6-dA		Unmod-
	R	***S***	***R***	***S***	***R***	***S***	**ified**
C5:G18	306	649	555	795	240	367	301
A6:T17	218	258	220	38257	195	159	242
C7:G16	369	390	318	655	400	350	356

HBI-Distance

Base pair	B[*c*]Ph-N^6-dA		B[*a*]P-N^6-dA		DB[*a,l*]P-N^6-dA		Unmod-
	R	***S***	***R***	***S***	***R***	***S***	**ified**
C5:G18	698	817	15979	1661	695	742	924
A6:T17	997	2525	2770	345561	1032	1580	1143
C7:G16	1218	867	2200	1364	835	797	991

Table S2. Ensemble average values with standard deviations (given in parentheses) for the structural properties at the lesion site.

	B[*c*]Ph-*N*^{6}-dA		**B[*a*]P-*N*^{6}-dA**		**DB[*a,l*]P-*N*^{6}-dA**		**Unmodified**	
	R	***S***	***R***	***S***	***R***	***S***		
Buckle (°)	−31.7 (7.8)	41.8 (8.4)	−36.0 (7.5)	36.4 (8.1)	−26.7 (7.5)	29.2 (7.5)	4.4 (11.1)	
Propeller (°)	24.9 (8.6)	−40.6 (10.0)	37.0(10.1)	−48.2 (12.4)	20.6 (7.4)	−28.0 (8.2)	−6.8 (9.1)	
Twist (°)	20.0 (7.9)	4.7 (10.5)	4.6 (6.3)	11.2 (14.8)	15.0 (7.6)	−13.2 (6.0)	32.0 (5.4)	
Roll (°)	11.7 (8.3)	21.7 (7.9)	9.4 (7.7)	17.9 (9.6)	7.6 (8.0)	21.8 (5.8)	−4.2 (6.9)	
Rise (Å)	7.4 (0.4)	7.3 (0.5)	8.8 (0.6)	6.4 (0.6)	7.6 (0.4)	7.1 (0.4)	3.3 (0.5)	3.4 (0.4)
Tilt (°)	−8.7 (5.2)	−14.7 (4.7)	−14.8 (4.6)	−21.3 (5.7)	−10.6 (5.8)	−16.5 (4.1)	0.4 (5.7)	1.8 (5.1)
Opening (°)	10.9 (8.8)	19.3 (10.1)	6.5 (8.9)	−15.4 (12.1)	12.8 (7.8)	13.1 (6.9)	2.2 (6.2)	

Table S3. Ensemble average values with standard deviations (given in parentheses) for groove widths (Å) [a]

Minor groove width (Å)	**B[*c*]Ph-N^6-dA**		**B[*a*]P-N^6-dA**		**DB[*a,l*]P-N^6-dA**		**Unmodified**
	R	***S***	***R***	***S***	***R***	***S***	
P6-P21	6.8 (1.4)	6.9 (1.6)	7.7 (1.7)	5.2 (1.8)	7.5 (3.1)	6.6 (2.0)	6.0 (1.7)
P7-P20	11.1 (2.1)	9.9 (1.6)	13.1 (2.2)	9.5 (1.3)	9.6 (3.4)	9.6 (1.6)	8.1 (2.7)
P8-P19	13.3 (1.9)	12.0 (2.0)	17.6 (1.5)	11.5 (1.8)	13.4 (1.9)	13.5 (1.9)	7.9 (2.6)
P9-18	10.2 (1.7)	9.2 (1.7)	12.2 (1.6)	12.8 (1.5)	9.5 (1.9)	13.1 (2.4)	5.3 (1.7)
P10-P17	8.3 (1.5)	9.6 (1.8)	8.0 (1.5)	9.2 (1.6)	7.6 (1.9)	9.1 (1.9)	6.4 (1.9)

Major groove width (Å)	**B[*c*]Ph-N^6-dA**		**B[*a*]P-N^6-dA**		**DB[*a,l*]P-N^6-dA**		**Unmodified**
	R	***S***	***R***	***S***	***R***	***S***	
P3-P17	17.1 (1.6)	16.1 (1.6)	17.9 (1.2)	17.7 (1.8)	17.2 (1.9)	16.4 (1.6)	14.4 (1.7)
P4-P16	17.5 (1.6)	17.5 (1.7)	16.6 (1.4)	18.6 (1.9)	17.2 (1.5)	18.7 (1.4)	13.4 (2.0)
P5-P15	17.5 (1.7)	17.6 (1.6)	16.7 (1.5)	16.0 (2.6)	16.8 (1.9)	17.8 (1.9)	14.5 (1.9)
P6-P14	16.9 (1.8)	16.4 (2.0)	17.1 (1.3)	17.0 (2.3)	15.9 (1.7)	17.6 (1.8)	14.4 (1.9)

[a]Definitions for minor groove and major groove widths are given in Figure S4.

Table S4. Ensemble average van der Waals interaction energies (kcal/mol) between PAH aromatic rings and nearby bases [a].

VDW interaction energy (kcal/mol)	**B[*c*]Ph-N^6-dA**		**B[*a*]P-N^6-dA**		**DB[*a,l*]P-N^6-dA**	
	R	***S***	***R***	***S***	***R***	***S***
C5− PAH	−4.4 (1.0)	−0.5 (0.1)	−6.2 (0.8)	−0.3 (0.1)	−5.7 (1.0)	−0.6 (0.1)
A6− PAH	−2.2 (1.2)	−3.3 (1.1)	−1.6 (0.6)	−2.0 (0.5)	−3.0 (1.1)	−2.9 (1.1)
C7− PAH	−0.3 (0.1)	−2.4 (0.7)	−0.2 (0.0)	−2.4 (0.7)	−0.3 (0.1)	−2.2 (0.7)
G16− PAH	−0.6 (0.1)	−8.1 (0.8)	−0.6 (0.1)	−9.1 (0.8)	−0.8 (0.2)	−9.8 (1.1)
T17− PAH	−6.6 (0.6)	−7.0 (0.7)	−8.5 (0.7)	−7.7 (0.9)	−9.0 (0.7)	−8.9 (0.7)
G18− PAH	−6.9 (1.2)	−0.5 (0.1)	−6.4 (1.0)	−1.3 (0.5)	−9.3 (1.0)	−0.5 (0.1)
Total	**−21.2 (2.0)**	**−21.8 (1.8)**	**−23.5 (1.6)**	**−21.6 (1.7)**	**−28.2 (2.1)**	**−25.1 (1.8)**

[a] Standard deviations are given in parentheses.

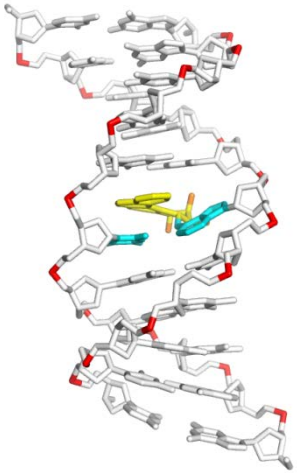
Table S5. Distortion energy (kcal/mol)[a]

	MM-PBSA energy	Distortion energy
***R*-B[*c*]Ph-N^6-dA**	−1074.4	13.0
***S*-B[*c*]Ph-N^6-dA**	−1073.0	14.4
***R*-B[*a*]P-N^6-dA**	−1072.7	14.7
***S* -B[*a*]P-N^6-dA**	−1069.6	17.8
R-DB[*a,l*]P-N^6-dA	−1076.6	10.8
***S* -DB[*a,l*]P-N^6-dA**	−1072.5	14.9
Unmodified	−1087.4	0

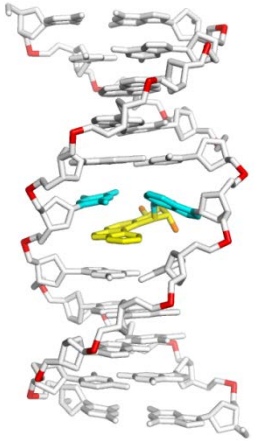
[a] Only the central 3-mers were considered. See Materials and Methods for details.

Table S6. Box size and numbers of waters added to the MD simulation starting model for each system

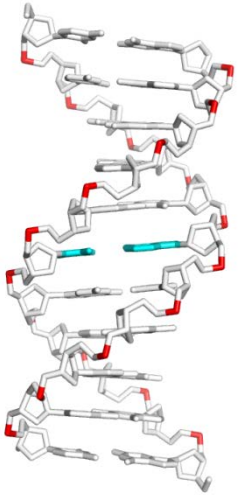
System	Box size (Å^3)	Number of water added
R **-B[*c*]Ph-*N*^{6}-dA**	51 x 49 x 69	3924
S **-B[*c*]Ph-*N*^{6}-dA**	49 x 51 x 72	4079
R**-B[*a*]P-*N*^{6}-dA**	55 x 58 x 68	5252
S **-B[*a*]P-*N*^{6}-dA**	58 x 64 x 60	5369
R **-DB[*a,l*]P-*N*^{6}-dA**	51 x 48 x 74	4157
S **-DB[*a,l*]P-*N*^{6}-dA**	49 x 51 x 69	3911
Unmodified duplex	50 x 51 x 67	3872



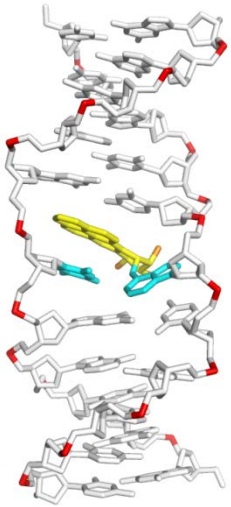
***R*-B[*c*]Ph-*N*^{6}-dA**



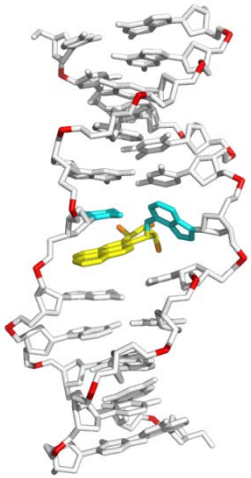
***S*-B[*c*]Ph-*N*^{6}-dA**



Unmodified

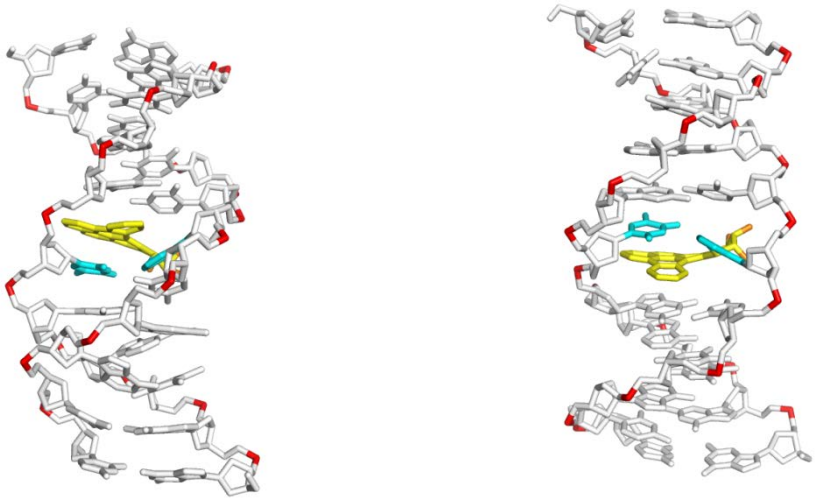


***R*-B[*a*]P-*N*^{6}-dA**



***S*-B[*a*]P-*N*^{6}-dA**

Figure S1. (First of two pages)



***R*-DB[*a,l*]P-N^6-dA** ***S*-DB[*a,l*]P-N^6-dA**

Figure S1. The initial models (See Methods) for the *R* and *S* B[*c*]Ph-N^6-dA, B[*a*]P-N^6-dA, and DB[*a,l*]P-N^6-dA adducts and their unmodified control in the CAC sequence context (shown in Figure 1D). The color codes for the modified duplexes are: PAHs, yellow, A6*(A6):T17 base pair, cyan; the rest of the DNA is white, except for the phosphate atoms, which are red. Hydrogen atoms and pendant phosphate oxygen atoms in the DNA duplexes are not displayed for clarity.

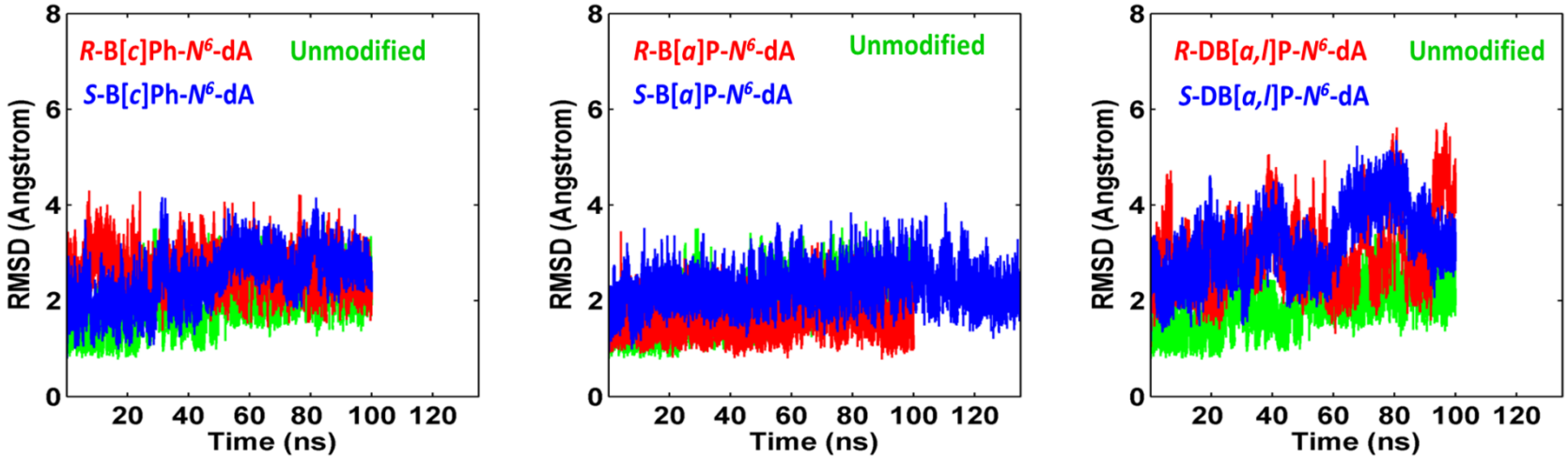


Figure S2. Time dependence of RMSDs for the *R* (red) and *S* (blue) PAH containing adducts and the unmodified control duplex (green). Ensemble average values and standard deviations are given. Only the central 9-mers is considered in this analysis. The ensemble averages and standard deviations over the 100.0 ns MD simulations are 2.6 ± 0.4 Å, 2.5 ± 0.5 Å, 1.6 ± 0.3 Å, 2.2 ± 0.4 Å, 3.0 ± 0.8 Å, 3.2 ± 0.7 Å and 1.8 ± 0.5Å, for the *R* -B[*c*]Ph-N^6-dA, *S* -B[*c*]Ph-N^6-dA, *R*-B[*a*]P-N^6-dA, *S* -B[*a*]P-N^6-dA, *R* -DB[*a,l*]P-N^6-dA, *S* -DB[*a,l*]P-N^6-dA and the unmodified control, respectively. Note that for the *S* -B[*a*]P-N^6-dA adduct, the trajectory from 45 ns to 135 ns is analyzed.

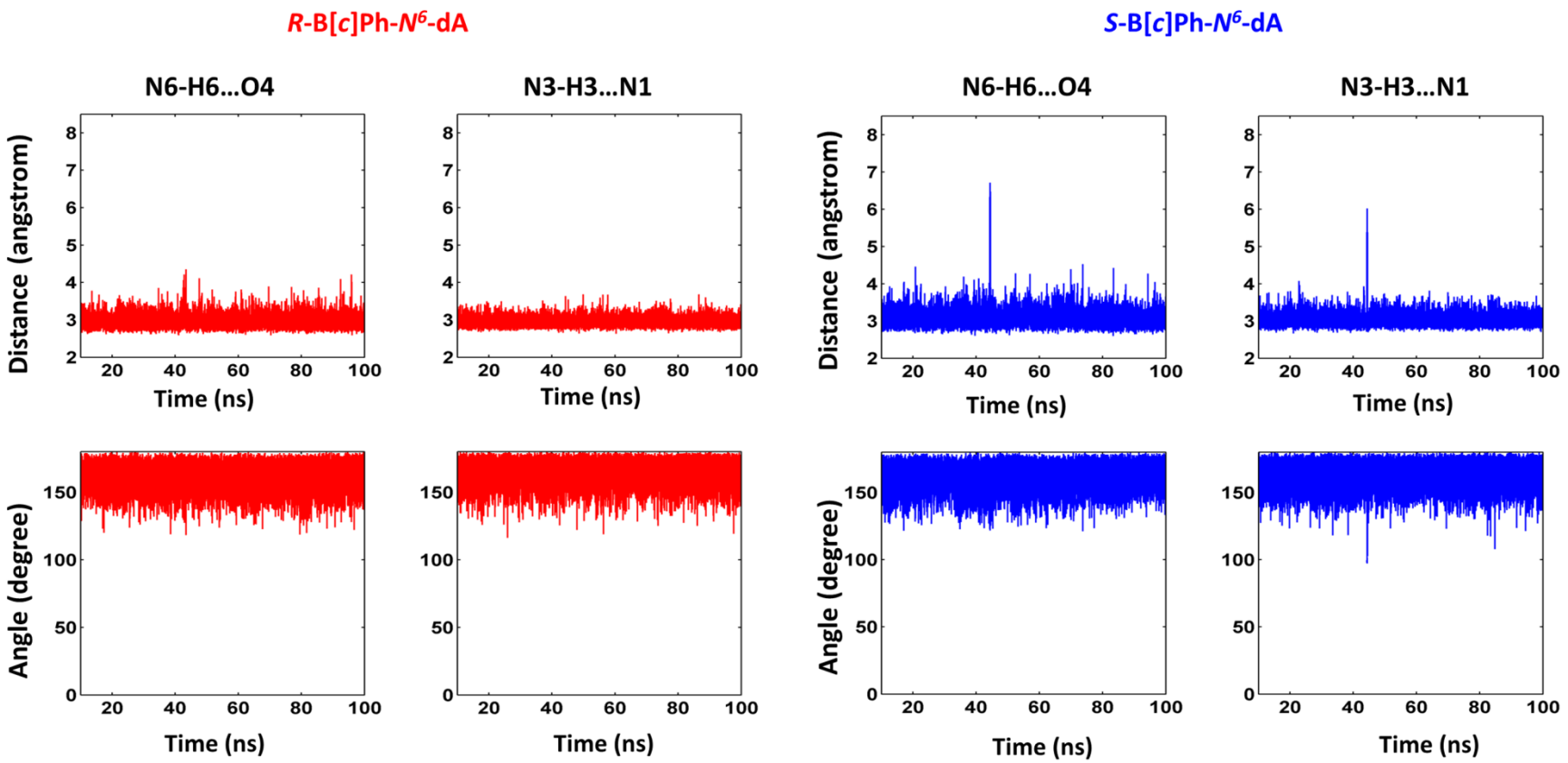


Figure S3. (First of four pages)

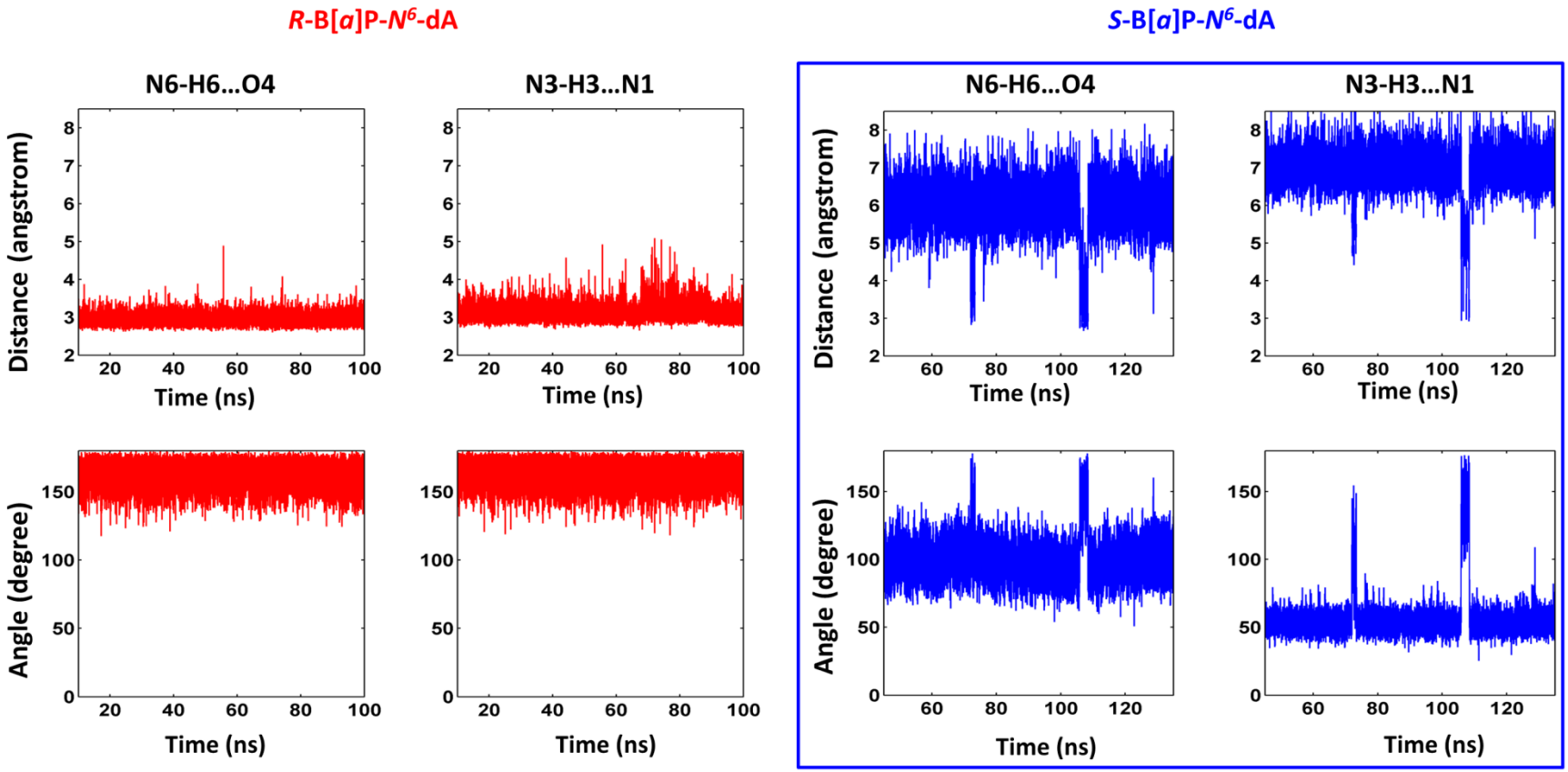


Figure S3. (Second of four pages)

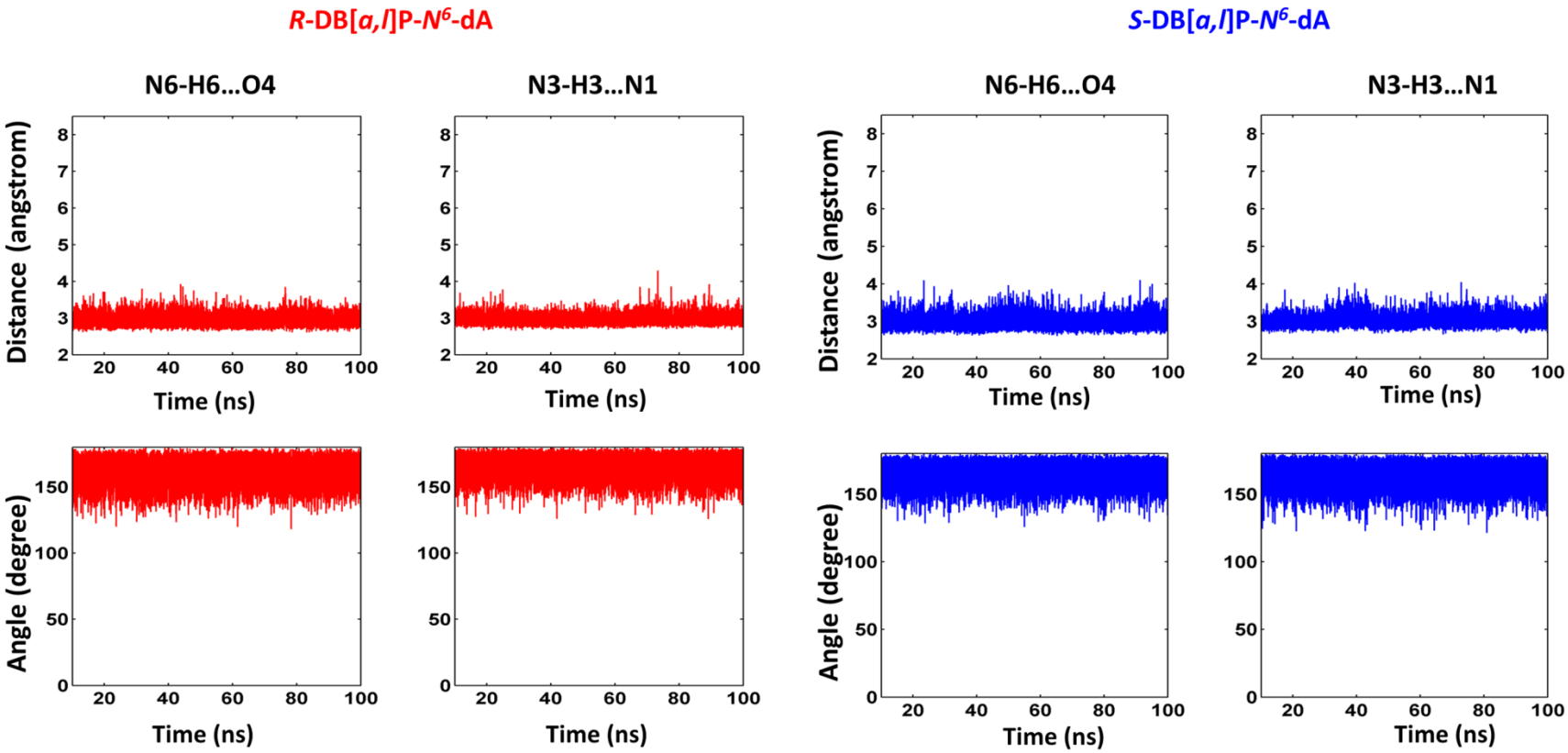


Figure S3. (Third of four pages)

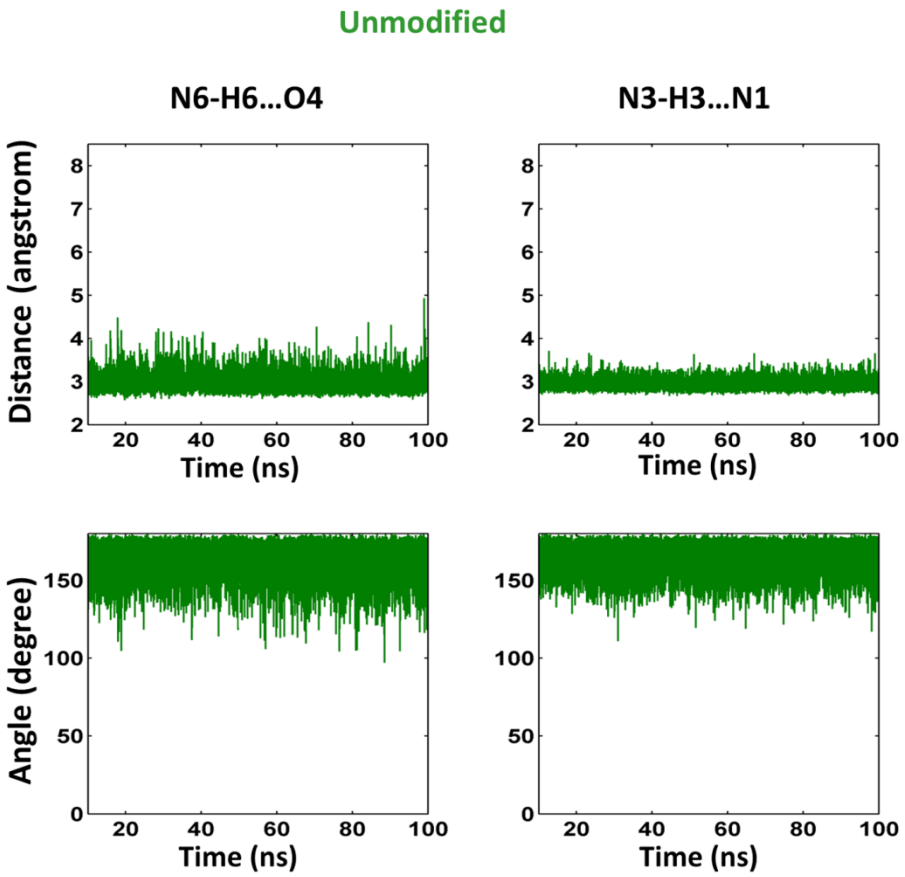
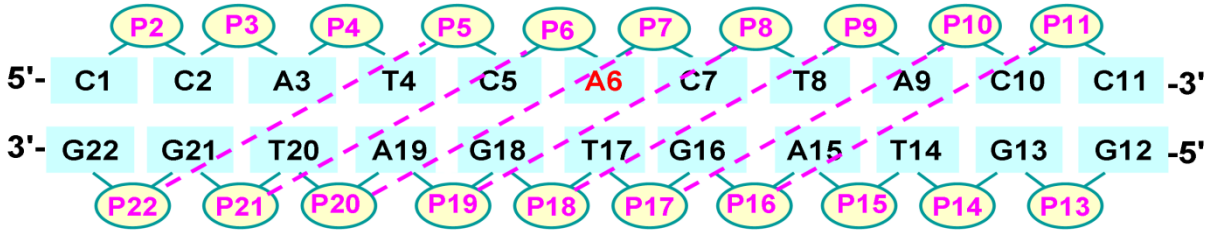


Figure S3. Time dependences of hydrogen bond distances and angles at A6*:T17 for lesion-containing adducts and at analogous A6:T17 site for the unmodified control. Note that the *S* B[*a*]P-N^6-dA with a blue box, is the only one displaying very dynamic hydrogen bond disturbance.

(A)



(B)

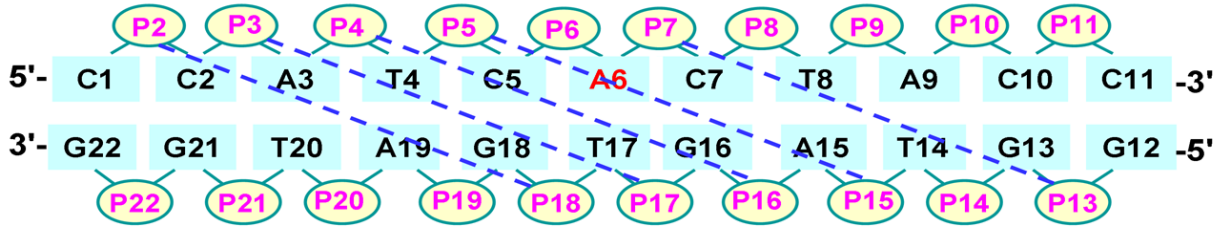


Figure S4. Definitions for minor groove (A) and major groove (B) widths.

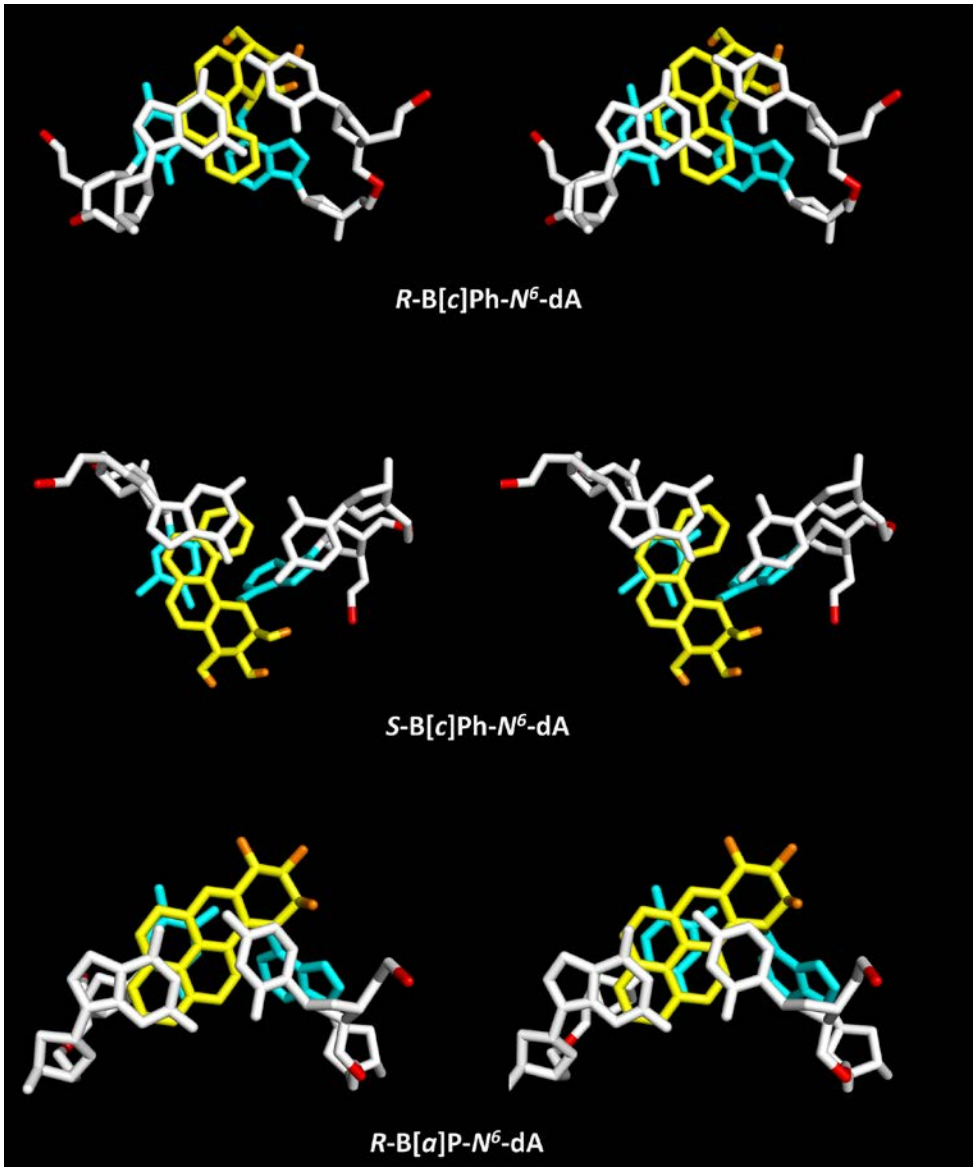


Figure S5. (First of two pages)

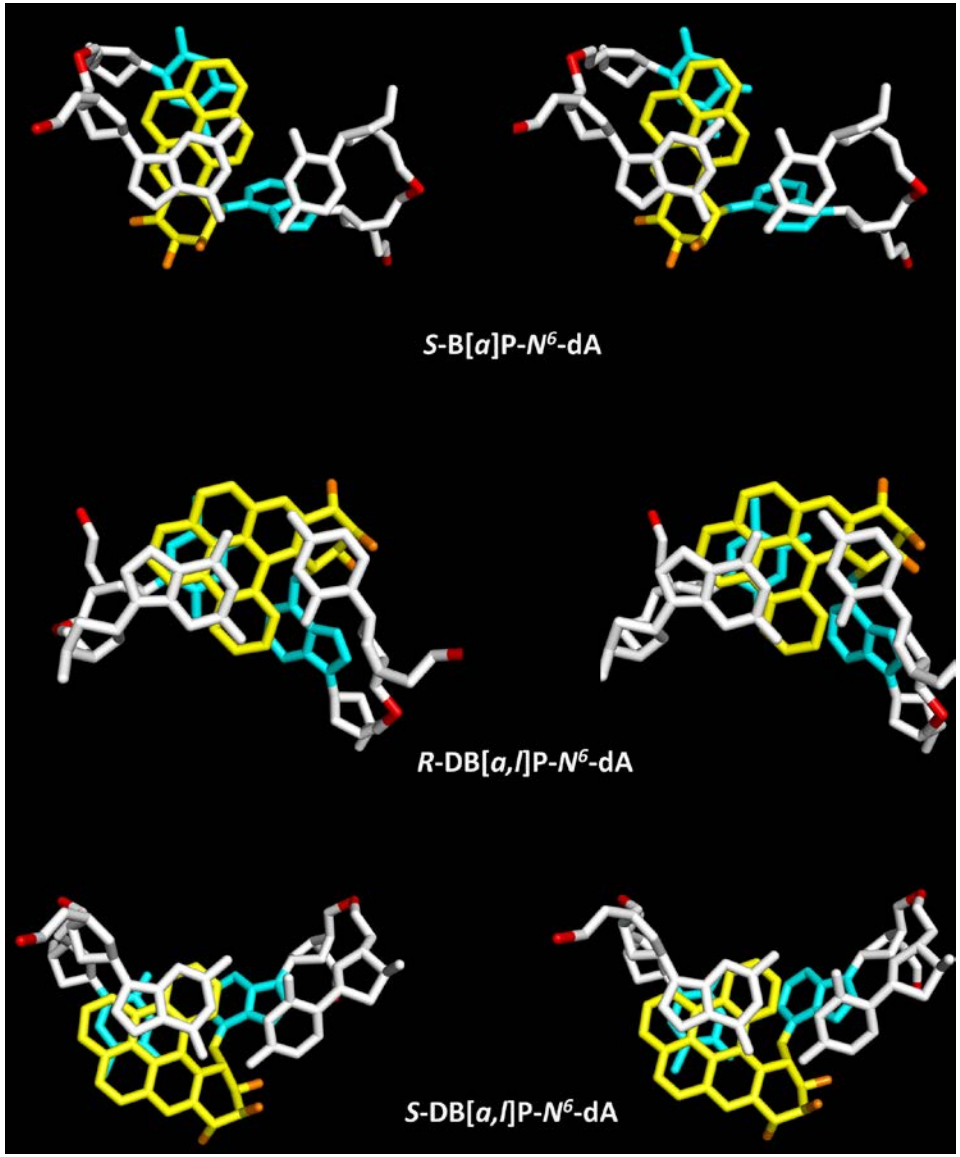


Figure S5. Stereo views looking down the helix axis of the intercalation pockets showing the stacking interactions.

References

1. Case, D. A., Darden, T. A., Cheatham III, T. E., Simmerling, C. L., Wang, J., Duke, R. E., Luo, R., Merz, K. M., Pearlman, D. A., Crowley, M., Walker, R. C., Zhang, W., Wang, B., Hayik, S., Roitberg, A., Seabra, G., Wong, K. F., Paesani, F., Wu, X., Brozell, S., Tsui, V., Gohlke, H., Yang, L., Tan, C., Mongan, J., Hornak, V., Cui, G., Beroza, P., Mathews, D. H., Schafmeister, C., Ross, W. S., and Kollman, P. A. (2006) AMBER 9, University of California, San Francisco, CA.
2. Jorgensen, W. L., Chandreskhar, J., Madura, J. D., Imprey, R. W., and Klein, M. L. (1983) Comparison of simple potential functions for simulating liquid water.*J. Chem. Phys. 79*, 926-935.
3. Darden, T., York, D., and Pedersen, L. (1993) Particle mesh Ewald: an *N* log(*N*) method for Ewald sums in large systems.*J. Chem. Phys. 98*, 10089-10092.
4. Essmann, U., Perera, L., Berkowitz, M. L., Darden, T., Lee, H., and Pederson, L. G. (1995) A smooth particle mesh Ewald method.*J. Chem. Phys. 103*, 8577-8593.
5. Ryckaert, J. P., Ciccotti, G., and Berendsen, H. J. C. (1977) Numerical integration of cartesian equations of motion of a system with constraints: molecular dynamics of *n*-alkanes.*J. Comp. Phys. 23*, 327-341.
6. Berendsen, H. J. C., Postma, J. P. M., van Gunsteren, W. F., DiNola, A., and Haak, J. R. (1984) Molecular dynamics with coupling to an external bath.*J. Chem. Phys. 81*, 3684-3690.
7. Simmerling, C., Elber, R., and Zhang, J. (1995), pp 241-265, Netherlands: Kluwer.
8. Hingerty, B. E., Figueroa, S., Hayden, T. L., and Broyde, S. (1989) Prediction of DNA structure from sequence: a build-up technique.*Biopolymers 28*, 1195-1222.
9. Saenger, W. (1983), In *Principles of Nucleic Acid Structure* p226, Springer-Verlag, New York.
10. Calladine, C. R., and Drew, H. R. (1997), Morgan Kaufmann.