Electronic Supplementary Material (ESI) for Toxicology Research.

Effect of base sequence context on the conformational heterogeneity of aristolactam-I adducted DNA: structural and energetic insights into sequence-dependent repair and mutagenicity

Preetleen Kathuria,[1] Purshotam Sharma[1,2] and Stacey D. Wetmore[1]*

[1]Department of Chemistry and Biochemistry, University of Lethbridge, 4401 University Drive West, Lethbridge, Alberta Canada T1K 3M4

[2]Centre for Computational Sciences, Central University of Punjab, Bathinda, Punjab India 151001

Supporting Information

(13 Pages)

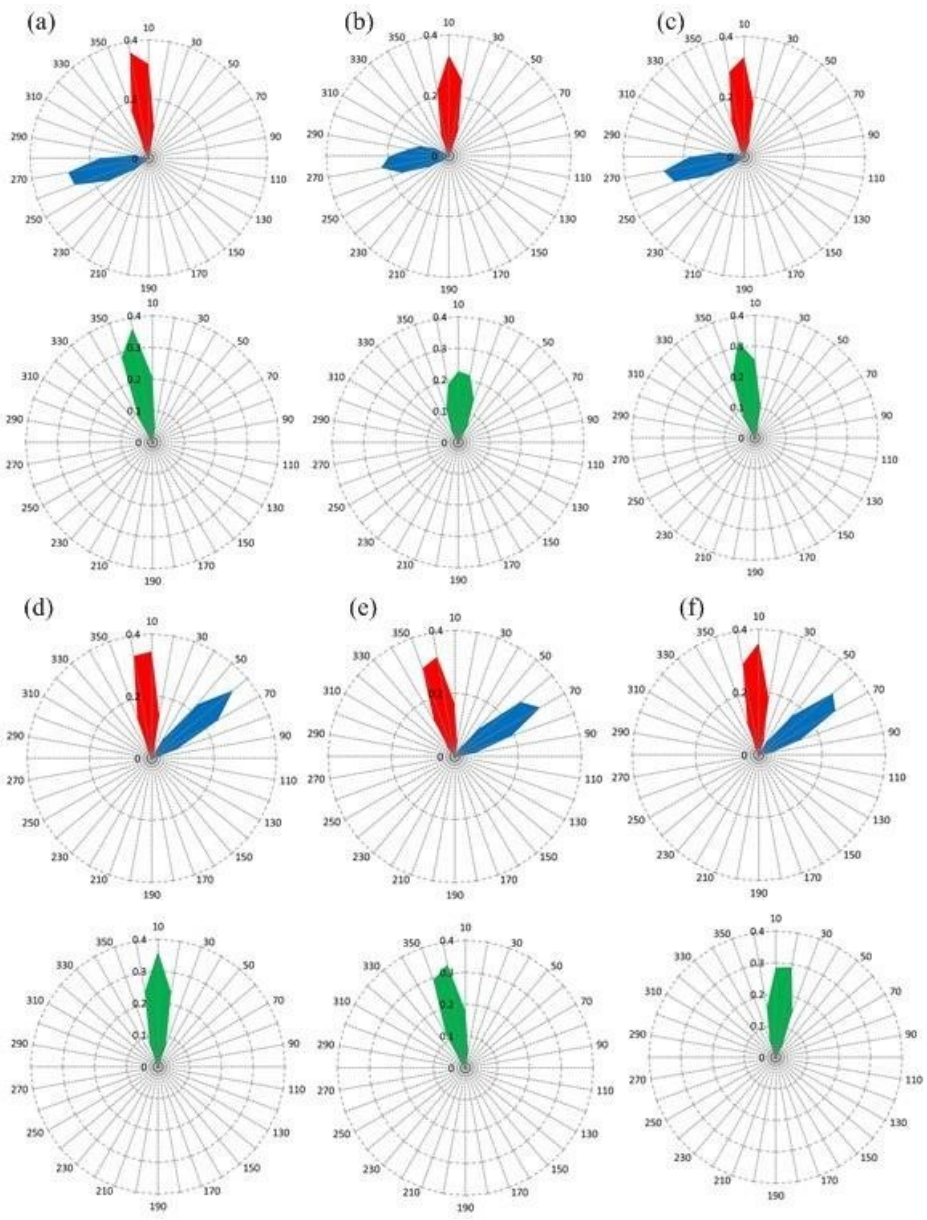

Figure S1. Radar plots for the probability distributions of χ (blue), θ (red) and ϕ (green) for *anti* (a, b and c) or *syn* (d, e and f) ALI-N^6-dA in the GXG sequence in the base-displaced intercalated (a and d), 5'–intercalated (b and e) or 3'–intercalated (c and f) adducted DNA conformations.

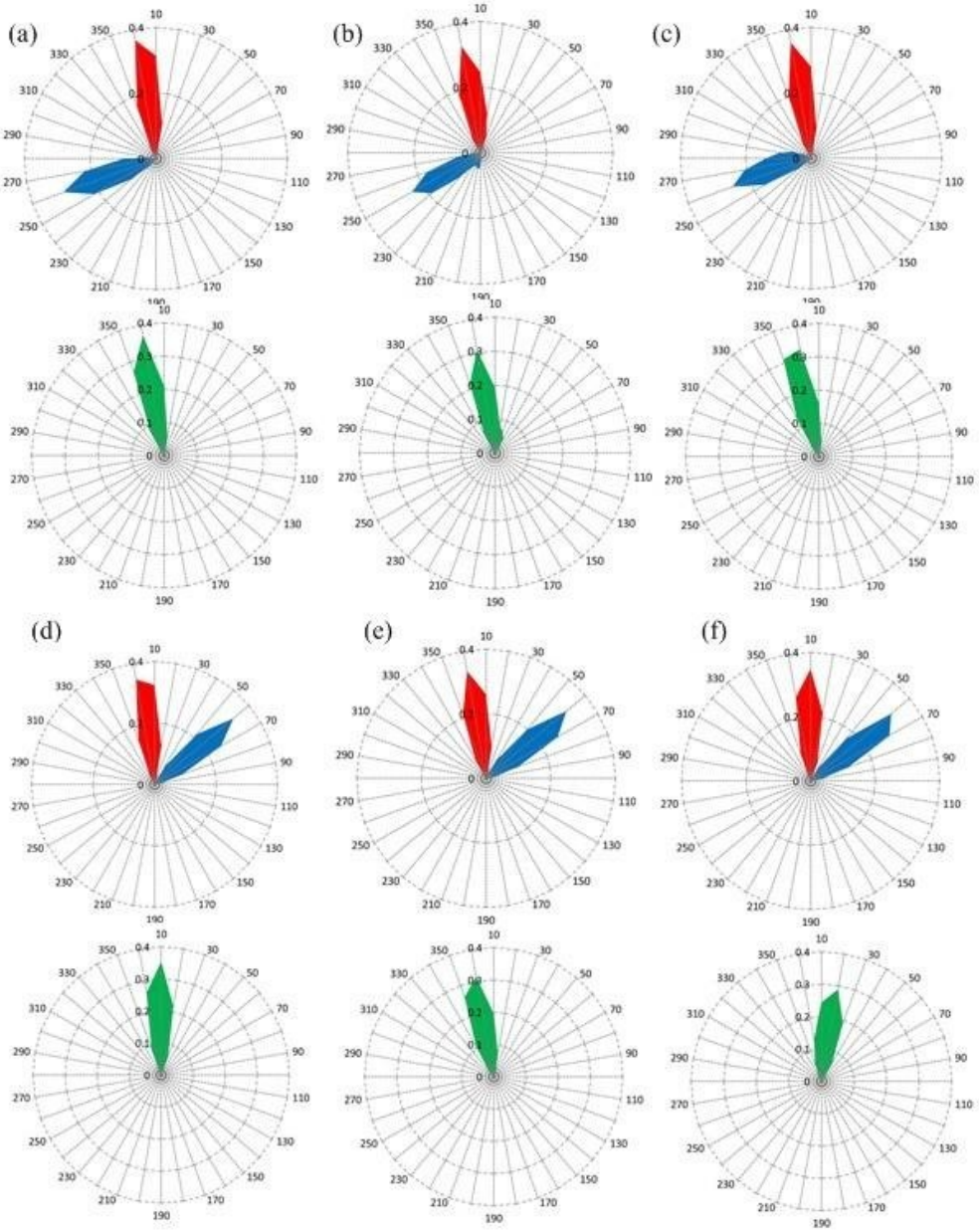

Figure S2. Radar plots for the probability distributions of χ (blue), θ (red) and ϕ (green) for *anti* (a, b and c) or *syn* (d, e and f) ALI-N^6-dA in the CXG sequence in the base-displaced intercalated (a and d), 5'–intercalated (b and e) or 3'–intercalated (c and f) adducted DNA conformations.

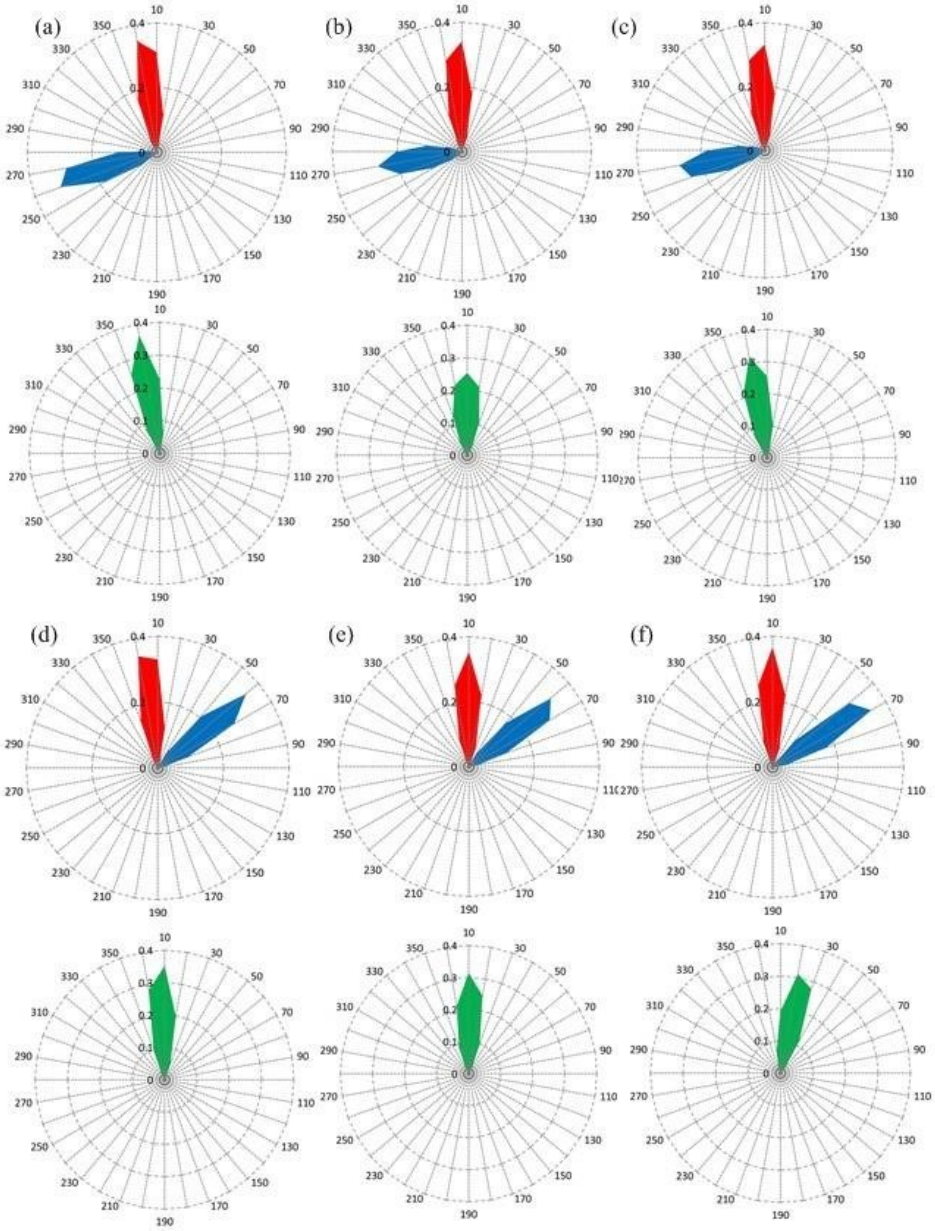

Figure S3. Radar plots for the probability distributions of χ (blue), θ (red) and ϕ (green) for *anti* (a, b and c) or *syn* (d, e and f) ALI-N^6-dA in the GXC sequence in the base-displaced intercalated (a and d), 5'–intercalated (b and e) or 3'–intercalated (c and f) adducted DNA conformations.

Table S1: RMSD of all heavy atoms in the trimer containing the damaged base pair and the 5′ and 3′–flanking bases in the base-displaced intercalated adducted DNA conformations during 20 and 320 ns simulations in the CXC sequence context.

Conformer	RMSD (Å)	
	20 ns	**320 ns**
anti base-displaced intercalated	0.82±0.2	0.96±0.3
syn base-displaced intercalated	1.09±0.2	1.21±0.3

Table S2: Backbone RMSD for each adducted DNA conformation in different sequence contexts.

Sequence Context	Conformers	RMSD (Å)
GXG	*anti* base-displaced	1.82±0.33
	anti 5′–intercalated	4.11±1.51
	anti 3′–intercalated	3.33±0.63
	syn base-displaced	2.05 ±0.37
	syn 5′–intercalated	3.73 ±0.77
	syn 3′–intercalated	2.21±0.53
CXG	*anti* base-displaced	1.90±0.39
	anti 5′–intercalated	2.24±0.46
	anti 3′–intercalated	3.11±0.75
	syn base-displaced	2.19±0.46
	syn 5′–intercalated	2.67±0.47
	syn 3′–intercalated	2.21±0.55
GXC	*anti* base-displaced	2.04 ±0.40
	anti 5′–intercalated	2.94±0.91
	anti 3′–intercalated	2.71±0.58
	syn base-displaced	2.04±0.41
	syn 5′–intercalated	3.01±0.76
	syn 3′–intercalated	2.89±0.72

Table S3: The hydrogen-bonding occupancies in the trimers composed of the damaged base pair and the 3′ and 5′–flanking base pairs in different sequence contexts.[a,b]

Sequence Context	Conformation	Base Pair	H-Bond	Occupancy (%)
CXC	*anti* base-displaced	5′–C(5):G(18)	O2•••H2–N2	99.9
			N3•••H1–N1	99.9
			N4–H4•••O6	98.9
		3′–C(7):G(16)	O2•••H2–N2	99.9
			N3•••H1–N1	99.9
			N4–H4•••O6	98.2
	syn base-displaced	5′–C(5):G(18)	O2•••H2–N2	99.7
			N3•••H1–N1	99.9
			N4–H4•••O6	98.9
		3′–C(7):G(16)	O2•••H2–N2	99.8
			N3•••H1–N1	99.9
			N4–H4•••O6	98.9
	anti 5′–intercalated	5′–C(5):G(18)	O2•••H2–N2	99.9
			N3•••H1–N1	99.9
			N4–H4•••O6	99.1
		3′–C(7):G(16)	N3•••H1–N1	99.9
			N4–H4•••O6	99.1
			O2•••H2–N2	96.9
		3′–C(7):T(17)	N4–H4•••O4	45.8
	syn 5′–intercalated	5′–C(5):G(18)	O2•••H2–N2	99.9
			N3•••H1–N1	99.9
			N4–H4•••O6	98.9
		3′–C(7):T(17)	O4•••H4–N4	83.7
			N3•••H3–N3	74.4
	anti 3′–intercalated	5′–C(5):T(17)	N4–H4•••O4	36.5
		3′–C(7):G(16)	O2•••H2–N2	99.7
			N3•••H1–N1	99.9
			N4–H4•••O6	98.6
	syn 3′–intercalated	5′–C(5):G(18)	O2•••H2–N2	99.5
			N3•••H1–N1	99.1
			N4–H4•••O6	97.1
		3′–C(7):G(16)	O2•••H2–N2	99.5
			N3•••H1–N1	99.9
			N4–H4•••O6	98.7

GXG	*anti* base-displaced	5′–G(5):C(18)	N2–H2•••O2	99.9
			N1–H1•••N3	99.9
			O6•••N4–H4	99.2
		3′–G(7):C(16)	N2–H2•••O2	99.7
			N1–H1•••N3	99.9
			O6•••N4–H4	98.9
	syn base-displaced	5′–G(5):C(18)	N2–H2•••O2	99.9
			N1–H1•••N3	99.9
			O6•••N4–H4	99.3
		3′–G(7):C(16)	N2–H2•••O2	99.8
			N1–H1•••N3	99.9
			O6•••N4–H4	99.1
	anti 5′–intercalated	5′–G(5):C(18)	N2–H2•••O2	99.9
			N1–H1•••N3	99.9
			O6•••N4–H4	99.9
		3′–G(7):C(16)	N2–H2•••O2	91.6
			N1–H1•••N3	92.5
			O6•••N4–H4	93.5
	syn 5′–intercalated	5′–G(5):C(18)	N2–H2•••O2	99.8
			N1–H1•••N3	99.9
			O6•••N4–H4	98.4
		3′–G(7):T(17)	O2•••N1–H1	47.2
	anti 3′–intercalated	5′–G(5):C(18)	N2–H2•••O2	34.1
			N1–H1•••N3	34.0
			O6•••N4–H4	32.5
		3′–G(7):C(16)	N2–H2•••O2	99.7
			N1–H1•••N3	99.2
			O6•••N4–H4	97.9
	syn 3′–intercalated	5′–G(5):C(18)	N2–H2•••O2	99.6
			N1–H1•••N3	99.7
			O6•••N4–H4	96.1
		3′–G(7):C(16)	N2–H2•••O2	99.6
			N1–H1•••N3	99.9
			O6•••N4–H4	99.2

CXG	*anti* base-displaced	5′–C(5):G(18)	O2•••H2–N2	99.9
			N3•••H1–N1	99.9
			N4–H4•••O6	98.8
		3′–G(7):C(16)	N2–H2•••O2	99.8
			N1–H1•••N3	99.9
			O6•••N4–H4	98.7
	syn base-displaced	5′–C(5):G(18)	O2•••H2–N2	99.7
			N3•••H1–N1	99.9
			N4–H4•••O6	98.9
		3′–G(7):C(16)	N2–H2•••O2	99.8
			N1–H1•••N3	99.9
			O6•••N4–H4	99.1
	anti 5′–intercalated	5′–C(5):G(18)	O2•••H2–N2	99.9
			N3•••H1–N1	99.9
			N4–H4•••O6	98.7
		3′–G(7):C(16)	N2–H2•••O2	98.1
			N1–H1•••N3	98.7
			O6•••N4–H4	98.6
	syn 5′–intercalated	5′–C(5):G(18)	O2•••H2–N2	99.9
			N3•••H1–N1	99.9
			N4–H4•••O6	98.9
		3′–G(7):C(16)	N2–H2•••O2	94.5
			N1–H1•••N3	95.6
			O6•••N4–H4	95.5
	anti 3′–intercalated	5′–C(5):G(18)	O2•••H2–N2	44.6
			N3•••H1–N1	44.5
			N4–H4•••O6	50.2
		3′–G(7):C(16)	N2–H2•••O2	99.9
			N1–H1•••N3	99.9
			O6•••N4–H4	98.8
	syn 3′–intercalated	5′–C(5):G(18)	O2•••H2–N2	99.5
			N3•••H1–N1	99.4
			N4–H4•••O6	98.8
		3′–G(7):C(16)	N2–H2•••O2	99.7
			N1–H1•••N3	99.9

			O6•••N4–H4	99.1
GXC	*anti* base-displaced	5′–G(5):C(18)	N2–H2•••O2	99.8
			N1–H1•••N3	99.9
			O6•••N4–H4	99.3
		3′–C(7):G(16)	O2•••H2–N2	99.9
			N3•••H1–N1	99.9
			N4–H4•••O6	98.3
	syn base-displaced	5′–G(5):C(18)	N2–H2•••O2	99.8
			N1–H1•••N3	99.9
			O6•••N4–H4	99.4
		3′–C(7):G(16)	O2•••H2–N2	99.8
			N3•••H1–N1	99.9
			N4–H4•••O6	99.1
	anti 5′–intercalated	5′–G(5):C(18)	N2–H2•••O2	99.8
			N1–H1•••N3	99.9
			O6•••N4–H4	99.0
		3′–C(7):G(16)	O2•••H2–N2	95.3
		3′–C(7):T(17)	N4–H4•••O4	52.1
			N3•••H2–N2	39.7
	syn 5′–intercalated	5′–G(5):C(18)	N2–H2•••O2	99.8
			N1–H1•••N3	99.9
			O6•••N4–H4	98.6
		3′–C(7):G(16)	O2•••H2–N2	80.3
		3′–C(7):T(17)	N4–H4•••O4	89.7
			N3•••H3–N3	80.6
	anti 3′–intercalated	5′–G(5):C(18)	N2–H2•••O2	99.2
			N1–H1•••N3	99.3
			O6•••N4–H4	97.9
		3′–C(7):G(16)	O2•••H2–N2	99.8
			N3•••H1–N1	99.9
			N4–H4•••O6	98.4
	syn 3′–intercalated	5′–G(5):C(18)	N2–H2•••O2	99.9
			N1–H1•••N3	99.9
			O6•••N4–H4	99.5
		3′–C(7):G(16)	O2•••H2–N2	99.9
			N3•••H1–N1	99.9

	N4–H4•••O6	98.5

[a]Hydrogen-bond distance cut-off within 3.4 Å heavy atom separation and 120˚ X–H–X angle.[b]Nucleotides in the 11-mer DNA are numbered starting from the 5′–side of each strand, with the strand containing the lesion numbered first.

Table S4: Average pseudostep parameters,[a] minor groove width,[b] and standard deviations for each adducted DNA conformer in different sequence contexts.

Sequence Context	Conformer	Shift (Å)	Slide (Å)	Rise (Å)	Tilt (˚)	Roll (˚)	Twist (˚)	Minor Groove (Å)
GXG	unmodified	0.1±0.7	–1.6±1.3	6.8±0.4	–1.9±5.0	5.4±7.0	65.0±6.0	7.2±1.5
	anti base-displaced	–1.5±0.7	1.3±1.1	6.7±0.3	0.7±4.8	–4.4±5.5	57.1±6.0	7.7±1.2
	anti 5′–intercalated	–1.9±1.1	1.6±1.8	9.5±0.6	–19.3±6.5	–1.1±9.8	46.2±11.3	16.3±4.5
	anti 3′–intercalated	–1.6±0.9	2.1±1.0	10.1±0.4	–4.2±5.3	–4.4±8.2	46.3±7.1	10.5±1.8
	syn base-displaced	–1.4±0.9	0.7±1.1	6.7±0.3	2.5±4.9	6.1±6.1	51.4±6.3	9.1±1.7
	syn 5′–intercalated	–1.7±1.2	0.9±1.6	9.6±0.8	–9.1±9.3	–3.8±10.4	54.4±11.8	14.6±3.3
	syn 3′–intercalated	–1.5±1.0	0.9±1.2	10.0±0.7	–8.9±5.5	6.6±8.3	50.4±7.7	11.5±1.7
CXG	unmodified	0.0±0.8	–1.6±1.0	6.7±0.4	0.1±5.2	11.3±8.4	62.9±6.3	7.7±1.4
	anti base-displaced	–1.6±0.7	1.0±1.0	6.8±0.3	–1.5±5.0	–3.8±5.9	56.5±5.1	7.4±1.4
	anti 5′–intercalated	–2.1±1.0	0.5±1.5	9.4±0.5	–17.0±6.9	–2.0±9.6	38.7±14.6	10.2±3.1
	anti 3′–intercalated	–0.9±1.5	1.9±1.4	9.1±0.8	–6.2±7.9	–3.2±13.3	48.9±12.1	12.9±2.8
	syn base-displaced	–1.9±0.9	0.6±1.1	6.8±0.4	1.2±5.3	6.4±6.9	48.8±5.9	8.7±1.7
	syn 5′–intercalated	–2.2±1.1	–0.2±1.0	9.8±0.6	–16.6±6.9	–1.3±9.0	50.5±8.0	7.9±1.7
	syn 3′–intercalated	–1.6±1.2	1.4±1.5	10.0±0.7	–9.5±6.0	6.3±9.5	47.2±8.3	11.2±2.0
GXC	unmodified	0.2±0.7	–1.4±1.1	6.6±0.4	–0.9±5.1	3.1±7.0	66.9±6.1	6.9±1.5
	anti base-displaced	–1.4±0.7	0.9±0.9	6.8±0.3	–0.5±4.9	–5.1±5.6	57.8±4.7	7.8±1.3
	anti 5′–intercalated	–1.5±0.9	1.0±1.0	9.0±0.5	–14.5±6.3	–1.5±8.4	51.5±7.0	14.5±3.8
	anti 3′–intercalated	–1.6±1.2	1.3±1.2	10.0±0.5	–9.9±6.4	–2.6±9.4	48.3±9.4	10.1±1.8
	syn base-displaced	–1.6±0.8	–0.2±1.2	6.8±0.3	0.8±4.9	5.9±6.3	54.4±6.0	8.4±1.5
	syn 5′–intercalated	–1.5±1.0	–0.9±1.2	8.7±0.5	–7.7±6.5	7.4±10.6	45.3±10.3	10.1±2.3
	syn 3′–intercalated	–1.7±1.0	2.1±0.9	10.3±0.5	–9.3±5.0	1.7±8.8	48.6±8.0	10.4±1.8
CXC	unmodified	0.3±0.8	–1.3±1.0	6.5±0.4	1.9±5.2	8.6±7.8	63.1±6.0	7.2±1.5
	anti base-displaced	–1.6±0.7	0.8±0.8	6.9±0.3	–2.1±5.0	–5.2±5.9	56.8±4.8	7.6±1.2
	anti 5′–intercalated	–2.2±1.1	0.5±1.0	9.0±0.7	–14.9±6.7	–0.7±9.0	45.9±9.3	9.9±2.0
	anti 3′–intercalated	–0.5±1.2	1.4±1.2	9.1±1.0	–7.2±7.0	–1.5±9.1	54.5±9.2	11.5±2.7
	syn base-displaced	–1.9±0.9	0.0±1.2	6.9±0.4	0.7±5.1	7.7±8.1	51.2±6.1	7.7±1.7
	syn 5′–intercalated	–1.0±1.5	–1.0±1.7	8.9±0.7	–10.1±8.4	5.1±10.4	53.0±8.4	11.6±1.7
	syn 3′–intercalated	–1.7±1.1	1.5±1.4	10.1±0.6	–11.1±5.4	7.5±9.5	47.3±8.3	9.5±2.0

[a]Calculated using the 3′ and 5′–flanking base pairs. [c]Calculated as the distance between the P atom of the 7th and the 20th residue minus 5.8 Å to account for the van der Waals radii of the two phosphate groups.

Table S5: Interaction energies between the adduct and the opposing thymine for the base-displaced intercalated conformations of ALI-N^6-dA adducted DNA in different sequence contexts derived from 20 ns MD simulations (kJ mol^{-1}).

Sequence Context	Conformation	$E_{vdW} + E_{elec}$
GXG	*anti* base-displaced	–18.9±3.8
	syn base-displaced	–10.5±4.4
CXG	*anti* base-displaced	–18.1±4.2
	syn base-displaced	–8.4±4.6
GXC	*anti* base-displaced	–17.6±4.2
	syn base-displaced	–10.1±3.3
CXC	*anti* base-displaced	–15.5±5.0
	syn base-displaced	–8.0±4.6