SUPPLEMENTARY DISCUSSION

QM data fitting.

As shown in the **Supplementary Table 12** refined parmbsc1 parameters fit very well high-level QM data. The *syn-anti* equilibrium, which was non-optimal in parmbsc0, is now well reproduced (**Supplementary Fig. 26**). The fitting to sugar puckering profile was improved by increasing the East barrier, and by displacing the North and South minima to more realistic regions (**Supplementary Table 12** and **Supplementary Fig. 27**). Additionally, parmbsc1 provides ε and ζ conformational map almost indistinguishable from the CCSD(T)/CBS results in solution (**Supplementary Fig. 28**), with errors in the estimates of relative BI/BII stability and transition barrier equal to 0.2 and 0.0 kcal mol^{-1} respectively.

Force-field benchmark simulations.

It is not our purpose here to perform a comprehensive comparison of parmbsc1 with previous force-fields. This would require the analysis of >100 structures with up to six other force-fields, clearly out of the scope of this work. We performed, however, a first critical evaluation of the most used force-fields using the well-known Drew Dickerson dodecamer as reference. We tested parmbsc0[1–3], parmbsc0-OL1[4] (ε and ζ corrections from Šponer's group), parmbsc0-OL4[5] (χ corrections), parmbsc0-OL1+OL4[4,5], CHARMM36[6], and a modified parmbsc0 developed by mixing corrected χ values and scaled-down van der Waals interactions (parmbsc0-CG, Cheng-Garcia)[7]. In all cases simulations were extended for at least 1 μs under identical simulation conditions. The value of this benchmark must not be overestimated, since different behavior may be found for other DNA sequences or conformations, but it can be useful to obtain an approximate idea of the range of error expected in parmbsc1 with respect to other modern force-fields. Results are summarized in **Supplementary Table 2** and **Supplementary Figs. 29–31**. All the force-fields are able to maintain the general B-like

conformation in the central part of the duplex. However, significant distortions are found in the terminal pairs for parmbsc0, parmbsc0-OL1 (ε and ζ corrections), and CHARMM36, which show large openings (**Supplementary Fig. 29**) and very frequent fraying, with the formation of non-canonical interactions. The distortion induced by the opening of the terminal C-G pairs is especially dramatic in CHARMM36 simulations (**Supplementary Fig. 29**), but it is not negligible for parmbsc0[8] and parmbsc0-OL1, where aberrant *trans* Watson-Crick contacts involving a cytosine in *syn*, are dominant (**Supplementary Fig. 30**). It is clear that duplexes are flexible and reversible opening and closing of terminal base pair should exist, as found for example in parmbsc1 simulations (**Supplementary Fig. 30**). However, detailed analysis of new NMR spectra (**Supplementary Fig. 31**) shows that there are just minor differences between terminal and interior base pairs, which mean that open states should be short-lived, and not prevalent as in CHARMM36 simulations. Furthermore, no NMR evidence exists (**Supplementary Fig. 31**) supporting the existence of stable unusual contacts involving terminal pairs, or the prevalence of non-*anti* conformations, which are observed in parmbsc0, parmbsc0-OL1 or CHARMM36 simulations.

The introduction of χ corrections removes the excessive fraying of terminal pairs, preserving better the integrity of the entire helix in parmbsc1, parmbsc0-OL4[8], parmbsc0-CG (Cheng-Garcia, and parmbsc0-OL1+OL4 (ε, ζ, and χ corrections together) trajectories (**Supplementary Figs. 29** and **30**). The duplex sampled from parmbsc0-CG calculations is however far from the experimental structures: RMSd around 4 Å (compared to values clearly below 2.0 Å for parmbsc1 simulations), strong under-twisting, poor groove geometry and incorrect description of the BI/BII equilibrium (**Supplementary Table 2**). The sequence dependence of the helical properties, which is clear for the rest of bsc0-based force-fields, is also lost here (**Supplementary Fig. 29**).

Parmbsc0-OL4 and parmbsc0-OL1+OL4 provide reasonable representations of the DDD geometry. However, the use of parmbsc1 leads to clear improvements in all structural

descriptors. Thus, parmbsc1 balances better the sugar puckering (see **Supplementary Fig. 29**), leads to a better balance of BI/BII states (**Supplementary Table 2**), improves very significantly the average roll which is now very close to the NMR estimates, avoiding the excess of roll found in other calculations (**Supplementary Table 2** and **Supplementary Fig. 29**). Parmbsc1 improves very clearly the average twist and its sequence-dependence (RMSd difference between NMR and parmbsc1 twist profiles is 1.9 º, compared with 3.7 º for parmbsc1-OL1+OL4, or 5.6 º for CHARMM36. Not surprisingly, the improvement in twist, roll and puckering is reflected in much more realistic groove dimensions. For example the average difference in groove widths is only 0.3 Å between parmbsc1 and NMR values, while for the parmbsc0-OL1+OL4 force-field error is above 1 Å. In summary, at least for DDD, parmbsc1 provide results of better quality than those obtained with the most recent force-fields for DNA available.

The effect of ionic strength and the nature of counterion.

To evaluate potential differences in simulations arising from the ionic strength we performed additionally 2 μs simulations of DDD with extra salt: Na^+Cl^- 150 mM, and 500 mM. These additional calculations were performed using the same conditions outlined previously, showing results that are quite independent on the exact choice (in the 0–500 mM range) of the added extra salt (**Supplementary Fig. 25**).

SUPPLEMENTARY REFERENCES

1. Pérez, A. *et al. Biophys. J.***92,** 3817–3829 (2007).
2. Cornell, W.D. *et al. J. Am. Chem. Soc.***117,** 5179–5197 (1995).
3. Cheatham III, T.E., Cieplak, P. & Kollman, P.A. *J. Biomol. Struct. Dyn*.**16,** 845–862 (1999).
4. Zgarbová, M. *et al. J. Chem. Theory Comput.***9,** 2339–2354 (2013).
5. Krepl, M. *et al. J. Chem. Theory Comput.***8,** 2506–2520 (2012).

6. Hess, B., Kutzner, C., Van Der Spoel, D. & Lindahl, E. *J. Chem. Theory Comput.***4,** 435–447 (2008).

7. Cheng, A.A., Garcia, A.E. *Proc. Natl. Acad. Sci. USA***110**, 16820–25 (2013).

8. Zgarbová, M., Otyepka, M., Šponer, J., Lankaš, F. & Jurečka, P. *J. Chem. Theory Comput.***10,** 3177–3189 (2014).

SUPPLEMENTARY TABLES

Supplementary Table 1. DNA sequences used for validation of the parmbsc1 force-field. The nature of the structure, the origin of the starting conformation and the length of the production trajectories are also reported. The validation set is divided in several blocks separated in the table by double lines (from top to bottom): i) Normal B-DNA structures (including mismatches, epigenetic modifications and polymeric sequences); ii) very large oligomers; iii) Complexes of DNA with proteins or drugs; iv) Unusual DNA structures; v) dynamic transitions.; parmbsc1 validation; and vi) parmbsc1 benchmarking.

Sequence	Family	Origine / PDB id	Length (ns)
$d(CGCGAATTCGCG)_2$	B-DNA	1BNA, 1NAJ	1x 800 2x 1000 1x 12001x 10000
$d(CCATACaATACGG)_2$	B-DNA mismatch AA	Fiber	500
$d(CCATACgATACGG)_2$	B-DNA mismatch GG	Fiber	500
$d(CGCGA_{5mC}GTCGCG)_2$	B-DNA 5methylC	Fiber	250
$d(CGCGA_{5hmC}GTCGCG)_2$	B-DNA 5hydroxy-methylC	Fiber	250
$d(CGCGT_{5mC}GACGCG)_2$	B-DNA 5methylC	Fiber	500
$d(CGCGACGTCGCG)_2$	B-DNA	Fiber	500
$d(CGCGTCGACGCG)_2$	B-DNA,	Fiber	500
$d(GCCTATAAACGCCTATAA)_2$	B-DNA	Fiber	1000
$d(CTAGGTGGATGACTCATT)_2$	B-DNA	Fiber	1000
$d(CACGGAACCGGTTCCGTG)_2$	B-DNA	Fiber	1000
$d(GGCGCGCACCACGCGCGG)_2$	B-DNA	Fiber	1000
$d(GCCGAGCGAGCGAGCGGC)_2$	B-DNA	Fiber	1000
$d(GCCTAGCTAGCTAGCTGC)_2$	B-DNA	Fiber	1000
$d(GCTGCGTGCGTGCGTGGC)_2$	B-DNA	Fiber	1000
$d(GCGATCGATCGATCGAGC)_2$	B-DNA	Fiber	1000
$d(GCGAGGGAGGGAGGGAGC)_2$	B-DNA	Fiber	1000
$d(GCGCGGGCGGGCGGGCGC)_2$	B-DNA	Fiber	1000
$d(GCGGGGGGGGGGGGGGGC)_2$	B-DNA	Fiber	1000
$d(GCGTGGGTGGGTGGGTGC)_2$	B-DNA	Fiber	1000
$d(CTCGGCGCCATC)_2$	B-DNA	2HKB	590
$d(CCTCTGGTCTCC)_2$	B-DNA	2K0V	590

$d(CGCATGCTACGC)_2$	B-DNA	2L8Q	590
$d(GGATATATCC)_2$	B-DNA	2LWG	590
$d(GCGCATGCTACGCG)_2$	B-DNA	2M2C	590
$d(CCTCAGGCCTCC)_2$	B-DNA	2NQ1	590
$d(CGCGAAAAAACG)_2$	B-DNA (A-track)	1D89	200
$d(GGCAAAAAACGG)_2$	B-DNA (A-track)	1FZX	200
$d(GCAAAATTTTGC)_2$	B-DNA (A-track)	1RVH	200
$d(CTTTTAAAAG)_2$	B-DNA (A-track)	1SK5	200
$d(AGGGGCCCCT)_2$	B-DNA (A-track)	440D	200
$d(GGCAAGAAACGG)_2$	B-DNA (A-track)	1G14	1000
$d(CGATCGATCG)_2$	B-DNA crystal	1D23	32x 2000
$d(ATGGATCCATAGACCAGAACATGATGTTCTCA)_2$	B-DNA 32mer	Fiber	1000
$d(CGCGATTGCCTAACGAGTACTCGTTAGGCAATCGCG)_2$	B-DNA 36mer	Fiber	2x 300
$d(CGCGATTGCCTAACGGACAGGCATAGACGTCTATGCCTGTCCGTTAGGCAATCGCG)_2$	B-DNA 56mer	Fiber	1x 290 1x 500
$d(CGTGGCGGCAGTAGCGCGGTGGTCCCACCTGACCCCATGCCGAACTCAGAAGTGCG)_2$	B-DNA 56mer	Fiber	300
$d(CGCCGGCAGTAGCCGAAAAAATAGGCGCGCGCTCAAAAAAATGCCCCATGCCGCGC)_2$	B-DNA 56mer	Fiber	1x 360 1x 440 1x 500
$d(ATCTTTGCGGCAGTTAATCGAACAAGACCCGTGCAATGCTATCGACATCAAGGCCTATCGCTATTACGGGGTTGGGAGTCAATGGGTTCAGGATGCAGGTGAGGAT)_2$	106-mer circle 10 turns (reg A)	Fiber	100
$d(ATCTTTGCGGCAGTTAATCGAACAAGACCCGTGCAATGCTATCGACATCAAGGCCTATCGCTATTACGGGGTTGGGAGTCAATGGGTTCAGGATGCAGGTGAGGAT)_2$	106-mer circle 10 turns (reg B)	Fiber	100
$d(ATCTTTGCGGCAGTTAATCGAACAAGACCCGTGCAATGCTATCGACATCAAGGCCTATCGCTATTACGGGGTTGGGAGTCAATGGGTTCAGGATGCAGGTGAGGAT)_2$	106-mer circle 10 turns (reg C)	Fiber	100
$d(ATCTTTGCGGCAGTTAATCGAACAAGACCCGTGCAATGCTATCGACATCAAGGCCTATCGCTATTACGGGGTTGGGAGTCAATGGGTTCAGGATGCAGGTGAGGAT)_2$	106-mer circle 9 turns	Fiber	50
$d(ATCTTGGCAGTTAATCGAACAAGACCCGTGCAATGCTATCGACATCAAGGCCTATCGTTACGGGGTTGGGAGTCAATGGGTTCAGGATGCAGGTGAGGAT)_2$	100-mer circle 9 turns	Fiber	100
147mer nucleosome	DNA-histones	1KX5	500
DNA:HU complex	DNA-HU protein	1P71	1000
DNA:HU complex	DNA-HU protein	1P71 (without mismatches and flipped bases)	1000

DNA:TRP repressor	DNA-repressor	1TRO	1000
DNA:leucine zipper	DNA-transc factor	2DGC	1000
DNA:P22 c2	DNA-represor	3JXC	1000
$d(CGCAAATTTGCG)_2$-distamycin	DNA-mG binder	2DND	700
$d(CTTTTCGAAAAG)_2$-Hoescht	Drug cooperativity	1QSX	10x 10
$d(CGTACG)_2$-daunomycin	DNA-intercalator	1D11	600
$d(GGGG)_4$	PS quadruplex	352D (without Thymine loops)	440
$d(GGGG)_4$	APS quadruplex	156D (without Thymine loops)	440
$d(T•A•T)_{10}$	PS triplex	Fiber	440
$d(G•G•C)_{10}$	PS triplex	Fiber	440
$d(G•G•C)_{10}$	APS triplex	Fiber	440
$d(ATATATATATAT)_2$	H-duplex	1GQU	720
$d(CGATATATATAT)_2$	H-duplex	2AF1	400
d(AAGGGTGGGTGTAAGTGTGGGTGGGT)	G_quadruplex	2LPW	5000
d(AGGGTTAGGGTTAGGGTTAGGG)	G-loop quadruplex(HTQ)	1KF1	1000
$d(GGGGTTTTGGGG)_2$	G quadruplex (OxyQ)	1JRN	1000
$d(CCGGTACCGG)_4$	Holliday Junction	1DCW	1000
$d(CGCGCGCGCG)_2$	Z-DNA, duplex	1I0T	2x 385
d(GCGAAGC)	Hairpinfold (REXMD)	1PQT	1000
$d(CGCGAATTCGCG)_2$	A-form in ethanol	1BNA	200
$d(CGCGAATTCGCG)_2$	A to B transition (H_2O)	1BNA	5x40
$d(GGCGCC)_2$	DNA unfolding (Pyridine)	1P25	400
$d(CGCGAATTCGCG)_2$	DDD, 0.15M NaCl	1BNA	2000
$d(CGCGAATTCGCG)_2$	DDD, 0.5M NaCl	1BNA	3000
$d(CGCGAATTCGCG)_2$	parmBSC0	1BNA	1500
$d(CGCGAATTCGCG)_2$	parmBSC0-OL1	1BNA	1500
$d(CGCGAATTCGCG)_2$	parmBSC0-OL4	1BNA	1500
$d(CGCGAATTCGCG)_2$	parmBSC0-OL1-OL4	1BNA	1500
$d(CGCGAATTCGCG)_2$	parmBSC0-Cheng-Garcia	1BNA	1500
$d(CGCGAATTCGCG)_2$	CHARMM36	1BNA	1500
$d(CGCGAATTCGCG)_2$	DDD, Amber GPU	1BNA	100
$d(CGCGAATTCGCG)_2$	DDD, Amber CPU	1BNA	100
$d(CGCGAATTCGCG)_2$	DDD, Gromacs GPU	1BNA	100
$d(CGCGAATTCGCG)_2$	DDD, Gromacs CPU	1BNA	100

Supplementary Table 2. MD-averaged helical parameters (on 1.2 μs simulation time) of Drew-Dickerson dodecamer in parmbsc1 simulations (and, as a control, other modern force-fields) compared with the NMR and X-ray estimates. [a]

	Twist	Roll	Slide	Rise	Shift	Tilt	BI(%)	Major groove width	Minor groove width
Parmbsc1	**34.3±5.4**	**1.5±5.4**	**-0.3±0.5**	**3.3±0.3**	**0.0±0.8**	**0.0±4.5**	**77**	**11.9±1.7**	**5.4±1.2**
Parmbsc0	32.8±5.8	2.7±5.8	-0.4±0.6	3.3±0.3	0.0±0.7	0.0±4.3	84	12.9±1.8	3.9±1.2
OL1	33.3±5.7	2.7±5.9	-0.2±0.6	3.3±0.3	0.0±0.7	0.0±4.4	83	12.2±1.4	6.1±1.3
OL4	33.3±6.4	2.6±5.9	-0.1±0.6	3.3±0.3	0.0±0.7	0.0±4.5	85	12.1±1.4	6.5±1.3
OL1+OL4	33.0±6.1	2.8±5.7	-0.3±0.6	3.3±0.3	0.0±0.7	0.0±4.3	86	12.4±1.5	6.0±1.2
C36 [d]	34.5±11	5.1±8.8	0.8±1.0	3.6±0.8	-0.1±1.1	0.9±8.0	66	10.5±1.5	8.3±1.7
Cheng-Garcia(CG)	32.5±3.4	1.5±5.2	-1.7±0.5	3.4±0.3	0.0±0.4	0.0±4.3	100	15.3±1.6	5.5±0.9
X-ray [b]	**35.2±0.6**	**-0.7±1.1**	**0.1±0.1**	**3.3±0.1**	**-0.1±0.1**	**-0.4±0.9**		**11.2±0.1**	**4.6±0.3**
NMR [c]	**35.6±0.8**	**1.6±1.0**	**-0.3±0.1**	**3.2±0.1**	**0.0±0.1**	**0.0±0.7**	**73[e]**	**11.9±0.3**	**4.7±0.3**

[a] Translational parameters and groove widths are in Å, while rotational parameters are in degrees. Note that for MD trajectories the standard deviations are computed from sequence-averages and time-averages. [b] X-ray mean values and standard deviations were obtained averaging the following structures (PDB id): 1BNA[1], 2BNA[2], 7BNA[3] and 9BNA[4]. [c] NMR mean values and standard deviations were obtained by averaging over the ensemble of structures contained in the PDB id 1NAJ[5].[d] These average values are contaminated by the opening of terminal base pairs (note large standard deviations in roll and twist). [e] Average value of BI population taken by averaging direct NMR estimates[6,7]. See also Supplementary Discussion and **Supplementary Figs. 29-31** for a discussion on the relative performance of parmbsc1 with respect to other force-fields.

1. Drew, H.R. *et al*. *Proc. Natl. Acad. Sci. U. S. A.***78,** 2179–2183 (1981).
2. Drew, H.R., Samson, S. & Dickerson, R.E. *Proc. Natl. Acad. Sci. U. S. A.***79,** 4040–4044 (1982).
3. Holbrook, S.R. *et al*. *Acta Crystallogr.,Sect.B***41,** 255–262 (1985).
4. Westhof, E. *J. Biomol. Struct. Dyn.***5,** 581–600 (1987).
5. Wu, Z., Delaglio, F., Tjandra, N., Zhurkin, V.B. & Bax, A.*J. Biomol. NMR***26,** 297–315 (2003).
6. Tian, Y., Kayatta, M., Shultis, K., Gonzalez, A., Mueller, L.J. & Hatcher, M.E., *J. Phys. Chem. B***113**, 2596–2603 (2008).
7. C. D. Schwieters, C.D. & Clore, G.M. *Biochemistry***46**, 1152–1166 (2007).

Supplementary Table 3. Ability of MD-ensembles obtained from parmbsc0 and parmbsc1 force fields to reproduce NMR observables for Drew-Dickerson dodecamer. The first block correspond to residual dipolar couplings Q-factor, $q = \sqrt{\sum\left(RDC_{calc} - RDC_{exp}\right)^2} \Big/ \sqrt{\sum RDC_{exp}^2}$, where RDC_{exp} has been determined using PALES[1], and the second block to NOEs (146 restraints).

	NMR	**X-ray**	**Fiber model B-DNA**	**Fiber model A-DNA**	**BSC1**	**BSC0**
Bicelles, 1NAJ [a], 129 RDCs	0.17	0.49	0.51	0.87	0.32	0.36
Bicelles, 1DUF [b], 204 RDCs	0.23	0.53	0.66	0.92	0.34	0.38
Sum of violations (A)	0.01	10.0	7.6	42.01	0.4	2.6
Largest violation (A)	0.01	1.0	0.4	1.3	0.2	1.3
Num. of violated restraints	1	35	36	84	2	5

[a] Data taken from ref. 2. [b] Data taken from ref. 3.

1. Zweckstetter, M. *Nat. Protoc.*, ***3***, 679-690 (2008).
2. Wu, Z., Delaglio, F., Tjandra, N., Zhurkin, V.B. & Bax, A.*J. Biomol. NMR***26,** 297–315 (2003).
3. Tjandra, N., Tate, S. I., Ono, A., Kainosho, M. & Bax, A. *J. Am. Chem. Soc.***122,** 6190–6200 (2000).

Supplementary Table 4. Different metrics showing the quality of parmbsc1 simulations for B-DNA duplexes.[a]

DNA seq or PDB id	Ref	RMSd	RMSd/bp	% H-bond	Avg. twist	Avg. roll
1BNA (12mer)	C	2.1 / 1.7	0.18 / 0.17	96 / 98	35.6 / 34.3	2.8 / 1.5
1NAJ (12mer)	N	1.7 / 1.4	0.15 / 0.15	96 / 98	35.6 / 34.3	2.8 / 1.5
CCATACgATACGG[b]	N	2.9 / 2.3	0.22 / 0.21	91 / 91	33.5 / 34.2	8.8 / 1.6
CCATACaATACGG[c]	N	3.3 / 3.1	0.26 / 0.28	93 / 94	33.7 / 34.1	2.7 / 2.5
CGCGACGTCGCG	F	2.0 / 1.5	0.17 / 0.15	98 / 99	34.8 / 34.6	3.1 / 2.0
CGCGTCGACGCG	F	2.6 / 1.5	0.22 / 0.16	97 / 99	34.1 / 34.5	3.4 / 2.3
GCGAGGGAGGGAGGGAGC	F	2.7 / 2.3	0.15 / 0.15	97 / 99	33.5 / 33.3	2.5 / 2.9
GCGCGGGCGGGCGGGCGC	F	2.3 / 2.0	0.13 / 0.13	97 / 99	33.7 / 33.7	2.8 / 3.3
GCGGGGGGGGGGGGGGGC	F	3.0 / 2.7	0.17 / 0.17	98 / 99	32.8 / 32.6	3.0 / 3.5
GCGTGGGTGGGTGGGTGC	F	2.2 / 1.9	0.12 / 0.12	97 / 99	33.1 / 33.0	2.7 / 3.2
GCCGAGCGAGCGAGCGGC	F	2.9 / 2.4	0.17 / 0.15	98 / 99	34.7 / 34.5	2.1 / 2.6
GCCTAGCTAGCTAGCTGC	F	2.2 / 1.9	0.13 / 0.12	97 / 98	34.3 / 34.2	1.6 / 2.1
GCTGCGTGCGTGCGTGGC	F	2.2 / 2.0	0.13 / 0.13	97 / 98	32.6 / 34.5	2.3 / 2.8
GCGATCGATCGATCGAGC	F	2.0 / 1.8	0.11 / 0.12	97 / 98	34.8 / 34.7	1.9 / 2.3
GCCTATAAACGCCTATAA	F	2.9 / 2.8	0.17 / 0.18	94 / 97	34.7 / 34.4	1.6 / 2.0
CTAGGTGGATGACTCATT	F	3.3 / 2.9	0.18 / 0.18	94 / 97	30.9 / 31.8	1.2 / 4.6
CACGGAACCGGTTCCGTG	F	3.0 / 2.9	0.17 / 0.18	95 / 97	34.6 / 33.8	2.7 / 2.0
GGCGCGCACCACGCGCGG	F	3.4 / 2.7	0.19 / 0.17	96 / 98	33.2 / 34.4	3.5 / 2.4
1D89 (12mer)	C	2.3 / 1.9	0.19 / 0.19	93 / 98	35.6 / 33.9	3.0 / 1.7
1FZX (12mer)	N	1.8 / 1.7	0.16 / 0.18	95 / 96	33.9 / 33.8	2.4 / 2.3
1RVH (12mer)	N	1.9 / 1.7	0.16 / 0.17	98 / 98	33.9 / 34.0	2.2 / 2.6
1SK5 (10mer)	C	2.1 / 1.8	0.21 / 0.23	93 / 97	34.2 / 34.3	1.7 / 1.7
CGATATATATATCG	F	1.9 / 1.6	0.16 / 0.17	96 / 97	34.4 / 34.4	2.9 / 1.7
2HKB (12mer)	N	1.8 / 1.7	0.15 / 0.17	96 / 97	34.1 / 33.8	2.3 / 2.6
2K0V (12mer)	N	2.4 / 2.1	0.20 / 0.22	95 / 96	33.9 / 33.5	2.2 / 1.9
2L8Q (12mer)	N	1.9 / 1.5	0.16 / 0.16	95 / 97	34.4 / 34.1	2.7 / 2.5
2LWG (10mer)	N	1.8 / 1.5	0.18 / 0.19	98 / 99	34.5 / 34.6	2.4 / 1.5
2M2C (14mer)	N	2.5 / 2.3	0.18 / 0.20	96 / 97	34.4 / 34.0	2.7 / 2.5

[a] The reference structures used for comparison were taken from X-ray crystallography (C), NMR (N) or fiber (F) data, as available. Except otherwise mentioned, all the duplexes were self-complementary and only one strand is noted. For structures available in the Protein Data Bank we display only the PDB code. RMSd are in Å and average rotational parameters are in degrees. Note that the first value in each cell corresponds to a sequence average considering the complete oligomer, while the second value in each cell was computed excluding the terminal residues. [b] Structure containing a G:G mismatch. The NMR structure used as reference was solved after parmbsc1 was derived[1]. [c] Same than [b] but containing an A:A mismatch.

1. Rossetti, G., Dans, P.D. *et al. Nucleic Acids Res.***43**, 4309-4321 (2015).

Supplementary Table 5. Long oligomers RMSd, helical parameters, and bending (reported herein as % of shortening) values, for all the residues or excluding the terminal ones, with respect to the ideal helix built using average dinucleotide X-ray helical parameters.

	Seq1[c]	**Seq2a**	**Seq2b**	**Seq3**	**Seq4a**	**Seq4b**
RMSd	4.4±1.3	4.2±1.5	4.3±1.3	6.7±2.8	7.2±2.7	7.4±2.7
RMSd (no ends)	4.2±1.2	4.0±1.4	4.1±1.2	6.4±2.6	6.9±2.6	7.0±2.5
RMSd / bp[a]	0.14	0.12	0.12	0.12	0.14	0.13
RMSd / bp (no ends)	0.14	0.12	0.12	0.12	0.13	0.13
Avg. twist (º)	34.9±7.3	35.0±5.3	34.5±5.4	34.2±5.6	34.8±5.3	34.3±5.8
Avg. roll (º)	2.1±8.4	1.5±5.8	1.7±5.8	2.2±5.7	1.7±5.8	2.0±6.0
Avg. slide (Å)	-0.4±0.7	-0.2±0.5	-0.3±0.6	-0.4±0.6	-0.2±0.5	-0.3±0.5
Shortening[b]	4±2 (16)	5±2 (20)	5±2 (17)	6±3 (18)	6±3 (23)	6±3(21)

[a] Values per base pair are indicated to avoid size-inconsistency. [b] Note that for helix shortening the maximum shortening percentages are reported in bracket.
[c] Seq1: ATGGATCCATAGACCAGAACATGATGTTCTCA in TIP3P water;
Seq2a: CGCGATTGCCTAACGAGTACTCGTTAGGCAATCGCG in SPCE water;
Seq2b: idem Seq2a in TIP3P water;
Seq3: CGCCGGCAGTAGCCGAAAAAATAGGCGCGCGCTCAAAAAAATGCCCCATGCCGCGC in TIP3P water;
Seq4a: CGCGATTGCCTAACGGACAGGCATAGACGTCTATGCCTGTCCGTTAGGCAATCGCG in SPCE water;
Seq4b: idem Seq4a in TIP3P water.

Supplementary Table 6. Statistic of NOE restraints violations for different nucleic acids (include: normal duplexes, hairpins, quadruplexes, and A-tracks).[a]

Structure (PDB id)	Number Restraints	Average Violation	Largest Violation	Number violations
1NAJ	146	*0.0001*	*0.01*	*1*
		0.003	2	1
2LPW	938	*0.0006*	*0.1*	*12*
		0.07[b]	7.0	45
1PQT	94	*0.01*	*0.1*	*3*
		0.01	0.1	2
1G14	218	*0.01*	*0.2*	*33*
		0.05	0.9	44
1RVH	446	*0.02*	*0.3*	*50*
		0.03	0.8	56
2LWG	415	*0.01*	*0.5*	*28*
		0.03	1.4	38
2K0V	634	*0.05*	*1.9*	*83*
		0.12	2.5	129
2L8Q	172	*0.0005*	*0.09*	*1*
		0.001	0.26	1
2M2C	296	*0.15*	*3.3*	*54*
		0.13	3.1	50
2NQ1	870	*0.02*	*1.3*	*111*
		0.09	3.9	162

[a] For each PDB entry we show the number of experimental restraints, the average deviation (A), the maximum deviation (A), and the number of restraint violations. In each cell NMR results are reported in italic, *i.e*., the values obtained when experimental restraints were enforced to solve the structure; while the MD results obtained using parmbsc1 simulations are reported with normal characters. [b] Since the NOE deviations were larger than usual for this hairpin, calculations were repeated using parmbsc0 and CHARMM36 force-fields, finding 73 and 64 violations respectively.

Supplementary Table 7. Quality factor (Q-factor), $q = \sqrt{\sum\left(RDC_{calc} - RDC_{exp}\right)^2} \Big/ \sqrt{\sum RDC_{exp}^2}$, for the agreement between observed and predicted residual dipolar couplings (RDCs), using both experimental NMR structures and parmbsc1 MD simulations. [a]

Structure	Alignment Method	Number RDCs	Q-factor (NMR)	Q-factor (MD)
1NAJ	Bicelles	204	0.23	0.34
2LPW	Bicelles	57	0.25	0.54
1PQT	Pf1	29	0.11	0.41
1RVH	Pf1	72	0.13	0.27
2LWG	Pf1	46	0.18	0.29

[a] Note that lower Q-factor indicates better agreement. Typically data sets include both C-H and N-H dipolar couplings. The alignment media used to record NMR RDCs is indicated in all the cases. RDCs were back-calculated from the MD simulations using PALES.

Supplementary Table 8. Statistic of NOE violations for different nucleic acids, for oligomers solved after parmbsc1 development. NOE restraints here are determined using the full matrix relaxation and are more accurate than those typically found in the literature (rough data available upon request). [a]

Duplex	Number restraints	Average violation	Largest violation	Number violations[b]	Rfactor$_{2\alpha}$[c]
GG mismatch	246	*0.004*	*0.090*	*73|15|0*	*0.204*
		0.012	0.302	64|36|7	0.172
AA mismatch	230	*0.003*	*0.160*	*64|6|1*	*0.290*
		0.006	0.083	51|27|0	0.292
A***C***GT control	208	*0.006*	*0.046*	*85| 29|0*	*0.261*
		0.022	0.123	106|79|12	0.250
A***5mC***GT[d]	102	*0.034*	*0.205*	*57|49|14*	*0.197*
		0.035	0.189	60|45|18	0.243
A***5hmC***GT[e]	216	*0.004*	*0.045*	*63|18|0*	*0.232*
		0.014	0.218	86|57|2	0.236

[a] Note that the comparisons are made between metrics obtained for the NMR ensemble (the set of structures refined by imposing NMR restraints) in italics, and those coming from the unbiased MD trajectory in roman. [b] To define "number of violations" we used three criteria: i) the distances given by the flat well limits (left value in the cell), ii) the boundaries of the "contact" are extended by ±0.2 Å (middle value), and finally iii) the upper-limit is multiplied by 1.25 (right value in the cell). [c] The global quality factor Rfactor$_{2\alpha}$[1, 2] take values around 0.6 and 0.7 for B and A-DNA respectively. The sequences considered here are reported in **Supplementary Table 1**.[d] 5mC stands for 5-methyl-cytosine. [e] 5hmC stands for 5-hydroxymethyl-cytosine.

1. Gonzalez, C., Rullmann, J.A.C., Bonvin, A., Boelens, R. & Kaptein, R. *J. Magn. Reson.***91,** 659–664 (1991).
2. Gronwald, W. *et al. J. Biomol. NMR***17,** 137–151 (2000).

Supplementary Table 9. Different metrics of DNA flexibility in the Cartesian space for the Drew-Dickerson dodecamer simulation using parmbsc0 and parmbsc1 force-fields.

Metrics	Parmbsc1	Parmbsc0
Entropy all heavy [a]	2.14	2.14
	2.00	*2.00*
Entropy backbone	1.16	1.15
	1.11	*1.10*
First three eigenvalues [b]	176,127,102	204,135,104
Eigenvalues 10, 20 and 30	20,8,4	23,9,4
Self-similarity (10 eigenvalues)[c]	0.89	0.94
Similarity parmbsc1/parmbsc0[d]	0.81	
Relative similarity[e]	0.89	
Energy weighted similarity	0.88	
Relative weighted similarity	0.93	

[a] Entropies in kcal mol^{-1} K^{-1} are determined using Schlitter (roman) and Andrioacei-Karplus (italics) for the entire 1.2 μs simulations. [b] Eigenvalues (in $Å^2$) are computed by diagonalization of the covariance matrix and ordered according to their contribution to the total variance. [c] Self-similarity is computed by comparing the first and second halves of the same trajectory. [d] Similarity and weighted similarity indexes are computed using the Hess matrix[1], or following reference[2]. [e] Relative similarities are computed from absolute similarities and self-similarities as described elsewhere[3].

1. Hess, B. *Phys. Rev. E***62,** 8438 (2000).
2. Pérez, A. *et al. J. Chem. Theory Comput.***1,** 790–800 (2005).
3. Orozco, M., Pérez, A., Noy, A. & Luque, F.J. *Chem. Soc. Rev.***32,** 350–364 (2003).

Supplementary Table 10. Sequence-dependent dinucleotide force constants associated with the deformation of a single helical degree of freedom.[a]

bps	Twist	Tilt	Roll	Shift	Slide	Rise
AA	0.028	0.037	0.020	1.72	2.13	7.64
	0.036	*0.045*	*0.023*	*1.68*	*2.91*	*9.33*
	0.043	**0.044**	**0.022**	**2.45**	**3.56**	**9.47**
	(0.092)	(0.100)	(0.049)	(3.98)	(6.16)	(21.75)
AC	0.036	0.038	0.023	1.28	2.98	8.83
	0.047	*0.045*	*0.027*	*1.54*	*3.67*	*10.44*
	0.034	**0.034**	**0.025**	**1.55**	**3.33**	**8.31**
	(0.073)	(0.111)	(0.080)	(2.94)	(6.37)	(23.86)
AG	0.028	0.037	0.019	1.40	1.78	7.04
	0.031	*0.049*	*0.025*	*1.54*	*2.78*	*9.73*
	0.036	**0.045**	**0.022**	**2.00**	**2.82**	**9.35**
	(0.064)	(0.149)	(0.096)	(3.21)	(7.19)	(29.50)
AT	0.031	0.035	0.022	1.05	3.77	9.34
	0.031	*0.033*	*0.024*	*1.24*	*4.10*	*9.23*
	0.032	**0.032**	**0.023**	**1.21**	**3.49**	**7.32**
	(0.070)	(0.166)	(0.055)	(3.17)	(10.69)	(25.55)
CA	0.015	0.025	0.016	1.05	1.80	6.30
	0.028	*0.028*	*0.016*	*0.77*	*2.69*	*7.66*
	0.032	**0.027**	**0.018**	**1.60**	**2.19**	**6.71**
	(0.043)	(0.082)	(0.048)	(3.73)	(2.40)	(18.24)
CC	0.026	0.042	0.020	1.43	1.57	7.86
	0.032	*0.049*	*0.021*	*1.50*	1.78	*9.59*
	0.030	**0.043**	**0.021**	**1.53**	**1.74**	**8.96**
	(0.041)	(0.119)	(0.064)	(2.43)	(3.54)	(30.31)
CG	0.014	0.026	0.016	1.05	1.91	6.11
	0.024	*0.032*	*0.016*	*1.10*	*2.47*	*7.61*
	0.032	**0.024**	**0.017**	**1.82**	**2.48**	**6.64**
	(0.047)	(0.068)	(0.050)	(1.59)	(3.30)	(14.16)
GA	0.024	0.038	0.020	1.32	1.88	8.48
	0.034	*0.045*	*0.023*	*1.40*	*2.66*	*10.08*
	0.040	**0.041**	**0.024**	**2.27**	**3.40**	**10.12**
	(0.071)	(0.087)	(0.046)	(6.54)	(2.78)	(22.82)
GC	0.022	0.036	0.026	1.18	2.59	9.47
	0.031	*0.043*	*0.025*	*1.32*	*3.19*	*11.16*
	0.027	**0.031**	**0.028**	**1.70**	**4.79**	**9.43**
	(0.055)	(0.082)	(0.082)	(3.35)	(6.24)	(25.86)
TA	0.018	0.019	0.015	0.64	1.25	6.08
	0.028	*0.025*	*0.015*	*0.50*	*2.16*	*7.47*
	0.036	**0.021**	**0.015**	**0.93**	**1.52**	**6.61**
	(0.052)	(0.148)	(0.029)	(3.86)	(2.35)	(21.91)

[a] Parmbsc0 (roman)[1], parmbsc1 (italics), and CHARMM27 (bold) force-fields are compared with stiffness values derived from inspection of the X-Ray structural variability of the different base pair steps (in brackets)[2]. Note that values for a particular base pair step are diagonal entries of its stiffness matrix. Values reported in the table are averages over all the equivalent steps. The rotational values are in kcal mol^{-1} deg^{-2} and translational ones are in kcal mol^{-1} $Å^{-2}$.

1. Perez, A., Lankas, F., Luque, F. J. & Orozco, M. *Nucleic Acids Res.***36,** 2379–2394 (2008).
2. Olson, W.K., Gorin, A.A., Lu, X.-J., Hock, L.M. & Zhurkin, V.B. *Proc. Natl. Acad. Sci.***95,** 11163–11168 (1998).

Supplementary Table 11. Elastic properties derived from atomistic MD simulations of three sequences of DNA.[a]

Persistence length							Other stiffness descriptors	
DNA	*Roll*	*Tilt*	*Isotropic*	*Dynamics*	*Static*	*Total*	*Torsion module*	*Stretch Module*
Seq3[b]	41±10	63±16	49±11	*63±1*	*566±150*	*57±2* 41±20 49±20	48±19 *101±9*	1,373±195 *1,857±22*
Seq4a	41±8	64±14	50±9	*71±1*	*608±150*	*64±2* 42±23 50±23	49±13 *102±10*	1,430±210 *1,567±42*
Seq4b	41±7	65±15	50±9	*71±1*	*310±44*	*57±2* 39±20 48±21	46±13 *107±12*	1,476±185 *1,832±45*
Avg.	41±14	64±26	50±17	*68±2*	*495±211*	*59±4* 41±30 49±30	47±26 *104±18*	1,426±341 *1,752±65*

[a] Persistence lengths and torsion modules are in nm, and stretch module are in pN. Values in roman correspond to 2 bp windows, while values in italic correspond approximately to one DNA turn windows[1]: (i) persistence lengths are calculated by linearly fitting the directional decay from 2 bp until 11 bp sub-fragments, and the static contributions come from the distribution of sequence-dependent static bends obtained through the MD average structure; (ii) stretch modulus are obtained by linearly fitting end-to-end variances of all central sub-fragments containing from 8 bp up to 16 bp to avoid the very long end-effect; (iii) torsional modulus is evaluated by averaging the 38 central sub-fragments containing 11 bp. Only the central 48-mer of the 56-mers was considered to minimize end-effects. Underlined values were obtained using a local implementation of Olson's Monte Carlo procedure[2], without additional corrections, or including (underlined with a curved line) partial variance corrections as discussed in Noy and Golestanian 2012[1].[b] See **Supplementary Table 5** for the definition of the sequences. As reference experimental estimates for persistence lengths are around 50 nm[3], for static persistence lengths are in the range of 200-1,500 nm[4, 5], for stretch modulus are around 1,100-1,500 pN[6, 7] and for torsion (twist) constants are in the range 80-120 nm[8, 9].

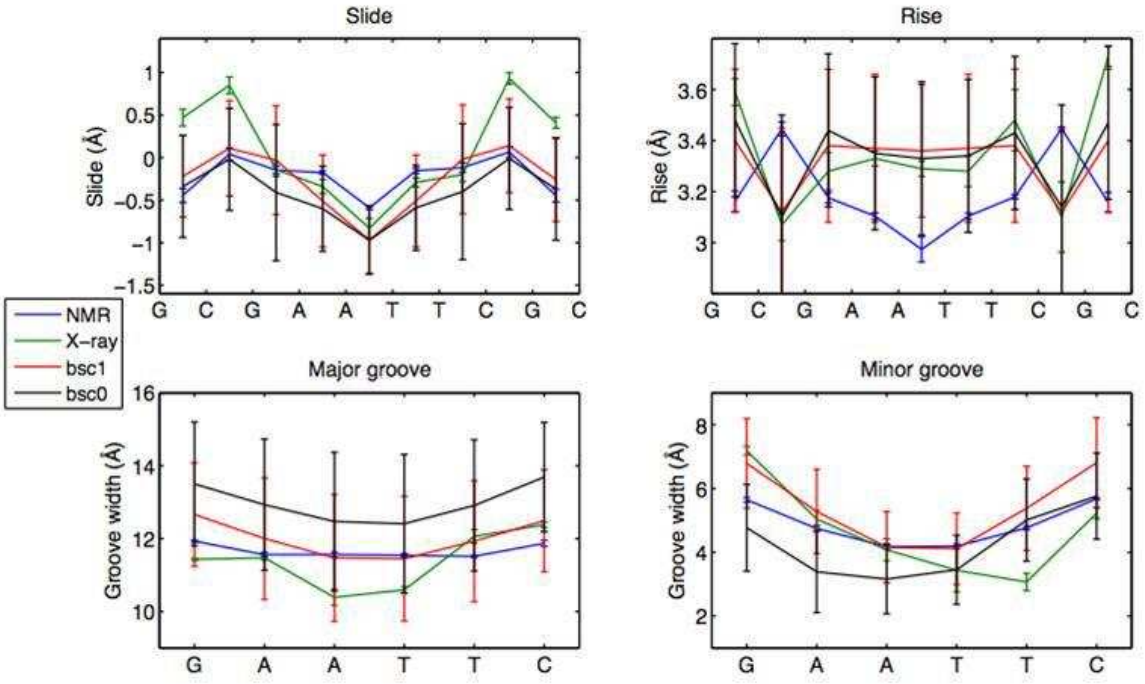
1. Noy, A. & Golestanian, R. *Phys. Rev. Lett.***109,** 228101 (2012).
2. Zheng, G., Czapla, L., Srinivasan, A.R. & Olson, W.K. *Phys. Chem. Chem. Phys.***12,** 1399–1406 (2010).
3. Mazur, A.K. & Maaloum, M. *Nucleic Acids Res.***42,** 14006-14012 (2014).
4. Smith, S.B., Finzi, L. & Bustamante, C. *Science***258,** 1122–1126 (1992).
5. Moukhtar, J. *et al. J. Phys. Chem. B***114,** 5125–5143 (2010).
6. Smith, S.B., Cui, Y. & Bustamante, C. *Science***271,** 795–799 (1996).
7. Gross, P. *et al. Nat. Phys.***7,** 731–736 (2011).
8. Strick, T.R., Allemand, J.-F., Bensimon, D., Bensimon, A. & Croquette, V. *Science***271,** 1835–1837 (1996).
9. Moroz, J.D. & Nelson, P. *Proc. Natl. Acad. Sci.***94,** 14418–14422 (1997).

Supplementary Table 12. Differences between QM and force-field estimates for the parameterized systems. Values refer to calculations performed in water.

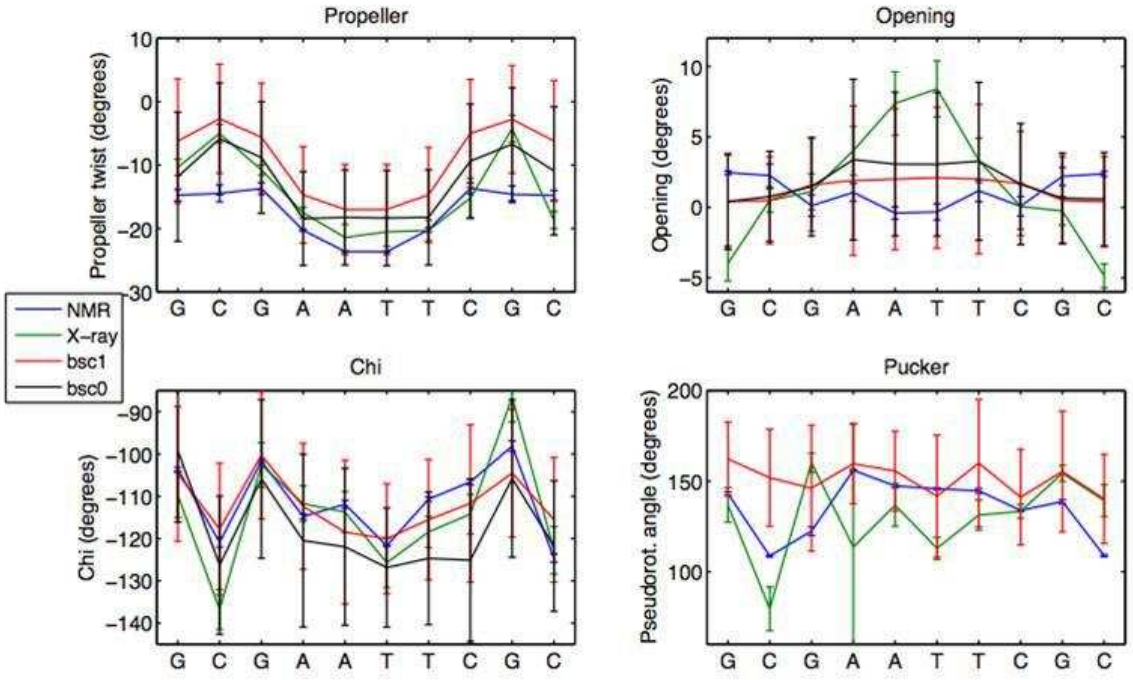
Torsion	Adenosine	Guanosine	Cytosine	Thymidine
Glycosidic torsion (χ)				
Geometries (°) [a]				
Anti	14 / 40	9 / 40	2.5 / 1	2.5 / 1
Barrier	1.5 / 11	2.5 / 15	13 / 10	11 / 11
Syn	7 / 32	2.5 / 30	12 / 30	–12 / 30
Energies (kcal mol^{-1}) [b]				
Anti/Syn	0.0 / –0.3	–0.4 / –0.6	–1.1 / 1.3	–0.8 / 1.7
Barrier [c]	0.3 / –2.0	0.0 / –2.1	–0.6 / –0.7	–0.9 / –1.2
Profile	0.3 / 2.5	1.2 / 2.8	0.9 / 4.0	0.9 / 3.9
Phase angle (P)				
Geometries (°) [a]				
North	10 / 30	10 / 10	10 / 40	0 / 10
East	0 / 10	0 / 0	10 / 10	0 / 10
South	0 / 0	10 / 10	0 / 0	0 / 0
Energies (kcal mol^{-1}) [b]				
North/South	–0.1 / –1.5	0.0 / –1.0	–0.6 / –1.6	0.5 / –0.5
East Barrier	–0.2 / 0.4	–0.5 /0.7	–0.1 / 1.2	–0.8 / 0.0
Profile	0.4 / 0.6	0.5 / 0.4	0.4/ 0.7	0.2 / 0.5

[a] Errors in the position of the minima and transition state when parmbsc1 (first number in the cell) or parmbsc0 (second number in the cell) values are compared with MP2 geometries. [b] Errors in the estimates of the relative stability and transition barrier when parmbs1 (first number in the cell) or parmbsc0 (second number in the cell) values are compared with single-point CCSD(T)/CBS results. [C] Energy values refer to barrier at χ around 120 degrees, note that the large barrier located at χ around 0 is very well reproduced at the parmbsc1 level, but very poorly at the parmbsc0 one (**Supplementary Fig. 26**).

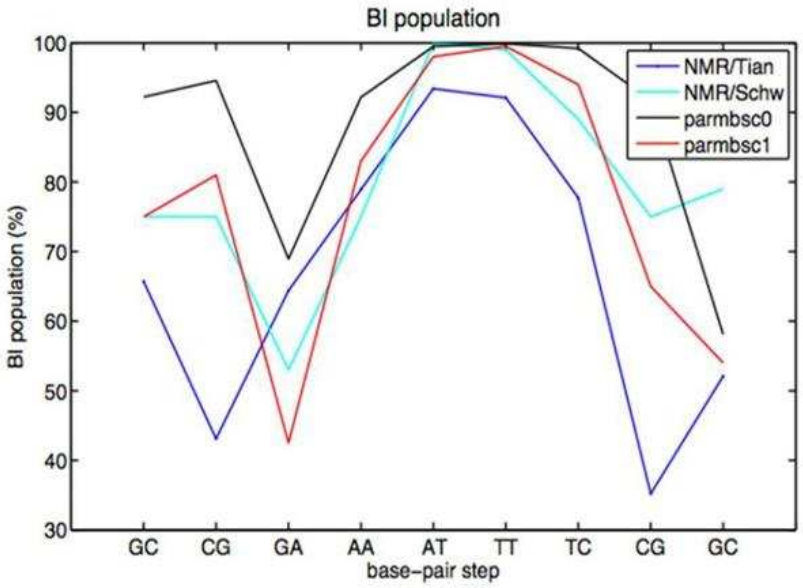
SUPPLEMENTARY FIGURES



Supplementary Figure 1| Helical parameters of DDD: Slide, Rise and grooves' width. Comparison of slide, rise, major and minor groove width average values per base-pair step coming from NMR structure pdb: 1NAJ (blue), X-ray structure pdb: 1BNA (green), 1 μs run using parmbsc0 force-field (black) and 1.2 μs run using parmbsc1 force-field.

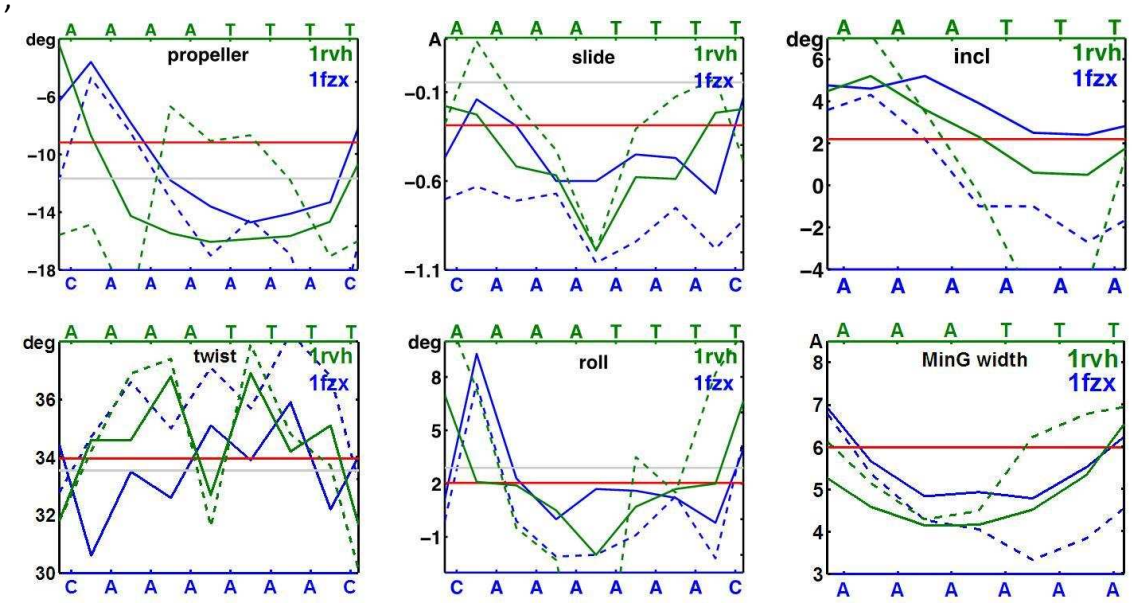


Supplementary Figure 2| Helical parameters per base-pair of DDD. Comparison of propeller twist, base opening, χ (chi) and pseudo-rotational angle (pucker) average values per base-pair step coming from NMR structure pdb:1NAJ (blue), X-ray structure pdb:1BNA (green), 1 μs run using parmbsc0 force-field (black), and 1.2 μs run using parmbsc1 force-field.

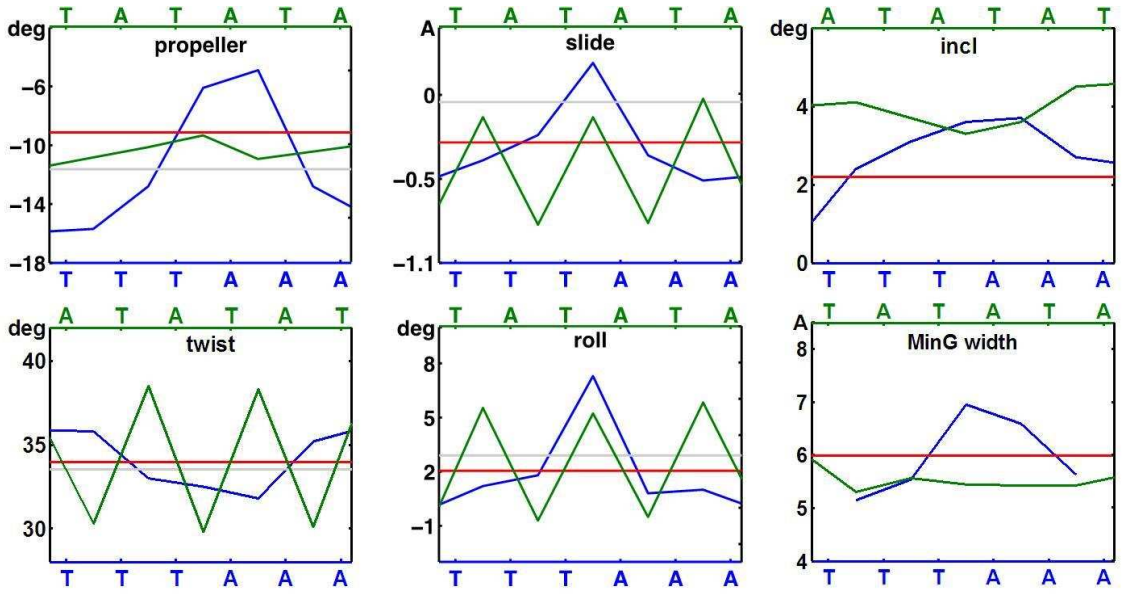


Supplementary Figure 3| BI/BII populations of DDD. Comparison of BI population percentage per base-pair step for DDD. Values coming from NMR/Tian *et al.*[1] (blue), NMR/ Schwieters *et al.*[2] (light blue), 1 µs run using parmbsc0 force-field (black) and 1.2 µs run using parmbsc1 force-field (red).

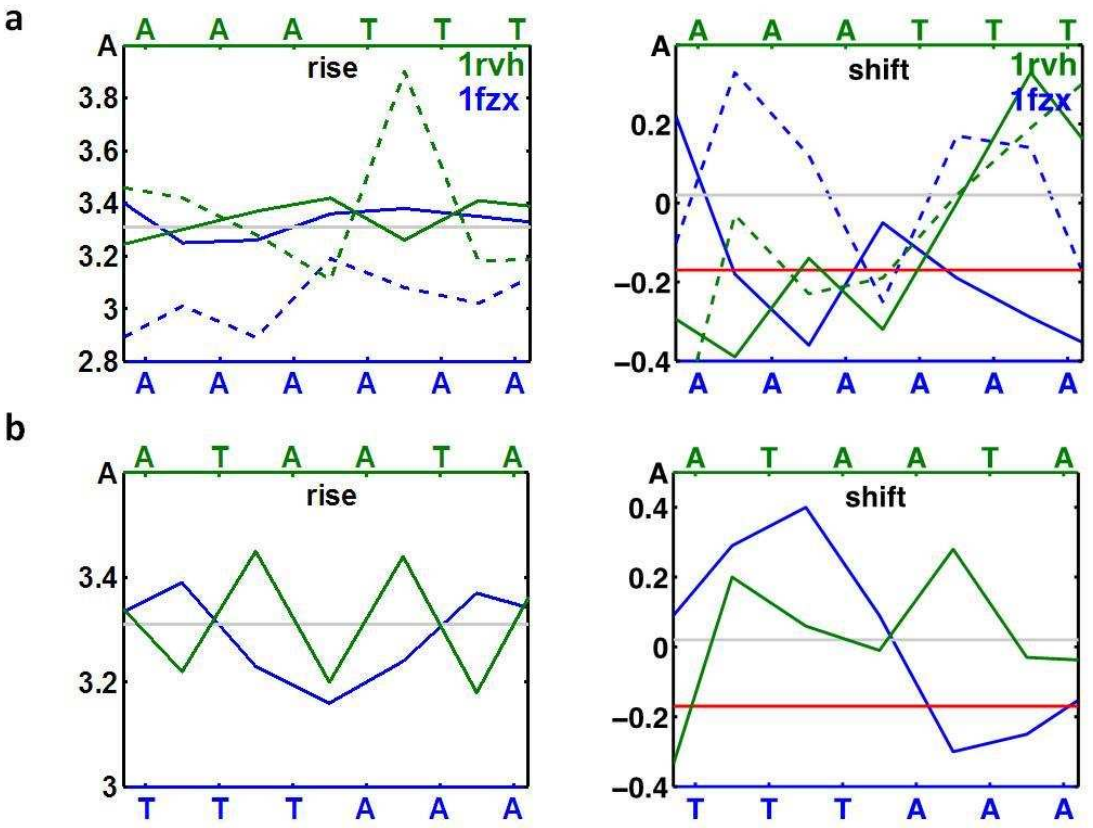
1. Tian, Y., Kayatta, M., Shultis, K., Gonzalez, A., Mueller, L.J., & Hatcher, M.E. *J. Phys. Chem. B***113**, 2596–2603 (2008).
2. Schwieters, C.D. & Clore, G.M., *Biochemistry***46**, 1152–1166 (2007).



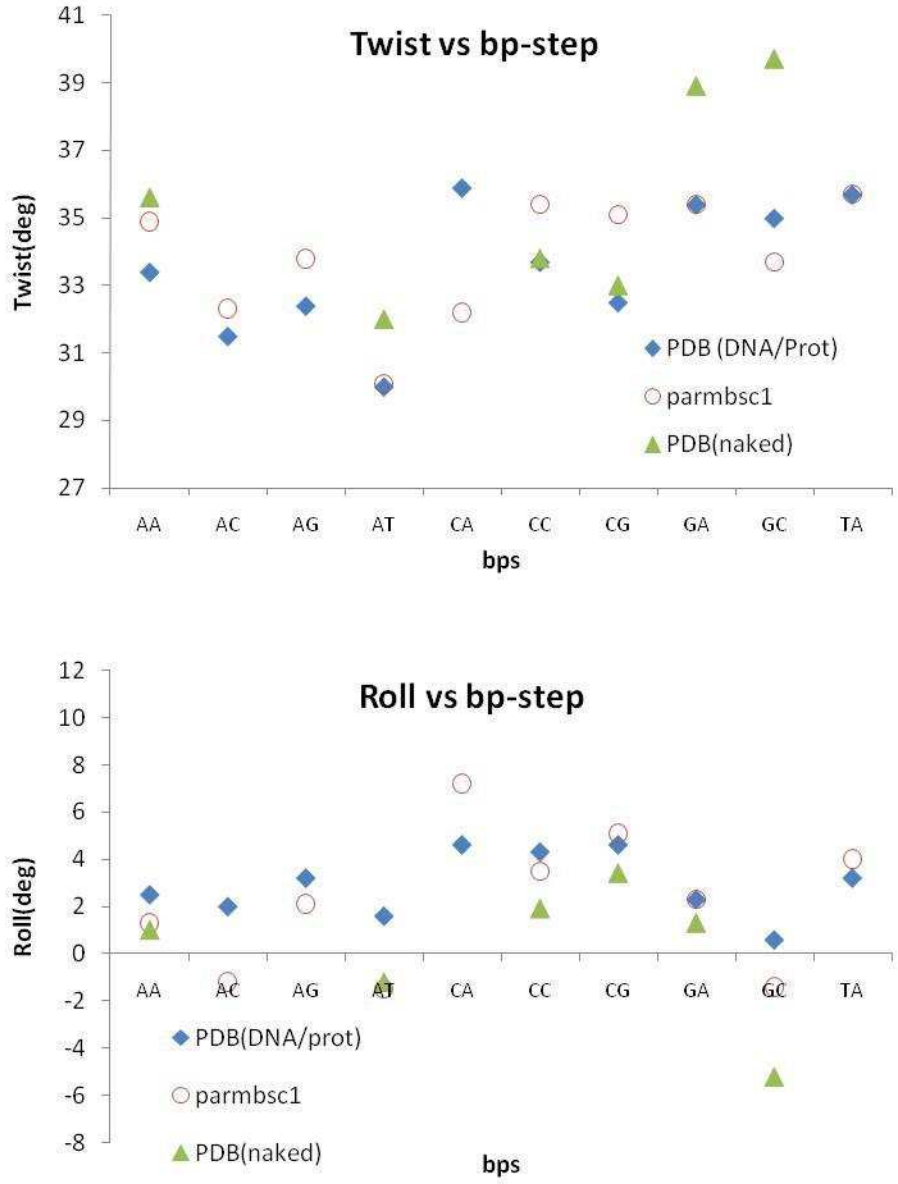
Supplementary Figure 4| Helical parameters of A-tract sequences: AATT and AAAA. Comparison in structural characteristics such as propeller twist, slide, inclination, twist, roll and minor groove width of values obtained using parmbsc1 force-field (full line) and experimental values (dashed lines) for AATT (pdb code:1RVH) (green) and AAAA (pdb code: 1FZX) (blue) sequences. Experimental average is represented with a grey line, while parmbsc1 average is represented with a red line.



Supplementary Figure 5| Helical parameters of A-tract sequences: ATAT and TTAA. Comparison in structural characteristics such as propeller twist, slide, inclination, twist, roll and minor groove width of values obtained using parmbsc1 force-field (full line) and experimental values (dashed lines) for ATAT (green) and TTAA (blue) sequences. Experimental average is represented with a grey line, while parmbsc1 average is represented with a red line.

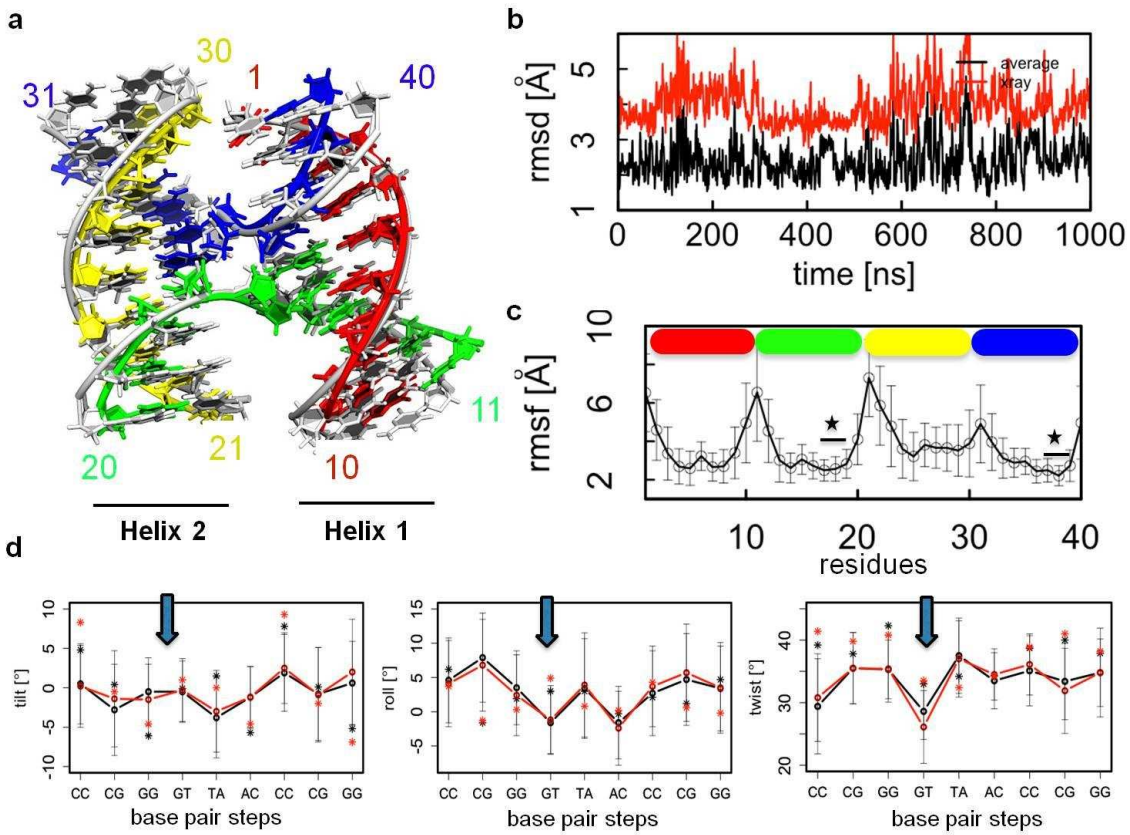


Supplementary Figure 6| Base-pair step helical parameters of A-tract sequences. Comparison in rise and shift of values obtained using parmbsc1 force-field (full line) and experimental values (dashed lines) for **(a)** AATT (pdb code:1RVH) (green) and AAAA (pdb code: 1FZX) (blue) and **(b)** ATAT (green) and TTAA (blue) sequences. Experimental average is represented with a grey line, while parmbsc1 average is represented with a red line.



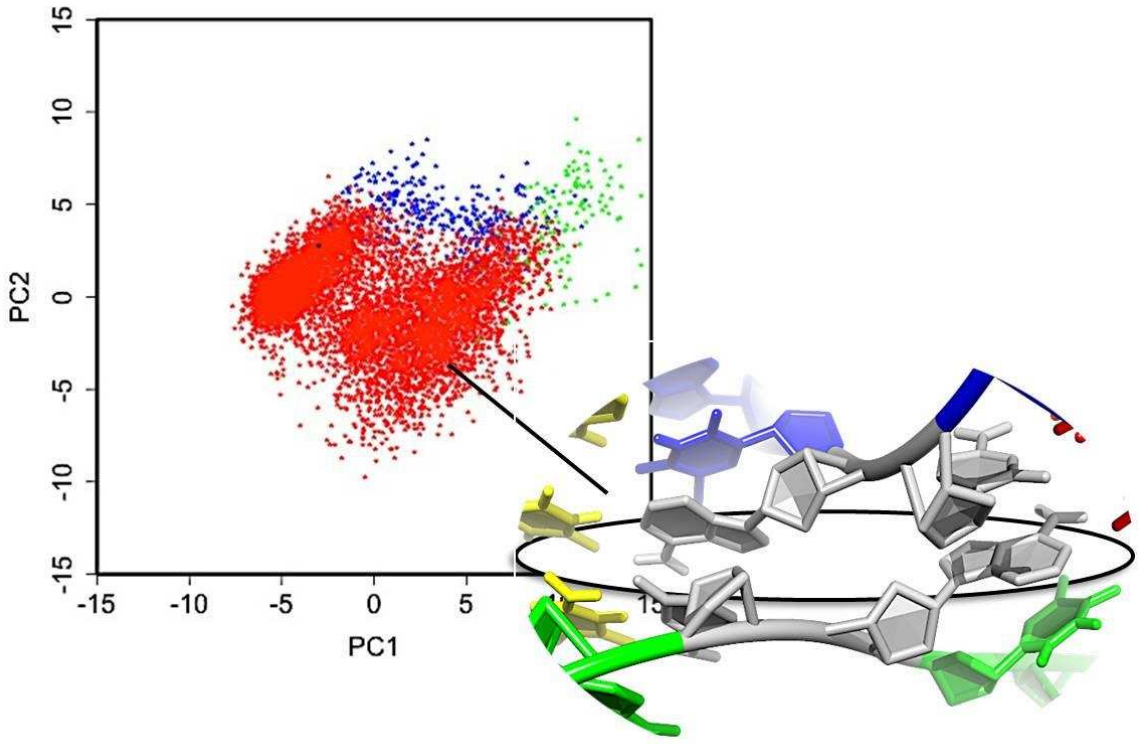
Supplementary Figure 7| Sequence-dependent variability of twist and roll. Comparison of DNA-protein complexes (blue), naked DNA (green) and parmbsc1 (red) values for twist (top) and roll (bottom) values per base-pair step. Values of DNA-protein complex come from analysis of 636 structures from PDB, while values of naked DNA come from analysis of 103 structures from PDB[1].

1. Dans, P.D., Pérez, A., Faustino, I., Lavery, R. & Orozco, M. *Nucleic Acids Res.***40**, 10668–10678 (2012).

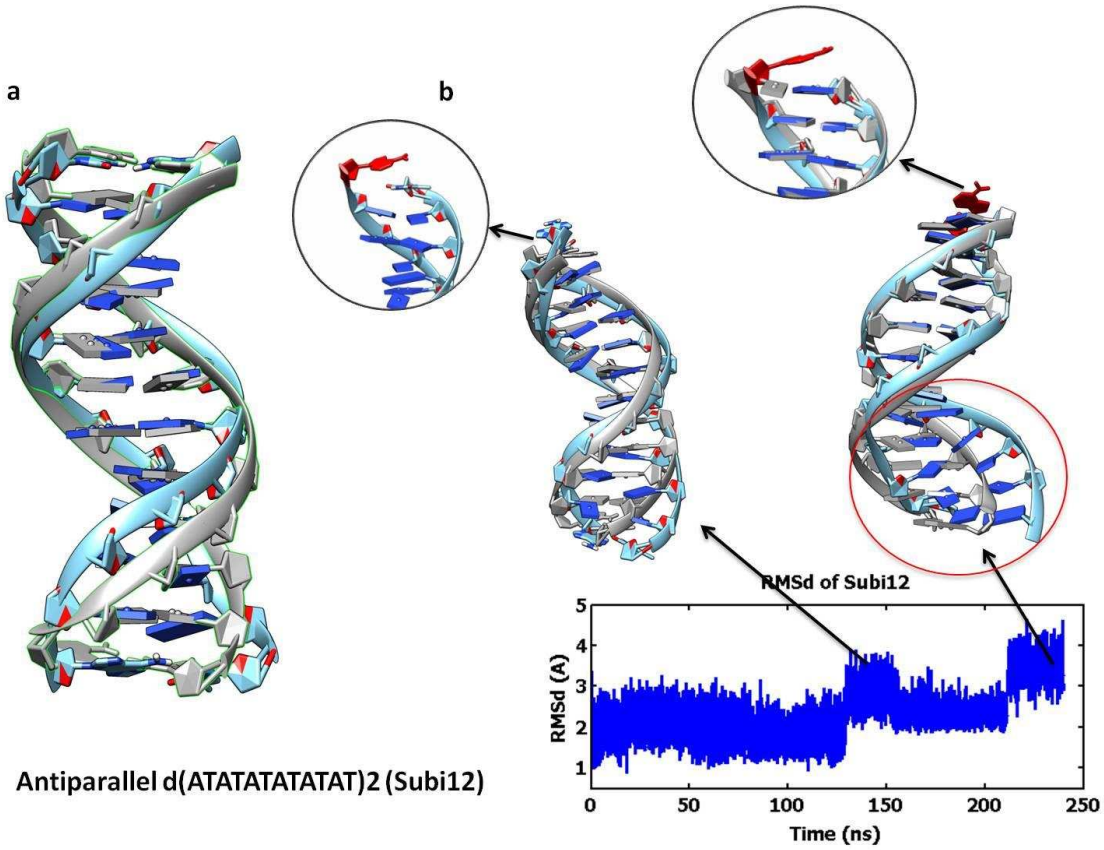


Supplementary Figure 8| Holliday junction structural features are close to x-ray (1DCW) structure. (**a**) Structural comparison of the time-averaged structure (in colors) with the x-ray reference structure (grey). (**b**) All heavy atoms RMSD and (**c**) per-residue RMSD from 1 µs MD simulation. X-ray structure was also taken as reference in the per-residue RMSD calculation. Note the higher RMSD values correspond to end strand bases. Starred residues are placed in the junction between helices. (**d**) Selected time-averaged helical parameters for the symmetric helices I and II. For experimental reference structures see ref. 1.

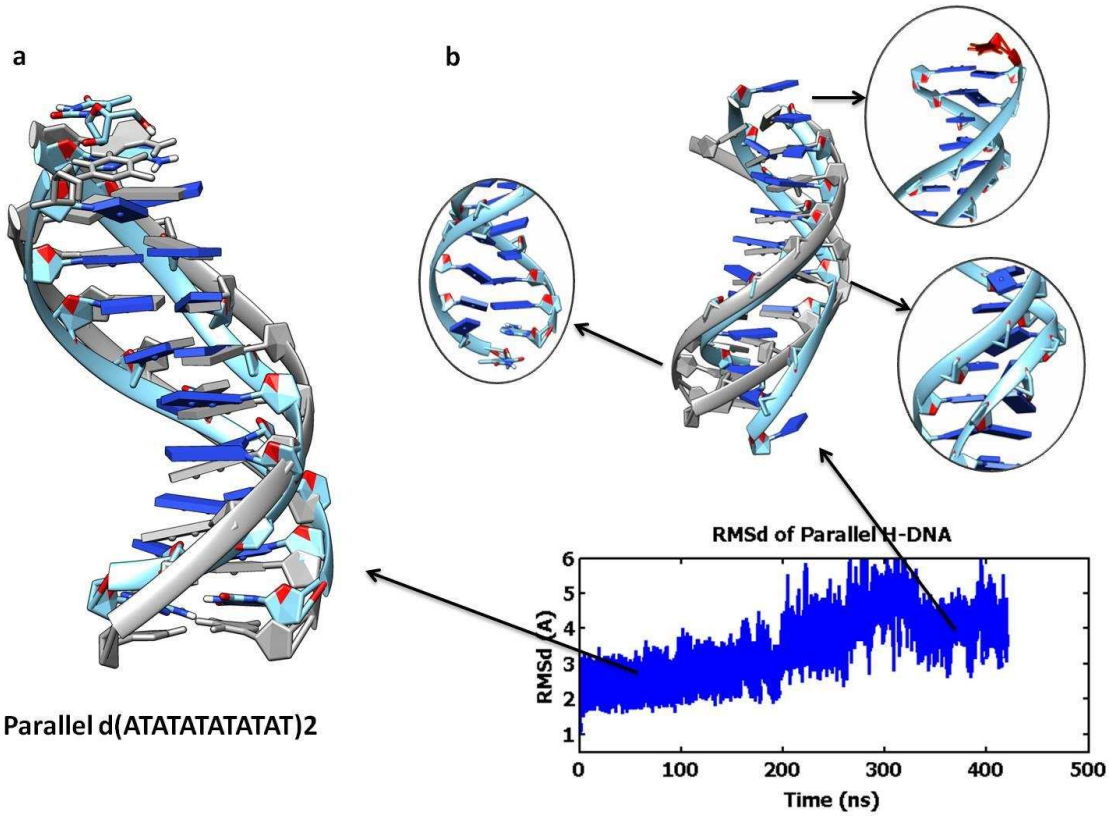
1. McKinney, S.A., Déclais, A.-C., Lilley, D.M.J. & Ha, T. *Nat. Struct. Mol. Biol.***10,** 93–97 (2003).



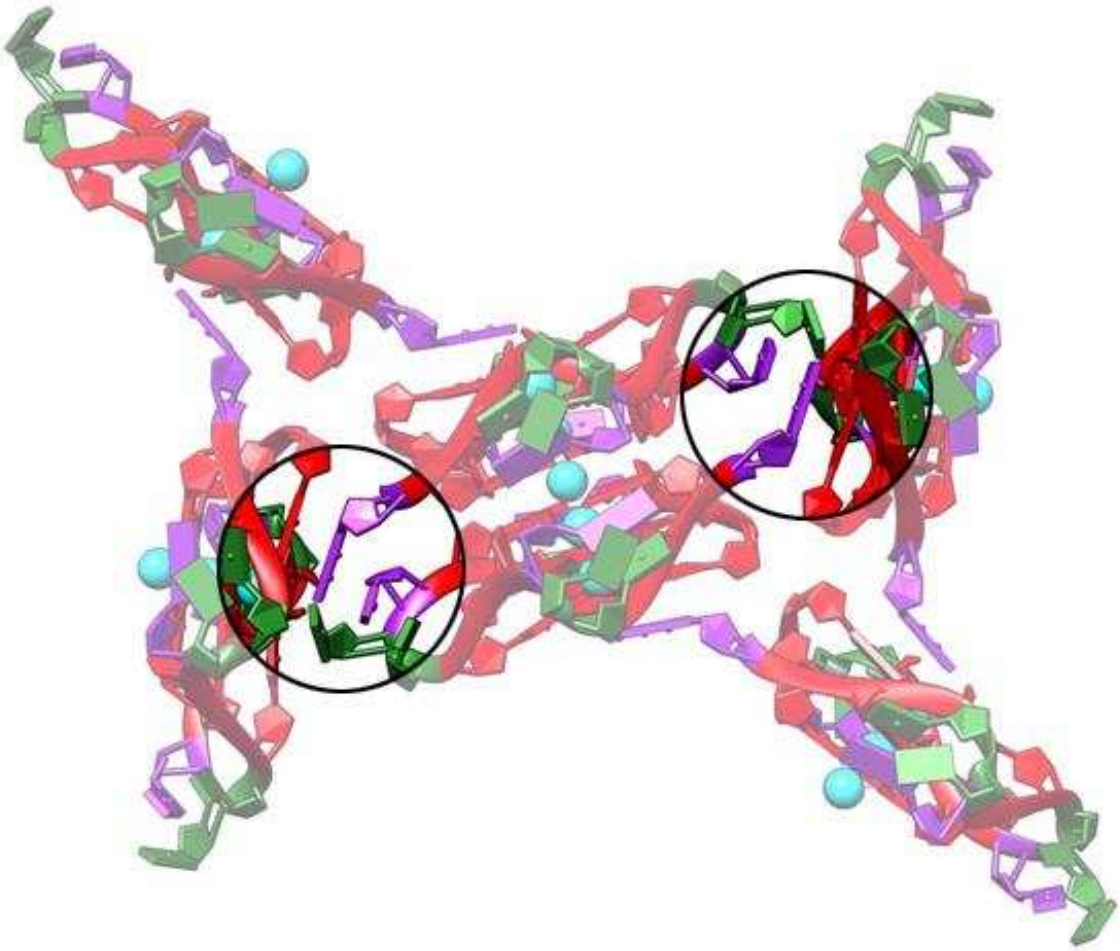
Supplementary Figure 9| Holliday junction PCA results. Projection to the first two PCA-eigenvectors based on the heavy atoms of junction bases (residues 16, 17, 36, and 37). The major conformation (in red) is present over ~95% of the simulation.



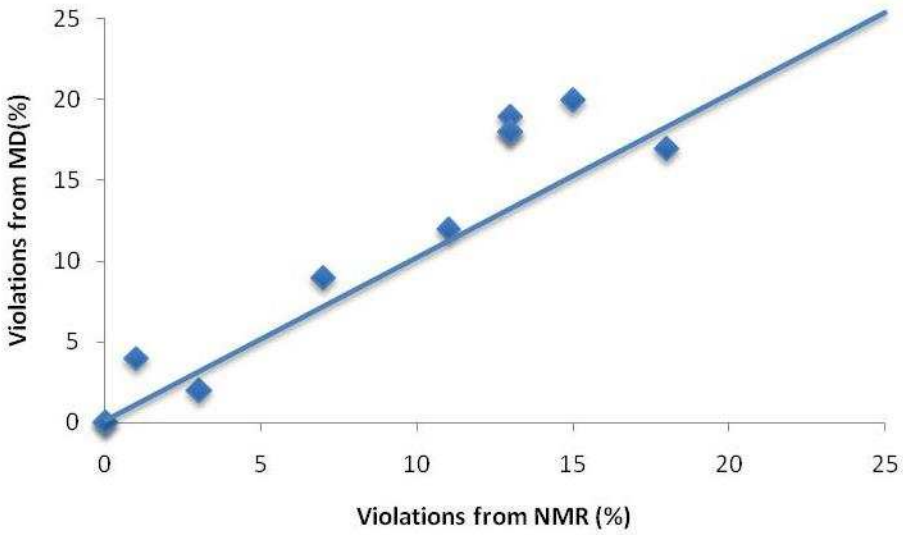
Supplementary Figure 10| Simulation antiparallel of H-DNA. (a) Comparison of experimental structure (made from pdb code: 1GQU) (grey) with the last snapshot of a 250 ns run using parmbsc1 (light blue). Bellow is an illustration of the duplex sequence. **(b)** RMSd of the 250 ns run with several snapshots plotted along the trajectory (light blue) compared with the experimental structure (grey) with highlighted distortions in the duplex.



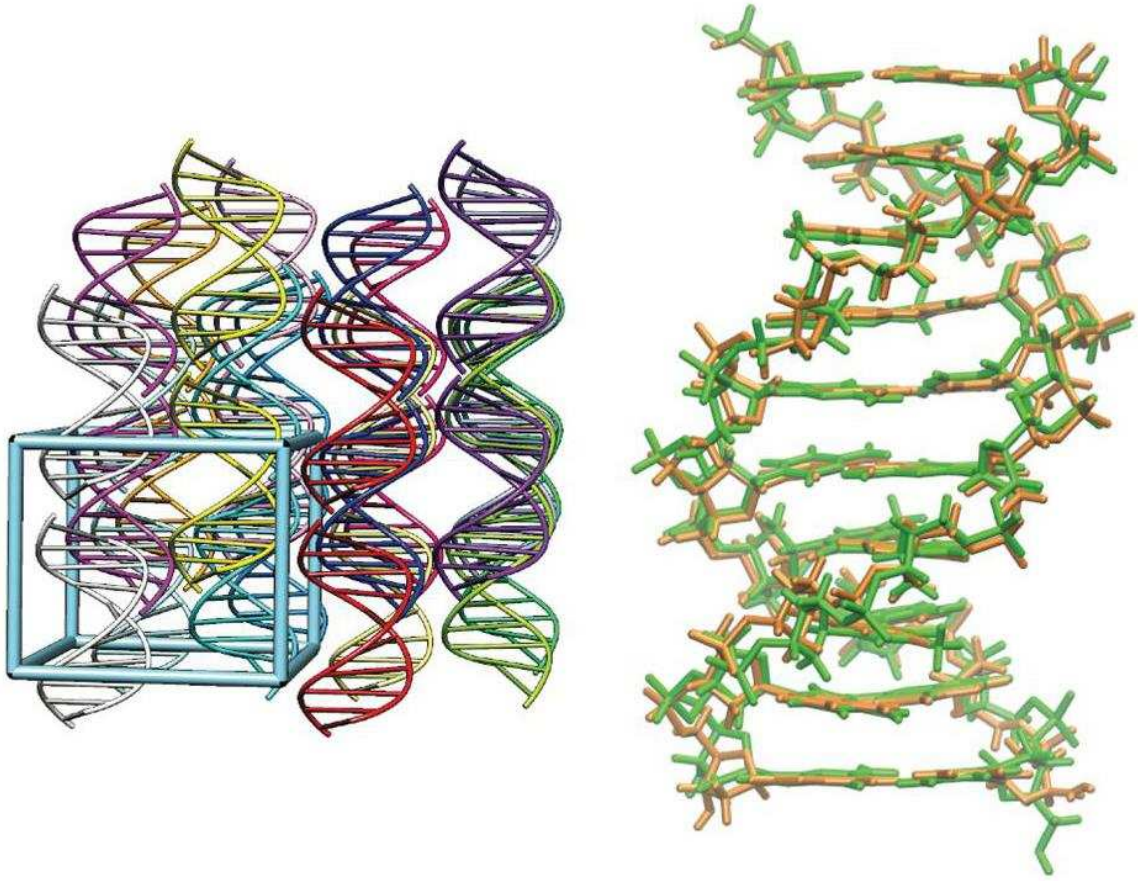
Supplementary Figure 11| Simulation of parallel H-DNA. (a) Comparison of experimental structure (grey) with a snapshot from a 400 ns run using parmbsc1 (light blue). **(b)** RMSd of the 400 ns run with several snapshots plotted along the trajectory (light blue) compared with the experimental structure (grey) with highlighted sever distortions in the duplex.



Supplementary Figure 12| Crystal packing of Human Talomeric Quadruplex (HTQ). Crystal packing of HTQ quadruplex (pdb code: 1KF1) showing interactions between loops' bases and other crystal units. Loop residues stacked to the neighboring units are highlighted in the circles.



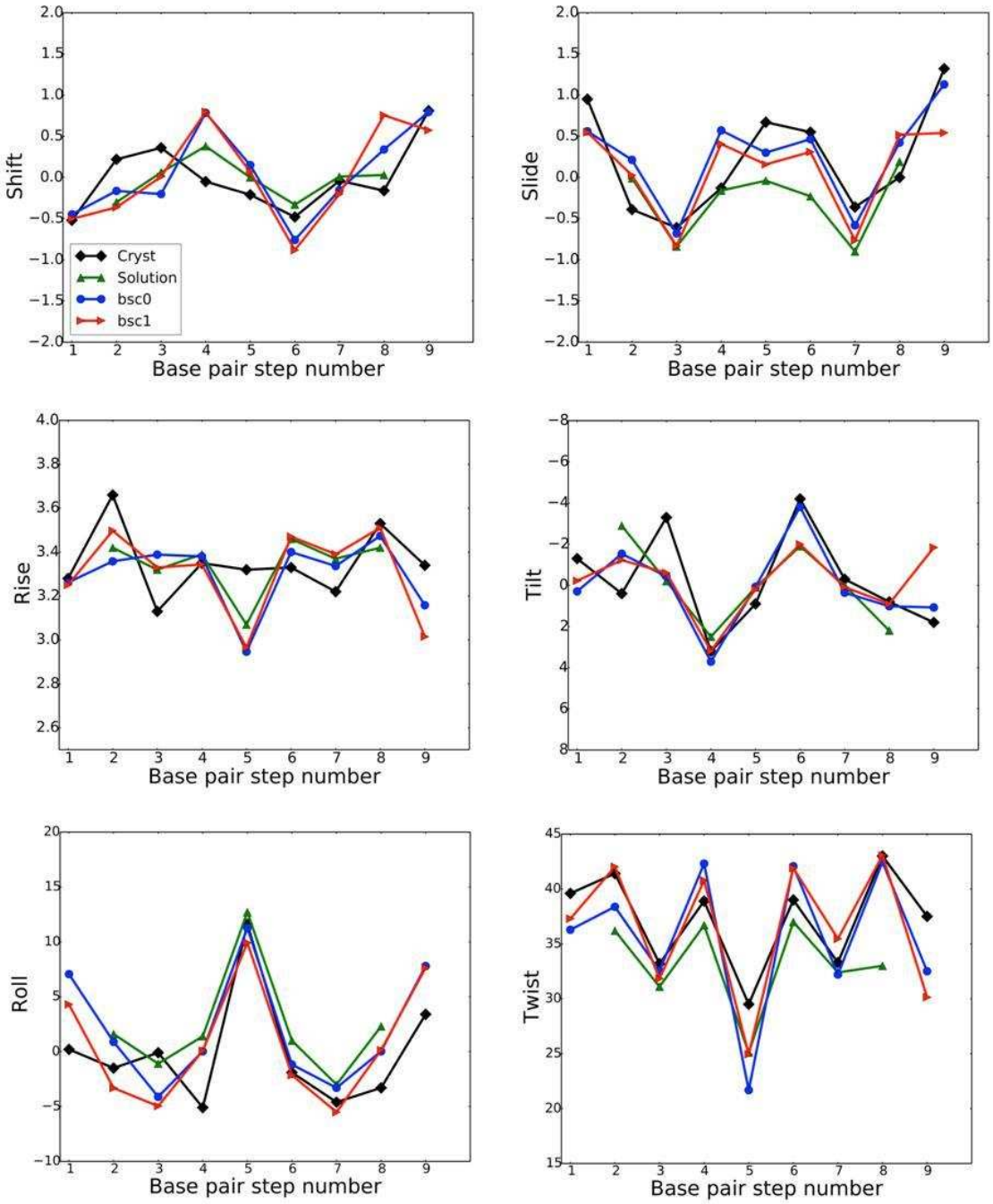
Supplementary Figure 13| Correlation between the number of violations in NOE restraints found in MD-parmbsc1 trajectories and corresponding NMR models. See **Supplementary Table 7** for details on structures.



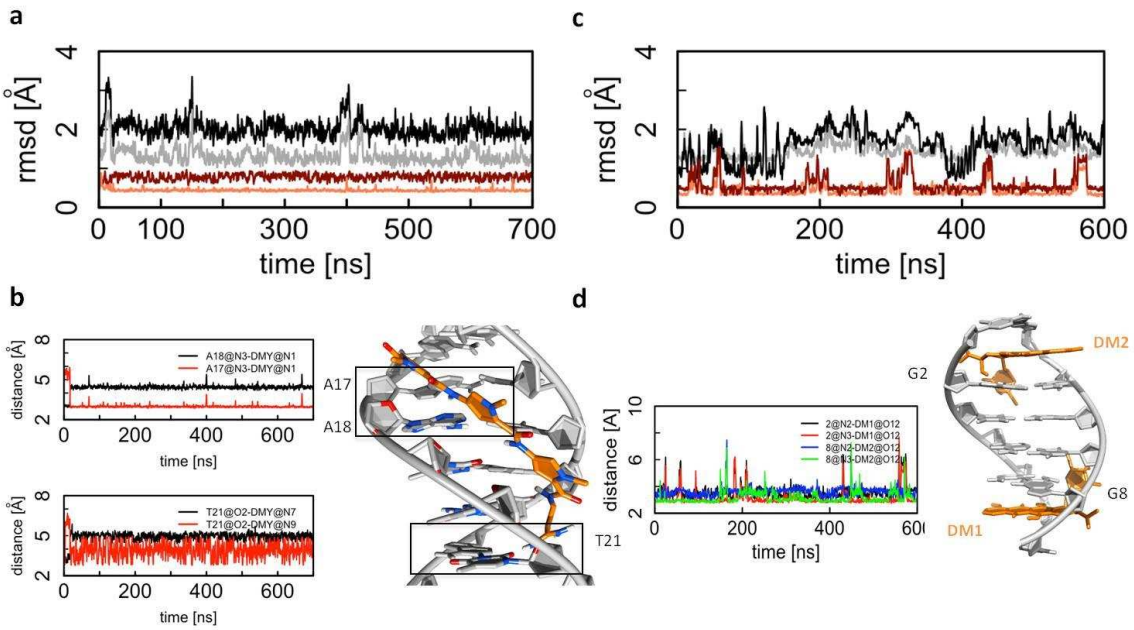
Supplementary Figure 14| Representation of the crystal structure simulation of a B-DNA duplex (PDB: 1D23). The **s**imulation box used in the crystal simulations is shown on the left, while comparison between the best-fit average structure from parmbsc1 simulations (orange) and the crystal structure (green) are shown on the right. Note that the RMS deviation for all DNA heavy atoms of the simulation average structure (compared to the PDB structure) is 0.70 Å. This can be compared to 0.77 Å for a crystal simulation using parmbsc0, and 1.83 Å for a solution simulation also using parmbsc0[1].

1. Liu, C., Janowski, P.A. & Case, D.A. *Biochim. Biophys. Acta (BBA)-General Subj.***1850**, 1059–1071 (2014).

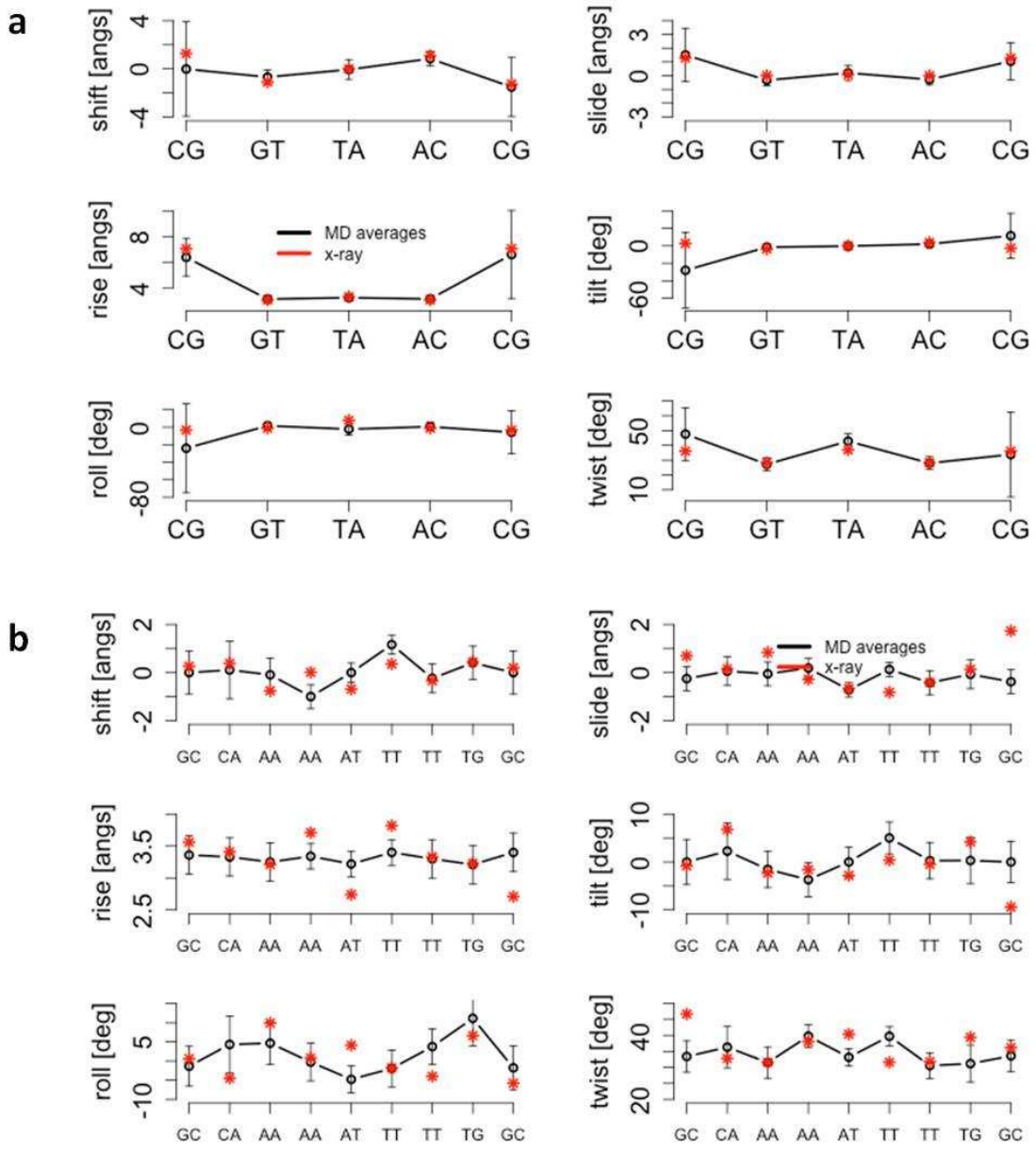
.



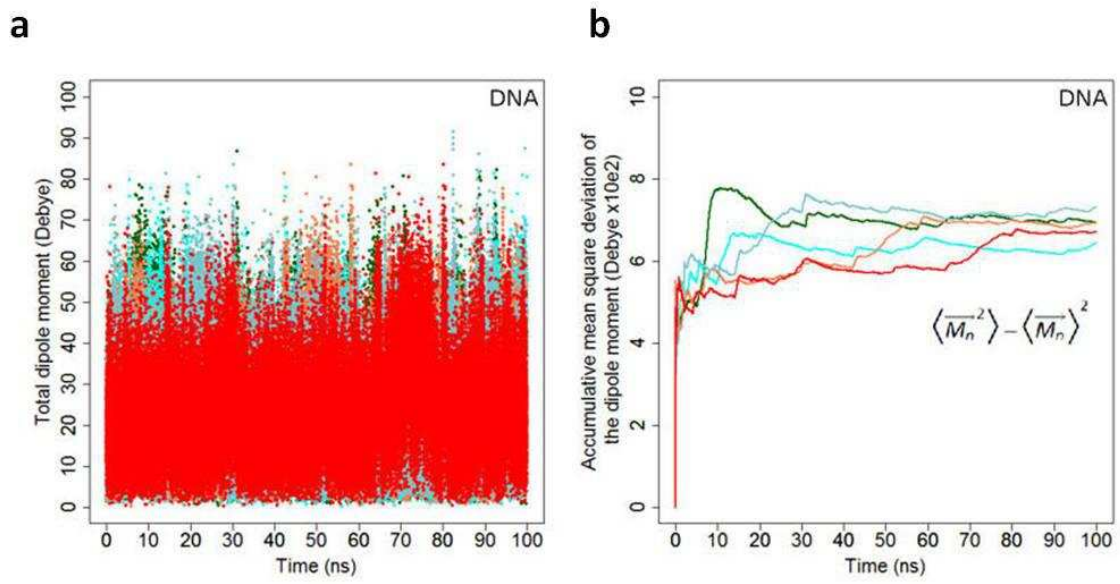
Supplementary Figure 15| Helicoidal analysis of a simulation of a B-DNA duplex (PDB: 1D23) within crystal environment. Helical parameters comparing results from simulation using parmbsc0 (blue) and parmbsc1 (red) force-fields, a simulation in solution (green) and the crystal structure (black).



Supplementary Figure 16| Representative stability properties in drug-DNA complexes with parmbsc1. RMSD (**a**) and representative distance between the distamycin A and the closest residues. (**b**) RMSD plots relative to x-ray (PDB id: 2DND), and MD-average structures for DNA (black and grey respectively) and distamycin A (red and orange respectively). Original contacts with the DNA are rapidly replaced by neighboring atoms keeping distamycin A within the minor groove. RMSD (**c**) and representative distances between the first daunomycin (PDB id: 1D11) and the closest guanine. (**d**) Second daunomycin's RMSd values are similar. Stabilizing interactions (h-bonds) between the N3 of guanine (residues 2 and 8 respectively) and a hydroxyl group in the daunomycin were stable along time.



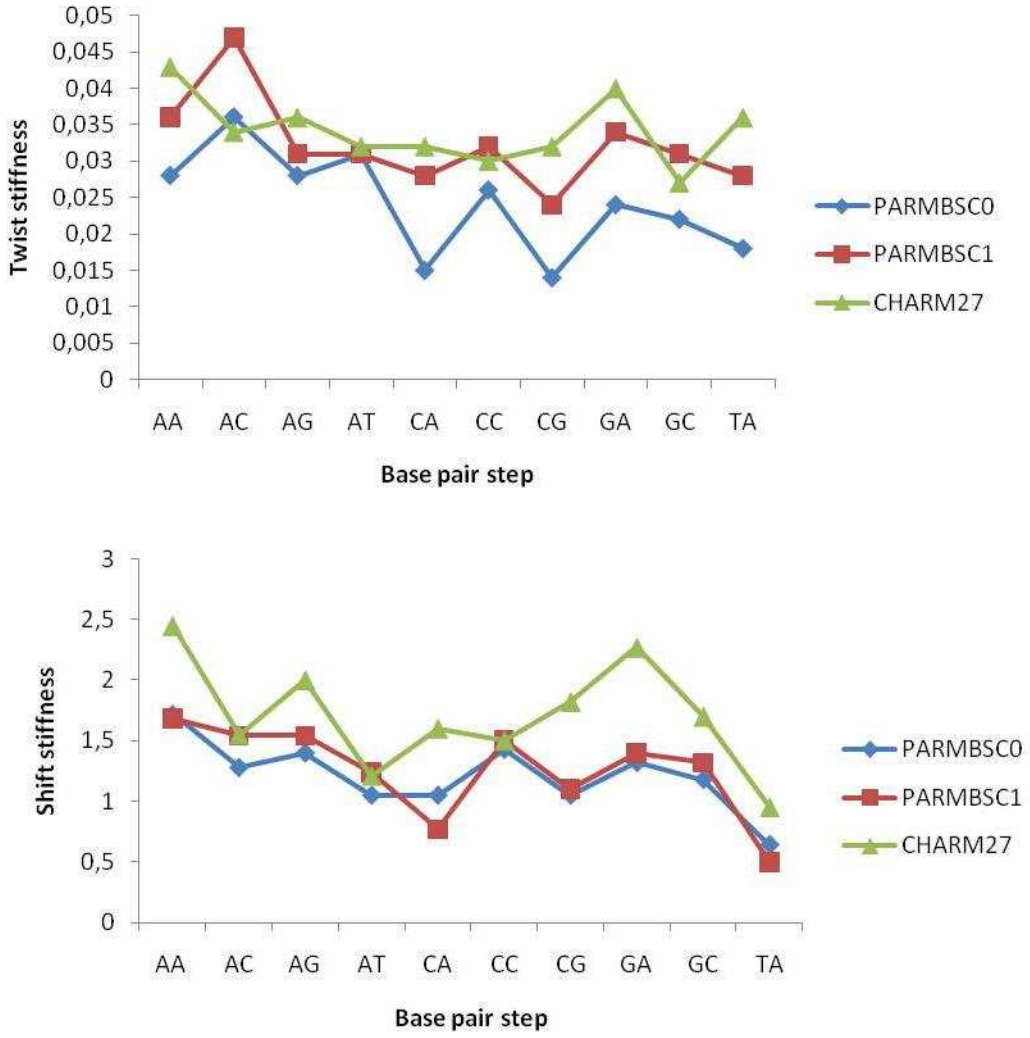
Supplementary Figure 17| Representative helical base pair step parameters in drug-DNA complexes. Time-averaged values associated to the DNA in complex with daunomycin (**a**) and distamycin A (**b**) in black compared with the original values from the X-ray structures (red, PDB id: 1D11 and 2DND for daunomycin and distamycin respectively).



Component	Volume (nm^3)	Dielectric constant parmbsc1 (ε_r)	Dielectric constant AFM (ε_r)
DNA	10.1 ± 0.2	8.0 ± 0.3	8.5 ± 1.4
$DNA_{sugar-phosphate}$	3.4 ± 0.2	19.9 ± 1.1	
$DNA_{sugar-base}$	7.6 ± 0.2	2.1 ± 0.2	

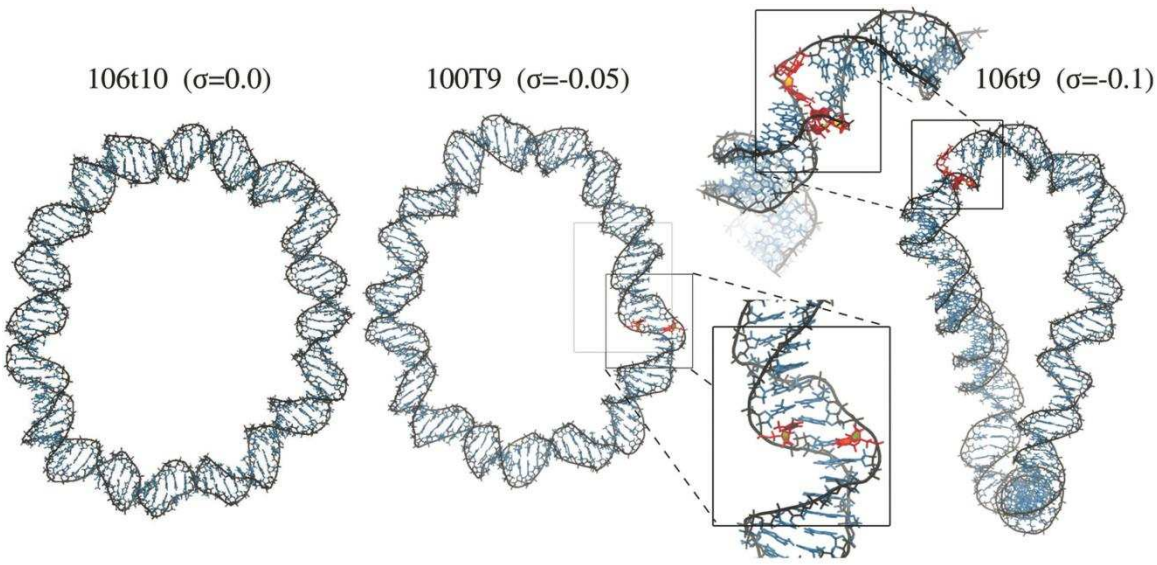
Figure 18| DNA dielectric constant. (**a**) Total dipole moment over time for 5 different replicas (100 ns each) taken from the microsecond long DDD simulation. (**b**) Accumulative mean square deviation of the dipole moment for the five replicas showing fairly good convergence after 30–40 ns. Values of whole DNA, sugar and phosphate groups, and sugar and base contributions are shown in the table below. See ref. 1 for the detailed procedure followed herein.

1. Cuervo, A., Dans, P. D.*et al. Proc. Natl. Acad. Sci.***111,** E3624–E3630 (2014).

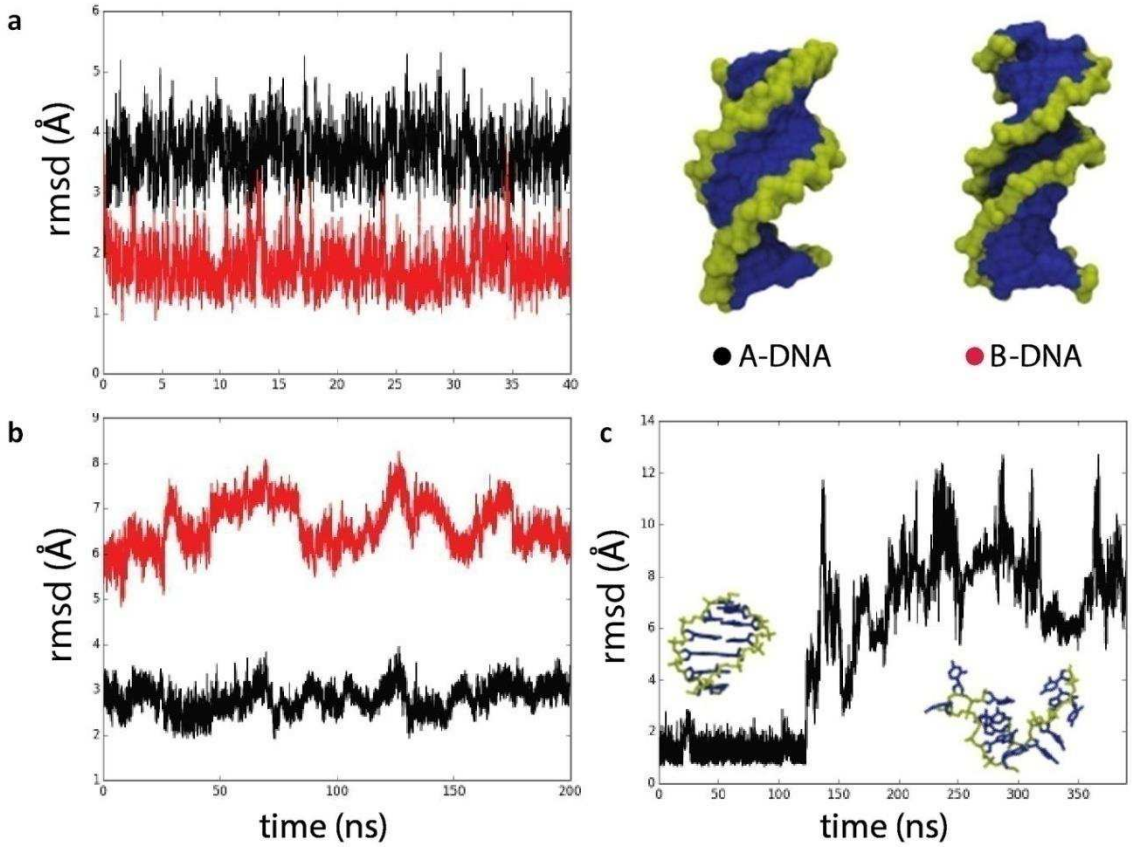


Supplementary Figure 19| Sequence dependent helical deformability. Variability of Twist (top) and Shift (bottom) stiffness constants for 10 unique base-steps. Parmbsc0 and CHARMM27 values are taken from ref 1.

1. Perez, A., Lankas, F., Luque, F.J. & Orozco, M. *Nucleic Acids Res.***36,** 2379–2394 (2008).

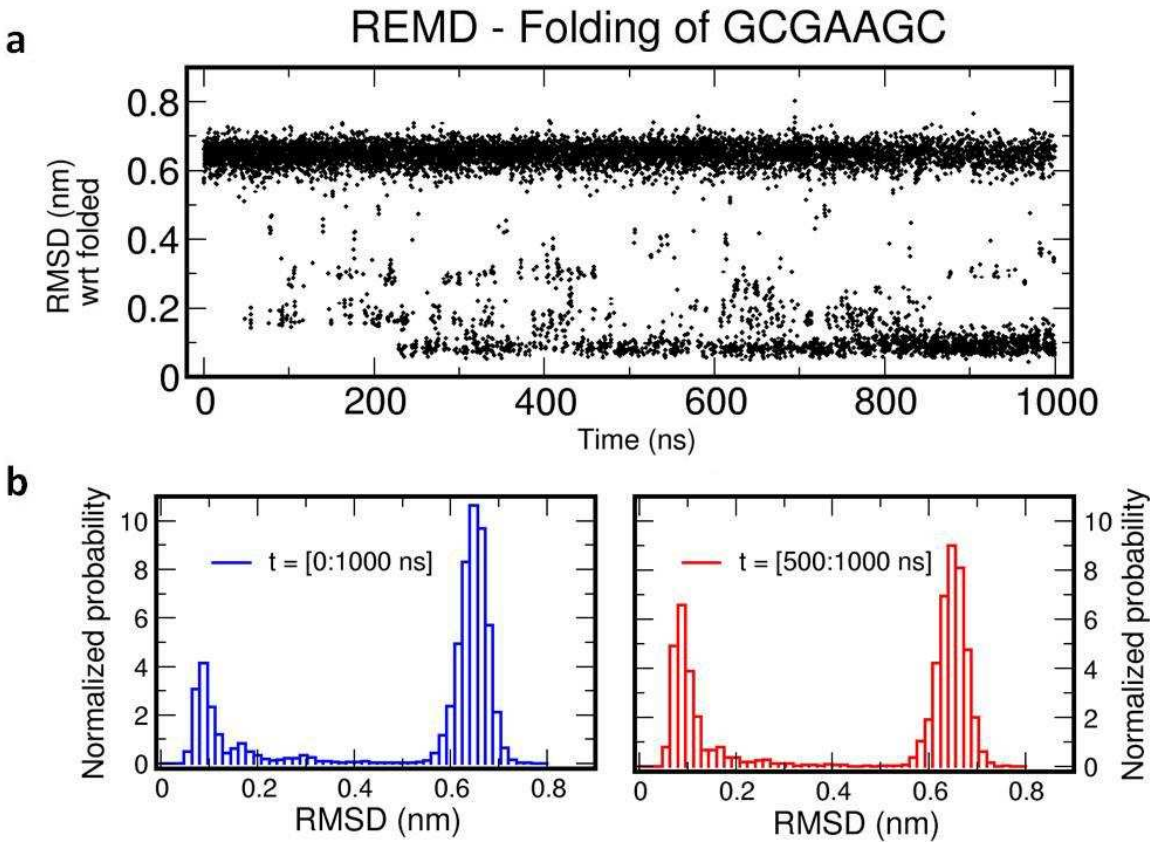


Supplementary Figure 20| Analysis of DNA minicircles. Final frames of the minicircles MD simulations. The secondary structure of the relaxed loop with 106 bp and 10 helical turns (106t10) remains intact, while the 2 negatively supercoiled circles show significant denaturalization. The 100 bp circle with 9 turns (100t9) presents 2 adjacent pyrimidine base-flipping towards the major groove, and the 106 bp circle with 9 turns (106t9), denature over multiple consecutive base pairs.



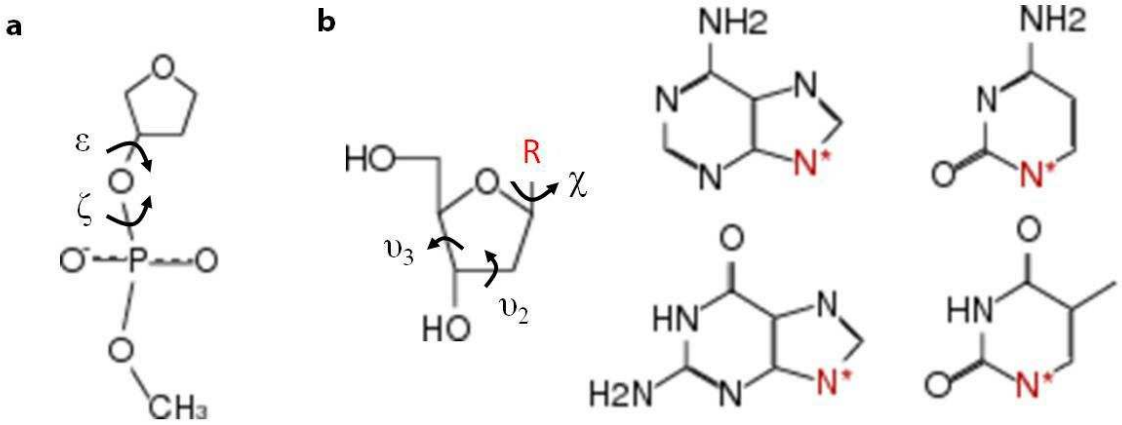
Supplementary Figure 21| MD simulations of conformational changes. (**a**) A to B transition simulation of DDD, where A-DNA form is presented in black with B-DNA in red. (**b**) Simulation of DDD in mixture of water and ethanol (see refs. 1 y 2 for additional discussion). (**c**) Unfolding of $d(GGCGGC)_2$ in 4 M pyridine water solution[3].

1. Soliva, R., Luque, F.J., Alhambra, C. & Orozco, M. *J. Biomol. Struct. Dyn.***17,** 89–99 (1999).
2. Ivanov, V.I., Minchenkova, L.E., Minyat, E.E., Frank-Kamenetskii, M.D. & Schyolkina, A.K. *J. Mol. Biol.***87,** 817–833 (1974).
3. Perez, A. & Orozco, M. *Angew. Chemie Int. Ed.***49,** 4805–4808 (2010).

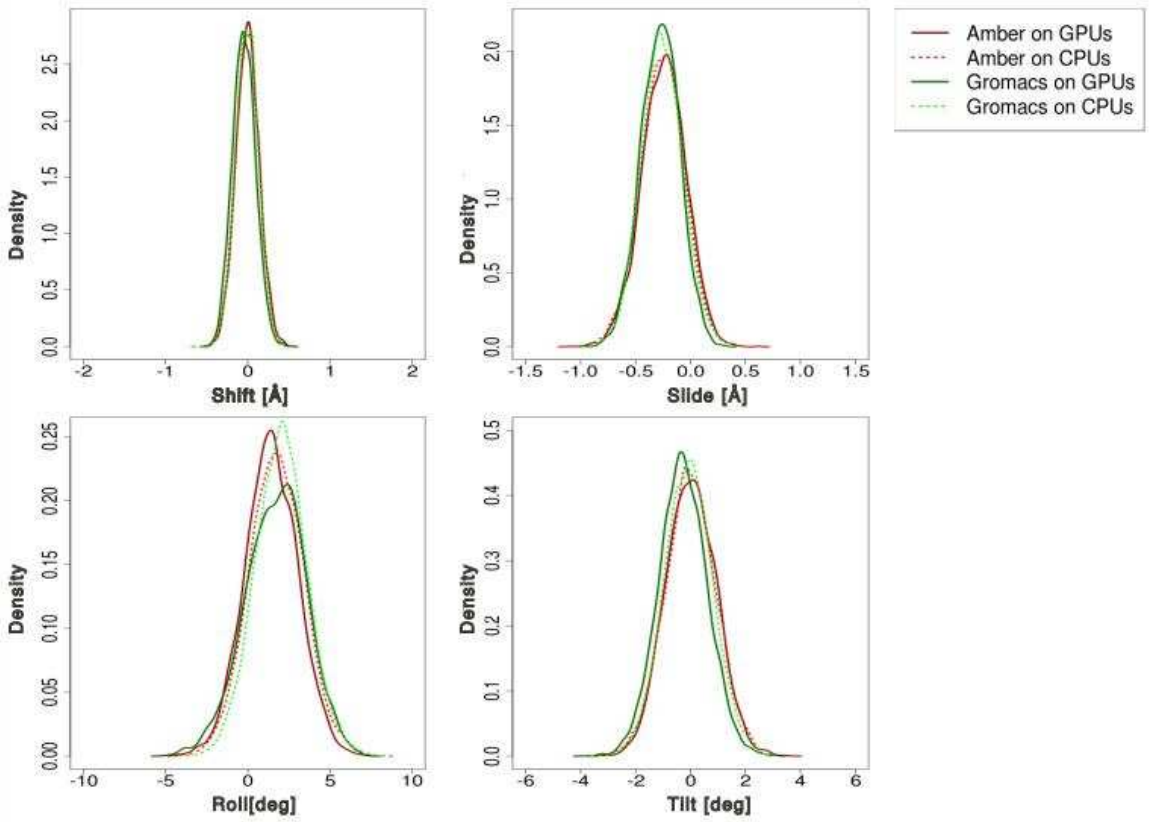


Supplementary Figure 22| Hairpin folding. Replica exchange MD (REMD) simulations of the folding of the small hairpin d(GCGAAGC) in water using parmbsc1 force-field. (**a**) RMSD with the respect to the folded state. (**b**) Probabilities of RMSDs in whole (blue) and second part (red) of microsecond runs of REMD. Structures are clearly recognizing the folded conformation and keeping it. For technical details see reference 1.

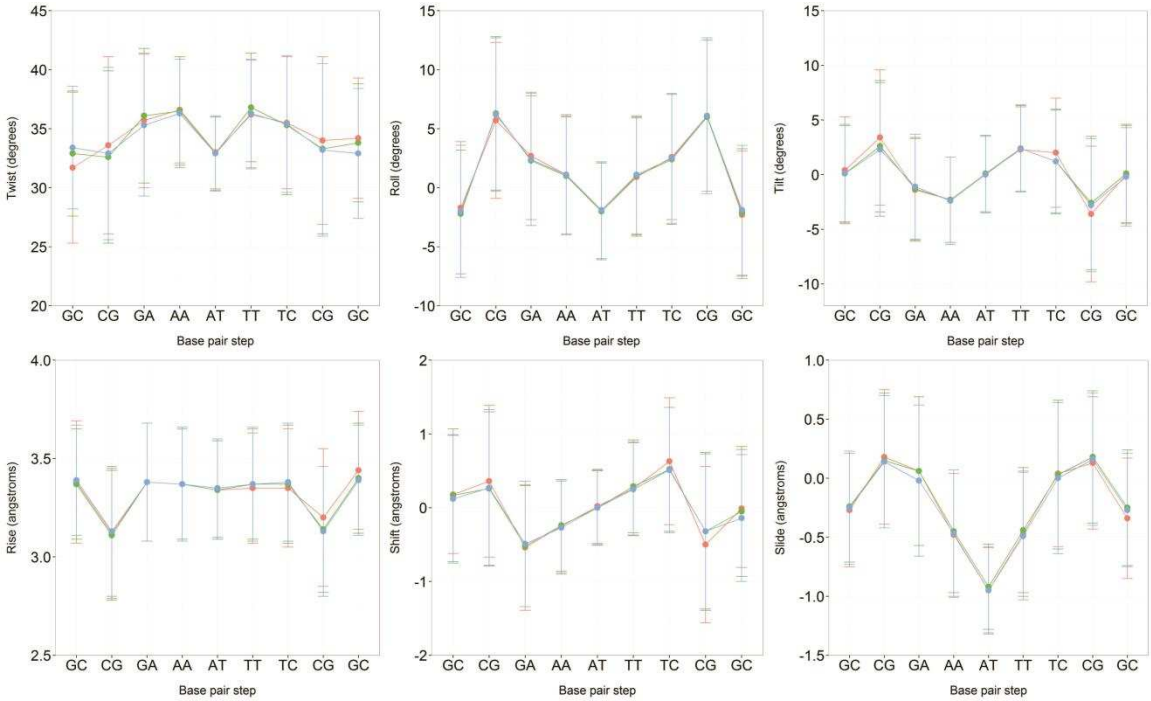
1. Portella, G., Orozco, M. *Angewandte chemie Int. Ed.***49**, 7673–7676 (2010).



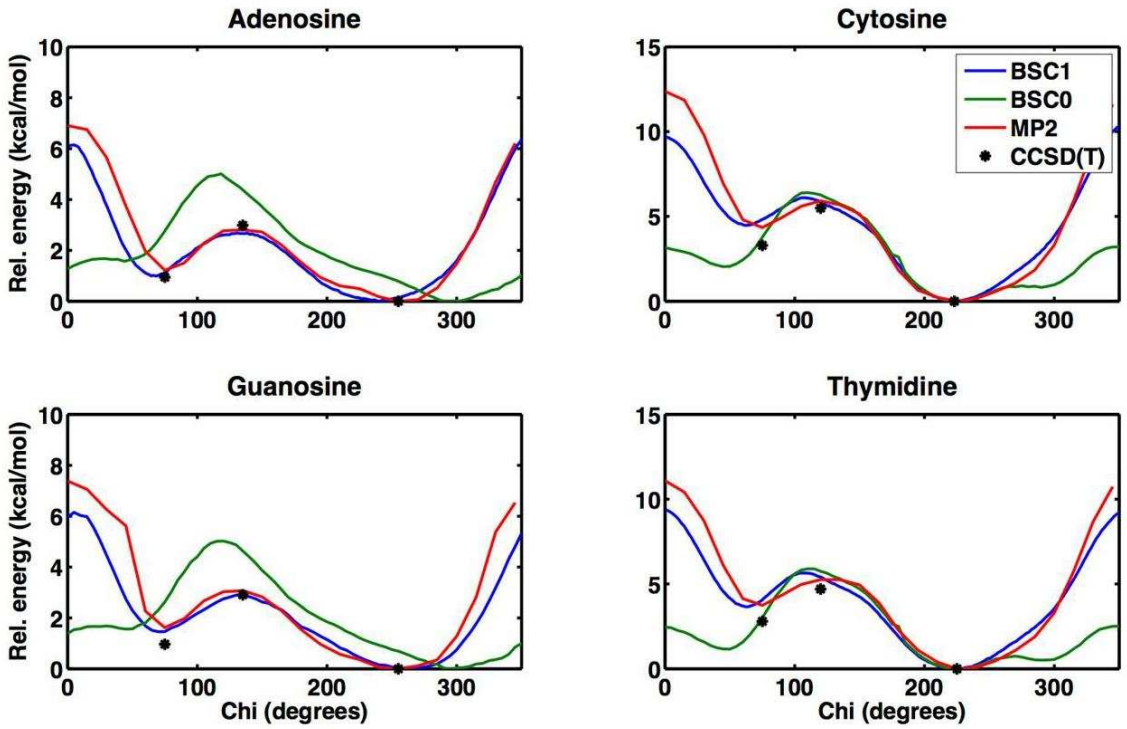
Supplementary Figure 23| **Model compounds used in QM optimization.** (**a**) Compound used for ε/ζ parameterization. (**b**) Compounds used for χ and sugar puckering parameterizations, where R represents the base, shown on the right.



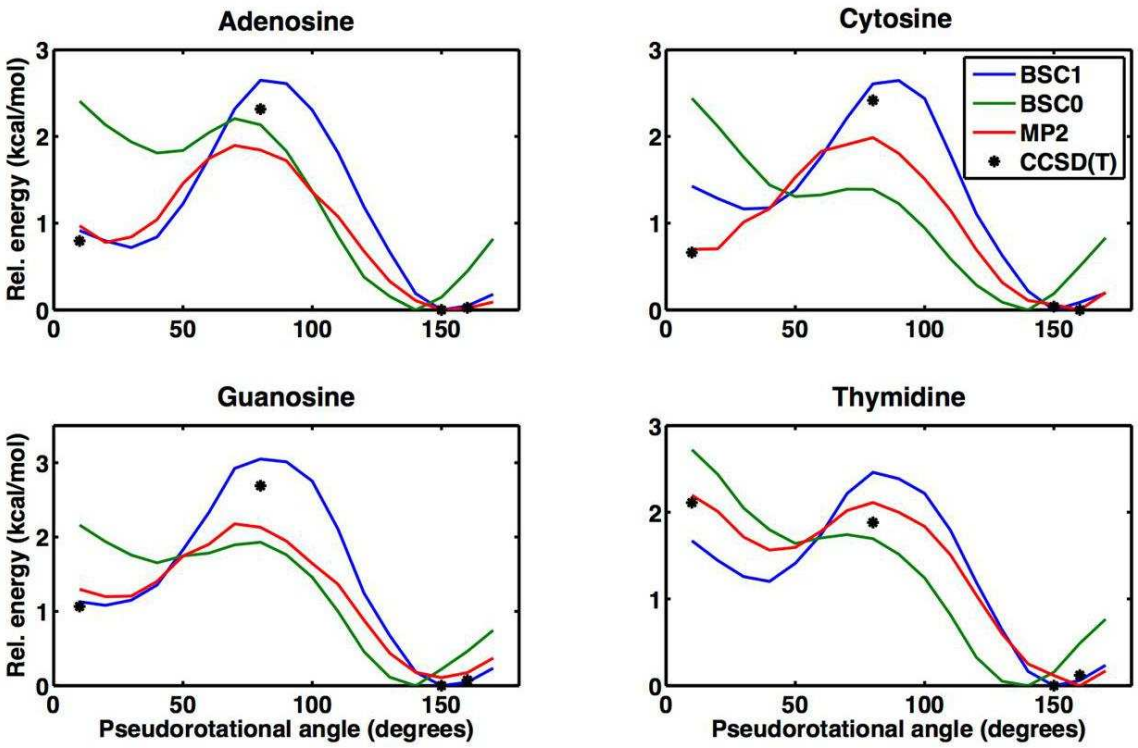
Supplementary Figure 24| Using DDD to compare different simulation engines. Normalized distributions of the helical parameters shift, slide, roll and tilt are shown for the four MD simulations (AMBER vs GROMACS, and GPU vs CPU codes). Due to the shortness of the simulation runs (100 ns), slight differences in roll angle can be detected using different MD engines.



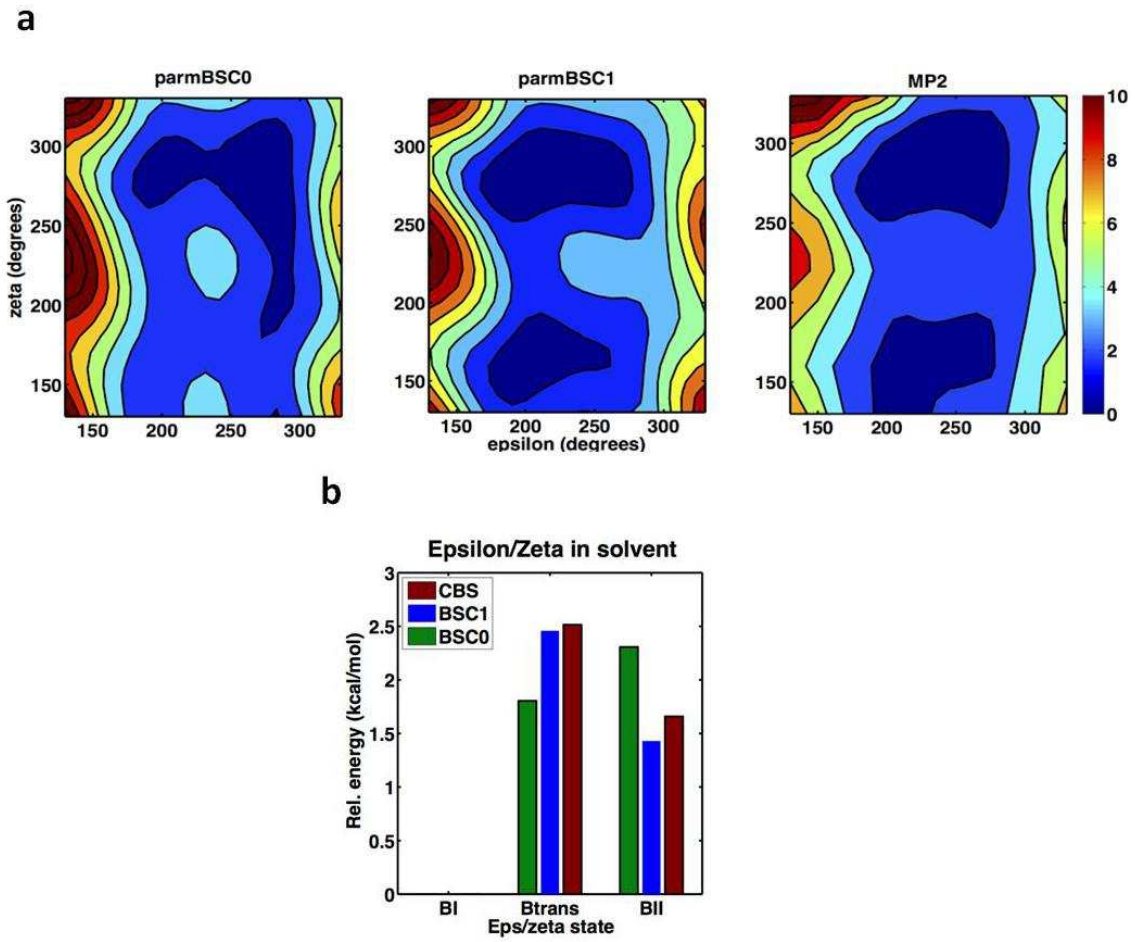
Supplementary Figure 25| Variation of helical parameters along the sequence for 2 µs of MD simulation of DDD with added salt (NaCl) concentrations: minimum Na^+ for neutrality (green), 150 mM (red) and 500 mM (blue). PME was used in all the cases.



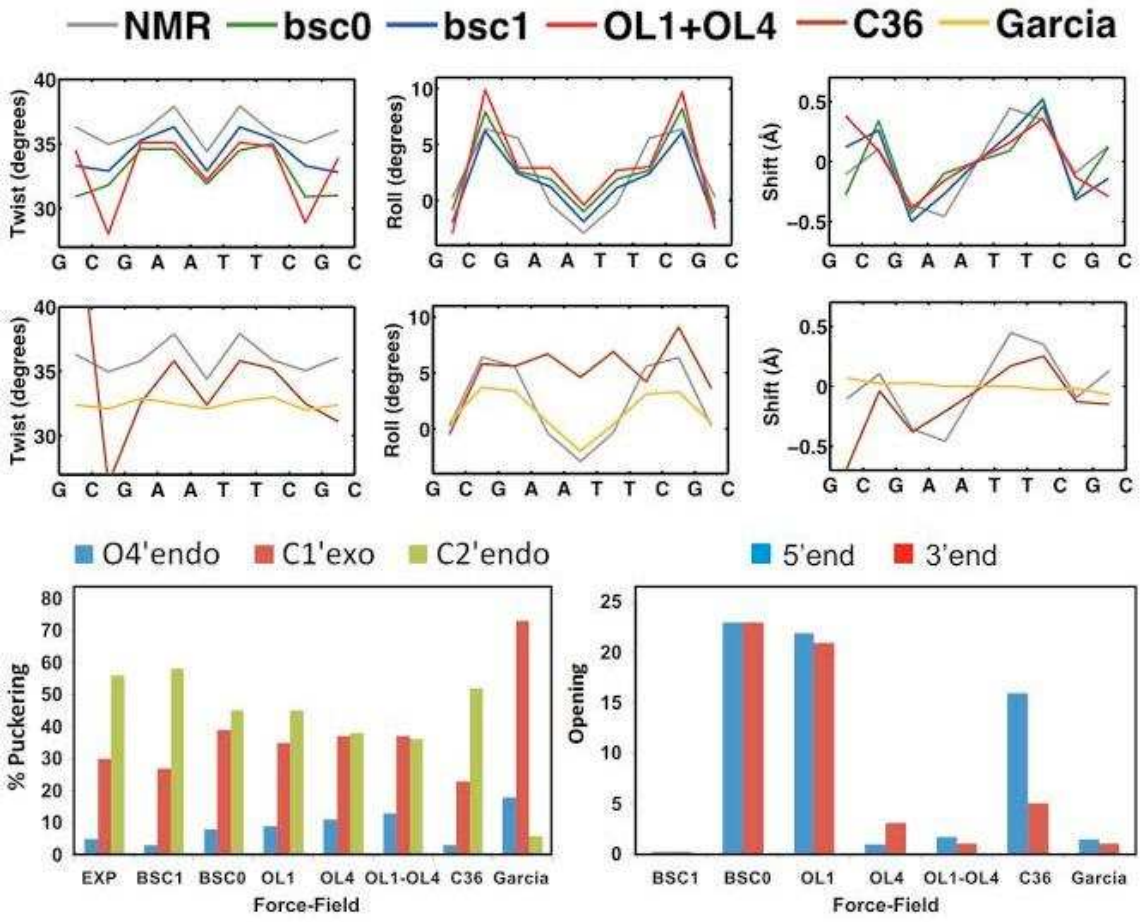
Supplementary Figure 26| Profiles of χ (chi) dihedral for 4 DNA bases in solution. Comparison of profiles obtained from QM using MP2/aug-cc-pVDZ (red) method with solvent corrections (Supplementary Notes), and PMF profiles using parmbsc0 (green) and parmbsc1 (blue) force-fields. Complete basis set (CBS) values for specific points are represented with a black dot.



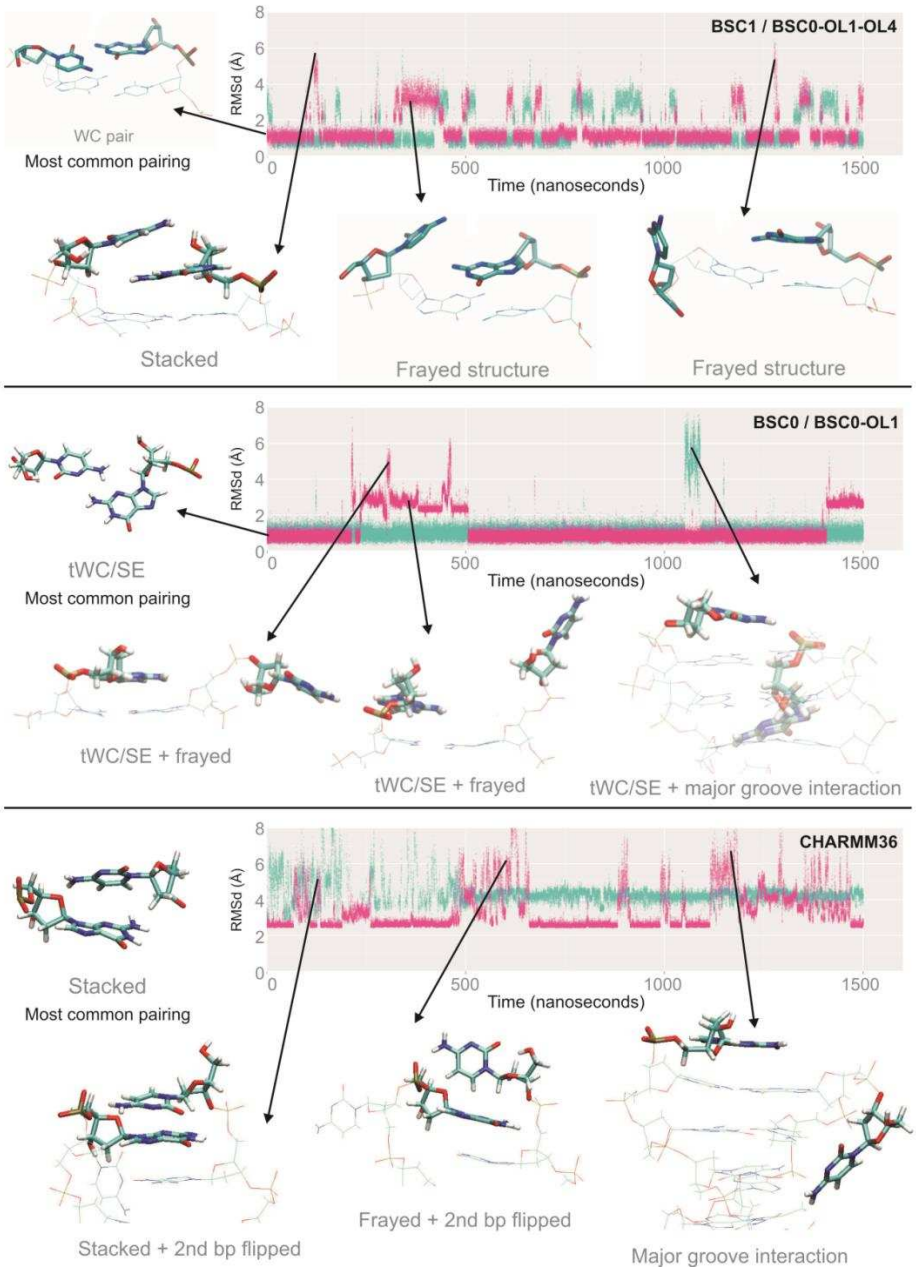
Supplementary Figure 27| Profiles of pseudorotational angle for 4 DNA bases in solution. Comparison of profiles obtained from QM using MP2/aug-cc-pVDZ (red) method with solvent corrections (Supplementary Notes), and PMF profiles using parmbsc0 (green) and parmbsc1 (blue) force-fields. Complete basis set (CBS) values for specific points are represented with a black dot.



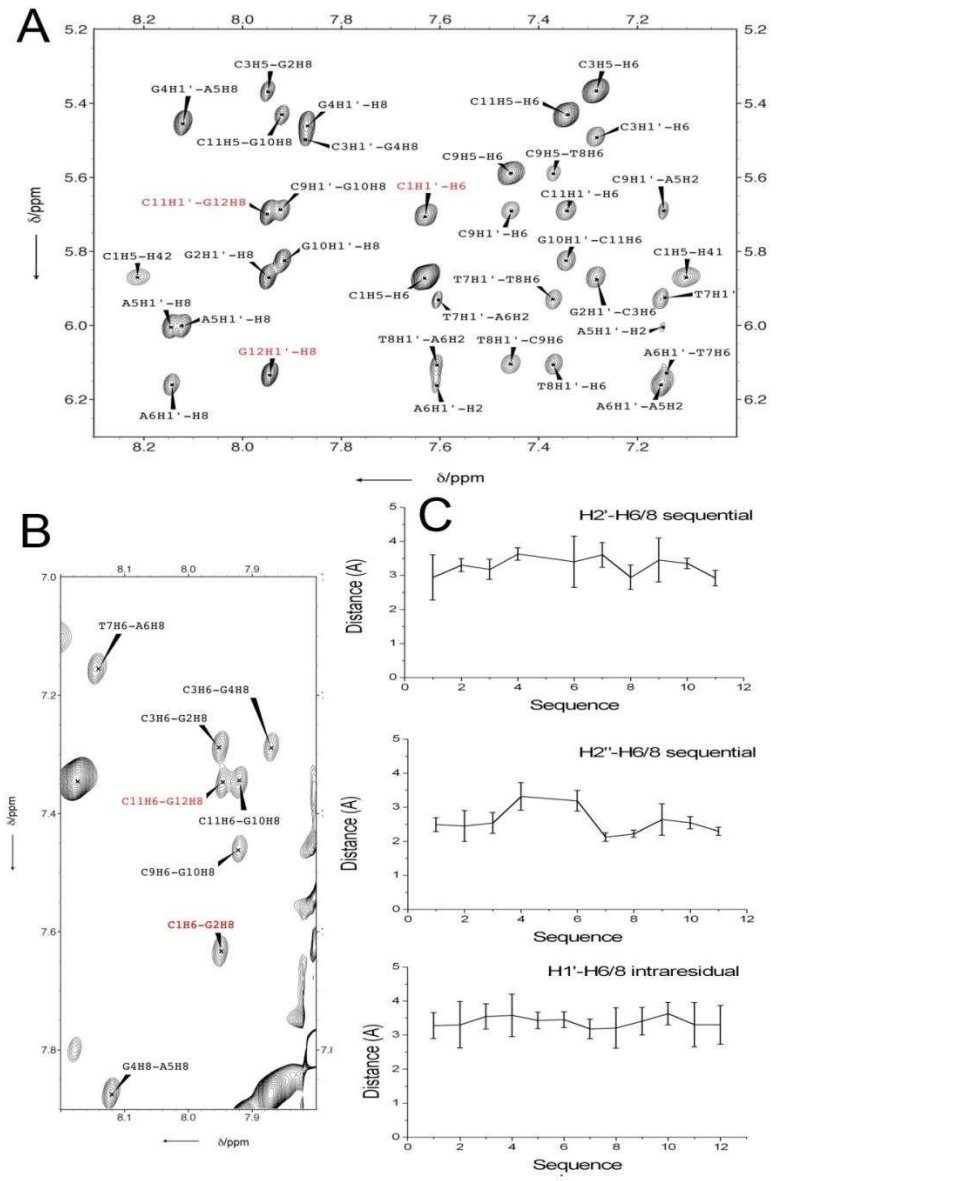
Supplementary Figure 28| ε/ζ (epsilon/zeta) profiles in solution. (**a**) Contour profiles of epsilon/zeta from QM calculations using MP2/aug-cc-pVDZ method (right), and PMF profiles using parmbsc0 (left) and parmbsc1 (middle) force-fields. Energies are given in kcal mol^{-1} and the color bar goes from blue (0 kcal mol^{-1}) to red (10 kcal mol^{-1}). (**b**) Values at key points of the profile comparing parmbsc0 (green), parmbsc1 (blue) and complete basis set (CBS) (dark red) values.



Supplementary Figure 29| Structural characteristics of DDD in MD simulations with different force-fields. First row variation of key helical coordinates along sequence in parmbsc0, parmbsc1 and parmbsc0-OL1+OL4 (those force-fields providing the best average parameters in **Supplementary Table 2**). Second raw correspond to force-fields providing less accurate average values in **Supplementary Table 2** (CHARMM36 and parmbsc0-Cheng-Garcia). In these two rows only the 10 mer segment is shown (to avoid dramatic scale bias in case of fraying of terminal bases), and only NMR results are used as reference (to make more clear the plots; note that nearly identical profiles are obtained from X-Ray (see **Fig. 1**)). The third row corresponds to the distribution of sugar puckering (taking as experimental reference the average of NMR and X-Ray structures) and the average opening at the terminal basis. The superior behavior of parmbsc1 is evident in all plots, as well the prevalence of fraying artifacts for some of the force-field, and the presence of non-negligible distortions in CHARMM36 and parmbsc0-CG trajectories, even for the central portion of the helix.



Supplementary Figure 30| Details of the evolution of the terminal base pairs. RMSd of the terminal base pairs (C1:G24 in pink and G12:C13 in cyan) along 1.5 μs of MD trajectories. Firs row: profiles for a force-field showing no fraying artifacts (but indeed frequent short-living openings) such a parmbsc1 (parmbsc0-OL1+OL4 and parmbsc0-OL4 provide similar profiles, while parmbsc0-CG (Cheng-Garcia) shows completely frozen terminal base pairs). Second row: profile for a force-field like parmbsc0 which suggest fraying and the formation of unusual contacts (parmbsc0-OL1 provides identical profiles) with tWC pairing and *syn* nucleotides. Third row: profiles obtained for CHARMM36, where despite the center of the duplex is well conserved terminal Watson-Crick pairings are mostly lost and substituted by a myriad of alternative contacts. In all cases structures sampled along specific time frames are shown.



Supplementary Figure 31| NOE data on the terminal base steps of DDD. A) H1´-aromatic region of the NOESY spectra of DDD (mixing time 200 ms, buffer conditions 125 mM NaCl, 25 mM sodium phosphate, pH 7, T = 25 º C). Some relevant cross-peaks involving terminal residues are labelled in red colour. B) Aromatic-aromatic region of the NOESY spectra (same experimental conditions). Note that NOE intensities involving terminal residues (i.e. C1H6-G2H8, C11H6-G12H8 in red) are not significantly lower than those involving central residues, indicating that the terminal bases remain stacked on top of their neighbours. C) Some experimental distances obtained from a full relaxation matrix analysis of the NOE data *vs* sequence. Sequential H2´-H6/8 and H2”-H6/8 do not exhibit dramatic changes for the terminal base steps, indicating that the fraying effect in these residues is not significant under these experimental conditions. All intra-residual H1´-H6/8 distances, including the terminal base residues, are around 3-4 Å, characteristic of glycosidic angle conformation in *anti*.