Supplementary Materials and Methods: DNA Sequence and Methylation Prescribe the Inside-Out Conformational Dynamics and Bending Energetics of DNA Minicircles

Authors: Jejoong Yoo[1*], Sangwoo Park[3], Christopher Maffeo[2], Taekjip Ha[3,4*], Aleksei Aksimentiev[2,5*]

Affiliations:

[1] Department of Physics, Sungkyunkwan University, Suwon 16419, Republic of Korea.

[2] Department of Physics and the Center for the Physics of Living Cells, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA.

[3] Department of Biophysics and Biophysical Chemistry, Johns Hopkins University, Baltimore, MD 21205

[4] Howard Hughes Medical Institute, Baltimore, MD 21218

[5] Beckman Institute for Advanced Science and Technology, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA.

*Correspondence to: Email: jejoong@skku.edu (J.Y.); aksiment@illinois.edu (A.A.); tjha@jhu.edu (J.H.)

Table of Contents

Supplementary Methods

Sources of genomic data

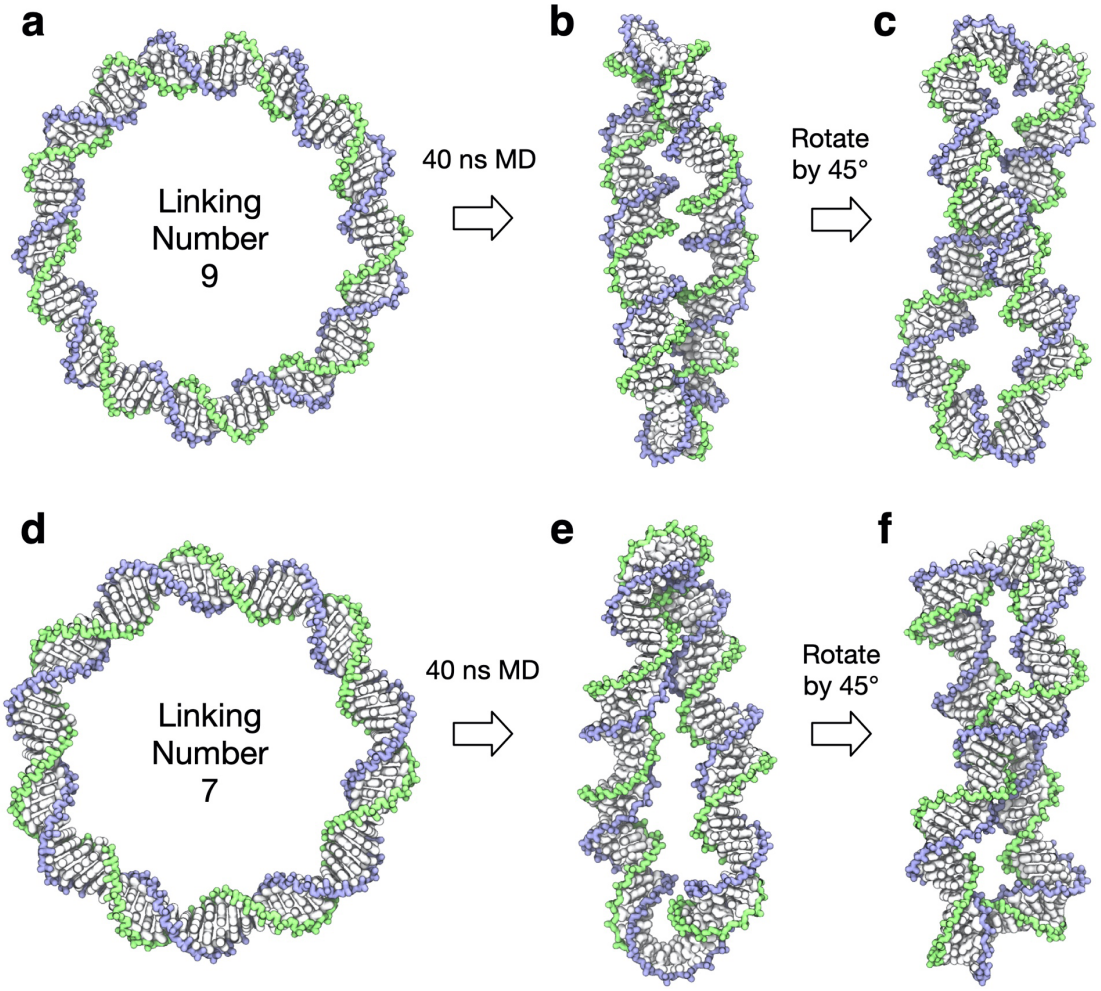
Data	Source
Nucleosome occupancy	Ref. (1)
Yeast genome and TSS	S288C_reference_genome_R64-2-1_20150113
Human genome, TSS, and CpG island data	GRCh38 assembly
DNA methylation for human B cell	ENCFF355UVU.bed from ENCODE data
DNA methylation for human HeLa cell	ENCFF601NBW.bed from ENCODE data
DNA methylation for human H1 cell	ENCFF064GJQ.bed from ENCODE data
DNA methylation for human HepG2 cell	ENCFF804NTQ.bed from ENCODE data
DamID	DataS1_Clone.14.1N.OE in GSE69423
Hi-C	GM12878 data in GSE63525

Supplementary Table 1. List of DNA minicircle simulations.

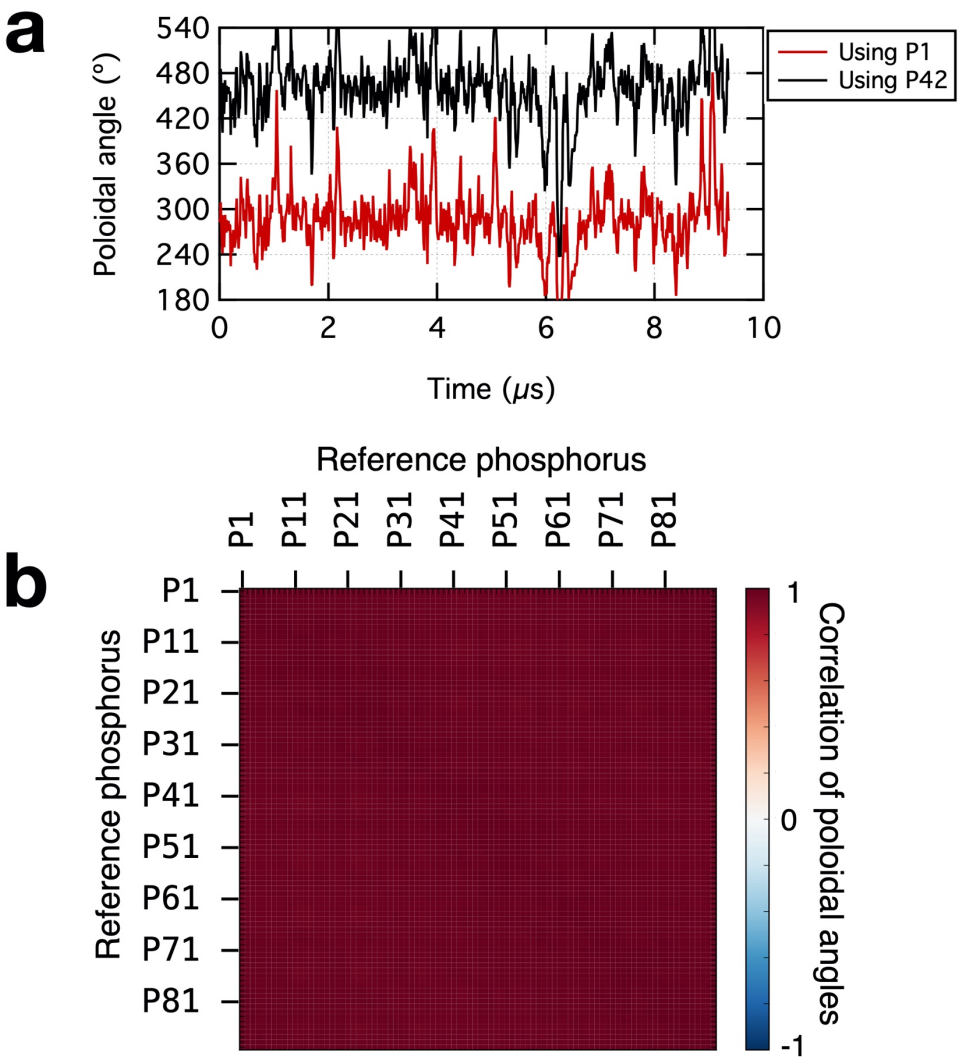
ID	Category	Name	DNA strand I's sequence (5' to 3') Note that "m" indicates methylcytosine	Simulation time (μs)	
				Equilibrium	Torque
1	NGO derivatives	NGO	cagaatccgtgctagtacctcaatatagactccctccggtgccga ggccgctcaattggtcgtaggactatcctcacctccaccgtttca	13.0	4.4
2		NGOme	cagaatccgtgctagtacctcaatatagactccctcmggtgcmga ggcmgctcaattggtmgtaggactatcctcacctccaccgtttca	9.5	3.4
3		NGOat	taaaatttatattaatattttaatataaatttttttaatattaa aattatttaattaattataaaattatttttattttattatttta	4.0	–
4		NGOgc	cggggcccgcgccggcgccccggcgcgggccccccccggcgccgg ggccgcccggccggccgcggggccgccccccgccccgccgccccg	4.0	–
5		NGOtcga	cagaatccgtgctagtacctcaatatagactcccttcgatgtcga ggtcgatcaattggtcgaaggactatcctcacctccaccgtttca	0.7	3.4
6		NGOtmga	cagaatccgtgctagtacctcaatatagactcccttmgatgtmga ggtmgatcaattggtmgaaggactatcctcacctccaccgtttca	0.8	–
7		NGO-shift	tccgtgctagtacctcaatatagactccctccggtgccgaggccg ctcaattggtcgtaggactatcctcacctccaccgtttcacagaa	0.4	3.3
8		NGOtcga-shift	tccgtgctagtacctcaatatagactcccttcgatgtcgaggtcg atcaattggtcgaaggactatcctcacctccaccgtttcacagaa	0.4	3.2
9		NGOtcga1	tccgtgctagtacctcaatatagactcccttcgatgccgaggccg ctcaattggtcgtaggactatcctcacctccaccgtttcacagaa	0.3	3.4
10		NGOtcga2	tccgtgctagtacctcaatatagactccctccggtgtcgaggccg ctcaattggtcgtaggactatcctcacctccaccgtttcacagaa	0.4	3.3
11		NGOtcga3	tccgtgctagtacctcaatatagactccctccggtgccgaggtcg atcaattggtcgtaggactatcctcacctccaccgtttcacagaa	0.3	3.4
12		NGOtcga4	cagaatccgtcagaataggaaccagtgctacgactcaacagatag cccattcgagctcgttgttgaggatggaccagtgctctttagtga	0.3	2.0
13	601 derivatives	601LH-90bp	ctggagaatcccggtgccgaggccgctcaattggtcgtagacagc tctagcaccgcttaaacgcacgtacgcgctgtcccccgcgtttta	0.4	3.1
14		601LHme-90bp	ctggagaatccmggtgcmgaggcmgctcaattggtmgtagacagc tctagcacmgcttaaamgcamgtamgmgctgtccccmgmgtttta	0.3	3.7
15		601RH-90bp	taaacgcacgtacgcgctgtcccccgcgttttaaccgccaagggg attactccctagtctccaggcacgtgtcagatatatacatcctgt	0.6	3.2
16		601RHme-90bp	taaamgcamgtamgmgctgtccccmgmgttttaacmgccaagggg attactccctagtctccaggcamgtgtcagatatatacatcctgt	0.9	3.1
17	TA repeats	AA(ta)$_4$t repeat	(aatatatatat)$_8$at	1.2	4.1
18		AT(ta)$_4$t repeat	(attatatatat)$_8$at	0.2	–
19		TT(ta)$_4$t repeat	(tttatatatat)$_8$at	0.4	–
20		CC(ta)$_4$t repeat	(cctatatatat)$_8$at	0.5	–
21		CG(ta)$_4$t repeat	(cgtatatatat)$_8$at	0.9	–
22		MG(ta)$_4$t repeat	(mgtatatatat)$_8$at	0.4	–
23		GC(ta)$_4$t repeat	(gctatatatat)$_8$at	0.8	–
24		GG(ta)$_4$t repeat	(ggtatatatat)$_8$at	0.4	–
25		TA repeat	(ta)$_{45}$	3.3	–
26	CG repeats	AA(cg)$_4$c repeat	(aacgcgcgcgc)$_8$gc	1.1	3.8
27		AT(cg)$_4$c repeat	(atcgcgcgcgc)$_8$gc	0.7	3.7
28		TA(cg)$_4$c repeat	(tacgcgcgcgc)$_8$gc	1.4	3.6
29		TT(cg)$_4$c repeat	(ttcgcgcgcgc)$_8$gc	0.3	2.8
30		GG(cg)$_4$c repeat	(ggcgcgcgcgc)$_8$gc	1.4	4.3

31		GC(cg)$_4$c repeat	(gccgcgcgcgc)$_8$gc	0.3	3.7
32		CC(cg)$_4$c repeat	(cccgcgcgcgc)$_8$gc	0.4	–
33	Tracts	G21 tract	(g)$_{21}$(cg)$_{34}$c	0.5	2.0
34		G9 tract	(g)$_9$(cg)$_{40}$c	0.8	–
35		A21 tract	(a)$_{21}$(ta)$_{34}$t	0.9	2.1
36		A9 tract	(a)$_9$(ta)$_{40}$t	0.5	–
37	CpG periodics	TCGA periodic	(cgctcgacgcg)$_8$cg	0.5	5.7
38		TMGA periodic	(cgctmgacgcg)$_8$cg	0.5	5.7
39		ACGT periodic	(cgcacgtcgcg)$_8$cg	0.5	3.4
40		AMGT periodic	(cgcamgtcgcg)$_8$cg	0.5	3.4
41		CCGG periodic	(cgcccggcgcg)$_8$cg	0.4	3.7
42		CMGG periodic	(cgccmggcgcg)$_8$cg	0.6	3.5
Total simulation time (μs)				55.7	100.5

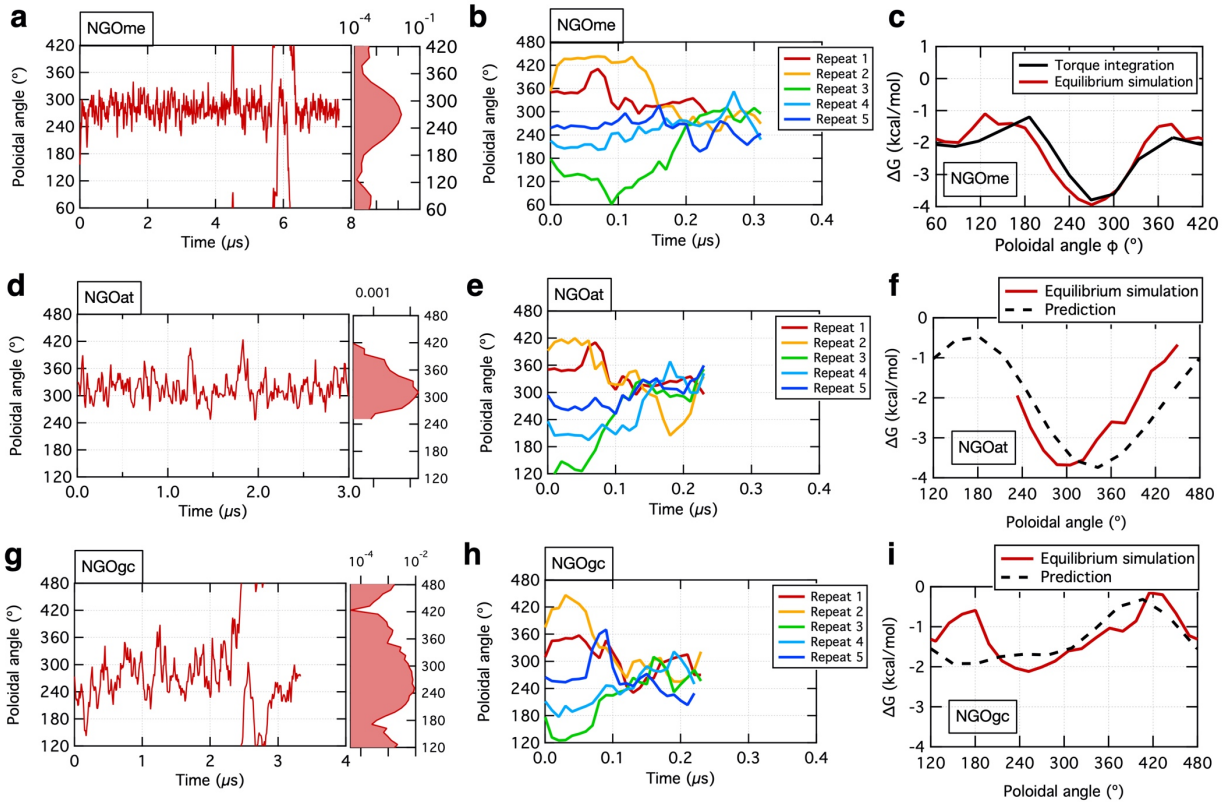
Supplementary Figures



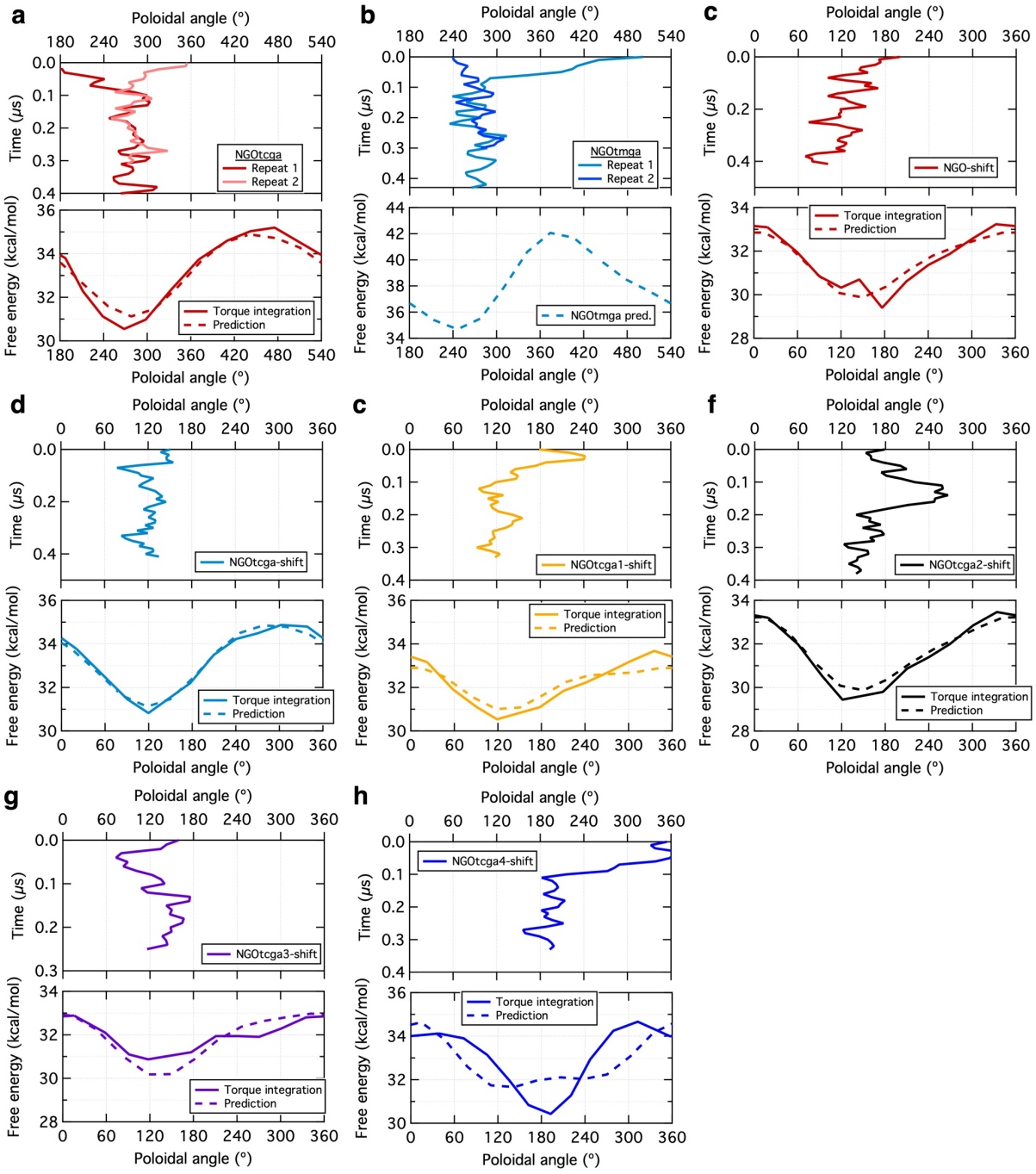
Supplementary Figure 1. Supercoiling of 90-bp DNA minicircles. a–f, Simulations of the NGO minicircle with a linking number of 9 (**a** to **c**) and 7 (**d** to **f**). Panel **a** and **d** show the initial conformations in an ideal circular geometry. Panels **b** and **c** show the supercoiled conformation from two different viewpoints for the linking number of 9. Panels **e** and **f** show the supercoiled conformations for the linking number of 7. Supplementary Movies 4 and 5 illustrate these 40 ns trajectories. During the 40-ns simulations, base pairs were protected as described in Supplementary Methods.



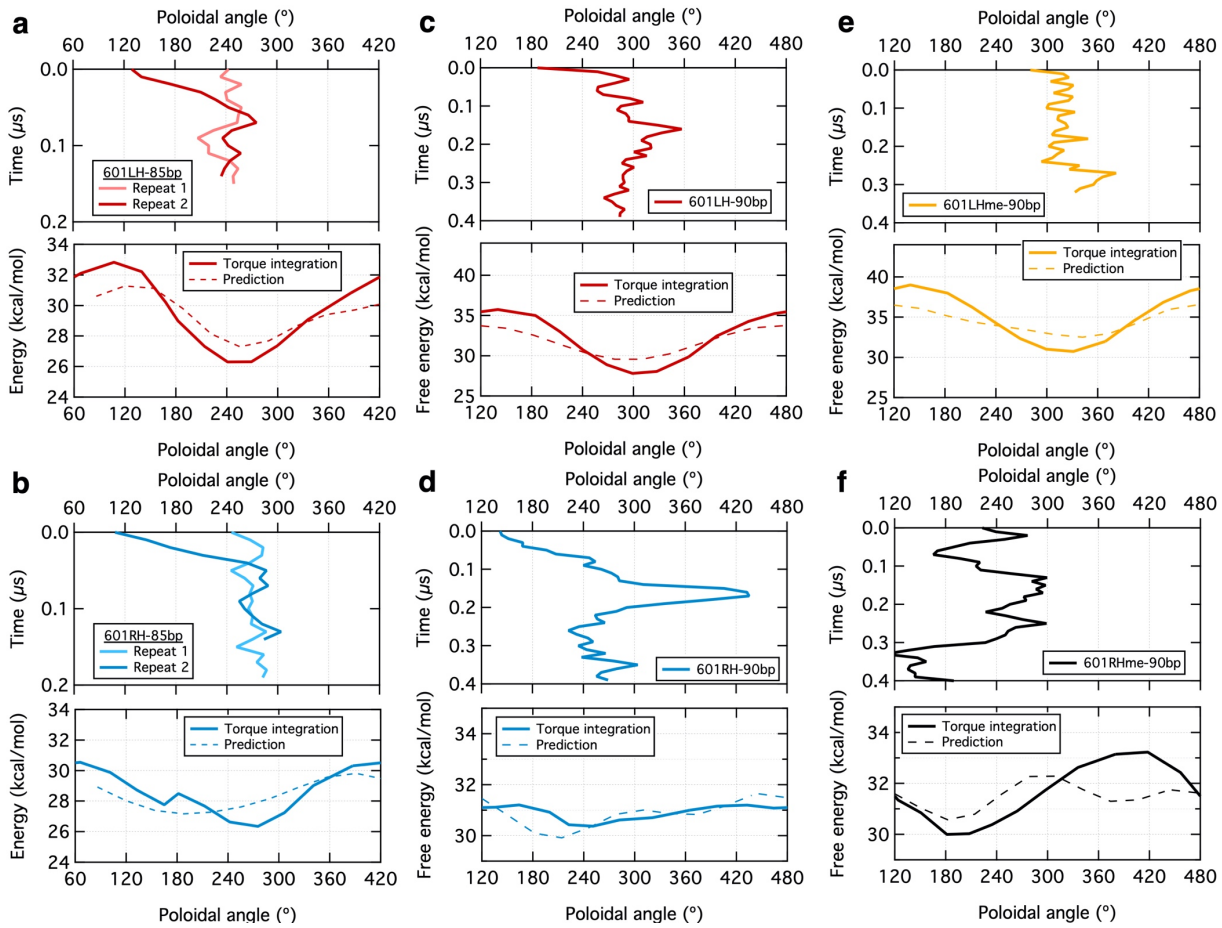
Supplementary Figure 2. Correlations of the poloidal angle in a DNA minicircle. a, Poloidal angle as a function of simulation time for two reference phosphorus atoms, P1 and P42 of strand I. Note the high temporal correlation of the two poloidal angle values; the Pearson correlation of the two plots is 0.97. **b**, Pearson correlations of all possible pairs of the 90 possible reference angles (P1 to P90 of strand I).



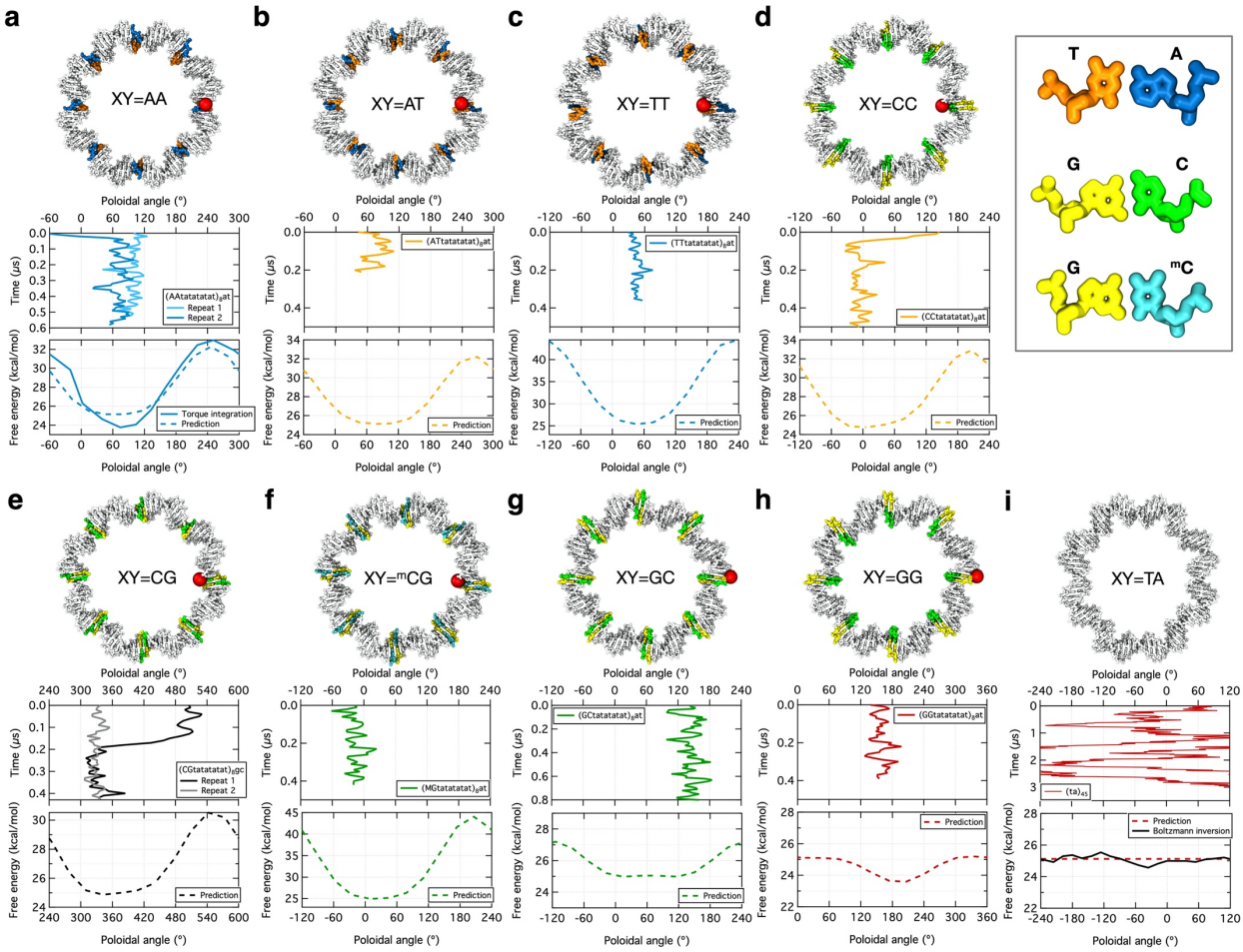
Supplementary Figure 3. Preferential poloidal angle of the NGOme, NGOat, and NGOgc minicircles. **a**, 10-ns block averages of ϕ in the 7.8-μs unrestrained simulation of the NGOme minicircle. RHS: Normalized probability density function in log scale, $\log P(\phi)$. **b,** Time evolutions of ϕ in five independent simulations of the NGOme minicircle starting from different initial ϕ value. **c,** Free energy of the NGO minicircle as a function of ϕ obtained by the torque integration (black) and Boltzmann inversion (red) methods. The later was computed as $-k_B T \log P(\phi)$ where $P(\phi)$ is the probability function for the trajectory shown in panel **a**. The free energy curves were vertically aligned to have approximately the same value at the free energy minimum. e. **d–f**, Same plots as **a–c**, respectively, for the NGOat minicircle. Note that the free energy landscape (black dashed line) here was obtained not by the torque integration but by the elastic model, Eq. 9. **g–i**, Same plots as **d–f**, respectively, for the NGOgc minicircle.



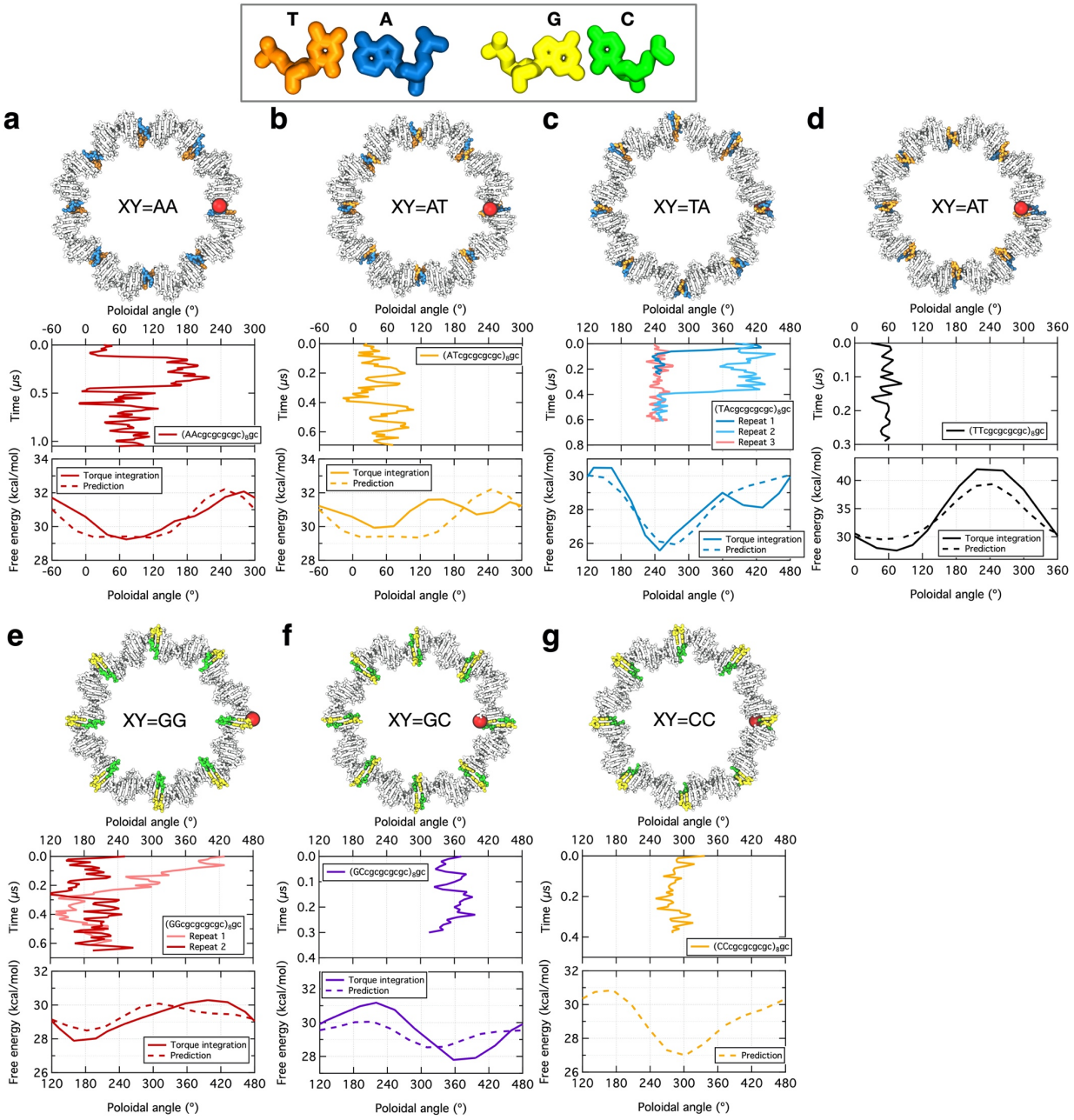
Supplementary Figure 4. Preferential poloidal angle of NGO-derivative DNA minicircles. a–h, Data for DNA minicircle ID 5 to 12, see Supplementary Table 1 for sequences. The top plot of each panel shows the values of the poloidal angle ϕ as a function of simulation time starting from a random initial value. The bottom of each panel shows the free energy of a minicircle as a function of ϕ obtained by the torque-integration method (solid line) and predicted by our elastic model, Eq. 9 (dashed line). Note that the free energy landscapes were vertically aligned to facilitate visual comparison.



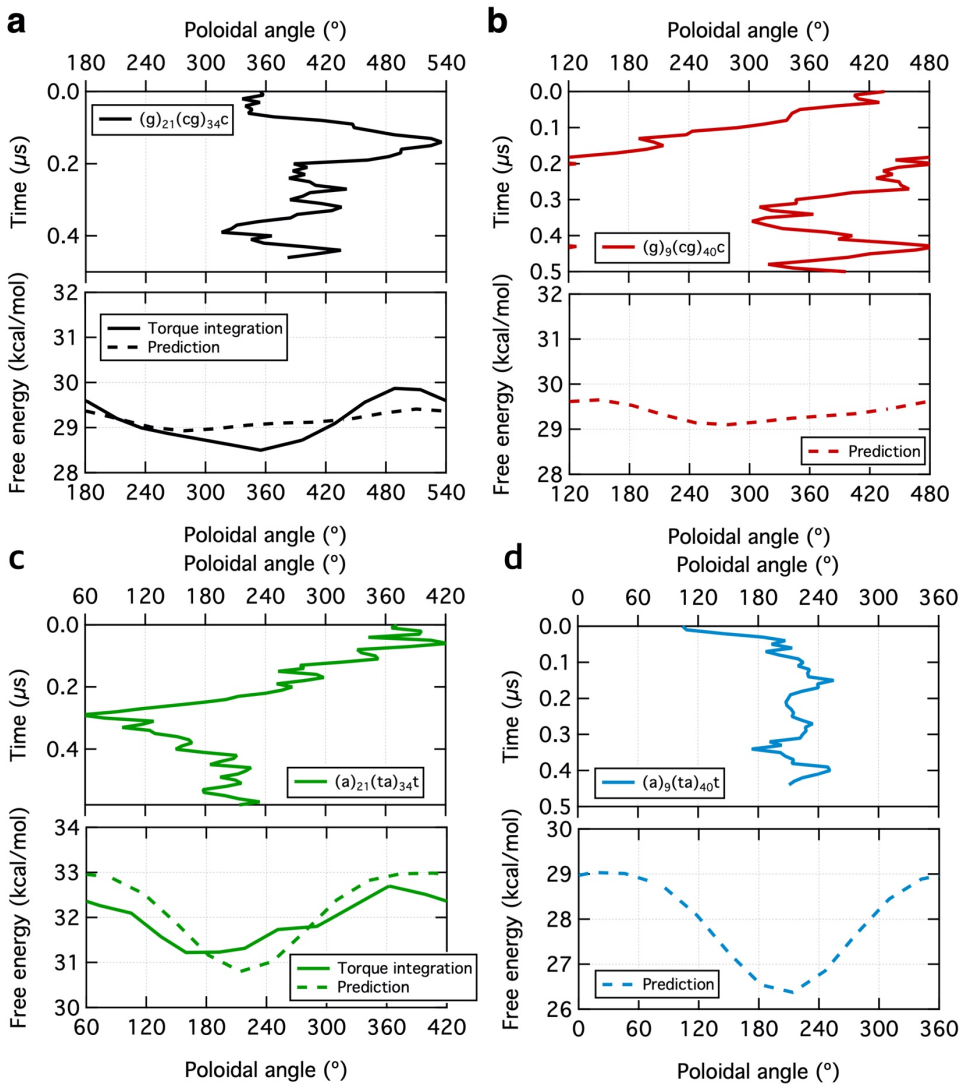
Supplementary Figure 5. Preferential poloidal angle of 601-derivative DNA minicircles. a,**b**, Data for 85-bp 601-derivative DNA minicircles in Fig. 1f and 1g, respectively. **c**–**f**, Data of 90-bp 601-derivative DNA minicircle ID 13 to 16, see Supplementary Table 1 for DNA sequences. The top plot of each panel shows the values of the poloidal angle ϕ as a function of simulation time starting from a random initial value. The bottom of each panel shows the free energy of a minicircle as a function of ϕ obtained by the torque-integration method (solid line) and predicted by our elastic model, Eq. 9 (dashed line). Note that the free energy landscapes were vertically aligned to facilitate visual comparison.



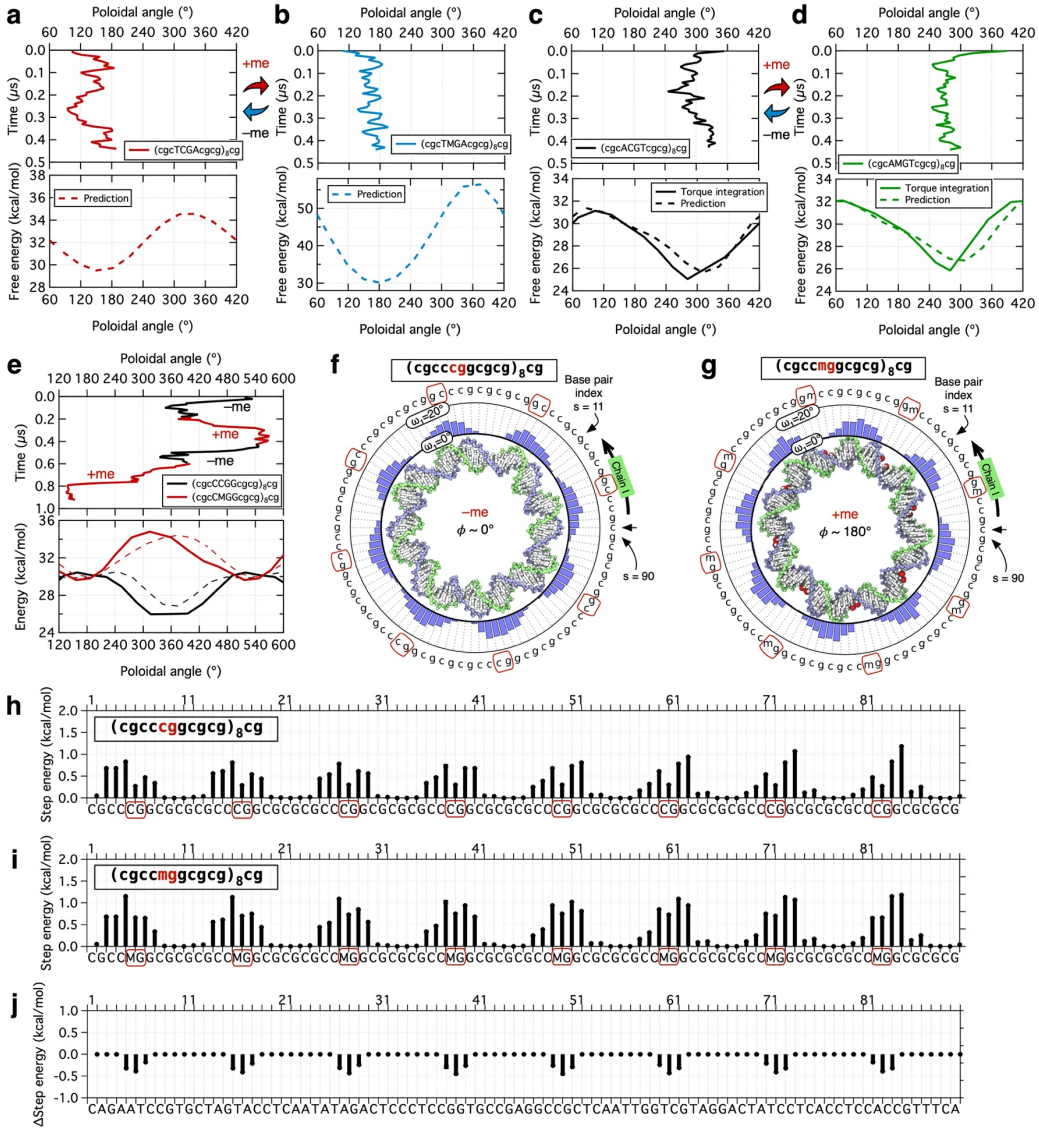
Supplementary Figure 6. Preferential poloidal angle of the TA-repeat minicircles, (XYtatatatat)$_8$at. **a**–**i**, Data for DNA minicircle ID 17 to 25, see Supplementary Table 1 for nucleotide sequences. The top image of each panel shows the equilibrium conformation of the minicircle, highlighting the location of the P1 atom using a red sphere. The type of X and Y nucleotides are color-coded as specified in the inset; all other nucleotides are shown in white. The middle plot of each panel shows the values of ϕ as a function of simulation time starting from a random initial value. The bottom of each panel shows the free energy of a minicircle as a function of ϕ obtained by the torque-integration method (solid line) and predicted by our elastic model, Eq. 9 (dashed line). Note that the free energy landscapes were vertically aligned to facilitate visual comparison. In panel i, the P1 atom is not shown because this particular sequence, (ta)$_{45}$, does not have any particular equilibrium poloidal angle.



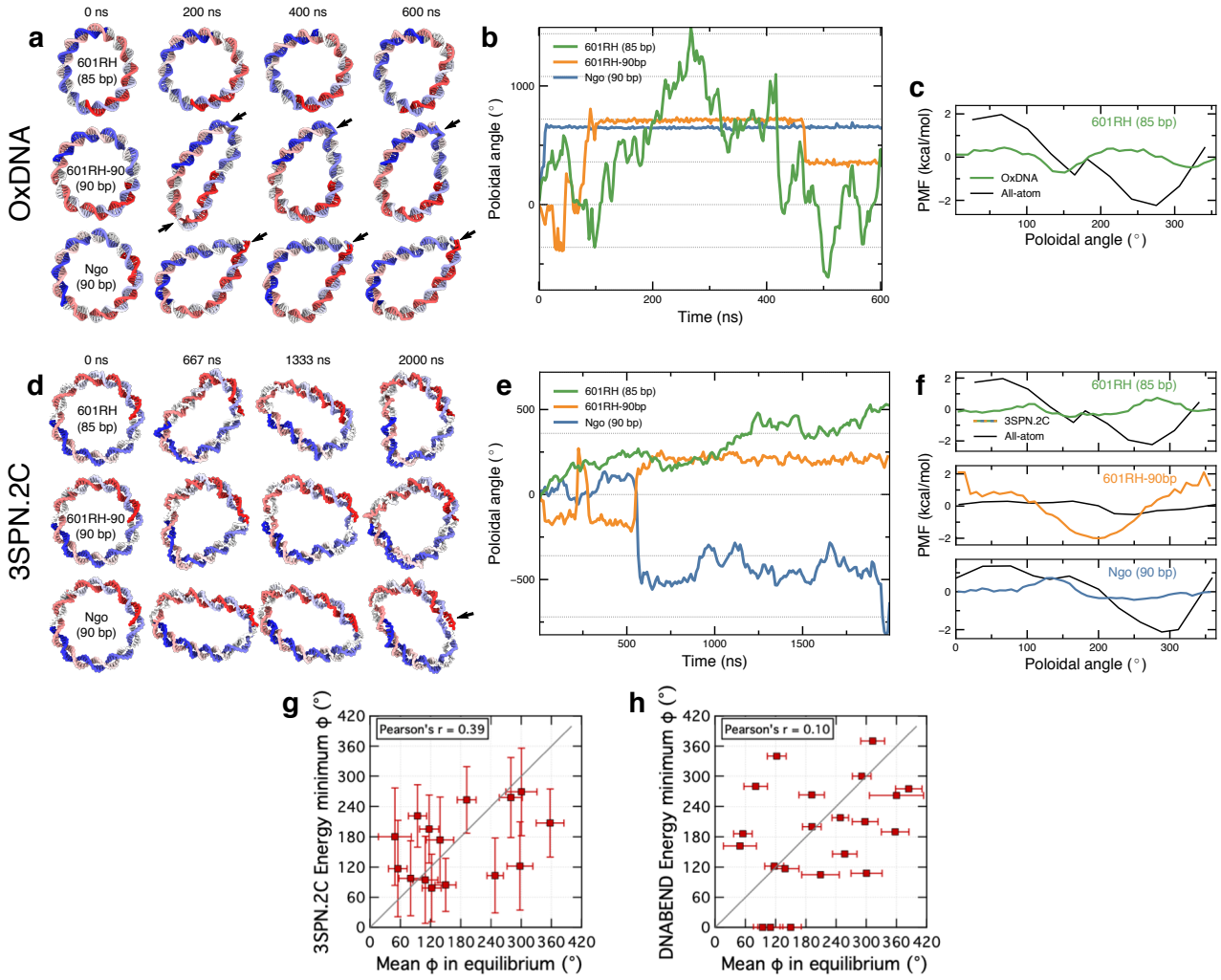
Supplementary Figure 7. Preferential poloidal angle of the CG-repeat minicircles, (XYcgcgcgcgc)$_8$gc. **a**–**g**, Data for DNA minicircle ID 26 to 32, see Supplementary Table 1 for nucleotide sequences. The top image of each panel shows the equilibrium conformation of the minicircle, highlighting the location of the P1 atom using a red sphere. The type of X and Y nucleotides are color-coded as specified in the inset; all other nucleotides are shown in white. The top plot of each panel shows the values of ϕ as a function of simulation time starting from a random initial value. The bottom of each panel shows the free energy of a minicircle as a function of ϕ obtained by the torque-integration method (solid line) and predicted by our elastic model, Eq. 9 (dashed line). Note that the free energy landscapes were vertically aligned to facilitate visual comparison.



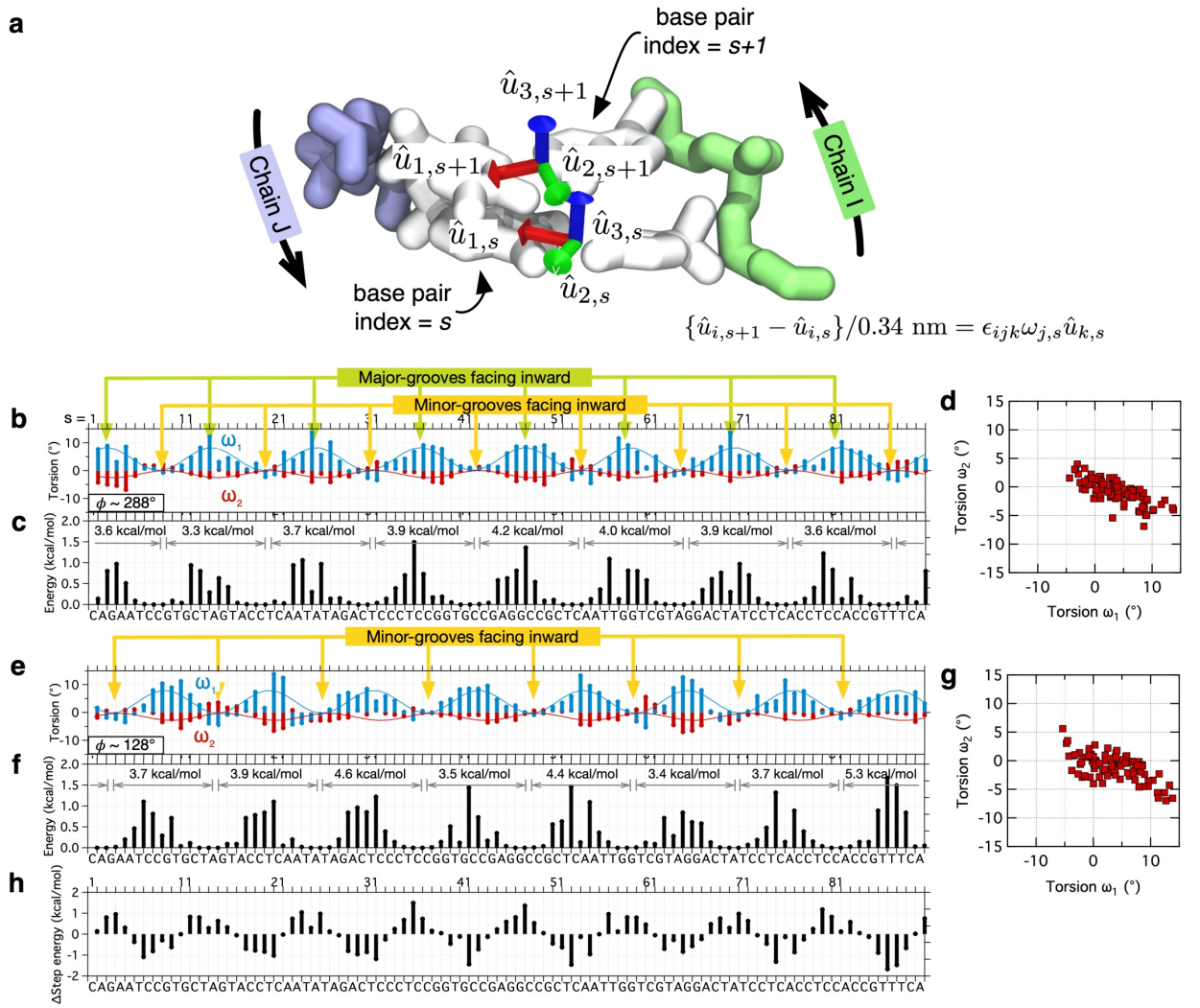
Supplementary Figure 8. Preferential poloidal angle for minicircles containing poly-(dA:dT) or poly-(dG:dC) tracts. a–d, Data for DNA minicircle ID 33 to 36, see Supplementary Table 1 for nucleotide sequences. The top plot of each panel shows the values of the poloidal angle ϕ as a function of simulation time starting from a random initial value. The bottom of each panel shows the free energy of a minicircle as a function of ϕ obtained by the torque-integration method (solid line) and predicted by our elastic model, Eq. 9 (dashed line). Note that the free energy landscapes were vertically aligned to facilitate visual comparison.



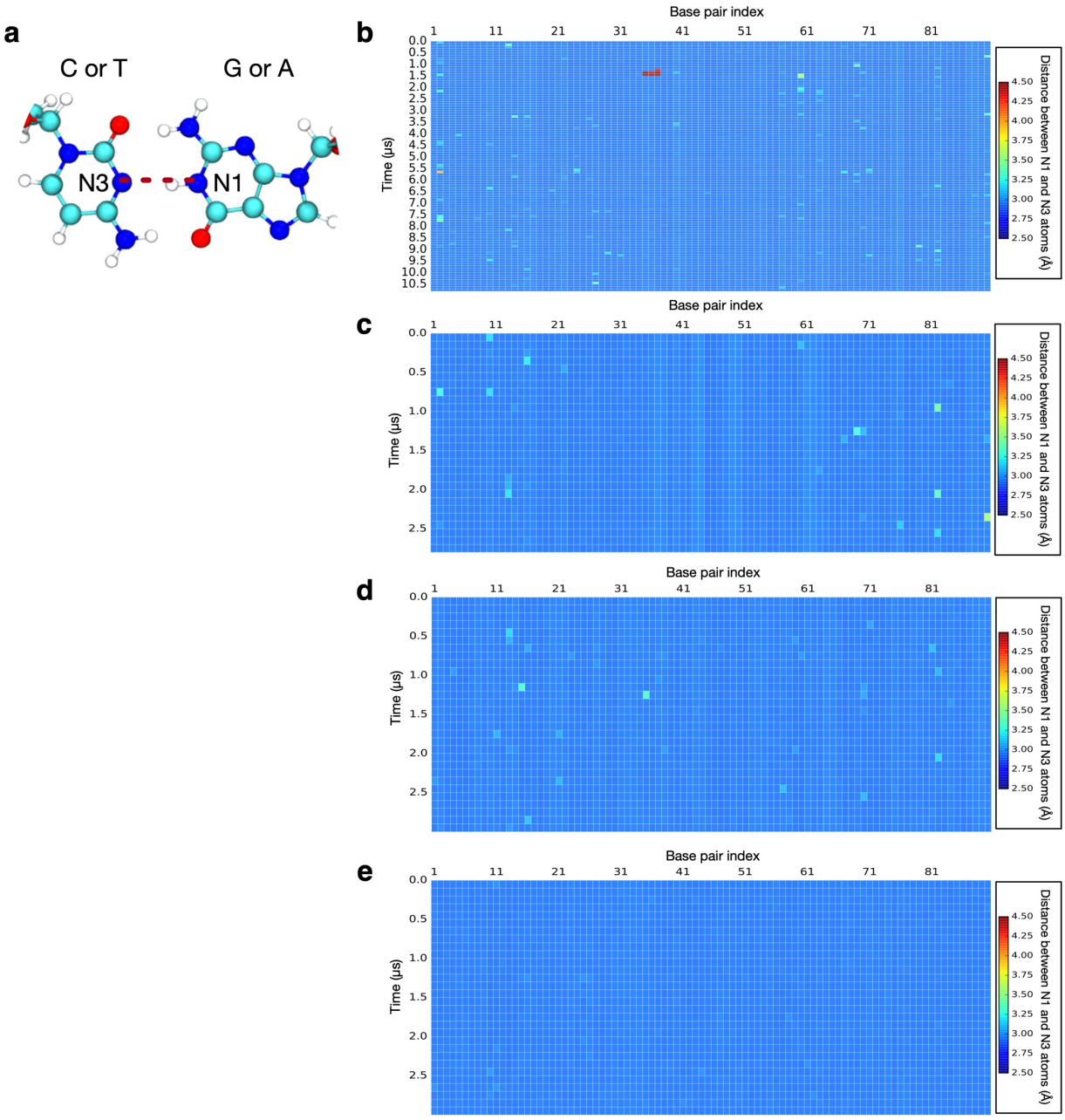
Supplementary Figure 9. Preferential poloidal angle of the minicircles with periodic CpG sites. a-d, Data for DNA minicircle ID 37 to 40, see Supplementary Table 1 for nucleotide sequences. The top plot of each panel shows the values of the poloidal angle ϕ as a function of simulation time starting from a random initial value. The bottom of each panel shows the free energy of a minicircle as a function of ϕ obtained by the torque-integration method (solid line) and predicted by our elastic model, Eq. 9 (dashed line). Note that the free energy landscapes were vertically aligned to facilitate visual comparison. **e**, (top) Free-equilibration trajectories illustrating the change of the poloid angle upon methylation of the (cgcc<u>cgg</u>cgcg)$_8$cg minicircle, sequence ID 41 and 42. (bottom) Methylation-dependent free energy landscapes computed by the torque-integration method (solid) and using the elastic model (dashed line). **f,g**, Energy minimum conformations of the minicircle without (panel f) and with (panel g) CpG methylation. Red spheres in panel g indicate the methyl groups. Bar graphs in the outside show the distributions of ω_1. **h,i**, Distribution of step energy values without (panel h) and with (panel i) CpG methylation at $\phi = 0°$. **j**, Difference in step energy values for the data shown in panels h and i.



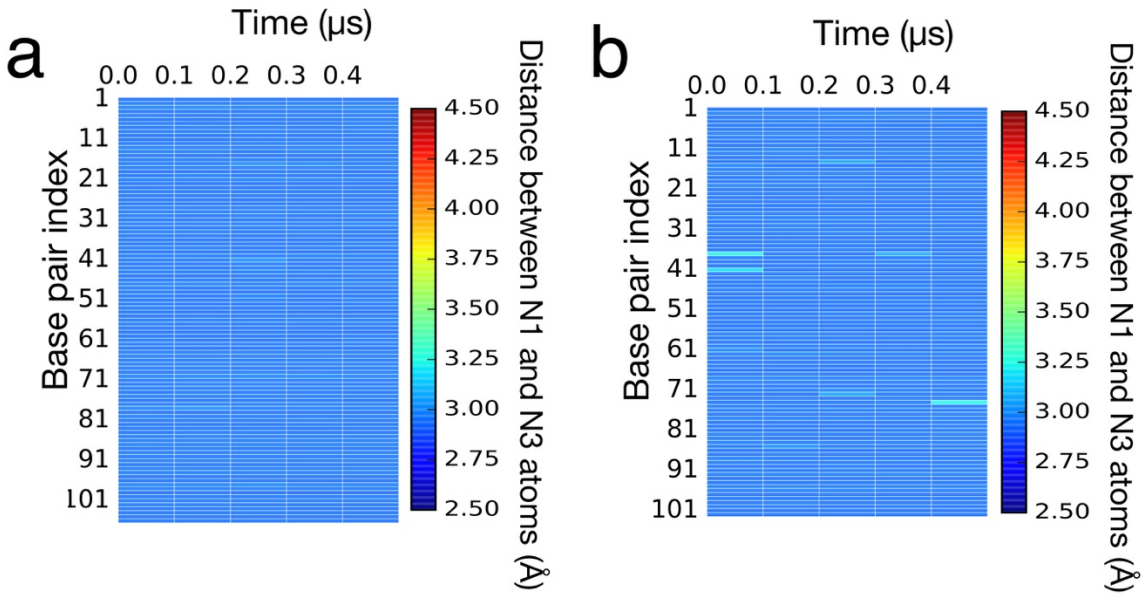
Supplementary Figure 10. Comparisons of coarse-grained models with all-atom MD results. a, Snapshots of select oxDNA minicircles depicted at regular intervals during 600-ns simulations. One strand is colored red to white, while the other strand is colored white to blue. Kinks in the B-DNA configuration are emphasized by small black arrows. **b**, Time series of the poloidal angle of the oxDNA minicircles shown in panel **a**. Dashed gray lines in the background are drawn every 360°. **c**, Comparison of the potential of mean force for the poloidal angle obtained by brute force sampling using oxDNA (green) and all-atom MD simulation (black) for the 85-bp 601RH sequence. **d**, Snapshots of select 3SPN.2C minicircles depicted at regular intervals during 2000-ns simulations. Partial unwinding of the minicircle in one snapshot is emphasized with a black arrow. **e**, Time series of the poloidal angle of the 3SPN.2C minicircles depicted in panel **d**. **f**, Comparison of the potential of mean force for the poloidal angle obtained by brute force sampling using 3SPN.2C (green, orange, blue) and all-atom MD simulation (black) for the sequences in panels **d** and **e**. **g**, Mean poloidal angles observed in unrestrained all-atom MD simulations and the 3SPN.2C coarse-grained simulations. **h**, Mean poloidal angles observed in unrestrained all-atom MD simulations and the minimum-energy poloidal angles of the DNABEND model. Error bars in panel **g** and **h** indicate the standard deviations.



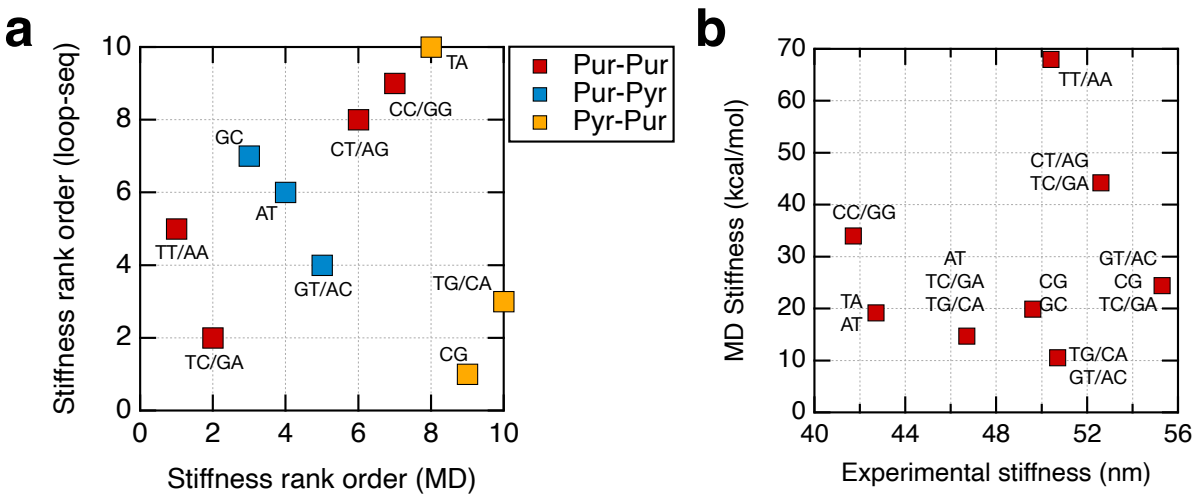
Supplementary Figure 11. Step-wise torsion angle and elastic energy in a minicircle. a, Local bending of a dinucleotide step of base pair s and $s+1$ described by the Frenet-Serrat formalism. For a base pair index s, we define an orthonormal triad of unit vectors $\hat{u}_{1,s}$, $\hat{u}_{2,s}$, and $\hat{u}_{3,s}$. Generalized torsions between base pairs s and $s+1$ are $\omega_{1,s}$, $\omega_{2,s}$, and $\omega_{3,s}$. Note that $\omega_{1,s}$ and $\omega_{2,s}$ denote two bending modes. **b,c,d,** Step-wise distributions of time-averaged $\omega_{1,s}$ and $\omega_{2,s}$ (**b**) and $\bar{E}_s$ (**c**) and the correlation between $\omega_{1,s}$ and $\omega_{2,s}$ (**d**) in the simulations of the NGO construct at $\phi = 288°$. **e,f,g,** Same plots as panels **b,c,d**, respectively, at $\phi = 128°$. **h**, Differences in step energy values for the data shown in panels **c** and **h**. Blue solid lines in panels b and e show $\omega_{1,s}^{\mathrm{circle}}(\phi)$ using Eq. 7.



Supplementary Figure 12. Stability of 90-bp NGO, NGOme, NGOat, and NGOgc minicircles. a, Red dashed line shows the N1–N3 distance. **b–e**, Changes of the N1–N3 distance during the simulation of the NGO (b), NGOme (c), NGOat (d), and NGOgc (e) minicircles. 100-ns block averages are shown as a heat map.



Supplementary Figure 13. Stability of 106-bp and 102-bp DNA minicircles. a,**b**, Changes in the hydrogen bond distances, defined as the distance between N1 and N3 atoms, during the simulation of the 106-bp (**a**) and 102-bp (**b**) minicircles of the sequence taken from Ref. (2). 100-ns block averages of the distance are shown as a heat map. To be consistent with Ref. (2), the simulations were run in 0.1 M NaCl at 300 K using the AMBER ff99bsc0 force field; the linking number of each minicircle was 9. The only difference from Ref. (2) was our custom equilibration strategy and the CUFIX correction for cation–phosphate interactions.



Supplementary Figure 14. Comparison of our D-WCL stiffness constants to experimentally derived values. a, Dinucleotide stiffness rank order extracted from the loop-seq experiment (3,4) versus the stiffness rank order extracted from the computational analysis of the DNA minicircles. Note that this comparison is qualitative (the lower the rank order, the stiffer the dinucleotide step) because of the differences in the underlying physical models and the lack of physical units in the experimental analysis. Red, blue, and yellow symbols indicate purine–purine (equivalently, pyrimidine–pyrimidine), purine–pyrimidine, and pyrimidine–purine dinucleotide steps, respectively. **b**, Dinucleotide stiffness extracted from our computational analysis of DNA minicircles versus the dinucleotide stiffness extracted from looping experiments of Geggier and Vologodskii (5). Note that the experimental values reflect stiffness constants averaged over two or three dinucleotide step types, as indicated by the text near each symbol.

Supplementary Movies

Supplementary Movie 1. The 150 ns trajectory of the 85 bp 601LH minicircle. Yellow spheres indicate the phosphate groups that are in contact with the histone core complex in the crystal structure.

Supplementary Movie 2. The 150 ns trajectory of the 85 bp 601RH minicircle. Yellow spheres indicate the phosphate groups that are in contact with the histone core complex in the crystal structure.

Supplementary Movie 3. The 11 microsecond trajectory of the 90 bp NGO minicircle with a linking number of 8. Red sphere indicates the phosphate that was used as a reference to measure the poloidal angle.

Supplementary Movie 4. The 40 ns trajectory of the 90 bp NGO minicircle with a linking number of 9.

Supplementary Movie 5. The 40 ns trajectory of the 90 bp NGO minicircle with a linking number of 7.

Supplementary Movie 6. An illustration of the minicircle that changes the poloidal angle from 0 to 360 degree. The poloidal angle is defined as the angle between two red lines: one that connects the center of the entire minicircle to the center of the DNA helix and the other that connects the center of the DNA helix to the reference phosphate.

Supplementary Movie 7. The 2 microsecond trajectory of the 90-bp TA-repeat minicircle with a linking number of 8. Red sphere indicates the phosphate that was used as a reference to measure the poloidal angle.

Supplementary Movie 8. The 200-ns trajectory of the $(cgccCGgcgcg)_8cg$ minicircle with a linking number of 8. Cytosine bases that are targets for methylation are shown in pink.

Supplementary Movie 9. The 200-ns trajectory of the $(cgcc^{m}CGgcgcg)_8cg$ minicircle with a linking number of 8. Methyl groups of methylcytosine bases are shown in red spheres.

Supplementary Movie 10. The 800-ns MD trajectory of the 85 bp 601LH minicircle using parmbsc1. One of two replica simulations.

Supplementary Movie 11. The 800-ns MD trajectory of the 85 bp 601LH minicircle using parmbsc1. The second replica simulation of the system shown in Movie S10.

Supplementary references

1. Kaplan, N., Moore, I.K., Fondufe-Mittendorf, Y., Gossett, A.J., Tillo, D., Field, Y., LeProust, E.M., Hughes, T.R., Lieb, J.D., Widom, J. *et al.* (2009) The DNA-encoded nucleosome organization of a eukaryotic genome. *Nature*, **458**, 362-366.
2. Sutthibutpong, T., Matek, C., Benham, C., Slade, G.G., Noy, A., Laughton, C., JP, K.D., Louis, A.A. and Harris, S.A. (2016) Long-range correlations in the mechanics of small DNA circles under topological stress revealed by multi-scale simulation. *Nucleic Acids Res*, **44**, 9121-9130.
3. Basu, A., Bobrovnikov, D.G., Cieza, B., Qureshi, Z. and Ha, T. (2020) Deciphering the mechanical code of genome and epigenome. *bioRxiv*, 2020.2008.2022.262352.
4. Basu, A., Bobrovnikov, D.G., Qureshi, Z., Kayikcioglu, T., Ngo, T.T.M., Ranjan, A., Eustermann, S., Cieza, B., Morgan, M.T., Hejna, M. *et al.* (2021) Measuring DNA mechanics on the genome scale. *Nature*, **589**, 462-467.
5. Geggier, S. and Vologodskii, A. (2010) Sequence dependence of DNA bending rigidity. *Proc Natl Acad Sci U S A*, **107**, 15421-15426.