PNAS
www.pnas.org

Supplementary Information for

Torsional stress can regulate the unwrapping of two outer half superhelical turns of nucleosomal DNA

Hisashi Ishida and Hidetoshi Kono

Corresponding Author: Hisashi Ishida

E-mail: ishida.hisashi@qst.go.jp

This PDF file includes:

Supplementary text

Figures S1 to S10

SI References

Supplementary Methods

Atomic model

We used the same atomic model used in the previous study (1), in which a crystal structure of a nucleosome with α-satellite palindromic 147 bp DNA (PDB code: 1kx5), was used but the histone tails were truncated to facilitate the unwrapping of the DNA. Consequently, residues 46 to 132 in H3, 25 to the C-terminal in H4, 16 to 114 in H2A, and 32 to 121 in H2B were used. As the nucleosome core particle contains two copies of each octamer and the DNA has two ends, we distinguish them with suffixes "a" and "b" for the octamer and "1" and "2" for the ends of the DNA. In this study, ~30 base pairs from each end of the DNA were taken into account. Nucleotides from −73 A to −39 A of chain I and nucleotides from 73 T to 39 T of chain J were assigned to DNA_1, and nucleotides from 73 T to 39 T of chain I and nucleotides from −73 A to −39 A of chain J were assigned to DNA_2 according to the nucleotide sequence in 1kx5. (see Fig. 1.)

Energy minimization

Energy minimization of the system was carried out to optimize the positions of the atoms and alleviate unfavorable interactions in the system. Simulations were carried out using an MD simulation program called SCUBA (2-8) with the AMBER *ff*14SB (9), bsc1 (10) and *ff*99ions08 (11) force-fields for the octamer, DNAs, and ions, respectively. Energy minimization was performed in a vacuum by assuming the distance-dependent dielectric constant of $4.0r$ with the value of r in Angstrom. Harmonic restraints with a force constant of 10 kcal/mol/Å^2 were applied to all the heavy atoms of the octamer and the DNA. Non-bonded interactions were evaluated with a cut-off radius of 12 Å. The system was minimized for 500 steps using steepest descent followed by 5,000 steps of

conjugate gradient.

MD simulations of the systems in water

After the energy minimization, the structure of nucleosome was placed in an aqueous medium. The system was placed in a rectangular box of ~125 Å × 245 Å × 155 Å, in which all the atoms of the nucleosome were separated more than 30 Å, 75 Å, and 25 Å from the edges of the box in the x, y, and z-directions (Fig. S1*B*). This box size was set large enough to observe the unwrapping of two half superhelical turns of nucleosomal DNA from either end of the DNA in the y-direction. ~148,000 TIP3P water molecules (12) were added to surround the system. In total, the system comprised ~465,000 atoms. To neutralize the charges of the system, sodium ions were placed in the box, and then additional sodium and chloride ions were added at a concentration of 150 mM NaCl. In total, the system comprised ~465,000 atoms.

The system was equilibrated at a constant pressure of one bar and a temperature of 300 K for 10 ns. The dielectric constant used was 1.0 and the van der Waals interactions were evaluated with a cut-off radius of 9 Å. The particle-particle particle-mesh (PPPM) method (13, 14) was used for the electrostatic interactions for the direct space cutoff of 9 Å. The Langevin dynamics algorithm was utilized to control the temperature and pressure of the system. The coupling times for the temperature and pressure control were both set at 2 ps^{-1}. The SHAKE algorithm (15, 16) was used to constrain all the bond lengths involving hydrogen atoms. The leap-frog algorithm with a time step of 2 fs was used throughout the simulation to integrate the equations of motion. The system was first heated from 0 K to 300 K within 1 ns during which the molecules and sodium ions were fixed with decreasing restraints and the water molecules were allowed

to move. After these restraints were removed, the system was equilibrated for 10 ns at a constant pressure of one bar and a temperature of 300 K with no restraint.

Generation of a series of conformations from wrapped to unwrapped DNA of the nucleosome using Adaptively Biased Molecular dynamics simulation

To generate a series of configurations of the unwrapping of the nucleosome from tightly wrapped to unwrapped, the adaptively biased molecular dynamics (ABMD) method (17) combined with the multiple walker method (18) was implemented in SCUBA. The equations of motion used in the ABMD method are expressed as (17):

$$\begin{aligned} m_a \frac{d^2\mathbf{r}_a}{dt^2} &= \mathbf{F}_a - \frac{\partial}{\partial \mathbf{r}_a} U\left[t \mid \sigma(\mathbf{R})\right], \\ \frac{\partial U(t \mid d)}{\partial t} &= \frac{k_B T}{\tau_F} K\left[d - \sigma(\mathbf{R})\right], \end{aligned} \tag{S1}$$

where $\mathbf{R} \equiv (\mathbf{r}_1, ..., \mathbf{r}_N)$ are the coordinates of the nucleosome core particle (NCP), N is the number of atoms in the NCP. d is the reaction coordinate, and $\sigma(\mathbf{R})$ is a function to give the value of the reaction coordinate. k_B is the Boltzmann constant, T is the constant temperature, τ_F is the flooding time scale, and K is the kernel which has distribution around the reaction coordinate. The first equation is for atom a, with an additional force coming from the biasing potential $U(t|d)$ with an ordinary atomic force of $\boldsymbol{F_a}$. The second equation is the time-evolving equation of the biasing potential. Details of the ABMD algorithm used in SCUBA are given in (5).

The ABMD simulation was carried out at a constant volume and a temperature of 300K with 6 walkers (replicas) of the nucleosome which were selected from the equilibration simulation. The reaction coordinate d was defined as the DNA near-end-to-end distance, or the distance between the center of mass of the 3rd base pair from each end of the DNA,

−71 C:G 71 in DNA_1, and 71 G:C −71 in DNA_2. The value of the reaction coordinate in the initial structure was 70.7 Å. The resolution of the reaction coordinate, Δd, in Eq. (1) was set at 1.0 Å. The relaxation time for the free-energy profile, τ, was set at 1,000 ps at the start and gradually decreased down to 100 ps until the simulation time of 15 ns at d = ~160 Å. Then τ, was set at ~200−800 ps. The structures were monitored approximately every 500 ps, and the replicas which did not show unwrapping of the DNA or which had abnormally deformed DNA at one end of the DNA were discarded. The discarded replicas were replaced by ones which showed unwrapping of the DNA and the simulation was started by assigning random velocities to the atoms and using a larger value of τ. This process was repeated until the unwrapping of the DNA had been completed. In the end, this process took ~30 ns. Then a series of 90 conformations at intervals of about 2.0 Å with d between 64 and 242 Å were selected from the conformational ensembles of the replicas. The distance between one end of DNA and the image of the periodically neighboring other end of the DNA was more than 22 Å at d = 242 Å, confirming that the ends of the DNA did not artificially interact with each other in the periodic boundary condition.

Umbrella sampling simulations under torsional stress and no stress

Using the final 90 conformations in the ABMD simulation, umbrella sampling simulations were carried out to obtain the free energy for the unwrapping of nucleosomal DNA with torsional stress and no stress. In the umbrella sampling, the reaction coordinate was divided into 90 windows with a width of 2 Å, which covered the entire range of d between 64 to 242 Å. The umbrella potential for each window used is a harmonic function with a force constant of 0.2 kcal/(mol $Å^2$). The weighted histogram

analysis method (WHAM) (19, 20) was used to refine the free-energy landscape from the sampled trajectories in the umbrella sampling simulations. With the WHAM approach, the unbiased probability distribution $P(\mathbf{R})$ is calculated from the biased probability distribution of the sampled coordinates as:

$$P(\mathbf{R}) = \sum_{i=1}^{N_{win}} n_i(\mathbf{R}) P_i^{(b)}(\mathbf{R}) \times \left[\sum_{j=1}^{N_{win}} n_j(\mathbf{R}) \exp\left(\left[F_j - V_j(\mathbf{R}) \right] / k_B T \right) \right]^{-1}, \quad \text{(S2)}$$

where $\mathbf{R}$ is the atomic coordinates, N_{win} is the number of windows, $n_i(\mathbf{R})$ is the number of data points in the i-th window, $P_i^{(b)}(\mathbf{R})$ is a biased probability from the raw data obtained in the umbrella sampling simulation, $V_j(\mathbf{R})$ is the biasing potential in the j-th window, k_B is the Boltzmann constant, and T is the constant temperature. In this study, $V_j(\mathbf{R})$ was selected to be a harmonic potential, which has the form:

$$V_i(\mathbf{R}) = k_i^{umb} \left(d(\mathbf{R}) - d_i^{\mathrm{fix}} \right)^2, \quad \text{(S3)}$$

where $d(\mathbf{R})$ is the distance between the centers of mass of the two terminal nucleotides of nucleosomal DNA. d_i^{fix} is a fixed distance to maintain $d(\mathbf{R})$ within the range of 64 Å to 242 Å with intervals of 2 Å (i = 1, ..., N_{win} = 90). k_i^{umb} is an arbitrary harmonic force constant, which was set at 0.2 kcal/mol/Å^2.

The umbrella sampling simulation was carried out for 100 ns. The trajectory for the last 50 ns was used for the analysis. The conformation of the nucleosome for the analysis was stored every 1 ps.

The coefficient F_j is defined by:

$$F_j = -k_B T \ln \left\{ \sum_{\text{windows}} P(\mathbf{R}) \exp\left(\left[-V_j(\mathbf{R}) \right] / k_B T \right) \right\}, \quad (j = 1, ..., N_{win}) \quad \text{(S4)}$$

where the summation includes all the coordinates of $\mathbf{R}$, which were sampled in any windows. By iterating Eqs. (S2) and (S4) to achieve self-consistency (using a tolerance

of 10^{-8}), the relative free energy $F(\mathbf{R})$ at a given $\mathbf{R}$ is obtained as:

$$F(\mathbf{R}) = -k_B T \ln P(\mathbf{R}). \quad \text{(S5)}$$

To visualize the free-energy profile, the dimension of $\mathbf{R}$ in Eq. (S5) was reduced to 1 or 2 dimensions by defining an appropriate coordinate (reaction coordinate). The reaction coordinate could be simply selected to be $d(\mathbf{R})$; however, the near-end-to-end distance of DNA, $d(\mathbf{R})$, may not be a good reaction coordinate, as it is not a direct indicator of how far the unwrapping of nucleosomal DNA has progressed from the initial state. Instead, the total number of base pairs unwrapped from the octamer would be a good indicator for interpreting the process of the unwrapping of the DNA from a structural viewpoint. Consequently, the first reaction coordinate, R_1, was defined as $d(\mathbf{R})$, and the second reaction coordinate, R_2, was defined as the total number of unwrapped base pairs. Unwrapping was defined when the center of mass of a base pair was shifted outward by more than 4 Å from the crystal structure. The probability of the trajectories on R_2, $P(R_2)$, can be written as:

$$P(R_2) = \int P(\mathbf{R}')\delta(R_2 - R_2'(\mathbf{R}'))d\mathbf{R}', \quad \text{(S6)}$$

where $\delta(X)$ is the Dirac delta-function, and $R_2' = R_2'(\mathbf{R})$. The free-energy profile in 1 dimension has the same form as Eq. (S5):

$$F(R_2) = -k_B T \ln P(R_2). \quad \text{(S7)}$$

To describe the changes in a physical quantity, A, such as the distance between atoms along R_2, the averaged quantity at R_2, $\bar{A}(R_2)$, is calculated by weighing the unbiased probability of the quantity $A(R)$ as:

$$\bar{A}(R_2) = \frac{\int A(\mathbf{R}')P(\mathbf{R}')\delta(R_2 - R_2'(\mathbf{R}'))d\mathbf{R}'}{P(R_2)}. \quad \text{(S8)}$$

The root mean square deviation (rmsd) from $\bar{A}(R_2)$ is

$$\sqrt{\sigma^2(R_2)} = \sqrt{\overline{A^2}(R_2) - \bar{A}(R_2)^2} \,. \tag{S9}$$

The probability with regard to R_1 and R_2, $P(R_1, R_2)$ is expressed as

$$P(R_1, R_2) = \int P(\mathbf{R}')\delta(R_1 - R_1'(\mathbf{R}'))\delta(R_2 - R_2'(\mathbf{R}'))d\mathbf{R}' \,. \tag{S10}$$

Then the relationship between R_1 and R_2 is expressed as

$$\overline{R_2}(R_1) = \int R_2 \cdot P(R_1, R_2)dR_2 \,. \tag{S11}$$

Torsional stress in the umbrella sampling simulations

In this study, in addition to the umbrella sampling simulation without torsional stress, umbrella sampling simulations under positive and negative torsional stress were carried out. The effect of the torsional stress on the DNA was incorporated as follows in a similar way to the procedure used by (21) (22) (see also Fig. 1*B* and *C*);

$$V_i^{stress} = k_{1i}\left(\phi_1 - \phi_1^0\right)^2 + k_{2i}\left(\phi_2 - \phi_2^0\right)^2 \tag{S12}$$

for the i-th window (i=1, …, N_{win} =90). The torsional angles of ϕ_1 is calculated using the positions of two pairs of atoms. The first pair is the C1' atom of −71 C (chain I, at the third base-pair from the end of DNA_1) and the C1' atom of 71 G (chain J, at the third base-pair from the end of DNA_1) and the other pair is the phosphorus atom of −39 A (chain I) and the phosphorus atom of −39 A (chain J) located opposite the dyad. The axis of torsion was defined as the line passing through the two midpoints between the two C1' atoms and between the two phosphorus atoms. Then, the torsional angle of ϕ_1 is defined as the angle between a plane which includes the two C1' atoms and the axis, and another plane which includes the two phosphorus atoms and the axis.

The torsional angle of ϕ_2 was calculated using the positions of two pairs of atoms.

The first pair is the C1' atom of 71 G (chain I, at the third base-pair from the end of DNA_2) and the C1' atom of −71 C (chain J, at the third base-pair from the end of DNA_2), and the other pair is the phosphorus atom of −39 A (chain I) and the phosphorus atom of −39 A (chain J) located opposite the dyad. The torsional angle of ϕ_2 is defined in a similar way to the torsional angle of ϕ_1.

The torsional restraint on the C1' atoms of the third base-pair from the end of the DNA is to avoid the potential problem that imposing torsional force on the first or second base-pairs may break the weaker base pairs and the torsional force may not be applied properly. Both ϕ_1^0 and ϕ_2^0 were set at 0° and 180° for the generation of positive and negative stress, respectively. As appropriate values of the force constants k_{1i} and k_{2i} to produce torque within 10 to 20 pN·nm were not known in advance, the values of k_{1i} and k_{2i} were adjusted for the first ~25 ns. Since then both of the force constants k_{1i} and k_{2i} were set at $2.0 \times 10^{-2} \times (i - 1) + 1.5$ kcal/(mol rad^2) for the positive stress, and at $1.5 \times 10^{-2} \times (i - 1) + 0.50$ kcal/(mol rad^2) for the negative stress. The torsional angles of ϕ_1 and ϕ_2, the torque applied to each C1' atom, and the base-pair step parameter of *Twist* in DNA_1 and DNA_2 against d^{fix} during the last 50 ns of the umbrella sampling simulations are shown in Fig. 1*D*, *E*, and *F*, respectively. It should be noted that the torsional restraint is not included in the WHAM calculation. The change in the free energy in this study is not from no-stress to positive or negative stress but from wrapped to unwrapped states under torsional stress.

The standard of the mean (SEM) in the free-energies

The values of the free energy at *d* were estimated using sufficient sampling data (more than 1% of 50,000 data points per window = 250 data points per 1 Å). The standard of

the mean (SEM) in the free-energies were determined by calculating the free energies for five 10-ns non-overlapping segments from 50 to 100 ns in the umbrella sampling simulations. Each free energy which was calculated from a 10 ns trajectory was aligned to the final free energy which was calculated from the 50 ns trajectory so that the average of the free energy coincides with the average of the final free energy. To calculate the average of the free energy, the values of the free energy at *d* which were estimated using sufficient sampling data (more than 1% of 10,000 data points per window = 50 data points per 1 Å) for each segment in the window were used. The standard deviation for each system was calculated using the equation, $\sigma = \sqrt{\sum_{k=1}^{n}\left(f_k(d) - \bar{f} \right)^2 / (n-1)}$, where $f_k(d)$ is the value of the aligned free-energy at *d* and $\bar{f}$ is the average and *n* is the number of segments at *d* for each system (n is almost always 5 except at *d* where sampling data is scarce). The SEM was calculated using the equation $\sigma / \sqrt{n}$.

Conformational entropy of the DNA

The conformational entropies of the two ends of $DNA_{1\ or\ 2}$ were calculated using the quasiharmonic approximation (23) as follows:

$$S_{conf} = 0.5\ k_B \ln \det [\mathbf{1}+(k_B T e^2/\hbar^2)\boldsymbol{\sigma}], \qquad \text{(S13)}$$

where *e* is *Euler*'s number, $\hbar$ is Planck's constant divided by 2π. $\boldsymbol{\sigma} = < \boldsymbol{x}\, \boldsymbol{x}^{\mathrm{T}} >$ represents the mass-weighted covariance matrix, where $\boldsymbol{x}$ is the coordinates of the first to 35th base pairs from the end of $DNA_{1\ or\ 2}$ (2226 atoms).

For the calculation of each covariance matrix for $DNA_{1\ or\ 2}$, $\boldsymbol{x}$ was best-fit in the reference coordinates. Each of the coordinates of $DNA_{1\ or\ 2}$ in the initial structure in the *i*-th window of the umbrella sampling simulation was used as the reference coordinates

for the best-fit at d_i^{fix}. The trajectories in the *i*-th window of the umbrella sampling simulation where the desired position of d_i^{fix} was set in Eq. (S3) were used for the calculation of the conformational entropies against d^{fix}. Hereafter, d_i^{fix} is referred to as d^{fix} unless it is specifically mentioned otherwise. It should be noted that the conformational entropies calculated by the quasiharmonic approximation can be markedly overestimated because of the anharmonicity in protein dynamics (5).

d^{fix}, not d, was used to avoid further overestimation of the conformational entropy. This happens when the asymmetric unwrapping of the DNA in one direction ($bp_1 >> bp_2$) and the asymmetric unwrapping of the DNA in the other direction ($bp_1 << bp_2$) occurs at the same value of d because both of these conformations are sampled at the same value of d. In this case, the quasi-harmonic approximation treats the two distinct distributions as one distribution by a broad energy potential, which leads to overestimation of the conformational entropy. The conformational entropy was also calculated against the average of bp in each window. The average of bp, not the individual value of bp, was used to keep the number of sampled conformations the same for all d^{fix}.

Free energies for the small- and large-scale asymmetric unwrapping of the DNA

Considering different unwrapping paths, we can estimate the contribution of the DNA – H3 and DNA – H2A-H2B interactions in changing the free energy (Fig. 4).

From the free energy for the small-scale asymmetric unwrapping of the DNA (Fig. 4*B*), the free energy for unwrapping one end of the DNA from H3, $\Delta G_{\mathrm{H3\_first}}^{\mathrm{small}}$, can be estimated from the increase in the free energy from $bp_{\mathrm{total}} = 0$ to 13 (roughly corresponding to the state from $S0$ to $sAS1$). Then $\Delta G_{\mathrm{H3\_first}}^{\mathrm{small}}$ was estimated to be 3.6, 1.9, and 2.4 under

positive, negative, and no stress, respectively. This indicates that *sAS*1 is in a higher free energy state under positive stress than under negative and no stress because breaking the stable interaction between the twisted minor groove of the DNA at SHL = ± 6.5 and the H3 αN-helix requires a high free energy (see. Fig. S3*D*).

Similarly from the free energy for the small-scale asymmetric unwrapping of the DNA, the free energy for unwrapping the DNA on the other side from H3, $\Delta G^{small}_{H3\_second}$, can be estimated from the increase in the free energy from bp_{total} = 13 to 26 (roughly corresponding to the state from *sAS*1 to *S*2). Then $\Delta G^{small}_{H3\_second}$ was estimated to be 3.6, 3.0, and 2.9 kcal/mol under positive, negative, and no stress, respectively. The high free energy of 3.6 kcal/mol under positive stress indicates that the DNA – H3 αN-helix interaction on the other side was also stable under positive stress (see. Fig. S3*E*).

The free energies for the subsequent unwrapping of the $DNA_{1\ or\ 2}$ from H2A-H2B, $\Delta G^{small}_{H2A/H2B\_first}$ and $\Delta G^{small}_{H2A/H2B\_second}$, were estimated from the increase in the free energy from bp_{total} = 26 to 39 (roughly corresponding to the state from *S*2 to *sAS*3) and bp_{total} = 39 to 52 (roughly corresponding to the state from *sAS*3 to *S*4), respectively. $\Delta G^{small}_{H2A\text{-}H2B\_first}$ was estimated to be 1.7, 12.5, and 5.3 under positive, negative, and no stress, while $\Delta G^{small}_{H2A\text{-}H2B\_second}$ was estimated to be 3.5, 1.9, and 4.6 under positive, negative, and no stress, respectively.

From the free energy for the large-scale asymmetric unwrapping of the DNA (Fig 4*C*), the free energy for continuously unwrapping the DNA from H2A-H2B, $\Delta G^{large}_{H2A\text{-}H2B\_first}$ was estimated from the maximum increase in the free energy from bp_{total} = 13 to 32 (roughly corresponding to the state from *sAS*1 to *lAS*2). Then $\Delta G^{large}_{H2A\text{-}H2B\_first}$ was

estimated to be 5.6, ND, and 12.1 kcal/mol under positive, negative, and no stress, respectively.

$\Delta G^{large}_{H3\_second}$ was estimated from the difference in the free energies at the state of *lAS*2 and *sAS*3. Then $\Delta G^{large}_{H3\_second}$ was estimated to be -0.3, no data, and -4.0 kcal/mol under positive, negative, and no stress, respectively.

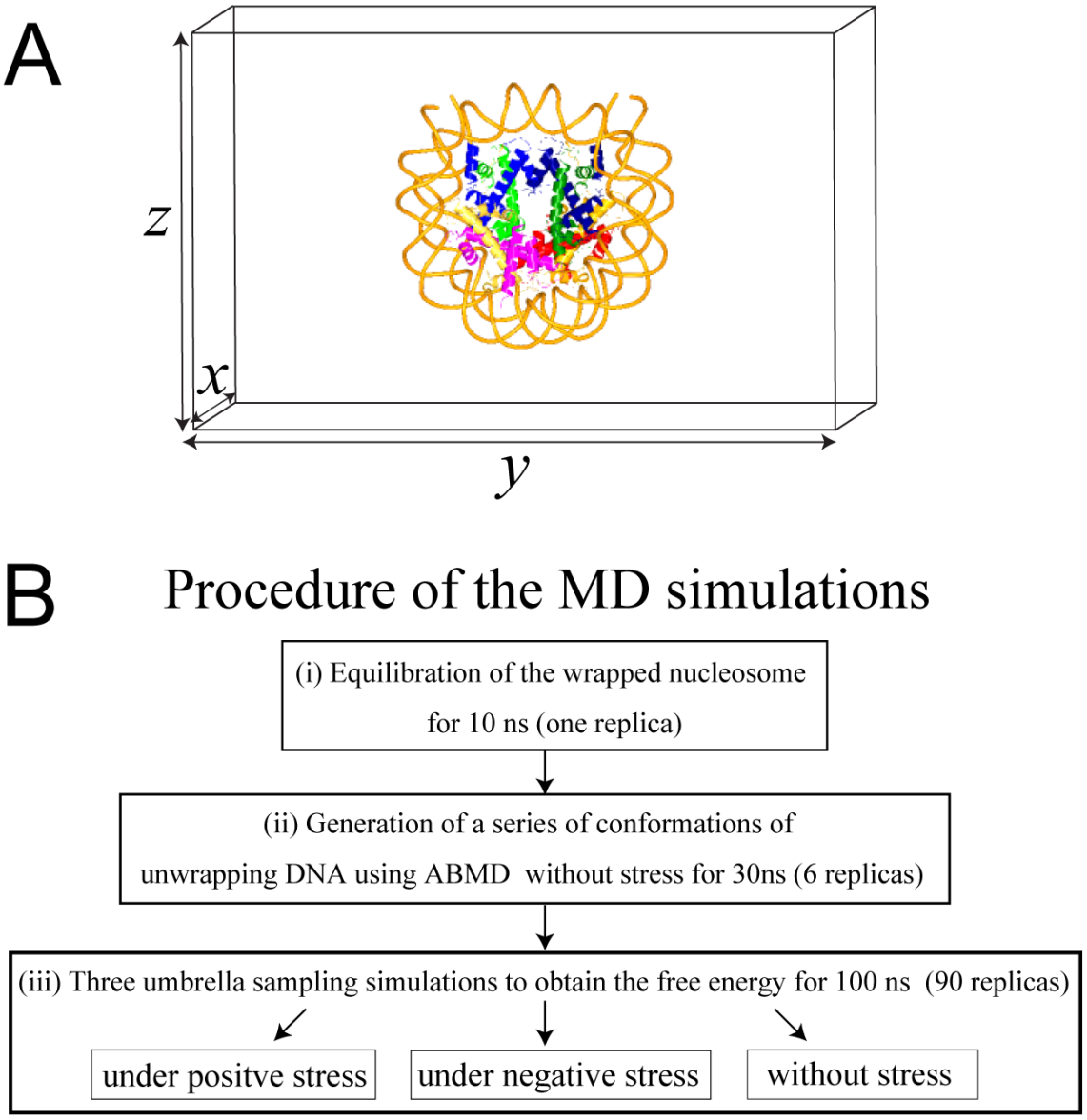


Figure S1 Setup and procedure of the simulations.

(*A*) Simulation setup: a nucleosome particle composed of a histone core and a 147-bp DNA is located in a rectangular box of ~125 Å × 245 Å × 155 Å.

(*B*) The procedure of the simulations for the free energy analysis.

Schematic diagram to show the procedure of the simulations for the free energy analysis. (i) The system (one replica) was equilibrated at a constant pressure of one bar and a temperature of 300 K for 10 ns. (ii) To generate a series of configurations of the unwrapping of the nucleosome from tightly wrapped to unwrapped, the adaptively biased molecular dynamics (ABMD) method (17) combined with the multiple walker method

(18) was carried out for ~30 ns without stress with 6 walkers (replicas) of the nucleosome which were selected from the equilibration simulation. (iii) Using a series of 90 conformations in the ABMD simulation, umbrella sampling simulations were carried out for 100 ns under positive, negative and no stress. The free energy for the unwrapping of nucleosomal DNA was calculated using the WHAM algorithm (19, 20).

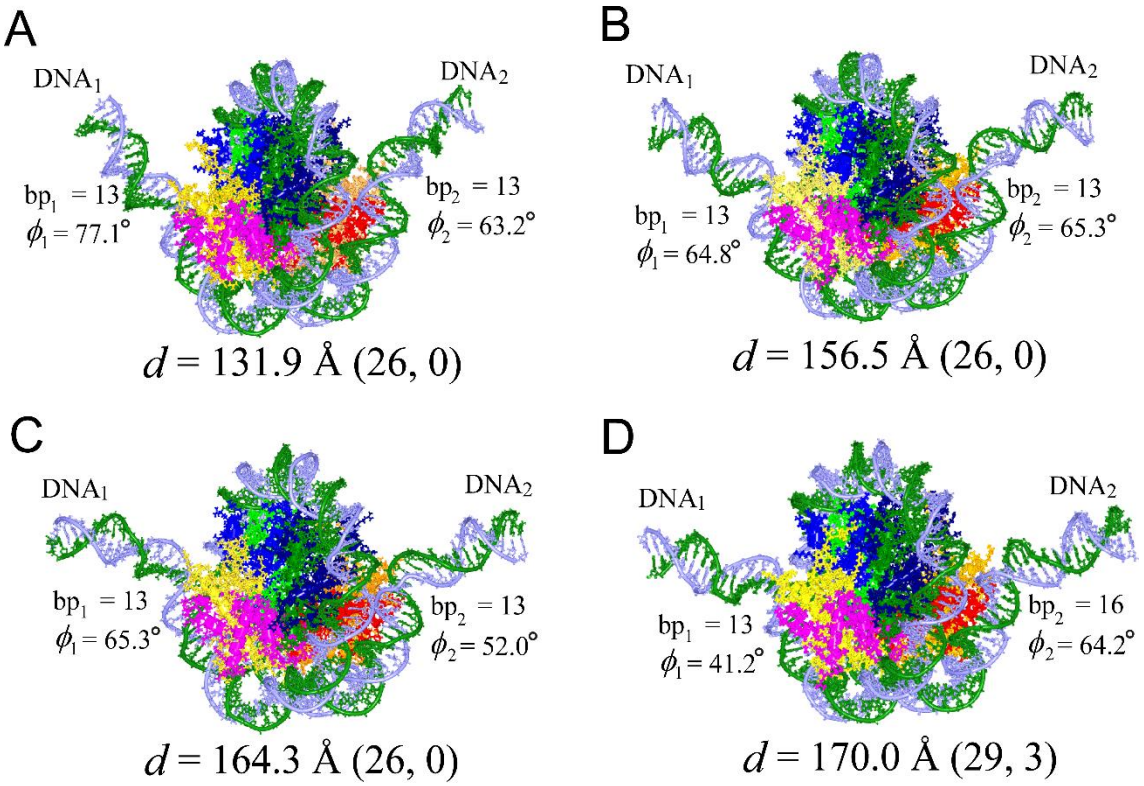
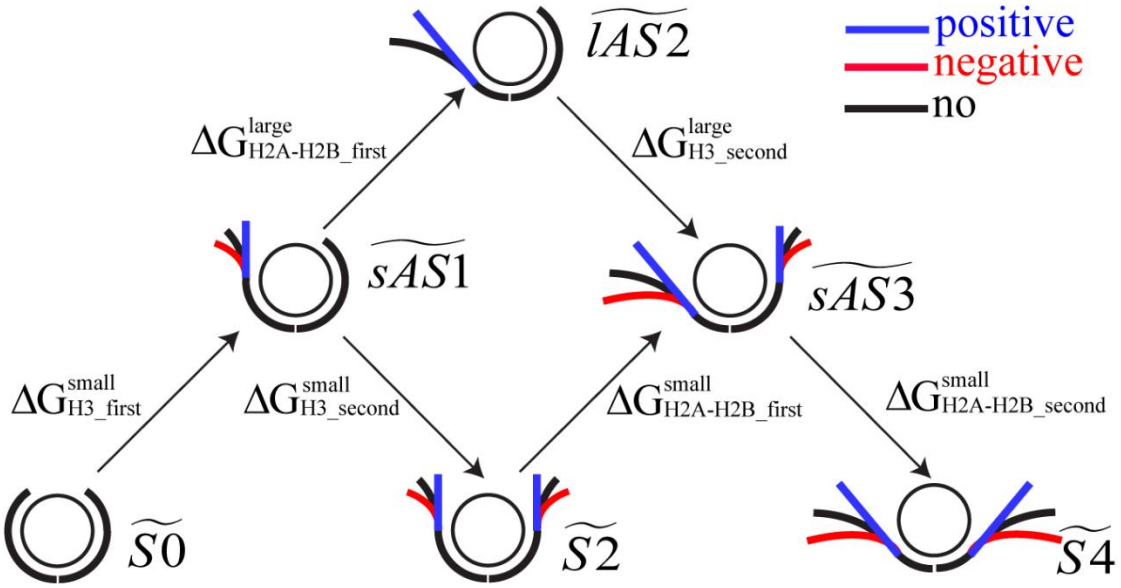


Figure S2 Four representative conformations observed at *d* = (*A*)131.9, (*B*)156.5, (*C*)164.3 and (*D*)170.0 Å within the state of $S2$ under negative stress.

These conformations are located in the range of *d* from 130 to 170 Å where the state of $S2$ is exclusively dominant (the right-hand column of Fig. 3). bp_{total} and bp_{diff} are shown in parentheses as (bp_{total}, bp_{diff}) with *d*. The free energies at *d* = 132, 157, 165, and 170 Å near (*A*-*D*) are 7.0, 11.4, 13.4, and 15.0 kcal/mol, respectively. DNA_2 was observed to be deformed in (*C*), where a maximum force of 27 pN was observed at *d* = 164 Å (Fig. 3). The conformation in (*C*) is the same as that at (bp_{total}, bp_{diff}) = (26,0) under negative stress in the left-hand column of Fig. 3.

The conformational change of the DNA by the extension of *d* from 132 to 157 Å is mainly bending of the DNA (See *A* to *B*). After *d* = 157 Å DNA_2 was stretched and partially deformed due to retained contact between DNA and H2A-H2B at *d* = 164 Å, where the maximum force of 27 pN was observed (*C*). Further unwrapping of DNA_2 from bp_2 = 13 to 16 overcame the DNA – H2A-H2B interaction, resulting in the disappearance of the deformation at *d* = 157 Å (*D*).



large-scale asymmetric unwrapping

$$\Delta G^{large}_{H2A\text{-}H2B\_first} = \begin{cases} 5.6 \\ ND \\ 12.1 \end{cases} \qquad \Delta G^{large}_{H3\_second} = \begin{cases} -0.3 \\ ND \\ -4.0 \end{cases}$$

small-scale asymmetric unwrapping

$$\Delta G^{small}_{H3\_first} = \begin{cases} 3.6 \\ 1.9 \\ 2.4 \end{cases} \quad \Delta G^{small}_{H3\_second} = \begin{cases} 3.6 \\ 3.0 \\ 2.9 \end{cases} \quad \Delta G^{small}_{H2A\text{-}H2B\_first} = \begin{cases} 1.7 \\ 12.5 \\ 5.3 \end{cases} \quad \Delta G^{small}_{H2A\text{-}H2B\_second} = \begin{cases} 3.5 \\ 1.9 \\ 4.6 \end{cases}$$

Figure S3 Interpretation of the contribution of the interactions between the DNA and histones from the free energies for the small- and large-scale unwrapping of the DNA.

For the small-scale asymmetric unwrapping of the DNA, $\Delta G^{small}_{H3\_first}$ is the free energy for unwrapping one end of the DNA from H3. $\Delta G^{small}_{H3\_second}$ is the free energy for unwrapping the other end of the DNA from H3. $\Delta G^{small}_{H2A\text{-}H2B\_first}$ and $\Delta G^{small}_{H2A\text{-}H2B\_second}$ are the free energies for unwrapping the DNA from H2A-H2B on one and the other sides of the DNA, respectively. For the large-scale asymmetric unwrapping of the DNA, $\Delta G^{large}_{H2A\text{-}H2B\_first}$ is the free energy for continuously unwrapping the DNA from H2A-H2B. $\Delta G^{large}_{H3\_second}$ is the free energy for unwrapping the DNA from H3 on the other side of the DNA. The arrows show the relative free energy for each ΔG. The unit is kcal/mol.

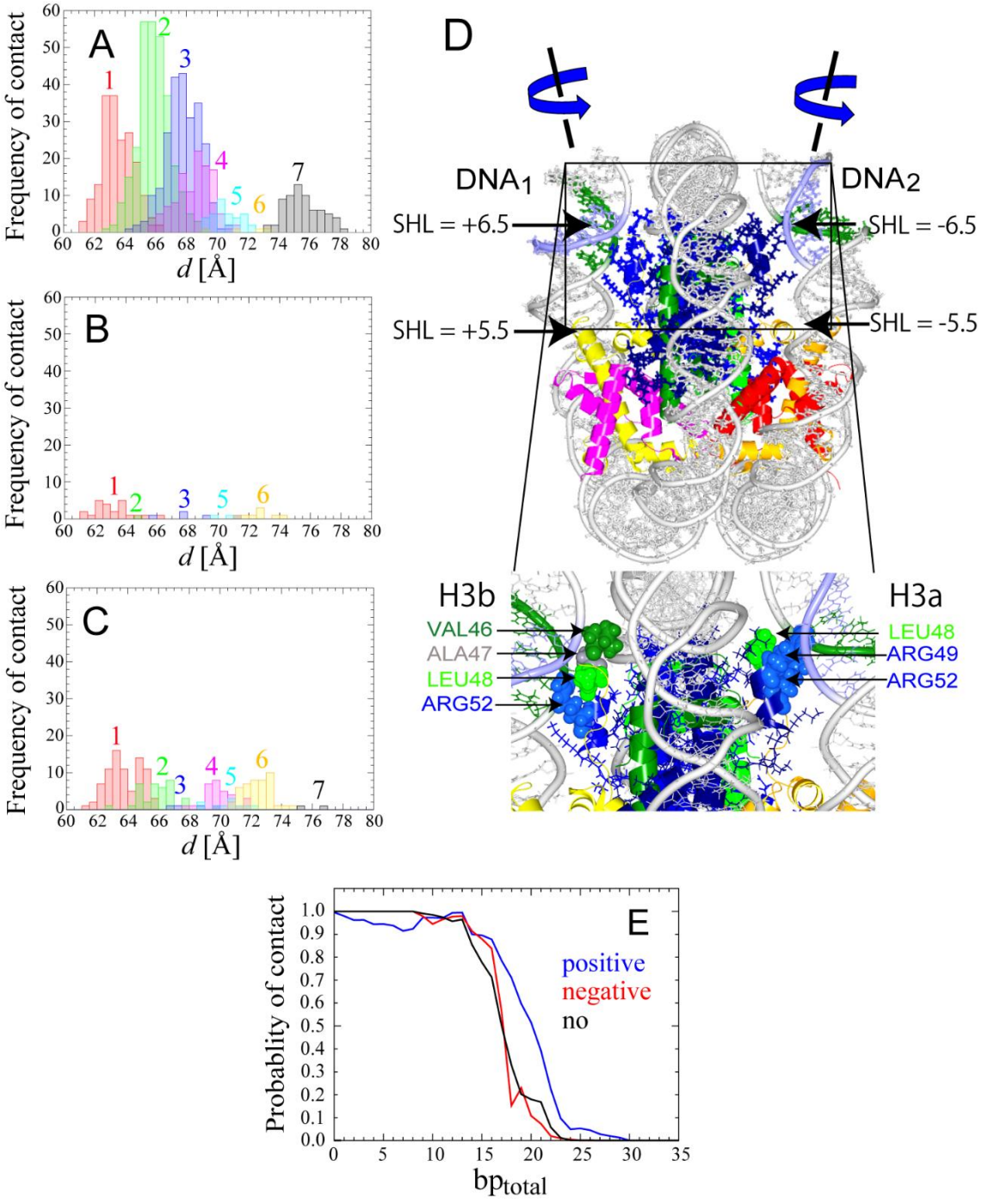


Figure S4 Contact between the minor grove DNA at SHL = ±6.5 and the H3 αN-helix.

The frequency of the simultaneous specific contact between the minor grove DNA at SHL = ±6.5 and the H3 αN-helix (residues 46-57) at both ends of the DNA is shown as a histogram under (*A*) positive (*B*) negative, and (*C*) no stress. Specific contact was counted as 1 when any pair of atoms from 70 T to 66 G in chain I of DNA_1 (the strand of DNA_1 in thin purple in (*D*)) and the H3 αN-helix in H3a were located within a distance

of 4 Å of each other, OR any pair of atoms from 70 T to 66 G in chain J of DNA_2 (the strand of DNA_2 in green in (*D*)) and the H3 αN-helix in H3b were located within a distance of 4 Å of each other. Simultaneous specific contact was counted as 1 when the specific contact was observed at DNA1 and DNA2 at the same time. The frequency of the simultaneous specific contact was analyzed at d_{fix} = 64, 66, 68, 70, 72, 74, and 76 Å. The histogram from trajectories at d_{fix} = 64, 66, 68, 70, 72, 74, and 76 Å are labeled as 1, 2, 3, 4, 5, 6, and 7. The frequency of the simultaneous contact was zero at $d_{fix} \geq 78$ Å. A representative conformation of the nucleosome with the simultaneous specific contact is shown in (*D*). The residues of the αN helix of H3, Val46, Ala47, Leu48, Arg49, and Arg52, which are in direct contact with the minor groove of DNA_1 (the strand of DNA_1 in thin purple) and DNA_2 (the strand of DNA_2 in green) are labeled. The probability of contact between the minor grove of the DNA at SHL = ±6.5 and the H3 αN-helix in the small-scale asymmetric unwrapping is shown against bp_{total} in (*E*). Here, contact was counted as 1 when any pair of atoms from −70 A:70 T, −69 A:69 T, −68 T:68 A, −67 A:67 T, and −66 T:66 A of DNA_1 (the double strands of DNA_1 in thin purple and green in (*D*)) and the H3 αN-helix in H3a were located within a distance of 4 Å of each other, OR when any pair of atoms from 70 T:−70 A, 69 T:−69 A, 68 A:−68 T, 67 T:−67 A and 66 A:−66 T of DNA_2 (the double strands of DNA_2 in green and thin purple in (*D*)) and the H3 αN-helix in H3b were located within a distance of 4 Å of each other. The probability of contact in the small-scale asymmetric unwrapping was calculated from (the number of conformations with $bp_{diff} \leq 13$ which had contact at bp_{total}) / (the total number of conformations with $bp_{diff} \leq 13$ at bp_{total}). The analysis used trajectories of 50 ns at intervals of 100 ps (500 conformations), and was carried out using Eq. (S8) without the weight of the probability distribution to evenly consider the conformations.

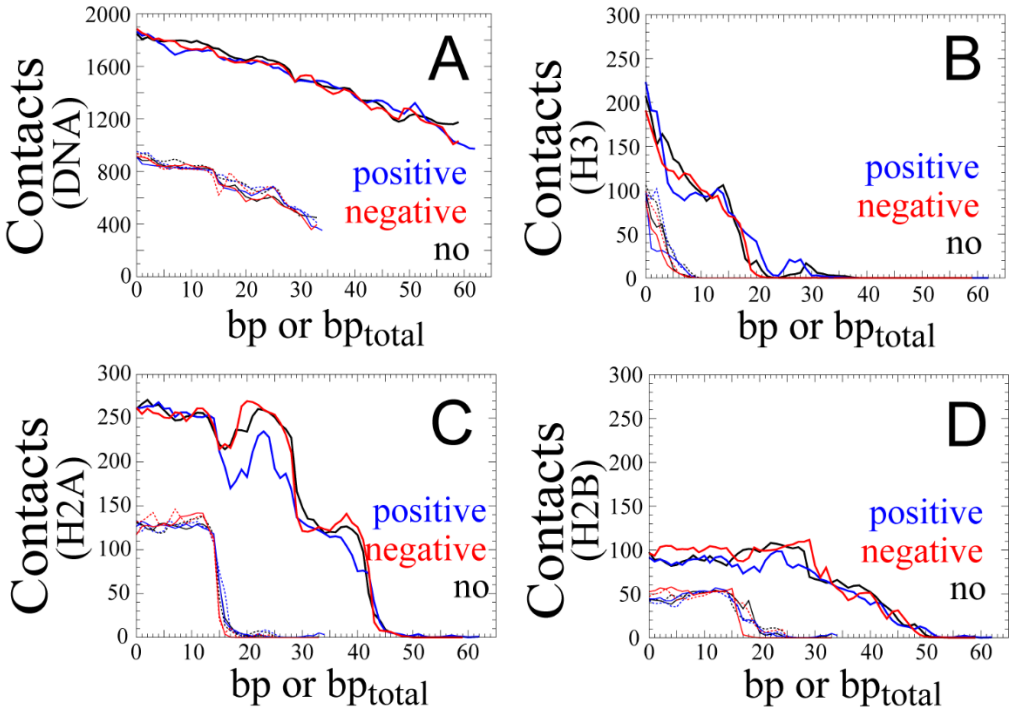
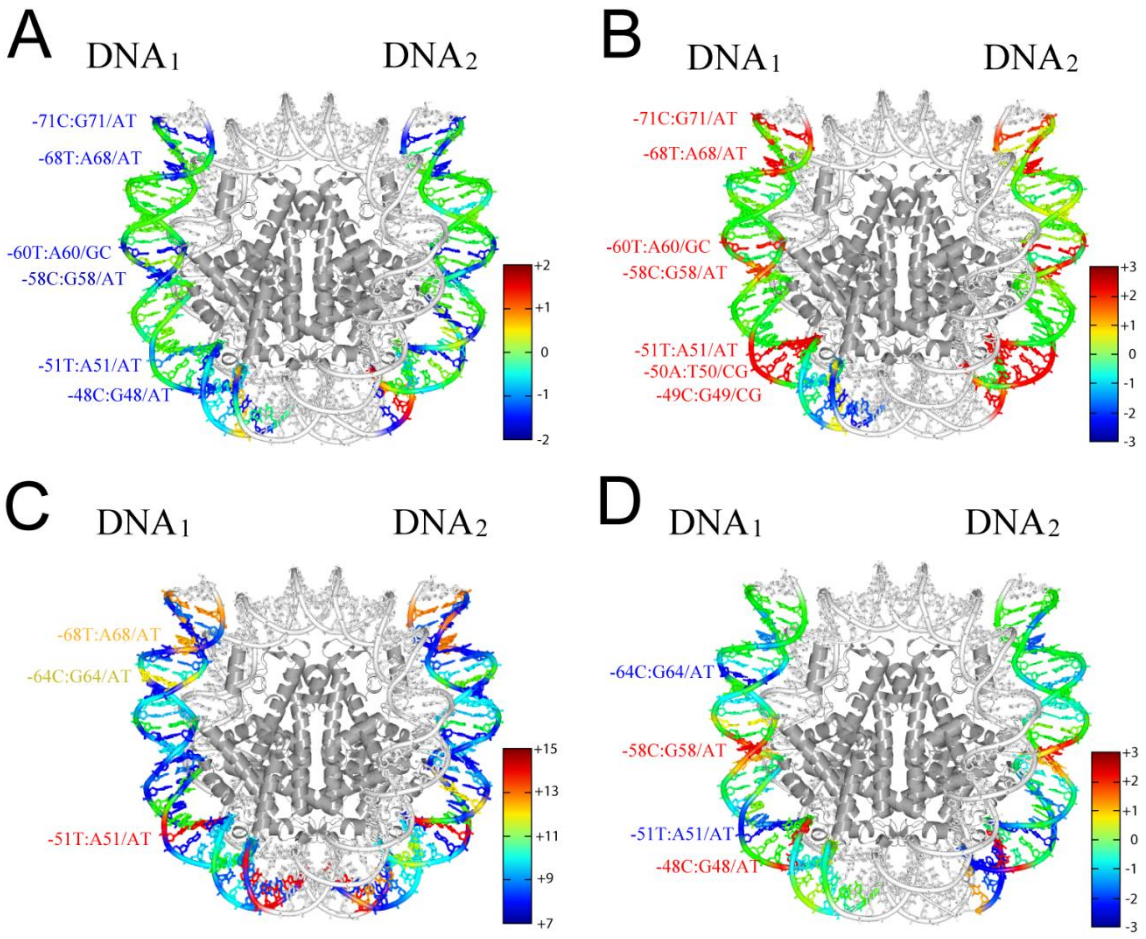
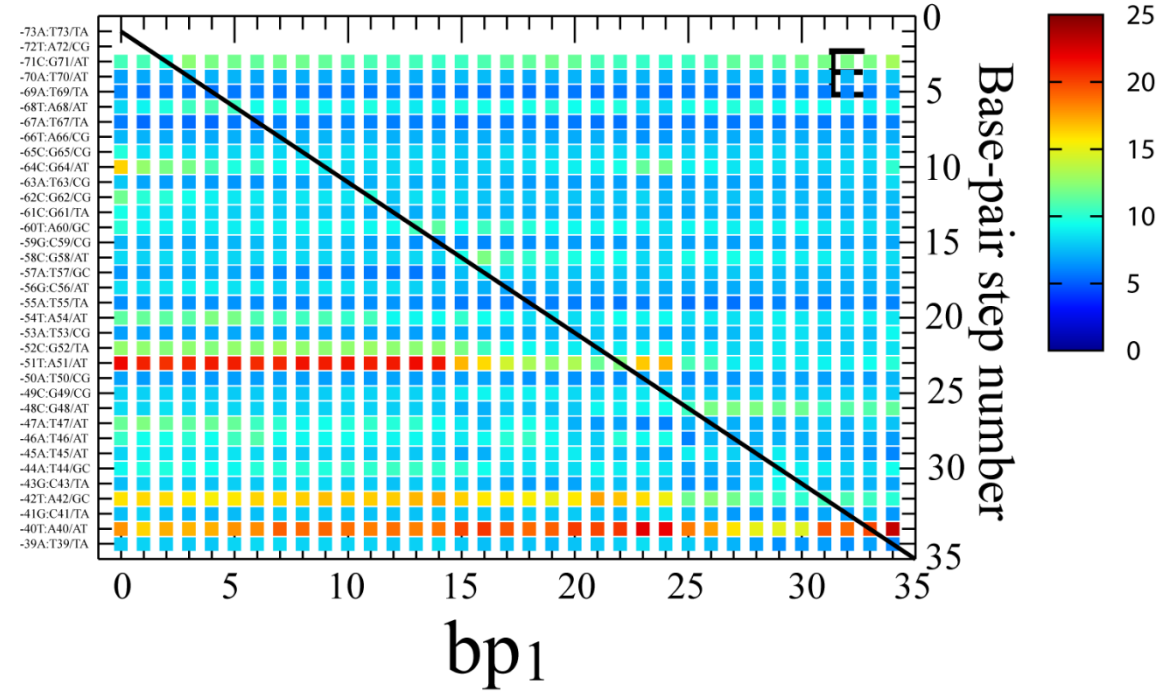
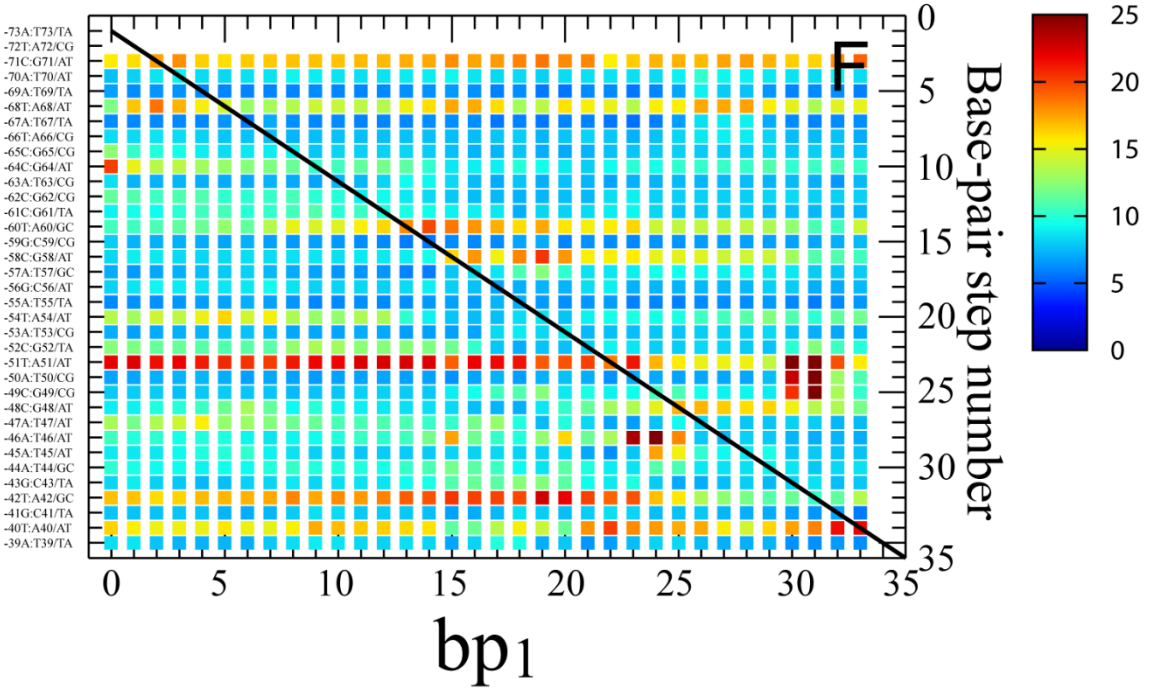


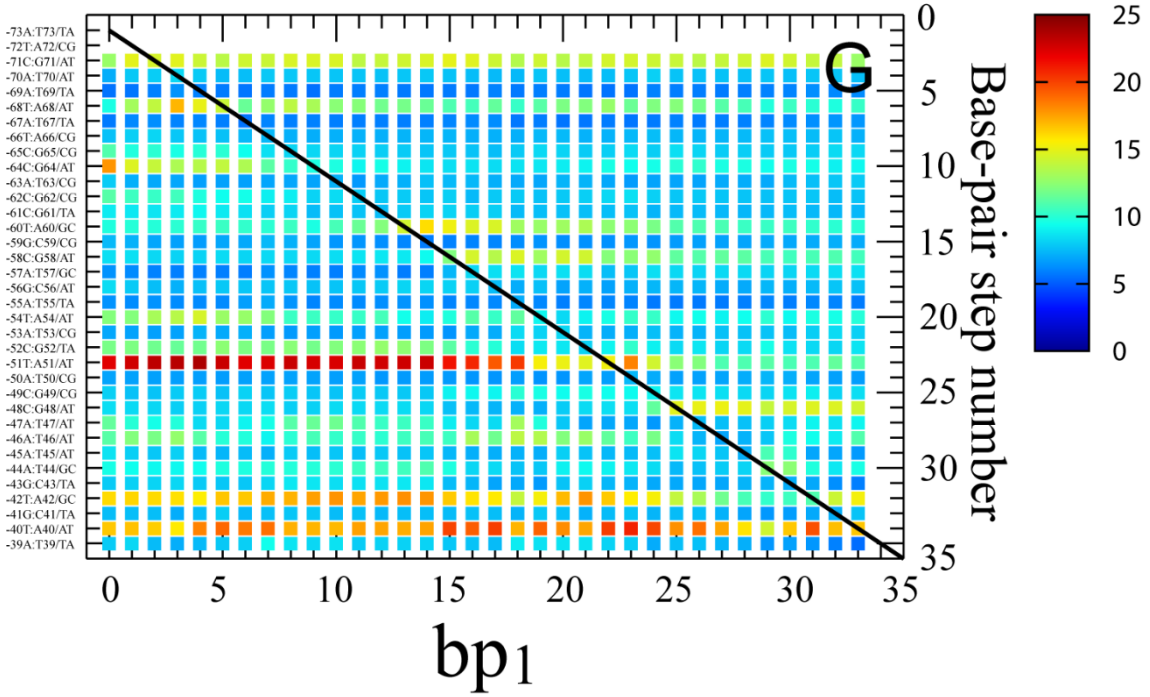
Figure S5 The number of contacts between the DNA and the octamer against bp or bp_{total}. Upper three lines are against bp_{total}. Lower six lines are against bp. It should be noted that the number of contacts against ($bp_{total} = bp_1 + bp_2$) is not equal to the number of contacts against bp_1 plus the number of contacts against bp_2.

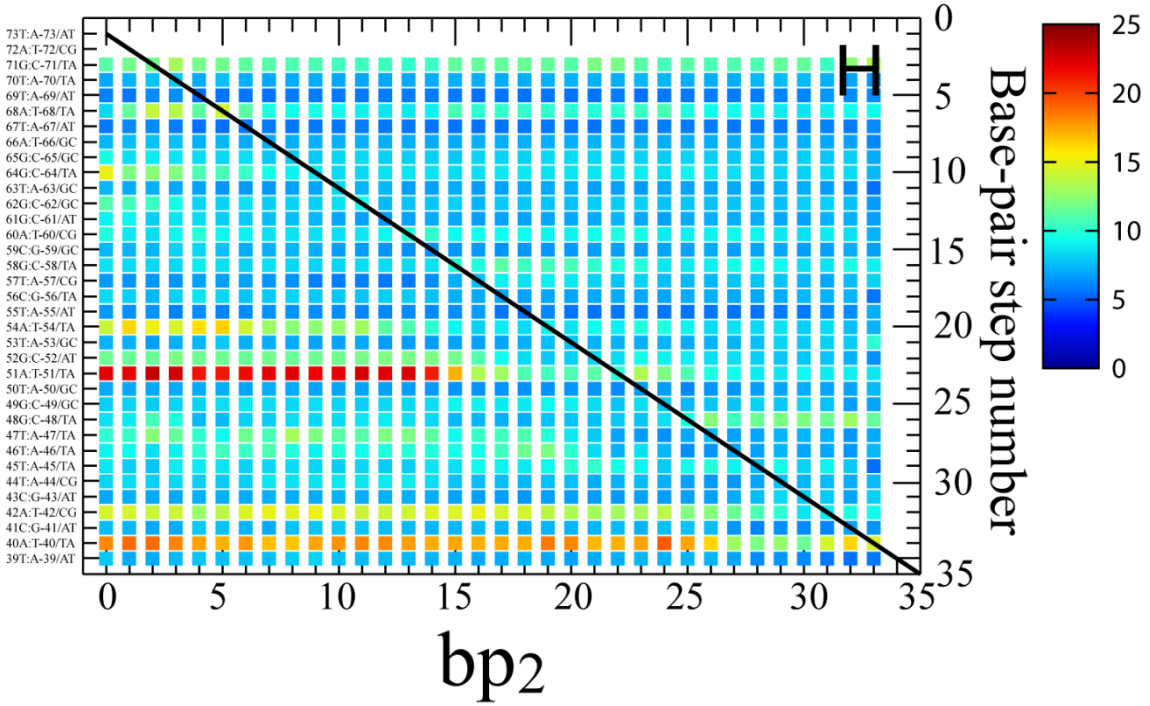
(*A*) Total number of contacts between the DNA and the octamer. (*B*) The number of contacts between the base pairs, −70 A:70 T, −69 A:69 T, −68 T:68 A, −67 A:67 T, and −66 T:66 A, and H3a (thick line) and between the base pairs, 70 T:−70 A, 69 T:−69 A, 68 A:−68 T, 67 T:−67 A, and 66 A:−66 T, and H3b (thin line). These base pairs mainly interact with the N-terminal α-helix of H3. (*C*) The number of contacts between the base pairs, −61 C:61 G, −60 T:60 A, −59 G:59 C, −58 C:58 G, and −57 A:57 T, and H2Aa (thick line) and between the base pairs, 61 G:−61 C, 60 A:−60 T, 59 C:−59 G, 58 G:−58 C, and 57 T:−57 A, and H2Ab (thin line). These base pairs mainly interact with the L2 loop of H2A. (*D*) The number of contacts between the base pairs, −56 G:56 C, −55 A:55 T, −54 T:54 A, −53 A:53 T, and −52 C:52 G, and H2Ba (thick line) and between the 56 C:−56 G, 55 T:−55 A, 54 A:−54 T, 53 T:−53 A, and 52 G:−52 C, and H2Bb (thin line). These base pairs mainly interact with the L1 loop of H2B. Unwrapping of the nucleotides in $DNA_{1 \text{ or } 2}$, T −68, A −67, T −66 in (*B*), A 60, C 59, G 58 in (*C*), and A −55, T −54, A −53 in (*D*) occurs at bp = 6−8, 14−16, and 19−21 on either end of the DNA, respectively. This analysis was carried out using Eq. (S8) without the weight of the probability distribution to evenly consider the conformations.


A
DNA1
DNA2
-71C:G71/AT
-68T:A68/AT
-60T:A60/GC
-58C:G58/AT
-51T:A51/AT
-48C:G48/AT
+2
+1
0
-1
-2
B
DNA1
DNA2
-71C:G71/AT
-68T:A68/AT
-60T:A60/GC
-58C:G58/AT
-51T:A51/AT
-50A:T50/CG
-49C:G49/CG
+3
+2
+1
0
-1
-2
-3
C
DNA1
DNA2
-68T:A68/AT
-64C:G64/AT
-51T:A51/AT
+15
+13
+11
+9
+7
D
DNA1
DNA2
-64C:G64/AT
-58C:G58/AT
-51T:A51/AT
-48C:G48/AT
+3
+2
+1
0
-1
-2
-3


-73A:T73/TA
-72T:A72/CG
-71C:G71/AT
-70A:T70/AT
-69A:T69/TA
-68T:A68/AT
-67A:T67/TA
-66T:A66/CG
-65C:G65/CG
-64C:G64/AT
-63A:T63/CG
-62C:G62/CG
-61C:G61/TA
-60T:A60/GC
-59G:C59/CG
-58C:G58/AT
-57A:T57/GC
-56G:C56/AT
-55A:T55/TA
-54T:A54/AT
-53A:T53/CG
-52C:G52/TA
-51T:A51/AT
-50A:T50/CG
-49C:G49/CG
-48C:G48/AT
-47A:T47/AT
-46A:T46/AT
-45A:T45/AT
-44A:T44/GC
-43G:C43/TA
-42T:A42/GC
-41G:C41/TA
-40T:A40/AT
-39A:T39/TA
0
5
10
15
20
25
30
35
bp1
Base-pair step number
25
20
15
10
5
0


-73A:T73/TA
-72T:A72/CG
-71C:G71/AT
-70A:T70/AT
-69A:T69/TA
-68T:A68/AT
-67A:T67/TA
-66T:A66/CG
-65C:G65/CG
-64C:G64/AT
-63A:T63/CG
-62C:G62/CG
-61C:G61/TA
-60T:A60/GC
-59G:C59/CG
-58C:G58/AT
-57A:T57/GC
-56G:C56/AT
-55A:T55/TA
-54T:A54/AT
-53A:T53/CG
-52C:G52/TA
-51T:A51/AT
-50A:T50/CG
-49C:G49/CG
-48C:G48/AT
-47A:T47/AT
-46A:T46/AT
-45A:T45/AT
-44A:T44/GC
-43G:C43/TA
-42T:A42/GC
-41G:C41/TA
-40T:A40/AT
-39A:T39/TA
0
5
10
15
20
25
30
35
bp1
Base-pair step number
25
20
15
10
5
0


-73A:T73/TA
-72T:A72/CG
-71C:G71/AT
-70A:T70/AT
-69A:T69/TA
-68T:A68/AT
-67A:T67/TA
-66T:A66/CG
-65C:G65/CG
-64C:G64/AT
-63A:T63/CG
-62C:G62/CG
-61C:G61/TA
-60T:A60/GC
-59G:C59/CG
-58C:G58/AT
-57A:T57/GC
-56G:C56/AT
-55A:T55/TA
-54T:A54/AT
-53A:T53/CG
-52C:G52/TA
-51T:A51/AT
-50A:T50/CG
-49C:G49/CG
-48C:G48/AT
-47A:T47/AT
-46A:T46/AT
-45A:T45/AT
-44A:T44/GC
-43G:C43/TA
-42T:A42/GC
-41G:C41/TA
-40T:A40/AT
-39A:T39/TA
G
0
5
10
15
20
25
30
35
Base-pair step number
25
20
15
10
5
0
bp1


73T:A-73/AT
72A:T-72/CG
71G:C-71/TA
70T:A-70/TA
69T:A-69/AT
68A:T-68/TA
67T:A-67/AT
66A:T-66/GC
65G:C-65/GC
64G:C-64/TA
63T:A-63/GC
62G:C-62/GC
61G:C-61/AT
60A:T-60/CG
59C:G-59/GC
58G:C-58/TA
57T:A-57/CG
56C:G-56/TA
55T:A-55/AT
54A:T-54/TA
53T:A-53/GC
52G:C-52/AT
51A:T-51/TA
50T:A-50/GC
49G:C-49/GC
48G:C-48/TA
47T:A-47/TA
46T:A-46/TA
45T:A-45/TA
44T:A-44/CG
43C:G-43/AT
42A:T-42/CG
41C:G-41/AT
40A:T-40/TA
39T:A-39/AT
H
0
5
10
15
20
25
30
35
Base-pair step number
25
20
15
10
5
0
bp2

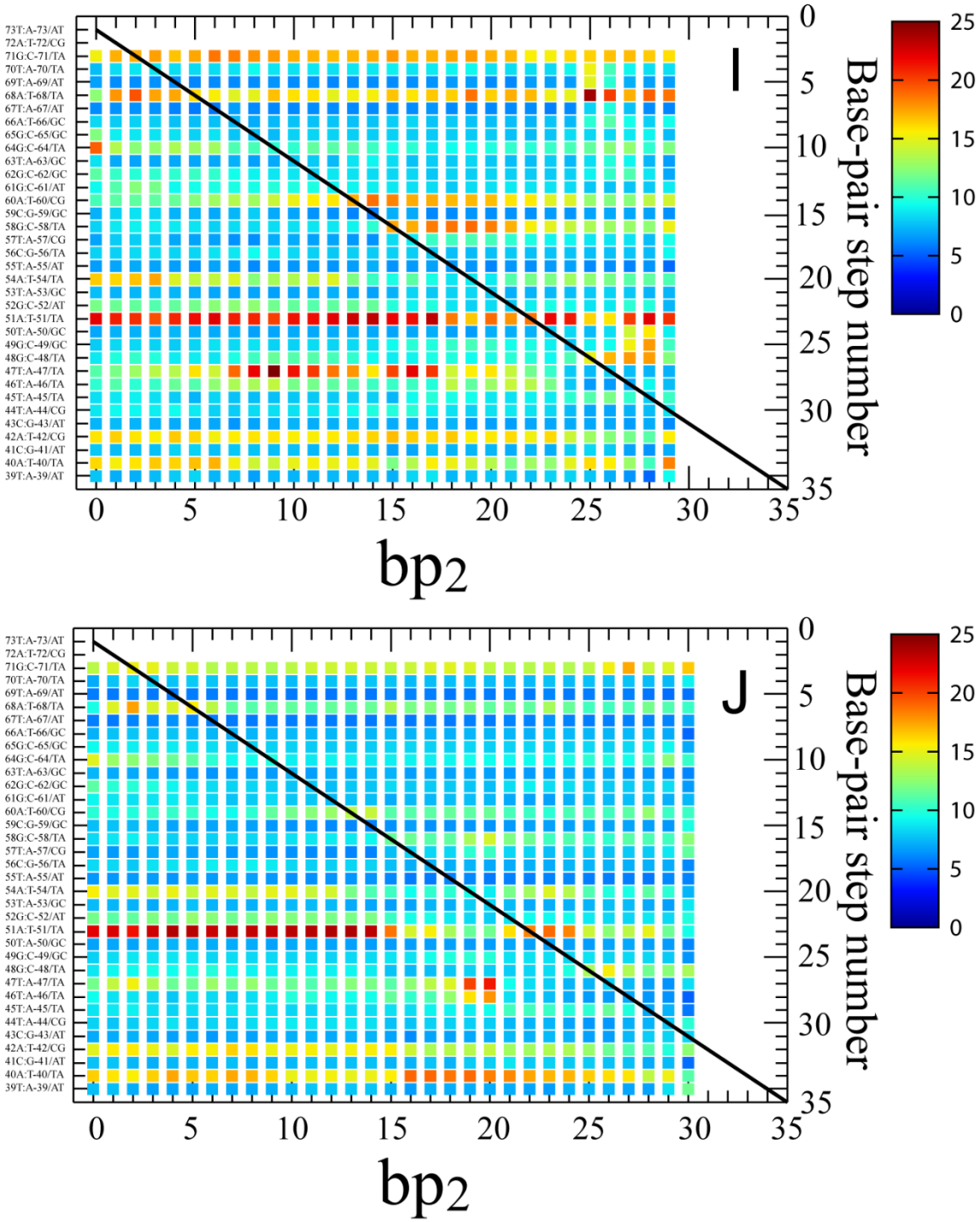


Figure S6 Changes in the local bending angles of the DNA against bp in unwrapping processes.

The differences in averaged values of the local bending angles of the DNA under between positive and no stress (*A*), and under between negative and no stress (*B*) are plotted on the DNA in color. The averaged value of the local bending angles at the base-

pair step of X-th – Y-th / (X+1)-th – (Y–1)-th is represented on the X-th – Y-th base-pair. The base-pair steps which were significantly flexible are labeled in the same color as the color in bending angle. The values of any bending angles more than 2.0 and less than −2.0 are set at 2.0 and −2.0, respectively, in (*A*). The values of any bending angles more than 3.0 and less than −3.0 are set at 3.0 and −3.0, respectively, in (*B*). (*C*) The averaged value of the local bending angles of the DNA before unwrapping under no stress is represented on the DNA. The values of any bending angles more than 15.0 and less than 7.0 are set at 15.0 and 7.0, respectively. (*D*) The differences between the averaged values of the bending angles before and after unwrapping under no stress is plotted on the DNA in color. The values of any bending angles more than 3.0 and less than −3.0 are set at 3.0 and −3.0, respectively.

The local bending angles of the DNA_1 and DNA_2 are plotted against bp_1 and bp_2 under (*E*) and (*H*) positive, (*F*) and (*I*) negative and (*G*) and (*J*) no stress, respectively. The base-pair steps and their number from the edge of the $DNA_{1\text{ or }2}$ are shown on the left and right sides, respectively. The diagonal line is drawn to show that the positions of the angles in the lower triangular region are in the wrapped part of the DNA while those in the upper triangular region are in the unwrapped part of the DNA.

For (*A*), the averaged values of the local bending angles after unwrapping in $DNA_{1\text{ or }2}$ under positive and no stress were calculated from data in the upper triangular region in (*E* and *G*) and (*H* and *J*), respectively. For (*B*), the averaged values of the local bending angles after unwrapping in $DNA_{1\text{ or }2}$ under negative and no stress were calculated from data in the upper triangular region in (*F* and *G*) and (*I* and *J*), respectively. This analysis was carried out using Eq. (S8) without the weight of the probability distribution to evenly consider the conformations.

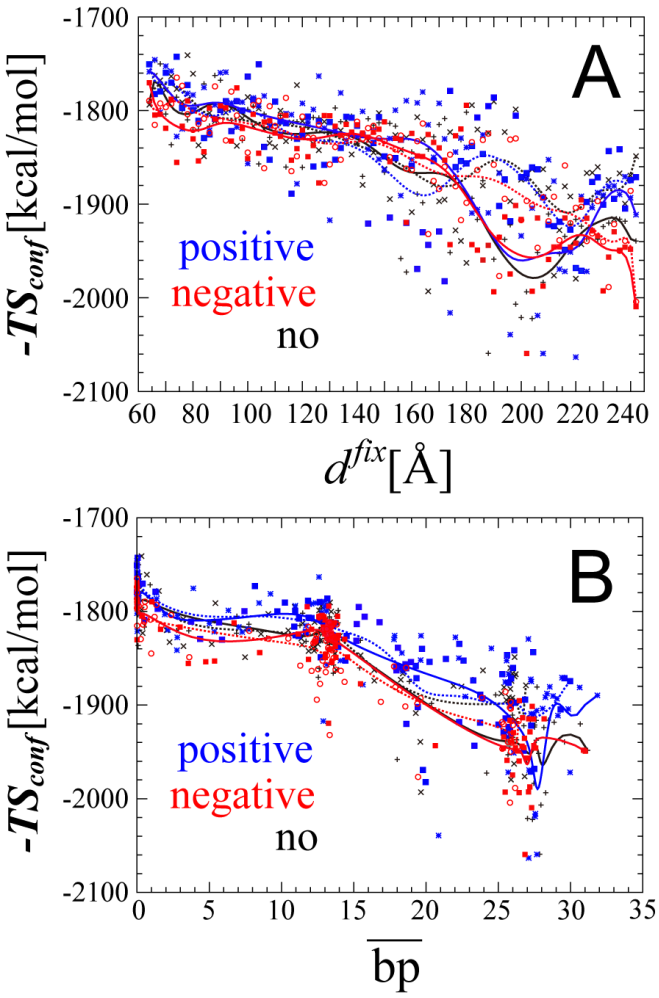


Figure S7 The conformational entropies of DNA against bp.

The conformational entropies of DNA_1 (solid line) and DNA_2 (dotted line) are shown in blue, red, and black under positive, negative, and no stress, respectively, against (*A*) d^{fix} (or the replica number) and (*B*) the average of bp. The conformational entropies were calculated according to Eq. (S13). For each d^{fix}, the average of bp was calculated in each window (according to Eq. (S8) without the weight of the probability distribution). Then the values of $-TS_{conf}$ were plotted against the average of bp with closed square/open circle ($DNA_{1\text{ or }2}$ under positive stress), plus/cross ($DNA_{1\text{ or }2}$ under negative stress) and asterisk/open square ($DNA_{1\text{ or }2}$ under no stress), respectively. Finally, the curves of $-TS_{conf}$ were smoothed using a Bézier curve. The reason why d^{fix} (not d) and the average of bp (not the individual value of bp) were used is to avoid overestimation of the conformational entropy, and keep the number of sampled conformations the same for all d^{fix} (see the SI "Conformational entropy of the DNA").

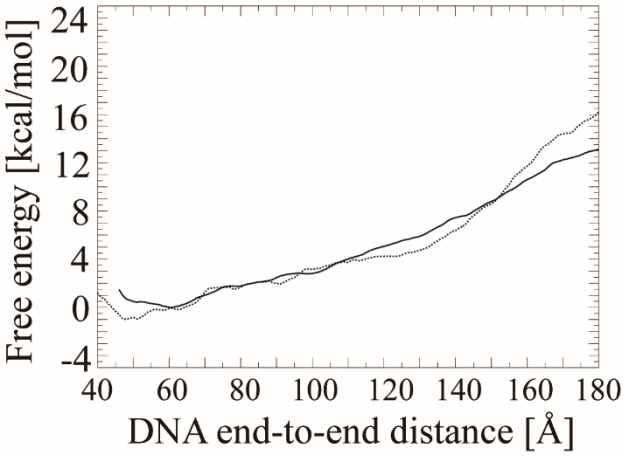


Figure S8 The difference in the free energies of unwrapping the DNA under no stress between two independent simulations.

The free energy curve under no stress at a concentration of 150 mM NaCl in this study (solid line) and the free energy curve under no stress at a concentration of 120 mM NaCl which was obtained in a previous study (8) (dotted line) were compared to check the reproducibility of the free energy curve. The reaction coordinate was defined as the DNA end-to-end distance between two phosphorous atoms of T 73 in chain I and T 73 in chain J to follow the definition used in the previous study. To meet the reaction coordinate, the transformation of the reaction coordinate in the previous study was carried out according to Eq. (S6). The free energy curves were compared in the range of the reaction coordinate from 46 to 180 Å as a smaller simulation box of 150 Å × 150 Å × 150 Å was used in the previous study. The standard deviation between them was calculated to be 0.40 kcal/mol.

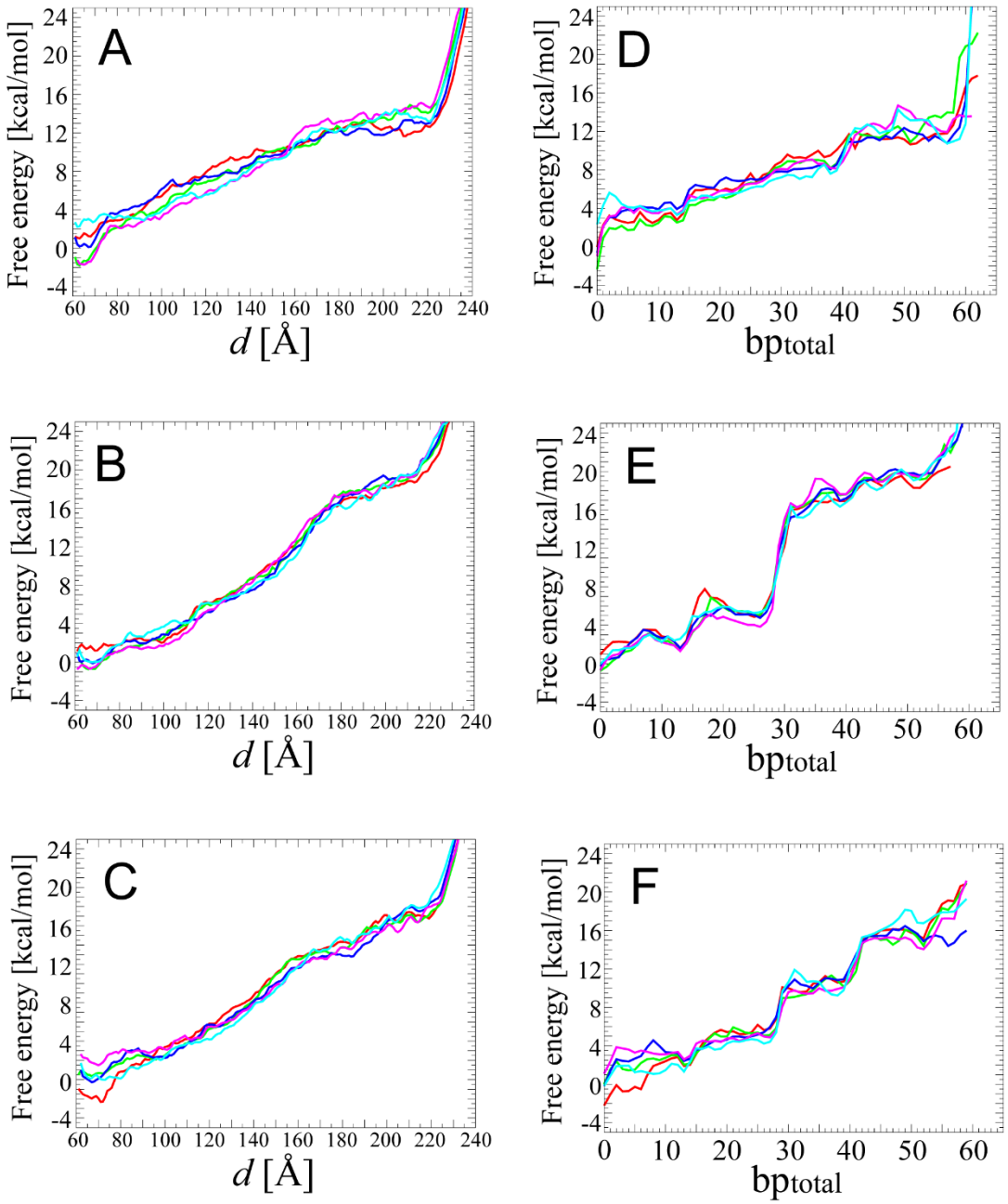


Figure S9 The free energy curves at intervals of 10 ns for the last 50 ns in the umbrella sampling simulation. (*A*) positive, (*B*) negative, and (*C*) no stress with regard to *d*, and under (*D*) positive, (*E*) negative, and (F) no stress with regard to bp_{total}, respectively. The free energy curves from 50 to 60, 60 to 70, 70 to 80, 80 to 90, and 90 to 100 ns are shown in red, green, blue, magenta, and cyan, respectively. The standard deviations for (*A*), (*B*), and (*C*) were calculated to be 0.91, 0.54, and 0.70 kcal/mol, respectively. The standard deviations for (*D*), (*E*), and (*F*) were calculated to be 0.98, 0.65, and 0.90 kcal/mol, respectively.

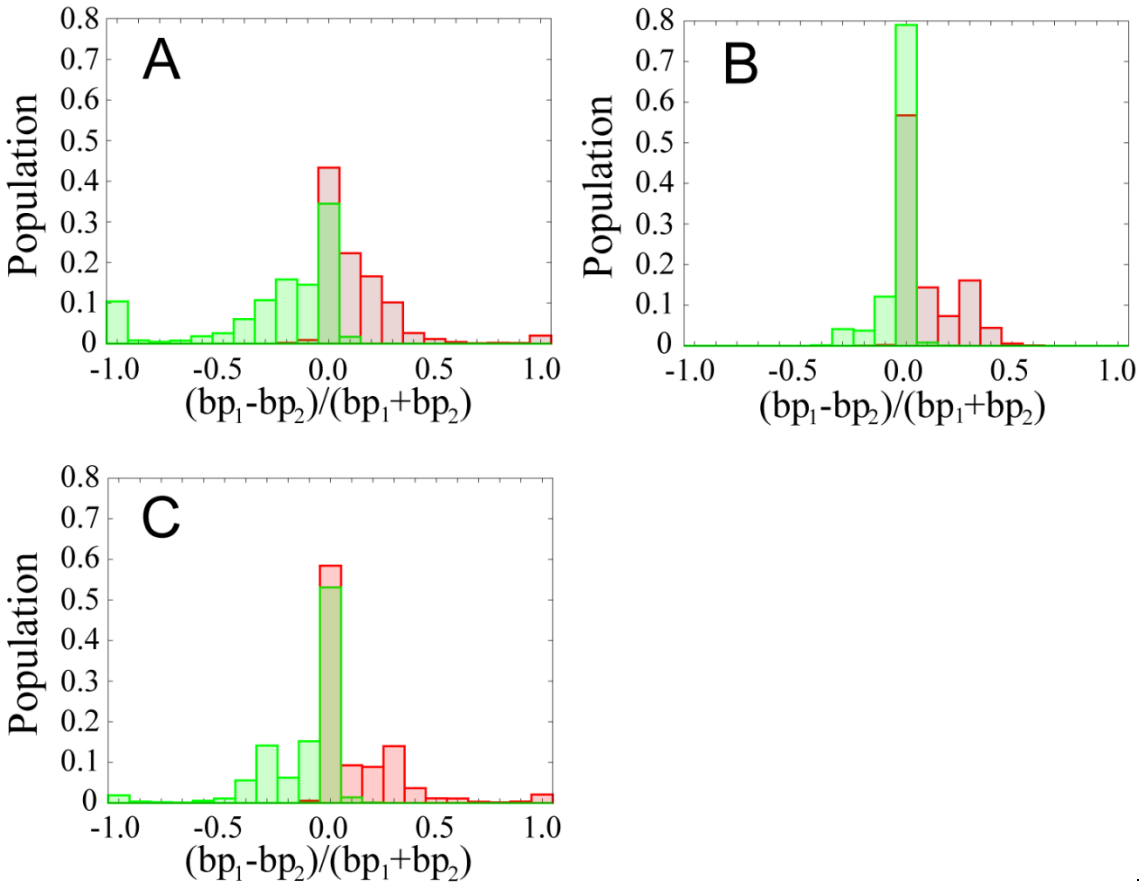


Figure S10 The population of the asymmetric unwrapping of the DNA.

The population of the nucleosome with the value of $(bp_1 - bp_2) / (bp_1 + bp_2)$ is shown as a histogram under (*A*) positive (*B*) negative, and (*C*) no stress. The conformations were classified into two groups; when the average of $(bp_1 - bp_2)$ in the conformations in a window was positive (negative) the conformations were classified into group positive (negative), respectively. The population of groups positive and negative is shown in brown and green, respectively. As we are interested in the population of the large asymmetric unwrapping of the DNA, the conformations with bp_{total} $(= bp_1 + bp_2) \geq 20$ were included in the analysis. This classification was carried out for all the windows. The population at –0.X and +0.X (X = 0, 1, 2,…, 9) in the histogram is the population in the range of (bp1 – bp2) / (bp1 + bp2) from more than –0.X – 0.1 to less than or equal to –0.X, and from equal to or more than +0.X to less than +0.X + 0.1. The population at –1.0 and 1.0 in the histogram is the population of the complete symmetric unwrapping of the DNA for $bp_1 = 0$ and $bp_2 = 0$, respectively.

References

1. Davey CA, Sargent DF, Luger K, Maeder AW, & Richmond TJ (2002) Solvent mediated interactions in the structure of nucleosome core particle at 1.9 A resolution. *J. Mol. Biol.* 319:1097-1113.
2. Ishida H & Hayward S (2008) Path of nascent polypeptide in exit tunnel revealed by molecular dynamics simulation of ribosome. *Biophys. J.* 95:5962-5973.
3. Ishida H (2010) Branch migration of Holliday junction in RuvA tetramer complex studied by umbrella sampling simulation using a path-search algorithm. *J. Comput. Chem.* 31:2317-2329.
4. Ishida H & Matsumoto A (2014) Free-energy landscape of reverse tRNA translocation through the ribosome analyzed by electron microscopy density maps and molecular dynamics simulations. *PLOS ONE* 9:e101951.
5. Ishida H (2014) Essential function of the N-termini tails of the proteasome for the gating mechanism revealed by molecular dynamics simulations. *Proteins* 82:1985-1999.
6. Ishida H & Matsumoto A (2016) Mechanism for verification of mismatched and homoduplex DNAs by nucleotides-bound MutS analyzed by molecular dynamics simulations. *Proteins* 84:1287-1303.
7. Ishida H & Kono H (2017) H4 Tails Potentially Produce the Diversity in the Orientation of Two Nucleosomes. *Biophys J.* 113:978-990.
8. Kono H, Sakuraba S, & Ishida H (2018) Free energy profiles for unwrapping the outer superhelical turn of nucleosomal DNA. *PLoS Comput Biol.* 14:e1006024.
9. Maier JA*, et al.* (2015) ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB. *J. Chem. Theory Comput* 11:3696-3713.
10. Ivani I*, et al.* (2016) Parmbsc1: a refined force field for DNA simulations. *Nat. Methods* 13:55-58.
11. Joung IS & Cheatham TE (2008) Determination of alkali and halide monovalent ion parameters for use in explicitly solvated biomolecular simulations. *J. Phys. Chem. B* 112:9020-9041.
12. Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, & Klein ML (1983) Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* 79:926-935.
13. Hockney RW & Eastwood JW eds (1988) *Computer simulation using particles* (Adam Hilger, New York).
14. Deserno M & Holm C (1998) How to mesh up Ewald sums. I. A theoretical and numerical comparison of various particle mesh routines. *J. Chem. Phys.* 109:7678-7693.
15. Ryckaert J, Cicotti G, & Berendsen HJC (1977) Numerical integraton of the Cartesian equations of motion of a system with constraints: molecular dynamics of *n*-alkanes. *J. Comput. Phys.* 23:327-341.
16. van Gunsteren WF & Berendsen HJC (1977) Algorithms for macromolecular dynamics and constraint dynamics. *Mol. Phys.* 34:1311-1327.

17. Babin V, Roland C, & Sagui C (2008) Adaptively biased molecular dynamics of free energy calculations. *J. Chem. Phys.* 128:134101.
18. Raiteri P, Laio A, Gervasio FL, Micheletti C, & Parrinello M (2006) Efficient Reconstruction of Complex Free Energy Landscapes by Multiple Walkers Metadynamics. *J. Phys. Chem. B.* 110:3533-3539.
19. Kumar S, Bouzida D, Swendsen RH, Kollman PA, & Rosenberg JM (1992) The weighted histogram analysis method for free energy calculations on biomolecules. I. The method. *J. Comp. Chem.* 13:1011-1021.
20. Souaille M & Roux B (2001) Extension to the weighted histogram analysis method: combining umbrella sampling with free energy calculations. *Computer Physics Communications* 135(1):40-57.
21. Kannan S & Kai Kohlhoff aMZ (2006) B-DNA Under Stress: Over- and Untwisting of DNA during Molecular Dynamics Simulations. *Biophys J.* 91:2956-2965.
22. Liebl K & Zacharias M (2017) Unwinding Induced Melting of Double-Stranded DNA Studied by Free Energy Simulations. *J. Phys. Chem. B.* 121:11019-11030.
23. Schlitter J (1993) Estimation of Absolute and Relative Entropies of Macromolecules Using the Covariance Matrix. *Chem. Phys. Lett.* 215:617-621.