Free energy methods

To calculate accurate free energy maps, we used the recently developed Adaptively Biased Molecular Dynamics (ABMD) method [28,29], with supplementary equilibrium umbrella sampling runs [29,40]. In addition, Steered MD (SMD) [30] simulations were used to study select transition pathways. Because the ABMD and SMD methods have already been described extensively in the literature, only a very brief discussion of these methods is given here.

ABMD is a nonequilibrium MD method that belongs to the general category of “adaptively-biased” simulations, that are based on umbrella sampling with a history-dependent biasing potential [31–35]. These methods all provide an elegant way of computing the free energy [36] — or potential of mean force (PMF) — of a collective variable $\sigma(r_1, \ldots, r_N)$ defined to be a smooth function of the atomic positions $r_1, \ldots, r_N$:

$$ f(\xi) = -k_B T \ln p(\xi), $$

where $k_B$ is the Boltzmann constant, $T$ is the temperature, and

$$ p(\xi) = \langle \delta[\xi - \sigma(r_1, \ldots, r_N)] \rangle, $$

is the probability density of the collective variable (the angular brackets denote an ensemble average). The ABMD method as currently implemented in AMBER v.10-12 [37–39] is characterized by a linear scaling in time and two control parameters (the kernel width $\Delta \xi$ and the flooding timescale $\tau_F$, as discussed below).

The ABMD method is formulated in terms of the following equations:

$$ m_a \frac{d^2 r_a}{dt^2} = F_a - \frac{\partial}{\partial r_a} U[t|\sigma(r_1, \ldots, r_N)], $$

$$ \frac{\partial U(t|\xi)}{\partial t} = \frac{k_B T}{\tau_F} G[\xi - \sigma(r_1, \ldots, r_N)], $$

where the first one contains the Newton’s equations that govern ordinary molecular dynamics (MD) (temperature and pressure regulation terms are not shown), augmented with an additional force coming from the time-dependent biasing potential $U(t|\xi)$ (with $U(t = 0|\xi) = 0$), whose time evolution is given by the second equation. In the following, we refer to $\tau_F$ as flooding timescale and to $G(\xi)$ as the kernel (in analogy to the kernel density estimator widely used in statistics [62]). The kernel should be positive definite ($G(\xi) > 0$) and symmetric ($G(-\xi) = G(\xi)$). It may be viewed as a smoothed Dirac delta function. For large enough $\tau_F$ and small enough kernel width, the biasing potential $U(t|\xi)$ converges towards $-f(\xi)$ as $t \to \infty$ [34,35].

To have a more efficient flattening of the effective free energy surface we use the “multiple walkers” [35,63] extension of the ABMD method that amounts to carrying out several different MD simulations (replicas), biased by the same effective potential, which is “flattened” by all the replicas simultaneously.

We note that in order to both estimate the error and improve the accuracy of the free energy maps obtained from ABMD simulations, follow-up equilibrium umbrella sampling runs may be necessary. Such
a procedure makes use of the biasing potential \( U(\xi) \) as is. With such runs, one calculates the biased probability density:

\[
p^B(\xi) = \langle \delta [\xi - \sigma(r_1, \ldots, r_N)] \rangle_B.
\]  

(5)

The idea here is that if, as a result of an ABMD run, \( f(\xi) + U(\xi) \) is a constant, then the biased probability density \( p^B(\xi) \) will be flat. In a typical simulation, however, this is not the case. Using the results of the follow-up equilibrium runs, one can correct the calculated free energy surface via:

\[
f(\xi) = -U(\xi) - k_B T \ln p^B(\xi).
\]  

(6)

These procedures ultimately can yield high accuracy free energy curves, and have already been applied to a variety of different biomolecular systems [28, 29, 40–47].

SMD [30] is another nonequilibrium MD method that takes use of a time-dependent biasing potential and falls into the category of nonequilibrium driven simulations. This method is particularly useful for examining select pathways and mechanisms between two equilibrium states, as well as estimating the transition rates for these reactions [48]. The SMD method is based on the following notions. Consider a thermodynamic process that changes the system via a control parameter \( \xi \), that varies from \( \xi_0 \) to \( \xi_t \) over a time interval \( t \). Because of the second law of thermodynamics, the average work \( W \) performed on the system cannot be less than the free energy difference:

\[
\Delta F = f(\xi_t) - f(\xi_0) \leq W,
\]  

(7)

with the equality holding if the process is reversible. Thus, a nonequilibrium process can readily provide us with an upper bound for \( \Delta F \). In principle, the Jarzynski [64] equality that holds even when the transformation is irreversible can be used to estimate \( \Delta F \) if several finite-time nonequilibrium simulations are performed Although it is possible to improve the estimates of the calculated free energy differences using bidirectional methods such as Bennet’s acceptance ratio method [65–67], the computational cost of even a rough estimate of \( \Delta F \) may be quite substantial, and so SMD is often used as a qualitative simulation tool for the study of simulation pathways [43].

In a practical SMD setting, a restraining harmonic potential is typically added to the potential energy:

\[
U_{\text{SMD}}(\xi) = \frac{K}{2} \left[ \sigma(r_1, \ldots, r_N) - \xi \right]^2,
\]  

(8)

so that the system can be “steered” towards the states with the desired value of the collective variable \( \sigma(r_1, \ldots, r_N) \) by “pulling” on the other end of the spring \( \xi(t) \). This harmonic potential should be evaluated in the stiff-spring limit, with \( K \gtrsim U_{\text{max}}/(\delta \xi)^2 \), where \( U_{\text{max}} \) is the highest PMF barrier to be explored and \( \delta \xi \) sets the desired precision in \( \xi \). This stiff-spring approximation ensures that the PMF as a function of \( \xi \) of the unbiased original system is well represented by the free energy of the system with the harmonic restraint. The latter can be conveniently thought of as an approximation of the delta function \( \delta(\xi - \sigma(r_1, \ldots, r_N)) \) in the partition function. Generally, a number of non-equilibrium “pulling” and “pushing” simulations is then carried out by changing the values of \( \xi \) in a prescribed way.

The SMD method is especially easy to implement with one-dimensional collective variables: one simply repeats the same run with different initial configurations and/or random seeds, and collects the work. The Jarzynski equality (or bidirectional estimates if both forward and reverse runs are performed) may then be used to estimate the free energy difference.

Turning to systems described by more than one collective variable, the formalism additionally requires one to average over all possible paths that join the points \( \xi(0) \) and \( \xi(t) \). This is rather difficult to achieve for larger systems, and therefore only selected pathways are typically examined in a practical setting. For these systems, SMD can be also used as a tool to examine and compare pathways associated with different transition mechanisms in a qualitative way. One can always “hand-pick” pathways that appear to be the most interesting. Particular attention is typically given to the pathways that require the least amount...
of work, since these tend to dominate the sum for $\Delta F$ in the Jarzynski equality. Another approach for investigating mechanisms is to steer the system along selected pathways over the two-dimensional free energy landscapes, which have been selected by other means such as the Least Free Energy Path (LFEP) method [55].

**Collective Variables**

Both the ABMD and SMD methods require that one or more collective variables be specified. For the ABMD runs, we calculated free energy maps as a function of two collective variables, and used SMD to steer the system through the corresponding phase space. In addition, we carried out other SMD runs with different collective variables.

As a first set, we considered the collective variables of handedness [42] ($H$) and the radius of gyration ($R_g$). Handedness is a natural choice for investigating the B-Z transition, since B- and Z-DNA are right- and left-handed helical structures. Our definition of handedness is based on our previous work on polyproline peptides, and it discriminates between left- ($H < 0$) and right-handed ($H > 0$) helical structures [42]. For the DNA double helix, the position of the phosphorus (P) atoms of the backbone phosphate groups was found to be a good choice for the definition of handedness. In brief, the definition of handedness for a portion of DNA between the base pairs $n$ and $m$ makes use of a sequence of P atoms: $P_{n1}, P_{n2}, P_{n+11}, P_{n+12}, \ldots, P_{m1}, P_{m2}$, where the upper index indicates the strand number (1 or 2, labeled arbitrarily) and the lower index indicates the base-pair number labeled in the $5' \rightarrow 3'$ direction of strand 1. Note that this definition of handedness is independent of the labeling of the strands. Supplementary Figure 1 (right) shows the P atoms involved in the definition of handedness of a DNA segment between base pairs $n$ and $m$; the red and purple balls in this figure are the first and last elements in the sequence. Since the $5'$ terminal is missing the backbone phosphate group, the first and last element of the sequence are ignored in the definition of handedness. Therefore, the sequence used in the definition starts from $P_{11}$ instead of $P_{12}$, and ends with $P_{N1}$ instead of $P_{N2}$ ($N$ being the total number of base pairs). The position of these P atoms then defines the handedness via

$$H(p_1p_2p_3 \ldots p_n) = \sum_{i=1}^{n-3} H(p_i p_{i+1} p_{i+2} p_{i+3}),$$

in which each $p_i$ is a point in the sequence discussed above, and

$$H(ABCD) = \frac{AB}{|AB|} \times \frac{CD}{|CD|} \times \frac{EF}{|EF|}.$$  

In this last equation, the points $A, B, C, D$ define the vectors $\overrightarrow{AB}$ and $\overrightarrow{CD}$ and the midpoints of these vectors, called $E$ and $F$, in turn form the vector $\overrightarrow{EF}$. Supplementary Figure 1 illustrates this definition for the first term of the sum in the relation (9). The cross product of the unit vectors of $\overrightarrow{AB}$ and $\overrightarrow{CD}$ defines the (purple) vector whose dot product with the unit vector of $\overrightarrow{EF}$ forms the first term of the sum in the definition of handedness.

Even though the collective variable of handedness is enough to differentiate between B- and Z-DNA, many possible intermediate structures are characterized by the same value of $H$. To reduce this degeneracy, we also consider the free energy associated with the radius of gyration $R_g$ of the heavy atoms, which is defined as:

$$R_g = \sqrt{\sum_a \frac{m_a}{m_\Sigma} (r_a - R_\Sigma)^2}.$$  

Here, $R_\Sigma = \sum_a (m_a/m_\Sigma) r_a$ is the center of mass; $m_\Sigma = \sum_a m_a$, and the sums run over all atoms except hydrogen. The radius of gyration helps to discriminate between intermediate conformations with varying
compactness. We note that the collective variables \((H, R_g)\) also proved useful for the investigation of chiral transitions in other systems [42].

The collective variables of handedness and radius of gyration are very useful to study the B-Z transition. Notice that in contrast, the conventional variables to describe the conformations of DNA, do not necessarily make for practical or even relevant collective variables to be used in the biasing scheme. First of all, a relevant collective variable for this transition must describe the break of symmetry involved in the transition. This break of symmetry is described by the conventional base-pair step variable of twist. Handedness as defined here is closely related to the conventional step parameter twist, that is the in-plane counter-clockwise rotation of a base pair around the center of its C6-C8 axis with respect to the other base pair (see Ref. [53] for the technical details of the definitions of the conventional parameters as implemented in the 3DNA package [53, 54]). Twist is a useful parameter in identifying right-handed B-DNA from left-handed Z-DNA since it takes a positive (negative) value for right- (left-)-handed structures; in average about 35 and -30 degrees for B- and Z-DNA, respectively [68]. However, twist cannot be measured meaningfully if the base pair is broken. Handedness of a step, as defined here is the same as \(\sin(\text{twist})\) when both base pairs are intact and tilt, shift, and slide parameters of the step are negligible. We need the \text{sine} function in order to convert this into a quantity that may be summed in a meaningful manner and thereby remove the periodicity associated with the angles. However, there are two more important differences: (i) We define handedness based on the phosphate groups instead of the bases. Using bases is problematic due to the possibility of having broken base pairs. In this situation the twist becomes meaningless, as it is defined based on the planes of the bases. However, a phosphate-based twist is a meaningful quantity even in the presence of a completely broken base pair (the orientation of phosphate groups correlates with the overall twist). Furthermore, defining handedness based on the bases, in the context of ABMD and SMD simulations, means applying forces directly on bases that will often result in base pair disruption that may not be necessary for twisting the DNA. (ii) Handedness is a variable that combines twist with other step parameters including tilt, slide, shift, and rise. In the standard definition of twist \((\omega)\), the planes of two base pairs are aligned to remove the contribution of other step parameters \(\text{(i.e., shift, slide, rise, roll, and tilt)}\) before computing their twist (see Ref. [53]). However, our definition of handedness includes contributions from these other parameters (except roll). For instance, if tilt, shift, and slide parameters are negligible in a particular step, then \(H \approx \sin(\omega)\) numerically (ignoring the use of phosphate groups vs bases). If the step is tilted, one can write \(H \approx \sin(\omega) \cos(\tau/2)\) in which \(\tau\) is the tilt angle. If the tilt is small but there is a non-negligible shift \((Dx)\) or slide \((Dy)\), \(H \approx \sin(\omega) \frac{Dz}{\sqrt{Dx^2 + Dy^2 + Dz^2}}\) in which \(Dz\) is the rise step parameter.

Comparing to other conventional parameters, we note that handedness is perhaps more closely related to “helical twist”, \(\Omega\) (as implemented in the 3DNA package [53, 54]) in that combines \(\omega\) and \(\Gamma\) (net bending angle) according to [53] \(2 \cos(\Omega) = \cos(\omega) (1 + \cos(\Gamma)) - (1 - \cos(\Gamma))\) in which \(\Gamma = \frac{\tau^2 + \rho^2}{2}\), \(\rho\) being the roll angle. For \(\Gamma = 0\), step twist and helical twist are identical and the same as \(\sin^{-1}(H)\) (if slide and shift are small), but for \(\Gamma \neq 0\), both \(\Omega\) and \(\sin^{-1}(H)\) are smaller than \(\omega\) (in magnitude). Indeed, if slide, shift, and roll are negligible, the relation between \(\Omega\) and \(\sin^{-1}(H)\) as functions of \(\omega\) and \(\tau\) is very subtle: although their definitions are not identical, their difference is often negligible, numerically smaller than 4°, for any \(|\tau| < 60°\). Both handedness and helical twist are superior to step twist as a collective variable to induce a B-Z transition, because they not only induce the conformational change associated with the twist, but also preserve the double helix from bending-like deformations that characterize neither B- nor Z-DNA. Conventional step twist, on the other hand, is highly degenerate, as a single value of twist can describe many different structures. The bending-related degeneracies of handedness lie more in the intermediate regions (small \(H\) values). Particularly for SMD simulations, a set of collective variables that has less degeneracies around the target state and more degeneracies around the intermediate or transition states is considered more practical, since it allows the system to explore different transition pathways but the system will eventually arrive at the desired target. In this context, handedness and helical twist are favored over the step parameter twist. Finally, we note that handedness as defined here is also closely
related to the twist angle as defined in the Ref. [69] that is based on the position of coarse-grained sugar beads of a step \((A, B, C, D)\) and their midpoints \((E, F)\) in a coarse-grained model of DNA.

In addition to handedness, we can define other collective variables specifically associated with the B-Z transition. Since one of the major differences between the two DNA conformers is the \textit{anti} versus \textit{syn} conformation of the purine nucleotides in B- and Z-DNA, respectively, the glycosidic torsion angle \((\chi)\) which characterizes the relative base/sugar orientation may also be used to distinguish between the two DNA helices. This torsion angle is defined by \(O'_1-C'_1-N_1-C_2\) for pyrimidines and \(O'_1-C'_1-N_9-C_4\) for purines, as shown in the diagram of Supplementary Fig. 2a. Supplementary Figures 2b,c also show the bonds involved in the definition of \(\chi\) for some typical C-G base pairs in B- and Z-DNA conformations.

A torsion angle is called \textit{syn} if it is between \(-90^\circ\) and \(90^\circ\) and called \textit{anti} if it is outside this range. However in well-known DNA conformations \((i.e., A-, B-, Z-DNA)\) the torsion angle \(\chi\) (when expressed in the \(-180^\circ\) to \(180^\circ\) range) is always positive (negative) when in the \textit{syn} \((\text{anti})\) conformation, so that the sum of \(\sin(\chi)\) may be used to calculate the number of nucleotides in a \textit{syn} and \textit{anti} configuration, \(i.e.,\)

\[
X(i_1i_2\ldots i_n) = \sum_{k=1}^{n} \sin \chi_k.
\] (12)

in which \(i_k\) indicates the residue number. For A-DNA and B-DNA, the pyrimidine and purine nucleotides are both characterized by \(-160^\circ < \chi < -100^\circ\) \((-0.98 < \sin(\chi) < -0.34)\). In Z-DNA, the pyrimidine nucleotides are characterized by \(-160^\circ < \chi < -140^\circ\) \((-0.64 < \sin(\chi) < -0.34)\), while for the purine nucleotides, \(50^\circ < \chi < 65^\circ\) \((0.77 < \sin(\chi) < 0.91)\). For a single base pair, \(X\) is simply defined as \(X = \sin(\chi_{\text{purine}}) + \sin(\chi_{\text{pyrimidine}})\) which ranges from about \(-1.96\) to \(-0.68\) for a right-handed \(A-\) or B-DNA-like structure and ranges from about \(0.13\) to \(0.57\) for a left-handed Z-DNA.

In our simulations, we used the collective variable \(X\) to characterize the single base pair associated with the B-Z junction. It turns out to be fruitful to investigate \(X\) in conjunction with another collective variable \(N_{WC}\), which indicates if a given base pair has the correct Watson-Crick hydrogen bonding. Here \(N_{WC}\) for a given base pair is defined as:

\[
N_{WC} = \frac{1 - R^6}{1 - R^6}.
\] (13)

in which \(R\) is the RMSD (in Å) of the heavy atoms of guanine \((i.e., O_6,N_1,N_2)\) atoms and cytosine \((i.e., N_4,N_3,O_2)\) bases involved in the hydrogen bonding of a Watson-Crick base pair with respect to an ideal reference structure of a typical Watson-Crick pair. With this definition, \(N_{WC} \approx 1\) when the correct Watson-Crick base pair is formed, and \(N_{WC} \approx 0\) if the base pair is broken or not a Watson-Crick pair configuration.

Finally, to study the B-Z transition by means of SMD simulations, we define two additional collective variables which quantify the number of base pairs in a B or Z configuration. We define \(N_B\) as:

\[
N_B = \sum_{i=1}^{n-1} \frac{1 - (R_i/2)^6}{1 - (R_i/2)^6}.
\] (14)

Here, \(n\) indicates the number of the base pairs, the index \(i\) runs over the steps for which step \(i\) includes the base pairs \(i\) and \(i + 1\), and \(R_i\) is the RMSD in Å of all the heavy-atom positions in step \(i\) from a given reference B-DNA step. Here, the collective variable \(N_B\) is similarly defined using a Z-DNA reference configuration. Approximately, \(N_B\) and \(N_Z\) track the number of base pairs in B- and Z-DNA conformations.

**Simulation Details**

Simulations were carried out for the DNA sequences \((CG)_6\) \((a\ notation\ for\ d(5'-CGCGCGCGCGCG-3')_2)\), \((CG)_4\), and \((CG)_3A(CG)_3\ ,\ using\ an\ implicit\ water\ model\ based\ on\ the\ generalized\ Born\ approxima-
tion [49, 50] with surface area (GB/SA). The simulations used the ff99 version of the Cornell et al force field [51]. The leap-frog algorithm with a 1 fs timestep was used along with Langevin dynamics and a cutoff of 32 Å for the nonbonded interactions. Initial configurations were generated using the LEaP program of the AMBER v.9 simulation package [52] for B-DNA (for all three sequences) and 3DNA package [53, 54] for Z-DNA (for (CG)6 and (CG)4 sequences). These initial geometries were optimized using energy minimization (steepest descent method) and constraint MD at 30, 60, 90, . . ., 300 K, each stage for 5 ps. All the heavy atoms were restrained using a harmonic constant of 100 kcal/(mol Å2).

The constraint on the heavy atoms was then removed by decreasing the harmonic constant to 90, 60, 30, 20, 10, 5, 1 and 0 kcal/mol, each for 5 ps. To generate the initial configurations for B-Z and Z-Z junctions we used a different procedure based on the SMD method that will be explained below. All the MD simulations were run using the SANDER program of AMBER v.11 package.

For the (CG)6 duplex simulations, we ran 100 ns of regular MD at 300 K for both B- and Z-DNA. Both structures were stable during these runs, except for the two base pairs at the ends which were constrained slightly to avoid the breaking and fraying of the hydrogen bonds. This was achieved by constraining $N_{WC}$ to be unity for the first and last base pairs, using a harmonic constant of 10 kcal/mol. The average of these trajectories were used to construct reference B- and Z-DNA structures.

Next, we ran $(H, R_g)$ multiple walker ABMD simulations for the (CG)6 duplex using 24 replicas all at 300 K in three steps with $\tau_F = 1, 5, 25$ ps for 50, 25, 25 ns, respectively. The kernel width was set to $4\Delta \xi = 0.5$ Å for these simulations with $U$ being 1 for handedness and Å for radius of gyration. The initial structures were randomly selected from the last 50 ns of both B- and Z-DNA regular MD trajectories. The completed ABMD simulations were followed by 10 ns of umbrella correction runs. The kernel estimator method [62] was then used to estimate the biased probability density to estimate the error of the ABMD free energy maps and to correct the free energy estimates. For these runs, the $N_{WC}$‘s of all 12 base pairs were slightly constrained to take on a value of unity using a 1 kcal/mol harmonic constant. This constraint was used to avoid having broken base pairs persist for a long time, but was flexible enough to allow for the breaking or extrusion of base pairs over relatively short periods of time.

SMD simulations were performed in two different settings. The first involved moving in the $(H, R_g)$ free energy space, so as to obtain the LFEPs in this space. The second simulation made use of the $(N_B, N_Z)$ space, moving between B-DNA (11, 0) to Z-DNA (0, 11). Both simulations lasted 100 ns, and used a harmonic constant of 25 kcal/mol/Å2 (with $U$ being Å for $R_g$ and one for the rest). The $(H, R_g)$ based SMD simulation used the same constraints on $N_{WC}$ as the ABMD simulations.

For the (CG)4 duplex simulations, we ran 10 ns of regular MD at 300 K for both B- and Z-DNA. Both structures were stable over this runtime. The two end base pairs were constrained similarly to the (CG)6 runs. The equilibrated B-DNA conformation was used as an initial structure in a SMD simulation using $N_B$ and $N_Z$ as the collective variables. However, these collective variables were only defined on the base pairs 5 to 8 (i.e., moving from (3,0) to (0,3) in the $(N_B, N_Z)$ space via a time constraint of the form $\frac{1}{4}k[(N_B-3(1-t/T))^2+(N_Z-3t/T)^2]$ ) with the harmonic constant chosen to be 25 kcal/mol. The base pairs 1 to 4 were constrained by setting the collective variable $N_B$ to be around 3, using the same harmonic constant. In order to create the final B-Z junction, the resulting structure from the SMD runs was further constrained for 10 ns using $N_B = 3$ on the base pairs 1 to 4 and $N_Z = 3$ on the base pairs 5 to 8 using a harmonic constant of 25 kcal/mol. This made the initial structure for the (CG)4 B-Z junction between base pairs 4 and 5.

We then ran multiple walker ABMD simulations for the three (CG)4 structures obtained before: pure B-DNA, pure Z-DNA and the B-Z junction, using the collective variables $(X, N_{WC})$ defined on the base pair 4, using 24 replicas all at 300 K in 3 steps using $\tau_F = 1, 5, 25$ ps each running for 10ns. A kernel width of $4\Delta \xi = 0.2$ Å for $X$ and $4\Delta \xi = 1$ Å for $N_{WC}$ was used. The initial structures were randomly selected from the final 1 ns equilibration trajectories of B-DNA, Z-DNA, and B-Z junction simulations. The ABMD simulations were followed by 10 ns of umbrella correction runs.

For the (CG)3A(CG)3 duplex simulations, we ran 10 ns of regular unbiased MD at 300 K for the
initial structure of B-DNA. The average of this trajectory was used to make reference structures for B-DNA segments (segment I: base pairs 1 to 6 and segment II: base pairs 8 to 13). The reference structures for Z-DNA segments were made by averaging the equilibrium (CG)$_6$ trajectories of Z-DNA using base pairs 1 to 6 of (CG)$_6$ for segment I of (CG)$_3$A(CG)$_3$ and base pairs 7 to 12 of (CG)$_6$ for segment II of (CG)$_3$A(CG)$_3$. These reference structures were needed for our SMD simulations to generate B-Z and Z-Z junctions. The equilibrated B-DNA conformation was used as an initial structure for a 20 ns long SMD simulation using the $N_B$ and $N_Z$ collective variables defined only on segment II (base pairs 8 to 13) going from (5,0) to (0,5) with a harmonic constant of 25 kcal/mol. Segment I (base pairs 1 to 6) was constrained by keeping the collective variable $N_B$ around 5, with the same harmonic constant. The B-Z junction made above, was used in another 20 ns long SMD simulation to make a Z-Z junction in a similar process. Here $N_Z$ of segment II was kept around 5, and $(N_B, N_Z)$ of segment I was steered from (5,0) to (0,5) using the same harmonic constant as used above for all constraints. The final B-Z and Z-Z structures were equilibrated for a further 10 ns using different constraints on segments I ($N_B = 5$ for B-Z and $N_Z = 5$ for Z-Z) and II ($N_Z = 5$ for both B-Z and Z-Z).

Finally, we ran multiple walker ABMD simulations for three (CG)$_3$A(CG)$_3$ systems B-DNA, Z-Z and the B-Z junctions as obtained above with two distance collective variables: one defined between the O$_4$ atom of thymine residue and N$_6$ atom of adenine residue and the other defined between the N$_3$ atom of thymine residue and N$_1$ atom of adenine residue. We used 24 replicas all at 300 K in 3 steps using $\tau_F =$1, 5, 25 ps each running for 10ns. A kernel width of $4\Delta \xi = 0.5$ Å was used. The initial structures were randomly selected from the last 1 ns of B-DNA, Z-Z, and B-Z equilibration trajectories. The ABMD simulations were followed by 10 ns of umbrella correction runs.
References

62. Silverman BW (1986) Density Estimation for Statistics and Data Analysis. Monographs on statistics and applied probability. Chapman and Hall.

63. 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 110: 3533 – 3539.

64. Jarzynski C (1997) Nonequilibrium equality for free energy differences. Phys Rev Lett 78: 2690-2693.

65. Crooks GE (2000) Path-ensemble averages in systems driven far from equilibrium. Phys Rev E 61: 2361–2366.

66. Shirts MR, Bair E, Hooker G, Pande VyS (2003) Equilibrium free energies from nonequilibrium measurements using maximum-likelihood methods. Phys Rev Lett 91: 140601.

67. Kosztin I, Barz B, Janosi L (2006) Calculating potentials of mean force and diffusion coefficients from nonequilibrium processes without Jarzynski’s equality. J Chem Phys 124: 064106.

68. Sinden R (1994) DNA Structure and Function. Academic Press.

69. Tepper HL, Voth GA (2005) A coarse-grained model for double-helix molecules in solution: Spontaneous helix formation and equilibrium properties. The Journal of Chemical Physics 122: 124906.
Supplementary Figure 1. Definition of handedness. Phosphorus atoms (shown as balls) used to define the collective variable of handedness \( H \) for a segment between two DNA base pairs \( n \) and \( m \). Here, \( P_{ij} \) represents the phosphorus atom of the backbone phosphate group of the \( j^{th} \) base pair of strand \( i \). Labeling of the base pairs is in the 5'\text{ - }3' direction for strand 1, and the opposite for strand 2. This labeling of the helices is of course arbitrary, and may be interchanged, leaving \( H \) unchanged. Red, blue, and green arrows on the right represent the vectors involved in the definition of the first term of the sum in Eq. (9). The arrows on the left represent the unit vectors corresponding to those on the right. On the left, the cross product of the red and blue vectors gives the purple vector whose dot product with the green vector gives the first term of the sum (which is positive in this example corresponding to a right-handed turn).
**Supplementary Figure 2. Definition of glycosidic torsion angle and its different conformers.** (a) Schematic of the bonds (red lines) involved in the definition of the glycosidic torsion angle $\chi$, that involves $O_4' - C_1' - N_1 - C_2$ in pyrimidines and $O_4' - C_1' - N_9 - C_4$ in purines. (b),(c) Typical C-G base pairs for (b) B-DNA and (c) Z-DNA conformations. The bonds involved in the glycosidic dihedral angles for $\chi$ are highlighted for both the purine and pyrimidine nucleotides. The hydrogen atoms are not shown.

**Supplementary Table 1. Step parameters for different DNA conformations**

| system    | conformation | shift (Å) | slide (Å) | rise (Å) | tilt (°) | roll (°) | twist (°) |
|-----------|--------------|-----------|-----------|----------|----------|----------|-----------|
| (CG)$_4$  | B            | 0.15 (0.85) | 0.41 (2.32) | 3.68 (0.51) | 0.49 (3.61) | 3.50 (8.58) | 21.63 (11.05) |
|           | Z            | 0.09 (0.46)  | 2.66 (3.22) | 3.74 (0.51) | 0.43 (3.28) | -0.80 (12.13) | -23.34 (20.55) |
|           | B-Z          | 0.37 (1.75)  | 0.34 (2.01) | 3.20 (1.53) | -0.62 (6.61) | -1.51 (11.62) | B:28.12 (14.59) |
|           |              |           |           |          |          |          | Z:-24.45 (26.23) |
| (CG)$_3$A(CG)$_3$ | B            | 0.02 (0.42)  | 0.75 (0.36) | 3.41 (0.30) | 0.34 (5.55) | 5.84 (6.46) | 30.68 (5.27) |
|           | Z            | 0.11 (0.58)  | 2.00 (3.06) | 3.43 (0.38) | -0.15 (5.86) | 3.57 (3.92) | -26.23 (21.85) |
|           | B-Z          | 0.57 (1.13)  | 0.93 (0.97) | 3.08 (1.52) | 1.23 (5.86) | 1.96 (5.46) | B:32.21 (7.01) |
|           |              |           |           |          |          |          | Z:-21.79 (29.89) |

Average and standard deviation (bracketed) of six translational and rotational step parameters of different conformations of DNA, studied here. These quantities were measured using only the steps with WC base-pairing. Twist was estimated separately for the B- and Z-DNA segments of B-Z junctions.
Supplementary Figure 3. Comparison of two left-handed DNA conformations. Licorice (top) and surface (bottom) representations of (a) a left-handed structure with Watson-Crick-type backbone direction, observed in some of the SMD and ABMD simulations and (b) conventional Z-DNA. Note that we only show 8 base pairs, discarding one step from each end of a (CG)$_6$ DNA for illustrative purposes. The surface representation reveals the difference of the backbone direction (relative to the bases) for the two DNA types. Despite the similar WC-type direction between the observed structure and Z[WC]-DNA model, there are differences. For instance, there is no additional hydrogen-bonding between the guanine base and the adjacent phosphate group (as proposed in the Z[WC]-DNA model) and there is no Z-DNA-like zig-zag behavior.
Supplementary Figure 4. Behavior of the glycosidic torsion angle along two B-Z DNA transition paths. Glycosidic torsion angle (χ) as a function of time as obtained from SMD simulations of a 12 bp B-Z DNA transition for the purines (left column) and pyrimidines (right column). Panels (a) and (b) show results for the stretch-collapse mechanism and (c) and (d) for the base-extrusion mechanism. The glycosidic angles for the 3rd (red), 6th (blue) and 10th (black) nucleotides are plotted. In the stretch-collapse mechanism, there is no particular order for the change in χ: the guanines flip in the time frame between 40 and 80 ns, and the cytosines flip at some time between 25 and 50 ns and then slowly come back to the anti conformation. By contrast, in the base-extrusion zipper mechanism the jump from anti to syn for the guanines takes place in a zipper-like fashion while cytosines more or less stay in the anti conformation. Panels (e), (f) summarize the average behavior of χ for the stretch-collapse (solid line) and base extrusion (dashed lines) mechanisms. The average is taken over all 10 non-terminal purine (e) and pyrimidine (f) nucleotides.
Supplementary Figure 5. Correlation between the disruption of base pairs and flipping of glycosidic torsion angles. Time series of the $N_3^G - N_1^G$ distance (top) and glycosidic torsion angle of the guanine nucleotide (bottom) as obtained from SMD simulations for the base-extrusion mechanism (see Supplementary Fig.4). The $N_3^G - N_1^G$ distance monitors base-extrusion and $\chi_G$ monitors flipping events. The time series corresponding to the 3rd (red), 6th (blue) and 10th (black) base pairs are plotted. In some cases, our snapshots catch the base pairs in a completely extruded conformation (base pair 3) but sometimes the extrusion occurs faster than the saved time frames, and we just observe a half-way extruded conformation (base pair 6) but the flipping occurs at the peak of the extrusion, in all cases.