

UC Irvine

UC Irvine Previously Published Works

Title

Force-Field-Dependent DNA Breathing Dynamics: A Case Study of Hoogsteen Base Pairing in A6-DNA.

Permalink

<https://escholarship.org/uc/item/0m53j0xs>

Journal

Journal of Chemical Information and Modeling, 62(24)

Authors

Stone, Sharon

Ray, Dhiman

Andricioaei, Ioan

Publication Date

2022-12-26

DOI

10.1021/acs.jcim.2c00519

Peer reviewed

Force-Field-Dependent DNA Breathing Dynamics: A Case Study of Hoogsteen Base Pairing in A6-DNA

Sharon Emily Stone, Dhiman Ray, and Ioan Andricioaei*



Cite This: *J. Chem. Inf. Model.* 2022, 62, 6749–6761



Read Online

ACCESS |



Metrics & More

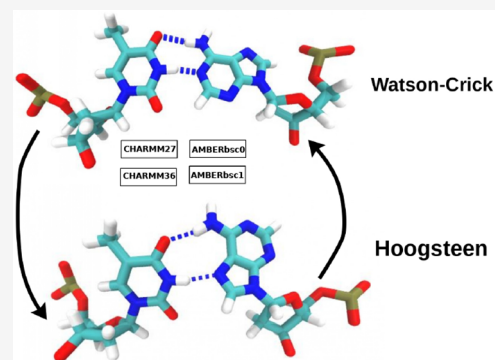


Article Recommendations



Supporting Information

ABSTRACT: The Hoogsteen (HG) base pairing conformation, commonly observed in damaged and mutated DNA helices, facilitates DNA repair and DNA recognition. The free energy difference between HG and Watson–Crick (WC) base pairs has been computed in previous studies. However, the mechanism of the conformational transition is not well understood. A detailed understanding of the process of WC to HG base pair transition can provide a deeper understanding of DNA repair and recognition. In an earlier study, we explored the free energy landscape for this process using extensive computer simulation with the CHARMM36 force field. In this work, we study the impact of force field models in describing the WC to HG base pairing transition using meta-eABF enhanced sampling, quasi-harmonic entropy calculation, and nonbonded energy analysis. The secondary structures of both base pairing forms and the topology of the free energy landscapes were consistent over different force field models, although the relative free energy, entropy, and the interaction energies tend to vary. The relative stability of the WC and HG conformations is dictated by a delicate balance between the enthalpic stabilization and the reduced entropy of the structurally rigid HG structure. These findings highlight the impact that subtleties in force field models can have on accurately modeling DNA base pair dynamics and should stimulate further computational investigations into other dynamically important motions in DNA.



The relative stability of the WC and HG conformations is dictated by a delicate balance between the enthalpic stabilization and the reduced entropy of the structurally rigid HG structure. These findings highlight the impact that subtleties in force field models can have on accurately modeling DNA base pair dynamics and should stimulate further computational investigations into other dynamically important motions in DNA.

INTRODUCTION

Based on the pioneering X-ray crystallography results of Rosalind Franklin¹ in 1953, Watson and Crick proposed the structure of duplex DNA,² a model later popularized as the Watson–Crick (WC) structure. By using the bases in their most plausible tautomeric form, they determined complementary base pairs (bps) between adenine and thymine (A–T) as well as guanine and cytosine (G–C).² When the following WC base pairings are assembled in a right-handed helix structure, the overall geometry is referred to as B-DNA. Almost 50 years after this discovery, Abrescia et al. have reported a different kind of double-helical structure, from the regulatory regions of DNA, involving a more uncommon Hoogsteen (HG) base pairing conformation.³ In relatively recent nuclear magnetic relaxation (NMR) $R_{1\rho}$ experiments of free B-DNA, Nikolova et al. showed that both A–T and protonated G–C (G–C⁺) WC bps spontaneously formed the HG structure, with a population and lifetime of 0.1–1% and 0.3–1.5 ms, respectively.⁴ Additionally, HG A–T bps are more abundant compared with G–C⁺ bps under solution conditions.

The HG base pairing form was initially observed in isolated nucleotide crystals in the late 1950s using heavy atom X-ray diffraction.⁵ HG structures form when the purine base (adenine or guanine) is flipped over 180° with respect to the glycosidic bond, resulting in a *syn* conformation (as opposed to *anti* in WC), resulting in the shortening of the helix diameter (Figure 1). This structural change creates a chemical

environment different from canonical WC bps and plays a key role in binding to proteins^{6,7} and to small molecule drugs.⁸

HG bps have also been observed in other scenarios, including damaged DNA,^{9–11} DNA duplex conjugated with the DNA repair machinery,^{6,12,13} and other diverse DNA–protein interactions.^{4,6,14–16} DNA repair processes that do not go up to completion can lead to cytotoxic damage, potentially resulting in cancer or cell death. An example is the methylation of DNA, which has been shown to cause blocking in DNA replication. Recent work by Xu et al. showed that WC bps are vulnerable to methylation in double-stranded DNA, and the sensitivity to methylation increases for the HG bp.⁷ Alternatively, a methylated adenine base is more prone to form a HG bp as a result of stabilizing steric interactions.¹⁷ Additionally, HG bps have been hypothesized to play an important role in DNA repair mechanisms. For instance, the crystal structure of polymerase- ϵ (hPol- ϵ) revealed a replication mechanism dominated by HG bps.¹⁸ The structure of hPol- ϵ encourages replication via the N²-adducted G, which forms a

Special Issue: Advancing Women in Chemistry

Received: May 1, 2022

Published: September 1, 2022



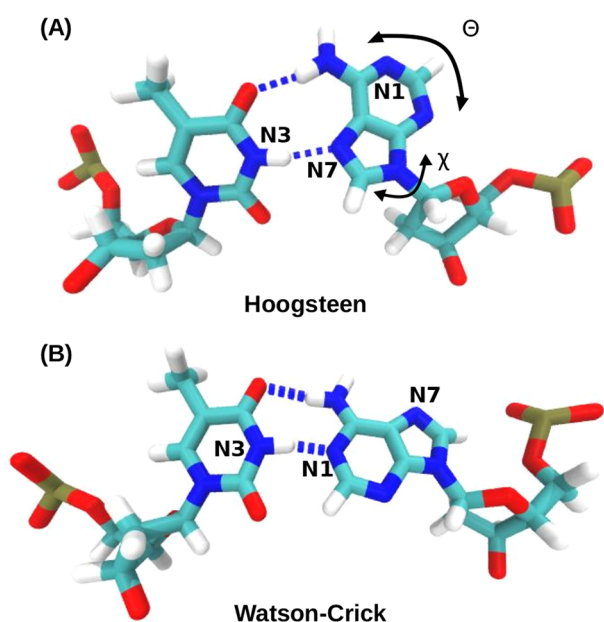


Figure 1. Conformation of (A) HG and (B) WC A-T bps in B-DNA. In the HG bp, the adenine base (on the right side) is flipped over 180° with respect to the glycosidic bond, which results in hydrogen bonding with the nitrogen in the five-membered ring, and shorter helix diameter. The N1, N3, and N7 nitrogen atoms, referred to in the text, are also marked in the figure.

syn conformation in DNA, and in turn facilitates a binding mechanism through the minor-groove adducts.¹⁸ The HG bp stabilizes the tumor suppressor p53-DNA complex and enhances its binding affinity.¹² Binding affinity of small molecules (like intercalation agents and anticancer drugs such as echinomycine) to nucleic acids is also increased in the vicinity of a HG bp.¹⁹ These examples demonstrate the biochemical significance of the HG configuration in nucleic acids and encourage detailed studies to understand the mechanism of formation of such noncanonical base pairing structures.

Molecular dynamics (MD) simulation can be used to understand the mechanism of biomolecular processes in atomistic detail. The HG base pairing has also been thoroughly investigated using classical MD simulation over the past two decades. The initial set of works by Cubero et al. focused on the structure and stability of the antiparallel helix solely consisting of HG bps²⁰ and the WC-HG junctions.²¹ The energetics of HG base pairing have also been explored using extensive quantum chemical studies by Hobza and co-workers,^{22–25} leading to the discovery that in gas phase, the HG bp is more stable than the WC conformation.

From experimental NMR relaxation studies, the time scale of the transition between WC and HG base pairing process has been reported to be 50–250 ms, which is beyond the reach of conventional MD simulation with currently available computing facilities. Therefore, the process of transition between these two conformations could only be studied using enhanced sampling methods including umbrella sampling,²⁶ metadynamics,^{27,28} and meta-eABF.²⁹ Path sampling techniques such as transition path sampling (TPS)³⁰ have also been employed to sample the conformational space of the WC to HG bp transition. Yang et al. computed the free energy landscape of the WC to HG transition in A6-DNA, using umbrella sampling

and multiple walker well-tempered metadynamics, and captured multiple transition pathways and observed the spontaneous changes in base flipping.^{31,32} Similarly, studies from our group used a fast converging method called meta-eABF²⁹ to obtain the potential of mean force (PMF) for the same transition process with CHARMM36 force field parameters.³³ Our free energy landscape revealed a stable open bp conformation adjacent to the WC base paired state. The mechanistic analysis using Markov state modeling^{34,35} also revealed a melted bp state to be the primary intermediate in the transition.³³ This highly stable extra-helical conformation is inconsistent with the conventional idea about the stability of DNA and provokes further discussion on the accuracy of force fields in describing the base pairing intermediates and transition pathways. Path sampling studies by Vreede et al.³⁶ and Hooft et al.³⁷ also revealed that partial opening of the A-T bp is necessary for the conformational switching, but no stable open base paired intermediate conformation was observed.

The force field is an important component of MD simulations, since the accuracy of our results is dependent on the accuracy of the force field model. The force field is the functional form and the parameter set used to represent the potential energy function of a biomolecule in the context of classical MD simulation.³⁸ Gradients of this potential energy function are used to propagate dynamical trajectories by solving Newton's equation of motion.³⁹ For nucleic acids, such as DNA and RNA, the commonly used nonpolarizable force fields are the CHARMM (CHARMM27, CHARMM36) and AMBER (AMBERbsc0, AMBERbsc1) family models.⁴⁰ Different models have different refinements in order to improve the description of nucleic acid properties. For example, the CHARMM27 model has refined backbone dihedrals, sugar puckering, and glycosidic linkage to match the results of model compounds computed with *ab initio* calculations.⁴¹ But it can model the groove width, population of the BI backbone in DNA, and opening of WC bp with limited accuracy. For CHARMM36, improvements were focused on refining backbone dihedral angles, sugar puckering, and the 2-hydroxyl dihedral for RNA parameters.⁴⁰ In the AMBER family, AMBERbsc0 was focused on refining the backbone conformations (α and γ dihedrals), in order to lower overpopulated α gauche+ and γ trans conformations.^{42,43} Additionally, the AMBERbsc1 model was constructed with AMBERbsc0, but with refined sugar puckering and χ , ϵ , and ζ dihedrals.⁴⁴ These refinements have become an essential process in order to obtain accurate results in the microsecond time scale. Comparative studies of the common nucleic acid force fields have been performed to evaluate the mesoscale properties of DNA⁴⁵ and to reproduce NMR structures of the Drew–Dickerson dodecamer (DDD) system.⁴³ Minhas et al. found the CHARMM27 model to have the most stable trajectories and best agreement to the experimental data for describing the mesoscale properties of multimicrosecond simulations of long DNA fragments (40 bps) in a cubic solvation box of TIP3P water.⁴⁵ Additionally, the CHARMM27 model had the best agreement with X-ray diffraction and NMR results compared with CHARMM37, AMBERbsc0, and AMBERbsc1 models. While the study revealed both AMBER family models to have good agreement with the experimental data, the CHARMM36 model had unstable trajectories in the microsecond time scale and produced irreversible fluctuations.⁴⁵

Whether the choice of force field can affect the pathways and free energy landscape of the conformational transition between the WC and HG base pairing forms is not yet clear. But such understanding is vital, considering the important physiological role of the HG base pairing in DNA repair and replication. In the current work, we address this question by comparing the underlying free energy landscape of WC to HG transition, and the energetic and entropic stabilization of individual base pairing states for the four different force fields: CHARMM27, CHARMM36, AMBERbsc0, and AMBERbsc1. Although newer nucleic acid force fields are available,^{43,46} our choice of force fields was inspired by the recent comparative study by Minhas et al.⁴⁵ and MD studies of HG bp formation in the literature. The free energy surface and associated thermodynamic properties of the HG conformation of the A16-T9 bp in A_c-DNA⁴ are explored through enhanced sampling simulation using the advanced meta-eABF approach. The potentials of mean force (PMF) obtained from the meta-eABF simulations are used to compare the force field models in terms of the minimum free energy pathway for the WC-HG transition. Presence of HG and WC bps is observed in all the four models, although the relative stability of each base pairing configuration can be somewhat different. To obtain a fundamental rationale behind this discrepancy, we computed the quasi-harmonic entropy and the root-mean-square fluctuations (RMSF) of the two base pairing conformations from equilibrium trajectories for all force field combinations. We observe that indeed there is a large variation of the free energy, enthalpy, and entropy of HG formation among the different force fields, although the predicted structures of the HG and WC conformation are almost identical. Based on our findings, we explain the main distinctions between each force field in describing the molecular details of WC and HG base pairing conformations and compare our findings with experimental results.

COMPUTATIONAL METHODS

System Preparation and Equilibration. We performed our simulations on the A6-DNA fragment (PDB ID: SUZF)⁴⁷ reported in the NMR relaxation study by Nikolova et al.⁴ This structure includes six A-T bps in a 12 bp sequence (CGATTTTGGC). The topology files for the CHARMM and AMBER family force fields were prepared using Visual Molecular Dynamics (VMD)⁴⁸ and AmberTools (tleap)⁴⁹ respectively. The initial structure of A6-DNA was placed in a solvated box with TIP3P water and 22 Na⁺ ions to neutralize the overall system. CHARMM-TIP3P force field⁵⁰ is used to model the water and ions when the CHARMM family force field is used for nucleic acids, and the TIP3P model⁵¹ is used for water and ions in the simulations with AMBER force field. We did not add additional salt in the system, as previous computational studies under neutralizing ion concentration could most accurately reproduce the experimental population ratio between WC and HG conformations.³² Also, the NMR experiments were performed at 25 mM ionic strength, which should correspond to 10 ions in our simulation box while we already have more than 20 neutralizing Na⁺ ions. A water padding of 17 Å was included in each direction. Each solvated system comprised of ~26,000 atoms. In case of the CHARMM family force fields, a 12 Å cutoff was used for nonbonded forces along with a switching function with switch distance of 10 Å. For AMBER family force fields, a 9 Å cutoff was used for nonbonded forces without any switching function, following the recommended protocol.⁵² The structures were minimized

for 50,000 steps using the conjugate gradient algorithm, followed by a slow heating to reach 298 K temperature using a velocity rescaling thermostat with periodically updating the temperature at a rate of 1 K/ps. During the slow heating, harmonic restraints of 3 kcal mol⁻¹ Å⁻² were applied on the heavy atoms. These restraints were then removed over a period of 1.2 ns, at a rate of 0.5 kcal mol⁻¹ per 200 ps. Then, the system was placed in a NVE ensemble for 3 ns with harmonic restraint with force constant of 0.1 kcal mol⁻¹ Å⁻², followed by an unrestrained NPT simulation for 10 ns. The terminal bps were restrained with a force constant of 0.05 kcal mol⁻¹ Å⁻² for all equilibration, production, and enhanced sampling simulations. The temperature was kept constant at 298 K using the Langevin thermostat with a coupling constant 1 ps⁻¹. The pressure was controlled to be at 1 atm, using the Nosè–Hoover Langevin piston with piston decay of 50 fs and oscillation period of 100 fs.^{53,54} An identical protocol was used for all four force field models. All MD simulations reported in this study have been performed using the NAMD 2.14 software package.⁵⁵ The input files for all simulations are provided with the manuscript (see [Data and Software Availability](#) section).

Enhanced Sampling Simulation. To accelerate the rare transitions between the WC and HG form, we performed enhanced sampling simulations using a combination of metadynamics²⁷ and extended-system adaptive biasing force (eABF),⁵⁶ commonly known as the meta-eABF approach.²⁹ Meta-eABF simulations were performed using the colvars module⁵⁷ patched with the in NAMD 2.14 simulation package. The two-dimensional (2D) PMF was calculated along the following collective variables (CV): the glycosidic angle χ and a pseudodihedral angle Θ of the A16-T9 bp.³¹ The Θ angle characterizes the base flipping into the major and minor grooves of DNA, while the χ dihedral angle characterizes the rotation of the adenine base with respect to the deoxyribose sugar (O4', C1', N9, and C4). A pictorial representation of these two torsion angles are provided in ref 32. We started the meta-eABF simulations from the end points of NPT equilibration and propagated for ~200–400 ns, until the PMFs were sufficiently converged. Gaussian hills of height 0.06 kcal/mol and width 3 CV unit (15°) were deposited every 2 ps for the metadynamics part. Additionally, the ABF bias was applied against the average force after 1000 samples were collected for the corresponding bin of width 5°. The MEPSA program⁵⁸ was used to obtain the minimum free energy pathway from the 2D energy landscape obtained from enhanced sampling simulation. The pathway is predicted as a smooth curve joining various nodes on the free energy surface using the “node-by-node” search algorithm,⁵⁸ which bears similarity with the Dijkstra algorithm.⁵⁹ The start and the end point of the interpolated pathway is manually chosen to be, respectively, the minimum of the WC basin and the minimum energy point approximately at $\chi = 180^\circ$. This is done to avoid any spurious interpolations resulting from the inability of the algorithm to take into consideration the periodicity of the collective variable.

Equilibrium Simulations for Entropy and Energy Calculation. Structures of the WC and HG base pairing conformations were sampled from the meta-eABF trajectories based on the hydrogen-bond donor–acceptor distances. A total of eight structures for four different force fields were used as starting points for 100 ns unbiased MD simulations with identical simulation setup as the meta-eABF simulation. The

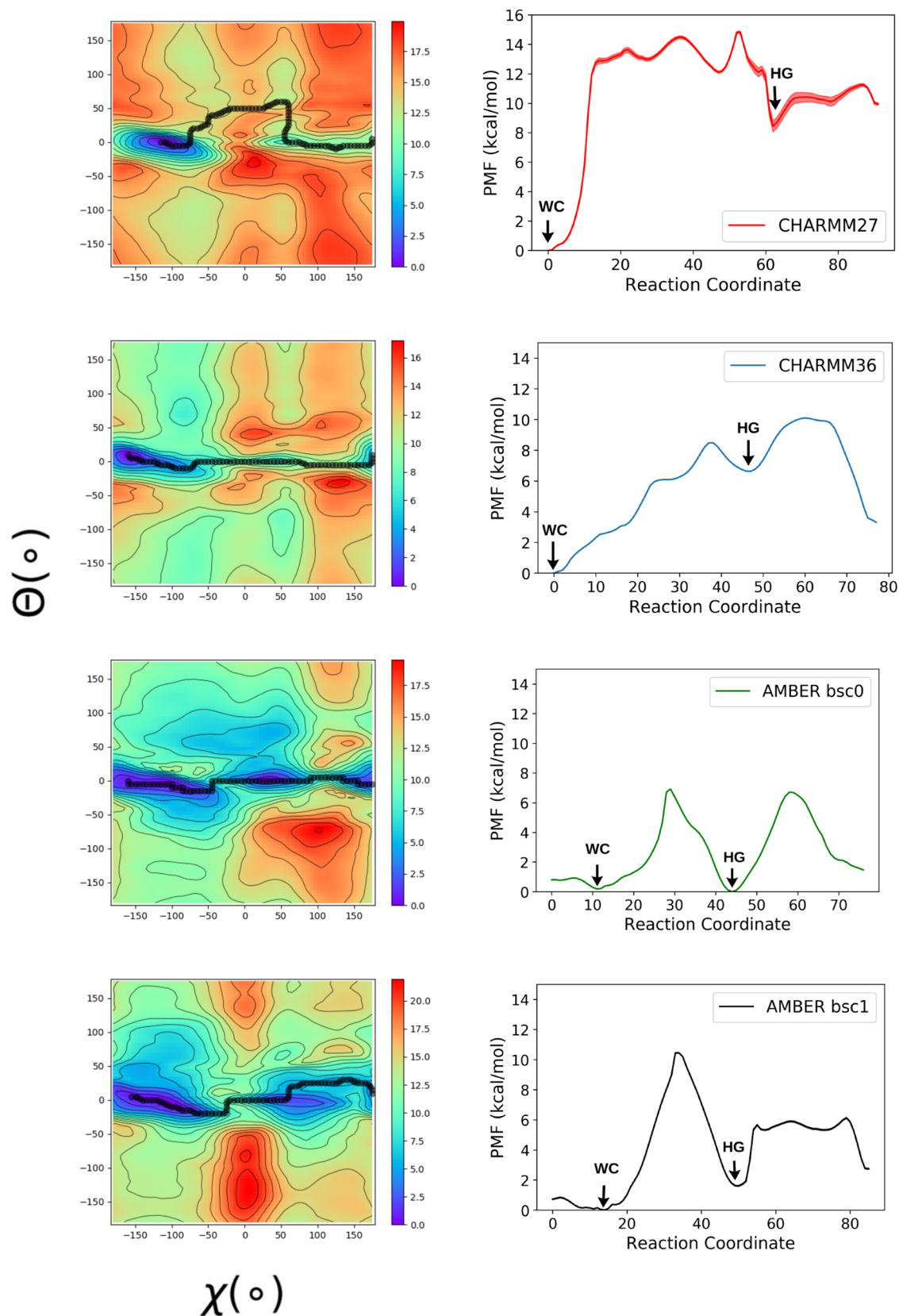


Figure 2. Free energy surface of the A16-T9 bp in A6-DNA was obtained using meta-eABF method for different force field models. The error bars obtained from the last 50 ns of the simulation using MEPSA software are indicated in the one-dimensional plots. But they are too small to distinguish from the plot, except for the CHARMM27 force field.

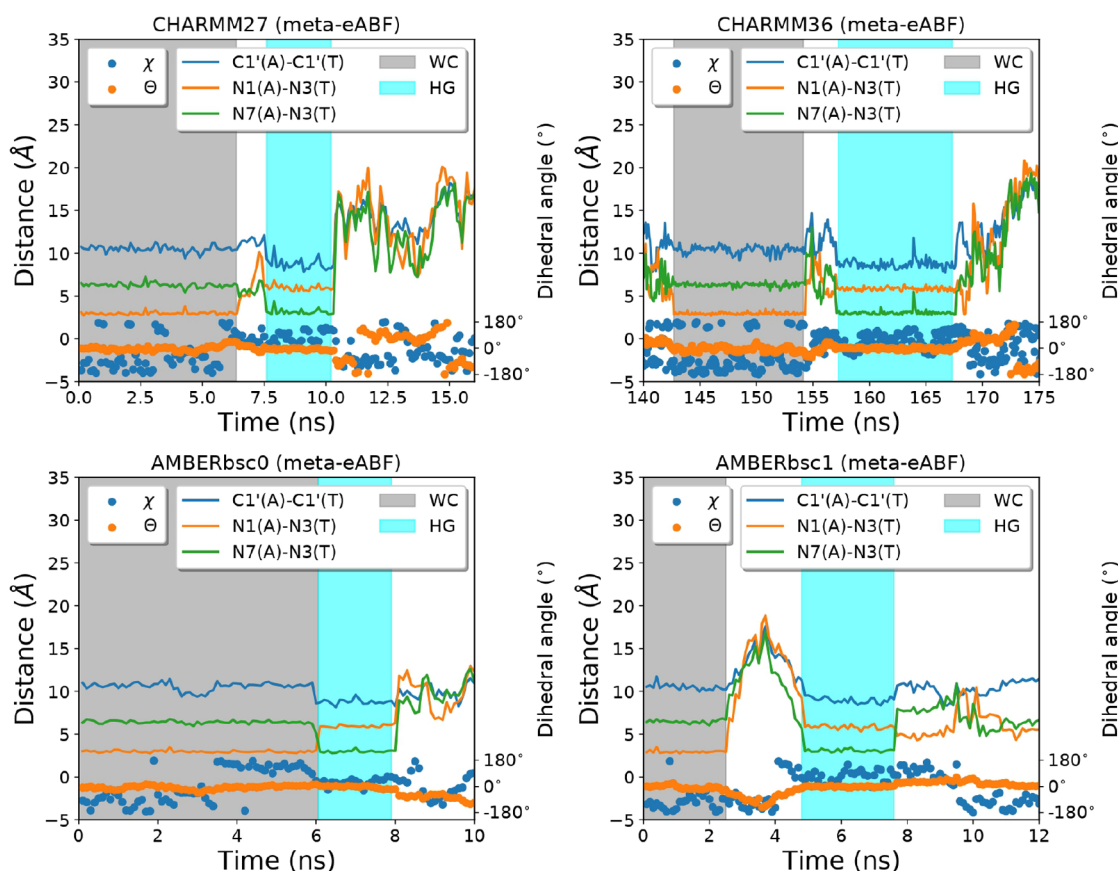


Figure 3. Representative regions of the meta-eABF trajectories showing transitions between WC and HG conformations. The hydrogen-bond donor–acceptor distance, the helix diameter measured as the C1'–C1' distance, and the two torsion angles χ and θ are depicted as a function of time.

coordinates were saved every 10 ps. The hydrogen-bonding donor distances between atoms N1(A)–N3(T) and N7(A)–N3(T) were monitored to ensure WC and HG conformations remained intact throughout the simulation (see [Supporting Information](#)). These trajectories were used to calculate the root-mean-square fluctuation (RMSF) and quasi-harmonic entropy.⁶⁰ Five 50 ns trajectory segments were extracted from the last 90 ns of each trajectory using overlapping time windows: 10–60 ns, 20–70 ns, 30–80 ns, 40–90 ns, and 50–100 ns. The entropy and RMSF were computed from each segment, and the mean and 95% confidence intervals are reported for the five data points.

The RMSF of all heavy atoms were computed using GROMACS 2018.1 software⁶¹ with the `gmx rmsf` tool. Details of the calculation of quasi-harmonic entropy from the covariance matrix of atomic fluctuation is provided in the original paper.⁶⁰ In brief, the entropy is given by

$$S = k_B \sum_i^{3N-6} \frac{\hbar\omega_i/k_B T}{\exp(\hbar\omega_i/k_B T) - 1} - \ln(1 - \exp(\hbar\omega_i/k_B T)) \quad (1)$$

where k_B is Boltzmann constant, \hbar is $\hbar/2\pi$ where \hbar is Planck's constant, and T is absolute temperature. The frequencies ω_i are obtained as $\omega_i = \sqrt{k_B T/\lambda_i}$, for $i = 1, 2, 3, \dots, 3N-6$, where λ_i are eigenvalues of the mass weighted covariance matrix σ obtained from MD simulation. The matrix σ is given by $\mathbf{M}^{1/2}\sigma'\mathbf{M}^{1/2}$, where \mathbf{M} is the mass-matrix (diagonal matrix

containing the masses of the different atoms) and σ' is the covariance matrix of atomic coordinates. The elements of σ' are given by

$$\sigma'_{ij} = \langle (x_i - \langle x_i \rangle)(x_j - \langle x_j \rangle) \rangle \quad (2)$$

where x_i is the i th component of the $3N$ dimensional Cartesian coordinate space of N atoms, and $\langle \rangle$ indicates ensemble average. The number of atoms involved in the calculation (N) equals to the number of heavy atoms in the DNA. The entropy is dominated by slow vibrational modes, so we excluded the fast moving hydrogen degrees of freedom from our analysis. The covariance matrices and their eigenvalues were computed using the `gmx covar` module of GROMACS 2018.1 package.⁶¹ The entropy was then evaluated using eq 1 by an in house python code. Nonbonded interaction energies between bases and nucleotides were calculated using the `NAMDenergy` module of VMD.⁶²

RESULTS AND DISCUSSIONS

Free Energy Surfaces. Similar to our previous study,³³ the meta-eABF approach could produce the PMFs within a few hundreds of nanoseconds, at significantly less computational cost compared to other approaches,^{31,32} facilitating the comparative study between multiple force fields. The free energy surfaces obtained for different models are shown in [Figure 2](#). In all models, the WC base paired configurations were present near $\Theta \sim 0^\circ$ and $\chi \sim -100^\circ$, and the HG form

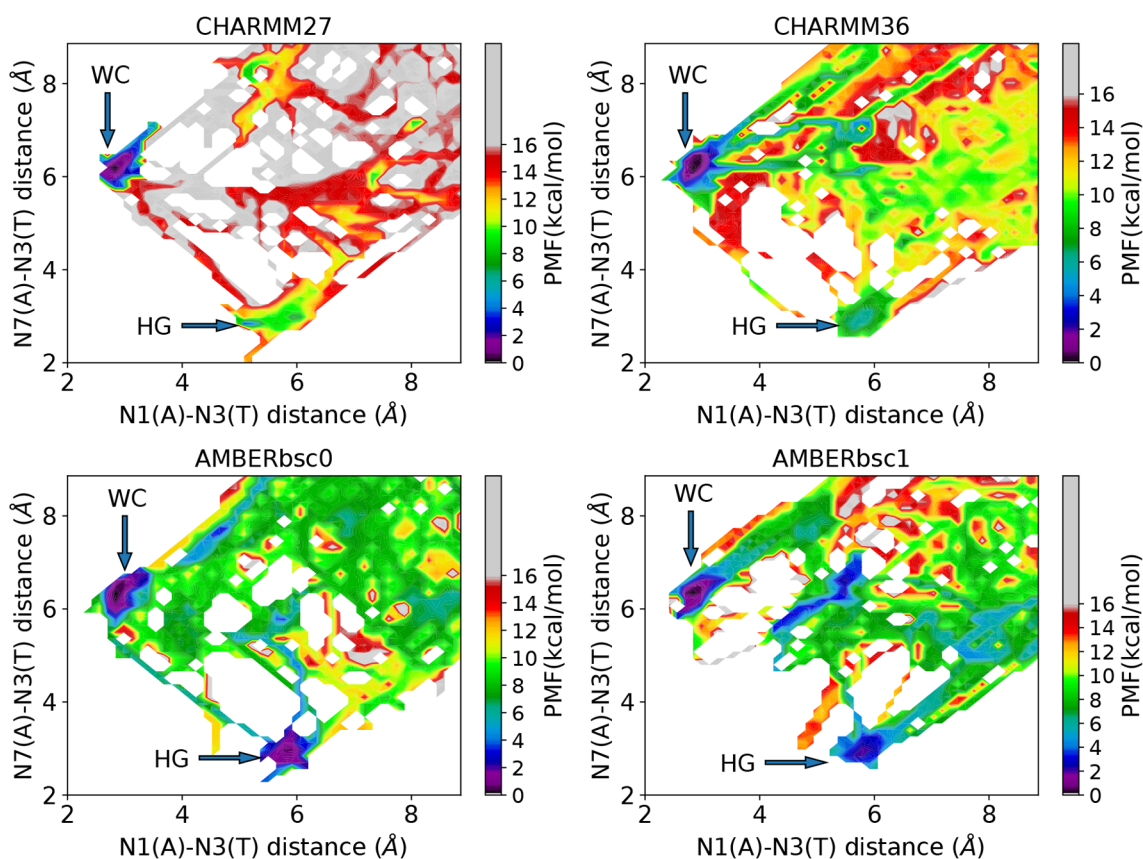


Figure 4. Free energy landscape of the conformational switching between WC and HG bp configurations, projected onto the two hydrogen-bond donor–acceptor distances, for different force field models.

near $\Theta \sim 0^\circ$ and $\chi \sim 60^\circ$. The free energy difference between the WC and HG bp ($\Delta G_{WC \rightarrow HG}$) was determined to be ~ 8 kcal/mol for CHARMM27, ~ 6 kcal/mol for CHARMM36, ~ 1 kcal/mol for AMBERbsc1, and ~ 0 kcal/mol for AMBERbsc0. When comparing the NMR value, 3.3 kcal/mol, the free energy difference is overestimated in the CHARMM family force fields, underestimated in both AMBER family models, and closest in agreement to the CHARMM36 and AMBERbsc1 force field. These discrepancies are much larger compared to the results obtained by Yang et al.^{31,32} using more expensive umbrella sampling²⁶ and multiple-walker well tempered metadynamics²⁸ simulations and can possibly indicate an artifact of the meta-eABF approach which converges faster but only provides a qualitatively accurate free energy landscape for systems with high energy barrier.⁶³ Also, the uncertainty of the free energy landscapes, we computed, is at least ~ 1 kcal/mol as apparent from the convergence plots (see [Supporting Information](#)).

All the four free energy landscapes indicate that the opening of the bps through the major groove region is more energetically favorable compared to the minor groove region, especially when starting from a WC structure, which is the most prevalent form in solvated DNA duplex. This pathway is in agreement with the previous experimental⁴ and simulation studies.^{33,36} Except for CHARMM36, all three force fields show that the base paired structures (WC or HG) are more stable than the open base paired configurations. In CHARMM36 we can observe an open bp state in the major groove region near the WC form which is lower in free energy in comparison to the HG state. This result is consistent with

the previous work, where a very stable open base paired intermediate state between WC and HG was captured using CHARMM36 force field and Markov state modeling.³³

Additionally, for the majority of models, the minimum free energy pathway reflected in the transition mechanism is a direct 180° flip-over of the adenine glycosidic angle, whereas the CHARMM27 force field shows the minimum free energy pathway involves a slight opening of the bp, indicating a possible extra-helical mechanism. From previous analysis of the T9-A4' bp in sequence 5'-CGATTTTTTGGC-3', it was concluded that the extra-helical mechanism, where the purine base flips out of the DNA helix first, was preferred for the HG to WC bp transition.³⁶ Although this mechanism was not the most energetically favorable in CHARMM36, AMBERbsc0, and AMBERbsc1, the metastable open bp conformations observed in those force fields allows for the possibility of conformational switching between WC and HG form via an extra-helical intermediate. For all the force fields, shallow free energy regions are detectable in the major groove ($\Theta > 0$) and minor grooves ($\Theta < 0$). In the CHARMM36 and AMBERbsc1 force fields, the open bp state, adjacent to the WC minima, is clearly distinguishable.

In our previous study,³³ we suggested that the dihedral angle based reaction coordinates may not be able to uniquely distinguish between WC and HG states, as some structures which do have the necessary hydrogen bonds for bp formation can also be assigned to either of these two categories just from the χ and θ torsion angle values. Here, we propose to use the hydrogen-bond donor–acceptor distances to identify the base paired forms. In WC base pairing, an N–H–N type hydrogen

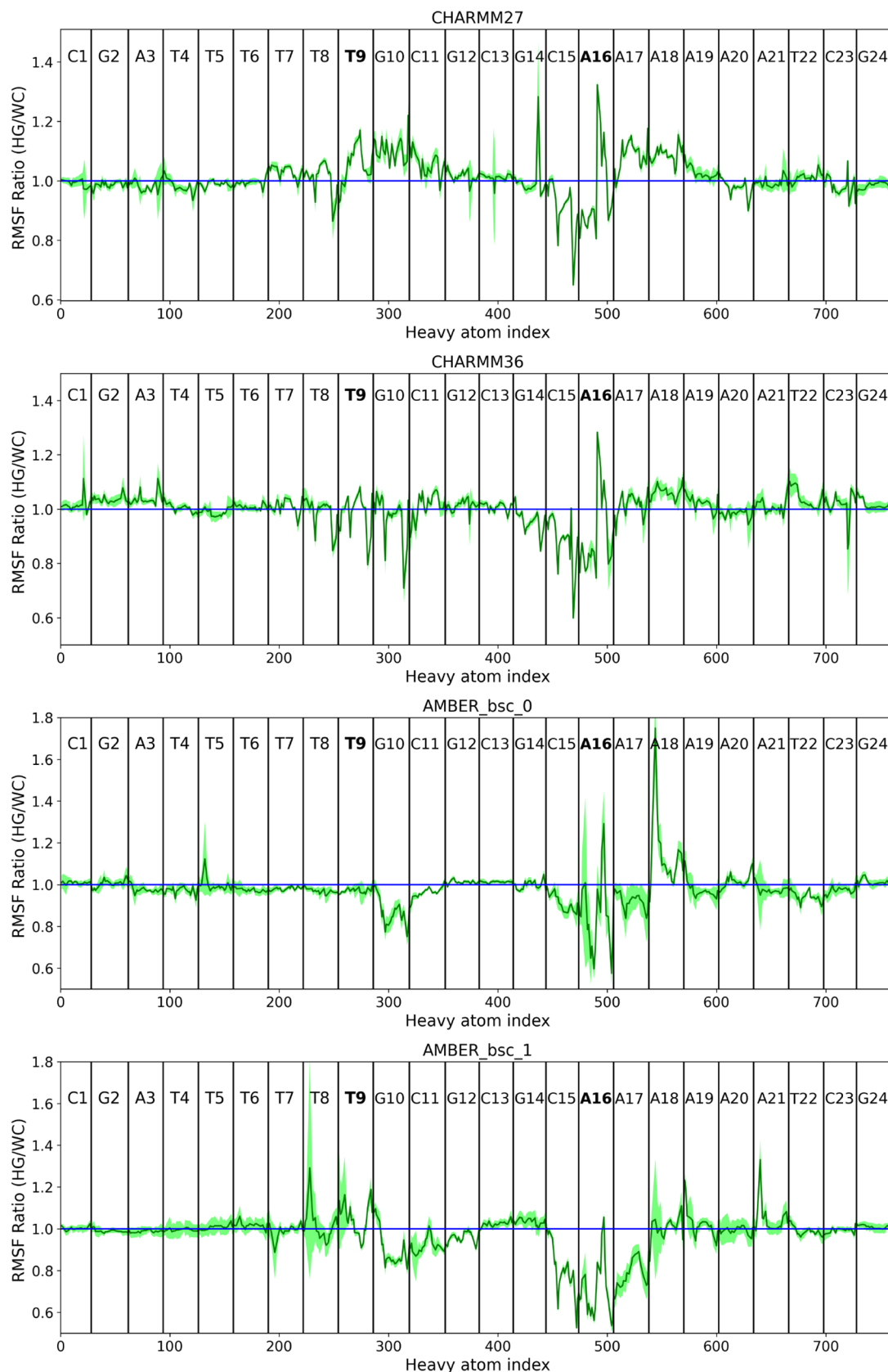


Figure 5. RMSF ratio of the DNA duplexes involving the HG to WC form of the A16-T9 bp. The location of all the bps are indicated by vertical lines. The A16 and the T9 nucleotides have been labeled in bold font. The error bar computed as the 95% confidence interval of the five 50 ns trajectory segments is indicated as the light green shading.

bond forms between the N1 nitrogen of adenine and the N3 nitrogen of thymine nucleotides (N1(A)-N3(T)). In the HG form, the N3(T) forms a hydrogen bond with N7(A) instead. So, by measuring the distance between the N1(A)-N3(T) atom pair and N7(A)-N3(T) atom pair the WC and HG states can be uniquely identified. This is reflected in the time series plots of dihedral angles and hydrogen bond donor–acceptor distances in Figure 3. We projected our free energy landscapes from Figure 2 in the hydrogen bond donor–acceptor distance space (Figure 4). All four models clearly distinguish the WC and HG states as free energy minima in the configurational space. The relative free energies of the HG bps are in slightly better agreement with the experimental results (CHARMM27: ~ 6 kcal/mol, AMBERbsc1: ~ 2 kcal/mol, AMBERbsc0: ~ 1 kcal/mol, CHARMM36: ~ 6 kcal/mol), reinforcing our proposition that in χ - θ based representation, the WC and HG states get contaminated by other states with higher/lower free energy and provide a poorer estimate of the relative free energy. One should note that these distance based coordinates may not be the optimal choice as a collective variable for the application of the biasing force, as we indeed need to force the transition of the glycosidic torsion angle to go from one base pairing conformation to the other. The hydrogen-bond donor–acceptor distance based CVs are useful to distinguish between WC and HG conformations during the postprocessing of the MD trajectories. In this sense, the primary significance of Figure 4 is that it shows the free energy landscape of the process of switching between WC and HG base pairing along the H-bond donor–acceptor distance collective variables which better distinguish the two basins as opposed to traditionally used torsion angle based CV. It was necessary to perform reweighting on the meta-eABF trajectories biased along χ - θ coordinates, as biasing these distances directly may disrupt the base paired structure of the DNA duplex making it difficult to revisit the base paired states during the simulation. This leads to significant noise in the free energy landscape in Figure 3 as the 2D space in the distance based CVs is not uniformly explored. Nevertheless, the energy values we observe specifically at the WC and the HG minima are slightly more reliable than what we observe in Figure 2 because of the better ability of these CVs to distinguish these states.

Root Mean Square Fluctuations and Absolute Entropy. The RMSF of heavy atoms are different at the A16 base and the nearby bps when comparing both WC and HG conformations (Figure 5). For instance, in both the CHARMM family force fields, the A16 base experienced more fluctuations in the HG form than the WC conformation. The opposite pattern was observed in the AMBER family models, where the fluctuations at the A16 base were larger in the WC form. In all force fields, there are also fluctuations occurring near the A16 and T9 base, which indicate the impact of the HG form on the entropy of the B-DNA. The large change in fluctuations could occasionally be observed in individual atoms in neighboring bases, for example, in the A18 base in case of AMBERbsc0 force field (Figure 5). Such fluctuations can be a result of long-range and short-range correlated motions of the atoms, but it is difficult to pinpoint from the current results. A similar collective motion between a base flipping out and the neighboring bases was observed by Lavery and co-workers in their “saloon door” mechanism, although we do not observe any base flip out motion in our unbiased trajectories. It is more interesting to look at the increase/decrease of fluctuations at

an atomic resolution. The change in fluctuation depicted in Figure 5 in residue-specific information is shown in Figure 6 as

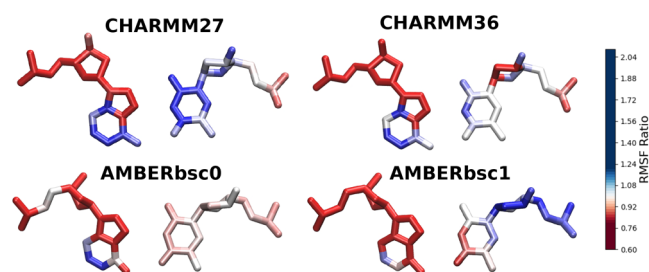


Figure 6. Atoms of the A16-T9 bp colored according to the RMSF ratio between the WC and HG states. The blue color indicates RMSF ratio >1 , and red color indicates <1 .

an atom-specific scale to understand which exact atoms are gaining and losing fluctuations. One common theme is apparent for all force fields: The base regions (particularly the six-membered ring of the adenine base) becomes more flexible in the HG form than in the WC form, while the backbone region becomes slightly more rigid in the HG form. The effect is not very consistent for the thymine base (Figure 6). This indicates that the apparent rigidity of the HG bp is not equally contributed by all the atoms.

In Figure 7, the quasi-harmonic entropy of the DNA duplex involving one HG conformation is found to be lower than that of a full WC DNA duplex, for all four force fields, possibly due to the increase of rigidity caused by the shrinkage of the helix diameter at the HG state. The extent by which the HG base pairing form is entropically destabilized varied between different force field models. The entropy difference of HG and WC bps is lowest in the CHARMM27 force field and is highest in AMBERbsc1, in contrast to the free energy results that indicate a lower free energy difference in AMBER force fields compared to CHARMM force fields. To understand the role of entropy and enthalpy in the stabilization of the WC base pairing form over the HG, we performed nonbonded interaction energy calculations. We found that the interaction energy between two bases is stronger for the HG base pairing form than the WC base pairing form, for all force field models (Figure 8). Although this seems counterintuitive, previous studies of Gould and Kollman show that in the crystalline state, the HG base pairing form is energetically more stable than the WC form.⁶⁴ This stabilization is much higher for the AMBER family force fields than the CHARMM family force fields. When we look at the interaction energy of the bp with the rest of the DNA duplex, the CHARMM family force fields show a positive (destabilizing) change in the interaction energy going from WC to HG, while a negative (stabilizing) change is observed for the AMBER family force fields. We also computed the relative “stacking” energy between the two base pairing forms. Although it is necessary to employ high-level quantum mechanical approaches to calculate the accurate π stacking energy of a given bp,^{22,65,66} we are limited to the use of nonpolarizable classical force fields for the current study. We therefore use the nonbonding interaction energy between the conjugated aromatic rings between the A16-T9 bp and the two adjacent bps (i.e., A17-T8 and C15-G10). We refer to this as the “base stacking” energy in the rest of the manuscript, to differentiate from the “ π stacking” energy. The difference of base stacking energy between the WC and HG forms in the

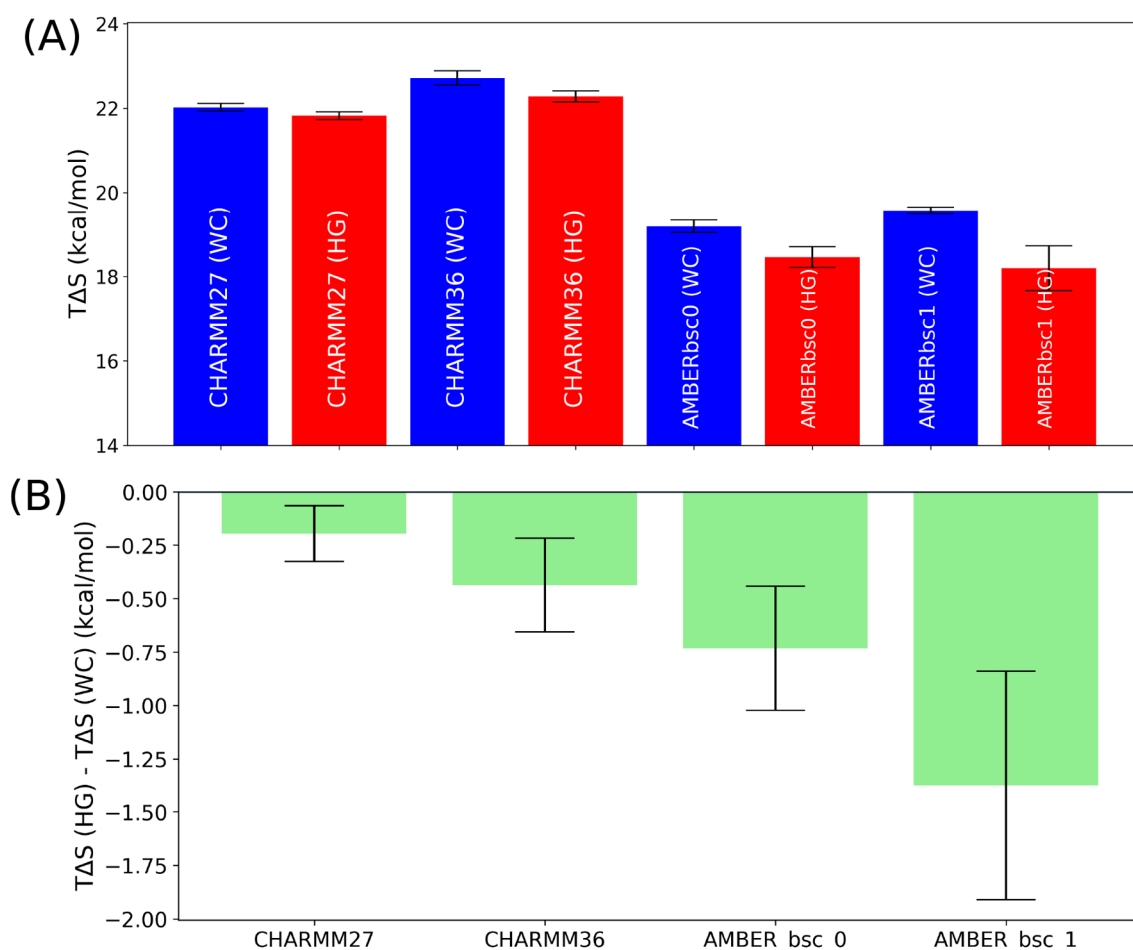


Figure 7. (A) The absolute value and (B) the relative value of the quasi-harmonic entropy of the HG bp with respect to the WC bp for different force fields.

CHARMM family force fields is almost negligible. But the stacking energy for the AMBER family force fields is stronger (more negative) in the HG form compared to the WC form (Figure 8). These results clearly suggest that the two types of force fields treat the HG base pairing very differently in terms of energy costs. The positive (or less negative) energy cost of forming a HG bp in CHARMM force fields results in a higher free energy difference between the WC and HG bps despite a lower entropic cost compared to AMBER force fields.

Nevertheless, the structural properties of the WC and HG bps are fairly consistent throughout the four force field models. The two torsion angles and various interatomic distances between the A16 and T9 bases are reported in Tables 1 and 2, respectively. Except for a wider fluctuation and a shifted location of the HG minima in the $\chi - \theta$ space for the AMBERbsc0 force field, the results of different force fields are in agreement with each other. One should also consider that the results in Tables 1 and 2 were averaged over multiple frames of a single MD trajectory. Due to the correlated nature of the data points, the true uncertainty is likely higher than the reported standard deviation, indicating that any apparent disagreement of structural parameters between force fields is insignificant. The time evolution of the H-bond donor-acceptor distances, the helix diameter (measures as the distance between the C1' atoms), and the torsion angles for the 100 ns equilibrium trajectories are provided in the

Supporting Information. It is clear from these plots that no WC to HG transition or base opening takes place during the simulation, confirming that these structures indeed represent minima in the free energy landscape.

CONCLUSIONS

We performed enhanced sampling and equilibrium MD simulations to understand the role of the force field model in predicting the properties related to the formation of the HG base pairing in the A6-DNA duplex. Although the equilibrium structures of the HG and WC base pairing forms were consistent in the different force field models, the relative free energy, entropy, and interaction energy of the HG bp change drastically when a different force field model was used. The relative stability of the two base pairing forms is determined by a fine balance of the entropy cost of forming the structurally rigid HG bp and the gain in interaction enthalpy over the WC structure. The topology of the underlying free energy landscape of the conformational transition between WC and HG when DNA breathes is more or less consistent between the different force field models. However, the predicted free energy costs and the barriers varied widely. Any quantitative comparison at the level of free energy should be performed with caution, given the approximate nature of the PMFs obtained from meta-eABF simulation. Multimicrosecond well-tempered metadynamics simulations can produce more

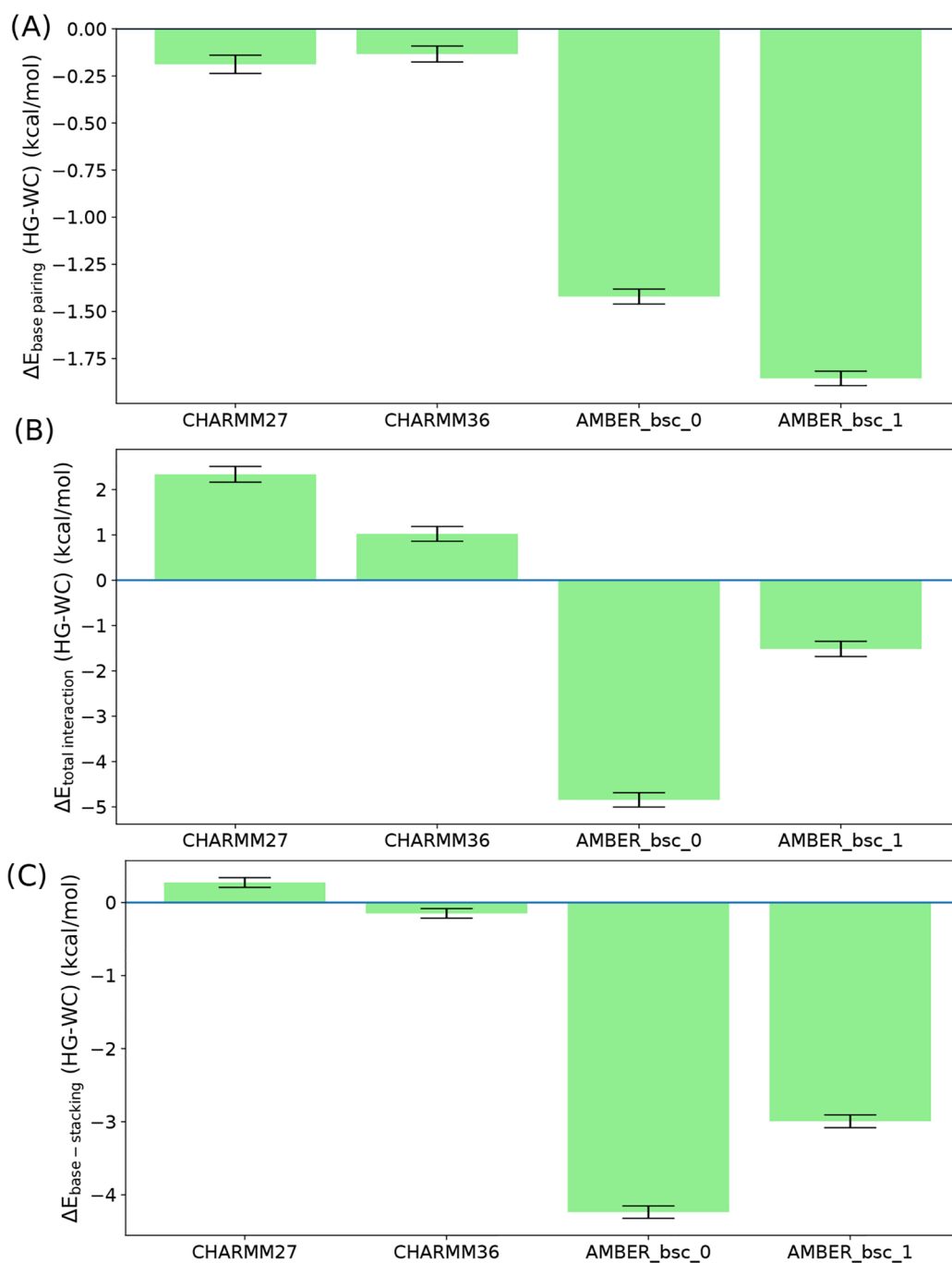


Figure 8. Nonbonded interaction energies in the HG bp relative to the WC bp, for the following interactions: (A) between the A16 and T9 bases, (B) between the bp with the rest of the DNA duplex, and (C) the stacking energy defined as the interaction energy between the π conjugated region of the bp with the two adjacent bps.

Table 1. Values of the Glycosidic Angle χ and the Base Flip-out Angle θ for the WC and HG Conformation for the Four Different Force Fields^a

force field	WC	HG	WC	HG
	χ (°)	χ (°)	θ (°)	θ (°)
CHARMM27	-109.2 ± 13.1	46.7 ± 10.3	-2.2 ± 3.9	0.0 ± 2.4
CHARMM36	-113.0 ± 14.8	45.2 ± 10.7	-0.4 ± 3.8	-0.2 ± 2.3
AMBERbsc0	-113.3 ± 20.1	26.8 ± 22.8	-5.9 ± 5.2	-0.3 ± 2.2
AMBERbsc1	-110.3 ± 15.7	61.0 ± 9.3	-5.3 ± 6.0	-1.3 ± 2.0

^aThe results are reported as mean ± standard deviation of the structures sampled in the 100 ns unbiased MD simulation.

Table 2. Distances between the Hydrogen Bond Donor–Acceptors (N1–N3 and N1–N7), and the C1' Atoms of the A16 and T9 Nucleotide, for the WC and HG Conformation for the Four Different Force Fields^a

force field	WC	HG	WC	HG	WC	HG
	N1(A)-N3(T) (Å)	N1(A)-N3(T) (Å)	N7(A)-N3(T) (Å)	N7(A)-N3(T) (Å)	C1(A)-C1(A) (Å)	C1(A)-C1(A) (Å)
CHARMM27	2.96 ± 0.13	5.90 ± 0.25	6.31 ± 0.15	3.12 ± 0.30	10.60 ± 0.28	8.88 ± 0.51
CHARMM36	2.96 ± 0.13	5.87 ± 0.21	6.29 ± 0.16	3.08 ± 0.22	10.60 ± 0.29	8.85 ± 0.40
AMBERbsc0	2.96 ± 0.12	5.84 ± 0.17	6.48 ± 0.15	3.02 ± 0.16	10.56 ± 0.29	9.03 ± 0.31
AMBERbsc1	2.97 ± 0.12	5.89 ± 0.17	6.51 ± 0.14	3.05 ± 0.14	10.59 ± 0.29	9.04 ± 0.30

^aThe results are reported as mean ± standard deviation of the structures sampled in the 100 ns unbiased MD simulation.

accurate free energy landscapes, but can become prohibitively expensive when trying to compare the results between different force fields. Importantly, the collective variables used for biasing are not optimal for distinguishing the two conformations and more sophisticated artificial intelligence driven order parameters, such as those obtained using spectral gap optimization⁶⁷ and harmonic linear discriminant analysis⁶⁸ may provide a better free energy difference.

It should also be taken into consideration that ergodicity, on the theoretical level, can be proven only as a necessary condition, but never a sufficient one. This means that faster convergence in any simulations does not always guarantee that the converged result is the ensemble average, which is, presumably, the experimental result. However, the experimental result is, at the same time, prone to its statistical inaccuracy. Therefore, despite the fact that our results do not agree perfectly with the previous computational studies with the AMBER bsc1 force field (which until now produced better agreement between simulation and experimental free energy difference), the question still remains on the match between MD and bulk experiments. Given this understanding, we here merely demonstrate that the choice of the force field can impact the results of the free energy landscape of DNA of breathing dynamics. When an identical protocol is used, the free energy landscapes from different force fields are comparable, provided that they visit all relevant configurations in the process (see the ergodicity argument above), which we can capture from our free energy landscapes. We chose meta-eABF because it quickly converges to a free energy surface that allows us to make a qualitative comparison between the different force fields in terms of their ability to describe the process in question. Alongside, we expect that both CHARMM and AMBER family force fields lead to a similar degree of exploration of the configurational space when subjected to identical enhanced sampling protocols.

Apart from the relative populations of the WC and HG base paired conformations, the transition kinetics between these two state have also been obtained from NMR relaxation experiments⁴ (exchange rates between 4 and 20 s⁻¹). But we restricted our analysis only to the thermodynamic properties, as calculating accurate kinetics for a multimillisecond process with different force fields is computationally extremely expensive. Nevertheless, in our previous work, we employed Markov state modeling to compute the kinetics of this conformational switching within 2 orders of magnitude agreement with the experimental data, from the CHARMM36 force field.³³

Based on the findings of the current study, we could only comment on the fact that the relative stability of the WC and HG base pairing conformation can be affected by the choice of the force fields. But, due to the quality of the agreement with experimental populations being poor in all force fields, it

remains difficult to conclusively suggest which force field optimally describes this very important conformational switching in DNA bps. Our results rather highlight the diversity of the predictions of molecular properties of WC and HG bps from different force fields and facilitates the assessment of the reliability of future computer simulation studies on DNA breathing dynamics.

■ SPECIAL ISSUE NOTE

The authors declare that they are committed to including, supporting and advancing women in Chemistry and scientific research in general. The authors are aware of the multitude of challenges faced by women in progressing their career in scientific research, particularly in academia, with only a small fraction of the leadership positions held by women. The lead author of this paper, S.E.S., is a first generation female college student from a historically underrepresented community and has first-hand experience of the enormous difficulties involved. The research group of I.A. have hosted many female scientists who proceeded to become successful independent scientists. The authors are not complacent, but aware that much more is needed to be done. All authors pledge to undertake conscious efforts in removing the barriers and creating an inclusive and supportive environment for women in science.

■ DATA AND SOFTWARE AVAILABILITY

The open-source package NAMD 2.14⁵⁵ is used for all MD simulations. The simulation input files and analysis scripts are available from the github repository: <https://github.com/dhimanray/DNA-HG-FF.git>.

■ ASSOCIATED CONTENT

SI Supporting Information

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/acs.jcim.2c00519>.

Additional plots of the stability of the HG and WC structures, convergence of entropy calculations, and the absolute values of interaction energies (PDF)

■ AUTHOR INFORMATION

Corresponding Author

Ioan Andricioaei – Department of Chemistry, University of California Irvine, Irvine, California 92697, United States; Department of Physics and Astronomy, University of California Irvine, Irvine, California 92697, United States; Email: andricio@uci.edu

Authors

Sharon Emily Stone – Department of Chemistry, University of California Irvine, Irvine, California 92697, United States

Dhiman Ray – Department of Chemistry, University of California Irvine, Irvine, California 92697, United States;
orcid.org/0000-0002-7103-0886

Complete contact information is available at:
<https://pubs.acs.org/10.1021/acs.jcim.2c00519>

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

This work was supported partially by the National Science Foundation (NSF) via grant MCB 2028443 awarded to I.A. S.E.S. acknowledges support through Pfizer - La Jolla Academic Industrial Relations (AIR) Diversity Research Fellowship in Chemistry for Undergraduates. The authors thank the University of California Irvine High Performance Computing (HPC) facility and the Triton Shared Computing Cluster (TSCC) in the San Diego Supercomputer Center (SDSC) for providing the computational resources.

REFERENCES

- (1) Franklin, R. E.; Gosling, R. G. Molecular Configuration in Sodium Thymonucleate. *Nature* **1953**, *171*, 740–741.
- (2) Watson, J. D.; Crick, F. H. C. Molecular Structure of Nucleic Acids: A Structure for Deoxyribose Nucleic Acid. *Nature* **1953**, *171*, 737–738.
- (3) Abrescia, N. G.; Thompson, A.; Huynh-Dinh, T.; Subirana, J. A. Crystal structure of an antiparallel DNA fragment with Hoogsteen base pairing. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 2806–2811.
- (4) Nikolova, E. N.; Kim, E.; Wise, A. A.; O'Brien, P. J.; Andricioaei, I.; Al-Hashimi, H. M. Transient Hoogsteen base pairs in canonical duplex DNA. *Nature* **2011**, *470*, 498–502.
- (5) Hoogsteen, K. IUCr, The structure of crystals containing a hydrogen-bonded complex of 1-methylthymine and 9-methyladenine. *Acta Crystallogr.* **1959**, *12*, 822–823.
- (6) Kitayner, M.; Rozenberg, H.; Rohs, R.; Suad, O.; Rabinovich, D.; Honig, B.; Shakked, Z. Diversity in DNA recognition by p53 revealed by crystal structures with Hoogsteen base pairs. *Nature Structural & Molecular Biology* **2010**, *17*, 423–429.
- (7) Xu, Y.; Manghrani, A.; Liu, B.; Shi, H.; Pham, U.; Liu, A.; Al-Hashimi, H. M. Hoogsteen base pairs increase the susceptibility of double-stranded DNA to cytotoxic damage. *J. Biol. Chem.* **2020**, *295*, 15933–15947.
- (8) Jiang, Y.-L.; Liu, Z.-P. Metallo-Organic G-Quadruplex Ligands in Anticancer Drug Design. *Mini-Reviews in Medicinal Chemistry* **2010**, *10*, 726–736.
- (9) Patel, D. J.; Shapiro, L.; Kozlowski, S. A.; Gaffney, B. L.; Jones, R. A. Covalent carcinogenic O6-methylguanosine lesions in DNA structural studies of the O6meGA and O6meG.G interactions in dodecanucleotide duplexes. *J. Mol. Biol.* **1986**, *188*, 677–692.
- (10) Singh, U. S.; Moe, J. G.; Reddy, G. R.; Weissenel, J. P.; Marnett, L. J.; Stone, M. P. Proton NMR of an oligodeoxynucleotide containing a propanodeoxyguanosine adduct positioned in a (CG)₃ frameshift hotspot of Salmonella typhimurium hisD3052: Hoogsteen base-pairing at pH 5.8. *Chem. Res. Toxicol.* **1993**, *6*, 825–836.
- (11) Yang, H.; Zhan, Y.; Fenn, D.; Chi, L. M.; Lam, S. L. Effect of 1-methyladenine on double-helical DNA structures. *FEBS Lett.* **2008**, *582*, 1629–1633.
- (12) Golovenko, D.; Bräuning, B.; Vyas, P.; Haran, T. E.; Rozenberg, H.; Shakked, Z. New Insights into the Role of DNA Shape on Its Recognition by p53 Proteins. *Structure* **2018**, *26*, 1237–1250.
- (13) Sathyamoorthy, B.; Shi, H.; Zhou, H.; Xue, Y.; Rangadurai, A.; Merriman, D. K.; Al-Hashimi, H. M. Insights into Watson–Crick/Hoogsteen breathing dynamics and damage repair from the solution structure and dynamic ensemble of DNA duplexes containing m1A. *Nucleic Acids Res.* **2017**, *45*, 5586–5601.
- (14) Rice, P. A.; Yang, S.-w.; Mizuuchi, K.; Nash, H. A. Crystal Structure of an IHF-DNA Complex: A Protein-Induced DNA U-Turn. *Cell* **1996**, *87*, 1295–1306.
- (15) Aishima, J.; Gitti, R. K.; Noah, J. E.; Gan, H. H.; Schlick, T.; Wolberger, C.; Hoogsteen, A. base pair embedded in undistorted B-DNA. *Nucleic Acids Res.* **2002**, *30*, 5244–5252.
- (16) Bunting, K. A.; Roe, S. M.; Headley, A.; Brown, T.; Savva, R.; Pearl, L. H. Crystal structure of the Escherichia coli dcm very-short-patch DNA repair endonuclease bound to its reaction product-site in a DNA superhelix. *Nucleic Acids Res.* **2003**, *31*, 1633–1639.
- (17) Zhou, H.; Kimsey, I. J.; Nikolova, E. N.; Sathyamoorthy, B.; Grazioli, G.; McSally, J.; Bai, T.; Wunderlich, C. H.; Kreutz, C.; Andricioaei, I.; Al-Hashimi, H. M. m1A and m1G disrupt A-RNA structure through the intrinsic instability of Hoogsteen base pairs. *Nature Structural & Molecular Biology* **2016**, *23*, 803–810.
- (18) Nair, D. T.; Johnson, R. E.; Prakash, S.; Prakash, L.; Aggarwal, A. K. Replication by human DNA polymerase- ϵ occurs by Hoogsteen base-pairing. *Nature* **2004**, *430*, 377–380.
- (19) Xu, Y.; McSally, J.; Andricioaei, I.; Al-Hashimi, H. M. Modulation of Hoogsteen dynamics on DNA recognition. *Nat. Commun.* **2018**, *9*, 1473.
- (20) Cubero, E.; Abrescia, N. G.; Subirana, J. A.; Luque, F. J.; Orozco, M. Theoretical Study of a New DNA Structure: The Antiparallel Hoogsteen Duplex. *J. Am. Chem. Soc.* **2003**, *125*, 14603–14612.
- (21) Cubero, E.; Luque, F. J.; Orozco, M. Theoretical Study of the Hoogsteen–Watson–Crick Junctions in DNA. *Biophys. J.* **2006**, *90*, 1000–1008.
- (22) Šponer, J.; Leszczynski, J.; Hobza, P. Nature of nucleic acid-base stacking: Nonempirical ab initio and empirical potential characterization of 10 stacked base dimers. Comparison of stacked and H-bonded base pairs. *J. Phys. Chem.* **1996**, *100*, 5590–5596.
- (23) Šponer, J.; Sabat, M.; Burda, J. V.; Leszczynski, J.; Hobza, P. Interaction of the adenine-thymine Watson–Crick and adenine-adenine reverse-Hoogsteen DNA base pairs with hydrated group IIa (Mg²⁺, Ca²⁺, Sr²⁺, Ba²⁺) and IIb (Zn²⁺, Cd²⁺, Hg²⁺) metal cations: Absence of the base pair stabilization by metal-induced polarization effects. *J. Phys. Chem. B* **1999**, *103*, 2528–2534.
- (24) Ryjáček, F.; Engkvist, O.; Vacek, J.; Kratochvíl, M.; Hobza, P. Hoogsteen and stacked structures of the 9-methyladenine...1-methylthymine pair are populated equally at experimental conditions: Ab initio and molecular dynamics study. *J. Phys. Chem. A* **2001**, *105*, 1197–1202.
- (25) Kratochvíl, M.; Šponer, J.; Hobza, P. Global minimum of the adenine/dotthymine base pair corresponds neither to Watson–Crick nor to Hoogsteen structures. Molecular dynamic/quenching/AMBER and ab initio beyond Hartree-Fock studies. *J. Am. Chem. Soc.* **2000**, *122*, 3495–3499.
- (26) Torrie, G.; Valleau, J. Nonphysical sampling distributions in Monte Carlo free-energy estimation: Umbrella sampling. *J. Comput. Phys.* **1977**, *23*, 187–199.
- (27) Laio, A.; Parrinello, M. Escaping free-energy minima. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 12562–6.
- (28) Barducci, A.; Bussi, G.; Parrinello, M. Well-Tempered Metadynamics: A Smoothly Converging and Tunable Free-Energy Method. *Phys. Rev. Lett.* **2008**, *100*, 020603.
- (29) Fu, H.; Zhang, H.; Chen, H.; Shao, X.; Chipot, C.; Cai, W. Zooming across the Free-Energy Landscape: Shaving Barriers, and Flooding Valleys. *J. Phys. Chem. Lett.* **2018**, *9*, 4738–4745.
- (30) Dellago, C.; Bolhuis, P. G.; Csajka, F. S.; Chandler, D. Transition path sampling and the calculation of rate constants. *J. Chem. Phys.* **1998**, *108*, 1964–1977.
- (31) Yang, C.; Kim, E.; Pak, Y. Free energy landscape and transition pathways from Watson–Crick to Hoogsteen base pairing in free duplex DNA. *Nucleic Acids Res.* **2015**, *43*, 7769–7778.
- (32) Yang, C.; Kim, E.; Lim, M.; Pak, Y. Computational Probing of Watson–Crick/Hoogsteen Breathing in a DNA Duplex Containing N1-Methylated Adenine. *J. Chem. Theory Comput.* **2019**, *15*, 751–761.

- (33) Ray, D.; Andricioaei, I. Free Energy Landscape and Conformational Kinetics of Hoogsteen Base Pairing in DNA vs. RNA. *Biophys. J.* **2020**, *119*, 1568–1579.
- (34) Pande, V. S.; Beauchamp, K.; Bowman, G. R. Everything you wanted to know about Markov State Models but were afraid to ask. *Methods* **2010**, *52*, 99–105.
- (35) Husic, B. E.; Pande, V. S. Markov State Models: From an Art to a Science. *J. Am. Chem. Soc.* **2018**, *140*, 2386–2396.
- (36) Vreede, J.; Pérez de Alba Ortíz, A.; Bolhuis, P. G.; Swenson, D. W. H. Atomistic insight into the kinetic pathways for Watson–Crick to Hoogsteen transitions in DNA. *Nucleic Acids Res.* **2019**, *47*, 11069.
- (37) Hooft, F.; Pérez De Alba Ortíz, A.; Ensing, B. Discovering Collective Variables of Molecular Transitions via Genetic Algorithms and Neural Networks. *J. Chem. Theory Comput.* **2021**, *17*, 2294–2306.
- (38) Frenkel, D.; Smit, B. *Understanding Molecular Simulation*, 2nd ed.; Academic Press: London, 2002.
- (39) Vanommeslaeghe, K.; Mackerell, A. D. CHARMM additive and polarizable force fields for biophysics and computer-aided drug design. *Biochimica et Biophysica Acta (BBA) - General Subjects* **2015**, *1850*, 861–871.
- (40) Salsbury, A. M.; Lemkul, J. A. Recent developments in empirical atomistic force fields for nucleic acids and applications to studies of folding and dynamics. *Curr. Opin. Struct. Biol.* **2021**, *67*, 9–17.
- (41) Foloppe, N.; Mackerell, A. D., Jr. All-Atom Empirical Force Field for Nucleic Acids: I. Parameter Optimization Based on Small Molecule and Condensed Phase Macromolecular Target Data. *J. Comput. Chem.* **2000**, *21*, 86–104.
- (42) Pérez, A.; Marchán, I.; Svozil, D.; Spöner, J.; Cheatham, T. E.; Laughton, C. A.; Orozco, M. Refinement of the AMBER force field for nucleic acids: Improving the description of α/γ conformers. *Biophys. J.* **2007**, *92*, 3817–3829.
- (43) Galindo-Murillo, R.; Robertson, J. C.; Zgarbová, M.; Šponer, J.; Otyepka, M.; Jurečka, P.; Cheatham, T. E. Assessing the Current State of Amber Force Field Modifications for DNA. *J. Chem. Theory Comput.* **2016**, *12*, 4114–4127.
- (44) Ivani, I.; Dans, P. D.; Noy, A.; Pérez, A.; Faustino, I.; Hospital, A.; Walther, J.; Andrio, P.; Goñi, R.; Balaceanu, A.; Portella, G.; Battistini, F.; Gelpli, J. L.; González, C.; Vendruscolo, M.; Laughton, C. A.; Harris, S. A.; Case, D. A.; Orozco, M. Parmbsc1: a refined force field for DNA simulations. *Nature Methods* **2016**, *13*, 55–58.
- (45) Minhas, V.; Sun, T.; Mirzoev, A.; Korolev, N.; Lyubartsev, A. P.; Nordenskiöld, L. Modeling DNA Flexibility: Comparison of Force Fields from Atomistic to Multiscale Levels. *J. Phys. Chem. B* **2020**, *124*, 38–49.
- (46) Liebl, K.; Zacharias, M. Tumuc1: A New Accurate DNA Force Field Consistent with High-Level Quantum Chemistry. *J. Chem. Theory Comput.* **2021**, *17*, 7096–7105.
- (47) Mohideen, K.; Muhammad, R.; Davey, C. A. Perturbations in nucleosome structure from heavy metal association. *Nucleic Acids Res.* **2010**, *38*, 6301–6311.
- (48) Humphrey, W.; Dalke, A.; Schulten, K. VMD: Visual molecular dynamics. *J. Mol. Graphics* **1996**, *14*, 33–38.
- (49) Case, D. A.; Cheatham, T. E.; Darden, T.; Gohlke, H.; Luo, R.; Merz, K. M.; Onufriev, A.; Simmerling, C.; Wang, B.; Woods, R. J. The Amber biomolecular simulation programs. *J. Comput. Chem.* **2005**, *26*, 1668–1688.
- (50) Boonstra, S.; Onck, P. R.; Van Der Giessen, E. CHARMM TIP3P Water Model Suppresses Peptide Folding by Solvating the Unfolded State. *J. Phys. Chem. B* **2016**, *120*, 3692–3698.
- (51) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **1983**, *79*, 926–935.
- (52) AMBER file and force field support (NAMD 2.14 User's Guide). <https://www.ks.uiuc.edu/Research/namd/2.14/ug/node13.html> (accessed July 11, 2020).
- (53) Martyna, G. J.; Tobias, D. J.; Klein, M. L. Constant pressure molecular dynamics algorithms. *J. Chem. Phys.* **1994**, *101*, 4177–4189.
- (54) Feller, S. E.; Zhang, Y.; Pastor, R. W.; Brooks, B. R. Constant pressure molecular dynamics simulation: The Langevin piston method. *J. Chem. Phys.* **1995**, *103*, 4613–4621.
- (55) Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kalé, L.; Schulten, K. Scalable molecular dynamics with NAMD. *J. Comput. Chem.* **2005**, *26*, 1781–1802.
- (56) Fu, H.; Shao, X.; Chipot, C.; Cai, W. Extended Adaptive Biasing Force Algorithm. An On-the-Fly Implementation for Accurate Free-Energy Calculations. *J. Chem. Theory Comput.* **2016**, *12*, 3506–3513.
- (57) Fiorin, G.; Klein, M. L.; Hénin, J. Using collective variables to drive molecular dynamics simulations. *Mol. Phys.* **2013**, *111*, 3345–3362.
- (58) Marcos-Alcalde, I.; Setoain, J.; Mendieta-Moreno, J. I.; Mendieta, J.; Gómez-Puertas, P. MEPSA: minimum energy pathway analysis for energy landscapes. *Bioinformatics* **2015**, *31*, 3853–3855.
- (59) Dijkstra, E. W. A note on two problems in connexion with graphs. *Numerische Mathematik* **1959**, *1*, 269–271.
- (60) Andricioaei, I.; Karplus, M. On the calculation of entropy from covariance matrices of the atomic fluctuations. *J. Chem. Phys.* **2001**, *115*, 6289.
- (61) Abraham, M. J.; Murtola, T.; Schulz, R.; Pall, S.; Smith, J. C.; Hess, B.; Lindahl, E. Gromacs: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **2015**, *1–2*, 19–25.
- (62) NAMD Energy Plugin, Version 1.4; University of Illinois at Urbana-Champaign: Urbana, IL. <https://www.ks.uiuc.edu/Research/vmd/plugins/namdenergy/>.
- (63) Fu, H.; Chen, H.; Wang, X.; Chai, H.; Shao, X.; Cai, W.; Chipot, C. Finding an optimal pathway on a multidimensional free-energy landscape. *J. Chem. Inf. Model.* **2020**, *60*, 5366–5374.
- (64) Gould, I. R.; Kollman, P. A. Theoretical Investigation of the Hydrogen Bond Strengths in Guanine-Cytosine and Adenine-Thymine Base Pairs. *J. Am. Chem. Soc.* **1994**, *116*, 2493–2499.
- (65) Hobza, P.; Šponer, J. Toward true DNA base-stacking energies: MP2, CCSD(T), and complete basis set calculations. *J. Am. Chem. Soc.* **2002**, *124*, 11802–11808.
- (66) Marek, P. H.; Szatyłowicz, H.; Krygowski, T. M. Stacking of nucleic acid bases: optimization of the computational approach—the case of adenine dimers. *Structural Chemistry* **2019**, *30*, 351–359.
- (67) Tiwary, P.; Berne, B. J. Spectral gap optimization of order parameters for sampling complex molecular systems. *Proc. Natl. Acad. Sci. U.S.A.* **2016**, *113*, 2839–2844.
- (68) Mendels, D.; Piccini, G.; Parrinello, M. Collective Variables from Local Fluctuations. *J. Phys. Chem. Lett.* **2018**, *9*, 2776–2781.