7100–7108 *Nucleic Acids Research, 2016, Vol. 44, No. 15* *Published online 12 July 2016*
doi: 10.1093/nar/gkw607

Free energy analysis and mechanism of base pair stacking in nicked DNA

Florian Häse and Martin Zacharias*

Physik-Department T38, Technische Universität München, James-Franck-Strasse 1, D-85748 Garching, Germany

Received May 20, 2016; Revised June 22, 2016; Accepted June 23, 2016

ABSTRACT

The equilibrium of stacked and unstacked base pairs is of central importance for all nucleic acid structure formation processes. The stacking equilibrium is influenced by intramolecular interactions between nucleosides but also by interactions with the solvent. Realistic simulations on nucleic acid structure formation and flexibility require an accurate description of the stacking geometry and stability and its sequence dependence. Free energy simulations have been conducted on a series of double stranded DNA molecules with a central strand break (nick) in one strand. The change in free energy upon unstacking was calculated for all ten possible base pair steps using umbrella sampling along a center-of-mass separation coordinate and including a comparison of different water models. Comparison to experimental studies indicates qualitative agreement of the stability order but a general overestimation of base pair stacking interactions in the simulations. A significant dependence of calculated nucleobase stacking free energies on the employed water model was observed with the tendency of stacking free energies being more accurately reproduced by more complex water models. The simulation studies also suggest a mechanism of stacking/unstacking that involves significant motions perpendicular to the reaction coordinate and indicate that the equilibrium nicked base pair step may slightly differ from regular B-DNA geometry in a sequence-dependent manner.

INTRODUCTION

The thermodynamic stability of double stranded (ds)DNA is of major importance for understanding the molecular details of gene expression and replication. Apart from sterical interactions determined by the nucleic acid backbone DNA stability is the result of two major interactions between nucleic acid base pairs, the pairing between complementary bases in opposite strands mediated by hydrogen bonding and the stacking between adjacent base pairs (1).

Quantifying the contribution of base stacking interactions has been of considerable interest for several decades. Early experimental approaches investigated base pair interactions in dinucleoside phosphates with various methods such as optical spectroscopic techniques (2), nuclear magnetic resonance (NMR) spectroscopy (3) and self-diffusion NMR (4). More recently base stacking interactions were measured in DNA fragments with a higher number of base pairs as well. Thermal denaturation methods (5), beacon kinetics (6) and mechanical unzipping (7), were used to estimate stacking free energies in single-stranded DNA. Stacking free energies in dsDNA could be obtained from experiments on nicked DNA fragments (8), characterized by a single strand break in one strand due to removal of a phosphate group (Figure 1).

Estimating stacking free energies for nucleic acid structures has also been of interest to the theoretical community. Accurate modeling of stacking interactions in computational studies is of particular importance as realistic descriptions of stacking interactions in molecular mechanics simulations are critical for performing realistic simulations on the stability and folding of nucleic acid structures. Stacking free energy contributions in early investigations were estimated with Poisson–Boltzmann calculations and free energy surface area computations (9). A number of molecular dynamics (MD) simulations were carried out to determine stacking free energies from unrestrained simulations (10), in umbrella sampling simulations (11), steered MD simulations (12) and one-step free energy perturbations (13). In addition, free energy simulations have also been employed to study blunt end stacking of DNA, however, limited to a single base pair step (14,15). The majority of the calculated stacking free energies from these simulations indicated the tendency of a stacking overestimation in molecular mechanics simulations.

Stacking interactions between bases in vacuum are due to the electronic molecular structure and can be calculated using quantum mechanical methods (16–18). However, in solution the stacking equilibrium (i.e. the stacking free energy) is determined by both nucleoside–nucleoside and nucleoside–solvent interactions. Hence, it is not surpris-

*To whom correspondence should be addressed. Tel: +49 89 289 12335; Fax: +49 89 289 12444; Email: martin.zacharias@ph.tum.de

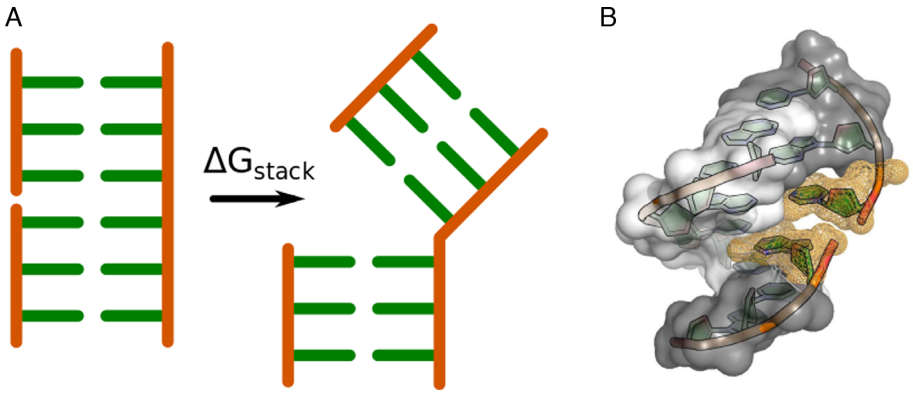

Figure 1. Nicked DNA fragments used for calculating nucleic acid stacking free energies. (**A**) Base pair stacking free energies can be determined from the stacked-unstacked equilibrium of nicked DNA. (**B**) Investigated DNA fragments consisted of six nucleic acid base pairs with a backbone nick between the third and the fourth base pair. Nicks were generated by removing the phosphate group at the center of one strand.

ing that pure quantum mechanical calculations in vacuum are only partially able to reproduce the sequence dependence of stacking in solution (16–18). However, quantum mechanical calculations were used as corrections to stacking free energies obtained from MD simulations (19–21), but also showed an overestimation of base stacking. Theoretical investigations on the influence of the solvent model on stacking free energies revealed a high sensitivity of computed stacking free energies to the solvent description used in the calculations (22). In fact, it could recently be shown that using a particular water model for explicit solvent water molecules in MD studies can result in more realistic stacking free energy predictions for dinucleoside steps (23).

Most simulation studies on nucleobase stacking were performed on isolated nucleobases or dinucleoside pairs (18,23–24). However, formation of dsDNA involves significant contributions from cross-stacking (interactions of complete base pairs) and the stacking geometry in a single stranded dinucleoside may also significantly differ from the geometry in B-DNA. Hence, studies on single strand dinucleotides may give only limited insight into the stability and formation of base pair stacks in dsDNA.

The equilibrium between stacked and unstacked structures of nicked DNA has been systematically investigated in experiments (8). This study allowed the determination of accurate stacking free energy parameters for all of the ten canonical base pair steps in B-DNA. The theoretical study of the same setup in free energy simulations offers the unique opportunity to directly compare stacking free energies based on molecular mechanics simulations with experimental findings for the same basic process.

In the present study umbrella sampling free energy simulations coupled with Hamiltonian replica exchange (H-REUS) were used to calculate the free energy of stacking/unstacking of nicked DNA for all 10 canonical base pair steps and comparing five different models for explicit water. In H-REUS simulations harmonic restraints are applied to several replica of the studied system to enhance sampling along a particular reaction coordinate. Conformational exchanges between these replica are allowed to further enhance sampling. Simulations were performed on small dsDNA fragments (six base pairs) with a central nick but using terminal restraints which mimic an embedding in longer DNA and allow extensive sampling to obtain converged stacking free energies.

Five different water models were employed to investigate the influence of explicit solvent water molecule parametrizations on nucleic acid base pair stacking free energies. We simulated DNA fragments in TIP3PF water (25), which is a modified TIP3P water model applied in many studies on DNA and other biomolecules (26–28), as well as in SPCE water (29), which includes a self-energy correction to the simple point charge (SPC) model. Both of these water models are three point water models and thus computationally cheaper than water models with more sites. We also employed the TIP4P model (30), which is a four site water model developed to closely reproduce experimental thermodynamic and structural data, as well as the TIP5P water model (31), with which the density anomaly of liquid water can be accurately simulated but which was shown to be structurally problematic for RNA simulations (32). Furthermore we also applied the recently developed optimal point charge (OPC) water model (33). Unlike the other water models, OPC was developed to accurately reproduce the electrostatic properties of water with three charges and four sites resulting in realistic predictions of liquid water properties.

To study the dynamics of nicked DNA compared to regular DNA the H-REUS simulations, totalling 61 μs, were supplemented with 8 μs of unrestrained simulations on nicked and regular DNA fragments. Comparing computed stacking free energies to experimental findings reported in (8) revealed a significant dependence of stacking free energies obtained from computational studies on the water model employed during the simulations which are of relevance for simulation studies on nucleic acid folding. We observed a tendency of calculated stacking free energies better agreeing with experimental results for more complex water models with more sites. The simulations also indicate a mechanism of unstacking at nicked sites which involves strong fluctuations perpendicular to the helical axis upon unstacking. We found that DNA fragments unstack via a combined slide and twist movement. Additionally, we ob-

served differences in the degree of base fraying of base pairs at the nicked site during unstacking.

MATERIALS AND METHODS

Initial structures for all ten canonical base pair steps (AA, AC, CC, CA, CG, GC, AT, TA, TC, CT) were generated with the NAB tool of the AmberTools15 software package (34). DNA fragments consisted of six nucleic acid base pairs with the sequences d(5′-CGXYCG/5′CGZWCG) in regular B-form, where all ten canonical base pair steps were inserted at positions XY and their complementary bases at positions ZW. A nick was introduced between X and Y by manually deleting the backbone phosphorus P and the attached oxygens OP1 and OP2. The TLEAP tool of the AmberTools15 software package was used to add hydrogens to the backbone oxygens O3′ and O5′.

Simulation protocols

Each of the generated DNA fragments was simulated in truncated octahedron periodic boxes with at least 10 Å distance to the box boundaries. The AMBER ff14 force field was used to model the DNA fragments (35). This set of parameters is based on the AMBER ff99 force field (36,37), supplemented with the bsc0 parameters (38) and was used in recent dinucleoside studies (23). DNA fragments were set up in explicit solvent with up to five different water models: TIP3PF (25), SPCE (29), OPC (33), TIP4P (30) and TIP5P (31). For each of the generated systems the charge was neutralized by adding sodium ions. Each system was minimized for 5000 steps followed by 200 ps equilibration in the NPT ensemble (constant temperature and pressure) at a temperature of 310 K to reproduce experimental conditions (8). Long range electrostatic interactions were calculated with the Particle-Mesh Ewald method (39).

H-REUS simulations were carried out to obtain the unstacking potentials of mean force (PMFs) for all generated DNA fragments in explicit solvent parametrized by different water models. The distance between centers-of-mass of heavy atoms in the deoxyriboses were used as a reaction coordinate, where a distance of 10 Å corresponds to stacked DNA conformations. A set of 25 umbrella windows biased by harmonic potentials with force constants of 8 kcal/mol·Å^2 was generated with an equidistant spacing of 0.6 Å used to sample distances from 10.0 Å up to 24.4 Å. Sampling was enhanced by allowing configurational exchanges between neighboring umbrella windows every 1 ps. With the applied umbrella window spacing exchange rates above 25% were achieved (Supplementary Figure S1). To prevent base fraying during the H-REUS simulations weak terminal restraints were applied to mimic embedding in longer DNA structures (Figure 2A). The total simulation time per DNA fragment for all umbrella windows was 2.65 μs with 50 ns to 150 ns per umbrella window (Supplementary Table S1).

For comparison we also conducted unrestrained MD simulations in the NVT ensemble at reference temperatures of 310 K for a total of 1 μs per DNA sequence per water model in accordance with the H-REUS set-up. Obtained heavy atom root mean square deviations are shown in Supplementary Figure S2. To prevent base fraying at the DNA termini weak restraints were applied to keep B-form DNA geometries at the termini (Figure 2A).

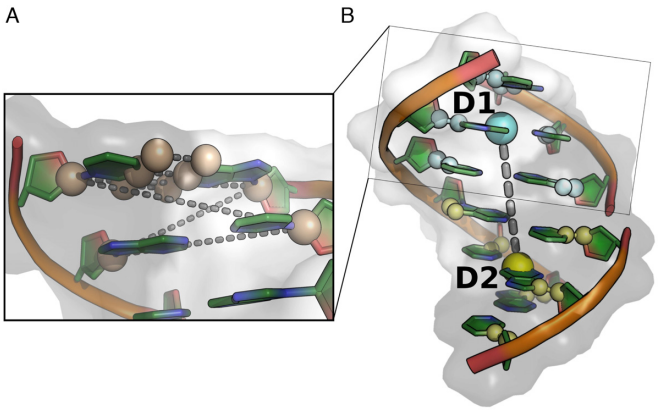

Figure 2. Restraints and reaction coordinate for calculating stacking free energies. (**A**) Restraints on the DNA geometry applied for all conducted MD simulations. Only atom distance restraints at one DNA terminus are depicted but were applied at both termini (see main text for details). (**B**) Center-of-mass distance between heavy atoms in the three base pairs upstream and downstream of the nicked site used as a reaction coordinate along which H-REUS simulations were conducted (see main text for details).

Restraints and reaction coordinate

The unstacking pathway was constructed by defining the distance between two particular centers-of-mass (Figure 2B), one above the nicked site (D1) and one below (D2). Both points, D1 and D2, were chosen to be the centers-of-mass of the C1′ and the N1 atoms in adenine and thymine and the C1′ and N9 atoms in guanine and cytosine respectively in the base pairs upstream and downstream of the nicked site. Equilibrium distances in B-DNA without backbone nick were determined as trajectory averages from the 1 μs unrestrained MD simulations and are reported for some DNA sequences in Supplementary Figure S3. Comparisons to distances obtained from unrestrained MD simulations confirmed the sensitivity of the chosen reaction coordinate on DNA unstacking.

Terminal base fraying in the simulated DNA fragments was prevented in all conducted simulations and B-form DNA geometries were kept by applying weak distance restraints on particular heavy atoms in base pairs upstream and downstream of the nicked base pair step. In particular, we restrained the distances *d*(N4,O6), *d*(N3,N1) and *d*(O2,N1) in the conjugate G-C base pairs at the DNA termini. Furthermore we introduced pairwise restraints between the C1’ atoms in bases upstream and downstream of the nicked base pair step. All restraints are highlighted in Figure 2A.

All restraints were chosen to be harmonic at distances of 0.1 Å below and 0.1 Å above B-DNA reference distances with force constants of 2.0 kcal/mol·Å^2. While these restraints keep the termini on average in a near B-DNA geometry (mimicing embbeding in a longer dsDNA) it still allows conformational fluctuations similar to fluctuations in regular B-DNA.

Obtaining stacking free energies and structural information

PMFs for the unstacking of all DNA fragments and water models were computed from reaction coordinate trajectories of all windows of the H-REUS simulations by employing the weighted histogram analysis method (WHAM) (40), at a temperature of 310 K to resemble experimental conditions (8), with 150 bins. The WHAM equations were iterated until free energy values in all bins did not change by more than 0.001 kcal/mol during two consecutive iterations.

Unstacking free energies $\Delta G_{\text{unstack}}$ were calculated from the obtained PMFs by comparing probabilities of being in the stacked conformation p_{stacked} to probabilities of being in the unstacked conformation $p_{\text{unstacked}}$ (Equation 1). Probabilities were computed by integrating Boltzmann factors along the corresponding reaction coordinate domain. DNA fragments were considered to be stacked up to distances of 16 Å and unstacked for larger distances. This choice of the stacking cut-off is further discussed in the Supplementary Section S1.4 and illustrated in Supplementary Figure S4.

$$\Delta G_{\text{unstack}} = k_{\text{B}} T \ln \left[\frac{p_{\text{stacked}}}{p_{\text{unstacked}}} \right],$$

$$\Delta G_{\text{unstack}} = k_{\text{B}} T \ln \left[\frac{\int_{r_0}^{r_{\text{cut}}} \exp\left(-\beta \,\text{PMF}\,(r)\right) dr}{\int_{r_{\text{cut}}}^{\infty} \exp\left(-\beta \,\text{PMF}\,(r)\right) dr} \right]. \quad (1)$$

Structural features of DNA unstacking were determined with the CURVES tool for nucleic acid structure analysis (41). We calculated several intra base-pair and inter base-pair DNA mechanics parameters at the nicked sites for all conformations sampled in the unrestrained simulations and the H-REUS simulations. In particular we calculated tilt, twist, slide, rise and opening to investigate the prevalent movements for DNA unstacking. Computed DNA mechanics parameters obtained from simulations of nicked DNA were compared to parameters computed in unrestrained MD simulations of regular DNA fragments to identify differences between nicked and regular fragments.

RESULTS AND DISCUSSION

Comparison of nicked and regular DNA dynamics

Structure and dynamics of nicked and regular DNA were compared in 1 μs unrestrained MD simulations on four different DNA fragments with central AT, TA, CG and GC base pair steps and either central nick or regular backbone in explicit OPC solvent. During the 1 μs simulation time a few transient unstacking events for nicked AT and TA sequences and no unstacking events for nicked CG or GC sequences were observed. These events were identified by computing root mean square deviations (RMSD) of heavy atoms in simulated DNA fragments with respect to native B-DNA geometries as generated by the NAB tool of the AmberTools15 software package (34). Obtained RMSD trajectories are depicted in Supplementary Figure S2. We conclude a qualitatively more favorable stacking for CG or GC base pair steps compared to central AT and TA steps.

The sampled distances between the centers-of-mass of the segments upstream and downstream of the center used as reaction coordinate distance for the subsequent free energy simulations (Figure 2B) were very similar for nicked and regular DNA fragments (Supplementary Figure S3). Averages and standard deviations of helical base pair parameters (tilt, twist, roll, slide, rise, shift) calculated with CURVES (41) are reported in Table 1. On average for the four investigated base pair steps the helical geometry of the DNA remained close to B-form in case of a nick in one strand but the deviations are significantly larger for AT and TA cases compared to dsDNA with central CG or GC steps (Table 1). While average values only change modestly and stay within the calculated standard deviations for all DNA mechanics parameters we observe a significant increase in the standard deviations of all helical parameters for nicked DNA sequences compared to regular DNA. In general, standard deviations increase more for CG than for GC and more for TA than for AT, indicating that GC is more tightly bound than CG and AT is more tightly bound than TA. This observation is in agreement with the higher stacking free energy computed for GC and AT described below.

Umbrella sampling free energy calculation of stacking at nicked DNA

H-REUS simulations employing a distance reaction coordinate between centers-of-mass of DNA segments flanking the central base pair step (see 'Materials and Methods' section) with a nick in one strand were carried out for different DNA sequences. The stacking/unstacking equilibrium of a central CG step was investigated with five water models (TIP3PF, SPCE, TIP4P, OPC and TIP5P). The PMF along the reaction coordinate were calculated with the weighted histogram analysis method (WHAM, see 'Materials and Methods' section). In standard umbrella sampling simulations the application of a penalty potential to limit the sampling to a small regime along the reaction coordinate in each umbrella window can result in trapping of conformations in locally stable states. Frequent exchange attempts of sampled conformations between umbrella windows using H-REUS allows to overcome barriers. Use of the H-REUS technique with frequent exchanges between neighboring umbrella simulation windows (Supplementary Figure S1) significantly enhanced the convergence of the free energy simulations. Convergence was carefully controlled using both splitting of the sampled trajectories in each umbrella windows into four block and calculating the PMFs from these blocks and performing a statistical analysis according to Zhu and Hummer (42) via block averaging (43). Details are provided in Supplementary Figure S5 in the Supplementary Section S1.5. In order to achieve convergence simulation times of up to 150 ns per umbrella window were required (Supplementary Table S1) amounting to 2.65 μs total simulation time for each calculated PMF in the present study.

The calculated PMFs for the CG case for all five investigated water models are depicted in Figure 3. The PMF minima are located at approximately 10.5 Å. The PMFs rise up to distances of about 16 Å (i) and continue relatively flat (ii) until they rise again significantly at about 21 Å (iii). Similar

Table 1. Averages and standard deviations of DNA helical parameters obtained from unrestrained simulations (1 μs)

	CG		GC		AT		TA	
Parameter	Regular	Nicked	Regular	Nicked	Regular	Nicked	Regular	Nicked
Twist [°]	32.3 ± 2.2	33.2 ± 3.1	32.7 ± 2.7	33.1 ± 2.7	33.4 ± 3.1	30.7 ± 17.1	31.1 ± 2.5	26.4 ± 9.2
Tilt [°]	0.3 ± 2.1	−0.3 ± 2.2	0.9 ± 2.2	0.1 ± 2.0	1.2 ± 2.1	−3.2 ± 12.1	0.2 ± 2.2	1.9 ± 5.3
Roll [°]	2.6 ± 3.2	1.7 ± 3.4	3.8 ± 3.1	2.7 ± 3.2	2.7 ± 3.5	5.3 ± 9.7	3.4 ± 3.8	6.1 ± 6.3
Slide [Å]	−0.1 ± 0.4	−0.3 ± 0.4	−0.3 ± 0.5	−0.7 ± 0.5	0.0 ± 0.4	0.5 ± 1.0	−0.4 ± 0.4	0.5 ± 0.9
Rise [Å]	3.3 ± 0.1	3.3 ± 0.1	3.4 ± 0.2	3.4 ± 0.1	3.4 ± 0.2	3.2 ± 0.8	3.3 ± 0.1	3.5 ± 0.9
Shift [Å]	0.0 ± 0.3	−0.1 ± 0.4	0.0 ± 0.3	0.0 ± 0.3	0.1 ± 0.3	0.1 ± 1.4	−0.2 ± 0.3	−0.5 ± 1.1

Helical parameters were calculated using the CURVES tool for nucleic acid structure analysis (41). Reported are values for twist, tilt, roll, slide, rise and shift computed as averages of the three central base pair steps (nicked step and adjacent steps) from simulations with and without central nick in OPC water (for central base pair steps CG, GC, AT and TA).

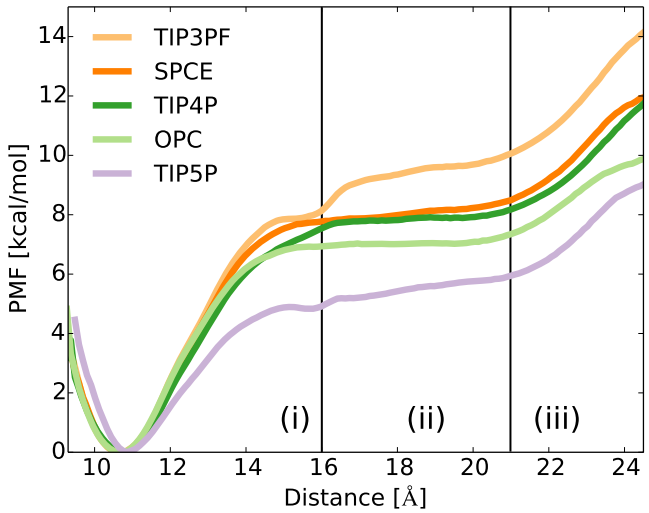

Figure 3. Computed potentials of mean force (PMFs) for the CG-nick system for all five investigated water models. All PMFs were obtained from 25 window H-REUS simulations. The reaction coordinate was chosen to be the distance between centers of mass of base pairs above and below the nick as explained in Figure 2. Region (i) indicates base pair unstacking at the nick site followed by further base pair separation in region (ii) until in region (iii) the DNA backbone is stretched.

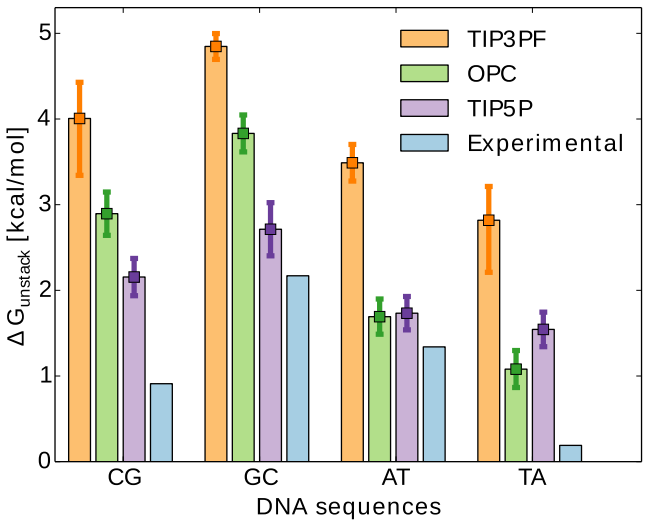

Figure 4. Computed unstacking free energies $\Delta G_{\text{unstack}}$ for four different DNA sequences (central base pair step indicated on the x-axis) compared to experimentally obtained stacking free energies (8). Results were obtained from integrating unstacking PMF. Molecular dynamics (MD) simulations were set up with the indicated water model. DNA was considered unstacked at a reaction coordinate distance of 16 Å. PMFs were computed in 25 window H-REUS simulations.

characteristic distances were observed for each water model and also for other DNA sequences. We relate these observations to base pair unstacking/stacking at the nicked site in region (i), indicated in Figure 3, further base pair separation in the flat free energy region (ii) and stretching of the DNA backbone in region (iii) at large center-of-mass distances between the DNA segments flanking the central nick.

Despite the PMF similarities among the water models crucial differences can be observed regarding the free energy penalty of unstacking. The rise of the PMFs in region (i) differs by up to 80% between TIP3PF and TIP5P water models. Furthermore a general trend of the penalty in region (i) with the water model can be observed. While less complex water models with only three sites (TIP3PF and SPCE) result in high penalties the lowest penalty is encountered for the five site water model (TIP5P).

For comparison the stacking/unstacking at the central nick was also investigated for the central GC, AT and TA cases using TIP3PF, OPC and TIP5P water models. With the obtained PMFs we computed stacking free energies $\Delta G_{\text{unstack}}$ for the four DNA sequences according to Equation 1. Based on the shape of the PMFs (Figure 3) we chose the cut-off distance r_{cut} to be 16 Å. Details about the influence of a different cut-off distance on the stacking free energies are discussed in the Supplementary Section S1.4, see Supplementary Figure S4 and Supplementary Table S3. Stacking free energies for three different water models (TIP3PF, OPC, TIP5P) applied to four different central DNA sequences (GC, CG, AT, TA) are shown in Figure 4 and compared to experimental values obtained from (8).

In all simulations with all water models stacking free energies are overestimated compared to experimental findings. TIP3PF shows the least agreement with experimental results and overestimates stacking free energies for all four sequences by more than 2 kcal/mol. Note, since we found for the CG case that the choice of SPCE or TIP4P gave unstacking penalties similar to TIP3PF the other three cases were only investigated using the most promising water models (OPC and TIP5P) and TIP3PF for comparison.

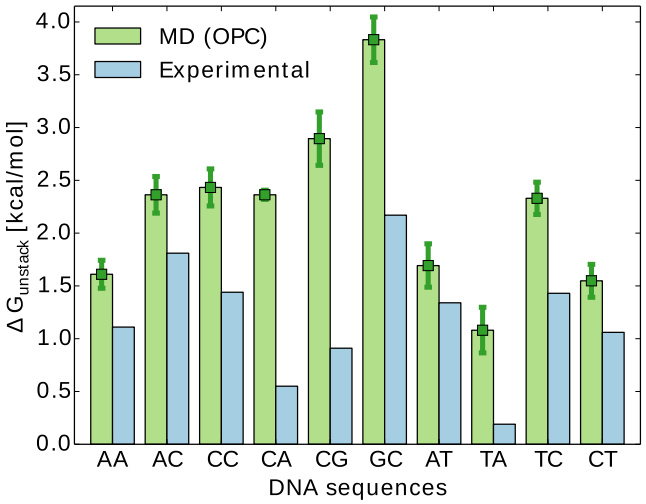

Figure 5. Computed unstacking free energies $\Delta G_{unstack}$ for all of the 10 investigated DNA sequences compared to experimentally obtained stacking free energies (8). Results were obtained from integrating unstacking PMF. DNA was considered unstacked at a reaction coordinate distance of 16 Å. PMFs were computed in 25 window H-REUS simulations.

Best agreement of calculated stacking free energies with experiment are achieved with the TIP5P water model for CG and GC and with the OPC model for the central AT and TA base pair steps. Calculated stacking free energies for both water models deviate from experimental values by <1 kcal/mol respectively. Differences between stacking free energies obtained with different water models are about 2 kcal/mol for all four sequences.

Since the computational demand for calculations using OPC as a four site water model is less than for TIP5P and since it has been specifically designed for improved description of solvation effects (33) the unstacking/stacking equilibrium of the remaining six conjugate base pair steps was investigated with the OPC water model. Details on the runtime are provided in the Supplementary Table S2. As a general trend the calculated free energies for unstacking the central nicked DNA are larger than the corresponding experimental data (8) in all cases (Figure 5). However, for 7 out of 10 cases the calculated free energies deviate <1 kcal/mol from the experimental results. In particular, the results for pyrimidine/purine steps (the least stable steps in experiment) deviate most from experiment (up to 2 kcal/mol). Also, the GC step is significantly more stable in the simulation as compared to experiment.

Mechanism of base pair unstacking

In order to characterize the helical motion of the central base pair step during the umbrella sampling unstacking/stacking simulations the average helical parameters and fluctuations for all 25 windows of the H-REUS simulations were calculated (Figure 6). We focus on the CG case, however, similar results were found for the other base pair steps (Supplementary Figure S6). At small reaction coordinate distances the averages of all investigated inter base-pair parameters are close to the B-DNA reference values. However, differences in the standard deviations are noticeably similar to the observation reported above in case of unrestrained simulations. The central rise and tilt only show small standard deviations (fluctuations) at reaction coordinate distances smaller than 14 Å while significantly higher fluctuations can be observed for twist and slide at the same reaction coordinate distance. Since distances smaller than 14 Å were identified to correspond to stacked DNA geometries we conclude that the prevalent movements initiating unstacking are sliding and twisting rather than tilting. This observation could be made for all ten canonical base pairs (Supplementary Figure S6). Representative conformations of the unstacking process obtained from H-REUS simulations are presented in Supplementary Figure S7.

Hence, rupture of the base pair stack has a strong component perpendicular to the helical axis of the construct. It occurs by a sliding and twisting fluctuation perpendicular to the pulling direction due to the umbrella potential independent of the water model used in the simulations. Note, that similar observations have been made by Maffeo *et al.* (15) in simulation studies on blunt end stacking of DNA molecules.

In order to characterize the degree of fraying of base pairs upon stacking disruption at the central nick we recorded the degree of opening of conjugate base pairs at the nicked site in all umbrella sampling windows. Base pair opening at the nicked site was defined as the sum of opening of the conjugate base pairs adjacent to the nicked site. For the cases with central AT and TA base pair steps more fraying (base pair opening) was observed compared to the GC and CG cases (Figure 7). Similar trends were observed for all three water models used for the four cases.

We found significantly higher opening for AT and TA base pairs than for CG and GC base pairs upon unstacking. With values of more than 120° AT and TA base pairs are prone to fraying as opposed to GC and CG with value of around 40°. This difference in base pair opening indicates that AT and TA are less tightly bound in the DNA geometry and have the tendency to be solvated.

CONCLUSION

In order to investigate the mechanism and thermodynamics of base pair stacking and unstacking MD simulations and free energy simulations were performed on DNA molecules containing a central nick in one DNA strand. The simulations resemble an experimental setup for measuring base pair stacking free energies in nicked DNA reported by Protozanova *et al.* (8) and allow direct comparison of calculated free energies with experiment.

Similar to the simulations the experimental measurements were performed at low salt concentrations because the analysis involved electrophoresis experiments. There is, however, one parameter that is the reaction coordinate distance at which conformations are counted as unstacked or stacked that may differ in experiment and simulation. We selected here a distance about 6 Å larger than the reaction coordinate distance for a minimum energy stacked state that characterizes a full unstacking of the central stack as transition distance. Larger distances will affect the calculated free energies only slightly. However, smaller distances may not be compatible with experiment because the detection of

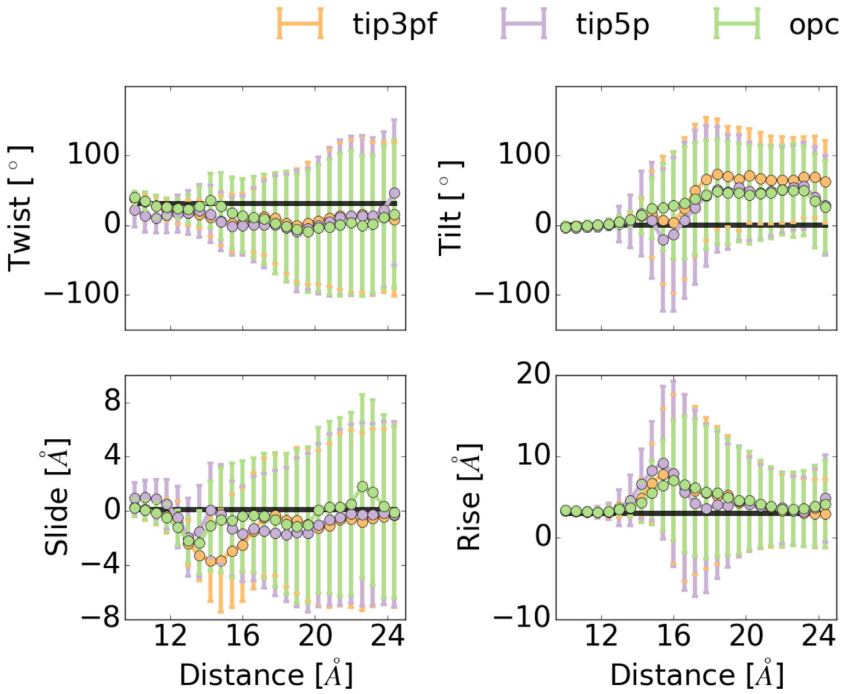

Figure 6. Selected base pair helical parameters for the central CG step during unstacking calculated for different water models. Presented are trajectory averages and standard deviations of twist, tilt, slide and rise of the nick enclosing base pairs. B-DNA reference values are depicted as black solid lines. While tilt and rise are close to equilibrium B-DNA values with a small standard deviations at reaction coordinate distances smaller than 12.5 Å, twist and slide show significant deviations with higher standard deviations even at small distances, indicating that sliding and twisting are the prevalent movements for unstacking for all three investigated water models.

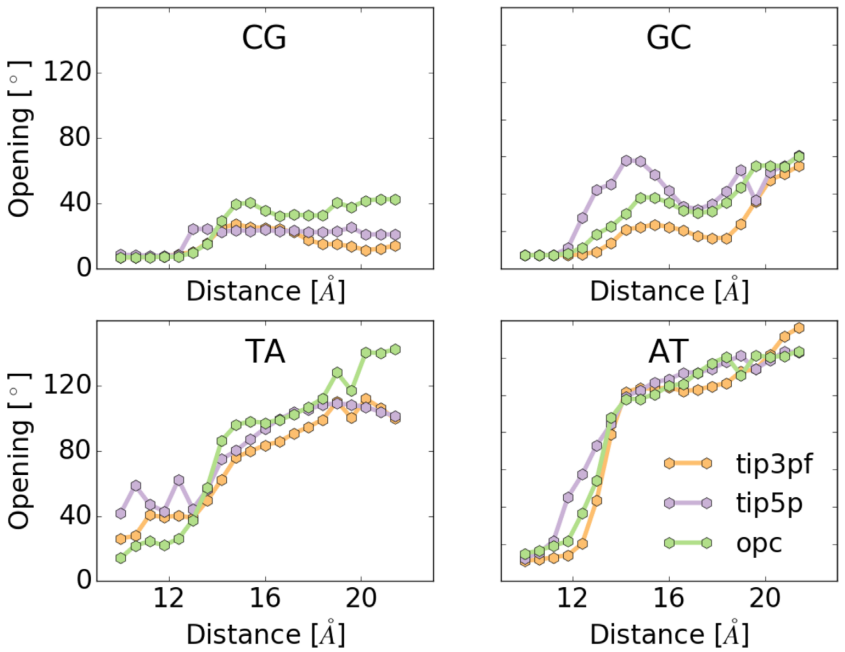

Figure 7. Base pair opening standard deviations of base pairs at the nicked site upon unstacking for four different central base pair steps and three different water models. TA and AT sequences show significantly higher opening standard deviations than CG and GC regardless of the water model indicating a higher degree of fraying at the nicked site upon unstacking along the reaction coordinate distance. With OPC water the highest increase in base pair opening standard deviations upon unstacking among all water models is encountered consistently for all DNA sequences.

unstacking in the experiment is based on the bending of the DNA which becomes significant only at distances greater than 16 Å.

Stacking free energies for different base pair sequences in explicit solvent with different water models could be computed from the H-REUS simulations. The calculated free energies for base pair unstacking at a central nick are in general larger than corresponding experimental data and this difference showed a significant dependence on the water model. Deviations from experimental results were smaller the more sites were used to parametrize water molecules. With only four sites and three charges and stacking free energy deviations of less than 1 kcal/mol from experimental results for seven out of ten canonical base pair sequences we found that the OPC water model might be the currently best compromise between physical accuracy and computational demand. Note, that Mobley *et al*. found significant differences in calculated solvation free energies for various organic compounds (44) using different water models. This is in line with our observation of a significant dependence of the effective stacking free energies which differ for each water model mostly because of different solvation properties of the water models. Indeed, the OPC water model has been designed specifically to improve the agreement between calculated and experimental solvation free energies for organic compounds. Also, in line with our results are simulation studies by Bergonzo and Cheetham (45) on small RNA molecules that indicated a much better agreement of sampled short RNA conformations with experimental results when using the OPC versus TIP3P or TIP4PEw water models (mostly due to incorrect stacking for the latter models).

Furthermore, we found structural evidence for prevalent unstacking movements and indications that free energies obtained from DNA unstacking do not only consist of base pair stacking contributions but can also partially include nucleobase solvation free energies due to base pair fraying at the nicked site. It is likely that this may also affect the interpretation of the experimentally derived stacking free energies. The experimental result may also include contributions due to fraying of the terminal unstacked base pair at least for AT rich sequences and not only exclusively the unstacking/stacking equilibrium of the base pair step at the nick.

Another issue tackled in the current study is the question if a base pair step at a nicked site is identical to a base pair step in regular B-DNA. Our extensive comparative unrestrained MD simulations indicate that overall the DNA at the nick site resembles B-DNA but deviations from B-DNA geometry are observed especially for TA and AT base pair step cases. In addition, the observed helical fluctuations are larger than in regular B-DNA. Both the small average deviations and the increased fluctuations may contribute to a slightly altered stacking at a nicked site relative to the stacking in regular B-DNA. Hence, in addition to the fraying effect, the stacking/unstacking equilibrium at a nicked site may not exactly represent the free energy of stacking/unstacking in regular B-DNA because of the difference in stacking geometry at the reference (stacked) state.

Understanding the molecular mechanism of unstacking events at nicked sites is important for folding processes of nucleic acids or enzymatic processes such as ligation or polymerization starting at a nicked site. The analysis of helical parameters and fluctuations indicate that the tension created in DNA due to an umbrella pulling potential to unstack the DNA results in increased fluctuations especially of twist and slide indicating increased motions perpendicular to the helical axis prior to disruption of stacking (manifested in an increase of rise). Hence, the base pairs at a nicked site open toward the solvent more readily in directions perpendicular to the helical axis by a slide/twist motion.

The simulations conducted in this study highlight the sensitivity of interactions among nucleobases on the parametrization of solvent water molecules and suggests to use more expensive water models for more accurate results. We found that OPC provides a reasonable trade-off between agreement with experimental data and computational demand. It remains to be seen if water models can be further improved to allow even more accurate MD simulations on nucleic acids. This is an important prerequisite for using molecular mechanics simulations for studying structure formation or folding and refolding processes of DNA and RNA molecules.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We thank the DFG (German Research Foundation) grant SFB (collaborative Research Center) 863 (project A10) for financial support.

FUNDING

Sonderforschungsbereich 863 project A10 of the Deutsche Forschungsgemeinschaft (DFG) [SFB863, project A10]; LRZ (Leibniz Supercomputer Center) [pr48po]. Funding for open access charge: DFG [SFB863, project A10].
Conflict of interest statement. None declared.

REFERENCES

1. Saenger,W. (1984) *Principles of Nucleic Acid Structure*. Springer-Verlag, NY.
2. Warshaw,M.M. and Tinoco,I. (1966) Optical properties of sixteen dinucleoside phosphates. *J. Mol. Biol.*, **20**, 29–38.
3. Chan,S.I. and Nelson,J.H. (1969) Proton magnetic resonance studies of ribose dinucleoside monophosphates in aqueous solution. I. Nature of the base-stacking interaction in adenylyl(3'-5')adenosine. *J. Am. Chem. Soc.*, **91**, 168–183.
4. Stokkeland,I. and Stilbs,P. (1985) A multicomponent self-diffusion NMR study of aggregation of nucleotides, nucleosides, nucleic acid bases and some derivatives in aqueous solution with divalent metal ions added. *Biophys. Chem.*, **22**, 65–75.
5. Guckian,K.M., Schweitzer,B.A., Ren,R.X.F., Sheils,C.J., Paris,P.L., Tahmassebi,D.C. and Kool,E.T. (1996) Experimental Measurement of Aromatic Stacking Affinities in the Context of Duplex DNA. *J. Am. Chem. Soc.*, **118**, 8182–8183.
6. Aalberts,D.P., Parman,J.M. and Goddard,N.L. (2003) Single-stranded stacking free energy from DNA beacon kinetics. *Biophys. J.*, **84**, 3212–3217.

7. Huguet,J.M., Bizarro,C.V., Forns,N., Smith,S.B., Bustamante,C. and Ritort,F. (2010) Single-molecule derivation of salt dependent base-pair free energies in DNA. *Proc. Natl. Acad. Sci. U.S.A.*, **107**, 15431–15436.
8. Protozanova,E., Yakovchuk,P. and Frank-Kamenetskii,M.D. (2004) Stacked-unstacked equilibrium at the nick site of DNA. *J. Mol. Biol.*, **342**, 775–785.
9. Friedman,R.A. and Honig,B. (1995) A free energy analysis of nucleic acid base stacking in aqueous solution. *Biophys. J.*, **69**, 1528–1535.
10. Norberg,J. and Nilsson,L. (1994) Stacking-unstacking of the dinucleoside monophosphate guanylyl-3',5'-uridine studied with molecular dynamics. *Biophys. J.*, **67**, 812–824.
11. Norberg,J. and Nilsson,L. (1995) Stacking free energy profiles for all 16 natural ribodinucleoside monophosphates in aqueous solution. *J. Am. Chem. Soc.*, **117**, 10832–10840.
12. Konrad,M.W. and Bolonick,J.I. (1996) Molecular dynamics simulation of DNA stretching is consistent with the tension observed for extension and strand separation and predicts a novel ladder structure. *J. Am. Chem. Soc.*, **118**, 10989–10994.
13. Oostenbrink,C. and van Gunsteren,W.F. (2005) Efficient calculation of many stacking and pairing free energies in DNA from a few molecular dynamics simulations. *Chem. Eur. J.*, **11**, 4340–4348.
14. Wang,R., Kuzuya,A., Liu,W. and Seeman,N.C. (2010) Blunt-ended DNA stacking interactions in a 3-helix motif. *Chem. Commun.*, **46**, 4905–4907.
15. Maffeo,C., Luan,B. and Aksimentiev,A. (2012) End-to-end attraction of duplex DNA. *Nucleic Acid Res.*, **40**, 3812–3821.
16. Gu,J., Wang,J. and Leszczynski,J. (2011) Stacking and H-bonding patterns of dGpdC and dGpdCpdG: performance of the M05-2X and M06-2X Minnesota density functionals for the single strand DNA. *Chem. Phys. Lett.*, **512**, 108–112.
17. Šponer,J., Šponer,J.E., Mládek,A., Jurečka,P., Banás,P. and Otyepka,M. (2013) Nature and magnitude of aromatic base stacking in DNA and RNA: Quantum chemistry, molecular mechanics and experiment. *Biopolymers*, **99**, 978–988.
18. Mak,C.H. (2016) Unraveling base stacking driving forces in DNA. *J. Phys. Chem. B*, doi:10.1021/acs.jpcb.6b01934.
19. Florián,J., Šponer,J. and Warshel,A. (1999) Thermodynamic parameters for stacking and hydrogen bonding of nucleic acid bases in aqueous solution: ab initio/langevin dipoles study. *J. Phys. Chem. B*, **103**, 884–892.
20. Kabeláč,M. and Hobza,P. (2001) Potential energy and free energy surfaces of all ten canonical and methylated nucleic acid base pairs: molecular dynamics and quantum chemical ab initio studies. *J. Phys. Chem. B*, **105**, 5804–5817.
21. Banás,P., Mládek,A., Otyepka,M., Zgarbová,M., Jurečka,P., Svozil,D., Lankas,F. and Šponer,J. (2012) Can we accurately describe the structure of adenine tracts in B-DNA? Reference quantum-chemical computations reveal overstabilization of stacking by molecular mechanics. *J. Chem. Theory Comput.*, **8**, 2448–2460.
22. Norberg,J. and Nilsson,L. (1998) Solvent influence on base stacking. *Biophys. J.*, **74**, 394–402.
23. Brown,R.F., Andrews,C.T. and Elcock,A.H. (2015) Stacking free energies of all DNA and RNA nucleoside pairs and dinucleoside-monophosphates computed using recently revised AMBER parameters and compared with experiment. *J. Chem. Theory Comput.*, **11**, 2315–2328.
24. Jafilan,S., Klein,L., Hyun,C. and Florián,J. (2012) Intramolecular base stacking of dinucleoside monophosphate anions in aqeuous solution. *J. Phys. Chem. B*, **116**, 3613–3618.
25. Price,D.J. and Brooks,C.L. (2004) A modified TIP3P water potential for simulation with Ewald summation. *J. Chem. Phys.*, **121**, 10096–10103.
26. Pérez,A., Luque,F.J. and Orozco,M. (2007) Dynamics of B-DNA on the microsecond time scale. *J. Am. Chem. Soc.*, **129**, 14739–14745.
27. Liebl,K., Drsata,T., Lankas,F., Lipfert,J. and Zacharias,M. (2015) Explaining the striking difference in twist-stretch coupling between DNA and RNA: a comparative molecular dynamics analysis. *Nucleic Acid Res.*, **43**, 10143–10156.
28. Hocking,H.G., Häse,F., Madl,T., Zacharias,M., Rief,M. and Žoldák,G. (2015) A compact native 24-residue supersecondary structure derived from the villin headpiece subdomain. *Biophys. J.*, **108**, 678–686.
29. Berendsen,H.J.C., Grigera,J.R. and Straatsma,T.P. (1987) The missing term in effective pair potentials. *J. Phys. Chem.*, **91**, 6269–6271.
30. Jorgensen,W.L.A., Chandrasekhar,J., Madura,J.D., Impey,R.W. and Klein,M.L. (1983) Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys*, **79**, 926–935.
31. Mahoney,M.W. and Jorgensen,W.L.A. (2000) A five-site model for liquid water and the reproduction of the density anomaly by rigid, nonpolarizable potential functions. *J. Chem. Phys.*, **112**, 8910–8922.
32. Kührovk,P., Otyepka,M., Šponer,J. and Banks,P. (2014) Are waters around RNA more than just a solvent?—An insight from molecular dynamics simulations. *J. Chem. Theory Comput.*, **10**, 401–411.
33. Izadi,S., Anandakrishnan,R. and Onufriev,A.V. (2014) Building water models: a different approach. *J. Phys. Chem. Lett.*, **5**, 3863–3871.
34. Case,D.A., Berryman,J.T., Betz,R.M., Cerutti,D.S., Cheatham,T.E., Darden,T.A., Duke,R.E., Giese,T.J., Gohlke,H., Goetz,A.W. *et al.* (2015) *AMBER 15*. University of California, San Francisco, CA.
35. Maier,J.A., Martinez,C., Kasavajhala,K., Wickstrom,L., Hause,K.E. and Simmerling,C. (2015) ff14SB: improving the accuracy of protein side chain and backbone parameters from ff99SB. *J. Chem. Theory Comput.*, **11**, 3696–3713.
36. Cornell,W.D., Cleplak,P., Bayly,C.I., Gould,I.R., Merz,K.M., Ferguson,D.M., Spellmeyer,D.C., Fox,T., Caldwell,J.W. and Kollman,P.A. (1995) A second generation force field for the simulation of proteins, nucleic acids, and inorganic molecules. *J. Am. Chem. Soc.*, **117**, 5179–5197.
37. Cheatham,T.E., Cieplak,R. and Kollman,P.A. (1999) A modified version of the cornell et al. Force field with improved sugar pucker phases and helical repeat. *J. Biomol. Struct. Dyn.*, **16**, 845–862.
38. Pérez,A., Marchán,I., Svozil,D., Sponer,J., Cheatham,T.E., Laughton,C.A. and Orozco,M. (2007) Refinement of the AMBER force field for nucleic acids: improving the description of α/γ conformers. *Biophys. J.*, **92**, 3817–3829.
39. Essman,U., Perera,L., Berkowitz,M.L., Darden,T., Lee,H. and Pederson,L. (1995) A smooth particle mesh Ewald method. *J. Chem. Phys.*, **103**, 8577–8593.
40. Grossfield,A. (2013) *WHAM: the weighted histogram analysis method*. University of Rochester, http://membrane.urmc.rochester.edu/content/wham.
41. Lavery,R., Moakher,M., Maddocks,J.H., Petkeviciute,D. and Zakrzewska,K. (2009) Conformational analysis of nucleic acids revisited: Curves+. *Nucleic Acid Res.*, **37**, 5917–5929.
42. Zhu,F. and Hummer,G. (2012) Convergence and error estimation in free energy calculations using the weighted histogram analysis method. *J. Comput. Chem.*, **33**, 453–465.
43. Hess,B. (2002) Determining the shear viscosity of model liquids from molecular dynamics simulations. *J. Chem. Phys.*, **116**, 209–217.
44. Mobley,D.L., Bayly,C.I., Cooper,M.D., Shirts,M.R. and Dill,K.A. (2009) Small molecule hydration free energies in explicit solvent: an extensive test of fixed-charge atomistic simulations. *J. Chem. Theory Comput.*, **5**, 350–358.
45. Bergonzo,C. and Cheatham,T. (2015) Improved force field parameters lead to a better description of RNA structure. *J. Chem. Theory Comput.*, **11**, 3969–3972.