Biochimica et Biophysica Acta 1850 (2015) 1059–1071

Contents lists available at ScienceDirect

Biochimica et Biophysica Acta

journal homepage: www.elsevier.com/locate/bbagen

All-atom crystal simulations of DNA and RNA duplexes ☆

Chunmei Liu [a,b], Pawel A. Janowski [b], David A. Case [b,*]

[a] The College of Chemistry and Molecular Engineering, Zhengzhou University, Zhengzhou, Henan Province 450001, PR China
[b] Dept. of Chemistry and Chemical Biology, Rutgers University, Piscataway, NJ 08854, USA

ARTICLE INFO

Article history:
Received 30 June 2014
Received in revised form 12 September 2014
Accepted 13 September 2014
Available online 26 September 2014

Keywords:
Molecular dynamics
Nucleic acids
Crystal

ABSTRACT

Background: Molecular dynamics simulations can complement experimental measures of structure and dynamics of biomolecules. The quality of such simulations can be tested by comparisons to models refined against experimental crystallographic data.
Methods: We report simulations of DNA and RNA duplexes in their crystalline environment. The calculations mimic the conditions for PDB entries 1D23 [d(CGATCGATCG)2] and 1RNA [(UUAUAUAUAUAUAA)2], and contain 8 unit cells, each with 4 copies of the Watson–Crick duplex; this yields in aggregate 64 μs of duplex sampling for DNA and 16 μs for RNA.
Results: The duplex structures conform much more closely to the average structure seen in the crystal than do structures extracted from a solution simulation with the same force field. Sequence-dependent variations in helical parameters, and in groove widths, are largely maintained in the crystal structure, but are smoothed out in solution. However, the integrity of the crystal lattice is slowly degraded in both simulations, with the result that the interfaces between chains become heterogeneous. This problem is more severe for the DNA crystal, which has fewer inter-chain hydrogen bond contacts than does the RNA crystal.
Conclusions: Crystal simulations using current force fields reproduce many features of observed crystal structures, but suffer from a gradual degradation of the integrity of the crystal lattice.
General significance: The results offer insights into force-field simulations that test their ability to preserve weak interactions between chains, which will be of importance also in non-crystalline applications that involve binding and recognition.
This article is part of a Special Issue entitled Recent developments of molecular dynamics.

1. Introduction

RNA and DNA molecules play an important role in many biological processes, and an understanding of their structure and dynamics is indispensable for a complete understanding of their function. Molecular dynamics simulations can offer a detailed complement to experiment, and nucleic acid simulations in a crystal environment have long been used to test simulation methods [1,2]. A "modern" era began in the mid 1990s with the introduction of Ewald-based methods to simulate long-range electrostatic interactions [3]. At about that time, simulations of Watson–Crick paired duplexes in both solution and in the crystal lattice provided evidence that simulations remained stable for a few nanoseconds without the need for artificial restraints [4–7], and that, not surprisingly, the duplex structure in the crystal lattice environment more closely resembles the experimental crystal structures than do simulations in a solution environment. Subsequent studies in a crystalline lattice employed longer time scales and different force fields, reaching broadly similar conclusions [8–13]. Advances in computing speed have spurred a new round of biomolecular crystal simulations (mostly for proteins [14–21]), that use larger simulation cells and pay increased attention to the properties of the crystal lattice, as well as to the structural characteristics of individual chains.

Here we report results of molecular dynamics simulations of duplexes of DNA (PDB ID: 1D23 [22]) and RNA (PDB ID: 1RNA [23,24]) for 2.0 and 0.5 μs; both crystals are in the P212121 space group, with four duplexes per unit cell. The periodic unit in the simulations is a "supercell" containing 8 unit cells, so the simulations contain 32 copies of each duplex. Parallel simulations of a single duplex in water (neutralized by Na^+ counterions) are reported for comparison to the crystal simulations. Since a large number of solution simulations of DNA and RNA helices have previously been reported [3,25–32], our primary emphasis here is on the properties of the crystal lattice and the ability of current simulation methods to reproduce such behavior. The

☆ This article is part of a Special Issue entitled Recent developments of molecular dynamics.

* Corresponding author. Tel.: +1 848 445-5885.
E-mail address: case@biomaps.rutgers.edu (D.A. Case).

http://dx.doi.org/10.1016/j.bbagen.2014.09.018
0304-4165/

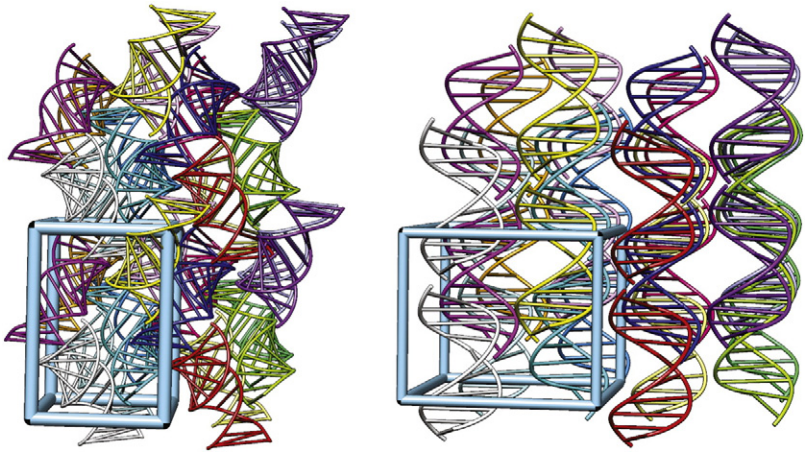

Fig. 1. Two three dimension views of the RNA (left) and DNA (right) packing in supercells. The supercell contains 8 unit cells in a 2 × 2 × 2 arrangement; each unit cell comprises four RNA/DNA molecules. The cyan box represents one unit cell. Image generated using Chimera.

simulation results explore the details of stability of the duplexes and crystal packing effects in μs dynamics; this time scale has been explored in earlier solution simulations of DNA duplexes [30,31].

The main results of this paper are as follows: first, the average duplex structure from the crystal simulation more closely resembles the experimental crystal structure than the average structure in a solution simulation. Some details of the crystal structure, such as groove widths, are averaged out in solution simulations but not in the crystal, whereas other sequence-dependent variations are maintained. These latter include the characteristic alternating pattern of base pair roll/twist in alternating A–U sequences in RNA and the BI/BII pattern in the backbone torsion angles ($\varepsilon - \xi$) in DNA [33–35]. As in previous simulations, the backbone conformation presents more of a challenge to simulations than does base positioning. However, for both DNA and RNA, the integrity of the crystal lattice is slowly degraded, and a full equilibration is not achieved. The interchain contacts that hold these lattices together are less extensive than is typical for protein crystals, and are likely to be influenced by ions in ways that are poorly represented here. A partial "melting" of the crystal lattice takes place at seemingly random points in the supercell, such that the 32 equivalent interfaces become rather heterogeneous, with some quite close to their original (experimental) configurations and others are displaced in ways that are irreversible on the time scales sampled here.

2. Models and computational methods

Coordinates for the initial state of the RNA and DNA came from PDB entries: 1RNA and 1D23. Crystal structures of the RNA and DNA crystal were constructed according to space group P212121, where four symmetry-related copies of the duplex make up one unit cell. We constructed a supercell of 2 × 2 × 2 unit cells by using the *PropPDB* module of the Amber12 package [36], yielding an orthogonal box of size 68.22 × 89.22 × 98.22 Å for 1RNA, and 77.86 × 79.26 × 66.60 Å for 1D23. Views of the supercells are shown in Fig. 1. For the DNA crystal simulation, Mg^{2+} ions were used as counterions: there are two $Mg(H_2O)_6^{2+}$ complexes per asymmetric unit in the 1D23 crystal structure, and 7 additional $Mg(H_2O)_6^{2+}$ ions per duplex were added to neutralize the 18 phosphate groups. The RNA supercell system was neutralized with 26 Na^+ ions per duplex. All counterions and water molecules were added to the supercell using the *AddToBox* program in Amber. The DNA system for the solution simulation contained one DNA duplex and 5045 water molecules together with 18 Na^+ ions in a truncated octahedral box with periodic boundary conditions; the RNA simulation had 7694 waters and 26 Na^+ ions. Further simulation details are given in Table 1.

We carried out crystal simulations using the parm99 force field [37] with bsc0 modifications [38] for both duplexes, and added the chiOL3 modifications [39] for RNA. The water model used was TIP3P [40] for DNA and TIP4P/Ew [41] for RNA; monovalent ion parameters were taken from Ref. [42] and Mg^{2+} parameters from Ref. [43]. Simulations of nucleic acids in solution show only small conformational dependencies

Table 1
Details of the simulations. Simulation speed is measured on a single NVIDIA GTX780 card for the RNA simulations, and on a Quadro K5000 card for the DNA simulations. Ions are Mg^{2+} for the DNA crystal and Na^+ for the other simulations. Atom counts do not include the "extra points" in the TIP4P/Ew waters.

	DNA	RNA	DNA (soln)	RNA (soln)
Number of waters	8578	12300	5045	7694
Number of ions	288	832	18	26
Total number of atoms	46246	65892	15785	23988
Mean internal pressure, atm	30	97	1.0	1.0
Density, g/cm^3	1.433	1.446	1.044	1.034
Simulation speed, ns/day	16.7	21.5	24.3	26.9
Time analyzed, μs	2.0	0.5	0.5	1.0

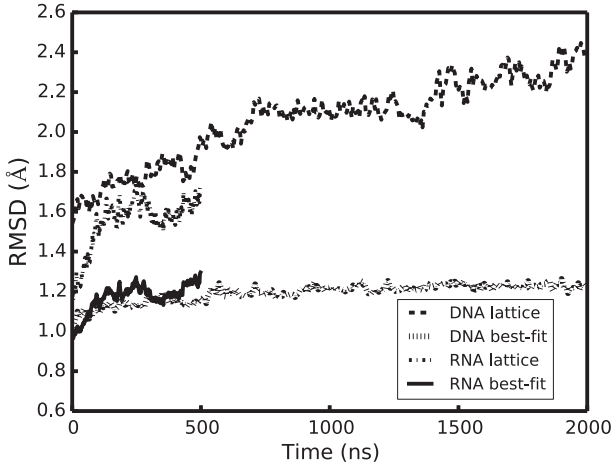

Fig. 2. Positional RMSDs of all heavy atoms for RNA and DNA relative to the initial X-ray structure in the course of simulation.

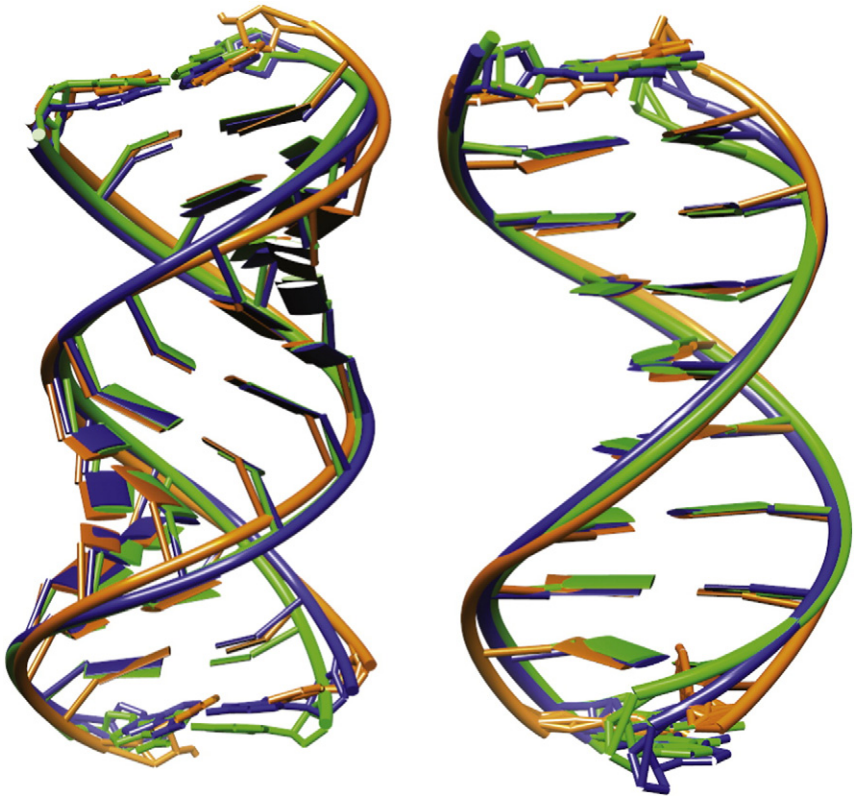

Fig. 3. Superpositions of the solution average structure (orange) and the best-fit average structure (blue) for RNA (left) and DNA (right) versus the deposited crystal structures (green).

on the water model used. As an example, we compared our TIP4P/Ew simulation of 1RNA in solution with results from Ref. [25], where a TIP3P water model was used. Average value twist, roll, helical twist, propeller twist, slide, x-displacement and helical rise were essentially the same in the two simulations. We do not yet have similar comparisons for crystal simulations.

The minimization, equilibration, and production dynamics were performed using the *pmemd* module of Amber12 [36,44,45]. The SHAKE algorithm was applied to all bonds involving hydrogen atoms. Force calculations were performed with periodic boundary conditions, a 9.0 Å cutoff on real space interactions, a homogeneity assumption for long-range Lennard–Jones contributions, and smooth particle-mesh Ewald electrostatics. An energy minimization procedure was used first to remove any bad contacts in the starting conformation. Next, the system was equilibrated at the experimental data collection temperatures of 308 K for RNA and 273 K for DNA, with successive restraints on the RNA/DNA atoms of 10.0, 1.0 and 0.1 kcal/(mol-Å^2) for a total of 40 ns. The volume was kept fixed at the experimental value, and the system pressure monitored. The number of water molecules was adjusted by trial and error to obtain a simulation with an external pressure near 1 atm. Finally, unrestrained production dynamics were propagated at a 2 fs time step for 2.0 μs for DNA and 0.5 μs for RNA. A total of 4000 equally-spaced snapshots for DNA and 2500 for RNA were saved for subsequent analysis. Some details of the simulation are collected in Table 1. The solution simulations followed the general procedures of the "ABC" simulations [27–29,46], and placed a single duplex in a truncated octahedron, neutralized by Na^+ ions, which were equilibrated in a similar fashion.

Results were analyzed using Amber *cpptraj* module, and figures were prepared by Chimera and VMD molecular visualization programs. Helical parameters throughout the trajectories were monitored with the Curves+ program [47]. The BI and BII configurations in DNA were characterized by the angles ε and ξ of the DNA backbone or by the angle difference $(\varepsilon - \xi)$, which is $-90°$ for BI and $+90°$ for BII phosphates [35]. Interfaces were analyzed using the PISA program [48].

Table 2
Root mean square deviations (Å) from the deposited crystal structures for 1RNA and 1D23. The statistics in each box are heavy atoms, backbone atoms, base atoms RMSD of average structure against the crystal structure. The values in parentheses exclude the terminal residues.

	All heavy atoms	Backbone	Base	All heavy atoms (solution)
Best-fit (RNA)	0.78 (0.71)	0.87	0.59	1.38 (1.12)
Lattice (RNA)	0.93	1.04	0.74	~
Best-fit (DNA)	0.77 (0.64)	0.89	0.56	1.83 (1.34)
Lattice (DNA)	1.12	1.23	0.95	~

3. Analysis of duplex structures and comparison with solution simulations

The root mean square deviations (RMSDs) from the crystal structure for all heavy atoms (of all 32 duplexes) are shown in Fig. 2. For the curves labeled "best-fit", an optimal translation and rotation movement to fit the crystal coordinates was determined for each duplex in each snapshot; these values reach an apparent plateau of about 1.2 Å in about 200 ns of simulation. For the curves labeled "lattice", the duplexes were superimposed on the crystal configuration using only the lattice symmetry parameters, with no fitting (see Ref. [19] for a detailed explanation of these statistics). These latter values reflect both deviations of individual duplexes from the crystal configuration and the degradation of the lattice itself. The latter continues for the entire trajectory, and will be analyzed in Section 4.

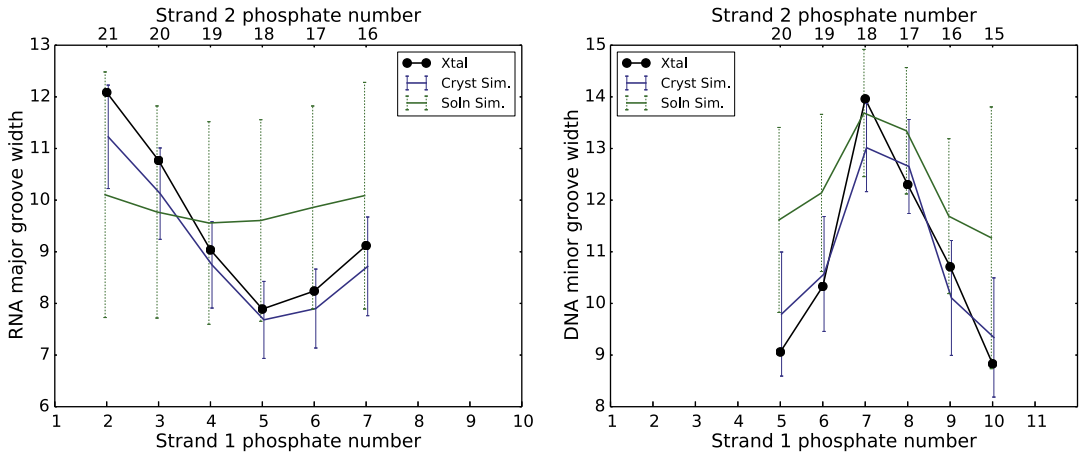

Fig. 4. Plots of major-groove width for RNA and minor-groove width for DNA in crystal and solution simulations. Widths in Å are defined by the distance between the phosphate atoms shown at the top and bottom. Vertical bars give the standard deviations of the fluctuations seen in the simulations.

Table 3
Average α and γ angles (degrees) in crystal and solution simulations. The data in parentheses are standard deviations.

	α (cryst.)	α (cryst. sim. ave)	α (solution sim. ave)	γ (cryst.)	γ (cryst. sim. ave)	γ (solution sim. ave)
A11	155	279 (29.9)	281 (15.1)	178	73 (23.7)	64 (12.6)
U24	210	282 (12.9)	282 (12.4)	127	64 (9.5)	63 (9.4)

The best-fit and solution average structures are compared to the experimental crystal configuration in Fig. 3, and some statistics are given in Table 2. As expected, and as seen in earlier crystal simulations, the average structure in solution deviates much more from the crystal configuration than does the average structure from the crystal simulation. The crystal simulation deviations are larger in the backbone for both DNA and RNA than for the bases. Some details of these differences are provided in the following sections.

3.1. Groove widths

The A- and B-form helices for RNA and DNA have very different geometries, and can be monitored as groove "widths" and "depths". In Fig. 4, we use a simple measure groove width, based on phosphate–phosphate distances from one chain to its complementary strand. For RNA, the sequence dependence of the widths of the major groove is suppressed in solution, leading to a more regular pattern of distances than are seen in the crystal. Figs. 3 and 4 show that the major groove widens slightly compared with crystal structure. This behavior of solution simulations has been discussed earlier, in connection with attempts to characterize how NMR constraints determine the details of A-form helices in RNA [49]. Furthermore, fluctuations of the major groove in the crystal simulation are much smaller than those in the solution simulation.

For DNA, the minor groove width is the greatest at phosphates P7/P8, not only in crystal structure but also in the crystal and solution simulations. This is related to the presence of a BII phosphate conformation at positions P7 and P17 in the crystal structure. The larger minor groove widths at positions P10 and P20 in the solution are related to the larger fluctuations in the terminal residues that are characteristic of solution simulations. In addition, the wide minor groove at P7 is consistent with the large positive roll deformation (Fig. 7, below) which compresses the major groove and then widens the minor groove [50]. (Because of the way steps are numbered, a wide minor groove separation between P7 and P18 should be correlated with a positive roll angle at position P5.)

3.2. Backbone conformations

The backbone of the 1RNA structure contains two kinks (at A23/U24 and U10/A11), which divide the whole duplex into three roughly equal regions. The kinks result from changes in the α and γ backbone torsions from their typical values near 280 and 70, respectively. Table 3 lists the average values in both simulations, showing a reversion in the simulation back to the canonical A-type values. This reversion suggests that crystal packing forces are not sufficient (at least for the simulation analyzed here) to retain the unusual crystal configuration.

The DNA backbone generally has two major sub-forms, BI and BII, that can be characterized by the ϵ and ξ backbone torsion angles [34,

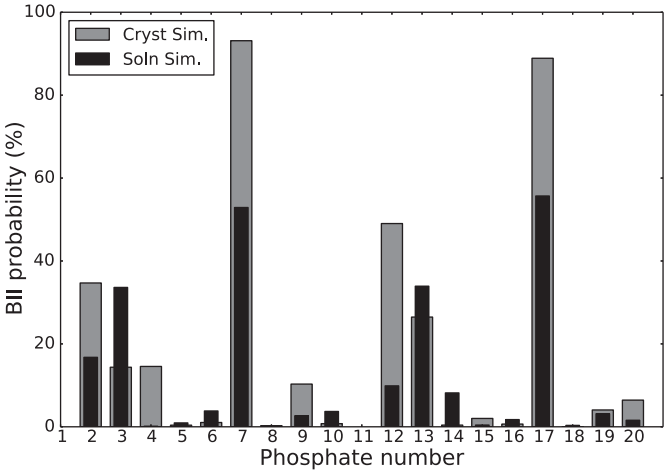

Fig. 5. Conformational substates (BII) probability in the crystal and solution simulation along the sequence: P2 to P10 are in strand 1 and P12 to P20 are in strand 2. In the crystal structure, P2, P7, P12, and P17 have a BII conformation, and all others have a BI conformation.

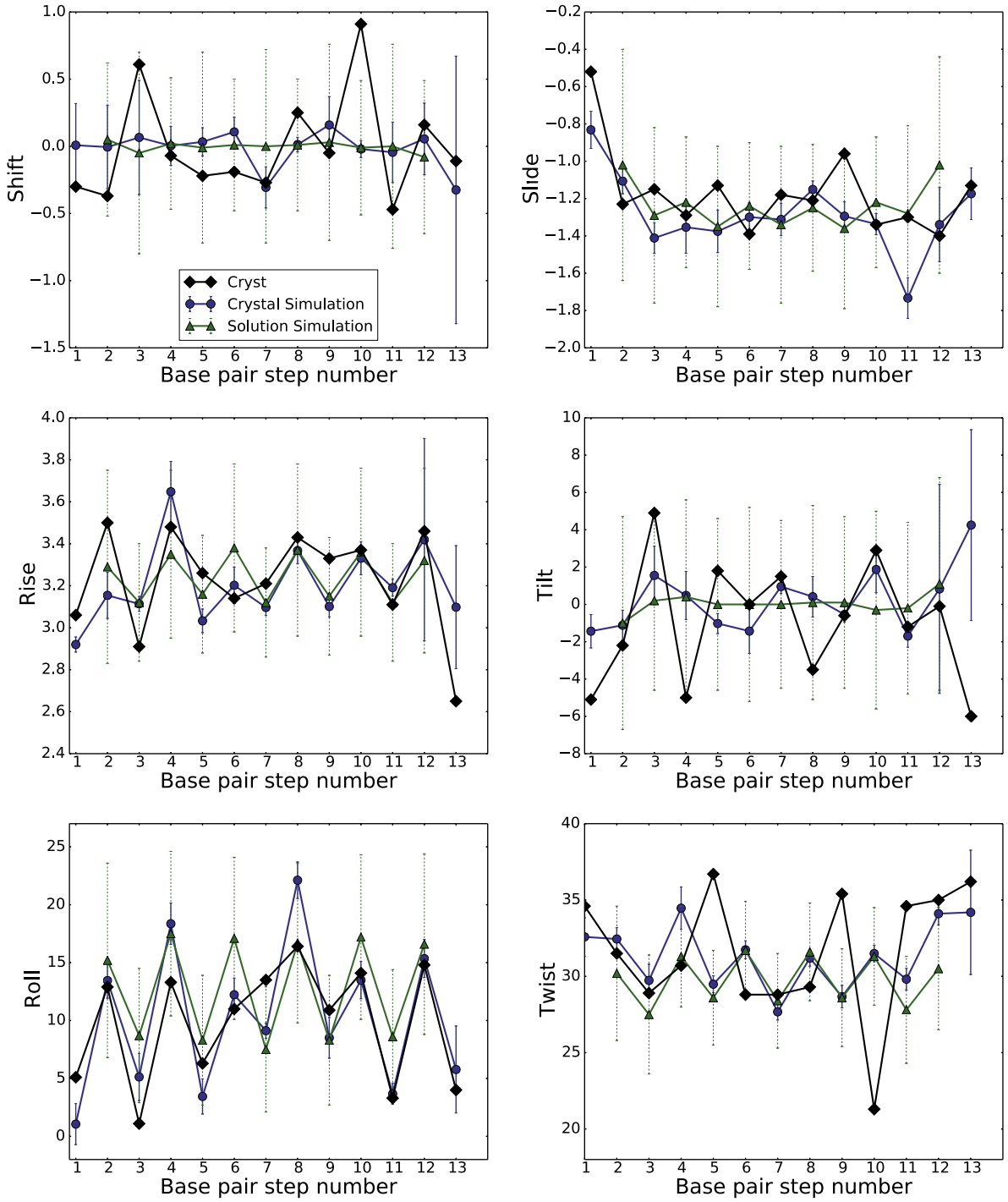

Fig. 6. RNA base pair step parameters, translational parameters are in angstroms (Å) and rotational parameters are in degrees (°).

35,51]. The range of ϵ and ξ torsion angles for BI structure varies from 120° to 210° and from 235° to 295°, respectively, while for BII these vary from 210° to 300° and from 150° to 210°, respectively. The difference $(\varepsilon - \xi)$ is near +90° for BII and near –90° for BI. In the crystal structure, P2, P7, P12, and P17 have a BII conformation, and all others have a BI conformation. Fig. 5 shows the fraction of BII conformation in the crystal and solution simulations. Both P7 and P17 largely maintain a BII conformation, whereas P2 and P12 revert to a mixture of BI and BII conformations. The percentages of the BII conformation at P2, P7, P12 and P17 are much smaller in the solution simulation.

3.3. Base pair step parameters

Nucleic acid helices are commonly characterized by six helicoidal parameters (tilt, roll, twist, shift, slide, and rise) that describe the configuration of base steps [52]; simulation results are given in Figs. 6 and 7. There

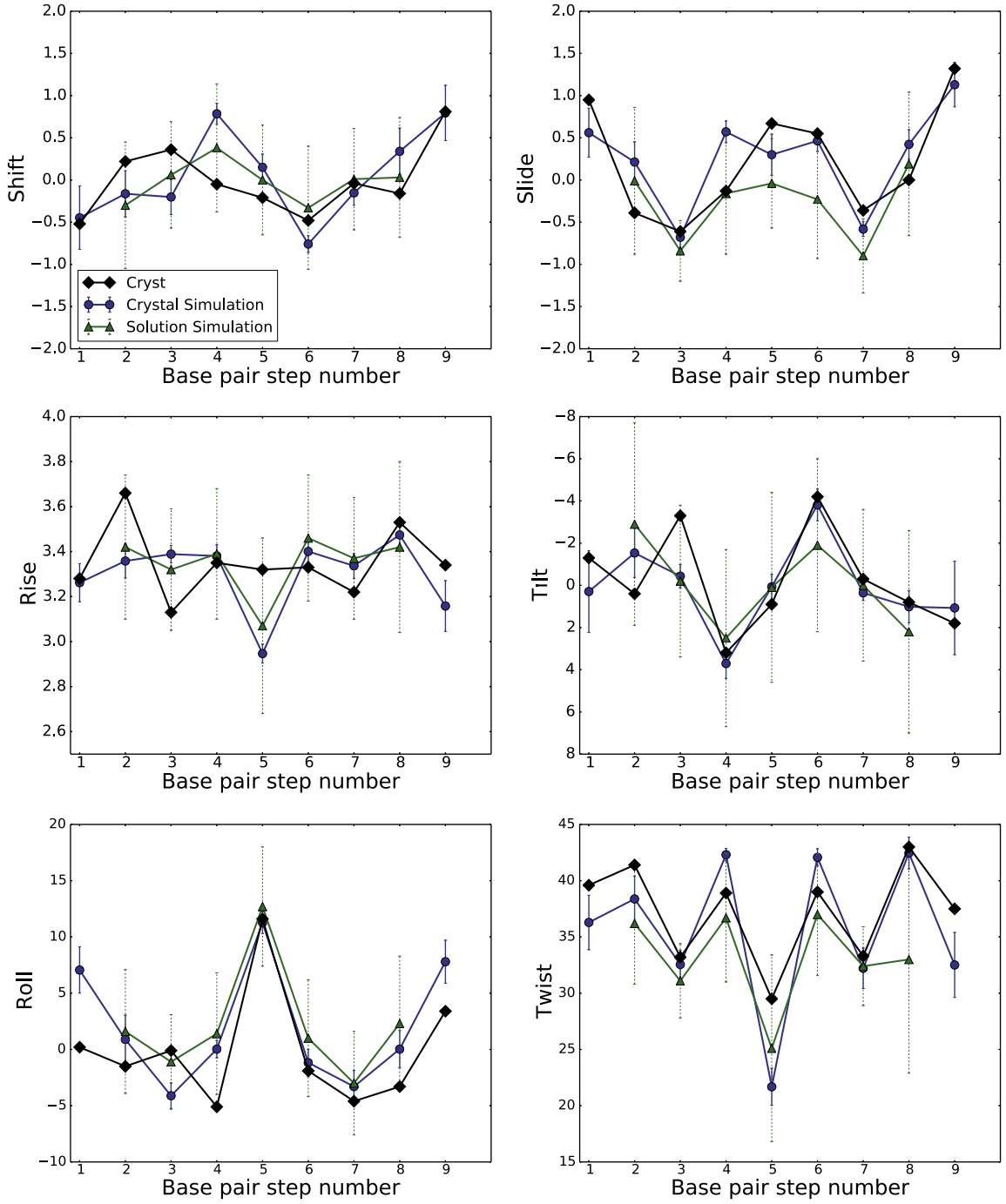

Fig. 7. DNA base pair step parameters, translational parameters are in angstroms (Å) and rotational parameters are in degrees (°).

are a number of observations that can be made. First, the shift and tilt parameters are evened out in solution for RNA; this is much less true for DNA, and is less true for the other base-step parameters. Solution and crystal simulations are in close agreement for the rise, roll and helical twist parameters, especially for DNA. The fluctuations vary little between steps in the crystal simulation, except for the more flexible shift, rise, tilt and twist of the end base pair A13·U16/A14·U15 step. Second, fluctuations in base step parameters in solution (green dashed bars) are much greater than those in the crystal simulation (blue solid bars).

Table 4
The average twist (in degrees) for the RNA and DNA average structures. Twist values were calculated using the program Curves+.

	RNA	DNA
Exp. crystal	31.7	37.3
Sim. crystal	31.4	35.7
Sim. solution	29.8	32.3

The twist of nucleic acid helices is an "emergent" property of force fields, that is difficult to associate with particular interactions or backbone torsion angle preferences. The Amber potentials lead to slightly undertwisted helices in solution simulations [32,53]. There are two unusual twist values (36.7 and 21.3) in the RNA duplex at steps 5 and 10 shown in Fig. 6, which is caused by two distinguished

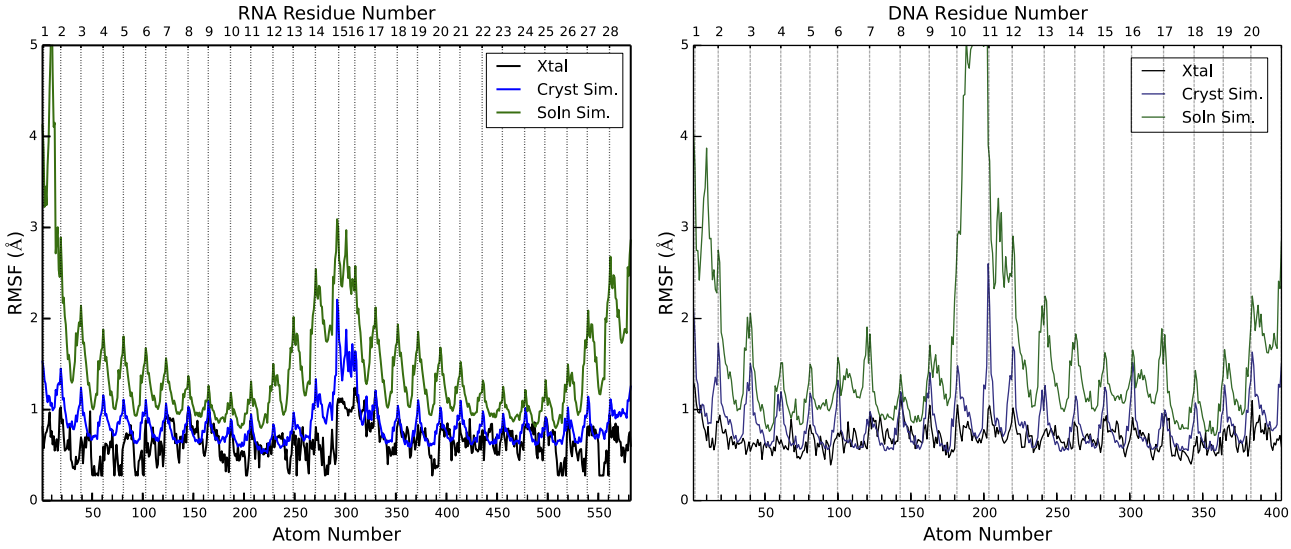

Fig. 8. Root-mean-square fluctuations as function of heavy atom number for RNA (left) and DNA (right). The left half of each figure represents chain A, and the right half, chain B. Light vertical lines identify the phosphate atoms (excluding residues U1A15 for RNA and C1C11 for DNA). B-factors in 1RNA/1D23 (PDB entries) were converted to fluctuations using Eq. (1).

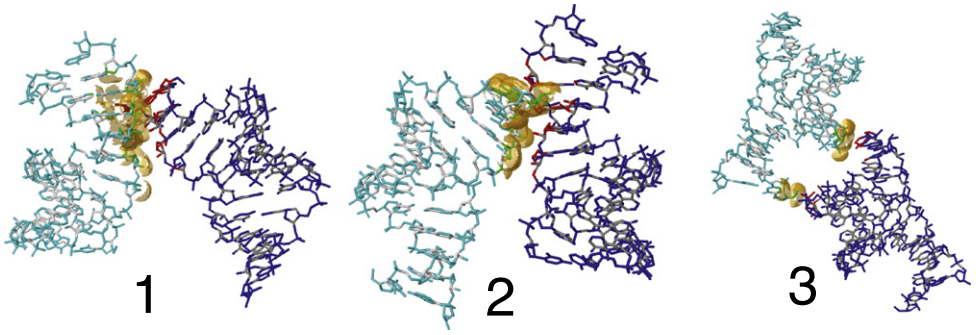

Fig. 9. Lattice contacts, showing the orientation of one RNA asymmetric unit relative to its neighbors. The contacts are shown as defined by PISA.

Table 5
Hydrogen bonds, van der Waals contacts and interactions between symmetry-related duplexes in RNA. The data in parentheses are standard deviations. Interface areas were computed by PISA.

Type	Interactions	Area (Å^2)	Atom 1	Symmetry operator	Atom 2	Xtal. dist.	Cryst. sim. dist.
1	Minor groove–minor groove	311.5	O2/U1	$-x+1, y-1/2, -z+1/2$	O2′/A11	3.01	2.78 (0.20)
			O2′/A28		O2/U18	3.02	2.80 (0.19)
			O3′/A28		O2′/U18	3.00	3.10 (0.23)
			O2′/U1		OP1/U12	3.30	4.28 (0.55)
			O2′/U2		O2′/U10	2.46	4.02 (0.80)
			O2′/A3		O2′/U20	3.20	3.86 (0.86)
2	Minor groove–minor groove	304.3	O2/U15	$x-1/2, -y+1/2, -z$	O2′/A25	2.67	3.31 (0.96)
			O2′/A14		O2/U4	2.77	2.99 (0.39)
			O3′/A14		O2′/U4	3.01	3.69 (1.36)
3	Major groove–major groove	138.8	C6/U15	$-x+1/2, -y+1, z-1/2$	O2′/U22	3.17	4.44 (0.94)
			C5/U15		O3′/U22	2.80	3.75 (0.69)
			O3′/A9		O5′/U1	3.53	4.10 (1.11)
			O3′/A9		C5′/U1	3.30	4.26 (1.10)
			OP1/U10		C5′/U1	3.32	4.19 (1.21)

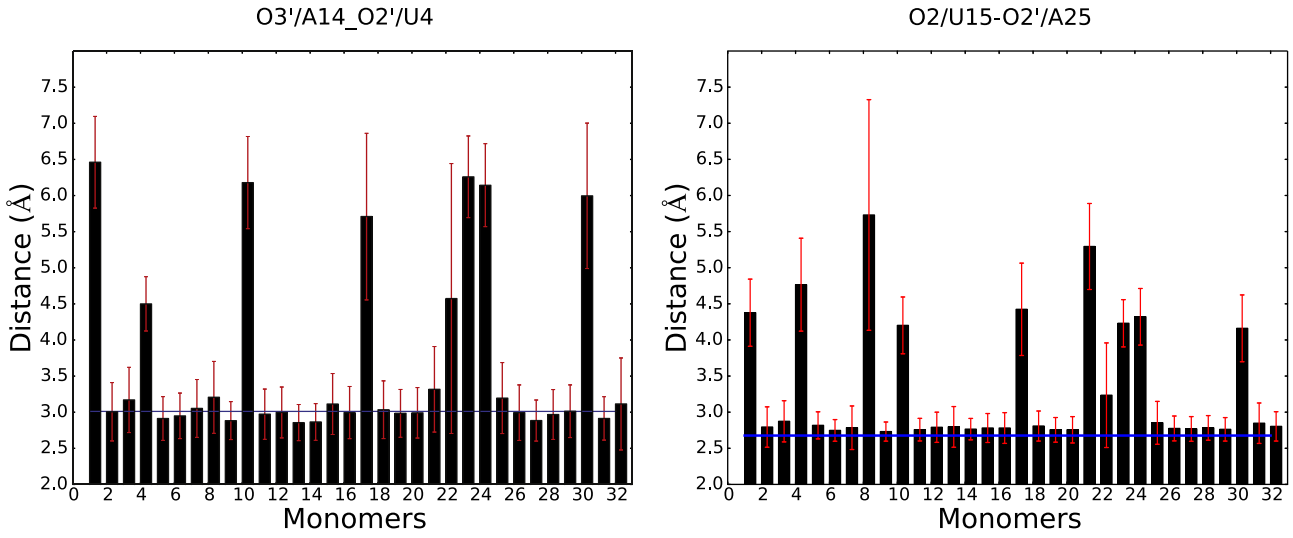

Fig. 10. The average distances of O3′/A14–O2′/U4 and O2/U15–O2′/A25 for the 32 duplexes in RNA supercell for interface 2 (red lines are standard deviations, blue lines are the distance in crystal structure, and black bars give the average distances for every duplex).

kinks dividing the whole duplex into three parts. However, in our simulation, these two unusual twist values in kink regions become 29.5 (28.7 in solution) for step 5 and 31.6 (31.5 in solution) for step 10 in average structures. Table 4 shows that the average twist is larger in the crystal simulation than in solution; it seems likely that the coaxial stacking interfaces (discussed below) serve to prevent the relaxation to an undertwisted state that is typically seen in solution simulations with this force field [3,25–29,32]. Some steps in the solution simulations show evidence of bimodal distributions in the twist [54]; no such behavior is seen in the DNA simulations reported here. This may be due to sequence or crystal-packing effects, and further analysis of these points will be a subject of future study.

3.4. Fluctuations about the average structure

Simulated B-factors can be calculated from an MD trajectory using Eq. (1) and compared to crystallographic B-factors:

$$B = \frac{8\pi^2}{3}\left\langle \mu^2 \right\rangle \tag{1}$$

where μ is the root mean square fluctuation about a mean position. In Fig. 8, the simulated and experimental B-factors for RNA and DNA are reported. For both solution and crystal simulations, the fluctuations are calculated by superimposing each duplex onto the crystal conformation using an optimal rotation and translation movement. The phosphate groups show the highest deviations from the average structure, and the low points between these peaks represent base atoms.

As expected, the fluctuations in solution (green curves) are higher than those in the crystal simulation (blue curves), not only at the chain termini (where large fluctuations and fraying take place in solution), but also in the centers of the duplexes. For RNA, the fluctuations in the crystal simulation are close to experiment except near the chain termini. This is also the case for the base atoms in the DNA (i.e. at the low points between peaks in Fig. 8), but the sugar-phosphate backbone atoms have larger fluctuations in the simulation. Note that the fluctuations reported here do not include the effects of the lattice distortion, which is fairly large for both RNA and DNA, as shown in Figs. 2 and 15; inclusion of these effects would result in fluctuations that are much larger than those shown in Fig. 8 (data not shown).

4. Contacts and packing interactions

Two of the features that distinguish the present simulations from earlier ones are the size of the supercell used and the length of the simulations. We have 32 copies of each RNA or DNA duplex, and hence 32 copies of each of the interactions between duplexes. This allows us to analyze the nature of these interactions, and the ability of current force fields and simulation protocols to correctly describe the behavior of these relatively weak interactions.

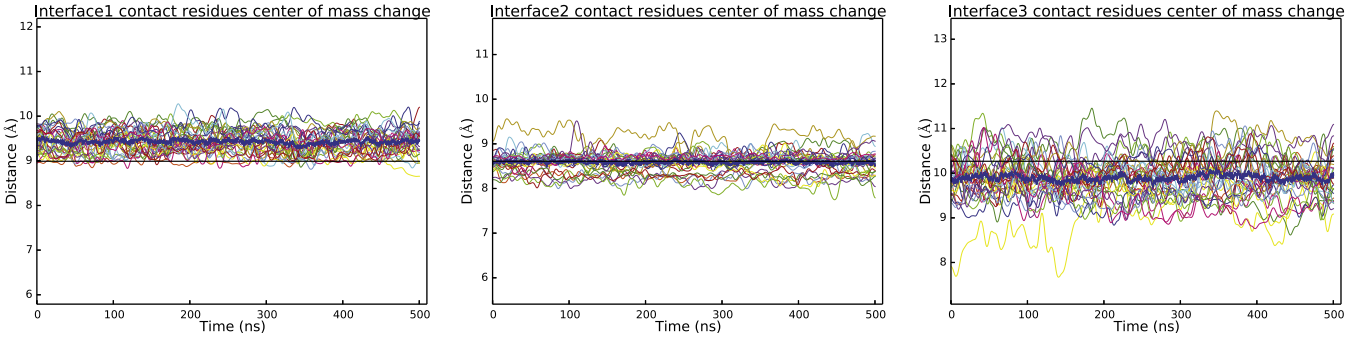

Fig. 11. Distances of center of mass between the interface residues for RNA, for each of the 32 copies of each interface. Black lines show the value from the X-ray structure. Bold blue lines are the average distances between interface centers of mass.

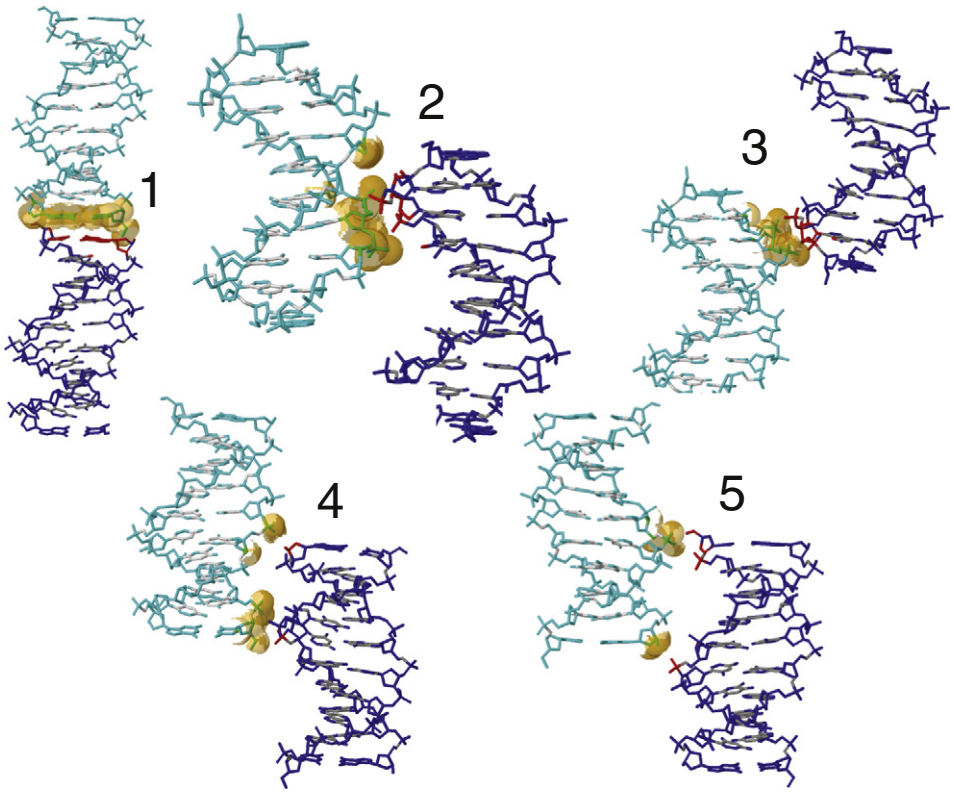

Fig. 12. Same as Fig. 9, but for DNA.

4.1. Crystal interfaces for RNA

In a crystal, each nucleic acid molecule makes several contacts with its neighboring molecules. Generally, the size of the pairwise interface between two neighboring molecules can be characterized by the number of contacts or the buried interface area. *PDBe* PISA (Proteins, Interfaces, Structures and Assemblies) shows three interfaces with nearest neighbors for RNA, shown in Fig. 9 and Table 5. Two interfaces (1 and 2 in Fig. 9) involved in the minor grooves of the symmetry related helices face each other at their ends; hydrogen bonds stabilize both contacts. The third interface places the major grooves of two duplexes face to face with each other. A few van der Waals contacts and CH–O hydrogen bonds (among residues U1, U22 and A9, U10, U15) stabilize this interface. The total surface area buried by the three listed interfaces is 1516 Å^2 (twice the total surface area of the three interfaces listed in Table 5), since in the crystal each duplex participates in two independent copies of each interface, (e.g. one on the "left" and one on the "right"). The total surface area for an isolated duplex is 5134 Å^2, so that 29% of the surface area is buried, and 71% is accessible to water and ions. This buried interface area is smaller than for many proteins:

Table 6
van der Waals contacts and interactions between symmetry-related helices in DNA. The data in parentheses are standard deviations.

Type	Interactions	Area (Å^2)	Atom 1	Symmetry operator	Atom 2	Xtal. dist.	Cryst. sim. dist.
1	Terminal–terminal	199.9	N3/C11	$x, y, z - 1$	N3/G20	3.49	3.51 (0.23)
			O3′/G10		O5′/C1	3.37	3.63 (0.89)
2	Backbone–minor groove	154.2	OP1/A3	$x - 1/2, -y + 1/2, -z$	O4′/A17	3.65	4.77 (1.08)
			P/A3		P/A17	6.49	7.59 (1.53)
			P/T4		P/A17	4.65	5.79 (1.93)
			OP1/T4		OP2/A17	3.88	5.83 (1.33)
			OP2/T4		O5′/A17	3.79	4.62 (0.96)
3	Backbone–backbone	102.7	OP1/A13	$x - 1/2, -y + 1/2, -z + 1$	OP2/C9	3.90	5.24 (1.39)
			O3′/G12		O5′/T8	3.93	5.23 (1.45)
			P/A13		P/C9	5.15	6.65 (1.29)
			P/G12		P/T8	8.25	7.12 (1.08)
4	Minor groove–major groove	76.1	P/G6	$-x + 1/2, -y, z - 1/2$	O3′/G20	3.74	4.29 (1.05)
			OP2/G6		O3′/G20	2.68	3.37 (1.00)
			P/G6		P/G20	8.07	8.36 (1.22)
			P/G12		P/G6	6.28	7.71 (1.10)
5	Major groove–major groove	76.0	P/G16	$-x + 1/2, -y + 1, z - 1/2$	O5′/C1	3.40	4.63 (1.14)
			O3′/G10		P/G16	3.93	3.96 (0.77)
			P/G16		P/G2	6.29	8.09 (1.03)
			P/G10		P/G16	8.36	7.57 (0.96)

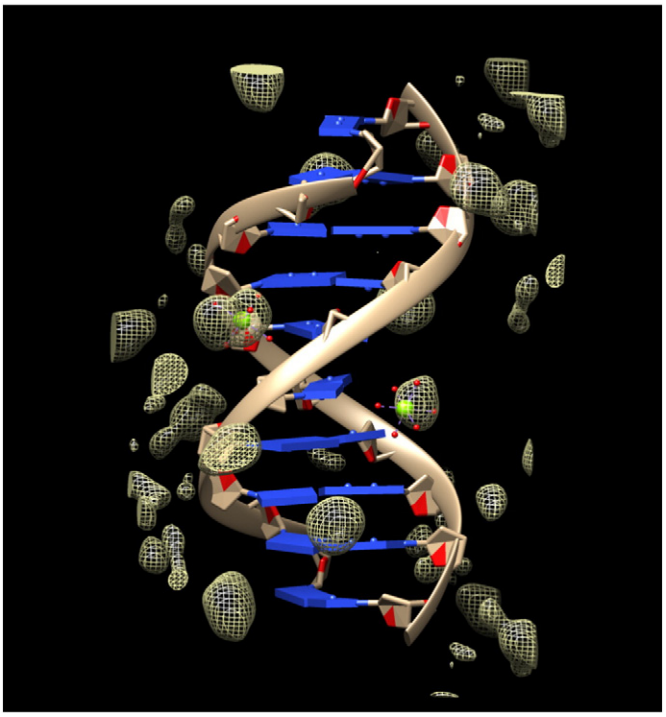

Fig. 13. Cartoon view of the 1D23 crystal structure, showing the two identified Mg^{2+} ions as green spheres with attached water molecules (smaller red spheres). The mesh structure shows the Mg^{2+} distribution from the simulation.

for example, the triclinic form of hen lysozyme (PDB code 4LZT) has 3030 Å^2 (or 46%) of its surface area in contact with other proteins. The relatively small contact area for RNA (and for DNA, as discussed below) may increase the likelihood that the contacts may degrade during an imperfect simulation.

The first and second interfaces for RNA involve hydrogen bond interactions, whereas the third interface for RNA and all interfaces for DNA are characterized by less-specific van der Waals contacts. Most of the contact distances become larger in the simulation for both RNA and DNA; exceptions include most of the hydrogen bonds in RNA interfaces 1 and 2; the longer (and presumably weaker) CH–O interactions in RNA interface 3 are less-faithfully maintained; this interface also has a much smaller buried surface area in the crystal.

The averages shown in Table 5 hide a large amount of heterogeneity among the 32 copies of each interface present in the supercell. As an example, Fig. 10 shows the variation of two of the hydrogen bonds in interface 2 for each copy. It is clear that these interactions are well-maintained in about 22–24 copies, and completely broken in 8–10 copies. It is likely that longer simulations would result in more copies becoming distorted, leading to progressive distortion of the lattice, as is seen more clearly in the longer DNA simulation discussed below. The average distance of the molecular centers of mass of the duplexes interacting across interface 1 increases by 0.5 Å; across interface 3 the distance decreases by about the same amount; across interface 2 it stays almost the same as in the crystal (Fig. 11). Individual interfaces can deviate from the average by up to about 1 Å, indicative of significant heterogeneity.

4.2. Crystal interfaces for DNA

The five interfaces for DNA are shown in Fig. 12 and Table 6. In interface 1 the decamer helices are stacked one on top of another along the *z*-axis leading to co-axial stacking. The second and third interfaces are along the *x* axis with P3, P4 facing P17, P18, and P8 and P9 facing P12 and P13, respectively. For the fourth and fifth interfaces, P6 and P16 are in close proximity to the inter-helix gaps of neighboring duplexes. The surface area buried by these contacts is 1218 Å^2, which is 31% of the total surface area of an isolated duplex.

Hydrated Mg^{2+} frequently mediates intermolecular contacts between adjacent DNA molecules. Evidence from molecular dynamics simulations by Hartmann [35] supports that Mg^{2+} bound to the DNA major groove have an effect on DNA structure and dynamics; hydrated Mg^{2+} often forms a stable intra-strand cross-link between the two purines and increases the BII population. Li et al. [55] have argued that the binding of Mg^{2+} rigidifies DNA compared to Na^+.

Two $Mg(H_2O)_6^{2+}$ clusters are identified in the DNA deposited structure, located at G16·A17 in the minor groove and at G6·A7 in the major groove, adjacent to two of the BII sites at P7 and P17; these are shown in Fig. 13. These sites are also highly occupied with Mg^{2+} in the simulation. Additional ion sites must be present to neutralize the total charge, but were not modeled in the deposited structure. Other relatively well-localized sites for hydrated magnesium are evident in the simulation, both in the grooves and near phosphate positions (see Fig. 13). Almost all of the Mg^{2+} ions remain coordinated to six water molecules throughout the simulation: for 0.58% of the time, the ions have 5 waters in the first coordination shell. All of the ions have approximately the same diffusion constant (about 3×10^{-9} cm^2/s), so that the ions initially placed in crystallographic density have the same mobility as those randomly added to achieve neutrality. A fuller analysis of ion distributions probably requires additional simulations, preferably with less lattice distortion than that seen here, since magnesium ions are involved with bridging interactions between duplexes as well.

The analysis of the deformations in the distances between the centers of mass between the interface residues (Fig. 14) implicates all interfaces in crystal lattice deformations. The coaxial stacking (interface 1) shows small fluctuations and a compression by about 0.5 Å, whereas the other interfaces increase in average length by up to 0.5 Å, and individual copies exhibit much larger variations from the mean. These results are consistent with the larger average values and fluctuations for individual contacts that are shown in Table 6. Deviations from the mean behavior are much greater for DNA than for RNA (compare Figs. 11 and 14.)

The displacements of the centers of mass from their ideal positions (averaged over time) for each duplex in the supercell are shown in Fig. 15. The average distance of each point from its ideal position is 0.37 Å for RNA and 1.77 Å for DNA. (For comparison, in a recent simulation of lysozyme, the mean deviations of the proteins from their ideal lattice positions was 0.27 Å.) Such large translations of some duplexes in the supercell correspond to a significant loss of crystal symmetry, which can also be seen in Fig. 2.

5. Conclusions

We have performed crystal simulations of 1D23 and 1RNA in the "supercell" with 32 copies of each duplex, as well as solution simulation of isolated duplexes. As has been seen in earlier studies, the average duplex structures from crystal simulation match the experimental structure much more closely than the average structures from solution simulations with the same force field. There is a general tendency for the solution simulation to flatten the variations in helicoidal parameters seen in the crystal (and, presumably, stabilized by crystal packing interactions.) Fluctuations about the average structure in the crystal simulation are in good agreement with refined B-factors for base atoms, but are somewhat higher for the sugar-phosphate backbone atoms; fluctuations in solution are considerably higher than in the crystal lattice.

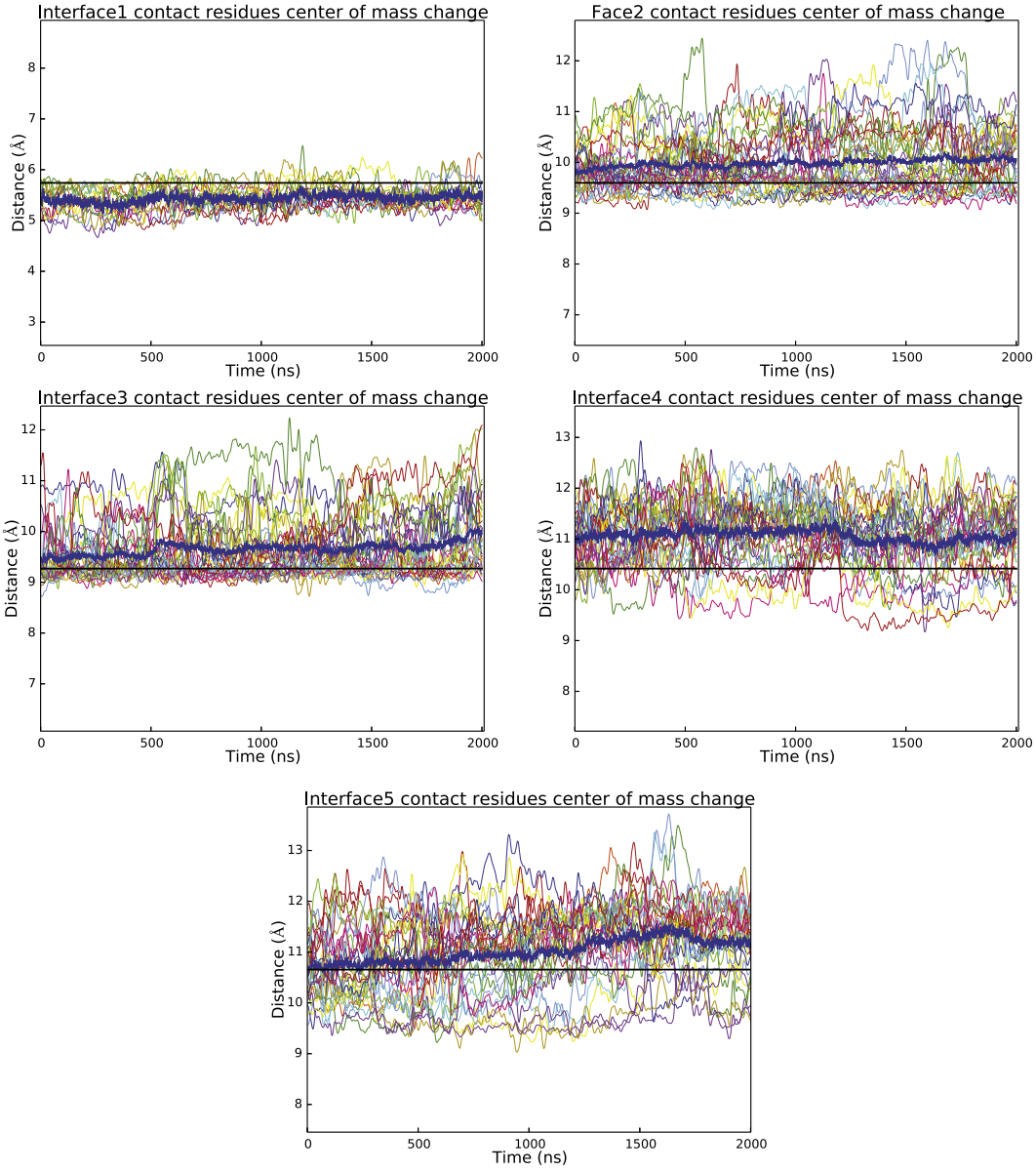

Fig. 14. Distances between the centers of mass between the interface residues for DNA. Black lines show the value from the X-ray structure. Bold blue lines are the average distances between interface centers of mass.

On the other hand, the contacts that stabilize the crystal lattice are not extensive in these crystals, especially for DNA, and our simulations show a progressive deformation of the lattice, such that the duplexes move away from the "proper" positions in the supercell by 1 to 2 Å. This is more true for DNA than for RNA, perhaps by virtue of having fewer hydrogen bonds between duplexes to stabilize the lattice. Another factor may simply be the shorter simulation time for RNA. The results of the present work represent the largest all-atom crystal "supercell" simulations of nucleic acid to date, and provide a clearer understanding of the structural and dynamical parameters for RNA and DNA crystals. They also offer insights into a new challenge for force-field simulations in crystals and non-crystalline applications with weak interactions between chains. Future studies are planned to investigate dependencies on force fields, water models, and methods of representing ion interactions.

Acknowledgements

This work was supported by NIH grant GM103297. We thank Tom Cheatham for the useful discussions.

References

[1] H.S. Kim, B.J. Mhin, C.W. Yoon, C.X. Wang, K.S. Kim, A theoretical study of a Z-DNA crystal: structure of counterions, water and DNA molecules, Bull. Kor. Chem. Soc. 12 (1991) 214–219.
[2] A.D. MacKerell Jr., J. Wiórkiewicz-Kuczera, M. Karplus, An all-atom empirical energy function for the simulation of nucleic acids, J. Am. Chem. Soc. 117 (1995) 11946–11975.
[3] T.E. Cheatham, D.A. Case, Twenty-five years of nucleic acid simulations, Biopolymers 99 (2013) 969–977.
[4] T.E. Cheatham, J.L. Miller III, T. Fox, T.A. Darden, P.A. Kollman, Molecular dynamics simulations on solvated biomolecular systems: the particle mesh Ewald method leads to stable trajectories of DNA, RNA and proteins, J. Am. Chem. Soc. 117 (1995) 4193–4194.

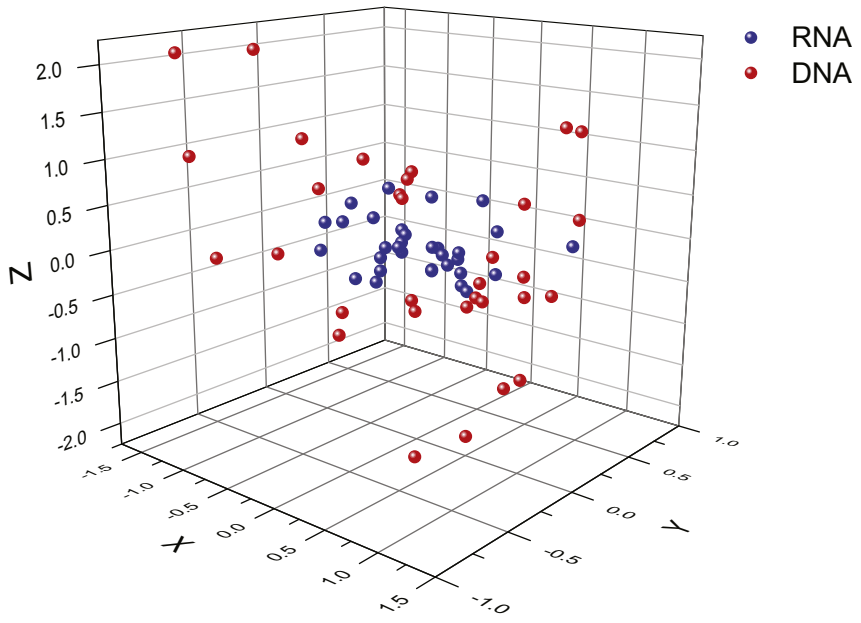

Fig. 15. The center of mass position for each of the 32 duplexes, where the origin represents their position in an ideal lattice.

[5] D.M. York, W. Yang, H. Lee, T. Darden, L.G. Pedersen, Toward the accurate modelling of DNA: the importance of long-range electrostatics, J. Am. Chem. Soc. 117 (1995) 5001–5002.
[6] H. Lee, T. Darden, L. Pedersen, Accurate crystal molecular dynamics using particle-mesh-Ewald: RNA dinucleotides — ApU and GpC, Chem. Phys. Lett. (1995) 229–235.
[7] H. Lee, T.A. Darden, L.G. Pedersen, Molecular dynamics simulations studies of a high resolution Z-DNA crystal, J. Chem. Phys. 102 (1995) 3830–3834.
[8] V. Babin, J. Baucom, T.A. Darden, C. Sagui, Molecular dynamics simulations of DNA with polarizable force fields: convergence of an ideal B-DNA structure to the crystallographic structure, J. Phys. Chem. B 110 (2006) 11571–11581.
[9] V. Babin, J. Baucom, T.A. Darden, C. Sagui, Molecular dynamics simulations of polarizable DNA in crystal environment, Int. J. Quantum Chem. 106 (2006) 3260–3269.
[10] Z. Gong, Y. Xiao, Y. Xiao, RNA stability under different combinations of Amber force fields and solvation models, J. Biomol. Struct. Dyn. 28 (2010) 431–441.
[11] V. Babin, D. Wang, R.B. Rose, C. Sagui, Binding polymorphism in DNA bound state of the Pdx1 homeodomain, PLoS Comput. Biol. 9 (2013) (e1003160).
[12] A. Kuzmanic, B. Zagrovic, Dependence of protein crystal stability of residue charge states and ion content of crystal solvent, Biophys. J. 106 (2014) 677–686.
[13] L.S. Ahlstrom, O. Miyashita, Packing interface energetics in different crystal forms of the lambda Cro dimer, Proteins 82 (2014) 1128–1141.
[14] Z. Hu, J. Jiang, S.I. Sandler, Water in hydrated orthorhombic lysozyme crystal: insight from atomistic simulations, J. Chem. Phys. 129 (2008) 075105.
[15] Z. Hu, J. Jiang, Molecular dynamics simulations for water and ions in protein crystals, Langmuir 24 (2008) 4215–4223.
[16] Z. Hu, J. Jiang, Assessment of biomolecular force fields for molecular dynamics simulations in a protein crystal, J. Comput. Chem. 31 (2009) 371–380.
[17] D.S. Cerutti, I. Le Trong, R.E. Stenkamp, T.P. Lybrand, Dynamics of the streptavidin-biotin complex in solution and in its crystalline lattice: distinct behavior revealed by molecular simulations, J. Phys. Chem. B 113 (2009) 6971–6985.
[18] D.S. Cerutti, P.L. Freddolino, R.E. Duke Jr., D.A. Case, Simulations of a protein crystal with a high resolution X-ray structure: evaluation of force fields and water models, J. Phys. Chem. B 114 (2010) 12811–12824.
[19] P.A. Janowski, D.S. Cerutti, J. Holton, D.A. Case, Peptide crystal simulations reveal hidden dynamics, J. Am. Chem. Soc. 135 (2013) 7938–7948.
[20] A. Kuzmanic, N.S. Pannu, B. Zagrovic, X-ray refinement significantly underestimates the level of microscopic heterogeneity in biomolecular crystals, Nat. Commun. 5 (2014) 3220.
[21] Y. Xue, N.R. Skrynnikov, Ensemble MD simulations restrained via crystallographic data: accurate structure leads to accurate dynamics, Proc. Natl. Acad. Sci. U. S. A. 23 (2014) 488–507.
[22] K. Grzeskowiak, K. Yanagi, G.G. Prive, R.E. Dickerson, The structure of B-helical C-G-A-T-C-G-A-T-C-G and comparison with C-C-A-A-C-G-T-T-G-G. The effect of base pair reversals, J. Biol. Chem. 266 (1991) 8861.
[23] A.C. Dock-Bregeon, B. Chevrier, A. Podjarny, D. Moras, J.S. De Bear, G.R. Gough, P.T. Gilham, J.E. Johnson, High resolution structure of the RNA duplex [U(U-A)6A]2, Nature 335 (1988) 375–379.
[24] A.C. Dock-Bregeon, B. Chevrier, A. Podjarny, J. Johnson, J.S. De Bear, G.R. Gough, P.T. Gilham, D. Moras, Crystallographic structure of an RNA helix: [U(U-A)6A]2, J. Mol. Biol. 209 (1989) 459–474.
[25] I. Besseová, M. Otykpka, K. Réblová, J. Sponer, Dependence of A-RNA simulations on the choice of force field and salt strength, Phys. Chem. Chem. Phys. 11 (2009) 10701–10711.
[26] I. Besseova, P. Banas, P. Kuhrova, M. Kosinova, M. Otyepka, J. Sponer, Simulations of A-RNA duplexes. The effect of sequence, solute force field, water model, and salt concentration, J. Phys. Chem. B 116 (2012) 9899–9916.
[27] S.B. Dixit, D.L. Beveridge, D.A. Case, T.E. Cheatham III, E. Giudice, F. Lankas, R. Lavery, J.H. Maddocks, R. Osman, H. Sklenar, K.M. Thayer, P. Varmai, Molecular dynamics simulations of the 136 unique tetranucleotide sequences of DNA oligonucleotides. II. Sequence context effects on the dynamical structures of the 10 unique dinucleotide steps, Biophys. J. 89 (2005) 3721–3740.
[28] R. Lavery, K. Zakrzewska, D. Beveridge, T.C. Bishop, D.A. Case, T. Cheatham, S. Dixit III, B. Jayaram, F. Lankas, C. Laughton, J.H. Maddocks, A. Michon, R. Osman, M. Orozco, A. Perez, N. Spackova, J. Sponer, A systematic molecular dynamics study of nearest-neighbor effects on base pair and base pair step conformations and fluctuations in B-DNA, Nucleic Acids Res. 38 (2010) 299–313.
[29] D.L. Beveridge, T.E. Cheatham III, M. Mezei The, ABCs of molecular dynamics simulations on B-DNA, circa 2012, J. Biosci. 37 (2012) 379–397.
[30] M. Pasi, J.H. Maddocks, D. Beveridge, T.C. Bishop, D.A. Case, T. Cheatham III, B. Jayaram, F. Lankas, C. Laughton, J. Mitchel, R. Osman, M. Orozco, D. Petkeviciute, N. Spackova, J. Sponer, K. Zakrzewska, R. Lavery, microABC: a systematic microsecond molecular dynamics study of tetranucleotide sequence effects in B-DNA, Nucleic Acids Res. (2014) (in press).
[31] A. Perez, F.J. Luque, M. Orozco, Dynamics of B-DNA on the microsecond time scale, J. Am. Chem. Soc. 129 (2007) 14739–14745.
[32] A. Pérez, F.J. Luque, M. Orozco, Frontiers in molecular dynamics simulations of DNA, Acc. Chem. Res. 45 (2011) 196–205.
[33] S. Teletchea, B. Hartmann, J. Kozelka, Discrimination between BI and BII conformational substates of B-DNA based on sugar-base interproton distances, J. Biomol. Struct. Dyn. 21 (2004) 489–494.
[34] B. Heddi, N. Foloppe, N. Bouchemal, E. Hantz, B. Hartmann, Quantification of DNA BI/BII backbone states in solution. Implications for DNA overall structure and recognition, J. Am. Chem. Soc. 128 (2006) 9170–9177.
[35] M. Guéroult, O. Boittin, O. Mauffret, C. Etchbest, B. Hartmann, Mg^{2+} in the major groove modulates B-DNA structure and dynamics, PLoS One 7 (2012) e41704.
[36] D.A. Case, T.E. Cheatham, T. Darden III, H. Gohlke, R. Luo, K.M. Merz Jr., A. Onufriev, C. Simmerling, B. Wang, R. Woods, The Amber biomolecular simulation programs, J. Comput. Chem. 26 (2005) 1668–1688.
[37] J. Wang, P. Cieplak, P.A. Kollman, How well does a restrained electrostatic potential (RESP) model perform in calculating conformational energies of organic and biological molecules? J. Comput. Chem. 21 (2000) 1049–1074.
[38] A. Perez, I. Marchan, D. Svozil, J. Sponer, T.E. Cheatham, C.A. Laughton, M. Orozco, Refinement of the AMBER force field for nucleic acids: improving the description of alpha/gamma conformers, Biophys. J. 92 (2007) 3817–3829.
[39] M. Zgarbova, M. Otyepka, J. Sponer, A. Mladek, P. Banas, T.E. Cheatham, P. Jurecka, Refinement of the Cornell et al. nucleic acids force field based on reference quantum chemical calculations of glycosidic torsion profiles, J. Chem. Theory Comput. 7 (2011) 2886–2902.
[40] W.L. Jorgensen, J. Chandrasekhar, J. Madura, M.L. Klein, Comparison of simple potential functions for simulating liquid water, J. Chem. Phys. 79 (1983) 926–935.
[41] H.W. Horn, W.C. Swope, J.W. Pitera, J.D. Madura, T.J. Dick, G.L. Hura, T. Head-Gordon, Development of an improved four-site water model for biomolecular simulations: TIP4P-Ew, J. Chem. Phys. 120 (2004) 9665–9678.

[42] I.S. Joung, T.E. Cheatham III, Determination of alkali and halide monovalent ion parameters for use in explicitly solvated biomolecular simulations, J. Phys. Chem. B 112 (2008) 9020–9041.
[43] J. Åqvist, Ion–water interaction potentials derived from free energy perturbation simulations, J. Phys. Chem. 94 (1990) 8021–8024.
[44] R. Salomon-Ferrer, D.A. Case, R.C. Walker, An overview of the Amber biomolecular simulation package, WIREs Comput. Mol. Sci. 3 (2013) 198–210.
[45] R. Salomon-Ferrer, A.W. Götz, D. Poole, S. Le Grand, R.C. Walker, Routine microsecond molecular dynamics simulations with AMBER on GPUs. 2. Explicit solvent particle mesh Ewald, J. Chem. Theory Comput. 9 (2013) 3878–3888.
[46] D.L. Beveridge, G. Barreiro, K.S. Byun, D.A. Case, T.E. Cheatham, S.B. Dixit, E. Giudice, F. Lankas, R. Lavery, J.H. Maddocks, Molecular dynamics simulations of the 136 unique tetranucleotide sequences of DNA oligonucleotides. I. Research design and results on d(CpG) Steps, Biophys. J. 87 (2004) 3799–3813.
[47] R. Lavery, M. Moakher, J.H. Maddocks, D. Petkeviciute, K. Zakrzewska, Conformational analysis of nucleic acids revisited: Curves+, Nucleic Acids Res. 37 (2009) 5917–5959.
[48] E. Krissinel, K. Henrick, Inference of macromolecular assemblies from crystalline state, J. Mol. Biol. 372 (2007) 774–797.
[49] B.S. Tolbert, Y. Miyazaki, S. Barton, B. Kinde, P. Starck, R. Singh, A. Bax, D.A. Case, M.F. Summers, Major groove width variations in RNA structures determined by NMR and impact of C-13 residual chemical shift anisotropy and H-1-C-13 residual dipolar coupling on refinement, J. Biomol. NMR 47 (2010) 205–219.
[50] D.R. Mack, T.K. Chiu, R.E. Dickerson, Intrinsic bending and deformability at the T-A step of CCTTTAAAGG: a comparative analysis of T-A and A-T steps within A-tracts, J. Mol. Biol. 312 (2001) 1037–1049.
[51] M. Poncin, B. Hartmann, R. Lavery, Conformational substates in B-DNA, J. Mol. Biol. 226 (1992) 775–794.
[52] R.E. Dickerson, Definitions and nomenclature of nucleic acid structure parameters, J. Biomol. Struct. Dyn. 6 (1989) 627–634.
[53] I. Faustino, A. Pérez, M. Orozco, Toward a consensus view of duplex RNA flexibility, Biophys. J. 99 (2010) 1876–1885.
[54] P.B. Dans, A. Pérez, I. Faustino, R. Lavery, M. Orozco, Exploring polymorphisms in B-DNA helical conformations, Nucleic Acids Res. 40 (2012) 10668–10678.
[55] W. Li, L. Nordenskiöld, Y. Mu, Sequence-specific Mg^{2+}–DNA interactions: a molecular dynamics simulation study, J. Phys. Chem. B 115 (2011) 14713–14720.