THE JOURNAL OF PHYSICAL CHEMISTRY Letters

Perspective

pubs.acs.org/JPCL

Molecular Dynamics Simulations of Nucleic Acids. From Tetranucleotides to the Ribosome

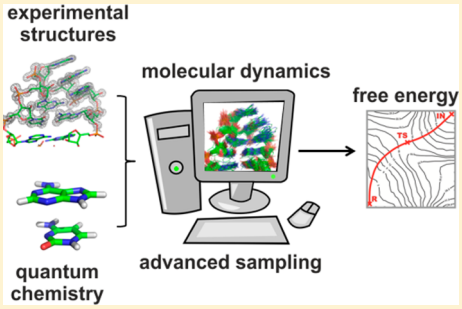
Jiří Šponer,*[,†,‡] Pavel Banáš,[§] Petr Jurečka,[§] Marie Zgarbová,[§] Petra Kührová,[§] Marek Havrila,[†,‡] Miroslav Krepl,[†] Petr Stadlbauer,[†] and Michal Otyepka[§]

[†]Institute of Biophysics, Academy of Sciences of the Czech Republic, Královopolská 135, 612 65 Brno, Czech Republic

[‡]CEITEC − Central European Institute of Technology, Campus Bohunice, Kamenice 5, 625 00 Brno, Czech Republic

[§]Regional Centre of Advanced Technologies and Materials, Department of Physical Chemistry, Faculty of Science, Palacký University, tř. 17 listopadu 12, 771 46 Olomouc, Czech Republic

ABSTRACT: We present a brief overview of explicit solvent molecular dynamics (MD) simulations of nucleic acids. We explain physical chemistry limitations of the simulations, namely, the molecular mechanics (MM) force field (FF) approximation and limited time scale. Further, we discuss relations and differences between simulations and experiments, compare standard and enhanced sampling simulations, discuss the role of starting structures, comment on different versions of nucleic acid FFs, and relate MM computations with contemporary quantum chemistry. Despite its limitations, we show that MD is a powerful technique for studying the structural dynamics of nucleic acids with a fast growing potential that substantially complements experimental results and aids their interpretation.



With the advance of fast computers, explicit solvent molecular dynamics (MD) simulations of nucleic acids (Figure 1) have become widespread.[1−7] Whereas early simulations were executed on a time scale of 1 ns, recent studies have reported microsecond-scale simulations, with millisecond time scales in sight.[8,9] Contemporary simulations have investigated a broad spectrum of nucleic acid systems, ranging from canonical helices,[10,11] tetranucleotides,[12−15] hairpin loops,[16−19] quadruplexes,[1,20,21] and RNA motifs,[22−29] through medium-sized systems (ribozymes, riboswitches, rRNA segments, etc.),[30−43] and up to the entire ribosome[44−46] and nucleosome.[47,48]

MD directly captures structural dynamics and indirectly may assess free energies. MD aims to obtain new information about nucleic acids that is not accessible by experiment to help interpret experiments and provide predictions. Simulations should attempt to reproduce known experimental data, although straightforward comparison is not always possible due to a lack of unambiguous experiments equivalent to the simulation (see below). A textbook example is the elusive thermodynamics and kinetics of BI/BII DNA backbone substates.[49,50]

Standard simulations can capture only the motions that a real molecule could achieve on the same time scale.

MD provides unprecedented temporal and spatial resolution of molecular motion as the position of each atom (including solvent) can be monitored within a subpicosecond-scale window. The main drawbacks relate to (i) the description of the

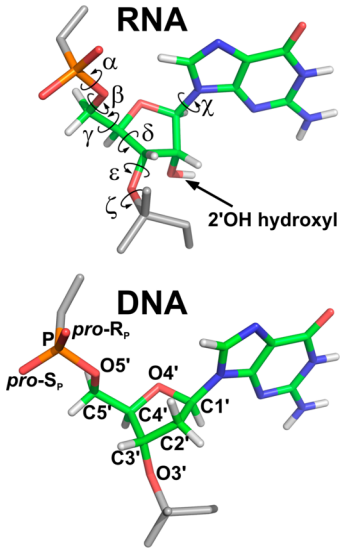


Figure 1. Structures of RNA and DNA nucleotides with depicted torsion angles and atom labels. The labeling of atoms and torsion angles is identical for both DNA and RNA nucleotides (except that the 2′-OH is missing in DNA). The 2′-OH group of the RNA nucleotide stabilizes the C3′-endo sugar pucker, favoring the A-form helix, makes RNA chemically reactive, and (most importantly) promotes key hydrogen bond donor/acceptor interactions, dramatically increasing the conformational versatility of RNA.

Received: March 19, 2014
Accepted: May 5, 2014
Published: May 5, 2014

   dx.doi.org/10.1021/jz500557y | J. Phys. Chem. Lett. 2014, 5, 1771−1782

structure−energy relation (potential energy function) using a simple parametrized molecular mechanics (MM) force field (FF) and (ii) the time scale of the simulations. The reliability of the FFs can be indirectly assessed by comparing simulation results with experimental data (Figure 2). Direct insight into the physical limitations of the MM can be obtained from the highest-accuracy quantum chemistry (QM) computations (Figure 3).[51−55]

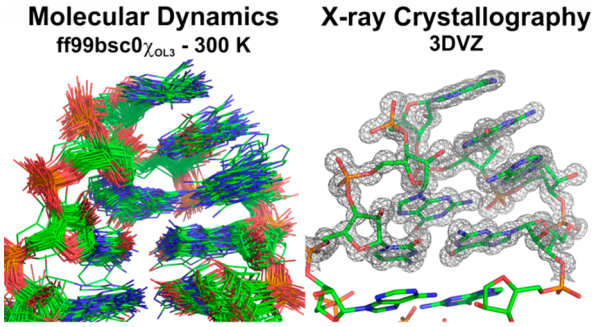


Figure 2. Comparison of theoretical (left) and experimental (right) structures of a GAGA tetraloop from the tip of the 23S rRNA sarcin−ricin loop (SRL) domain. The electron density map (1.0 σ) of the high-resolution X-ray structure (PDB code 3DVZ) shows a clear, albeit static and averaged, picture of the tetraloop. MD simulation (overlay of 50 snapshots) monitors the dynamics with picosecond-scale time resolution. In contrast to NMR, MD simulation follows the time evolution of a single molecule and shows the molecular configurations directly. Structure propagation is described from a given starting configuration and is usually not long enough to reach ergodicity, which is necessary for equilibrium description of the studied system. The MD picture may be affected by the quality of the chosen FF (see the text). This particular simulation agrees with the starting X-ray structure in all details.[17]

Large amounts of simulation data are difficult to process without loss of useful information. Thus, effective data mining techniques should be used to avoid unwanted loss of important features. Many early studies analyzed the interactions between solvent and nucleic acids, such as long-residency water bridges and highly occupied cation-binding sites.[4−6,22−26,30−32,36] However, because such analyses are tedious, many contemporary studies have omitted them.

For every scientific tool, reproducibility and assessment of the limitations are important. The QM literature has paid an enormous amount of attention to verification of data through benchmark databases.[53,54] QM studies are easily reproducible because they report molecular structures and energies. Likewise, deposited atomic coordinates and electron densities allow assessment of X-ray structures. However, rigorous quality assessment of simulations is difficult because of their stochastic nature and overwhelming amount of data. There are currently no established validation standards for nucleic acid simulations. This simplifies publications but also makes the results of simulations prone to overinterpretation. While many MD papers have provided careful description of the simulation protocols and report details of the behavior, some studies have presented limited analyses or failed to clearly define which FF version was used. Whereas the trend nowadays is to write short papers, the nature of simulation data does not lend itself to excessive compaction. Studies not documenting structural details may hide major inaccuracies, and their biochemical interpretations may be based on simulations progressively deviating from the native state.

Although different simulation tasks may have different requirements with regard to the accuracy of the MM description, assessment of the limitations is always essential to interpret the data properly.

Simulations of small systems should reproduce all atomistic details. Although recent FF refinements have enabled all signature interactions in UNCG and GNRA RNA tetraloops (Figure 4) to be reproduced for the first time,[17−19] none of the existing FFs can reproduce the populations of different structures of RNA tetranucleotides known from benchmark NMR experiments (Figure 4).[12−15] This highlights two limitations; (i) absolute accuracy of the MM description is not yet guaranteed even for simple systems, and (ii) different systems (and even different parts of a given system) may be described by MM with different accuracy, depending on the system-specific balance of different energy contributions. Thus, quadruplex DNA (G-DNA) stems are described by MM better than their

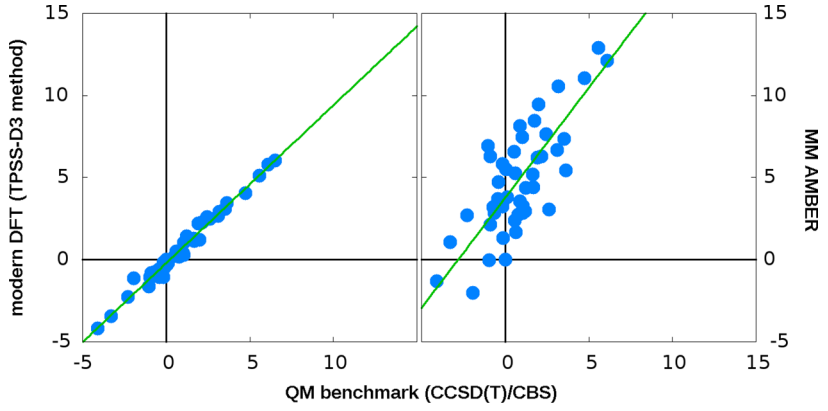


Figure 3. Relative energies (on the potential energy surface, in kcal/mol) of 46 RNA backbone families; a canonical A-RNA conformation is used as the reference point. Benchmark calculations (x-axis) are obtained using the CCSD(T) method with a complete atomic orbital basis set, which is an established gold standard in contemporary computational chemistry. Modern dispersion-corrected density functional theory DFT-D3 (left) closely reproduces the benchmark values. In contrast, pair-additive AMBER MM (right) is much less accurate. The MM energy variations are exaggerated, which means that the MM potential energy surface has insufficient flexibility.[54]

dx.doi.org/10.1021/jz500557y | J. Phys. Chem. Lett. 2014, 5, 1771−1782

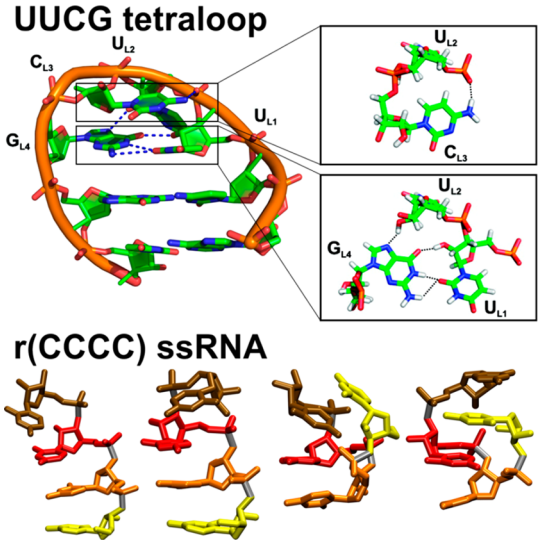


Figure 4. Success of the MM description may differ for different systems. Whereas MD with the latest FF refinements can reproduce all signature interactions of the UUCG RNA tetraloop (top), MD populations of conformational substates of r(CCCC) ssRNA remain inconsistent with benchmark NMR measurements (bottom). The figure shows the four most populated r(CCCC) substates, but whereas the two on the left agree with experiment, the other two are spurious.[12−15,17−19]

single-stranded loops.[20] It is usually more difficult to achieve a fully balanced description for larger systems. Fortunately, we often do not need absolute accuracy. Thus, although a key apical sugar−base H-bond is usually (incorrectly) replaced by a water bridge in simulations of RNA kink-turns, the simulations still capture their biochemically relevant nonharmonic and anisotropic flexibility (Figure 5).[56] The flexibility of kink-turns cannot be studied by any other technique, and derivation of the flexibility of nucleic acids is one area where MD simulations can complement static experimental structures.[4,25,43] Similarly, MD simulations do not reproduce details of the S-turn and dinucleotide platform in the sarcin−ricin loop (SRL) RNA motif (Figure 6).[28] Nevertheless, MD provides reliable insights into the

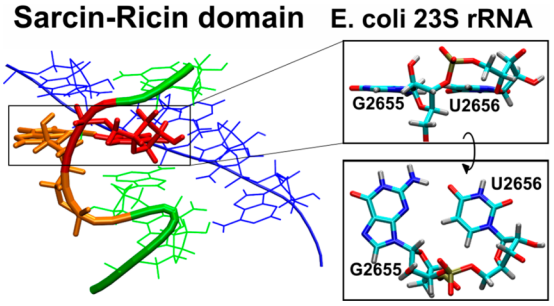


Figure 6. RNA S-turn (left, in color) is an important part of the sarcin−ricin motif. It involves a dinucleotide GpU platform (right), sugar double flip-over, highly specific backbone conformations with anomalous high *anti*-χ and low-β dihedral angles, base−phosphate interactions, and other noncanonical features. There are many highly conserved SRL motifs in the ribosome and some other large RNAs. Their 3D structures conserve every single H-bond and backbone detail, dictating the SRL consensus sequence. The SRL motif poses a particular challenge for MM.[28]

effects of various base substitutions in the SRL.[28] Simulations of G-DNA stems are affected by a lack of polarization associated with binding of multiple closely spaced ions in their channel.[53,57] However, MD can provide valuable insights into the properties of G-DNA, many of which are unobtainable by experiment. MD-based free-energy computations correctly predicted favorable syn−anti distributions of nucleotides in antiparallel G-DNA stems.[58] Nevertheless, the FF used was not sufficiently accurate to predict the relative energy balance between parallel and antiparallel stems.[55] Although simulations of the hairpin ribozyme

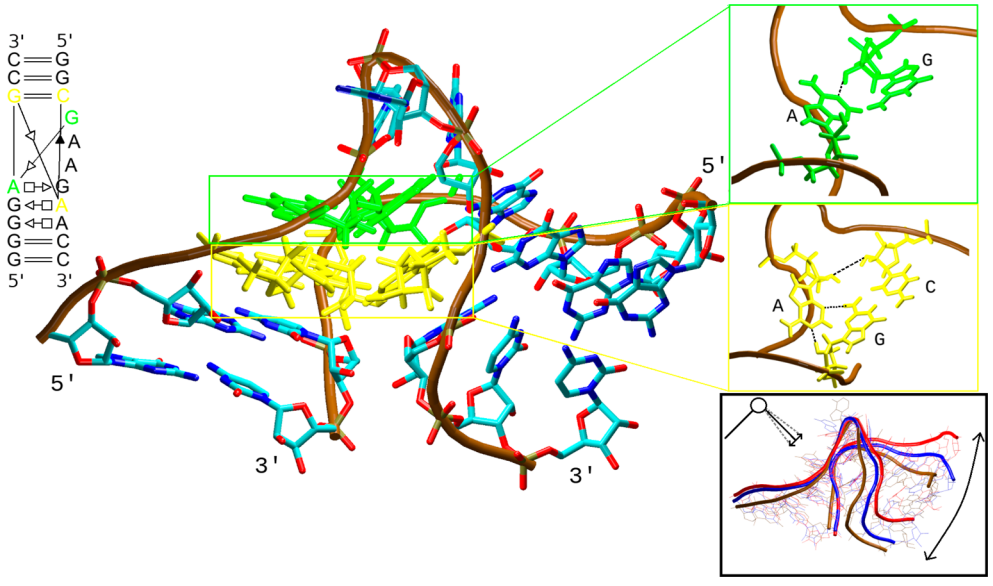


Figure 5. 2D (standard annotation) and 3D structures of a typical kink-turn. The kink-turn is a common tertiary structural motif in RNA that creates a sharp bend between two RNA helices. The essential features of kink-turns include a three-nucleotide bulge, a signature sugar−base interaction between a 2′-OH group from the first bulged nucleotide and N1 of a conserved adenine in the first AG base pair (green inset) and an A-minor interaction (most common tertiary interaction in RNA; yellow) between the stems. Although the signature interaction is not fully stable in simulations, the overall description is satisfactory. MD has revealed that kink-turns can function as flexible passive molecular hinges (black inset). Water-mediated fluctuations of the A-minor interaction substantially contribute to the flexibility.[25,56]

dx.doi.org/10.1021/jz500557y | J. Phys. Chem. Lett. 2014, 5, 1771−1782

with older FFs resulted in structural collapse (ladder-like structure) of one of the stems, this peripheral structural degradation did not affect the active site.[59] Thus, minor structural imperfections or local structural deviations that are well separated from the rest of the structure often do not invalidate the results but need to be judged case by case. For example, investigations of the catalytic centers of ribozymes have higher "atomistic" accuracy requirements than studies of the overall structural flexibility of folded RNAs. It would be futile to expect atomistic accuracy in challenging simulations of the overall dynamics of the ribosome[3,44−46] because unavoidable uncertainties in the experimental (starting) structures are likely to be more significant than the FF errors. Nevertheless, atomistic simulations of the entire ribosome can provide unique insights that complement experiments and coarse-grained models of the large-scale dynamics of the ribosome.[3,60] For example, biased targeted MD simulations suggested a plausible pathway of tRNA accommodation into the ribosome starting from a partially bound A/T conformation to the final A/A state.[44] The prediction agreed with the accommodation corridor inferred from biochemical studies. Subsequent microsecond-scale studies provided estimates of diffusion coefficients of tRNA movements in the ribosome.[61] Explicit solvent MD also characterized coupling between movement of the L1 stalk of the large ribosomal subunit and tRNA bound at the ribosomal exit site.[46] Also noteworthy is utilization of MD simulations in flexible fitting of atomic structures into electron microscopy maps.[45]

Interpretation of Simulations. A straightforward way to understand the meaning of MD results is to view the simulation as a hypothetical single-molecule solution experiment. The studied system is initially fixed in an exactly defined starting geometry (a single configuration of atomic coordinates, usually experimental structure). It is then immersed in a solvent box and undergoes room-temperature dynamics. Although simulations are very short compared to actual biochemical processes, they can provide useful information about the properties of the simulated molecules. However, unrestrained standard simulations can only capture the dynamics of a single molecule for the particular simulation time scale and starting structure used. If an unbiased simulation deviates substantially from an established experimental structure over a few tens of nanoseconds, then the simulation likely does not visualize the genuine unfolding pathway (cf. experimental unfolding rate constants) but rather reflects either MM inaccuracy or some high-energy strain in the starting structure. Such behavior should be interpreted with care because, in principle, a FF that well reproduces the properties of real molecules should not lead to unfolding of established experimental structures in standard simulations. A key requirement for a reliable FF is its ability to describe stable conformational dynamics of native structures of biomolecules that are well supported by experimental methods, such as X-ray crystallography and/or NMR. For most systems, standard simulations should show just local thermal fluctuations near the X-ray or NMR structures. Although the thermodynamic equilibrium includes minor populations of unfolded species (or single strands when duplexes are studied), real single-molecule folding and refolding events can hardly occur on the nanosecond to microsecond time scale. In simulations, larger changes can be seen only with techniques accelerating conformational sampling (see below). When unfolding in simulations is caused by FF imbalances or inaccuracies in the starting structures, it may not accurately reflect real processes.

Starting Structure. The quality of the starting geometry critically affects the simulation because (i) the simulation time scale may be insufficient to overcome barriers to conformational changes that remove potential deficiencies of the starting structures and (ii) the MM description may not be robust enough to find the relevant structure.[4,62] For example, if a starting RNA structure has incorrectly determined the χ glycosidic angle (syn instead of anti, or vice versa, a rather common problem in refinement of X-ray structures of large RNAs),[56] the simulation is unlikely to repair this inaccuracy. Similarly, incorrectly placed Mg^{2+} can bias subsequent simulations. It is because of slow exchange of its first-shell ligands (~1.5 μs for water)[63] and poor FF description of divalents.[4] Simulations may help to detect inaccuracies in experimental structures as rapid structural changes often take place shortly after the simulation starts. For large systems, the starting structure may cause problems even if it is basically correct from the experimental point of view. The unprocessed experimental coordinate file is always high in MM potential energy, and hence, careful equilibration is needed. However, elimination of the excess energy for complicated systems with dense networks of molecular interactions may derail the simulation even after careful equilibration. Our unpublished data for protein−RNA complexes indicated that it may take several microseconds to determine whether the distortions propagate or the structure relaxes back.

Previously, we attempted to elucidate the catalytic role of the C75 nucleotide of the hepatitis delta virus ribozyme (HDVr).[30] The simulations were based on an available X-ray structure of the HDVr precursor inactivated by U75 substitution. This precursor structure was consistent with canonical C75 acting as a general base.[64] Indeed, the simulations identified structures with C75 poised to act as the general base,[30] and this mechanism was also found to be energetically feasible by QM/MM calculations.[65] However, key biochemical data suggested that C75 is N3-protonated before the cleavage and acts as a general acid.[66] We tried many unsuccessful simulations with protonated C75 to find an appropriate arrangement for the C75 general acid mechanism. The issue was resolved by a new X-ray structure that was consistent with the general acid role of C75.[67] In the earlier structure, the catalytic center was reconformed by the inactivating C75U substitution, and our simulation time scale was not sufficient to alleviate the initial structural bias.

The most conservative task for simulations is to investigate known experimental structures and their straightforward modifications (such as base substitutions,[23,28] different protonation states,[33,59] etc.). The most ambitious tasks are structure predictions and folding studies,[7] which also represent the ultimate test of MM performance.[12,14,15] Enhanced sampling methods applied with inaccurate FFs will sample spurious free-energy minima.[12−15,17]

Simulation versus Experimental Conditions. It is often assumed that solution simulations mimic solution experiments. However, this is an oversimplification. Simulations are generally too short to capture true thermodynamic equilibrium. Obviously, if the simulations start from a configuration that well represents the equilibrium ensemble, typically the experimental structure, then the simulations sample various configurations belonging to this ensemble. Sampling of other substates (large conformational transitions) is usually not observed as the molecules remain trapped near the starting structure by the potential energy wells. If we consider that a simple event such as base pair breathing occurs on a millisecond time scale,[68] the time limitation of unbiased simulations is clearly apparent.

dx.doi.org/10.1021/jz500557y | J. Phys. Chem. Lett. 2014, 5, 1771−1782

Still, the essentially unlimited time and coordinate resolution of simulations can give unique information not accessible by experiment, provided that the FF is adequate. In some cases, the short time scale of simulations can even be advantageous. Functional geometries of some recurrent RNA building blocks (kink-turns, 5′-UAA/5′-GAN internal loops, reverse kink-turns)[25,69,70] are stabilized by additional interactions in the context of the ribosome. These RNA moduli are remodeled (have different tertiary structure) in the functional state compared to their global thermodynamic equilibrium in isolation. The evolutionary reason for this is that the remodeled (functional) structures provide some salient topological or molecular recognition features that can be utilized by the surrounding ribosomal elements to achieve specific functions. The native (functional) structures of these building blocks are not captured by solution NMR methods, which instead show the nonfunctional equilibrium structure. In contrast, simulations can probe the properties of these RNAs in their relevant topologies by subtracting them from their full-context X-ray structures and then simulating them in isolation, provided that the simulations keep the functional fold sufficiently long. Obviously, such investigations assume that the neglected tertiary interactions do not make the dynamics irrelevant. However, this is usually a plausible assumption. The remodeled functional topologies typically correspond to quite well defined local free-energy minima. The intrinsic flexibilities and dynamics of these building blocks (in their folded topologies) and their surroundings are carefully tuned by evolution to be mutually adapted and complementary. In other words, the surrounding elements cooperate with the intrinsic properties of the remodeled RNA motifs.

We suggest that the best way to understand simulation studies is to imagine how a given molecule would behave under experimental conditions (including the starting structure) and time scale *identical* to the computer experiment. Then, it becomes clear where the simulation conditions do not match the experiment (and vice versa). Taking an initial single configuration of the biomolecule is not equivalent to any experimental condition. However, whereas it is often considered a disadvantage, it can also be used to test various modifications of the starting structures.[23,28,33] The molecule is surrounded by a finite amount of solvent and ions under periodic boundary conditions. To speed up the simulations, the solvent boxes are usually quite small, and the ion atmosphere does not form the true bulk background.[71] A typical sized box (number of waters) corresponds to a large solute concentration of ∼5−50 mM. However, periodicity prevents any contact of the solute molecules in different boxes provided that the box size is sufficiently large. These conditions are also different from any experimental setup and should be considered when comparing simulations with experiments. The above explanations should in no case be taken as an argument against the simulation technique because also all experimental data are affected by the specific conditions of the techniques used. An experimental technique equivalent to the simulation conditions (and having the same resolution) would be invaluable and widely used. However, when interpreting the results, the simulation conditions need to be considered carefully to ensure that the right information is obtained about the properties of the studied molecules. A similar appraisal should be made of the techniques used to enhance sampling or derive free energies (see below), carefully considering the impact of each specific method. Nevertheless, the primary physical chemistry limitation of the MD technique is the FF. The MM description is the source of differences between modeled and real molecules, even under hypothetically identical simulation and experimental conditions.[54,55]

Standard Simulations versus Enhanced Sampling. Whereas final structures of folded nucleic acids can be well characterized by atomistic experiments, folding of nucleic acids is less well understood. There are many qualitative models of folding, but they are all generally difficult to interpret at the level of atomistic structure description. This paves the way for atomistic simulations,[16] although microsecond-scale unrestrained simulations do not allow the full folding pathway to be simulated. There are many methods to increase sampling, for example, replica exchange molecular dynamics (REMD)[7,72] and metadynamics.[7,73] These methods accelerate conformational changes but cannot fully replace standard unbiased simulations or provide a true temporal picture of the folding.

Temperature REMD (T-REMD) consists of multiple independent simulations (replicas) running at different temperatures, which periodically attempt exchange in the temperature space (Figure 7). Replicas at low temperature travel to higher

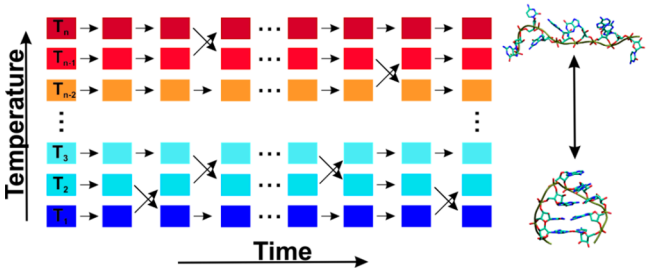


Figure 7. Principle of T-REMD. Several replicas simulated at different temperatures attempt to exchange their temperatures so that the system might overcome the enthalpic barrier at high temperatures while ensuring canonical sampling at low temperatures.

temperatures, where they can more easily cross barriers and then travel back, effectively increasing the sampling space of each replica.[14−16,18,19,74] The correct canonical sampling at each temperature is ensured by using the Metropolis algorithm for replica exchange attempts assuming Boltzmann weights based on the difference in potential energies of the exchanging replicas. The high frequency of exchange attempts results in equilibrium between the simulations at all temperatures and increases effectiveness of the sampling, particularly for the low-temperature replica.[75] T-REMD accelerates crossing over enthalpic barriers but cannot increase sampling of entropy-driven processes and therefore changes the entropy/enthalpy balance.[76] Some conformational transitions may not be facilitated by the increased temperature within the simulation time scale due to increased entropy of the unfolded state.[76] Whereas fully converged T-REMD should in principle guarantee correct (unbiased) thermodynamic populations at the reference replica, any suggestions about the kinetics of the folding processes should be treated with caution because of discontinuities introduced by fast switching among the individual replicas.[76] In fact, in contrast to sufficiently long continuous conventional simulations, the discontinuous nature of T-REMD does not allow realistic modeling of the kinetic pathways of folding and unfolding.[76] The efficacy of T-REMD declines with the size of the system because as the latter is increased, a larger number of replicas is required to cover the

dx.doi.org/10.1021/jz500557y | J. Phys. Chem. Lett. 2014, 5, 1771−1782

same T-REMD temperature range, increasing the computer demands. Increasing the number of replicas or reducing the temperature differences, ΔT, limits the motions of replicas across the temperature space. Solvent fluctuations not related to solute conformation may also significantly contribute to the acceptance ratio of the exchange attempts, which may affect the efficacy of explicit solvent T-REMD simulations for large systems.[77] Ideal T-REMD simulations would guarantee no transfer of errors from higher-temperature replicas to the reference low-temperature replica. However, thermostats producing noncanonical ensembles (with spurious fluctuations in the solvent energy) disrupt the balance between replicas at different temperatures, leading to the acceptance of incorrect structures at low temperature.[78] Although some recent T-REMD studies for small systems (RNA hairpins) have shown promising results,[18,19] other studies have demonstrated convergence difficulties even for tetranucleotides.[14,15]

In alternative Hamiltonian REMD (H-REMD), part of the Hamiltonian is scaled or altered, for example, by weighted bias potential flattening of the MM dihedral terms or by modulating the potential energy surface by restraints.[79−82] Replicas with an unbiased Hamiltonian collect an unbiased ensemble of structures. T-REMD and H-REMD can be combined to give two-dimensional REMD, where the replicas move through both temperature and biased potential energy space.[7,15,83] Nevertheless, further work is needed to establish the applicability of REMD for large RNAs. The folding principles of small and large nucleic acids may differ considerably. T-REMD can be sped up using an implicit solvent; however, implicit solvent is a severe approximation for nucleic acids.[84,85]

Metadynamics involves applying a history-dependent biasing potential to a chosen collective variable in order to expel the simulated system away from potential energy minima. The choice of collective variable (radius of gyration, root-mean-square deviation from a given structure, number of hydrogen bonds, etc.) represents the Achilles heel of metadynamics and critically affects the quality of sampling.[86]

MD simulation is not equivalent to any experimental technique.

To fully prove folding, REMD must start from an unfolded state.[18] Simulations initiated from the folded state (irrespective of the method used) may not accurately capture the folding pathway as folding and unfolding pathways generally differ. Metadynamics simulations always require prior structural insight, which limits their usefulness for predicting the folding pathway. As noted elsewhere,[87] the picture of unfolding of the 15-TBA quadruplex depicted by T-REMD (~20 μs cumulative time)[88] and a metadynamics run (~100 ns)[89] appears to differ. A challenging problem for all methods is to ensure adequate equilibrium sampling of the unfolded state.[15,16] For example, a folding event of an unfolded DNA single strand into an intramolecular quadruplex with three G-quartets can in principle start from 4096 distinct substates of anti−syn orientations of its 12 guanines. Different initial anti−syn substates may result in different atomistic folding routes with diverse off- and on-pathway intermediates. It is difficult to imagine that any presently available method would provide converged sampling of the G-DNA unfolded state.[87]

Insights into folding can also be obtained by standard simulations.[16,87,90,91] Spontaneous folding of the d(GCGAAGC) hairpin has been achieved in microsecond-scale simulations.[16] Although not reaching thermodynamic equilibrium, folding events occurred in several trajectories. Other trajectories, however, remained trapped in stable off-pathway intermediates. This suggests, at the resolution of individual molecules, that folding occurs via multiple pathways. Standard simulations and free-energy computations have identified possible intermediates in the formation of tetramolecular quadruplexes by investigating a series of two-, three-, and four-stranded intermediates.[90] Structures involved in the late stages of folding of intramolecular quadruplexes were studied by first initiating partial unfolding of folded structures by no-salt simulations and then folding them back by adding salt, resembling a hypothetical single-molecule microsecond-scale "stop-flow" experiment.[87] Folding of quadruplexes has been suggested to be a highly multipathway process with a myriad of potential off- and on-pathway structures.

Cornell et al. AMBER Nucleic Acid FFs. The most widely used nucleic acid FFs are based on the Cornell et al. AMBER 1995 parametrization, abbreviated as ff94.[92] It was later modified to give versions ff98[93] and ff99[94] (reviewed in ref 1). Success of this FF stems from parametrization of the electrostatic term by fitting the charges to reproduce the electrostatic potential around the nucleic acid building blocks. This fixed charge model well describes the electrostatic component of base stacking and base pairing because the bases are generally flat and rigid, except for the occasional amino group pyramidalization not included in the parametrization.[95] It also well describes interactions involving the 2′-OH group of ribose.[96] However, the fixed charge model gives a worse description of the backbone.[53,54] The nucleic acid backbone is generally flexible, anionic, and populates diverse conformational classes (different combinations of the backbone dihedrals). Their balanced description is important in simulations.[53] Fixed charge models cannot fully capture energy differences between backbone families, which we consider the main limitation of pair-additive nucleic acid MM (Figure 3).[53,54] The final performance of biomolecular FFs is tuned by the backbone dihedral terms. However, these basically unphysical intramolecular torsional functions usually do not fully compensate for all of the missing electronic structure contributions. An improved backbone description may in the future be achieved either by well calibrated polarization FFs[97] or by QM computations.[55,84] The physical chemistry differences between the classical pair-additive MM model and real molecules (captured by high-quality QM methods) are considerable.[55]

Tuning of the α, γ, and χ Dihedrals of the Cornell et al. FF to Stabilize Nucleic Acid Simulations. Long simulations have revealed that the original (ff94−ff99) version does not provide sufficiently stable trajectories.[17,59,98] This has led to two key refinements of its torsional potential to tackle the major structural problems.[17,52,98]

The 2007 Parmbsc0 version reparameterized the α/γ dihedrals to prevent collapse of the B-DNA due to accumulation of irreversible non-native γ-trans backbone dihedral states (cf. Figure 4 in ref 98). Omission of the bsc0 correction in DNA simulations inevitably leads to robust artifacts on a scale of dozens of nanoseconds. Such perturbations have occasionally been misinterpreted as novel structures.

The 2010 χ_{OL3} (or χ_{OL}) version[52] reparameterized the χ dihedral to suppress anti to high-anti χ shifts in RNA, which (among other problems) lead to sudden irreversible transitions of RNAs into entirely untwisted ladder-like structures (see Figure 8 in ref 59). The RNA ladders usually form more slowly

dx.doi.org/10.1021/jz500557y | J. Phys. Chem. Lett. 2014, 5, 1771−1782

than the γ-trans structure of B-DNA.[17] Therefore, the majority (but not all) of older RNA studies are unaffected. The χ imbalance becomes detrimental in long and folding simulations. Collapse of a folded riboswitch within 1 μs has been reported (see Figure 2 in ref 38), while for small hairpins, the artifact can build up even much more quickly.[17] The RNA χ_{OL3} modification should preferably be used with bsc0.[17] The χ_{OL3} version also improves the description of the RNA syn region and the syn−anti balance.[17] A similar performance has been achieved by the χ reparameterization of Yildirim et al. (χ_{YL}), originally intended to correct the syn−anti balance based on NMR data.[99] We have later shown that it also eliminates the ladders, although the high-anti region is subtly overpunished, reducing the inclination of A-RNA.[11]

The challenges faced in MM description of nucleic acid backbones are highlighted by CCSD(T) benchmark QM studies of the relative energies of all known DNA and RNA backbone families (almost 70 different conformers).[53,54] Whereas modern dispersion-corrected DFT methods can excellently reproduce the benchmark QM data, the errors obtained by MM description, despite trying different charge distributions, are almost an order of magnitude larger (Figure 3). Modern DFT-D3-level QM computations on seven different folds of two-quartet G-DNA stems have demonstrated that the relative energies of different G-DNA folds predicted by MM and QM differ substantially.[55] The main source of the differences is the description of the backbone and base−backbone coupling. Although careful parametrization of the backbone torsions can possibly further improve the FF, it is likely that the use of torsions to tune the FF is not unlimited.[1] Unless targeting major simulation artifacts (the bsc0 and χ_{OL3} versions), adjustments of the FF require tedious testing. It might be difficult to decide whether a modification is useful due to either a lack of appropriate experimental data or ambiguity of the results (improving some features and worsening others). The contemporary FFs well describe canonical helices and correctly fold (some of) the smallest RNA hairpins. However, they may not be as successful in folding more complex nucleic acids.

The MM force field approximation is the main source of physical chemistry difference between simulated and real molecules.

The above-described FF adjustments illustrate the complexity of tuning the FF. Although bsc0 has improved B-DNA simulations, relatively long-living γ-trans states still occur. They may be filtered out during postprocessing of the simulation data.[100] It would be easy to completely eliminate the γ-trans states in B-DNA. However, the bsc0 correction already eliminates native γ-trans states of G-DNA loops[20] and overcorrects γ-trans in A-RNA.[11] Also, the fact that a 0.2−0.3 kcal/mol imbalance in a single χ torsional term (which is the accuracy of the best QM methods) may push the RNA structure into a completely incorrect global minimum highlights the delicacy of tuning.[59] Making MM ideal for one specific target feature may worsen its properties elsewhere.[101] In fact, it is fortuitous that the bsc0 and χ_{OL3} corrections were feasible because it was not a priori guaranteed that viable corrections (not leading to substantial undesired side effects) existed. χ_{OL3} cannot be used for DNA because it was not possible to simultaneously tune the χ torsion for DNA and RNA, leading to a split in DNA and RNA FFs. Despite using rigorous QM data, tuning of FFs ultimately remains an empirical process that attempts to find a compromise in describing multiple structures and features. Any changes should preferably be small. The more radical the change, the more likely it will produce undesired side effects.[101] The balance of MM highly relies on compensation of errors, and even a physically meaningful modification might worsen the behavior of complex folded systems due to loss of error compensation. MM modifications resemble the development, testing, and use of drugs. They require extensive testing at the limits of available computers, and even very useful modifications may occasionally worsen simulation performance for specific systems. In some cases, it may take years before a FF variant is fully assessed.

Other Refinements of the Cornell et al. FF. There are more subtle modifications of the Cornell et al. dihedral terms. χ_{OL4} is the DNA counterpart of RNA χ_{OL3} that improves the behavior of antiparallel G-DNA.[21] AMBERTOR attempted to reparameterize the remaining $\beta\varepsilon\zeta$ dihedrals for RNA.[102] However, it seems to severely distort A-RNA (see Figure 12 in ref 53). $\varepsilon\zeta_{OL1}$ (obviously combined with the key bsc0 refinement) tunes the BI/BII populations in B-DNA,[50] reduces the RMSD of simulated B-DNA duplexes, and visibly improves the sequence dependence of the B-DNA helical twist. It increases twist in those steps where bsc0 simulations predict too low of a twist and vice versa. However, its full-scale testing for folded RNAs is not yet completed. Another recent modification changes the balance between stacking, H-bonding, and solvation.[19] It is based on (i) rescaling van der Waals (vdW) parameters of the nucleobases and (ii) using different, empirically adjusted vdW combination rules for base−water interactions. Although its main aim was to weaken stacking and eliminate the stacking overstabilization, it also strengthens base pairing. Although the unequal combination rules may look at first glance unphysical (they are not supported by the AMBER12 code but are possible in some other codes), they are a valid option for empirically tuning the MM. The vdW modification is accompanied by adjustment of the χ torsion closely resembling the χ_{OL3} and χ_{YL} refinements in the anti to high-anti region to prevent the ladder-like structures. It also tunes the syn−anti balance (compared to ff94−ff99). The modification can achieve very fast folding of 8-mer RNA tetraloops via T-REMD from an unfolded state.[19] Bsc0χ_{OL3} also allows folding of 10-mer RNA tetraloops in T-REMD simulations, albeit on a much longer time scale.[18] Nevertheless, the reported improvement of tetraloop structures by the vdW modification is based on comparison with simulations of the same group performed with the ff94 FF. Such comparison is somewhat incomplete as the ff94 version is almost 20 years old and has been repeatedly shown by others not to be suitable for description of RNAs. In addition, the structural improvement may be primarily due to the conventional correction of the χ torsion. The rescaled vdW version does not appear to bring structural improvements with respect to the current default AMBER ff99bsc0χ_{OL3} version (abbreviated as ff10, ff12, or ff14 RNA FF in recent AMBER code versions). Broader testing of the rescaled vdW version needs to be performed.

CHARMM FF. CHARMM is another MM potential that was systematically developed for biomacromolecular simulations, including nucleic acids. Its older version CHARMM27[103] has recently been reparameterized to CHARMM36 by correcting the 2′-OH ribose torsional potential in an attempt to suppress rather rapidly developing instabilities in base-paired regions.[51] However, as yet, long time scale testing has not been performed

dx.doi.org/10.1021/jz500557y | J. Phys. Chem. Lett. 2014, 5, 1771−1782

for folded RNAs. The specific CHARMM36 reparameterization of the ε and ζ torsions and sugar pucker enhances its ability to capture local variations of B-DNA.[49] The first version of polarizable DNA CHARMM MM has been released.[97] Although polarizable MM will certainly need some time for tuning, it could represent a turning point in the development of biomolecular FFs.

Water and Ions. A common misunderstanding in the assessment of MD studies revolves around the inclusion of ions and salt because of the sensitivity of nucleic acid experiments to salt conditions. Some readers expect that MD studies should use the same salt conditions as experiments. However, this requirement is not valid because experimental and simulation conditions are not equivalent (see the above section entitled Simulation versus Experimental Conditions). In experiments, the type of ion may affect the overall free-energy balance between different G-DNA folds at thermodynamic equilibrium. However, in the microsecond time window investigated by MD, there is no systematic difference between NaCl and KCl G-DNA simulations as both ions stabilize G-DNA.[87] The simulated molecules remain locked around the starting structures and do not sense other alternative folds unreachable on the simulation time scale, irrespective of their thermodynamics stability. Similarly, when simulating native structures of folded RNAs, the exact salt conditions are unlikely to have a dramatic effect because the simulations do not reach the folding process. We have conducted several studies comparing different salt conditions but have never detected any systematic differences attributable to the specific salt conditions.[1,11,70,87] This does not rule out the possibility that salt conditions could matter in some simulations, but a visible effect is improbable on simulation time scales. Most nucleic acid simulations behave well when using a net-neutral monovalent cation condition, which with present simulation protocols (box sizes and time scales) provides a sufficient concentration of cations in the vicinity of the molecule.[1] Note that the unfortunate combination of ion parameters and vdW combination rules may lead to salt-clustering. This problem has, however, been eliminated.[104] Divalent ions are difficult to describe by MM because real ions induce large polarization and charge-transfer effects that are not includable in their MM description using vdW spheres with a point charge of 2+ in their centers. To understand the FF limitations for divalent ions, the reader should consult the articles on QM studies listed elsewhere.[1,4,5] The water model may affect nucleic acid simulations more than the ionic conditions (but much less than the solute FF), and there is undoubtedly coupling between the solute and solvent FFs.[11,21,70]

Riboswitches As Another Example. The complex relationship between the starting structure, FF, simulation time scale, and actual biochemical processes can be illustrated by considering studies of riboswitches, that is, modular regulatory segments of mRNAs. Upon binding a ligand to their highly conserved aptamer domains with large specificity, riboswitches regulate protein expression via large-scale conformational changes in their more variable expression domains. The full process occurs on a time scale that is far beyond the reach of unbiased simulations, and in some cases, the control is exerted via kinetics rather than thermodynamics.[105]

MD can be used to investigate conformational changes in an aptamer domain initiated by removal of the native ligand from the active site of the experimental structure of the aptamer holo form. Such structural changes should be describable by MM with sufficient accuracy. However, the full conformational change does not usually occur on an achievable time scale as microsecond simulation can only overcome a free-energy barrier of ~10 kcal/mol (Figure 8). The conformational space

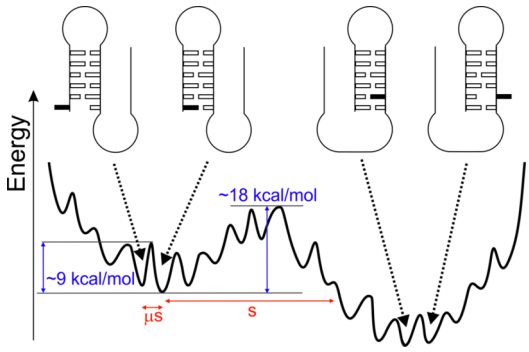


Figure 8. Rugged potential energy surface of nucleic acids. Whereas local conformational states covering base pair opening and interhelical flexibility are separated by barriers of ~9 kcal/mol that might be overcome by MD on a microscale time scale, rebuilding of secondary structures cannot be addressed by unrestrained simulations.

of RNA is very rugged.[106−109] This is illustrated by standard simulations of SAM-II[39,40] and preQ$_1$[38] riboswitches, which reported subtle differences in the conformational behavior of the holo form with ligand and after removal of the ligand on a time scale of hundreds of nanoseconds up to a few microseconds. Such studies usually capture local effects of noncognate ligand binding on the structural dynamics of the aptamer (in comparison with cognate ligand binding)[40,45] and characterize structural dynamics of the apo form of the aptamer in cases where the corresponding experimental structure is available.[38] Still, attention should be paid to a potential FF bias, as demonstrated in an analysis of the preQ$_1$ riboswitch, where only the ff99bsc0χ_{OL3} resulted in stable long trajectories.[38] Some of the results reported in the riboswitch simulation literature reflect MM inaccuracies.

Because unbiased MD may only sample local conformations and cannot overcome barriers, other specialized techniques, like nonequilibrium sampling represented, for example, by pulling or flooding techniques, have been developed.[41] Such studies are still based on atomistic MM. However, as noted above, the accuracy of the specialized techniques critically depends on the choice of a collective variable along which the conformational dynamics is studied.[86] Pulling MD has been used to evaluate the thermodynamic stability of the stem bearing a Shine−Dalgarno sequence in both the apo and holo form of an adenine riboswitch. The computations distinguished between a ligand-binding-dependent conformational switch and ligand-induced aptamer preorganization mechanisms of its action.[41] Obviously, large conformational changes of riboswitches can be studied by considerably more approximate methods, such as knowledge-based potentials (Go-model) and so forth.[110]

We have provided a brief overview of contemporary atomistic simulations of nucleic acids, focusing on simulations

Enhanced sampling methods facilitate conformational transitions but may bias the apparent folding process.

dx.doi.org/10.1021/jz500557y | J. Phys. Chem. Lett. 2014, 5, 1771−1782

of small- and medium-sized RNAs. This Perspective was not intended to give a full literature review but rather articulates our own experiences and expectations of future developments in the research field. Due to space restrictions, we could not discuss all of the relevant literature. Instead, we concentrated on explaining the limitations of the simulations, namely, the MM approximation and limited time scale, as they are often overlooked in contemporary literature. We also tried to explain the relation (and differences) between simulations and experiments. We show that MD can provide unique insights into biomolecular structure and dynamics that cannot be obtained by existing experimental techniques. Thus, despite extensively discussing the limitations, we remain convinced that MD is a powerful technique for studying the structural dynamics of nucleic acids with a fast growing potential, capable of substantially complementing experimental methods and aiding interpretation of their results.

AUTHOR INFORMATION

Corresponding Author

*E-mail: sponer@ncbr.muni.cz.

Notes

The authors declare no competing financial interest.

Biographies

Jiří Šponer is a head of the Department of Structure and Dynamics of Nucleic Acids at the Institute of Biophysics, Academy of Sciences of the Czech Republic, Brno (http://ibp.cz/en/departments/structure-and-dynamics-of-nucleic-acids) and Professor at the Central European Institute of Technology, Masaryk University, Brno and Palacký University, Olomouc. His research is focused on computational studies of nucleic acids.

Pavel Banáš is an Associate Professor of Physical Chemistry at the Department of Physical Chemistry (http://fch.upol.cz/en/) and junior researcher at the Regional Centre of Advanced Technologies and Materials (RCPTM), Palacký University, Olomouc. His research is focused on computational studies of structural dynamics of noncoding RNAs and of RNA catalysis.

Petr Jurečka is an Associate Professor of Physical Chemistry at the Department of Physical Chemistry and senior researcher of the Regional Centre of Advanced Technologies and Materials (RCPTM), Palacký University, Olomouc. His research is focused on structural dynamics of DNA and RNA, parameter development, and intermolecular interactions.

Marie Zgarbová is an Assistant Professor at the Department of Physical Chemistry and junior researcher at the Regional Centre of Advanced Technologies and Materials, Palacký University, Olomouc. Her research is focused on computational studies of the structure and dynamics of nucleic acids.

Petra Kührová is an Assistant Professor at the Department of Physical Chemistry and junior researcher at the Regional Centre of Advanced Technologies and Materials (RCPTM), Palacký University, Olomouc. Her research is focused on structural dynamics of biomolecules.

Marek Havrila is a junior researcher and Ph.D. student at the Institute of Biophysics, Academy of Sciences of the Czech Republic, Brno. His research is focused on molecular dynamics simulations and bioinformatics studies of RNA.

Miroslav Krepl is a junior researcher and Ph.D. student at the Institute of Biophysics, Academy of Sciences of the Czech Republic, Brno. His research is focused on molecular dynamics simulations of small RNA molecules and protein/RNA complexes.

Petr Stadlbauer is a junior researcher and Ph.D. student at the Institute of Biophysics, Academy of Sciences of the Czech Republic, Brno. His research is primarily focused on MD simulations of structural dynamics and folding of G-quadruplexes.

Michal Otyepka is a head of the Department of Physical Chemistry at the Palacký University, Olomouc and senior researcher at the Regional Centre of Advanced Technologies and Materials (RCPTM), Palacký University, Olomouc. His research is focused on modeling of biomacromolecules, complex molecular systems, and hybrid materials.

ACKNOWLEDGMENTS

This work was supported by the Czech Science Foundation (Grant Number P208/12/1878). Institutional support was obtained through the project CEITEC - Central European Institute of Technology (CZ.1.05/1.1.00/02.0068) from the European Regional Development Fund, the Operational Program Research and Development for Innovations − European Regional Development Fund (CZ.1.05/2.1.00/03.0058), and the Operational Program Education for Competitiveness − European Social Fund (CZ.1.07/2.3.00/20.0017) of the Ministry of Education, Youth and Sports of the Czech Republic.

REFERENCES

(1) Sponer, J.; Cang, X. H.; Cheatham, T. E. Molecular Dynamics Simulations of G-DNA and Perspectives on the Simulation of Nucleic Acid Structures. *Methods* **2012**, *57*, 25−39.

(2) Perez, A.; Luque, F. J.; Orozco, M. Frontiers in Molecular Dynamics Simulations of DNA. *Acc. Chem. Res.* **2011**, *45*, 196−205.

(3) Sanbonmatsu, K. Y. Computational Studies of Molecular Machines: The Ribosome. *Curr. Opin. Struct. Biol.* **2012**, *22*, 168−174.

(4) Ditzler, M. A.; Otyepka, M.; Sponer, J.; Walter, N. G. Molecular Dynamics and Quantum Mechanics of RNA: Conformational and Chemical Change We Can Believe In. *Acc. Chem. Res.* **2009**, *43*, 40−47.

(5) McDowell, S. E.; Spackova, N.; Sponer, J.; Walter, N. G. Molecular Dynamics Simulations of RNA: An In Silico Single Molecule Approach. *Biopolymers* **2007**, *85*, 169−184.

(6) Auffinger, P.; Hashem, Y. Nucleic Acid Solvation: From Outside to Insight. *Curr. Opin. Struct. Biol.* **2007**, *17*, 325−333.

(7) Kara, M.; Zacharias, M. Theoretical Studies of Nucleic Acids Folding. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2014**, *4*, 116−126.

(8) Dror, R. O.; Dirks, R. M.; Grossman, J. P.; Xu, H.; Shaw, D. E. Biomolecular Simulation: A Computational Microscope for Molecular Biology. *Annu. Rev. Biophys.* **2012**, *41*, 429−452.

(9) Salomon-Ferrer, R.; Goetz, A. W.; Poole, D.; Le Grand, S.; Walker, R. C. Routine Microsecond Molecular Dynamics Simulations with AMBER on GPUs. 2. Explicit Solvent Particle Mesh Ewald. *J. Chem. Theory Comput.* **2013**, *9*, 3878−3888.

(10) Lavery, R.; Zakrzewska, K.; Beveridge, D.; Bishop, T. C.; Case, D. A.; Cheatham, T.; Dixit, S.; Jayaram, B.; Lankas, F.; Laughton, C.; et al. 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.* **2010**, *38*, 299−313.

(11) Besseova, I.; Banas, P.; Kuehrova, P.; Kosinova, P.; Otyepka, M.; Sponer, J. Simulations of A-RNA Duplexes. The Effect of Sequence, Solute Force Field, Water Model, and Salt Concentration. *J. Phys. Chem. B* **2012**, *116*, 9899−9916.

(12) Yildirim, I.; Stern, H. A.; Tubbs, J. D.; Kennedy, S. D.; Turner, D. H. Benchmarking AMBER Force Fields for RNA: Comparisons to NMR Spectra for Single-Stranded r(GACC) Are Improved by Revised χ Torsions. *J. Phys. Chem. B* **2011**, *115*, 9261−9270.

(13) Tubbs, J. D.; Condon, D. E.; Kennedy, S. D.; Hauser, M.; Bevilacqua, P. C.; Turner, D. H. The Nuclear Magnetic Resonance of CCCC RNA Reveals a Right-Handed Helix, and Revised Parameters

for AMBER Force Field Torsions Improve Structural Predictions from Molecular Dynamics. *Biochemistry* **2013**, *52*, 996−1010.
(14) Henriksen, N. M.; Roe, D. R.; Cheatham, T. E. Reliable Oligonucleotide Conformational Ensemble Generation in Explicit Solvent for Force Field Assessment Using Reservoir Replica Exchange Molecular Dynamics Simulations. *J. Phys. Chem. B* **2013**, *117*, 4014−4027.
(15) Bergonzo, C.; Henriksen, N. M.; Roe, D. R.; Swails, J. M.; Roitberg, A. E.; Cheatham, T. E. Multidimensional Replica Exchange Molecular Dynamics Yields a Converged Ensemble of an RNA Tetranucleotide. *J. Chem. Theory Comput.* **2013**, *10*, 492−499.
(16) Portella, G.; Orozco, M. Multiple Routes to Characterize the Folding of a Small DNA Hairpin. *Angew. Chem., Int. Ed.* **2010**, *49*, 7673−7676.
(17) Banas, P.; Hollas, D.; Zgarbova, M.; Jurecka, P.; Orozco, M.; Cheatham, T. E.; Sponer, J.; Otyepka, M. Performance of Molecular Mechanics Force Fields for RNA Simulations: Stability of UUCG and GNRA Hairpins. *J. Chem. Theory Comput.* **2010**, *6*, 3836−3849.
(18) Kuehrova, P.; Banas, P.; Best, R. B.; Sponer, J.; Otyepka, M. Computer Folding of RNA Tetraloops? Are We There Yet? *J. Chem. Theory Comput.* **2013**, *9*, 2115−2125.
(19) Chen, A. A.; Garcia, A. E. High-Resolution Reversible Folding of Hyperstable RNA Tetraloops Using Molecular Dynamics Simulations. *Proc. Natl. Acad. Sci. U.S.A.* **2013**, *110*, 16820−16825.
(20) Fadrna, E.; Spackova, N.; Sarzynska, J.; Koca, J.; Orozco, M.; Cheatham, T. E.; Kulinski, T.; Sponer, J. Single Stranded Loops of Quadruplex DNA as Key Benchmark for Testing Nucleic Acids Force Fields. *J. Chem. Theory Comput.* **2009**, *5*, 2514−2530.
(21) Krepl, M.; Zgarbova, M.; Stadlbauer, P.; Otyepka, M.; Banas, P.; Koca, J.; Cheatham, T. E.; Jurecka, P.; Sponer, J. Reference Simulations of Noncanonical Nucleic Acids with Different Chi Variants of the AMBER Force Field: Quadruplex DNA, Quadruplex RNA, and Z-DNA. *J. Chem. Theory Comput.* **2012**, *8*, 2506−2520.
(22) Reblova, K.; Spackova, N.; Sponer, J. E.; Koca, J.; Sponer, J. Molecular Dynamics Simulations of RNA Kissing-Loop Motifs Reveal Structural Dynamics and Formation of Cation−Binding Pockets. *Nucleic Acids Res.* **2003**, *31*, 6942−6952.
(23) Reblova, K.; Spackova, N.; Stefl, R.; Csaszar, K.; Koca, J.; Leontis, N. B.; Sponer, J. Non-Watson−Crick Basepairing and Hydration in RNA Motifs: Molecular Dynamics of 5S rRNA Loop E. *Biophys. J.* **2003**, *84*, 3564−3582.
(24) Auffinger, P.; Bielecki, L.; Westhof, E. Symmetric K^+ and Mg^{2+} Ion-Binding Sites in the 5S rRNA Loop E Inferred from Molecular Dynamics Simulations. *J. Mol. Biol.* **2004**, *335*, 555−571.
(25) Razga, F.; Koca, J.; Sponer, J.; Leontis, N. B. Hinge-Like Motions in RNA Kink-Turns: The Role of the Second A-Minor Motif and Nominally Unpaired Bases. *Biophys. J.* **2005**, *88*, 3466−3485.
(26) Singh, A.; Sethaphong, L.; Yingling, Y. G. Interactions of Cations with RNA Loop−Loop Complexes. *Biophys. J.* **2011**, *101*, 727−735.
(27) Romanowska, J.; Reuter, N.; Trylska, J. Comparing Aminoglycoside Binding Sites in Bacterial Ribosomal RNA and Aminoglycoside Modifying Enzymes. *Proteins: Struct., Funct., Bioinf.* **2013**, *81*, 63−80.
(28) Havrila, M.; Reblova, K.; Zirbel, C. L.; Leontis, N. B.; Sponer, J. Isosteric and Nonisosteric Base Pairs in RNA Motifs: Molecular Dynamics and Bioinformatics Study of the Sarcin−Ricin Internal Loop. *J. Phys. Chem. B* **2013**, *117*, 14302−14319.
(29) Aytenfisu, A. H.; Spasic, A.; Seetin, M. G.; Serafini, J.; Mathews, D. H. A Modified Amber Force Field Correctly Models the Conformational Preference for Tandem GA Pairs in RNA. *J. Chem. Theory Comput.* **2014**, *10*, 1292−1301.
(30) Krasovska, M. V.; Sefcikova, J.; Reblova, K.; Schneider, B.; Walter, N. G.; Sponer, J. Cations and Hydration in Catalytic RNA: Molecular Dynamics of the Hepatitis Delta Virus Ribozyme. *Biophys. J.* **2006**, *91*, 626−638.
(31) Rhodes, M. M.; Reblova, K.; Sponer, J.; Walter, N. G. Trapped Water Molecules Are Essential to Structural Dynamics and Function of a Ribozyme. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 13380−13385.
(32) Lee, T.-S.; Lopez, C. S.; Giambasu, G. M.; Martick, M.; Scott, W. G.; York, D. M. Role of Mg^{2+} in Hammerhead Ribozyme Catalysis from Molecular Simulation. *J. Am. Chem. Soc.* **2008**, *130*, 3053−3064.
(33) Banas, P.; Walter, N. G.; Sponer, J.; Otyepka, M. Protonation States of the Key Active Site Residues and Structural Dynamics of the glmS Riboswitch As Revealed by Molecular Dynamics. *J. Phys. Chem. B* **2010**, *114*, 8701−8712.
(34) Xin, Y.; Hamelberg, D. Deciphering the Role of Glucosamine-6-Phosphate in the Riboswitch Action of glmS Ribozyme. *RNA* **2010**, *16*, 2455−2463.
(35) Lee, T.-S.; Giambasu, G. M.; Harris, M. E.; York, D. M. Characterization of the Structure and Dynamics of the HDV Ribozyme in Different Stages Along the Reaction Path. *J. Phys. Chem. Lett.* **2011**, *2*, 2538−2543.
(36) Veeraraghavan, N.; Ganguly, A.; Chen, J.-H.; Bevilacqua, P. C.; Hammes-Schiffer, S.; Golden, B. L. Metal Binding Motif in the Active Site of the HDV Ribozyme Binds Divalent and Monovalent Ions. *Biochemistry* **2011**, *50*, 2672−2682.
(37) Villa, A.; Woehnert, J.; Stock, G. Molecular Dynamics Simulation Study of the Binding of Purine Bases to the Aptamer Domain of the Guanine Sensing Riboswitch. *Nucleic Acids Res.* **2009**, *37*, 4774−4786.
(38) Banas, P.; Sklenovsky, P.; Wedekind, J. E.; Sponer, J.; Otyepka, M. Molecular Mechanism of preQ1 Riboswitch Action: A Molecular Dynamics Study. *J. Phys. Chem. B* **2012**, *116*, 12721−12734.
(39) Kelley, J. M.; Hamelberg, D. Atomistic Basis for the On−Off Signaling Mechanism in SAM-II Riboswitch. *Nucleic Acids Res.* **2010**, *38*, 1392−1400.
(40) Doshi, U.; Kelley, J. M.; Hamelberg, D. Atomic-Level Insights into Metabolite Recognition and Specificity of the SAM-II Riboswitch. *RNA* **2012**, *18*, 300−307.
(41) Di Palma, F.; Colizzi, F.; Bussi, G. Ligand-Induced Stabilization of the Aptamer Terminal Helix in the add Adenine Riboswitch. *RNA* **2013**, *19*, 1517−1524.
(42) Jung, S.; Schlick, T. Interconversion between Parallel and Antiparallel Conformations of a 4H RNA junction in Domain 3 of Foot-and-Mouth Disease Virus IRES Captured by Dynamics Simulations. *Biophys. J.* **2014**, *106*, 447−458.
(43) Reblova, K.; Sponer, J.; Lankas, F. Structure and Mechanical Properties of the Ribosomal L1 Stalk Three-Way Junction. *Nucleic Acids Res.* **2012**, *40*, 6290−6303.
(44) Sanbonmatsu, K. Y.; Joseph, S.; Tung, C.-S. Simulating Movement of tRNA into the Ribosome During Decoding. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 15854−15859.
(45) Trabuco, L. G.; Villa, E.; Mitra, K.; Frank, J.; Schulten, K. Flexible Fitting of Atomic Structures into Electron Microscopy Maps Using Molecular Dynamics. *Structure* **2008**, *16*, 673−683.
(46) Trabuco, L. G.; Schreiner, E.; Eargle, J.; Cornish, P.; Ha, T.; Luthey-Schulten, Z.; Schulten, K. The Role of L1 Stalk−tRNA Interaction in the Ribosome Elongation Cycle. *J. Mol. Biol.* **2010**, *402*, 741−760.
(47) Roccatano, D.; Barthel, A.; Zacharias, M. Structural Flexibility of the Nucleosome Core Particle at Atomic Resolution Studied by Molecular Dynamics Simulation. *Biopolymers* **2007**, *85*, 407−421.
(48) Portella, G.; Battistini, F.; Orozco, M. Understanding the Connection between Epigenetic DNA Methylation and Nucleosome Positioning from Computer Simulations. *PLoS Comput. Biol.* **2013**, *9*, e1003354.
(49) Hart, K.; Foloppe, N.; Baker, C. M.; Denning, E. J.; Nilsson, L.; MacKerell, A. D. Optimization of the CHARMM Additive Force Field for DNA: Improved Treatment of the BI/BII Conformational Equilibrium. *J. Chem. Theory Comput.* **2011**, *8*, 348−362.
(50) Zgarbova, M.; Luque, F. J.; Sponer, J.; Cheatham, T. E.; Otyepka, M.; Jurecka, P. Toward Improved Description of DNA Backbone: Revisiting Epsilon and Zeta Torsion Force Field Parameters. *J. Chem. Theory Comput.* **2013**, *9*, 2339−2354.
(51) Denning, E. J.; Priyakumar, U. D.; Nilsson, L.; Mackerell, A. D. Impact of 2′-Hydroxyl Sampling on the Conformational Properties of

RNA: Update of the CHARMM All-Atom Additive Force Field for RNA. *J. Comput. Chem.* **2011**, *32*, 1929−1943.
(52) Zgarbova, M.; Otyepka, M.; Sponer, J.; Mladek, A.; Banas, P.; Cheatham, T. E.; Jurecka, P. Refinement of the Cornell et al. Nucleic Acids Force Field Based on Reference Quantum Chemical Calculations of Glycosidic Torsion Profiles. *J. Chem. Theory Comput.* **2011**, *7*, 2886−2902.
(53) Sponer, J.; Mladek, A.; Sponer, J. E.; Svozil, D.; Zgarbova, M.; Banas, P.; Jurecka, P.; Otyepka, M. The DNA and RNA Sugar−Phosphate Backbone Emerges as the Key Player. An Overview of Quantum-Chemical, Structural Biology and Simulation Studies. *Phys. Chem. Chem. Phys.* **2012**, *14*, 15257−15277.
(54) Mladek, A.; Banas, P.; Jurecka, P.; Otyepka, M.; Zgarbova, M.; Sponer, J. Energies and 2′-Hydroxyl Group Orientations of RNA Backbone Conformations. Benchmark CCSD(T)/CBS Database, Electronic Analysis, and Assessment of DFT Methods and MD Simulations. *J. Chem. Theory Comput.* **2014**, *10*, 463−480.
(55) Sponer, J.; Mladek, A.; Spackova, N.; Cang, X. H.; Cheatham, T. E.; Grimme, S. Relative Stability of Different DNA Guanine Quadruplex Stem Topologies Derived Using Large-Scale Quantum-Chemical Computations. *J. Am. Chem. Soc.* **2013**, *135*, 9785−9796.
(56) Krepl, M.; Reblova, K.; Koca, J.; Sponer, J. Bioinformatics and Molecular Dynamics Simulation Study of L1 Stalk Non-Canonical rRNA Elements: Kink-Turns, Loops, and Tetraloops. *J. Phys. Chem. B* **2013**, *117*, 5540−5555.
(57) Gkionis, K.; Kruse, H.; Platts, J. A.; Mladek, A.; Koca, J.; Sponer, J. Ion Binding to Quadruplex DNA Stems. Comparison of MM and QM Descriptions Reveals Sizable Polarization Effects Not Included in Contemporary Simulations. *J. Chem. Theory Comput.* **2014**, *10*, 1326−1340.
(58) Cang, X. H.; Sponer, J.; Cheatham, T. E. Explaining the Varied Glycosidic Conformational, G-Tract Length and Sequence Preferences for Anti-Parallel G-Quadruplexes. *Nucleic Acids Res.* **2011**, *39*, 4499−4512.
(59) Mlynsky, V.; Banas, P.; Hollas, D.; Reblova, K.; Walter, N. G.; Sponer, J.; Otyepka, M. Extensive Molecular Dynamics Simulations Showing That Canonical G8 and Protonated A38H$^+$ Forms Are Most Consistent with Crystal Structures of Hairpin Ribozyme. *J. Phys. Chem. B* **2010**, *114*, 6642−6652.
(60) Trylska, J. Coarse-Grained Models to Study Dynamics of Nanoscale Biomolecules and Their Applications to the Ribosome. *J. Phys.: Condens. Matter* **2010**, *22*, 453101.
(61) Whitford, P. C.; Onuchic, J. N.; Sanbonmatsu, K. Y. Connecting Energy Landscapes with Experimental Rates for Aminoacyl-tRNA Accommodation in the Ribosome. *J. Am. Chem. Soc.* **2010**, *132*, 13170−13171.
(62) Hashem, Y.; Auffinger, P. A Short Guide for Molecular Dynamics Simulations of RNA Systems. *Methods* **2009**, *47*, 187−197.
(63) Bleuzen, A.; Pittet, P. A.; Helm, L.; Merbach, A. E. Water Exchange on Magnesium(II) in Aqueous Solution: A Variable Temperature and Pressure O-17 NMR Study. *Magn. Reson. Chem.* **1997**, *35*, 765−773.
(64) Ke, A.; Zhou, K.; Ding, F.; Cate, J. H. D.; Doudna, J. A. A Conformational Switch Controls Hepatitis Delta Virus Ribozyme Catalysis. *Nature* **2004**, *429*, 201−205.
(65) Banas, P.; Rulisek, L.; Hanosova, V.; Svozil, D.; Walter, N. G.; Sponer, J.; Otyepka, M. General Base Catalysis for Cleavage by the Active-Site Cytosine of the Hepatitis Delta Virus Ribozyme: QM/MM Calculations Establish Chemical Feasibility. *J. Phys. Chem. B* **2008**, *112*, 11177−11187.
(66) Das, S. R.; Piccirilli, J. A. General Acid Catalysis by the Hepatitis Delta Virus Ribozyme. *Nat. Chem. Biol.* **2005**, *1*, 45−52.
(67) Chen, J.-H.; Yajima, R.; Chadalavada, D. M.; Chase, E.; Bevilacqua, P. C.; Golden, B. L. A 1.9 Angstrom Crystal Structure of the HDV Ribozyme Precleavage Suggests both Lewis Acid and General Acid Mechanisms Contribute to Phosphodiester Cleavage. *Biochemistry* **2010**, *49*, 6508−6518.
(68) Gueron, M.; Leroy, J.-L. Studies of Base Pair Kinetics by NMR Measurement of Proton Exchange. In *Methods of Enzymology*; Thomas, L. J., Ed.; Academic Press: San Diego, CA, 1995; Vol. *261*; pp 383−413.
(69) Reblova, K.; Strelcova, Z.; Kulhanek, P.; Besseova, I.; Mathews, D. H.; Van Nostrand, K.; Yildirim, I.; Turner, D. H.; Sponer, J. An RNA Molecular Switch: Intrinsic Flexibility of 23S rRNA Helices 40 and 68 5 ′-UAA/5 ′-GAN Internal Loops Studied by Molecular Dynamics Methods. *J. Chem. Theory Comput.* **2010**, *6*, 910−929.
(70) Sklenovsky, P.; Florova, P.; Banas, P.; Reblova, K.; Lankas, F.; Otyepka, M.; Sponer, J. Understanding RNA Flexibility Using Explicit Solvent Simulations: The Ribosomal and Group I Intron Reverse Kink-Turn Motifs. *J. Chem. Theory Comput.* **2011**, *7*, 2963−2980.
(71) Chen, A. A.; Draper, D. E.; Pappu, R. V. Molecular Simulation Studies of Monovalent Counterion-Mediated Interactions in a Model RNA Kissing Loop. *J. Mol. Biol.* **2009**, *390*, 805−819.
(72) Sugita, Y.; Okamoto, Y. Replica-Exchange Molecular Dynamics Method for Protein Folding. *Chem. Phys. Lett.* **1999**, *314*, 141−151.
(73) Laio, A.; Parrinello, M. Escaping Free-Energy Minima. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 12562−12566.
(74) Kannan, S.; Zacharias, M. Folding of a DNA Hairpin Loop Structure in Explicit Solvent Using Replica-Exchange Molecular Dynamics Simulations. *Biophys. J.* **2007**, *93*, 3218−3228.
(75) Sindhikara, D. J.; Emerson, D. J.; Roitberg, A. E. Exchange Often and Properly in Replica Exchange Molecular Dynamics. *J. Chem. Theory Comput.* **2010**, *6*, 2804−2808.
(76) Beck, D. A. C.; White, G. W. N.; Daggett, V. Exploring the Energy Landscape of Protein Folding Using Replica-Exchange and Conventional Molecular Dynamics Simulations. *J. Struct. Biol.* **2007**, *157*, 514−523.
(77) Periole, X.; Mark, A. E. Convergence and Sampling Efficiency in Replica Exchange Simulations of Peptide Folding in Explicit Solvent. *J. Chem. Phys.* **2007**, *126*, 014903.
(78) Rosta, E.; Buchete, N.-V.; Hummer, G. Thermostat Artifacts in Replica Exchange Molecular Dynamics Simulations. *J. Chem. Theory Comput.* **2009**, *5*, 1393−1399.
(79) Fukunishi, H.; Watanabe, O.; Takada, S. On the Hamiltonian Replica Exchange Method for Efficient Sampling of Biomolecular Systems: Application to Protein Structure Prediction. *J. Chem. Phys.* **2002**, *116*, 9058−9067.
(80) Affentranger, R.; Tavernelli, I.; Di Iorio, E. E. A Novel Hamiltonian Replica Exchange MD Protocol to Enhance Protein Conformational Space Sampling. *J. Chem. Theory Comput.* **2006**, *2*, 217−228.
(81) Liu, P.; Kim, B.; Friesner, R. A.; Berne, B. J. Replica Exchange with Solute Tempering: A Method for Sampling Biological Systems in Explicit Water. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 13749−13754.
(82) Kannan, S.; Zacharias, M. Enhanced Sampling of Peptide and Protein Conformations Using Replica Exchange Simulations with a Peptide Backbone Biasing-Potential. *Proteins: Struct., Funct., Bioinf.* **2007**, *66*, 697−706.
(83) Sugita, Y.; Kitao, A.; Okamoto, Y. Multidimensional Replica-Exchange Method for Free-Energy Calculations. *J. Chem. Phys.* **2000**, *113*, 6042−6051.
(84) Kruse, H.; Havrila, M.; Sponer, J. QM Computations on Complete Nucleic Acids Building Blocks: Analysis of the Sarcin−Ricin RNA Motif Using DFT-D3, HF-3c, PM6-D3H and MM Approaches. *J. Chem. Theory Comput.* **2014**, DOI: 10.1021/ct500183w.
(85) Gaillard, T.; Case, D. A. Evaluation of DNA Force Fields in Implicit Solvation. *J. Chem. Theory Comput.* **2011**, *7*, 3181−3198.
(86) Barducci, A.; Bonomi, M.; Parrinello, M. Metadynamics. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2011**, *1*, 826−843.
(87) Stadlbauer, P.; Krepl, M.; Cheatham, T. E.; Koca, J.; Sponer, J. Structural Dynamics of Possible Late-Stage Intermediates in Folding of Quadruplex DNA Studied by Molecular Simulations. *Nucleic Acids Res.* **2013**, *41*, 7128−7143.
(88) Kim, E.; Yang, C.; Pak, Y. Free-Energy Landscape of a Thrombin-Binding DNA Aptamer in Aqueous Environment. *J. Chem. Theory Comput.* **2012**, *8*, 4845−4851.
(89) Limongelli, V.; De Tito, S.; Cerofolini, L.; Fragai, M.; Pagano, B.; Trotta, R.; Cosconati, S.; Marinelli, L.; Novellino, E.; Bertini, I.;

Randazzo, A.; Luchinat, C.; Parrinello, M. The G-Triplex DNA. *Angew. Chem., Int. Ed.* **2013**, *52*, 2269−2273.
(90) Stefl, R.; Cheatham, T. E.; Spackova, N.; Fadrna, E.; Berger, I.; Koca, J.; Sponer, J. Formation Pathways of a Guanine-Quadruplex DNA Revealed by Molecular Dynamics and Thermodynamic Analysis of the Substates. *Biophys. J.* **2003**, *85*, 1787−1804.
(91) Sorin, E. J.; Rhee, Y. M.; Pande, V. S. Does Water Play a Structural Role in the Folding of Small Nucleic Acids? *Biophys. J.* **2005**, *88*, 2516−2524.
(92) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. A Second Generation Force Field for the Simulation of Proteins, Nucleic Acids, and Organic Molecules. *J. Am. Chem. Soc.* **1995**, *117*, 5179−5197.
(93) Cheatham, T. E.; Cieplak, P.; Kollman, P. A. A Modified Version of the Cornell et al. Force Field with Improved Sugar Pucker Phases and Helical Repeat. *J. Biomol. Struct. Dyn.* **1999**, *16*, 845−862.
(94) Wang, J. M.; Cieplak, P.; Kollman, P. A. How Well Does a Restrained Electrostatic Potential (RESP) Model Perform in Calculating Conformational Energies of Organic and Biological Molecules? *J. Comput. Chem.* **2000**, *21*, 1049−1074.
(95) Hobza, P.; Sponer, J. Structure, Energetics, and Dynamics of the Nucleic Acid Base Pairs: Nonempirical Ab Initio Calculations. *Chem. Rev.* **1999**, *99*, 3247−3276.
(96) Sponer, J.; Zgarbova, M.; Jurecka, P.; Riley, K. E.; Sponer, J. E.; Hobza, P. Reference Quantum Chemical Calculations on RNA Base Pairs Directly Involving the2 ′-OH Group of Ribose. *J. Chem. Theory Comput.* **2009**, *5*, 1166−1179.
(97) Savelyev, A.; MacKerell, A. D. All-Atom Polarizable Force Field for DNA Based on the Classical Drude Oscillator Model. *J. Comput. Chem.* **2014**, DOI: 10.1002/jcc.23611.
(98) Perez, A.; Marchan, I.; Svozil, D.; Sponer, J.; Cheatham, T. E.; Laughton, C. A.; Orozco, M. Refinenement of the AMBER Force Field for Nucleic Acids: Improving the Description of Alpha/Gamma Conformers. *Biophys. J.* **2007**, *92*, 3817−3829.
(99) Yildirim, I.; Stern, H. A.; Kennedy, S. D.; Tubbs, J. D.; Turner, D. H. Reparameterization of RNA χ Torsion Parameters for the AMBER Force Field and Comparison to NMR Spectra for Cytidine and Uridine. *J. Chem. Theory Comput.* **2010**, *6*, 1520−1531.
(100) Drsata, T.; Perez, A.; Orozco, M.; Morozov, A. V.; Sponer, J.; Lankas, F. Structure, Stiffness and Substates of the Dickerson−Drew Dodecamer. *J. Chem. Theory Comput.* **2013**, *9*, 707−721.
(101) Huang, J.; MacKerell, A. D. CHARMM36 All-Atom Additive Protein Force Field: Validation Based on Comparison to NMR Data. *J. Comput. Chem.* **2013**, *34*, 2135−2145.
(102) Yildirim, I.; Kennedy, S. D.; Stern, H. A.; Hart, J. M.; Kierzek, R.; Turner, D. H. Revision of AMBER Torsional Parameters for RNA Improves Free Energy Predictions for Tetramer Duplexes with GC and iGiC Base Pairs. *J. Chem. Theory Comput.* **2011**, *8*, 172−181.
(103) Foloppe, N.; MacKerell, A. D. All-Atom Empirical Force Field for Nucleic Acids: I. Parameter Optimization Based on Small Molecule and Condensed Phase Macromolecular Target Data. *J. Comput. Chem.* **2000**, *21*, 86−104.
(104) Joung, I. S.; Cheatham, T. E. Determination of Alkali and Halide Monovalent Ion Parameters for Use In Explicitly Solvated Biomolecular Simulations. *J. Phys. Chem. B* **2008**, *112*, 9020−9041.
(105) Batey, R. T. Structure and Mechanism of Purine-Binding Riboswitches. *Q. Rev. Biophys.* **2012**, *45*, 345−381.
(106) Chen, S.-J.; Dill, K. A. RNA Folding Energy Landscapes. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 646−651.
(107) Ma, H.; Proctor, D. J.; Kierzek, E.; Kierzek, R.; Bevilacqua, P. C.; Gruebele, M. Exploring the Energy Landscape of a Small RNA Hairpin. *J. Am. Chem. Soc.* **2006**, *128*, 1523−1530.
(108) Zhang, Q.; Al-Hashimi, H. M. Extending the NMR Spatial Resolution Limit for RNA by Motional Couplings. *Nat. Methods* **2008**, *5*, 243−245.
(109) Chen, S.-J. RNA Folding: Conformational Statistics, Folding Kinetics, and Ion Electrostatics. *Annu. Rev. Biophys.* **2008**, *37*, 197−214.
(110) Feng, J.; Walter, N. G.; Brooks, C. L. Cooperative and Directional Folding of the preQ1 Riboswitch Aptamer Domain. *J. Am. Chem. Soc.* **2011**, *133*, 4196−4199.