JCTC Journal of Chemical Theory and Computation

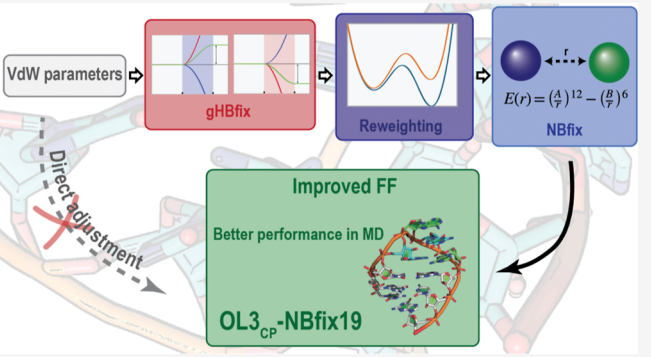
pubs.acs.org/JCTC Article

From gHBfix to NBfix: Reweighting-Driven Refinement of Hydrogen-Bond Interactions in RNA Force Fields

Vojtěch Mlýnský,[⊥] Petra Kührová,[⊥] Giovanni Bussi, Michal Otyepka, Jiří Šponer, and Pavel Banáš*

Cite This: https://doi.org/10.1021/acs.jctc.6c00924 Read Online

ABSTRACT: Understanding RNA structural dynamics is essential for elucidating its biological functions, and molecular dynamics (MD) simulations provide an important atomistic complement to experimental approaches. However, the predictive power of MD is fundamentally limited by the accuracy of the underlying empirical force fields (FFs), particularly in capturing the delicate balance of nonbonded interactions. Here, we present a systematic reparameterization strategy that replaces the external gHBfix19 hydrogen-bond (H-bond) correction potential with an equivalent and easy-to-use NBfix Lennard-Jones representation of its main thermodynamic effects within the state-of-the-art OL3 RNA FF. Using a quantitatively converged temperature replica-exchange MD ensemble of the GAGA tetraloop, we employed a reweighting-based optimization protocol to derive NBfix parameters that reproduce the thermodynamic effects of the original gHBfix19 terms. Sequential optimization of the individual gHBfix19 components proved essential to ensure stable and transferable parameter refinement. The resulting fully reformulated NBfix-based variant, termed $OL3_{CP}$-NBfix19, was validated on a representative set of RNA motifs, including tetranucleotides, A-form duplexes, and tetraloops. Across all tested systems, its performance is comparable to that of the reference gHBfix19 FF. By embedding the H-bond corrections directly into the standard nonbonded framework, the NBfix formulation eliminates external biasing potentials, simplifies practical deployment, and reduces computational overhead. Beyond this specific reparameterization, our results demonstrate a practical workflow for translating targeted H-bond corrections into native FF terms for efficient biomolecular simulations.



INTRODUCTION

Molecular dynamics (MD) simulations have become an essential tool for studying the structure, dynamics, and folding of nucleic acids at atomic resolution.[1−6] Over the past decades, continuous improvements in simulation algorithms, enhanced sampling techniques, and computational resources have substantially expanded the range of RNA systems accessible to MD simulations.[4,7] At the same time, the reliability and predictive power of these simulations remain critically dependent on the accuracy of the underlying force fields (FFs).[8−16] In particular, even subtle imbalances in the FF parameterization can lead to systematic errors in RNA conformational ensembles and folding thermodynamics, limiting the transferability of simulation results across different RNA motifs and environments.[11,15,17−19]

State-of-the-art RNA FFs derived within the AMBER framework have reached a level where major artifacts observed in early simulations, such as irreversible backbone distortions or ladder-like unfolding of A-RNA helices, have largely been eliminated.[8,20] However, extensive simulation studies have shown that current state-of-the-art RNA FFs are limited not by obvious deficiencies in canonical helices but rather by more subtle imbalances in nonbonded interactions that govern the thermodynamic stability and rugged free-energy landscapes of diverse RNA motifs.[11,15,17,19,21−27] Importantly, the manifestation of these imbalances is often system-dependent and may involve interdependent effects, complicating their identification and correction.[4,16,23,28−30] As a result, accurately reproducing the thermodynamic balance among competing conformational states remains challenging, particularly for benchmark systems such as short oligonucleotides and compact RNA motifs. To address these limitations, several targeted FF corrections have been proposed and systematically evaluated, aiming to selectively rebalance specific interactions while preserving the overall consistency of the parameterization.

Among them, the gHBfix approach introduced an additional, physically transparent correction acting directly on selected classes of hydrogen bonds (H-bonds).[11] By strengthening

Received: May 10, 2026
Revised: July 17, 2026
Accepted: August 3, 2026

under-stabilized base−base −NH···N− interactions and weakening overly favorable sugar−phosphate −OH···O− contacts, gHBfix19 was shown to significantly improve the stability of native RNA motifs and the folding behavior of challenging systems such as GNRA tetraloops (TLs).[11,24] Subsequent studies further demonstrated that gHBfix provides a flexible framework for fine-tuning H-bond interactions and can be systematically optimized using data-driven or semi-automated machine-learning approaches (gHBfix21).[19] Despite its conceptual clarity and demonstrated effectiveness, the gHBfix formalism also introduces practical limitations. The correction is implemented through additional distance-dependent potential terms that must be explicitly defined for each relevant donor−acceptor pair. This increases the complexity of simulation setup, complicates FF distribution, and introduces a non-negligible computational overhead that scales with the number of corrected interactions. While these costs are acceptable in FF development and benchmarking studies, they limit the routine use of gHBfix-augmented FFs in large-scale simulations or high-throughput applications.

A natural strategy to overcome these limitations is to transfer the physical effects encoded in gHBfix corrections into standard nonbonded FF terms that are natively supported by MD engines. Within the framework of the AMBER FF, this can be achieved using so-called NBfix modifications, which allow pair-specific adjustments of Lennard-Jones (LJ) parameters outside the conventional Lorentz−Berthelot combination rules.[31] In contrast to gHBfix, NBfix terms are straightforward to implement, do not require additional restraints, and incur no extra computational cost. However, a direct translation of gHBfix corrections into NBfix parameters is nontrivial, as it must preserve the delicate thermodynamic balance achieved by the original H-bond corrections without introducing new artifacts. A preliminary attempt to reproduce the effect of the gHBfix using pair-specific NBfix parameters was already reported in the original gHBfix work.[11] However, this direct mapping was only a proof of concept, because the effective (de)stabilization produced by NBfix depends not only on the modified atom types but also on the underlying electrostatics and structural context of a given H-bond. More broadly, pair-specific NBfix modifications have since been used in several studies to tune nucleic-acid interactions.[25,27,32−36] These developments illustrate the potential of NBfix-based corrections while emphasizing the need for a systematic strategy to derive parameters that reproduce the main thermodynamic effects of targeted H-bond corrections.

In this work, we present a systematic reparameterization of the gHBfix19 potential into a native NBfix-based representation. Building on a quantitatively converged temperature replica-exchange MD simulation (T-REMD) of the r(gcGAGAgc) TL (GAGA TL),[37] we employ a reweighting-based strategy[38,39] to identify optimal LJ parameter modifications that reproduce the thermodynamic effects of the individual gHBfix terms. By explicitly monitoring the statistical reliability of the reweighted ensembles, we ensure that the resulting NBfix parameters remain within the domain of validity of the reference FF. The derived NBfix parameterization replaces the original gHBfix19 corrections[11] while preserving the folding equilibrium of the reference GAGA TL. Beyond reweighting, we validate the new NBfix FF variant through extensive T-REMD simulations, first by substituting individual gHBfix terms with their NBfix counterparts and subsequently by replacing the entire gHBfix19 potential. Finally, we assess the basic performance of the resulting FF, denoted as $OL3_{CP}$-NBfix19, on a diverse set of benchmark RNA systems including five tetranucleotides (TNs), three A-RNA duplexes, and two TLs. Together, these results demonstrate that the physical effects of gHBfix can be robustly transferred into a computationally efficient NBfix-based formulation, providing a practical and production-ready RNA FF while retaining the benefits of targeted H-bond corrections.

METHODS

Reweighting

Reweighting[40] was applied to evaluate the thermodynamic consequences of removing each gHBfix19 term and substituting it with an NBfix-modified LJ interaction, allowing us to screen candidate parameterizations without rerunning simulations. All reweighting calculations were based on the equilibrated 298.996 K replica from the ref 25 μs-long T-REMD simulation of the GAGA TL initiated from the folded state, which was published in our recent study.[37] These simulations were performed with AMBER *ff*99bsc0χ_{OL3} RNA FF (i.e., OL3),[8,20,41,42] further adjusted by the phosphate oxygen van der Waals (vdW) correction (CP),[43] in combination with our reparameterization of the affected α, γ, δ, and ζ backbone torsions[9] and gHBfix19 H-bond correction,[11] henceforth denoted as $OL3_{CP}$-gHBfix19. The final 5 μs of this simulation were used as the training ensemble.

For each tested NBfix parameter set, the change in nonbonded energy was evaluated as the difference between the reference FF description (standard LJ interactions supplemented by gHBfix19 terms) and the target description in which the corresponding gHBfix19 corrections are removed and replaced by NBfix-modified LJ interactions. Formally, the nonbonded energy difference for a configuration x_i was computed as

$$\Delta E_{\mathrm{non}}(x_i) = \sum_n E_n^{\mathrm{gHBfix}}(x_i) - \sum_m \left(E_m^{\mathrm{NBfix}}(x_i) - E_m^{\mathrm{LJ}}(x_i)\right)$$

where the first term represents the total contribution of the gHBfix19 potentials acting on all relevant H-bond interactions n, and the second term corresponds to the difference between the NBfix-modified LJ interaction and the original LJ interaction for each affected atom-type pair m. The gHBfix energy terms were calculated as defined in our previous study,[11] while the LJ and NBfix interaction energies were evaluated using the standard 12-6 LJ functional form with either the original Lorentz−Berthelot parameters or the modified NBfix parameters, respectively.

The gHBfix19 potential comprises three energy terms: (i) an attractive correction that stabilizes base−base H-bonds by strengthening −NH···N− interactions between amino or imino H and N acceptors (AMBER H, NB, and NC atom types) by 1.0 kcal/mol, (ii) a repulsive correction that destabilizes sugar−phosphate H-bonds by weakening interactions between ribose 2′-OH hydrogens and phosphate bridging oxygens (−OH···bO− interactions; AMBER HO and OR atom types) by −0.5 kcal/mol, and (iii) an analogous repulsive correction applied to −OH···nbO− interactions between ribose 2′-OH hydrogens and phosphate nonbridging oxygens (AMBER HO and OP atom types), also by −0.5 kcal/mol.

The statistical weight of ith frame in the reweighted ensemble is then

https://doi.org/10.1021/acs.jctc.6c00924
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

$$w_i = \frac{\exp(\beta \Delta E_{non}(x_i))}{\sum_i \exp(\beta \Delta E_{non}(x_i))}$$

with $\beta = 1/k_B T$. Any observable $\langle A \rangle$ in the modified FF can be obtained by $\langle A \rangle = \sum_i w_i A_i$. In our procedure, we estimated the reweighted fraction of the folded state at ~298 K that was defined as a total weight of the snapshot having the ϵrmsd[44] difference from the native state conformation below 0.7. This quantity served as the primary target observable for matching the behavior of the $OL3_{CP}$-gHBfix19 reference ensemble. Because reweighting becomes unreliable when the target FF deviates too strongly from the control FF used to generate the training ensemble, the individual gHBfix19 terms were reparameterized separately. To quantify the effective statistical support of the reweighted ensemble, we employed the normalized exponential Shannon entropy, hereafter referred to as the normalized effective sample size P_{eff}:

$$P_{eff} = \frac{\exp(-\sum_i w_i \ln w_i)}{N}$$

where N is the number of snapshots in the training set. The P_{eff} values close to 1 indicate that a large fraction of the reference ensemble contributes meaningfully to the reweighted distribution, while values approaching zero signal that only a few frames dominate and the reweighting results are therefore unreliable.

Starting Structures and Simulation Setup

All reweighting calculations and validation simulations were based on the r(gcGAGAgc) 8-mer RNA TL (GAGA TL), which served as the primary reference system. The reference native state (i.e., the folded structure) was taken from the 1.04 Å resolution X-ray structure of the sarcin/ricin loop (PDB ID 1Q9A,[45] residues 2658−2663) and capped by one additional GC base pair, yielding the final r(gcGAGAgc) sequence. The equilibrated training ensemble used for reweighting was extracted from the final 5 μs of a previously published 25 μs-long T-REMD simulation initiated from the folded state and performed using the $OL3_{CP}$-gHBfix19 FF.[37] This ensemble corresponds to the 298.996 K replica and was shown to be fully converged with respect to folding−unfolding equilibrium.[37] All subsequent validations by T-REMD simulations in the present work were started directly from the final snapshots of this reference simulation, ensuring continuity between the training ensemble and newly generated trajectories.

To evaluate the general performance of the derived NBfix-based FF (hereafter denoted $OL3_{CP}$-NBfix19) beyond the reference TL, additional standard MD simulations were carried out on a diverse set of RNA systems, including five TNs, three A-RNA duplexes, and two TL motifs (Table 1). The TN set comprised r(AAAA), r(CAAU), r(CCCC), r(GACC), and r(UUUU) sequences constructed using the Nucleic Acid Builder module of AmberTools18.[46] The RNA duplex systems included (i) a decamer with the $r(GCACCGUUGG)_2$ sequence, excised from the crystal structure (PDB ID 1QC0;[47] 1QC0 duplex); (ii) an $r(UUAUAUAUAUAUAA)_2$ tetradecamer (PDB ID 1RNA;[48] 1RNA duplex); and (iii) a short $r(GGCC)_2$ duplex constructed by Nucleic Acid Builder. In addition, two TL systems were examined: (i) r(ggcacUUCGgugcc) 14-mer TL (UUCG TL) taken from the NMR ensemble (PDB ID 2KOC,[49] structure #13), and (ii) the GAGA TL prepared from the 1Q9A,[45] X-ray structure as described above.

Table 1. Overview of All Enhanced Sampling and Standard MD Simulations Performed in This Study[a]

method	replicas	system	FF	time [μs]
T-REMD[b]	64	GAGA TL	$OL3_{CP}$-gHBfix19	25
T-REMD	64	GAGA TL	$OL3_{CP}$-gHBfix19-$NBfix_{NH\cdots N}$	5
T-REMD	64	GAGA TL	$OL3_{CP}$-gHBfix19-$NBfix_{OH\cdots bO}$	5
T-REMD	64	GAGA TL	$OL3_{CP}$-gHBfix19-$NBfix_{OH\cdots nbO}$	5
T-REMD	64	GAGA TL	$OL3_{CP}$-NBfix19	10
REST2[c]	8	r(AAAA)	$OL3_{CP}$-gHBfix19	10
REST2[c]	8	r(CAAU)	$OL3_{CP}$-gHBfix19	10
REST2[c]	8	r(CCCC)	$OL3_{CP}$-gHBfix19	10
REST2[c]	8	r(GACC)	$OL3_{CP}$-gHBfix19	10
REST2[c]	8	r(UUUU)	$OL3_{CP}$-gHBfix19	10
REST2	8	r(AAAA)	$OL3_{CP}$-NBfix19	10
REST2	8	r(CAAU)	$OL3_{CP}$-NBfix19	10
REST2	8	r(CCCC)	$OL3_{CP}$-NBfix19	10
REST2	8	r(GACC)	$OL3_{CP}$-NBfix19	10
REST2	8	r(UUUU)	$OL3_{CP}$-NBfix19	10
PT-OPES	32	$r(GGCC)_2$	$OL3_{CP}$-gHBfix19	5
PT-OPES	32	$r(GGCC)_2$	$OL3_{CP}$-NBfix19	5
MD	1	r(AAAA)	$OL3_{CP}$-gHBfix19	10
MD	1	r(CAAU)	$OL3_{CP}$-gHBfix19	10
MD	1	r(CCCC)	$OL3_{CP}$-gHBfix19	10
MD	1	r(GACC)	$OL3_{CP}$-gHBfix19	10
MD	1	r(UUUU)	$OL3_{CP}$-gHBfix19	10
MD	1	1RNA duplex	$OL3_{CP}$-gHBfix19	10
MD	1	1QC0 duplex	$OL3_{CP}$-gHBfix19	10
MD	1	GAGA TL	$OL3_{CP}$-gHBfix19	10
MD	1	UUCG TL	$OL3_{CP}$-gHBfix19	10
MD	1	r(AAAA)	$OL3_{CP}$-NBfix19	10
MD	1	r(CAAU)	$OL3_{CP}$-NBfix19	10
MD	1	r(CCCC)	$OL3_{CP}$-NBfix19	10
MD	1	r(GACC)	$OL3_{CP}$-NBfix19	10
MD	1	r(UUUU)	$OL3_{CP}$-NBfix19	10
MD	1	1RNA duplex	$OL3_{CP}$-NBfix19	10
MD	1	1QC0 duplex	$OL3_{CP}$-NBfix19	10
MD	1	GAGA TL	$OL3_{CP}$-NBfix19	10
MD	1	UUCG TL	$OL3_{CP}$-NBfix19	10

[a]See the Methods section for a detailed description of the simulation protocols, studied systems, and applied FFs. [b]Data taken from ref 37. [c]Data taken from ref 67.

The T-REMD training ensemble was generated using the $OL3_{CP}$-gHBfix19 FF. All testing simulations were performed using either the reference $OL3_{CP}$-gHBfix19 FF or a newly derived $OL3_{CP}$-NBfix19 FF, where the gHBfix19 terms were replaced by the corresponding NBfix reparameterizations introduced in this work. To validate individual NBfix reparameterizations, additional validation T-REMD simulations were carried out using *hybrid* FF variants (i.e., $OL3_{CP}$-gHBfix19-$NBfix_{NH\cdots N}$, $OL3_{CP}$-gHBfix19-$NBfix_{OH\cdots bO}$, and $OL3_{CP}$-gHBfix19-$NBfix_{OH\cdots nbO}$), in which only a single interaction type originally modified by gHBfix19 was replaced by its NBfix counterpart, while all remaining gHBfix19 terms were retained. All systems underwent a standard equilibration protocol consisting of initial solvent minimization with

https://doi.org/10.1021/acs.jctc.6c00924
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

positional restraints on RNA heavy atoms, followed by constant-pressure equilibration to stabilize the system density. The RNA solute was then minimized in several stages with gradually decreasing restraints on the sugar−phosphate backbone. Heating to 298 K was performed in two phases, first under constant-volume and subsequently under constant-pressure conditions.

General MD Settings

Systems were solvated in explicit OPC water[50] using rectangular simulation boxes with a minimum solute-box distance of ∼10 Å. Monovalent ions were added to neutralize the systems and to achieve an ∼0.15 M KCl excess salt concentration. K^+ and Cl^- ion parameters were taken from Joung and Cheatham (K^+: R = 1.590 Å, ε = 0.2795 kcal/mol; Cl^-: R = 2.760 Å, ε = 0.0117 kcal/mol).[51] Long-range electrostatic interactions were treated with the particle mesh Ewald method[52,53] using a real-space cutoff of 10 Å and a grid spacing of 1 Å. Hydrogen mass repartitioning[54] was applied to all simulations, enabling a 4 fs integration time step. All standard MD simulations were performed for 10 μs (Table 1).

T-REMD Settings

All T-REMD[55] simulations employed 64 replicas distributed over a temperature range of ∼278−461 K, selected to yield an average exchange probability of approximately ∼25%. The simulations were carried out in the *NVT* ensemble under periodic boundary conditions. Temperature was regulated using Langevin dynamics with a friction coefficient of 2 ps^{-1}, and exchange attempts were carried out every 10 ps. Validation T-REMD simulations were run for 5 μs per replica when assessing individual NBfix substitutions, whereas the fully optimized $OL3_{CP}$-NBfix19 FF was simulated for 10 μs per replica.

REST2 Settings

Replica exchange with solute tempering (REST2)[56] simulations were performed at ∼298 K (the reference replica) with 8 replicas for all five TNs (Table 1). The scaling factor (λ) values ranged from 1 to 0.601700871, corresponding to the effective solute temperature range from ∼298 K to ∼500 K. Those values were chosen to maintain an exchange rate above 20%. REST2 simulations were performed with the AMBER18[46] GPU MD simulation engine (pmemd.cuda).[57] Further details about REST2 settings can be found elsewhere.[11]

PT-OPES Settings

Folding simulations (dimerization) of the short $r(GGCC)_2$ duplex were performed using the on-the-fly probability enhanced sampling (OPES)[58] method combined with parallel tempering (PT-OPES). The duplex was solvated in explicit OPC[50] water using a truncated octahedral simulation box with a minimum solute-box distance of ∼18 Å. Na^+ and Cl^- ions were added to neutralize the system and to achieve an excess salt concentration of approximately 1 M NaCl, matching the experimental conditions.[59] The Li and Merz ion parameters optimized for the OPC water model (Na^+: R = 1.467 Å, ε = 0.0296 kcal/mol; Cl^-: R = 2.360 Å, ε = 0.6788 kcal/mol) were used.[60] PT-OPES simulations were performed using the GPU-enabled version of GROMACS 2022 in combination with PLUMED 2.8.[61,62] All bonds involving hydrogen atoms were constrained using the LINCS algorithm.[63] The real-space cutoff for electrostatic interactions was set to 10 Å, and the temperature was controlled using the stochastic velocity-rescaling thermostat.[64] A total of 32 replicas were simulated, all initiated from the native duplex conformation and distributed over a temperature range of ∼278−391 K, resulting in an average replica-exchange acceptance probability of ∼20%. Each replica was simulated for 5 μs.

A combination of rmsd and εrmsd[44] (defined as $0.5 \times \text{rmsd}^2 + 2 \times \varepsilon\text{rmsd}^2$) was used as the biasing collective variable (CV). This metric has been shown to capture both global conformational rearrangements and local structural changes associated with the formation and disruption of base-stacking and base-pairing interactions.[65] Sampling was enhanced by depositing the bias every 1 ps using a free-energy barrier parameter (BARRIER) of 30 kJ/mol and a kernel width of 1.2 (in CV units). Frames with a CV value <8.9 relative to the native reference structure were classified as folded. This cutoff includes not only fully folded (native-like) conformations but also structures with weakened hydrogen bonding in individual GC base pairs. Consequently, conformations with frayed terminal base pairs were included in the folded ensemble. The analysis tools provided by the original OPES authors[58] were used to reconstruct free-energy profiles and estimate folding free energies (ΔG_{fold}) from the corresponding folded and unfolded populations at the individual temperatures. These values correspond to the monomer concentration used in the simulation box. In the present simulations, the total monomer (single strands) concentration was $c_M = 2/(N_A \times V) \approx 23.3$ mM, where V is the simulation box volume. Standard folding free energies ($\Delta G^\circ_{\text{fold}}$) were therefore calculated as

$$\Delta G^\circ_{\text{fold}}(T) = \Delta G_{\text{fold,box}}(T) + RT\ln(c_M/c^\circ)$$

where c° = 1 M is the standard-state concentration.

Conformational Analysis

Dominant conformations of five TNs sampled during REST2 simulations were identified using the density-based clustering algorithm of Rodriguez and Laio[66] combined with εrmsd.[44] Because closely related conformational states with frequent interconversions are not fully resolved by clustering alone, we complemented this analysis with a targeted εrmsd-based state assignment. Representative PDB structures of key conformations were used as references, and state populations were estimated from the trajectories using conservative εrmsd cutoffs. Detailed methodological descriptions are provided in our previous studies.[11,17,67]

Comparison between MD and NMR Data

The conformational ensembles of five TNs obtained from REST2 simulations were compared with previously published NMR experiments.[68−70] We analyzed separately four NMR observables, i.e., (i) backbone ^3J scalar couplings, (ii) sugar ^3J scalar couplings, (iii) nuclear Overhauser effect intensities (NOEs), and (iv) the absence of specific peaks in NOE spectroscopy (uNOEs). Scalar couplings were calculated from instantaneous dihedral angles using standard Karplus relationships. NOE and uNOE intensities were obtained by ensemble averaging over the reference (unbiased) trajectories from REST2 simulations, applying the conventional inverse sixth-power distance dependence and subsequent back-transformation to effective distances.[38,67]

For the observed quantities (scalar couplings and NOEs), agreement between simulation and experiment was quantified using a χ^2 metric. As in previous studies,[15,38,67] the agreement with uNOEs, which represent the absence of experimentally detectable NOE peaks and therefore correspond to inequality

https://doi.org/10.1021/acs.jctc.6c00924
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

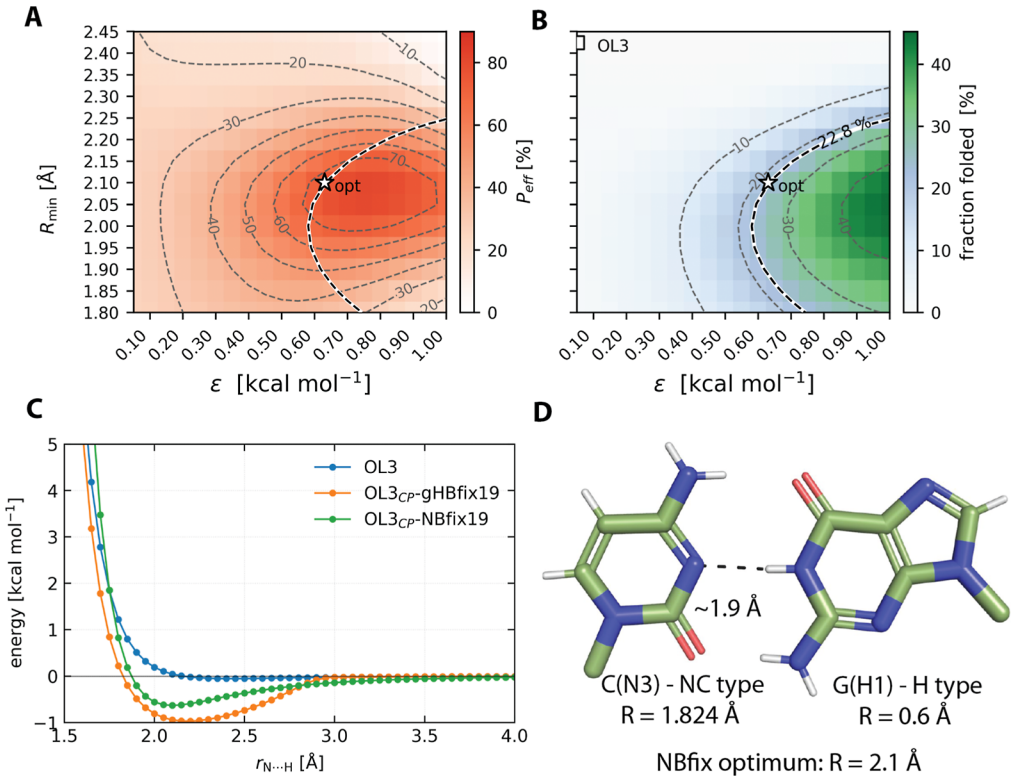


Figure 1. Reweighting-based optimization of NBfix parameters for representing the gHBfix19 −NH···N− correction through the donor-H/acceptor-N LJ pair. (A) Normalized effective sample size P_{eff} of the reweighted ensemble as a function of LJ radius R and well depth ε. (B) Reweighted folded-state fraction over the same range of parameters, with the dashed contour denoting the target folded-state population from the converged OL3$_{CP}$-gHBfix19 T-REMD simulation. The star indicates the optimal NBfix parameters, and the square the original OL3 values. (C) LJ interaction profiles for the donor-H/acceptor-N pair in OL3, OL3$_{CP}$-gHBfix19, and the optimized NBfix model. (D) Canonical GC base pair showing the AMBER vdW radii for the C(N3) and G(H1) atoms, their typical H-bond distance, and the optimized NBfix radius.

constraints rather than direct measurements, was evaluated using a one-sided penalty function that quantifies violations of the corresponding distance limits. Because this quantity does not follow a χ^2 distribution, we report it separately in this study as a uNOE violation score p. In contrast to earlier studies where individual contributions were combined by averaging over all NMR signals including the unobserved NOEs,[15,38,67] we here report the overall agreement between simulation and experiment S_{total} as the arithmetic mean of the four observable classes (backbone ^3J couplings, sugar ^3J couplings, NOEs, and uNOE constraints). This procedure avoids an artificial statistical bias arising from the typically larger number of uNOE constraints and ensures comparable contributions of the different experimental observables to the final agreement metric. Lower values of the individual χ^2 and p scores and of the overall metric S_{total} indicate better agreement between the simulated ensembles and experimental measurements.

RESULTS AND DISCUSSION

The goal of this study was to develop an NBfix-based parameterization that reproduces the effects of the gHBfix19 potential within the OL3$_{CP}$-gHBfix19 RNA FF. The gHBfix formalism provides a flexible and physically transparent way to fine-tune specific H-bond interactions, but its implementation is computationally demanding and requires additional restraint definitions for each donor−acceptor pair. To enable routine use in large-scale simulations, we sought to develop a methodology to transfer the physical effects of the gHBfix potential into a common NBfix formulation that modifies the LJ parameters for selected atom-type pairs outside the standard Lorentz−Berthelot combination rules. This approach aims to preserve the main thermodynamic effect of the reference correction, in this particular case the OL3$_{CP}$-gHBfix19 RNA FF, while achieving the same thermodynamic balance through a simpler and fully native nonbonded formulation compatible with standard AMBER FFs.

Reweighting-Driven Reparameterization of gHBfix19 into NBfix Terms

To identify NBfix parameters reproducing the effect of gHBfix19 corrections, we employed a reweighting approach based on an equilibrated ensemble of the GAGA TL obtained from T-REMD simulation. As the training set, we used the final 5 μs of the 25 μs T-REMD trajectory initiated from the folded state, which was previously shown to be fully converged with respect to folding−unfolding equilibrium.[37] This ensemble provided the donor−acceptor distance distributions required to evaluate all H-bond classes affected by the gHBfix19 terms.[11] For every tested NBfix parameter combination, we recalculated the nonbonded energies and reweighted the ensemble to estimate the resulting folded-state population and the statistical reliability of reweighting, expressed by the P_{eff} values (see Methods). This procedure allowed us to systematically explore the LJ parameter space (radius (R) and well depth (ε)) for each interaction type and to select optimal values that best reproduce the target thermodynamic balance provided by the reference OL3$_{CP}$-gHBfix19 FF while maintaining physical consistency with the parent OL3 FF.

https://doi.org/10.1021/acs.jctc.6c00924
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

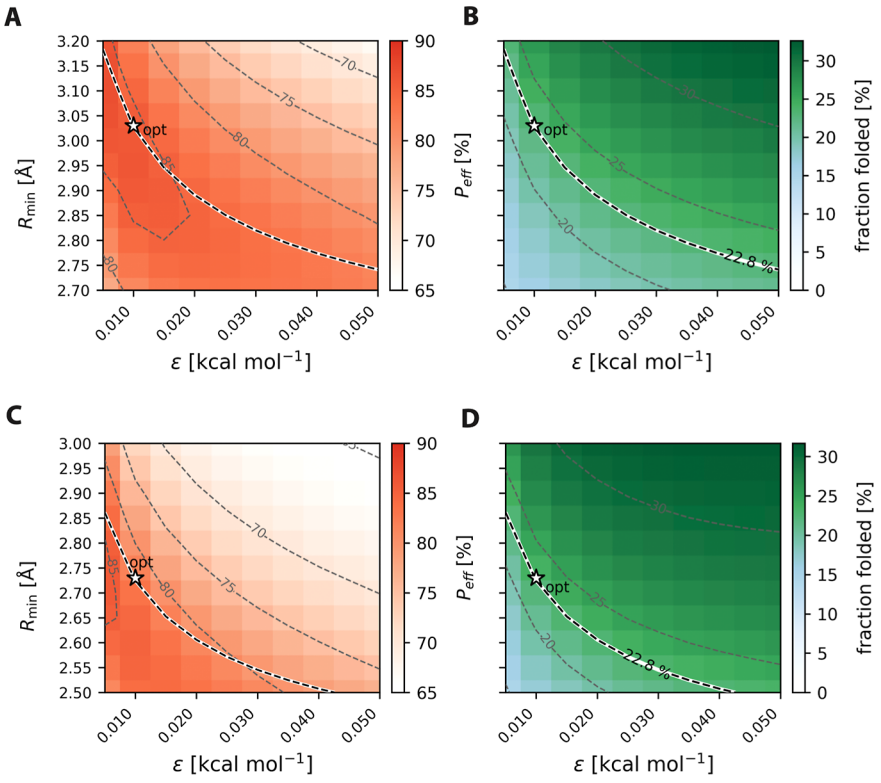


Figure 2. Reweighting-based optimization of NBfix parameters for representing the gHBfix19 −OH···(n)bO− terms through the donor-H/acceptor-O atom pair. Panels (A,C) show P_{eff} values of the reweighted ensembles as a function of R and ε for the −OH···bO− and −OH···nbO− terms, respectively. Panels (B,D) correspond to reweighted folded-state fractions, with dashed contours marking the target value obtained from the converged $OL3_{CP}$-gHBfix19 T-REMD simulation. Stars indicate optimal NBfix parameters; the original OL3 FF values (no modification) fall outside the plotted range.

The GAGA TL ensemble was selected as the training set because the converged T-REMD simulation samples a broad distribution of geometries for all three interaction classes targeted by gHBfix19. As shown in Figure S1 in the Supporting Information, −NH···N−, −OH···bO−, and −OH···nbO− contacts are all populated in this ensemble across multiple structural contexts, providing direct structural support for each part of the gHBfix19-to-NBfix19 mapping. Thus, the optimization was not intended to fit the properties of the GAGA TL as a single RNA motif but to derive NBfix LJ atom-pair parameters over a representative set of sampled gHBfix-relevant interaction geometries. At the same time, a single training ensemble cannot exhaustively cover all RNA sequence and motif contexts. For example, some base-pairing geometries, including canonical AU *cis* Watson−Crick/Watson−Crick interactions,[71] are not represented in the GAGA TL. The present work, therefore, establishes the feasibility of the reweighting-based gHBfix-to-NBfix mapping for gHBfix19 while suggesting that future extensions to more complex correction schemes, such as gHBfix21,[19] would benefit from broader and structurally richer training ensembles.

Having established the suitability and scope of the training ensemble, we reparameterized the gHBfix19 terms individually because the P_{eff} values of reweighting decrease with increasing deviation from the reference FF. The gHBfix19 potential was introduced to correct two key deficiencies of the OL3 FF: (i) the underestimated stability of base−base H-bonds, addressed by adding +1.0 kcal/mol stabilization to all −NH···N− interactions, and (ii) the overstabilized ribose−phosphate contacts, reduced by penalizing −OH···(n)bO− interactions by −0.5 kcal/mol.[11] Since the −NH···N− term plays a central role in stabilizing canonical and noncanonical base pairing, we addressed it first. Although the gHBfix19 −NH···N− correction acts on the donor-H···acceptor-N distance (i.e., the hydrogen···heavy-atom distance), its NBfix counterpart can be introduced either on the same donor-H/acceptor-N LJ pair or on the associated donor-N/acceptor-N heavy-atom LJ pair of a given H-bond. To identify which representation best reproduces the original gHBfix19 behavior, we performed two-dimensional reweighting scans over the LJ radius R and well depth ε for both alternatives using the equilibrated T-REMD ensemble of the GAGA TL.

The reweighting demonstrates that representing the gHBfix19 −NH···N− correction through the donor-H/acceptor-N LJ pair, rather than through the corresponding heavy-atom LJ pair (donor-N/acceptor-N), can simultaneously recover the correct folded-state population and maintain acceptable reliability (Figure 1). In contrast, tuning this heavy-atom LJ atom pair fails to recover the correct thermodynamic balance and yields negligible P_{eff} values of the reweighting (Figure S2 in the Supporting Information). Optimal parameters were selected to reproduce the folded fraction obtained with the reference $OL3_{CP}$-gHBfix19 FF while maximizing the reliability. The best-performing parameters were found at R = 2.1 Å and ε = 0.63 kcal/mol, compared to the default OL3 values of R = 2.424 Å and ε = 0.0517 kcal/mol. This represents a modest strengthening and slight contraction from the original OL3 FF values, i.e., a larger ε and slightly smaller R, consistent with the intended +1.0 kcal/mol stabilization introduced by the gHBfix19 potential. These

https://doi.org/10.1021/acs.jctc.6c00924
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

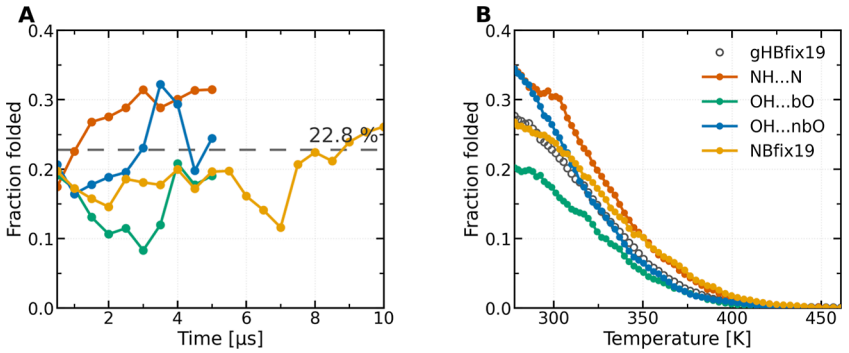


Figure 3. Validation of NBfix reparameterizations through 64-replica T-REMD simulations, either substituting the individual gHBfix19 terms with their corresponding NBfix or using the full NBfix model replacing all gHBfix19 corrections, i.e., the $OL3_{CP}$-NBfix19 FF. (A) Time evolution of the folded-state population at 298.996 K. (B) Folded-state fractions averaged over the final 2 μs of each simulation including the $OL3_{CP}$-gHBfix19 reference as a function of temperature.

parameters were applied to the LJ pair between H atoms of amino or imino groups (AMBER atom type H) and N acceptors capable of forming H-bonds (AMBER atom types NB and NC, corresponding to purine N7 and pyrimidine N3/N1 atoms, respectively). The comparison of the standard OL3 FF, reference $OL3_{CP}$-gHBfix19 FF, and the newly optimized potential (Figure 1) illustrates that the new NBfix parameterization successfully reproduces the intended stabilization of −NH···N− interactions without altering the overall shape of the nonbonded potential. A slight shift in the position of the LJ minimum is nevertheless observed for the fitted $OL3_{CP}$-NBfix19 FF relative to the original $OL3_{CP}$-gHBfix19 profile (Figure 1C). To assess potential site-specific effects of this shift, we calculated radial distribution functions $g(r)$ for −NH···N− interactions, evaluated using corresponding donor-H/acceptor-N distances, in canonical RNA duplexes. The resulting profiles are nearly identical, indicating no detectable differences in these interaction-distance distributions between the two FFs (Figure S3 in the Supporting Information).

In analogy to the −NH···N− reparameterization, we next examined the reparameterization of the −OH···bO− and −OH···nbO− gHBfix19 terms that penalized each of these interactions by −0.5 kcal/mol. The OL3 (AMBER) FF assigns zero LJ parameters to H atoms of 2′-OH groups, meaning that both the attractive and repulsive components of the LJ potential between hydrogen and oxygen are entirely missing and that the FF relies on the repulsion between the corresponding heavy atoms. This approximation removes the dispersion contribution of polar hydrogens, but it also eliminates their short-range steric repulsion, which has been shown to be associated with overstabilized ribose-phosphate interactions in RNA.[10,11,21,22,72] To better understand how the gHBfix19 penalty for −OH···bO− and −OH···nbO− contacts can be represented by NBfix terms, we complemented the original R/ε scans (Figure 2) by direct scans in the C_6/C_{12} parameter space (Figures S4 and S5 in the Supporting Information). These scans show that, along the contour reproducing the target folded-state population, P_{eff} increases toward the limit of a vanishing dispersion C_6 coefficient while retaining a finite repulsive C_{12} contribution. Thus, the optimal correction is effectively repulsive, suppressing the attractive dispersion component while preserving short-range steric repulsion. This limiting solution cannot be represented exactly by a finite pair of R and ε parameters in the standard 12-6 LJ form used by AMBER-like FFs. In the R/ε representation, the limit of zero C_6 with finite C_{12} corresponds to ε approaching zero while R diverges to infinity. Setting ε exactly to zero, therefore, removes not only the dispersion contribution but also the repulsive component. We therefore fixed ε at a small but finite value of 0.01 kcal/mol and optimized R within this regularized R/ε representation (Figure 2). This choice preserves a well-defined LJ potential, retains physically meaningful R values, and approximates the C_6/C_{12} optimum without driving the LJ radius to unrealistically large values. The resulting parameters, R = 3.03 Å for the hydrogen/oxygen pair of −OH···bO− interaction and R = 2.73 Å for the hydrogen/oxygen pair of −OH···nbO− interaction, reintroduce the missing short-range repulsion between 2′-OH hydrogens and phosphate oxygens while closely approximating the weakening effect originally imposed by the gHBfix19 potential.[11]

Validation of gHBfix19 Reparameterization via NBfix

Reweighting provides an efficient way to identify NBfix parameters that should reproduce the energetic effects of the gHBfix19 corrections, and high P_{eff} values indicate that the reweighted ensemble remains statistically representative of the reference simulation. Nevertheless, any such parameterization must be validated by explicit simulations. Because the three gHBfix19 terms (−NH···N−, −OH···bO−, and −OH···nbO−) were optimized individually, we performed three corresponding validation simulations in which each interaction type was selectively removed from the gHBfix19 and replaced by its optimized NBfix counterpart. For each case, we initiated a 64-replica T-REMD simulation of 5 μs per replica using the final ensemble from the original 25 μs-long $OL3_{CP}$-gHBfix19 T-REMD run as starting structures. In all three tests, the folded-state population remained close to the reference value of 22.8% obtained with $OL3_{CP}$-gHBfix19 FF. Replacing the −NH···N− term yielded a folded fraction of 30.4% ± 1.2% (mean and standard deviation of fraction folded over final 2 μs derived from the εrmsd-based definition with a cutoff of 0.7), while substitutions of the −OH···bO− and −OH···nbO− terms gave 17.4% ± 3.8% and 26.4% ± 5.5% folded populations, respectively (Figure 3; see S6−S8 in Supporting Information for time evolution of the main structural states). Thus, two parameterizations slightly increased the folded-state population relative to the reference, whereas the −OH···bO− substitution led to a modest decrease, all remaining within acceptable deviations for a partial term-by-term replacement.

As a final test of reweighting-driven NBfix parameter optimization, we carried out a T-REMD simulation in which all three gHBfix19 components were simultaneously removed

https://doi.org/10.1021/acs.jctc.6c00924
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

and replaced by their optimized NBfix parameterizations, yielding a fully gHBfix-free variant (the $OL3_{CP}$-NBfix19 FF). This test evaluates not only the accuracy of individual NBfix substitutions but also their collective behavior and potential cross-interactions that cannot be detected in isolated reweighting scans or single-term validation runs. Starting from the same equilibrated structures used in the partial tests, the simulation was propagated for 10 μs. Over the final 2 μs, the folded-state population stabilized at 23.8% ± 2.0%, which is in close agreement with the reference folded fraction of 22.8% obtained with the original $OL3_{CP}$-gHBfix19 FF (Figure 3; see Figure S9 in Supporting Information for time evolution of the main structural states). This confirms that the jointly applied NBfix terms do not introduce adverse cooperative effects and that the complete NBfix reformulation closely preserves the intended balance between folded and unfolded ensembles.

$OL3_{CP}$-NBfix19 FF Benchmarking across Representative RNA Motifs

To evaluate the transferability of the newly derived $OL3_{CP}$-NBfix19 FF beyond the reference training system, we assessed its performance across three representative classes of RNA motifs of increasing structural complexity: (i) short TNs, (ii) canonical A-form duplexes, and (iii) TL hairpins (see Methods and Table 1). While the benchmark set is not exhaustive, it encompasses key RNA structural motifs that should be sufficiently sensitive to detect major differences in the structural behavior predicted by the two tested FFs. The results obtained with the NBfix-based FF were systematically compared with those obtained using the reference $OL3_{CP}$-gHBfix19 FF.

We first examined a set of five RNA TNs (r(AAAA), r(CAAU), r(CCCC), r(GACC), and r(UUUU)), which represent highly flexible benchmark systems with NMR reference data. Standard MD simulations (10 μs-long) were initially performed for each TN using both $OL3_{CP}$-gHBfix19 and $OL3_{CP}$-NBfix19 FFs. Despite the relatively long simulation times for such small systems, these trajectories did not yield fully converged conformational ensembles. The simulations remained dominated by the initial A-form-like conformations together with partially stacked and intercalated states (Tables S1 and S2 in Supporting Information). Transitions between these states occurred on the hundreds-of-ns to μs time scale, and their relative populations varied substantially between independent trajectories (Figures S10−S14 in Supporting Information). These observations highlight the intrinsically slow conformational exchange of TN systems and indicate that even multi-μs conventional MD simulations are insufficient to reliably sample their equilibrium ensembles.

To achieve better conformational sampling, we therefore employed REST2 enhanced sampling simulations for all five TNs. REST2 substantially improved transitions among major conformational states and enabled a more reliable redistribution of populations. Both $OL3_{CP}$-NBfix19 and $OL3_{CP}$-gHBfix19 FFs produced qualitatively similar ensembles, including the native A-major states and the well-known spurious intercalated structures.[10,11,18,38,67,70,72,73] The corresponding conformer populations are summarized in Tables 2 and S3 in Supporting Information. However, the individual populations are not expected to be identical, since the two correction schemes are not formally equivalent. In particular, the −OH···bO− correction in gHBfix19 excludes the vicinal interaction between the 2′-OH group and the adjacent O3′ atom of the same ribose to minimize possible side effects, whereas the corresponding NBfix19 term acts on the atom-type pair and therefore also includes this vicinal 2′-OH···O3′ interaction. Such local differences can affect 2′-OH orientations and the stability of TN conformations, including

Table 2. Overview of REST2 Simulations of RNA TNs[a]

TNs	observable		$OL3_{CP}$-gHBfix19[b]	$OL3_{CP}$-NBfix19
r(AAAA)	clusters [%][c]	A-form	∼29	∼29
		intercalated	∼11	∼9
	NMR comparison (χ^2)[d]	^3J-backbone	0.66	0.66
		^3J-sugar	1.23	1.05
		NOE	1.08	1.20
	(p)	uNOE	1.14	1.07
		S_{total}	1.03	1.00
r(CAAU)	clusters [%][c]	A-form	∼33	∼50
		intercalated	∼51	∼31
	NMR comparison (χ^2)[d]	^3J-backbone	0.73	0.58
		^3J-sugar	1.14	0.92
		NOE	1.81	1.34
	(p)	uNOE	6.07	4.44
		S_{total}	2.44	1.82
r(CCCC)	clusters [%][c]	A-form	∼55	∼71
		intercalated	∼40	∼23
	NMR comparison (χ^2)[d]	^3J-backbone	0.44	0.35
		^3J-sugar	0.68	0.47
		NOE	1.54	1.27
	(p)	uNOE	3.22	1.98
		S_{total}	1.47	1.02
r(GACC)	clusters [%][c]	A-form	∼74	∼87
		intercalated	∼2	∼0
	NMR comparison (χ^2)[d]	^3J-backbone	0.39	0.34
		^3J-sugar	0.50	0.44
		NOE	1.13	1.20
	(p)	uNOE	0.34	0.01
		S_{total}	0.59	0.50
r(UUUU)[e]	clusters [%][c]	unstructured	∼3	∼16
		1−3/2−4 stack	∼38	∼31
	NMR comparison (χ^2)[d]	^3J-backbone	0.53	0.50
		^3J-sugar	0.87	0.80
		NOE	3.53	4.40
	(p)	uNOE	1.65	2.08
		S_{total}	1.65	1.95

[a]All simulations used 8 replicas and were run for 10 μs (the last 7 μs were used for data analysis). See the Methods section for a detailed description of the REST2 protocol, studied TNs, and applied FFs. [b]Data taken from ref 67. [c]Only main clusters (typically native A-form and spurious intercalated states) are given. See the Supporting Information for complete results from the clustering analysis. [d]The agreement between simulation and experiment was quantified using χ^2 values comparing calculated and experimental backbone ^3J scalar couplings, sugar ^3J scalar couplings, and NOEs and using the uNOE violation score p. The overall agreement metric S_{total} was calculated as the arithmetic mean of the four observable classes. See Methods for details. [e]The UUUU TN is expected to populate diverse conformations rather than a canonical A-RNA structure;[70] therefore, we report populations of unstructured states and 1−3/2−4 stacking arrangements.

https://doi.org/10.1021/acs.jctc.6c00924
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

intercalated states. This is reflected, for example, in the reduced intercalated populations of r(CAAU) and r(CCCC) in $OL3_{CP}$-NBfix19 compared to $OL3_{CP}$-gHBfix19. Importantly, these population shifts do not deteriorate the agreement with the available NMR observables, as reflected by comparable χ^2 and p values for the individual observable classes and by the overall S_{total} metric (Table 2). Overall, the $OL3_{CP}$-NBfix19 FF retains the main conformational effects of the reference $OL3_{CP}$-gHBfix19 FF for the TN set while preserving or improving the agreement with experimental data.

Next, we evaluated the stability of canonical A-form RNA helices using three duplex systems: the $r(GCACCGUUGG)_2$ decamer derived from the high-resolution crystal structure 1QC0,[47] the alternating $r(UUAUAUAUAUAUAA)_2$ duplex derived from the 1RNA[48] structure, and the short $r(GGCC)_2$ duplex. These structurally diverse duplexes were used to assess the stability and local dynamics of the native A-form structure. Standard 10 μs MD simulations were performed for the two longer duplexes and demonstrated that the $OL3_{CP}$-NBfix19 FF maintained stable A-form helices without detectable structural distortions. Global helical parameters and local base-pair geometries remained within the expected ranges for canonical A-RNA and were consistent with those obtained using the reference $OL3_{CP}$-gHBfix19 FF (Tables S4 and S5 in Supporting Information). Terminal base-pair opening (fraying) occurred occasionally at duplex ends but showed no systematic differences between the two tested FF variants (Table S5 in Supporting Information). These results indicate that the NBfix-based FF preserves the stabilizing effects of the gHBfix19 correction, which was previously shown to substantially reduce excessive base-pair fraying observed in earlier OL3 simulations lacking this correction.[11] Since standard MD simulations primarily probe local dynamics around the native A-form state, we additionally performed a preliminary enhanced-sampling study of the short $r(GGCC)_2$ duplex folding using 5 μs PT-OPES simulations to assess duplex stability over a broader conformational and temperature range. Although complete convergence of these simulations cannot be rigorously demonstrated, as indicated by the time evolution of folding and unfolding events in the demultiplexed replicas (Figures S15 and S16 in the Supporting Information), both the $OL3_{CP}$-NBfix19 and $OL3_{CP}$-gHBfix19 FFs yielded comparable free-energy profiles, with broadly consistent temperature dependences of standard ΔG°_{fold} (Figure 4). However, both FFs substantially overstabilized the folded state, placing the predicted melting temperature well above 400 K (Figure 4), in disagreement with the available experimental value of 308 K.[59] Thus, although the PT-OPES simulations reveal excessive stabilization of this short duplex by both FFs, they show no systematic difference between the gHBfix19 and NBfix19 formulations. The observed overstabilization suggests that the strengthening of base−base interactions introduced by both FFs may be excessive for this particular system. It may also reflect additional factors, such as sequence-dependent effects or the contribution of terminal residues, and a detailed analysis of these contributions is beyond the scope of the present work. Overall, the reference and newly derived FFs exhibited comparable performance across all three duplex systems.

Finally, we assessed the behavior of two TL systems representing compact RNA motifs with well-defined tertiary interactions: the GAGA TL used as the reference training system and the UUCG TL as an independent validation case

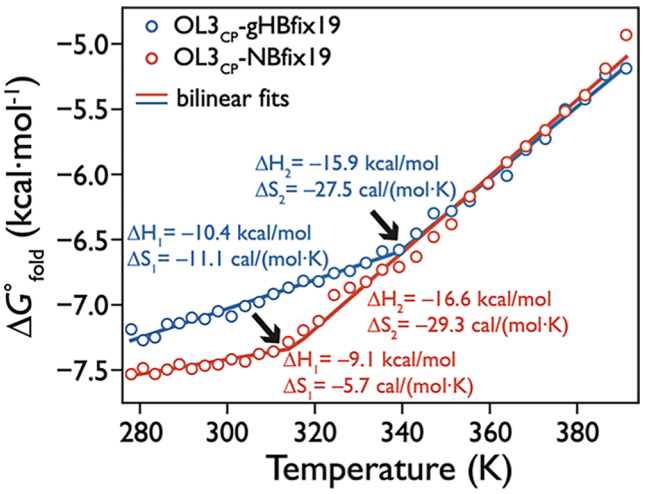


Figure 4. Temperature dependence of the standard folding free energy (ΔG°_{fold}) from PT-OPES simulations of the $r(GGCC)_2$ duplex. ΔG°_{fold} values were fitted using piecewise linear functions (bilinear fits) for both $OL3_{CP}$-gHBfix19 (blue) and $OL3_{CP}$-NBfix19 (red) FFs. Thermodynamic parameters were extracted separately from each linear regime, and the slope-change temperatures are indicated by black arrows.

(see Methods). In 10 μs standard MD simulations, the GAGA TL remained stably folded in both $OL3_{CP}$-NBfix19 and $OL3_{CP}$-gHBfix19 FFs, exhibiting low rmsd values and persistent sampling of the native loop conformation (Figures S17 and S18 in Supporting Information). This result is consistent with the T-REMD validation described above and confirms that the newly derived $OL3_{CP}$-NBfix19 FF preserves the thermodynamic balance of this reference system. In contrast, the UUCG TL lost its native loop conformation in both simulations. Disruption of the signature loop interactions occurred within the μs time scale in both simulations, leading to loss of the native loop conformation (Figures S19 and S20 in Supporting Information). The A-form stem containing five canonical base pairs remained stable in both simulations for the entire 10 μs-long time scale. The limited stability of the UUCG TL is consistent with previous studies showing that this motif remains challenging for the majority of current RNA FFs, including more recent parameterizations and polarizable models.[11,16,23] We also tested the latest machine-learning-assisted refinement of the OL3 FF incorporating experimental data, i.e., the OL3-vdW7-PAK version.[27] Rather surprisingly, it showed comparably unsatisfactory performance for the UUCG TL (see Supporting Information for full details). Notably, the original gHBfix19 correction itself was not expected to fully stabilize this motif; previous work indicated that only the later gHBfix21 version[19] approaches near-zero folding free energies for the 8-mer UUCG TL, whereas most currently available RNA FFs still predict substantially positive folding free energies for the native state.[16]

Taken together, these benchmark tests demonstrate that the $OL3_{CP}$-NBfix19 FF preserves the key structural and thermodynamic characteristics of the reference $OL3_{CP}$-gHBfix19 parameterization across diverse RNA motifs while providing a simplified and computationally efficient representation of the underlying H-bond corrections.

https://doi.org/10.1021/acs.jctc.6c00924
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

CONCLUDING REMARKS

In this work, we introduced a systematic strategy for translating externally defined H-bond correction potentials into native FF terms. Using a quantitatively converged T-REMD ensemble of the GAGA TL as a reference system, we employed a statistically controlled reweighting protocol to identify NBfix LJ parameters that reproduce the thermodynamic effects of the gHBfix19 corrections[11] within the $OL3_{CP}$ RNA FF. Reweighting approaches have previously been employed to derive correction terms for MD FFs (see, e.g., refs 19 and 74−77). In the present case, the task was facilitated by the availability of a previously optimized set of corrections that were obtained by trial-and-error optimization,[11] which we reimplemented here using a different functional form. This allowed us to use the preoptimized ensemble as a reference, thereby alleviating issues associated with insufficient statistical overlap in reweighting, which would otherwise require computationally demanding iterative procedures.[78] At the same time, the use of a standard LJ parameterization presents specific challenges. While LJ potentials are efficiently implemented in most MD codes, making the resulting parameterization readily transferable, their steep repulsive region at short distances reduces the statistical efficiency of the reweighting procedure compared to smoother H-bond correction potentials.[19] Because the statistical reliability of reweighting decreases with increasing deviation from the reference Hamiltonian, the reparameterization had to be carried out sequentially, addressing individual gHBfix19 components step by step. This approach is conceptually related to that proposed in ref 75, where reweighting was performed for individual amino acids and the combined effect of the corrections was evaluated a posteriori. The resulting fully reformulated variant, denoted $OL3_{CP}$-NBfix19, reproduces the folding equilibrium of the reference system and exhibits performance comparable to the original $OL3_{CP}$-gHBfix19 FF across a representative set of RNA motifs. Importantly, this variant can be readily implemented in standard MD packages such as AMBER[46] and GROMACS,[61] without the need for additional computationally expensive correction terms as required in ref 19.

More broadly, the present work highlights a practical strategy for addressing a long-standing challenge in FF development. Nonbonded interactions remain one of the major sources of residual imbalance in contemporary RNA FFs,[10,11,17,19,21,22,25,27,34,72,76,79−82] yet their direct reparameterization is notoriously difficult due to strong coupling between vdW parameters, electrostatics, and other FF terms. Targeted correction potentials such as gHBfix provide a powerful and physically transparent way to tune specific interaction classes without introducing widespread side effects, making them highly effective tools for diagnosing and refining problematic interactions. However, their reliance on additional external potentials and explicit donor−acceptor restraints complicate their routine use in large-scale production simulations. By demonstrating how the effects of such targeted corrections can be quantitatively transferred into native NBfix parameters, our approach bridges the gap between efficient FF tuning and practical production-level FF implementations.

The present analysis also provides a more specific insight into the physical origin of the problematic 2′-OH···phosphate contacts and, more generally, into the FF representation of polar hydrogens. The C_6/C_{12} scans indicate that these polar-hydrogen interactions are best described by suppressing the attractive dispersion component while retaining short-range steric repulsion. This limiting behavior is not naturally captured by the conventional R/ε LJ representation, in which both components are coupled. Thus, beyond yielding a practical NBfix19 parameterization, the present work highlights a more general limitation of the standard vdW description of polar hydrogens in AMBER-like FFs.

Finally, the workflow presented here establishes a general framework for reweighting-driven refinement of nonbonded FF parameters. While the present study focused on the relatively compact gHBfix19 correction scheme, the same strategy can, in principle, be applied to more complex correction frameworks or larger sets of interaction classes such as gHBfix21.[19] In such cases, automated optimization procedures and machine-learning-assisted parameter exploration will likely become necessary to efficiently explore the high-dimensional parameter space. Nevertheless, the present results demonstrate that targeted interaction corrections can be systematically translated into efficient native FF parameterizations while preserving the intended thermodynamic balance of RNA conformational ensembles.

ASSOCIATED CONTENT

Data Availability Statement

All raw MD simulation data are available through the IDA4SIMS database: https://next.ida.4sims.eu/IDA4SIMS.

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jctc.6c00924.

ParmEd input file for the NBfix19 modification of the $OL3_{CP}$ FF, details about additional simulations with the OL3-vdW7-PAK FF, and supporting tables and figures (PDF)

AUTHOR INFORMATION

Corresponding Author

Pavel Banáš − *Regional Center of Advanced Technologies and Materials, The Czech Advanced Technology and Research Institute (CATRIN), Palacký University Olomouc, 779 00 Olomouc, Czech Republic; IT4Innovations, VSB-Technical University of Ostrava, 708 00 Ostrava-Poruba, Czech Republic;* orcid.org/0000-0002-7137-8225; Email: pavel.banas@upol.cz

Authors

Vojtěch Mlýnský − *Institute of Biophysics of the Czech Academy of Sciences, 612 00 Brno, Czech Republic;* orcid.org/0000-0003-2769-1553

Petra Kührová − *Institute of Biophysics of the Czech Academy of Sciences, 612 00 Brno, Czech Republic; Regional Center of Advanced Technologies and Materials, The Czech Advanced Technology and Research Institute (CATRIN), Palacký University Olomouc, 779 00 Olomouc, Czech Republic;* orcid.org/0000-0003-1593-5282

Giovanni Bussi − *Scuola Internazionale Superiore di Studi Avanzati, SISSA, 34136 Trieste, Italy;* orcid.org/0000-0001-9216-5782

Michal Otyepka − *Regional Center of Advanced Technologies and Materials, The Czech Advanced Technology and Research Institute (CATRIN), Palacký University Olomouc, 779 00 Olomouc, Czech Republic; IT4Innovations, VSB-*

https://doi.org/10.1021/acs.jctc.6c00924
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

Technical University of Ostrava, 708 00 Ostrava-Poruba, Czech Republic; orcid.org/0000-0002-1066-5677

Jiří Šponer − *Institute of Biophysics of the Czech Academy of Sciences, 612 00 Brno, Czech Republic; Regional Center of Advanced Technologies and Materials, The Czech Advanced Technology and Research Institute (CATRIN), Palacký University Olomouc, 779 00 Olomouc, Czech Republic;* orcid.org/0000-0001-6558-6186

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jctc.6c00924

Author Contributions

[⊥]V.M. and P.K. contributed equally.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This work was supported by the Czech Science Foundation (grant number 23-05639S; to V.M. and J.Š.). This research also received the support of EXA4MIND, a European Union's Horizon Europe Research and Innovation programme under grant agreement No. 101092944 (V.M., M.O., and P.B.). Views and opinions expressed are, however, those of the author(s) only and do not necessarily reflect those of the European Union or the European Commission. Neither the European Union nor the granting authority can be held responsible for them. P.K., M.O., and P.B. were supported by the ERDF/ESF project TECHSCALE (No. CZ.02.01.01/00/22_008/0004587). G.B. acknowledges the Italian National Centre for HPC, Big Data, and Quantum Computing (grant No. CN00000013), founded within the Next Generation EU initiative. This work was also supported by the Ministry of Education, Youth and Sports of the Czech Republic through the e-INFRA CZ (ID: 90254).

■ REFERENCES

(1) Cheatham, T. E. Simulation and modeling of nucleic acid structure, dynamics and interactions. *Curr. Opin. Struct. Biol.* **2004**, *14* (3), 360−367.

(2) Sponer, J.; Banas, P.; Jurecka, P.; Zgarbova, M.; Kuhrova, P.; Havrila, M.; Krepl, M.; Stadlbauer, P.; Otyepka, M. Molecular Dynamics Simulations of Nucleic Acids. From Tetranucleotides to the Ribosome. *J. Phys. Chem. Lett.* **2014**, *5* (10), 1771−1782.

(3) Smith, L. G.; Zhao, J.; Mathews, D. H.; Turner, D. H. Physics-based all-atom modeling of RNA energetics and structure. *Wiley Interdiscip. Rev. RNA* **2017**, *8* (5), No. e1422.

(4) Sponer, J.; Bussi, G.; Krepl, M.; Banas, P.; Bottaro, S.; Cunha, R. A.; Gil-Ley, A.; Pinamonti, G.; Poblete, S.; Jurecka, P.; et al. RNA Structural Dynamics As Captured by Molecular Simulations: A Comprehensive Overview. *Chem. Rev.* **2018**, *118* (8), 4177−4338.

(5) Palermo, G.; Casalino, L.; Magistrato, A.; Andrew McCammon, J. Understanding the mechanistic basis of non-coding RNA through molecular dynamics simulations. *J. Struct. Biol.* **2019**, *206* (3), 267−279.

(6) Muscat, S.; Martino, G.; Manigrasso, J.; Marcia, M.; De Vivo, M. On the Power and Challenges of Atomistic Molecular Dynamics to Investigate RNA Molecules. *J. Chem. Theory Comput.* **2024**, *20*, 6992.

(7) Mlynsky, V.; Bussi, G. Exploring RNA structure and dynamics through enhanced sampling simulations. *Curr. Opin. Struct. Biol.* **2018**, *49*, 63−71.

(8) 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* (9), 2886−2902.

(9) Mlynsky, V.; Kuhrova, P.; Zgarbova, M.; Jurecka, P.; Walter, N. G.; Otyepka, M.; Sponer, J.; Banas, P. Reactive conformation of the active site in the hairpin ribozyme achieved by molecular dynamics simulations with epsilon/zeta force field reparametrizations. *J. Phys. Chem. B* **2015**, *119* (11), 4220−4229.

(10) Bergonzo, C.; Cheatham, T. E., 3rd Improved Force Field Parameters Lead to a Better Description of RNA Structure. *J. Chem. Theory Comput.* **2015**, *11* (9), 3969−3972.

(11) Kuhrova, P.; Mlynsky, V.; Zgarbova, M.; Krepl, M.; Bussi, G.; Best, R. B.; Otyepka, M.; Sponer, J.; Banas, P. Improving the Performance of the Amber RNA Force Field by Tuning the Hydrogen-Bonding Interactions. *J. Chem. Theory Comput.* **2019**, *15* (5), 3288−3305.

(12) Zhao, J.; Kennedy, S. D.; Berger, K. D.; Turner, D. H. Nuclear Magnetic Resonance of Single-Stranded RNAs and DNAs of CAAU and UCAAUC as Benchmarks for Molecular Dynamics Simulations. *J. Chem. Theory Comput.* **2020**, *16* (3), 1968−1984.

(13) Zhao, J.; Kennedy, S. D.; Turner, D. H. Nuclear Magnetic Resonance Spectra and AMBER OL3 and ROC-RNA Simulations of UCUCGU Reveal Force Field Strengths and Weaknesses for Single-Stranded RNA. *J. Chem. Theory Comput.* **2022**, *18* (2), 1241−1254.

(14) Liebl, K.; Zacharias, M. The development of nucleic acids force fields: From an unchallenged past to a competitive future. *Biophys. J.* **2023**, *122* (14), 2841−2851.

(15) Mlynsky, V.; Kuhrova, P.; Stadlbauer, P.; Krepl, M.; Otyepka, M.; Banas, P.; Sponer, J. Simple Adjustment of Intranucleotide Base-Phosphate Interaction in the OL3 AMBER Force Field Improves RNA Simulations. *J. Chem. Theory Comput.* **2023**, *19* (22), 8423−8433.

(16) Mlynsky, V.; Kührová, P.; Pykal, M.; Krepl, M.; Stadlbauer, P.; Otyepka, M.; Banás, P.; Sponer, J. Can We Ever Develop an Ideal RNA Force Field? Lessons Learned from Simulations of the UUCG RNA Tetraloop and Other Systems. *J. Chem. Theory Comput.* **2025**, *21* (8), 4183−4202.

(17) Kuhrova, P.; Best, R. B.; Bottaro, S.; Bussi, G.; Sponer, J.; Otyepka, M.; Banas, P. Computer Folding of RNA Tetraloops: Identification of Key Force Field Deficiencies. *J. Chem. Theory Comput.* **2016**, *12* (9), 4534−4548.

(18) Cesari, A.; Gil-Ley, A.; Bussi, G. Combining Simulations and Solution Experiments as a Paradigm for RNA Force Field Refinement. *J. Chem. Theory Comput.* **2016**, *12* (12), 6192−6200.

(19) Frohlking, T.; Mlynsky, V.; Janecek, M.; Kuhrova, P.; Krepl, M.; Banas, P.; Sponer, J.; Bussi, G. Automatic Learning of Hydrogen-Bond Fixes in the AMBER RNA Force Field. *J. Chem. Theory Comput.* **2022**, *18* (7), 4490−4502.

(20) Perez, A.; Marchan, I.; Svozil, D.; Sponer, J.; Cheatham, T. E.; Laughton, C. A.; Orozco, M. Refinement of the AMBER Force Field for Nucleic Acids: Improving the Description of Alpha/Gamma Conformers. *Biophys. J.* **2007**, *92* (11), 3817−3829.

(21) Yang, C.; Lim, M.; Kim, E.; Pak, Y. Predicting RNA Structures via a Simple van der Waals Correction to an All-Atom Force Field. *J. Chem. Theory Comput.* **2017**, *13* (2), 395−399.

(22) Tan, D.; Piana, S.; Dirks, R. M.; Shaw, D. E. RNA force field with accuracy comparable to state-of-the-art protein force fields. *Proc. Natl. Acad. Sci. U.S.A.* **2018**, *115* (7), E1346−E1355.

(23) Mrazikova, K.; Mlynsky, V.; Kuhrova, P.; Pokorna, P.; Kruse, H.; Krepl, M.; Otyepka, M.; Banas, P.; Sponer, J. UUCG RNA Tetraloop as a Formidable Force-Field Challenge for MD Simulations. *J. Chem. Theory Comput.* **2020**, *16* (12), 7601−7617.

(24) Mlynsky, V.; Janecek, M.; Kuhrova, P.; Frohlking, T.; Otyepka, M.; Bussi, G.; Banas, P.; Sponer, J. Toward Convergence in Folding Simulations of RNA Tetraloops: Comparison of Enhanced Sampling Techniques and Effects of Force Field Modifications. *J. Chem. Theory Comput.* **2022**, *18* (4), 2642−2656.

(25) Raguette, L. E.; Gunasekera, S. S.; Diaz Ventura, R. I.; Aminov, E.; Linzer, J. T.; Parwana, D.; Wu, Q.; Simmerling, C.; Nagan, M. C. Adjusting the Energy Profile for CH-O Interactions Leads to

Improved Stability of RNA Stem-Loop Structures in MD Simulations. *J. Phys. Chem. B* **2024**, *128* (33), 7921−7933.
(26) Puyo-Fourtine, J.; Zgarbová, M.; Kantin, L.; Jurecka, P.; Duboué-Dijon, E. Implicit Electronic Polarization via Charge Scaling in Amber OLx Nucleic Acid Force Fields Improves Ion Pairing without Backbone Reparametrization. *J. Chem. Theory Comput.* **2025**, *21* (17), 8583−8600.
(27) Kim, H.; Pak, Y. Ensemble-Based Precision Refinement of All-Atom Nucleic Acid Force Fields Guided by NMR NOE Pair-Distance Measurements. *J. Chem. Theory Comput.* **2025**, *21* (21), 11245−11258.
(28) Krepl, M.; Pokorna, P.; Mlynsky, V.; Stadlbauer, P.; Sponer, J. Spontaneous binding of single-stranded RNAs to RRM proteins visualized by unbiased atomistic simulations with a rescaled RNA force field. *Nucleic Acids Res.* **2022**, *50* (21), 12480−12496.
(29) Lemmens, T.; Mlynsky, V.; Sponer, J.; Pykal, M.; Banas, P.; Otyepka, M.; Krepl, M. The Kink-Turn 7 Motif: An Additional Test for RNA Force Field Performance. *J. Chem. Theory Comput.* **2025**, *21* (24), 12796−12809.
(30) Krepl, M.; Mlynsky, V.; Rusnakova, A.; Banas, P.; Otyepka, M.; Sponer, J. Destabilization of Structured RNAs by OPC and TIP4PD Water Models. *J. Chem. Theory Comput.* **2026**, *22* (4), 2013−2027.
(31) Yoo, J.; Aksimentiev, A. New tricks for old dogs: improving the accuracy of biomolecular force fields by pair-specific corrections to non-bonded interactions. *Phys. Chem. Chem. Phys.* **2018**, *20* (13), 8432−8449.
(32) Tucker, M. R.; Piana, S.; Tan, D.; LeVine, M. V.; Shaw, D. E. Development of Force Field Parameters for the Simulation of Single- and Double-Stranded DNA Molecules and DNA-Protein Complexes. *J. Phys. Chem. B* **2022**, *126* (24), 4442−4457.
(33) Li, Z. X.; Song, G.; Zhu, J. J.; Mu, J. X.; Sun, Y. T.; Hong, X. K.; Choi, T.; Cui, X. C.; Chen, H. F. Excited-Ground-State Transition of the RNA Strand Slippage Mechanism Captured by the Base-Specific Force Field. *J. Chem. Theory Comput.* **2024**, *20* (14), 6082−6097.
(34) Das, A. K.; Tuckerman, M. E.; Kirmizialtin, S. HB-CUFIX: Force field for accurate RNA simulations. *J. Chem. Phys.* **2025**, *162* (20), 200901.
(35) Kim, H.; Pak, Y. Free Energy-Based Refinement of DNA Force Field via Modification of Multiple Nonbonding Energy Terms. *J. Chem. Inf. Model.* **2025**, *65* (1), 288−297.
(36) Puthenpeedikakkal, A. M. K.; Cavender, C. E.; Smith, L. G.; Grossfield, A.; Mathews, D. H. Reparameterization of the Amber RNA Force Field Non-Bonded Terms. *bioRxiv* **2026**, 2026.05.18.725894.
(37) Kuhrova, P.; Mlynsky, V.; Frebort, I.; Frebortova, J.; Otyepka, M.; Sponer, J.; Banas, P. Toward Accurate RNA Folding Thermodynamics: Evaluation of Enhanced Sampling Methods for Force Field Benchmarking. *J. Chem. Theory Comput.* **2026**, *22* (10), 5247−5258.
(38) Bottaro, S.; Bussi, G.; Kennedy, S. D.; Turner, D. H.; Lindorff-Larsen, K. Conformational ensembles of RNA oligonucleotides from integrating NMR and molecular simulations. *Sci. Adv.* **2018**, *4* (5), No. eaar8521.
(39) Kofinger, J.; Rozycki, B.; Hummer, G. Inferring Structural Ensembles of Flexible and Dynamic Macromolecules Using Bayesian, Maximum Entropy, and Minimal-Ensemble Refinement Methods. *Methods Mol. Biol.* **2019**, *2022*, 341−352.
(40) Shen, T.; Hamelberg, D. A statistical analysis of the precision of reweighting-based simulations. *J. Chem. Phys.* **2008**, *129* (3), 034103.
(41) 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 (vol 117, pg 5179, 1995). *J. Am. Chem. Soc.* **1996**, *118* (9), 2309.
(42) 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* (12), 1049−1074.
(43) Steinbrecher, T.; Latzer, J.; Case, D. A. Revised AMBER Parameters for Bioorganic Phosphates. *J. Chem. Theory Comput.* **2012**, *8* (11), 4405−4412.
(44) Bottaro, S.; Di Palma, F.; Bussi, G. The role of nucleobase interactions in RNA structure and dynamics. *Nucleic Acids Res.* **2014**, *42* (21), 13306−13314.
(45) Correll, C. C.; Beneken, J.; Plantinga, M. J.; Lubbers, M.; Chan, Y. L. The common and the distinctive features of the bulged-G motif based on a 1.04 Å resolution RNA structure. *Nucleic Acids Res.* **2003**, *31* (23), 6806−6818.
(46) Case, D. A.; Ben-Shalom, I. Y.; Brozell, S. R.; Cerutti, D. S.; Cheatham, I. T. E.; Cruzeiro, V. W. D.; Darden, T. A.; Duke, R. E.; Ghoreishi, D.; Gilson, M. K.; et al. *AMBER 2018*; University of California: San Francisco, 2018.
(47) Klosterman, P. S.; Shah, S. A.; Steitz, T. A. Crystal structures of two plasmid copy control related RNA duplexes: An 18 base pair duplex at 1.20 A resolution and a 19 base pair duplex at 1.55 A resolution. *Biochemistry* **1999**, *38* (45), 14784−14792.
(48) Dockbregeon, A. C.; Chevrier, B.; Podjarny, A.; Johnson, J.; Debear, J. S.; Gough, G. R.; Gilham, P. T.; Moras, D. Crystallographic Structure of an Rna Helix - [U(Ua)6a]2. *J. Mol. Biol.* **1989**, *209* (3), 459−474.
(49) Nozinovic, S.; Furtig, B.; Jonker, H. R.; Richter, C.; Schwalbe, H. High-resolution NMR structure of an RNA model system: the 14-mer cUUCGg tetraloop hairpin RNA. *Nucleic Acids Res.* **2010**, *38* (2), 683−694.
(50) Izadi, S.; Anandakrishnan, R.; Onufriev, A. V. Building Water Models: A Different Approach. *J. Phys. Chem. Lett.* **2014**, *5* (21), 3863−3871.
(51) 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* (30), 9020−9041.
(52) Darden, T.; York, D.; Pedersen, L. Particle Mesh Ewald - An N.log(N) Method for Ewald Sums in Large Systems. *J. Chem. Phys.* **1993**, *98* (12), 10089−10092.
(53) Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. A Smooth Particle Mesh Ewald Method. *J. Chem. Phys.* **1995**, *103* (19), 8577−8593.
(54) Hopkins, C. W.; Le Grand, S.; Walker, R. C.; Roitberg, A. E. Long-Time-Step Molecular Dynamics through Hydrogen Mass Repartitioning. *J. Chem. Theory Comput.* **2015**, *11* (4), 1864−1874.
(55) Sugita, Y.; Okamoto, Y. Replica-exchange molecular dynamics method for protein folding. *Chem. Phys. Lett.* **1999**, *314* (1−2), 141−151.
(56) Wang, L.; Friesner, R. A.; Berne, B. J. Replica exchange with solute scaling: a more efficient version of replica exchange with solute tempering (REST2). *J. Phys. Chem. B* **2011**, *115* (30), 9431−9438.
(57) Mermelstein, D. J.; Lin, C.; Nelson, G.; Kretsch, R.; McCammon, J. A.; Walker, R. C. Fast and flexible gpu accelerated binding free energy calculations within the amber molecular dynamics package. *J. Comput. Chem.* **2018**, *39* (19), 1354−1358.
(58) Invernizzi, M.; Parrinello, M. Rethinking Metadynamics: From Bias Potentials to Probability Distributions. *J. Phys. Chem. Lett.* **2020**, *11* (7), 2731−2736.
(59) Xia, T. B.; SantaLucia, J. J. r.; Burkard, M. E.; Kierzek, R.; Schroeder, S. J.; Jiao, X. Q.; Cox, C.; Turner, D. H. Thermodynamic parameters for an expanded nearest-neighbor model for formation of RNA duplexes with Watson-Crick base pairs. *Biochemistry* **1998**, *37* (42), 14719−14735.
(60) Sengupta, A.; Li, Z.; Song, L. F.; Li, P.; Merz, K. M. Parameterization of Monovalent Ions for the OPC3, OPC, TIP3P-FB, and TIP4P-FB Water Models. *J. Chem. Inf. Model.* **2021**, *61* (2), 869−880.
(61) Abraham, M. J.; Murtola, T.; Schulz, R.; Páll, S.; Smith, J. C.; Hess, B.; Lindahl, E. GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **2015**, *1*, 19−25.

(62) Tribello, G. A.; Bonomi, M.; Branduardi, D.; Camilloni, C.; Bussi, G. PLUMED 2: New feathers for an old bird. *Comput. Phys. Commun.* **2014**, *185* (2), 604−613.
(63) Hess, B.; Bekker, H.; Berendsen, H. J. C.; Fraaije, J. G. E. M. LINCS: A Linear Constraint Solver for Molecular Simulations. *J. Comput. Chem.* **1997**, *18* (12), 1463−1472.
(64) Bussi, G.; Donadio, D.; Parrinello, M. Canonical Sampling through Velocity Rescaling. *J. Chem. Phys.* **2007**, *126* (1), 014101.
(65) Janecek, M.; Kührová, P.; Mlynsky, V.; Stadlbauer, P.; Otyepka, M.; Bussi, G.; Sponer, J.; Banás, P. Computer Folding of Parallel DNA G-Quadruplex: Hitchhiker's Guide to the Conformational Space. *J. Comput. Chem.* **2025**, *46* (1), No. e27535.
(66) Rodriguez, A.; Laio, A. Clustering by Fast Search and Find of Density Peaks. *Science* **2014**, *344* (6191), 1492−1496.
(67) Mlynsky, V.; Kuhrova, P.; Kuhr, T.; Otyepka, M.; Bussi, G.; Banas, P.; Sponer, J. Fine-Tuning of the AMBER RNA Force Field with a New Term Adjusting Interactions of Terminal Nucleotides. *J. Chem. Theory Comput.* **2020**, *16* (6), 3936−3946.
(68) 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 chi torsions. *J. Phys. Chem. B* **2011**, *115* (29), 9261−9270.
(69) 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* (6), 996−1010.
(70) Condon, D. E.; Kennedy, S. D.; Mort, B. C.; Kierzek, R.; Yildirim, I.; Turner, D. H. Stacking in RNA: NMR of Four Tetramers Benchmark Molecular Dynamics. *J. Chem. Theory Comput.* **2015**, *11* (6), 2729−2742.
(71) Leontis, N. B.; Westhof, E. Geometric Nomenclature and Classification of RNA Base Pairs. *RNA* **2001**, *7* (4), 499−512.
(72) Bergonzo, C.; Henriksen, N. M.; Roe, D. R.; Cheatham, T. E., 3rd Highly sampled tetranucleotide and tetraloop motifs enable evaluation of common RNA force fields. *RNA* **2015**, *21* (9), 1578−1590.
(73) Wales, D. J.; Yildirim, I. Improving Computational Predictions of Single-Stranded RNA Tetramers with Revised alpha/gamma Torsional Parameters for the Amber Force Field. *J. Phys. Chem. B* **2017**, *121* (14), 2989−2999.
(74) Norgaard, A. B.; Ferkinghoff-Borg, J.; Lindorff-Larsen, K. Experimental parameterization of an energy function for the simulation of unfolded proteins. *Biophys. J.* **2008**, *94* (1), 182−192.
(75) Li, D. W.; Brüschweiler, R. Iterative Optimization of Molecular Mechanics Force Fields from NMR Data of Full-Length Proteins. *J. Chem. Theory Comput.* **2011**, *7* (6), 1773−1782.
(76) Cesari, A.; Bottaro, S.; Lindorff-Larsen, K.; Banas, P.; Sponer, J.; Bussi, G. Fitting Corrections to an RNA Force Field Using Experimental Data. *J. Chem. Theory Comput.* **2019**, *15* (6), 3425−3431.
(77) Chen, J.; Liu, H.; Cui, X.; Li, Z.; Chen, H. F. RNA-Specific Force Field Optimization with CMAP and Reweighting. *J. Chem. Inf. Model.* **2022**, *62* (2), 372−385.
(78) Frohlking, T.; Bernetti, M.; Calonaci, N.; Bussi, G. Toward empirical force fields that match experimental observables. *J. Chem. Phys.* **2020**, *152* (23), 230902.
(79) 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* (42), 16820−16825.
(80) Li, Z.; Mu, J.; Chen, J.; Chen, H. F. Base-specific RNA force field improving the dynamics conformation of nucleotide. *Int. J. Biol. Macromol.* **2022**, *222* (Pt A), 680−690.
(81) Piomponi, V.; Froehlking, T.; Bernetti, M.; Bussi, G. Molecular Simulations Matching Denaturation Experiments for N -Methyladenosine. *ACS Cent. Sci.* **2022**, *8* (8), 1218−1228.
(82) Piomponi, V.; Krepl, M.; Sponer, J.; Bussi, G. Molecular Simulations to Investigate the Impact of N6-Methylation in RNA Recognition: Improving Accuracy and Precision of Binding Free Energy Prediction. *J. Phys. Chem. B* **2024**, *128* (37), 8896−8907.

https://doi.org/10.1021/acs.jctc.6c00924
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX