HHS Public Access

Author manuscript

*J Chem Theory Comput*. Author manuscript; available in PMC 2022 March 11.

Published in final edited form as:

*J Chem Theory Comput*. 2010 December 14; 6(12): 3836–3849. doi:10.1021/ct100481h.

# Performance of Molecular Mechanics Force Fields for RNA Simulations: Stability of UUCG and GNRA Hairpins

**Pavel Banáš**[†,‡], **Daniel Hollas**[†], **Marie Zgarbová**[†], **Petr Jurečka**[†], **Modesto Orozco**[§], **Thomas E. Cheatham III**[‖], **Jiří Šponer**[*,†,‡], **Michal Otyepka**[*,†,‡]

[†]Regional Centre of Advanced Technologies and Materials, Department of Physical Chemistry, Faculty of Science, Palacky UniVersity Olomouc, tr. 17. listopadu 12, 771 46 Olomouc, Czech Republic

[‡]Institute of Biophysics, Academy of Sciences of the Czech Republic, KraloVopolska 135, 612 65 Brno, Czech Republic

[§]Joint Research Program in Computational Biology, Institut de Recerca Biomédica and Barcelona Superocomputing Center, Baldiri i Reixac 10, Barcelona 08028, Spain, Jordi Girona 31, Barcelona 08028, Spain, and Departament de Bioquýmica i Biologýa Molecular, Facultat de Biologýa, AVga Diagonal 645 UniVersitat de Barcelona, Barcelona 08028, Spain

[‖]Departments of Medicinal Chemistry, Pharmaceutical Chemistry, and Pharmaceutics and Bioengineering, University of Utah, Salt Lake City, Utah 84112, United States

## Abstract

The RNA hairpin loops represent important RNA topologies with indispensable biological functions in RNA folding and tertiary interactions. 5′-UNCG-3′ and 5′-GNRA-3′ RNA tetraloops are the most important classes of RNA hairpin loops. Both tetraloops are highly structured with characteristic signature three-dimensional features and are recurrently seen in functional RNAs and ribonucleoprotein particles. Explicit solvent molecular dynamics (MD) simulation is a computational technique which can efficiently complement the experimental data and provide unique structural dynamics information on the atomic scale. Nevertheless, the outcome of simulations is often compromised by imperfections in the parametrization of simplified pairwise additive empirical potentials referred to also as force fields. We have pointed out in several recent studies that a force field description of single-stranded hairpin segments of nucleic acids may be particularly challenging for the force fields. In this paper, we report a critical assessment of a broad set of MD simulations of UUCG, GAGA, and GAAA tetraloops using various force fields. First, we utilized the three widely used variants of Cornell et al. (AMBER) force fields known as *ff*94, *ff*99, and *ff*99bsc0. Some simulations were also carried out with CHARMM27. The simulations reveal several problems which show that these force fields are not able to retain

[*]Corresponding author. Tel.: +420 585 634 756 (M.O.). Fax: +420 585 634 761 (M.O.). michal.otyepka@upol.cz (M.O.). Tel.: +420 541 517 133. sponer@ncbr.chemi.muni.cz.

Supporting Information **Available**: Library files with *ff*99$\chi$OL-DFT, *ff*99$\chi$OL, *ff*99bbsc0$\chi$OL-DFT, and *ff*99bsc0$\chi$OL force fields, a detailed description of how to include these force fields in the AMBER standard preparation tool *leap*, detailed information about shifted energy minima for *syn* region of glycosidic $\chi$ torsion in standard AMBER force fields, detailed description of the structural degradation of GAAA TL that appeared on a very large time scale, and a description of the simulations of short A-RNA stems including tables with helical parameters. This information is available free of charge via the Internet at http://pubs.acs.org.

all characteristic structural features (structural signature) of the studied tetraloops. Then we tested four recent reparameterizations of glycosidic torsion of the Cornell et al. force field (two of them being currently parametrized in our laboratories). We show that at least some of the new versions show an improved description of the tetraloops, mainly in the *syn* glycosidic torsion region of the UNCG tetraloop. The best performance is achieved in combination with the bsc0 parametrization of the $\alpha/\gamma$ angles. Another critically important region to properly describe RNA molecules is the *anti*/high-*anti* region of the glycosidic torsion, where there are significant differences among the tested force fields. The tetraloop simulations are complemented by simulations of short A-RNA stems, which are especially sensitive to an appropriate description of the *anti*/high-*anti* region. While excessive accessibility of the high-*anti* region converts the A-RNA into a senseless "ladder-like" geometry, excessive penalization of the high-*anti* region shifts the simulated structures away from typical A-RNA geometry to structures with a visibly underestimated inclination of base pairs with respect to the helical axis.

## Introduction

RNA is an unbranched, linear polymer composed of four nucleotide units, A, C, G, and U. RNA molecules are usually single-stranded and fold back upon themselves. The 2′-OH group of ribose, absent in DNA, is a powerful donor and acceptor of hydrogen bonds (H-bonds) that is involved in an astonishing repertoire of non-Watson–Crick (noncanonical) interactions. The noncanonical interactions are essential features of RNA three-dimensional structure, dynamics, function, and evolution. Folded RNA molecules typically form short antiparallel double helices by aligning Watson–Crick-complementary stretches of a sequence. These canonical RNA double helices alternate with regions of nucleotides not forming canonical base pairs, i.e., formally unpaired regions. The secondary (2D) structure depicts canonical regions of the folded RNA molecule through the display of parallel lines representing canonical duplex RNA. All of the remaining nucleotides are shown as unpaired loops in such 2D plots. Although these are called loops, these nominally unpaired regions are usually precisely structured via noncanonical interactions and are of the utmost importance for RNA structure and function. The 2D structures of loops can be formally classified as hairpin loops formed by a single-strand segment folded on itself to terminate a helix, internal loops having two strand segments that occur between two helices, and multihelix junctions consisting of multiple-strand segments.[1,2]

The most frequently observed and functionally important hairpin loops are tetraloops (TLs), which cap canonical helices with four loop bases, abbreviated as L1–L4 in this paper. TLs facilitate the backbone inversion required for the formation of secondary and tertiary structures.[3–7] Among all possible combinations,[8] YNMG and GNRA (Y stands for pyrimidine, N for any nucleotide, M for adenine or cytosine, and R for purine), TL families are the most abundant.[5] These TL families are exceptionally thermodynamically stable (namely, when the TL is closed by CG base pairs in the stem),[9] have well-defined structures, and are involved in many biologically relevant processes. In general, TLs initiate folding of RNA structures[3,4,10] and are important interaction sites for tertiary contacts.[11–13]

### UNCG Tetraloop.

The UNCG TLs (a subfamily of the YNMG family) nucleate RNA global folding.[3] This tetraloop displays poor binding to natural ligands except cations and is not involved in RNA/RNA interactions. Experimental structures of this loop display very limited structural variability.[14–17] The most stable of UNCG TLs (UUCG, see Figure 1) has been extensively studied by several authors. Sakata showed that the 2′-OH groups of $U_{L1}$, $C_{L3}$, and $G_{L4}$ and the amino group of $G_{L4}$ are responsible for the thermodynamic stability of the UUCG motif.[18] Later Williams and Hall studied the role of 2′-OH groups of all nucleobases through ribose to 2′-deoxyribose mutations. They concluded that the most significant effect was observed for $U_{L1}$(2′-OH) deletion.[19,20]

The first NMR experiments identified the *trans*-Watson–Crick/sugar-edge (*t*WS)[21] $G_{L4}/U_{L1}$ base pair with the $U_{L1}(O2')\cdots G_{L4}(O6)$ H-bond as a signature interaction of the UUCG TL.[22] The NMR structure further revealed extensive stacking interactions and $C_{L3}(N4)\cdots U_{L2}$(*pro*-$R_p$) base phosphate interaction type 7 (7BPh),[23] which are considered as the main source of the high thermodynamic stability.[17,22] The X-ray structures agreed well in the overall topology of the UUCG TL and unraveled two additional $U_{L2}$-$(O2')\cdots G_{L4}(N7)$ and $C_{L3}(O2')\cdots C_{L3}(O2)$ H-bonds.[14] The latest NMR experiments of Schwalbe et al. gained an ultra-high resolution of 0.25 Å for the loop region (0.3 Å for the stem region).[17] The structure confirmed the *t*WS $G_{L4}/U_{L1}$ base pair with its characteristic $U_{L1}(02')\cdots G_{L4}(O6)$ H-bond and 7BPh interaction between $C_{L3}(N4)$ and $U_{L2}$(*pro*-$R_p$). Sugars of $U_{L2}$ and $C_{L3}$ adopted the C2′-endo puckering in agreement with the X-ray structure.

The stability of UUCG TL was also extensively studied by molecular dynamics (MD). Miller and Kollman[24] observed the destabilization of the $U_{L1}(O2')\cdots G_{L4}(O6)$ H-bond in explicit solvent MD simulations with the AMBER Cornell et al. *ff*94 force field and argued that the $U_{L1}(O2')\cdots G_{L4}(O6)$ interaction cannot be considered as the main source of the exceptional thermodynamic stability of the UUCG TL.[25] However, as we will demonstrate below, the loss of the signature $U_{L1}(O2')\cdots G_{L4}(O6)$ H-bond in simulations was in fact due to the imperfectness of the force field. The exceptional thermodynamic stability and structural features of UUCG TL were also addressed in many recent theoretical studies including replica exchange molecular dynamics and umbrella sampling PMF calculation.[26–30]

### GNRA Tetraloops.

Contrary to UNCG, GNRA TLs primarily mediate RNA tertiary interactions. An analysis of X-ray structures shows striking geometrical conservation also for the GNRA TLs.[8] Structural adaptations in GNRA TL–TL receptor complexes typically includes changes of the TL receptor while the TL is stiff.[31] Williams and co-workers identified 21 examples of standard TLs with the GNRA-like topology in the 2.4 Å resolution X-ray structure of *Haloarcula marismortui* (*H.m.*) large ribosomal subunit.[8] Although they occur in variable contexts within the ribosomal subunit, they adopt virtually identical geometries. The study further identified many hairpin loops with nucleotide insertions, deletions, switches, or strand clips which also adopt very similar 3D structures. Nevertheless, for the GNRA TLs, other experimental methods furnish evidence supporting their conformational dynamics,

although in some cases the flexibility can also reflect error margins and inaccuracies in the experiments (see below). Note that even in the lower-resolution ribosomal X-ray structure data, refinement and noise inaccuracies cannot be ruled out as sources of error. For example, this may introduce *syn/anti* bias[32,33] and perhaps obscure the exact hairpin loop structures in some cases.

The first information about structural features of GNRA TL came from NMR[34] and lower resolution (~3 Å) X-ray structures.[35,36] The high-resolution structures of the sarcin/ricin loop (SRL) domain of the large ribosomal subunit[37–40] together with a structural analysis[8] of the large ribosomal subunit[41] and NMR experiments[42,43] furnished in-depth insight into common features of the native fold of GNRA TLs. They include the *trans* Hoogsteen/sugar-edge (*t*HS) $A_{L4}/G_{L1}$[21] base pair; three signature H-bonds, namely, the $G_{L1}$(N1/N2)•••$A_{L4}$(*pro*-$R_p$) 3BPh interaction, $G_{L1}$(N2)•••$A_{L4}$(N7), and $G_{L1}$(O2′)•••$R_{L3}$(N7) (Figure 2); and stacked $N_{L2}$, $R_{L3}$, and $A_{L4}$ bases. The backbone of GNRA TLs adopts classic U-turn topology. Contrary to UNCG TL, some structural variability of GNRA TLs is anticipated because, for instance, protein ribotoxin restrictocin binds an unfolded GNRA TL.[44] However, the majority of GNRA TLs (~80%) adopt the canonical structure.[37] The question whether ribotoxins induce the conformation change or capture a temporarily unstructured GNRA TL remains open. The dynamics of GNRA TL are the subject of intensive experimental[43,45–49] and theoretical studies.[50–52]

Due to their small size, TLs have been a genuine target for simulation studies.[19,26–30,50,53,54] The simulation studies in general indicate rather substantial flexibility of the TLs, which exceeds variability that is inferred from atomic resolution experiments (see above), suggesting that simulations can be affected by the quality of force fields, typically parametrized keeping in mind the representation of regular helices, not compact irregular structures.[55–57]

In the present study, we investigate the structural dynamics of three representatives (UUCG, GAGA, and GAAA) of the UNCG and GNRA TL families. The aim of the paper is two-fold: first, to get insights into the balance of forces in the TLs and, second, to better understand the performance of molecular mechanics force fields for these difficult systems. The TL simulations are supplemented by simulations of short A-RNA stems. We selected four widely used force fields for nucleic acids: three AMBER (Cornell et al.) force fields, *ff*94,[25] *ff*99,[58] and *ff*99bsc0,[59] and CHARMM27.[60] The *ff*99 and *ff*99bsc0 simulations were also performed at higher ionic strength using excess KCl salt to check the impact of ionic strength on the TL structure and dynamics.[61] Besides using the above established force fields, four recent reparameterizations (two of them from our laboratories) of $\chi$ glycosidic torsion of the AMBER force field are tested. They are combined with *ff*99 and *ff*99bsc0 force fields (see the Methods for details). These $\chi$ modifications were derived recently primarily on the basis of quantum-chemical (QM) computations but were not extensively tested in real simulations. Despite centering on high-level QM calculations, results from the reparameterizations differ considerably as different models and very different levels of QM computations were applied. Thus, in total, the performances of 12 RNA force field variants and combinations were considered (CHARMM27, *ff*94, *ff*99, *ff*99bsc0, and the last two in combination with four $\chi$ modifications).

## Methods

### Starting Structures.

The starting structure of UUCG TL was taken from the high-resolution NMR structure (PDB ID: 2KOC).[17] The GAAA TL was taken from the X-ray structure of a large ribosomal subunit of *Haloarcula maristumortui* (PDB ID, 1JJ2; mean resolution, 2.40 Å residues 800–813).[62] The GAGA TL was derived from the high-resolution X-ray structure determined at 1.04 Å resolution of the sarcin–ricin loop (SRL; PDB ID, 1Q9A; residues 2658–2663)[39] and capped by two additional C/G base pairs. Short A-RNA stems were built using NAB available from the AmberTools package.[63]

### AMBER Simulation Protocol.

We performed classical MD simulations using well established simulation protocols.[57,61,64] Missing hydrogen atoms were added by the LeaP module of the AMBER package on the basis of standard residue templates. Each system was neutralized by $Na^+$ counterions (radius = 1.868 Å and well depth = 0.00277 kcal/mol) and immersed for the MD simulation in a rectangular water box (TIP3P)[65] with a 10-Å-thick layer of water molecules (60 × 50 × 45 $Å^3$ for UUCG and GAAA and 40 × 45 × 50 $Å^3$ for GAGA systems). The RNA–solvent system was minimized prior to the AMBER simulation as follows. Minimization of the solute hydrogen atoms was followed by minimization of counterions and water molecules. Subsequently, the hairpin was frozen, and solvent molecules with counterions were allowed to move during a 10-ps-long MD run, the purpose of which is to relax the density of the system. After that, the nucleobases were allowed to relax in several minimization runs with decreasing force constants applied to the backbone phosphate atoms. After full relaxation, the system was slowly heated to 298.15 K over 100 ps using 2 fs time steps and NpT conditions using a weak-coupling scheme with a coupling time of 1 ps.[66] The simulations were carried out under periodic boundary conditions (PBC) in the NpT ensemble (298.15 K, 1 atm) with 2 fs time steps. The particle-mesh Ewald (PME) method[67,68] was used to calculate electrostatic interactions with a cubic spline interpolation and ~1 Å grid spacing, and a 10.0 Å cutoff was applied for Lennard-Jones interactions with automatic rebuilding of the buffered pair list when atoms moved more than 0.5 Å. The SHAKE algorithm was applied to fix all bonds containing hydrogen atoms. The SANDER module of AMBER 10.0[63] was used for simulations.

### AMBER Force Fields.

Standard AMBER force fields *ff*94,[25] *ff*99,[58] and *ff*99bsc0[59] were used for simulations. In addition, simulations were performed also with four variants of alternative profiles of the glycosidic $\chi$ torsion that were suggested recently as modifications of the *ff*99 force field:

**i.** The Ode et al.[69] $\chi$ parameters are based on quantum chemical profiles obtained with high-accuracy *in vacuo* MP2/aug-cc-pVTZ//HF/6-31+G(d,p) energy calculations on small model compounds. The force field has been suggested to be compatible with both *ff*99 and *ff*99bsc0 basic parametrizations, and the respective simulations are henceforth labeled as *ff*99$\chi_{ODE}$ and *ff*99bsc0$\chi_{ODE}$ in the present paper. Note that this force field has not been tested

in production runs so far except in our recent study on G-DNA quadruplexes, where it was shown to bring no advantage over the *ff*99 and *ff*99bsc0 force fields.

**ii.** Reparameterization against the lower-quality in vacuo QM profile (MP2/6-31G(d)//HF/6-31G(d) level) of ribonucleosides of Yildirim et al.[70] was performed. The force field has not been tested for RNA simulations so far, but it was shown to improve the *syn* vs *anti* balance in nucleoside simulations. Although the original paper does not acknowledge the latest *ff*99bsc0 parametrization and considers the $\chi$ parameters exclusively in the context of *ff*99, we decided to test its performance with both *ff*99 and *ff*99bsc0. The respective simulations are marked as *ff*99$\chi_{\mathrm{YIL}}$ and *ff*99bsc0$\chi_{\mathrm{YIL}}$.

**iii.** Reparameterization based on a high-quality dispersion-corrected[71] DFT QM profile (PBE/6-311++G(3df,3pd)/D-1.06–23//PBE/6-311++G(3df,3pd)/COSMO method) of deoxyribonucleosides in a continuum water environment (this work and Zgarbova et al., manuscript in preparation) labeled as *ff*99$\chi_{\mathrm{OL\text{-}DFT}}$ and *ff*99bsc0$\chi_{\mathrm{OL\text{-}DFT}}$ was performed. (The label OL stands for Olomouc, see affiliations.)

**iv.** Reparameterization based on the high-level QM profile (MP2/CBS//PBE/6-311++G(3df,3pd)/COSMO method) in continuum water considering weighted parameters for C2′-endo deoxyribose and C3′-endo ribose was performed; this variant is labeled *ff*99$\chi_{\mathrm{OL}}$ and *ff*99bsc0$\chi_{\mathrm{OL}}$ (this work and Zgarbova et al., manuscript in preparation).

The OL-DFT and OL parameter files are provided in the Supporting Information, while a full account of the parametrizations including extensive testing will be given separately (Zgarbova et al., manuscript in preparation). The OL force field should be considered as the final version; nevertheless, we also provide some results obtained with the preliminary OL-DFT version, as it provides important insights into the sensitivity of the results to the parametrization.

Note that the modified $\chi$ profiles are entirely independent of the recent *ff*99bsc0 reparameterization of the $\alpha/\gamma$ torsional profile, and therefore the *ff*99bsc0 force field is to be independently cited if applied together with any of the $\chi$ terms. The *ff*99bsc0 is essential, particularly for DNA, in modification of the preceding versions of the AMBER Cornell et al. force fields.

To assess effect of salt concentration on the stability of TLs, reference simulations under KCl salt excess ($c(K^+)$ ~ 0.45 mol/L, $c(Cl^-)$ ~ 0.22 mol/L) conditions and using the SPC/E water model[72] were carried out. Parameters for $K^+$ (radius, 1.593 Å well depth, 0.4297 kcal/mol) and $Cl^-$ (radius, 2.711 Å well depth, 0.012 kcal/mol)[73] were used.

### CHARMM Simulations.

MD simulations of selected systems were also carried out with the CHARMM all27 force field[60] with the NAMD[74] package (ver. 2.6) using the following protocol. To avoid any differences in starting geometries, the neutralized and solvated system prepared for AMBER simulations was used as a starting structure to prepare CHARMM27 topologies and

coordinates in the CHARMM[75] software package (ver. 34b2). The waters and counterions were minimized in 2500 steps and shaken by short NpT dynamics (100 ps) at 300 K and 1 atm. The system was minimized prior to simulation in 3000 steps and then slowly heated to 300 K over 100 ps using 1 fs time steps and NpT conditions using Langevin dynamics.[76,77] The simulation was produced under periodic boundary conditions in the NpT ensemble (300 K, 1 atm) with 1 fs time steps, because the 2 fs integration step produced considerably less stable trajectories for A-RNA stems. The particle-mesh Ewald method was applied to calculate electrostatic interactions (PME tolerance $10^{-6}$), and a 12.0 Å cutoff with an 8.0 Å switching distance was applied for Lennard-Jones interactions. The protocol applied performed well in test simulations on the B-DNA structure, in agreement with literature data.[78]

Table 1 summarizes all simulations analyzed in this study. The simulations were initially intended to be extended to 100 ns. However, some simulations were terminated earlier because of a major degradation of the TLs, i.e., an unfolding event in UUCG_charmm and the formation of a "ladder-like" structure in GAGA_bcs0, GAGA_99$\chi_{ODE}$, and GAGA_bsc0$\chi_{ODE}$ simulations. On the other hand, the simulations carried out with reasonably performing force fields were extended to 300 ns (*ff*99bsc0$\chi_{YIL}$ and *ff*99bsc0$\chi_{OL\text{-}DFT}$) or to 0.8–1.0 $\mu$s (*ff*99bsc0$\chi_{OL}$) to get better insight into the simulation behavior.

Analyses were performed using ptraj (from AmberTools package) and X3DNA.[79] H-bonds were analyzed using in-house software H-bonds (P. Banáš, http://fch.upol.cz/en/software/) using a 3.1 Å cutoff for the H-bond distance and 40° for the hydrogen–H-bond donor•••H-bond acceptor angle.

## Results

### Signature Interactions in the Tetraloops.

As explained in the Introduction, the UNCG and GNRA TLs are very precisely structured recurrent RNA motifs that adopt their native structure independently of their contexts. They therefore possess several characteristic (signature) structural features. For the UUCG TL, these include a *t*WS $G_{L4}/U_{L1}$ base pair, *syn* conformation of $G_{L4}$, and C2′-endo sugar puckers for $U_{L2}$ and $C_{L3}$ (Figure 1). There are four UUCG signature H-bonds (Figure 2): $U_{L1}$(O2′)•••$G_{L4}$(O6), $G_{L4}$-(N1)•••$U_{L1}$(O2), $C_{L3}$(N4)•••$U_{L2}$(*pro*-$R_p$), and $U_{L2}$(O2′)•••$G_{L4}$(N7). The latter one is seen only in approximately one-third of the high-resolution NMR structurally derived ensembles.[17] $U_{L1}$(O2) tends to form a bifurcated H-bond to $G_{L4}$(N1) and $G_{L4}$(N2) in some X-ray structures.[14–16] On the other hand, the distance between $G_{L4}$(N2) and $U_{L1}$(O2) is always larger than 3.3 Å in the high-resolution NMR structure[17] (see also Table 2). The $C_{L3}$(N4)•••$U_{L2}$(*pro*-$R_p$) H-bond corresponds to a type 7 base–phosphate interaction (7BPh[23]) between the $C_{L3}$ base and $U_{L2}$ phosphate.

The GNRA TLs include the *t*HS $A_{L4}/G_{L1}$ ("sheared") base pair[37] complemented by three H-bonds (Figure 2): $G_{L1}$(N2)•••$A_{L4}$(*pro*-$R_p$) (3BPh interaction, which is altered with the $G_{L1}$(N1/N2)•••$A_{L4}$(*pro*-$R_p$) 4BPh interaction in MD or some X-ray structures),

$G_{L1}$(N2)•••$A_{L4}$(N7), and $G_{L1}$(O2′)•••$R_{L3}$(N7).[8] The $N_{L2}$, $R_{L3}$, and $A_{L4}$ bases form a purine triple stack.

## UUCG Tetraloop Dynamics.

**AMBER Simulations: $\chi$ Reparameterizations Maintain Important Signature H-Bonds and Overall Integrity of the Tetraloop.—**In all UUCG TL simulations with the standard AMBER force fields (with standard $\chi$ torsion, i.e, UUCG_94, UUCG_99, UUCG_bsc0, UUCG_99SE, and UUCG_bsc0SE simulations), we observed a loss of the signature $U_{L1}$(O2′)•••$G_{L4}$(O6) H-bond immediately after the simulation started. This H-bond was replaced by the $U_{L1}$(O2′)•••$U_{L2}$(O5′) H-bond (Figure 3A). Despite the fact that the loops still stay, at first sight, locked close to the starting structure, these changes are clear signs of some force field imbalance. Considering the unambiguous structural data, this simulation development is not satisfactory.

Simultaneously with the disruption of the $U_{L1}$(O2′)•••$G_{L4}$(O6) H-bond, we observed significant propeller twisting (changing from −2° to ~−25°) of the $G_{L4}$/$U_{L1}$ *t*WS base pair and mainly a shift of *syn* $G_{L4}$ $\chi$ torsion from 60° to 40° in all above-mentioned simulations (Table 2). The same shifts of $\chi$ torsion toward lower values in the *syn* region were also observed for *syn* G+1 and A38H$^+$ nucleobases in *ff*99 MD simulations of the hairpin ribozyme (Supporting Information, Table S1).[80] This structural shift comes from the strain in $G_{L4}$ $\chi$ torsion, as the energy profile of guanosine $\chi$ torsion in AMBER *ff*99 (Supporting Information, Figure S1) shows that the minimum in the *syn* region equals 40°. On the basis of a structural analysis of the above-mentioned MD simulations, we suggest that the shift of $G_{L4}$ $\chi$ torsion is the primary source of perturbation of the signature interaction.

All simulations with reparameterized $\chi$ torsion prevented the shift of $G_{L4}$ $\chi$ torsion, and the signature $U_{L1}$-(O2′)•••$G_{L4}$(O6) H-bond was stable, except in some cases where it was disrupted due to some unrelated perturbations elsewhere in the structure (Table 2). Thus, UUCG_99$\chi_{ODE}$, UUCG_99$\chi_{OL}$, UUCG_99$\chi_{YIL}$, UUCG_bsc0$\chi_{OL\text{-}DFT}$, UUCG_bsc0$\chi_{OL}$, and UUCG_bsc0$\chi_{YIL}$ were the most stable trajectories keeping all signature H-bonds (Table 2). In other words, stable UUCG TL was observed in all simulations with modified $\chi$ torsion parameters, except for UUCG_99$\chi_{OL\text{-}DFT}$, showing an undesired $\alpha/\gamma$ flip of the $U_{L1}$ phosphate (see below) and UUCG_bsc0$\chi_{ODE}$ where a “ladder-like” artifact of $C_{S\text{-}1}$ and $U_{L1}$ $\chi$ torsions, described below, occurred. $C_{S\text{-}1}$ denotes stem cytosine at the 5′ side of the TL, i.e., C5 in the presented model of UUCG (Figure 1). This indicates that the modified $\chi$ torsion profiles locally improve sampling within the *syn* region.

The formation of the $U_{L1}$(O2′)•••$U_{L2}$(O5′) H-bond in simulations with an unmodified $\chi$ profile was likely partially facilitated by modest shifts of $\varepsilon$ and $\zeta$ torsions of $C_{S\text{-}1}$ ($\varepsilon$ from −126° to ~−150° and $\zeta$ from −80° to ~−60°) and $U_{L1}$ ($\varepsilon$ from −160° to ~−175° and $\zeta$ from −100° to ~−90°). This backbone adaptation occurred in entirely all AMBER UUCG simulations in the initial minimization and was irreversible. This shift of $\varepsilon$ and $\zeta$ torsions moved the $U_{L2}$(05′) oxygen closer to the $U_{L1}$(2′-OH) hydroxyl (from 3.4 Å in the NMR structure to ~3.0 Å in MD simulations), which supported the formation of the new $U_{L1}$(O2′)•••$U_{L2}$(05′) H-bond. However, we do not consider this backbone adaptation to be

the most crucial force field problem, since in the simulations with reparametrized $\chi$ torsions the original $U_{L1}(O2')\bullet\bullet\bullet G_{L4}(O6)$ H-bond remains stable despite the $\varepsilon/\zeta$ shift.

**ff99bsc0 Clearly Improves the Stability of γ Angle Distribution.—**In the advanced stages of simulations (on the tens of nanoseconds time scale), we evidenced further problems due to a disruption of the $C_{L3}(N4)\bullet\bullet\bullet U_{L2}$(*pro*-$R_p$) 7BPh interaction in UUCG_94, UUCG_99, and UUCG_99$\chi_{OL\text{-}DFT}$ trajectories. This was caused by $\alpha/\gamma$ flip of the $U_{L2}$ phosphate (Figure 3B). The flip involved a shift of $\alpha(U_{L2})$ from ~−160° to ~−50°, $\varepsilon(U_{L1}$ from ~−170° to ~−100°, and $\gamma(U_{L2})$ from ~50° to ~−170°. It further correlated with C3′-endo to C2′-endo $U_{L2}$ sugar repuckering. In all cases, the $U_{L2}$ phosphate remained distorted until the simulation ended. The $\gamma$ torsion of the $U_{L2}$ phosphate sampled the gauche(+) region in all *ff*99bsc0 simulations except for UUCG_bsc0$\chi_{ODE}$ and UUCG_bsc0$\chi_{OL}$, where we also observed weakly populated (population about ~5%) and fully reversible $\gamma$-*trans* substates. Thus, in contrast to *ff*99 simlations, the bsc0 correction prevents an irreversible $\alpha/\gamma$ flip of the $U_{L2}$ phosphate to $\gamma$-*trans.* The *ff*99bsc0 force field has been designed to prevent pathological $\gamma$-*trans* substates in B-DNA MD simulations.[59] It is worth noting that the native position of $\gamma(G_{L4})$ is *trans* due to a sharp bend of the RNA strand at the tip of UUCG TL. Interestingly, $\gamma(G_{L4})$ kept its native $\gamma$-*trans* orientation in all simulations with *ff*99bsc0, despite some expectations that *ff*99bsc0 may occasionally overcorrect the $\gamma$-*trans* substates.[61,64] Clearly, at least when starting simulations from the native structure, *ff*99bsc0 is superior to *ff*99 for the UUCG TL, as it prevents one undesired irreversible $\gamma$-trans flip while keeping the native $\gamma$-*trans* nucleotide stable.

**The Occurrence of High-anti Substates in Correlation with the Force Field Artifact of Forming a “Ladder-Like” Structure.—**An almost reversible disruption of the TL signature accompanied by a shift of $\chi$ torsions of $C_{S\text{-}1}$ and $U_{L1}$ from the *anti* to the high-*anti* region (from ~−150° to ~−90°) and breaking of the $G_{L4}(N1)\bullet\bullet\bullet U_{L1}(O2)$ and $U_{L1}$-$(O2')\bullet\bullet\bullet G_{L4}(O6)$ H-bonds was observed in UUCG_bsc0 and UUCG_bsC0$\chi_{ODE}$ simulations (Figure 4). The shift of $\chi$ torsions to the high-*anti* region corresponds to a recently discovered common force field artifact named a “ladder-like” structure of RNA stems, because the most characteristic feature of the “ladder-like” structure is a transition of $\chi$ torsion to the high-*anti* region with a value ~−85° (the exact value slightly depends on the system and force field).[80] In a fully developed “ladder-like” structure of a duplex, besides the shift of the $\chi$ torsion, the sugar puckering, $\varepsilon$ and $\zeta$ torsions, slide, twist, and peaks in the P–P radial distribution function are also modestly affected by the transition, which in addition is not reversible (see ref 80 for more details). In the present simulations, although the $C_{S\text{-}1}$ and $U_{L1}$ $\chi$ torsions later returned to the *anti* region, the signature $U_{L1}$-$(O2')\bullet\bullet\bullet G_{L4}(O6)$ H-bond was not fully stabilized and experienced fluctuations. As will be discussed below, the unmodified *ff*99 and $\chi_{ODE}$[69] parametrizations support the formation of the “ladder-like” artifact, while the remaining three $\chi$ reparameterizations appear to prevent its formation. It is entirely consistent with the behavior of UUCG simulations. The independent *ff*99bsc0 modification of $\alpha/\gamma$ dihedrals is neutral with respect to the “ladder-like” structure formation.

### CHARMM Simulations.

The MD simulation of UUCG TL carried out with a CHARMM27 force field showed complete melting during the first 10 ns. This is in full agreement with recently published simulation data from Deng and Cieplak.[28] The signature $U_{L1}$(O2′)•••$G_{L4}$(O6) and $G_{L4}$(N1)•••$U_{L1}$(O2) H-bonds were broken at ~0.5 ns, and $G_{L4}$ departed from its initial position. The $C_{L3}$ nucleobase unstacked from $U_{L1}$ at9.5ns, breaking its 7BPh$C_{L3}$(N4)•••$U_{L2}$(*pro*-$R_p$) H-bond. The ribose pucker of $C_{L3}$ switched from C2′-endo to C3′-endo at 20 ns, and $G_{L4}$ switched from a *syn* to an *anti* orientation at ~40 ns. Stem base pairs were also not stable and exhibited frequent breathing.

### GNRA Tetraloop Dynamics.

In simulations of GNRA TLs with the standard $\chi$ AMBER force fields, we typically observed a transition of the A-RNA stem bearing the GNRA TL to the underwound “ladder-like” structure (Figure 5). The “ladder-like” structure is a force field artifact occurring on the tens of nanoseconds time scale, which we first described in our study on hairpin ribozymes.[80] This transition was always preceded by an irreversible disruption of the GNRA signature. The “ladder-like” structure of the stem bearing GNRA TL occurred with standard AMBER force fields as well as with the $\chi$ torsion reparameterization of Ode et al.[69] (see Table 1). On the other hand, it has been prevented by the other three $\chi$ torsion reparameterizations ($\chi_{OL\text{-}DFT}$, $\chi_{OL}$, and $\chi_{YIL}$) and was also not observed in CHARMM simulations.

Although the formation of the “ladder-like” structure was shown primarily to be the force field artifact of A-RNA stems[80] (also see below), a detailed structural analysis, mainly the monitoring of $\chi$ torsions, revealed that in our present case the formation of a “ladder-like” structure in the stem was typically preceded by the structural degradation of GNRA TL. Thus, it appears that the loss of the TL integrity facilitates the “ladder-like” transition. Both issues are, however, most likely interconnected. More specifically, the loss of structural integrity of GNRA TL was mainly facilitated by a reversible flip of the $\beta/\gamma$ torsions of the $R_{L3}$ phosphate from the *trans/gauche*(+) to the *gauche*(+)/*trans* conformation (Figure 6). This $\beta/\gamma$ flip increased the buckle of the *t*HS $A_{L4}$/$G_{L1}$ base pair from ~5° to ~60° (Figure 6). Although the $\beta/\gamma$ flips are reversible, we suggest that the increased $A_{L4}$/$G_{L1}$ buckle causes some steric strain of the stem–loop junction and accelerates the structural degradations of the system including a transition of the stem to a “ladder-like” structure (when the force field does not prevent this “ladder-like” artifact).

As noted above, we did not observe any formation of a “ladder-like” structure in the simulations with $\chi_{OL\text{-}DFT}$, $\chi_{OL}$ and $\chi_{YIL}$ reparameterizations. However, simulations with *ff*99 (combining *ff*99 with either the $\chi_{OL\text{-}DFT}$, $\chi_{OL}$, or $\chi_{YIL}$ parametrization) exhibited serious distortion of GNRA TL caused by either the $R_{L3}$ $\beta/\gamma$ flip or less often by a flip of the α($A_{L2}$) torsion from *trans* to *gauche*, usually accompanied by the shift of $\gamma$($A_{L2}$) torsion from *gauche*(+) to *trans*. Furthermore, these flips caused disruption of GNRA signatures in four of the six *ff*99 simulations with $\chi_{OL\text{-}DFT}$, $\chi_{OL}$, and $\chi_{YIL}$ (GAGA_99$\chi_{YIL}$, GAGA_99$\chi_{OL}$, GAAA_99$\chi_{YIL}$, and GAAA_99$\chi_{OL\text{-}DFT}$). On the other hand, the simulations combining *ff*99bsc0 with $\chi_{OL\text{-}DFT}$, $\chi_{OL}$, or $\chi_{YIL}$ parametrizations

exhibited stable behavior of GNRA TL on the hundreds of nanoseconds time scale. It should be noted that both $R_{L3}$ $\beta/\gamma$ and $A_{L2}$ $\alpha/\gamma$ flips were still present in simulations with the bsc0 correction; however, these flips were reversible and short-lived and thus did not result in distortion of GNRA TL.

The $R_{L3}$ $\beta/\gamma$ flip, which seems to be the main source of GNRA TL destabilization in AMBER simulations, might be a consequence of imperfect force field parameters of $\beta$ and $\gamma$ torsions. Nonetheless, the $R_{L3}$ phosphate undergoing the $\beta/\gamma$ flip is positioned in proximity to the $A_{L2}$ phosphate (P–P distance in the X-ray structure is 5.9 and 5.8 Å in GAGA and GAAA, respectively) because of sharp inversion of the sugar–phosphate backbone path. Thus, the $R_{L3}$ $\beta/\gamma$ flip might also be alternatively caused by insufficiently compensated electrostatic repulsion between these two phosphates or some other imbalance of the intermolecular terms. This is in agreement with the fact that structural degradation is initiated by the flip of either $R_{L3}$ or $A_{L2}$ phosphate. However, the involvement of *ff*99bsc0 correction significantly attenuates these phosphate flips, although it is not able to completely eliminate them. Thus, it seems that both imperfect $\alpha/\beta/\gamma$ torsion parameters and an imbalance of the intermolecular terms can contribute to structural degradation of GNRA TL in the *ff*99 force field. Nevertheless, the present GNRA TL simulations are substantially improved when combining the bsc0 correction together with the $\chi_{OL\text{-}DFT}$, $\chi_{OL}$, or $\chi_{YIL}$ modification.

We extended the *ff*99bsc0$\chi_{OL}$ force field simulations of GAAA and GAGA systems to 0.8 and 1.0 $\mu$s, respectively, to test the performance of this force field on the microsecond time scale. We found that the GAGA_bsc0$\chi_{OL}$ simulation was entirely stable on the microsecond time scale. However, we observed conformational changes of the TL region in the GAAA_bsc0$\chi_{OL}$ simulation after 0.56 $\mu$s. The deformation of GAAA TL is related neither to the flip of the $R_{L3}$ and $A_{L2}$ phosphates nor to the formation of a “ladder-like” structure. It may be caused by some other force field imbalances which become visible on the microsecond time scale (see Supporting Information, Figure S2).

Neither the formation of a “ladder-like” conformation nor a $R_{L3}$ $\beta/\gamma$ flip was observed in simulations with CHARMM27. However, we observed local switches of $A_{L2}$ and $R_{L3}$ phosphates including a rapid fluctuation of $\alpha(A_{L2})$ between native *trans* and *gauche*(+), $\alpha(R_{L3})$ between *trans* and native *gauche*(+), and rapid switches of $\varepsilon$ and $\zeta$ in all four TL nucleobases that were accompanied by structural distortion of the GNRA TL in both CHARMM27 simulations. Nonetheless, the most distinctive feature of CHARMM27 simulations was the instability of the stem bearing the GNRA TL that exhibited extensive terminal base pair breathing followed by the disruption of base-pairing in the stem and subsequent unfolding of the structure, similar to what has been reported, for example, for stems in simulations of kissing-loop complexes.[81,82]

### Simulations of A-RNA Stems.

The Supporting Information describes a set of simulations of short canonical A-RNA stems. These simulations illustrate rather visible differences between the modified $\chi$ parametrizations in the description of the canonical A-RNA structure, mainly a different inclination of base pairs with respect to the helical axis. Although not related directly to the

main topic of this paper, these A-RNA simulations provide further insight into the sensitivity of A-RNA simulations to force field parameters and help understanding of the simulation behavior of the TLs.

## Discussion

The 5′-UNCG-3′ and 5′-GNRA-3′ RNA tetraloops (TL) are the two most important classes of RNA hairpin loops. These thermodynamically very stable TLs belong to the most prominent RNA motifs, i.e., recurrent RNA building blocks with a precisely defined context-independent 3D structure. While the UNCG TLs play a key role in RNA folding, the GNRA TLs are involved in tertiary interactions and recognition processes.

As with each RNA motif, the UNCG and GNRA TLs are characterized by signature molecular interactions which define their native structure and, subsequently, following the isostericity principle, their consensus sequences.[33] The 3D signature of the studied UUCG TL includes the *t*WS $G_{L4}/U_{L1}$ base pair, *syn* conformation of $G_{L4}$, south C2′-endo pucker of $U_{L2}$ and $C_{L3}$, and four UUCG signature H-bonds: $U_{L1}(O2')\cdots G_{L4}(O6)$, $G_{L4}(N1)\cdots U_{L1}(O2)$, $C_{L3}(N4)\cdots U_{L2}(pro\text{-}R_p)$, and $U_{L2}(O2')\cdots G_{L4}(N7)$. The $U_{L1}(O2')\cdots G_{L4}(O6)$ and $C_{L3}(N4)\cdots U_{L2}(pro\text{-}R_p)$ H-bonds are unambiguous. The $U_{L2}(O2')\cdots G_{L4}(N7)$ H-bond is seen only in approximately one-third of the high-resolution NMR structure ensemble.[17] The GNRA TLs are structured with a *t*HS $A_{L4}/G_{L1}$ (“sheared”) base pair[37] complemented by three H-bonds: $G_{L1}(N1/N2)\cdots A_{L4}(pro\text{-}R_p)$, $G_{L1}(N2)\cdots A_{L4}(N7)$, and $G_{L1}(O2')\cdots R_{L3}(N7)$.[8] The GNRA signature further includes a $N_{L2}$, $R_{L3}$, and $A_{L4}$ triple base stack.

Although it cannot be ruled out that the TLs (especially the GNRA one) exhibit some structural dynamics or can be remodeled in some structural contexts, structural biology data as well structural bioinformatics convincingly show that the above-described signature interactions define the genuine native structures of these RNA TL classes.[8,37] Therefore, correct computational methods should be capable of reproducing the characteristic structures of UNCG and GNRA RNA TLs, identifying them as global minima, and dominantly sampling them. However, as noted in the literature, a correct force field description of nucleic acid hairpin loops may be a considerable challenge for the contemporary molecular mechanical force fields.[57] Hairpin loops are characterized by a complex mixture of different molecular interactions and noncanonical backbone conformations and are substantially exposed to the solvent.

The RNA TLs, due to their small size and biochemical importance, became a favorable target for simulation studies in the past several years. These studies were primarily concentrated on the folding of the TLs, using sophisticated enhanced sampling methods and massive large-scale parallel computations.[28,29,50] These impressive studies clearly demonstrated the basic capability of the simulation technique to correctly identify the stem base pairing and subsequently fold the structure. However, less attention has been paid to the exactness of the final or most stable structures identified as the native states. Closer inspection of the published data reveals that at least in some cases the predicted topology is not fully consistent with the native topology as known from structural biology.

In the present paper, we have considered a less ambitious but perhaps no less important task. We investigate the capability of the established force fields to keep the native topology of the UNCG and GNRA TLs. We analyze typical structural rearrangements seen on the ~100+ ns time scale and their force field dependence. Simultaneously, we use the TLs as model systems to test four recent attempts (two of them from our laboratories) to adjust the $\chi$ glycosidic torsion profile of the Cornell et al. force field, in addition to the basic *ff*99 and *ff*99bsc0 force field variants. The *ff*99bsc0 is the only viable AMBER force field for DNA simulations, while RNA was until now considered to be almost equivalently well described by all basic Cornell et al. force field variants.[57,59] The $\chi$ modifications were derived on the basis of QM computations using different model systems, different QM levels, and different overall philosophies of parametrization (see the Introduction and Methods and the Supporting Information for parameters). There are considerable differences among the four suggested $\chi$ glycosidic torsion profiles, while they also substantially differ from the original parametrization (supplementary Figure S1B, Supporting Information). We also performed a set of simulations on short A-RNA stems (Supporting Information). Although the A-RNA simulations are not directly related to the main topic of this paper, these A-RNA simulations provide insights into the sensitivity of A-RNA simulations to force field parameters and help with understanding the simulation behavior of the TLs. The A-RNA simulations indicate that adjusting the $\chi$ torsion has a visible effect on the calculated inclination and base pair roll of A-RNA helices (see Supporting Information, Table S3). The global (helical) structure parameter inclination and local (wedge) parameter roll are mathematically interconnected. They characterize the degree to which the base pairs in the helix adopt the A-RNA geometry having the base pairs inclined with respect to the global helical axis. Due to helical twisting, the inclination then leads to base pair roll in the local base pair step coordination frames.[83,84] The $\chi_{\text{OL-DFT}}$ and $\chi_{\text{OL}}$ variants of $\chi$ adjustment modestly reduce the inclination/roll values compared with simulations using the unmodified force field (see Supporting Information, Table S3). The $\chi_{\text{YIL}}$ adjustment leads to a qualitative reduction (and underestimation) of inclination/roll (see Supporting Information, Table S3). The impact of these effects on RNA simulations is under further investigation. The $\chi_{\text{ODE}}$ parametrization destabilizes the A-RNA by promoting the “ladder-like” structure. This behavior reflects the balance of *anti* and high-*anti* regions of the respective parametrizations.

For the tetraloops, we have obtained the following results. In all UUCG TL simulations with the standard AMBER force fields (with unmodified $\chi$ torsion), we observed a loss of the signature $U_{L1}(O2')\cdots G_{L4}(O6)$ H-bond immediately after the simulation start, i.e., within few picoseconds. This H-bond is replaced by the $U_{L1}(O2')\cdots U_{L2}(O5')$ H-bond (Figure 3A). Despite the fact that the loop subsequently remains close to the starting structure, the loss of the signature interaction is not in agreement with structural data. Simultaneously with the disruption of the $U_{L1}(O2')\cdots G_{L4}(O6)$ H-bond, we observed significant propeller twisting of the $G_{L4}/U_{L1}$ *t*WS base pair and mainly a shift of $G_{L4}$ $\chi$ *syn* torsion from 60° to 40°. We argued that the $G_{L4}$ $\chi$ torsion shift is the primary source of perturbation of the signature interaction. All studied modifications of $\chi$ torsions (i.e., $\chi_{\text{ODE}}$, $\chi_{\text{YIL}}$, $\chi_{\text{OL-DFT}}$, and $\chi_{\text{OL}}$) improve the behavior most likely because they provide a more realistic description of the *syn* region of G (Supporting Information, Figure S1B).

Further analysis indicates that the use of bsc0 correction improves the simulation behavior by stabilizing the observed distribution of the $\gamma$ backbone angles, mainly by preventing the undesired and irreversible $\gamma$-*trans* flip of the $U_{L2}$ phosphate. Interestingly, the native $\gamma$-*trans* flip of $G_{L4}$ is kept. Therefore, this TL is best described when using the *ff*99bsc0 basic parametrization with some of the $\chi$ torsion adjustments.

The most significant feature of GNRA simulations with the standard variants of the AMBER force field is a loss of the GNRA integrity on a scale of dozens of nanoseconds followed by a subsequent "ladder-like" conversion of the whole helical stem (Figure 6). Adding the suggested $\chi$ corrections (except of $\chi_{ODE}$) improves the behavior of the GNRA TL simulations and prevents larger degradations on the ~100 ns time scale, which is the typical time scale for presently published RNA simulations. The most likely reason why the three successful $\chi$ corrections improve the GNRA simulation behavior is the change of the profile in the high-*anti* region compared to that in the *anti* region (Supporting Information, Figure S1B). Compared with the basic *ff*99/*ff*99bsc0 parametrizations, $\chi_{OL\text{-}DFT}$ and $\chi_{OL}$ bring a modest penalty to the high-*anti* region, which however seems to be enough to prevent the forming of a "ladder-like" structure. The $\chi_{YIL}$ works in the same direction, but the high-*anti* penalty is much more vigorous. The $\chi_{ODE}$ rather supports the high-*anti* $\chi$ region, and that is why it does not prevent the ladder-like artifact. Note that despite the overall improvement the GNRA simulations exhibit some local dynamics which may indicate some more subtle imbalances. This will require further studies. We noticed reversible flips of $\beta/\gamma$ torsions of the $R_{L3}$ phosphate from the *trans/gauche*(+) to the *gauche*(+)/*trans* conformation (Figure 6) which are associated with a dramatically increased buckle of the $G_{L1}/A_{L4}$ base pair from ~5° to ~60°. Such base pair distortion may accelerate further undesired rearrangements of the TLs. The use of *ff*99bsc0 correction improves the simulation behavior by stabilizing the native conformation of TLs phosphates but does not completely prevent these flips. We need to keep in mind that some of the observed dynamical effects may be related to imbalances of the intermolecular terms of the force field. In such a case, the ability of the torsion angle adjustments to improve the simulations may be limited. As noted above, we suspect that the lack of a fully balanced description of the interphosphate repulsion may contribute to the observed backbone dynamics. A balanced description of some such effects may thus require the development of polarization force fields or at least some reparameterization of the solvation terms. Taken together, the GNRA TL is best described by the *ff*99bsc0 basic parametrization with $\chi_{YIL}$, $\chi_{OL\text{-}DFT}$, or $\chi_{OL}$ adjustments.

In summary, our data show that three of the $\chi$ glycosidic torsion profiles, namely, $\chi_{OL\text{-}DFT}$, $\chi_{OL}$, and $\chi_{YIL}$, improve the description of the TLs, especially when combined with the *ff*99bsc0 basic parametrization.[59] Mainly, they prevent the formation of the degrading "ladder-like" structures of RNA stems, which break down the simulated GNRA tetraloops. The "ladder-like" conformation is associated with an excessive high-*anti* shift of the $\chi$ torsion.[80] On the other hand, the $\chi$ glycosidic torsion reparameterization of Ode et al.[69] is much less suitable for RNA simulations, as it accelerates the formation of the ladder-like structures. All four $\chi$ modifications locally improve the description of the *syn* region, which stabilizes the UUCG TL simulations. Again, bsc0 is to be used as the basic force field for the UUCG simulation. The present results, however, should be taken as preliminary, and considerably more extensive tests on numerous different RNA and DNA systems are

under way. We would like to stress that although some of the $\chi$ torsion adjustments show significant potential for improving RNA simulations, mainly by preventing the “ladder-like” structure, it cannot be ruled out that we will in the future identify also a worsening of some other properties of the simulated molecules.

When assessing the significance of the results, we have to make a few cautionary notes. First, the $\chi$ torsion reparameterizations are applicable exclusively to RNA. We have tested them (not shown) also for B-DNA and DNA quadruplex loops, and they do not improve DNA simulations. In fact, it appears that fine-tuning of the $\chi$ torsion simultaneously for DNA and RNA is not possible, unless some other parameters are modified too. This is related to the second cautionary comment. When using simple analytical force fields, the description of the simulated system is unavoidably only approximate, despite careful parametrization procedures. Thus, the force field is inescapably physically inexact and incomplete. Therefore, the impact of adjustments of the individual torsions, although potentially improving the simulation performance, should not be overrated. The physically inexact torsional potentials are used to approximate the effective overall sum of many diverse physical contributions. The QM reparameterization of torsions does not per se guarantee that the simulations are subsequently improved, as the force field performance depends on the overall balance of all of the force field terms. Therefore, before any application of a modified force field, it must be carefully tested for relevant nucleic acid systems. It is therefore a rather unusual practice that the $\chi_{ODE}$ and $\chi_{YIL}$ parametrizations were made available without any testing.[69,70] In addition, it is well-known that improving one feature of the simulated systems may have undesired side effects elsewhere. This explains why the $\chi$ torsion adjustments tested here do not improve the behavior of DNA simulations.

For the sake of completeness, we also performed limited simulations using the CHARMM27 force field. The main dynamics that we noticed in GNRA TLs are phosphate flips and fluctuations, similar to those reported above for the AMBER simulations. Nevertheless, the simulated structures were later destabilized in their stem regions (base pair fluctuations and fraying), which ultimately also affected the TLs. Note that the simulations were carried out with quite short stems. Such reduced stability of the short stems is consistent with literature data.[81] The UNCG TL trajectory was unstable. This simulation result is identical to more extensive data reported and in more detail described by Deng and Cieplak.[28]

The 100+ ns simulations are sufficient for many useful applications, such as the very basic MD characterization of existing RNA structures. Note that perturbation of the TLs may cause bias in the overall assessment of the data even when the TL is not the primary focus of a given simulation study. Therefore, stabilization of the RNA TLs on this time scale is important. Work is in progress to investigate the TLs using much longer simulations and also using substantially larger RNA systems to prevent eventual end effects.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgment.

This study was supported by Grants CZ.1.05/2.1.00/03.0058, LC512, LC06030, and MSM6198959216 from the Ministry of Education of the Czech Republic; Grants 203/09/1476 and 203/09/H046 from the Grant Agency of the Czech Republic; Grant IAA400040802 from the Grant Agency of the Academy of Sciences of the Czech Republic; Grants AV0Z50040507 and AV0Z50040702 from the Academy of Sciences of the Czech Republic; and Student Project PrF_2010_025 of Palacký University.

## References

(1). Leontis NB; Westhof E Analysis of RNA motifs. Cur. Opin. Struct. Biol 2003, 13, 300–308.

(2). Mathews DH; Turner DH Prediction of RNA secondary structure by free energy minimization. Cur. Opin. Struct. Biol 2006, 16, 270–278.

(3). Tuerk C; Gauss P; Thermes C; Groebe DR; Gayle M; Guild N; Stormo G; Daubentoncarafa Y; Uhlenbeck OC; Tinoco I; Brody EN; Gold L CUUCGG Hairpins - Extraordinarily Stable RNA Secondary Structures Associated with Various Biochemical Processes. Proc. Natl. Acad. Sci 1988, 85, 1364–1368. [PubMed: 2449689]

(4). Uhlenbeck OC Nucleic-Acid Structure - Tetraloops and RNA Folding. Nature 1990, 346, 613–614. [PubMed: 1696683]

(5). Woese CR; Winker S; Gutell RR Architecture of Ribosomal-RNA - Constraints on the Sequence of Tetra-Loops. Proc. Natl. Acad. Sci 1990, 87, 8467–8471. [PubMed: 2236056]

(6). Wolters J The Nature of Preferred Hairpin Structures in 16s-Like Ribosomal-RNA Variable Regions. Nucleic Acids Res. 1992, 20, 1843–1850. [PubMed: 1374559]

(7). Bevilacqua PC; Blose JM Structures, kinetics, thermo-dynamics, and biological functions of RNA hairpins. Annu. Rev. Phys. Chem 2008, 59, 79–103. [PubMed: 17937599]

(8). Hsiao C; Mohan S; Hershkovitz E; Tannenbaum A; Williams LD Single nucleotide RNA choreography. Nucleic Acids Res. 2006, 34, 1481–1491. [PubMed: 16531589]

(9). Blose JM; Proctor DJ; Veeraraghavan N; Misra VK; Bevilacqua PC Contribution of the Closing Base Pair to Exceptional Stability in RNA Tetraloops: Roles for Molecular Mimicry and Electrostatic Factors. J. Am. Chem. Soc 2009, 131, 8474–8484. [PubMed: 19476351]

(10). Varani G Exceptionally Stable Nucleic-Acid Hairpins. Annu. Rev. Biophys. Biomed 1995, 24, 379–404.

(11). Marino JP; Gregorian RS; Csankovszki G; Crothers DM Bent Helix Formation between RNA Hairpins with Complementary Loops. Science 1995, 268, 1448–1454. [PubMed: 7539549]

(12). Chauhan S; Woodson SA Tertiary interactions determine the accuracy of RNA folding. J. Am. Chem. Soc 2008, 130, 1296–1303. [PubMed: 18179212]

(13). Jaeger L; Michel F; Westhof E Involvement of a GNRA Tetraloop in Long-Range Tertiary Interactions. J. Mol. Biol 1994, 236, 1271–1276. [PubMed: 7510342]

(14). Ennifar E; Nikulin A; Tishchenko S; Serganov A; Nevskaya N; Garber M; Ehresmann B; Ehresmann C; Nikonov S; Dumas P The crystal structure of UUCG tetraloop. J. Mol. Biol 2000, 304, 35–42. [PubMed: 11071808]

(15). Tishchenko S; Nikulin A; Fomenkova N; Nevskaya N; Nikonov O; Dumas P; Moine H; Ehresmann B; Ehresmann C; Piendl W; Lamzin V; Garber M; Nikonov S Detailed analysis of RNA-protein interactions within the ribosomal protein S8-rRNA complex from the archaeon Methanococcus jannaschii. J. Mol. Biol 2001, 311, 311–324. [PubMed: 11478863]

(16). Carter AP; Clemons WM; Brodersen DE; Morgan-Warren RJ; Wimberly BT; Ramakrishnan V Functional insights from the structure of the 30S ribosomal subunit and its interactions with antibiotics. Nature 2000, 407, 340–348. [PubMed: 11014183]

(17). Nozinovic S; Furtig B; Jonker HRA; 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, 683–694. [PubMed: 19906714]

(18). Sakata T; Hiroaki H; Oda Y; Tanaka T; Ikehara M; Uesugi S Studies on the Structure and Stabilizing Factor of the CUUCGG Hairpin Rna Using Chemically Synthesized Oligonucleotides. Nucleic Acids Res. 1990, 18, 3831–3839. [PubMed: 1695732]

(19). Williams DJ; Hall KB Unrestrained stochastic dynamics simulations of the UUCG tetraloop using an implicit solvation model. Biophys. J 1999, 76, 3192–3205. [PubMed: 10354444]

(20). Williams DJ; Boots JL; Hall KB Thermodynamics of 2′-ribose substitutions in UUCG tetraloops. RNA 2001, 7, 44–53. [PubMed: 11214179]

(21). Leontis NB; Westhof E Geometric nomenclature and classification of RNA base pairs. RNA 2001, 7, 499–512. [PubMed: 11345429]

(22). Allain FHT; Varani G Structure of the P1 Helix from Group-I Self-Splicing Introns. J. Mol. Biol 1995, 250, 333–353. [PubMed: 7608979]

(23). Zirbel CL; Sponer JE; Sponer J; Stombaugh J; Leontis NB Classification and energetics of the base-phosphate interactions in RNA. Nucleic Acids Res. 2009, 37, 4898–4918. [PubMed: 19528080]

(24). Miller JL; Kollman PA Theoretical studies of an exceptionally stable RNA tetraloop: Observation of convergence from an incorrect NMR structure to the correct one using unrestrained molecular dynamics. J. Mol. Biol 1997, 270, 436–450. [PubMed: 9237909]

(25). Cornell WD; Cieplak P; Bayly CI; Gould IR; Merz KM; Ferguson DM; Spellmeyer DC; Fox T; Caldwell JW; Kollman PA A second generation force field for the simulation of proteins, nucleic acids, and organic molecules. J. Am. Chem. Soc 1995, 117, 5179–5197.

(26). Villa A; Widjajakusuma E; Stock G Molecular dynamics simulation of the structure, dynamics, and thermostability of the RNA hairpins uCACGg and cUUCGg. J. Phys. Chem. B 2008, 112, 134–142. [PubMed: 18069816]

(27). Riccardi L; Nguyen PH; Stock G Free-Energy Landscape of RNA Hairpins Constructed via Dihedral Angle Principal Component Analysis. J. Phys. Chem. B 2009, 113, 16660–16668. [PubMed: 20028141]

(28). Deng NJ; Cieplak P Free Energy Profile of RNA Hairpins: A Molecular Dynamics Simulation Study. Biophys. J 2010, 98, 627–636. [PubMed: 20159159]

(29). Garcia AE; Paschek D Simulation of the pressure and temperature folding/unfolding equilibrium of a small RNA hairpin. J. Am. Chem. Soc 2008, 130, 815–817. [PubMed: 18154332]

(30). Zuo GH; Li WF; Zhang J; Wang J; Wang W Folding of a Small RNA Hairpin Based on Simulation with Replica Exchange Molecular Dynamics. J. Phys. Chem. B 2010, 114, 5835–5839. [PubMed: 20392088]

(31). Qin PZ; Feigon J; Hubbell WL Site-directed spin labeling studies reveal solution conformational changes in a GAAA tetraloop receptor upon Mg2+-dependent docking of a GAAA tetraloop. J. Mol. Biol 2005, 351, 1–8. [PubMed: 15993422]

(32). Reblova K; Razga F; Li W; Gao HX; Frank J; Sponer J Dynamics of the base of ribosomal A-site finger revealed by molecular dynamics simulations and Cryo-EM. Nucleic Acids Res. 2010, 38, 1325–1340. [PubMed: 19952067]

(33). Stombaugh J; Zirbel CL; Westhof E; Leontis NB Frequency and isostericity of RNA base pairs. Nucleic Acids Res. 2009, 37, 2294–2312. [PubMed: 19240142]

(34). Heus HA; Pardi A Structural Features That Give Rise to the Unusual Stability of RNA Hairpins Containing GNRA Loops. Science 1991, 253, 191–194. [PubMed: 1712983]

(35). Pley HW; Flaherty KM; Mckay DB 3-Dimensional Structure of a Hammerhead Ribozyme. Nature 1994, 372, 68–74. [PubMed: 7969422]

(36). Scott WG; Finch JT; Klug A The Crystal-Structure of an all-RNA Hammerhead Ribozyme - a Proposed Mechanism for RNA Catalytic Cleavage. Cell 1995, 81, 991–1002. [PubMed: 7541315]

(37). Correll CC; Swinger K Common and distinctive features of GNRA tetraloops based on a GUAA tetraloop structure at 1.4 angstrom resolution. RNA 2003, 9, 355–363. [PubMed: 12592009]

(38). Correll CC; Wool IG; Munishkin A The two faces of the Escherichia coli 23 S rRNA sarcin/ricin domain: The structure at 1.11 angstrom resolution. J. Mol. Biol 1999, 292, 275–287. [PubMed: 10493875]

(39). Correll CC; Beneken J; Plantinga MJ; Lubbers M; Chan YL The common and the distinctive features of the bulged-G motif based on a 1.04 angstrom resolution RNA structure. Nucleic Acids Res. 2003, 31, 6806–6818. [PubMed: 14627814]

(40). Correll CC; Munishkin A; Chan YL; Ren Z; Wool IG; Steitz TA Crystal structure of the ribosomal RNA domain essential for binding elongation factors. Proc. Natl. Acad. Sci 1998, 95, 13436–13441. [PubMed: 9811818]

(41). Ban N; Nissen P; Hansen J; Moore PB; Steitz TA The complete atomic structure of the large ribosomal subunit at 2.4 angstrom resolution. Science 2000, 289, 905–920. [PubMed: 10937989]

(42). Szewczak AA; Moore PB; Chan YL; Wool IG The Conformation of the Sarcin Ricin Loop from 28s Ribosomal-RNA. Proc. Natl. Acad. Sci 1993, 90, 9581–9585. [PubMed: 8415744]

(43). Szewczak AA; Moore PB The Sarcin Ricin Loop, a Modular RNA. J. Mol. Biol 1995, 247, 81–98. [PubMed: 7897662]

(44). Yang XJ; Gerczei T; Glover L; Correll CC Crystal structures of restrictocin-inhibitor complexes with implications for RNA recognition and base flipping. Nat. Struct. Biol 2001, 8, 968–973. [PubMed: 11685244]

(45). Jucker FM; Heus HA; Yip PF; Moors EHM; Pardi A A network of heterogeneous hydrogen bonds in GNRA tetraloops. J. Mol. Biol 1996, 264, 968–980. [PubMed: 9000624]

(46). Menger M; Eckstein F; Porschke D Dynamics of the RNA hairpin GNRA tetraloop. Biochemistry 2000, 39, 4500–4507. [PubMed: 10757999]

(47). Xia TB Taking femtosecond snapshots of RNA conformational dynamics and complexity. Curr. Opin. Chem. Biol 2008, 12, 604–611. [PubMed: 18824128]

(48). Johnson JE; Hoogstraten CG Extensive Backbone Dynamics in the GCAA RNA Tetraloop Analyzed Using C-13 NMR Spin Relaxation and Specific Isotope Labeling. J. Am. Chem. Soc 2008, 130, 16757–16769. [PubMed: 19049467]

(49). Zhao L; Xia TB Direct revelation of multiple conformations in RNA by femtosecond dynamics. J. Am. Chem. Soc 2007, 129, 4118–4119. [PubMed: 17373794]

(50). Sorin EJ; Engelhardt MA; Herschlag D; Pande VS RNA simulations: Probing hairpin unfolding and the dynamics of a GNRA tetraloop. J. Mol. Biol 2002, 317, 493–506. [PubMed: 11955005]

(51). Spackova N; Sponer J Molecular dynamics simulations of sarcin-ricin rRNA motif. Nucleic Acids Res. 2006, 34, 697–708. [PubMed: 16456030]

(52). Depaul AJ; Thompson EJ; Patel SS; Haldeman K; Sorin EJ Equilibrium conformational dynamics in an RNA tetraloop from massively parallel molecular dynamics. Nucleic Acids Res. 2010, 38, 4856–4867. [PubMed: 20223768]

(53). Bowman GR; Huang XH; Yao Y; Sun J; Carlsson G; Guibas LJ; Pande VS Structural insight into RNA hairpin folding intermediates. J. Am. Chem. Soc 2008, 130, 9676–9678. [PubMed: 18593120]

(54). Ferner J; Villa A; Duchardt E; Widjajakusuma E; Wohnert J; Stock G; Schwalbe H NMR and MD studies of the temperature-dependent dynamics of RNA YNMG-tetraloops. Nucleic Acids Res. 2008, 36, 1928–1940. [PubMed: 18272534]

(55). Fadrna E; Spackova N; Stefl R; Koca J; Cheatham TE; Sponer J Molecular dynamics simulations of guanine quadruplex loops: Advances and force field limitations. Biophys. J 2004, 87, 227–242. [PubMed: 15240460]

(56). Fadrna E; Spackova N; Sarzynska J; Koca J; Orozco M; Cheatham TE; 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. [PubMed: 26616629]

(57). Ditzler MA; Otyepka M; Sponer J; Walter NG Molecular Dynamics and Quantum Mechanics of RNA: Conformational and Chemical Change We Can Believe In. Acc. Chem. Res 2010, 43, 40–47. [PubMed: 19754142]

(58). Wang JM; Cieplak P; Kollman PA 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.

(59). Perez A; Marchan I; Svozil D; Sponer J; Cheatham TE; Laughton CA; Orozco M Refinenement of the AMBER force field for nucleic acids: Improving the description of alpha/gamma conformers. Biophys. J 2007, 92, 3817–3829. [PubMed: 17351000]

(60). MacKerell AD; Banavali NK All-atom empirical force field for nucleic acids: II. Application to molecular dynamics simulations of DNA and RNA in solution. J. Comput. Chem 2000, 21, 105–120.

(61). Besseova I; Otyepka M; Reblova K; Sponer J Dependence of A-RNA simulations on the choice of the force field and salt strength. Phys. Chem. Chem. Phys 2009, 11, 10701–10711. [PubMed: 20145814]

(62). Klein DJ; Schmeing TM; Moore PB; Steitz TA The kink-turn: a new RNA secondary structure motif. EMBO J. 2001, 20, 4214–4221. [PubMed: 11483524]

(63). Case DA; Darden TA; Cheatham TE III; Simmerling CL; Wang J; Duke RE; Luo R; Walker RC; Zhang W; Merz KM; Roberts BP; Wang B; Hayik S; Roitberg A; Seabra G; Kolossváry I; Wong KF; Paesani F; Vanicek J; Wu X; Brozell SR; Steinbrecher T; Gohlke H; Cai Q; Ye X; Wang J; Hsieh M-J; Cui G; Roe DR; Mathews DH; Seetin MG; Sagui C; Babin V; Luchko T; Gusarov S; Kovalenko A; Kollman PA AMBER 10; University of California: San Francisco, 2009.

(64). Banas P; Jurecka P; Walter NG; Sponer J; Otyepka M Theoretical studies of RNA catalysis: Hybrid QM/MM methods and their comparison with MD and QM. Methods 2009, 49, 202–216. [PubMed: 19398008]

(65). Jorgensen WL; Chandrasekhar J; Madura JD; Impey RW; Klein ML Comparison of Simple Potential Functions for Simulating Liquid Water. J. Chem. Phys 1983, 79, 926–935.

(66). Berendsen HJC; Postma JPM; Vangunsteren WF; Dinola A; Haak JR Molecular-Dynamics with Coupling to an External Bath. J. Chem. Phys 1984, 81, 3684–3690.

(67). 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, 10089–10092.

(68). Essmann U; Perera L; Berkowitz ML; Darden T; Lee H; Pedersen LG A Smooth Particle Mesh Ewald Method. J. Chem. Phys 1995, 103, 8577–8593.

(69). Ode H; Matsuo Y; Neya S; Hoshino T Force Field Parameters for Rotation Around chi Torsion Axis in Nucleic Acids. J. Comput. Chem 2008, 29, 2531–2542. [PubMed: 18470965]

(70). Yildirim I; Stern HA; Kennedy SD; Tubbs JD; Turner DH Reparameterization of RNA chi Torsion Parameters for the AMBER Force Field and Comparison to NMR Spectra for Cytidine and Uridine. J. Chem. Theory Comput 2010, 6, 1520–1531. [PubMed: 20463845]

(71). Riley KM; Pitonak M; Jurecka P; Hobza P, Stabilization and Structure Calculations for Noncovalent Interactions in Extended Molecular Systems Based on Wave Function and Density Functional Theories. Chem. Rev 2010, in press.

(72). Berendsen HJC; Grigera JR; Straatsma TP The Missing Term in Effective Pair Potentials. J. Phys. Chem 1987, 91, 6269–6271.

(73). Joung IS; Cheatham TE Determination of alkali and halide monovalent ion parameters for use in explicitly solvated biomolecular simulations. J. Phys. Chem. B 2008, 112, 9020–9041. [PubMed: 18593145]

(74). Phillips JC; Braun R; Wang W; Gumbart J; Tajkhorshid E; Villa E; Chipot C; Skeel RD; Kale L; Schulten K Scalable molecular dynamics with NAMD. J. Comput. Chem 2005, 26, 1781–1802. [PubMed: 16222654]

(75). Brooks BR; Bruccoleri RE; Olafson DJ; States DJ; Swaminathan S; Karplus M CHARMM: A Program for Macromolecular Energy, Minimization, and Dynamics Calculations. J. Comput. Chem 1983, 4, 187–217.

(76). Martyna GJ; Tobias DJ; Klein ML Constant-Pressure Molecular-Dynamics Algorithms. J. Chem. Phys 1994, 101, 4177–4189.

(77). Feller SE; Zhang YH; Pastor RW; Brooks BR Constant-Pressure Molecular-Dynamics Simulation - the Langevin Piston Method. J. Chem. Phys 1995, 103, 4613–4621.

(78). Perez A; Lankas F; Luque FJ; Orozco M Towards a molecular dynamics consensus view of B-DNA flexibility. Nucleic Acids Res. 2008, 36, 2379–2394. [PubMed: 18299282]

(79). Lu XJ; Olson WK 3DNA: a versatile, integrated software system for the analysis, rebuilding and visualization of three-dimensional nucleic-acid structures. Nat. Protoc 2008, 3, 1213–1227. [PubMed: 18600227]

(80). Mlynsky V; Banas P; Hollas D; Reblova K; Walter NG; 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. [PubMed: 20420375]

(81). Reblova K; Fadrna E; Sarzynska J; Kulinski T; Kulhanek P; Ennifar E; Koca J; Sponer J Conformations of flanking bases in HIV-1 RNA DIS kissing complexes studied by molecular dynamics. Biophys. J 2007, 93, 3932–3949. [PubMed: 17704156]

(82). Sarzynska J; Reblova K; Sponer J; Kulinski T Conformational transitions of flanking purines in HIV-1 RNA dimerization initiation site kissing complexes studied by Charmm explicit solvent molecular dynamics. Biopolymers 2008, 89, 732–746. [PubMed: 18412127]

(83). Sponer J; Kypr J Different Intrastrand and Interstrand Contributions to Stacking Account for Roll Variations at the Alternating Purine-Pyrimidine Sequences in a-DNA and a-Rna. J. Mol. Biol 1991, 221, 761–764. [PubMed: 1719210]

(84). Bhattacharyya D; Bansal M A Self-Consistent Formulation for Analysis and Generation of Non-Uniform DNA Structures. J. Biomol. Struct. Dyn 1989, 6, 635–653. [PubMed: 2619932]

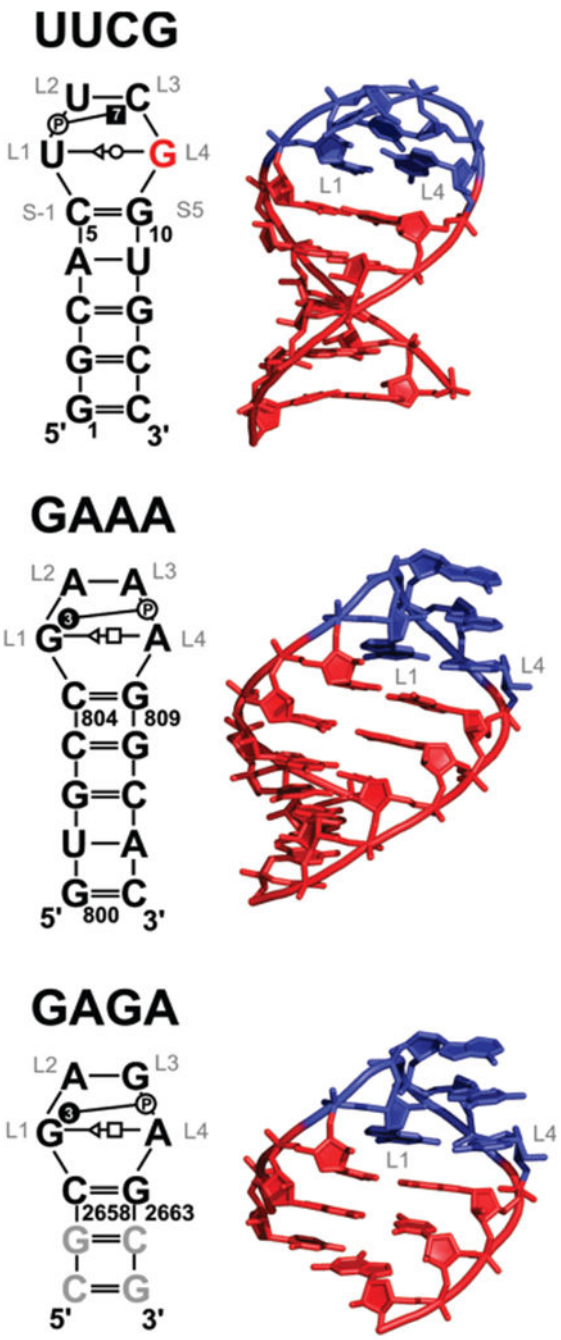


**Figure 1.**

(Left) Secondary structures of the studied systems with base pairing and base–phosphate interactions annotated according to the standard classifications.[21,23] $G_{L4}$ of the UUCG tetraloop having *syn* orientation is highlighted in red. The modeled GC pairs in the GAGA system are shown in gray. The loop residues are labeled as L1–L4 to avoid context numbering. For instance, U6 of UUCG is labeled as $U_{L1}$. (Right) Three-dimensional structures of studied systems. The A-RNA stem part is shown in red, while the tetraloop nucleotides are in blue.

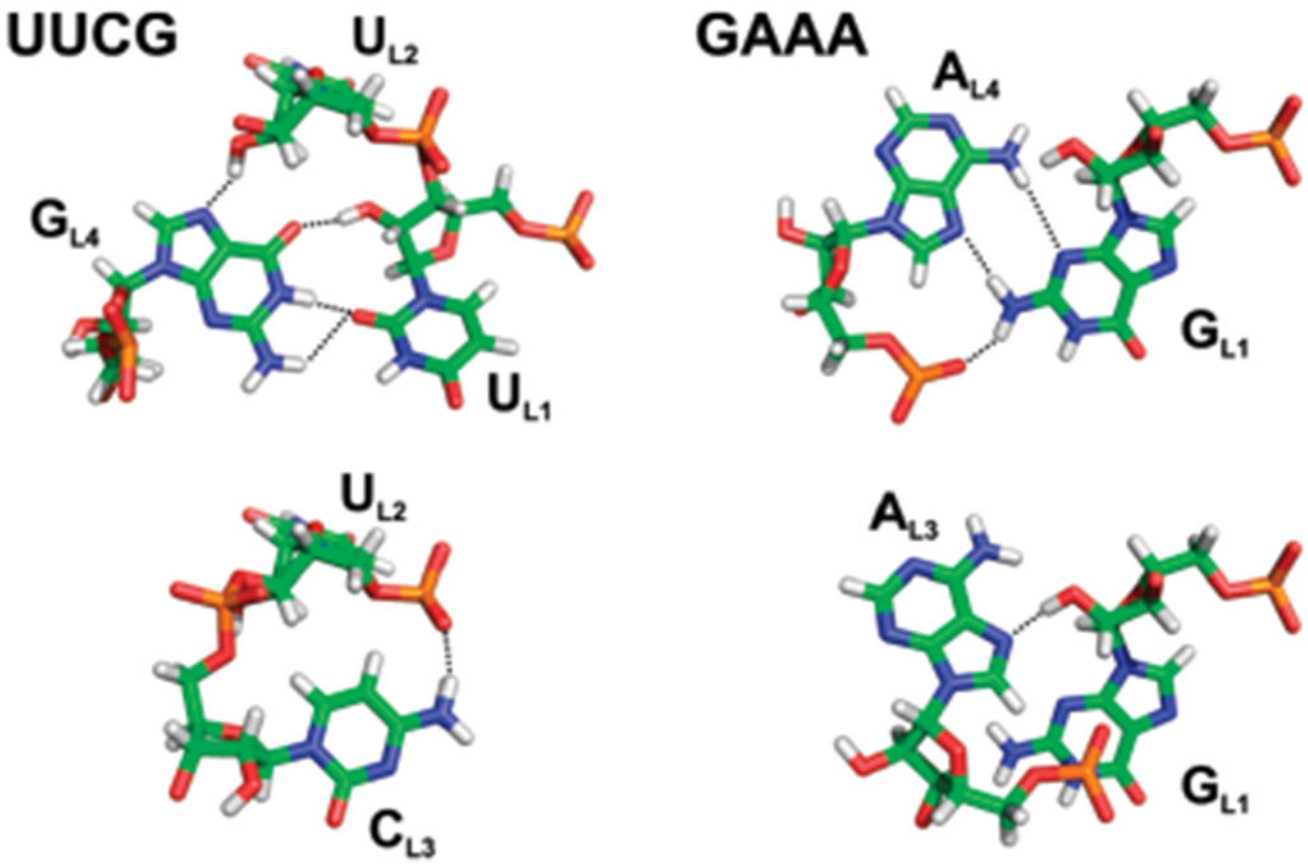


**Figure 2.**
Signature H-bonds (black dashed lines) of UUCG and GAAA tetraloops on the left and right sides, respectively.

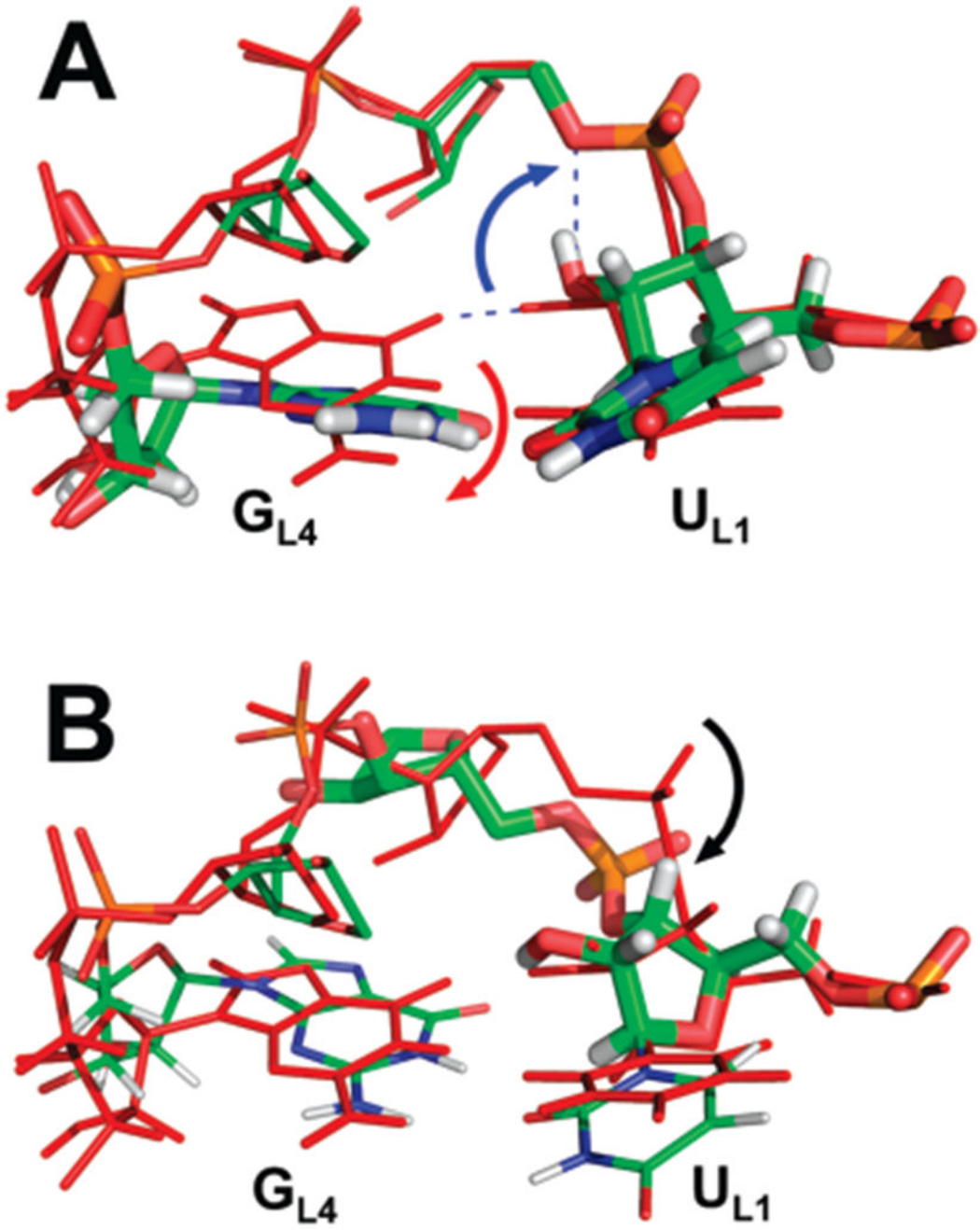


**Figure 3.**

The MD snapshots (colored by atom types) compared with high-resolution NMR structure (in red) showing structural problems seen in simulations of the UUCG tetraloop (some atoms are not shown for clarity, and important parts are shown as sticks). (A) The disruption of the $U_{L1}$(O2′)•••$G_{L4}$(O6) H-bond and formation of a new $U_{L1}$(O2′)•••$U_{L2}$(O5′) H-bond observed in all MD simulations with standard $\chi$ profiles are highlighted by the blue arrow, while the simultaneous decrease of $\chi$ of $G_{L4}$ leading to a change in the $U_{L1}/G_{L4}$ propeller is

shown by the red arrow. (B) The $U_{L2}$ phosphate $\alpha/\gamma$ flip is depicted by the black arrow. See the text for full details.

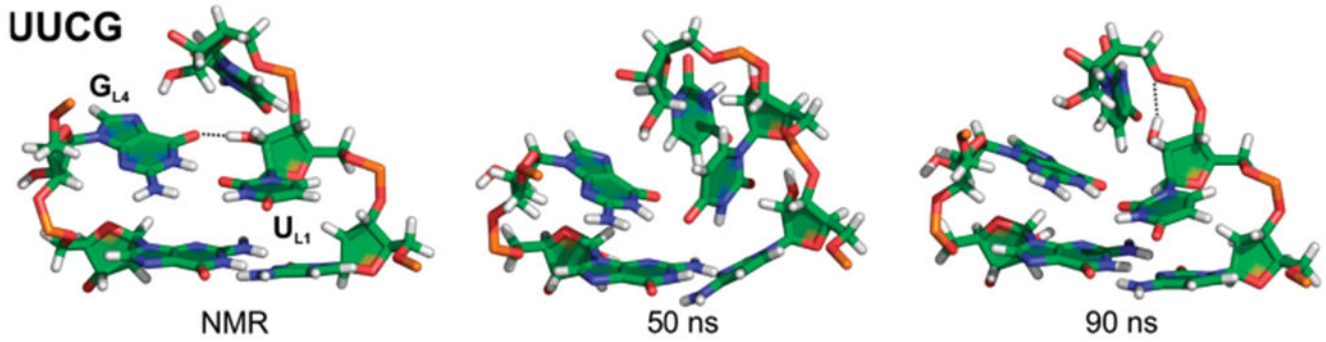


**Figure 4.**

Structures of UUCG TL at the beginning of the UUCG_bsc0 simulation, at 50 ns and at 90 ns, showing the distortion of the UUCG tetraloop. $C_{L3}$ is not shown, for clarity.

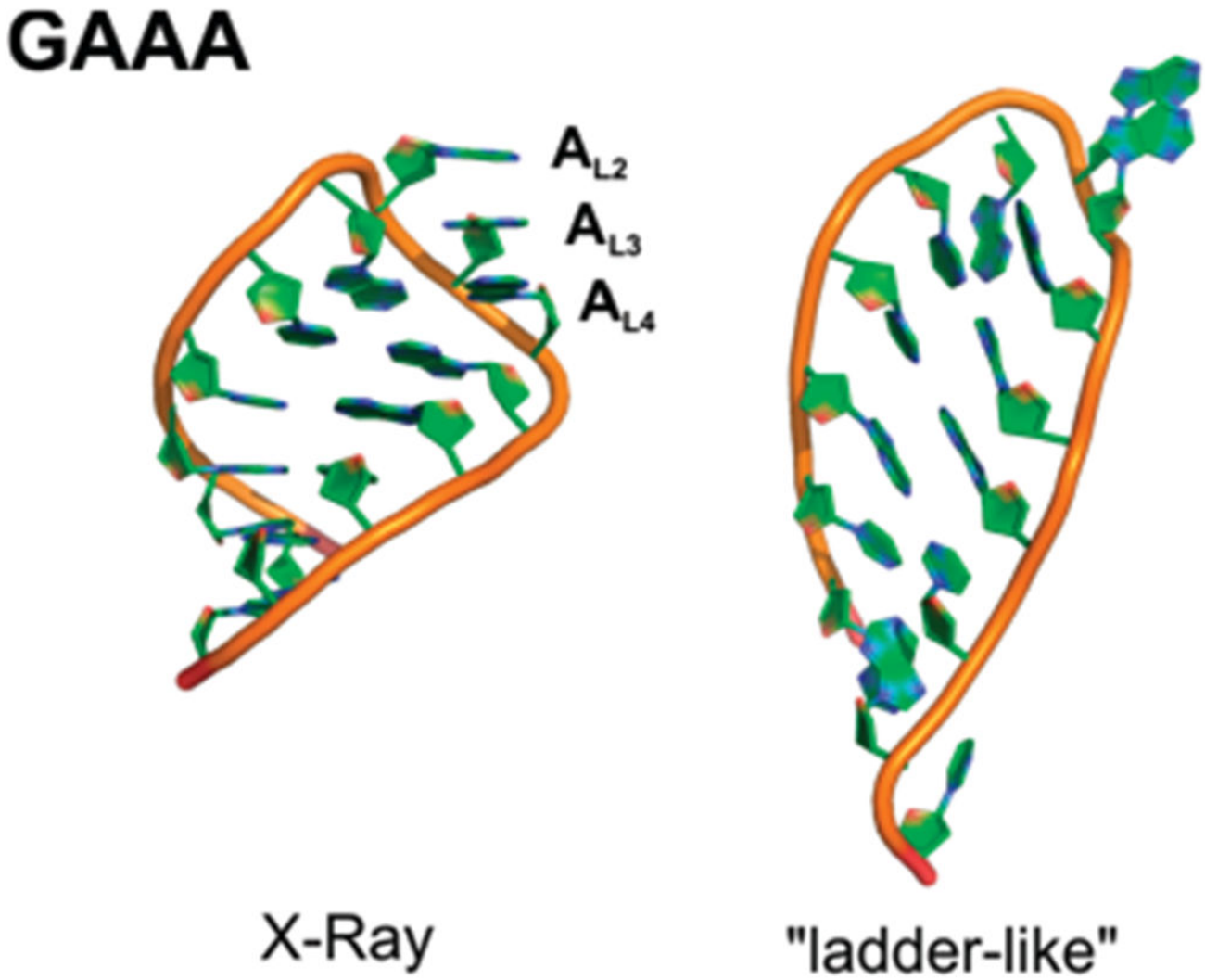


**Figure 5.**

Ladder-like conformer as observed in a simulation of the GNRA tetraloop with AMBER force fields, unless the $\chi$ torsion profile is appropriately modified. Initial geometry is on the left, and “ladder-like” conformer is on the right.

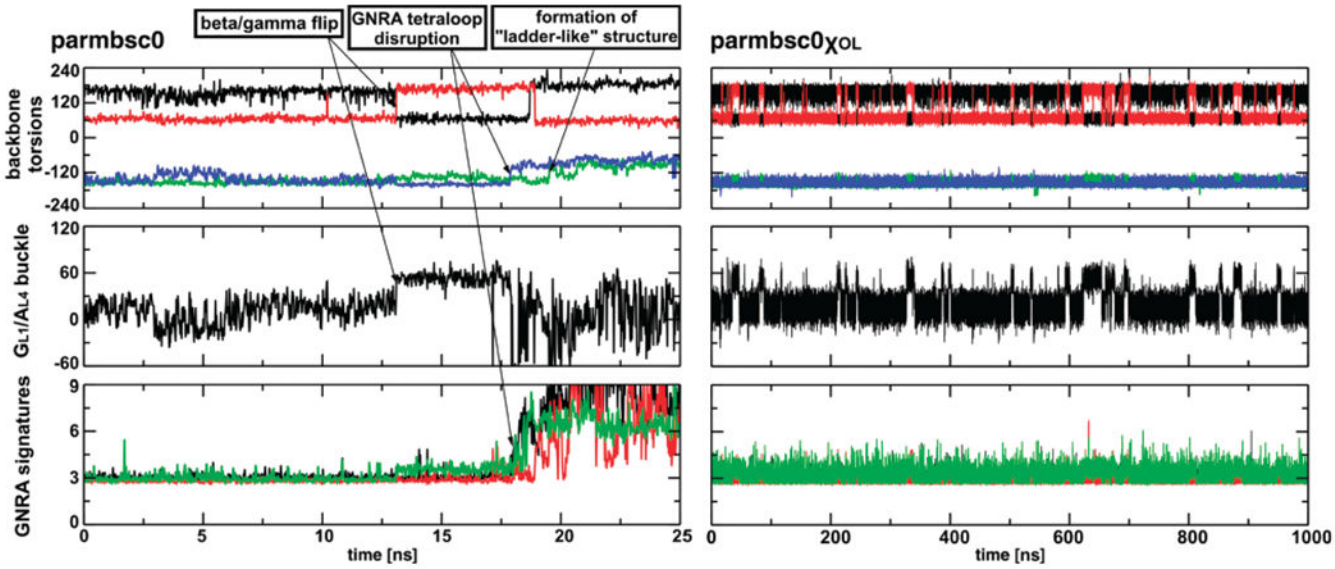


**Figure 6.**

Typical progression of GNRA tetraloop AMBER simulations. Left: simulations without the $\chi$ correction or with Ode et al.'s correction illustrated by a 25 ns GAGA_bsc0 simulation. Right: simulations with $\chi_{\text{OL-DFT}}$, $\chi_{\text{OL}}$, and $\chi_{\text{YIL}}$ variants illustrated by a 1 $\mu$s GAGA_bsc0$\chi_{\text{OL}}$ simulation. The upper graphs present a time evolution of the $\beta(G_{L3})$ torsion (black line), $\gamma(G_{L3})$ torsion (red line), and mean $\chi$ torsion averaged over either stem nucleobases (green line) or the GNRA tetraloop (blue line). The middle graph shows the $G_{L1}/A_{L4}$ buckle, and the lower graph presents the GNRA tetraloop signature H-bonds: $G_{L1}(N2)\bullet\bullet\bullet A_{L4}(N7)$, $G_{L1}(N2)\bullet\bullet\bullet A_{L4}(pro\text{-}R_p)$, and $G_{L1}(O2')\bullet\bullet\bullet G_{L3}(N7)$ in black, red, and green, respectively.

**Table 1.**

Overview of MD Simulations of TL Systems Carried Out

| label | force field | duration (ns) | the first appearance of "ladder-like" structure[80] (ns)[a] |
|---|---|---|---|
| UUCG_94 | *ff*94 | 50 | NO |
| UUCG_99 | *ff*99 | 100 | NO |
| UUCG_bsc0 | *ff*99bsc0 | 100 | NO |
| UUCG_99$\chi_{ODE}$ | *ff*99$\chi_{ODE}$ | 100 | NO |
| UUCG_bsc0$\chi_{ODE}$ | *ff*99bsc0$\chi_{ODE}$ | 100 | NO |
| UUCG_99$\chi_{YIL}$ | *ff*99$\chi_{YIL}$ | 100 | NO |
| UUCG_bsc0$\chi_{YIL}$ | *ff*99bsc0$\chi_{YIL}$ | 300 | NO |
| UUCG_99$\chi_{OL\text{-}DFT}$ | *ff*99$\chi_{OL\text{-}DFT}$ | 100 | NO |
| UUCG_bsc0$\chi_{OL\text{-}DFT}$ | *ff*99bsc0$\chi_{OL\text{-}DFT}$ | 300 | NO |
| UUCG_99$\chi_{OL}$ | *ff*99$\chi_{OL}$ | 100 | NO |
| UUCG_bsc0$\chi_{OL}$ | *ff*99bsc0$\chi_{OL}$ | 800 | NO |
| UUCG_charmm | CHARMM27 | 50 | NO |
| UUCG_99SE[b] | *ff*99, KCl SE | 100 | NO |
| UUCG_bsc0SE[b] | *ff*99bsc0, KCl SE | 100 | NO |
| GAAA_99 | *ff*99 | 50 | NO |
| GAAA_bsc0 | *ff*99bsc0 | 100 | NO |
| GAAA_99$\chi_{ODE}$ | *ff*99$\chi_{ODE}$ | 100 | 85 |
| GAAA_bsc0$\chi_{ODE}$ | *ff*99bsc0$\chi_{ODE}$ | 100 | NO |
| GAAA_99$\chi_{YIL}$ | *ff*99$\chi_{YIL}$ | 100 | NO |
| GAAA_bsc0$\chi_{YIL}$ | *ff*99bsc0$\chi_{YIL}$ | 300 | NO |
| GAAA_99$\chi_{OL\text{-}DFT}$ | *ff*99$\chi_{OL\text{-}DFT}$ | 100 | NO |
| GAAA_bsc0$\chi_{OL\text{-}DFT}$ | *ff*99bsc0$\chi_{OL\text{-}DFT}$ | 300 | NO |
| GAAA_99$\chi_{OL}$ | *ff*99$\chi_{OL}$ | 100 | NO |
| GAAA_bsc0$\chi_{OL}$ | *ff*99bsc0$\chi_{OL}$ | 800 | NO |
| GAAA_charmm27 | CHARMM27 | 100 | NO |
| GAAA_99K$^+$[c] | *ff*99, K$^+$ | 100 | 95 |
| GAAA_99SE[b] | *ff*99, KCl SE | 100 | NO |
| GAGA_99 | *ff*99 | 100 | 36 |
| GAGA_bcs0[d] | *ff*99bsc0 | 25 | 20 |
| GAGA_99$\chi_{ODE}$[d] | *ff*99$\chi_{ODE}$ | 15 | 5 |
| GAGA_bsc0$\chi_{ODE}$[d] | *ff*99bsc0$\chi_{ODE}$ | 25 | 21 |
| GAGA_99$\chi_{YIL}$ | *ff*99$\chi_{YIL}$ | 100 | NO |
| GAGA_bsc0$\chi_{YIL}$ | *ff*99bsc0$\chi_{YIL}$ | 300 | NO |
| GAGA_99$\chi_{OL\text{-}DFT}$ | *ff*99$\chi_{OL\text{-}DFT}$ | 100 | NO |
| GAGA_bsc0$\chi_{OL\text{-}DFT}$ | *ff*99bsc0$\chi_{OL\text{-}DFT}$ | 300 | NO |

| label | force field | duration (ns) | the first appearance of "ladder-like" structure[80] (ns)[a] |
|---|---|---|---|
| GAGA_99$\chi_{OL}$ | *ff*99$\chi_{OL}$ | 100 | NO |
| GAGA_bsc0$\chi_{OL}$ | *ff*99bsc0$\chi_{OL}$ | 1000 | NO |
| GAGA_charmm27 | CHARMM27 | 100 | NO |
| GAGA_99SE[b] | *ff*99, KCl SE | 100 | 50 |
| GAGA_bsc0SE[b] | *ff*99bsc0, KCl SE | 100 | NO |

[a] "NO" means not observed.

[b] Simulations in excess of KCl salt.

[c] Simulation under minimal salt conditions with $Na^+$ ions replaced by $K^+$.

[d] Simulations were terminated because a "ladder-like" structure was irreversibly formed.

**Table 2.**

Basic Structural Characteristics of UUCG TL and H-Bond Populations Calculated from MD Simulations[a]

| Structures or simulations | $G_{L4}$(N1)•••$U_{L1}$(O2) (Å) | $C_{L3}$(N4)•••$U_{L2}$(*pro*-$R_p$) (Å) | $U_{L2}$(O2′)•••$G_{L4}$(N7) (Å) | $U_{L1}$(O2′)•••$G_{L4}$(O6) (Å) | $U_{L1}$ (O2′)•••$U_{L2}$(O5′) (Å) | $G_{L4}$ $\chi$ (deg) | *t*SW $U_{L1}$/$G_{L4}$ propeller (deg) |
|---|---|---|---|---|---|---|---|
| NMR | 2.7 ± 0.1 | 2.9 ± 0.1 | 2.9 ± 0.1 | 2.6 ± 0.1 | 3.4 ± 0.1 | 58 ± 4 | −4.3 ± 4.5 |
| X-ray | 3.0 ± 0.1 | 2.9 ± 0.2 | 4.0 ± 0.5 | 2.7 ± 0.3 | 3.8 ± 0.2 | 60 ± 1 | −7.8 ± 7.0 |
| UUCG_94 | 88% | 8% | 0% | 68% | 8% | 45 ± 11 | −21 ± 11 |
| UUCG_99 | 87% | 9% | 2% | 65% | 8% | 48 ± 13 | −22 ± 11 |
| UUCG_bsc0 | 55% | 68% | 10% | 14% | 41% | 42 ± 18 | −32 ± 32 |
| UUCG_99$\chi_{ODE}$ | 90% | 72% | 54% | 94% | 4% | 83 ± 13 | −1 ± 13 |
| UUCG_bsc0$\chi_{ODE}$ | 74% | 66% | 50% | 70% | 18% | 75 ± 17 | −14 ± 22 |
| UUCG_99$\chi_{YIL}$ | 95% | 76% | 41% | 87% | 6% | 65 ± 15 | 1 ± 11 |
| UUCG_bsc0$\chi_{YIL}$ | 92% | 77% | 46% | 89% | 6% | 69 ± 17 | 2 ± 11 |
| UUOG_99$\chi_{OL\text{-}DFT}$ | 92% | 36% | 28% | 83% | 7% | 67 ± 14 | −8 ± 13 |
| UUCG_bsc0$\chi_{OL\text{-}DFT}$ | 93% | 70% | 50% | 88% | 9% | 76 ± 14 | −1 ± 11 |
| UUCG_99$\chi_{OL}$ | 93% | 68% | 42% | 80% | 15% | 64 ± 15 | −5 ± 12 |
| UUCG_bsc0$\chi_{OL}$ | 92% | 71% | 49% | 85% | 10% | 72 ± 17 | −3 ± 12 |
| UUCG_99SE | 82% | 55% | 8% | 15% | 53% | 38 ± 15 | −25 ± 12 |
| UUCG_bsc0SE | 79% | 59% | 6% | 7% | 57% | 34 ± 13 | −29 ± 11 |

[a]NMR values were averaged from a set of 20 structures taken from PDB 2KOC.[17] X-ray values were averaged from X-ray structures 1F7Y (res. 2.8 Å),[14] 1I6U (res. 2.6 Å),[15] and 1FJG (res. 3.0 Å).[16] Some of the values are presented as average ± standard deviation. H-bond populations are calculated from respective MD simulations of UUCG TL (see Methods).