



HAL
open science

Multiscale Modelling of Phosphate... π Contacts in RNA U-turns Exposes Differences Between Quantum-ChemicalM and AMBER Force Field Descriptions

Klaudia Mráziková, Holger Kruse, Vojtěch Mlýnský, Pascal Auffinger, Jirí Sponer

► **To cite this version:**

Klaudia Mráziková, Holger Kruse, Vojtěch Mlýnský, Pascal Auffinger, Jirí Sponer. Multiscale Modelling of Phosphate... π Contacts in RNA U-turns Exposes Differences Between Quantum-ChemicalM and AMBER Force Field Descriptions. Journal of Chemical Information and Modeling, In press. hal-03857441

HAL Id: hal-03857441

<https://cnrs.hal.science/hal-03857441>

Submitted on 17 Nov 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Multiscale modelling of phosphate... π contacts in RNA U-turns reveals AMBER force-field deficiencies

Klaudia Mrazikova^{1,2}, Holger Kruse¹, Vojtech Mlynsky¹, Pascal Auffinger^{3*}, Jiri Sponer^{1*}

¹ Institute of Biophysics of the Czech Academy of Sciences, Královopolská 135, 612 65 Brno, Czech Republic

² National Centre for Biomolecular Research, Faculty of Science, Masaryk University, Kamenice 5, 625 00 Brno, Czech Republic

³ Architecture et Réactivité de l'ARN, Université de Strasbourg, Institut de Biologie Moléculaire et Cellulaire du CNRS, Strasbourg, 67084, France

*corresponding authors, e-mails: p.auffinger@ibmc-cnrs.unistra.fr, sponer@ncbr.muni.cz

Abstract

Phosphate... π , also called anion... π , contacts occur between nucleobases and phosphate OP oxygens in r(GNRA) and r(UNNN) U-turn motifs (N = A,G,C,U; R = A,G). We investigated these contacts in detail by using state-of-the-art quantum chemical methods (QM) to characterize some of their physico-chemical properties and to evaluate the ability of the AMBER force field (AFF) to describe these contacts. We found that AFF interaction energies of phosphate... π contacts calculated for model dimethyl phosphate...nucleobase systems are less stabilizing in comparison with double-hybrid DFT methods and that the minimum contact distances are stretched for all nucleobase systems. This distance stretch is also observed in large-scale AFF computations on several r(gcGNRAgc) tetraloop hairpins when compared to QM/MM. Further, classical molecular dynamics (MD) simulations of these tetraloop hairpins confirm this distance stretch and reveal shifted OP2/nucleobase positions when compared to experimental data extracted from high-resolution X-ray/cryo-EM structures (≤ 2.5 Å) of r(GNRA) tetraloops using the WebFR3D bioinformatic tool. We propose that discrepancies between QM and AFF are caused by a combination of missing polarization, too large AFF Lennard-Jones (LJ) radii of nucleobase carbon atoms and exaggerated short-range repulsion due to an approximate r^{-12} LJ repulsive term. We put these results in regard with those obtained in earlier investigations on lone pair... π contacts occurring in CpG Z-steps. Charge-transfer calculations do not support any significant $n \rightarrow \pi^*$ donation effects and hence this label is inappropriate. We also investigated thiophosphate... π contacts for which we calculated less stabilizing interaction energies than for the phosphate... π contacts. We thus challenge suggestions that the experimentally observed enhanced thermodynamic stability of phosphorothioated r(GNRA) tetraloops can be straightforwardly explained by larger London dispersion.

Introduction

In nucleic acids, phosphate... π also called anion... π contacts^{1,2} are part of the ubiquitous U-turn motifs found in r(GNRA) tetraloop hairpins and tRNA anticodon loops with r(UNNN) sequences. In these motifs, an OP2...G/U contact with a ≈ 3.0 Å average distance is formed between the phosphate group of the third nucleotide and the first nucleobase (Figure 1).³⁻⁶ This phosphate... π contact is one of the two main signature interactions that define a U-turn, the second being a hydrogen bond between the Watson-Crick edge imino/amino groups of the first G/U nucleobase and the OP2 atom of the fourth nucleotide.^{6,7} The latter interaction is classified as base...phosphate or BPh interaction of type 3, 4 or 5 depending on the Watson-Crick N-H group it interacts with.⁸ Several studies discussed the possible stabilization role of phosphate... π contacts in r(GNRA) tetraloop hairpins.^{9,10}

The r(GNRA) tetraloops that represent a subcategory of U-turn motifs play major biological roles. They initiate RNA folding and are involved in tertiary contacts with other RNA segments and proteins.¹¹⁻¹⁵ Despite their high thermodynamic stability^{11,16} and the apparent simplicity of their four-nucleotide structure, MD simulations of r(GNRA) hairpins remain challenging.¹⁷⁻²⁸ Classical force fields (FFs) used in MD simulations, like the Cornell et al. AMBER force field²⁹ and its subsequent modifications use pairwise-additive potentials and fixed point-charges. The most frequently used version of the AMBER force field (AFF) for RNA is ff99bsc0 χ_{OL3} , abbreviated as OL3.³⁰ Despite reasonable success in reproducing nucleic acid fine structural details,²⁸ Kùhrová et al.²⁴ reported OL3 imbalances that either destabilize folded state of a r(gcGAGAgc) tetraloop or stabilize the unfolded ensemble. These imbalances were attributed to the underestimation of the base pair hydrogen bond strength, stabilization of non-native ribose...phosphate hydrogen bonds and overstabilization of BPh interactions. The recently introduced “generalized hydrogen bond fix”, called gHBfix, resulted in a notable improvement in the folding behavior of the r(gcGAGAgc) tetraloop during MD simulations by inducing a substantial stabilization of base pair hydrogen bonds and a weakening of non-native ribose...phosphate hydrogen bonds.^{19,31} Several correction terms including various variants of gHBfix and NBfix (non-bonded fix) potentials tested on converged simulations of r(GAGA) tetraloop lead to improvement of the estimated free energy balance between folded and unfolded states.²⁰ However, it is possible that these AFF modifications compensate for imprecisions in other AFF terms, which may lead to overfitting and limited transferability of the parametrization.

FF imbalances are affecting all systems and are not limited to r(GNRA) tetraloops. Recently, short-range imbalances in the AFF description of lp... π contacts between the (deoxy)ribose O4' atom and guanine in CpG Z-steps were described.³²⁻³⁴ It was suggested that these imbalances contribute to an improper dynamical description of r(UNCG) tetraloops and Z-DNA helices during MD simulations.^{33,34} Standard van der Waals radii for nucleobase sp^2 carbons were questioned,³⁵ and a revision of Lennard-Jones (LJ) parameters of nucleobase atoms in AFF was advocated.

Herein, we investigate the energetics of phosphate... π contacts by using quantum mechanics (QM) calculations with a dispersion-corrected double-hybrid density-functional approximation (DHDF-D3) and Symmetry Adapted Perturbation Theory (SAPT) analysis and compare these results to those obtained by the classical AMBER force field (AFF). For that purpose, we focus on phosphate... π contacts as they occur in the two ubiquitous r(GNRA) and r(UNNN) U-turn motifs, because a search for phosphate... π contacts in X-ray/cryo-EM experimental structures with resolutions ≤ 2.5 Å revealed that U-turn motifs starting with G/U nucleobases prevail, next to rare U-turns starting with a protonated cytosine.^{7,36} Based on QM and AFF calculations we show that the absence of polarization in the AFF leads to under-stabilizing phosphate... π interaction energies. Moreover, too large AFF Lennard-Jones radii combined with overestimated short-range repulsion leads to overestimated phosphate... π distances. We discuss possible implications of these imbalances for MD simulations of r(GNRA) tetraloops. We also address the physical nature and differences of phosphate... π and lp... π contacts from a QM and AFF perspective and interrogate the $n \rightarrow \pi^*$ character often attributed to these interactions.

Finally, we investigate the effect of a phosphorothioate substitution at the phosphate... π position. Phosphorothioated nucleic acids are frequently used in therapeutics and nanotechnologies

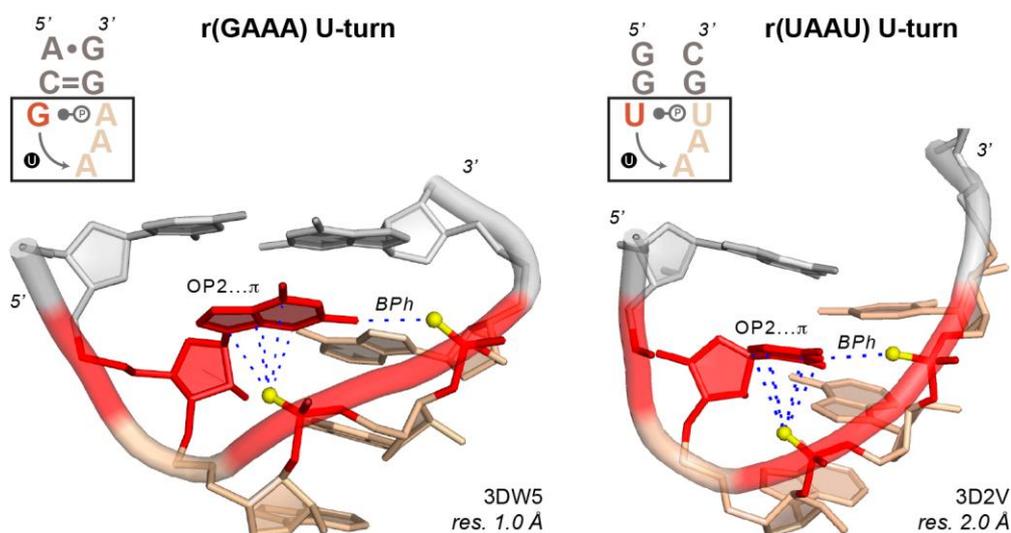


Figure 1. Two U-turn tetraloop motifs characterized by a phosphate...guanine/uracil (OP2... π) contact and a base...phosphate (BPh)⁸ interaction closing the loop. These contacts are visualized by blue dashed lines. The OP2 atoms of nt3 (i.e., third nucleotide of the U-turn motif) and nt4 are shown as yellow spheres. For clarity, all non-relevant OP atoms were hidden. For the secondary structures, symbols according to the Leontis and Westhof nomenclature were used.¹³¹ In the secondary structure, the U-turn is symbolized by a circled “U”. The 1st nucleobase and the phosphate of the 3rd nucleotide (nt3) involved in a phosphate... π contact are in red as well as the OP2 atoms involved in the (nt1...nt4)base...phosphate (BPh) hydrogen bonds. The other three tetraloop nucleotides are in wheat.

and were uncovered in DNA and RNA of prokaryotes and eukaryotes.^{9,37–42} At least two studies^{9,43} reported enhanced stabilization of r(GNRA) tetraloops through melting and calorimetry experiments after thio-substitution of the (nt3)OP2 (nt3 = third nucleotide of the U-turn motif) atom that is involved in the phosphate... π contact (Figure 1). The experimentally determined difference in ΔG is $-2.9 \text{ kcal.mol}^{-1}$ (37°C) in ref. ⁴³ and $-4.7 \text{ kcal.mol}^{-1}$ (25 °C) in ref. ⁹. This experimental difference in stability is puzzling given that the native and phosphorothioated loops were found to have identical X-ray structures.⁹ Indeed, Egli et al. presented high-resolution X-ray studies of several phosphorothioated r(GAGA) hairpins that showed that their structure is similar to that of the native tetraloop even when the (nt3)OP2 atom was substituted by a neutral methylphosphonate. A theoretical study of the neomycin sensing riboswitch coupled with NMR data showed similarly that the U-turn structure, in this case a r(UNNN) motif, was not altered by a phosphorothioate substitution.⁴¹ Clear insight on the underlying physico-chemistry of these interactions could not be gained from the preceding investigations.^{9,41,43} Although we complement current knowledge on these interactions and document differences between QM and AFF, our calculations cannot explain the experimental gain in stabilities observed upon phosphorothioation in RNA.

Methods

WebFR3D searches of phosphate... π contacts in PDB structures

The WebFR3D tool^{44,45} was used to find U-turn motifs with OP2...nucleobase contacts in RNA structures from the PDB. For that purpose, an ensemble of X-ray and cryo-EM structures with resolution ≤ 2.5 Å using the 3.221 set of non-redundant structures (March 2022)⁴⁶ was used for the analysis. A general search on phosphate... π contacts in RNA and specifically on U-turn motifs confirmed that the r(GNRA) and r(UNNN) sequences were highly represented among all RNA motifs with a phosphate... π contact.

For r(GNRA) and r(UNNN) U-turn motifs, the following search constraints were used: *i*) all nucleotides are consecutive, *ii*) the dihedral angle of the glycosidic bond is in *anti* for all nucleotides, *iii*) a phosphate... π contact between the (nt1)3'-face and the (nt3)OP2 atom is present (Figure 1), and *iv*) a (nt1...nt4)BPh interaction of type 3, 4 or 5 is present.⁸ Additionally, for r(nGNRA) tetraloop hairpins, the first stem nucleotides form a Watson-Crick base pair to ensure retrieving canonical hairpins. For r(UNNN), no constraints were placed on the closing base pairs given the paucity of r(UNNN) tetraloop hairpins. More details are given in Figure S1 and S2. Specifics on WebFR3D phosphate... π and lp... π contact searches are to be found in Zirbel and Auffinger.⁴⁷

Potential energy surface (PES) calculations

Preparation of model system structures. The initial stacked phosphate...guanine model used for the phosphate... π calculations was extracted from a high-resolution crystallographic structure of a sarcin-ricin hairpin containing an r(cGAG₂₆₆₁Ag) tetraloop hairpin (PDBid: 4NLF; res. 1.0 Å⁴⁸). The phosphate moiety of residue 2661 was manually transformed into either dimethyl phosphate (DMP) or dimethyl thiophosphate (DMTP). The G nucleobase was replaced by A, C or U with Jmol⁴⁹ to create all DMP/DMTP...nucleobase models.

QM geometry optimizations. The geometries of all DMP/DMTP...nucleobase dimers were optimized at the B3LYP-D3(BJ)/def2-QZVP level of theory.⁵⁰⁻⁵³ We applied harmonic penalty function restraints to preserve the monomers orientation observed in the experimental structure using an in-house optimizer *Xopt*^{54,55} coupled to TURBOMOLE V.7.3⁵⁶ (Table S1). An energy change threshold of 10^{-7} E_h was used for the SCF and geometry convergence. The DFT quadrature grid *m4* was employed.⁵⁷ The resolution of identity approximation for Coulomb and HF-exchange integrals (RI-JK) was used to accelerate the computations.⁵⁸ Since the anionic DMP tended to form hydrogen bonds with nucleobase edges during optimization, we used the COSMO⁵⁹ implicit solvation model with a dielectric constant ϵ that was set to infinity to keep the DMP probe over the nucleobase. The atomic radii used for the cavity construction are: 2.00 Å (C), 1.83 Å (N), 1.72 Å (O), 2.11 Å (P), 2.16 Å (S) and 1.30 Å (H). These optimized structures were used as starting structures for the potential energy surface (PES) mapping (see below).

QM interaction energy calculations. The single-point energy values for the construction of the potential energy surfaces (PES) were computed by using the double-hybrid density functional approximation DSD-BLYP⁶⁰ with DFT-D3(BJ)^{51,61} correction and the minimally-augmented ma-def2-QZVPP^{53,62} basis set using ORCA V.4.2.⁶³ This method was used for its high accuracy proven in benchmark studies^{64,65} and also since dispersion-corrected double-hybrid DFT methods are recommended for the description of non-covalent anion... π interactions.⁶⁶ The DFT quadrature grid *GRID5* was used. The RI approximation for Coulomb integrals (RI-J) and COSX (chain-of-spheres exchange) approximation for exchange integrals with the accurate *GRIDX6* grid were employed with automatic construction of auxiliary basis set (AutoAux keyword).^{67,68} A tight energy change threshold for SCF convergence was used (*TightSCF* keyword). Coefficient values for the DSD-BLYP density functional were taken from the original publication.⁶⁰

The reference FNO-CCSD(T) computations with extrapolation to the complete basis set (CBS) were performed using PSI4 v1.4⁶⁹ as:

$$E_{CCSD(T)/CBS} = E_{HF/aQ} + E_{MP2/CBS(aT,aQ)} + E_{FNO-CCSD(T)/juntZ} - E_{MP2/juntZ}$$

where aT , aQ and $junTZ$ stand for the aug-cc-pVTZ, aug-cc-pVQZ and jun-cc-pVTZ basis set, respectively.^{70–72} The $MP2/CBS$ represents Halkier's 2-point extrapolation scheme for correlation energies.⁷³ The $FNO-CCSD(T)$ was used together with density fitting using standard auxiliary basis sets and frozen core approximation.

SAPT calculations. Symmetry Adapted Perturbation Theory (SAPT) analysis was performed at the recommended SAPT2+(3) δ MP2 “gold-level”^{74–76} with jun-cc-pVTZ basis set^{70–72} using PSI4 v1.4.⁷⁷ To estimate the charge-transfer (CT) contribution the same level of theory was used but without the δ MP2 correction as it does not apply in that case. The SAPT scheme divides the interaction energy in electrostatic, exchange-repulsion, induction, and London dispersion components. The induction energy includes polarization and charge transfer. These terms are neglected by classical point-charge force-fields but polarization is taken into account in polarizable force-fields for which, however, obstacles still remain that limit their general usage for nucleic acids.^{78–81} As reported further in the text, charge transfer contributions are negligible for present phosphate... π interactions and thus induction energy corresponds to polarization in this case (see **Section 2** in the **Supporting Information**).

AFF interaction energy calculations. The AMBER force field (AFF) calculations were carried out for DMP...nucleobase models by using the in-house *bff* program.⁸² Only non-bonded terms of the AFF were used for calculating interaction energies. The Lennard-Jones (LJ) parameters for OL3 were taken from the AmberTool library files of AMBER16.⁸³ Note that they are identical to original Cornell et al. AMBER parameters.²⁹ Partial charges were calculated using the AMBER14 RESP^{84,85} module of Antechamber for non-interacting monomers (DMP/DMTP and nucleobase) in the geometry-optimized complex.⁸⁶ The electrostatic potential used for the RESP calculations was computed at the HF/6-31G* level of theory using the Merz-Singh-Kollman (MK) population analysis employing Gaussian 09.⁸⁷ The AFF(CP) notation refers to alternative Lennard-Jones parameters for phosphate OP1/OP2 and O3'/O5' atoms that were taken from Steinbrecher, Case et al.⁸⁸ (**Table S2**); “CP” signifies “Case phosphates”.

Implicit solvent calculations. For DHDF-D3 calculations, a conductor-like polarizable continuum model (CPCM) was used with the COSMO-type ϵ function, Gaussian charge scheme with a scaled vdW cavity and larger grid, i.e., 770 Lebedev points for each scaled vdW sphere.^{59,89–91} The ORCA V.4.2 software was used. For AFF, the generalized Born (GB) implicit solvent model was used.^{92,93} The dielectric constant ϵ was set to the default value of ϵ for water environments (80.4 for COSMO and 78.5 for GB); for biomolecular environments, $\epsilon=4.0$ was used. All the other parameters were set as noted in previous sections.

Minimum Interaction Energy distance (*minIEd*) and Interaction Energy (*IE*) surfaces. Single-point energy calculations were performed at different positions of the (DMP/DMTP)OP2/SP2 atoms over each of the four nucleobases. This procedure is similar to that previously used for dimethyl ether (DME) scans over nucleobases.^{34,35} Starting with the optimized DMP/DMTP...nucleobase geometry, DMP/DMTP was shifted in horizontal and vertical directions with respect to the nucleobase. For the horizontal displacements, three positions from the pyrimidine ring centroid to each of the pyrimidine ring atoms and two positions from the imidazole ring centroid to each of the imidazole ring atoms with an additional position between C4 and C5 atoms were defined (**Figure S3b**). Vertical interaction energy scans were calculated for each of these positions. Altogether, ≈ 580 (29 horizontal x 20 vertical displacements) and ≈ 380 (19 horizontal x 20 vertical displacements) single-point energies were calculated for purine (G, A) and pyrimidine (U, C) nucleobase scans, respectively. The *minIEd* surfaces were constructed from minimum-energy distances of the vertical scans while *IE* surfaces were created from the interaction energies at the minimum-energy distances (**Figure S3** and **Section 1.1** in the **Supporting Information**; see also ref. ³⁴).

QM/MM and AFF geometry optimizations for r(GNRA) tetraloops

Initial structures for QM/MM and AFF (MM) geometry optimizations of r(GNRA) hairpins were extracted from the four r(gcGAAAgc), r(gcGAGAgc), r(gcGCAAgc) and r(gcGCGAgc) MD simulations (see below). Three snapshots (s1, s2 and s3) were taken from each simulation.

A SPC/E⁹⁴ water sphere with a 40 Å radius from the RNA centroid of mass was added to the tetraloops. This was followed by the addition of seven K⁺ counterions for RNA charge neutralization with the AMBER16⁸³ tLEaP module. Prior to the hairpin geometry optimizations, short 10-20 ps water and K⁺ equilibrations were performed for each structure.

For QM/MM calculations, the QM region included the whole RNA hairpin while solvent molecules were treated by MM. The additive QM/MM scheme using the electrostatic embedding with the point-charge approximation option was used as implemented in the Sander module of AMBER14.⁸⁴ This option considers QM electron density polarization by MM point charges. An in-house modified version of Sander was used to couple the QM/MM module of AMBER14^{84,95} with TURBOMOLE V7.3.^{56,96} A hybrid-DFT PBEh-3c method,⁹⁷ employing double- ζ valence-polarized def2-mSVP basis set, D3(BJ)^{51,52} dispersion correction and geometrical counterpoise (gCP) correction⁹⁸ for BSSE, was used with the resolution-of-identity (RI) approximation for Coulomb integrals.⁵⁸ For MM the state-of-the-art OL3³⁰ AFF was used with the phosphate OP1/OP2 and O3'/O5' LJ parameters by Steinbrecher, Case et al.⁸⁸ This AFF version is abbreviated as OL3CP and was also used for molecular dynamics (MD) simulations as described below. The limited-memory Broyden-Fletcher-Goldfarb-Shanno (L-BFGS)⁹⁹ quasi-Newton algorithm with a convergence threshold for gradient norm of 10⁻⁴ kcal·mol⁻¹·Å⁻¹ was used for geometry optimizations.

MD simulations of GNRA tetraloops

Classical MD simulations were performed for the four 8-mer r(GNRA) tetraloop sequences mentioned above with the AMBER18¹⁰⁰ software and the OL3CP FF^{29,30,88,101-103} (AMBER library file of this FF version can be found in the Supporting Information of ref. ²⁴) with the gHBfix19 potential.^{19,20,33} The gHBfix19 correction stabilizes all –NH...N– interactions by 1.0 kcal·mol⁻¹ while weakening the sugar–phosphate interactions by 0.5 kcal·mol⁻¹. Initial geometries for MD simulations of r(GNRA) tetraloops were taken from X-ray structures: r(gc₂₄₁₁GAAAgc) from PDBid 1M90;¹⁰⁴ r(gc₂₆₅₈GAGAgc) from PDBid 1Q9A;¹⁰⁵ r(gc₅₇₆GCGAgc) from PDBid 1JJ2¹⁰⁶ and from the PDBid 1ZIH¹⁰⁷ (model1) NMR structure for r(gcGCAAgc). For r(gcGAGAgc), one of the stem base pairs was manually modified from a non-canonical A•G to a canonical G=C base pair as described in ref. ²⁴. Tetraloops were solvated using a cubic box of ≈50 Å edge containing about 3000 OPC¹⁰⁸ water molecules. Simulations were run at T = 298 K with a ≈0.15 M KCl salt concentration using the Joung-Cheatham ion parameters¹⁰⁹ and with the hydrogen mass repartitioning scheme that allows using a 4-fs integration time step.¹¹⁰ One 5 μs long trajectory was obtained for each of the four analyzed GNRA sequences in NVT ensemble with SHAKE restraints by using the *pmemd.cuda* module of AMBER18 (for other details about the tetraloop simulation protocol, see Supporting Information of ref. 34). Each simulation comprises 1.25·10⁹ simulation steps; every 100th ps was used for analysis. Each simulation takes around 10 days using one GeForce Turing RTX2080 Ti GPU card.

Analysis of OP2...guanine contacts in MD simulations of GNRA tetraloops

To obtain consistent average OP2...guanine distances from MD simulations, several criteria were set to avoid accounting for snapshots with a tetraloop structure deviating too far from the native conformation. These criteria are: *i*) a (nt3...nt1)OP2...G distance ≤ 3.8 Å, *ii*) projection point of the OP2 atom on the G plane within either 1.5 Å from pyrimidine ring centroid or 1.25 Å from imidazole ring centroid, and *iii*) the angle between the nt1 (G) and nt4 (A) planes ≤ 45°. In this and previous database studies⁶ a stricter 3.5 Å cutoff was used for experimental structures. However, given thermal fluctuations in MD simulations and a tetraloop tendency to unfold and generate longer OP2...G distances, we chose to expand the OP2...G distance cutoff to 3.8 Å for MD analyses.

Results and Discussion

A WebFR3D exploration of phosphate... π contacts in X-ray/cryo-EM structures shows that most of the U-turns start with G or U

We searched the PDB with WebFR3D^{44,45} to find phosphate... π contacts in an ensemble of X-ray and cryo-EM structures with resolution ≤ 2.5 Å using the 3.221 set of non-redundant structures (March 2022).⁴⁶ This search led to 601 hits with an OP...nucleobase contact that divide into 64 (OP1) and 537 (OP2) contacts. These hits involving all nucleotide combinations show that OP atoms can stack on any nucleobase. Among them, we found 311 r(NNNN) motifs with U-turn characteristics, i.e., with a (nt3...nt1)OP2...nucleobase contact, a (nt1...nt4)BPh hydrogen bond involving the Watson-Crick edge (3BPh, 4BPh, 5BPh) and all bases in *anti*. This implies that U-turns are the most frequent r(NNNN) motifs involving phosphate... π contacts in RNA. This U-turn ensemble comprises 207 r(GNNN) and 104 r(UNNN) while r(CNNN) and r(ANNN) U-turns were not detected although rare r(C⁺NNN) U-turns do exist.^{6,7} For r(CNNN), five hits were found with a 7BPh hydrogen bond. Thus, r(GNNN) and r(UNNN) are the dominant sequences forming U-turns in RNA. This led us to focus on the characteristics of phosphate... π contacts involving G/U nucleobases.

Since in this study we perform MD simulations of four r(gcGNRAgc) hairpins, we further refined our PDB exploration by extracting r(nGNRA) U-turns closed by a Watson-Crick pair, i.e., tetraloop hairpins that involve an OP2... π contact since we could isolate only one r(GNNN) with OP1... π contact that seemed artefactual. No OP1... π contacts for r(UNNN) were found. The searches shown in [Figure S1](#) and [S2](#) isolated 155 r(nGNRA) hairpins and 104 r(UNNN) U-turns, respectively.

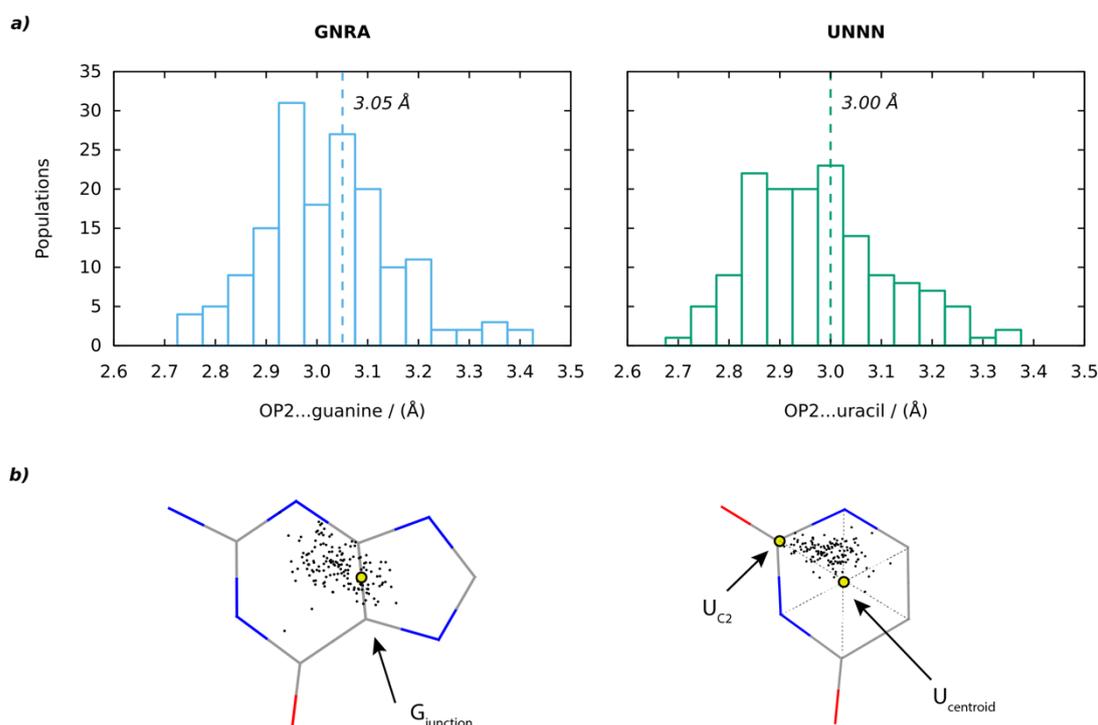


Figure 2. OP2...guanine and OP2...uracil distance histograms (a) and OP2 projection points on guanine and uracil (b) in r(nGNRA) and r(UNNN) X-ray/cryo-EM structures (res. ≤ 2.5 Å) obtained by a WebFR3D^{44,45} search ([Figure S1](#) and [S2](#)). The three idealized G_{junction}, U_{centroid} and U_{c2} positions that have been used for further analysis are shown on the 2D structures.

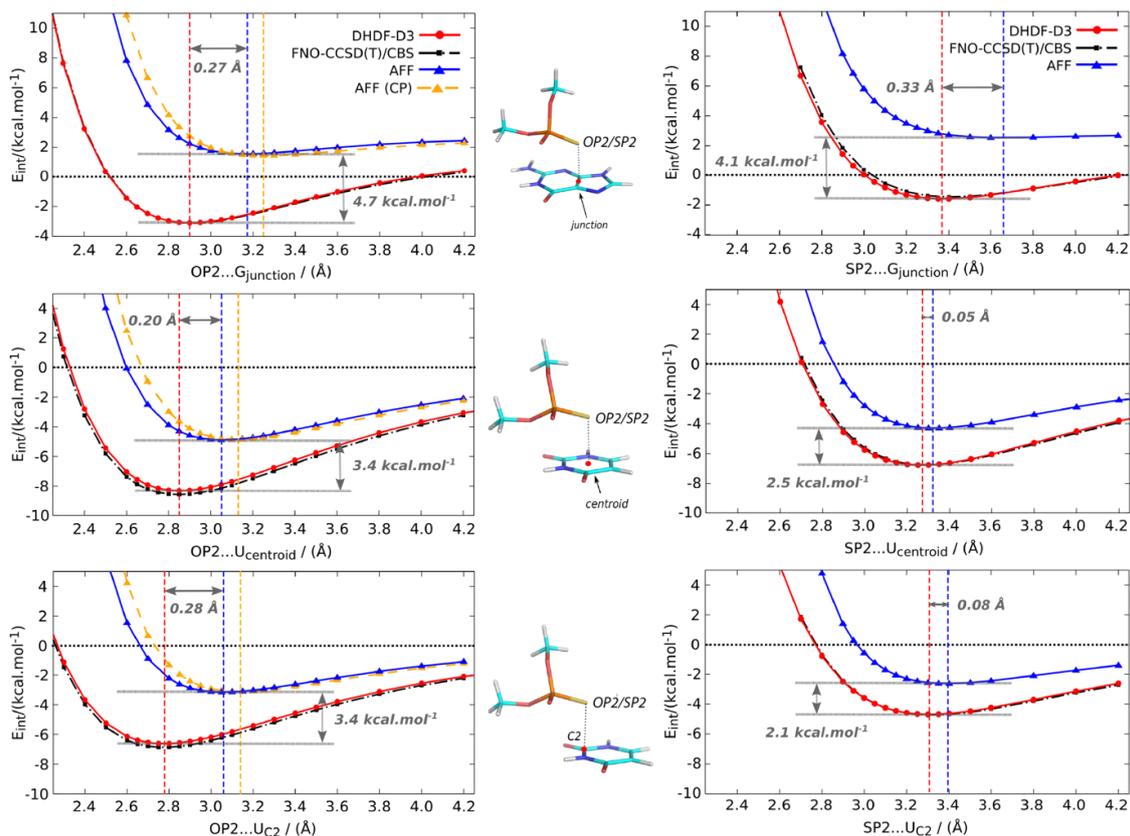


Figure 3. (Left) Vertical interaction energy scans for OP2...G_{junction} (top), OP2...U_{centroid} (middle) and U_{C2} (bottom) using DSD-BLYP-D3(BJ)/ma-def2-QZVPP (DHDF-D3), AMBER force field (AFF) and the QM reference FNO-CCSD(T)/CBS methods. AFF(CP) uses the “Case phosphates”⁸⁸ (Table S2). Minima positions are marked by vertical dashed lines (for QM reference the minimum is identical to DHDF-D3) and listed in Table 1. (Middle) Dimethyl phosphate (DMP) and dimethyl thiophosphate (DMTP) models used for the vertical interaction energy scans with OP2/SP2 atoms stacking with G/U nucleobases. Note the different energy ranges used for guanine and uracil scans. (Right) Same as Left but with the (DMP)OP2 atom replaced by the (DMTP)SP2 atom.

To gather experimental reference points, we calculated OP2...guanine/uracil (OP2...G/U) distances perpendicular to nucleobase plane and projection points of OP2 on G/U for the U-turn ensembles composed of 155 hits of r(nGNRAn) and 104 hits of r(UNNN). In r(nGNRAn) tetraloop hairpins, the (nt3)OP2 atom stacks with the region close to the junction of the guanine pyrimidine and imidazole rings (C4, C5 atoms) with a ≈ 3.05 Å average OP2...G distance (Figure 2). For r(UNNN) U-turns, the OP2 atom stacks with the region between the ring centroid and the C2 atom with a ≈ 3.0 Å average OP2...U distance. We chose to use three idealized positions, G_{junction} for guanine and U_{centroid}, U_{C2} for uracil, in further calculations. These points are not located at the experimental average of the OP2 positions but represent a reasonable compromise for vertical energy scan calculations.

DHDF-D3 calculations for OP2...guanine/uracil contacts

We first calculated DSD-BLYP-D3(BJ)/def2-QZVPP (DHDF-D3) interaction energy scans in gas phase for OP2...G and OP2...U contacts at idealized positions derived from X-ray/cryo-EM structures (Figure 2). We validated DHDF-D3 by FNO-CCSD(T)/CBS, which is the gold standard QM reference for non-covalent interactions (Figure 3).¹¹¹ The DHDF-D3 and FNO-CCSD(T)/CBS curves overlap almost perfectly, assessing that the DHDF-D3 method is appropriate for these calculations as also demonstrated earlier for lp... π contacts.³⁴

The OP2...G_{junction} interaction energy is attractive with a minimum of -3.1 kcal.mol⁻¹ (Table 1). Yet, the OP2...U_{centroid} (-8.3 kcal.mol⁻¹) and OP2...U_{C2} (-6.6 kcal.mol⁻¹) minima are significantly deeper compared to that of guanine. Minimum interaction energy distances (*minIEd*) are also slightly shorter

Table 1. Minimum interaction energy distance (*minIEd*) and interaction energy (*IE*) values for OP2...G/U and SP2...G/U model systems. These values are taken from the G_{junction} , U_{centroid} and U_{C_2} vertical scans (Figure 3). The *minIEd* and *IE* values are in Å and kcal.mol⁻¹, respectively.

	DHDF-D3		AFF		AFF (CP) ^a	
	<i>minIEd</i>	<i>IE</i>	<i>minIEd</i>	<i>IE</i>	<i>minIEd</i>	<i>IE</i>
OP2/SP2						
G_{junction}	2.90/3.37	-3.1/-1.6	3.17/3.66	+1.6/+2.6	3.25	+1.4
U_{centroid}	2.85/3.27	-8.3/-6.8	3.05/3.66	-4.9/-4.3	3.13	-4.9
U_{C_2}	2.78/3.31	-6.6/-4.7	3.06/3.39	-3.2/-2.6	3.14	-3.1
Average^b	+0.5	+1.7	+0.5	+0.7		

^aWith the slightly enlarged CP Lennard Jones parameters from Steinbrecher, Case et al.⁸⁸

^bAverage value for SP2 minus OP2

for U_{centroid} (2.85 Å) and U_{C_2} (2.78 Å) compared to G_{junction} (2.90 Å). The QM optimal contact distances are shorter than those derived from the statistical survey of PDB structures (≈ 3.0 Å) described above (Figure 3).

To complement these data, we explored the OP2...G/U *minIEd* and *IE* surfaces (see Supporting Information). The surfaces show that *minIEd* for OP2...G is 2.72 Å (Figure S4) and therefore slightly shorter than the *minIEd* value shown in Table 1 suggesting that the experimental position may not exactly correspond to the lowest *minIEd* position.

Like for single interaction energy curves (Figure 3), *IE* surfaces show more stabilizing energies for uracil vs. guanine (Figure S10), which is similar to early results by Egli and Sarkhel¹ who reported a stronger interaction energy for OP2...U compared to OP2...G. However, these authors also reported a destabilizing OP2...G interaction at the DFT/6-31G* level. We hypothesize that their results are biased by an inadequate level of theory neglecting London dispersion and using a too small basis set.¹¹² An update of these QM data is not provided in a recent investigation on a r(GAGA) system.⁹ This justifies present investigations using more accurate theory levels. Recent calculations for r(GAAA) and r(GGAG) hairpins found values in the -3.1 to -4.0 kcal.mol⁻¹ range for DMP...G models extracted from X-ray structures that are in overall agreement with the Table 1 data.¹⁰

SAPT energy decomposition shows that induction significantly contributes to the stabilization of the DMP...G/U systems

As shown by SAPT analysis, the induction energy significantly contributes to the stabilization of the OP2...G/U contacts and its relative importance increases with distance (Figure 4). In other words, at the studied distance range, the strength of each stabilizing interaction energy term decreases with increasing intermonomer distance, yet the induction decays more slowly than London dispersion and electrostatics and thus becomes the most important stabilizing contribution at longer distances (Figure 4b). This is an issue for classical force-fields such as AFF that do not take the induction/polarization term into account as discussed below. In the following, we will use the equivalent induction or polarization terms when discussing QM or AFF data (see Methods: SAPT calculations). Note that above 2.8 Å, the electrostatic component becomes positive for OP2... G_{junction} leading to less stabilizing total interaction energies while below 2.8 Å, it is attractive, likely due to charge-penetration effects.¹¹³ The electrostatic term contributes to the large calculated difference between OP2...G and OP2...U.

AFF energy curves are shifted to higher and even destabilizing energies when compared to DHDF-D3

A principal aim of this study is to evaluate how well AFF agree with QM calculations. Surprisingly, the curves obtained with the AFF are systematically shifted to higher interaction energies by more than +3 kcal.mol⁻¹ (Figure 3 and Table 1). For guanine, the AFF interaction energy curve becomes even destabilizing with a minimum at +1.6 kcal.mol⁻¹. For uracil, the AFF interaction energies are similarly shifted (+3.4 kcal.mol⁻¹) when compared to DHDF-D3. However, the AFF interaction energy is attractive

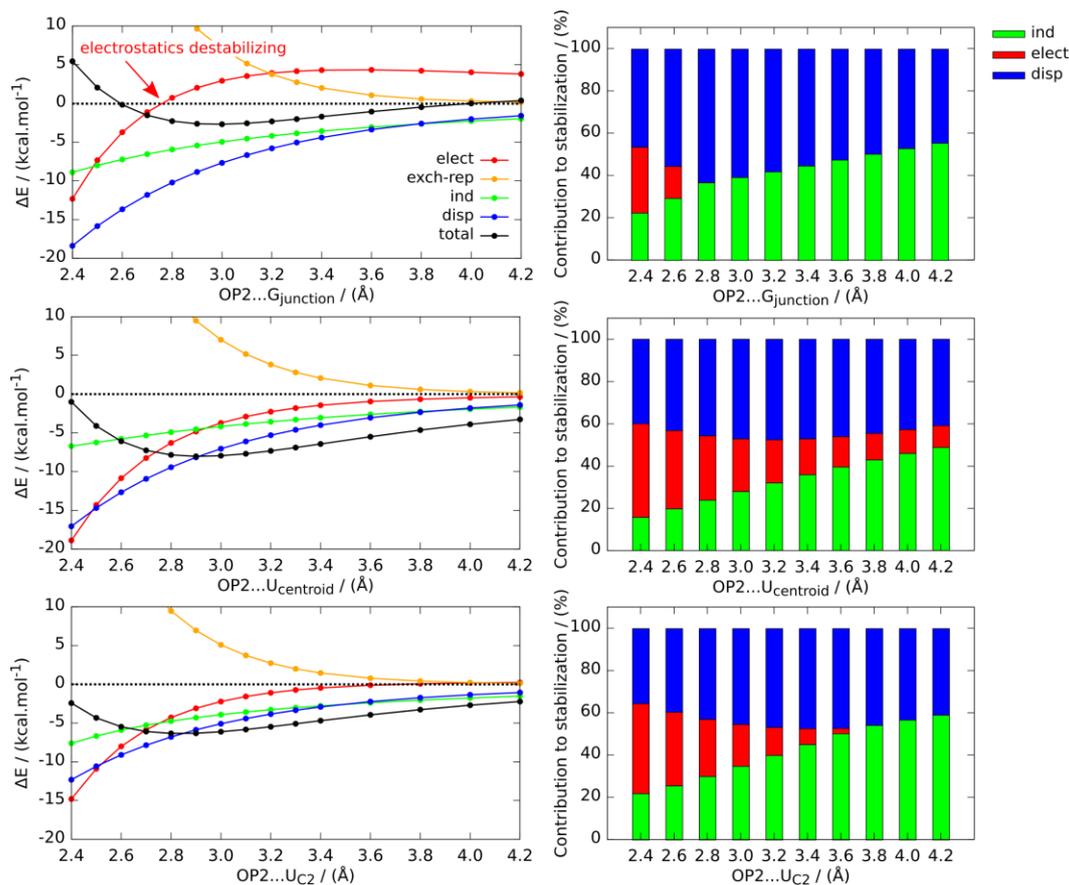


Figure 4. SAPT analysis (SAPT2+(3) δ MP2/jun-cc-pVTZ) for OP2...G_{junction} (top), OP2...U_{centroid} (middle) and U_{C2} atom (bottom). **(Left)** SAPT interaction energy terms electrostatics (elect), exchange-repulsion (exch-rep), induction/polarization (ind) and London dispersion (disp) for vertical scans. **(Right)** Contribution of the stabilizing terms (electrostatics, induction and London dispersion) to the stabilizing energy at specific phosphate... π distances showing the importance of specific interaction energy terms in the stabilization of the system. Note that for some distance ranges the electrostatics is positive, i.e. destabilizing. Destabilizing contributions are not included in the graph.

at the U_{centroid} (−4.9 kcal.mol^{−1}) and the U_{C2} (−3.2 kcal.mol^{−1}) atom positions (Figure 3 and Table 1). This strong deviation between QM and AFF might affect to an unknown extent the balance between folded and unfolded states of r(GNRA) hairpins while having possibly a smaller impact on r(UNNN) U-turns for which the interaction energies remain attractive.

The AFF phosphate... π minimum contact distances (*minIEd*) are longer when compared to DHDF-D3

The inter-monomer distance perpendicular to the aromatic plane of the nucleobase that corresponds to the minimum interaction energy distance (*minIEd*) is an important descriptor of the phosphate... π interaction. For OP2...G_{junction}, the AFF *minIEd* is longer by +0.27 Å when compared to DHDF-D3 (Figure 3 and Table 1). Similarly, over the U_{centroid}/U_{C2} positions the AFF *minIEd*s are longer by +0.20/+0.28 Å.

At short-range (< 2.8 Å), AFF is too repulsive due to the very steep r^{-12} Lennard-Jones (LJ) potential term, as already reported for Ip... π contacts.³⁴ This may lead to large spurious forces between atomic pairs during MD simulations when short distances are sampled or to a decreased sampling in this distance range. In addition, the shift of the AFF (3.05–3.17 Å; Table 1) compared to the DHDF-D3 minima (2.8–2.9 Å) leads to a sampling of conformations with extended OP2... π distances facilitated by the relatively flat AFF potential profile. Both above-mentioned effects are expected to be at the origin of the lengthened MD simulation average distances discussed below.

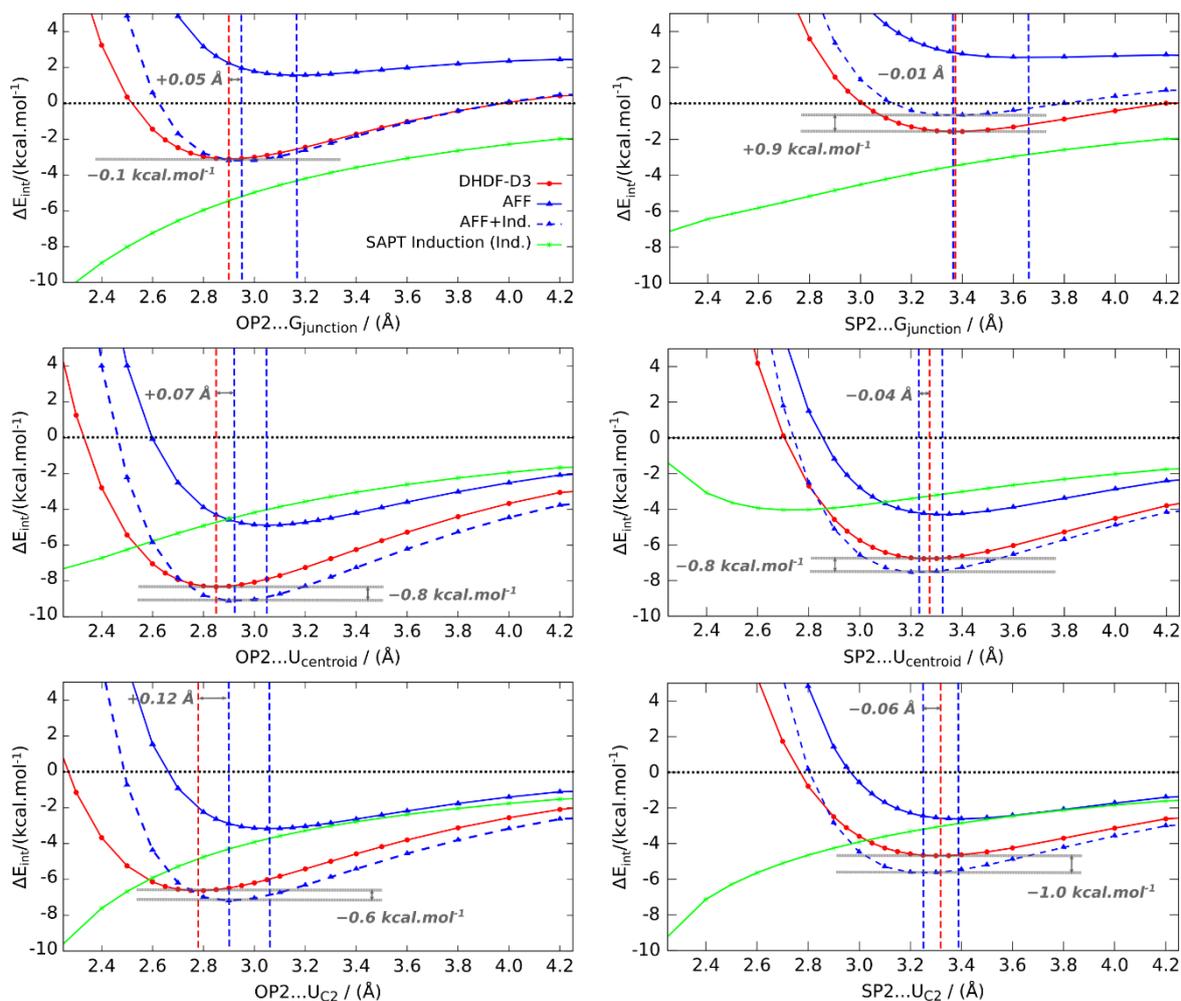


Figure 5. DHDF-D3 and AFF vertical interaction energy scans and AFF+Ind curves for OP2/SP2...G_{junction} (top), OP2/SP2...U_{centroid} (middle) and OP2/SP2...U_{c2} (bottom). AFF+Ind (blue dashed curves) corresponds to AFF with added SAPT induction/polarization term (corresponding to polarization in AFF; green curve); see also (Figure 4). The AFF+Ind curves serve for investigation purposes only and not as corrections for AFF.

The AFF overestimates *minIE*s on whole surfaces by 0.2-0.5 Å and 0.1-0.3 Å for guanine and uracil, respectively (Figure S4 and S5). For a detailed discussion on *minIE*d surfaces see Section 1 in Supporting Information.

“CP” modified Lennard-Jones parameters slightly worsen AFF phosphate... π description

The alternative “CP” LJ parameters for phosphates⁸⁸ with “enlarged” oxygen atoms were reported to improve the folding behavior of a r(GAGA) tetraloop in MD simulations due to weakening of the base...phosphate hydrogen bonds (BPh) that overstabilize unfolded conformations.²⁴ Since these modifications are used for MD simulations of r(GNRA) tetraloop hairpins, we evaluated their effect on the phosphate... π contacts.^{19,31} We found that they deepen the AFF vs. DHDF-D3 difference for minimum-energy distances by +0.08 Å (Figure 3 and Table 1) while barely modifying interaction energies at minima. Thus, we conclude that these differences are non-significant in the present context and will not significantly affect phosphate... π contacts in MD simulation trajectories.

Adding the SAPT induction term to the AFF interaction energy improves the agreement between DHDF-D3 and AFF methods

It is well appreciated that classical non-additive force fields do not capture polarization effects, although it has been advocated that errors in the force fields tend to cancel-out at equilibrium

distance.¹¹⁴ This is not the case for the phosphate... π interactions investigated in this study that reveal a large discrepancy between DHDF-D3 and AFF interaction energy curves (**Figure 3**). Our SAPT calculations show that the induction or polarization term contributes significantly to the stabilization of these systems (**Figure 4b**). Thus, to check if adding polarization to the AFF energy curves improves the agreement of AFF with QM, we simply added the SAPT induction term to the AFF curves. The newly generated “AFF+Ind” curves show a better agreement with the DHDF-D3 ones over all distances (**Figure 5**), the largest discrepancies being observed for the short-range distances, below the minima.

A significant improvement is also observed for all the AFF+Ind distance minima that are now closer to the DHDF-D3 ones (**Table 1** and **Figure 5**). For AFF+Ind, the OP2...U_{C2} distance becomes smaller than the OP2...U_{centroid} in accordance with DHDF-D3. This is because induction/polarization is more important above the C2 atom compared to centroid and thus the addition of the SAPT induction term has a larger effect at this position (**Figure 5**). This topic is also discussed in **Section 1.4** of the **Supporting Information**.

Implicit solvent effects on phosphate... π interactions

Gas phase single-point energy calculations on small models are insufficient for a complete description of intermolecular interactions as they appear in biomolecular systems. Among others, they neglect dielectric screening effects brought by the solvent and the rest of the biomolecule. Gas phase calculations of charged non-covalently bound systems are known to overestimate the importance of electrostatics and polarization.^{115,116} While this also affects neutral molecules, the impact is much greater for charged systems.

To evaluate dielectric screening effects, we analyzed the OP2...G_{junction} contact with two implicit solvent models, i.e. the conductor-like screening model (COSMO)⁵⁹ for DHDF-D3 and the generalized Born (GB)^{92,93} for AFF (**Figure S16**). We consider biomolecular environment conditions with $\epsilon = 4.0$ more relevant than the solvent conditions with $\epsilon \approx 80.0$ because neither the phosphate nor the nucleobase forming the phosphate... π contact is fully exposed to the solvent. As expected, interaction energy differences between DHDF-D3/COSMO and AFF/GB are reduced compared to gas phase calculations. At the average experimental distance (3.05 Å; **Figure 2**) the AFF/GB interaction energy is overestimated by +1.0 kcal.mol⁻¹ (“biomolecule”) and 1.2 kcal.mol⁻¹ (“water”), which is much smaller than the +4.6 kcal.mol⁻¹ gas phase value at the same distance. However, the principal AFF issues like interaction energy differences or exaggerated short-range repulsion persists for implicit solvent calculations.

The COSMO model considers only dielectric effects and as such it is a crude approximation of true solvent effects. The DHDF-D3/COSMO curve for water fails to show a minimum around 3.0 Å. Thus, we chose gas phase calculations as the main tool to uncover AFF limitations although the entire neglect of solvent effects likely overrates the AFF vs. DHDF-D3 energy differences.

r(GNRA) tetraloop geometry optimizations with AFF show longer OP2...guanine distances when compared to equivalent QM/MM calculations

In a further tentative to include solvent effects, we investigated phosphate... π contacts in GNRA tetraloops surrounded by explicit solvent using a hybrid QM/MM approach where the whole RNA molecule is in the QM region (PBEh-3c DFT method). The RNA-solvent interactions are handled through electrostatic embedding with point-charge approximation, which offers charge screening and explicit treatment of RNA-solvent hydrogen bonding at the MM level.

In principle, comparison of QM/MM and AFF (i.e., MM) optimizations should reveal systematic differences and non-equivalencies of the QM and AFF potential energy surfaces (PES's). However, the RNA tetraloop PES's contain many local energy minima¹¹⁷ necessitating a huge number of calculations to characterize the PES through QM/MM optimizations.^{118,119} To limit this issue, we optimized QM/MM and AFF starting geometries derived from MD simulation snapshots with various OP2...G distances (from 2.79 to 3.55 Å). The r(GNRA) tetraloops reached different microstates after QM/MM and AFF geometry optimizations which hampered a direct QM/MM and AFF comparison. Thus, as described elsewhere, we re-optimized the QM/MM-optimized geometries with AFF (QM/MM-to-AFF

Table 2. Comparison of OP2...G distances (in Å) of GNRA tetraloops after QM/MM, AFF (MM) geometry optimizations and QM/MM-to-AFF re-optimizations.

	START (MD snapshot)	QM/MM	AFF	QM/MM-to-AFF	$d_{(QM/MM-to-AFF)-QM/MM}$
GAAA-s1	3.68	3.83	3.81	3.87	+0.04
GAAA-s2	2.79	2.95	3.03	3.08	+0.13
GAAA-s3	3.21	2.95	3.09	3.03	+0.08
GAGA-s1	3.41	3.74	3.51	3.80	+0.07
GAGA-s2	3.09	3.13	3.18	3.25	+0.12
GAGA-s3	2.90	2.91	2.94	2.99	+0.08
GCAA-s1	3.28	3.07	3.17	3.20	+0.13
GCAA-s2	3.16	3.02	3.09	3.12	+0.10
GCAA-s3	3.00	2.99	3.01	3.10	+0.11
GCGA-s1	2.97	3.07	3.06	3.14	+0.07
GCGA-s2	3.15	3.30	3.37	3.38	+0.08
GCGA-s3	3.55	3.27	3.16	3.35	+0.08
average	3.22	3.19	3.20	3.28	+0.09

computations).¹²⁰ These optimizations resulted in reasonably equivalent QM/MM and AFF-optimized geometries allowing a rough comparison between QM/MM and AFF PES's. After QM/MM optimizations, OP2...G distances ranged between 2.91 and 3.83 Å. We observed a systematic elongation of the OP2...G distances after the QM/MM-to-AFF re-optimizations (**Table 2**). The same result was obtained for a few QM/MM geometry optimizations with the B3LYP-D3/def2-TZVP method that is more accurate than the PBEh-3c one. This confirms that imbalances between AFF and QM descriptions for small gas phase models are transferred to the fully solvated RNA system. Additional information is in **Section 4** of the **Supporting Information**.

r(GNRA) phosphate... π contacts in MD simulations vs. X-ray/cryo-EM structures

To further evaluate the AFF performance in structural description of phosphate... π contacts, we performed MD simulations of the four r(gcGAAAgc), r(gcGAGAgc), r(gcGCAAgc) and r(gcGCGAgc)

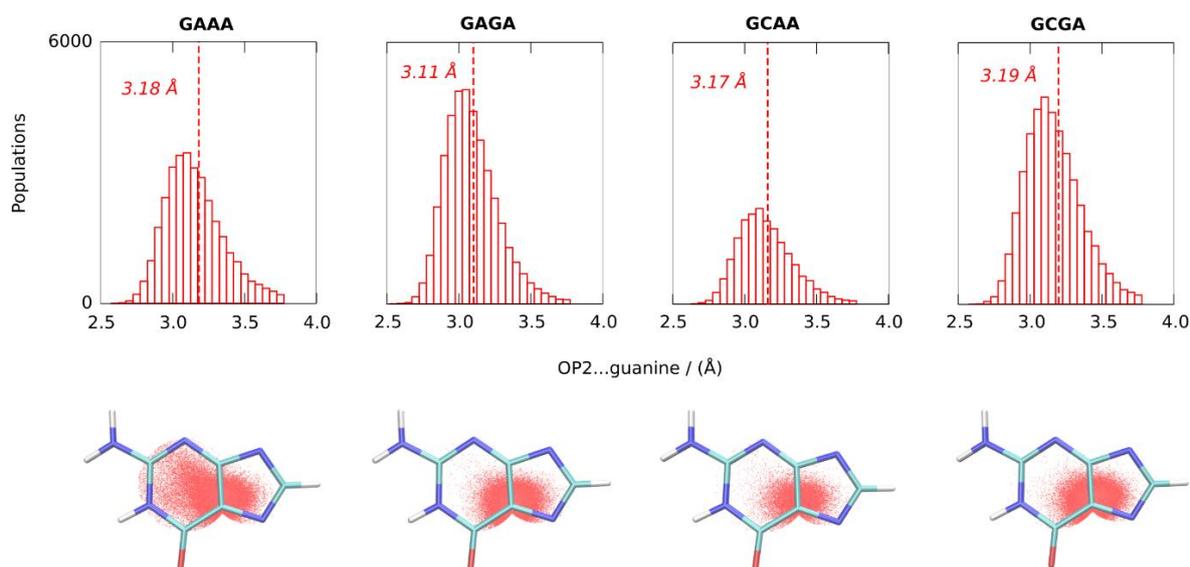


Figure 6. OP2...guanine distance histograms (top) and projection points of OP2 atoms on the guanine nucleobase plane (bottom) derived from MD simulations of r(gcGAAAgc), r(gcGAGAgc), r(gcGCAAgc) and r(gcGCGAgc) tetraloops. The vertical lines mark average distance values. For time-dependence of OP2...G distances in MD trajectories see **Figure S17**. See **Methods** for criteria used to select the analyzed snapshots.

tetraloops. The average OP2...G distances (Figure 6) are overestimated by 0.06 to 0.14 Å in MD simulations when compared to the ≈ 3.05 Å experimental value derived from an r(nGNRA) ensemble of X-ray/cryo-EM structures (Figure 2). Yet, the MD data agree well with the AFF results obtained for the DMP...G_{junction} model (3.17 Å; Table 1). However, as expected these values are longer than the 2.90 Å DHDF-D3 value.

MD ensembles reveal that the OP2 atoms are mostly located above the junction of guanine rings close to the C5 atom (Figure 6). In this region, the AFF interaction energy for the phosphate... π contact is destabilizing (Figure 3 and S10). This suggests that other forces in the r(GNRA) systems counterbalance these destabilizing interactions and establish a dynamical equilibrium. Yet, the shift to the C5 atom might be an AFF artifact since in experimental structures the OP2 atom is closer to the C4 atom (Figure 2). The source of this deviation remains unclear.

MD simulations of r(GNRA) tetraloops explore also non-native or partially unfolded conformations. In one of these conformations, the (nt3)OP2 atom is flipped and the phosphate... π contact is lost (Figure S17 and S18). Similar non-native substates could not be identified in GNRA tetraloops of X-ray/cryo-EM structures with resolution ≤ 2.5 Å. Thus, we suspect that these conformations are an AFF artifact. These substates were excluded from the OP2...guanine distance analysis shown in Figure 6 but are shown in Figure S18.

Similarities and differences between phosphate... π and lp... π contacts: a QM perspective

As mentioned in the Introduction, phosphate... π and lp... π are oxygen... π contacts that are recurrent in certain nucleic acid motifs. Both are weak to moderately attractive non-covalent interactions driven mostly by London dispersion (Figure 3 and refs. ^{34,41}). They occur in two of the most important and most stable RNA tetraloop folds, namely r(GNRA) and r(UNCG) tetraloops, and are usually placed in two different categories, namely anion... π for the U-turn phosphate... π contact and lp... π in Z-steps.⁶ In this section, we address their similarities and differences by focusing on the OP2/O4'...G_{junction} contact.

DHDF-D3 interaction energy curves show that for the OP2...G_{junction} the interaction energy differs only by -0.7 kcalmol⁻¹ from that of O4'...G_{junction} while *minIEd* is smaller by -0.23 Å (Figure 7).

Due to the presence of a negative charge on the phosphate group, the SAPT electrostatic and induction components differ between phosphate... π and lp... π while exchange-repulsion and London dispersion are comparable (Figure 8 and refs. ^{34,35}). Yet, the “*electrostatics + induction*” sum results in very similar potentials (Figure 9). The same is true for the *effective dispersion* (“*London dispersion + exchange-repulsion*” that is analogue to the classical LJ potential). Figure 9 suggests that differences between OP2...G_{junction} and O4'...G_{junction} in short-range regions stem from the “*effective dispersion*” term while above 3.2 Å they are related to the “*electrostatics + induction*” term.

The similar “*electrostatics + induction*” component for both OP2...G_{junction} and O4'...G_{junction} contacts indicates that the negative charge of the phosphate group does *not* play a key role in the stabilization of these phosphate... π contacts. However, given the negative charge of the phosphate group we wondered about the magnitude of an eventual charge-transfer component. We calculated that the phosphate... π charge-transfer part of the SAPT induction term is small (≈ 0.4 kcal·mol⁻¹ representing 3.5 % of stabilization energy for OP2...G_{junction} at 3.0 Å; see **Section 2** in the **Supporting Information**). This means that the SAPT induction term correspond mainly to polarization. Similarly negligible charge-transfer contributions were calculated for lp... π interactions.^{34,35} In addition, we evaluated the n-> π^* charge-transfer using COVPs (complementary occupied-virtual orbital pairs)¹²¹ based on the ALMO-EDA¹²² energy decomposition scheme. We found that the charge-transfer and therefore the n-> π^* character of the interaction is negligible (**section 4** in the **Supporting Information**). Thus, our calculations do not support the assumption that OP2...G interactions as observed in r(GNRA) U-turns are strongly influenced by the charge on the phosphate group and that there is any significant n-> π^* orbital donation effect. Thus, in U-turns, the “anion... π ” naming might be misleading because the anionic nature of the phosphate group does not play a major role in the OP2... π interaction and in the stabilization of the U-turn motifs. This is further supported by a recent X-ray structure (PDBid: 7JJF) showing that the substitution of the stacked anionic phosphate by a neutral methylphosphonate (MePO₃) does not perturb the loop structure, i.e., the methyl group occupies the position of the OP2 atom and leads to an “Me... π ” contact.⁹ However, we do not exclude that the anion... π terminology might be meaningful for chemical systems in aprotic solvents or in very specific biological environments and with anions with a more localized charge.^{123,124}

This parallels lp... π contacts that similarly involve a short oxygen... π stacking and for which no significant n-> π^* charge-transfer contacts could be found.³⁵ Therefore, in r(GNRA) as well as r(UNCG) and CpG Z-step motifs, the short contact is more shaped by the structural context that can lead through compression effects to significantly shortened distances than by strong intramolecular forces involving the lone pairs carried by the oxygen atom.

Similarities and differences between phosphate... π and lp... π contacts: an AFF perspective

From an AFF perspective, we observed that LJ and electrostatic potentials both substantially contribute to non-bonded interaction in phosphate... π systems (**Figure S9**) whereas for lp... π the electrostatic term contributes only marginally.³⁴ While for phosphate... π the AFF interaction energies are notably underestimated due to the lack of polarization in the force-field, they appear closer to DHDF-D3 for lp... π likely because the polarization contribution is much smaller in lp... π than

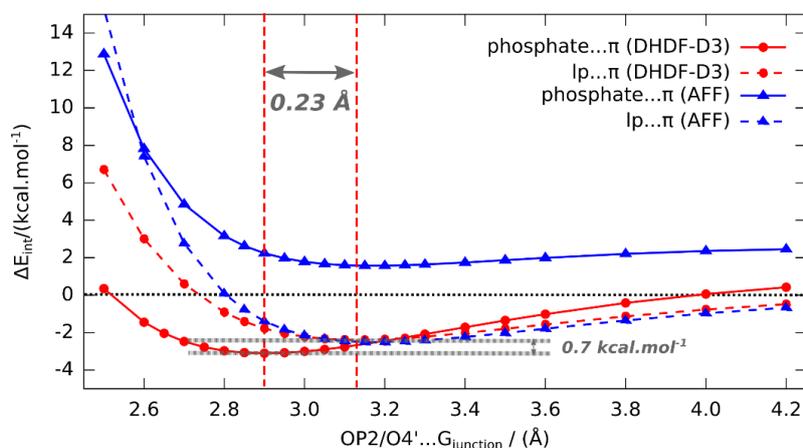


Figure 7. DHDF-D3 vs. AFF interaction energy curves for phosphate... π and lp... π contacts at the position of oxygen atom (OP2 or O4') above the G_{junction}. Data for lp... π are taken from ref. 34.

phosphate... π (Figure 6 and ref. 35). Therefore, the error due to the absence of polarization is minimized in lp... π contacts.

Short-range repulsion associated with a $1/r^{12}$ Lennard-Jones component is a concern for all interactions. Yet, it is probably a more significant issue for Z-DNA and r(UNCG) tetraloop systems containing CpG Z-steps with lp... π contacts that appear to be vertically compressed by the overall conformational Z-step context.^{34,35} The compression effect can lead to O4'... π distances in the 2.8–2.9 Å range that are on average shorter than the OP2... π interactions observed in U-turns.

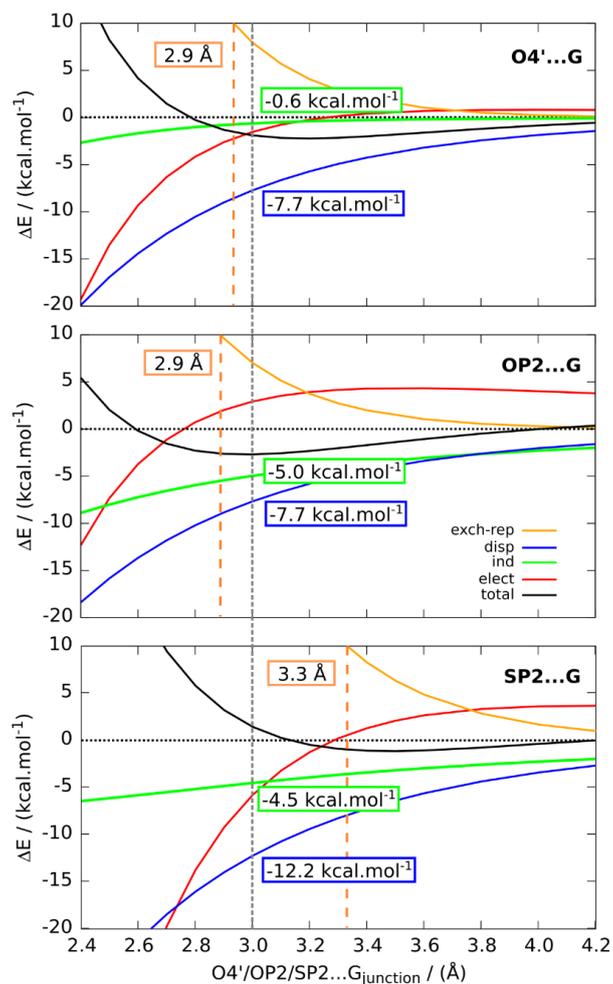


Figure 8. O4'/OP2/SP2... G_{junction} SAPT analysis. SAPT interaction energy terms electrostatics (elect), exchange-repulsion (exch-rep), induction/polarization (ind) and London dispersion (disp) for vertical scans are shown.

For phosphate... π contacts, the AFF interaction energy depends strongly on the position of OP2 over the nucleobase and can even be positive (Figure S10 and S11), in contrast with lp... π contacts where interaction energies are negative over all nucleobases. While the strongest interaction energy for lp... π contacts is $-4.2 \text{ kcal.mol}^{-1}$,³⁴ phosphate... π shows a wider range of values (from $+2.6$ to $-11.6 \text{ kcal.mol}^{-1}$ for G and U, respectively; Figure S10 and S11). This extended range is in part caused

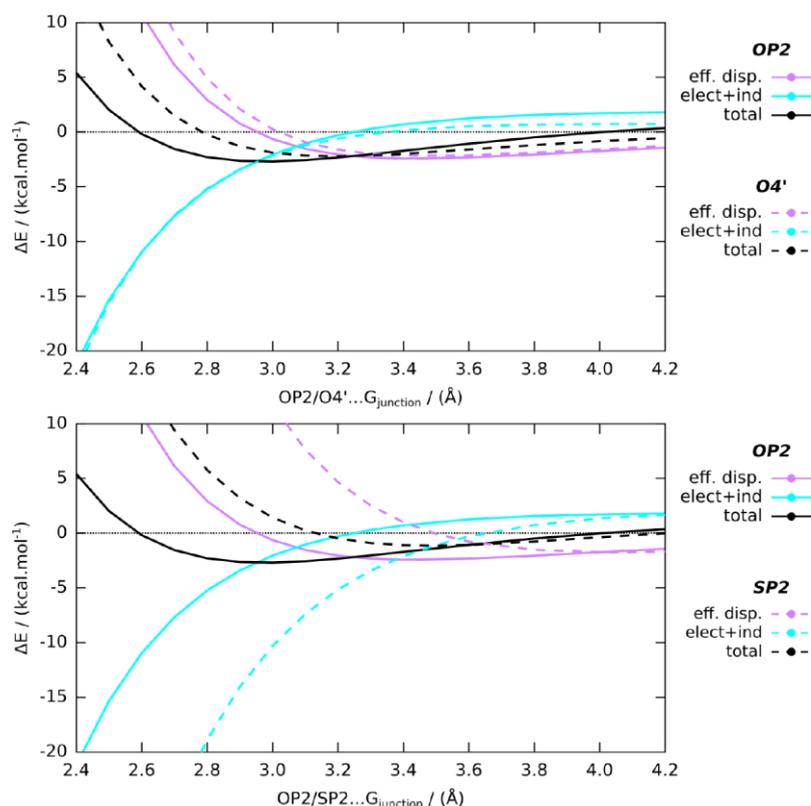


Figure 9. OP2/O4'...G_{junction} and OP2/SP2...G_{junction} SAPT analysis comparison. The SAPT interaction energy terms exchange-repulsion (exch-rep) and London dispersion (disp) from Figure 8 are summed up to a term called „effective dispersion“ (eff. disp.), a term analogue to the classical Lennard-Jones term. Electrostatics (elect) and induction (ind) are summed up into "elect+ind".

by the gas-phase treatment of the anionic system where unscreened electrostatics is exaggerated (see *Implicit solvent effects on phosphate... π interaction* and refs.^{115,116}).

As suggested earlier, AFF description of lp... π contacts could be improved by decreasing LJ radii for some of the nucleobase atoms and especially the strongly polarized carbons.³⁴ Such a modification would shift the short-range repulsive region to shorter distances, resulting in a better AFF description at the experimentally observed positions. However, the interaction energy gap between AFF and DHDF-D3 curves would be still present, as shown in Figure S14 where the AFF curves with decreased LJ R_{\min} values for the C4-C5 atoms are displayed. Based on our preliminary data, we suggest that a combined re-parametrization of nucleobase Lennard-Jones parameters along with a correction for the missing polarization are needed to improve the AFF description of phosphate... π -containing systems. Re-parameterization of LJ parameters is a complex problem although their revision is likely vital.³⁴ Yet, it might even be necessary to replace the LJ term by a more accurate van der Waals model starting by the introduction of a Buckingham-like potential that would be more accurate at short-range distances. While the inclusion of explicit polarization is possible through a Drude oscillator model it comes with many challenges of its own.^{125,126} Protocols with a Drude model implemented in CHARMM have been recently applied on RNA hairpins including r(GAAA) tetraloops to generate 500 ns long MD simulations with promising results.⁸¹ However, the phosphate... π contacts were not described and the stability of the system on the microsecond timescale can currently not be robustly evaluated. It should also be noted that the AMBER force-field is currently less advanced in including polarization than CHARMM. As an *ad-hoc* solution for phosphate... π contacts in GNRA tetraloop simulations in AFF, a gHbfix-like biasing function¹⁹ stabilizing the interaction between an OP2 atom and the junction of guanine rings could be added to the AFF potential. It could improve the stabilization of phosphate... π contacts in GNRA tetraloops and subsequently prevent the occurrence of unfolded states of GNRA structures in MD simulations, which is currently overestimated.²⁰

To summarize this part, the AFF imbalance in $lp\cdots\pi$ contacts is dominantly found in the short-range repulsion region. It is affecting MD simulations because the overall structural context of the Z-steps compresses the $lp\cdots\pi$ contacts. On the other hand, the phosphate $\cdots\pi$ contacts are less compressed and AFF MD simulations are probably more affected by the absence of explicit polarization.

Thiophosphate $\cdots\pi$ (SP2 $\cdots\pi$) interaction energies are less stabilizing than phosphate $\cdots\pi$ (OP2 $\cdots\pi$)

Sulfur has a large 1.80 Å vdW radius, as estimated by Bondi, that is 0.25 Å larger than the 1.55 Å oxygen vdW radius.¹²⁷ Thus, it seems at first sight difficult to replace the stacked OP2 atom by an SP2 atom in a r(GNRA) environment characterized by already short OP2 $\cdots\pi$ contacts. It seems also difficult to rationalize the gain in stability of r(GNRA) hairpins through phosphorothioation observed experimentally on different systems^{9,43} unless the larger and more polarizable sulfur atom establishes stronger interactions with the nucleobase than the smaller oxygen atom. However, at the DHDF-D3 level, the DMTP \cdots G/U interaction energies are on average less stabilizing by +1.7 kcal.mol⁻¹ than those calculated for the DMP \cdots G/U models (Table 1 and Figure 3).

SAPT shows that induction contribution to stabilization of SP2 \cdots G stacking is by \approx 15 % smaller compared to OP2 \cdots G, which is in accordance with earlier results for SP2/OP2 \cdots U stacking.⁴¹ The sulfur atom has a larger polarizability than oxygen, leading to both an increased London dispersion and increased exchange-repulsion component in SAPT (Figure 8). However, both terms compensate each other as also observed in the OP2 \cdots G curve (Figure 9). Thus, it appears that QM interaction energies that suggest a lesser stabilization of the SP2 $\cdots\pi$ contact are unable on their own to rationalize the experimental Δ G data that show a greater stability of tetraloops or U-turn motifs upon phosphorothioation.^{9,43} These data stress the limits of *in vacuo* QM calculations. Indeed, the thermodynamic stability of the folded state is always determined with respect to the unfolded reference state in solution and single-point energy calculations do not provide any definitive insights into the effects of substitutions on unfolded states and of the respective phosphate and thiophosphate desolvation contributions to the Δ G of folding.¹²⁸ Thus, a suggestion by Pallan et al.⁹ about London dispersion being responsible for the increased thermodynamic stability of phosphorothioated r(GAGA) tetraloop appears oversimplified.

AFF+ind improves the description of thiophosphate $\cdots\pi$ contacts

Like for OP2 systems, AFF interaction energy curves for SP2 systems are shifted to higher energies when compared to DHDF-D3 curves and the SP2 \cdots G_{junction} interaction energies are even destabilizing over the entire distance range. Moreover, the SP2 \cdots G_{junction} interaction energy is underestimated by +4.1 kcal.mol⁻¹ at minimum while the *minIE*d is stretched by 0.33 Å (Table 1 and Figure 3). For SP2 \cdots U, the situation differs slightly. AFF *minIE*d's are only marginally stretched (by less than 0.10 Å) while interaction energies are less stabilizing (by +2.5 and +2.1 kcal.mol⁻¹ for SP2 \cdots U_{centroid} and SP2 \cdots U_{C2}, respectively; Table 1 and Figure 3). It also appears that the SP2 \cdots G_{junction} curve is almost flat above 3.4 Å such that a minimum can difficultly be deduced.

Although the SP2 $\cdots\pi$ *minIE*d and *IE* surfaces for AFF better agree with DHDF-D3 surfaces when compared to phosphate $\cdots\pi$ contacts (Figure S4, S5, S8 and S9), a comparable and systematic shift to larger distances and less stabilizing interaction energies is calculated. As demonstrated for OP2 \cdots G/U, this shift can be corrected to a significant extent by adding the SAPT induction term to generate "AFF+Ind" curves (Figure 5). This demonstrates that for phosphorothioated systems, polarization is also a key component that should be considered in further force-field developments and that the development of polarizable force-fields for RNA should be accelerated.^{79,126,129}

In the phosphorothioated neomycin-sensing riboswitch U-turn, contradictory demands on sulfur LJ parameters were highlighted.⁴¹ It was found that for a proper description of SP2 \cdots U stacking, the sulfur atom should be smaller, but it should be simultaneously larger for another U-turn signature interaction, the base-phosphate N3-H3 \cdots SP2 hydrogen bond (5BPh).⁸ It is likely a result of the large

anisotropy of the SP2 atom, which is not taken into consideration by the isotropic AMBER atom types.¹³⁰ Thus, it might be impossible with current methods to describe both interactions with one set of LJ parameters. This suggests the need to develop different atom types for atoms involved in hydrogen bonds and π -stacking.

Conclusion

We investigated phosphate... π contacts occurring in r(nGNRAn) hairpins and r(UNNN) U-turn motifs. Using the WebFR3D tool to search the PDB, we found that in U-turns, phosphate... π involve almost exclusively OP2...G/U contacts while OP1...G/U contacts are rare and probably of artefactual origin.

QM calculations showed that phosphate... π interactions are attractive with moderate to weak interaction energies (-8 to -3 kcal.mol⁻¹) and short contact distances (≈ 2.9 Å). SAPT analysis stressed the importance of the induction term in the stabilization of these contacts. We hypothesized that the absence of polarization (induction) in the classical AFF might be at the origin of significant discrepancies between DHDF-D3 and AFF data reflected by a systematic shift to less stabilizing interaction energies for OP2...U and even unrealistically destabilizing interaction energies for OP2...G. By adding the induction component calculated by SAPT analysis to the AFF interaction energy curves, we obtained "AFF+Ind" curves which show a significantly better agreement over all distances with the DHDF-D3 curves for the OP2...G/U but also for the phosphorothioated SP2...G/U systems.

We also found that the AFF OP2...G/U distances are systematically elongated by > 0.2 Å and that the addition of the SAPT induction component to AFF ("AFF+Ind") reduced this discrepancy to ≤ 0.1 Å. In addition to the lack of polarization, the AFF repulsive r^{-12} part of the LJ potential known to overestimate the short-range repulsion and with the probably too large Lennard-Jones (LJ) radii of nucleobase carbon atoms could play a role. Reducing carbon LJ R_{\min} values improves the agreement between the AFF and DHDF-D3 minimum energy distance as well as the short-range component although the exaggerated short-range repulsion is an inherent feature of the r^{-12} part of the LJ potential and is just shifted to shorter distances. Moreover, the large interaction energy gap, which is an important issue for phosphate... π contacts, remains intact and still needs adjustments. By using QM/MM and AFF geometry optimizations of the r(gcGNRAgc) tetraloop hairpin, we found that AFF distances are overestimated for all the tested structures, supporting the results from interaction energy scans. MD simulation results revealed the same trend in a shift towards longer contact distances when compared with X-ray/cryo-EM structures.

It is difficult to confidently attribute MD simulation issues for r(GNRA) hairpins to any of the above AFF imbalances. However, the under-stabilizing or even destabilizing interaction energy for the OP2...G/U system may cause local structural inaccuracies within the free-energy basin of the folded state and therefore affect the balance between folded and unfolded tetraloops.

Regarding thiophosphate... π contacts, we observed that SP2...G/U interaction energies are less stabilizing compared to OP2...G/U ones. We suggest that the stability boost of r(GNRA) tetraloops upon SP2 substitution may be the result of a combination of many factors such as solvation and is not entirely attributable to QM energy terms although these contributions are important to calculate. Thus, other factors than interaction energy components are likely responsible for the enhanced stabilization of phosphorothioated U-turn motifs reported in experimental studies.

Associated Content

Supporting Information

The *minIEd* and *IE* surfaces: procedures for generating the *minIEd* and *IE* surfaces, DHDF-D3 and AFF *minIEd* surfaces for OP2...nucleobase and SP2...G/U, DHDF-D3 and AFF *minIEd* surfaces with AFF electrostatics subtraction, regions with no AFF minima; charge transfer effects; additional information to QM/MM calculations (PDF)

Small models: optimized geometries of DM(T)P...nucleobase models in XYZ format, LJ parameters and partial charges used for *bff* calculations, geometries and DHDF-D3 and AFF interaction energies used for calculations of *minIEd* and *IE* surfaces in XYZ format; GNRA tetraloops: PDB files of starting and optimized geometries of GNRA tetraloops used for QM/MM and AFF calculations (ZIP)

Acknowledgement

This work was supported by the project SYMBIT reg. number CZ.02.1.01/0.0/0.0/15_003/0000477 financed by the ERDF (KM, HK, JS) and by Czech Science Foundation project number 20-16554S (VM, JS). KM and PA thank Craig Zirbel for providing the distance histograms and projection points derived from WebFR3D searches and for comments on the manuscript.

References

- (1) Egli, M.; Sarkhel, S. Lone Pair-Aromatic Interactions: To Stabilize or Not to Stabilize. *Acc. Chem. Res.* **2007**, *40* (3), 197–205. <https://doi.org/10.1021/ar068174u>.
- (2) Rather, I. A.; Wagay, S. A.; Ali, R. Emergence of Anion- π Interactions: The Land of Opportunity in Supramolecular Chemistry and Beyond. *Coord. Chem. Rev.* **2020**, *415*, 213327. <https://doi.org/10.1016/j.ccr.2020.213327>.
- (3) Jucker, F. M.; Pardi, A. GNRA Tetraloops Make a U-Turn. *RNA* **1995**, *1* (2), 219–222.
- (4) Gutell, R. R.; Cannone, J. J.; Konings, D.; Gautheret, D. Predicting U-Turns Ribosomal RNA with Comparative Sequence Analysis. *J. Mol. Biol.* **2000**, *300* (4), 791–803. <https://doi.org/10.1006/jmbi.2000.3900>.
- (5) Auffinger, P.; Westhof, E. An Extended Structural Signature for the tRNA Anticodon Loop. *Rna* **2001**, *7* (3), 334–341. <https://doi.org/10.1017/S1355838201002382>.
- (6) D’Ascenzo, L.; Leonarski, F.; Vicens, Q.; Auffinger, P. Revisiting GNRA and UNCG Folds: U-Turns versus Z-Turns in RNA Hairpin Loops. *RNA* **2017**, *23* (3), 259–269. <https://doi.org/10.1261/rna.059097.116>.
- (7) Gottstein-Schmidtke, S. R.; Duchardt-Ferner, E.; Groher, F.; Weigand, J. E.; Gottstein, D.; Suess, B.; Wöhnert, J. Building a Stable RNA U-Turn with a Protonated Cytidine. *RNA* **2014**, *20* (8), 1163–1172. <https://doi.org/10.1261/rna.043083.113>.
- (8) Zirbel, C. L.; Sponer, J. E.; Sponer, J.; Stombaugh, J.; Leontis, N. B. Classification and Energetics of the Base-Phosphate Interactions in RNA. *Nucleic Acids Res.* **2009**, *37* (15), 4898–4918. <https://doi.org/10.1093/nar/gkp468>.
- (9) Pallan, P. S.; Lybrand, T. P.; Schlegel, M. K.; Harp, J. M.; Jahns, H.; Manoharan, M.; Egli, M. Incorporating a Thiophosphate Modification into a Common RNA Tetraloop Motif Causes an Unanticipated Stability Boost. **2020**. <https://doi.org/10.1021/acs.biochem.0c00685>.
- (10) Esmaeeli, R.; Piña, M. D. L. N.; Frontera, A.; Pérez, A.; Bauzá, A. Importance of Anion- π Interactions in RNA GAAA and GGAG Tetraloops: A Combined MD and QM Study. *J. Chem. Theory Comput.* **2021**, *17* (10), 6624–6633. <https://doi.org/10.1021/acs.jctc.1c00756>.
- (11) Varani, G. Exceptionally Stable Nucleic Acid Hairpins. *Annu. Rev. Biophys. Biomol. Struct.* **1995**, *24*, 379–404. <https://doi.org/10.1146/annurev.bb.24.060195.002115>.
- (12) Chauhan, S.; Woodson, S. A. Tertiary Interactions Determine the Accuracy of RNA Folding. *J. Am. Chem. Soc.* **2008**, *130* (4), 1296–1303. <https://doi.org/10.1021/ja076166i>.
- (13) Fiore, J. L.; Nesbitt, D. J. An RNA Folding Motif: GNRA Tetraloop-Receptor Interactions. *Q. Rev. Biophys.* **2013**, *46* (3), 223–264. <https://doi.org/10.1017/S0033583513000048>.
- (14) Hall, K. B. Mighty Tiny. *Rna* **2015**, *21* (4), 630–631. <https://doi.org/10.1261/rna.050567.115>.
- (15) D’Ascenzo, L.; Vicens, Q.; Auffinger, P. Identification of Receptors for UNCG and GNRA Z-Turns and Their Occurrence in rRNA. *Nucleic Acids Res.* **2018**, *46* (15), 7989–7997. <https://doi.org/10.1093/nar/gky578>.
- (16) Sheehy, J. P.; Davis, A. R.; Znosko, B. M. Thermodynamic Characterization of Naturally Occurring RNA Tetraloops. *RNA* **2010**, *16* (2), 417–429. <https://doi.org/10.1261/rna.1773110>.
- (17) Sorin, E. J.; Engelhardt, M. A.; Herschlag, D.; Pande, V. S. RNA Simulations: Probing Hairpin Unfolding and the Dynamics of a GNRA Tetraloop. *J. Mol. Biol.* **2002**, *317* (4), 493–506.

<https://doi.org/10.1006/jmbi.2002.5447>.

- (18) Banas, P.; Hollas, D.; Zgarbova, M.; Jurecka, P.; Orozco, M.; Cheatham III, T. E.; Sponer, J.; Otyepka, M. Performance of Molecular Mechanics Force Fields for RNA Simulations: Stability of UUCG and GNRA Hairpins. *J. Chem. Theory Comput.* **2010**, *6* (12), 3836–3849. <https://doi.org/10.1021/ct100481h>.
- (19) Kührová, P.; Mlýnský, V.; Zgarbová, M.; Krepl, M.; Bussi, G.; Best, R. B.; Otyepka, M.; Šponer, J.; Banáš, 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. <https://doi.org/10.1021/acs.jctc.8b00955>.
- (20) Mlýnský, V.; Janeček, M.; Kührová, P.; Fröhlking, T.; Otyepka, M.; Bussi, G.; Banáš, P.; Šponer, J. Towards Convergence in Folding Simulations of RNA Tetraloops: Comparison of Enhanced Sampling Techniques and Effects of Force Field Corrections. *J. Chem. Theory Comput.* **2022**, 2021.11.30.470631. <https://doi.org/10.1021/acs.jctc.1c01222>.
- (21) Kuhrova, P.; Banas, P.; Best, R. B.; Sponer, J.; Otyepka, M. Computer Folding of RNA Tetraloops? Are We There Yet? *J. Chem. Theory Comput.* **2013**, *9* (4), 2115–2125. <https://doi.org/10.1021/ct301086z>.
- (22) Chen, A. A.; Garcia, A. E. High-Resolution Reversible Folding of Hyperstable RNA Tetraloops Using Molecular Dynamics Simulations. *Proc. Natl. Acad. Sci.* **2013**, *110* (42), 16820–16825. <https://doi.org/10.1073/pnas.1309392110>.
- (23) Haldar, S.; Kuhrova, P.; Banas, P.; Spiwok, V.; Sponer, J.; Hobza, P.; Otyepka, M. Insights into Stability and Folding of GNRA and UCG Tetraloops Revealed by Microsecond Molecular Dynamics and Well-Tempered Metadynamics. *J. Chem. Theory Comput.* **2015**, *11* (8), 3866–3877. <https://doi.org/10.1021/acs.jctc.5b00010>.
- (24) 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. <https://doi.org/10.1021/acs.jctc.6b00300>.
- (25) Miner, J. C.; Chen, A. A.; García, A. E. Free-Energy Landscape of a Hyperstable RNA Tetraloop. *Proc. Natl. Acad. Sci. U. S. A.* **2016**, *113* (24), 6665–6670. <https://doi.org/10.1073/pnas.1603154113>.
- (26) Bottaro, S.; Banas, P.; Sponer, J.; Bussi, G. Free Energy Landscape of GAGA and UUCG RNA Tetraloops. *J. Phys. Chem. Lett.* **2016**, *7* (20), 4032–4038. <https://doi.org/10.1021/acs.jpcllett.6b01905>.
- (27) Zgarbová, M.; Jurečka, P.; Banáš, P.; Havrila, M.; Šponer, J.; Otyepka, M. Noncanonical α/γ Backbone Conformations in RNA and the Accuracy of Their Description by the AMBER Force Field. *J. Phys. Chem. B* **2017**, *121* (11), 2420–2433. <https://doi.org/10.1021/acs.jpccb.7b00262>.
- (28) Sponer, J.; Bussi, G.; Krepl, M.; Banas, P.; Bottaro, S.; Cunha, R. A.; Gil-Ley, A.; Pinamonti, G.; Poblete, S.; Jurečka, P.; Walter, N. G.; Otyepka, M. RNA Structural Dynamics as Captured by Molecular Simulations: A Comprehensive Overview. *Chem. Rev.* **2018**, *118* (8), 4177–4338. <https://doi.org/10.1021/acs.chemrev.7b00427>.
- (29) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. A Second Generation Force Field for the Simulation of Proteins, Nucleic Acids, and Organic Molecules. *J. Am. Chem. Soc.* **1995**, *117* (19), 5179–5197. <https://doi.org/10.1021/ja00124a002>.
- (30) Zgarbová, M.; Otyepka, M.; Šponer, J.; Mládek, A.; Banáš, P.; Cheatham III, T. E.; Jurečka, 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. <https://doi.org/10.1021/ct200162x>.
- (31) Fröhlking, T.; Mlýnský, V.; Janeček, M.; Kührová, P.; Krepl, M.; Banáš, P.; Šponer, J.; Bussi, G. Automatic Learning of Hydrogen-Bond Fixes in an AMBER RNA Force Field. *arXiv* **2022**, 1–38. <https://doi.org/arXiv:2201.04078>.
- (32) D’Ascenzo, L.; Leonarski, F.; Vicens, Q.; Auffinger, P. “Z-DNA like” Fragments in RNA: A Recurring Structural Motif with Implications for Folding, RNA/Protein Recognition and Immune Response. *Nucleic Acids Res.* **2016**, *44* (12), 5944–5956. <https://doi.org/10.1093/nar/gkw388>.
- (33) Mráziková, K.; Mlýnský, V.; Kührová, P.; Pokorná, P.; Kruse, H.; Krepl, M.; Otyepka, M.; Banáš, P.; Šponer, J. UUCG RNA Tetraloop as a Formidable Force-Field Challenge for MD Simulations. *J. Chem. Theory Comput.* **2020**, *16*, 7601–7617.
- (34) Mráziková, K.; Šponer, J.; Mlýnský, V.; Auffinger, P.; Kruse, H. Short-Range Imbalances in the AMBER Lennard-Jones Potential for (Deoxy)Ribose···Nucleobase Lone-Pair··· π Contacts in Nucleic Acids. *J. Chem. Inf. Model.* **2021**, *61*, 5644–5657. <https://doi.org/10.1021/acs.jcim.1c01047>.
- (35) Kruse, H.; Mrazikova, K.; D’Ascenzo, L.; Sponer, J.; Auffinger, P. Short but Weak: The Z-DNA Lone-Pair··· π Conundrum Challenges Standard Carbon Van Der Waals Radii. *Angew. Chemie - Int. Ed.* **2020**, *59* (38), 16553–16560. <https://doi.org/10.1002/anie.202004201>.
- (36) Krepl, M.; Vögele, J.; Kruse, H.; Duchardt-Ferner, E.; Wöhnert, J.; Sponer, J. An Intricate Balance of Hydrogen Bonding, Ion Atmosphere and Dynamics Facilitates a Seamless Uracil to Cytosine Substitution in the U-Turn of the Neomycin-Sensing Riboswitch. *Nucleic Acids Res.* **2018**, *46* (13), 6528–6543. <https://doi.org/10.1093/nar/gky490>.
- (37) Madsen, M.; Gothelf, K. V. Chemistries for DNA Nanotechnology. *Chem. Rev.* **2019**. <https://doi.org/10.1021/acs.chemrev.8b00570>.
- (38) Hunsicker-Wang, L. M.; Vogt, M. J.; Hoogstraten, C. G.; Cospers, N. J.; Davenport, A. M.; Hendon, C. H.; Scott, R. A.; Britt, R. D.; DeRose, V. J. Spectroscopic Characterization of Mn²⁺ and Cd²⁺ Coordination to Phosphorothioates in the Conserved A9 Metal Site of the Hammerhead Ribozyme. *J. Inorg. Biochem.* **2022**, *230* (November 2021), 111754. <https://doi.org/10.1016/j.jinorgbio.2022.111754>.
- (39) Saran, R.; Huang, Z.; Liu, J. Phosphorothioate Nucleic Acids for Probing Metal Binding, Biosensing and Nanotechnology. *Coord. Chem. Rev.* **2021**, *428*, 213624. <https://doi.org/10.1016/j.ccr.2020.213624>.
- (40) Hyjek-Składanowska, M.; Vickers, T. A.; Napiórkowska, A.; Anderson, B. A.; Tanowitz, M.; Crooke, S. T.; Liang, X. H.; Seth, P. P.; Nowotny, M. Origins of the Increased Affinity of Phosphorothioate-Modified Therapeutic Nucleic Acids for Proteins. *J. Am. Chem. Soc.* **2020**, *142* (16), 7456–7468. <https://doi.org/10.1021/jacs.9b13524>.
- (41) Zhang, Z.; Vögele, J.; Mráziková, K.; Kruse, H.; Cang, X.; Wöhnert, J.; Krepl, M.; Šponer, J. Phosphorothioate Substitutions in RNA Structure Studied by Molecular Dynamics Simulations, QM/MM Calculations, and NMR Experiments. *J. Phys. Chem. B* **2021**, *125* (3), 825–840. <https://doi.org/10.1021/acs.jpccb.0c10192>.
- (42) Zheng, Y. Y.; Wu, Y.; Begley, T. J.; Sheng, J. Sulfur Modification in Natural RNA and Therapeutic Oligonucleotides. *RSC Chem. Biol.* **2021**, *2* (4), 990–1003. <https://doi.org/10.1039/d1cb00038a>.
- (43) Horton, T. E.; Maderia, M.; DeRose, V. J. Impact of Phosphorothioate Substitutions on the

- Thermodynamic Stability of an RNA GAAA Tetraloop: An Unexpected Stabilization. *Biochemistry* **2000**, 39 (28), 8201–8207. <https://doi.org/10.1021/bi000141d>.
- (44) Sarver, M.; Zirbel, C. L.; Stombaugh, J.; Mokdad, A.; Leontis, N. B. FR3D: Finding Local and Composite Recurrent Structural Motifs in RNA 3D Structures. *J. Math. Biol.* **2008**, 56 (1–2), 215–252. <https://doi.org/10.1007/s00285-007-0110-x>.
- (45) Petrov, A. I.; Zirbel, C. L.; Leontis, N. B. WebFR3D - A Server for Finding, Aligning and Analyzing Recurrent RNA 3D Motifs. *Nucleic Acids Res.* **2011**, 39 (SUPPL. 2), 50–55. <https://doi.org/10.1093/nar/gkr249>.
- (46) Leontis, N. B.; Zirbel, C. L. Nonredundant 3D Structure Datasets for RNA Knowledge Extraction and Benchmarking. *RNA 3D Struct. Anal. Predict.* **2012**, 27, 319–334. https://doi.org/10.1007/978-3-642-25740-7_13.
- (47) Zirbel, C. L.; Auffinger, P. Lone Pair...pi Contacts and Structure Signatures of r(UNCG) Tetraloops, Z-Turns and Z-Steps: A WebFR3D Survey. **2022**.
- (48) Ennifar, E.; Micura, R.; Kosutic, M. 2'-Trifluoromethylthio-Modified RNA: Synthesis, X-Ray Structure, Thermodynamic Analysis, and Applications. <https://doi.org/10.2210/PDB4NLF/PDB>.
- (49) Jmol: An Open-Source Java Viewer for Chemical Structures in 3D.
- (50) Becke, A. D. Density-Functional Thermochemistry. III. The Role of Exact Exchange. *J. Chem. Phys.* **1993**, 98 (7), 5648–5652. <https://doi.org/10.1063/1.464913>.
- (51) Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. A Consistent and Accurate Ab Initio Parametrization of Density Functional Dispersion Correction (DFT-D) for the 94 Elements H-Pu. *J. Chem. Phys.* **2010**, 132 (15), 154104. <https://doi.org/10.1063/1.3382344>.
- (52) Grimme, S.; Ehrlich, S.; Goerigk, L. Effect of the Damping Function in Dispersion Corrected Density Functional Theory. *J. Comput. Chem.* **2011**, 32 (7), 1456–1465. <https://doi.org/10.1002/jcc.21759>.
- (53) Weigend, F.; Ahlrichs, R. Balanced Basis Sets of Split Valence, Triple Zeta Valence and Quadruple Zeta Valence Quality for H to Rn: Design and Assessment of Accuracy. *Phys. Chem. Chem. Phys.* **2005**, 7 (18), 3297. <https://doi.org/10.1039/b508541a>.
- (54) Kruse, H.; Sponer, J. Towards Biochemically Relevant QM Computations on Nucleic Acids: Controlled Electronic Structure Geometry Optimization of Nucleic Acid Structural Motifs Using Penalty Restraint Functions. *Phys. Chem. Chem. Phys.* **2015**, 17 (2), 1399–1410. <https://doi.org/10.1039/C4CP04680C>.
- (55) Kruse, H. Xopt <https://github.com/hokru/xopt> (development version).
- (56) Ahlrichs, R.; Bär, M.; Häser, M.; Horn, H.; Kölmel, C. Electronic Structure Calculations on Workstation Computers: The Program System Turbomole. *Chem. Phys. Lett.* **1989**, 162 (3), 165–169. [https://doi.org/10.1016/0009-2614\(89\)85118-8](https://doi.org/10.1016/0009-2614(89)85118-8).
- (57) Treutler, O.; Ahlrichs, R. Efficient Molecular Numerical Integration Schemes. *J. Chem. Phys.* **1995**, 102 (1), 346–354. <https://doi.org/10.1063/1.469408>.
- (58) Weigend, F. A Fully Direct RI-HF Algorithm: Implementation, Optimised Auxiliary Basis Sets, Demonstration of Accuracy and Efficiency. *Phys. Chem. Chem. Phys.* **2002**, 4 (18), 4285–4291. <https://doi.org/10.1039/b204199p>.
- (59) Klamt, A.; Schuurmann, G. COSMO: A New Approach to Dielectric Screening in Solvents with Explicit Expressions for the Screening Energy and Its Gradient. *J. Chem. Soc., Perkin Trans. 2*

- 1993, No. 5, 799–805. <https://doi.org/10.1039/P29930000799>.
- (60) Kozuch, S.; Gruzman, D.; Martin, J. M. L. DSD-BLYP: A General Purpose Double Hybrid Density Functional Including Spin Component Scaling and Dispersion Correction. *J. Phys. Chem. C* **2010**, *114* (48), 20801–20808. <https://doi.org/10.1021/jp1070852>.
- (61) Grimme, S.; Ehrlich, S.; Goerigk, L. Effect of the Damping Function in Dispersion Corrected Density Functional Theory. *J. Comput. Chem.* **2011**, *32* (7), 1456–1465. <https://doi.org/10.1002/jcc.21759>.
- (62) Zheng, J.; Xu, X.; Truhlar, D. G. Minimally Augmented Karlsruhe Basis Sets. *Theor. Chem. Acc.* **2011**, *128* (3), 295–305. <https://doi.org/10.1007/s00214-010-0846-z>.
- (63) Neese, F. Software Update: The ORCA Program System, Version 4.0. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2018**, *8* (1), 4–9. <https://doi.org/10.1002/wcms.1327>.
- (64) Kruse, H.; Mladek, A.; Gkionis, K.; Hansen, A.; Grimme, S.; Sponer, J. Quantum Chemical Benchmark Study on 46 RNA Backbone Families Using a Dinucleotide Unit. *J. Chem. Theory Comput.* **2015**, *11* (10), 4972–4991. <https://doi.org/10.1021/acs.jctc.5b00515>.
- (65) Goerigk, L.; Grimme, S. Efficient and Accurate Double-Hybrid-Meta-GGA Density Functionals-Evaluation with the Extended GMTKN30 Database for General Main Group Thermochemistry, Kinetics, and Noncovalent Interactions. *J. Chem. Theory Comput.* **2011**, *7* (2), 291–309. <https://doi.org/10.1021/ct100466k>.
- (66) Spicher, S.; Caldeweyher, E.; Hansen, A.; Grimme, S. Benchmarking London Dispersion Corrected Density Functional Theory for Noncovalent Ion- π Interactions. *Phys. Chem. Chem. Phys.* **2021**, *23* (20), 11635–11648. <https://doi.org/10.1039/d1cp01333e>.
- (67) Stoychev, G. L.; Auer, A. A.; Neese, F. Automatic Generation of Auxiliary Basis Sets. *J. Chem. Theory Comput.* **2017**, *13* (2), 554–562. <https://doi.org/10.1021/acs.jctc.6b01041>.
- (68) Neese, F.; Wennmohs, F.; Hansen, A.; Becker, U. Efficient, Approximate and Parallel Hartree-Fock and Hybrid DFT Calculations. A “chain-of-Spheres” Algorithm for the Hartree-Fock Exchange. *Chem. Phys.* **2009**, *356* (1–3), 98–109. <https://doi.org/10.1016/j.chemphys.2008.10.036>.
- (69) Smith, D. G. A.; Burns, L. A.; Simmonett, A. C.; Parrish, R. M.; Schieber, M. C.; Galvelis, R.; Kraus, P.; Kruse, H.; Di Remigio, R.; Alenaizan, A.; James, A. M.; Lehtola, S.; Misiewicz, J. P.; Scheurer, M.; Shaw, R. A.; Schriber, J. B.; Xie, Y.; Glick, Z. L.; Sirianni, D. A.; O’Brien, J. S.; Waldrop, J. M.; Kumar, A.; Hohenstein, E. G.; Pritchard, B. P.; Brooks, B. R.; Schaefer, H. F.; Sokolov, A. Y.; Patkowski, K.; DePrince, A. E.; Bozkaya, U.; King, R. A.; Evangelista, F. A.; Turney, J. M.; Crawford, T. D.; Sherrill, C. D. Psi4 1.4: Open-Source Software for High-Throughput Quantum Chemistry. *J. Chem. Phys.* **2020**, *152* (18), 184108. <https://doi.org/10.1063/5.0006002>.
- (70) Dunning, T. H. Gaussian Basis Sets for Use in Correlated Molecular Calculations. I. The Atoms Boron through Neon and Hydrogen. *J. Chem. Phys.* **1989**, *90* (2), 1007–1023. <https://doi.org/10.1063/1.456153>.
- (71) Kendall, R. A.; Dunning, T. H.; Harrison, R. J. Electron Affinities of the First-Row Atoms Revisited. Systematic Basis Sets and Wave Functions. *J. Chem. Phys.* **1992**, *96* (9), 6796–6806. <https://doi.org/10.1063/1.462569>.
- (72) Papajak, E.; Truhlar, D. G. Convergent Partially Augmented Basis Sets for Post-Hartree-Fock Calculations of Molecular Properties and Reaction Barrier Heights. *J. Chem. Theory Comput.* **2011**, *7* (1), 10–18. <https://doi.org/10.1021/ct1005533>.

- (73) Halkier, A.; Helgaker, T.; Jørgensen, P.; Klopper, W.; Koch, H.; Olsen, J.; Wilson, A. K. Basis-Set Convergence in Correlated Calculations on Ne, N₂, and H₂O. *Chem. Phys. Lett.* **1998**, *286* (3–4), 243–252. [https://doi.org/10.1016/S0009-2614\(98\)00111-0](https://doi.org/10.1016/S0009-2614(98)00111-0).
- (74) Hohenstein, E. G.; Sherrill, C. D. Density Fitting of Intramonomer Correlation Effects in Symmetry-Adapted Perturbation Theory. *J. Chem. Phys.* **2010**, *133* (1), 0–12. <https://doi.org/10.1063/1.3451077>.
- (75) Hohenstein, E. G.; Sherrill, C. D. Efficient Evaluation of Triple Excitations in Symmetry-Adapted Perturbation Theory via Second-Order Møller-Plesset Perturbation Theory Natural Orbitals. *J. Chem. Phys.* **2010**, *133* (10). <https://doi.org/10.1063/1.3479400>.
- (76) Parker, T. M.; Burns, L. A.; Parrish, R. M.; Ryno, A. G.; Sherrill, C. D. Levels of Symmetry Adapted Perturbation Theory (SAPT). I. Efficiency and Performance for Interaction Energies. *J. Chem. Phys.* **2014**, *140* (9). <https://doi.org/10.1063/1.4867135>.
- (77) Parrish, R. M.; Burns, L. A.; Smith, D. G. A.; Simmonett, A. C.; DePrince, A. E.; Hohenstein, E. G.; Bozkaya, U.; Sokolov, A. Y.; Di Remigio, R.; Richard, R. M.; Gonthier, J. F.; James, A. M.; McAlexander, H. R.; Kumar, A.; Saitow, M.; Wang, X.; Pritchard, B. P.; Verma, P.; Schaefer, H. F.; Patkowski, K.; King, R. A.; Valeev, E. F.; Evangelista, F. A.; Turney, J. M.; Crawford, T. D.; Sherrill, C. D. Psi4 1.1: An Open-Source Electronic Structure Program Emphasizing Automation, Advanced Libraries, and Interoperability. *J. Chem. Theory Comput.* **2017**, *13* (7), 3185–3197. <https://doi.org/10.1021/acs.jctc.7b00174>.
- (78) Cisneros, G. A.; Karttunen, M.; Ren, P.; Sagui, C. Classical Electrostatics for Biomolecular Simulations. *Chem. Rev.* **2014**, *114* (1), 779–814. <https://doi.org/10.1021/cr300461d>.
- (79) Jing, Z.; Liu, C.; Cheng, S. Y.; Qi, R.; Walker, B. D.; Piquemal, J.-P.; Ren, P. Polarizable Force Fields for Biomolecular Simulations: Recent Advances and Applications. **2019**, *48*, 371–394. <https://doi.org/10.1146/annurev-biophys-070317-033349>.
- (80) Zhang, C.; Lu, C.; Jing, Z.; Wu, C.; Piquemal, J. P.; Ponder, J. W.; Ren, P. AMOEBA Polarizable Atomic Multipole Force Field for Nucleic Acids. *J. Chem. Theory Comput.* **2018**, *14* (4), 2084–2108. <https://doi.org/10.1021/acs.jctc.7b01169>.
- (81) Sengul, M. Y.; MacKerell Jr, A. D. Accurate Modeling of RNA Hairpins Through the Explicit Treatment of Electronic Polarizability with the Classical Drude Oscillator Force Field. *J. Comput. Biophys. Chem.* **2022**, *21* (4), 461–471.
- (82) Kruse, H. BrnoFF <https://github.com/hokru/BrnoFF>.
- (83) Case, D. A.; Betz, R. M.; Cerutti, D. S.; Cheatham III, T. E.; Darden, T. A.; Duke, R. E.; Giese, T. J.; Gohlke, H.; Goetz, A. W.; Homeyer, N.; Izadi, S.; Janowski, P.; Kaus, J.; Kovalenko, A.; Lee, T. S.; LeGrand, S.; Li, P.; Lin, C.; Luchko, T.; Luo, R.; Madej, B.; Mermelstein, D.; Merz, K. M.; Monard, G.; Nguyen, H.; Nguyen, H. T.; Omelyan, I.; Onufriev, A.; Roe, D. R.; Roitberg, A.; Sagui, C.; Simmerling, C. L.; Botello-Smith, W. M.; Swails, J.; Walker, R. C.; Wang, J.; Wolf, R. M.; Wu, X.; Xiao, L.; Kollman, P. A. AMBER 2016. University of California, San Francisco 2016.
- (84) Case, D. A.; Babin, V.; Berryman, J. T.; Betz, R. M.; Cai, Q.; Cerutti, D. S.; Cheatham III, T. E.; Darden, T. A.; Duke, R. E.; Gohlke, H.; Goetz, A. W.; Gusarov, S.; Homeyer, N.; Janowski, P.; Kaus, J.; Kolossvary, I.; Kovalenko, A.; Lee, T. S.; LeGrand, S.; Luchko, T.; Luo, R.; Madej, B.; Merz, K. M.; Paesani, F.; Roe, D. R.; Roitberg, A.; Sagui, C.; Salomon-Ferrer, R.; Seabra, G.; Simmerling, C. L.; Smith, W.; Swails, J.; Walker, R. C.; Wang, J.; Wolf, R. M.; Wu, X.; Kollman, P. A. AMBER 14. University of California, San Francisco 2014.
- (85) Bayly, C. I.; Cieplak, P.; Cornell, W. D.; Kollman, P. A. A Well-Behaved Electrostatic Potential

- Based Method Using Charge Restraints for Deriving Atomic Charges: The RESP Model. *J. Phys. Chem.* **1993**, *97* (40), 10269–10280. <https://doi.org/10.1021/j100142a004>.
- (86) Wang, J.; Wang, W.; Kollman, P. A.; Case, D. A. Automatic Atom Type and Bond Type Perception in Molecular Mechanical Calculations. *J. Mol. Graph. Model.* **2006**, *25* (2), 247–260. <https://doi.org/10.1016/j.jmglm.2005.12.005>.
- (87) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J. A.; Peralta Jr., J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, J. M.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, O.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. Gaussian 09, Revision D.01. Gaussian, Inc., Wallingford CT 2009.
- (88) Steinbrecher, T.; Latzer, J.; Case, D. A. Revised AMBER Parameters for Bioorganic Phosphates. *J. Chem. Theory Comput.* **2012**, *8* (11), 4405–4412. <https://doi.org/10.1021/ct300613v>.
- (89) Barone, V.; Cossi, M. Quantum Calculation of Molecular Energies and Energy Gradients in Solution by a Conductor Solvent Model. *J. Phys. Chem. A* **1998**, *102* (11), 1995–2001. <https://doi.org/10.1021/jp9716997>.
- (90) York, D. M.; Karplus, M. A Smooth Solvation Potential Based on the Conductor-like Screening Model. *J. Phys. Chem. A* **1999**, *103* (50), 11060–11079. <https://doi.org/10.1021/jp992097l>.
- (91) Garcia-Ratés, M.; Neese, F. Effect of the Solute Cavity on the Solvation Energy and Its Derivatives within the Framework of the Gaussian Charge Scheme. *J. Comput. Chem.* **2020**, *41* (9), 922–939. <https://doi.org/10.1002/jcc.26139>.
- (92) Hawkins, G. D.; Cramer, C. J.; Truhlar, D. G. Pairwise Solute Descreening of Solute Charges from a Dielectric Medium. *Chem. Phys. Lett.* **1995**, *246* (1–2), 122–129. [https://doi.org/10.1016/0009-2614\(95\)01082-K](https://doi.org/10.1016/0009-2614(95)01082-K).
- (93) Hawkins, G. D.; Cramer, C. J.; Truhlar, D. G. Parametrized Models of Aqueous Free Energies of Solvation Based on Pairwise Descreening of Solute Atomic Charges from a Dielectric Medium. *J. Phys. Chem.* **1996**, *100* (51), 19824–19839. <https://doi.org/10.1021/jp961710n>.
- (94) Berendsen, H. J. C.; Grigera, J. R.; Straatsma, T. P. The Missing Term in Effective Pair Potentials. *J. Phys. Chem.* **1987**, *91* (24), 6269–6271. <https://doi.org/10.1021/j100308a038>.
- (95) Götz, A. W.; Clark, M. A.; Walker, R. C. An Extensible Interface for QM/MM Molecular Dynamics Simulations with AMBER. *J. Comput. Chem.* **2014**, *35* (2), 95–108. <https://doi.org/10.1002/jcc.23444>.
- (96) TURBOMOLE V7.3 2018, a Development of University of Karlsruhe and Forschungszentrum Karlsruhe GmbH, 1989-2007, TURBOMOLE GmbH, since 2007.
- (97) Grimme, S.; Brandenburg, J. G.; Bannwarth, C.; Hansen, A. Consistent Structures and Interactions by Density Functional Theory with Small Atomic Orbital Basis Sets. *J. Chem. Phys.* **2015**, *143* (5). <https://doi.org/10.1063/1.4927476>.
- (98) Kruse, H.; Grimme, S. A Geometrical Correction for the Inter- and Intra-Molecular Basis Set

- Superposition Error in Hartree-Fock and Density Functional Theory Calculations for Large Systems. *J. Chem. Phys.* **2012**, *136* (15), 154101. <https://doi.org/10.1063/1.3700154>.
- (99) Liu, D. C.; Nocedal, J. On the Limited Memory BFGS Method for Large Scale Optimization. *Math. Program.* **1989**, *45*, 503–528.
- (100) Case, D. A.; Ben-Shalom, I. Y.; Brozell, S. R.; Cerutti, D. S.; Cheatham, III, T. E.; Cruzeiro, V. W. D.; Darden, T. A.; Duke, R. E.; Ghoreishi, D.; Gilson, M. K.; Gohlke, H.; Goetz, A. W.; Greene, D.; Harris, R.; Homeyer, N.; Huang, Y.; Izadi, S.; Kovalenko, A.; Kurtzman, T.; Lee, T. S.; LeGrand, S.; Li, P.; Lin, C.; Liu, J.; Luchko, T.; Luo, R.; Mermelstein, D.; Merz, K. M.; Miao, Y.; Monard, G.; Nguyen, C.; Nguyen, H.; Omelyan, I.; Onufriev, A.; Pan, F.; Qi, R.; Roe, D. R.; Roitberg, A.; Sagui, C.; Schott-Verdugo, S.; Shen, J.; Simmerling, C. L.; Smith, J.; Salomon-Ferrer, R.; Swails, J.; Walker, R. C.; Wang, J.; Wei, H.; Wolf, R. M.; Wu, X.; Xiao, L.; York, D. M.; Kollman, P. A. AMBER 2018. University of California, San Francisco 2018.
- (101) Wang, J.; 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. [https://doi.org/https://doi.org/10.1002/1096-987X\(200009\)21:12<1049::AID-JCC3>3.0.CO;2-F](https://doi.org/https://doi.org/10.1002/1096-987X(200009)21:12<1049::AID-JCC3>3.0.CO;2-F).
- (102) Perez, A.; Marchan, I.; Svozil, D.; Sponer, J.; Cheatham III, T. E.; Laughton, C. A.; Orozco, M. Refinement of the AMBER Force Field for Nucleic Acids: Improving the Description of α/γ Conformers. *Biophys. J.* **2007**, *92* (11), 3817–3829. <https://doi.org/10.1529/biophysj.106.097782>.
- (103) Mlýnský, V.; Kührová, P.; Zgarbová, M.; Jurečka, P.; Walter, N. G.; Otyepka, M.; Šponer, J.; Banáš, P. Reactive Conformation of the Active Site in the Hairpin Ribozyme Achieved by Molecular Dynamics Simulations with ϵ/ζ Force Field Reparametrizations. *J. Phys. Chem. B* **2015**, *119* (11), 4220–4229. <https://doi.org/10.1021/jp512069n>.
- (104) Hansen, J. L.; Schmeing, T. M.; Moore, P. B.; Steitz, T. A. Structural Insights into Peptide Bond Formation. *Proc. Natl. Acad. Sci. U. S. A.* **2002**, *99* (18), 11670–11675. <https://doi.org/10.1073/pnas.172404099>.
- (105) 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 Research*. December 1, 2003, pp 6806–6818. <https://doi.org/10.1093/nar/gkg908>.
- (106) Klein, D. J.; Schmeing, T. M.; Moore, P. B.; Steitz, T. A. The Kink-Turn: A New RNA Secondary Structure Motif. *EMBO J.* **2001**, *20* (15), 4214–4221. <https://doi.org/10.1093/emboj/20.15.4214>.
- (107) Jucker, F. M.; Heus, H. A.; Yip, P. F.; Moors, E. H. M.; Pardi, A. A Network of Heterogeneous Hydrogen Bonds in GNRA Tetraloops. *J. Mol. Biol.* **1996**, *264* (5), 968–980. <https://doi.org/10.1006/jmbi.1996.0690>.
- (108) Izadi, S.; Anandakrishnan, R.; Onufriev, A. V. Building Water Models: A Different Approach. *J. Phys. Chem. Lett.* **2014**, *5* (21), 3863–3871. <https://doi.org/10.1021/jz501780a>.
- (109) Joung, I. S.; Cheatham III, 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. <https://doi.org/10.1021/jp8001614>.
- (110) 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. <https://doi.org/10.1021/ct5010406>.

- (111) Burns, L. A.; Marshall, M. S.; Sherrill, C. D. Appointing Silver and Bronze Standards for Noncovalent Interactions: A Comparison of Spin-Component-Scaled (SCS), Explicitly Correlated (F12), and Specialized Wavefunction Approaches. *J. Chem. Phys.* **2014**, *141* (23). <https://doi.org/doi:http://dx.doi.org/10.1063/1.4903765>.
- (112) Kruse, H.; Goerigk, L.; Grimme, S. Why the Standard B3LYP/6-31G* Model Chemistry Should Not Be Used in DFT Calculations of Molecular Thermochemistry: Understanding and Correcting the Problem. *J. Org. Chem.* **2012**, *77* (23), 10824–10834. <https://doi.org/10.1021/jo302156p>.
- (113) Gryn'ova, G.; Corminboeuf, C. Steric "Attraction": Not by Dispersion Alone. *Beilstein J. Org. Chem.* **2018**, *14*, 1482–1490. <https://doi.org/10.3762/bjoc.14.125>.
- (114) Zgarbova, M.; Otyepka, M.; Sponer, J.; Hobza, P.; Jurecka, P. Large-Scale Compensation of Errors in Pairwise-Additive Empirical Force Fields: Comparison of AMBER Intermolecular Terms with Rigorous DFT-SAPT Calculations. *Phys. Chem. Chem. Phys.* **2010**, *12* (35), 10476–10493. <https://doi.org/10.1039/c002656e>.
- (115) Sirianni, D. A.; Zhu, X.; Sitkoff, D. F.; Cheney, D. L.; Sherrill, C. D. The Influence of a Solvent Environment on Direct Non-Covalent Interactions between Two Molecules: A Symmetry-Adapted Perturbation Theory Study of Polarization Tuning of π - Π Interactions by Water. *J. Chem. Phys.* **2022**, *156* (19). <https://doi.org/10.1063/5.0087302>.
- (116) Kruse, H.; Sponer, J.; Auffinger, P. Comment on "Evaluating Unexpectedly Short Non-Covalent Distances in X-Ray Crystal Structures of Proteins with Electronic Structure Analysis." *J. Chem. Inf. Model.* **2019**, *59* (9), 3605–3608. <https://doi.org/10.1021/acs.jcim.9b00473>.
- (117) Chakraborty, D.; Collepardo-Guevara, R.; Wales, D. J. Energy Landscapes, Folding Mechanisms, and Kinetics of RNA Tetraloop Hairpins. *J. Am. Chem. Soc.* **2014**, *136* (52), 18052–18061. <https://doi.org/10.1021/ja5100756>.
- (118) Ryde, U. How Many Conformations Need to Be Sampled to Obtain Converged QM/MM Energies? The Curse of Exponential Averaging. *J. Chem. Theory Comput.* **2017**, *13* (11), 5745–5752. <https://doi.org/10.1021/acs.jctc.7b00826>.
- (119) Gkionis, K.; Kruse, H.; Sponer, J. Derivation of Reliable Geometries in QM Calculations of DNA Structures: Explicit Solvent QM/MM and Restrained Implicit Solvent QM Optimizations of G-Quadruplexes. *J. Chem. Theory Comput.* **2016**, *12* (4), 2000–2016. <https://doi.org/10.1021/acs.jctc.5b01025>.
- (120) Pokorna, P.; Kruse, H.; Krepl, M.; Sponer, J. QM/MM Calculations on Protein-RNA Complexes: Understanding Limitations of Classical MD Simulations and Search for Reliable Cost-Effective QM Methods. *J. Chem. Theory Comput.* **2018**, *14* (10), 5419–5433. <https://doi.org/10.1021/acs.jctc.8b00670>.
- (121) Khaliullin, R. Z.; Bell, A. T.; Head-Gordon, M. Analysis of Charge Transfer Effects in Molecular Complexes Based on Absolutely Localized Molecular Orbitals. *J. Chem. Phys.* **2008**, *128* (18). <https://doi.org/10.1063/1.2912041>.
- (122) Horn, P. R.; Mao, Y.; Head-gordon, M. Probing Non-Covalent Interactions with a Second Generation Energy Decomposition Analysis Using Absolutely Localized Molecular Orbitals. *Phys. Chem. Chem. Phys.* **2016**, *18*, 23067–23079. <https://doi.org/10.1039/C6CP03784D>.
- (123) Zhao, Y.; Cotellet, Y.; Liu, L.; López-Andarias, J.; Bornhof, A. B.; Akamatsu, M.; Sakai, N.; Matile, S. The Emergence of Anion- π Catalysis. *Acc. Chem. Res.* **2018**, *51* (9), 2255–2263. <https://doi.org/10.1021/acs.accounts.8b00223>.
- (124) Wang, D. X.; Wang, M. X. Exploring Anion- Π Interactions and Their Applications in

- Supramolecular Chemistry. *Acc. Chem. Res.* **2020**, *53* (7), 1364–1380. <https://doi.org/10.1021/acs.accounts.0c00243>.
- (125) Lemkul, J. A.; Huang, J.; Roux, B.; Mackerell, A. D. An Empirical Polarizable Force Field Based on the Classical Drude Oscillator Model: Development History and Recent Applications. *Chem. Rev.* **2016**, *116* (9), 4983–5013. <https://doi.org/10.1021/acs.chemrev.5b00505>.
- (126) Lemkul, J. A.; MacKerell Jr, A. D. Polarizable Force Field for RNA Based on the Classical Drude Oscillator. *J. Comput. Chem.* **2018**, *39* (32), 2624–2646. <https://doi.org/10.1002/jcc.25709>.
- (127) Bondi, A. Van Der Waals Volumes and Radii. *J. Phys. Chem.* **1964**, *68* (3), 441–451. <https://doi.org/10.1021/j100785a001>.
- (128) Školáková, P.; Badri, Z.; Foldynová-Trantírková, S.; Ryneš, J.; Šponer, J.; Fojtová, M.; Fajkus, J.; Marek, R.; Vorlíčková, M.; Mergny, J. L.; Trantírek, L. Composite 5-Methylations of Cytosines Modulate i-Motif Stability in a Sequence-Specific Manner: Implications for DNA Nanotechnology and Epigenetic Regulation of Plant Telomeric DNA. *Biochim. Biophys. Acta - Gen. Subj.* **2020**, *1864* (9), 129651. <https://doi.org/10.1016/j.bbagen.2020.129651>.
- (129) Wei, H.; Qi, R.; Wang, J.; Cieplak, P.; Duan, Y.; Luo, R. Efficient Formulation of Polarizable Gaussian Multipole Electrostatics for Biomolecular Simulations. *J. Chem. Phys.* **2020**, *153* (11). <https://doi.org/10.1063/5.0019560>.
- (130) Šponer, J.; Leszczynski, J.; Hobza, P. Thioguanine and Thiouracil: Hydrogen-Bonding and Stacking Properties. *J. Phys. Chem. A* **1997**, *101* (49), 9489–9495. <https://doi.org/10.1021/jp9720404>.
- (131) Leontis, N. B.; Westhof, E. Geometric Nomenclature and Classification of RNA Base Pairs. *RNA* **2001**, *7*, 499–512.