



Published in final edited form as:

Curr Opin Struct Biol. 2021 April ; 67: 9–17. doi:10.1016/j.sbi.2020.08.003.

Recent developments in empirical atomistic force fields for nucleic acids and applications to studies of folding and dynamics

Alexa M Salsbury¹, Justin A Lemkul^{1,2}

¹Department of Biochemistry, Virginia Tech, Blacksburg, VA 24061, United States

²Center for Drug Discovery, Virginia Tech, Blacksburg, VA 24061, United States

Abstract

Nucleic acids play critical roles in carrying genetic information, participating in catalysis, and preserving chromosomal structure. Despite over a century of study, efforts to understand the dynamics and structure-function relationships of DNA and RNA at the atomic level are still ongoing. Molecular dynamics (MD) simulations augment experiments by providing atomistic resolution and quantitative relationships between structure and conformational energy. Steady advancements in computer hardware, software, and atomistic force fields (FFs) over 40 years have facilitated new discoveries. Here, we review nucleic acid FF development with emphasis on recent refinements that have improved descriptions of important nucleic acid properties. We then discuss several key examples of successes and challenges in modeling nucleic acid structure and dynamics using the latest FFs.

Nonpolarizable nucleic acid force field development

Nucleic acid dynamics are governed by many interconnected properties, namely base stacking, hydrogen bonding, interactions with water and ions, sugar puckering, glycosidic torsion (χ) rotation, and phosphodiester backbone conformational sampling. For accurate modeling in MD simulations, a FF must be well balanced to capture these effects. Backbone properties are particularly challenging as each torsion in the phosphodiester backbone (Figure 1 a) is coupled. Accurate electrostatics are also key for modeling interactions of highly charged nucleic acids with water and ions. As such, most FF development and refinement has focused on accurately modeling these interactions.

Most FFs applied in MD simulations to date have relied on a pairwise-additive approximation to treating electrostatics, neglecting explicit electronic polarization and instead assigning partial atomic charges assuming a mean-field polarization effect. The two most popular nonpolarizable FF families for simulating nucleic acids are AMBER and CHARMM. Here, we review key milestones in the development of each FF.

Corresponding author: Lemkul, Justin A (jalemkul@vt.edu).

Conflict of interest statement

Nothing declared.

The first all-atom nucleic acid FF in the AMBER lineage, ff94, was introduced in 1995 by Cornell et al. [1] A key feature of ff94 was the assignment of partial charges by fitting to the quantum mechanical (QM) electrostatic potential (ESP) around model compounds [1]. The use of symmetry constraints in parametrization, referred to as restrained ESP (RESP) fitting [1], is central to the success of AMBER nucleic acid FFs and was used in subsequent refinement [2]. The ff98 [3] and ff99 [4] AMBER parameter sets added DNA-specific torsional parameters and improved duplex geometric properties. ff99 sufficiently described hydrogen bonding, base stacking, and base pairing on short time scales (<10 ns) [5–8]. However, longer simulations (>50 ns) with ff99 revealed incorrect modeling of γ -*trans* backbone conformers that resulted in B-DNA collapse and non-native RNA helical twist [9]. The parmbsc0 parameter set reparametrized the α and γ dihedrals to reduce non-native α/γ conformations and helical twisting, at the expense of overpenalizing γ -*trans* conformations that do occur in nucleic acids [10,11]. Another major issue in the ff99 FF lineage was the adoption of ‘ladder-like’ RNA structures that were attributed to incorrect χ sampling. The ff99bsc0_{OL3} FF was developed to prevent incorrect χ shifts and resulting non-native RNA helical twist, but the γ -*trans* penalty still destabilizes many noncanonical nucleic acid structures [11]. The AMBER parmbsc1 FF was built upon parmbsc0 by refining sugar puckering and χ , ϵ , and ζ dihedrals [12], and showed convincing improvement over parmbsc0 in exhaustive simulations of A-DNA, B-DNA, and Z-DNA. Different AMBER FFs may be combined with various water models to simulate specific nucleic acid systems [12,13]. Examples of such efforts will be described below. Further information on AMBER DNA and RNA FF parametrization and validation has been reviewed elsewhere [2,13,14].

The CHARMM FF is also widely used in simulating nucleic acids. The first parameter set, CHARMM22 (C22), was introduced in 1995 [15]. Subnanosecond MD simulations of DNA duplexes were stable, however, C22 overstabilized A-form DNA [16–18]. Foloppe *et al.* subsequently developed the CHARMM27 (C27) FF in 2000 [19,20] by refining backbone dihedrals, sugar puckering, and the glycosidic linkage to reproduce macromolecular target data from model compounds [19]. C27 yielded an improved model for B-DNA and RNA hairpins, and early simulations (<5 ns) yielded expected behaviors in hydrogen bonding, backbone dihedrals, and sugar puckering [20]. Limitations of C27 included modeling properties like groove width, overpopulation of the BI backbone substate in DNA (Figure 1b), and opening of Watson-Crick base pairs in RNA, prompting the development of CHARMM36 (C36) [21,22]. For DNA, C36 optimization focused on ϵ and ζ backbone dihedral angles as well as deoxyribose sugar puckering [21], while optimization of the RNA parameters focused on 2'-hydroxyl dihedral sampling [23]. More recently, longer C36 simulations have revealed that this FF may still suffer from aberrant base-pair opening [23] and possibly RNA helix fraying [2].

As is clear from the progressive refinement of the AMBER and CHARMM FFs, achieving adequate balance in the FF, especially in critical backbone conformational sampling, remains a challenge. Despite these shortcomings, both the AMBER and CHARMM nucleic acid FFs have been applied to investigations of many diverse structures and have provided important insights on their dynamics. Still, targeted improvements can be made. Recently, Kůhrová *et al.* developed and applied a generalized interaction-specific variant (gHBfix), that can fine-tune select nonbonded interactions like hydrogen bonding to improve the

accuracy of simulated RNA structural properties [24]. This strategy was used to improve upon ff99bsc0x_{OL3} and has been extended to improve hydrogen bonds involving terminal nucleotides (tHBfix) [24,25]. Both gHBfix and tHBfix have improved the conformational ensembles of RNA tetranucleotides and tetraloops in the AMBER FF and are promising strategies to improve nucleic acid FF performance without extensive refinement that may introduce new imbalances to the model [24,25]. Such select tuning can, in a sense, be viewed as another type of approximation of an electrostatics model that includes directionality, that is, induced polarization.

Polarizable force fields for nucleic acids

Polarizable FFs that explicitly account for electronic polarization have recently emerged as promising tools for modeling nucleic acids. The AMOEBA FF [26] is a multipole-induced dipole model in which each atom is assigned dipole and higher-order multipole moments. Inductive effects are calculated between ‘polarization groups’ that interact via permanent electrostatics and induced dipoles. The classical Drude oscillator FF [27] represents electronic degrees of freedom via negatively charged auxiliary particles (Drude oscillators) that are connected to non-hydrogen atoms via harmonic springs.

Both the AMOEBA [28] and Drude [29–32] FFs have complete parameter sets for nucleic acids, and the underlying approach to parametrization of both was similar. Parameter fitting in the AMOEBA FF targeted the QM ESP of model compounds and conformational energetics of sugar puckering and other dihedral rotations [28,33]. A key element of the AMOEBA FF parametrization was the inclusion of solvent effects in the QM calculations using a polarizable continuum model (PCM). The Drude FF targeted similar quantities (molecular dipole and polarizability, conformational energetics, base stacking, and hydrogen bonding) but did not utilize PCM in its QM calculations. Instead, the solvent response is incorporated into the FF by targeting scaled molecular polarizabilities and explicit water interactions with model compounds. The current Drude FF for nucleic acids is denoted Drude-2017 [30–32] and is a revision of Drude-2013 [29], the first polarizable FF for DNA. Extended simulations (>200 ns) with Drude-2013 revealed instability in duplex DNA that was attributed to systematically weak base-stacking interactions that were rectified in the Drude-2017 parameter set by explicitly targeting QM stacking energies [31], yielding stable microsecond simulations of A-DNA, B-DNA, and Z-DNA duplexes [32].

Both the AMOEBA and Drude nucleic acid FFs have been validated in microsecond simulations of duplex DNA and RNA, as well as prototypical RNA secondary structure motifs [32,34–37]. Both of these FFs yield good agreement with experimental properties such as base pairing, base stacking, backbone torsions, sugar puckering, groove geometries, and the conversion between A-form and B-form DNA in response to changes in solvent [28]. The main limitation in using polarizable FFs is reduced simulation speed; however, the implementation of AMOEBA in Tinker-HP, an MPI-enabled parallel package for MD simulations [38], and in the graphical processing unit (GPU)-accelerated OpenMM software [39], has made these polarizable simulations more accessible. Similarly, the Drude model is available in a wide range of high-performance simulation software, including CHARMM [40], NAMD [41], GROMACS [42], and OpenMM [43]. The use of GPU hardware greatly

accelerates polarizable simulations and further reduces barriers to their routine implementation in MD studies on nucleic acids.

Simulations of water and ion interactions with nucleic acids

Hydration patterns around nucleic acids are of interest because water molecules help stabilize DNA and RNA structure and modulate ion binding [44]. Canonical A-form RNA is stabilized by a water-mediated hydrogen bond between the 2'-hydroxyl group and the nitrogenous bases; release of this water with concomitant rotation of the 2'-hydroxyl toward backbone atoms is associated with flexibility and the adoption of noncanonical geometries [45]. Recently, Bergonzo and Grishaev assessed explicit and implicit solvent models in simulations of a model RNA helix. They showed that judicious inclusion of NMR restraints with MD simulations can improve accuracy, regardless of the choice of solvent representation, though for unbiased simulations, the choice of water model is particularly important for achieving agreement with NMR observables [46]. Their study indicates that the FFs applied are robust and demonstrates that combining modern FFs and experimental data can help refine experimental structures to improve descriptions of nucleic acid dynamics [46]. Another interesting simulation study focused on nucleobase-water interactions to understand the contributions of lone pair- π and OH- π interactions, which are generally considered weak, in stabilizing RNA structural motifs [47]. Using the C36 FF and QM calculations, Kalra *et al.* showed that these interactions are uniformly favorable and are related to the ESP of all four nucleobases, a principle that may be broadly applicable to highly ordered nucleic acid structures [47].

Like water, monovalent and divalent ions play an integral role in governing the folding and dynamics of nucleic acids. Accurately modeling ion interactions with nucleic acids in MD simulations is of critical importance and has been a persistent challenge in the field. One particular aspect of this challenge is delineating the roles of divalent metal ions like Mg^{2+} , which often have specific structural roles in folded RNA, from general screening effects of monovalent ions. Fischer *et al.* showed that Mg^{2+} and Na^+ interactions differ at high and low concentrations and in simulations of helical and complex folded RNA [48]. At lower concentrations, both ions showed increased direct interactions, but overall, Mg^{2+} ions better stabilized these RNA structures via close, water-mediated interactions at experimentally determined binding sites. These findings demonstrate the importance of salt concentration and Mg^{2+} :RNA interactions in MD simulations. Further, Nguyen *et al.* recently employed a coarse-grained RNA model with a reference interaction site model (RISM) for Mg^{2+} :phosphate interactions while treating monovalent cation effects implicitly [49]. With this approach, they were able to calculate folding free energies of RNA in agreement with experimental findings, particularly due to strong interactions of Mg^{2+} with partially folded RNA intermediates. This study serves as a strong foundation for future theoretical work on ion-induced nucleic acid folding [49].

The development of the Drude FF explicitly included fine-tuning of ion interactions with nucleic acids against QM target data, including common monovalent (Li^+ , Na^+ , K^+ , and Rb^+) and divalent (Mg^{2+}) ions [50,51]. As a result, MD simulations with the Drude FF have reproduced experimental solution X-ray scattering profiles for duplex DNA in solutions of

Li⁺, Na⁺, K⁺, and Rb [52], as well as proposing a mechanism by which water-mediated interactions between solvated cations and DNA perturb minor groove widths [52]. Further, mutual polarization between nucleobases and water was critical to computing accurate free-energy barriers to DNA base flipping [53]. Together, these findings suggest that subtle effects of ions and base hydration are modeled accurately using this polarizable FF and the Drude model may serve as the basis for new insights into nucleic acid structure and dynamics.

RNA folding simulations

Exploring RNA folding with MD simulations has been a longstanding challenge, as a reliable description of folding pathways and the native-state conformational ensembles depends on balance among base-base, base-backbone, and base-water interactions. Further, noncanonical hydrogen-bonding interactions are prevalent in many RNA motifs that are not often targeted during FF development. Here, we focus on simulations of RNA tetraloops (e.g. the four-nucleotide sequences that comprise the loops of stem-loop structures) and tetramers as representative examples of particularly challenging systems for MD simulations.

In 2013, Chen and Garcia investigated the folding of hyperstable r(UUCG), r(GCAA), and r(CUUG) RNA tetraloops [54]. Their simulations employed a modified ff99 parameter set in which they (1) refined the Lennard-Jones radii to better balance base stacking and base-water interactions, and (2) adjusted the χ dihedral parameters to achieve better *syn/anti* balance. With this modified FF, replica-exchange simulations recovered the native fold of the three tetraloops within 1–3 Å RMSD of the experimental structures. This result highlights the critical role of balanced interactions in nucleic acid FFs (and any FF, in general). Despite this success in folding RNA, subsequent investigations revealed that the modified parameters led to overly strong base-base hydrogen bonding and distorted structures [55,56], indicating that while the proposed refinements were suitable for RNA folding, they do not represent a transferable nucleic acid FF.

RNA tetramers are also difficult to model in MD simulations. Condon *et al.* carried out a systematic analysis of ff99-derived FFs in systems of r(AAAA), r(CAAU), r(GACC), and r(UUUU) using MD simulations and NMR spectroscopy [57]. Each of the FFs produced long-lived structures with incorrect, intercalated base stacking, suggesting an intrinsic inaccuracy in base hydration that leads to nonspecific collapse of the oligonucleotide chain. Bergonzo and Cheatham approached the same question of tetramer folding using a modified version of parmbsc0 with an additional χ refinement [10] and reoptimized Lennard-Jones radii [58] on phosphate oxygens [59]. Their simulations systematically evaluated different water models in lieu of the conventional TIP3P that is often used with AMBER FFs, with the OPC model [60] producing the best results. As such, their approach was to rebalance RNA-water and intra-RNA interactions to improve the description of tetramer conformational ensembles. Their modified parmbsc0+OPC model was suitable for r(GACC) but not for r(CCCC). Thus, the results of these studies indicate that there are intrinsic imbalances in several AMBER FFs that may be remedied in some cases by minor

refinement or use of different water models, but as yet there is not a general solution for producing RNA tetramer ensembles that agree with NMR data.

In validating the AMOEBA nucleic acid FF, Zhang *et al.* also simulated r(CAAU), r(GACC), and r(AAAA) tetramers, finding that their conformers showed only small deviations from NMR data and few incidences of inter-calation [28]. These findings suggest that use of a polarizable FF may alleviate some of the imbalances in the AMBER FFs noted above as the electronic structure of the oligonucleotide, particularly the bases, respond to changes in stacking, solvation, and hydrogen bonding in ways that nonpolarizable FFs cannot. It may also be possible that the success of the AMOEBA FF is due to more comprehensive parameter fitting and use of higher-quality QM target data. Further investigations with different polarizable models, like the Drude FF, should be undertaken to determine if polarizable FFs represent a general solution to issues faced in simulating RNA tetramers.

Simulations of highly ordered nucleic acids

Noncanonical nucleic acids serve functional roles in many biological processes and are often implicated in disease states. In the context of MD simulations, these noncanonical structures serve as stringent FF tests since their stabilizing interactions are different from those found in canonical nucleic acids normally targeted for FF parametrization and validation. As such, improved accuracy in modeling four-way junctions [61], G-quadruplexes (GQs) [36], i-motifs [62], hairpins [63,64], pseudoknots [47], and riboswitches [65] (Figure 2) reflects the maturity of modern nucleic acid FFs. GQs are an important example of noncanonical nucleic acids to focus on here because (1) they form in DNA and RNA, (2) their highly ordered structures are governed by interactions with water and ions, and (3) they have been the subject of many simulation studies with both additive and polarizable FFs.

GQ sequences self-associate into four-stranded helical structures that are characterized by stacked tetrads, each consisting of a square planar arrangement of Hoogsteen hydrogen-bonded guanines [66]. GQs typically have three stacked guanine tetrads, resulting in twelve inward-facing carbonyl oxygen atoms, four in each tetrad (Figure 3). This configuration forms a strongly electronegative channel that requires cation stabilization [67]. Broadly speaking, interactions with other proteins, ions, and water govern the folded topology and dynamics of GQs, which in turn influence their function [68–70]. MD simulations with AMBER, CHARMM, and Drude FFs have provided valuable insights into GQs [35–37,70–79]. For example, simulations provided the first description of the rigid tetrad core and destabilization in the absence of coordinated ions [78], identified possible GQ folding intermediates [79], and proposed folding pathways [73]. In general, recent simulations have yielded stable GQ structures in reasonable agreement to experimental data but have not provided complete descriptions of ion coordination and ion-ion energetics in the tetrad core, shortcomings that have been largely attributed to the nonpolarizable nature of the FFs that have typically been employed [75–77]. Specifically, charge-charge repulsion between core ions in nonpolarizable simulations hinders descriptions of ion binding and coordination in the tetrad core [77]. However, the inclusion of dipole-dipole interactions between ions can smooth electrostatic repulsion and stabilize ion binding [75]. While one simulation study showed it is possible to find system-specific combinations of water models and ion

parameters to improve descriptions of GQ-ion interactions [71], use of a polarizable FF may be a more general approach to overcoming these longstanding challenges.

We have recently employed the Drude FF to investigate DNA and RNA GQs [35–37], finding that the polarizable model improves descriptions of GQ-ion interactions and that the nucleobase dipole moments in GQs differ from canonical DNA and RNA [35–37]. The variation in nucleobase dipole moments is a function of ion type and coordination, hydration, and primary and secondary structure, all which influence conformational sampling and core energetics [35–37]. For example, in studying the stacked *c-kit2/c-kit** GQs (Figure 2b), we showed that the interplay of the two GQs and upstream and downstream regions modulates secondary structure, ion sampling, and core-ion interaction energies [36]. Though these systems were large (238 323 atoms), simulations of 500 ns were computationally tractable through the use of GPU acceleration. However, extending the time scales of these simulations further is important as other studies have shown that unbiased simulations of telomeric GQs may not be converged within even 10 μ s [72,74]. The combination of including sufficient flanking (non-GQ) nucleotides, increasing time scales, and achieving adequate sampling, is an emerging challenge in GQ simulations. Meeting this challenge requires careful attention to FF parametrization and reporting possible deficiencies in models, the latter of which requires high-quality experimental data for these systems. Despite these challenges, recent simulations of GQs have benefited from improved FF descriptions of water, ions, and polarization; therefore, their successes mark considerable progress in nucleic acid simulations.

Conclusions and outlook

The quality and scope of nucleic acid simulations have progressed dramatically over the past four decades, providing important atomistic insights into DNA and RNA structure and dynamics. These successes have been driven by the development, assessment, and progressive refinement of atomistic FFs, which have made use of increasing repositories of experimental and QM target data (Figure 4). Employing polarizable FFs is a promising approach to overcoming longstanding issues regarding hydration and ion interactions with nucleic acids. However, more investigations of diverse nucleic acid structures are needed to further assess the quality of these polarizable models. Further, recent efforts combining experiments and simulations show promise for improving the accuracy of nucleic acid FFs. To achieve better agreement with experimental data, future work may require well-characterized model systems from experimental and computational perspectives. Finally, multi- μ s simulations show that sampling is still an issue in some nucleic acid simulations, so extending simulation time scales and exploring efficient enhanced sampling methods are an important future endeavor. Together, these efforts will drive improvements to atomistic FFs and ultimately our understanding of nucleic acid structure, dynamics, and function.

Acknowledgements

This work was supported by the National Institutes of Health (grant R35GM133754), the Thomas F. and Kate Miller Jeffress Memorial Trust (Bank of America, Trustee), USDA-NIFA (project number VA-160092), and The American Association of University Women (American Dissertation Fellowship).

References

1. Cornell WD, Cieplak P, Bayly CI, Gould IR, Merz KM, Ferguson DM, Spellmeyer DC, Fox T, Caldwell JW, Kollman PA: A second generation force field for the simulation of proteins, nucleic acids, and organic molecules. *J Am Chem Soc* 1995, 117:5179–5197.
2. Šponer J, Bussi G, Krepl M, Banáš P, Bottaro S, Cunha RA, Gil-Ley A, Pinamonti G, Poblete S, Jurek P et al.: RNA structural dynamics as captured by molecular simulations: a comprehensive overview. *Chem Rev* 2018, 118: 4177–4338. [PubMed: 29297679]
3. Cheatham TE, Cieplak P, Kollman PA: A modified version of the Cornell et al. force field with improved sugar pucker phases and helical repeat. *J Biomol Struct Dyn* 1999, 16:845–862. [PubMed: 10217454]
4. Wang J, Cieplak P, Kollman PA: How well does a restrained electrostatic potential (RESP) model perform in calculating conformational energies of organic and biological molecules? *J Comput Chem* 2000, 21:1049–1074.
5. Réblová K, Špačková N, Kocá J, Leontis NB, Šponer J: Long-residency hydration, cation binding, and dynamics of loop E/Helix IV rRNA-L25 protein complex. *Biophys J* 2004, 87:3397–3412. [PubMed: 15339800]
6. Auffinger P, Bielecki L, Westhof E: The Mg²⁺ binding sites of the 5S rRNA Loop E motif as investigated by molecular dynamics simulations. *Chem Biol* 2003, 10:551–561. [PubMed: 12837388]
7. Auffinger P, Bielecki L, Westhof E: Symmetric K⁺ and Mg²⁺ ion-binding sites in the 5S rRNA loop inferred from molecular dynamics simulations. *J Mol Biol* 2004, 335:555–571. [PubMed: 14672663]
8. Réblová K, Špačková N, Štefl R, Csaszar K, Kocá J, Leontis NB, Šponer J: Non-watson-crick basepairing and hydration in RNA motifs: molecular dynamics of 5S rRNA Loop E. *Biophys J* 2003, 84:3564–3582. [PubMed: 12770867]
9. Pérez A, Marchán I, Svozil D, Šponer J, Cheatham TE, Laughton CA, Orozco M: Refinement of the AMBER force field for nucleic acids: improving the description of α/γ conformers. *Biophys J* 2007, 92:3817–3829. [PubMed: 17351000]
10. Zgarbová M, Otyepka M, Šponer J, Mládek A, Banáš P, Cheatham TE, Jurek P, Refinement of the Cornell et al.: Nucleic acids force field based on reference quantum chemical calculations of glycosidic torsion profiles. *J Chem Theory Comput* 2011, 7:2886–2902. [PubMed: 21921995]
11. Zgarbová M, Jurek 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, 121:2420–2433. [PubMed: 28290207]
12. Ivani I, Dans PD, Noy A, Pérez A, Faustino I, Hospital A, Walther J, Andrio P, Goñi R, Balaceanu A et al.: Parmbsc1: a refined force field for DNA simulations. *Nat Methods* 2015, 13:55–58. [PubMed: 26569599]
13. Galindo-Murillo R, Robertson JC, Zgarbová M, Šponer J, Otyepka M, Jurek P, Cheatham TE: Assessing the current state of amber force field modifications for DNA. *J Chem Theory Comput* 2016, 12:4114–4127. [PubMed: 27300587]
14. Krepl M, Zgarbová M, Stadlbauer P, Otyepka M, Banáš P, Kocá J, Cheatham TE, Jurek P, Šponer J: Reference simulations of noncanonical nucleic acids with different χ variants of the AMBER force field: Quadruplex DNA, Quadruplex RNA, and Z-DNA. *J Chem Theory Comput* 2012, 8:2506–2520. [PubMed: 23197943]
15. MacKerell AD, Wiórkiewicz-Kuczera J, Karplus M, MacKerell AD: An all-atom empirical energy function for the simulation of nucleic acids. *J Am Chem Soc* 1995, 117:11946–11975.
16. Yang L, Pettitt BM: B to A transition of DNA on the nanosecond time scale. *J Phys Chem* 1996, 100:2564–2566.
17. Pastor N, Pardo L, Weinstein H: Does TATA matter? A structural exploration of the selectivity determinants in its complexes with TATA box-binding protein. *Biophys J* 1997, 73:640–652. [PubMed: 9251783]
18. MacKerell AD: Influence of magnesium ions on duplex DNA structural, dynamic, and solvation properties. *J Phys Chem B* 1997, 101:646–650.

19. Foloppe N, MacKerell AD: All-atom empirical force field for nucleic acids: I. Parameter optimization based on small molecule and condensed phase macromolecular target data. *J Comput Chem* 2000, 21:86–104.
20. MacKerell AD, Banavali NK: All-atom empirical force field for nucleic acids: II. application to molecular dynamics simulations of DNA and RNA in solution. *J Comput Chem* 2000, 21:105–120.
21. Hart K, Foloppe N, Baker CM, Denning EJ, Nilsson L, MacKerell AD: Optimization of the CHARMM additive force field for DNA: improved treatment of the BI/BII conformational equilibrium. *J Chem Theory Comput* 2012, 8:348–362. [PubMed: 22368531]
22. Denning EJ, Priyakumar UD, Nilsson L, MacKerell AD: Impact of 2'-hydroxyl sampling on the conformational properties of RNA: update of the CHARMM all-atom additive force field for RNA. *J Comput Chem* 2011, 32:1929–1943. [PubMed: 21469161]
23. Minhas V, Sun T, Mirzoev A, Korolev N, Lyubartsev AP, Nordenskiöld L: Modeling DNA flexibility: comparison of force fields from atomistic to multiscale levels. *J Phys Chem B* 2020, 124:38–49. [PubMed: 31805230]
24. Kührová P, Mlýnský V, Zgarbová M, Krepl M, Bussi G, Best RB, 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, 15:3288–3305. [PubMed: 30896943]
25. Mlýnský V, Kührová P, Kühn T, Otyepka M, Bussi G, Banáš P, Šponer J: Fine-tuning of the AMBER RNA Force field with a new term adjusting interactions of terminal nucleotides. *J Chem Theory Comput* 2020, 16:3936–3946. [PubMed: 32384244]
26. Ponder JW, Wu C, Ren P, Pande VS, Chodera JD, Schnieders MJ, Haque I, Mobley DL, Lambrecht DS, Distasio RA et al.: Current status of the AMOEBA polarizable force field. *J Phys Chem B* 2010, 114:2549–2564. [PubMed: 20136072]
27. Lemkul JA, Huang J, Roux B, MacKerell AD: An empirical polarizable force field based on the classical drude oscillator model: development history and recent applications. *Chem Rev* 2016, 116:4983–5013. [PubMed: 26815602]
28. Zhang C, Lu C, Jing Z, Wu C, Piquemal J-P, Ponder JW, Ren P: AMOEBA polarizable atomic multipole force field for nucleic acids. *J Chem Theory Comput* 2018, 14:2084–2108. [PubMed: 29438622]
29. Savelyev A, MacKerell AD: All-atom polarizable force field for DNA based on the classical drude oscillator model. *J Comput Chem* 2014, 35:1219–1239. [PubMed: 24752978]
30. Lemkul JA, MacKerell AD: Polarizable force field for RNA based on the classical drude oscillator. *J Comput Chem* 2018, 39:2624–2646. [PubMed: 30515902]
31. Lemkul JA, MacKerell AD: Polarizable force field for DNA based on the classical drude oscillator: I. Refinement using quantum mechanical base stacking and conformational energetics. *J Chem Theory Comput* 2017, 13:2053–2071. [PubMed: 28399366]
32. Lemkul JA, MacKerell AD: Polarizable force field for DNA based on the classical drude oscillator: II. Microsecond molecular dynamics simulations of duplex DNA. *J Chem Theory Comput* 2017, 13:2072–2085. [PubMed: 28398748]
33. Shi Y, Xia Z, Zhang J, Best R, Wu C, Ponder JW, Ren P: Polarizable atomic multipole-based AMOEBA force field for proteins. *J Chem Theory Comput* 2013, 9:4046–4063. [PubMed: 24163642]
34. Lemkul JA, MacKerell AD: Polarizable force field for DNA based on the classical drude oscillator:II. Microsecond molecular dynamics simulations of duplex DNA. *J Chem Theory Comput* 2017, 13:2072–2085. [PubMed: 28398748]
35. Salsbury AM, Lemkul JA: Molecular dynamics simulations of the c-kit1 promoter G-quadruplex: importance of electronic polarization on stability and cooperative ion binding. *J Phys Chem B* 2019, 123:148–159. [PubMed: 30525627]
36. Salsbury AM, Dean TJ, Lemkul JA: Polarizable molecular dynamics simulations of two c-kit oncogene promoter G-quadruplexes: effect of primary and secondary structure on loop and ion sampling. *J Chem Theory Comput* 2020, 16:3430–3444. [PubMed: 32307997]

37. Lemkul JA: Same fold, different properties: polarizable molecular dynamics simulations of telomeric and TERRA G-quadruplexes. *Nucleic Acids Res* 2020, 48:561–575. [PubMed: 31807754]
38. Lagardère L, Jolly LH, Lipparini F, Aviat F, Stamm B, Jing ZF, Harger M, Torabifard H, Cisneros GA, Schnieders MJ et al.: Tinker-HP: a massively parallel molecular dynamics package for multiscale simulations of large complex systems with advanced point dipole polarizable force fields. *Chem Sci* 2018, 9:956–972. [PubMed: 29732110]
39. Eastman P, Swails J, Chodera JD, McGibbon RT, Zhao Y, Beauchamp KA, Wang L-P, Simmonett AC, Harrigan MP, Stern CD et al.: OpenMM 7: rapid development of high performance algorithms for molecular dynamics. *PLoS Comput Biol* 2017, 13: e1005659. [PubMed: 28746339]
40. Lamoureux G, Roux B: Modeling induced polarization with classical drude oscillators: theory and molecular dynamics simulation algorithm. *J Chem Phys* 2003, 119:3025–3039.
41. Jiang W, Hardy DJ, Phillips JC, MacKerell AD, Schulten K, Roux B: High-performance scalable molecular dynamics simulations of a polarizable force field based on classical drude oscillators in NAMD. *J Phys Chem Lett* 2011, 2:87–92. [PubMed: 21572567]
42. Lemkul JA, Roux B, van der Spoel D, MacKerell AD: Implementation of extended lagrangian dynamics in GROMACS for polarizable simulations using the classical drude oscillator model. *J Comput Chem* 2015, 36:1473–1479. [PubMed: 25962472]
43. Huang J, Lemkul JA, Eastman PK, MacKerell AD: Molecular dynamics simulations using the drude polarizable force field on GPUs with OpenMM: implementation, validation, and benchmarks. *J Comput Chem* 2018, 39:1682–1689. [PubMed: 29727037]
44. Lipfert J, Doniach S, Das R, Herschlag D: Understanding nucleic acid–ion interactions. *Annu Rev Biochem* 2014, 83:813–841. [PubMed: 24606136]
45. Denning EJ, MacKerell AD: Intrinsic contribution of the 2′-hydroxyl to RNA conformational heterogeneity. *J Am Chem Soc* 2012, 134:2800–2806. [PubMed: 22242623]
46. Bergonzo C, Grishaev A: Accuracy of MD solvent models in RNA structure refinement assessed via liquid-crystal NMR and spin relaxation data. *J Struct Biol* 2019, 207:250–259. [PubMed: 31279068]
47. Kalra K, Gorle S, Cavallo L, Oliva R, Chawla M: Occurrence and stability of lone pair-and OH-interactions between water and nucleobases in functional RNAs. *Nucleic Acids Res* 2020, 48:5825–5838. [PubMed: 32392301]
48. Fischer NM, Polêto MD, Steuer J, van der Spoel D: Influence of Na⁺ and Mg²⁺ ions on RNA structures studied with molecular dynamics simulations. *Nucleic Acids Res* 2018, 46:4872–4882. [PubMed: 29718375]
49. Nguyen HT, Hori N, Thirumalai D: Theory and simulations for RNA folding in mixtures of monovalent and divalent cations. *Proc Natl Acad Sci U S A* 2019, 116:21022–21030. [PubMed: 31570624]
50. Savelyev A, MacKerell AD: Balancing the interactions of ions, water, and DNA in the drude polarizable force field. *J Phys Chem B* 2014, 118:6742–6757. [PubMed: 24874104]
51. Lemkul JA, MacKerell AD: Balancing the interactions of Mg²⁺ in aqueous solution and with nucleic acid moieties for a polarizable force field based on the classical drude oscillator model. *J Phys Chem B* 2016, 120:11436–11448. [PubMed: 27759379]
52. Savelyev A, MacKerell AD: Differential deformability of the DNA minor groove and altered BI/BII backbone conformational equilibrium by the monovalent ions Li⁺, Na⁺, K⁺, and Rb⁺ via water-mediated hydrogen bonding. *J Chem Theory Comput* 2015, 11:4473–4485. [PubMed: 26575937]
53. Lemkul JA, Savelyev A, MacKerell AD: Induced polarization influences the fundamental forces in DNA base flipping. *J Phys Chem Lett* 2014, 5:2077–2083. [PubMed: 24976900]
54. Chen AA, García AE: High-resolution reversible folding of hyperstable RNA tetraloops using molecular dynamics simulations. *Proc Natl Acad Sci U S A* 2013, 110:16820–16825. [PubMed: 24043821]
55. Bergonzo C, Henriksen NM, Roe DR, Cheatham TE: Highly sampled tetranucleotide and tetraloop motifs enable evaluation of common RNA force fields. *RNA* 2015, 21:1578–1590. [PubMed: 26124199]

56. Havrila M, Zgarbova M, Jurecka P, Banas P, Krepl M, Otyepka M, Šponer J: Microsecond-scale MD simulations of HIV-1 DIS kissing-loop complexes predict bulged-in conformation of the bulged bases and reveal interesting differences between available variants of the AMBER RNA force fields. *J Phys Chem B* 2015, 119:15176–15190. [PubMed: 26548477]
57. Condon DE, Kennedy SD, Mort BC, Kierzek R, Yildirim I, Turner DH: Stacking in RNA: NMR of four tetramers benchmark molecular dynamics. *J Chem Theory Comput* 2015, 11:2729–2742. [PubMed: 26082675]
58. Steinbrecher T, Latzer J, Case DA: Revised AMBER parameters for bioorganic phosphates. *J Chem Theory Comput* 2012, 8:4405–4412. [PubMed: 23264757]
59. Bergonzo C, Cheatham TE: Improved force field parameters lead to a better description of RNA structure. *J Chem Theory Comput* 2015, 11:3969–3972. [PubMed: 26575892]
60. Izadi S, Anandakrishnan R, Onufriev AV: Building water models: a different approach. *J Phys Chem Lett* 2014, 5:3863–3871. [PubMed: 25400877]
61. Wheatley EG, Pieniazek SN, Mukerji I, Beveridge DL: Molecular dynamics of a DNA holliday junction: the inverted repeat sequence d(CCGGTACCGG)₄. *Biophys J* 2012, 102:552–560. [PubMed: 22325278]
62. Wolski P, Nieszporek K, Panczyk T: G-Quadruplex and i-motif structures within the telomeric DNA duplex. A molecular dynamics analysis of protonation states as factors affecting their stability. *J Phys Chem B* 2019, 123:468–479. [PubMed: 30589547]
63. Deng NJ, Cieplak P: Free energy profile of RNA hairpins: a molecular dynamics simulation study. *Biophys J* 2010, 98:627–636. [PubMed: 20159159]
64. Bottaro S, Banáš P, Šponer J, Bussi G: Free energy landscape of GAGA and UUCG RNA tetraloops. *J Phys Chem Lett* 2016, 7:4032–4038. [PubMed: 27661094]
65. Suddala KC, Price IR, Dandpat SS, Janecek M, Kührová P, Šponer J, Banáš P, Ke A, Walter NG: Local-to-global signal transduction at the core of a Mn²⁺ sensing riboswitch. *Nat Commun* 2019, 10:1–16. [PubMed: 30602773]
66. Gellert M, Lipsett MN, Davies DR: Helix formation by guanylic acid. *Proc Natl Acad Sci U S A* 1962, 48:2013–2018. [PubMed: 13947099]
67. Todd AK, Haider SM, Parkinson GN, Neidle S: Sequence occurrence and structural uniqueness of a G-quadruplex in the human c-Kit promoter. *Nucleic Acids Res* 2007, 35:5799–5808. [PubMed: 17720713]
68. Bhattacharyya D, Mirihana Arachchilage G, Basu S: Metal cations in G-quadruplex folding and stability. *Front Chem* 2016, 4:38. [PubMed: 27668212]
69. Gray RD, Chaires JB: Linkage of cation binding and folding in human telomeric quadruplex DNA. *Biophys Chem* 2011, 159:205–209. [PubMed: 21764207]
70. Haider S, Parkinson GN, Neidle S: Molecular dynamics and principal components analysis of human telomeric quadruplex multimers. *Biophys J* 2008, 95:296–311. [PubMed: 18375510]
71. Havrila M, Stadlbauer P, Islam B, Otyepka M, Šponer J: Effect of monovalent ion parameters on molecular dynamics simulations of G-quadruplexes. *J Chem Theory Comput* 2017, 13:3911–3926. [PubMed: 28657760]
72. Zhang X, Xu CX, Di Felice R, Šponer J, Islam B, Stadlbauer P, Ding Y, Mao L, Mao ZW, Qin PZ: Conformations of human telomeric G-quadruplex studied using a nucleotide-independent nitroxide label. *Biochemistry* 2016, 55:360–372. [PubMed: 26678746]
73. Havrila M, Stadlbauer P, Kührová P, Banáš P, Mergny J-L, Otyepka M, Šponer J: Structural dynamics of propeller loop: towards folding of RNA G-quadruplex. *Nucleic Acids Res* 2018, 46:8754–8771. [PubMed: 30165550]
74. Islam B, Stadlbauer P, Gil-Ley A, Pérez-Hernández G, Haider S, Neidle S, Bussi G, Banas P, Otyepka M, Šponer J: Exploring the dynamics of propeller loops in human telomeric DNA quadruplexes using atomistic simulations. *J Chem Theory Comput* 2017, 13:2458–2480. [PubMed: 28475322]
75. Šponer J, Bussi G, Stadlbauer P, Kührová P, Banáš P, Islam B, Haider S, Neidle S, Otyepka M: Folding of guanine quadruplex molecules—funnel-like mechanism or kinetic partitioning? An overview from MD simulation studies. *Biochim Biophys Acta - Gen Subj* 2017, 1861:1246–1263. [PubMed: 27979677]

76. Fadrná E, Špačková N, Štefl R, Kořánek J, Cheatham TE, Šponer J: Molecular dynamics simulations of guanine quadruplex loops: advances and force field limitations. *Biophys J* 2004, 87:227–242. [PubMed: 15240460]
77. Gkionis K, Kruse H, Platts JA, Mládek A, Kořánek J, Šponer J: Ion binding to quadruplex DNA stems. Comparison of MM and QM descriptions reveals sizable polarization effects not included in contemporary simulations. *J Chem Theory Comput* 2014, 10:1326–1340. [PubMed: 26580197]
78. Špačková N, Berger I, Šponer J: Nanosecond molecular dynamics simulations of parallel and antiparallel guanine quadruplex DNA molecules. *J Am Chem Soc* 1999, 121:5519–5534.
79. Štefl R, Cheatham TE, Špačková N, Fadrná E, Berger I, Kořánek J, Šponer J: Formation pathways of a guanine-quadruplex DNA revealed by molecular dynamics and thermodynamic analysis of the substates. *Biophys J* 2003, 85:1787–1804. [PubMed: 12944293]

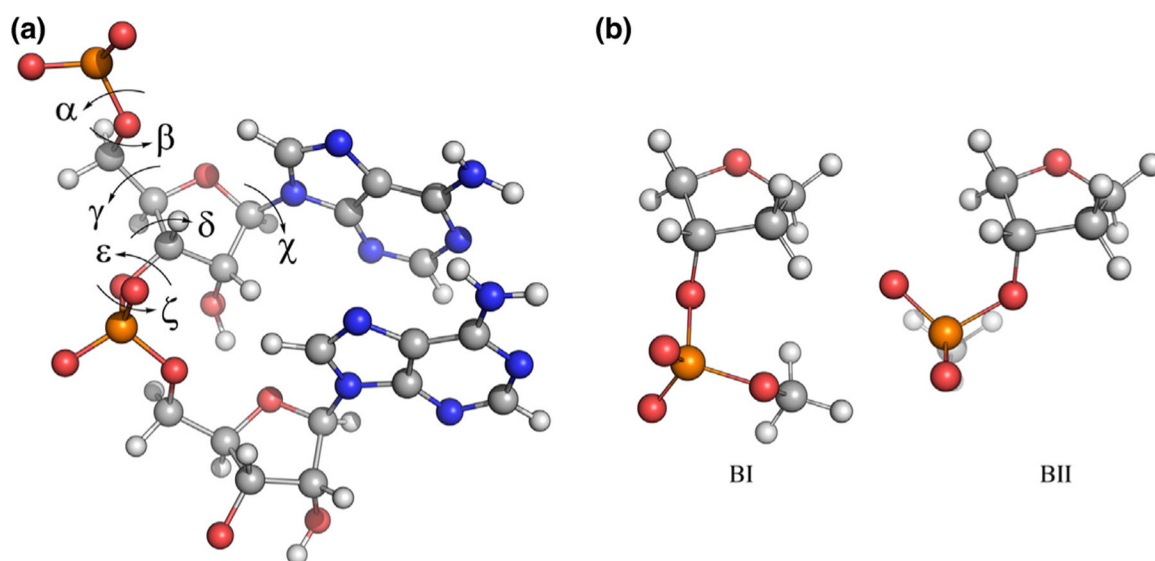


Figure 1. Key dihedrals and conformations in nucleic acids. **(a)** r(AA) dinucleotide with backbone and glycosidic torsions labeled. **(b)** BI and BII substates of DNA as represented by a model compound encompassing the deoxyribofuranose ring and 3'-methylphosphate group.

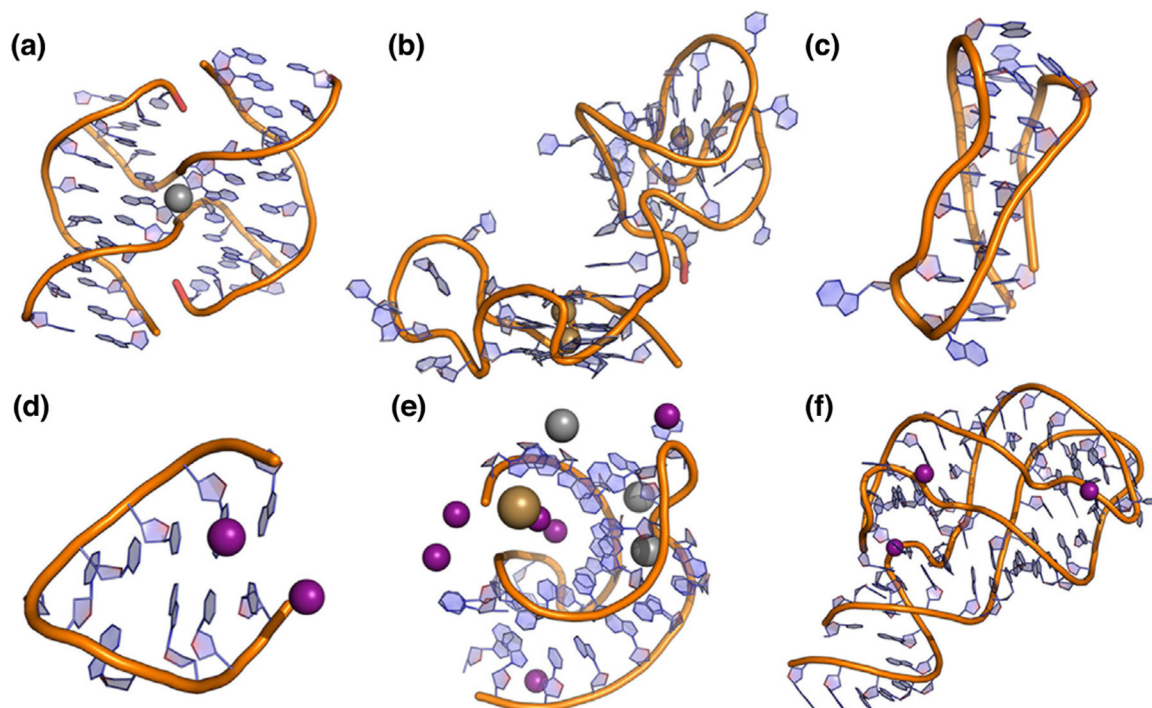


Figure 2. Examples of the diverse types of nucleic acids studied with MD simulations. Highly ordered DNA structures studied include the (a) Holliday junction structure of d(CCGGTACCGG), (b) linked *c-kit2/c-kit⁺* GQs, and (c) human telomeric i-motif. RNA structures include (d) UUCG hairpins, (e) viral RNA pseudoknots, and (f) add adenine riboswitch. Ions participating in these structures include K⁺ (tan), Na⁺ (grey), and Mg²⁺ (purple).

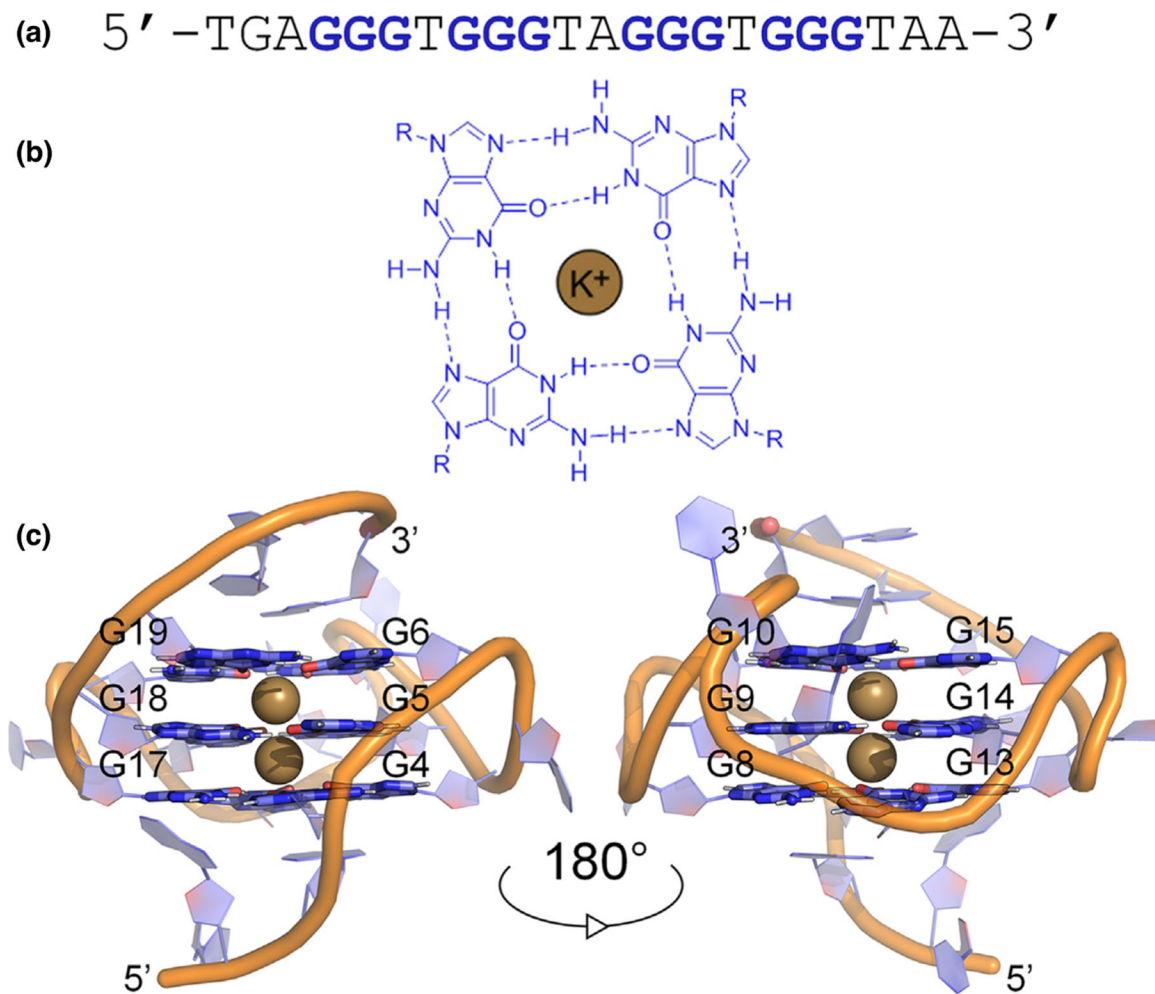


Figure 3. Details of the structural hierarchy of the major GQ structure formed in the human c-MYC promoter. (a) Guanine-rich sequences with four runs of at least three guanines self-associate into tetrads, consisting of a square planar arrangement of (b) Hoogsteen hydrogen-bonded guanines. (c) Monovalent metal ions, typically K^+ (tan), stabilize the tetrad core and the highly ordered structure.

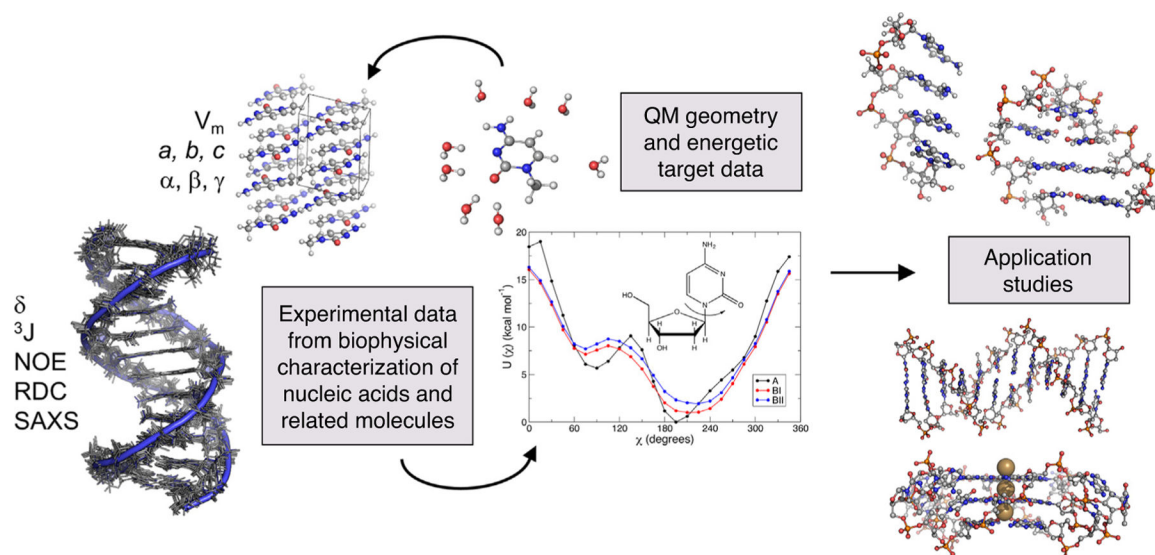


Figure 4. Overview of nucleic acid FF development and applications. Refinement of FFs is driven by comparisons to available experimental data on nucleic acids and suitable model compounds, as well as QM calculations that describe intermolecular interactions, electrostatic properties, and conformational energetics, and so on. Application studies on canonical and noncanonical nucleic acid structures can be performed after sufficiently rigorous FF parametrization and validation toward a greater understanding of nucleic acid structure-function relationships.