On the importance of accounting for nuclear quantum effects in ab initio calibrated force fields in biological simulations

Leonid Pereyaslavets[a,1], Igor Kurnikov[a], Ganesh Kamath[a], Oleg Butin[a], Alexey Illarionov[a], Michael Olevanov[a], Michael Levitt[b,1], Roger D. Kornberg[b], and Boris Fain[a]

[a]InterX Inc., Berkeley, CA 94710; and [b]Department of Structural Biology, Stanford University School of Medicine, Stanford, CA 94305

In many important processes in chemistry, physics, and biology the nuclear degrees of freedom cannot be described using the laws of classical mechanics. At the same time, the vast majority of molecular simulations that employ wide-coverage force fields treat atomic motion classically. In light of the increasing desire for and accelerated development of quantum mechanics (QM)-parameterized interaction models, we reexamine whether the classical treatment is sufficient for a simple but crucial chemical species: alkanes. We show that when using an interaction model or force field in excellent agreement with the "gold standard" QM data, even very basic simulated properties of liquid alkanes, such as densities and heats of vaporization, deviate significantly from experimental values. Inclusion of nuclear quantum effects via techniques that treat nuclear degrees of freedom using the laws of classical mechanics brings the simulated properties much closer to reality.

nuclear quantum effect | path integral molecular dynamics | force field | ab initio | alkanes

Historically, simulation of biomedically important systems has been done with the empirical force fields (FFs) pioneered 50 years ago (1, 2). As computer power has increased and quantum mechanical energy calculations have become more accurate than experiment, FFs derived from this quantum mechanical data have become feasible. In future quantum mechanics (QM) parameterized FFs are likely to be much more important than empirical FFs as they can be applied to new systems without needing experimental data for calibration. Being able to accurately reproduce experimental data with QM-based FFs is essential. The aim of this paper is to assess the need to include nuclear quantum effects (NQEs) in simulations. Here we show that they must be included if a QM parameterized FF is to correctly reproduce bulk properties.

Molecular mechanics (3, 4) bridges the precise calculation of interactions of a few atoms at rest by ab initio methods (QM) (4, 5) with the larger numbers of atoms and dynamic simulations over long time scales essential to predict measured bulk properties of matter. The starting component of molecular mechanics is an effective Newtonian interaction model, or FF (4, 6). In this work we employ a FF parameterized by agreement with QM-derived energies, forces, and monomer properties (we expand on this choice in *Discussion*). Once the interaction energies and forces have been adequately described, the quantities of interest are obtained by sampling the potential surface of a large molecular ensemble via various means, usually Monte Carlo or Newtonian molecular dynamics (MD) (3, 4, 7). The sampling machinery encompasses many critical parts such as proper maintenance of ensemble temperature and pressure, careful integration and truncation practices, enhanced sampling techniques that overcome energy barriers, limitations of machine precision, and many others (3, 7–13).

In Monte Carlo and MD sampling, the Born–Oppenheimer approximation (14) decouples the electron degrees of freedom from nuclear motion: the former are treated as instantaneous and the latter, conventionally, as classical Newtonian point masses.

Atoms and nuclei, however, do not behave classically (3, 15–18) and quantum uncertainty in atomic position may significantly impact the structure and properties of the ensemble. The importance of these NQEs is widely known and accepted by researchers working on, for example, path integral MD (PIMD) and ab initio MD (15), as well as by the community working on water models. The focus on water is natural since water, due to its importance, ubiquity, and anomalies is the traditional starting point for FF development. However, water simulations can actually obscure the importance of NQEs because the various effects oppose and thus partially cancel each other (19). At this time the widely used FFs and simulation packages rarely, if ever, consider NQEs. There are two reasons for this. First, these FFs are empirical and thus fitted to reproduce experimental results; they therefore contain the NQE corrections implicitly. Second, as the vast majority of investigations of NQE have concentrated primarily on water (15, 20) or systems with extreme physical conditions that enhance NQE (21, 22), the errors caused by leaving them out are thought to be minor compared with the accuracy of the FFs themselves, and hence with the expected accuracy of predictions made for biological systems at room temperature and pressure.

Significance

In molecular modeling the motion of nuclei, especially hydrogen, cannot be described using the laws of classical mechanics. The importance of nuclear quantum effects has long been appreciated by the ab initio molecular dynamics and by the water simulation communities. However, the vast majority of simulations of biological systems performed at ambient conditions treat atomic motion classically. Even in the new-generation force fields parameterized from quantum mechanics these effects are thought to be minor compared with other inaccuracies at room temperature and pressure. We show that a force field in excellent agreement with quantum mechanical energies and forces will not produce acceptably inaccurate predictions at ambient conditions unless the nuclear motion and interaction are accounted for in the simulation.

Valuable exceptions are the works of Martyna and coworkers, who investigated NQEs for weakly interacting species: the noble gases helium and xenon (17) and alkanes (16, 23). These researchers discovered that the NQE, as modeled by PIMD (17, 23), significantly alters the molar volume of weakly interacting liquid alkanes (16, 23). Since alkanes are literally the backbone of all organic molecules (16), this work emphatically suggests that the inclusion of NQE should not be limited to strongly interacting systems. Some of the effects may, in fact, be more prominent in weaker interactions such as those occurring in nonpolar environments. However, as Martyna and coworkers used a relatively simple empirical FF which did not closely agree with QM energies, the question of whether simulations with a quantum mechanically parameterized FF must include NQE remained open.

The ArrowFF is InterX's medium-precision polarizable FF. Agreement with QM properties and energies is only one of many several guiding principles in ArrowFF design. Some of the other considerations are, for example, transferability, economy of parameters, and computational tractability. Consequently, for some functional groups and atoms, we allow the ArrowFF to deviate from QM benchmarking in a carefully restrained way.

However, as the aim of this study is to highlight the contribution NQEs make to a QM-faithful FF, we will narrow our focus to a chemical subset ArrowFF describes very well: the alkanes. Alkanes are generally easier to describe than many other chemical species because they are fairly isotropic, nonpolar, and have low three-body interaction energy. In fact, even coarse-grained united-atom representations (24, 25) describe alkanes fairly well. This does not narrow our conclusions because alkanes are ubiquitous and important in biological molecules and the drugs that target them. We first confirm that the ArrowFF does indeed represent alkanes faithfully, making them a perfect system on which to demonstrate the contribution of NQE for prediction of ensemble properties. We then show that although the ArrowFF of alkanes is in excellent agreement with high-level QM, NQE must be included for accurate calculation of very basic bulk properties such as densities and heats of vaporization.

Methods

We use the ArrowFF with the following features: nonbonded terms are dipole-polarizable multipolar exchange-repulsion and electrostatics that suitably describe charge penetration and delocalization, as well as Tang–Toennies-damped (26) spherical C6 and C8 dispersion terms. The functional form of the bonded interactions is taken from MMFF94 (27), with force constants and equilibrium values fitted to QM energies at the MP2/aug-cc-pVTZ level of theory. Multibody effects are modeled by polarization; there are no explicit three-body terms [e.g., Axilrod–Teller (28)] for alkanes. More details can be found in refs. 29–31. The ArrowFF alkanes' parameters are not specifically designed for alkanes but are a part of the FF broadly covering protein and ligand functional groups. Nevertheless, as we show below, the ArrowFF describes the alkanes particularly well.

We model NQEs via PIMD (13, 16), which is based on the Feynman path integral formulation of QM (32). In PIMD, a quantum system is mapped onto an extended classical system consisting of multiple replicas of the original classical system with harmonic springs linking adjacent replicas of each atom of the system (17). As the number of replicas increases, the ensemble averages of the observables of the extended classical system converge to those of the quantum system (3, 17, 33). Simulation is performed on an isothermal–isobaric ensemble (NPT) modeled using a Nosé–Hoover chain thermostat (12) and a Martyna–Tuckerman–Tobias–Klein barostat (34). Electrostatic interactions are truncated with multipolar PME (35) at the cutoff of 9 A. A multiple time step technique (36) with primitive integration is used to speed up PIMD calculations. Bonded and interreplica harmonic interactions, which are easier to compute, are sampled with the time step of 0.25 fs, while the more expensive nonbonded interactions are sampled with the time step of 2 fs. Multiple time steps and PIMD are implemented in our Arbalest MD simulation software. In Arbalest the nonbonded interactions are computed efficiently on NVIDIA graphical processing units.

To ensure that the simulation discrepancies do not arise from the imprecision of the FF or the underlying QM calculations, we require both a high level of theory and a good agreement between FF and QM energies. We calculate QM dimer energies at the highest QM level reasonable for large-scale calculations: MP2/CBS [based on Helgaker two-point extrapolation

(37) from aug-cc-pVTZ to aug-cc-pVQZ basis sets] + CCSD(T)/aug-cc-pVTZ - MP2/aug-cc-pVTZ basis sets. This level of theory is known as the "gold standard" (38) and is commonly used as a benchmark in the computational chemistry community. The ArrowFF is fitted to these gold-standard energy values. In Fig. 1A we plot energy values for the ArrowFF with corresponding QM energies for a diverse set of alkane dimers. In Fig. 1 B and C we show the agreement between ArrowFF and QM along the dissociation curves for the minima of methane–methane and ethane–ethane dimers.

As seen from Fig. 1 A–C the agreement between ArrowFF and QM dimer energies is excellent everywhere, including the high-energy regions of significant electron overlap.

A direct proof that many-body ArrowFF energies also agree with their QM counterparts is difficult because precise determination of QM energies of large clusters is computationally costly. For alkanes, however, we can infer this indirectly. Using a high level of theory [the level of theory for the trimer is HF(aug-cc-pV5Z) + df-MP2/(aug-cc-pVQZ ->aug-cc-pV5Z) + CCSD(T)/(aug-cc-pVTZ ->aug-cc-pVQZ)-MP2/(aug-cc-pVTZ ->aug-cc-pVQZ)] we calculate the nonadditive trimerization energy of the optimal methane trimer to be 0.014 kcal/mol. This small value for the optimal trimer (much less than, for example, the intermolecular dimerization zero-point vibration energy of 0.26 kcal/mol) means that many-body corrections arise almost exclusively from the nonadditivity of the induction term. The agreement of this term with QM is illustrated in Fig. 1D, where we display the correspondence of the QM and ArrowFF induction energies along a dissociation curve for the methane–methane dimer. We therefore believe that ArrowFF reproduces the molecular interactions in bulk liquid alkanes well enough for the current investigation.

Results

Having shown that our alkane model is accurate, we proceed to calculate the densities and heats of vaporization of several alkane systems. Fig. 2 A and B show the deviations between the calculated and experimental densities for alkanes and water using $n = 1, 4, 8,$ or 16 PIMD beads. Although the results have not converged fully with 16 beads, we consider this range of bead numbers (n) sufficient for the conclusions of this paper. As a precaution, we computed methane properties with 32 and 64 beads with a further reduced time step. The results presented in *SI Appendix* confirm that truncating at 16 beads is reasonable.

The computed classical ($n = 1$) densities deviate by as much as 11% from the experimental values. This is a disturbingly large discrepancy: the community traditionally expects predicted densities
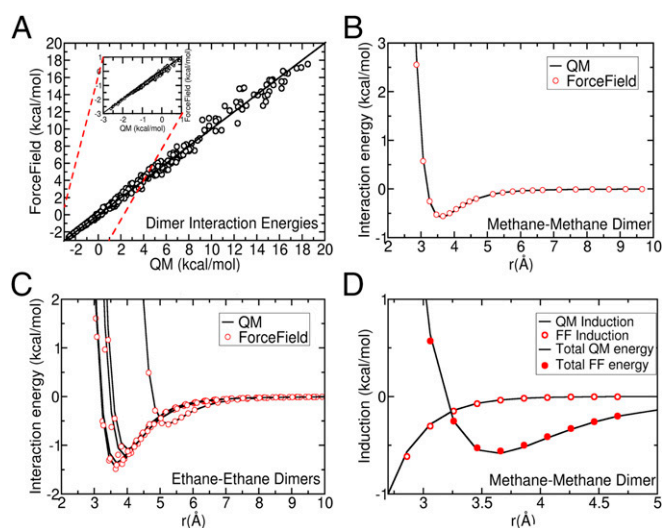
Fig. 1. *A* shows that dimer ArrowFF interaction energies agree well with the corresponding QM energies for a representative set of four alkanes combined in six alkane dimer species (2,526 poses total). The *Inset* magnifies the important attractive region. The agreement between ArrowFF and QM is good as shown by (*B*) the dissociation curves for the minimal methane dimer, (*C*) the dissociation curves of five local minima of the ethane–ethane dimer, and (*D*) the induction energy of the methane dimer along the dissociation curve.
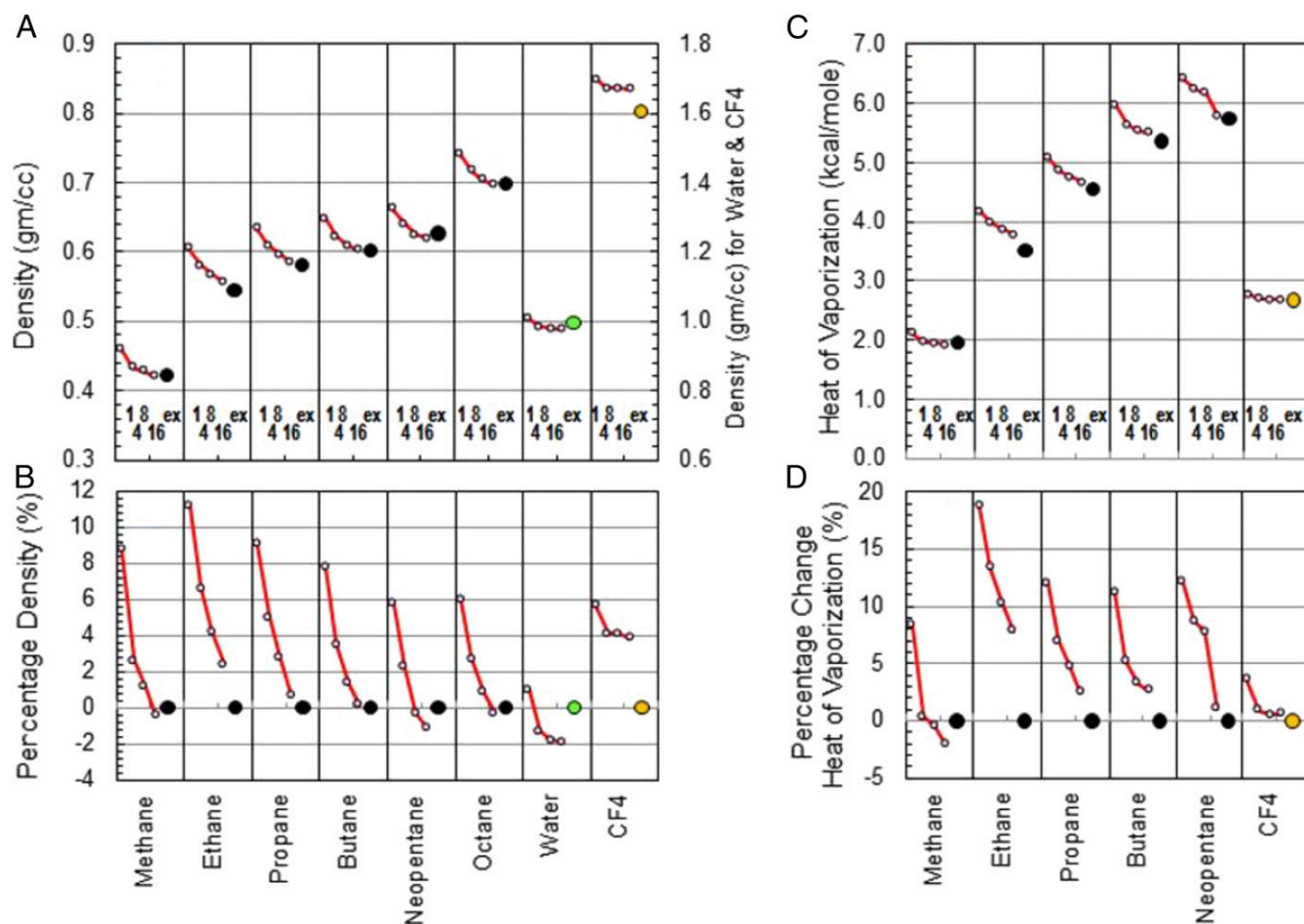
Fig. 2. *A* and *B* show the densities and their percentage errors for representative alkanes (methane, ethane, propane, butane, octane, and neopentane) as well as for water and CF$_4$. *C* and *D* show the heats of vaporization values and their percentage errors for methane, ethane, propane, butane, neopentane, and CF$_4$. The error decreases as the number of PIMD beads increases (1, 4, 8, and 16). The experimental value is shown as a large black dot for alkanes, a green dot for water, and an orange dot for CF$_4$. The simulation temperatures for the systems were 112, 184, 231, 272, 298, 298, 298, and 145 K for methane, ethane, propane, butane, neopentane, octane, water, and CF$_4$, respectively. The simulation box is a 32-Å cube; each system was run for 500 ps at a pressure of 1 bar. The error bars for density are within 0.01 g/cm^3 and for heat of vaporization within 0.02 kcal/mol and smaller than the symbol size in the figures.

to be at most within 3–4% of experiment (24, 39, 40). Similar deviations are seen in heat of vaporization values (Fig. 2 *C* and *D*). It is unlikely that deficiencies in the ArrowFF are responsible for the errors because the models reproduce the QM energies well (Fig. 1). It is also unlikely that we have significant bond-length errors as we use the minimized QM ground-state geometries.

Once the NQEs are incorporated by having more than one bead per atom, the discrepancies gradually decrease to below 2%, which is very acceptable for a purely QM parameterization. Note that due to several opposing subeffects mentioned above the systematic NQE shift for water is much smaller than that for alkanes (Fig. 2 *A* and *B*).

The magnitudes of the density reductions we find here as the number of beads increases are consistent with the changes in effective volume observed by Martyna and coworkers (16) and Martyna et al. (23) with Amber95 (41): ~3% for octane and ~9% for butane. The volume changes reported by Martyna and coworkers (16) and Martyna et al. (23) also differ slightly between Amber95 and CHARMM22, thus making the imperfect agreement with the ArrowFF here reasonable.

To confirm that the above results were caused by the hydrogen-related NQE we damped the effects by substituting hydrogen with a much heavier fluorine. The FF parameters for CF$_4$ were constructed via exactly the same procedure we use for parameterization of alkanes

and, indeed, all other atom types. The model accuracy for CF$_4$ (Fig. 3 *A* and *B*) is good and similar in quality to the alkane models. The simulation results (Fig. 2) show that in the classical approximation of just one bead the density of CF$_4$ is much closer to experiment than the density of methane. Also, as expected, the calculated density of CF$_4$ does not decrease appreciably with the number of PIMD beads. Fig. 3 *C* and *D* show the radial distribution functions (RDFs) for CH$_4$ and CF$_4$ for classical ($n = 1$) and PIMD simulations ($n = 16$). Again, the structure of CH$_4$ is significantly altered by PIMD, whereas the structure of CF$_4$ remains almost the same. To ensure that our conclusions are not influenced by the differences in the center-of-mass fluctuations between CH$_4$ and CF$_4$ we also constructed a fictitious CH$_4$ molecule where the weight of the central carbon was increased to 84 a.u. The results, presented in *SI Appendix*, do not alter our conclusions. The much smaller deviations from experiment and the relative insensitivity to the number of beads in CF$_4$ compared with CH$_4$ affirms our assertion that NQEs are the primary cause of deviation from experiment of alkane density and heat of vaporization values.

Discussion

We have shown that NQEs significantly change ensemble predictions made by molecular mechanics simulations of biological systems. We have further shown that when the FF is a good fit to QM energies these changes are needed to bring the results of
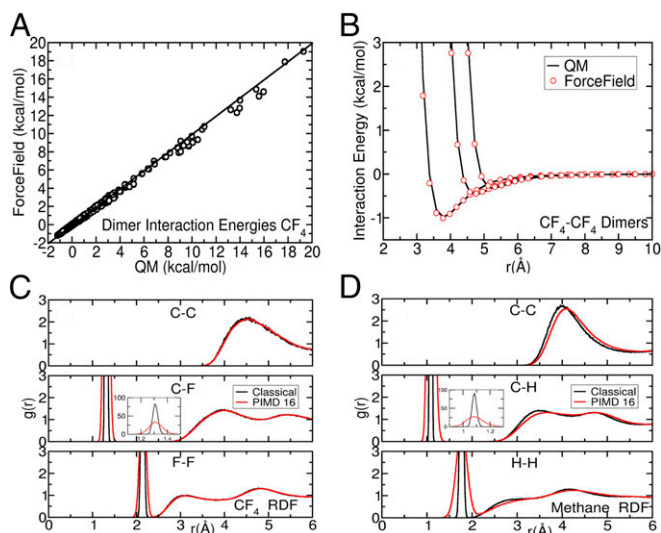
Fig. 3. *A* shows agreement between ArrowFF and QM energies for CF_4 dimers. *B* shows that ArrowFF and QM energies are very similar along the dissociation curves of the three minimum energy dimer geometries of CF_4 (similar to Fig. 1*B*). We also show the RDFs for (*C*) CF_4 and (*D*) CH_4. The RDF of CH_4 is significantly shifted when we include NQE with 16 beads, whereas that of CF_4 remains largely unchanged.

calculations closer to experimental values. The important conclusion from these results is that QM-parameterized FFs, carefully parameterized to reproduce accurate QM data, give substantial errors unless NQEs are included.

Whether or not one should include the NQE in simulations, via PIMD, centroid MD (42, 43), or other available methods is contingent on several factors. Clearly, if one is interested in investigating the isotope effect, one has no choice but to include NQEs in the simulation. The recommendation to the practitioner seeking predictive values for ensemble properties (e.g., energies of solvation, etc.) depends on the choice of parameterization paradigm, the computational priorities, and the available expertise. Properly implementing PIMD is nontrivial (15–17, 23), but much guidance is available, and so the primary drawback of PIMD is computational cost. Recently, several promising bead truncation methods (44–48) reduce the amount of computational overhead needed to sufficiently capture the NQE corrections to approximately four times that of the classical simulation. This leaves the nature of the FF parameterization as the main factor in deciding whether to include NQE in the simulations.

All current wide-coverage FFs infer some of their parameters empirically by requiring that the computed ensemble properties agree with experimental values such as heats of evaporation, densities, solvation energies, dielectric constants, and so on (24, 39, 49–51). If one follows this parameterization paradigm, one can fold the NQE into the framework of fitting to experiment as is done for other unaccounted-for effects. After all, the corrections to computed properties due to NQEs are of the same order of magnitude as

other frequently neglected effects such as polarization, penetration effect, functional form choices, three-body dispersion, multipole truncation, and so on (6, 52). For example, the approach taken by the AMOEBA FF (51, 53, 54) of getting some parameters (e.g., electrostatics) from QM calculations and others (van der Waals, vdW) from judicious fitting to experimental values makes good sense. Such implicit inclusion of NQEs (and other factors) into the vdW parameterization is reminiscent of how the first-generation classical FFs [e.g., GROMOS (24, 55), Merck (27, 40), GAFF (50), OPLS (25, 39, 40, 56), and CFF (8, 57)] implicitly included polarization by increasing partial charges in the liquid phase (58, 59).

Although the experimentally fitted FFs can and do achieve good agreement with the training set measurements, they encounter serious transferability issues. One of the reasons for this is the difficulty of properly choosing which parameters to adjust and by how much. The problem is illustrated by Hobza and coworkers (60), Sherrill et al. (61), and Demerdash et al. (62), who showed that although the agreement of the total energy for parameterized molecules is rather good, the individual components can contain significant deviations from their QM counterparts. These deviations then trigger a chain of transferability failures and total energy errors once the model is extended to new chemical species or even to pair interactions not considered in the training set.

The other choice of FF development is to design the FF to agree either in toto or componentwise with some or all of the following QM derived quantities: the intramolecular energies ("bonded terms") (63), the intermolecular two-body and higher energies ("nonbonded terms") (6, 63), nonadditive energies (e.g., polarization) (52), and the properties of single molecules (e.g., multipole moments, ESP-fitted charges) (6, 64) and small clusters. FF parameterization by QM calculations started in 1970s (65–68) and because of dramatic increases in computational resources has become fully feasible today. As QM computations are limited only by molecular size, the QM parameterized approach can be directly applied to molecules that have not had their properties measured, or, especially, have not yet been synthesized. Furthermore, recent computational advances, for example symmetry adapted perturbation theory (SAPT) (69, 70) or density functional theory (DFT)-SAPT (71–73) help the transferability of FFs based on QM data. DFT-SAPT decomposes the interactions into individual components which can then be accurately fitted by suitable FF functional forms and then subsequently extended by their corresponding appropriate transferability rules. Because of these advantages the current generation of advanced FFs (18, 29–31, 51, 53, 54, 74, 75) are increasingly adopting the QM–FF correspondence as a guiding principle. To aid in this development, we have shown here that if an FF is parameterized solely by fitting ab initio QM data (or is itself solely ab initio) (47), NQEs must be accounted for in the simulation.

1. Lifson S, Warshel A (1968) Consistent force field for calculations of conformations, vibrational spectra, and enthalpies of cycloalkane and n-alkane molecules. *J Chem Phys* 49:5116–5129.
2. Levitt M, Lifson S (1969) Refinement of protein conformations using a macromolecular energy minimization procedure. *J Mol Biol* 46:269–279.
3. Allen MP, Tildesley DJ (1987) *Computer Simulation of Liquids* (Oxford Univ Press, New York).
4. Cramer CJ (2004) *Essentials of Computational Chemistry: Theories and Models* (Wiley, Chichester, UK), 2nd Ed.
5. Jensen F (2007) *Introduction to Computational Chemistry* (Wiley, Chichester, UK), 2nd Ed.
6. Stone A (2013) *The Theory of Intermolecular Forces* (Oxford Univ Press, Oxford), 2nd Ed.
7. Frenkel D, Smit B (2001) *Understanding Molecular Simulation: From Algorithms to Applications* (Academic, New York), p 664.

8. Levitt M, Hirshberg M, Sharon R, Daggett V (1995) Potential energy function and parameters for simulations of the molecular dynamics of proteins and nucleic acids in solution. *Comput Phys Commun* 91:215–231.
9. Darden T, York D, Pedersen L (1993) Particle mesh Ewald: An N·log(N) method for Ewald sums in large systems. *J Chem Phys* 98:10089–10092.
10. Boateng HA, Todorov IT (2015) Arbitrary order permanent Cartesian multipolar electrostatic interactions. *J Chem Phys* 142:034117.
11. Nose S (1991) Constant temperature molecular dynamics methods. *Prog Theor Phys Suppl* 103:1–46.
12. Martyna GJ, Klein ML, Tuckerman M (1992) Nosé–Hoover chains: The canonical ensemble via continuous dynamics. *J Chem Phys* 97:2635–2643.
13. Martyna GJ, Tobias DJ, Klein ML (1994) Constant pressure molecular dynamics algorithms. *J Chem Phys* 101:4177–4189.

Pereyaslavets et al.

14. Born M, Oppenheimer JR (1927) Zur Quantentheorie der Molekeln. *Ann Phys* 389: 457–484.
15. Ceriotti M, et al. (2016) Nuclear quantum effects in water and aqueous systems: Experiment, theory, and current challenges. *Chem Rev* 116:7529–7550.
16. Balog E, Hughes AL, Martyna GJ (1999) Constant pressure path integral molecular dynamics studies of quantum effects in the liquid state properties of n-alkanes. *J Chem Phys* 112:870–880.
17. Tuckerman ME, Berne BJ, Martyna GJ, Klein ML (1993) Efficient molecular dynamics and hybrid Monte Carlo algorithms for path integrals. *J Chem Phys* 99:2796–2808.
18. Stern HA, Berne BJ (2001) Quantum effects in liquid water: Path-integral simulations of a flexible and polarizable ab initio model. *J Chem Phys* 115:7622–7628.
19. Habershon S, Markland TE, Manolopoulos DE (2009) Competing quantum effects in the dynamics of a flexible water model. *J Chem Phys* 131:024501.
20. Kuharski RA, Rossky PJ (1985) A quantum mechanical study of structure in liquid H_2O and D_2O. *J Chem Phys* 82:5164–5177.
21. Morales MA, McMahon JM, Pierleoni C, Ceperley DM (2013) Nuclear quantum effects and nonlocal exchange-correlation functionals applied to liquid hydrogen at high pressure. *Phys Rev Lett* 110:065702.
22. Errea I, et al. (2015) High-pressure hydrogen sulfide from first principles: A strongly anharmonic phonon-mediated superconductor. *Phys Rev Lett* 114:157004.
23. Martyna GJ, Hughes A, Tuckerman ME (1999) Molecular dynamics algorithms for path integrals at constant pressure. *J Chem Phys* 110:3275–3290.
24. Horta BAC, et al. (2016) A GROMOS-compatible force field for small organic molecules in the condensed phase: The 2016H66 parameter set. *J Chem Theory Comput* 12: 3825–3850.
25. Jorgensen WL, Madura JD, Swenson CJ (1984) Optimized intermolecular potential functions for liquid hydrocarbons. *J Am Chem Soc* 106:6638–6646.
26. Tang KT, Toennies JP (1984) An improved simple model for the van der Waals potential based on universal damping functions for the dispersion coefficients. *J Chem Phys* 80:3726–3741.
27. Halgren TA (1996) Merck molecular force field. I. Basis, form, scope, parameterization, and performance of MMFF94. *J Comput Chem* 17:490–519.
28. Axilrod BM, Teller E (1943) Interaction of the van der Waals type between three atoms. *J Chem Phys* 11:299–300.
29. Donchev AG, Galkin NG, Pereyaslavets LB, Tarasov VI (2006) Quantum mechanical polarizable force field (QMPFF3): Refinement and validation of the dispersion interaction for aromatic carbon. *J Chem Phys* 125:244107.
30. Donchev AG, et al. (2008) Assessment of performance of the general purpose polarizable force field QMPFF3 in condensed phase. *J Comput Chem* 29:1242–1249.
31. Kamath G, et al. (2016) Prediction of cyclohexane-water distribution coefficient for SAMPL5 drug-like compounds with the QMPFF3 and ARROW polarizable force fields. *J Comput Aided Mol Des* 30:977–988.
32. Feynman RP, Hibbs AR (1965) *Quantum Mechanics and Path Integrals* (McGraw-Hill, New York).
33. Chandler D, Wolynes PG (1981) Exploiting the isomorphism between quantum theory and classical statistical mechanics of polyatomic fluids. *J Chem Phys* 74:4078–4095.
34. Martyna GJ, Tuckerman ME, Tobias DJ, Klein ML (1995) Explicit reversible integrators for extended systems dynamics. *Mol Phys* 87:1117–1157.
35. Giese TJ, Panteva MT, Chen H, York DM (2015) Multipolar Ewald methods, 1: Theory, accuracy, and performance. *J Chem Theory Comput* 11:436–450.
36. Tuckerman M, Berne BJ, Martyna GJ (1992) Reversible multiple time scale molecular dynamics. *J Chem Phys* 97:1990–2001.
37. Halkier A, Helgaker T, Jørgensen P, Klopper W, Olsen J (1999) Basis-set convergence of the energy in molecular Hartree–Fock calculations. *Chem Phys Lett* 302:437–446.
38. Burns LA, Marshall MS, Sherrill CD (2014) 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* 141:234111.
39. Jorgensen WL, Maxwell DS, Tirado-Rives J (1996) Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids. *J Am Chem Soc* 118:11225–11236.
40. Kaminski G, Jorgensen WL (1996) Performance of the AMBER94, MMFF94, and OPLS-AA force fields for modeling organic liquids. *J Phys Chem* 100:18010–18013.
41. Cornell WD, et al. (1995) A second generation force field for the simulation of proteins, nucleic acids, and organic molecules. *J Am Chem Soc* 117:5179–5197.
42. Cao J, Voth GA (1994) The formulation of quantum statistical mechanics based on the Feynman path centroid density. IV. Algorithms for centroid molecular dynamics. *J Chem Phys* 101:6168–6183.
43. Hone TD, Rossky PJ, Voth GA (2006) A comparative study of imaginary time path integral based methods for quantum dynamics. *J Chem Phys* 124:154103.
44. Vishnevskiy YV, Tikhonov D (2016) Quantum corrections to parameters of interatomic distance distributions in molecular dynamics simulations. *Theor Chem Acc* 135:88.
45. Poltavsky I, Tkatchenko A (2016) Modeling quantum nuclei with perturbed path integral molecular dynamics. *Chem Sci (Camb)* 7:1368–1372.
46. Markland TE, Manolopoulos DE (2008) An efficient ring polymer contraction scheme for imaginary time path integral simulations. *J Chem Phys* 129:024105.
47. Marsalek O, Markland TE (2016) Ab initio molecular dynamics with nuclear quantum effects at classical cost: Ring polymer contraction for density functional theory. *J Chem Phys* 144:054112.
48. Ceriotti M, Manolopoulos DE (2012) Efficient first-principles calculation of the quantum kinetic energy and momentum distribution of nuclei. *Phys Rev Lett* 109: 100604.
49. Mackerell AD, Jr (2004) Empirical force fields for biological macromolecules: Overview and issues. *J Comput Chem* 25:1584–1604.
50. Wang J, Wolf RM, Caldwell JW, Kollman PA, Case DA (2004) Development and testing of a general amber force field. *J Comput Chem* 25:1157–1174.
51. Ponder JW, et al. (2010) Current status of the AMOEBA polarizable force field. *J Phys Chem B* 114:2549–2564.
52. Khoruzhii O, et al. (2014) Polarizable force fields for proteins. *Protein Modelling*, ed Gamble A (Springer, Cham, Switzerland), pp 91–134.
53. Wang LP, et al. (2013) Systematic improvement of a classical molecular model of water. *J Phys Chem B* 117:9956–9972.
54. Wu JC, Chattree G, Ren P (2012) Automation of AMOEBA polarizable force field parameterization for small molecules. *Theor Chem Acc* 131:1138.
55. Oostenbrink C, Villa A, Mark AE, van Gunsteren WF (2004) A biomolecular force field based on the free enthalpy of hydration and solvation: The GROMOS force-field parameter sets 53A5 and 53A6. *J Comput Chem* 25:1656–1676.
56. Kaminski GA, Friesner RA, Tirado-Rives J, Jorgensen WL (2001) Evaluation and reparameterization of the OPLS-AA force field for proteins via comparison with accurate quantum chemical calculations on peptidet. *J Phys Chem B* 105:6474–6487.
57. Levitt M (1983) Molecular dynamics of native protein. I. Computer simulation of trajectories. *J Mol Biol* 168:595–617.
58. Leontyev I, Stuchebrukhov A (2011) Accounting for electronic polarization in non-polarizable force fields. *Phys Chem Chem Phys* 13:2613–2626.
59. Leontyev IV, Stuchebrukhov AA (2014) Polarizable molecular interactions in condensed phase and their equivalent nonpolarizable models. *J Chem Phys* 141:014103.
60. Zgarbová M, Otyepka M, Sponer J, Hobza P, Jurecka P (2010) 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* 12: 10476–10493.
61. Sherrill CD, et al. (2009) Assessment of standard force field models against high-quality ab initio potential curves for prototypes of π-π, CH/π, and SH/π interactions. *J Comput Chem* 30:2187–2193.
62. Demerdash O, Mao Y, Liu T, Head-Gordon M, Head-Gordon T (2017) Assessing many-body contributions to intermolecular interactions of the AMOEBA force field using energy decomposition analysis of electronic structure calculations. *J Chem Phys* 147: 161721.
63. Allinger NL (1976) Calculation of molecular structure and energy by force-field methods. *Advances in Physical Organic Chemistry*, eds Gold V, Bethell D (Academic, New York), Vol 13, pp 1–82.
64. Bayly CI, Cieplak P, Cornell WD, Kollman PA (1993) A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges: The RESP model. *J Phys Chem* 97:10269–10280.
65. Lie GC, Clementi E (1976) Calculation of the second virial coefficients for water using a recent "*ab initio*" potential. *J Chem Phys* 64:5308–5309.
66. Matsuoka O, Clementi E, Yoshimine M (1976) CI study of the water dimer potential surface. *J Chem Phys* 64:1351–1361.
67. McDonald IR, Klein ML (1978) Intermolecular potentials and the simulation of liquid water. *J Chem Phys* 68:4875–4877.
68. Clementi E, Habitz P (1983) A new two-body water-water potential. *J Phys Chem* 87: 2815–2820.
69. Rybak S, Jeziorski B, Szalewicz K (1991) Many-body symmetry-adapted perturbation theory of intermolecular interactions. H2O and HF dimers. *J Chem Phys* 95:6576–6601.
70. Szalewicz K (2012) Symmetry-adapted perturbation theory of intermolecular forces. *Wiley Interdiscip Rev Comput Mol Sci* 2:254–272.
71. Williams HL, Chabalowski CF (2001) Using Kohn–Sham orbitals in symmetry-adapted perturbation theory to investigate intermolecular interactions. *J Phys Chem A* 105: 646–659.
72. Misquitta AJ, Szalewicz K (2002) Intermolecular forces from asymptotically corrected density functional description of monomers. *Chem Phys Lett* 357:301–306.
73. Misquitta AJ, Podeszwa R, Jeziorski B, Szalewicz K (2005) Intermolecular potentials based on symmetry-adapted perturbation theory with dispersion energies from time-dependent density-functional calculations. *J Chem Phys* 123:214103.
74. Babin V, Leforestier C, Paesani F (2013) Development of a "first principles" water potential with flexible monomers: Dimer potential energy surface, VRT spectrum, and second virial coefficient. *J Chem Theory Comput* 9:5395–5403.
75. Medders GR, Babin V, Paesani F (2014) Development of a "first-principles" water potential with flexible monomers. III. Liquid phase properties. *J Chem Theory Comput* 10:2906–2910.