Box size effects are negligible for solvation free energies of neutral solutes

Sreeja Parameswaran · David L. Mobley

Abstract Hydration free energy calculations in explicit solvent have become an integral part of binding free energy calculations and a valuable test of force fields. Most of these simulations follow the conventional norm of keeping edge length of the periodic solvent box larger than twice the Lennard-Jones (LJ) cutoff distance, with the rationale that this should be sufficient to keep the interactions between copies of the solute to a minimum. However, for charged solutes, hydration free energies can exhibit substantial box size-dependence even at typical box sizes. Here, we examine whether similar size-dependence affects hydration of neutral molecules. Thus, we focused on two strongly polar molecules with large dipole moments, where any size-dependence should be most pronounced, and determined how their hydration free energies vary as a function of simulation box size. In addition to testing a variety of simulation box sizes, we also tested two LJ cut-off distances, 0.65 and 1.0 nm. We show from these simulations that the calculated hydration free energy is independent of the box-size as well as the LJ cut-off distance, suggesting that typical hydration free energy calculations of neutral compounds indeed need not be particularly concerned with finite-size effects as long as standard good practices are followed.

Keywords Hydration free energy · Box size · Free energy calculation

S. Parameswaran · D. L. Mobley
Department of Chemistry, University of New Orleans, 2000 Lakeshore Drive, New Orleans, LA 70148, USA

D. L. Mobley (✉)
Department of Pharmaceutical Sciences and Department of Chemistry, University of California, Irvine, 147 Bison Modular, Irvine, CA 92697, USA
e-mail: dmobley@uci.edu

Introduction

Solvation free energy calculations based on classical molecular simulations are of considerable interest to test force fields, help guide pharmaceutical drug discovery, and compute other physical properties of interest. Thus, a large number of tests have focused on computing hydration free energies of both ions [3, 6] and neutral molecules [7]. Most commonly, these calculations are done via a thermodynamic transformation approach. In this so-called alchemical approach, a solute is taken from the state in which it interacts fully with solvent, to a noninteracting state, via a series of nonphysical intermediate states [11]. One key part of this transformation involves modifying the solute electrostatics, either by turning off the solute's electrostatics interactions with its environment, or by turning off its charges entirely, and computing the associated free energy change, commonly known as the charging free energy.

However, we now know that alchemical calculations of solvation of ions are affected in subtle ways by several algorithmic issues which profoundly impact the computed free energies, requiring analytical or semi-analytical corrections [3, 6]. For example, calculations done under periodic boundary conditions where the electrostatic interactions are defined by a periodic lattice sum method like Ewald summation introduce several artifacts due to the limited size of the periodic simulation cell, compared to the bulk experimental systems we model, and require

corrections for missing ionic interactions between the solute and distant solvent. While these issues are now well understood for hydration free energies of ions, and are beginning to be understood for binding free energy calculations of charged ligands [9, 10], we were concerned that hydration free energy calculations of neutral compounds might encounter similar errors. For instance, neutral solutes simulated under periodic boundary conditions could potentially be missing dipole interactions with the other copies of the solute which, though smaller in magnitude, might still be substantial.

Hence, our interest here is determining whether profound system size-dependence effects observed in calculations of ionic hydration free energies [3, 6] also adversely affect hydration free energy calculations for neutral solutes. Thus, our study directly tests this empirically by calculating hydration free energies for representative solutes with high dipole moments, and a simple dipole, at a variety of simulation box sizes ranging from smaller than typical, to extremely large.

Theory

Alchemical free energy calculations work by computing the difference between the desired two end states (here, the solute in water and the solute in gas) along a nonphysical (alchemical) path. A series of intermediate thermodynamic states are introduced to make this practical. The free energy of each of these states is computed using the Multistate Bennett Acceptance Ratio (MBAR) [1, 12]. The computed free energy has an uncharging contribution, corresponding to the free energy of turning off the electrostatic interactions between the solute and the solvent, and a non-polar contribution, where the solute-solvent Lennard-Jones (LJ) interactions are turned off using soft core potentials [2] in the absence of molecular charges. Depending on the details, the electrostatic contribution may instead involve the free energy of turning off the solute's partial charges entirely, in which case an additional set of calculations computes the free energy of turning the solute internal electrostatic interactions back on in the gas phase as here. The total hydration free energy is $\Delta G_{hyd} = \Delta G_{chg,vac} - \Delta G_{chg} - \Delta G_{LJ}$, where ΔG_{chg} denotes the free energy of turning off the electrostatics in water, $\Delta G_{chg,vac}$ denotes the same quantity for vacuum, and ΔG_{LJ} denotes the free energy of turning off the solute-water LJ interactions in water.

Solvation free energy calculations are typically done by introducing the molecule in a nanoscale explicit solvent computational box simulated under periodic boundary conditions. Ideally the solvent should be a bulk system of

macroscopic size and the electrostatic interactions should be treated under non-periodic boundary conditions. Since it is not practical to use a macroscopic simulation box, the conventional norm is to keep the edge length of the box greater than solute size plus twice the LJ cut-off distance. As noted, this results in substantial finite-size effects for alchemical calculations of hydration free energies of charged solutes, but in principle could also result in similar effects for highly polar solutes. Thus, these finite-size effects have seen substantial interest as they pertain to binding and solvation of ionic solutes, and various correction schemes have been proposed [9, 10]. In principle some of these corrections, especially the correction term for missing solute-solvent interactions due to system's periodicity, could apply to solvation of neutral molecules as well, especially for relatively polar solutes [9, 10]. Here we are interested in the overall effect of the solvent box size on the calculated free energy for polar molecules as well as on the uncharged and LJ contributions to the calculated free energies, to test whether these issues have an appreciable effect at typical simulation box sizes [3, 6].

Results and discussion

Here, we essentially find no significant box size-dependence for hydration free energy calculations of neural, polar solutes at typical simulation box sizes. Below, we detail these null results.

We selected two polar molecules, Compound 1 (Fig. 1a) and Compound 2 (Fig. 1b), to investigate the effect of the box size of the system on hydration free energies. Compound 1, which is a derivative of anthraquinone, has a dipole moment of approximately 6 Debye, while Compound 2, nitrobenzene, has a dipole moment of approximately 5 Debye. We focused deliberately on molecules with high dipole moments as we expect that the leading-order term in any box-size dependent correction would scale with the molecular dipole moment. We also investigated a simple dipole, with dipole moment 0.5 Debye, (Fig. 1c) which just has two carbon atoms connected by a single bond.

(a) Compound 1   (b) Compound 2   (c) Simple Dipole
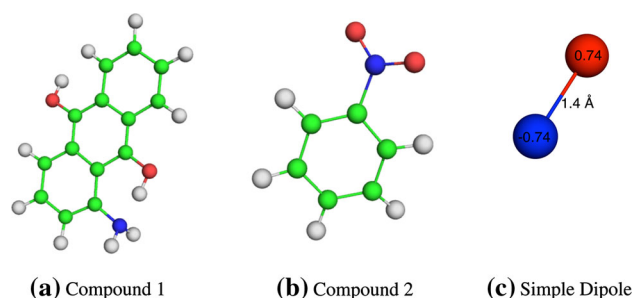
Fig. 1 Structure of compounds under study

The simulations were set up as discussed in the Materials and Methods Section. Briefly, the molecules were solvated in a rhombic dodecahedral water box for a variety of different simulation cell sizes, from very small, to typical, to large. The free energy of solvation in water was calculated using alchemical transformation methods [11]. In this approach, the parameter λ controls the progress of the alchemical transformation—qualitatively, it controls the strength of interactions between the solute and its environment and (depending on setup) the strength of any internal non-bonded interactions. Calculations at each box size involved 20 separate simulations with 20 different λ values . We used two separate λ values, the first one to modify the solute charges and the second one to modify the LJ interactions between the solute and its environment. Free energy differences between the fully charged, full LJ state and the uncharged, noninteracting state are computed using MBAR [1, 12] and summed. We also computed the free energy of modifying the solute internal interactions in the gas phase using the same procedure. The hydration free energy was obtained by calculating the difference in the free energy of the molecule in water relative to the free energy in the gas phase, as in our standard approach [8].

Figure 2 shows the calculated hydration free energy for the three molecules as a function of the box edge length of our periodic rhombic dodecahedral simulation box. For Compounds 1 and 2 the image distance or box edge (d) that determines the box volume was in the range of approximately 2–9 nm. For the simple dipole the box edge was in the range of 1.6–4 nm. A typical box edge for hydration free energy simulations is 3 nm.

We used the bootstrap method [11] to obtain the statistical uncertainty in the derivative of free energy with respect to the box-edge length. We conducted 10000 bootstrap trials, where each one consisted of constructing a new set of results by randomly selecting from the original results, with replacement. For each new set, we computed the slope of a best-fit line, which measures the derivative of the free energy with respect box-edge length. The uncertainty was taken as the standard deviation of the slope over 10000 trials. This data is presented in Table 1. The average value of the derivative for all the lines in Fig. 2 is insignificant. Based on the average value of the derivatives

Table 1 Derivative of the free energy difference with respect to the box-edge length obtained using bootstrap sampling

|	LJ cut-off = 0.65 nm $\Delta G_{hyd}/\Delta d$ (kJ/mol/nm)	LJ cut-off = 1.0 nm $\Delta G_{hyd}/\Delta d$ (kJ/mol/nm)
Compound 1	−0.13 ± 0.06	−0.01 ± 0.07
Compound 2	0.10 ± 0.07	−0.03 ± 0.04
Simple dipole	−0.05 ± 0.07	−0.21 ± 0.11
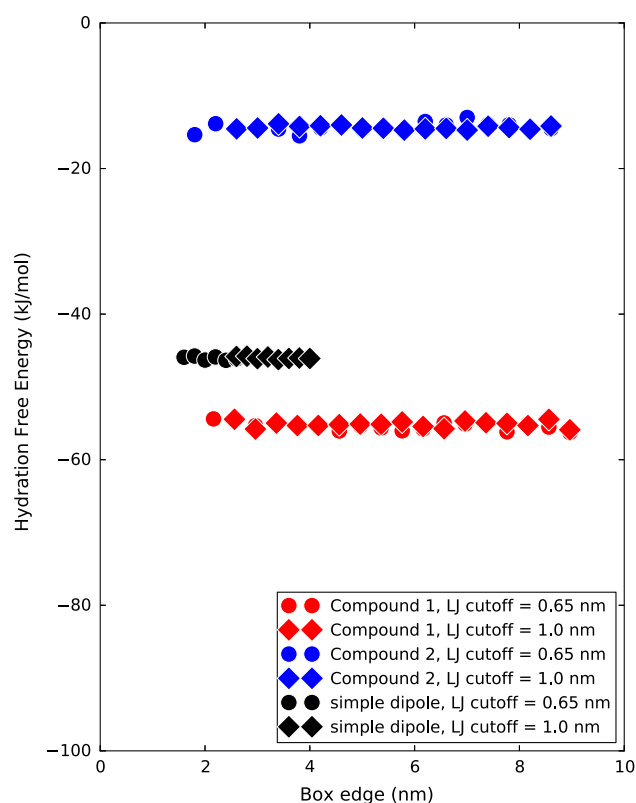
Fig. 2 Calculated hydration free energy for Compounds 1, 2 and the simple dipole as a function of the solvent box size

in Table 1 we can conclude that the free energy simulations are independent of the solvent box size and that if there is any edge-dependence it is less than 0.2 kJ/mol/nm over the box edge lengths considered.

We also conducted free energy calculations using two different LJ cut-off distances and, as shown in Fig. 2, we find that the hydration free energy is independent of this parameter, as it should be for a homogeneous system as here, when a long-range LJ dispersion correction is included [14].

In order to further investigate whether there is a box size dependence for the unchanging and LJ contributions of the free energy for the different alchemical states we closely looked at those energies for different solvent box sizes, and found no significant size-dependence. These results are presented in Figs. 3 and 4. The electrostatic contribution of hydration free energy in Fig. 4 was obtained from the first four alchemical states where the λ_{chg} values changes from 0 (full solute partial charges) to 1.0 (no solute partial charges). The LJ contribution of hydration free energy was obtained from the alchemical states where the λ_{LJ} values changes from 0 (full interaction) to 1 (no interactions) with the electrostatic interactions already turned off. From Figs. 3 and 4 its clearly evident that the charging and LJ contributions of free energy are independent of the size of the simulation box.
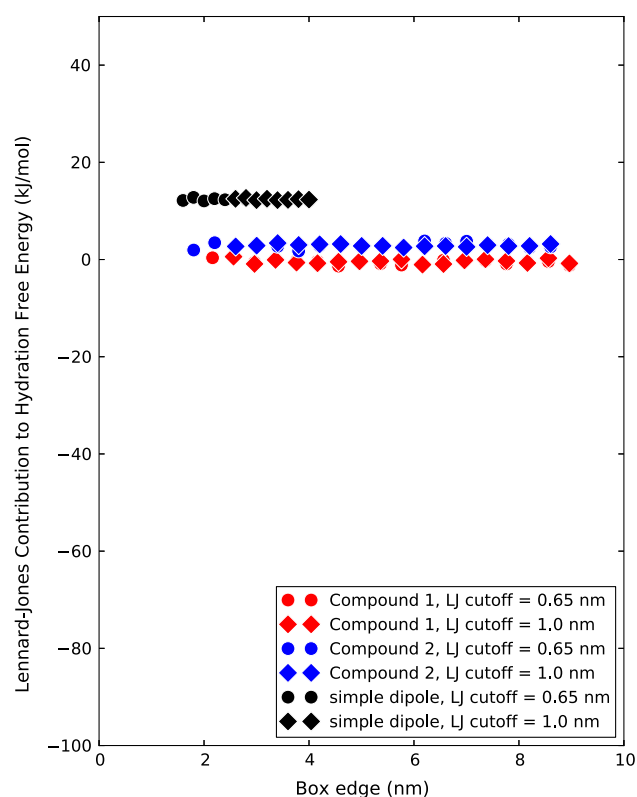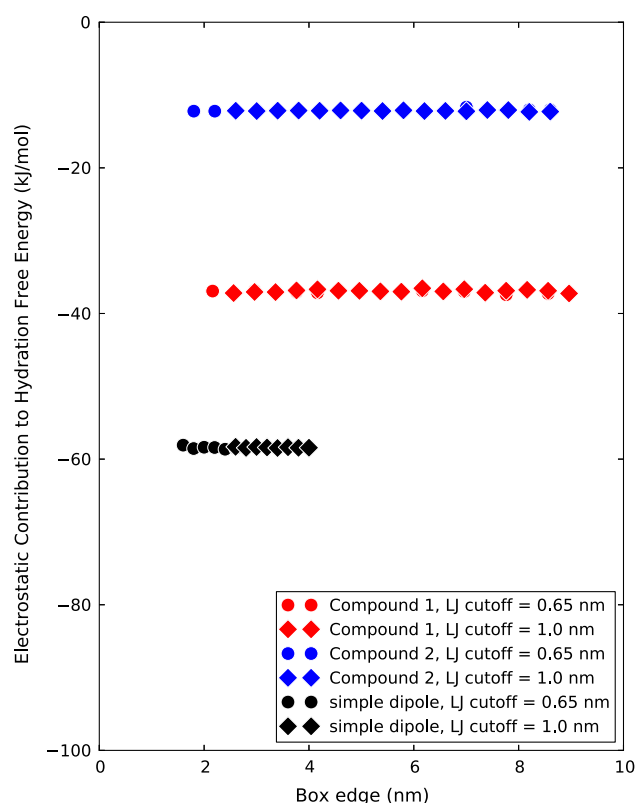
Fig. 3 Lennard-Jones interaction energy for Compounds 1, 2 and the simple dipole as a function of the solvent box size

Fig. 4 Electrostatic interaction energy for Compounds 1, 2 and the simple dipole as a function of the solvent box size

Conclusions

Here, motivated by recent work which found profound finite-size effects in calculations of hydration and binding free energies of ionic solutes or ligands, we looked for similar effects on hydration of neutral solutes. We found that box sizes for typical simulations used for solvation calculations are adequate and do not cause any finite-size effects, at least at the level of statistical precision for our calculations. Thus, our work suggests that standard best practices are adequate for hydration of neutral solutes, even those with substantial dipole moments.

Materials and methods

The GROMACS 4.5.3 software package was used for all simulations. Explicit solvent molecular dynamics simulations were performed with TIP3P [5] water and general AMBER forcefield (GAFF) [16] small molecule parameters as assigned by ANTECHAMBER package [15]. AM1-BCC [4] partial charges were used for the molecules. The molecules were solvated in a rhombic dodecahedral water box starting with a box edge length of 2.2 nm for Compound 1 and 1.8 nm for Compound 2. The box edge

lengths were increased at increments of 0.4 nm up to approximately 9 nm for both compounds. Two different LJ cut-off distances, 0.65 and 1.0 nm, were used in our simulations. The corresponding cut-off distances for the short-range neighbor list was set to 0.7 and 1.2 nm.

The hydration free energy calculations involved several simulations at different alchemical λ values as described elsewhere [13]. In these simulations we used two separate λ values, one to control the modification of solute partial charges and the other to control the modification of LJ interactions. Specifically λ_{chg} was set to [0.0 0.25 0.5 0.75 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0] and λ_{LJ} was set to [0.0 0.0 0.0 0.0 0.0 0.05 0.1 0.2 0.3 0.4 0.5 0.6 0.65 0.7 0.75 0.8 0.85 0.9 0.95 1.0]. At each λ value, the starting structure was minimized using steepest descent for 1500 steps. The minimized structure was run through a constant volume equilibration step consisting of 10 ps followed by a 100 ps of constant pressure equilibration. The production part of the simulation were run at each λ value for 5 ns for Compound 1 and simple dipole and 50 ns for Compound 2, at constant volume. Langevin dynamics was used for temperature control and the reference temperature was 300 K. Particle Mesh Ewald (PME) method was used for long-range electrostatic interactions with a grid space of 0.1 nm. All other protocols were as described previously[8].

Supporting information

The calculated free energies and the box-edge lengths for the three compounds we studied are presented as supplementary tables.

References

1. Bennett CH (1976) Efficient estimation of free energy differences from Monte Carlo data. J Comput Phys 22(2):245–268
2. Beutler TC, Mark AE, van Schaik RC, Gerber PR, van Gunsteren WF (1994) Avoiding singularities and numerical instabilities in free energy calculations based on molecular simulations. Chem Phys Lett 222(6):529–539
3. Hünenberger PH, Reif MM (2011) Single-ion solvation: experimental and theoretical approaches to elusive thermodynamic quantities, vol 3, 1st edn. Royal Society of Chemistry, London
4. Jakalian A, Bush BL, Jack DB, Bayly CI (2000) Fast, efficient generation of high-quality atomic charges. AM1-BCC model: I. Method. J Comput Chem 21(2):132–146
5. Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML (1983) Comparison of simple potential functions for simulating liquid water. J Chem Phys 79(2):926–935
6. Kastenholz MA, Hünenberger PH (2006) Computation of methodology-independent ionic solvation free energies from molecular simulations. II. The hydration free energy of the sodium cation. J Chem Phys 124(22):224501
7. Mobley DL, Dumont É, Chodera JD, Dill KA (2007) Comparison of charge models for fixed-charge force fields: small-molecule hydration free energies in explicit solvent. J Phys Chem B 111(9):2242–2254
8. Mobley DL, Liu S, Cerutti DS, Swope WC, Rice JE (2011) Alchemical prediction of hydration free energies for SAMPL. J Comput Aided Mol Des 26(5):551–562
9. Reif MM, Oostenbrink C (2013) Net charge changes in the calculation of relative ligand-binding free energies via classical atomistic molecular dynamics simulation. J Comput Chem 35(3):227–243
10. Rocklin GJ, Mobley DL, Dill KA, Hünenberger PH (2013) Calculating the binding free energies of charged species based on explicit-solvent simulations employing lattice-sum methods: An accurate correction scheme for electrostatic finite-size effects. J Chem Phys 139(18):184103
11. Shirts M, Mobley D (2013) An introduction to best practices in free energy calculations. In: Monticelli L, Salonen E (eds.) Biomolecular Simulations, Methods in Molecular Biology, vol. 924, pp. 271–311. Humana Press
12. Shirts MR, Chodera JD (2008) Statistically optimal analysis of samples from multiple equilibrium states. J Chem Phys 129(12):124,105
13. Shirts MR, Mobley DL, Chodera JD (2007) Alchemical Free Energy Calculations: Ready for Prime Time?, vol. 3, chap. 4, pp. 41–59. Elsevier
14. Shirts MR, Mobley DL, Chodera JD, Pande VS (2007) Accurate and efficient corrections for missing dispersion interactions in molecular simulations. J Phys Chem B 111(45):13,052–13,063
15. Wang J, Wang W, Kollman PA, Case DA (2006) Automatic atom type and bond type perception in molecular mechanical calculations. J Mol Graph Model 25(2):247–260
16. Wang J, Wolf RM, Caldwell JW, Kollman PA, Case DA (2004) Development and testing of a general AMBER force field. J Comput Chem 25(9):1157–1174