UC Irvine

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

Authors
Parameswaran, Sreeja
Mobley, David L

Peer reviewed

Box Size Effects Are Negligible For Solvation Free Energies of Neutral Solutes

Sreeja Parameswaran and 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 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 Lennard-Jones cutoff distances, 0.65 nm 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 Lennard-Jones 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

David L. Mobley
Department of Pharmaceutical Sciences
and Department of Chemistry
147 Bison Modular
University of California, Irvine
Irvine, CA 92697
Tel.: 949-824-6383
Fax: 949-824-2949
E-mail: dmobley@mobleylab.org

David L. Mobley · Sreeja Parameswaran
Department of Chemistry
University of New Orleans
2000 Lakeshore Drive
New Orleans, LA 70148

1 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 [6, 3] 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 [6, 3]. 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

(a) Compound 1   (b) Compound 2   (c) Simple Dipole

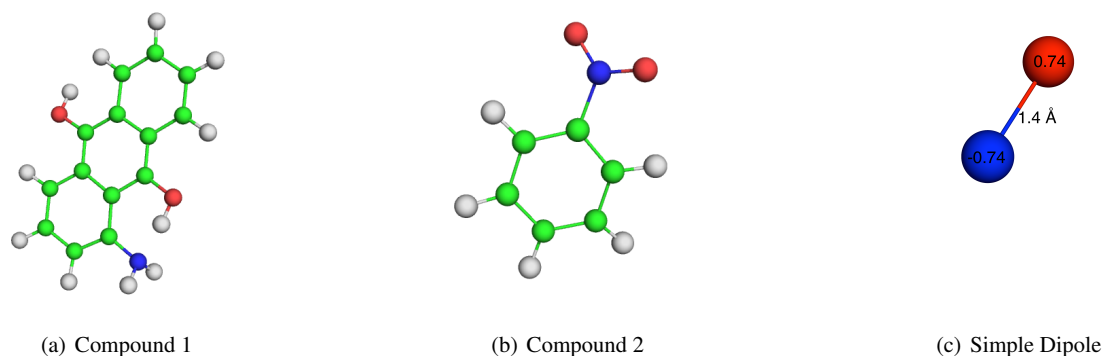**Fig. 1** Structure of compounds under study

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 [10, 9], 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 [6, 3] 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.

2 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 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 Lennard-Jones 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 Lennard-Jones cutoff 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 [10, 9]. 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 [10, 9]. 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 Lennard-Jones contributions to the calculated free energies, to test whether these issues have an appreciable effect at typical simulation box sizes [6, 3].
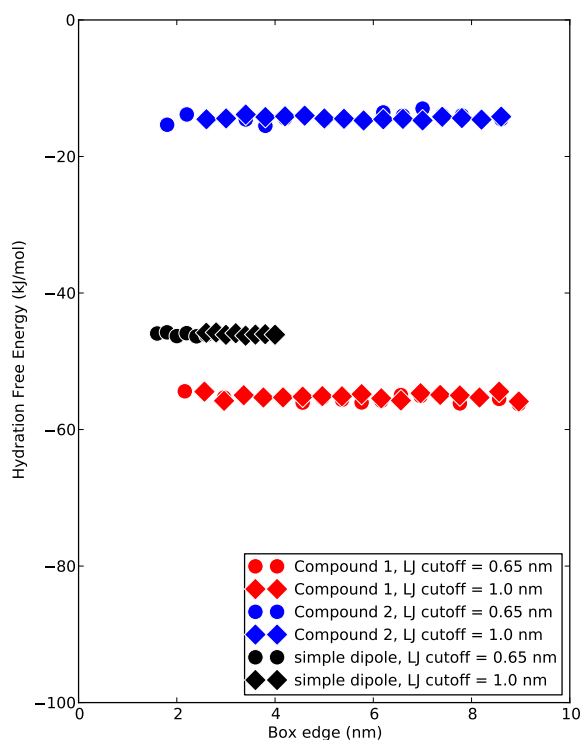
Fig. 2 Calculated hydration free energy for Compound 1, 2 and the simple dipole as a function of the solvent box size.

3 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 (Figure 1A) and Compound 2 (Figure 1B), to investigate the effect of the box size of the solvent 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, (Figure 1C) which just has two carbon atoms connected by a single bond.

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 ap-

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	LJ cut-off = 1.0 nm
	$\Delta G_{hyd}/\Delta d$ (kJ/mol/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

proach, 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 Lennard-Jones (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 nm to 9 nm. For the simple dipole the box edge was in the range of 1.6 nm to 4 nm. A typical box edge for 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 Figure 2 is insignificant. Based on the average value of the derivatives 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 Lennard-Jones cut off distances and, as shown in Figure 2, we find that the hydration free energy is indepen-
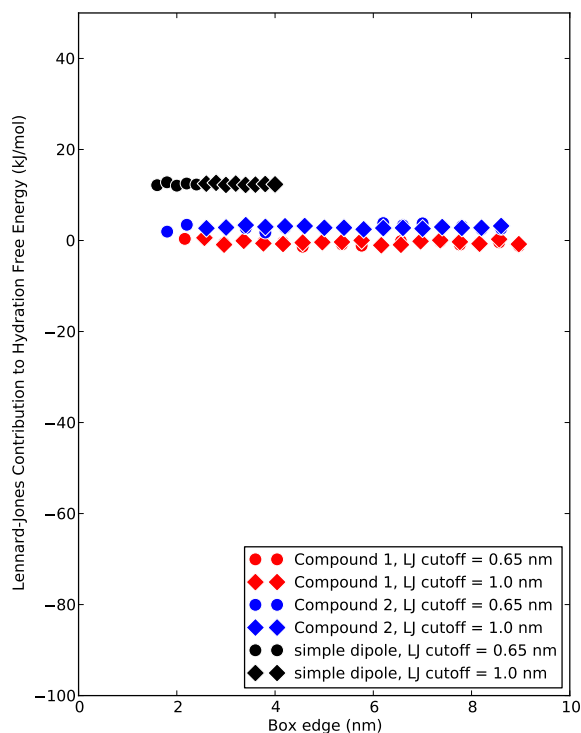
Fig. 3 Lennard-Jones interaction Energy for Compounds 1, 2 and the simple dipole as a function of the solvent box size
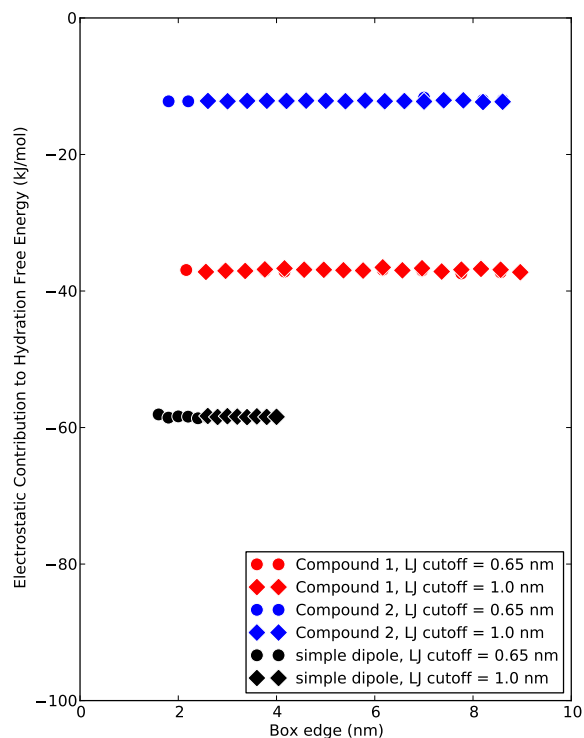

Fig. 4 Electrostatic interaction Energy for Compounds 1, 2 and the simple dipole as a function of the solvent box size

dent of this parameter, as it should be for a homogeneous system as here, when a long-range Lennard-Jones dispersion correction is included [14].

In order to further investigate whether there is a box size dependence for the unchanging and Lennard-Jones 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 Figures 3 and 4. The electrostatic contribution of hydration free energy in Figure 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 Lennard-Jones 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 Figures 3 and 4 its clearly evident that the charging and Lennard-Jones contributions of free energy are independent of the size of the simulation box.

4 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 Lennard-Jones cut-off distances 0.65 nm 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 nm 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 Lennard-Jones 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] 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 100ps 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].

5 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, C.H.: Efficient estimation of free energy differences from monte carlo data. Journal of Computational Physics **22**(2), 245 – 268 (1976)

2. Beutler, T.C., Mark, A.E., van Schaik, R.C., Gerber, P.R., van Gunsteren, W.F.: Avoiding singularities and numerical instabilities in free energy calculations based on molecular simulations. Chemical Physics Letters **222**(6), 529 – 539 (1994)

3. Hünenberger, P.H., Reif, M.M.: Single-ion Solvation: Experimental and Theoretical Approaches to Elusive Thermodynamic Quantities, vol. 3, 1 edn. Royal Society of Chemistry, London (2011)

4. Jakalian, A., Bush, B.L., Jack, D.B., Bayly, C.I.: Fast, efficient generation of high-quality atomic charges. am1-bcc model: I. method. Journal of Computational Chemistry **21**(2), 132–146 (2000)

5. Jorgensen, W.L., Chandrasekhar, J., Madura, J.D., Impey, R.W., Klein, M.L.: Comparison of simple potential functions for simulating liquid water. The Journal of Chemical Physics **79**(2), 926–935 (1983)

6. Kastenholz, M.A., Hünenberger, P.H.: Computation of methodology-independent ionic solvation free energies from molecular simulations. ii. the hydration free energy of the sodium cation. The Journal of Chemical Physics **124**(22), 224501 (2006)

7. Mobley, D.L., Dumont, É., Chodera, J.D., Dill, K.A.: Comparison of charge models for fixed-charge force fields: small-molecule hydration free energies in explicit solvent. The Journal of Physical Chemistry B **111**(9), 2242–2254 (2007). PMID: 17291029

8. Mobley, D.L., Liu, S., Cerutti, D.S., Swope, W.C., Rice, J.E.: Alchemical prediction of hydration free energies for sampl. Journal of computer-aided molecular design (2011). J Comput Aided Mol Des. 2011 Dec 24.

9. Reif, M.M., Oostenbrink, C.: Net charge changes in the calculation of relative ligand-binding free energies via classical atomistic molecular dynamics simulation. Journal of Computational Chemistry (2013)

10. Rocklin, G.J., Mobley, D.L., Dill, K.A., Hünenberger, P.H.: 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. The Journal of Chemical Physics **139**(18), 184103 (2013)

11. Shirts, M., Mobley, D.: An introduction to best practices in free energy calculations. In: L. Monticelli, E. Salonen (eds.) Biomolecular Simulations, *Methods in Molecular Biology*, vol. 924, pp. 271–311. Humana Press (2013)

12. Shirts, M.R., Chodera, J.D.: Statistically optimal analysis of samples from multiple equilibrium states. jcp **129**(12), 124,105 (2008)

13. Shirts, M.R., Mobley, D.L., Chodera, J.D.: Alchemical Free Energy Calculations: Ready for Prime Time?, vol. 3, chap. 4, pp. 41–59. Elsevier (2007)

14. Shirts, M.R., Mobley, D.L., Chodera, J.D., Pande, V.S.: Accurate and efficient corrections for missing dispersion interactions in molecular simulations. The Journal of Physical Chemistry B **111**(45), 13,052–13,063 (2007)

15. Wang, J., Wang, W., Kollman, P.A., Case, D.A.: Automatic atom type and bond type perception in molecular mechanical calculations. Journal of Molecular Graphics and Modelling **25**(2), 247 – 260 (2006)

16. Wang, J., Wolf, R.M., Caldwell, J.W., Kollman, P.A., Case, D.A.: Development and testing of a general amber force field. Journal of Computational Chemistry **25**(9), 1157–1174 (2004)