

Charge-Changing Perturbations and Path Sampling via Classical Molecular Dynamic Simulations of Simple Guest–Host Systems

Christoph Öhlknecht, Jan Walther Perthold, Bettina Lier, and Chris Oostenbrink*

Cite This: *J. Chem. Theory Comput.* 2020, 16, 7721–7734

Read Online

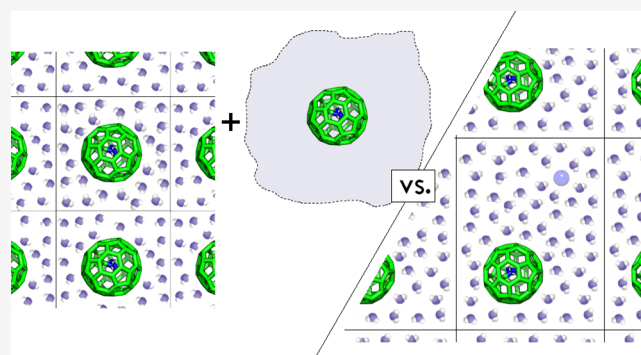
ACCESS |

Metrics & More

Article Recommendations

Supporting Information

ABSTRACT: Currently, two different methods dominate the field of biomolecular free-energy calculations for the prediction of binding affinities. Pathway methods are frequently used for large ligands that bind on the surface of a host, such as protein–protein complexes. Alchemical methods, on the other hand, are preferably applied for small ligands that bind to deeply buried binding sites. The latter methods are also widely known to be heavily artifacted by the representation of electrostatic energies in periodic simulation boxes, in particular, when net-charge changes are involved. Different methods have been described to deal with these artifacts, including postsimulation correction schemes and instantaneous correction schemes (e.g., co-alchemical perturbation of ions). Here, we use very simple test systems to show that instantaneous correction schemes with no change in the system net charge lower the artifacts but do not eliminate them. Furthermore, we show that free energies from pathway methods suffer from the same artifacts.



INTRODUCTION

One of the major challenges in modern computational biochemistry is the accurate and reliable calculation of affinities between binding partners.^{1–8} The most accurate methods can be separated into two different, well-established classes, alchemical modifications and path sampling.⁹ Each approach has its individual strengths and shortcomings. Path-sampling methods rely on sampling physical pathways of the binding and unbinding process between receptors and ligands. This is computationally rather expensive as the entire path should be sampled completely and preferably reversibly. Also, it can be a problem if the ligand is buried within its receptor, requiring large conformational changes upon binding or unbinding, therefore making it even harder to simulate the pathway explicitly.^{10,11} Alchemical methods, on the other hand, can be used to obtain the binding affinity of a single molecule by using the double-decoupling method (DDM) in which the guest molecule is decoupled from its surroundings both when it is free in solution and when it is bound to the host.¹² Free energies for deeply buried ligands are often less biased when calculated with the DDM approach since there is no need to sample the physical binding pathway explicitly. However, particularly for guest molecules bearing a net charge, the free energies for decoupling are typically rather large. To compute the difference between two large free energies (the guest bound to the host and free in solution), the individual terms need to be obtained with sufficient precision, possibly requiring a large number of intermediate states and long simulations. Because of these different properties, a method has been described recently to combine

the advantages of alchemical and pathway methods.¹³ However, a remaining very prominent shortcoming of alchemical methods is that free energies are artifacted in the case of charged ligands. This is connected to the way electrostatic interactions are simplified in molecular simulations. The details have extensively been discussed previously.^{14–20}

In short, due to finite-size effects, the calculated free energies become dependent on system parameters such as the cutoff radius (cutoff schemes) or the box shape and size (lattice-summation methods). These system-dependent artifacts translate directly into the calculated charging free energies. They have been studied quite extensively, and multiple strategies have been proposed to correct for them *a posteriori* or to avoid them altogether. Relevant methods can be split into two groups, post-simulation correction schemes and instantaneous methods.

Post-simulation charge-correction terms have been shown to work for lattice summation (LS) methods like Ewald summation or particle-particle-particle-mesh (P3M)^{21–24} and cutoff schemes with reaction field (RF) electrostatics.^{25–27} The corresponding terms can be evaluated from a combination of numerical and analytical models.^{14–19} These approaches require

Received: July 10, 2020

Published: November 2, 2020



analyses of the MD-generated trajectories to estimate the size of the artifacts. Through application of correction terms, independence of system-related parameters can be achieved. It has been shown before^{17,19,20} that reliable results can be obtained, approximating macroscopic, fully coulombic systems.

Instantaneous correction schemes are built on the fact that the dominant errors are dependent on the net-charge change of the system.^{18,28} Here, the alchemical perturbation of the charged moiety is simultaneously performed with a counter-alchemical charge perturbation of a remote molecule. A very prominent method is the so-called co-alchemical ion approach where the remote molecule is represented by a counter-ion.^{29–34} A similar method couples both legs of a thermodynamic cycle in the same simulation box.^{18,35,36} Here, the charges of the bound ligand are diminished while they are simultaneously installed in the unbound ligand at a distant position of the bound complex. Both methods assure that the net charge of the entire system does not change during the alchemical change. As a consequence, dominant errors arising from finite-size effects might become negligible. Also, a hybrid method between post-simulation corrections and instantaneous correction schemes was described.³⁷ Here, corrections were imposed on the forces on-the-fly during the MD simulation.

In contrast, electrostatic artifacts have not been widely discussed for path-sampling methods. In fact, these methods are used next to instantaneous correction schemes to avoid artifacts due to charge changes. Although pathway methods seem to be slightly better suited for charged ligands,³⁸ it has been pointed out in the past that these approaches are also affected by finite-size effects.¹⁷ The motivation of the underlying work is twofold. First, the effectiveness of the co-alchemical ion approach was compared to an alchemical approach in combination with a post-simulation correction scheme.²⁰ Free energies were calculated using both LS and RF methods. Both were calculated for complexes in pure water and in a saline solvent with a high ionic strength. Second, free energies and associated artifacts calculated from the alchemical double-decoupling method¹² were compared with those calculated from path sampling, using RF electrostatics. Note that, within this study, the effect of different methods and protocols on the finite-size effects was studied but not the effect of the size-dependence on the artifacts itself. It has been shown before using (i) the buckyball systems that are also used in the current work¹⁹ and (ii) real protein systems¹⁷ that the applied post-simulation correction scheme does not only bring different approximate electrostatics schemes in agreement but also eliminates the size dependence of the artifacts.

To avoid problems arising from insufficient sampling, relatively simple systems were used to study the size of the artifacts for the different methods. We used two simple, oligoatomic ligands (guests) with opposite charges. We studied their charging or binding free energies with respect to C60 fullerenes (buckyballs, hosts) derivatized with different chemical groups to mimic a wide variety of chemical properties.^{19,39}

METHODS

MD-Simulations. To study free energies of a guest binding to a host, relatively simple systems were used.³⁹ Two oppositely charged guests were considered, acetate (ACE) and methylammonium (MAM). Host molecules with low flexibility and high symmetry were used to focus on electrostatic artifacts rather than insufficient sampling: a C60 fullerene (buckyball) and three derivatives thereof with different physical properties;

an un-derivatized buckyball with an apolar cavity (CAPO); a buckyball containing a covalently bound amide group, representing a neutral polar cavity with hydrogen-bonding capability (CHB); a buckyball containing a covalently bound methylammonium group, representing a positively charged cavity (CPOS); a buckyball containing a covalently bound carboxylate group, representing a negatively charged cavity (CNEG). In comparison to a realistic model of a buckyball, all C–C bonds were artificially kept at a 0.2 nm length. All simulations were performed with the GROMOS11 program.⁴⁰

For all simulations, cubic boxes under periodic boundary conditions were used. For the description of all molecular interactions, a modified version of the GROMOS 53A6 force field was used as described before.¹⁹ Water was treated explicitly and implemented by means of the three-site simple point charge (SPC) model.⁴¹ The equations of motion were integrated using the leap-frog scheme, using a timestep of 2 fs.⁴² Bond vibrations were constrained using the SHAKE algorithm⁴³ with a relative geometric tolerance of 10^{-4} . The center of mass translation of the computational box was removed every 2 ps. The temperature was maintained at 300 K by weak coupling using a coupling time of $\tau_T = 0.1$ ps. The edges of the computational box were kept constant (see the length in the individual paragraphs). Artifacts in the free energies between alchemical transformations and co-alchemical transformations were compared. This was done by simulating the alchemical charging process with and without co-perturbation of an ion that bore the same charge as the guest. To be able to compare artifacts in the free energies between alchemical transformations and path-sampling methods, we opened all buckyballs to allow for efficient sampling of the binding path. The methods are explained in the following paragraphs. Note that, in all simulations regardless of the electrostatics, the self-energy term^{44–46} of the ligand was included in the electrostatic potential. This term accounts for electrostatic interactions between excluded atoms (via the reaction field or the periodic copies) and between periodic copies of the atoms themselves in lattice sum methods. Inclusion of this term in the electrostatic potential is in contrast to previous studies^{14,15,19} but resembles real case studies better.

(Co-)alchemical Transformations in the Closed Hosts. All simulations in the closed buckyballs (2 guests, 4 hosts) were executed using a cutoff scheme with reaction field (RF) or lattice summation (LS) electrostatics using the particle–particle–particle mesh (P³M) method for the electrostatic interactions. In simulations with the RF scheme, the relative permittivity of the reaction-field was set to a value $\epsilon_{RF} = 66.6$, as appropriate for the SPC water model.⁴⁷ In these simulations, all nonbonded interactions were truncated at a charge-group cutoff distance of 1.4 nm, calculated at every timestep based on a pairlist that was updated every timestep. The P³M algorithm was applied using tinfoil boundary conditions,²¹ a spherical hat charge-shaping function of width 1.0 nm, a triangular-shaped cloud assignment function, a finite-difference (FD) scheme of order two, and a grid spacing of 0.10 and 0.12 nm for the unbound and bound systems, respectively. In simulations with LS electrostatics, Lennard–Jones interactions were truncated at a group-based cutoff distance of 1.0 nm. Real-space electrostatics and Lennard–Jones interactions were calculated at every timestep based on a pairlist that was updated every timestep. For all simulations, cubic boxes with an edge length of 4.5 nm (bound state) or 3.25 nm (unbound state) were used. All simulations were executed in two different solvents, pure water and 0.5 M sodium chloride. In the case of the

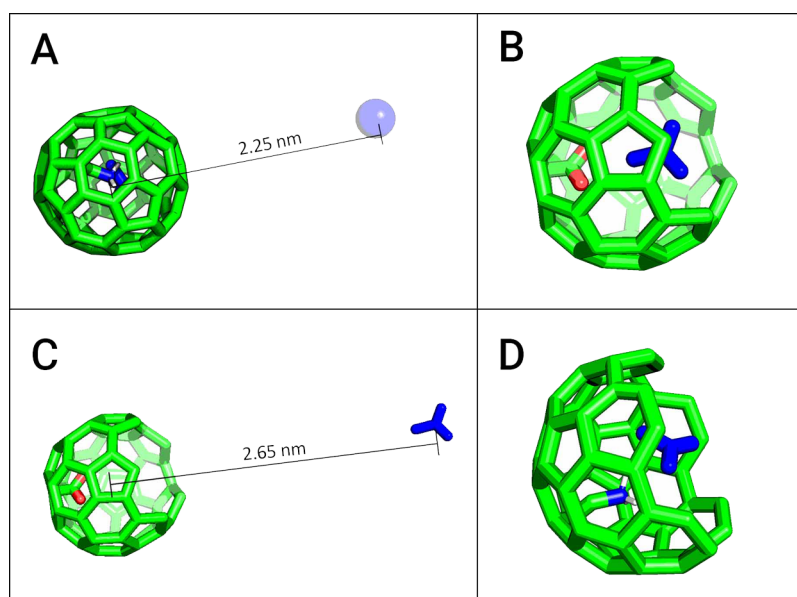


Figure 1. Representation of the different systems used in this study. The hosts are represented in green; the guests in blue colour. (A) Closed buckyball, here shown with a distant Na ion; (B) open buckyball with the ligand in the bound state; (C) open buckyball with the ligand in the unbound state; (D) more opened buckyball as used for the systems with highly attractive forces between the guest and host molecules (ACE in CPOS, MAM in CNEG).

coalchemical transformations, an additional sodium ion was inserted bearing the same charge as the ligand (see Figure 1). Positively and negatively charged sodium ions were used to keep the Lennard–Jones interactions of the ions the same between all systems. The charge of this ion was co-perturbed with the ligand along λ to keep the net charge of the system constant (see Table 1). To avoid any direct interactions with

Table 1. Coalchemical Strategy for the Simulations in the Closed Buckyballs^a

cavity	lig. pert.	charge change	ion pert.	system charge
CAPO	DUM → MAM	0 → 1	NA ⁺ → NA ⁰	1 → 1
CAPO	DUM → ACE	0 → -1	NA ⁻ → NA ⁰	-1 → -1
CHB	DUM → MAM	0 → 1	NA ⁺ → NA ⁰	1 → 1
CHB	DUM → ACE	0 → -1	NA ⁻ → NA ⁰	-1 → -1
CPOS	DUM → MAM	0 → 1	NA ⁺ → NA ⁰	2 → 2
CPOS	DUM → ACE	0 → -1	NA ⁻ → NA ⁰	0 → 0
CNEG	DUM → MAM	0 → 1	NA ⁺ → NA ⁰	0 → 0
CNEG	DUM → ACE	0 → -1	NA ⁻ → NA ⁰	-2 → -2

^aTo keep the net charge of the system constant, positively or negatively charged sodium ions were uncharged while perturbing a dummy molecule (DUM) into the fully interacting ligand (acetate, ACEm or methylammonium, MAM). Four hosts were considered: a buckyball with an apolar cavity (CAPO) and buckyballs with hydrogen-bonding capabilities (CHB), with a positively charged cavity (CPOS), and a negatively charged cavity (CNEG).

the guest–host complex, this coalchemical ion was restrained at 2.25 nm from the center of the buckyball using a harmonic distance restraint with a force constant of 1000 kJ mol⁻¹ nm⁻² in x , y , and z directions. In case of the unbound guests, due to smaller box sizes, the distance restraints of the coalchemical ions were set to 1.625 nm from the center of ACE and to 1.605 nm from the center of MAM using force constants of 10,000 and 11,000 kJ mol⁻¹ nm⁻², respectively. The charge perturbations of the ligands or the charge perturbation of the ligands together

with the coupled ions were performed along 11 equidistant λ -points between 0 and 1 using 500,000 simulation steps each. No soft-core interaction function was applied to the perturbed electrostatic interactions. Free energies of charging were calculated from the sampled trajectories using the Bennett acceptance ratio method (BAR).⁴⁸ Standard deviations were calculated from 50 bootstraps of data that was sampled every n steps from the trajectories where n was estimated considering the statistical inefficiency from autocorrelation analysis.⁴⁹

Alchemical Transformations and Path Sampling in the Opened Hosts. To be able to compare free energies from path-sampling methods and alchemical methods, we opened buckyballs to allow for sampling of the physical binding paths. All buckyballs were opened by removing a ring of five carbon atoms. To allow for more efficient sampling of the pathways in the case of highly attractive forces, i.e., to facilitate the entrance of water molecules into the buckyball cavity, we removed five additional carbon atoms for the two systems with opposite charges (ACE/CPOS and MAM/CNEG). Exemplary illustrations of these opened bucky balls are given in Figure 1. For the calculation of the electrostatic interactions, a cutoff scheme with RF electrostatics was used. All simulations were performed in cubic boxes with an edge length of 5.3 nm (path sampling), 4.5 nm (alchemical transformations in the bound state), or 3.25 nm (alchemical transformations in the unbound state). The relative permittivity of the reaction field was set to a value $\epsilon_{\text{RF}} = 66.6$, as appropriate for the SPC water model.⁴⁷ Nonbonded interactions were truncated at a charge-group cutoff distance of 1.4 nm, calculated at every timestep based on a pairlist that was updated every timestep. The alchemical perturbations of Lennard–Jones and Coulomb parameters of the ligands were performed along 11 equidistant λ -points between 0 and 1 using a soft-core interaction function⁵⁰ with parameters $\alpha^{\text{LJ}} = 1.51$ and $\alpha^{\text{CB}} = 0.50$ nm². Free energies were calculated from the sampled trajectories using BAR.⁴⁸ Standard deviations were calculated from 50 bootstraps of the data, considering the statistical inefficiency from autocorrelation

analysis.⁴⁹ In the case of the bound state, a distance restraint (with a force constant of 1500 kJ mol⁻¹ nm⁻²) with respect to the center of the buckyball (with a reference distance of 0 nm) was used to prevent the ligand from sampling the whole simulation box. The free-energy calculations using path-sampling simulations were performed using 51 equidistant umbrella windows with 500,000 steps (ACE/CAPO, ACE/CHB, MAM/CHB, MAM/CPOS), 750,000 steps (ACE/CNEG, MAM/CAPO), 1,500,000 steps (MAM/CNEG), or 2,500,000 (ACE/CPOS) each. The distances between the host and guest molecules were restrained using a force constant of 1500 kJ/mol⁻¹ nm⁻². The distances ranged from 0 to 2.65 nm (λ -points between 0 and 1), and free energies were calculated from the sampled trajectories using also the BAR method.⁴⁸ The restrained raw free energies from both alchemical perturbations, and path sampling need to be corrected for the standard state volume V^\ominus to render them comparable to each other. The reason is that the restrained guests and hosts are not present at a standard-state concentration of 1 mol/L. These correction terms account for the removal of the distance restraints and bringing the binding partners from their available volume in the restrained bound and decoupled or unbound state, respectively, to the standard state volume. The correction for the alchemical binding free energies in the bound state reads⁵¹

$$\Delta G_{\text{raw}}^{\ominus, \text{alch}} = \Delta G_{\text{raw}}^{\text{res, alch}} + RT \ln \frac{V^\ominus}{(2\pi RT/K)^{3/2}} \quad (1)$$

where $\Delta G_{\text{raw}}^{\text{res, alch}}$ is the restrained binding free energy from instalment of a restrained guest in the host, $\Delta G_{\text{raw}}^{\ominus, \text{alch}}$ is the corrected standard-state free energy, $V^\ominus = 1.661 \text{ nm}^3$, and K is the force constant of the harmonic distance restraint. Mind that the denotation “raw” here means that these free energies still need to be corrected for electrostatic artifacts (as described in the following paragraphs).

The correction for the free energies of binding from path sampling reads

$$\Delta G_{\text{raw}}^{\ominus, \text{path}} = \Delta G_{\text{raw}}^{\text{res, path}} + RT \ln \frac{V^\ominus}{V^{\text{unb}}} \quad (2)$$

where $\Delta G_{\text{raw}}^{\ominus, \text{path}}$ and $\Delta G_{\text{raw}}^{\text{res, path}}$ are the standard state and restrained free energies of binding from path sampling. $V^{\text{unb}} = \frac{4}{3}\pi[(d_u^{\text{max}})^3 - (d_u^{\text{min}})^3]$ is the volume of the sampled sphere in the unbound state. d_u^{max} and d_u^{min} are the maximum and minimum distances, respectively, between guest and host molecules with respect to their center of masses at $\lambda = 1$. This gives the volume of the hollow sphere in the end state.⁵²

Free-Energy Correction Terms. The artifacts in the raw free energies ΔG_{raw} from all methods were analyzed by using a post-simulation correction scheme.^{19,20} In case of artifacts being present, the total correction estimate ΔG_{cor} can be used to yield methodology-independent values ΔG as

$$\Delta G = \Delta G_{\text{raw}} + \Delta G_{\text{cor}} \quad (3)$$

ΔG_{cor} is a combination of multiple free-energy corrections

$$\Delta G_{\text{cor}} = \Delta G_{\text{pol}} + \Delta G_{\text{dir}} + \Delta G_{\text{dsm}} \quad (4)$$

A comprehensive explanation of these correction terms and a guideline for their application was given in our previous work.²⁰ In the current manuscript, we focus on the corrections for simulations performed with a group-based cutoff scheme with reaction field contributions to account for long-range

interactions or with a lattice-sum method. Previously, similar correction schemes have been proposed for straight cutoffs or atom-based cutoff schemes with reaction-field contributions.¹⁵ In the next paragraphs, only a general overview over the used methodologies is given. We provide example files for the simulations and the calculations of the corrections in a GitHub repository.⁵³

Solvent Polarization. Dependent on the effective electrostatic interaction function and associated parameters used, simulations are artifacted by a spurious solvent polarization around the moiety with perturbed atom charges. When cutoff truncation is applied, the solvent outside the cutoff radius is non-interacting with the charged atoms of the solute. When an LS scheme is applied, the solvent molecules are not only polarized by the solute in the reference box but also by the solute in the periodic copies, leading to an effect of underhydration. Also, the usage of a solvent model with an inaccurate dielectric permittivity leads to spurious polarization in close proximity of the perturbed atom sites. The free-energy correction for the wrong solvent polarization ΔG_{pol} can be deduced from a set of continuum-electrostatics calculations and reads

$$\Delta G_{\text{pol}} = \Delta G_{\text{chg, P(S)}^{\text{macro}}} - \Delta G_{\text{chg, P(S)}^{\text{sim}}} \quad (5)$$

where P denotes the charge-perturbed atoms, and S is the solvent. $\Delta G_{\text{chg, P(S)}^{\text{macro}}}$ and $\Delta G_{\text{chg, P(S)}^{\text{sim}}}$ are free energies derived under fully coulombic and macroscopic conditions or for a periodic system with effective electrostatic interactions (RF or LS) to mimic simulation conditions, respectively. The individual electrostatic free energies can be calculated from the environment-generated electrostatic potentials ϕ derived from the continuum-electrostatics calculations under individual charge states $\lambda \in [0,1]$:

$$\Delta G_{\text{chg, P(S)}^{\text{env}}} = \sum_{i=1}^{N_p} \Delta q_i \int_0^1 [\phi(r_i; \lambda; \text{env}; \epsilon_{\text{sol}}) - \phi(r_i; \lambda; \text{env}; \epsilon_0)] d\lambda \quad (6)$$

where $\text{env} \in (\text{macro}, \text{sim})$, $i = 1 \dots N_p$ are the charge-perturbed atoms, $\Delta q_i = q_i^{\lambda=1} - q_i^{\lambda=0}$ is the total partial charge change of atom i , λ is a coupling parameter, ϵ_{sol} is the solvent dielectric permittivity, and $\epsilon_0 = 1$ is the vacuum dielectric permittivity. However, as the electrostatic potential shows a linear dependence with the charge state along λ , eq 6 can be simplified using the trapezoidal rule to

$$\begin{aligned} \Delta G_{\text{chg, P(S)}^{\text{env}}} &= \sum_{i=1}^{N_p} (1/2) \Delta q_i [\phi(r_i; \lambda = 0; \text{env}; \epsilon_{\text{sol}}) \\ &+ \phi(r_i; \lambda = 1; \text{env}; \epsilon_{\text{sol}})] \\ &- (1/2) \Delta q_i [\phi(r_i; \lambda = 0; \text{env}; \epsilon_0) \\ &+ \phi(r_i; \lambda = 1; \text{env}; \epsilon_0)] \end{aligned} \quad (7)$$

For the calculation of ΔG_{pol} , the program *dGslv_pbsolv*, which is included in the GROMOS++ simulation package⁴⁰ was used. This program employs a finite difference (FD) Poisson equation solver^{54–56} capable of handling periodic boundary conditions in combination with a fast Fourier transform (FFT) Poisson equation solver capable of handling RF schemes.^{19,57,58}

The dielectric boundary was based on atomic radii that were derived from the minimum of the Lennard–Jones interaction energy between the solute-atoms and SPC water molecule. The Lennard–Jones interactions were taken from a GROMOS

53A6 force field.⁵⁹ Hydrogen atoms were assigned a radius of 0.05 nm. The solvent dielectric permittivity ϵ_{sol} was set to 66.6 for the calculation of $\Delta G_{\text{chg}, \text{P}(\text{S})}^{\text{sim}}$ to account for the dielectric permittivity of SPC water⁴⁷ and to 78.4 for the calculation of $\Delta G_{\text{chg}, \text{P}(\text{S})}^{\text{macro}}$ to account for the dielectric permittivity of water. The calculations of $\Delta G_{\text{chg}, \text{P}(\text{S})}^{\text{macro}}$ and $\Delta G_{\text{chg}, \text{P}(\text{S})}^{\text{sim}}$ were performed on discrete, equidistant configurations that were extracted from all λ -generated trajectories between $\lambda = 0$ and $\lambda = 1$ every 100,000 steps (using eq 6). Note that these results could have been approximated using eq 7 (see Table S4 in the Supporting Information). For the trajectories generated from path sampling, only snapshots generated every 500,000 steps at $\lambda = 0$ and $\lambda = 1$ were considered (using eq 7).

In these continuum-electrostatic calculations, both the ligand and host were treated as polyatomic charges within solvent-excluded space. This is important to calculate the potential that a polyatomic charge distribution exerts on the ligand under nonperiodic versus periodic boundary conditions. The solvent, however, is represented as a continuum. Hence, the offset in the potential arising from the discrete solvent representation under periodic boundary conditions compared to the nonperiodic case cannot be captured by these calculations. It is thus necessary to add an additional correction term for the discrete solvent molecules. This will be discussed in the next paragraph.

Potential from Discrete Solvent Molecules. The solvent-generated potential at the sites of the perturbed atoms deviates from the “real” potential in a macroscopic environment.

There are two possible conventions to calculate the absolute value of this potential under periodic boundary conditions. Either the zero of the electrostatic potential is calculated (i) as an average of the potential over the exterior and the interior of the discrete solvent molecules (P-convention) or (ii) as an average of the potential only over the exterior of the solvent molecules (M-convention). The electrostatic potentials generated at the sites of the perturbed atoms are noted ϕ_{P} and ϕ_{M} . These are different in size, and the difference $\phi_{\text{P}} - \phi_{\text{M}}$ is referred to as the exclusion potential of the solvent model. After elimination of finite-size and/or cutoff effects, the solvent-generated potential in an MD-generated trajectory will resemble the potential calculated from the P-convention, if an LS or RF scheme is applied.⁶⁰ It was shown before^{60–63} that to approximate true coulombic interactions with the zero at infinity, the M-convention should be used.

The corresponding free-energy correction is proportional to the molecular density of water inside the box (LS) or within the cutoff radius (RF) and reads

$$\Delta G_{\text{dsm}}(\text{LS}) = -N_{\text{A}}(6\epsilon_0)^{-1}\gamma_{\text{s}}\Delta QN_{\text{S}}V_{\text{B}}^{-1} \quad (8)$$

for the LS scheme and

$$\Delta G_{\text{dsm}}(\text{RF}) = -N_{\text{A}}(6\epsilon_0)^{-1}\frac{2(\epsilon_{\text{RF}} - 1)}{2\epsilon_{\text{RF}} + 1}\gamma_{\text{s}} \times \sum_{i=1}^n \Delta q_i \langle N_{\text{S}}(R_{\text{C},i}) \rangle V_{\text{C}}^{-1} \quad (9)$$

for the RF scheme where N_{A} is the Avogadro constant, ϵ_0 is the vacuum dielectric permittivity, ϵ_{RF} is the reaction field dielectric permittivity, γ_{s} is the quadrupole-moment trace of the water model, ΔQ is the net-charge change in the system, Δq_i is the net-charge change of the perturbed atom i , N_{S} is the number of solvent molecules in the box, $\langle N_{\text{S}}(R_{\text{C},i}) \rangle$ is the

average number of solvent molecules in the cutoff sphere of the perturbed atom i , V_{B} is the volume of the computational box, and V_{C} is the volume of the cutoff sphere. For the calculation of $\langle N_{\text{S}}(R_{\text{C},i}) \rangle$, a radial distribution function between the center of mass of the perturbed atoms and the van der Waals interaction sites of the solvent molecules was calculated over the full trajectories at $\lambda = 1$ (alchemical transformations) or at $\lambda = 0$ and $\lambda = 1$ (path transformations).

Note that the quadrupole estimate discussed above gives only a correct estimate for a solvent in the so-called orientational-disorder limit (ODL), which is an idealized situation based on the absence of intermolecular electrostatic interactions. In a simulation where the solvent is polarized and certainly not in an ODL situation, polarization-dependent components also contribute. These are taken care of in the continuum-electrostatics calculations (ΔG_{pol}).

Direct Nonsolvent Interactions. In computer simulations, the calculated electrostatic interactions between the individual charge groups are artifacted due to the applied effective electrostatic interactions scheme. When cutoff truncation is applied in the simulation, interactions beyond a particular cutoff are neglected. If an RF contribution is added, only the effect of a homogeneous medium outside the cutoff distance on the interactions within the cutoff is included. In simulations with the LS scheme, the perturbed atoms are interacting with the periodic copies of the host also. Analogous to the correction scheme for the solvent polarization, the correction term for the artifacted direct interactions ΔG_{dir} can be deduced from the two electrostatic free energies

$$\Delta G_{\text{dir}} = \Delta G_{\text{chg}, \text{P}(\text{R})}^{\text{macro}} - \Delta G_{\text{chg}, \text{P}(\text{R})}^{\text{sim}} \quad (10)$$

where $\Delta G_{\text{chg}, \text{P}(\text{R})}^{\text{macro}}$ is the contribution to the free energy of charging of the perturbed atoms due to the remaining (unperturbed) atoms, and $\Delta G_{\text{chg}, \text{P}(\text{R})}^{\text{sim}}$ is the corresponding free-energy contribution that also includes self-term contributions, an RF contribution for the polarization between excluded atoms or interactions with periodic copies in LS simulations. The first is calculated under full coulombic electrostatic interactions in a macroscopic environment, the latter with effective electrostatic interactions in a periodic environment.

The charging free energies of the perturbed atoms due to the remaining atoms (P(R)) and the self-term can be calculated via reanalysis of the simulated trajectories

$$\Delta G_{\text{chg}, \text{P}(\text{R})}^{\text{env}} = \int_0^1 [\langle E_{\text{elec}}^{\text{PR}} \rangle_{\lambda}^{\text{env}} - \langle E_{\text{elec}}^{\text{R}} \rangle_{\lambda}^{\text{env}}] d\lambda \quad (11)$$

where $\langle E_{\text{elec}}^{\text{PR}} \rangle_{\lambda}^{\text{env}}$ and $\langle E_{\text{elec}}^{\text{R}} \rangle_{\lambda}^{\text{env}}$ are electrostatic energies sampled in the trajectories defined by λ . Note that, for the calculation of $\langle E_{\text{elec}}^{\text{PR}} \rangle_{\lambda}^{\text{env}}$, the solute–solute interactions from trajectories generated at $\lambda = 0$ and $\lambda = 1$ were reanalyzed with full charges ($\lambda = 1$). For the calculation of $\langle E_{\text{elec}}^{\text{R}} \rangle_{\lambda}^{\text{env}}$, the charges of the perturbed atoms were zeroed to extract only the energies between the remaining atoms.

Prior to analysis, all explicit solvent molecules were deleted from the sampled configurations. To extract the energies, the configurations sampled at individual λ -points were reanalyzed with full charges ($\lambda = 1$) for the perturbed and the remaining atoms (PR). Then, the same configurations were reanalyzed with full charges for the remaining atoms but eliminated charges for the perturbed atoms (R). The same configurations were used as in the calculations of ΔG_{pol} . Error estimates on the sum of ΔG_{pol} and ΔG_{dir} are reported as the standard error of the mean

Table 2. Raw Alchemical Free Energies ($\Delta G_{\text{raw}}^{\text{alch}}$), Correction Terms ($\Delta G_{\text{pol}} + \Delta G_{\text{dir}}$, ΔG_{dsm}), and Corrected Alchemical Free Energies (ΔG^{alch}) for both Guests Acetate (ACE) and Methylammonium (MAM) in all Four Closed Hosts with an Apolar Cavity (CAPO), with Hydrogen-Bonding Capability (CHB), with a Positively-Charged Cavity (CPOS), with a Negatively-Charged Cavity (CNEG) and as Free in Solution^a

LIG	HOST	scheme	$\Delta G_{\text{raw}}^{\text{alch}}$	$\Delta G_{\text{pol}}^{\text{alch}} + \Delta G_{\text{dir}}^{\text{alch}}$	$\Delta G_{\text{dsm}}^{\text{alch}}$	ΔG^{alch}
4	CAPO	water, RF	-137.6 ± 0.2	0.6 ± 0.2	66.0	-71.0 ± 0.3
		water, LS	-151.5 ± 0.4	4.3 ± 0.2	76.3	-70.9 ± 0.4
		saline, RF	-138.7 ± 0.3	1.9 ± 0.4	64.3	-72.6 ± 0.5
		saline, LS	-152.7 ± 0.3	6.5 ± 0.3	74.8	-71.4 ± 0.4
	CHB	water, RF	-181.2 ± 0.1	1.5 ± 0.2	64.3	-115.4 ± 0.2
		water, LS	-195.6 ± 0.2	4.3 ± 0.3	76.4	-114.9 ± 0.4
		saline, RF	-182.7 ± 0.4	2.1 ± 0.4	64.6	-116.0 ± 0.6
		saline, LS	-195.2 ± 0.4	6.6 ± 0.4	74.9	-113.7 ± 0.6
	CPOS	water, RF	-358.6 ± 0.3	-5.0 ± 0.3	66.3	-297.2 ± 0.4
		water, LS	-373.6 ± 0.2	0.9 ± 0.1	76.3	-296.4 ± 0.2
		saline, RF	-361.0 ± 0.2	0.6 ± 0.5	65.1	-295.3 ± 0.5
		saline, LS	-374.3 ± 0.5	3.0 ± 0.4	74.8	-296.5 ± 0.6
	CNEG	water, RF	-18.9 ± 0.1	6.8 ± 0.1	66.1	54.0 ± 0.1
		water, LS	-32.7 ± 0.3	10.9 ± 0.1	76.3	52.8 ± 0.3
		saline, RF	-20.4 ± 0.7	5.8 ± 0.3	65.5	50.9 ± 0.8
		saline, LS	-32.7 ± 0.4	10.9 ± 0.3	74.9	53.1 ± 0.5
	free	water, RF	-374.8 ± 0.3	-3.4 ± 0.1	79.0	-299.2 ± 0.3
		water, LS	-383.1 ± 0.4	1.4 ± 0.1	81.0	-300.6 ± 0.4
		saline, RF	-365.7 ± 0.3	-8.5 ± 0.3	77.6	-296.6 ± 0.4
		saline, LS	-384.2 ± 0.4	7.4 ± 0.2	79.6	-297.3 ± 0.4
MAM	CAPO	water, RF	-81.5 ± 0.2	1.8 ± 0.1	-66.4	-146.0 ± 0.2
		water, LS	-76.8 ± 0.3	4.4 ± 0.2	-76.3	-148.7 ± 0.4
		saline, RF	-80.7 ± 0.3	-3.9 ± 0.3	-65.2	-149.8 ± 0.4
		saline, LS	-75.0 ± 0.6	-0.9 ± 0.2	-74.8	-150.7 ± 0.6
	CHB	water, RF	-116.1 ± 0.1	0.4 ± 0.2	-66.4	-182.1 ± 0.2
		water, LS	-109.2 ± 0.3	4.2 ± 0.2	-76.3	-181.3 ± 0.4
		saline, RF	-115.7 ± 0.3	-4.7 ± 0.3	-65.3	-185.7 ± 0.4
		saline, LS	-109.0 ± 0.2	2.4 ± 0.6	-74.9	-181.5 ± 0.6
	CPOS	water, RF	47.0 ± 0.3	4.9 ± 0.3	-67.1	-15.2 ± 0.4
		water, LS	54.6 ± 0.4	8.0 ± 0.2	-76.3	-13.7 ± 0.4
		saline, RF	47.0 ± 0.2	0.9 ± 0.3	-65.9	-18.0 ± 0.4
		saline, LS	51.6 ± 0.2	4.5 ± 0.3	-74.9	-18.8 ± 0.4
	CNEG	water, RF	-319.8 ± 0.1	-3.6 ± 0.2	-66.2	-389.6 ± 0.2
		water, LS	-312.1 ± 0.2	0.6 ± 0.1	-76.3	-387.9 ± 0.2
		saline, RF	-318.1 ± 0.7	-7.2 ± 0.6	-65.0	-390.3 ± 0.9
		saline, LS	-311.2 ± 0.4	-1.2 ± 0.4	-74.9	-387.3 ± 0.6
	free	water, RF	-248.1 ± 0.3	4.9 ± 0.2	-79.1	-332.1 ± 0.4
		water, LS	-251.8 ± 0.3	1.3 ± 0.1	-81.1	-331.6 ± 0.3
		saline, RF	-232.1 ± 0.4	-19 ± 0.1	-77.6	-328.7 ± 0.4
		saline, LS	-251.9 ± 0.4	2.8 ± 0.2	-79.7	-328.8 ± 0.4

^aResults are reported for two solvents, pure water and 0.5 M saline. Two different methods for electrostatic energies were used, a cutoff-scheme with reaction-field (RF) and lattice-summation (LS) electrostatics. All values are reported in kJ/mol.

over the selected configurations. We have previously shown that the sum of these correction terms is fairly independent of time due to compensating effects in the individual terms.²⁰

Corrections for Binding Free Energies. From the corrected alchemical free energies, physically more relevant free energies of charging (closed hosts) or binding free energies (opened hosts) were calculated by cycle closure of the perturbations in the bound and unbound state. For the free energies that were calculated from (co)alchemical transformations, $\lambda = 0$ and $\lambda = 1$ denote the dummy state versus the state where the ligand is fully interacting. Here, to close the thermodynamic cycle, ΔG_{cor} has to be calculated for both simulations in the bound and unbound states. In the bound state, the following electrostatic interactions are considered between the solutes:

(i) intramolecular interactions within the ligand (L) and
(ii) intermolecular interactions between the ligand and host atoms (H). In the unbound state, only intramolecular interactions are present:

$$\Delta \Delta G^{\text{alch}} = \Delta G_{\text{raw}}^{\text{alch, bound}} + \Delta G_{\text{cor}}^{\text{alch, bound}} [L(L, H)] - \Delta G_{\text{raw}}^{\text{alch, unbound}} - \Delta G_{\text{cor}}^{\text{alch, unbound}} [L(L)] \quad (12)$$

For the free energies that were calculated from path sampling, $\lambda = 0$ and $\lambda = 1$ denote the bound versus unbound state and cycle closure can be achieved from analysis of one system only. However, this has some important consequences on the calculation of the correction terms: (i) while the corrections for alchemical free energies were calculated over all

Table 3. Raw Coalchemical Free Energies ($\Delta G_{\text{raw}}^{\text{coalch}}$), Correction Terms ($\Delta G_{\text{pol}} + \Delta G_{\text{dir}}$, ΔG_{dsm}), and Corrected Coalchemical Free Energies (ΔG^{coalch}) for both Guests Acetate (ACE) and Methylammonium (MAM) in all Four Closed Hosts with an Apolar Cavity (CAPO), with Hydrogen-Bonding Capability (CHB), with a Positively-Charged Cavity (CPOS), with a Negatively-Charged Cavity (CNEG) and Free in Solution^a

LIG	HOST	scheme	$\Delta G_{\text{raw}}^{\text{coalch}}$	$\Delta G_{\text{pol}}^{\text{coalch}} + \Delta G_{\text{dir}}^{\text{coalch}}$	$\Delta G_{\text{dsm}}^{\text{coalch}}$	ΔG^{coalch}
ACE	CAPO	water, RF	675.3 ± 2.7	4.1 ± 1.4	-10.4	669.0 ± 3.8
		water, LS	669.5 ± 1.1	2.3 ± 1.5	0.0	671.7 ± 1.6
		saline, RF	678.7 ± 1.6	5.4 ± 1.3	-9.9	674.2 ± 2.3
		saline, LS	676.0 ± 1.8	0.8 ± 2.2	0.0	676.8 ± 2.5
	CHB	water, RF	631.7 ± 0.9	3.7 ± 1.3	-10.3	625.2 ± 1.3
		water, LS	625.0 ± 1.2	2.6 ± 1.6	0.0	627.6 ± 1.7
		saline, RF	637.0 ± 1.4	5.9 ± 1.8	-10.1	632.9 ± 2.0
		saline, LS	632.7 ± 2.3	1.7 ± 1.6	0.0	634.4 ± 3.3
	CPOS	water, RF	455.9 ± 2.3	-2.8 ± 1.4	-10.1	443.0 ± 3.3
		water, LS	443.0 ± 1.2	1.9 ± 1.3	0.0	444.9 ± 1.7
		saline, RF	461.2 ± 1.2	2.2 ± 2.5	-9.6	453.8 ± 1.7
		saline, LS	454.7 ± 2.1	0.3 ± 1.6	0.0	454.9 ± 3.0
	CNEG	water, RF	792.3 ± 2.6	11.4 ± 1.3	-9.9	793.8 ± 3.7
		water, LS	791.4 ± 0.9	2.6 ± 1.2	0.0	793.9 ± 1.3
		saline, RF	797.0 ± 2.6	8.5 ± 1.5	-10.4	795.1 ± 3.7
		saline, LS	795.5 ± 1.7	3.3 ± 1.7	0.0	798.8 ± 2.4
	free	water, RF	439.4 ± 2.8	1.0 ± 0.9	-0.8	439.5 ± 4.0
		water, L	442.0 ± 3.4	0.3 ± 1.5	0.0	442.2 ± 4.8
		saline, RF	451.7 ± 2.3	-1.8 ± 1.3	-0.7	449.1 ± 3.3
		saline, LS	453.7 ± 1.8	-2.4 ± 1.2	0.0	451.4 ± 2.5
MAM	CAPO	water, RF	314.3 ± 2.4	4.0 ± 1.4	9.3	327.6 ± 3.4
		water, LS	324.9 ± 1.3	1.5 ± 0.5	0.0	326.5 ± 1.8
		saline, RF	315.5 ± 1.2	0.3 ± 1.7	9.2	325.0 ± 1.7
		saline, LS	324.9 ± 1.1	0.6 ± 2.8	0.0	325.6 ± 1.6
	CHB	water, RF	279.6 ± 1.1	3.4 ± 1.7	9.3	292.2 ± 1.6
		water, LS	290.5 ± 2.1	1.3 ± 1.5	0.0	291.8 ± 3.0
		saline, RF	280.2 ± 1.3	5.2 ± 3.4	9.5	294.9 ± 1.8
		saline, LS	292.2 ± 2.4	3.8 ± 2.3	0.0	296.1 ± 3.4
	CPOS	water, RF	440.6 ± 2.7	10.0 ± 1.5	8.7	459.3 ± 3.8
		water, LS	456.5 ± 2.5	3.5 ± 0.7	0.0	460.0 ± 3.5
		saline, RF	440.5 ± 1.3	7.7 ± 0.6	8.5	456.7 ± 1.8
		saline, LS	457.6 ± 1.6	-1.7 ± 1.4	0.0	456.0 ± 2.3
	CNEG	water, RF	78.4 ± 2.7	-3.1 ± 1.3	9.6	84.9 ± 3.8
		water, LS	83.4 ± 1.8	0.8 ± 1.4	0.0	84.2 ± 2.5
		saline, R	79.5 ± 1.0	-1.3 ± 1.5	9.6	87.9 ± 1.4
		saline, LS	86.9 ± 1.3	0.3 ± 1.7	0.0	87.1 ± 1.8
	free	water, RF	140.6 ± 1.4	-0.3 ± 1.9	0.6	140.9 ± 2.0
		water, LS	140.3 ± 2.6	0.6 ± 1.4	0.0	140.8 ± 3.7
		saline, RF	148.6 ± 0.7	-2.1 ± 0.7	0.6	147.2 ± 1.0
		saline, LS	151.5 ± 2.3	-3.3 ± 1.5	0.0	148.3 ± 3.3

^aResults are reported for two solvents, pure water and 0.5 M saline. Two different methods for electrostatic energies were used, a cutoff-scheme with reaction-field (RF) and lattice-summation (LS) electrostatics. All values are reported in kJ/mol.

simulated λ -points, the corrections for path sampling were calculated only for the end points $\lambda = 0$ and $\lambda = 1$. These correspond to the situations where the ligand is fully bound to or dissociated from the host (with distances 0 or 2.65 nm); (ii) in the case of the alchemical correction terms, the electrostatic energies were recalculated for the perturbed atom sites only. In the case of correction terms for the path-sampling method, the electrostatic energies were recalculated for all atoms of both binding partners; (iii) as the electrostatic energies were calculated using a cutoff scheme with a 1.4 nm cutoff distance, the dissociated binding partners are believed to be exempt of any interaction at $\lambda = 1$ (with a distance of 2.65 nm). However, applying the correction terms $\Delta G_{\text{pol}} + \Delta G_{\text{dir}}$ would eliminate the cutoff scheme in favor of a fully macroscopic situation. To ensure that the corrected energies at $\lambda = 1$ still

correspond to a fully unbound situation, the corrections were calculated for ligand (L) and host (H) atoms individually where both charges and Lennard–Jones radii of the respective other binding partner were set to zero. From that, it follows that both ligand and host molecules are only interacting with themselves ($L(L)$ and $H(H)$)

$$\begin{aligned} \Delta \Delta G^{\text{path}} = & \Delta G_{\text{raw}}^{\text{path}} + \Delta G_{\text{cor}}^{\text{path}, \lambda=0}[L(L, H)] \\ & + \Delta G_{\text{cor}}^{\text{path}, \lambda=0}[H(L, H)] - \Delta G_{\text{cor}}^{\text{path}, \lambda=1}[L(L)] \\ & - \Delta G_{\text{cor}}^{\text{path}, \lambda=1}[H(H)] \end{aligned} \quad (13)$$

This ensures that at $\lambda = 1$, both terms, ΔG_{pol} and ΔG_{dir} correspond to a fully dissociated situation in the corrected binding free energies.

Table 4. Raw and Corrected Free Energies of Charging ($\Delta\Delta G_{\text{raw}}$, $\Delta\Delta G$) Using Different Methodologies with Acetate (ACE) in the Closed Buckyballs^a

LIG	HOST	scheme	$\Delta\Delta G_{\text{raw}}$	$\Delta\Delta G$	$\Delta\Delta G_{\text{cor}}$	
ACE	CAPO	water, alchem, RF	237.1 ± 0.4	228.2 ± 0.4	9.0 ± 0.2	
		water, alchem, LS	231.6 ± 0.6	229.8 ± 0.6	1.8 ± 0.2	
		water, coalch, RF	235.9 ± 3.9	229.5 ± 5.5	6.4 ± 1.7	
		water, coalch, LS	227.3 ± 3.6	229.5 ± 5.1	2.3 ± 2.1	
		saline, alchem, RF	226.9 ± 0.4	224.0 ± 0.6	2.9 ± 0.5	
		saline, alchem, LS	231.5 ± 0.5	225.8 ± 0.6	5.7 ± 0.4	
		saline, coalch, RF	227.0 ± 2.8	225.0 ± 4.0	1.9 ± 1.8	
		saline, coalch, LS	222.3 ± 2.5	225.4 ± 3.5	3.2 ± 2.5	
		CHB	water, alchem, RF	193.6 ± 0.3	183.8 ± 0.4	9.8 ± 0.2
			water, alchem, LS	187.5 ± 0.4	185.7 ± 0.6	1.8 ± 0.3
			water, coalch, RF	192.3 ± 2.9	185.6 ± 4.2	6.7 ± 1.6
			water, coalch, LS	182.8 ± 3.6	185.4 ± 5.1	2.6 ± 2.2
	saline, alchem, RF		183.0 ± 0.5	180.6 ± 0.7	2.4 ± 0.5	
	saline, alchem, LS		189.0 ± 0.6	183.5 ± 0.7	5.5 ± 0.4	
	saline, coalch, RF		185.4 ± 2.7	183.8 ± 3.9	1.6 ± 2.2	
	saline, coalch, LS		179.0 ± 2.9	183.1 ± 4.1	4.1 ± 2.0	
	CPOS		water, alchem, RF	16.2 ± 0.4	1.9 ± 0.5	14.3 ± 0.3
			water, alchem, LS	9.5 ± 0.4	4.2 ± 0.4	5.3 ± 0.1
			water, coalch, RF	16.5 ± 3.6	4.5 ± 5.2	13.0 ± 1.7
			water, coalch, LS	0.8 ± 3.6	2.6 ± 5.1	1.9 ± 2.0
		saline, alchem, RF	4.7 ± 0.4	1.3 ± 0.6	3.4 ± 0.6	
		saline, alchem, LS	10.0 ± 0.6	0.8 ± 0.7	9.2 ± 0.4	
		saline, coalch, RF	9.5 ± 2.6	4.7 ± 3.7	4.8 ± 2.8	
		saline, coalch, LS	0.9 ± 2.8	3.5 ± 3.9	2.6 ± 2.0	
	CNEG	water, alchem, RF	355.9 ± 0.3	353.1 ± 0.3	2.8 ± 0.1	
		water, alchem, LS	351.2 ± 0.5	353.5 ± 0.5	4.8 ± 0.1	
		water, coalch, RF	352.9 ± 3.8	354.3 ± 5.4	1.3 ± 1.6	
		water, coalch, LS	349.1 ± 3.5	351.7 ± 5.0	2.6 ± 1.9	
		saline, alchem, RF	345.3 ± 0.8	347.5 ± 0.9	2.2 ± 0.4	
		saline, alchem, LS	351.6 ± 0.6	350.3 ± 0.6	1.2 ± 0.4	
		saline, coalch, RF	345.3 ± 3.5	346.0 ± 5.0	0.7 ± 2.0	
		saline, coalch, LS	341.8 ± 2.5	347.4 ± 3.5	5.6 ± 2.1	

^aThe effects of the total artifacts in the free energies of charging ($\Delta\Delta G_{\text{cor}} = \Delta\Delta G_{\text{raw}} - \Delta\Delta G_{\text{raw}}$, see eq 4) are reported in the last column. All values are reported in kJ/mol.

Mind that for the calculations of ΔG_{pol} , a fully solvated system with implicit solvent is considered, while all electrostatic energies for the calculation of ΔG_{dir} are vacuum energies. While the solvent was not explicitly mentioned in eqs 12 and 13, an implicit solvent is present for the calculations of ΔG_{pol} and polarization effects are considered from both ligand and host molecules in the cases of $L(LH)$ and $H(LH)$ or only one molecule in the cases of $L(L)$ and $H(H)$.

RESULTS AND DISCUSSION

(Co-)Alchemical Transformations in the Closed Host.

For all systems, the alchemical and coalchemical raw free energies were corrected using the post-simulation free-energy correction from eq 4. To simplify further discussions, all calculated artifacts terms are reported as two terms: (i) a group $\Delta G_{\text{pol}} + \Delta G_{\text{dir}}$ that corrects wrong electrostatic energies due to a spurious polarization of the solvent and incorrect vacuum-energies between the solute atoms due to the use of an incorrect solvent model, cutoff truncation or periodicity; (ii) a term ΔG_{dsm} that corrects for the convention to sum the potentials over individual solvent point charges in a periodic environment. These terms are reported together with raw free energies and corrected free energies in Tables 2 and 3. The method used for the calculation of the electrostatic energies

(RF or LS) has an impact on the raw free energies (column ΔG_{raw}). These methodology-dependent artifacts were successfully eliminated in the corrected free energies (column ΔG). In the case of the alchemical free energies, the presence of 0.5 M saline (mobile ions) had only a negligible effect on both raw free energies and corrected free energies for the bound systems (Table 2). However, the effect of mobile ions on the free systems is more pronounced. The same effect can be observed in the coalchemical free energies (Table 3), as these systems contain a charge-perturbed mobile ion free in solution. Independent of the method used for calculation of the electrostatic energies, the presence of mobile ions has a non-observable effect on the electrostatic artifacts and the connected size of the correction terms.

The coalchemical ion approach does not eliminate artifacts from cutoff effects (RF) or periodicity (LS) in a satisfying way (column $\Delta G_{\text{pol}} + \Delta G_{\text{dir}}$ in Table 3). This was expected, as both perturbed ligands are not embedded in the same environment: a ligand bound to a solvent-excluded cavity in the host versus a fully solvated ion. Furthermore, in the case of LS electrostatics, both perturbed moieties are not only interacting with their own periodic replicates but also with the periodic replicates of the respective coalchemical partner (ligand with periodic ion/ion with periodic ligand). This represents an additional artifact

Table 5. Raw and Corrected Free Energies of Charging ($\Delta\Delta G_{\text{raw}}$, $\Delta\Delta G$) Using Different Methodologies with Methylammonium (MAM) in the Closed Buckyballs^a

LIG	HOST	scheme	$\Delta\Delta G_{\text{raw}}$	$\Delta\Delta G$	$\Delta\Delta G_{\text{cor}}$	
MAM	CAPO	water, alchem, RF	166.6 ± 0.4	186.0 ± 0.4	19.4 ± 0.2	
		water, alchem, LS	174.9 ± 0.4	182.9 ± 0.5	8.0 ± 0.2	
		water, coalch, RF	173.7 ± 2.8	186.7 ± 3.9	13.0 ± 2.4	
		water, coalch, LS	184.1 ± 2.9	185.6 ± 4.1	1.5 ± 1.5	
		saline, alchem, RF	151.4 ± 0.5	178.9 ± 0.6	27.5 ± 0.3	
		saline, alchem, LS	176.9 ± 0.7	178.1 ± 0.7	1.1 ± 0.3	
		saline, coalch, RF	166.9 ± 1.4	177.8 ± 2.0	10.9 ± 1.8	
		saline, coalch, LS	173.4 ± 2.5	177.3 ± 3.7	3.9 ± 3.2	
		CHB	water, alchem, RF	132.0 ± 0.3	150.0 ± 0.4	18.0 ± 0.3
			water, alchem, LS	142.6 ± 0.4	150.3 ± 0.5	7.7 ± 0.2
			water, coalch, RF	139.0 ± 1.8	151.3 ± 2.6	12.3 ± 2.5
			water, coalch, LS	149.7 ± 3.3	150.9 ± 4.8	1.3 ± 2.1
	saline, alchem, RF		116.4 ± 0.5	143.0 ± 0.6	26.6 ± 0.3	
	saline, alchem, LS		142.9 ± 0.4	147.3 ± 0.7	4.4 ± 0.6	
	saline, coalch, RF		131.6 ± 1.5	147.8 ± 2.1	16.2 ± 3.5	
	saline, coalch, LS		140.7 ± 3.3	147.8 ± 4.7	7.1 ± 2.7	
	CPOS		water, alchem, RF	295.0 ± 0.4	316.9 ± 0.6	21.8 ± 0.4
			water, alchem, LS	306.4 ± 0.5	317.9 ± 0.5	11.5 ± 0.2
			water, coalch, RF	300.0 ± 3.0	318.4 ± 4.3	18.3 ± 2.4
			water, coalch, LS	315.7 ± 3.6	319.2 ± 5.1	3.5 ± 1.6
		saline, alchem, RF	279.1 ± 0.4	310.7 ± 0.6	31.6 ± 0.3	
		saline, alchem, LS	303.5 ± 0.4	310.0 ± 0.6	6.5 ± 0.4	
		saline, coalch, RF	291.9 ± 1.5	309.6 ± 2.1	17.6 ± 0.9	
		saline, coalch, LS	306.1 ± 2.8	307.7 ± 4.0	1.6 ± 2.1	
	CNEG	water, alchem, RF	-71.7 ± 0.3	-57.6 ± 0.4	14.2 ± 0.3	
		water, alchem, LS	-60.3 ± 0.4	-56.2 ± 0.4	4.1 ± 0.1	
		water, coalch, RF	-62.2 ± 3.0	-56.0 ± 4.3	6.1 ± 2.3	
		water, coalch, LS	-57.4 ± 3.2	-56.7 ± 4.5	0.8 ± 2.0	
		saline, alchem, RF	-86.0 ± 0.8	-61.7 ± 1.0	24.4 ± 0.6	
		saline, alchem, LS	-59.3 ± 0.6	-58.5 ± 0.7	0.8 ± 0.4	
		saline, coalch, RF	-69.1 ± 1.2	-59.3 ± 1.7	9.8 ± 1.7	
		saline, coalch, LS	-64.7 ± 2.6	-61.1 ± 3.8	3.5 ± 2.3	

^aThe effects of the total artifacts in the free energies of charging ($\Delta\Delta G_{\text{cor}} = \Delta\Delta G_{\text{raw}} - \Delta\Delta G_{\text{raw}}$, see eq 4) are reported in the last column. All values are reported in kJ/mol.

in the coalchemical ion strategy. However, the discrete solvent effect is reduced in the case of RF or completely eliminated in the case of LS electrostatics (ΔG_{dsm} in Table 3). This was expected, as according to eqs 8, and 9 this artifact is proportional to the solvent density around the perturbed moieties. In the case of LS electrostatics, the solvent interactions of both coalchemical partners is calculated over the entire simulation box, while in the case of RF, these interactions are calculated over a sphere with a defined cutoff radius that was defined a priori. Here, the difference in the solvent density between both cutoff spheres of the coalchemical partners depends mainly on the volume of the solvent-excluded host cavity. In the case of simulations of the unbound ligand, the value for ΔG_{dsm} is roughly 0 for the lack of a solvent-excluded cavity.

From both raw and corrected free energies, physically more meaningful contributions of the charging process to the binding free energies were calculated from the raw and the corrected values (Tables 4 and 5). The size of the electrostatic artifacts contained in the raw free energies ($\Delta\Delta G_{\text{cor}} = \Delta\Delta G - \Delta\Delta G_{\text{raw}}$) diminished compared to the artifacts contained in the alchemical raw free energies ($\Delta G - \Delta G_{\text{raw}}$ in Tables 2 and 3). It follows that a major part of the artifacts is eliminated upon application of a thermodynamic cycle. However, in neither case (RF vs LS, alchemical vs coalchemical, water vs saline) do the artifacts

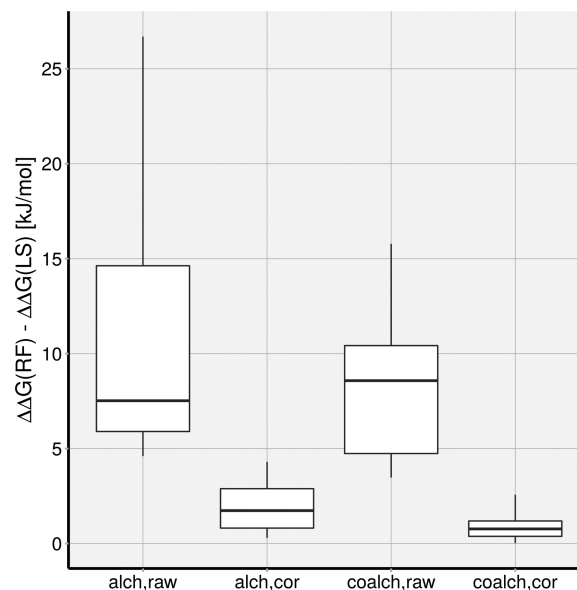


Figure 2. Comparison of the methodological differences between free energies from RF and LS methods (y-axis), for raw and corrected data. Each box plot represents 16 datapoints for the different systems in different solvents.

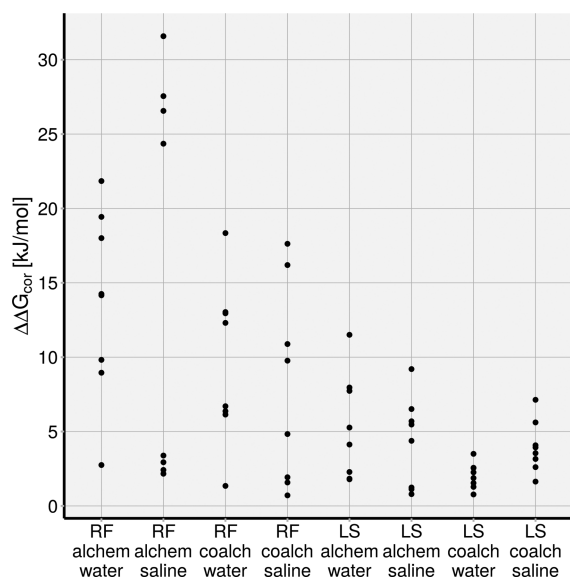


Figure 3. Comparison of the total size of the artifacts for the eight different systems (2 ligands, 4 hosts) using different methods. The y-axis represents the total correction estimates for the charging free energies.

disappear completely. Figure 2 compares the spread in the differences between free energies from RF and LS methods for the different systems in different solvents. Application of correction terms reduces the differences between both methods by a considerable amount. This shows the requirement for post-simulation correction terms to achieve methodological independence and is true for both alchemical and coalchemical free energies.

Figure 3 compares the size of the artifacts between the different methods (RF/LS, alchemic/coalchemic, water/saline). In total, the spread of the artifacts seems to be bigger for the RF methods compared to the LS method and for the alchemical perturbation scheme compared to the coalchemical scheme. The impact of the solvent (water vs saline) seems to be negligible. Further, the artifacts reported as $\Delta G_{\text{pol}} + \Delta G_{\text{dir}}$ and ΔG_{dsm} were averaged over the different systems (ACE/MAM, CAPO/CHB/CPOS/CNEG, see Table S1). The terms $\Delta G_{\text{pol}} + \Delta G_{\text{dir}}$ are not significantly different for coalchemical free energies compared to their alchemical counterpart. The ΔG_{dsm} term, however, is significantly lower for free energies that were calculated using LS electrostatics compared to the RF method. However, this is true for the test systems where the water molecule densities for the guest/host systems were roughly the same compared to the free guests, while the densities within the cutoff radii differ by a

Table 6. Raw Alchemical Free Energies ($\Delta G_{\text{raw}}^{\ominus, \text{alch}}$), Correction Terms ($\Delta G_{\text{pol}} + \Delta G_{\text{dir}}$, ΔG_{dsm}), and Corrected Alchemical Free Energies ($\Delta G^{\ominus, \text{alch}}$) for both Guests Acetate (ACE) and Methylammonium (MAM) in all Four Opened Hosts with an Apolar Cavity (CAPO), with Hydrogen-Bonding Capability (CHB), with a Positively-Charged Cavity (CPOS), with a Negatively-Charged Cavity (CNEG), and as Free in Solution^a

LIG	HOST	$\Delta G_{\text{raw}}^{\ominus, \text{alch}}$	$\Delta G_{\text{pol}} + \Delta G_{\text{dir}}$	ΔG_{dsm}	$\Delta G^{\ominus, \text{alch}}$
ACE	CAPO	-214.7 ± 1.0	-0.4 ± 0.1	68.1	-147.0 ± 1.0
	CHB	-233.3 ± 1.9	-0.1 ± 0.1	68.3	-165.2 ± 1.9
	CPOS	-371.2 ± 1.6	-0.5 ± 0.2	68.4	-303.3 ± 1.6
	CNEG	-108.7 ± 5.7	-0.3 ± 0.1	68.7	-40.3 ± 5.7
	free	-357.7 ± 0.7	-3.7 ± 0.0	79.1	-282.3 ± 0.7
MAM	CAPO	-138.1 ± 4.1	-2.5 ± 0.2	-68.4	-209.0 ± 4.1
	CHB	-146.2 ± 3.8	-1.7 ± 0.1	-68.7	-216.7 ± 3.8
	CPOS	28.1 ± 1.7	-5.0 ± 0.4	-68.7	-45.6 ± 1.7
	CNEG	-248.2 ± 7.2	0.1 ± 0.0	-8.1	-316.3 ± 7.2
	free	-230.0 ± 1.3	-3.8 ± 0.1	-79.1	-312.9 ± 1.3

^aFree energies in the bound state are corrected for the standard states. All values are reported in kJ/mol.

Table 7. Raw Binding Free Energies ($\Delta G_{\text{raw}}^{\ominus, \text{path}}$) from Path Sampling (Using RF Electrostatics)^a

LIG	HOST	$\Delta G_{\text{raw}}^{\ominus, \text{path}}$	state	$\Delta G_{\text{pol}} + \Delta G_{\text{dir}}$	ΔG_{dsm}	$\Delta G^{\ominus, \text{path}}$
ACE	CAPO	139.1 ± 0.8	bound	-0.3 ± 0.1	66.9	
			unbnd	-0.4 ± 0.1	76.3	129.9 ± 0.8
	CHB	124.8 ± 1.2	bound	1.4 ± 0.2	67.4	
			unbnd	0.3 ± 0.1	77.8	115.6 ± 1.2
	CPOS	-13.9 ± 1.1	bound	-0.4 ± 0.1	0.0	
			unbnd	-1.1 ± 0.3	7.2	-20.4 ± 1.1
CNEG	236.3 ± 1.8	bound	7.1 ± 0.6	138.0		
		unbnd	-1.5 ± 0.2	140.2	242.7 ± 1.9	
MAM	CAPO	96.9 ± 1.0	bound	-0.6 ± 0.2	-68.3	
			unbnd	-0.5 ± 0.1	-73.9	102.4 ± 1.0
	CHB	91.9 ± 0.9	bound	-0.9 ± 0.3	-66.3	
			unbnd	-0.5 ± 0.1	-73.6	98.8 ± 1.0
	CPOS	257.5 ± 1.7	bound	6.6 ± 0.5	-138.7	
			unbnd	-1.3 ± 0.2	-141.8	268.5 ± 1.8
CNEG	-10.9 ± 1.4	bound	-0.5 ± 0.1	1.4		
		unbnd	-1.7 ± 0.4	-6.8	-1.5 ± 1.5	

^aCorrection terms ($\Delta G_{\text{pol}} + \Delta G_{\text{dir}}$ and ΔG_{dsm}) are reported for the bound and unbound states in two separate lines. The corrected binding free energies ($\Delta G^{\ominus, \text{path}}$) are reported in the last column. All values are reported in kJ/mol.

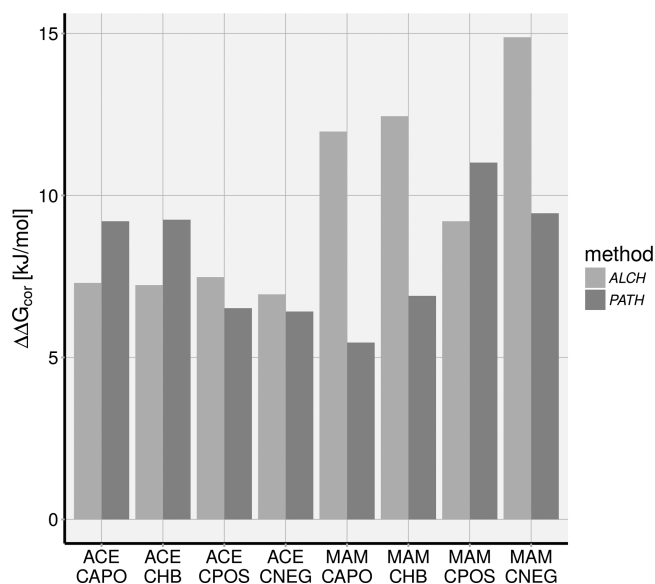


Figure 4. Comparison of the total size of the artifacts for the 8 different systems (2 ligands, 4 hosts) between free energies from alchemical perturbations vs path sampling. The y-axis represents the total correction estimates for the binding free energies.

greater amount. In the case of coalchemical free energies calculated with LS electrostatics, the ΔG_{dsm} term drops out completely, thereby reducing the total artifacts. However, note that, for the coalchemical approach with LS electrostatics, in a solvent with mobile charges (as recommended by Chen et al.²⁸), the corrections still count to 4 kJ/mol, roughly the accepted accuracy of alchemical methods in drug design.^{1,64}

Alchemical Transformations and Path Sampling in the Opened Hosts. The raw free energies and corresponding standard-state corrections for both the alchemical and the path sampling approaches are reported in Table S1.

Electrostatic artifacts in the raw free energies from simulations of the open buckyballs were analyzed for both methods: alchemical and path sampling. The artifacts for the alchemical DDM method were calculated analogously to the method used for the closed buckyballs, and the reporting is similar. All values for the correction terms are on a comparable scale relative to the closed buckyballs (compare Tables 2 and 6). The exact size of the individual correction terms calculated using both eqs 6 and 7 shown in Table S4. A root-mean-square difference between the two approaches of 0.27 kJ/mol (maximum deviation of 0.7 kJ/mol) was observed. In the case of path sampling, the artifacts were calculated

for $\lambda = 0$ (bound state) and $\lambda = 1$ (unbound state). These were used to correct raw free energies from path sampling (Table 7). As expected, raw binding free energies from alchemical modifications and path sampling were not in agreement with each other using the RF method. The free energies from path-sampling were also artifacted, with the size and sign of the artifacts mainly depending on the charges of the individual molecules. Eliminating these artifacts by applying the correction terms to the free energies from both methodologies gave consistent results (Table S2). The total size of the artifacts depends on the charges of the molecules.

The total artifacts for both alchemical transformations and path sampling are on a very comparable scale on average but not for the individual systems (Figure 4). The corrections applied to pathway methods are not less compared to those applied to alchemical methods. Mind that, in the case of alchemical perturbations, these values report the artifacts for the perturbed molecule. In the case of path sampling, these report the artifacts for both binding partners (compare eqs 12 and 13). Root-mean-square differences (RMSD) between free energies from alchemical and pathway methods were calculated for the raw data and for the corrected data (Table 8). The RMSD decreased from 6.4 to 2.4 kJ/mol, indicating much smaller deviations between the corrected values compared to the raw values. Both groups of errors (from cutoff artifacts and the summation scheme) are equally important, dependent on the charge configuration in the system. Table S3 reports the individual contributions to the artifacts. This is an important consideration since ΔG_{dsm} would be zero for path sampling in case a LS scheme would have been used (as long as the water molecule density would not change between bound state vs. unbound state). However, artifacts would still play a role, arising from spurious polarization and electrostatic energies arising from the periodicity. To estimate the size of the artifacts for pathway methods using LS electrostatics, the conformations that were generated using RF electrostatics were analyzed using LS conditions (Table 9, also see values for LS in Table S1). The estimated values of $\Delta G_{\text{cor}}^{\text{path}}$ for LS schemes are smaller than under RF conditions but still amount to about 5–6 kJ/mol for ligands and hosts of equally signed net charges.

CONCLUSIONS

Two methodologies were studied in terms of their capabilities to eliminate methodology-dependent artifacts in the electrostatic free energies: (i) a coalchemical-ion approach where the elimination of a charged ligand is coupled to the charging process of a distant ion and (ii) explicit sampling of the dissociation path

Table 8. Comparison of Raw and Corrected Binding Free Energies (Using RF Electrostatics) for Alchemical and Path-Sampling Methods with Root-Mean-Square Differences (RMSD) between them^a

LIG	HOST	$\Delta\Delta G_{\text{raw}}^{\ominus, \text{alch}}$	$\Delta G_{\text{raw}}^{\ominus, \text{path}}$	$\Delta\Delta G^{\ominus, \text{alch}}$	$\Delta G^{\ominus, \text{path}}$	$\Delta G_{\text{cor}}^{\text{alch}}$	$\Delta G_{\text{cor}}^{\text{path}}$
ACE	CAPO	143.1 ± 1.2	139.1 ± 0.8	135.3 ± 1.2	129.9 ± 0.8	7.3 ± 0.1	9.2 ± 0.8
	CHB	124.4 ± 2.0	124.8 ± 1.2	117.2 ± 2.0	115.6 ± 1.2	7.2 ± 0.1	9.2 ± 1.2
	CPOS	−13.5 ± 1.7	−13.9 ± 1.1	−21.0 ± 1.8	−20.4 ± 1.1	7.5 ± 0.2	6.5 ± 1.1
	CNEG	249.0 ± 5.7	236.3 ± 1.8	242.0 ± 5.7	242.7 ± 1.9	7.0 ± 0.1	6.4 ± 1.9
MAM	CAPO	91.9 ± 4.3	96.9 ± 1.0	103.9 ± 4.3	102.4 ± 1.0	12.0 ± 0.2	5.5 ± 1.0
	CHB	83.8 ± 4.0	91.9 ± 0.9	96.2 ± 4.0	98.8 ± 1.0	12.5 ± 0.1	6.9 ± 1.0
	CPOS	258.0 ± 2.1	257.5 ± 1.7	267.2 ± 2.2	268.5 ± 1.8	9.2 ± 0.4	11.0 ± 1.8
	CNEG	−18.3 ± 7.3	−10.9 ± 1.4	−3.4 ± 7.3	−1.5 ± 1.5	14.9 ± 0.0	9.4 ± 1.5
	RMSD			6.4		2.4	

^aThe total artifacts in the binding free energies are reported for alchemical energies and free energies from path sampling. All values are reported in kJ/mol.

Table 9. Estimated Artifacts ($\Delta G_{\text{pol}} + \Delta G_{\text{dir}}$) for Path-Sampling Methods under LS Electrostatics^a

LIG	HOST	state	$\Delta G_{\text{pol}} + \Delta G_{\text{dir}}$	$\Delta G_{\text{cor}}^{\text{path}}$
ACE	CAPO	bound	-1.1	
		unbnd	-0.1	-1.0
	CHB	bound	0.2	
		unbnd	0.0	0.2
	CPOS	bound	-0.4	
		unbnd	-1.0	0.6
CNEG	bound	3.6		
	unbnd	-1.4	5.0	
MAM	CAPO	bound	-1.3	
		unbnd	-0.2	-1.1
	CHB	bound	-1.6	
		unbnd	-0.1	-1.5
	CPOS	bound	4.1	
		unbnd	-1.5	5.6
CNEG	bound	-0.5		
	unbnd	-0.7	0.2	

^aHere, only artifacts arising from the spurious polarization and spurious direct interactions ($\Delta G_{\text{pol}} + \Delta G_{\text{dir}}$) were considered. The total artifacts $\Delta G_{\text{cor}}^{\text{path}}$ were calculated from the differences between bound and unbound states on the configurations generated with RF electrostatics. All values are reported in kJ/mol.

to avoid a change in the net charge of the system. Both methods were proven insufficient to eliminate methodology-dependent artifacts completely. Using RF electrostatics, errors still arise from spurious polarization of the solvent, wrong solute–solute interactions, and the use of a summation scheme over point charges, a necessity arising from periodic boundary conditions. These errors remain due to the fact that the situations for a distant ion versus the bound ligand or the unbound ligand versus the bound ligand are different in terms of electrostatic potentials and solvent densities. Using LS electrostatics, artifacts due to the summation over solvent point charges are successfully eliminated. However, a spurious solvent polarization and wrong solute–solute interactions still contribute to artifacted potentials. As a consequence, the application of a post-simulation correction scheme (as the one used within this study, which is based on tools provided within the GROMOS++ simulation package)²⁰ is highly recommended to achieve consistent results, independent of the applied methodology.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/acs.jctc.0c00719>.

Tables with individual contributions of the correction terms to the charging free energy (Table S1), standard state corrections for alchemical and path sampling methods (Table S2), individual contributions of the correction terms to the binding free energies from alchemical and path sampling methods (Table S3) (PDF)

■ AUTHOR INFORMATION

Corresponding Author

Chris Oostenbrink – Institute of Molecular Modeling and Simulation, University of Natural Resources and Life Sciences, Vienna 1190, Austria; orcid.org/0000-0002-4232-2556; Email: chris.oostenbrink@boku.ac.at

Authors

Christoph Öhlknecht – Institute of Molecular Modeling and Simulation, University of Natural Resources and Life Sciences, Vienna 1190, Austria; orcid.org/0000-0003-1847-1719

Jan Walther Perthold – Institute of Molecular Modeling and Simulation, University of Natural Resources and Life Sciences, Vienna 1190, Austria; orcid.org/0000-0002-8575-0278

Bettina Lier – Institute of Molecular Modeling and Simulation, University of Natural Resources and Life Sciences, Vienna 1190, Austria; orcid.org/0000-0002-8032-0084

Complete contact information is available at: <https://pubs.acs.org/doi/10.1021/acs.jctc.0c00719>

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

J.W.P. and B.L. are recipients of DOC Fellowships of the Austrian Academy of Sciences (ÖAW) at the Institute for Molecular Modeling and Simulation at the University of Natural Resources and Life Sciences, Vienna (grant nos. 24987 and 25743). The computational results presented have been achieved in part using the Vienna Scientific Cluster (VSC).

■ REFERENCES

- (1) Shirts, M. R.; Mobley, D. L.; Brown, S. P. In *Drug Design: Structure- and Ligand-Based Approaches*; Merz, K. M., Jr., Ringe, D., Reynolds, C. H., Eds.; Cambridge University Press: 2010; pp. 61–86.
- (2) Hou, T.; Wang, J.; Li, Y.; Wang, W. Assessing the Performance of the MM/PBSA and MM/GBSA Methods. 1. The Accuracy of Binding Free Energy Calculations Based on Molecular Dynamics Simulations. *J. Chem. Inf. Model.* **2011**, *51*, 69–82.
- (3) Gumbart, J. C.; Roux, B.; Chipot, C. Standard Binding Free Energies from Computer Simulations: What Is the Best Strategy? *J. Chem. Theory Comput.* **2013**, *9*, 794–802.
- (4) Hansen, N.; van Gunsteren, W. F. Practical Aspects of Free-Energy Calculations: A Review. *J. Chem. Theory Comput.* **2014**, *10*, 2632–2647.
- (5) Duan, L.; Liu, X.; Zhang, J. Z. H. Interaction Entropy: A New Paradigm for Highly Efficient and Reliable Computation of Protein–Ligand Binding Free Energy. *J. Am. Chem. Soc.* **2016**, *138*, 5722–5728.
- (6) Cournia, Z.; Allen, B.; Sherman, W. Relative Binding Free Energy Calculations in Drug Discovery: Recent Advances and Practical Considerations. *J. Chem. Inf. Model.* **2017**, *57*, 2911–2937.
- (7) Fu, H.; Gumbart, J. C.; Chen, H.; Shao, X.; Cai, W.; Chipot, C. BFEE: A User-Friendly Graphical Interface Facilitating Absolute Binding Free-Energy Calculations. *J. Chem. Inf. Model.* **2018**, *58*, 556–560.
- (8) Cabeza de Vaca, I.; Qian, Y.; Vilseck, J. Z.; Tirado-Rives, J.; Jorgensen, W. L. Enhanced Monte Carlo Methods for Modeling Proteins Including Computation of Absolute Free Energies of Binding. *J. Chem. Theory Comput.* **2018**, *14*, 3279–3288.
- (9) de Ruiter, A.; Oostenbrink, C. Advances in the calculation of binding free energies. *Curr. Opin. Struct. Biol.* **2020**, *61*, 207–212.
- (10) de Ruiter, A.; Oostenbrink, C. Protein–Ligand Binding from Distancefield Distances and Hamiltonian Replica Exchange Simulations. *J. Chem. Theory Comput.* **2013**, *9*, 883–892.
- (11) Villarreal, O. D.; Yu, L.; Rodriguez, R. A.; Chen, L. Y. Computing the binding affinity of a ligand buried deep inside a protein with the hybrid steered molecular dynamics. *Biochem. Biophys. Res. Commun.* **2017**, *483*, 203–208.
- (12) Gilson, M. K.; Given, J. A.; Bush, B. L.; McCammon, J. A. The statistical-thermodynamic basis for computation of binding affinities: a critical review. *Biophys. J.* **1997**, *72*, 1047–1069.

- (13) Cruz, J.; Wickstrom, L.; Yang, D.; Gallicchio, E.; Deng, N. Combining Alchemical Transformation with a Physical Pathway to Accelerate Absolute Binding Free Energy Calculations of Charged Ligands to Enclosed Binding Sites. *J. Chem. Theory Comput.* **2020**, *16*, 2803–2813.
- (14) 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. *J. Chem. Phys.* **2006**, *124*, 224501.
- (15) Reif, M. M.; Hünenberger, P. H. Computation of methodology-independent single-ion solvation properties from molecular simulations. III. Correction terms for the solvation free energies, enthalpies, entropies, heat capacities, volumes, compressibilities, and expansivities of solvated ions. *J. Chem. Phys.* **2011**, *134*, 144103.
- (16) Reif, M. M.; Hünenberger, P. H.; Oostenbrink, C. New Interaction Parameters for Charged Amino Acid Side Chains in the GROMOS Force Field. *J. Chem. Theory Comput.* **2012**, *8*, 3705–3723.
- (17) 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 latticesum methods: An accurate correction scheme for electrostatic finite-size effects. *J. Chem. Phys.* **2013**, *139*, 184103.
- (18) Lin, Y.-L.; Aleksandrov, A.; Simonson, T.; Roux, B. An Overview of Electrostatic Free Energy Computations for Solutions and Proteins. *J. Chem. Theory Comput.* **2014**, *10*, 2690–2709.
- (19) Reif, M. M.; Oostenbrink, C. Net Charge Changes in the Calculation of Relative Ligand-Binding Free Energies via Classical Atomistic Dynamics Simulation. *J. Comput. Chem.* **2014**, *35*, 227–243.
- (20) Öhlnknecht, C.; Lier, B.; Petrov, D.; Fuchs, J.; Oostenbrink, C. Correcting electrostatic artifacts due to net-charge changes in the calculation of ligand binding free energies. *J. Comput. Chem.* **2020**, *41*, 986–999.
- (21) Hockney, R.; Eastwood, J. *Computer Simulation Using Particles*; CRC Press: 1988.
- (22) Darden, T.; York, D.; Pedersen, L. Particle-mesh Ewald: An $N \log(N)$ method for Ewald sums in large systems. *J. Chem. Phys.* **1993**, *98*, 10089–10092.
- (23) Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. A smooth particle mesh Ewald method. *J. Chem. Phys.* **1995**, *103*, 8577–8593.
- (24) Hünenberger, P. H. Lattice-sum methods for computing electrostatic interactions in molecular simulations. *AIP Conf. Proc.* **1999**, *492*, 17–83.
- (25) Barker, J. A.; Watts, R. O. Monte Carlo studies of the dielectric properties of water-like models. *Mol. Phys.* **1973**, *26*, 789–792.
- (26) Lee, F. S.; Warshel, A. A local reaction field method for fast evaluation of long-range electrostatic interactions in molecular simulations. *J. Chem. Phys.* **1992**, *97*, 3100–3107.
- (27) Tironi, I. G.; Sperb, R.; Smith, P. E.; van Gunsteren, W. F. A generalized reaction field method for molecular dynamics simulations. *J. Chem. Phys.* **1995**, *102*, 5451–5459.
- (28) Chen, W.; Deng, Y.; Russell, E.; Wu, Y.; Abel, R.; Wang, L. Accurate Calculation of Relative Binding Free Energies between Ligands with Different Net Charges. *J. Chem. Theory Comput.* **2018**, *14*, 6346–6358.
- (29) Dixit, S. B.; Chipot, C. Can Absolute Free Energies of Association Be Estimated from Molecular Mechanical Simulations? The Biotin-Streptavidin System Revisited. *J. Phys. Chem. A* **2001**, *105*, 9795–9799.
- (30) Donnini, S.; Mark, A. E.; Juffer, A. H.; Villa, A. Incorporating the effect of ionic strength in free energy calculations using explicit ions. *J. Comput. Chem.* **2005**, *26*, 115–122.
- (31) Zhou, R.; Das, P.; Royyuru, A. K. Single Mutation Induced H3N2 Hemagglutinin Antibody Neutralization: A Free Energy Perturbation Study. *J. Phys. Chem. B* **2008**, *112*, 15813–15820.
- (32) Wallace, J. A.; Shen, J. K. Charge-leveling and proper treatment of long-range electrostatics in all-atom molecular dynamics at constant pH. *J. Chem. Phys.* **2012**, *137*, 184105.
- (33) Chen, W.; Wallace, J. A.; Yue, Z.; Shen, J. K. Introducing Titratable Water to AllAtom Molecular Dynamics at Constant pH. *Biophys. J.* **2013**, *105*, L15–L17.
- (34) Chen, Y.; Roux, B. Constant-pH Hybrid Nonequilibrium Molecular Dynamics–Monte Carlo Simulation Method. *J. Chem. Theory Comput.* **2015**, *11*, 3919–3931.
- (35) Gapsys, V.; Michielsens, S.; Peters, J. H.; de Groot, B. L.; Leonov, H. *Calculation of Binding Free Energies*. In: Kukol A. (eds) *Molecular Modeling of Proteins. Methods in Molecular Biology (Methods and Protocols)*; Humana Press: New York, NY, 2015; Vol. 1215.
- (36) Ekimoto, T.; Yamane, T.; Ikeguchi, M. Elimination of Finite-Size Effects on Binding Free Energies via the Warp-Drive Method. *J. Chem. Theory Comput.* **2018**, *14*, 6544–6559.
- (37) Reif, M. M.; Oostenbrink, C. Toward the correction of effective electrostatic forces in explicit-solvent molecular dynamics simulations: restraints on solvent-generated electrostatic potential and solvent polarization. *Theor. Chem. Acc.* **2015**, *134*, 2.
- (38) Deng, N.; Cui, D.; Zhang, B. W.; Xia, J.; Cruz, J.; Levy, R. Comparing alchemical and physical pathway methods for computing the absolute binding free energy of charged ligands. *Phys. Chem. Chem. Phys.* **2018**, *20*, 17081–17092.
- (39) Oostenbrink, C. Efficient free energy calculations on small molecule host-guest systems: A combined linear interaction energy/one-step perturbation approach. *J. Comput. Chem.* **2009**, *30*, 212–221.
- (40) Schmid, N.; Christ, C. D.; Christen, M.; Eichenberger, A. P.; van Gunsteren, W. F. Architecture, implementation and parallelisation of the GROMOS software for biomolecular simulation. *Comput. Phys. Commun.* **2012**, *183*, 890–903.
- (41) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; Hermans, J. *Intermolecular Forces*; Springer: Dordrecht, 1981; Vol. 14; pp. 341–342.
- (42) Hockney, R. W. The Potential Calculation And Some Applications. *Methods Comput. Phys.* **1970**, *9*, 136–211.
- (43) Ryckaert, J.-P.; Ciccoti, G.; Berendsen, H. J. C. Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *J. Comput. Phys.* **1977**, *23*, 327–341.
- (44) Redlack, A.; Grindlay, J. The Electrostatic Potential in a Finite Ionic Crystal. *Can. J. Phys.* **1972**, *50*, 2815–2825.
- (45) Hummer, G.; Pratt, L. R.; García, A. E. Free Energy of Ionic Hydration. *J. Phys. Chem.* **1996**, *100*, 1206–1215.
- (46) Heinz, T. N.; Hünenberger, P. H. Combining the lattice-sum and reaction-field approaches for evaluating long-range electrostatic interactions in molecular simulations. *J. Chem. Phys.* **2005**, *123*, No. 034107.
- (47) Glättli, A.; Daura, X.; van Gunsteren, W. F. Derivation of an improved simple point charge model for liquid water: SPC/A and SPC/L. *J. Chem. Phys.* **2002**, *116*, 9811–9828.
- (48) Bennett, C. H. Efficient estimation of free energy differences from Monte Carlo data. *J. Comput. Phys.* **1976**, *22*, 245–268.
- (49) Chodera, J. D.; Swope, W. C.; Pitera, J. W.; Seok, C.; Dill, K. A. Use of the Weighted Histogram Analysis Method for the Analysis of Simulated and Parallel Tempering Simulations. *J. Chem. Theory Comput.* **2007**, *3*, 26–41.
- (50) 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. *Chem. Phys. Lett.* **1994**, *222*, 529–539.
- (51) Boresch, S.; Tettinger, F.; Leitgeb, M.; Karplus, M. Absolute Binding Free Energies: A Quantitative Approach for Their Calculation. *J. Phys. Chem. B* **2003**, *107*, 9535–9551.
- (52) Perthold, J. W.; Oostenbrink, C. Simulation of Reversible Protein–Protein Binding and Calculation of Binding Free Energies Using Perturbed Distance Restraints. *J. Chem. Theory Comput.* **2017**, *13*, 5697–5708.
- (53) Example input files for simulations and corrections are available at [christophohlknecht / supplementary_material_buckyballs](https://github.com/christophohlknecht/supplementary_material_buckyballs); https://github.com/christophohlknecht/supplementary_material_buckyballs.git. Accessed: 2020-09–20

(54) Hünenberger, P. H.; McCammon, J. A. Ewald artifacts in computer simulations of ionic solvation and ion–ion interaction: A continuum electrostatics study. *J. Chem. Phys.* **1999**, *110*, 1856–1872.

(55) Davis, M. E.; Madura, J. D.; Luty, B. A.; McCammon, J. A. Electrostatics and diffusion of molecules in solution: simulations with the University of Houston Brownian dynamics program. *Comput. Phys. Commun.* **1991**, *62*, 187–197.

(56) Madura, J. D.; Briggs, J. M.; Wade, R. C.; Davis, M. E.; Luty, B. A.; Ilin, A.; Antosiewicz, J.; Gilson, M. K.; Bagheri, B.; Scott, L. R.; McCammon, J. A. Electrostatics and diffusion of molecules in solution: simulations with the University of Houston Brownian Dynamics program. *Comput. Phys. Commun.* **1995**, *91*, 57–95.

(57) Peter, C.; van Gunsteren, W. F.; Hünenberger, P. H. Solving the Poisson equation for solute–solvent systems using fast Fourier transforms. *J. Chem. Phys.* **2002**, *116*, 7434–7451.

(58) Peter, C.; van Gunsteren, W. F.; Hünenberger, P. H. A fast-Fourier transform method to solve continuum-electrostatics problems with truncated electrostatic interactions: Algorithm and application to ionic solvation and ion–ion interaction. *J. Chem. Phys.* **2003**, *119*, 12205–12223.

(59) Oostenbrink, C.; Villa, A.; Mark, A.; van Gunsteren, W. F. 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.* **2004**, *25*, 1656–1676.

(60) Kastenholz, M. A.; Hünenberger, P. H. Computation of methodology-independent ionic solvation free energies from molecular simulations. I. The electrostatic potential in molecular liquids. *J. Chem. Phys.* **2006**, *124*, 124106.

(61) Åqvist, J.; Hansson, T. Analysis of Electrostatic Potential Truncation Schemes in Simulations of Polar Solvents. *J. Phys. Chem. B* **1998**, *102*, 3837–3840.

(62) Vorobjev, Y. N.; Hermans, J. A Critical Analysis of Methods of Calculation of a Potential in Simulated Polar Liquids: Strong Arguments in Favor of “Molecule-Based” Summation and of Vacuum Boundary Conditions in Ewald Summation. *J. Phys. Chem. B* **1999**, *103*, 10234–10242.

(63) Hünenberger, P. H.; Reif, M. M. *Single-Ion Solvation; Theoretical and Computational Chemistry Series; The Royal Society of Chemistry*: 2011; pp. P001–664.

(64) Wang, L.; Wu, Y.; Deng, Y.; Kim, B.; Pierce, L.; Krilov, G.; Lupyan, D.; Robinson, S.; Dahlgren, M. K.; Greenwood, J.; Romero, D. L.; Masse, C.; Knight, J. L.; Steinbrecher, T.; Beuming, T.; Damm, W.; Harder, E.; Sherman, W.; Brewer, M.; Wester, R.; Murcko, M.; Frye, L.; Farid, R.; Lin, T.; Mobley, D. L.; Jorgensen, W. L.; Berne, B. J.; Friesner, R. A.; Abel, R. Accurate and Reliable Prediction of Relative Ligand Binding Potency in Prospective Drug Discovery by Way of a Modern Free-Energy Calculation Protocol and Force Field. *J. Am. Chem. Soc.* **2015**, *137*, 2695–2703.