

Net Charge Changes in the Calculation of Relative Ligand-Binding Free Energies via Classical Atomistic Molecular Dynamics Simulation

Maria M. Reif and Chris Oostenbrink*

The calculation of binding free energies of charged species to a target molecule is a frequently encountered problem in molecular dynamics studies of (bio-)chemical thermodynamics. Many important endogenous receptor-binding molecules, enzyme substrates, or drug molecules have a nonzero net charge. Absolute binding free energies, as well as binding free energies relative to another molecule with a different net charge will be affected by artifacts due to the used effective electrostatic interaction function and associated parameters (e.g., size of the computational box). In the present study, charging contributions to binding free energies of small oligoatomic ions to a series of model host cavities functionalized with different chemical groups are calculated with classical atomistic molecular dynamics simulation. Electrostatic interactions are treated using a lattice-summation scheme or a cutoff-truncation scheme with Barker–Watts reaction-field correction, and the simulations are conducted in boxes of different edge lengths. It is illustrated that the charging free energies

of the guest molecules in water and in the host strongly depend on the applied methodology and that neglect of correction terms for the artifacts introduced by the finite size of the simulated system and the use of an effective electrostatic interaction function considerably impairs the thermodynamic interpretation of guest–host interactions. Application of correction terms for the various artifacts yields consistent results for the charging contribution to binding free energies and is thus a prerequisite for the valid interpretation or prediction of experimental data via molecular dynamics simulation. Analysis and correction of electrostatic artifacts according to the scheme proposed in the present study should therefore be considered an integral part of careful free-energy calculation studies if changes in the net charge are involved. © 2013 The Authors Journal of Computational Chemistry Published by Wiley Periodicals, Inc.

DOI: 10.1002/jcc.23490

Introduction

The calculation of binding free energies is a standard task in the thermodynamic analysis of multicomponent molecular systems involving an association reaction between two system constituents, as, for example, an enzyme and a substrate, a receptor and a drug, or a nanocage and a guest compound. Physics-based approaches to compute binding free energies rely on statistical mechanics, which expresses the free energy as the natural logarithm of the system partition function (multiplied by the negative of the thermal energy, $-k_B T$, where k_B is Boltzmann's constant). The underlying configurational ensembles can be generated by, for example, molecular dynamics (MD) simulation. A wealth of methodological improvements, along with increased computational resources allow (in principle) the accurate calculation of binding free energies, as extensively reviewed in the case of protein–ligand association.^[1–9] However, if conducted without a proper eye on all potential pitfalls, binding free energies may be spuriously affected by limitations of MD simulations, such as, for example, an inadequate force-field description, approximations or/and assumptions in the free-energy calculation methodology, insufficient configurational sampling, or spurious configurational sampling due to the use of an effective electrostatic interaction function. These points are briefly discussed in turn below.

First, besides intrinsic deficiencies of classical force fields such as, for example, the neglect or mean-field treatment of electronic polarizability^[10,11] and the use of effective interac-

tion energy functions^[12–14] with empirical parameters, additional problems arise if the system under consideration involves molecular species for which no force-field parameters are available. For instance, standard (bio-)molecular force fields may not provide parameterizations of certain metal ions, cofactors, or drug molecules. Ideally, the corresponding parameters should be parameterized against experimental data using a strategy consistent with the parameterization of the used force field. In practice, however, they are either inferred based on chemical intuition and comparison with similar compounds or taken from automatized parameterization protocols.^[15,16] In addition, although the solvent representation in most (bio-)molecular force fields is already highly simplistic (rigid three-

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

M. M. Reif C. Oostenbrink

Institute for Molecular Modeling and Simulation, University of Natural Resources and Life Sciences Vienna, Muthgasse 18, 1190, Wien, Austria
Fax: (+43) 1 476548309

E-mail: chris.oostenbrink@boku.ac.at

Contract/grant sponsor: Austrian Science Fund (FWF); Contract/grant number: M1281-N17; Contract/grant sponsor: Vienna Science and Technology Fund (WWTF); Contract/grant number: LS08-QM03; Contract/grant sponsor: European Research Council (ERC); Contract/grant number: 260408

© 2013 The Authors Journal of Computational Chemistry Published by Wiley Periodicals, Inc.

site models^[17]), its structural characteristics may be relinquished for the sake of computational savings, the solvent then being modeled implicitly and the solvent-generated electrostatic potential computed via numerical or empirical (generalized Born) solutions of the Poisson-Boltzmann equation.^[18–20]

Second, because they rely on a thorough characterization of the phase space of the system, simulations involving free-energy calculations are computationally expensive, which is why a number of approximate methods are sometimes applied. For instance, the free energy of charging a neutral particle may be estimated from an electrostatic linear-response approximation^[21,22] or cumulant expansions at the endpoints of thermodynamic integration (TI).^[23–25] Similarly, the free energy of growing the van der Waals envelope of a particle is sometimes approximated using physics-based^[26,27] or empirical^[21] relationships. Furthermore, assumptions in the ansatz of free-energy calculation methods, such as, for example, sufficient overlap of the phase-space distribution functions in different states of relative free-energy calculations,^[28] or electrostatic linear response^[22,29] may limit the scope of their applicability. Lastly, discretization errors in numerical free-energy calculation methods, for example, the window width in potential of mean force calculations^[30,31] or the integration method in TI,^[32,33] limit the precision of the obtained results, although usage of optimal methods for statistical analysis [e.g., Bennett acceptance ratio (BAR)^[34,35] or multi-state BAR^[36,37] approaches] may lead to significant gains in computational efficiency and statistical certainty.

Third, the phase space accessible to the system should be sampled exhaustively and according to the Gibbs measure appropriate for the desired thermodynamic ensemble, for example, canonical Boltzmann weighting in the case of simulations at constant particle number, temperature and volume. However, exhaustive sampling of phase space is complicated by the sheer number of possible configurations, growing exponentially with the system size, and by energy barriers higher than $k_B T$, usually not amenable to transitions in plain MD simulation. Enhanced sampling methods can be used to improve coverage of the relevant phase space. A widely-used technique to address this problem involves the alteration of the potential energy function, for example, through local^[38,39] or nonlocal^[40,41] biasing, or more complex smoothening procedures,^[42,43] along with subsequent reweighting of the sampled configurations to the Gibbs measure corresponding to the unaltered potential energy function.

Finally, even if the phase space accessible to the system is sampled exhaustively and according to the Gibbs measure appropriate for the desired thermodynamic ensemble, the sampled configurations might not be representative of the real (experimental) situation because of an approximate or incorrect calculation of interatomic interactions. This is generally the case for electrostatic interactions which, due their long-range nature, are treated in an effective manner during MD simulations.^[44–49] Ensuing artifacts become strongly apparent in the configurational sampling of systems involving charged particles or in free-energy calculations involving the change of the net charge of the system (charging free energy calculations), and have been reviewed extensively.^[48–56] For instance, if electrostatic interactions are calculated via lattice-

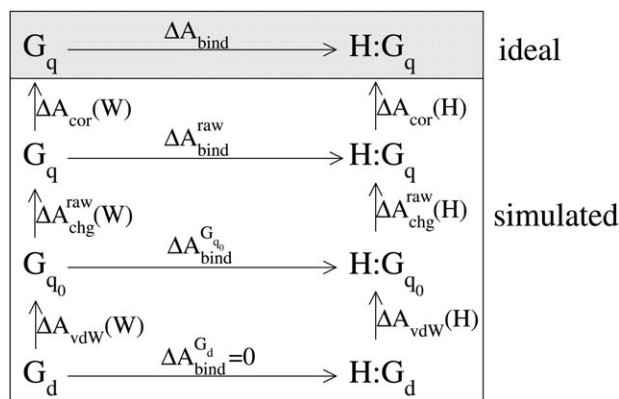


Figure 1. Thermodynamic cycle illustrating the connection between the binding free energy $\Delta A_{\text{bind}}^{\text{G}_d}$ of a noninteracting (dummy) guest molecule G_d , the binding free energy $\Delta A_{\text{bind}}^{\text{G}_{q_0}}$ of a neutral guest molecule G_{q_0} , the raw binding free energy $\Delta A_{\text{bind}}^{\text{raw}}$ of a charged guest molecule G_q (full atomic partial charges), and the corrected binding free energy ΔA_{bind} of the latter guest to a host molecule H . Corresponding complexes formed by the host and the bound guest molecules are denoted $H:G_d$, $H:G_{q_0}$, $H:G_q$, and $H:G_q$, respectively. The free energies of growing the van der Waals envelope of the guest molecule, $\Delta A_{\text{vdW}}(E)$, raw free energies of charging, $\Delta A_{\text{chg}}(E)$, and correction terms, $\Delta A_{\text{cor}}(E)$, of the guest in environment E (either water W or the host molecule H) are associated with the reversible work of creating neutral van der Waals particles (G_d mutated into G_{q_0}), installing partial charges (G_{q_0} mutated into G_q) and applying corrections for approximate-electrostatics, summation, and finite-size artifacts (G_q represented in the simulated situation associated with these artifacts versus in the ideal situation, i.e., a macroscopic nonperiodic system with Coulombic electrostatic interactions). The differences $\Delta \Delta A_{\text{chg}}^{\text{raw}} = \Delta A_{\text{chg}}^{\text{raw}}(H) - \Delta A_{\text{chg}}^{\text{raw}}(W)$ and $\Delta \Delta A_{\text{chg}} = \Delta A_{\text{chg}}(H) - \Delta A_{\text{chg}}(W)$ are the raw and corrected charging contributions to the binding free energy, that is, the raw and corrected binding free energies of the charged guest G_q relative to its neutralized analog G_{q_0} [eqs. (1) and (21)].

summation (LS) over a periodic system in charging free energy calculations, the orientational polarization of the environment of the particle to be charged will be affected by the influence of the periodic copies of this particle, which is an inappropriate contribution if actually a truly nonperiodic system is to be described. The magnitude of the introduced errors may be strongly dependent on the parameters of the system or the interaction function (e.g., the box-edge length), giving rise to so-called methodology-dependent charging free energies.^[54] It has been shown before how charging free energies of monoatomic^[54,57,58] and polyatomic^[59–61] ions in infinitely dilute aqueous solution can be corrected for these errors, such that methodology-independent values are obtained.

The goal of the present article is to address the last point above for model systems representative of a protein-ligand complex in aqueous solution, that is, to present a correction scheme for the charging of polyatomic ions in a low-dielectric cavity functionalized with different chemical groups (section “Simulated guest-host systems”), such that the raw charging free energy $\Delta A_{\text{chg}}^{\text{raw}}(H)$ of a ligand bound to a host molecule can be corrected to a methodology-independent value $\Delta A_{\text{chg}}(H)$ (Fig. 1). Comparison with the corresponding raw or corrected charging free energies in bulk water, $\Delta A_{\text{chg}}^{\text{raw}}(W)$ or $\Delta A_{\text{chg}}(W)$, respectively, yields the raw or corrected binding free energies of the charged ligand to the host molecule relative to a neutralized analog of the ligand, in the following denoted as $\Delta \Delta A_{\text{chg}}^{\text{raw}}$ and $\Delta \Delta A_{\text{chg}}$, respectively (Fig. 1). The

Table 1. Abbreviations used throughout the text for the names of guest molecules and chemical groups functionalizing the host cavity.

	Charge (<i>e</i>)	Abbreviation
Guest		
Methylammonium ion	1.0	MAM
Acetate ion	−1.0	ACE
Host functionalization		
−	0	CAPO
Formamide	0	CHB
Methylammonium	1.0	CPOS
Formate	−1.0	CNEG

possible occurrence of methodology-dependent artifacts (caused by the use of an approximate electrostatic interaction function, an improper summation scheme and simulated systems of finite size) in $\Delta\Delta A_{\text{chg}}^{\text{raw}}$ directly impairs calculations of the (absolute) binding free energy of a charged ligand and of relative binding free energies between ligands of different net charge. The value obtained for $\Delta\Delta A_{\text{chg}}^{\text{raw}}$ is not representative of a macroscopic nonperiodic system with Coulombic electrostatic interactions, and only $\Delta\Delta A_{\text{chg}}$ allows a meaningful comparison to or prediction of experimental data measured in systems of macroscopic extent (Fig. 1). This issue was, however, not duly appreciated in previous work. Examples from the authors' own research include, for example, the calculation of ligand binding free energies^[25,62,63] or redox potentials.^[64]

On the long term, increases in computational power as currently mainly driven by graphics processing unit-based electrostatic interaction calculation^[65–67] and advances in multiscale simulation methodologies targeted to an improved representation of electrostatic interactions^[68,69] may eventually allow for the simulation of macroscopic nonperiodic systems with Coulombic electrostatic interactions, or electrostatic interactions truncated at sufficiently large distances, such that an adequate representation of experimental bulk systems is achieved. Before such techniques have become state of the art, however, a scheme that corrects for methodology-induced artifacts will prove valuable in the calculation of binding free energies of charged ligands to (bio-)macromolecular host compounds.

Methods

Simulated guest-host systems

A previously described^[70] simplified guest-host system was used to assess the size of methodology-induced artifacts in the calculation of (relative) charging free energies (Table 1). Two oppositely-charged guest molecules, methylammonium (MAM) and acetate (ACE), binding to a host C60 molecule (buckyball), or derivatives thereof, were considered. In comparison to a realistic buckyball model, here all C–C bonds were artificially extended to 0.2 nm. For the host molecules, four variants were used: (i) an empty, apolar C60 cavity (CAPO); (ii) a C60 cavity containing a covalently-bound amide group, representing a neutral polar cavity with hydrogen-bonding capability (CHB); (iii) a C60 cavity containing a covalently-bound

methylammonium group, representing a positively-charged cavity (CPOS); (iv) a C60 cavity containing a covalently-bound carboxylate group, representing a negatively-charged cavity (CNEG). Because of the high symmetry and low flexibility in these systems, the observed artifacts can be expected to be solely due to methodological aspects rather than, in addition, insufficient sampling. Figure 2a provides a graphical illustration of an example guest-host complex used in the present study. GROMOS molecular topology^[71] files for the eight guest-host complexes are provided as supporting information.

The raw charging component $\Delta\Delta A_{\text{chg}}^{\text{raw}}$ of binding free energies of charged guests MAM and ACE to host molecules CAPO, CHB, CPOS, and CNEG was calculated with MD simulation according to a thermodynamic cycle (Fig. 1) involving the free energies of guest-charging in water $\Delta A_{\text{chg}}^{\text{raw}}(\text{W})$ and in the host cavity $\Delta A_{\text{chg}}^{\text{raw}}(\text{H})$,

$$\Delta\Delta A_{\text{chg}}^{\text{raw}} = \Delta A_{\text{chg}}^{\text{raw}}(\text{H}) - \Delta A_{\text{chg}}^{\text{raw}}(\text{W}). \quad (1)$$

The MD simulations involved cubic computational boxes containing one guest molecule at multiple charge states q_G varying from 0.0 to the full charges $Q_G = 1.0e$ or $-1.0e$ for MAM and ACE, respectively, in either pure water or the hydrated host cavity. Each system was simulated in four different box sizes of edge lengths L_{ss} , L_s , L_m , and L_l , of about 2.46–2.53, 2.90, 3.21–3.25, or 3.80–3.81 nm, respectively, differing by the number of water molecules. Furthermore, two different methods were used to calculate electrostatic interactions (section “MD simulations”), namely either LS^[72,73] or molecule-based cutoff truncation in combination with a Barker–Watts reaction-field correction^[74] (BM). Simulations with the LS scheme were performed in boxes of edge length L_{ss} , L_s , L_m , and L_l , and are in the following referred to as LS,ss, LS,s, LS,m, and LS,l, respectively. Simulations with the BM scheme were performed in boxes of edge length L_m and L_l , and are in the following referred to as BM,m and BM,l, respectively.

MD simulations

All MD simulations were performed either with a modified version of the GROMOS96 program^[71] or with the GROMOS11 program.^[75] The former was exclusively used for free-energy calculations in simulations using the particle-particle-particle-mesh (P³M) method^[72,73] for the treatment of electrostatic interactions. Water was represented by means of the three-site simple point charge (SPC) model.^[76] Host and guest molecules were described with the GROMOS 53A6 force-field parameter set as in the previous study of Ref. [70]. For CNEG, the appropriate GROMOS improper dihedral type 1 (reference value of 0°) was used for the improper dihedral angle in the formate group rather than an erroneous type 2 (reference value of 35° as in Ref. [70]).

All simulations were carried out under periodic boundary conditions (PBC) based on cubic computational boxes. The equations of motion were integrated using the leap-frog scheme^[77] with a timestep of 2 fs. The solute bond-length distances and the rigidity of the water molecules were enforced by application of the SHAKE algorithm^[78] with a relative

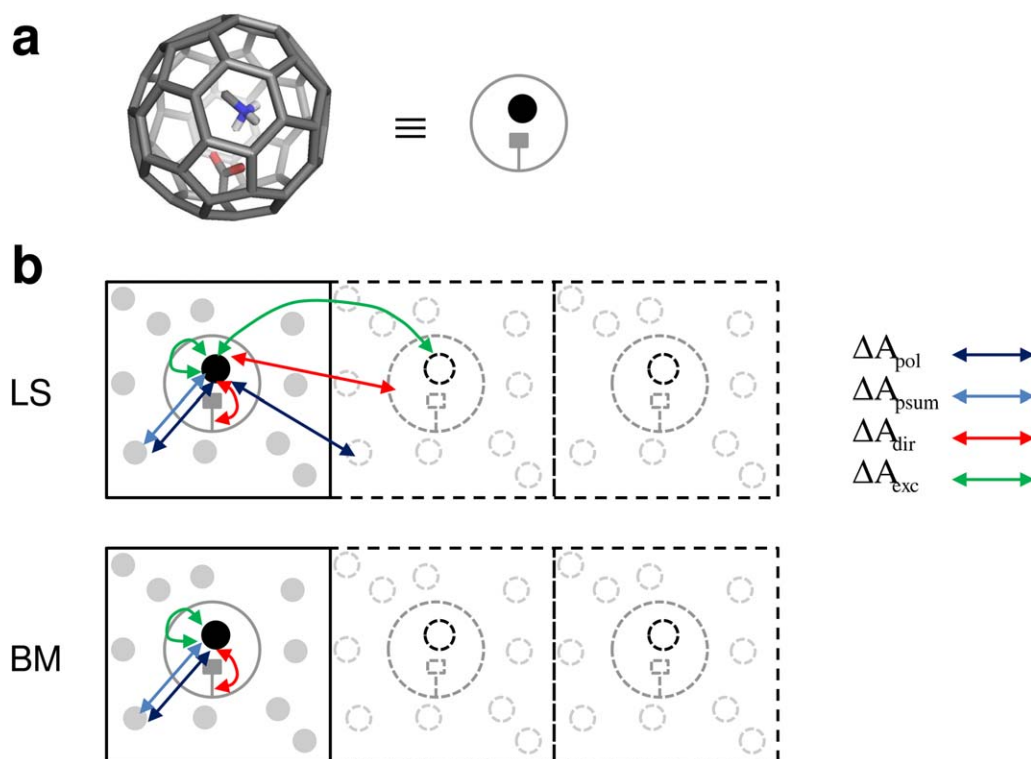


Figure 2. a) Stick representation of the guest-host complex MAM-CNEG. All guest-host complexes used in the present study are similarly composed of a noncovalently-bound oligoatomic ion (guest) and an artificial buckyball molecule possibly functionalized with a covalently-bound chemical group (host), as explained in section “Simulated guest-host systems.” In the following, they are depicted by a simplified schematic where the ion is drawn as a filled black circle and the buckyball is drawn as the gray surrounding structure. b) Graphic illustration of the meaning of correction terms ΔA_{pol} , ΔA_{psum} , ΔA_{dir} , and ΔA_{exc} [eqs. (10), (11), (14), (15), (16), (17), (18)] to be applied to the raw charging free energy $\Delta A_{\text{chg}}^{\text{raw}}$ [eq. (7)], for charging of a guest molecule in a solvated host to get a corrected charging free energy ΔA_{chg} [eq. (20)] for the LS and BM electrostatic schemes. The guest, host, and water molecules are depicted in black, dark gray, and light gray colors, respectively. Periodic copies of the computational system are depicted with dashed lines. Arrows labeled with the above correction terms depict the concerned interactions, that is, guest-solvent interactions (ΔA_{pol} , ΔA_{psum}), guest-host interactions (ΔA_{dir}), and environment-mediated guest-guest interactions (ΔA_{exc}). [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

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 to a heat bath^[79] using a coupling time of 0.1 ps and the volume was kept constant. Electrostatic interactions were either calculated using LS based on the P³M algorithm with tinfoil boundary conditions,^[72,73] or using the BM scheme.^[74] The LS scheme was applied with^[72,80,81] 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 about 0.08–0.12 nm in the different systems. The self-energy term^[23,73,82–84] of the guest molecule was not included in the electrostatic potential at the guest atom sites to be consistent with the previously developed correction scheme for single-ion solvation free energies.^[54,57] In simulations with the LS scheme, Lennard–Jones interactions were truncated at an atom-based cutoff distance $R_C = 1.0$ nm. Both real-space electrostatic and Lennard–Jones interactions were calculated at each timestep based on a pairlist updated at each timestep. The BM scheme was applied with a value $\epsilon_{\text{BW}} = 66.6$ for the relative permittivity of the dielectric continuum surrounding the cutoff sphere, as appropriate^[85] for the SPC water model. Here, too, the self-energy term^[84,86] of the guest molecule was not included in the calculation of the electrostatic potential at the guest atom

sites to be consistent with previous work.^[54,57] In simulations with the BM scheme, electrostatic and Lennard–Jones (van der Waals) interactions were truncated at a charge-group cutoff distance $R_C = 1.4$ nm, and calculated at each timestep based on a pairlist that was updated at each timestep.

The calculation of the charging component to the binding free energy was performed in two TI procedures^[87] considering the free energies of charging the guest molecules in water and in the host cavity (Fig. 1). All production simulations for the free-energy calculations were preceded by an equilibration period of 0.3 ns and lasted 1 ns. The configurations sampled along these simulations were written to file every 0.3 ps for subsequent analysis, whereas the corresponding energetic data was written every timestep.

Structural properties

The configurations sampled in simulations LS,ss, LS,l, and BM,l of the fully charged guests MAM and ACE in hydrated host molecules CAPO, CHB, CPOS, and CNEG were analyzed in terms of the solvent radial distribution function $g(r)$ around the buckyball center of mass, and the solvent radial polarization $P(r)$ around the charged guest. These functions were calculated as

$$g(r) = (4\pi\eta_w\Delta r r^2)^{-1}\langle N(r; \Delta r) \rangle, \quad (2)$$

and

$$P(r) = \mu_w\eta_w\tilde{g}(r)\langle \tilde{N}^{-1}(r; \Delta r)M(r; \Delta r) \rangle, \quad (3)$$

where $\langle \dots \rangle$ denotes ensemble averaging, η_w is the solvent number density, $N(r; \Delta r)$ is the number of water molecules j for which $r - \Delta r/2 < r_j \leq r + \Delta r/2$ (\mathbf{r}_j denoting all possible minimum-image vectors connecting the center of mass of the 60 buckyball carbon atoms to the oxygen atom of any periodic copy of water molecule j), $\tilde{N}(r; \Delta r)$ is the number of water molecules j for which $r - \Delta r/2 < r_j \leq r + \Delta r/2$ (\mathbf{r}_j here denoting all possible minimum-image vectors connecting the MAM nitrogen atom or the ACE carboxylate carbon atom to the oxygen atom of any periodic copy of water molecule j),

$$\tilde{g}(r) = (4\pi\eta_w\Delta r r^2)^{-1}\langle \tilde{N}(r; \Delta r) \rangle, \quad (4)$$

$\Delta r = 0.01$ nm is the bin width, $\mu_w = 0.0473$ e nm is the molecular dipole moment of the SPC water model^[76] and $M(r; \Delta r)$ is defined as

$$M(r; \Delta r) = \sum_{j, r - \Delta r/2 < r_j \leq r + \Delta r/2} r_j^{-1} \mathbf{r}_j \cdot \mathbf{e}_j, \quad (5)$$

\mathbf{e}_j being a unit vector along the dipole moment of molecule j . For systems MAM-CAPO, MAM-CHB, MAM-CPOS, ACE-CAPO, ACE-CHB, and ACE-CNEG, $P(r)$ was compared to a radial continuum-electrostatics analog, here approximated by the Born polarization $P_{\text{Born}}(r)$ around a charge of $Q_G' = +1$ e (MAM-CAPO, MAM-CHB), $+2$ e (MAM-CPOS), -1 e (ACE-CAPO, ACE-CHB), or -2 e (ACE-CNEG) centered at the MAM nitrogen or the ACE carboxylate carbon atom,

$$P_{\text{Born}}(r) = (4\pi)^{-1}r^{-2}(1 - \epsilon_S^{-1})Q_L', \quad (6)$$

where $\epsilon_S = 66.6$ is the relative dielectric permittivity of the SPC water model.^[85] Equation (6) is an approximation because in the considered systems the charge Q_G' is actually spread out over multiple atom sites. In addition, the dielectric permittivity around the highly-charged MAM-CPOS and ACE-CNEG systems may be lower than the bulk value of 66.6.

Free-energy calculations

Raw charging free energy. Raw charging free energies $\Delta A_{\text{chg}}^{\text{raw}}(\text{W})$ and $\Delta A_{\text{chg}}^{\text{raw}}(\text{H})$ [eq. (1)] were calculated with the TI approach^[87] along progressive installation of λ -dependent intermolecular electrostatic interactions $U_{\text{elec}}^{\text{inter}}(\mathbf{x}; \lambda)$,

$$\Delta A_{\text{chg}}^{\text{raw}}(\text{E}) = \int d\lambda \left(\frac{\partial U_{\text{elec}}^{\text{inter}}(\mathbf{x}; \lambda)}{\partial \lambda} \right)_{\lambda, \text{E}}, \quad (7)$$

where the subscript E denotes the environment of the guest (either W or H), $\mathbf{x} = \{\mathbf{x}_G, \mathbf{x}_E\}$ denotes the $3N$ -dimensional coordinate vector of the system containing N_G guest atoms and N_E water and host atoms, λ denotes the scaling parameter of the

TI procedure, and $\langle \dots \rangle_{\lambda, \text{E}}$ denotes ensemble averaging over configurations sampled during a simulation where the guest is in environment E, and guest partial atom charges are scaled by λ . The term $U_{\text{elec}}^{\text{inter}}(\mathbf{x}; \lambda)$ in eq. (7) was calculated based on the sampled configurations as

$$U_{\text{elec}}^{\text{inter}}(\mathbf{x}; \lambda) = (4\pi\epsilon_0)^{-1} \lambda \sum_{i=1}^{N_G} \sum_{j=1}^{N_E} q_i q_j \psi_{ij}(\mathbf{x}), \quad (8)$$

where ϵ_0 is the permittivity of vacuum and q_i is the partial charge of atom i . $\psi_{ij}(\mathbf{x})$ is the effective pairwise electrostatic interaction function which describes the implementation of the particular electrostatics scheme.^[71] The exact details of this function depend on the implementation in a simulation program, and can be found elsewhere.^[71,86] Note that the electrostatic interaction energy $U_{\text{elec}}^{\text{inter}}(\mathbf{x}; \lambda)$ is exempt of intramolecular contributions.

During the simulations of a given charge state, the partial charges of the guest atoms were scaled by λ . For all free-energy estimates, eleven charge states ($\lambda = 0.0, 0.1, \dots, 0.9, 1.0$) were used and, the integral in eq. (7) was evaluated numerically with the trapezoidal rule. Statistical error estimates on ensemble averages pertaining to particular λ -values were obtained from block averaging.^[88] Errors on the free-energy values were calculated by numerical integration (trapezoidal rule) of the individual errors and amounted to about 0.2–1.9 kJ mol⁻¹.

Free-energy correction terms. The raw charging free energies $\Delta A_{\text{chg}}^{\text{raw}}(\text{E})$ [eq. (7)] were used to calculate corresponding methodology-independent values $\Delta A_{\text{chg}}(\text{E})$ as^[54,57,58]

$$\Delta A_{\text{chg}}(\text{E}) = \Delta A_{\text{chg}}^{\text{raw}}(\text{E}) + \Delta A_{\text{cor}}(\text{E}), \quad (9)$$

where $\Delta A_{\text{cor}}(\text{E})$ is a free-energy correction for the various methodology-dependent errors committed during the simulation. These errors have been discussed extensively for the case of monoatomic^[54,56,58] and polyatomic^[61] single-ion hydration. In simulations with the LS and BM schemes, they arise from:

- i. The deviation of the solvent polarization around the charged solute from the polarization in an ideal macroscopic system with Coulombic electrostatic interactions. This is a consequence of the use of a finite (microscopic) system during the simulation, for example, a computational box simulated under PBC, of the use of approximate (non-Coulombic in the limit of infinite system sizes) electrostatic interaction functions, for example, involving cutoff truncation, and of the use of a solvent model with an inaccurate dielectric permittivity. The corresponding correction term is here denoted ΔA_{pol} . Note that in previous work,^[61] this correction term was denoted ΔG_{A+B+D} , or split up into three terms^[54,57,58] ΔG_A , ΔG_B , and ΔG_D .
- ii. The deviation of the solvent-generated electrostatic potential at the atom sites of the charged solute as calculated from the simulated trajectory from the "correct"

electrostatic potential. This is a consequence of the possible application of an inappropriate summation scheme for the contributions of individual solvent atomic charges to this potential, that is, summing over individual point charges ("P-summation"^[89]) versus summing over charges within individual solvent molecules ("M-summation"^[89]), as well as of the possible presence of a constant offset in this potential, due to the presence of an interfacial potential at the surface of the solute along with the constraint of vanishing average potential in the LS scheme. The corresponding correction term is here denoted ΔA_{psum} . Note that in previous work,^[54,57,58,61] this correction term was denoted ΔG_{C_1} .

- iii. The spurious calculation of guest-host interactions with an effective electrostatics scheme rather than with Coulombic electrostatic interactions. The corresponding correction term is denoted ΔA_{dir} . Note that it is only pertinent to systems containing the host moiety, that is, it only occurs in $\Delta A_{\text{cor}}(\text{H})$.
- iv. The presence of electrostatic interactions between excluded atoms in the guest molecule. The corresponding correction term is denoted ΔA_{exc} . Note that it is absent in the case of monoatomic guest compounds and that it was not explicitly listed in Ref. [61], because there these interactions were considered an integral part of the environment-induced electrostatic potential at the solute atoms.

Figure 2b illustrates the interactions addressed by the aforementioned correction terms ΔA_{pol} , ΔA_{psum} , ΔA_{dir} , and ΔA_{exc} for the case of a guest molecule binding noncovalently to a solvated host. In the following, the calculation of these terms is explained.

As in previous work,^[61] ΔA_{pol} can be deduced from the results of two continuum-electrostatics calculations,

$$\Delta A_{\text{pol}} = \Delta A_{\text{chg}}^{\text{CB}} - \Delta A_{\text{chg}}^{\text{LS}} \quad (10)$$

for the LS scheme and

$$\Delta A_{\text{pol}} = \Delta A_{\text{chg}}^{\text{CB}} - \Delta A_{\text{chg}}^{\text{BM}} \quad (11)$$

for the BM scheme, where $\Delta A_{\text{chg}}^{\text{CB}}$ is the charging free energy of the guest molecule in a macroscopic nonperiodic system with Coulombic electrostatic interactions based on the experimental solvent permittivity (CB). $\Delta A_{\text{chg}}^{\text{LS}}$ and $\Delta A_{\text{chg}}^{\text{BM}}$ are the charging free energies of the guest molecule in a periodic system with LS or BM electrostatic interactions based on the model solvent permittivity, respectively. Here, the relative dielectric permittivity is set to 66.6 in the calculation of $\Delta A_{\text{chg}}^{\text{LS}}$ and $\Delta A_{\text{chg}}^{\text{BM}}$ as appropriate for the SPC water model^[85] and to 78.4 in the calculation of $\Delta A_{\text{chg}}^{\text{CB}}$, as appropriate for water,^[56] to account for the inaccurate dielectric permittivity of the SPC water model. As the guest molecules considered in this study are essentially rigid, $\Delta A_{\text{chg}}^{\text{LS}}(W)$ and $\Delta A_{\text{chg}}^{\text{BM}}(W)$ are essentially configuration-independent. Also, the rotational and translational sampling of the guest molecules in the host cavities does not lead to significant fluctuations in ΔA_{pol} (data not shown). Therefore, the calculations of $\Delta A_{\text{chg}}^{\text{CB}}$, $\Delta A_{\text{chg}}^{\text{LS}}$, and $\Delta A_{\text{chg}}^{\text{BM}}$ were only performed based on a sin-

gle structure, taken as the final solute configuration of the simulation at guest charge states $\lambda \in \{0.0, 0.2, 0.4, 0.6, 0.8, 1.0\}$ in the system with box-edge length L_b , as

$$\Delta A_{\text{chg}}^{\text{X}} = \sum_{i=1}^{N_G} q_i \int d\lambda \phi^{\text{inter}}(\mathbf{r}_i; \lambda, \text{X}), \quad (12)$$

where $\text{X} = \text{CB}, \text{LS}, \text{or BM}$, and $\phi^{\text{inter}}(\mathbf{r}_i; \lambda, \text{X})$ is the environment-generated electrostatic potential evaluated at guest atom site i and charge state λ for the given boundary conditions and electrostatics scheme X . For both the LS and BM scheme, $\phi^{\text{inter}}(\mathbf{r}_i; \lambda, \text{CB})$ was evaluated using the FD Poisson equation solver of Refs. [90–92] with the appropriate boundary conditions, and solvent permittivity. The FD solver was also used to evaluate $\phi^{\text{inter}}(\mathbf{r}_i; \lambda, \text{LS})$. Because the FD solver does not offer the option of using the BM scheme, a combination with the fast Fourier transform (FFT) Poisson equation solver of Refs. [93,94] was used to evaluate $\phi^{\text{inter}}(\mathbf{r}_i; \lambda, \text{BM})$ as

$$\begin{aligned} \phi^{\text{inter}}(\mathbf{r}_i; \lambda, \text{BM}) &= \phi^{\text{inter}}(\mathbf{r}_i; \lambda, \text{LS}(\text{FD})) - \phi^{\text{inter}}(\mathbf{r}_i; \lambda, \text{LS}(\text{FFT})) \\ &+ \phi^{\text{inter}}(\mathbf{r}_i; \lambda, \text{BM}(\text{FFT})), \end{aligned} \quad (13)$$

where the first and the last two terms on the right-hand side are calculated based on the FD and FFT Poisson equation solvers, respectively. This is done to enhance cancellation of grid-discretization and boundary-smoothing errors in the two different Poisson equation solvers. In both algorithms, the grid spacing was set to about 0.02 nm and the convergence threshold for the electrostatic free energy was set to 10^{-6} kJ mol⁻¹. A van der Waals envelope was used to define the guest-host system, where the atomic radii were based on distances at the minimum of the Lennard–Jones potential between the different solute atoms and the oxygen atom of a SPC water molecule^[76] using the Lennard–Jones interaction parameters of the GROMOS 53A6 force-field parameter set,^[95] reduced by an estimate^[96] for the radius of a water molecule (0.14 nm). Polar hydrogen atoms were treated differently^[53,55] and assigned an atomic radius of 0.05 nm. Instead of using eq. (12) to evaluate $\Delta A_{\text{chg}}^{\text{X}}$, less computationally intensive but more approximate approaches can be used, which are discussed in Appendix section "Calculation of ΔA_{pol} ".

The term ΔA_{psum} corrects for the atom-based summation scheme implied by the LS and BM schemes in comparison to a proper molecule-based summation scheme.^[89] In the present study, it is calculated as

$$\Delta A_{\text{psum}}^{\text{LS}} = -N_A(6\epsilon_0)^{-1} \gamma_w Q_G N_w L^{-3} \quad (14)$$

for the LS scheme and

$$\Delta A_{\text{psum}}^{\text{BM}} = -N_A(6\epsilon_0)^{-1} \frac{2(\epsilon_{\text{BW}} - 1)}{2\epsilon_{\text{BW}} + 1} \gamma_w Q_G \langle N_w(R_C) \rangle [(4/3)\pi R_C^3]^{-1} \quad (15)$$

for the BM scheme, where N_A is Avogadro's constant, γ_w is the quadrupole-moment trace of the water model relative to its single van der Waals interaction site, N_w is the number of

water molecules, ϵ_{BW} the reaction-field permittivity, R_C the cut-off distance, L^3 the box volume (here, a constant box-edge length L is adopted because all simulations were performed at constant volume), and $\langle N_w(R_C) \rangle$ is the average number of water molecules present within the cutoff sphere around the center of mass of the 60 buckyball carbon atoms for in-host charging, or around the MAM nitrogen or ACE carboxylate carbon atoms in the case of in-water charging. For the SPC water model,^[76,89] $\gamma_w = 0.0082 e \text{ nm}^2$. Equation (15) differs from the equation used in Ref. [61] for the calculation of ΔA_{psum} . The corresponding equation used in Ref. [61], as well a derivation of eq. (15) and a comparison in terms of root-mean-square deviations of corrected charging free energies are reported in Appendix section "Calculation of ΔA_{psum} ".

The additional contribution to ΔA_{psum} related to the quadrupole-moment trace of the guest and host molecules is implicitly accounted for through the presence of these molecules in the continuum-electrostatics calculations [eq. (12)], because the boundary conditions in the FD Poisson solver calculations under PBC enforce zero average electrostatic potential over the computational box, which is equivalent to the situation in the MD simulations. On the contrary, in the FD Poisson solver calculations under nonperiodic boundary conditions (NPBC), there is no such shifting of the average electrostatic potential. That means that the correction ΔA_{pol} , which involves a difference between calculations under NPBC and PBC [eqs. (10) and (11)] also corrects for the spurious vanishing average (over the computational box) electrostatic potential contribution due to the guest and host quadrupole moments in the MD simulations. Note that this implicit inclusion of the solute-associated ΔA_{psum} correction in ΔA_{pol} was not recognized in previous work.^[61]

Similar to Ref. [61], the term ΔA_{C_2} was neglected in the present study because this term is proportional to the ratio of the guest volume to the box volume, that is, its magnitude is very small for the systems considered here.

The term ΔA_{dir} , to be applied only to raw charging free energies of the guest in the host, was calculated as

$$\Delta A_{\text{dir}}^{\text{LS}} = \Delta A_{\text{chg,G(H)}}^{\text{CB}} - \Delta A_{\text{chg,G(H)}}^{\text{raw,LS}} \quad (16)$$

and

$$\Delta A_{\text{dir}}^{\text{BM}} = \Delta A_{\text{chg,G(H)}}^{\text{CB}} - \Delta A_{\text{chg,G(H)}}^{\text{raw,BM}}, \quad (17)$$

where $\Delta A_{\text{chg,G(H)}}^{\text{CB}}$, $\Delta A_{\text{chg,G(H)}}^{\text{raw,LS}}$, and $\Delta A_{\text{chg,G(H)}}^{\text{raw,BM}}$ are charging free energies of the guest due to the host calculated with Coulombic electrostatic interactions in a nonperiodic system and with effective electrostatic interactions (LS or BM) under PBC, respectively. The guest-host complex configurations sampled in the explicit-water MD simulations at all guest charge states q_G (section "Raw charging free energy") were used to extract $\Delta A_{\text{chg,G(H)}}^{\text{CB}}$, $\Delta A_{\text{chg,G(H)}}^{\text{raw,LS}}$, and $\Delta A_{\text{chg,G(H)}}^{\text{raw,BM}}$ in postanalyses under NPBC (guest-host complex in vacuum described with Coulombic electrostatic interactions) and under PBC (guest-host complex described with LS or BM electrostatic interactions), respectively (section "Solute and solvent contributions to the free energy of charging").

The term ΔA_{exc} was calculated as

$$\Delta A_{\text{exc}}(E) = -(4\pi\epsilon_0)^{-1} \sum_{i=1}^{N_G} \sum_{j>i}^{N_G} \int d\lambda \langle q_i q_j [\tilde{\psi}_{ij}(\mathbf{x}_G) - r_{ij}^{-1}(\mathbf{x}_G)] \rangle_{\lambda,E}, \quad (18)$$

where $\tilde{\psi}$ is a modified LS or BM electrostatic interaction function exempt of self term. The integrand of eq. (18) corresponds to minus the first term in eq. (20) of Ref. [61] and corrects for the presence of electrostatic interactions between excluded atoms (first and second covalently-bound neighbors) in the Hamiltonian used in the present simulations. Electrostatic interactions between excluded atoms are equal to the normal interaction function applied to nonexcluded atoms, but reduced by the Coulombic contribution. As a result, excluded atoms may be viewed to interact via a term that depends on the representation of the surroundings of the solute, that is, periodic copies of the computational box in the case of the LS scheme and a continuous medium of homogeneous relative dielectric permittivity in the case of the BM scheme. Therefore, in the present study, interactions between excluded atoms are regarded as contributing in a methodologically-dependent way to the charging free energy. Application of ΔA_{exc} removes the corresponding contribution.

Given the above correction terms, the charging free energy $\Delta A_{\text{chg}}(E)$ is calculated according to eq. (9) as the sum of the raw charging free energy $\Delta A_{\text{chg}}^{\text{raw}}(E)$, and the correction terms ΔA_{pol} [eqs. (10) and (11)], ΔA_{psum} [eqs. (14) and (15)], ΔA_{dir} [eqs. (16) and (17)] and $\Delta A_{\text{exc}}(E)$ [eq. (18)] as

$$\Delta A_{\text{chg}}(W) = \Delta A_{\text{chg}}^{\text{raw}}(W) + \Delta A_{\text{pol}} + \Delta A_{\text{psum}} + \Delta A_{\text{exc}}(W), \quad (19)$$

in the case of charging in water, and as

$$\Delta A_{\text{chg}}(H) = \Delta A_{\text{chg}}^{\text{raw}}(H) + \Delta A_{\text{pol}} + \Delta A_{\text{psum}} + \Delta A_{\text{dir}} + \Delta A_{\text{exc}}(H), \quad (20)$$

in the case of charging in the host. These charging free energies were calculated for guests MAM and ACE in water and host molecules CAPO, CHB, CPOS, and CNEG and yield the corrected charging component $\Delta\Delta A_{\text{chg}}$ of binding free energies of the guest to the host,

$$\Delta\Delta A_{\text{chg}} = \Delta A_{\text{chg}}(H) - \Delta A_{\text{chg}}(W). \quad (21)$$

Solute and solvent contributions to the free energy of charging

The trajectories of guest-host complexes were reanalyzed to obtain the raw free energy of charging the guest molecule due to the host and periodic host copies $\Delta A_{\text{chg,G(H)}}^{\text{raw}}$ as

$$\Delta A_{\text{chg,G(H)}}^{\text{raw}} = \int d\lambda [U_{\text{elec}}^{\text{tot}}(\lambda, \text{GH}) - U_{\text{elec}}^{\text{tot}}(\lambda, \text{G}) - U_{\text{elec}}^{\text{tot}}(\lambda, \text{H})], \quad (22)$$

where $U_{\text{elec}}^{\text{tot}}(\lambda, \text{GH})$, $U_{\text{elec}}^{\text{tot}}(\lambda, \text{G})$, and $U_{\text{elec}}^{\text{tot}}(\lambda, \text{H})$ are the total electrostatic energies sampled in trajectories pertaining to

guest charge states defined by λ using modified interaction parameters with full guest and host charges, full guest and zeroed host charges and zeroed guest and full host charges, respectively. The corresponding raw free energies of charging the guest molecule due to the solvent and periodic solvent copies $\Delta A_{\text{chg,G(W)}}^{\text{raw}}$ are calculated as

$$\Delta A_{\text{chg,G(W)}}^{\text{raw}} = \Delta A_{\text{chg}}^{\text{raw}}(\text{H}) - \Delta A_{\text{chg,G(H)}}^{\text{raw}}, \quad (23)$$

where $\Delta A_{\text{chg}}^{\text{raw}}(\text{H})$ is given by eq. (7). Corrected values $\Delta A_{\text{chg,G(H)}}$ and $\Delta A_{\text{chg,G(W)}}$ are calculated as the sum of the raw charging free energies and the correction term ΔA_{dir} [eqs. (16) and (17)],

$$\Delta A_{\text{chg,G(H)}} = \Delta A_{\text{chg,G(H)}}^{\text{raw}} + \Delta A_{\text{dir}}, \quad (24)$$

in the case of $\Delta A_{\text{chg,G(H)}}^{\text{raw}}$, and as the sum of the raw charging free energies and the correction terms ΔA_{pol} , ΔA_{psum} , and ΔA_{exc} [eqs. (10), (11), (14), (15), and (18)],

$$\Delta A_{\text{chg,G(W)}} = \Delta A_{\text{chg,G(W)}}^{\text{raw}} + \Delta A_{\text{pol}} + \Delta A_{\text{psum}} + \Delta A_{\text{exc}}, \quad (25)$$

in the case of $\Delta A_{\text{chg,G(W)}}^{\text{raw}}$.

Results

Application of correction terms for spurious solvent polarization and wrong dielectric permittivity of the solvent model (ΔA_{pol}) [eqs. (10) and (11)], improper electrostatic potential summation (ΔA_{psum}) [eqs. (14) and (15)], effective guest-host direct electrostatic interactions (ΔA_{dir}) [eqs. (16) and (17)] and the presence of electrostatic interactions between excluded solute atoms in the Hamiltonian (ΔA_{exc}) [eq. (18)] to raw charging free energies $\Delta A_{\text{chg}}^{\text{raw}}$ [eq. (7)] yields corrected values ΔA_{chg} [eqs. (19) and (20)] reported in Table 2 for charging of guests MAM and ACE in water or in complex with the hydrated host molecules CAPO, CHB, CPOS, or CNEG (Table 1). This is illustrated for four different sizes of the computational box (L_{ss} , L_s , L_m and L_l) and the two different electrostatics schemes (LS and BM) considered (section "MD simulations"). While the root-mean-square deviations for $\Delta A_{\text{chg}}^{\text{raw}}$ are 6.6, 6.8, 6.7, 8.1, and 9.0 kJ mol⁻¹ for charging of MAM in CAPO, CHB, CPOS, CNEG, and water, respectively, and 12.4, 12.3, 11.8, 13.9, and 10.2 kJ mol⁻¹ for charging of ACE in CAPO, CHB, CPOS, CNEG, and water, respectively, they are reduced to 0.8, 1.0, 1.5, 2.1, and 1.4 kJ mol⁻¹ for MAM and 0.5, 0.6, 1.6, 2.5, and 0.6 kJ mol⁻¹ for ACE in corresponding corrected free energies ΔA_{chg} . Complexes MAM-CNEG and ACE-CNEG exhibit the largest rmsd values in corrected charging free energies (2.1 and 2.5 kJ mol⁻¹). Notably, for these complexes, the ΔA_{chg} value from simulation LS_{ss} has higher magnitudes by up to 6.0 kJ mol⁻¹ (MAM-CNEG) or up to 8.0 kJ mol⁻¹ (ACE-CNEG) compared to simulations LS_s, LS_m, LS_l, BM_m, and BM_l. These deviations might be due to the inability of the continuum-electrostatics representation to capture short-range artifacts in solvent structure.

For all guest-host complexes, in simulations LS_{ss}, $g(r)$ shows spurious density fluctuations (overestimated height of the first and second peaks; Fig. 3) and $P(r)$ shows marked underpolarization

in comparison to the Born polarization $P_{\text{Born}}(r)$ (due to the large influence of periodic solute copies, a consequence of the small edge length of the computational box; Fig. 4). As discussed in Appendix section "Calculation of ΔA_{psum} ", the former artifact might cause the ΔA_{psum} correction [eq. (14)] to be inadequate. The long-range regime of the latter artifact is corrected by the ΔA_{pol} correction [eqs. (10) and (11)]. However, short-range artifacts in the solvent polarization, which affect solvation shell structure in the vicinity of the guest-host complex, cannot be represented by a continuum-electrostatics description of the solvent and are thus not captured by the ΔA_{pol} correction [eqs. (10) and (11)]. Such artifacts appear especially pronounced in simulation LS_{ss} of complexes MAM-CNEG and ACE-CNEG. This is evidenced by the relatively large deviation of the short-range $P(r)$ in simulation LS_{ss} from the corresponding polarization in simulations LS_l and BM_l in the fully-charged complex ACE-CNEG (Fig. 4) and in complex MAM-CNEG containing the uncharged guest molecule (data not shown). Note furthermore that in simulations BM_m and BM_l, $P(r)$ shows marked cutoff artifacts at the cutoff distance of 1.4 nm (a dip in the case of MAM-containing complexes and a crest in the case of ACE-containing complexes) and just before the cutoff distance (overpolarization in the case of MAM-containing complexes). These peculiarities, arising from molecule-based cutoff truncation in an explicit-solvent system, are also not captured by the continuum-electrostatics analog of $P(r)$ for the BM scheme^[94,97] and can thus not be remedied by ΔA_{pol} .

For the systems considered in the present study, the magnitude of correction terms ΔA_{pol} , ΔA_{psum} , and ΔA_{dir} (CPOS and CNEG only) is very large (on the order of 50–200 kJ mol⁻¹). For hydration in pure water, ΔA_{pol} is always negative (independent of the sign of the guest charge) to account for the underhydration of the guest molecule caused by the presence of neighboring periodic copies (LS scheme), or the omission of guest-solvent interactions outside the cutoff sphere (BM scheme). In contrast, ΔA_{pol} is positive for MAM-charging in CNEG and ACE-charging in CPOS because in these complexes the initial state of the TI procedure contains a charged guest-host complex, whereas the final state contains a neutral complex, that is, the electrostatic potential sampled at the guest atom sites is spurious in the initial rather than in the final state. As is the case for hydration in pure water, ΔA_{psum} is negative for the charging of cations (MAM) and positive for the charging of anions (ACE). Because of the absence of electrostatic guest-host interactions, ΔA_{dir} is zero for the charging in CAPO. As the LS and BM electrostatic interaction functions used in the present study are reduced in comparison to the Coulombic component (presence of self- and reaction-field terms^[71,86]), ΔA_{dir} is negative for charging of MAM in CHB and the oppositely-charged CNEG and of ACE in CHB and the oppositely-charged CPOS. Likewise, it is positive for charging of the guests in the like-charged hosts, that is, MAM in CPOS and ACE in CNEG. In comparison to the other correction terms, ΔA_{exc} is of rather small magnitude (0.1–0.4 kJ mol⁻¹), because it is a short-range electrostatic interaction reduced by the short-range singularity associated with the Coulombic

Table 2. Charging free energies ΔA_{chg} of the guest molecules MAM and ACE in hydrated host molecules CAPO, CHB, CPOS, and CNEG, as well as in water, computed in cubic computational boxes of edge length L containing N_w water molecules using LS or BM electrostatic interactions (sections "Simulated guest-host systems" and "MD simulations"). Values obtained with the LS scheme in boxes of edge lengths L_{ss} , L_s , L_m , and L_l are labeled LS,ss, LS,s, LS,m, and LS,l, respectively, and values obtained with the BM scheme in boxes of edge lengths L_m and L_l are labeled BM,m and BM,l, respectively. The charging free energy ΔA_{chg} [eqs. (19) and (20)] is calculated as a sum of the raw charging free energy $\Delta A_{\text{chg}}^{\text{raw}}$ [eq. (7)], and the correction terms ΔA_{pol} , ΔA_{psum} , ΔA_{dir} , and ΔA_{exc} [eqs. (10), (11), (14), (15), (16), (17), (18)]. Error estimates on the raw charging free energies refer to the statistical uncertainty obtained from block averaging.^[88]

Guest	Host/water	Scheme	N_w	L (nm)	$\Delta A_{\text{chg}}^{\text{raw}}$ (kJ mol ⁻¹)	ΔA_{pol} (kJ mol ⁻¹)	ΔA_{psum} (kJ mol ⁻¹)	ΔA_{dir} (kJ mol ⁻¹)	ΔA_{exc} (kJ mol ⁻¹)	ΔA_{chg} (kJ mol ⁻¹)	
MAM	CAPO	LS,ss	471	2.46	-4.4 ± 0.2	-71.2	-75.4	0.0	-0.2	-151.2	
		LS,s	780	2.90	-12.4 ± 0.2	-62.4	-76.3	0.0	-0.1	-151.2	
		LS,m	1095	3.24	-17.6 ± 0.2	-56.8	-76.9	0.0	-0.1	-151.4	
		LS,l	1792	3.81	-24.7 ± 0.2	-49.2	-77.1	0.0	-0.1	-151.1	
		BM,m	1095	3.24	-8.5 ± 0.2	-72.3	-68.9	0.0	-0.3	-150.0	
		BM,l	1792	3.81	-9.7 ± 0.2	-71.3	-67.8	0.0	-0.3	-149.1	
	CHB	LS,ss	469	2.46	-38.7 ± 0.8	-70.5	-75.4	-0.7	-0.2	-185.5	
		LS,s	780	2.90	-47.3 ± 0.9	-61.8	-76.3	-0.4	-0.1	-185.9	
		LS,m	1095	3.24	-53.2 ± 0.8	-56.3	-77.1	-0.3	-0.1	-187.0	
		LS,l	1781	3.81	-59.7 ± 0.9	-48.9	-77.3	-0.2	-0.1	-186.2	
		BM,m	1095	3.24	-44.1 ± 0.3	-71.2	-69.0	-0.9	-0.3	-185.5	
		BM,l	1781	3.80	-44.3 ± 0.3	-70.2	-68.0	-0.9	-0.3	-183.7	
	CPOS	LS,ss	473	2.46	115.7 ± 0.5	-209.7	-75.4	155.8	-0.2	-13.8	
		LS,s	780	2.90	111.2 ± 0.5	-184.6	-76.3	133.3	-0.1	-16.5	
		LS,m	1090	3.23	107.6 ± 0.6	-168.6	-76.9	120.0	-0.1	-18.0	
		LS,l	1772	3.80	102.8 ± 0.6	-146.4	-77.1	102.6	-0.1	-18.2	
		BM,m	1090	3.23	122.2 ± 0.5	-212.8	-69.1	142.4	-0.3	-17.6	
		BM,l	1772	3.80	119.7 ± 0.6	-209.9	-68.2	142.4	-0.3	-16.3	
	CNEG	LS,ss	473	2.46	-234.2 ± 1.8	69.9	-75.4	-157.9	-0.2	-397.8	
		LS,s	780	2.90	-245.4 ± 1.8	61.5	-76.3	-134.6	-0.1	-394.9	
		LS,m	1095	3.25	-252.6 ± 1.6	56.0	-76.5	-120.5	-0.1	-393.7	
		LS,l	1785	3.81	-261.2 ± 1.9	48.6	-77.2	-102.9	-0.1	-392.8	
		BM,m	1095	3.25	-248.9 ± 0.4	70.8	-68.3	-145.1	-0.3	-391.8	
		BM,l	1785	3.81	-248.7 ± 0.4	69.8	-67.7	-145.1	-0.3	-392.0	
	Water	LS,ss	533	2.54	-172.6 ± 0.5	-75.8	-77.8	0.0	-0.2	-326.4	
		LS,s	800	2.90	-181.8 ± 0.5	-66.7	-78.3	0.0	-0.1	-326.9	
		LS,m	1091	3.21	-188.1 ± 0.5	-60.4	-78.4	0.0	-0.1	-327.0	
		LS,l	1827	3.80	-197.8 ± 0.5	-51.3	-79.2	0.0	-0.1	-328.4	
		BM,m	1091	3.21	-173.7 ± 0.5	-77.4	-76.6	0.0	-0.3	-328.0	
		BM,l	1827	3.80	-175.5 ± 0.5	-75.7	-79.2	0.0	-0.3	-330.7	
	ACE	CAPO	LS,ss	468	2.46	-74.5 ± 0.2	-70.8	75.4	0.0	-0.3	-70.2
			LS,s	780	2.90	-84.5 ± 0.2	-62.1	76.3	0.0	-0.2	-70.5
			LS,m	1095	3.24	-90.8 ± 0.3	-56.5	76.8	0.0	-0.1	-70.6
			LS,l	1792	3.81	-99.1 ± 0.3	-49.0	77.1	0.0	-0.1	-71.1
			BM,m	1095	3.24	-65.9 ± 0.3	-71.7	68.4	0.0	-0.4	-69.6
			BM,l	1792	3.81	-66.2 ± 0.3	-70.7	67.4	0.0	-0.4	-69.9
		CHB	LS,ss	471	2.46	-117.8 ± 0.5	-70.2	75.4	-0.2	-0.3	-113.1
			LS,s	780	2.90	-128.0 ± 0.6	-61.7	76.3	-0.1	-0.2	-113.7
			LS,m	1090	3.24	-134.1 ± 0.6	-56.3	76.9	-0.1	-0.1	-113.7
			LS,l	1784	3.81	-142.7 ± 0.6	-48.8	77.2	-0.1	-0.1	-114.5
			BM,m	1090	3.24	-109.5 ± 0.4	-71.0	68.5	-0.3	-0.3	-112.6
			BM,l	1784	3.81	-110.0 ± 0.4	-70.0	67.6	-0.3	-0.3	-113.0
CPOS		LS,ss	474	2.47	-288.7 ± 0.3	70.0	75.5	-157.6	-0.3	-301.1	
		LS,s	780	2.90	-301.9 ± 0.3	61.6	76.3	-134.5	-0.2	-298.7	
		LS,m	1095	3.24	-310.1 ± 0.3	56.1	76.5	-120.5	-0.1	-298.1	
		LS,l	1781	3.80	-319.8 ± 0.3	48.8	77.3	-103.0	-0.1	-296.8	
		BM,m	1095	3.24	-290.5 ± 0.3	70.8	68.3	-144.8	-0.3	-296.5	
		BM,l	1781	3.80	-289.3 ± 0.3	69.8	67.9	-144.9	-0.3	-296.8	
CNEG		LS,ss	479	2.47	37.5 ± 0.4	-208.9	75.5	155.8	-0.3	59.5	
		LS,s	780	2.90	30.2 ± 0.4	-184.6	76.3	133.7	-0.2	55.4	
		LS,m	1095	3.24	25.0 ± 0.4	-168.2	76.7	120.0	-0.1	53.4	
		LS,l	1788	3.81	18.0 ± 0.4	-146.1	77.2	102.5	-0.1	51.5	
		BM,m	1095	3.24	55.4 ± 0.4	-212.9	68.4	143.2	-0.3	53.8	
		BM,l	1788	3.81	53.3 ± 0.4	-210.1	67.7	143.2	-0.3	53.8	
Water		LS,ss	526	2.53	-300.8 ± 0.6	-76.0	77.6	0.0	-0.3	-299.5	
		LS,s	800	2.90	-310.6 ± 0.7	-66.6	78.3	0.0	-0.2	-299.1	
		LS,m	1132	3.25	-318.6 ± 0.7	-59.7	78.9	0.0	-0.1	-299.5	
		LS,l	1826	3.81	-327.0 ± 0.7	-51.2	79.1	0.0	-0.1	-299.2	
		BM,m	1132	3.25	-299.8 ± 0.6	-77.2	77.0	0.0	-0.3	-300.3	
		BM,l	1826	3.81	-301.5 ± 0.6	-75.6	79.0	0.0	-0.3	-298.4	

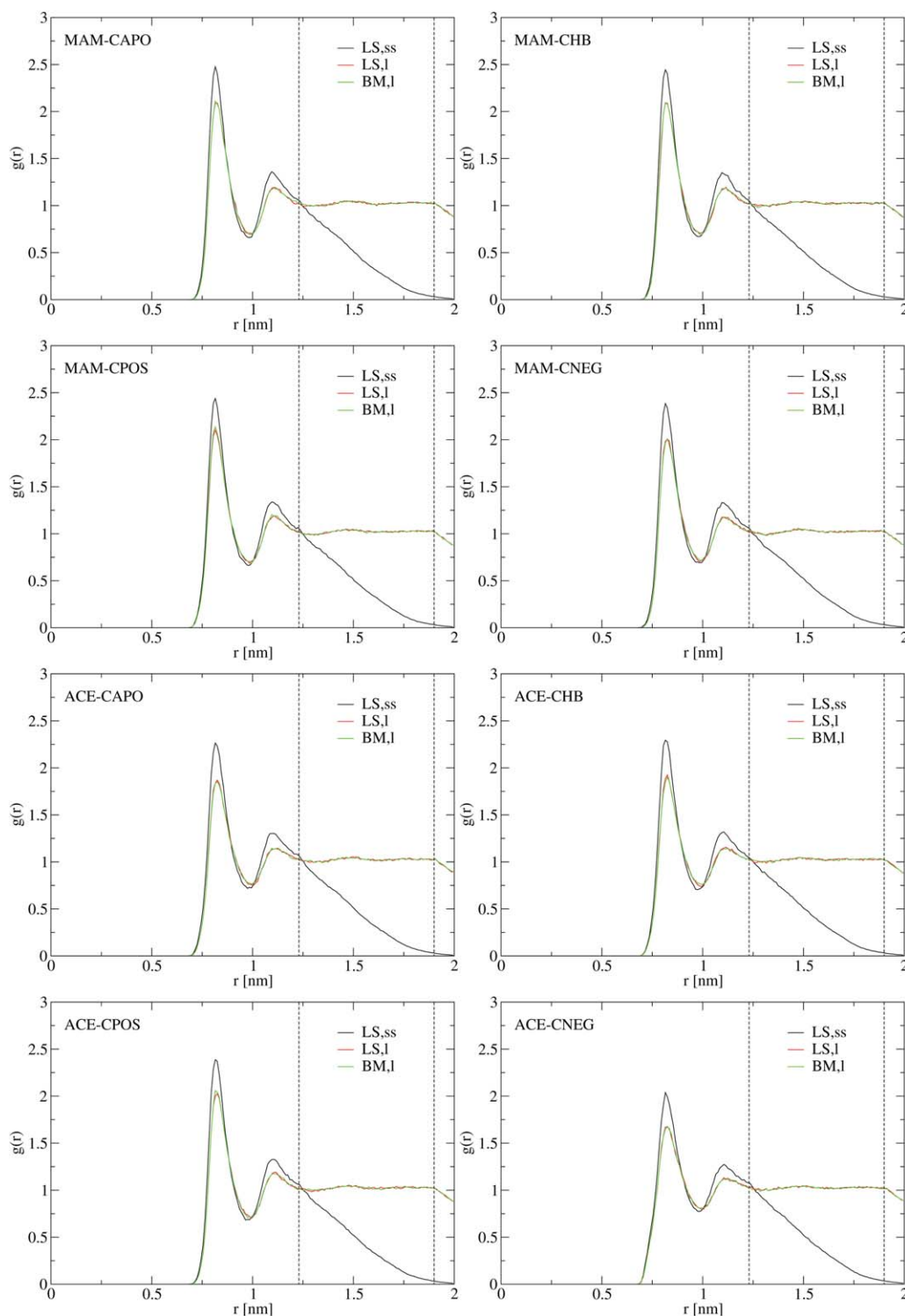


Figure 3. Radial distribution $g(r)$ [eq. (2)] of water oxygen atoms around the center of mass of the 60 buckyball carbon atoms, evaluated from simulations LS,ss, LS,l, and BM,l for systems containing guest molecules MAM or ACE in hydrated host molecules CAPO, CHB, CPOS, or CNEG. The vertical dashed lines indicate $L_{ss}/2$ and $L_l/2$, that is, the threshold beyond which $g(r)$ decays due to box-corner artifacts.

component. For the LS scheme, the magnitude of ΔA_{exc} decreases with increasing box-edge length due to decreasing periodicity artifacts. For the BM scheme, ΔA_{exc} is independent of box-edge length.

Table 3 reports raw free energies of charging the guest molecule due to the host and periodic host copies (LS scheme

only), $\Delta A_{\text{chg,G(H)}}^{\text{raw}}$ [eq. (22)], or due to the solvent and periodic solvent copies (LS scheme only), $\Delta A_{\text{chg,G(W)}}^{\text{raw}}$ [eq. (23)], and corresponding corrected values $\Delta A_{\text{chg,G(H)}}$ and $\Delta A_{\text{chg,G(W)}}$ [eqs. (24) and (25)]. $\Delta A_{\text{chg,G(H)}}$ differs from $\Delta A_{\text{chg,G(H)}}^{\text{raw}}$ in that it is exempt of interaction of the guest with periodic host copies (LS scheme) or of reaction-field terms (BM scheme) and

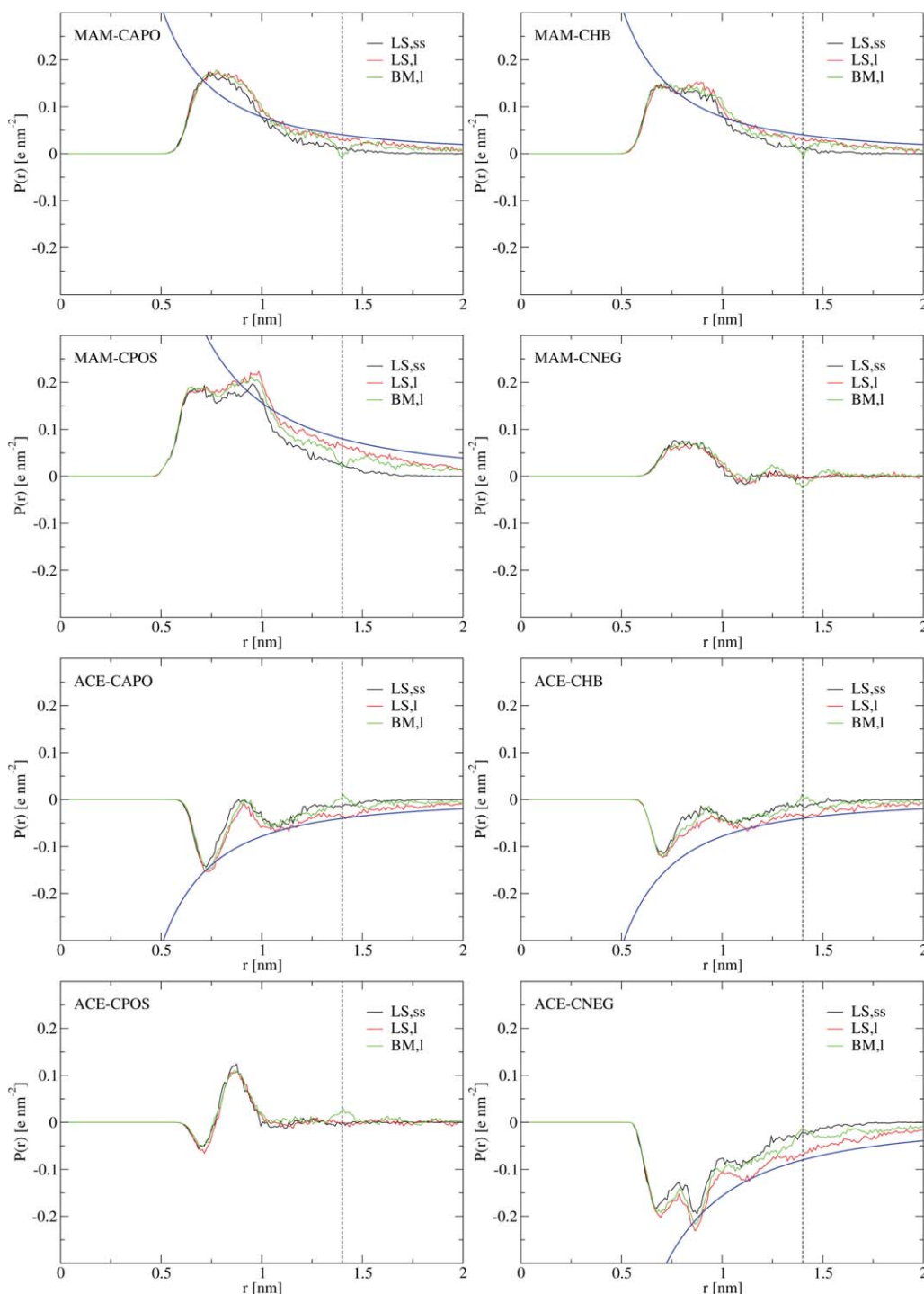


Figure 4. Radial polarization $P(r)$ [eq. (3)] of water molecules around the MAM nitrogen atom or the ACE carboxylate carbon atom in hydrated host molecules CAPO, CHB, CPOS, or CNEG, evaluated from simulations LS,ss, LS,l, and BM,l. The blue line depicted for systems MAM-CAPO, MAM-CHB, MAM-CPOS, ACE-CAPO, ACE-CHB, and ACE-CNEG is the Born polarization $P_{\text{Born}}(r)$ [eq. (6)] according to a system where the total solute (guest and host) charge is centered at the MAM nitrogen or the ACE carboxylate carbon atom. The vertical dashed line indicates the cutoff distance $R_c = 1.4$ nm used in simulations BM,l.

corrected to have Coulombic electrostatic interactions between the guest and the host within the central computational box (ΔA_{dir} ; Table 2). $\Delta A_{\text{chg,G(W)}}$ differs from $\Delta A_{\text{chg,G(W)}}^{\text{raw}}$ in that it is corrected for all solvent-associated artifacts, that is, spurious solvent polarization and wrong dielectric permittivity of the solvent model, improper electrostatic potential summation and the presence of electrostatic interactions between

excluded atoms (ΔA_{pol} , ΔA_{psum} , ΔA_{exc} ; Table 2). It can be seen that application of the correction terms may cause considerable shifts in the ratio of $\Delta A_{\text{chg,G(W)}}$ and $\Delta A_{\text{chg,G(H)}}$. In particular, the relative dominance of the components reverses in the case of MAM-CHB and MAM-CPOS, that is, while interactions with the host dominate $\Delta A_{\text{chg,G}}^{\text{raw}}$, those with the solvent dominate ΔA_{chg} (Table 3). Moreover, in the case of ACE-CPOS, the

Table 3. Charging free-energy contributions due to the solvent, $\Delta A_{\text{chg,G(W)}}$ [eq. (25)], and due to the host molecule, $\Delta A_{\text{chg,G(H)}}$ [eq. (24)], of the guest molecules MAM and ACE in hydrated host molecules CAPO, CHB, CPOS, and CNEG (section "MD simulations"). Values obtained with the LS scheme in boxes of edge lengths L_{ss} , L_s , L_m , and L_l are labeled LS,ss, LS,s, LS,m, and LS,l, respectively, and values obtained with the BM scheme in boxes of edge lengths L_m and L_l are labeled BM,m and BM,l, respectively (section "Simulated guest-host systems"). The charging free energies $\Delta A_{\text{chg,G(W)}}$ and $\Delta A_{\text{chg,G(H)}}$ are calculated as the sum of the raw charging free energy $\Delta A_{\text{chg,G(W)}}$ [eq. (23)] and the correction terms ΔA_{pol} , ΔA_{psum} , and ΔA_{exc} [eqs. (10), (11), (14), (15), and (18)] and as the sum of the raw charging free energy $\Delta A_{\text{chg,G(H)}}$ [eq. (22)] and the correction term ΔA_{dir} [eqs. (16) and (17)], respectively (section "Solute and solvent contributions to the free energy of charging").

Guest	Host	Scheme	$\Delta A_{\text{chg,G(W)}}^{\text{raw}}$ (kJ mol ⁻¹)	$\Delta A_{\text{chg,G(H)}}^{\text{raw}}$ (kJ mol ⁻¹)	$\Delta A_{\text{chg,G(W)}}$ (kJ mol ⁻¹)	$\Delta A_{\text{chg,G(H)}}$ (kJ mol ⁻¹)
MAM	CAPO	LS,ss	-4.4	0.0	-151.2	0.0
		LS,s	-12.4	0.0	-151.2	0.0
		LS,m	-17.6	0.0	-151.4	0.0
		LS,l	-24.7	0.0	-151.1	0.0
		BM,m	-8.5	0.0	-150.0	0.0
		BM,l	-9.7	0.0	-149.1	0.0
	CHB	LS,ss	-1.8	-36.9	-147.9	-37.6
		LS,s	-9.9	-37.4	-148.1	-37.8
		LS,m	-15.0	-38.2	-148.5	-38.5
		LS,l	-21.9	-37.8	-148.2	-38.0
		BM,m	-6.4	-37.7	-146.9	-38.6
		BM,l	-7.4	-36.9	-145.9	-37.8
	CPOS	LS,ss	-48.5	164.2	-333.8	320.0
		LS,s	-75.1	186.3	-336.1	319.6
		LS,m	-91.6	199.2	-337.2	319.2
		LS,l	-114.5	217.3	-338.1	319.9
		BM,m	-54.7	176.9	-336.9	319.3
		BM,l	-57.0	176.7	-335.4	319.1
CNEG	LS,ss	57.3	-291.5	51.6	-449.4	
	LS,s	69.2	-314.6	54.3	-449.2	
	LS,m	76.2	-328.8	55.6	-449.3	
	LS,l	85.2	-346.4	56.5	-449.3	
	BM,m	55.4	-304.3	57.6	-449.4	
	BM,l	55.7	-304.4	57.5	-449.5	
ACE	CAPO	LS,ss	-74.5	0.0	-70.2	0.0
		LS,s	-84.5	0.0	-70.5	0.0
		LS,m	-90.8	0.0	-70.6	0.0
		LS,l	-99.1	0.0	-71.1	0.0
		BM,m	-65.9	0.0	-69.6	0.0
		BM,l	-66.2	0.0	-69.9	0.0
	CHB	LS,ss	-71.1	-46.7	-66.2	-46.9
		LS,s	-81.1	-46.9	-66.7	-47.0
		LS,m	-87.2	-46.9	-66.7	-47.0
		LS,l	-95.4	-47.3	-67.1	-47.4
		BM,m	-62.8	-46.7	-65.6	-47.0
		BM,l	-63.1	-46.9	-65.8	-47.2
	CPOS	LS,ss	-16.1	-272.6	129.1	-430.2
		LS,s	-6.4	-295.5	131.3	-430.0
		LS,m	-0.4	-309.7	132.1	-430.2
		LS,l	7.2	-327.0	133.2	-430.0
		BM,m	-5.2	-285.3	133.6	-430.1
		BM,l	-4.0	-285.3	133.4	-430.2
	CNEG	LS,ss	-134.5	172.0	-268.2	327.8
		LS,s	-163.8	194.0	-272.3	327.7
		LS,m	-182.5	207.5	-274.1	327.5
		LS,l	-207.1	225.1	-276.1	327.6
		BM,m	-129.0	184.4	-273.8	327.6
		BM,l	-131.1	184.4	-273.8	327.6

signs of $\Delta A_{\text{chg,G(W)}}^{\text{raw}}$ and $\Delta A_{\text{chg,G(W)}}$ are different for simulations LS,ss, LS,s, LS,m, BM,m, and BM,l. Corresponding uncorrected values are slightly negative, that is, indicative of favorable ACE-

solvent interactions, whereas the corrected values are positive, indicative of a solvent polarization unfavorable for interactions with an anion (solvent polarized by the positively-charged functional group in CPOS, which is located closer to the solvent than is the ACE ion). For complexes MAM-CNEG, ACE-CPOS, and ACE-CNEG, the raw and corrected free energies of charging due to the host molecule, $\Delta A_{\text{chg,G(H)}}^{\text{raw}}$ and $\Delta A_{\text{chg,G(H)}}$, are the dominant components in the charging contributions to the respective binding free energies $\Delta\Delta A_{\text{chg}}^{\text{raw}}$ and $\Delta\Delta A_{\text{chg}}$, whereas for complex ACE-CHB the charging contributions $\Delta A_{\text{chg,G(W)}}^{\text{raw}}$ and $\Delta A_{\text{chg,G(W)}}$ due to the solvent are dominant.

The most drastic change in contributions to the binding free energy occurs with MAM-CPOS and is effected by the large value of ΔA_{pol} (-146.4 to -212.8 kJ mol⁻¹; Table 2) in combination with the fact that ΔA_{psum} also has a negative sign (cation charging) [eqs. (14) and (15)]. The least change in contributions to the binding free energy occurs with ACE-CHB and is effected by ΔA_{pol} and ΔA_{psum} approximately canceling each other (-48.8 to -71.0 kJ mol⁻¹ for the former versus 67.6 to 77.2 kJ mol⁻¹ for the latter; Table 2) by virtue of the positive sign of ΔA_{psum} in the case of anion charging [eqs. (14) and (15)]. Note, in this context, that in system MAM-CPOS, ΔA_{pol} corrects for spurious polarization around charges +1 *e* and +2 *e* in the initial and final states of the TI, respectively. Thus, considering, for example, the LS scheme, ΔA_{pol} approximately evaluates to three times the correction for artificial periodicity in the case of charging a single monovalent ion in a box of edge length L ,^[23,54,92] namely to $3N_A(8\pi\epsilon_0)^{-1}L^{-1}\xi_{\text{LS}}(1-\epsilon_S)^{-1}$, where the factor three arises from the proportionality of this correction to the square of the ionic charge (Appendix section "Calculation of ΔA_{pol} ") [eq. (A3)].

The raw charging free energies $\Delta A_{\text{chg}}^{\text{raw}}(\text{H})$ and $\Delta A_{\text{chg}}^{\text{raw}}(\text{W})$, as well as the corrected data $\Delta A_{\text{chg}}(\text{H})$ and $\Delta A_{\text{chg}}(\text{W})$ may be used to calculate raw and corrected charging contributions to the binding free energies, $\Delta\Delta A_{\text{chg}}^{\text{raw}}$ and $\Delta\Delta A_{\text{chg}}$, respectively [eqs. (1) and (21)]. For the corrected, that is, methodology-independent data, this can be done for all possible combinations of system sizes or/and electrostatics schemes used in the simulations of charging in water and in the host molecule. In practice, binding free energies are often calculated using computational boxes that are smaller for the in-water than for the in-host simulations. Table 4 reports the uncorrected data $\Delta\Delta A_{\text{chg}}^{\text{raw}}$ for such a situation (in-water charging in small box size, here L_{ss} for the LS scheme and L_m for the BM scheme; in-host charging in large box size, here L_l for the LS and BM scheme) and for those situations where approximate cancellation of periodicity-induced artifacts is expected to occur (in-water and in-host charging in boxes of equal size). Note, however, that the latter cancellation is of greater relevance for the LS scheme, because raw charging free energies obtained from simulations with the BM scheme are less sensitive to system size.^[89] The averages $\overline{\Delta\Delta A_{\text{chg}}}$ of corrected values $\Delta\Delta A_{\text{chg}}$ [eq. (21)] over all combinations of box sizes used for in-water and in-host charging, along with associated root-mean-square deviations are also provided. Values obtained for $\overline{\Delta\Delta A_{\text{chg}}^{\text{raw}}}$ based on L_{ss} for in-water charging and L_l for in-host charging differ by -28.1, -28.1, -35.2, and -21.0 kJ mol⁻¹ for MAM binding to

Table 4. Raw charging contributions $\Delta\Delta A_{\text{chg}}^{\text{raw}}$ [eq. (1)] to binding free energies of guest molecules MAM and ACE to hydrated host molecules CAPO, CHB, CPOS, and CNEG based on values for $\Delta A_{\text{chg}}^{\text{raw}}(\text{H})$ and $\Delta A_{\text{chg}}^{\text{raw}}(\text{W})$ calculated in four different system sizes (boxes of edge lengths L_{ss} , L_s , L_m , and L_i section “Simulated guest-host systems”) using LS or BM electrostatic interactions (Table 2). Only a subset of the 16 (LS scheme-based) or four (BM scheme-based) possible combinations is reported. $L(\text{W})$ and $L(\text{H})$ denote the box-edge lengths used for simulations of in-water and in-host charging, respectively. For comparison, averages $\overline{\Delta\Delta A_{\text{chg}}}$ of corrected values $\Delta\Delta A_{\text{chg}}$ [eq. (21)] over all 16 combinations of box sizes $\{L_i\} = L_{\text{ss}}$, L_s , L_m and L_i in the case of the LS scheme, over all four combinations of box sizes $\{L_i\} = L_m$ and L_i in the case of the BM scheme, or over the union of the two latter sets (denoted LS+BM) used for in-water and in-host charging, along with associated root-mean-square deviations (rmsd) are also provided.

Scheme	Guest		MAM				ACE			
	Host		CAPO	CHB	CPOS	CNEG	CAPO	CHB	CPOS	CNEG
	$L(\text{W})$	$L(\text{H})$				$\Delta\Delta A_{\text{chg}}^{\text{raw}}(\text{kJ mol}^{-1})$				
LS	L_{ss}	L_i	147.9	112.9	275.4	−88.6	201.7	158.1	−19.0	318.8
LS	L_{ss}	L_{ss}	168.2	133.9	288.3	−61.6	226.3	183.0	12.1	338.3
LS	L_s	L_s	169.4	134.5	293.0	−63.6	226.1	182.6	8.7	340.8
LS	L_m	L_m	170.5	134.9	295.7	−64.5	227.8	184.5	8.5	343.6
LS	L_i	L_i	173.1	138.1	300.6	−63.4	227.9	184.3	7.2	345.0
BM	L_m	L_i	164.0	129.4	293.4	−75.0	233.6	189.8	10.5	353.1
BM	L_m	L_m	165.2	129.6	295.9	−75.2	233.9	190.3	9.3	355.2
BM	L_i	L_i	165.8	131.2	295.2	−73.2	235.3	191.5	12.2	354.8
						$\overline{\Delta\Delta A_{\text{chg}}}(\text{kJ mol}^{-1})$				
LS	$\{L_i\}$	$\{L_i\}$	176.0	141.0	310.6	−67.6	228.7	185.6	0.7	354.3
BM	$\{L_i\}$	$\{L_i\}$	179.8	144.8	312.4	−62.6	229.6	186.6	2.7	353.2
LS+BM	$\{L_i\}$	$\{L_i\}$	176.7	141.8	310.9	−66.6	228.9	185.8	1.1	354.1
						rmsd ($\Delta\Delta A_{\text{chg}}$) (kJ mol ^{−1})				
LS	$\{L_i\}$	$\{L_i\}$	0.8	0.9	1.9	2.0	0.4	0.5	1.6	3.0
BM	$\{L_i\}$	$\{L_i\}$	1.4	1.6	1.5	1.4	1.0	1.0	1.0	1.0
LS+BM	$\{L_i\}$	$\{L_i\}$	1.8	1.9	2.0	2.8	0.6	0.8	1.7	2.7

CAPO, CHB, CPOS, and CNEG, respectively, and by −27.0, −27.5, −19.7, and −35.5 kJ mol^{−1} for ACE binding to CAPO, CHB, CPOS, and CNEG, respectively, from corresponding data for $\overline{\Delta\Delta A_{\text{chg}}}$ (LS scheme), and values obtained for $\Delta\Delta A_{\text{chg}}^{\text{raw}}$ based on L_m for in-water charging and L_i for in-host charging differ by −15.8, −15.4, −19.0, and −12.4 kJ mol^{−1} for MAM binding to CAPO, CHB, CPOS, and CNEG, respectively, and by 4.0, 3.2, 7.8, and −0.1 kJ mol^{−1} for ACE binding to CAPO, CHB, CPOS, and CNEG, respectively, from corresponding data for $\overline{\Delta\Delta A_{\text{chg}}}$ (BM scheme). The majority of these deviations are non-negligible, and it is thus essential to correct raw charging contributions to binding free energies. Note that box-edge length dependence is more pronounced for simulations with the LS scheme, because here the system-size parameter crucially determines the magnitude of artificial periodicity artifacts.

If both the in-water and in-host charging simulations are conducted in boxes of identical edge length, the deviations are significantly reduced for the LS scheme, that is, they evaluate to −2.9, −2.9, −10.0, and 4.2 kJ mol^{−1} for MAM binding to CAPO, CHB, CPOS, and CNEG, respectively, and to −0.8, −1.3, 6.5, and −9.3 kJ mol^{−1} for ACE binding to CAPO, CHB, CPOS, and CNEG, respectively, based on simulations in boxes of edge length L_i , the best agreement with $\overline{\Delta\Delta A_{\text{chg}}}$ thus being achieved for complexes containing the apolar CAPO and CHB host molecules. Note that simulations in equisized boxes do not lead to an improvement for the BM scheme, where, using data pertaining to edge length L_i , the deviations of $\Delta\Delta A_{\text{chg}}^{\text{raw}}$ from $\overline{\Delta\Delta A_{\text{chg}}}$ are −14.0, −13.6, −17.2, and −10.6 kJ mol^{−1} for MAM binding to CAPO, CHB, CPOS, and CNEG, respectively, and 5.7, 4.9, 9.5, and 1.6 kJ mol^{−1} for ACE binding to CAPO, CHB, CPOS, and CNEG, respectively.

The averages $\overline{\Delta\Delta A_{\text{chg}}}$ differ for simulation data pertaining to solely the LS or BM scheme by 0.9–5.0 kJ mol^{−1}. Overall, the averages $\overline{\Delta\Delta A_{\text{chg}}}$ based on the BM scheme data differ on average by 2.4 kJ mol^{−1} from the LS scheme data, the agreement between the two different electrostatics schemes being better for ACE-containing complexes (average absolute difference 1.3 kJ mol^{−1}) than for MAM-containing complexes (average absolute difference 3.6 kJ mol^{−1}). This might be due to favorable cancelation of artifacts in the ACE-containing complexes, as well as the more pronounced cutoff artifacts in $P(r)$ and the continuum-electrostatics-based correction scheme insufficiently capturing the pronounced overpolarization within the cutoff sphere for the MAM-containing complexes (Fig. 4). In comparison to the polarization in a homogeneous dielectric medium, approximated here by the Born polarization $P_{\text{Born}}(r)$ [eq. (6)] around a charge of +1 *e* (MAM-CAPO, MAM-CHB), +2 *e* (MAM-CPOS), −1 *e* (ACE-CAPO, ACE-CHB), or −2 *e* (ACE-CNEG) centered at the MAM nitrogen or the ACE carboxylate carbon atom, hydration shell peaks in $P(r)$ appear more pronounced for MAM in comparison to ACE in neutral host cavities CAPO and CHB and significantly broader for MAM in comparison to ACE in host cavities CAPO, CHB and the like-charged functionalized one (CPOS in the case of MAM, CNEG in the case of ACE). A less pronounced water radial polarization around anionic in comparison to cationic solutes was also observed before in the context of the hydration of monoatomic ions and can be drawn back to a decreased orientational freedom of water molecules around cations.^[58]

The corrected charging contributions (entailing all possible combinations of box sizes for in-water and in-host charging) show rmsd values within 2.5 kJ mol^{−1} for all complexes except

MAM-CNEG (2.8 kJ mol^{-1}) and ACE-CNEG (2.7 kJ mol^{-1}). As discussed above, the spread in $\Delta\Delta A_{\text{chg}}$ for these systems may be drawn back to the inability of the continuum-electrostatics approximation to capture short-range artifacts in solvent structure which appear to be very strong for the LS,ss simulations of these complexes.

Both raw and corrected charging contributions to the binding free energy, $\Delta\Delta A_{\text{chg}}^{\text{raw}}$ and $\Delta\Delta A_{\text{chg}}$ (Table 4), obey intuitive reasoning in that they are least favorable for the like-charged guest-host complexes (MAM-CPOS, ACE-CNEG), considerably less unfavorable for the apolar host cavity (CAPO) and the host cavity allowing hydrogen bonding (CHB) and least unfavorable for the oppositely-charged guest-host complexes (MAM-CNEG, ACE-CPOS). In comparison to charging the guest in water, binding to the host is, however, only favorable in the case of MAM-CNEG ($\overline{\Delta\Delta A_{\text{chg}}} = -66.6 \text{ kJ mol}^{-1}$; Table 4). The charging of guest molecule ACE is basically indifferent toward pure water or host CPOS environments ($\overline{\Delta\Delta A_{\text{chg}}} = 1.1 \text{ kJ mol}^{-1}$; Table 4), which can probably be explained in terms of water being an extremely good solvent for anion solvation because the hydrogen atoms of the water molecule can approach anions very closely.^[98–101]

Altogether, as it can significantly alter the charging contribution to binding free energies and thus crucially change the interpretation or prediction of experimental data, analysis of possible electrostatic artifacts and application of required correction terms appears very important and should be considered an integral part of careful free-energy calculation studies if changes in the net charge are involved. Note that for more complex guest-host systems (e.g., a drug-receptor complex) it might be necessary to take into account the possible flexibility of the molecules, giving rise to time-dependent correction terms.

Conclusion

The calculation of binding free energies of charged species to a target molecule is a frequently encountered problem in MD studies of (bio-)chemical thermodynamics. A number of important endogenous receptor-binding molecules (e.g., glutamate, acetylcholine), enzyme substrates (e.g., superoxide anion, lysine) or drug molecules (e.g., aspirin, proguanil) have a non-zero net charge. Absolute binding free energies, as well as binding free energies relative to another molecule with a different net charge will be affected by artifacts due to the used effective electrostatic interaction function and associated parameters (e.g., size of the computational box). This is increasingly being recognized in the field of free-energy simulations. Independently from the authors' work, Rocklin et al. proposed a very similar correction scheme.^[102] In the present study, charging contributions to binding free energies of either of two ionic guest molecules, MAM and ACE, to functionalized buckyball-like host cavities were calculated with classical atomistic MD simulation. Electrostatic interactions were treated using a LS scheme or a BM scheme, and the simulations were conducted in boxes of four different edge lengths. It was illustrated that: (i) the charging free energies of the guest mole-

cules in water and in the host molecule strongly depend on the applied methodology; (ii) the charging free energies of the guest molecules in water and in the host molecule obtained from the LS scheme present a non-negligible dependence on the edge length of the simulation box; (iii) considering the investigated systems, error cancellation in computed charging contributions to binding free energies is only approximately guaranteed for systems with an apolar cavity (zero host charges) if corresponding in-water and in-host charging simulations are performed with the LS scheme in equisized boxes; (iv) neglect of correction terms for the artifacts introduced by the finite size of the simulated system and the use of an effective electrostatic interaction function considerably impairs the thermodynamic interpretation of guest-host interactions, and in particular the relative contributions of the solvent and the host compound; (v) application of correction terms for spurious solvent polarization and wrong dielectric permittivity of the solvent model, improper electrostatic potential summation, effective guest-host direct electrostatic interactions, and the presence of electrostatic interactions between excluded solute atoms in the Hamiltonian yields consistent results for the charging contribution to binding free energies. In particular, rmsd values over 20 results lie within 2.5 kJ mol^{-1} for all systems except MAM-CNEG and ACE-CNEG. For these systems, the spread might be drawn back to strong artifacts in solvent configurational sampling in a very small computational box using the LS scheme for the treatment of electrostatic interactions that are not captured by the continuum-electrostatics-based correction procedure.

As long as simulations of macroscopic nonperiodic systems with Coulombic electrostatic interactions, or electrostatic interactions truncated at sufficiently large distances, such that an adequate representation of experimental bulk systems is achieved, are out of reach, the proposed correction scheme for the charging contribution to binding free energies is a crucial step in obtaining thermodynamically sensible results for the free energy of binding of charged ligands.

Acknowledgments

The authors thank Philippe Hünenberger for stimulating discussions and note that the present study and the work of Rocklin et al., 2013 evolved independently.

APPENDIX

Calculation of ΔA_{pol}

Equation (12) describes calculation of a charging free energy via TI of the environment-generated electrostatic potential monitored at the guest atoms along a scaling parameter λ which grows the guest atomic partial charges from zero to the full charge state. This integration is commonly performed numerically using a finite number of intermediate guest charge states. In the present work, six charge states were used (section "Free-energy correction terms"). Evaluation of ΔA_{pol} requires conduction of two

Table A1. Averages $\overline{\Delta A_{\text{chg}}}$ and root-mean-square deviations (rmsd) of corrected values ΔA_{chg} [eq. (20) and Table 2], for charging of the guest molecules MAM and ACE in hydrated host molecules CAPO, CHB, CPOS, and CNEG (section "MD simulations"). The correction term ΔA_{psum} was evaluated using eq. (14) for the LS scheme and either eq. (A4) or (15) for the BM scheme. The volume V_{vdW} of the host cavity was approximated by a sphere of radius R_{vdW} [eq. (A4)]. For a given approach and choice of R_{vdW} averages $\overline{\text{rmsd}}$ over the individual ΔA_{chg} rmsd values are provided.

Guest	Host	Equations (14) and (A4), $R_{\text{vdW}} = 0.5 \text{ nm}$		Equations (14) and (A4), $R_{\text{vdW}} = 0.6 \text{ nm}$		Equations (14) and (A4), $R_{\text{vdW}} = 0.7 \text{ nm}$		Equations (14) and (15)	
		$\overline{\Delta A_{\text{chg}}}$ (kJ mol ⁻¹)	rmsd (kJ mol ⁻¹)	$\overline{\Delta A_{\text{chg}}}$ (kJ mol ⁻¹)	rmsd (kJ mol ⁻¹)	$\overline{\Delta A_{\text{chg}}}$ (kJ mol ⁻¹)	rmsd (kJ mol ⁻¹)	$\overline{\Delta A_{\text{chg}}}$ (kJ mol ⁻¹)	rmsd (kJ mol ⁻¹)
MAM	CAPO	-152.1	1.3	-151.5	0.4	-150.6	0.9	-150.6	0.8
	CHB	-187.0	1.4	-186.4	0.7	-185.5	1.1	-185.6	1.0
	CPOS	-18.1	2.5	-17.5	1.9	-16.6	1.5	-16.7	1.5
	CNEG	-395.3	1.8	-394.7	1.6	-393.8	2.1	-393.8	2.1
ACE	CAPO	-68.7	2.7	-69.3	1.8	-70.2	0.6	-70.3	0.5
	CHB	-111.9	2.7	-112.5	1.8	-113.4	0.6	-113.4	0.6
	CPOS	-296.5	3.3	-297.1	2.5	-298.0	1.6	-298.0	1.6
	CNEG	56.1	2.9	55.5	2.5	54.6	2.5	54.6	2.5
$\overline{\text{rmsd}}$ (kJ mol ⁻¹)			2.3		1.7		1.4		1.3

[eq. (10)] or four [eq. (11)] such TIs. However, three alternative computationally less intensive approaches may be conceived of:

- As the electrostatic potential at the guest atoms varies linearly with the guest charge state, the TI can be cut down to involve only the initial and final state, that is, eq. (12) may be simplified to

$$\Delta A_{\text{chg}}^X = \sum_{i=1}^{N_G} (1/2) q_i [\phi^{\text{inter}}(\mathbf{r}_i; \lambda = 0, X) + \phi^{\text{inter}}(\mathbf{r}_i; \lambda = 1, X)], \quad (\text{A1})$$

where the factor (1/2) arises from numerical integration using the trapezoidal rule and the electrostatic potentials are evaluated at $\lambda = 0$ (uncharged guest) and $\lambda = 1$ (fully charged guest). ΔA_{pol} then follows from eqs. (10) and (11).

- A linear variation of the electrostatic potential at the guest atoms with the guest charge state allows in principle a linear-response approximation according to

$$\Delta A_{\text{chg}}^X = (1/2) \left\{ \sum_{i=1}^{N_G} q_i \phi^{\text{inter}}(\mathbf{r}_i; \lambda = 1, X) + Q_H \frac{\sum_{i=1}^{N_G} [\phi^{\text{inter}}(\mathbf{r}_i; \lambda = 1, X) - \phi^{\text{inter}}(\mathbf{r}_i; \lambda = 0, X)]}{N_G} \right\}, \quad (\text{A2})$$

where the factor (1/2) arises from the assumption of linear response, and the second term in the curly brackets accounts for a possibly nonzero host charge Q_H . In the latter case, eq. (A2) is very approximate because it assumes that the host charge is concentrated at one point and experiences the average of electrostatic potentials monitored at the guest atom sites. ΔA_{pol} then follows from eqs. (10) and (11).

- A crude approximation to ΔA_{pol} for the LS scheme may be obtained through usage of the analytical formula per-

tinents to the case of solvation of a single nonpolarizable monoatomic ion of radius R_I in a box of edge length L filled with a homogeneous dielectric medium of permittivity ϵ_S .^[54,92] In the present case of a solvated guest-host system, this formula may be rewritten

$$\Delta A_{\text{pol}}^{\text{LS}} = N_A [(Q_G + Q_H)^2 - Q_H^2] (8\pi\epsilon_0)^{-1} (1 - \epsilon_S^{-1}) L^{-1} \cdot \left[\xi_{\text{LS}} + \frac{4\pi}{3} \left(\frac{R_{\text{vdW}}}{L} \right)^2 - \frac{16\pi^2}{45} \left(\frac{R_{\text{vdW}}}{L} \right)^5 \right], \quad (\text{A3})$$

where $\xi_{\text{LS}} = -2.837297$.^[82,83,103] Equation (A3) approximates the guest-host complex by a cavity of radius R_{vdW} containing a point charge of magnitudes Q_H and $Q_G + Q_H$ in the initial and final states, respectively.

As an example, the case of MAM charging in CPOS in system LS,I is considered. Equation (12), using six charge states as described in section "Free-energy correction terms" yields $\Delta A_{\text{pol}} = -146.4 \text{ kJ mol}^{-1}$ (Table 2). Equation (A1) yields the same result, $\Delta A_{\text{pol}} = -146.4 \text{ kJ mol}^{-1}$. Equation (A2) yields $\Delta A_{\text{pol}} = -147.3 \text{ kJ mol}^{-1}$, with contributions of -97.9 and $-49.4 \text{ kJ mol}^{-1}$ from the first and second term in curly brackets, respectively. Using $Q_G = 1 e$, $Q_H = 1 e$, $R_{\text{vdW}} = 0.5 \text{ nm}$, $L = 3.80 \text{ nm}$ and $\epsilon_S = 66.6$, eq. (A3) yields $-149.4 \text{ kJ mol}^{-1}$. Whereas eq. (A1) is numerically accurate (due to a linear charging profile the initial and final points of the charging curve are sufficient to calculate the charging free energy), eqs. (A2) and (A3) rely on the afore mentioned approximations, which causes the resulting ΔA_{pol} value to be less accurate. However, both of the two latter estimates deviate less than $k_B T$ from the numerically accurate ΔA_{pol} result.

Calculation of ΔA_{psum}

In Ref. [61], ΔA_{psum} was calculated as eq. (14) for the LS scheme and using

$$\Delta A_{\text{psum}}^{\text{BM}} = -N_A (6\epsilon_0)^{-1} \frac{2(\epsilon_{\text{BW}} - 1)}{2\epsilon_{\text{BW}} + 1} \gamma_w Q_G N_w \frac{1 - [(4/3)\pi R_C^3]^{-1} V_{\text{vdW}}}{L^3 - V_{\text{vdW}}} \quad (\text{A4})$$

for the BM scheme. V_{vdW} is an estimate for the volume of the solute cavity (guest molecule in the case of in-water charging; host molecule in the case of in-host charging). Here, the volume of the host cavity was approximated by a sphere of radius R_{vdW} (Table A1).

In the present study, a slightly different equation was used for ΔA_{psum}^{BM} [eq. (15)] (section "Free-energy correction terms"). Equation (15) derives from expressing the product of the fraction f_c of the cutoff sphere occupied by water with the water number density η_w in the original expression^[54,89] for ΔA_{psum}^{BM} in the case of hydration of a single monoatomic ion,

$$\Delta A_{psum}^{BM} = -N_A(6\epsilon_o)^{-1} \frac{2(\epsilon_{BW}-1)}{2\epsilon_{BW}+1} \gamma_w Q_G f_c \eta_w, \quad (A5)$$

as

$$\begin{aligned} f_c \eta_w &= \{1 - V_{vdW}[(4/3)\pi R_C^3]^{-1}\} \langle N_w(R_C) \rangle \{(4/3)\pi R_C^3 - V_{vdW}\}^{-1} \\ &= \langle N_w(R_C) \rangle [(4/3)\pi R_C^3]^{-1}. \end{aligned} \quad (A6)$$

Similarly, eq. (14) derives from expressing the product of the fraction f_b of the computational box occupied by water with the water number density η_w in the original expression^[54,89] for ΔA_{psum}^{LS} in the case of hydration of a single monoatomic ion,

$$\Delta A_{psum}^{LS} = -N_A(6\epsilon_o)^{-1} \gamma_w Q_G f_b \eta_w, \quad (A7)$$

as

$$f_b \eta_w = (1 - V_{vdW}L^{-3}) N_w (L^3 - V_{vdW})^{-1} = N_w L^{-3}. \quad (A8)$$

Note that the expressions for the water number density used in eqs. (14) and (15), that is, $N_w L^{-3}$ and $\langle N_w(R_C) \rangle [(4/3)\pi R_C^3]^{-1}$, respectively, might not be adequate for simulations LS,ss, because of spurious water density fluctuations in this very small system that are not captured by a global (box-volume associated) density measure. In particular, for simulations LS,ss, the height of the first and second hydration shell peaks around the buckyball cavity is overestimated by up to 18 and 15% for MAM and ACE, respectively, in hydrated host molecules CAPO, CHB, CPOS, and CNEG in comparison to simulations LS,l and BM,l (Fig. 3). Equation (15) seems advantageous in comparison to eq. (A4) because it yields greater methodological independence, as evidenced by the root-mean-square deviations (rmsd) of corrected charging free energies ΔA_{chg} [eq. (20)] for charging the guest molecules in the host, reported in Table A1, and is independent of the volume of the solute. Three different choices for R_{vdW} were investigated, namely $R_{vdW} = 0.5, 0.6,$ or 0.7 nm. Distances of 0.5 and 0.7 nm from the center of mass of the 60 buckyball carbon atoms correspond to the average location of the 60 buckyball carbons atoms and the value where the water oxygen radial distribution function starts to deviate from zero, respectively. Overall, the cavity volume-independent eqs. (14) and (15) achieve lower rmsd values, the average rmsd over all systems being 1.3 kJ mol^{-1} , as

compared to 2.3, 1.7, and 1.4 kJ mol^{-1} for the three above choices of radii in eq. (A4).

Keywords: computer simulation · molecular dynamics · free-energy calculations · charging free energies · electrostatic artifacts

How to cite this article: M. M. Reif, C. Oostenbrink. *J. Comput. Chem.* **2014**, *35*, 227–243. DOI: 10.1002/jcc.23490.

- [1] T. Simonson, G. Archontis, M. Karplus, *Acc. Chem. Res.* **2002**, *35*, 430.
- [2] M. K. Gilson, H. X. Zhou, *Annu. Rev. Biophys. Biomol. Struct.* **2007**, *36*, 21.
- [3] Y. Deng, B. Roux, *J. Phys. Chem. B* **2009**, *113*, 2234.
- [4] J. Michel, J. W. Essex, *J. Comput. Aided Mol. Des.* **2010**, *24*, 639.
- [5] T. Steinbrecher, A. Labahn, *Curr. Med. Chem.* **2010**, *8*, 767.
- [6] J. D. Chodera, D. L. Mobley, M. R. Shirts, R. W. Dixon, K. Branson, V. S. Pande, *Curr. Opin. Struct. Biol.* **2011**, *21*, 150.
- [7] A. de Ruiter, C. Oostenbrink, *Curr. Opin. Chem. Biol.* **2011**, *15*, 547.
- [8] D. L. Mobley, P. V. Klimovich, *J. Chem. Phys.* **2012**, *137*, 230901/1.
- [9] J. C. Gumbart, B. Roux, C. Chipot, *J. Chem. Theory Comput.* **2013**, *9*, 794.
- [10] N. Gresh, G. A. Cisneros, T. A. Darden, J.-P. Piquemal, *J. Chem. Theory Comput.* **2007**, *3*, 1960.
- [11] P. Söderhjelm, *Theor. Chem. Acc.* **2012**, *131*, 1159/1.
- [12] J. Åqvist, *J. Comput. Chem.* **1996**, *17*, 1587.
- [13] B. R. Morgan, F. Massi, *J. Chem. Theory Comput.* **2010**, *6*, 1884.
- [14] X. Lu, Q. Cui, *J. Phys. Chem. B* **2013**, *117*, 2005.
- [15] A. W. Schüttelkopf, D. M. F. van Aalten, *Acta Crystallogr.* **2004**, *D60*, 1355.
- [16] A. K. Malde, L. Zuo, M. Breeze, M. Stroet, D. Poger, P. C. Nair, C. Oostenbrink, A. E. Mark, *J. Chem. Theory Comput.* **2011**, *7*, 4026.
- [17] N. Agarwal, M. P. Alam, C. Chakravarty, *J. Phys. Chem. B* **2011**, *115*, 6935.
- [18] S. Genheden, U. Ryde, *J. Chem. Theory Comput.* **2011**, *7*, 3768.
- [19] S. Genheden, U. Ryde, *Proteins* **2012**, *80*, 1326.
- [20] H. K. Srivastava, G. N. Sastry, *J. Chem. Inf. Model.* **2012**, *52*, 3088.
- [21] J. Åqvist, C. Medina, J.-E. Samuelsson, *Protein Eng.* **1994**, *7*, 385.
- [22] J. Åqvist, T. Hansson, *J. Phys. Chem.* **1996**, *100*, 9512.
- [23] G. Hummer, L. R. Pratt, A. E. Garcia, *J. Phys. Chem.* **1996**, *100*, 1206.
- [24] G. Hummer, A. Szabo, *J. Chem. Phys.* **1996**, *105*, 2004.
- [25] A. de Ruiter, C. Oostenbrink, *J. Chem. Theory Comput.* **2012**, *8*, 3686.
- [26] F. M. Floris, M. Selmi, A. Tani, J. Tomasi, *J. Chem. Phys.* **1997**, *107*, 6353.
- [27] G. Graziano, *Chem. Phys. Lett.* **2009**, *479*, 56.
- [28] X. Liu, C. Wang, J. Wang, Z. Li, H. Zhao, R. Luo, *Phys. Chem. Chem. Phys.* **2013**, *15*, 129.
- [29] C. S. Babu, P.-K. Yang, C. Lim, *J. Biol. Phys.* **2002**, *28*, 95.
- [30] J. E. Basner, C. Jarzynski, *J. Phys. Chem. B* **2008**, *112*, 12722.
- [31] I. Buch, S. K. Sadiq, G. De Fabritis, *J. Chem. Theory Comput.* **2011**, *7*, 1765.
- [32] M. Jorge, N. M. Garrido, A. J. Queimada, I. G. Economou, E. A. Macedo, *J. Chem. Theory Comput.* **2010**, *6*, 1019.
- [33] A. de Ruiter, S. Boresch, C. Oostenbrink, *J. Comput. Chem.* **2013**, *34*, 1024.
- [34] C. H. Bennet, *J. Comput. Phys.* **1976**, *22*, 245.
- [35] M. R. Shirts, V. S. Pande, *J. Chem. Phys.* **2005**, *122*, 144107/1.
- [36] M. R. Shirts, J. D. Chodera, *J. Chem. Phys.* **2008**, *129*, 124105/1.
- [37] V. Cruz, V. Ramos, J. Martínez-Salazar, *J. Phys. Chem. B* **2012**, *116*, 469.
- [38] G. M. Torrie, J. P. Valleau, *J. Comput. Phys.* **1977**, *23*, 187.
- [39] J. P. Valleau, G. M. Torrie, In *Modern Theoretical Chemistry*, Vol. 5; B. J. Berne, Ed.; Plenum Press: New York, **1977**; pp. 169–194.
- [40] C. Bartels, M. Karplus, *J. Comput. Chem.* **1997**, *18*, 1450.
- [41] E. Piccinini, M. Ceccarelli, F. Affinito, R. Brunetti, C. Jacoboni, *J. Chem. Theory Comput.* **2008**, *4*, 173.
- [42] K.-K. Han, *Phys. Lett. A* **1992**, *165*, 28.
- [43] C. D. Christ, W. F. van Gunsteren, *J. Chem. Phys.* **2008**, *128*, 174112/1.

- [44] H. J. C. Berendsen, In *Computer Simulation of Biomolecular Systems, Theoretical and Experimental Applications*, Vol. 2; W. F. van Gunsteren, P. K. Weiner, A. J. Wilkinson, Eds.; ESCOM Science Publishers B.V.: Leiden, Netherlands, **1993**; pp. 161–181.
- [45] M. E. Davis, J. A. McCammon, *Chem. Rev.* **1990**, *90*, 509.
- [46] R. M. Levy, E. Gallicchio, *Annu. Rev. Phys. Chem.* **1998**, *49*, 531.
- [47] T. Darden, L. Perera, L. Li, L. Pedersen, *Structure* **1999**, *7*, R55.
- [48] P. H. Hünenberger, U. Börjesson, R. D. Lins, *Chimia* **2001**, *55*, 861.
- [49] P. Koehl, *Curr. Opin. Struct. Biol.* **2006**, *16*, 142.
- [50] T. E. Cheatham, III, B. R. Brooks, *Theor. Chem. Acc.* **1998**, *99*, 279.
- [51] C. Sagui, T. A. Darden, *Annu. Rev. Biophys. Biomol. Struct.* **1999**, *28*, 155.
- [52] J. Norberg, L. Nilsson, *Acc. Chem. Res.* **2002**, *35*, 465.
- [53] M. A. Kastenholz, P. H. Hünenberger, *J. Phys. Chem. B* **2004**, *108*, 774.
- [54] M. A. Kastenholz, P. H. Hünenberger, *J. Chem. Phys.* **2006**, *124*, 224501/1.
- [55] M. M. Reif, V. Kräutler, M. A. Kastenholz, X. Daura, P. H. Hünenberger, *J. Phys. Chem. B* **2009**, *113*, 3112.
- [56] P. H. Hünenberger, M. M. Reif, *Single-Ion Solvation: Experimental and Theoretical Approaches to Elusive Thermodynamic Quantities*, 1st ed.; Royal Society of Chemistry (Theoretical and Computational Chemistry Series): London, UK, **2011**.
- [57] M. M. Reif, P. H. Hünenberger, *J. Chem. Phys.* **2011**, *134*, 144103/1.
- [58] M. M. Reif, P. H. Hünenberger, *J. Chem. Phys.* **2011**, *134*, 144104/1.
- [59] S. Garde, G. Hummer, M. E. Paulaitis, *J. Chem. Phys.* **1998**, *108*, 1552.
- [60] K. P. Jensen, W. L. Jorgensen, *J. Chem. Theory Comput.* **2006**, *2*, 1499.
- [61] M. M. Reif, P. H. Hünenberger, C. Oostenbrink, *J. Chem. Theory Comput.* **2012**, *8*, 3705.
- [62] J. Dolenc, C. Oostenbrink, J. Koller, W. F. van Gunsteren, *Nucleic Acids Res.* **2005**, *33*, 725.
- [63] A. de Ruiter, C. Oostenbrink, *J. Chem. Theory Comput.* **2013**, *9*, 883.
- [64] M. van den Bosch, M. Swart, J. G. Snijders, H. J. C. Berendsen, A. E. Mark, C. Oostenbrink, W. F. van Gunsteren, G. W. Canters, *ChemBioChem* **2005**, *6*, 738.
- [65] M. J. Harvey, G.-De Fabritiis, *J. Chem. Theory Comput.* **2009**, *5*, 2371.
- [66] D. J. Hardy, J. E. Stone, K. Schulten, *Parallel Comput.* **2009**, *35*, 164.
- [67] R. Yokota, J. P. Bardhan, M. G. Knepley, L. A. Barba, T. Hamada, *Comput. Phys. Commun.* **2011**, *182*, 1271.
- [68] M. Cascella, M. A. Neri, P. Carloni, M. Dal Peraro, *J. Chem. Theory Comput.* **2008**, *4*, 1378.
- [69] M. Masella, D. Borgis, P. Cuniasse, *J. Comput. Chem.* **2013**, *34*, 1112.
- [70] C. Oostenbrink, *J. Comput. Chem.* **2009**, *30*, 212.
- [71] W. F. van Gunsteren, S. R. Billeter, A. A. Eising, P. H. Hünenberger, P. Krüger, A. E. Mark, W. R. P. Scott, I. G. Tironi, *Biomolecular Simulation: The GROMOS96 Manual and User Guide*; Verlag der Fachvereine: Zürich, Switzerland, **1996**.
- [72] R. W. Hockney, J. W. Eastwood, *Computer Simulation Using Particles*; McGraw-Hill: New York, **1981**.
- [73] P. H. Hünenberger, In *Simulation and Theory of Electrostatic Interactions in Solution: Computational Chemistry, Biophysics, and Aqueous Solution*; G. Hummer, L. R. Pratt, Eds.; American Institute of Physics: New York, **1999**; pp. 17–83.
- [74] J. A. Barker, R. O. Watts, *Mol. Phys.* **1973**, *26*, 789.
- [75] N. Schmid, C. D. Christ, M. Christen, A. P. Eichenberger, W. F. van Gunsteren, *Comput. Phys. Commun.* **2012**, *183*, 890.
- [76] H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, J. Hermans, In *Intermolecular Forces*; B. Pullman, Ed.; Reidel: Dordrecht, Netherlands, **1981**; pp. 331–342.
- [77] R. W. Hockney, *Methods Comput. Phys.* **1970**, *9*, 136.
- [78] J.-P. Ryckaert, G. Ciccotti, H. J. C. Berendsen, *J. Comput. Phys.* **1977**, *23*, 327.
- [79] H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, A. di Nola, J. R. Haak, *J. Chem. Phys.* **1984**, *81*, 3684.
- [80] B. A. Luty, I. G. Tironi, W. F. van Gunsteren, *J. Chem. Phys.* **1995**, *103*, 3014.
- [81] P. H. Hünenberger, *J. Chem. Phys.* **2000**, *113*, 10464.
- [82] A. Redlack, J. Grindlay, *Can. J. Phys.* **1972**, *50*, 2815.
- [83] B. R. A. Nijboer, T. W. Ruijgrok, *J. Stat. Phys.* **1988**, *53*, 361.
- [84] T. N. Heinz, P. H. Hünenberger, *J. Chem. Phys.* **2005**, *123*, 034107/1.
- [85] A. Glättli, X. Daura, W. F. van Gunsteren, *J. Chem. Phys.* **2002**, *116*, 9811.
- [86] M. Christen, P. H. Hünenberger, D. Bakowies, R. Baron, R. Bürgi, D. P. Geerke, T. N. Heinz, M. A. Kastenholz, V. Kräutler, C. Oostenbrink, C. Peter, D. Trzesniak, W. F. van Gunsteren, *J. Comput. Chem.* **2005**, *26*, 1719.
- [87] J. G. Kirkwood, *J. Chem. Phys.* **1935**, *3*, 300.
- [88] M. P. Allen, D. J. Tildesley, *Computer Simulation of Liquids*; Oxford University Press: New York, **1987**.
- [89] M. A. Kastenholz, P. H. Hünenberger, *J. Chem. Phys.* **2006**, *124*, 124106/1.
- [90] M. E. Davis, J. D. Madura, B. A. Luty, J. A. McCammon, *Comput. Phys. Commun.* **1991**, *62*, 187.
- [91] J. D. Madura, J. M. Briggs, R. C. Wade, M. E. Davis, B. A. Luty, A. Ilin, J. Antosiewicz, M. K. Gilson, B. Bagheri, L. R. Scott, J. A. McCammon, *Comput. Phys. Commun.* **1995**, *91*, 57.
- [92] P. H. Hünenberger, J. A. McCammon, *J. Chem. Phys.* **1999**, *110*, 1856.
- [93] C. Peter, W. F. van Gunsteren, P. H. Hünenberger, *J. Chem. Phys.* **2002**, *116*, 7434.
- [94] C. Peter, W. F. van Gunsteren, P. H. Hünenberger, *J. Chem. Phys.* **2003**, *119*, 12205.
- [95] C. Oostenbrink, A. Villa, A. E. Mark, W. F. van Gunsteren, *J. Comput. Chem.* **2004**, *25*, 1656.
- [96] P. Ferrara, A. Caflich, *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 10780.
- [97] M. Bergdorf, C. Peter, P. H. Hünenberger, *J. Chem. Phys.* **2003**, *119*, 9129.
- [98] S. Rajamani, T. Ghosh, S. Garde, *J. Chem. Phys.* **2004**, *120*, 4457.
- [99] A. Grossfield, *J. Chem. Phys.* **2005**, *122*, 024506/1.
- [100] D. L. Mobley, A. E. Barber, II, C. J. Fennell, K. A. Dill, *J. Phys. Chem. B* **2008**, *112*, 2405.
- [101] D. L. Mobley, J. R. Baker, A. E. Barber, II, C. J. Fennell, K. A. Dill, *J. Phys. Chem. B* **2011**, *115*, 15145.
- [102] G. J. Rocklin, D. L. Mobley, K. A. Dill, P. H. Hünenberger, *J. Chem. Phys.* doi: 10.1063/1.4826261.
- [103] G. Placzek, B. R. A. Nijboer, L. van Hove, *Phys. Rev.* **1951**, *82*, 392.

Received: 13 September 2013

Accepted: 13 October 2013

Published online on 19 November 2013