JCTC Journal of Chemical Theory and Computation

Article

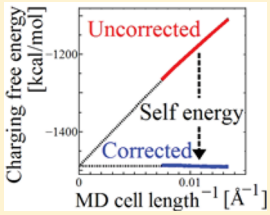
pubs.acs.org/JCTC

Finite-Size Effect on the Charging Free Energy of Protein in Explicit Solvent

Toru Ekimoto,[†] Nobuyuki Matubayasi,[§,∥] and Mitsunori Ikeguchi*[,†]

[†]Graduate School of Medical Life Science, Yokohama City University, 1-7-29 Suehiro-cho, Tsurumi-ku, Yokohama 230-0045, Japan
[§]Division of Chemical Engineering, Graduate School of Engineering Science, Osaka University, Toyonaka, Osaka 560-8531, Japan
[∥]Elements Strategy Initiative for Catalysts and Batteries, Kyoto University, Katsura, Kyoto 615-8520, Japan

ABSTRACT: The finite-size effect in periodic system is examined for the charging free energy of protein in explicit solvent over a variety of charged states. The key to the finite-size correction is the self-energy, which is defined as the interaction energy of the solute with its own periodic images and the neutralizing background. By employing the thermodynamic-integration method with systematically varied sizes of the unit cell of molecular dynamics (MD) simulations, we show for ubiquitin that the self-energy corrects the finite-size effect on the charging free energy within 1 kcal/mol at total charges of $-5e$, $-1e$, neutral, and $+1e$ and within 5 kcal/mol even for a highly charged state with $+8e$. We then sought the additional correction from the solvation effect using the numerical solution to the Poisson equation of the protein with implicit solvent. This correction reduces the cell-size dependence of the charging free energy at $+8e$ to 3 kcal/mol and is well expressed as the self-energy divided by the dielectric constant of solvent water.



INTRODUCTION

The electrostatic interaction is a key component for determining protein stability in solvent. The solvent effect on the stability is governed by the free energy of solvation. The electrostatic component of the free energy is quantified by the charging free energy, which is the free energy change for turning on the electrostatic interaction of protein with solvent (water). In molecular dynamics (MD) simulations, one difficulty lies in the treatment of the electrostatic interaction over long ranges. This difficulty manifests itself in particular for protein systems because, in typical setups, the total simulation system is comparable in size to the protein molecule.

The finite-size effect refers to the cell size dependence of the charging free energy at different cell sizes. This effect is one of the challenges of MD simulations, and it needs to be corrected to compare a computed energetic quantity with experimental results. Indeed, the finite-size effect has been reported with possible artifacts and corrections for a variety of cases including the charging free energy,[1−8] ion−ion interactions,[9] the binding free energy,[10] charge mutations,[11] orientation properties,[12] pressure,[13] stabilities of small peptide,[14] and membrane−water systems.[15] Hummer et al. proposed a correction scheme with the self-energy, which is defined as the interaction energy of the solute with its own periodic images and the neutralizing background. They showed for ions[16,17] and small molecule[18] at low total charges that the self-energy essentially cancels the finite-size effect.

Recently, the explicit treatment of pH conditions becomes possible in MD simulations with the development of the constant-pH method.[19] Indeed, pH controls the charged state of protein through acid dissociation constants (pK_a) over the residues, and its effect is desirable to be incorporated into atomistic simulations of protein. One of the important steps in the constant-pH method is the evaluation of the difference in the charging free energy of protein over variable charged states; a constant-pH simulation may then be problematic when the charging free energy has large error. In fact, the finite-size effect in the charging free energy is not negligible at all, as will be seen in the Results and Discussion section. Furthermore, the finite-size effect is considered to be roughly quadratic in charge. This is because the free energy depends quadratically on charge in Born's correction with the dielectric continuum model and Hummer et al.'s scheme described above. These methods are often observed to effectively remove the finite-size effect,[16−18] which implies in turn that the charging free energy may involve a quadratic dependence on the charge if not corrected. When a non-negligible correction exhibits a strong dependence on charge, its appropriate treatment is certainly necessary for conducting an MD with variable charged states.

In the present paper, we examine the finite-size effect on protein with explicit solvent over a variety of charged states. Using ubiquitin as a test protein, we calculate the charging free energy by the thermodynamic-integration method[20] with systematically varied sizes of the MD unit cell. In order to emphasize the effect of the protein charge in the finite-size effect, the structure of ubiquitin is fixed, and only the (total) charge is varied at $-5e$, $-1e$, neutral, $+1e$, and $+8e$, corresponding to pH conditions of 3, 6, 7, 10, and 11, respectively.

We base our analysis on the finite-size correction scheme formulated by Hummer et al.[16−18] Still, there are nontrivial issues when treating protein system. One issue is the system size. In Hummer et al.'s original treatment of ion and small

Received: September 18, 2014
Published: December 9, 2014

DOI: 10.1021/ct5008394
J. Chem. Theory Comput. 2015, 11, 215−223

molecule, the MD unit cell can be much larger in size than the solute diameter. In protein MD, in contrast, the system size is usually limited to twice the size of the protein due to the computational cost; the MD unit cell is typically prepared by placing several layers of water molecules around the protein solute. Another issue is the charge distribution. Hydrophilic residues tend to be at the surface of the protein, which means that the (partial) charges within the protein distribute over a spatial region that is comparable in size to the protein molecule and also the MD unit cell. The performance of the finite-size correction scheme is thus examined for spatially extended and diffuse distribution of charges.

When the solute of interest is ionic, the simulated system can be kept charge-neutral by either placing counterions or employing a neutralizing background.[1,22] The configuration of the counterions relative to the solute varies with the time evolution of MD and may require a long simulation to obtain convergent results for energetics, especially at low concentrations. Use of a neutralizing background is preferable in terms of computational time, in contrast, because the interaction with the neutralizing background is constant at a given solute configuration and cell volume. In fact, the neutralizing background is a built-in component of the Ewald and particle-mesh Ewald (PME) methods, and it is not required in these lattice-sum methods to explicitly place counterions for the purpose of neutralizing an MD system. We thus employ a neutralizing background in the present work to examine the finite-size effect on the charging free energy.

THEORETICAL BACKGROUND

The target quantity of the present development is the charging free energy at the limit of large system size $\Delta\mu(\infty)$; the finite-size correction[16−18] is necessary to determine it from the value $\Delta\mu(L)$ in a finite system with size L. The link between $\Delta\mu(\infty)$ and $\Delta\mu(L)$ is established by the thermodynamic cycle shown in Figure 1, and $\Delta\mu(\infty)$ is obtained from $\Delta\mu(L)$ as

$$\Delta\mu(\infty) = \Delta\mu(L) + \Delta E_{\mathrm{self}}(L) - \Delta F_{\mathrm{inter}}(L) \tag{1}$$

where L is the length of the cell when it is cubic and ΔE_{self} and $\Delta F_{\mathrm{inter}}$ are the correction terms described below. The finite-size effect originates from the intramolecular contribution ΔE_{self} of solute and the intermolecular contribution $\Delta F_{\mathrm{inter}}$ between solute and solvent (solvation effect), which reflect the difference of the electrostatic (free) energy in the finite system from that in the infinite system. In typical MD simulations, the finite system is constructed with an MD unit cell under the periodic boundary condition, and the electrostatic energy is handled by the Ewald (or PME) potential. In contrast, in the case of the infinite system, the periodic boundary condition is not imposed, and the electrostatic energy is handled by the bare Coulomb potential $(1/r)$. The self-energy ΔE_{self} is defined as the difference of the electrostatic energy of the solute in vacuo placed in a finite system with size L from that in the infinite system ($L = \infty$). It is equal to the interaction energy of the solute with its own periodic images and the neutralizing background. The intermolecular contribution $\Delta F_{\mathrm{inter}}$ is defined as

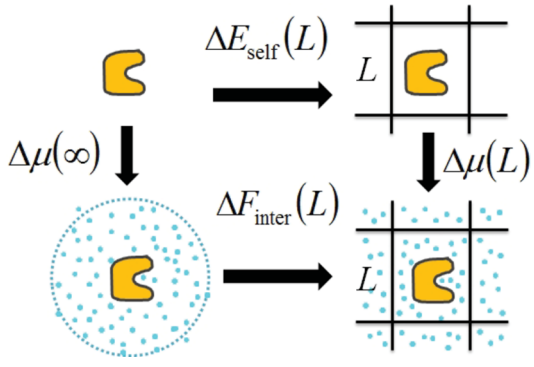


Figure 1. Schematic picture of the thermodynamic cycle (eq 1) connecting the charging free energy at the limit of the large system size and that in a finite system with size L. The orange object is the solute, and the blue dots are solvent molecules. The left and right portions represent the system in the thermodynamic limit and a finite system with periodicity, respectively, and the upper and lower portions have the solute in vacuo and in solution, respectively. The interaction involved in the figure refers only to the electrostatic component, and the van der Waals component is not included.

$$\Delta F_{\mathrm{inter}}(L) = F_{\mathrm{inter}}(L) - F_{\mathrm{inter}}(\infty) \tag{2}$$

where $F_{\mathrm{inter}}(\infty)$ and $F_{\mathrm{inter}}(L)$ are the electrostatic free energies of the solute placed in the infinite system ($L = \infty$) and the finite system with size L, respectively. However, these quantities are difficult to obtain in practice in all-atom computation. To resolve the difficulty, Hummer et al. proposed the use of the implicit solvent model.[16] In their method, $\Delta F_{\mathrm{inter}}$ is approximately given by

$$\Delta F_{\mathrm{imp}}(L) = F_{\mathrm{imp}}(L) - F_{\mathrm{imp}}(\infty) \tag{3}$$

where $F_{\mathrm{imp}}(\infty)$ and $F_{\mathrm{imp}}(L)$ are the electrostatic free energies of the solute in implicit solvent placed in the infinite and finite systems, respectively, and are calculated by the continuum dielectric model. By replacing $\Delta F_{\mathrm{inter}}$ by ΔF_{imp}, eq 1 is approximately given by

$$\Delta\mu(\infty) \sim \Delta\mu(L) + E_{\mathrm{self}}(L) - \Delta F_{\mathrm{imp}}(L) \tag{4}$$

CORRECTION METHOD

The charging free energy $\Delta\mu$ is defined by

$$e^{-\Delta\mu/k_BT} = \frac{\int d\Gamma\, e^{-(U_u+U_{uv}^{elec}+U_{uv}^{vdW}+U_v)/k_BT}}{\int d\Gamma\, e^{-(U_u+U_{uv}^{vdW}+U_v)/k_BT}} \tag{5}$$

where k_B is the Boltzmann constant, T is temperature, Γ denotes the system configuration, and the intramolecular energy of the solute and the solvent−solvent energy are U_u and U_v, respectively. Here, the electrostatic energy between the solute and solvent molecules is given by U_{uv}^{elec}, and the van der Waals contribution is U_{uv}^{vdW}. We employ the thermodynamic-integration (TI) method[20] to evaluate $\Delta\mu$ by introducing a set of intermediate states in which the partial charges of the solute are linearly scaled by the coupling parameter λ. The potential energy of the system is then $U_u + \lambda U_{uv}^{elec} + U_{uv}^{vdW} + U_v$, where U_{uv}^{elec} is the solute−solvent electrostatic energy at $\lambda = 1$, and eq 5 reduces to

$$\Delta\mu = \int_0^1 d\lambda \langle U_{uv}^{elec}(\Gamma)\rangle_\lambda \tag{6}$$

where $\langle A\rangle_\lambda$ is the ensemble average at each intermediate state that has the potential energy $U_u + \lambda U_{uv}^{elec} + U_{uv}^{vdW} + U_v$. The integrand of eq 6 was calculated by MD simulation for each intermediate state, and the numerical integration over λ was carried out by the trapezoidal rule.

DOI: 10.1021/ct5008394
J. Chem. Theory Comput. 2015, 11, 215−223

The self-energy ΔE_{self} was numerically calculated by subtraction of the bare Coulomb potential from the Ewald one:

$$\Delta E_{self}(L) = \lim_{r\to 0}\left[\frac{1}{2}\sum_{\alpha} q_\alpha^2 \varphi_{EW}(\mathbf{r}, L) - \frac{1}{2}\sum_{\alpha}\frac{q_\alpha^2}{|\mathbf{r}|}\right] + \left[\frac{1}{2}\sum_{\alpha,\beta} q_\alpha q_\beta \varphi_{EW}(\mathbf{r}_{\alpha\beta}, L) - \frac{1}{2}\sum_{\substack{\alpha,\beta \\ \alpha\neq\beta}}\frac{q_\alpha q_\beta}{|\mathbf{r}_{\alpha\beta}|}\right] \quad (7)$$

where α and β represent atoms within the solute molecule, q_i is the point charge on the i-th atom, and $\mathbf{r}_{\alpha\beta}$ is the vector connecting the α and β atoms. When the unit cell is cubic, the (smooth particle mesh) Ewald potential $\varphi_{EW}(\mathbf{r}_{\alpha\beta}, L)$[21] is given by

$$\begin{aligned}\frac{1}{2}\sum_{\alpha,\beta} q_\alpha q_\beta \varphi_{EW}(\mathbf{r}_{\alpha\beta}, L) &= \frac{1}{2}\sum_{\mathbf{n}}\sum_{\alpha,\beta} q_\alpha q_\beta \frac{\mathrm{erfc}(\kappa|\mathbf{r}_{\alpha\beta}+\mathbf{n}L|)}{|\mathbf{r}_{\alpha\beta}+\mathbf{n}|} \\ &+ \frac{1}{2\pi L^3}\sum_{\mathbf{m}\neq 0}\frac{\exp(-\pi^2\mathbf{m}^2/\kappa^2)}{\mathbf{m}^2}B(\mathbf{m})\tilde{S}(\mathbf{m})\tilde{S}(-\mathbf{m}) \\ &- \frac{\pi}{2\kappa^2 L^3}\left(\sum_{\alpha} q_\alpha\right)\left(\sum_{\beta} q_\beta\right)\end{aligned} \quad (8)$$

where $\mathrm{erfc}(x)$ is the complementary error function, κ is the screening factor, $\mathbf{n}$ is a vector consisting of integers, $\boldsymbol{L}$ is the length of the unit cell, $\mathbf{m}$ is the reciprocal lattice vector, $\tilde{S}(\mathbf{m})$ is the discrete Fourier transform of the structure factor $S(\mathbf{m}) = \sum_\alpha q_\alpha \exp(i\mathbf{m}\cdot\mathbf{r}_\alpha)$ approximately calculated by the cardinal B-spline interpolation of the complex exponential with grid charges mapped on discrete lattices,[21] $B(\mathbf{m})$ is a normalizing factor of the interpolation function,[21] and the third term of the right-hand side ensures that the average of the Ewald potential over the simulation cell is zero.[22,23]

Following the procedures of Hummer et al.,[17,18] we examine the performance of the correction scheme in two steps. First, only the self-energy correction is applied:

$$\Delta\mu_{corr} = \Delta\mu + \Delta E_{self} \quad (9)$$

where $\Delta\mu$ and ΔE_{self} are given by eqs 6 and 7, respectively. To evaluate the performance of eq 9 at each intermediate state of the TI, the λ dependence of the self-energy is also considered. When the partial charge of the solute is linearly scaled by λ, the self-energy at an intermediate state with λ is $\lambda^2\Delta E_{self}$, where ΔE_{self} is the self-energy at $\lambda = 1$ and is given by eq 7. The derivative of the scaled self-energy with respect to λ is then $2\lambda\Delta E_{self}$, and eq 9 can be written as

$$\Delta\mu_{corr} = \int_0^1 d\lambda(\langle U_{uv}^{elec}\rangle_\lambda + 2\lambda\Delta E_{self}) \quad (10)$$

where ΔE_{self} is the value at $\lambda = 1$. When the integrand $(\langle U_{uv}^{elec}\rangle_\lambda + 2\lambda\Delta E_{self})$ is independent of the cell size, $\Delta\mu_{corr}$ is also independent of cell size.

After the self-energy correction is examined with eq 9, the correction from the solvation effect is applied:

$$\Delta\mu_{full} = \Delta\mu_{corr} - \Delta F_{imp} \quad (11)$$

where ΔF_{imp} is given by eq 3, and is obtained by the numerical solution of the Poisson equation. The electrostatic free energy F_{imp} is calculated as $^1/_2\sum_i q_i\phi_i$, where the factor $^1/_2$ represents the linear response to the charging, and ϕ_i is the electrostatic potential at the i-th atom. The electrostatic potential is obtained both in a finite system with periodic boundary conditions and in an infinite system. All the calculations were conducted by the PBEQ module[24−26] in the program package CHARMM.[27] The dielectric constant was set to 92 so that it is close to that of the TIP3P water model (92 ± 5).[28−30] Although Hummer et al. proposed an analytical form of the ΔF_{imp} correction,[17,18] we did not apply it to the protein systems since the form was obtained for small and spherical ions.

COMPUTATIONAL DETAILS

Ubiquitin (PDB id: 1UBQ[31]) was employed as a test protein. The resolution is 1.8 Å, and modeling of missing residues was not necessary. The missing hydrogen atoms were inserted by the MMTSB tool[32] in the program package CHARMM, and the N- and C-termini were set to NH_3^+ and COO^-, respectively.

The simulations were conducted by placing the protein at the center of a cubic cell and adding layers of water molecules. Four different thicknesses of layers were prepared: 15, 20, 30, and 35 Å from the long axis of the protein. These cells are about 1.6, 1.9, 2.3, and 2.5 times larger than the protein molecule, and are hereafter called 15-layer, 20-layer, 30-layer, and 35-layer systems, respectively. The cell lengths are summarized in Table 1. In addition to the above four systems, 10-layer and 45-layer systems were also simulated for the +8e charged state; the cells were prepared with the layer thickness of 10 and 45 Å from the long axis of the protein and is about 1.4 and 3 times larger than the protein, respectively. The number of water molecules in each cell was set so that the water density is near 0.94 g/cm^3. Fifty-eight crystal water molecules are contained in the PDB structure and were used in the MD simulation. They were observed to move freely between the PDB sites and the protein outside, and thus, their initial positions have no effect on the results presented below.

Table 1. Cell Size and PME Mesh Size

	10-layer	15-layer	20-layer	30-layer	35-layer	45-layer
cell length [Å]	64.3	73.0	84.9	104.9	114.9	134.4
PME mesh size	$(60)^3$	$(70)^3$	$(80)^3$	$(96)^3$	$(108)^3$	$(126)^3$

The protein structure was prepared through two steps of minimizations in the 15-layer system without any restraint: the steepest descent method (2000 steps) as the first step and the conjugate gradient method (2000 steps) as the second step. After the minimizations, the structure of the protein was fixed. The water molecules were kept rigid both in the minimization and the MD simulation. To prepare the other systems, the minimized protein and water molecules within 5 Å of the surface were extracted. These extracted protein and water molecules were then immersed into the cell of interest, and water molecules were further added. Only the configurations of the water molecules were minimized by the steepest descent method (2000 steps) after adding water.

The net charge of the protein was modulated by the oxidation/reduction of charged residues. The modulated residues were chosen by the pK_a values calculated by the program H++.[33−36] We employed total charges of the protein

DOI: 10.1021/ct5008394
J. Chem. Theory Comput. 2015, 11, 215−223

of +8*e*, +1*e*, 0*e*, −1*e*, and −5*e* corresponding to pH = 3, 6, 7, 10, and 11, respectively. The net charge increased from neutral to +8*e* by protonating His68 and neutralizing Glu24, Glu51, C-terminal (COOH), Glu16, Glu64, Asp32, and Glu18. The net charge decreased to −5*e* by neutralizing the N-terminal (NH_2), Lys33, Lys6, Lys63, and Lys48. The neutralizations were handled by patches of the CHARMM22 force-field. Upon neutralization, the number of hydrogen atoms changed, while the other atoms were kept at the minimized structure. The structures with various net charges were immersed into the four (six when the charge is +8*e*) different layer systems.

All the MD simulations were performed using the MD program package MARBLE[37] with CHARMM22 force field[38,39] and TIP3P water model.[28] The isothermal−isobaric (*NPT*) ensemble was adopted with periodic boundary conditions. The temperature and pressure were set at 300 K and 1 atm. The rigid body simulation in the *NPT* ensemble was done by the extended-system approach[37,40] with the symplectic integrator[37,41] at 2 fs time step. The Lennard-Jones interaction was smoothly truncated by the switching function with the switching range 8−10 Å.[42] The electrostatic interaction was handled by the PME method.[21] The real-space cutoff, spline order, and the Ewald tolerance were 10 Å, 4, and 10^{-6}, respectively, and the reciprocal-space mesh size is listed in Table 1. No counterions were added in the solution systems since the neutralizing background involved in the Ewald potential (eq 8) mimics the effect of counterions.

The integrand of the thermodynamic-integration method was calculated from a series of snapshots generated by MD simulations. The integration of eq 6 was divided into 10 intermediate states with an equally spaced λ interval of 0.1; an interval of 0.05 was also examined at the +1*e* charged state in the 15-layer system as described in the first paragraph of the Results and Discussion section. At each intermediate state, a 2000-step minimization, a heating simulation from 10 to 300 K over 100 ps, and a 1 ns equilibration were consecutively executed before the production run. The sampling MD was done over 1 ns, and a snapshot was saved every 100 fs (1×10^4 snapshots in total). To examine the convergence of the integrand of eq 6, the block average was collected every 1000 snapshots, and its cumulative average (10 blocks in total) was analyzed.[43] The integration in eq 6 with respect to λ was conducted by the trapezoidal rule.

To see the contribution of the self-energy, furthermore, we also investigate the following proteins in the corresponding 10-layer systems. To address the effect of size, we treated trp-cage (PDB id: 1L2Y[45]) *Staphylococcus aureus* protein A (PDB id: 1BDD[46]), human lysozyme (PDB id: 2NWD[47]), *Staphylococcal* nuclease (SNase, PDB id: 3TP5[48]), HIV RNase H (PDB id: 3LP3[49]), and *Vibrio cholerae* FocA (VcFocA, PDB id: 3KLZ[50]); the first four are monomers and the last two are a dimer and a pentamer, respectively. The effect of charge was examined with the lead-dependent ribozyme (PDB id: 1LDZ[51]), the DNA complex of PhoB DNA-binding domain (PhoB-DNA, PDB id: 2Z33[52]), and the tRNA complex of *C. albicans* tRNA(His) guanylyltransferase (CaThg1-tRNA, PDB id: 3WC1[53]). The numbers of residues of these nine proteins are 20 (trp-cage), 60 (protein A), 130 (human lysozyme), 130 (SNase), 130 × 2 (RNaseH), 278 × 5 (VcFocA), 30 (ribozyme), 104 + 16 base pairs of DNA (PhoB-DNA), and 268 × 4 + 73 bases of RNA × 2 (CaThg1-tRNA). Missing residues in the N- and C-termini for SNase, RNaseH, VcFocA, and CaThg1-tRNA were not added, and the termini were set to NH_3^+ and COO^-, respectively. In the cases of SNase, VcFocA, and CaThg1, the missing residues in the loop region were capped by methylated C-terminus ($COOCH_3$) and acetylated N-terminus ($NCOCH_3$), respectively. Each of the solutes was immersed into a 10-layer system which was prepared with layer thickness of 10 Å from the long axis and was filled with water molecules so that the water density is near 0.94 g/cm^3. The cell lengths of these solution systems are 43.2 (trp-cage), 70.0 (protein A), 68.4 (human lysozyme), 69.4 (SNase), 79.7 (RNaseH), 110.8 (VcFocA), 75 (ribozyme), 89.9 (PhoB-DNA), and 131.6 Å (CaThg1-tRNA). In the solution system, the solute was minimized with the same procedures as for ubiquitin described above. The minimizations for ribozyme, PhoB-DNA, and CaThg1 were performed with CHARMM27 force field,[38,39,54,55] and the others were treated with CHARMM22.[38,39] After the minimizations, the structure of the solute was fixed, and the self-energy was calculated.

RESULTS AND DISCUSSION

First, the convergence behavior is examined. Figure 2 illustrates $\langle U_{uv}^{elec} \rangle_{\lambda=1}$ (integrand of eq 6) at the +8*e* charged state in the 15-layer system as a function of the simulation time of the sampling MD. The $\langle U_{uv}^{elec} \rangle_{\lambda=1}$ value in the 100 ps interval varies within ∼20 kcal/mol and the cumulative average is stable within 2 kcal/mol after 1 ns. The 95% error at 1 ns is 2 kcal/mol or about 0.1% in relative error. The errors at the other intermediate states are 0.8−3 kcal/mol. With the integration over λ, the 95% errors at the five charged states are 0.6−0.8 kcal/mol. This precision is sufficient for the following discussions and the sampling MD was thus restricted to 1 ns. The integrand in eq 6 is computed at discrete λ, and the effect of the discreteness is further examined for the +1*e* charged state in the 15-layer system.[56] The examination is carried out by comparing three schemes of integration. Two use the trapezoidal rule with λ intervals of 0.05 and 0.1 as described at the second last paragraph of the Computational Details section, and the third is an analytical integration with a fitting function. The difference in the integration values among the three schemes is less than the numerical error in Table 2. The number of intermediate states is thus sufficient for smooth integration.

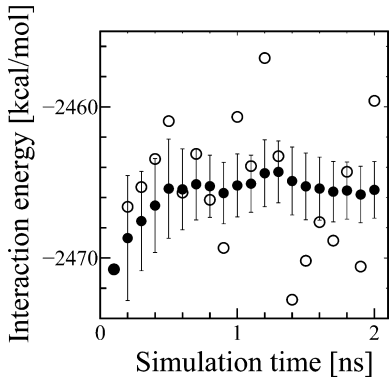


Figure 2. Solute−solvent interaction energy $\langle U_{uv}^{elec} \rangle_{\lambda=1}$ (integrand of eq 6) at the +8*e* charged state in the 15-layer system as a function of the simulation time of the sampling MD. The open and filled circles represent the averaged values over 100 ps intervals and the cumulative averages, respectively. The error bars are expressed at 95% confidence limit.

DOI: 10.1021/ct5008394
J. Chem. Theory Comput. 2015, 11, 215−223

Table 2. Charging Free Energy with the 95% Error

charge	10-layer	15-layer	20-layer	30-layer	35-layer	45-layer
neutral		−1231.7 ± 0.7	−1233.7 ± 0.8	−1235.0 ± 0.7	−1235.2 ± 0.7	
+1*e*		−1324.0 ± 0.8	−1327.9 ± 0.7	−1330.4 ± 0.8	−1331.5 ± 0.7	
+8*e*	−1061.6 ± 0.7	−1108.8 ± 0.6	−1149.2 ± 0.7	−1207.9 ± 0.7	−1229.9 ± 0.6	−1266.4 ± 0.7
−1*e*		−1298.9 ± 0.8	−1300.7 ± 0.7	−1302.9 ± 0.8	−1303.9 ± 0.8	
−5*e*		−1268.5 ± 0.7	−1284.9 ± 0.7	−1308.9 ± 0.7	−1317.3 ± 0.7	

We turn to the finite-size effect on the λ dependence of the integrand of eq 6. Figure 3 illustrates the λ dependence of

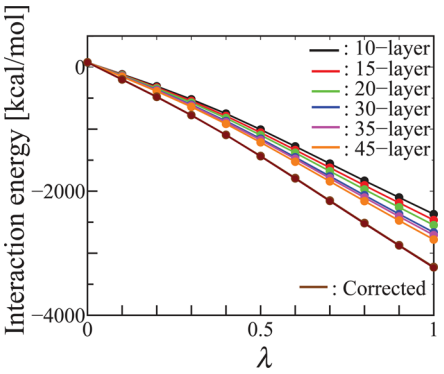


Figure 3. λ Dependence of the integrand of the thermodynamic-integration method $\langle U_{elec}\rangle_\lambda$ (eq 6) at the +8*e* charged state in the 10, 15, 20, 30, 35, and 45-layer systems. The 95% errors are smaller than the corresponding symbols, and the lines are drawn for an eye guide. The brown line refers to the corrected integrands (eq 10), and the lines from the different system sizes are not distinguishable within the resolution of the figure.

$\langle U_{uv}^{elec}\rangle_\lambda$ for the +8*e* charged state in the six different cell sizes. It is evident that the curves for the different cell sizes do not overlap with one another. This causes the system-size dependence of the charging free energy $\Delta\mu$ through the integration in eq 6. The $\Delta\mu$ values for the charged states examined in the present work are listed in Table 2. The dependence of $\Delta\mu$ on the cell size is larger when the net charge of the protein is larger. The cell-size dependence amounts to more than 100 kcal/mol when the net charge is −5*e* or +8*e*. A common tendency is for $\Delta\mu$ to increase in magnitude with the cell size, as observed for small-ion systems.[17]

In the Correction Method section, the correction to the computed free energy of charging is described in two steps. The first step is the self-energy correction given by eq 9, and the corrected free-energy value $\Delta\mu_{corr}$ is listed in Table 3 with the self-energy (eq 7); the error is common to the corresponding entry in Table 2. At the total charges of −5*e*, −1*e*, neutral, and +1*e*, the $\Delta\mu_{corr}$ for the different layers are in agreement within errors. At the highly charged state with +8*e*, $\Delta\mu_{corr}$ exhibits a cell size dependence of ∼5 kcal/mol. Given that the cell size dependence at +8*e* is more than 100 kcal/mol in Table 2, Table 3 shows that the self-energy correction essentially cancels the finite-size effect. To understand the cancellation of the cell size dependence, the self-energy correction is analyzed for the integrand of $\Delta\mu$. The integrand at each λ is corrected with eq 10, and the corrected integrand for the +8*e* charged state in the six different layers is illustrated in Figure 3. At each λ, the corrected integrands for the different cell sizes overlap within errors. This leads to the cancellation of the cell-size dependence by the self-energy correction.

The uncorrected free energy $\Delta\mu$ and the corrected free energy $\Delta\mu_{corr}$ are shown in Figure 4 against the cell size L. It is evident that the difference between $\Delta\mu$ and $\Delta\mu_{corr}$ reduces with the cell size. Note that the uncorrected and corrected values should be in agreement with each other in the limit of large system size. To examine the limiting behavior, in Figure 4, we also plot $\overline{\Delta\mu_{corr}} - \Delta E_{self}(L)$, where $\overline{\Delta\mu_{corr}}$ is the average of $\Delta\mu_{corr}$ computed with explicit solvent and listed in Table 3, and $\Delta E_{self}(L)$ is the self-energy at cell size L. This plot is shown since the $\Delta\mu$ calculation is prohibitive at very large L while $\Delta E_{self}(L)$ can be obtained essentially at any L. We cannot calculate the uncorrected free energy $\Delta\mu$ at large L, so instead we examine $\overline{\Delta\mu_{corr}} - \Delta E_{self}(L)$ to see how much the uncorrected free energy $\Delta\mu$ will change until the limit of large system size is reached. According to Figure 4, $\Delta\mu$ will vary by ∼80 and ∼250 kcal/mol at the charged states of −5*e* and +8*e*, respectively, with the change of the cell size from L = ∼100 Å to ∞. Although L = ∼100 Å is considered a large system in the state-of-art MD, it is still far from infinity with respect to the charging free energy.

It is further notable in Figure 4 that when the cell size L is not very large, the uncorrected free energy $\Delta\mu$ is more negative at the charged state of −5*e* than at +8*e*, while their order switches for the corrected value $\Delta\mu_{corr}$. This points to a possibility that the comparison of the stabilities over variable charged states is erroneous when the charging free energy is left uncorrected. The correction is necessary, particularly when the solute carries a large charge and different charged states are to be compared.

As noted in the Theoretical Background section, the finite-size correction is also necessary for the solvation effect. Indeed, the thermodynamic cycle closes only by incorporating the correction from the solvation effect $\Delta F_{inter}(L)$. In the present work, $\Delta F_{inter}(L)$ was approximated as the value $\Delta F_{imp}(L)$ (eq 3) from the continuum model described at the end of the

Table 3. Corrected Charging Free Energy with the Self-Energy in Parentheses

charge	10-layer	15-layer	20-layer	30-layer	35-layer	45-layer
neutral		−1236.6 (−4.9)	−1236.7 (−3.1)	−1236.5 (−1.5)	−1236.3 (−1.1)	
+1*e*		−1337.6 (−13.6)	−1338.2 (−10.3)	−1337.3 (−6.9)	−1337.4 (−5.9)	
+8*e*	−1492.3 (−430.7)	−1490.4 (−381.6)	−1489.9 (−340.6)	−1488.0 (−280.1)	−1486.8 (−256.9)	−1486.9 (−220.4)
−1*e*		−1309.4 (−10.5)	−1309.0 (−8.3)	−1308.7 (−5.9)	−1309.1 (−5.2)	
−5*e*		−1418.8 (−150.3)	−1419.6 (−134.7)	−1419.3 (−110.5)	−1418.7 (−101.3)	

DOI: 10.1021/ct5008394
J. Chem. Theory Comput. 2015, 11, 215−223

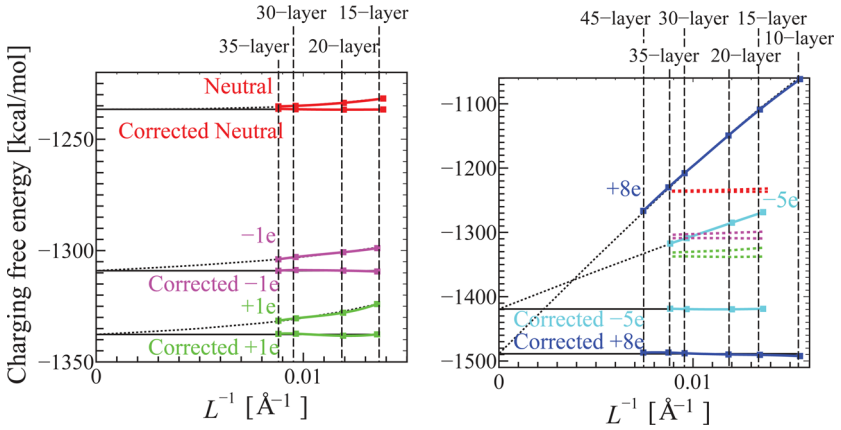


Figure 4. Cell-size dependence of the uncorrected (Table 2) and corrected charging free energies (Table 3) at the charged states of neutral and $\pm 1e$ (left panel) and of $+8e$ and $-5e$ (right panel). The solid lines connecting the numerical data are drawn for an eye guide. The straight, solid black lines represent the averages of $\Delta\mu_{corr}$ over the different layer systems $\overline{\Delta\mu_{corr}}$. The curved, dotted black lines refer to $\overline{\Delta\mu_{corr}} - \Delta E_{self}$, where ΔE_{self} is the self-energy (eq 7). In the right panel, the short dashed lines in red, pink, and green represent the same data as the solid lines in the corresponding colors in the left panel.

Correction Method section. The $\Delta F_{imp}(L)$ value is listed in Table 4 and is negligible when the protein charge is small. At the charged state of $-5e$, $\Delta F_{imp}(L)$ is less than ~2 kcal/mol and depends on L within ~1 kcal/mol. The correction with $\Delta F_{imp}(L)$ is thus not effective since its dependence on L is buried within the error shown in Table 2. In contrast, at the highly charged state of $+8e$, $\Delta F_{imp}(L)$ reduces the finite-size effect that persists after the self-energy correction. Figure 5 compares the free energy at $+8e$ corrected only with the self-energy $\Delta\mu_{corr}(L)$ (eq 9) and the free energy also corrected with the solvation effect $\Delta\mu_{full}(L)$ (eq 11). The $\Delta\mu_{full}(L)$ values are observed to be more positive than $\Delta\mu_{corr}(L)$ at all L, and the $\Delta\mu_{full}(L)$ values at the six different cell sizes are in agreement within ~3 kcal/mol. The reduced dependence of $\Delta\mu_{full}(L)$ on L shows that $\Delta F_{imp}(L)$ is an effective representation of the solvation effect. At the $+8e$ charged state, for example, $\Delta\mu_{corr}(L)$ in the 10-layer system deviates from that in the 30-layer by ~4 kcal/mol, and this deviation is half canceled by $\Delta F_{imp}(L)$ evaluated by the continuum model. This result is useful since 10 Å is a typical cutoff length in protein MD with explicit solvent and the minimum size of the layer is determined by the cutoff distance. These indicate that the correction from the solvation effect is effective to further reduce the finite-size effect at the highly charged states and is necessary especially at small cell systems.

It is then possible to use a simplified scheme of $\Delta F_{inter}(L)$ evaluation, and we will examine an alternative approach to $\Delta F_{inter}(L)$. The finite-size effect of the solvation effect originates from the modification of the solvation free energy due to perturbation of the solvent polarization affected by the interaction with periodic images of the solute.[2,4,17] This modification of the solvation free energy was evaluated analytically[2,4,17] in a single ion system with implicit solvent and was expressed in a Taylor expansion in terms of L^{-1} and ϵ^{-1}, where ϵ is the dielectric constant of the solvent. In the present work, this expression was not used, and we approximate $\Delta F_{inter}(L)$ by an analog of the leading term of the expansion as

$$\Delta F_{inter}(L) \sim \Delta F_{alter}(L) = \frac{\Delta E_{self}(L)}{\epsilon} \qquad (12)$$

where $\epsilon = 92$ (TIP3P) is used in the following calculations, and $\Delta E_{self}(L)$ is the self-energy (eq 7). To see the performance of eq 12, Figure 6 shows a correlation between $\Delta F_{alter}(L)$ and $\Delta F_{imp}(L)$ with the implicit solvent model (eq 3) for the $+8e$ charged state. The agreement between the two is evident, and $\Delta F_{alter}(L)$ can be used in place of $\Delta F_{imp}(L)$. Substituting $\Delta F_{alter}(L)$ into eq 1, the free energy corrected both for the self-energy and the solvation effect is approximated as

Table 4. Correction Due to the Solvation Effect As Evaluated from the Continuum Model Described at the End of the Correction Method Section

charge	10-layer	15-layer	20-layer	30-layer	35-layer	45-layer
neutral		−0.1	−0.1	0.0	0.0	
$+1e$		−0.1	−0.1	0.0	0.0	
$+8e$	−4.6	−4.1	−3.6	−3.0	−2.7	−2.3
$-1e$		−0.2	−0.2	−0.1	−0.1	
$-5e$		−2.1	−1.8	−1.4	−1.2	

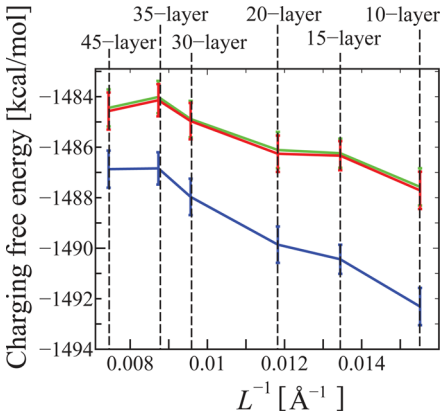


Figure 5. Cell-size dependence of the corrected charging free energy at the $+8e$ charged state in the six different cell sizes. The blue and red plots represent the charging free energy corrected only with the self-energy (eq 9) and the free energy also corrected with the solvation effect (eq 11), respectively. The green corresponds to the charging free energy corrected with the alternative approach to the solvation effect (eq 13). The lines are drawn for an eye guide.

DOI: 10.1021/ct5008394
J. Chem. Theory Comput. 2015, 11, 215−223

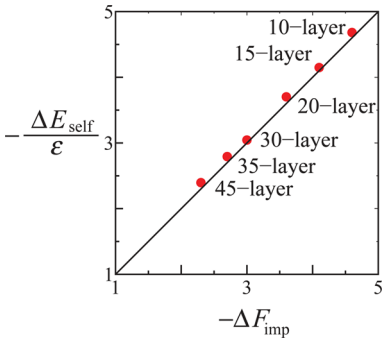


Figure 6. Correlation for the correction of the solvation effect between the continuum model $-\Delta F_{imp}$ (eq 3) and the alternative approach $-\Delta E_{self}/\epsilon$ (eq 12) at the $+8e$ charged state for the six cell sizes in the unit of kcal/mol. The correlation coefficient is 0.99. The diagonal line represents the ideal case of $\Delta E_{self}/\epsilon = \Delta F_{imp}$.

$$\Delta\mu_{full} = \Delta\mu + \left(1 - \frac{1}{\epsilon}\right)\Delta E_{self} \quad (13)$$

and Figure 5 shows eq 13 for the $+8e$ charged state at different cell sizes. Although eq 13 has been proposed in previous studies,[2,4,17,18] in fact, it was not applied to protein systems, as was so for the self-energy correction. Note that eq 12 is a simplified scheme of the correction of the solvation effect and is advantageous from computational standpoint since any calculations of the Poisson equation are not necessary.

The finite-size effect is further examined by the comparison between the present calculation with TI and the linear response approximation (LRA)[18,57,58] to the charging free energy. As shown in Figure 3, the λ dependence of the integrand $\langle U_{uv}^{elec}\rangle_\lambda$ is concave even after the self-energy correction. This indicates that the LRA provides a free energy that is too negative compared to the (exact) value from TI. The deviation from the TI value was ∼10% at the $+8e$ charged state for the simplest LRA given by $^1/_2\langle U_{uv}^{elec}\rangle_{\lambda=1}$. When the $\lambda = 0$ information is incorporated and the LRA is modified to $^1/_2\langle U_{uv}^{elec}\rangle_{\lambda=0} + (1/2)\langle U_{uv}^{elec}\rangle_{\lambda=1}$, the deviation reduces to ∼7%.

When the LRA is left uncorrected, the stability order of the charging free energy among the different charged states switches at all the cell sizes examined, as is so for the uncorrected TI in Figure 4. The L dependence of the order is common to that of TI when the modified LRA is used, in contrast, it is changed when the simplest LRA is used due to the coarse approximation.[59] In both cases of the simplest and modified LRAs, the self-energy correction reduces the L dependence of the charging free energy, and its performance is comparable to the case of TI. The stability order of the corrected LRA among the charged states, in addition, is common to those of the corrected TI. This indicates that the self-energy correction is effective and necessary to compare the charging free energies for different charged states even though the free energy is estimated by an approximate scheme.

Finally, the significance of the self-energy correction is discussed for a variety of proteins and complexes in Table 5. The number of residues in the solutes is between 20 and 1390, and the net charge is between $-144e$ and $+25e$; DNA and RNA have extremely negative charges. Our key conclusion is that the self-energy essentially cancels the finite-size effect, and thus the finite-size effect may be inferred from the value of the self-energy. As described at the end of the Computational Details section, the 10-layer system was employed to meet a practical requirement that a minimal cell size be used due to the limitation of computational cost. The self-energy in the 10-layer system is summarized in Table 5 and is in the range of 10^3 to 10^5 kcal/mol at extremely highly charged states. This strongly suggests that especially at highly charged states, the finite-size effect is not negligible and needs to be corrected in energetic analyses.

Table 5. Net Charge and the Self-Energy for Various Proteins and Complexes

	charge	self-energy [kcal/mol]
Trp-cage	$+1e$	−14.3
protein A	$-2e$	−31.7
human lysozyme	$+9e$	−491.1
SNase	$+5e$	−150.3
RNaseH	$-4e$	−85.8
VcFocA	$+25e$	−2129.3
ribozyme	$-29e$	−4782.9
PhoB-DNA	$-26e$	−3270.8
CaThg1-tRNA	$-144e$	−59786.5

CONCLUSIONS

The finite-size effect on the charging free energy of proteins has been examined with explicit solvent over a variety of charged states. We have calculated the charging free energy of ubiquitin by the thermodynamic-integration (TI) method with systematically varied sizes of the MD unit cell. Without any correction, the TI value varies by ∼160 kcal/mol as the cell length is increased to twice the size of that typically employed in protein MD. It has even been observed that the stability order quantified by the free energy may be reversed for two highly charged states. In order to correct the finite-size effect, we have adopted the correction scheme formulated by Hummer et al.[16−18] and assessed its performance. The correction consists of two parts: self-energy and solvation effect. The self-energy is defined as the interaction energy of the solute with its own periodic images and a neutralizing background. For the protein systems treated in the present work, the self-energy correction has been shown to reduce the finite-size effect to ∼1 kcal/mol. At the highest charged state of $+8e$, the performance deteriorates with ∼5 kcal/mol of the finite-size effect.

The correction for the solvation effect is further necessary in principle to close the thermodynamic cycle. We have found that the finite-size effect is essentially corrected by the self-energy. The correction of the solvation effect is effective at highly charged states, and we have also examined a simplified scheme using the self-energy and the dielectric constant of solvent. The present work has demonstrated the importance of the finite-size correction for the charging free energy of protein. The correction is especially important when different charged states are to be compared such as in a constant-pH MD; it will be desirable to conduct this calculation with the finite-size correction.

AUTHOR INFORMATION

Corresponding Author

*E-mail: ike@tsurumi.yokohama-cu.ac.jp.

Notes

The authors declare no competing financial interest.

DOI: 10.1021/ct5008394
J. Chem. Theory Comput. 2015, 11, 215−223

■ ACKNOWLEDGMENTS

This work is supported by the Platform for Drug Design, Informatics and Structural Life Science from the Ministry of Education, Culture, Sports, Science, and Technology of Japan (MEXT) to M.I.; by the JSPS (Japan Society for the Promotion of Science) Grants-in-Aid for Scientific Research (Nos. 23651202 and 26240045 to N.M. and No. 25291036 to M.I.); by the Elements Strategy Initiative for Catalysts and Batteries (MEXT) to N.M.; by Computational Materials Science Initiative, Theoretical and Computational Chemistry Initiative, and HPCI System Research Project (Project IDs: hp120093, hp130022, hp140156, and hp140214) of the Next-Generation Supercomputing Project to N.M.; by HPCI Strategic Programs for Innovative Research, Computational Life Science and Application in Drug Discovery and Medical Development (Project IDs: hp120309, hp130003, and hp140229) (MEXT) to M.I. We further thank Dr. Tsutomu Yamane of Yokohama City University for valuable discussions and comments.

■ REFERENCES

(1) Darden, T.; Pearlman, D.; Pedersen, L. G. *J. Chem. Phys.* **1998**, *109*, 10921.
(2) Figueirido, F.; Del Buono, G. S.; Levy, R. M. *J. Phys. Chem. B* **1997**, *101*, 5622.
(3) Sakane, S.; Ashbaugh, H. S.; Wood, R. H. *J. Phys. Chem. B* **1998**, *102*, 5673.
(4) Hünenberger, P. H.; McCammon, P. H. *J. Chem. Phys.* **1999**, *110*, 1856.
(5) Boresch, S.; Steinhauser, O. *J. Chem. Phys.* **2001**, *115*, 10793.
(6) Kastenholz, M. A.; Hünenberger, P. H. *J. Chem. Phys.* **2006**, *124*, 224501.
(7) Reif, M. M.; Hünenberger, P. H. *J. Chem. Phys.* **2011**, *134*, 144103.
(8) Lin, Y. L.; Aleksandrov, A.; Simonson, T.; Roux, B. *J. Chem. Theory Comput.* **2014**, *10*, 2690.
(9) Figueirido, F.; Del Buono, G. S.; Levy, R. M. *J. Chem. Phys.* **1995**, *103*, 6133.
(10) Rocklin, G. J.; Mobley, D. L.; Dill, K. A.; Hünenberger, P. H. *J. Chem. Phys.* **2013**, *139*, 184103.
(11) Morgan, B. R.; Massi, F. *J. Chem. Theory Comput.* **2010**, *6*, 1884.
(12) Smith, P. E.; Pettitt, B. M. *J. Chem. Phys.* **1996**, *105*, 4289.
(13) Bogusz, S.; Cheatham, T. E., III; Brooks, B. R. *J. Chem. Phys.* **1998**, *108*, 7070.
(14) Weber, W.; Hünenberger, P. H.; McCammon, J. A. *J. Phys. Chem. B* **2000**, *104*, 3668.
(15) Hub, J. S.; de Groot, B. L.; Grubmüller, H.; Groenhof, G. *J. Chem. Theory Comput.* **2014**, *10*, 381.
(16) Hummer, G.; Pratt, L. R.; García, A. E. *J. Phys. Chem.* **1996**, *100*, 1206.
(17) Hummer, G.; Pratt, L. R.; García, A. E. *J. Chem. Phys.* **1997**, *107*, 9275.
(18) Hummer, G.; Pratt, L. R.; García, A. E. *J. Phys. Chem. A* **1998**, *102*, 7885.
(19) Wallace, J. A.; Shen, J. K. *Methods Enzymol* **2009**, *466*, 455.
(20) Kollman, P. *Chem. Rev.* **1993**, *93*, 2395.
(21) Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. *J. Chem. Phys.* **1995**, *103*, 8577.
(22) Hummer, G.; Soumpasis, D. M. *J. Chem. Phys.* **1993**, *98*, 581.
(23) De Leeuw, S. W.; Perram, J. W.; Smith, E. R. *Proc. R. Soc. London A* **1980**, *373*, 27.
(24) Klapper, I.; Hagstrom, R.; Fine, R.; Sharp, K.; Honig, B. *Proteins* **1986**, *1*, 47.
(25) Nicholls, A.; Honig, B. *J. Comput. Chem.* **1991**, *12*, 435.
(26) The PBEQ module uses the iterative finite-difference relaxation algorithm to solve the linearlized Poisson−Boltzmann equation. The dielectric constant is set to 1 in the protein interior. The distance between grids is 0.5 Å, and the protein−solvent boundary is defined as the surface at contact with the spherical probe which mimics the water molecule and has a radius of 1.4 Å.
(27) Brooks, B. R.; Brooks, C. L., III; Mackerell, A. D., Jr.; Nilsson, L.; Petrella, R. J.; Roux, B.; Won, Y.; Archontis, G.; Bartels, C.; Boresch, S.; Caflisch, A.; Caves, L.; Cui, Q.; Dinner, A. R.; Feig, M.; Fischer, S.; Gao, J.; Hodoscek, M.; Im, W.; Kuczera, K.; Lazaridis, T.; Ma, J.; Ovchinnikov, V.; Paci, E.; Pastor, R. W.; Post, C. B.; Pu, J. Z.; Schaefer, M.; Tidor, B.; Venable, R. M.; Woodcock, H. L.; Wu, X.; Yang, W.; York, D. M.; Karplus, M. *J. Comput. Chem.* **2009**, *30*, 1545.
(28) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926.
(29) Lamoureux, G.; MacKerell, A. D., Jr.; Roux, B. *J. Chem. Phys.* **2003**, *119*, 5185.
(30) Khavrutskii, I. V.; Dzubiella, J.; McCammon, J. A. *J. Chem. Phys.* **2008**, *128*, 044106.
(31) Vijay-Kumar, S.; Bugg, C. E.; Cook, W. J. *J. Mol. Biol.* **1987**, *194*, 531.
(32) Feig, M.; Karaniolas, J.; Brooks, C. L., III *J. Mol. Graph. Model* **2004**, *5*, 377.
(33) H++. http://biophysics.cs.vt.edu/H++ (accessed Nov. 11, 2013).
(34) Anandakrishnan, R.; Aguilar, B.; Onufriev, A. V. *Nucleic Acids Res.* **2012**, *40*, W537.
(35) Myers, J.; Grothaus, G.; Narayanan, S.; Onufriev, A. *Proteins* **2006**, *63*, 928.
(36) Gordon, J. C.; Myers, J. B.; Forta, T.; Shoja, V.; Heath, L. S.; Onufriev, A. *Nucleic Acids Res.* **2005**, *33*, W368.
(37) Ikeguchi, M. *J. Comput. Chem.* **2004**, *25*, 529.
(38) MacKerell, A. D., Jr.; Feig, M.; Brooks, C. L., III. *J. Comput. Chem.* **2004**, *25*, 1400.
(39) MacKerell, A. D., Jr.; Bashford, D.; Bellott, M.; Dunbrack, R. L., Jr.; Evanseck, J. D.; Field, M. J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S.; JosephMccarthy, D.; Kuchnir, L.; Kuczera, K.; Lau, F. T. K.; Mattos, C.; Michnick, S.; Ngo, T.; Nguyen, D. T.; Prodhom, B.; Reiher, W. E., III; Roux, B.; Schlenkrich, M.; Smith, J. C.; Stote, R.; Straub, J.; Watanabe, M.; Wiórkiewicz-Kuczera, J.; Yin, D.; Karplus, M. *J. Phys. Chem. B* **1998**, *102*, 3586.
(40) Martyna, G. J.; Tobias, D. J.; Klein, M. L. *J. Chem. Phys.* **1994**, *101*, 4177.
(41) Miller, T. F., III; Eleftheriou, M.; Pattnaik, P.; Ndirango, A.; Newns, D.; Martyna, G. J. *J. Chem. Phys.* **2002**, *116*, 8649.
(42) Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. *J. Comput. Chem.* **1983**, *4*, 187.
(43) The computation of $\langle U_{uv}^{elec} \rangle$ was executed by the ERmod program[44] (http://sourceforge.net/projects/ermod).
(44) Sakuraba, S.; Matubayasi, N. *J. Comput. Chem.* **2014**, *35*, 1592.
(45) Neidigh, J. W.; Fesinmeyer, R. M.; Andersen, N. H. *Nat. Struct. Biol.* **2002**, *9*, 425.
(46) Gouda, H.; Torigoe, H.; Saito, A.; Sato, M.; Arate, Y.; Shimada, I. *Biochemistry* **1992**, *31*, 9665.
(47) Durek, T.; Torbeev, V. Y.; Kent, S. B. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 4846.
(48) Robinson, A. C.; Schlessman, J. L.; Garcia-Moreno, E. B. Unpublished work.
(49) Su, H. P.; Yan, Y.; Prasad, G. S.; Smith, R. F.; Daniels, C. L.; Abeywickrema, P. D.; Reid, J. C.; Loughran, H. M.; Kornienko, M.; Sharma, S.; Grobler, J. A.; Xu, B.; Sardana, V.; Allison, T. J.; Williams, P. D.; Darke, P. L.; Hazuda, D. J.; Munshi, S. *J. Virol.* **2010**, *84*, 7625.
(50) Waight, A. B.; Love, J.; Wang, D. N. *Nat. Struct. Mol. Biol.* **2010**, *17*, 31.
(51) Hoogstraten, C. G.; Legault, P.; Pardi, A. *J. Mol. Biol.* **1998**, *284*, 337.
(52) Yamane, T.; Okamura, H.; Ikeguchi, M.; Nishimura, Y.; Kidera, A. *Proteins* **2008**, *71*, 1970.
(53) Nakamura, A.; Nemoto, T.; Heinemann, I. U.; Yamashita, K.; Sonoda, T.; Komoda, K.; Tanaka, I.; Soll, D.; Yao, M. *Proc. Natl. Acad. Sci. U.S.A.* **2013**, *110*, 20970.

DOI: 10.1021/ct5008394
J. Chem. Theory Comput. 2015, 11, 215−223

(54) Foloppe, N.; MacKerell, A. D. *J. Comput. Chem.* **2000**, *21*, 86.
(55) MacKerell, A. D.; Banavali, N. *J. Comput. Chem.* **2000**, *21*, 105.
(56) We employed a polynomial function $A\lambda^4 + B\lambda^3 + C\lambda^2 + D\lambda + \langle U_{uv}^{elec}\rangle_{\lambda=0}$ to fit $\langle U_{uv}^{elec}\rangle_{\lambda}$. Among the cases treated in the present work, the choice of the integration scheme affects the +1*e* state in the 15-layer system most strongly, and the results for this case is illustrated in the following. The coefficients in the polynomial are $A = 637.14$, $B = -274.2$, $C = -1337.2$, $D = -1897.7$, and $\langle U_{uv}^{elec}\rangle_{\lambda=0} = 12.7$ in the units of kcal/mol, and the integration value is −1323.0 kcal/mol. The integration value is unchanged within 0.1 kcal/mol even when a higher-degree polynomial is used. When the trapezoidal rule is used, the integration value is −1323.5 ± 0.6 and −1324.0 ± 0.8 kcal/mol with the λ intervals of 0.05 and 0.1, respectively.
(57) Hummer, G.; Szabo, A. *J. Chem. Phys.* **1996**, *105*, 2004.
(58) Åqvist, J.; Medina, C.; Samuelsson, J.-E. *Protein Eng.* **1994**, *7*, 385.
(59) When the simplest LRA without the self-energy correction is employed, the order of the charging free energy is +1e < −1e < −5e < neutral < +8e at the 30-layer system or +1e < −1e < −5e < +8e < neutral at the 35-layer system. These orders are different from that calculated by the modified LRA and TI: +1e < −5e < −1e < neutral < +8e.

DOI: 10.1021/ct5008394
J. Chem. Theory Comput. 2015, 11, 215−223