Assessing the accuracy of approximate treatments of ion hydration based on primitive quasichemical theory

Benoît Roux[a] and Haibo Yu
Department of Biochemistry and Molecular Biology, University of Chicago, Chicago, Illinois 60637, USA

(Received 21 December 2009; accepted 6 May 2010; published online 15 June 2010)

Quasichemical theory (QCT) provides a framework that can be used to partition the influence of the solvent surrounding an ion into near and distant contributions. Within QCT, the solvation properties of the ion are expressed as a sum of configurational integrals comprising only the ion and a small number of solvent molecules. QCT adopts a particularly simple form if it is assumed that the clusters undergo only small thermal fluctuations around a well-defined energy minimum and are affected exclusively in a mean-field sense by the surrounding bulk solvent. The fluctuations can then be integrated out via a simple vibrational analysis, leading to a closed-form expression for the solvation free energy of the ion. This constitutes the primitive form of quasichemical theory (pQCT), which is an approximate mathematical formulation aimed at reproducing the results from the full many-body configurational averages of statistical mechanics. While the results from pQCT from previous applications are reasonable, the accuracy of the approach has not been fully characterized and its range of validity remains unclear. Here, a direct test of pQCT for a set of ion models is carried out by comparing with the results of free energy simulations with explicit solvent. The influence of the distant surrounding bulk on the cluster comprising the ion and the nearest solvent molecule is treated both with a continuum dielectric approximation and with free energy perturbation molecular dynamics simulations with explicit solvent. The analysis shows that pQCT can provide an accurate framework in the case of a small cation such as Li^+. However, the approximation encounters increasing difficulties when applied to larger cations such as Na^+, and particularly for K^+. This suggests that results from pQCT should be interpreted with caution when comparing ions of different sizes. © *2010 American Institute of Physics*. [doi:10.1063/1.3436632]

I. INTRODUCTION

Quasichemical theory (QCT) is an elegant statistical mechanical framework in which the influence of the solvent surrounding an ion can be partitioned into near and distant contributions. The general concept goes back to the method of "local configurations" formulated by Bethe, Guggenheim, Fowler, and Kirkwood.[1–5] More recently, QCT was further developed and extended by Pratt and co-workers in studies of the hydrophobic effect and ion solvation.[6–8]

In the case of ion solvation, QCT typically consists of defining a spherical inner shell region of radius R based on a distance criterion between the ion and the solvent molecule. Introducing step functions into the Boltzmann configurational integral to enforce this criterion, QCT then proceeds to carry out a strict book-keeping of the configurations with exactly n solvent molecules lying inside the inner shell region. Thermodynamic solvation properties can then be expressed in an expanded form as a finite sum over configurational integrals of clusters comprising the ion and the n inner solvent molecules under the average influence of the surrounding bulk.[6,7] When the radius of the inner region is chosen sufficiently small to correspond to the first coordination shell of the ion, the contribution from the clusters with a small number of solvent molecules dominates and the sum converges rapidly. This theoretical construct provides a rich framework to analyze and interpret the results from experimental measurements in the gas phase,[9] as well as from computer simulations.[10,11]

QCT adopts a particularly simple form if it is assumed that the clusters only undergo small thermal fluctuations around a well-defined energy minimum and are affected by the surrounding bulk solvent only in a mean-field sense based on a dielectric continuum approximation. The fluctuations can then be integrated out via a simple normal mode vibrational analysis, leading to closed-form expressions for the solvation free energy and the inner shell solvent occupancy probability. When combined with quantum chemistry electronic structure programs, this so- called "primitive" QCT (pQCT) provides a computationally tractable route to calculate the solvation free energy of ions in bulk liquids, while treating the interactions within the inner shell at an *ab initio* level. This strategy has been used to examine the hydration of Li^+, Na^+, and K^+.[12–14] pQCT has also been employed to discuss ion selectivity in proteins' binding sites,[15,16] but as it is a framework strictly designed to treat ion solvation in a bulk liquid, the significance of those studies is less clear.

While the overall numerical results from pQCT on ion hydration are encouraging, the accuracy of the approach and

[a]Electronic mail: roux@uchicago.edu.

its range of validity have not been fully characterized. It appears that the most probable coordination state predicted by pQCT for small ions such as Li^+ agrees well with the results from *ab initio* molecular dynamics (AIMD) simulations.[12] However, for larger ions such as K^+, pQCT seems to predict a lower optimal coordination number than the most probable coordination number observed in AIMD simulations.[14,17] Clearly, small ions bind water strongly, leading to tight and well-defined coordinated states that are relatively insensitive to the presence of the surrounding bulk solvent. Such small clusters are amenable to meaningful vibrational analysis. On the other hand, K^+ does not bind with water so strongly and its coordination shell is broad and very diffuse,[17] which appears to undermine the effectiveness of the cluster sum.

The situation concerning the accuracy of the absolute hydration free energies predicted by pQCT is less clear.[12–14] For Li^+, Rempe and Pratt reported a hydration free energy of −128 kcal/mol using a dielectric continuum for the outer region.[12] This was revisited by Merchant and Asthagiri,[11] who obtained −120.5 kcal/mol when treating the outer shell with a continuum dielectric and −115.1 kcal/mol with an outer shell of explicit TIP3P water molecules. For Na^+, Rempe and Pratt[13] reported a hydration free energy of −103 kcal/mol with a dielectric continuum, but then used a maximum entropy argument to replace this result by −68 kcal/mol. This was revisited by Merchant and Asthagiri,[11] who obtained a value of −96.1 kcal/mol using the dielectric continuum and −90.0 kcal/mol with explicit TIP3P. Varma and Rempe[15] later reported a value of −95.6 kcal/mol with a dielectric continuum (numerical value extracted from their Fig. 3). For K^+, Rempe *et al.*[14] reported a pQCT hydration free energy with a dielectric continuum of −71.5 kcal/mol. This was again revisited by Varma and Rempe,[15] who reported a value of −73.5 kcal/mol using a dielectric continuum (numerical value extracted from their Fig. 3). In another pQCT study, free energies of −93.3 kcal/mol for Na^+ and −71.8 kcal/mol for K^+ were reported by Varma and Rempe.[16]

While these studies show that the hydration free energies obtained by pQCT for those cations are clearly of the correct magnitude, they do not inform us of the accuracy of the pQCT framework as an approximation to the full statistical mechanical many-body configurational integral. In effect, this issue cannot be resolved by comparing the computational results with experimental estimates of the single ion absolute hydration free energy. Experimental measurements require extra thermodynamic assumptions, which lead to a considerable spread in the experimental values.[18] For example, the values of Marcus[19] are among the smallest, whereas those of Tissandier *et al.*[20] are among the largest; for Li^+ they differ by about 15 kcal/mol, and for Na^+ they differ by about 16 kcal/mol.

What is needed is a pure test of the accuracy of pQCT *in silico* for a given microscopic model. That is, a plain and direct comparison of computational results from pQCT with *exact* computational results from free energy perturbation molecular dynamics (FEP/MD) simulations with explicit solvent molecules. Surprisingly, such a comparison has never been reported.

The goal of this article is to test and illustrate the accuracy of pQCT to model the hydration of Li^+, Na^+, and K^+. Simple classical nonpolarizable models of water and ions are utilized. While potential functions incorporating the influence of electronic polarization are more advanced, the present models are sufficiently accurate for the purpose of illustrating the application of the pQCT. Section II summarizes the theoretical formulation of QCT and of pQCT. Then, details about the computations are given and the main results are discussed. The article closes with a brief conclusion.

II. ION SOLVATION AND CONSTRAINED COORDINATION SHELL

An ion of type i immersed in a liquid of N solvent molecules in a large volume V is considered. The solvent molecules are assumed to be chemically identical and, without loss of generality, it is assumed that the ion is fixed at the origin (in the following, the fixed ion is implicitly included in the notation but is omitted for clarity). The potential energy of the system is $U_i(\mathbf{X})$, where $\mathbf{X} \equiv \{\mathbf{x}_1, \ldots, \mathbf{x}_N\}$ represents the coordinates of the N solvent molecules. When the potential energy of the system is written as $U_0(\mathbf{X})$, it is implied that the interaction of the ion with the solvent molecules has been switched off (decoupled). The solvation free energy of the ion i is

$$e^{-\beta \Delta G_i} = \frac{\int d\mathbf{X} e^{-\beta U_i(\mathbf{X})}}{\int d\mathbf{X} e^{-\beta U_0(\mathbf{X})}}. \tag{1}$$

Equation (1) involves unconstrained configurational integrals. It is also of interest to express solvation properties in terms of the local configurations of the solvent molecules occupying the neighborhood of the ion. Let us define inner shell region as a spherical subvolume v within a distance R away from the ion and an outer region corresponding to the rest of the space. Let us then introduce a function, $n(\mathbf{X};R)$, to count for the number of solvent molecules within the inner region,

$$n'(\mathbf{X};R) = \sum_k H(|\mathbf{r}_k| - R), \tag{2}$$

where H is a Heaviside step function equal to 1 only when its argument is less than 0. For example, in the case of water, the kth molecule lies within the inner region when its oxygen at $\mathbf{r}_k$ is at a distance $|\mathbf{r}_k|$ smaller than the radius R. For any configuration, the function $n'(\mathbf{X};R)$ is equal to the number of solvent molecules in the inner region. Using the function $n'(\mathbf{X};R)$ to define the discrete Kroenecker delta function $\delta_{n,n'}$, one can write the probability to have exactly n solvent molecules in the inner region of ion i is defined as

$$P_i(n;R) = \frac{\int d\mathbf{X} \delta_{n,n'} e^{-\beta U_i(\mathbf{X})}}{\int d\mathbf{X} e^{-\beta U_i(\mathbf{X})}} = \langle \delta_{n,n'(\mathbf{X};R)} \rangle_{(U_i)}. \tag{3}$$

Because $\Sigma_n \delta_{n,n'} = 1$, the probability $P_i(n;R)$ is correctly normalized by construction (in the following, the dependence of P_i on R is not always explicitly written). Using the potential energy U_0, similar expressions can be derived for $P_0(n;R) = \langle \delta_{n,n'} \rangle_{(U_0)}$, the probability of finding exactly n solvent mol-

ecules inside a spherical subvolume centered on a decoupled (noninteracting) ion fixed at the origin.

Different expressions for the solvation free energy ΔG_i of Eq. (1) can be derived by exploiting the normalization condition and using the number of solvent molecules in the inner shell as a constraint. For example, it is possible to express in terms of constrained averages over the number of solvent molecules occupying the inner shell,

$$e^{-\beta\Delta G_i} = \frac{\langle \delta_{n,n'} \rangle_{(U_0)}}{\langle \delta_{n,n'} \rangle_{(U_i)}} \frac{\int d\mathbf{X} \delta_{n,n'} e^{-\beta U_i(\mathbf{X})}}{\int d\mathbf{X} \delta_{n,n'} e^{-\beta U_0(\mathbf{X})}} = \frac{P_0(n)}{P_i(n)} \frac{\int d\mathbf{X} \delta_{n,n'} e^{-\beta U_i(\mathbf{X})}}{\int d\mathbf{X} \delta_{n,n'} e^{-\beta U_0(\mathbf{X})}} = \frac{P_0(n)}{P_i(n)} e^{-\beta\Delta G_i(n)}, \tag{4}$$

where $\Delta G_i(n)$ is the solvation free energy of ion of type i with an inner shell constrained to contain exactly n solvent molecules. This expression is valid for any n as long as the ratio $P_0(n)/P_i(n)$ is well defined and can be most useful if there is a good overlap between these probabilities. Alternatively, it is possible to use a decoupled (noninteracting) ion as a reference state,

$$\begin{aligned} e^{-\beta\Delta G_i} &= \frac{\int d\mathbf{X} e^{-\beta U_i(\mathbf{X})}}{\int d\mathbf{X} e^{-\beta U_0(\mathbf{X})}} \\ &= \frac{\int d\mathbf{X} \left(\Sigma_n \delta_{n,n'}\right) e^{-\beta U_i(\mathbf{X})}}{\int d\mathbf{X} e^{-\beta U_0(\mathbf{X})}} \\ &= \sum_n \left(\frac{\int d\mathbf{X} \delta_{n,n'} e^{-\beta U_i(\mathbf{X})}}{\int d\mathbf{X} e^{-\beta U_0(\mathbf{X})}} \times \frac{\int d\mathbf{X} \delta_{n,n'} e^{-\beta U_0(\mathbf{X})}}{\int d\mathbf{X} \delta_{n,n'} e^{-\beta U_0(\mathbf{X})}} \right) \\ &= \sum_n \left(\langle \delta_{n,n'} \rangle_{(U_0)} \frac{\int d\mathbf{X} \delta_{n,n'} e^{-\beta U_i(\mathbf{X})}}{\int d\mathbf{X} \delta_{n,n'} e^{-\beta U_0(\mathbf{X})}} \right) \\ &= \sum_n P_0(n;R) e^{-\beta\Delta G_i(n)}. \end{aligned} \tag{5}$$

This expression is not limited by overlap of probabilities such as Eq. (4), but it requires the evaluation of $\Delta G_i(n)$ for multiple values of n. Equations (4) and (5) were also obtained by Merchant and Asthagiri[11] via a different route and an expression closely related to Eq. (4) was derived and used by Yu *et al.*[21] Lastly, one can also use an empty inner shell region occupied by $n=0$ solvent molecules with a decoupled (switched off) ion as a reference state,

$$\begin{aligned} e^{-\beta\Delta G_i} &= \frac{\int d\mathbf{X} e^{-\beta U_i(\mathbf{X})}}{\int d\mathbf{X} e^{-\beta U_0(\mathbf{X})}} \\ &= \frac{\int d\mathbf{X} \left(\Sigma_n \delta_{n,n'}\right) e^{-\beta U_i(\mathbf{X})}}{\int d\mathbf{X} e^{-\beta U_0(\mathbf{X})}} \\ &= \sum_n \frac{\int d\mathbf{X} \delta_{n,n'} e^{-\beta U_i(\mathbf{X})}}{\int d\mathbf{X} e^{-\beta U_0(\mathbf{X})}} \times \frac{\int d\mathbf{X} \delta_{0,n'} e^{-\beta U_0(\mathbf{X})}}{\int d\mathbf{X} \delta_{0,n'} e^{-\beta U_0(\mathbf{X})}} \\ &= \langle \delta_{0,n'} \rangle_{(U_0)} \sum_n \left(\frac{\int d\mathbf{X} \delta_{n,n'} e^{-\beta U_i(\mathbf{X})}}{\int d\mathbf{X} \delta_{0,n'} e^{-\beta U_0(\mathbf{X})}} \right) \\ &= P_0(0;R) \sum_n \frac{N!}{(N-n)!n!} \\ &\quad \times \left(\frac{\int_{\text{in}} d\mathbf{x}_1 \cdots d\mathbf{x}_n e^{-\beta W_i(\mathbf{x}_1,\ldots,\mathbf{x}_n;R)}}{\int_{\text{out}} d\mathbf{x}_1 \cdots d\mathbf{x}_n e^{-\beta W_0(\mathbf{x}_1,\ldots,\mathbf{x}_n;R)}} \right), \end{aligned} \tag{6}$$

where W_i and W_0 are potential of mean forces (PMFs) defined below. The quantity $-k_B T \ln[P_0(0;R)] \equiv \Delta G_{\text{hs}}(R)$ is the reversible work for creating an empty inner sphere of radius R in the bulk solvent (effectively, the solvation free energy of a hard sphere in water). It is worth pointing out that only a finite number of terms are expected to be nonzero and contribute to the sum in Eq. (6). Beyond some maximum number $N_{\max}$, it is physically impossible to insert additional solvent molecules into the finite volume of the inner shell. For this reason, it is preferable to describe Eq. (6) as a cluster sum rather than a cluster expansion (which suggests the existence of contributions to all order). The factor $1/n!$ comes about in Eq. (12) to avoid the overcounting configurations of n identical solvent molecules allowed to fluctuate throughout the entire inner shell.

The function W_i is the PMF for n tagged solvent molecules restricted to the inner shell surrounding the ion i under the influence of the $(N-n)$ others restricted to the outer region,

$$e^{-\beta W_i(\mathbf{x}_1,\ldots,\mathbf{x}_n;R)} \equiv \frac{\int_{\text{out}} d\mathbf{x}_{n+1} \cdots \int_{\text{out}} d\mathbf{x}_N e^{-\beta U_i}}{\int_{\text{out}} d\mathbf{x}_{n+1} \cdots \int_{\text{out}} d\mathbf{x}_N e^{-\beta U_0^*}}, \tag{7}$$

where U_0^* means that the n solvent molecules as well as the ion are decoupled from one another and from the surrounding. Likewise, W_0 is the PMF for the n solvent molecules in the absence of the ion,

$$e^{-\beta W_0(\mathbf{x}_1,\ldots,\mathbf{x}_n;R)} \equiv \frac{\int_{\text{out}} d\mathbf{x}_{n+1} \cdots \int_{\text{out}} d\mathbf{x}_N e^{-\beta U_0}}{\int_{\text{out}} d\mathbf{x}_{n+1} \cdots \int_{\text{out}} d\mathbf{x}_N e^{-\beta U_0^*}}. \tag{8}$$

By construction, $W_i(\mathbf{x}_1,\ldots,\mathbf{x}_n;R)$ represents the total reversible work for assembling the ion of type i and the n solvent molecules within the inner shell region in the configuration $\{\mathbf{x}_1,\ldots,\mathbf{x}_n\}$ and then solvating the entire complex in the bulk solvent while allowing no other solvent molecules to come within the inner region of radius R. It can be written as $W_i = u_i + \Delta W_i$, where u_i represents the vacuum potential energy and ΔW_i represents the solvation free energy of the ion-solvent complex. In the denominator of Eq. (6), the n solvent molecules are constrained to be in the outer region. Once they are far away in the bulk, the n solvent molecules become decoupled and $W_0(\mathbf{x}_1,\ldots,\mathbf{x}_n;R) \to \Sigma_i w_s(\mathbf{x}_i)$, which is a superposition of PMF for each individual solvent molecule. The single solvent molecule PMF can be related to, $\Delta\mu_s$, the excess chemical potential of a solvent molecule in the bulk liquid,

$$\begin{aligned} &\int_{\text{out}} d\mathbf{x}_1 \cdots d\mathbf{x}_n e^{-\beta W_0(\mathbf{x}_1,\ldots,\mathbf{x}_n;R)} \\ &\quad \to V^n \left(\int_{\text{out}} d\mathbf{x}_1 \delta(\mathbf{r}_1 - \mathbf{r}_1^*) e^{-\beta w(\mathbf{x}_1)} \right)^n \\ &\quad = V^n (e^{-\beta\Delta\mu_s} Z_s)^n, \end{aligned} \tag{9}$$

where Z_s is the configurational integral of a single solvent molecule in vacuum with its center of mass constrained at $\mathbf{r}_1^*$,

defined in terms of the internal potential energy of a single solvent molecule $u_s(\mathbf{x}_1)$,

$$Z_s = \int d\mathbf{x}_1 \delta(\mathbf{r}_1 - \mathbf{r}_1^*) e^{-\beta u_s(\mathbf{x}_1)} \tag{10}$$

(Z_s is equal to $8\pi^2$ for a rigid solvent molecule). Using these definitions, the solvation free energy in Eq. (6) is written as

$$e^{-\beta \Delta G_i} = P_0(0;R) \sum_n (\bar{\rho}_s)^n e^{n\beta\Delta\mu_s} K_i(n), \tag{11}$$

where $\bar{\rho}_s = (N/V)$ is the bulk solvent density [using $N!/(N-n)! \approx N^n$] and the factor $K_i(n)$ is the configurational integral of n solvent molecules restricted to the inner shell around the ion

$$K_i(n) = \frac{1}{n!} \frac{1}{(Z_s)^n} \int_{\text{in}} d\mathbf{x}_1 \cdots d\mathbf{x}_n e^{-\beta W_i(\mathbf{x}_1,\ldots,\mathbf{x}_n;R)}. \tag{12}$$

By construction, $K_i(n)$ for $n=0$ corresponds to the electrostatic solvation free energy ΔW_i of an ion embedded inside the inner subvolume and coordinated by zero solvent molecules. Using similar arguments, the normalized probability can be expressed as

$$P_i(n) = \frac{(\bar{\rho}_s)^n e^{n\beta\Delta\mu_s} K_i(n)}{\Sigma_m (\bar{\rho}_s)^m e^{m\beta\Delta\mu_s} K_i(m)}. \tag{13}$$

Equations (11) and (13) are introduced in the context of QCT of ion solvation.[6,7] QCT gets its name because the factors $K_i(n)$ in Eq. (11) assume the form of equilibrium constants for pseudochemical binding reactions for one ion and n solvent molecules. For example, hydration of Na^+ is conceptually described as $Na^+ + nH_2O \leftrightarrow Na(H_2O)_n^+$ in QCT, where the water molecules are treated like a reactant at concentration $\bar{\rho}_w$.

It is worth pointing out that the formal derivation leading to Eqs. (11) and (13) differs somewhat from the standard QCT literature.[6,7] Here, our starting point with Eq. (6) consists of inserting $\delta_{n,n'}$ in the configurational integral to sort out the number of solvent molecules in the inner-shell region, which is treated as a small open system in equilibrium with a bath of solvent molecules. A similar development was previously used to design a grand canonical Monte Carlo (GCMC) algorithm for the treatment of hydration in free energy calculations in finite subsystems submitted to an effective solvent boundary potential.[22–24] Interestingly, the present derivation of QCT bears also some similarities to the concept of the small system grand canonical ensemble considered by Reiss and co-workers,[25] revealing some unsuspected relationships between those two frameworks.

An alternative expression for the solvation free energy may also be obtained by noting that the normalized probability in Eq. (13) can be re-expressed as

$$P_i(n) = P_0(0;R)(\bar{\rho}_s)^n e^{n\beta\Delta\mu_s} K_i(n) e^{+\beta\Delta G_i}, \tag{14}$$

where the closed formed of the sum was taken from Eq. (11). Rearranging the terms yields the expression,

$$\Delta G_i = -k_B T \ln[K_i(n)(\bar{\rho}_s)^n] + k_B T \ln[P_i(n)] - k_B T \ln[P_0(0;R)] - n\Delta\mu_s, \tag{15}$$

which is valid for any n as long as the corresponding coordinated state occurs with nonzero probability. If all the quantities are treated strictly at the same level, then the transformation from Eq. (11) to Eq. (15) is merely formal and carries no new information. In practice, Eq. (15) may provide a useful approximation by allowing one to combine the $P_i(\tilde{n})$ extracted from computer simulations with the coefficient $K_i(\tilde{n})$ calculated from a cluster as in standard pQCT. The best results might be expected by choosing n equal to the average coordination number $\langle n \rangle$, or the most probably coordination number $\tilde{n}$. With the latter choice, the expression above is closely related to Eq. (9) from a recent review by Asthagiri *et al.*[8]

III. APPROXIMATION OF SMALL OSCILLATIONS

If the Boltzmann configurational integrals in Eq. (11) are approximated by using the assumption of small oscillations around the energy minimum, then one obtains a simple closed form expression for the solvation free energy of an ion that can be evaluated numerically. The treatment can be further simplified by ignoring the influence of the bulk solvent in the outer region on the geometry and the vibrational frequencies of the complex comprising the ion and the n solvent molecules, which can then be treated as an isolated cluster in vacuum using *ab initio* quantum chemistry approaches. The influence of the outer-shell contribution to the solvation free energy is then calculated separately by keeping the cluster frozen in the configuration corresponding to its energy minimum in vacuum. These approximations yield the so-called pQCT. In the following, we illustrate this strategy for computing the hydration free energy of monatomic ions using a classical force field.

Typically, pQCT is formulated in such a way to directly use the output of standard quantum chemistry programs. For this purpose, it is convenient to restore the momentum integrals in all configurational integrals, i.e., for each degree of freedom x,

$$\int dx e^{-\beta U} \rightarrow \frac{1}{h} \int dp \int dx e^{-\beta E}, \tag{16}$$

where E is the total kinetic and potential energy. This substitution does not affect any classical statistical mechanics results since all the configurational integrals appear as ratio and the momentum component cancels out (equilibrium properties such as the solvation free energy do not depend on the momenta and all the results are independent of the masses attributed to the particles). This makes it is possible to express the binding factor $K_i(n)$ in terms of traditional-looking vacuum partition functions Q for the translational (t), rotational (r), and vibrational (v) degrees of freedom for the complex (c), and the isolated ion (i), and water (w) molecule.

First, the cluster is characterized in vacuum, yielding the optimized geometry $\{\mathbf{x}^{\min}\}$, the energy minimum $u_c^{\min} \equiv u_c(\{\mathbf{x}^{\min}\})$. From this optimized geometry, the translational, rotational, and vibrational partition functions Q_t^c, Q_r^c, and Q_v^c

can be calculated. Similarly, an isolated water molecule is also characterized in vacuum, yielding the optimal geometry $\mathbf{x}_1^{\min}$, the energy minimum, $u_w^{\min} \equiv u_w(\{\mathbf{x}_1^{\min}\})$, as well as the translational, rotational and vibrational partition functions Q_t^w, Q_r^w, and Q_v^w. The solvation of the cluster, frozen in the optimized geometry $\{\mathbf{x}^{\min}\}$, is $\Delta W_c^{\min} \equiv \Delta W_c(\{\mathbf{x}^{\min}\})$. As will be shown, $\Delta W_c^{\min}$ represents a critical contribution to the resulting hydration free energies. With these approximations, the binding factors in pQCT are approximated by

$$K_i(n) = e^{-\beta[\Delta u_c^{\min} + \Delta W_c^{\min}]} \frac{(Q_t^c/V) Q_r^c Q_v^c}{(Q_t^i/V)[(Q_t^w/V) Q_r^w Q_v^w]^n}$$
$$= e^{-\beta[\Delta u_c^{\min} + \Delta W_c^{\min}]} K_n^{\mathrm{trv}}, \tag{17}$$

where $\Delta u_c^{\min} = [u_c^{\min} - n(u_w^{\min})]$ is the binding energy of the cluster in vacuum, and Q_t^c, Q_r^c, Q_v^c, Q_t^i, Q_t^w, Q_r^w, and Q_v^w are the vacuum partition functions for translation, rotation, and vibration of the complex and of the isolated ion and water molecule, respectively. The ground state energy of the ion provides only an offset constant that is omitted here for the sake of simplicity. The factor of $1/n!$ appearing in Eq. (12), introduced when the identical solvent molecules are allowed to fluctuate throughout the entire inner shell, is removed with the assumption of small oscillations because each solvent molecule fluctuates around a unique configuration. For the complex comprising 1 ion and n water molecules, there are $3+3n$ degrees of freedom and $3(3+3n)$ vibrational normal modes. (Vibrational modes 1–6 of a nonlinear molecule have zero frequency when the calculations are carried out with Cartesian coordinates.) There are three nonzero vibrational normal modes for the isolated water molecule. The translational partition functions are

$$Q_t^c/V = (2\pi M^c k_B T/h^2)^{3/2} \tag{18}$$

for the complex,

$$Q_t^i/V = (2\pi M^i k_B T/h^2)^{3/2} \tag{19}$$

for the isolated ion, and

$$Q_t^w/V = (2\pi M^w k_B T/h^2)^{3/2} \tag{20}$$

for the isolated water molecule. The volume factor V is removed from all translational partition functions. The rotational partition functions are

$$Q_r^c = (\pi I_a^c I_b^c I_c^c)^{1/2} (8\pi^2 k_B T/h^2)^{3/2} \tag{21}$$

for the complex and

$$Q_r^w = (\pi I_a^w I_b^w I_c^w)^{1/2} (8\pi^2 k_B T/h^2)^{3/2} \tag{22}$$

for the isolated water molecule. The vibrational partition functions are

$$Q_v^c = \Pi_{i=7}^{9n+3} (k_B T/hF_i^c) \tag{23}$$

for the complex and

$$Q_v^w = \Pi_{i=7}^{9} (k_B T/hF_i^w) \tag{24}$$

for an isolated water molecule. In the above expressions, h is Planck's constant, M^c, M^i, and M^w are the masses, I^c, I^i, and I^w are the moments of inertia and F_i is the vibrational frequency (it can be verified that all factors involving Planck's constant cancel out, consistent with classical statistical mechanics). The hydration free energy of the ion is

$$\Delta G_i = \Delta G_{hs}(R) - k_B T \ln\left[\sum_n e^{-\beta(\Delta u_c^{\min} + \Delta W_c^{\min} - n\Delta\mu_w)} (\bar{\rho}_w)^n K_n^{\mathrm{trv}}\right]$$
$$= \Delta G_{hs}(R) - k_B T\left[\sum_n e^{B_n}\right], \tag{25}$$

and the normalized probability for finding n water molecule within the inner shell of radius R is

$$P_i(n) = \frac{e^{B_n}}{\Sigma_m e^{B_m}}. \tag{26}$$

IV. COMPUTATIONAL DETAILS

pQCT is used to calculate the hydration free energy of Li^+, Na^+, and K^+ based on simple potential functions. The water molecules are modeled with the simple nonpolarizable TIP3P potential.[26] The bulk density of water is $\bar{\rho}_w = 0.0334$ Å^{-3}. The excess chemical potential (absolute hydration free energy) of a TIP3P water molecule in bulk water, $\Delta\mu_w$, was calculated using a step-by-step FEP/MD simulations protocol. Based on a number of previous estimates, $\Delta\mu_w$ is set to -6.0 kcal/mol (the uncertainty is on the order of 0.2 kcal/mol).[24,27] This comprises roughly 5.0 kcal/mol from the repulsive (Weeks–Chandler–Andersen[28]) Lennard-Jones (WCA-LJ) interaction, -2.8 kcal/mol from the dispersive van der Waals attraction, and -8.2 kcal/mol from charging electrostatics.

The ions are modeled as simple LJ particles with a fixed charge; equal to $E_{\min}((R_{\min}/r)^{12} - 2(R_{\min}/r)^6)$. The LJ parameters for the ions, $E_{\min}$ and $R_{\min}/2$, are $-0.002\,33$ kcal/mol and 1.2975 Å for Li^+, -0.0469 kcal/mol and 1.410 75 Å for Na^+,[29,30] and -0.0870 kcal/mol and 1.76375 Å for K^+.[31] Ion-water LJ parameters are defined with standard combination rules: $E_{\min}^{iw} = \sqrt{E_{\min}^{ww} E_{\min}^{ii}}$ and $R_{\min}^{iw} = (R_{\min}^{ww}/2 + R_{\min}^{ii}/2)$. The results from the K^+ model in TIP3P water are consistent with those obtained with a recent polarizable potential function as well as with MD simulations based on density function theory.[17]

All the simulations were carried as with the spherical solvent boundary potential (SSBP) (Ref. 31) to avoid any ambiguities in the absolute charging free energies, which are calculated as *real* free energies.[18,32] To gather statistics on the coordination structure of each ion, a 2 ns simulation with the ion fixed at the center of a sphere of 400 TIP3P water molecules was generated. A 2 ns simulation of a sphere of 400 TIP3P water molecules with no ion was also generated for analysis. The molecular dynamics (MD) trajectories were generated with a Langevin thermostat to ensure thermalization at a temperature of 300 K with a friction of 5/ps applied to all the water oxygens. A time step of 2 fs was used and the geometry of the TIP3P molecules was kept rigid using the SHAKE algorithm.[33] The MD trajectories were saved every 10 steps and 100 000 configurations were used to calculate the ion-solvent radial distribution function (RDF) $g(r)$, the running coordination number $N(r)$ (with a spacing of 0.025 Å), and the probability distributions $P_i(n;R)$ for different inner shell radius R. All the simulations were generated and

TABLE I. Summary of all data for Li^+. (Energies are in kcal/mol, and K_n^{trv} is in $Å^{3n}$. ΔW_c^{min} based on a dielectric continuum approximation for the outer region.)

					$\ln[P_i(n)]$	
n	ΔU_c^{min}	ΔW_c^{min}	$\ln[K_n^{trv}]$	B_n	pQCT	MD
0	⋯	−66.050	⋯	110.7924	−111.746	⋯
1	−34.120	−64.000	−2.613	148.6702	−73.868	⋯
2	−65.910	−62.760	−6.103	182.8973	−39.641	⋯
3	−92.902	−62.399	−12.936	207.2904	−15.248	−7.729
4	−114.469	−62.601	−20.775	222.5142	−0.024	−0.174
5	−123.563	−63.000	−26.946	218.8064	−3.732	−1.843
6	−132.330	−62.300	−34.254	211.5678	−10.971	−7.059

analyzed with the program CHARMM.[34]

The absolute hydration free energy of the K^+ model was calculated using a step-by-step FEP/MD simulations protocol[27] with 400 TIP3P water molecules in SSBP; the total simulation time for each calculation is 14.8 ns excluding equilibration. The calculated ΔG_K is equal to −80.6 kcal/mol, comprising of 5.1 kcal/mol from the repulsive WCA-LJ, −2.0 kcal/mol from the dispersive van der Waals attraction, and −83.7 kcal/mol from electrostatics. The hydration free energy of K^+ relative to Na^+ ($\Delta G_K - \Delta G_{Na}$) was calculated from FEP/MD and is equal to 18.0 kcal/mol. The hydration free energy of Na^+ relative to Li^+ ($\Delta G_{Na} - \Delta G_{Li}$) was calculated from FEP/MD and is equal to 22.7 kcal/mol. From these results, it is deduced that the absolute hydration free energies of the Na^+ and Li^+ models are −98.6 and −121.3 kcal/mol, respectively.

The hydration free energy of the frozen cluster was calculated using a continuum dielectric approximation for the bulk solvent in the outer region. The finite-difference Poisson–Boltzmann solver PBEQ[35,36] of the program CHARMM was used.[34] In those calculations, the dielectric constant of the outer bulk region was set to 80 and the dielectric constant of the region from which solvent is excluded was set to 1 (the salt concentration was set to 0 and no ionic screening was included). The dielectric=1 region is constructed as the union of the spherical inner shell region of radius R and of the region of space where water molecules would overlap with the repulsive core of any of the n water molecules surrounding the ion in the inner shell. In this construction, the radius of the water molecules in the inner shell was set to 1.88 Å. This value was empirically adjusted to reproduce the charging free energy of −8.1 kcal/mol for a single TIP3P water molecule in bulk.[27] Such Born-type atomic radii, empirically adjusted to reproduce FEP/MD simulations with explicit solvent, were introduced by Nina *et al.*[35] The molecular surface with a probe radius of 1.4 Å was used to define the cluster-bulk dielectric interface (including the re-entrant surface); using the van der Waals surface altered the results by less than 0.2 kcal/mol. A cubic grid of 140^3 points with a grid spacing of 0.15 Å was used in PBEQ.

The contribution from the outer region to the hydration free energy of the frozen cluster was also calculated using FEP/MD simulations with explicit solvent.[27] The FEP/MD simulations were carried out with SSBP in which the frozen clusters were embedded. For each system, the total number of water molecules is 400, as in the unconstrained MD simulations. For Li^+, 4 water molecules are in the inner region and 396 are in the outer region. For Na^+, 6 water molecules are in the inner region and 394 are in the outer region. All the atoms in the inner-shell region are kept fixed in the configuration optimized in vacuum. A spherical steep half-harmonic restraining potential with a force constant of 1000 kcal/mol $Å^2$ is imposed to prevent the water molecules from the outer region to enter into the inner-shell region. The radius of the inner-shell is 2.5 Å for Li^+ and 3.0 Å for Na^+. Following a step-by-step reversible work protocol established previously,[27] the hydration free energy of the frozen cluster was decomposed into three components, ΔW_c^{min}(elec), ΔW_c^{min}(dis), and ΔW_c^{min}(rep), corresponding to the electrostatic, WCA-LJ dispersive attraction, WCA-LJ

TABLE II. Summary of all data for Na^+. (Energies are in kcal/mol and K_n^{trv} is in $Å^{3n}$. ΔW_c^{min} based on a dielectric continuum approximation for the outer region.)

					$\ln[P_i(n)]$	
n	ΔU_c^{min}	ΔW_c^{min}	$\ln[K_n^{trv}]$	B_n	pQCT	MD
0	⋯	−54.990	⋯	92.240	−90.704	⋯
1	−24.815	−55.770	−1.712	120.2453	−62.698	⋯
2	−48.198	−56.730	−3.491	145.6663	−37.277	⋯
3	−68.838	−58.100	−8.041	164.5840	−18.360	⋯
4	−86.638	−59.700	−13.510	178.2045	−4.739	−4.874
5	−99.605	−61.200	−20.829	181.6956	−1.248	−1.486
6	−110.537	−62.601	−27.160	182.5930	−0.351	−0.294
7	−120.921	−60.500	−41.463	168.7238	−14.220	−3.877
8	−131.389	−57.200	−57.839	150.9089	−32.035	−7.929

TABLE III. Summary of all data for K^+. (Energies are in kcal/mol, and K_n^{trv} is in $Å^{3n}$. ΔW_c^{min} based on a dielectric continuum approximation for the outer region.)

					$\ln[P_i(n)]$	
n	ΔU_c^{min}	ΔW_c^{min}	$\ln[K_n^{trv}]$	B_n	pQCT	MD
0	⋯	−51.460	⋯	86.3191	−64.115	⋯
1	−18.925	−52.250	−0.938	105.1597	−45.274	⋯
2	−36.898	−53.500	−1.295	123.4768	−26.957	⋯
3	−53.036	−55.000	−4.326	136.5771	−13.857	−8.047
4	−67.302	−56.900	−8.165	146.4003	−4.033	−4.352
5	−78.776	−59.000	−13.461	150.4158	−0.018	−1.898
6	−91.005	−55.899	−32.215	133.5137	−16.920	−0.790
7	−104.188	−55.300	−43.665	129.7120	−20.722	−1.128
8	−115.315	−57.600	−48.051	134.3906	−16.043	−2.863
9					⋯	−6.156
10					⋯	−8.874

core repulsion, respectively. The total production FEP/MD simulation is 14.8 ns for each ion (excluding 7.4 ns equilibration).

All the vibrational analysis calculations were performed using the VIBRAN module of CHARMM;[34] the latter outputs all the normal model frequencies for the system. In the vibrational analysis, the TIP3P water molecules were treated as flexible, with a stiff harmonic constant of 10000 kcal/mol $Å^2$ for the O–H1, O–H2, and H1–H2 stretching. Test calculations indicated that the results did not depend on the force constant. All the information for the application of pQCT is summarized in Tables I–III. The free energy results are reported in Table V. Figure 1 shows the free energy for creating an empty cavity of radius R in pure bulk solvent, $\Delta G_{hs}(R)$ calculated from MD.

V. DISCUSSION

To carry calculations with pQCT, it is first necessary to choose the value of the inner shell radius R and then construct a set of energy-minimized clusters comprising the ion and n water molecules such that the farthest water oxygen lies within a distance R. A possible choice for the inner shell radius could be the position of the first minimum in the ion-water RDF. The ion-solvent RDFs are shown in Fig. 2. The first minima are at 2.66, 3.14, and 3.54 Å for Li^+, Na^+, and K^+, respectively. The average coordination numbers defined at the first minimum in the RDFs calculated from the MD trajectories are 4.18, 5.85, and 7.11 for Li^+, Na^+, and K^+, respectively. In a typical application of QCT, it may be advantageous to choose R as small as possible to limit the number of explicit water molecules that must be included in the cluster sum. In the context of pQCT, this is critical because the approximation of small oscillations used to integrate the thermal fluctuations becomes increasingly invalid when the anharmonic effects dominate for $n > 5–6$.[37]

The coordination probability $P_i(n;R)$ as a function of the inner-shell radius for the different ions, shown in Figs. 3–5, displays the accessible coordinated states of a hydrated ion. For Li^+, the most probable coordination state is $n=4$. Using the first minimum in the RDFs implies that coordinated states ranging from 4 to 6 should be incorporated into the cluster sum. For Na^+, this choice implies that coordinated states ranging from 4 to 8 should be incorporated into the

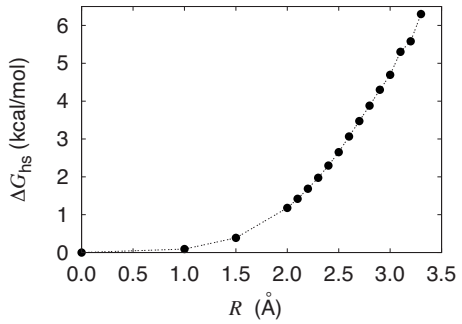

FIG. 1. Reversible work $\Delta G_{hs}(R)$ for creating an empty cavity of radius R in bulk water.

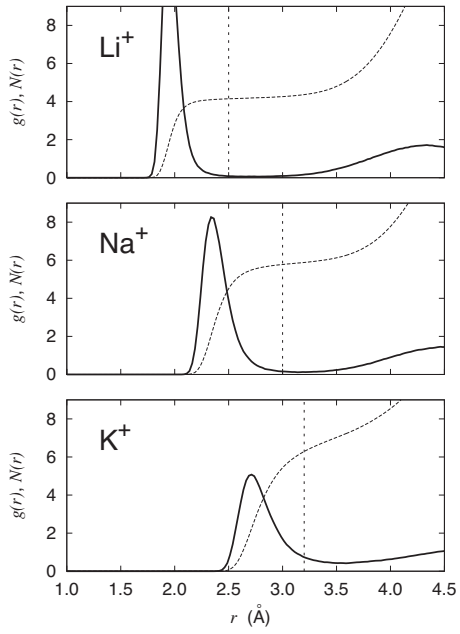

FIG. 2. Radial ion-solvent distribution function (RDF) $g(r)$ and running coordination number $N(r)$. In the $g(r)$, the minimum is at 2.66, 3.14, and 3.54 Å, and the maximum 1.94, 2.31, and 2.71 Å, for Li^+, Na^+, and K^+, respectively. The average coordination numbers based on the first minimum in $g(r)$ are 4.18, 5.85, and 7.11 for Li^+, Na^+, and K^+, respectively.

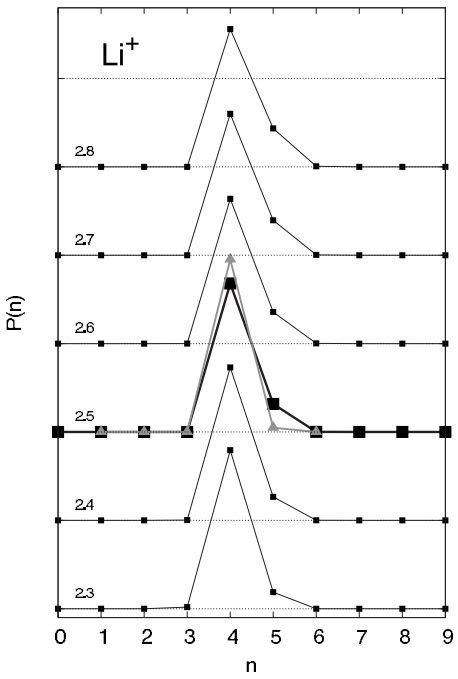

FIG. 3. Probability of occupancy with different inner shell radii around a Li^+ ion. The occupancy probability $P(n)$ for an inner shell radius of 2.5 Å chosen for pQCT (shown as a thicker line with black squares) is $P(3)=0.000\,44$, $P(4)=0.840\,34$, $P(5)=0.158\,36$, and $P(6)=0.000\,86$. For the sake of completeness, the strict frequencies extracted from 100 000 configurations are reported, with no considerations for statistical uncertainty. The occupancy probability calculated from pQCT is shown as a gray line with triangles (see also Table I).

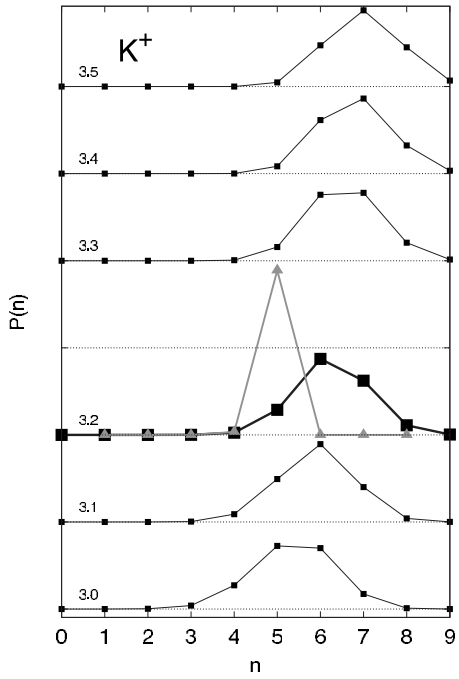

FIG. 5. Probability of occupancy with different inner shell radii around a K^+ ion. The occupancy probabilities $P(n)$ for an inner shell radius of 3.2 Å chosen for pQCT (shown as a thicker line with black squares) are $P(3)=0.000\,32$, $P(4)=0.012\,88$, $P(5)=0.149\,84$, $P(6)=0.454\,04$, $P(7)=0.323\,56$, $P(8)=0.057\,10$, $P(9)=0.002\,12$, $P(10)=0.000\,14$. For the sake of completeness, the strict frequencies extracted from 100 000 configurations are reported, with no considerations for statistical uncertainty. The occupancy probability calculated from pQCT is shown as a gray line with triangles (see also Table III).

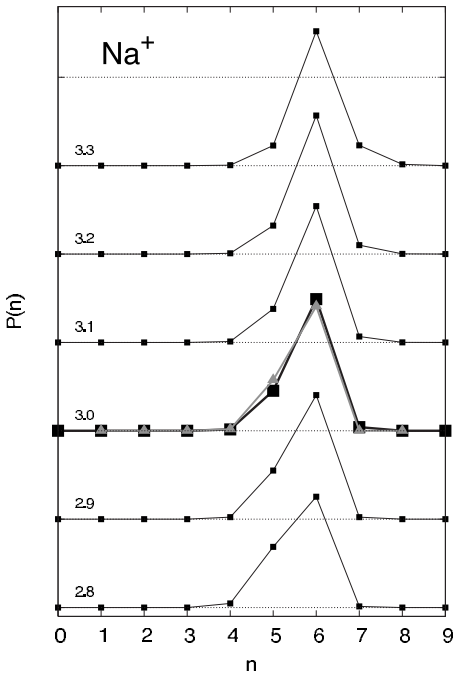

FIG. 4. Probability of occupancy with different inner shell radii around a Na^+ ion. The occupancy probabilities $P(n)$ for an inner shell radius of 3.0 Å chosen for pQCT (shown as a thicker line with black squares) are $P(3)=0.007\,64$, $P(4)=0.226\,28$, $P(5)=0.745\,00$, $P(6)=0.020\,72$, and $P(7)=0.000\,36$. For the sake of completeness, the strict frequencies extracted from 100,000 configurations are reported, with no considerations for statistical uncertainty. The occupancy probability calculated from pQCT is shown as a gray line with triangles (see also Table II).

cluster sum, i.e., ± 2 around the most probable coordinated state. However, for K^+, this implies that coordinated states ranging from 4 up to 10 should be incorporated into the cluster sum; i.e., ± 3 around the most probable coordinated state.

The final set of optimized clusters used in the pQCT calculations is shown in Fig. 6. To provide complementary information, a series of instantaneous n-coordinated states were extracted from the MD simulations and optimized in vacuum. The distribution in energy minimum and farthest distance for the optimized clusters is shown in Fig. 7. For all ions, it is clearly possible to optimize a cluster with a small number of water molecules ($n<5$) that remain in direct contact with the ion. With the current models, direct contact between the ion and the water molecule implies distances of 1.89, 2.26, and 2.62 Å for Li^+, Na^+, and K^+, respectively. However, when the number of water molecules in the cluster exceeds a certain number (5–6), additional water molecules often end up in the second hydration shell during optimization. In the case of K^+ for example, the ion-water distance jumps by about 1 Å for $n>7$ (Fig. 7). As a consequence, only a limited set of clusters can be optimized for a given choice of the inner shell radius. For Li^+, the "gap" where the water molecules start to be transferred to the second shell during the geometry optimization is somewhere between 2.2 and 4.0 Å. This strong demarcation supports the concept of a tight and well-defined first coordination shell for Li^+. For Na^+, the gap extends from 2.7 to 3.5 Å. For K^+, the gap extends from 3.2 to 3.5 Å, consistent with a first coordination

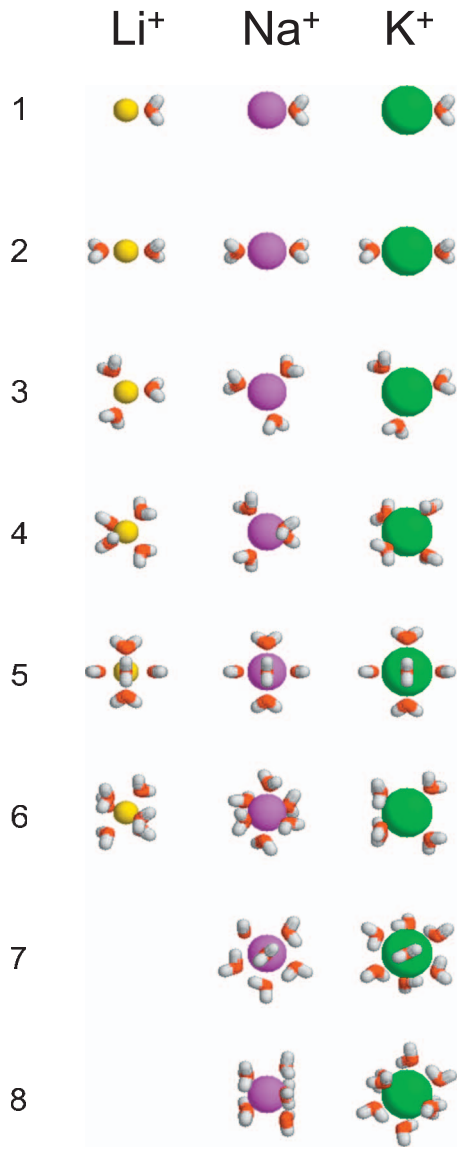

FIG. 6. Vacuum optimized clusters used in the application of pQCT.

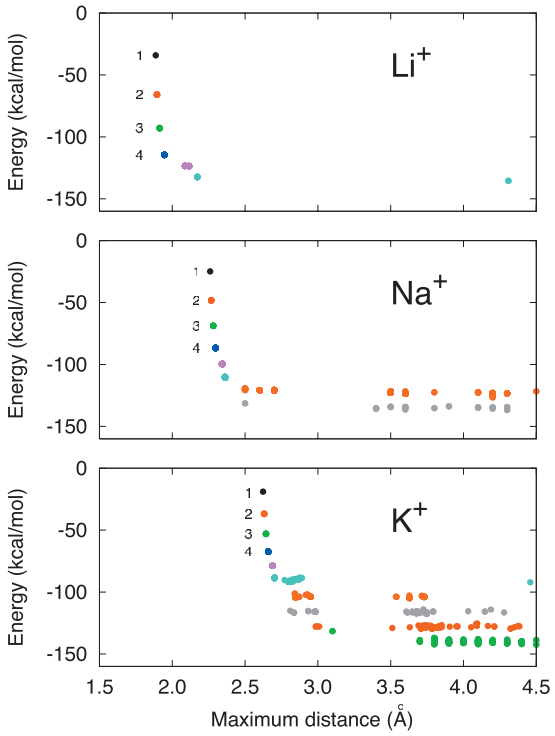

FIG. 7. Properties of the vacuum optimized clusters comprising of one ion and n water molecules used in the application of pQCT.

shell that is more difficult to define unambiguously. Examination of $P_i(n;R)$ as a function of the inner shell radius, together with the optimization of the small clusters, suggests that the reasonable choices for R are 2.5, 3.0, and 3.2 Å for Li^+, Na^+, and K^+, respectively. As observed in Fig. 2, those inner shell radii cut out a slice of the first peak in the RDFs. Ultimately, the inner shell radius must be chosen in such a way that the set of optimized clusters is consistent with the coordinated states observed in the MD simulations.

Estimates for the free energy contribution from the outer-shell region are given in Table IV. The main results about the hydration free energies are given in Table V. The hydration free energies for Li^+, Na^+, and K^+ calculated via Eq. (25) are −130.0, −104.3, and −84.1 kcal/mol. Those values, obtained by treating the outer region with a dielectric continuum are quite close to the values of −121.3, −98.6, and −80.6 obtained by explicit FEP/MD simulations. The deviation is slightly larger for Li^+, but it remains a moderate fraction of the total hydration free energy. One may note that there is a slight ambiguity in the choice of the inner-shell radius. For example, in the case of Na^+, the inner shell radius is set to 3.0 Å. However, examination of the optimized cluster properties shown in Fig. 7 indicates that even if the inner shell radius had been chosen as 3.5 Å, one would have to rely upon the same finite set of optimized clusters. Test calculations indicate that the alternative choice would shift the predicted hydration free energy by 2–3 kcal/mol for Na^+. As noted previously,[12] such change has moderate effect because the explicit water molecules in the cluster overlap with the surface of the inner sphere and determine the position of the dielectric interface.

The results are particularly sensitive to the treatment of the outer bulk solvent in the calculation of the solvation free energy $\Delta W_c^{\min}$ of the frozen optimized clusters. Only electrostatic effects from the outer region were included and no attempt to model the nonpolar component to $\Delta W_c^{\min}$ was made. Overall, the outer-region electrostatic contribution is on the order of −50 to −65 kcal/mol for the different ions (see Tables I–III). This represents more than one-half of the total hydration free energy for the ions. $\Delta W_c^{\min}$ was calculated using a continuum electrostatics approximation and it is important to note that the magnitude of this contribution is very sensitive to the position of the dielectric boundary between the cluster and the surrounding bulk solvent. Here, the dielectric boundary was constructed from the molecular surface of the cluster by attributing a Born-type radius of 1.88 Å to the water oxygen (i.e., the dielectric interface is 1.88 Å away from the water oxygen). This radius was empirically adjusted by matching the electrostatic contribution to the excess chemical potential of a TIP3P water molecule based on FEP/MD simulations with explicit solvent.[27] However, setting the oxygen-oxygen distance to 2.8 Å (i.e., the dielectric

TABLE IV. Free energy contribution from the outer-shell region (kcal/mol).

Contributions		$Li^+(water)_4$	$Na^+(water)_6$
FEP/MD	ΔW_c^{min}(rep)	11.8	14.6
	ΔW_c^{min}(dis)	−8.6	−11.7
	ΔW_c^{min}(rep)+ΔW_c^{min}(dis)	3.2	2.9
	ΔW_c^{min}(elec)	−57.8	−55.0
	ΔW_c^{min}	−54.6	−52.1
DC[a]	ΔW_c^{min}	−62.6	−62.6

[a]Dielectric continuum approximation for the outer region.

interface is located 2.8 Å away from the water oxygen), a more commonly accepted value for the size of a water molecule, would shift the solvation free energies by a considerable amount. For example, in the case of the cluster of Li^+ with four water molecules, the solvation of the frozen cluster would shift by more than 20 kcal/mol, from −62.6 to −40.7 kcal/mol, as a result of this physically motivated change. Such large variations are considerably reduced if the excess chemical potential of the single water molecule are also calculated using a continuum dielectric model with the same radius. In this case, varying the radius from 1.8 to 2.8 Å to construct the dielectric boundary results in variations in the hydration free energy of the ion on the order of 3 kcal/mol.

As observed previously, a consistent treatment of the atomic radius of the solvent molecule results in a more stable outcome from pQCT.[12–14] However, the observed consistency does not address more fundamental questions about the limitations of the dielectric continuum approximation in the treatment of the outer-shell region. In this regard, calculations based on FEP/MD simulations with explicit solvent molecules provide a critical point of reference. The present analysis was limited to the particular cluster, making the dominant contributions to the hydration free energy of Li^+ and Na^+. K^+ was not considered in this analysis due to the apparent difficulties in representing the diffuse coordination shell of this cation by pQCT. The FEP/MD calculations were carried out for the frozen cluster of one Li^+ coordinated by four water molecules and one Na^+ coordinated by six water molecules. The results are given in Table IV. For Li^+, the outer-shell contribution calculated from FEP/MD with explicit solvent molecules is −54.6 kcal/mol. This differs considerably from the value of −62.6 kcal/mol based on the dielectric continuum. As a result, the absolute hydration free energy of Li^+ shifts by 8 kcal/mol, yielding a value of −122.0 kcal/mol for Li^+. This result, based on pQCT with a treatment of the outer shell with explicit solvent, compares remarkably well with the results of −121.3 kcal/mol from direct FEP/MD. In the case of Na^+, the results are less good. The outer-shell contribution calculated from FEP/MD with explicit solvent molecules differs from the value based on the dielectric continuum by about 10 kcal/mol, and the absolute hydration free energy of Na^+ shifts from −104.3 to −93.8 kcal/mol. This estimate, based on pQCT with a treatment of the outer shell with explicit solvent, misses the exact results of −98.6 by about 5 kcal/mol. A plausible explanation for the difference between Li^+ and Na^+ is that the remaining approximations are more accurate for the former than the latter. Those are the treatment of fluctuations based on small harmonic oscillations around a single energy-minimized cluster, and the treatment of solvation of this cluster as a rigid entity kept frozen in its energy-minimized configuration. The concept of a very stiff first coordination shell comprising exactly n water molecules undergoing very small fluctuations is obviously a better approximation in the case of Li^+ than of Na^+.

The comparison with the results of FEP/MD also indicates that the dielectric continuum significantly overestimates the contribution to the charging free energy. In the case of Li^+, the dielectric continuum is too negative by about 5 kcal/mol, and for Na^+, it is too negative by almost 8 kcal/mol. Similar observations were made in the study by Asthagiri *et al.*;[9] the outer-sphere contribution from the dielectric continuum was also too negative by about 5–6 kcal/mol for Li^+ and Na^+.

The reason for the inaccuracy of the dielectric continuum is fairly clear. It is well understood that atomic radii used to construct the dielectric boundaries are empirical parameters that are context- and charge-dependent; atomic radii must be engineered carefully to yield quantitatively accurate results.[35,38] Utilizing the same atomic radius for a water molecule in the bulk, and in the first coordination shell of an ion in a dielectric representation is an approximation. The radius engineered to get a reasonable excess chemical potential for a neutral water molecule is actually too small when it is used to set the dielectric boundary for the water molecules around the charged cation. As a result, the solvation free energy from the outer shell on the frozen cluster is too negative by 5–8 kcal/mol when compared with the results of FEP/MD with explicit solvent molecules.

It is important to realize that the true inaccuracies from the dielectric continuum are considerably larger than the small variations in the pQCT hydration free energy observed when varying the radius of the water molecules used to construct the dielectric boundary results noted above. It has often been argued that consistency and robustness of the calculated free energy under variations of the solvent radius was generally indicative of the overall accuracy of pQCT.[12–14] The present analysis indicates that this argument is invalid and that consistency of pQCT results upon variations in empirical parameters does inform us about the accuracy of the results. It is also worth pointing out that the step-by-step free energy decomposition allows the determination of the nonpolar contributions to the solvation free energy provided from the outer region, ΔW_c^{min}(rep)+ΔW_c^{min}(dis). In most applications of pQCT,[9,11–14] such nonpolar contribution from the outer shell has been ignored and the contribution from the outer-shell region has been exclusively based on electrostatics. The present results show that including the nonpolar contribution is necessary to reach quantitative accuracy in the case of Li^+. More generally, this suggests that an accurate evaluation of all contributions to ΔW_c^{min} using FEP/MD simulations with explicit solvent molecules offers a powerful avenue to make the most out of the pQCT approximation.

Examination of the coordination structure predicted by pQCT provides important information about the nature of the approximations involved. From the MD simulations, the average coordination numbers $\langle n \rangle$ for the inner-shell radius R

TABLE V. Summary of data for all ions. [The ΔG_i (FEP/MD) results were obtained by calculating the absolute solvation free energy for K^+ (−80.6 kcal/mol), and relative free energies for Na^+ and Li^+ using FEP/MD simulations. The hydration free energy of K^+ relative to Na^+ was calculated from FEP/MD and is equal to 18.0 kcal/mol. The hydration free energy of Na^+ relative to Li^+ was calculated from FEP/MD and is equal to 22.7 kcal/mol.]

Contributions	Li^+	Na^+	K^+
$-k_BT \ln[\Sigma_n e^{B_n}]$	−132.6	−109.0	−89.6
$\Delta G_{hs}(R)$	2.7	4.7	5.6
ΔG_i (pQCT)	−129.9	−104.3	−84.1
ΔG_i (pQCT, FEP/MD)	−122.0	−93.8	⋯
ΔG_i (FEP/MD)	−121.3	−98.6	−80.6
$\Delta\Delta G_{ij}$ (pQCT)	25.6	20.3	
$\Delta\Delta G_{ij}$ (FEP/MD)	22.7	18.0	
$\Delta\Delta G_{ij}$ (MD/trv)	23.6	19.3	
$\Delta\Delta G_{ij}$ (pQCT, FEP/MD)	28.1	⋯	

are 4.16, 5.78, and 6.27 for Li^+, Na^+, and K^+, respectively. The corresponding average coordination numbers calculated from pQCT via Eq. (26) are 4.02, 5.70, and 4.98, respectively. For Li^+ and Na^+, the average coordination number is slightly underestimated by pQCT, but of reasonable accuracy. However, in the case of K^+, the discrepancy between pQCT and MD is more important. The correct average coordination number with an inner shell of 3.2 Å for the current K^+ model is 6.27 according to MD, whereas pQCT predicts a value of 4.98, which is lower by 1.3 water molecules. The coordination probabilities given in Tables I–III reveal that the coordination structure arising from pQCT overemphasizes a single or a few coordinate states for all ions. For Li^+ and Na^+, the most probable coordinate state, $n=4$ and $n=6$, respectively, are correctly reproduced by pQCT, although the distributions of coordinated states are slightly too narrow; this is made clearer by comparing the log of $P_i(n)$ for the ions. The situation is worse for K^+, which has a very diffuse coordination structure with a $P_i(n)$ extending up to $n=10$ (Fig. 5). Based on MD, the most probably state is $n=6$, but is incorrectly predicted to be $n=5$ by pQCT. Furthermore, the distribution $P_i(n)$ from pQCT for K^+ shown in Fig. 5 differs significantly from the MD result. It is somewhat surprising that the hydration free energy of K^+ is in relatively good agreement with the result of FEP/MD even though the most probable coordination state and the average coordination number are incorrectly predicted by pQCT. This suggests that the hydration free energy of ions is relatively insensitive to errors affecting the different terms in the cluster sum. In view of these limitations, it appears that the conclusions of previous pQCT studies that have sought to explain ion selectivity in protein binding sites primarily in terms of coordination numbers may need to be reassessed.[15,16] Interestingly, the inaccuracy in $P_i(n)$ does not considerably affect the magnitude of the hydration free energies. As first noted by Merchant and Asthagiri,[11] this is explained by the thermodynamic dominance of the low coordination number contributions to ΔG_i. For Li^+, the hydration free energy is essentially converged by truncating the cluster sum at $n=4$. Similarly, truncating the sum at $n=4$ underestimates the hydration free energy of Na^+ or K^+ by about only 3%. This shows that in the context of pQCT, the total hydration free energy is strongly dominated by the low coordination number.

While the prediction of absolute hydration free energies is important, the relative free energy between ions is often of greater significance in understanding the factors governing ion selectivity in biological systems.[29] In this regard, it may be interesting to combine coordination probabilities from MD simulations with a pQCT-like free energy calculation based on a translational/rotational/vibrational (trv) gas phase partition function for the small cluster. Equation (4) can be easily reformulated to express the relative free energy between ions i and j;[21] $[\Delta G_j - \Delta G_i] = [\Delta G_j(n) - \Delta G_i(n)] - k_BT \ln[P_i(n)/P_j(n)]$. Carrying this analysis with Eq. (15) would lead to equivalent results. Because the coordinate state $n=4$ is probably the one that is the best described by the small oscillation approximation, this suggests combining the observed coordinated state from straight MD to evaluate the coordination probabilities, with a small oscillation approximation (via K_n^{trv}) for a small cluster. As shown in Table V, the relative free energies with this combined MD/trv are slightly more accurate than those with pQCT. Nevertheless, it seems unlikely that pQCT can serve as a reliable framework for meaningful studies of biological systems, where ion selectivity is typically controlled by a few kcal/mol.

VI. CONCLUSION

The framework provided by QCT supports the general notion that ion solvation in complex liquids can largely be understood in terms of local interactions. Ion solvation is expressed as a series of cluster configurational integrals comprising of one ion and a small number of solvent molecules under the influence of the remaining bulk solvent. While it is intuitively reasonable to expect that the influence of more distant solvent molecules be understood with simple mean-field approximations, QCT provides a powerful conceptual framework to make those ideas more quantitative.[6,7,9–11] For example, the importance of local factors has also been highlighted to explain ion binding selectivity in proteins and ion channels,[29,30,39,40] and a QCT-like framework might help shed some light on the underlying microscopic mechanisms

for these systems. However, an extension of QCT designed to account for ion binding to protein sites would be required to pursue those ideas further.

While the framework provided by QCT is a mathematical rewriting of configurational integrals that is strictly exact, it is also of interest to derive approximate forms of the theory that are amenable to practical calculations. The so-called pQCT is one such approximate form, whereby it is assumed that the solvent molecules occupying the first coordination shell only undergo small thermal fluctuations around a well-defined energy minimum. pQCT provides a computationally tractable route with closed-form expressions to calculate the solvation free energy of ions in bulk liquids. Here, classical FEP/MD simulations were used to examine and illustrate the accuracy of pQCT when applied to ion hydration free energy.

Once the radius of the inner-shell region has been chosen, the three essential elements of pQCT are (1) geometry optimization of isolated clusters of one ion and n water molecules using energy minimization in vacuum, (2) vibrational analysis models of those optimized clusters in vacuum, and (3) solvation from the outer shell based on the cluster frozen in their optimized configuration. While (1) and (2) can be carried out using *ab initio* quantum mechanical approaches, (3) can be carried out either using a dielectric continuum for the outer region, or with free energy simulations with explicit solvent molecules. It seems likely that the latter option offers the best avenue to make the most of pQCT, as indicated by the spectacular agreement between the results from pQCT and the FEP/MD calculations in the case of Li^+. There are more uncertainties and inaccuracies when the outer-region contribution is determined from a simple continuum dielectric model.

In summary, the present analysis suggests that pQCT can serve as a reasonable and useful semiquantitative framework, although the results should be interpreted with caution. An important point to note is that pQCT does not present a uniform and controlled approximation when considering ions of different size. For example, while Li^+ is represented with quantitative accuracy, Na^+ is described less accurately, and K^+ encounters even more difficulties. The distribution of coordinated states also appears to encounter serious problems in the case of a large ion such as K^+. Such inaccuracies probably have their origin with the diffuse coordination structure of the larger cation, which is not represented as well by a small set of energy-minimized clusters.

These observations, obtained from computations based on classical potential functions, broadly outline the strengths and weaknesses of pQCT. Of course, the real practical use of pQCT is to provide a framework for treating the interactions within the inner shell at the *ab initio* level. A careful application of such quantum mechanical pQCT strategy (pQCT/QM) may yield useful thermodynamic solvation data for a wide range of molecular systems not readily available from experiments. Systems of interest may include various types of monatomic and molecular neutral or ionic solutes immersed in different kinds of complex liquids. Assessing the overall accuracy of an application of pQCT/QM to any arbitrary system is difficult. Unanticipated system-dependent problems are possible and the quantitative accuracy of the results from pQCT should not be taken for granted. In view of this uncertainty, the best and safest course of action is to test the validity of the framework for the system of interest by verifying if the results from pQCT are consistent with those from explicit solvent FEP/MD simulations in the context of classical potential functions.

ACKNOWLEDGMENTS

The work was funded by Grant No. GM-62342 from the National Institutes of Health (NIH). Helpful discussion with Sergei Y. Noskov is gratefully acknowledged.

[1] H. Bethe, Proc. R. Soc. London, Ser. A **150**, 552 (1935).
[2] E. Guggenheim, Proc. R. Soc. London, Ser. A **148**, 304 (1935).
[3] J. Kirkwood, J. Chem. Phys. **8**, 623 (1940).
[4] H. Bethe and J. Kirkwood, J. Chem. Phys. **7**, 578 (1939).
[5] R. Fowler and E. Guggenheim, Proc. R. Soc. London, Ser. A **174**, 189 (1940).
[6] L. R. Pratt Randall and R. Laviolette, Mol. Phys. **94**, 909 (1998).
[7] L. Pratt, T. Beck, and M. Paulatis, *The Potential Distribution Theorem and Models of Molecular Solutions* (Cambridge University Press, Cambridge, 2006).
[8] D. Asthagiri, P. D. Dixit, S. Merchant, M. E. Paulaitis, L. R. Pratt, S. B. Rempe, and S. Varma, Chem. Phys. Lett. **485**, 1 (2010).
[9] D. Asthagiri, L. Pratt, and H. Ashbaugh, J. Chem. Phys. **119**, 2702 (2003).
[10] D. M. Rogers and T. L. Beck, J. Chem. Phys. **129**, 134505 (2008).
[11] S. Merchant and D. Asthagiri, J. Chem. Phys. **130**, 195102 (2009).
[12] S. Rempe, L. R. Pratt, G. Hummer, J. D. Kress, R. L. Martin, and A. Redondo, J. Am. Chem. Soc. **122**, 966 (2000).
[13] S. Rempe and L. Pratt, Fluid Phase Equilib. **183–184**, 121 (2001).
[14] S. Rempe, D. Asthagiri, and L. Pratt, Phys. Chem. Chem. Phys. **6**, 1966 (2004).
[15] S. Varma and S. Rempe, Biophys. J. **93**, 1093 (2007).
[16] S. Varma and S. B. Rempe, J. Am. Chem. Soc. **130**, 15405 (2008).
[17] T. W. Whitfield, S. Varma, E. Harder, G. Lamoureux, S. B. Rempe, and B. Roux, J. Chem. Theory Comput. **3**, 2068 (2007).
[18] G. Lamoureux and B. Roux, J. Phys. Chem. B **110**, 3308 (2006).
[19] Y. Marcus, *Ion Solvation* (Wiley, Chichester, 1985).
[20] M. D. Tissandier, K. A. Cowen, W. Y. Feng, E. Gundlach, M. H. Cohen, A. D. Earhart, T. R. Tuttle, Jr., and J. V. Coe, J. Phys. Chem. A **102**, 9308 (1998).
[21] H. Yu, S. Y. Noskov, and B. Roux, J. Phys. Chem. B **113**, 8725 (2009).
[22] W. Im, S. Bernèche, and B. Roux, J. Chem. Phys. **114**, 2924 (2001).
[23] H. Woo, A. Dinner, and B. Roux, J. Chem. Phys. **121**, 6392 (2004).
[24] Y. Q. Deng and B. Roux, J. Chem. Phys. **128**, 115103 (2008).
[25] G. Soto-Campos, D. S. Corti, and H. Reiss, J. Chem. Phys. **108**, 2563 (1998).
[26] W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, and M. L. Klein, J. Chem. Phys. **79**, 926 (1983).
[27] Y. Deng and B. Roux, J. Phys. Chem. **108**, 16567 (2004).
[28] J. D. Weeks, D. Chandler, and H. C. Andersen, J. Chem. Phys. **54**, 5237 (1971).
[29] S. Noskov, S. Berneche, and B. Roux, Nature (London) **431**, 830 (2004).
[30] S. Noskov and B. Roux, J. Gen. Physiol. **129**, 135 (2007).
[31] D. Beglov and B. Roux, J. Chem. Phys. **100**, 9050 (1994).
[32] E. Harder and B. Roux, J. Chem. Phys. **129**, 234706 (2008).
[33] J. Ryckaert, G. Ciccotti, and H. Berendsen, J. Comput. Chem. **23**, 327 (1977).
[34] B. R. Brooks, C. L. Brooks III, A. D. Mackerell, Jr., L. Nilsson, R. J. Petrella, B. Roux, Y. Won, G. Archontis, C. Bartels, S. Boresch, A. Caflisch, L. Caves, Q. Cui, A. R. Dinner, M. Feig, S. Fischer, J. Gao, M. Hodoscek, W. Im, K. Kuczera, T. Lazaridis, J. Ma, V. Ovchinnikov, E. Paci, R. W. Pastor, C. B. Post, J. Z. Pu, M. Schaefer, B. Tidor, R. M. Venable, H. L. Woodcock, X. Wu, W. Yang, D. M. York, and M. Karplus, J. Comput. Chem. **30**, 1545 (2009).

[35] M. Nina, D. Beglov, and B. Roux, J. Phys. Chem. B **101**, 5239 (1997).
[36] W. Im, D. Beglov, and B. Roux, Comput. Phys. Commun. **111**, 59 (1998).
[37] H. Yu and B. Roux, Biophys. J. **97**, L15 (2009).
[38] B. Roux, H. Yu, and M. Karplus, J. Phys. Chem. **94**, 4683 (1990).
[39] S. Noskov, S. Berneche, and B. Roux, Biophys. Chem. **124**, 279 (2006).
[40] D. Asthagiri, L. Pratt, and M. Paulaitis, J. Chem. Phys. **125**, 024701 (2006).