Investigation of the dominant hydration structures among the ionic species in aqueous solution: Novel quantum mechanics/molecular mechanics simulations combined with the theory of...

ARTICLE *in* THE JOURNAL OF CHEMICAL PHYSICS · MARCH 2008

Impact Factor: 2.95 · DOI: 10.1063/1.2825600 · Source: PubMed

7 AUTHORS, INCLUDING:

Ryohei Kishi
Osaka University
110 PUBLICATIONS **1,948** CITATIONS
SEE PROFILE

Masayoshi Nakano
Osaka University
337 PUBLICATIONS **4,785** CITATIONS
SEE PROFILE

Investigation of the dominant hydration structures among the ionic species in aqueous solution: Novel quantum mechanics/molecular mechanics simulations combined with the theory of energy representation

Hideaki Takahashi,[1,a)] Hajime Ohno,[1] Toshihiko Yamauchi,[1] Ryohei Kishi,[1] Shin-ichi Furukawa,[1] Masayoshi Nakano,[1] and Nobuyuki Matubayasi[2]
[1]*Division of Chemical Engineering, Department of Materials Engineering Science, Graduate School of Engineering Science, Osaka University, Toyonaka, Osaka 560-8531, Japan*
[2]*Institute for Chemical Research, Kyoto University, Uji, Kyoto 611-0011, Japan*

(Received 29 May 2007; accepted 27 November 2007; published online 13 February 2008)

In the present work, we have performed quantum chemical calculations to determine preferable species among the ionic complexes that are present in ambient water due to the autodissociation of water molecule. First, we have formulated the relative population of the hydrated complexes with respect to the bare ion (H_3O^+ or OH^-) in terms of the solvation free energies of the relevant molecules. The solvation free energies for various ionic species (H_3O^+, $H_5O_2^+$, $H_7O_3^+$, $H_9O_4^+$ or OH^-, $H_3O_2^-$, $H_5O_3^-$, $H_7O_4^-$, $H_9O_5^-$), categorized as proton or hydroxide ion in solution, have been computed by employing the QM/MM-ER method recently developed by combining the quantum mechanical/molecular mechanical (QM/MM) approach with the theory of energy representation (ER). Then, the computed solvation free energies have been used to evaluate the ratio of the populations of the ionic complexes to that of the bare ion (H_3O^+ or OH^-). Our results suggest that the Zundel form, i.e., $H_5O_2^+$, is the most preferable in the solution among the cationic species listed above though the Eigen form ($H_9O_4^+$) is very close to the Zundel complex in the free energy, while the anionic fragment from water molecules mostly takes the form of OH^-. It has also been found that the loss of the translational entropy of water molecules associated with the formation of the complex plays a role in determining the preferable size of the cluster. © *2008 American Institute of Physics*. [DOI: 10.1063/1.2825600]

I. INTRODUCTION

Autoionization of water molecule or acidic (basic) substrate is a ubiquitous chemical event in the aqueous solution[1] and the resulting excess proton or hydroxide ion is known to play a key role in determining the pathway of reaction occurring in the solution or in the biological systems immersed in water. The population of proton in solution, often quantified by *p*H, can be readily obtained by experiments. However, theoretical determination of *p*H is known to be a demanding task.[2–5] The difficulty mainly arises from the fact that it requires much computational cost to compute free energy change associated with chemical reactions in condensed phases. Another situation that complicates the theoretical determination of *p*H is that ionic species categorized as proton or hydroxide ion would take several hydration structures in aqueous solution corresponding to the number of water molecules bonded directly to H_3O^+ or OH^- ions.[5,6] More specifically, proton takes the structure expressed as $H_3O^+(H_2O)_{n-1}$, while hydroxide ion is in the form of $OH^-(H_2O)_{m-1}$ in aqueous solution. Since the value of *p*H is dominated by the concentration of the most probable complex among the ionic species, it is of particular importance to know the size of the hydrated complex that has the largest population in the solution. The populations of these ionic complexes in solution are determined by the solvation free energies of the relevant molecules as well as the binding free energies for the complexes.

The molecular species $H_3O^+(H_2O)$ and $H_3O^+(H_2O)_3$ are often referred to as Zundel or Eigen cations, respectively. There have been a number of works that investigated the solvation of these molecules in aqueous media in an attempt to understand the properties of the excess proton in water.[6–11] It can be readily understood that the Zundel cation $H_5O_2^+$ embedded in bulk water is smoothly transformed to Eigen cation $H_9O_4^+$ without hydrogen-bond rearrangement around the Zundel cation. It clearly suggests that there is no distinction between the Zundel and Eigen cations as well as other cationic species listed above and they are no more than limiting concepts.[11] Hence, the excess proton can only be represented by the mixture of these limiting states. The same discussion also holds for anionic species. The empirical valence bond[12] (EVB) or multistate (MS)-EVB method[13–15] is one of the effective solutions that enable one to simulate such a situation. In the MS-EVB model, the limiting states are chosen as the basis to express the instantaneous distribution of the excess proton and the expansion coefficients for the basis are obtained by solving the eigenvalue problem with respect to a set of empirical potentials that defines the reactive system. The (MS-)EVB approach makes it possible to describe the realistic state of a proton in water, the intermediate state between Zundel and Eigen cations, for in-

a)Author to whom correspondence should be addressed. Electronic mail: takahasi@cheng.es.osaka-u.ac.jp.

stance. Of course, such description of the excess proton is valuable, however, it is also true that the concept of the limiting state helps us have a physical intuition on the proton solvation. Hence, in the present work, an emphasis will be placed on the static properties of the pivotal chemical species rather than the delocalized nature of proton that hops from one molecule to another.

The quantum chemical calculation is essential for the theoretical description of chemical processes such as bond dissociations or formations. The computational cost for the first-principles approach, however, increases in the order of N^3–N^4 even at the lowest level of theory.[16] Hence, we inevitably encounter a difficulty in the theoretical investigation for the reaction involving a large number of particles. Besides the problem in the quantum chemical calculations, it is also heavily demanding to compute free energy change for the process in a condensed system by utilizing the method of molecular simulations.[17] Therefore, we have to conquer simultaneously the difficulties arising from the execution of the quantum chemical calculation and the free energy calculation to study chemical reactions occurring in solution. The polarizable continuum model (PCM) incorporates the solvent effects in the self-consistent field (SCF) procedure in the quantum chemical calculation by modeling the solvent as a uniform dielectric continuum.[18] The method is quite efficient and has been extensively utilized to study solvation effects; however, it is obviously poor in representing the short-range interactions such as hydrogen bonds. The reference interaction site model (RISM) combined with the SCF calculation[19] (RISM-SCF) is a powerful integral equation method to compute simultaneously the electronic wave function of the solute and the set of site-site radial distribution functions for the solvent from which solvation free energy can be obtained. In the framework of the RISM-SCF, the solvent is considered as a structured environment in contrast to the PCM approach. However, the electric field due to the solvent is reduced to a set of effective charges placed on the nuclei of the solute in constructing the solvated Fock operator. In the Car-Parrinello approach,[20] the forces acting on the nuclei are fully determined by solving the Schrödinger equation for all electrons contained in the system. Although the approach is in principle robust in representing the solute-solvent interactions including chemical bondings, it is computationally intensive to accomplish the free energy calculation. Thus, there is still a need to develop an efficient and computationally modest approach for the study of chemical reactions in condensed systems.

The hybrid quantum mechanical/molecular mechanical (QM/MM) approach is a promising computational technique[21] that allows one to construct the electronic wave function of a solute under the influence of the environment described by classical mechanics. The QM/MM method has been successfully applied to various systems and the efficiency has been well established.[22] In a recent development, Takahashi *et al.* proposed a QM/MM approach[23–25] that employs the real-space grids (RS-QM/MM) to express the one-electron wave functions[26–28] for the Kohn-Sham density functional theory,[29,30] which can afford high efficiency in the parallel computations on a workstation cluster with common network facilities.[31] Further, Takahashi *et al.* developed a novel approach[32] by combining the RS-QM/MM approach with the theory of energy representation (QM/MM-ER) for the purpose of computing free energy change within a reasonable computational cost. In the framework of the theory of energy representation,[33–35] as differed from the conventional theory of solution,[36] the free energy change is expressed in terms of the distribution functions of the solute-solvent interaction potential energy instead of the spatial distribution of the solvent around the solute. It should be emphasized that the concept of the interaction site is no longer needed when one constructs the distribution functions on the energy coordinate. Hence, it is straightforward to take into account the spatially diffuse nature of the electron density of the solute, which is inherent in the quantum mechanical object, by utilizing the QM/MM-ER approach. In the previous papers, we examined the efficiency and the accuracy of the method by performing simulations for polar solutes in aqueous solution and found that the QM/MM-ER approach is adequate enough to reproduce a free energy change associated with a chemical reaction[37–44] and also a slight free energy difference between two possible conformations of a solute in aqueous solution, in substantial agreement with experimental results.[45]

In the present work, we determine the preferable forms for the hydrated ionic complexes in aqueous solution by utilizing the QM/MM-ER method. The organization of the paper is as follows. In Sec. II, we derive the general expression for the dissociation constant of water by considering the contributions from various types of ionic species, from which the population of each complex is determined. We also make a brief review for the QM/MM-ER approach. Section III is devoted to mentioning the computational details for the QM/MM-ER simulations as well as for the geometry optimizations of the solutes using first-principles calculations. In Sec. IV, the results and discussions are presented.

II. THEORY AND METHODOLOGY

Dissociation constant K_w of water can be decomposed into the contributions from ions with various types of hydrations in solution. Hence, we can determine the populations of the individual ionic species by formulating an expression for K_w given as contributions from these ions. In Sec. II A, we introduce the equilibrium constants K_n between ion (H_3O^+ or OH^-) and its hydrated forms in terms of the free energy difference between two states in solution. In Sec. II B, K_w is derived by using these equilibrium constants K_n. Section II C will be devoted to a brief review of the QM/MM-ER approach which is employed to compute solvation free energies of the related ions.

A. Equilibrium for ionic complex

Prior to the investigation for the population of the ionic species produced by the dissociation of water molecule, we consider here a chemical process where an ion makes a hydrated complex in solution,

$$A + (n-1)H_2O \leftrightarrow A(H_2O)_{n-1}, \tag{1}$$

where A stands for H_3O^+ or OH^- and n is a natural number ($n=1,2,3,\ldots$). Equation (1) represents an equilibrium between the hydrated complex of the ion A and the situation where the constituent molecules are separately present in solution. In the following we derive an equation for the equilibrium constant $K_n(A)$. At the equilibrium, the chemical potential of the left hand side is equal to that of the right; thus,

$$\mu(A) + (n-1)\mu(H_2O) = \mu(A(H_2O)_{n-1}), \tag{2}$$

where $\mu(X)$ denotes the chemical potential for the solute X [$X=A$, H_2O, $A(H_2O)_{n-1}$]. Further, $\mu(X)$ can be decomposed into the contributions

$$\mu(X) = \mu_{\text{trans}}(X) + \mu_{\text{vibrot}}(X) + \mu_{\text{solv}}(X) + E_{\text{ZPE}}(X) + E_0(X) + k_BT \ln[X], \tag{3}$$

where the free energy components in the right hand side are in order due to translation, vibration-rotation, solvation, zero-point vibrational energy, and electronic energy at isolation of the solute X. In Eq. (3), k_BT is the Boltzmann constant multiplied by temperature T and the brackets indicate the molar concentration. The translational free energy $\mu_{\text{trans}}(X)$ can be expressed by

$$\mu_{\text{trans}}(X) = k_BT \ln(\lambda(X)), \tag{4}$$

where $\lambda(X)$ is the third power of the thermal de Broglie wavelength and explicitly given by

$$\lambda(X) = \left(\frac{h^2}{2\pi m_X k_B T}\right)^{3/2}, \tag{5}$$

where h is the Planck constant and m_X is the mass of the solute X. By substituting Eq. (3) into Eq. (2), we obtain

$$\ln \frac{[A(H_2O)_{n-1}]}{[A][H_2O]^{n-1}} = \ln \frac{\lambda(A)\lambda(H_2O)^{n-1}}{\lambda(A(H_2O)_{n-1})} - \frac{\Delta\mu_n(A)}{k_BT}. \tag{6}$$

In Eq. (6), $\Delta\mu_n(A)$ is defined as the difference of the sum of the free energy components except for the first and the last term in Eq. (3); thus,

$$\begin{aligned}\Delta\mu_n(A) = {} & \{\mu_{\text{vibrot}}(A(H_2O)_{n-1}) + \mu_{\text{solv}}(A(H_2O)_{n-1}) \\ & + E_{\text{ZPE}}(A(H_2O)_{n-1}) + E_0(A(H_2O)_{n-1})\} \\ & - \{(n-1)(\mu_{\text{vibrot}}(H_2O) + \mu_{\text{solv}}(H_2O) \\ & + E_{\text{ZPE}}(H_2O) + E_0(H_2O)) + (\mu_{\text{vibrot}}(A) \\ & + \mu_{\text{solv}}(A) + E_{\text{ZPE}}(A) + E_0(A))\}.\end{aligned} \tag{7}$$

Then, the equilibrium constant $K_n(A)$ for Eq. (1) becomes

$$K_n(A) = \frac{[A(H_2O)_{n-1}]}{[A][H_2O]^{n-1}} = \frac{\lambda(A)\lambda(H_2O)^{n-1}}{\lambda(A(H_2O)_{n-1})} \exp\left(-\frac{\Delta\mu_n(A)}{k_BT}\right). \tag{8}$$

Equation (8) suggests that the probability that ion A makes a hydrated complex is affected not only by the free energy difference $\Delta\mu_n(A)$ given by Eq. (7) but also by the translational free energy of water molecules. As the number of water molecules involved in the complex increases, the free energy of complex formation increases since the translational degrees of freedom for the water molecules are lost through the formation of an aggregate.

B. Dissociation constant of water

In this subsection, we derive the general expression for the dissociation constant K_w of water which takes into consideration the existence of various kinds of ionic complexes under the equilibrium of Eq. (1). Then, K_w can be decomposed into contributions due to the individual dissociation reactions corresponding to the number of water molecules involved.

The ionic species referred to as *proton* H^+ stands for the molecular complexes having the form expressed by $H_3O^+(H_2O)_{n-1}$ in aqueous solution. Therefore, the concentration of proton $[H^+]$ can be given by the sum of those of the individual species,

$$[H^+] = \sum_{n=1} [H_3O^+(H_2O)_{n-1}]. \tag{9}$$

It should be noted that in the development of Eq. (9), we exclude the possibility that a bare H^+, which is a proton with no chemical bonds, exists in the solution. This assumption can be validated by the fact that ionic dissociation of a water molecule requires extremely high activation energy in gaseous phase. Analogous to Eq. (9), the concentration of the hydroxide ion OH^- is regarded as the sum of those of the hydrated complexes; thus,

$$[OH^-] = \sum_{m=1} [OH^-(H_2O)_{m-1}]. \tag{10}$$

By substituting Eqs. (9) and (10) into the definition of the dissociation constant K_w of water, we obtain

$$\begin{aligned}K_w \equiv {} & \frac{[H^+][OH^-]}{[H_2O]^2} \\ = {} & \frac{[H_3O^+][OH^-]}{[H_2O]^2} \sum_{n=1} K_n(H_3O^+)[H_2O]^{n-1} \\ & \times \sum_{m=1} K_m(OH^-)[H_2O]^{m-1},\end{aligned} \tag{11}$$

where the equilibrium constants $K_n(H_3O^+)$ and $K_m(OH^-)$ are explicitly given by Eq. (8). In Eq. (11), the term $H_0 \equiv [H_3O^+][OH^-]/[H_2O]^2$ expresses the equilibrium constant for the ionic dissociation where only two water molecules are involved. H_0 is written as

$$H_0 = \frac{\lambda(H_2O)^2}{\lambda(H_3O^+)\lambda(OH^-)} \exp\left(-\frac{\Delta G_0}{k_BT}\right). \tag{12}$$

where ΔG_0 is the free energy change for the reaction excluding the translational contribution. For clarity, we further make definitions that $I_n \equiv K_n(H_3O^+)[H_2O]^{n-1}$ and $J_m \equiv K_m(OH^-)[H_2O]^{m-1}$. From Eq. (8), it can be readily understood that I_n indicates the ratio of the concentration of $H_3O^+(H_2O)_{n-1}$ to that of H_3O^+ and, analogously, J_m is the ratio of $OH^-(H_2O)_{m-1}$ with respect to OH^-. By making use of these definitions, we rewrite Eq. (11) as

$$K_w = H_0 \sum_{n=1} I_n \sum_{m=1} J_m. \tag{13}$$

Equation (13) clearly shows that the dissociation constant K_w of water is given as the sum of the contributions from ionic species $H_3O^+(H_2O)_{n-1}$ and $OH^-(H_2O)_{m-1}$.

Here we introduce the specific dissociation constant $K_w(n,m)$ defined as

$$K_w(n,m) = H_0 I_n J_m / [H_2O]^{n+m-2} = K'_w(n,m)/[H_2O]^{n+m-2}, \tag{14}$$

where $K'_w(n,m)$ is given by $H_0 I_n J_m$. It is easy to see that $K_w(n,m)$ is the equilibrium constant corresponding to the specific dissociation reaction involving $(n+m)$ water molecules,

$$(n+m)H_2O \leftrightarrow H_3O^+(H_2O)_{n-1} + OH^-(H_2O)_{m-1}. \tag{15}$$

We define the dissociation free energy $\Delta G_{diss}(n,m)$ for Eq. (15) by using $K_w(n,m)$; thus,

$$\Delta G_{diss}(n,m) = -k_B T \ln K_w(n,m). \tag{16}$$

From Eqs. (13) and (14), K_w can be expressed in terms of $\Delta G_{diss}(n,m)$ as

$$\begin{aligned} K_w &= \sum_{n,m=1} K'_w(n,m) \\ &= \sum_{n,m=1} K_w(n,m)[H_2O]^{n+m-2} \\ &= \sum_{n,m=1} \exp\left(-\frac{\Delta G_{diss}(n,m)}{k_B T}\right)[H_2O]^{n+m-2} \\ &= \sum_{n,m=1} \exp\left(-\frac{\Delta G'_{diss}(n,m)}{k_B T}\right). \end{aligned} \tag{17}$$

In Eq. (17), $\Delta G'_{diss}(n,m)$ is defined as $\Delta G'_{diss}(n,m) = -kT \ln K'_w(n,m)$. K_w is given by the sum of the Boltzmann factors for the free energy changes $\Delta G'_{diss}(n,m)$. One should bear in mind that a slight difference in the free energy results in a significant difference in the population. For instance, the ratio of the Boltzmann factor is estimated to be ~5.3 at 300 K when the difference in $\Delta G'_{diss}(n,m)$ is only about 1 kcal/mol. Thus, the dominant ionic species specified by n and m govern the dissociation constant K_w.

C. QM/MM-ER approach

The computation of the solvation free energy $\mu_{solv}(X)$ in Eq. (3) is the key that determines the accuracy of $K_w(n,m)$ in Eq. (14), for which we perform a series of simulations referred to as the QM/MM-ER approach,[32] which combines the quantum chemical approach with the theory of energy representation. Within the framework of the standard theory of energy representation, however, it is assumed that the solute-solvent interaction is pairwise. Hence, a treatment is required when we combine the QM/MM with the method of energy representation since a quantum mechanical object inherently involves many-body effects in the interaction with surrounding solvent molecules. The point of the method is to divide the solvation free energy into the terms due to the two-body interactions and the remaining minor contribution due to the many-body effect. We briefly give the outline of the methodology below.

In the QM/MM approach, total energy E of the system is decomposed into three terms;[21] thus,

$$E = E_{QM} + E_{QM/MM} + E_{MM}, \tag{18}$$

where E_{QM} is the electronic energy of the QM subsystem including nuclear repulsion energy and E_{MM} is the contribution from the MM subsystem. $E_{QM/MM}$ in Eq. (18) indicates the interaction potential between the QM and MM subsystems. Then, the solvation free energy $\Delta\mu$ of the QM solute can be expressed as

$$\exp(-\beta\Delta\mu) = \frac{\int d\mathbf{X} \exp(-\beta(E_{dist} + E_{QM/MM}(n(\mathbf{r}),\mathbf{X}) + E_{MM}(\mathbf{X})))}{\int d\mathbf{X} \exp(-\beta E_{MM}(\mathbf{X}))}, \tag{19}$$

where $n(\mathbf{r})$ is the instantaneous electron density of the QM solute at position $\mathbf{r}$, $\mathbf{X}$ denotes collectively the set of molecular configurations $\{\mathbf{x}_i\}$ of the solvent, and E_{dist} is the distortion energy of the QM subsystem. E_{dist} is given by the energy E_{QM} subtracted by the energy of the solute at isolation. β in Eq. (19) stands for the reciprocal of the Boltzmann constant k_B multiplied by T. As was derived in the previous paper,[32] the solvation free energy $\Delta\mu$ can be decomposed into three terms; thus,

$$\Delta\mu = \bar{E} + \Delta\bar{\mu} + \delta\mu, \tag{20}$$

where $\bar{E}$ is the average of the distortion energy of the wave function in solution ($\bar{E} = \langle E_{dist} \rangle$), $\Delta\bar{\mu}$ is the solvation free energy of the solute with electron density fixed at its average distribution $\bar{n}(\mathbf{r})$ in solution, and $\delta\mu$ is the free energy change due to the electron density fluctuation around the density $\bar{n}(\mathbf{r})$. $\bar{E}$ and $\bar{n}(\mathbf{r})$ are simply the statistical averages and can be obtained through a QM/MM simulation. The free energy $\Delta\bar{\mu}$ can be expressed as

$$\exp(-\beta\Delta\bar{\mu}) = \frac{\int d\mathbf{X} \exp(-\beta(E_{QM/MM}(\bar{n}(\mathbf{r}),\mathbf{X}) + E_{MM}(\mathbf{X})))}{\int d\mathbf{X} \exp(-\beta E_{MM}(\mathbf{X}))}. \tag{21}$$

In the computation of $\Delta\bar{\mu}$, it should be noted that the solute-solvent interaction energy is pairwise additive since the electron density of the QM solute is fixed at $\bar{n}(\mathbf{r})$ and no longer corresponds to the solvent configuration. Hence, the standard version of theory of energy representation[33] can be applied to the computation of $\Delta\bar{\mu}$. The remaining minor contribution $\delta\mu$ due to the many-body effect of the QM object can be computed separately by the theory of energy representation by conducting additional QM/MM simulations.[32]

Within the framework of the theory of energy representation,[33–35] the distribution functions of the solute-solvent interaction energy, instead of the spatial distributions, serve as fundamental variables to determine excess chemical potential of the solute. The pure solvent (reference) system is defined as the solution where the solute-solvent interaction is switched off. The energy distribution function $\rho_0(\varepsilon)$ for the

pure solvent system is obtained by the solute insertion into the solvent as a test particle. On the other hand, the solution system is the solution of interest and the distribution function $\rho(\varepsilon)$ is constructed through a QM/MM simulation under the existence of the solute-solvent interaction. In terms of the energy distribution functions, the free energy $\Delta\bar{\mu}$ of Eq. (21) can be exactly given by

$$\Delta\bar{\mu} = -k_B T \int d\varepsilon \Bigg[(\rho(\varepsilon) - \rho_0(\varepsilon)) + \beta\omega(\varepsilon)\rho(\varepsilon) - \beta\left(\int_0^1 d\lambda\, \omega(\varepsilon;\lambda) \right)(\rho(\varepsilon) - \rho_0(\varepsilon)) \Bigg], \tag{22}$$

where the energy distribution functions are constructed with respect to the solute-solvent interaction energy $\nu(\tilde{n}, \mathbf{x}_i)$ between the QM solute with electron density $\tilde{n}(\mathbf{r})$ and the solvent molecule. In Eq. (22), $\omega(\varepsilon)$ is the indirect part of the solute-solvent potential of mean force and λ is the coupling parameter for the gradual insertion of the solute in the solvent. In the development of Eq. (22), it is assumed that the solute-solvent interaction is chosen so that the corresponding energy distribution varies linearly from $\rho_0(\varepsilon)$ to $\rho(\varepsilon)$ with respect to the variation of the coupling parameter λ. In the practical implementation, integration with respect to λ is performed approximately by adopting a hybrid functional of PY and HNC,[36] for which the correlation matrix $\chi_0(\varepsilon, \eta)$ on the energy coordinate must be constructed in the reference system. Space limitations do not allow us to give the explicit form of the functional in the energy representation. We refer the readers to the previous papers for a complete description.[32,34] The cavitation free energy is taken into account through the overlapping of the solvent molecules on the solute's exclusion volume in the reference system. Such overlaps give rise to a nonzero $\rho_0(\varepsilon)$ at the large energy coordinate where the distribution $\rho(\varepsilon)$ completely vanishes. Note that the value of the energy coordinate related to the exclusion volume is arbitrarily large and is physically meaningless since the overlap of molecules is quite unrealistic. However, it can be readily understood that the integrand of Eq. (22) is completely independent of the ε value itself for the region where $\rho(\varepsilon)=0$ is satisfied. Hence, the free energy due to the cavitation of the solute can be soundly computed in the theory of energy representation. In the present work, we neglect the free energy contribution $\delta\mu$ in Eq. (20) due to the electron density fluctuation since cancellation of error is expected to take place when we compute the free energy difference between the product and reactant states in a reaction.

The method of QM/MM-ER offers several advantages in the computation of the free energy changes in condensed phases. Since the solvent molecules with the same energy coordinate are grouped a unit, the convergence of the distribution function on the energy coordinate is much faster as compared with that of the spatial distribution functions. This benefits substantially the computation of the free energy change since the quantum chemical calculation generally requires much computational costs. In addition, the diffuse nature of the realistic electron density can be naturally taken into account as the interaction sites are no longer needed. The accuracy and efficiency of the QM/MM-ER approach were examined well in the previous works.[32,42,43,45]

III. COMPUTATIONAL DETAILS

This section consists of two subsections. In Sec. III A, computational details for the geometry optimizations of the hydrated complexes are presented. In Sec. III B, we describe the details for the QM/MM-ER simulations.

A. Geometry optimizations

Geometries have been optimized by using the GAUSSIAN 03 package[46] for the cationic molecules $[H_3O^+(H_2O)_{n-1}$ $(n=1-4)]$ and also for the anionic ones $[OH^-(H_2O)_{m-1}$ $(m=1-5)]$. The method used is the density functional theory (DFT) with the exchange functional developed by Becke and the correlation functional by Lee *et al.* (BLYP).[47,48] The correlation consistent basis set of Dunning, Jr.[49] with polarization functions augmented by diffuse-type orbitals referred to as aug-cc-pVDZ has been used for the linear combination of atomic orbitals expansions of the one-electron orbitals. The "VeryTight" option in the geometry optimization has been invoked along with the "Int=UltraFine" option in the DFT calculations. We have referred the geometries of H_3O^+, OH^-, and H_2O to those in our previous paper[32,41] obtained by the same manner. The zero-point vibrational energy $E_{ZPE}(X)$ and the free energy $\mu_{vibrot}(X)$ for the solute X in Eq. (3) have been computed by the frequency analysis in the GAUSSIAN 03 package. These calculations have been carried out for the solutes under the gas phase environment on the assumption that the solvation effects on these values will be negligible.

B. QM/MM-ER simulations

To determine the equilibrium constant $K_n(A)$ for ion A $(=H_3O^+$ or $OH^-)$ expressed as Eq. (8), we have to compute the solvation free energy $\mu_{solv}(X)$ in Eq. (3) which is the most time-consuming part in the present work. In the QM/MM-ER approach, the energy distribution functions, which serve as a fundamental variable to determine the free energy change, are obtained by a set of QM/MM simulations. The QM/MM simulations have been performed by conducting our original Kohn-Sham DFT code that utilizes the real-space grids to express the one-electron wave functions. The accuracy of the real-space QM/MM approach has been examined well in the previous work.[25]

We have regarded the cationic molecules $[H_3O^+(H_2O)_{n-1}$ $(n=1-4)]$ and the anionic ones $[OH^-(H_2O)_{m-1}$ $(m=1-5)]$ as the QM molecules in the QM/MM simulations. Each QM molecule has been placed at the center of the QM cell that has been spanned by equally spaced 64 grids in each direction where the grid spacing h has been set at 0.1518 Å. The geometries of the QM solutes have been kept fixed during the QM/MM simulations. The norm-conserving pseudopotentials in the Kleinman-Bylander forms[50] have been used as the effective core potentials for the valence electrons. The time-saving double-grid technique developed by Ono and Hirose[51] has been implemented to realize the rapid behavior of the pseudo-wave-functions near

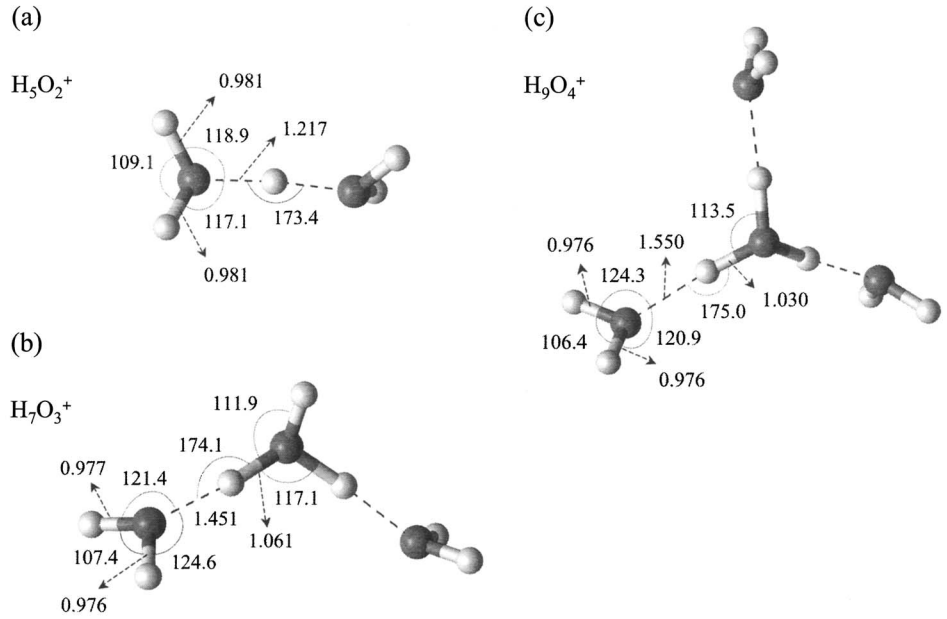

FIG. 1. Optimized geometries (BLYP/aug-cc-pVDZ) for the hydrated cationic complexes [(a) $H_5O_2^+$, (b) $H_7O_3^+$, and (c) $H_9O_4^+$]. The bond lengths are given in units of angstrom, while angles are in units of degree.

the atomic core regions, for which the grid spacing is set at $h/3$. The kinetic energy operators in the Kohn-Sham DFT have also been described in the real space by employing the fourth-order finite difference expression.[27] To eliminate the periodicity of the QM subsystem, Hartree potential has been computed by the method developed by Barnett and Landman.[52] The exchange and correlation energy has been estimated by the BLYP functional. The electronic wave functions have been updated after every five steps of the molecular dynamics (MD) of the MM subsystem to expedite the simulation to construct the average distribution $\tilde{n}(\mathbf{r})$.

The electronically static water molecules have been treated by classical models. Depending on the size of the QM solute, the solvent has been constructed from the 251–255 TIP4P water molecules[53] which are driven by the leapfrog algorithm[17] under the constant-*NVT* condition. In the TIP4P model, the interaction site of the Lennard-Jones (LJ) potential is placed at the O atom (σ=3.154 Å, ε =0.155 kcal/mol), and the negative point charge q_M is placed near the O atom and the positive point charges q_H at hydrogens (q_M=−1.04, q_H=0.52 in the unit of elementary charge). Specifically the LJ size parameters for the oxygen atoms in H_3O^+ and OH^- molecules included in the QM solutes are adjusted to 2.878 and 2.562 Å, respectively, so that each two-body interaction between the QM solute and a TIP4P molecule reproduce a result of an *ab initio* calculation.[41] Then, the LJ parameters between unlike atoms have been determined simply by the Lorentz-Berthelot mixing rule.[17] The molecular geometries of the MM molecules have been fixed during the simulations as well as the QM

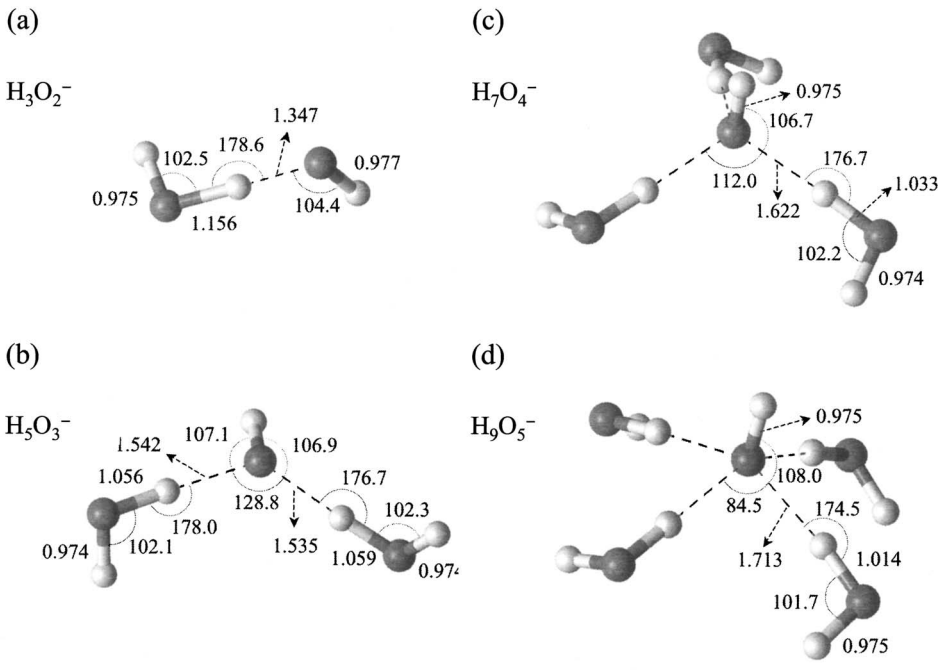

FIG. 2. Optimized geometries (BLYP/aug-cc-pVDZ) for the hydrated anionic complexes [(a) $H_3O_2^-$, (b) $H_5O_3^-$, (c) $H_7O_4^-$, and (d) $H_9O_5^-$]. The bond lengths are given in units of angstrom, while angles are in units of degree.

solutes, which has been realized by using the quaternions. The thermodynamic condition of the solution has been specified by $T=300$ K and $\rho=1.0$ g/cm^3. The time step for the MD simulation has been set at 1.0 fs, and momentum rescaling has been performed after every two MD steps for the temperature control. The long-range interaction between MM molecules due to the Coulomb interaction has been taken into account by the Ewald method.[54] The cut off distance of the LJ potential has been set at half of the box length. In constructing the instantaneous distribution functions, we have considered the solvent water molecules of which oxygens are within a sphere of radius of 9 Å, the center of which has been placed at the weight center of the QM solute. The compensation has been made by the Born equation[55] for the solvation free energy due to water molecules outside the sphere,

$$\mu_{\mathrm{Born}} = -16.59Z^2(1-\varepsilon^{-1})/r, \tag{23}$$

where Z, ε, and r are, respectively, the total charge of the QM subsystem in the unit of elementary charge, the dielectric constant of water at a given thermodynamic condition, and the radius of the sphere containing the ion in nanometers. For the solvation free energies of H_3O^+, OH^-, and H_2O, we have used our previous results given by the QM/MM-ER simulations.[41]

The computational procedures to construct the energy distribution functions are summarized as follows:

(1) At first 5 ps QM/MM simulation has been carried out to equilibrate the system and, subsequently, 50 ps simulation has been performed to obtain the average distortion energy $\bar{E}$ as well as the average distribution of the electron density $\tilde{n}(\mathbf{r})$.
(2) The QM/MM simulations for the QM solute with the average density $\tilde{n}(\mathbf{r})$ have been performed and the energy distribution functions have been constructed for the solution and the pure solvent systems to compute $\Delta\bar{\mu}$ of Eq. (21). The ensemble averages have been taken over 100 ps for the solution and 200 ps for the pure solvent systems after 5 ps equilibration.

IV. RESULTS AND DISCUSSION

A. Molecular geometry

Molecular geometries for the cationic and anionic complexes, which have been obtained by BLYP/aug-cc-pVDZ levels of theory, are depicted in Figs. 1(a)–1(c) and Figs. 2(a)–2(d), respectively.

Geometrical parameters for H_3O^+, OH^-, and H_2O are presented in the previous paper.[32,41] These parameters were determined by the same functional (BLYP) and the same basis set as those used in the present calculations. From Figs. 1(a) and 1(c), one can recognize that the cationic complexes, $H_5O_2^+$ and $H_9O_4^+$, are almost in the C_2 and C_3 symmetries, respectively. In $H_7O_3^+$ complex, it is found that two water molecules connected to OH bonds of H_3O^+ located at the center have almost the same relative structure with respect to H_3O^+. As for the anionic molecules, it is shown in Figs. 2(c) and 2(d) that $H_7O_4^-$ and $H_9O_5^-$ have the C_3 and C_4 symmetries, respectively, where the hydroxide anion locates on the principal axis of the symmetry. Two water molecules in $H_5O_3^-$ complex also have the same relative structure with respect to OH at the center, analogous to $H_7O_3^+$ cation. Despite the optimizations with tight convergence, the strict symmetries of the systems cannot be achieved because of the shallow nature of the potential energy surface near the minimums.

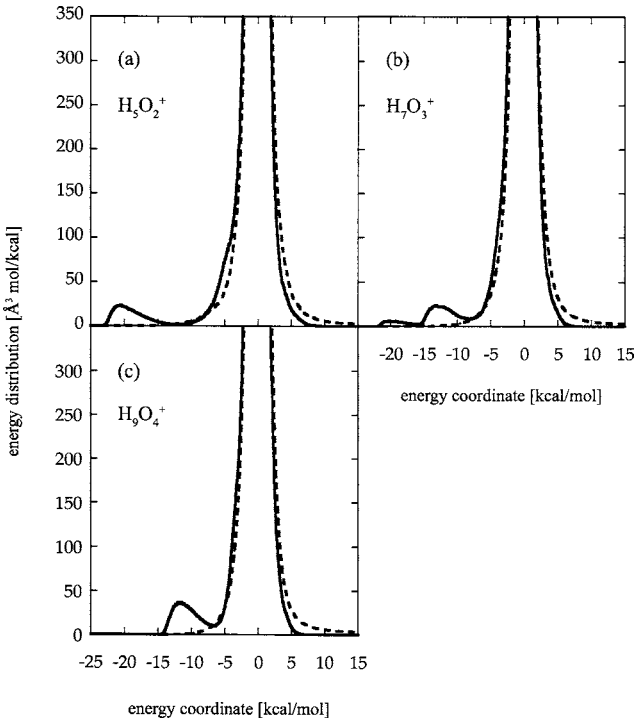

FIG. 3. Energy distribution functions for the hydrated cationic complexes [(a) $H_5O_2^+$, (b) $H_7O_3^+$, and (c) $H_9O_4^+$]. The real and broken lines are the distributions for the solution and the pure solvent systems, respectively. The distributions are normalized by the bulk number density ρ_{bulk}.

B. Dissociation free energy

We have computed the free energy components in Eq. (3) except for the first and last terms to determine the free energy difference $\Delta\mu_n$ in Eq. (7) which is directly related to the equilibrium constant K_n for the reaction represented by Eq. (2). The QM/MM-ER simulations have been utilized to compute solvation free energies μ_{solv} in Eq. (3) for the ionic solutes. The energy distribution functions $\rho(\varepsilon)$ for the solution and the reference systems have been presented in Figs. 3(a)–3(c) for the cationic solutes $H_5O_2^+$, $H_7O_3^+$, and $H_9O_4^+$.

In the energy distribution for $H_5O_2^+$ complex [Fig. 3(a)], one finds a distinct peak at −20 kcal/mol on the energy coordinate which corresponds to the ion-dipole interaction between the complex and solvent water molecules. Coordination of one water molecule to $H_5O_2^+$ molecule leads to the formation of $H_7O_3^+$, which brings about a significant change in the energy distribution function as shown in Fig. 3(b). A new peak arises at the energy coordinate of −13 kcal/mol, which can be attributed to the interaction between a solvent water and a H_2O molecule involved in the solute. Since the positive charge mainly resides in H_3O^+ group that locates at the center of $H_7O_3^+$, an OH bond of

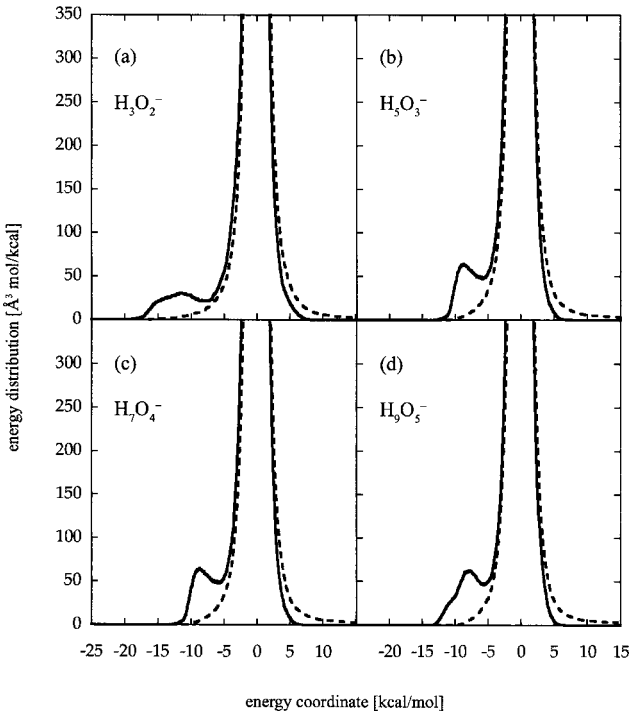

FIG. 4. Energy distribution functions for the hydrated anionic complexes [(a) $H_3O_2^-$, (b) $H_5O_3^-$, (c) $H_7O_4^-$, and (d) $H_9O_5^-$]. The real and broken lines are the distributions for the solution and the pure solvent systems, respectively. The distributions are normalized by the bulk number density ρ_{bulk}.

H_3O^+ can afford a large stabilization energy by making a direct hydrogen bond with a solvent water. However, two OH bonds are occupied by the solute water molecules; only one OH bond is available to form a hydrogen bond with solvent water molecules. As a consequence, the peak of the distribution at −20 kcal/mol becomes much lower as compared to that for $H_5O_2^+$. In $H_9O_4^+$ complex, the remaining dangling OH bond of H_3O^+ is occupied by another water molecule. Then the distribution on ~−20 kcal/mol completely vanished as expected from the above discussion. Here, we have to note a subtle point associated with the model used in the present calculations. The $H_7O_3^+$ complex solvated by one water molecule may be geometrically identical to the $H_9O_4^+$ complex as shown in Figs. 1(b) and 1(c). In other words, it is possible that $H_7O_3^+$ and $H_9O_4^+$ can be categorized into the same molecular species in water solution. If this is true, one should bear in mind that the free energy difference between these two states arises only from the difference in the representations of the interaction potentials. In the case that $H_7O_3^+$ is treated as a QM solute molecule, the interaction between $H_7O_3^+$ and H_2O is to be represented by QM/MM, while it is fully described quantum chemically for the case that $H_9O_4^+$ is treated as a solute.

In Figs. 4(a)–4(d), the energy distribution functions for the anionic solutes are presented. A notable feature found in these figures is that some distribution seen at the energy coordinate of −15 kcal/mol for $H_3O_2^-$ solute completely vanished in the distribution functions for the larger complexes. This may be attributed to the fact that the excess electron is populated over the complex system. The delocalization of the charge will weaken the solute-solvent interaction. Actually it can be clearly seen that the highest occupied molecular orbital of the anionic complex, in which excess charge may reside, is delocalized over the constituent molecules.

The energy distribution functions as well as the correlation matrices have been employed as inputs to Eq. (22) to evaluate the solvation free energies $\Delta\bar{\mu}$. The computed solvation free energies are summarized in Table I as well as the free energy components in Eq. (3) except for the first and last terms.

The zero-point vibrational energies E_{ZPE} and free energies μ_{vibrot} due to the vibrational and rotational motions of the solute molecules have been obtained by the frequency analysis in the GAUSSIAN package. It is shown that the absolute of the free energy μ_{vibrot} as well as E_{ZPE} increases as the size of the complex becomes larger. This is simply a direct consequence that vibrational degrees of freedom and the moments of inertia increase as the size of the cluster increases. A notable fact in the table is that the absolute of the solvation

TABLE I. Free energy components except for the first and last terms in Eq. (3) for water molecule and cationic $[H_3O^+(H_2O)_{n-1}, n=1-4]$ and anionic $[OH^-(H_2O)_{m-1}, m=1-5]$ complexes. μ_{vibrot}, E_{ZPE}, and μ_{solv} are given in units of kcal/mol, while E_0 are in a.u.

		μ_{vibrot}	E_{ZPE}	μ_{solv}	E_0
	H_2O	−2.69	12.85	−5.8[a]	−76.426 76
Cation	H_3O^+	−3.17	20.73	−93.1[b]	−76.694 08
	$H_5O_2^+$	−6.07	34.54	−77.1	−153.177 19
	$H_7O_3^+$	−9.41	50.03	−63.7	−229.640 43
	$H_9O_4^+$	−12.70	65.20	−56.5	−306.096 79
Anion	OH^-	−1.44	5.12	−94.4[b]	−75.806 67
	$H_3O_2^-$	−5.69	18.27	−70.0	−152.277 51
	$H_5O_3^-$	−9.43	34.18	−59.7	−228.737 83
	$H_7O_4^-$	−12.84	49.55	−52.9	−305.192 61
	$H_9O_5^-$	−15.56	65.15	−48.9	−381.634 00

[a]Reference 23. Free energy contribution due to the electron density fluctuation is subtracted from the total solvation free energy.
[b]Reference 32. Free energy contribution due to the electron density fluctuation is subtracted from the total solvation free energy.

free energy monotonically decreases as the size of the cluster increases, which will give serious effects on the dissociation constant given by Eq. (8). This result is similar to the one given by Born's equation [Eq. (23)] which suggests that the solvation free energy of a solute decreases in magnitude when the size of the cavity that contains the ion increases. By substituting these free energies into Eqs. (7) and (8), we have computed the values of $\{I_n\}$ and $\{J_m\}$. The results are summarized in Table II.

As noted in Sec. II B, I_n expresses the ratio of the number density of $H_3O^+(H_2O)_{n-1}$ to that of H_3O^+. It is found in the table that the population of the $H_3O^+(H_2O)_{n-1}$ complex for $n=2-4$ is much larger than that of H_3O^+; on the other hand, OH^- turned out to be the most probable among the anionic species in aqueous solution. It is obvious that $H_5O_2^+$ and OH^- are dominant among the cationic and the anionic species, respectively, which suggests that dissociation constant K_w of water is governed by these species as understood from Eq. (13). It is also found that the population of $H_7O_3^+$ is almost the same with that of $H_9O_4^+$. The free energy difference between these states is estimated to be within 1 kcal/mol. This result is quite reasonable provided that $H_7O_3^+$ is considered to be identical to $H_9O_4^+$ in water solution as described above. Here we have one of the major results in the present article that the Zundel form is more preferable among the possible cationic species. This conclusion is qualitatively consistent with that given by another work[9] which performed *ab initio* molecular dynamics simulations with gradient-corrected functionals. On the contrary, it was suggested that the Eigen form is preponderant by the MS-EVB simulations[15] or an experimental observation by vibrational spectroscopy.[56] The results obtained by neutron diffraction experiment followed by analyses using Monte Carlo simulations also imply the dominance of the Eigen species.[57] Such difference between the experiments and the computational results may be attributed mainly to the subtlety of the free energy difference between the Zundel and the Eigen complexes. Indeed the ratio of the population $(=I_2/I_3)$ in Table II is estimated as ~70 which corresponds to the free energy difference of ≃2.5 kcal/mol between the Zundel and the Eigen complexes. Figure 3 of Ref. 15 indicates that the free energy difference between these complexes is within 1 kcal/mol for the simulation using the later version of the MS-EVB potential. Also in Ref. 58, which performed QM/MM simulations for ice, it was concluded that the energy difference is less than a few kcal/mol. It may not be valid to state which estimation is most reliable since each method has its own drawbacks. The small free energy difference does not allow us to make a definite conclusion.

To make a detailed analysis of the factors which determine the preferable molecular configuration, we decompose the free energies related with the computation of I_n or J_m. We reformulate I_n as follows:

$$\begin{aligned} I_n &= K_n(H_3O^+)[H_2O]^{n-1} \\ &= \frac{[H_3O^+(H_2O)_{n-1}]}{[H_3O^+]} \\ &= \frac{\lambda(H_3O^+)\lambda(H_2O)^{n-1}}{\lambda(H_3O^+(H_2O)_{n-1})}[H_2O]^{n-1} \\ &\quad \times \exp\left(-\frac{\Delta\mu_n(H_3O^+)}{k_BT}\right) \\ &= \exp\left(-\frac{\Delta\mu_n(H_3O^+)+\Delta\mu_n^{\text{trans}}(H_3O^+)}{k_BT}\right), \end{aligned} \tag{24}$$

where we define

$$\Delta\mu_n^{\text{trans}}(H_3O^+) = -k_BT\ln\left(\frac{\lambda(H_3O^+)\lambda(H_2O)^{n-1}}{\lambda(H_3O^+(H_2O)_{n-1})}[H_2O]^{n-1}\right). \tag{25}$$

$\Delta\mu_n^{\text{trans}}(H_3O^+)$ in Eq. (24) can be regarded as translational free energy change associated with the formation of an aggregate for water molecules around H_3O^+. We further decompose the free energy $\Delta\mu_n(H_3O^+)$ into the contributions due to the binding free energy $\Delta\mu_n^{\text{bind}}(H_3O^+)$ and free energy difference $\Delta\mu_n^{\text{solv}}(H_3O^+)$ due to solvation; thus,

$$\Delta\mu_n(H_3O^+) = \Delta\mu_n^{\text{bind}}(H_3O^+) + \Delta\mu_n^{\text{solv}}(H_3O^+). \tag{26}$$

$\Delta\mu_n^{\text{bind}}(H_3O^+)$ in Eq. (26) is the free energy difference in the gas phase for the process where the constituent molecules make the complex $H_3O^+(H_2O)_{n-1}$. It can be explicitly expressed as

$$\begin{aligned} \Delta\mu_n^{\text{bind}}(H_3O^+) = \{&\mu_{\text{vibrot}}(H_3O^+(H_2O)_{n-1}) + E_{\text{ZPE}}(H_3O^+(H_2O)_{n-1}) + E_0(H_3O^+(H_2O)_{n-1})\} - \{(n-1)(\mu_{\text{vibrot}}(H_2O) \\ &+ E_{\text{ZPE}}(H_2O) + E_0(H_2O)) + (\mu_{\text{vibrot}}(H_3O^+) + E_{\text{ZPE}}(H_3O^+) + E_0(H_3O^+))\}. \end{aligned} \tag{27}$$

Then, from Eqs. (7), (26), and (27), the difference in the solvation free energy $\Delta\mu_n^{\text{solv}}(H_3O^+)$ can be given as

$$\Delta\mu_n^{\text{solv}}(H_3O^+) = \mu_{\text{solv}}(H_3O^+(H_2O)_{n-1}) - \{(n-1)\mu_{\text{solv}}(H_2O) + \mu_{\text{solv}}(H_3O^+)\}. \tag{28}$$

By substituting Eq. (26) into Eq. (24) we have,

$$I_n = \exp\left(-\frac{\Delta\mu_n^{\text{bind}}(H_3O^+)+\Delta\mu_n^{\text{solv}}(H_3O^+)+\Delta\mu_n^{\text{trans}}(H_3O^+)}{k_BT}\right). \tag{29}$$

TABLE II. Ratios of the number of densities of the cationic or anionic complexes to that of H_3O^+ or OH^-, respectively. I_n is for the cationic complexes in the form expressed by $H_3O^+(H_2O)_{n-1}$ $(n=1-4)$, while J_m is for anionic complexes $OH^-(H_2O)_{m-1}$ $(m=1-5)$.

n,m	1	2	3	4	5
I_n	1.0	2.5×10^{6}	3.5×10^{4}	8.2×10^{3}	⋯
J_m	1.0	1.5×10^{-4}	4.1×10^{-5}	2.0×10^{-6}	5.5×10^{-9}

The expressions similar to Eqs. (24)–(29) also hold for the anionic species $OH^-(H_2O)_{m-1}$. Thus, it is shown that the population of an ionic complex with respect to bare ion H_3O^+ or OH^- is determined by the interplay of three kinds of free energy difference associated with the formation of the complex in solution. In Figs. 5(a) and 5(b), we show the size dependence of these free energies for the cationic and anionic species, respectively.

For the cationic molecules $[H_3O^+(H_2O)_{n-1}]$, it can be easily recognized that the difference $\Delta\mu_n^{\mathrm{bind}}$ in the binding free energy decreases monotonically as the number n increases, while the difference $\Delta\mu_n^{\mathrm{solv}}$ in the solvation free energy increases. The magnitude of $\Delta\mu_n^{\mathrm{bind}}$ is substantially large; however, a large amount of the stabilization energy is canceled by the free energy difference $\Delta\mu_n^{\mathrm{solv}}$ with opposite sign. As a result, the sum of $\Delta\mu_n^{\mathrm{bind}}$ and $\Delta\mu_n^{\mathrm{solv}}$ amounts to -13 to -18 kcal/mol for $n=2-4$. It has also been found that the difference $\Delta\mu_n^{\mathrm{trans}}$ in the translational free energy gives non-negligible contribution to the total free energy difference. $\Delta\mu_n^{\mathrm{trans}}$ increases almost linearly with the system size and it becomes ~12.5 kcal/mol at $n=4$. Thus, it has been found that one cannot overlook the loss of the translational entropy associated with the formation of aggregate even under a high density and a low temperature condition. As a whole, it has been revealed that the free energy $\Delta\mu_n(H_3O^+)$ has the minimum at $n=2$ (i.e., $H_5O_2^+$). It should be noted, however, that there are no remarkable differences among the free energies $\Delta\mu_n(H_3O^+)$ for $n=2-4$. The free energy difference is no more than 3.4 kcal/mol at most. In particular, for the complexes of $n=3$ and 4 (i.e., $H_7O_3^+$ and $H_9O_4^+$, respectively), the difference in $\Delta\mu_n(H_3O^+)$ turned out to be only within 1 kcal/mol. It suggests that these two complexes are energetically almost comparable in solution and, hence the conversion between two hydration complexes may take place frequently in solution being driven by the dynamics of the surrounding solvent molecules. For the anionic species $[OH^-(H_2O)_{m-1}]$, one can recognize the trends similar to cationic systems in the behavior of the free energy differences $\Delta\mu_m^{\mathrm{bind}}(OH^-)$, $\Delta\mu_m^{\mathrm{solv}}(OH^-)$, and $\Delta\mu_m^{\mathrm{trans}}(OH^-)$ for the variation of the size parameter m. A major difference is that both $\Delta\mu_m^{\mathrm{bind}}(OH^-)$ and $\Delta\mu_m^{\mathrm{solv}}(OH^-)$ increased as compared with the cationic complex with the same cluster size. This may be attributed to the diffuse nature of the excess charge distributed over the complexes, which destabilize the solvation free energies. The behavior of the free energy difference $\Delta\mu_m^{\mathrm{trans}}(OH^-)$ is similar to that of cationic system. As a result, it has been found that the bare ion OH^- is the most stable in the solution among the possible ionic complexes as shown in Fig. 5(b).

The studies so far have revealed that the leading term in the series of $K_w'(n,m)$ defined by Eq. (14) is $K_w'(2,1)$. In other words, it has been demonstrated that ionic species $H_5O_2^+$ and OH^- dominate the value of the dissociation constant K_w of water. Here we examine the pK_w value obtained by the present work by making a comparison with the well-known experimental value of $pK_w=14$ at 300 K. Conventionally pK_w is defined as

$$pK_w = -\log K_w - 2\log\{H_2O\},$$

where $\{H_2O\}$ expresses the concentration of water in the unit of mol/l. In the present calculation, $-\log K_w'(2,1)$ has been obtained as 27.3 in which $\log H_0$ is estimated to be 33.7 from Eq. (12). By substituting the value of $\log K_w'(2,1)$ into the above equation, we have $pK_w=23.8$, which differs substantially from the experimental value. The dissociation free energy $\Delta G_{\mathrm{diss}}'(2,1)$ in Eq. (17) has been obtained as 37.5 kcal/mol, while the experimental value of $pK_w=14$ corresponds to the free energy difference of 24.0 kcal/mol. Hence, the difference in the dissociation free energy between the present work and the experiment has been estimated to be ~13 kcal/mol. In Ref. 4 the autodissociation of water under various conditions was studied by carrying out elaborate QM/MM simulations where up to 13 water molecules are treated as QM molecules in the free energy calculations. Their simulation faithfully reproduced the temperature dependence of pK_w experimentally determined. However, the computed dissociation free energies must be shifted by −9 kcal/mol to yield $pK_w=14$ at 300 K. Thus, the error in the present work is almost comparable to that found in Ref. 4. In another theoretical investigation[2] that computed pK_w over a wide range of thermodynamic condition, it was also shown that the absolutes of pK_w did not coincide with the experimental values, though their density or temperature de-

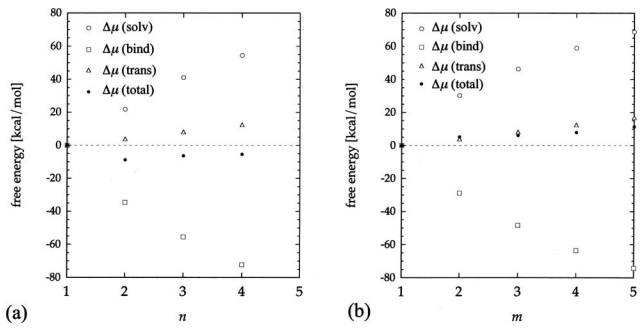

FIG. 5. Size dependences of the free energy difference $\Delta\mu_n$ given by Eq. (7) and its components for (a) cationic solutes $H_3O^+(H_2O)_{n-1}$ and for (b) anionic solutes $OH^-(H_2O)_{m-1}$.

pendence was quite consistent with the experimental observation. In a recent theoretical investigation[59] based on the PCM with the conductorlike screening model,[60] it was reported that the computed *p*H is still above the experimentally known value even for the cluster involving eight water molecules. Contrary to these results, *ab initio* molecular dynamics simulation,[3] which studied the dissociation of H_2O in water, substantially underestimated the dissociation free energy. It seems that the major reason for this is the finite-size effect as speculated by the authors. The rather small repeated cell, which contained only 32 water molecules, would artificially stabilize the charge-polarized state in the cell due to its periodicity. In conclusion, we speculate that the source of the error would be ascribed to some physical property that we did not take into consideration in the simulations, such as nuclear quantum effect of the proton. The underlying mechanism responsible for the error is now under investigation.

V. CONCLUSION

In the present work, we have addressed the issue of the dominant hydration forms among the ionic species provided as the dissociation products of a water molecule in aqueous solution. The population of each ion is described in terms of the solvation free energies of the constituent molecules which have been determined by utilizing the QM/MM-ER approach recently developed. Our simulations have revealed that the population of $H_5O_2^+$ identified as a proton coordinated by two water molecules is dominant among the hydrated cationic complexes with the form of $H_3O^+(H_2O)_{n-1}$ ($n=1-4$). The populations of the complexes for $n=2-4$ have been found to be at least 10^3 times as large as that of H_3O^+. However, the free energy difference between two complexes with $n=3$ and 4 is subtle and the conversion from one to another may be driven by the fluctuation of the surrounding water molecules. In contrast, it has been demonstrated that a bare hydroxide ion (OH^-) is much more preferable in solution among the hydrated ions in the form of $OH^-(H_2O)_{m-1}$ ($m=1-5$). The reason for this can be attributed to the excess charge being distributed over the complex which gives rise to the destabilization in the solvation free energies. For both the cationic and anionic species, we stress that entropy loss in the translational free energy associated with the formation of complex plays a role in determining the population of the complex. When an ionic cluster is formed by the condensation of three water molecules, the change in the translational free energy amounts to ~ 12.5 kcal/mol in ambient water. By utilizing the solvation free energies for the dominant ionic complexes, we have also estimated the dissociation constant pK_w and the corresponding free energy ΔG_{diss}. pK_w has been obtained as 23.8, which differs significantly from the well-known experimental value of $pK_w=14$. Then ΔG_{diss} has been computed as 37.5 kcal/mol, of which deviation from the experimental value is about 13 kcal/mol, which is almost comparable to the result given by recent elaborate QM/MM simulations. The error was also reported in other theoretical work. These facts lead us to propose that the source of the error would be ascribed to some physical property such as quantum effect of the proton that has not been taken into consideration in these calculations. It has also been demonstrated that the novel QM/MM approach we developed is successfully applied to the investigation of the chemical event in condensed system within modest computational costs.

ACKNOWLEDGMENTS

This work was supported by Grant-in-Aid for Scientific Research on Priority Areas (Nos. 18031022 and 18066010) from the Ministry of Education, Science, Sports and Culture of Japan and also by Grant-in-Aid for Japan Society for the Promotion of Science (JSPS) fellow (No. 1809170). This work was also supported by the Next Generation Super Computing Project, Nanoscience Program, MEXT, Japan.

[1] *Water, a Comprehensive Treatise*, edited by F. Franks (Plenum, New York, 1972), Vol. 1; (Plenum, New York, 1973), Vol. 2; (Plenum, New York, 1973), Vol. 3; (Plenum, New York, 1975), Vol. 4; (Plenum, New York, 1975), Vol. 5; (Plenum, New York, 1979), Vol. 6; (Plenum, New York, 1982), Vol. 7.
[2] H. Sato and F. Hirata, J. Phys. Chem. B **103**, 6596 (1999).
[3] B. L. Trout and M. Parrinello, Chem. Phys. Lett. **288**, 343 (1998).
[4] T. Yagasaki, K. Iwahashi, S. Saito, and I. Ohmine, J. Chem. Phys. **122**, 144504 (2005).
[5] P. L. Geissler, C. Dellago, D. Chandler, J. Hutter, and M. Parrinello, Science **291**, 2121 (2001).
[6] M. E. Tuckerman, D. Marx, and M. Parrinello, Nature (London) **417**, 925 (2002).
[7] M. E. Tuckerman, D. Marx, M. L. Klein, and M. Parrinello, Science **275**, 817 (1997).
[8] D. Marx, M. E. Tuckerman, J. Hutter, and M. Parrinello, Nature (London) **397**, 601 (1999).
[9] D. Asthagiri, L. R. Pratt, and J. D. Kress, Proc. Natl. Acad. Sci. U.S.A. **102**, 6704 (2005).
[10] G. A. Voth, Front. Biosci. **8**, 1384 (2003).
[11] J. M. J. Swanson, C. M. Maupin, H. Chen, M. K. Petersen, J. Xu, Y. Wu, and G. A. Voth, J. Phys. Chem. B **111**, 4300 (2007).
[12] J. Åqvist and A. Warshel, Chem. Rev. (Washington, D.C.) **93**, 2523 (1993).
[13] R. Vuilleumier and D. Borgis, Chem. Phys. Lett. **284**, 71 (1998).
[14] U. W. Schmitt and G. A. Voth, J. Phys. Chem. B **102**, 5547 (1998).
[15] T. J. F. Day, A. V. Soudackov, M. Cuma, U. W. Schmitt, and G. A. Voth, J. Chem. Phys. **117**, 5839 (2002).
[16] A. Szabo and N. S. Ostlund, *Modern Quantum Chemistry* (Macmillan, New York, 1982).
[17] D. Frenkel and B. Smit, *Understanding Molecular Simulation* (Academic, New York, 1996).
[18] J. Tomasi and M. Persico, Chem. Rev. (Washington, D.C.) **94**, 2027 (1994).
[19] S. Ten-no, F. Hirata, and S. Kato, J. Chem. Phys. **110**, 7443 (1994).
[20] R. Car and M. Parrinello, Phys. Rev. Lett. **55**, 2471 (1985).
[21] *Combined Quantum Mechanical and Molecular Mechanical Methods*, edited by J. Gao and M. A. Thompson (American Chemical Society, Washington, DC, 1998).
[22] *Combined QM/MM Calculations in Chemistry and Biochemistry*, edited by M. F. Ruiz-Lopez, special issue of J. Mol. Struct.: THEOCHEM 632 (2003).
[23] H. Takahashi, T. Hori, T. Wakabayashi, and T. Nitta, Chem. Lett. **2000**, 222.
[24] H. Takahashi, T. Hori, T. Wakabayashi, and T. Nitta, J. Phys. Chem. A **105**, 4351 (2001).
[25] H. Takahashi, T. Hori, H. Hashimoto, and T. Nitta, J. Cryst. Growth **22**, 1252 (2003).
[26] J. R. Chelikowsky, N. Troullier, and Y. Saad, Phys. Rev. Lett. **72**, 1240 (1994).
[27] J. R. Chelikowsky, N. Troullier, K. Wu, and Y. Saad, Phys. Rev. B **50**, 11355 (1994).
[28] X. Jing, N. Troullier, D. Dean, N. Binggeli, J. R. Chelikowsky, K. Wu, and Y. Saad, Phys. Rev. B **50**, 12234 (1994).
[29] P. Hohenberg and W. Kohn, Phys. Rev. **136**, B864 (1964).

[30] W. Kohn and L. Sham, Phys. Rev. **140**, A1133 (1965).
[31] T. Hori, H. Takahashi, and T. Nitta, J. Theor. Comput. Chem. **4**, 867 (2005).
[32] H. Takahashi, N. Matubayasi, T. Nitta, and M. Nakahara, J. Chem. Phys. **121**, 3989 (2004).
[33] N. Matubayasi and M. Nakahara, J. Chem. Phys. **113**, 6070 (2000).
[34] N. Matubayasi and M. Nakahara, J. Chem. Phys. **117**, 3605 (2002); **118**, 2446 (2003).
[35] N. Matubayasi and M. Nakahara, J. Chem. Phys. **119**, 9686 (2003).
[36] J. P. Hansen and I. R. McDonald, *Theory of Simple Liquids* (Academic, London, 1986).
[37] T. Hori, H. Takahashi, and T. Nitta, J. Comput. Chem. **24**, 209 (2002).
[38] H. Takahashi, S. Takei, T. Hori, and T. Nitta, J. Mol. Struct. **632**, 185 (2003).
[39] H. Takahashi, H. Hashimoto, and T. Nitta, J. Chem. Phys. **119**, 7964 (2003).
[40] T. Hori, H. Takahashi, and T. Nitta, J. Chem. Phys. **119**, 8492 (2003).
[41] H. Takahashi, W. Satou, T. Hori, and T. Nitta, J. Chem. Phys. **122**, 044504 (2005).
[42] H. Takahashi, Y. Kawashima, and T. Nitta, J. Chem. Phys. **122**, 124504 (2005).
[43] T. Hori, H. Takahashi, S. Furukawa, M. Nakano, and W. Yang, J. Phys. Chem. B **111**, 581 (2007).
[44] H. Takahashi, K. Tanabe, M. Aketa, R. Kishi, S. Furukawa, and M. Nakano, J. Chem. Phys. **126**, 084508 (2007).
[45] T. Hori, H. Takahashi, M. Nakano, T. Nitta, and W. Yang, Chem. Phys. Lett. **419**, 240 (2005).
[46] M. J. Frisch, G. W. Trucks, H. B. Schlegel *et al.*, GAUSSIAN 03, Revision B.05, Gaussian, Inc., Pittsburgh, PA, 2003.
[47] A. D. Becke, J. Chem. Phys. **98**, 5648 (1993).
[48] C. Lee, W. Yang, and R. G. Parr, Phys. Rev. B **37**, 785 (1988).
[49] T. H. Dunning, Jr., J. Chem. Phys. **90**, 1007 (1989).
[50] L. Kleinman and D. M. Bylander, Phys. Rev. Lett. **48**, 1425 (1982).
[51] T. Ono and K. Hirose, Phys. Rev. Lett. **82**, 5016 (1999).
[52] R. N. Barnett and U. Landman, Phys. Rev. B **48**, 2081 (1993).
[53] W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, and M. L. Klein, J. Chem. Phys. **79**, 926 (1983).
[54] P. Ewald, Ann. Phys. **64**, 253 (1921).
[55] M. Born, Z. Phys. **1**, 45 (1920).
[56] W. Amir, G. Gallot, F. Hache, S. Bratos, J.-C. Leicknam, and R. Vuilleumier, J. Chem. Phys. **126**, 034511 (2007).
[57] A. Botti, F. Bruni, S. Imberti, M. A. Ricci, and A. K. Soper, J. Chem. Phys. **121**, 7840 (2004); A. Botti, F. Bruni, M. A. Ricci, and A. K. Soper, *ibid.* **125**, 014508 (2006).
[58] C. Kobayashi, S. Saito, and I. Ohmine, J. Chem. Phys. **113**, 9090 (2000).
[59] J. Mrázek and J. V. Burda, J. Chem. Phys. **125**, 194518 (2006).
[60] A. Klamt and G. Schuurmann, J. Chem. Soc., Perkin Trans. 2 **1993**, 799.