

A linear-scaling self-consistent generalization of the multistate empirical valence bond method for multiple excess protons in aqueous systems

Feng Wang and Gregory A. Voth^{a)}

Department of Chemistry and Center for Biophysical Modeling and Simulation, University of Utah, Salt Lake City, Utah 84112-0850

(Received 24 November 2004; accepted 2 February 2005; published online 13 April 2005)

An extension to the multistate empirical valence bond (MS-EVB) method is presented in this paper that is capable of treating multiple excess protons within the context of molecular-dynamics simulation. The computational cost of the method scales linearly with respect to the number of excess protons. Calculations for a 0.44 M HCl systems are carried out to illustrate the multiproton extension of the MS-EVB method. A significant decrease in the Eigen-type H_9O_4^+ cation is observed in the contact ion-pair configuration formed between Cl^- and hydronium ions. © 2005 American Institute of Physics. [DOI: 10.1063/1.1881092]

I. INTRODUCTION

The significant advances in computer technology in the last 30 years have made molecular-dynamics (MD) simulations an increasingly powerful and effective tool to study the dynamical and structural properties of many interesting systems.^{1–3} Whereas *ab initio* MD approaches are still too expensive to be carried out on systems with more than a hundred or so atoms, various empirical force fields have been developed to make large scale MD simulation a reality.^{4–7} Of all the various systems being studied by computer simulations, active proton transfer^{8–11} is one of the most challenging fundamental problems for empirical force fields to describe.

It is well known that proton diffusion constant in water is significantly higher compared to other forms of cations.^{8,12} A number of mechanisms have been proposed for this fast proton diffusion, including the Hückel mechanism,¹³ in which the rate limiting step is the rotation of the H_3O^+ ; the Bernal and Fowler mechanism,¹⁴ in which a free water rotor in the first solvation shell is assumed; the Conway mechanism,¹⁵ which involves a “field-induced” rotation in the first solvation shell; the Eigen picture,¹⁶ in which “structural diffusion” is argued to be the rate determining step in bulk water; and the “Moses” mechanism,^{17,18} in which the bond cleavage at the second solvation shell is believed to be the precursor to proton translocation. The Grotthuss hopping process^{19,20} plays a crucial role in all of these mechanisms. Although recent work²¹ has suggested that none of them is completely correct, the Grotthuss hopping process is still believed to be a key step in proton diffusion.

The Grotthuss process is named after C. G. T. de Grotthuss who proposed a hydrogen–oxygen shuttling mechanism to explain electrolysis of water.²⁰ The first invocation of such a process to explain proton transport was believed to be done by Danneel in 1905.¹⁹ In the Grotthuss shuttling pro-

cess, an excess proton bonds to the oxygen atom of a water molecule and forms a valence bond with that oxygen atom. Meanwhile, one of the valence-bonded hydrogen atoms of the water molecule being bonded to by the proton leaves and becomes an excess proton. This process transforms the incoming proton into a bonded atom, and a hydrogen atom of the water molecule into an excess proton. In order to properly describe the Grotthuss process in MD simulations, the force field must be flexible enough to allow bonding topology and atom identity to change. This requirement renders most traditional empirical force fields unsuitable to treat excess proton solvation and transport dynamics in aqueous environments.

The multistate empirical valence bond (MS-EVB) model has enjoyed considerable success.^{22–36} (For a description of this model in relation to other models, the reader is referred to the introduction of Ref. 33.) This model might be regarded as a simulation version of the well-known resonance theory frequently used in organic chemistry.³⁷ For any given geometric arrangement of the $\text{H}_{2n+1}\text{O}_n^+$ cation, multiple bonding arrangements can be drawn. If we restrict ourselves only to the bonding topology that contains one hydronium and $n-1$ water molecules, n possible low-energy bonding configurations can be found for the $\text{H}_{2n+1}\text{O}_n^+$ system. In MS-EVB, each bonding topology is termed a basis state. The above-mentioned $\text{H}_{2n+1}\text{O}_n^+$ system can then be represented as a linear combination of these possible “basis states” with their relative weight determined by solving an eigenvalue-eigenvector problem for the EVB matrix. The diagonal elements of the EVB matrix are given by the “diabatic” energy of the system in that particular excess proton binding arrangement. The coupling term between different basis states is described as a constant plus a term that is determined by viewing the system as a Zundel-type complex, formed by allowing the two hydroniums in each basis state to share the same proton simultaneously, and solvated by the rest of the

^{a)}Electronic mail: voth@chem.utah.edu

water molecules in the system for that instantaneous configuration. More details about this model can be found in Refs. 24 and 33.

Although the MS-EVB model has proven to be very successful in describing single excess proton solvation and transport, a great challenge has been faced in expanding the methodology to describe systems containing more than one excess proton in the same periodic simulation cell. The straightforward way to use the model to describe multiproton systems is to form a large matrix in a MS-EVB product basis set of all the protons. This is very similar to the configuration-interaction (CI) procedure in electronic structure calculations.³⁸ This procedure will be called “big-matrix” approach in this paper. The order of the big matrix is approximately N^m , where N is the number of MS-EVB basis states accessible to each proton and m is the number of excess protons. In a typical bulk water simulation, N is close to 25. The MD propagation of systems containing more than two protons using the big-matrix approach would therefore require the extremely expensive solution of an eigenvalue problem with a matrix of order larger than 15 000 at each MD time step. This makes it infeasible to carry out converged big-matrix MS-EVB MD simulations in a reasonable amount of CPU time on current generation computer hardware. A different approach was required, which is the focus of the present work.

In this paper, a self-consistent iterative multistate empirical valence bond (SCI-MS-EVB) method is presented that solves the multiproton problem with a computational effort scaling linearly with respect to the number of excess protons. This methodology is described in Sec. II, while illustrative calculations are described in Sec. III for a 0.44 M HCl solution with both one HCl per periodic box and two HCl per periodic box. The results from these simulations are discussed in Sec. IV, while concluding remarks are given in Sec. V.

II. METHODOLOGY

A. The iterative method

The EVB model was first introduced by Warshel and co-worker,^{39,40} while the first application to proton transfer in water was done by Lobaugh and Voth.⁴¹ The work in this paper is based on the multistate generation by Schmitt and Voth²⁴ (MS-EVB) and later by Day *et al.*³³ (MS-EVB2). The SCI-MS-EVB method described in this paper is based on the MS-EVB2 model³³ parameters with minor modifications.

In the new SCI-MS-EVB method, the system is divided into EVB complexes, each consisting of a single proton. A single proton MS-EVB problem is solved for each EVB complex within an effective field of the other EVB complexes. The effective field represents atoms in the other EVB complexes by effective charges and van der Waals parameters according to the corresponding EVB vectors of that EVB complex.

Figure 1 depicts a two-proton system to illustrate our approach, with proton A in EVB complex A and proton B in EVB complex B. According to the formalism of MS-EVB2, the EVB matrix of proton A has as its diagonal elements

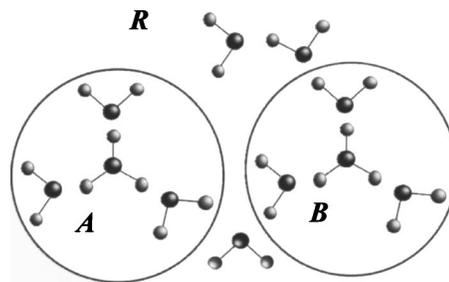


FIG. 1. A two-proton system, with proton A in EVB complex A (left circle) and proton B in EVB complex B (right circle). The water molecules that do not belong to either EVB complex are in region R.

$$\begin{aligned} \langle i|H|i\rangle = & V_{\text{H}_3\text{O}^+}^{\text{intra}} + \sum_{k=1}^{N_{\text{H}_2\text{O}}} V_{\text{H}_2\text{O}}^{\text{intra},k} + \sum_{k=1}^{N_{\text{H}_2\text{O}}} V_{\text{H}_3\text{O}^+, \text{H}_2\text{O}}^{\text{inter},k} \\ & + \sum_{k < k'}^{N_{\text{H}_2\text{O}}} V_{\text{H}_2\text{O}}^{\text{inter},kk'} \end{aligned} \quad (1)$$

and off-diagonal elements

$$\langle i|H|j\rangle = (V_{\text{const}}^{ij} + V_{\text{exchange}}^{ij})A(R_{\text{OO}}, R_{\text{OH}}), \quad (2)$$

where $|i\rangle$ and $|j\rangle$ form an orthonormal set of EVB basis states selected according to the procedures described in Ref. 33. Detailed definitions of various terms in Eqs. (1) and (2) can also be found in Ref. 33.

In the SCI-MS-EVB approach, the Hamiltonian for EVB complex A is further partitioned as follows:

$$H_A = H_{AA} + H_{AB'} + H_{AR}, \quad (3)$$

where H_{AA} is the contribution from particles interacting inside EVB complex A and $H_{AB'}$ describes the interaction between particles in EVB complex A and EVB complex B, where the prime indicates that particles in EVB complex B are seen by particles in EVB complex A as effective particles. H_{AR} describes the interactions between particles in EVB complex A and particles not contributing to any of the EVB complexes (see Fig. 1). It is evident that the first two terms in Eq. (1) and the first term in Eq. (2) only contribute to H_{AA} , while the rest of the terms contain contributions to all the three parts in Eq. (3). A detailed examination of Eqs. (1) and (2) reveals that these terms belong to three types of interactions: electrostatic, van der Waals, and short-range repulsion. Electrostatic interactions are expressed as

$$V_{\text{el}}(r) = \frac{q_s q_t}{r}, \quad (4)$$

where q_s and q_t are point charges. The van der Waals interactions can be expressed as

$$V_{\text{vdW}}(r) = 4\epsilon_{st} \left[\left(\frac{\sigma_{st}}{r_{st}} \right)^{12} - \left(\frac{\sigma_{st}}{r_{st}} \right)^6 \right] = \left(\frac{\alpha_{st}}{r_{st}} \right)^{12} - \left(\frac{\beta_{st}}{r_{st}} \right)^6, \quad (5)$$

where ϵ_{st} and σ_{st} are Lennard-Jones (LJ) parameters of particles s and t , while

$$\alpha_{st} = (4\varepsilon_{st})^{1/12} \sigma_{st}, \quad \beta_{st} = (4\varepsilon_{st})^{1/6} \sigma_{st}. \quad (6)$$

Repulsive interactions can be expressed as

$$V^{\text{rep}}(r_{st}) = B(1 - \tanh[b(r_{st} - d_{st}^0)]), \quad (7)$$

where B , b , d_{st}^0 are repulsive parameters defined in Ref. 33. In MS-EVB2, the repulsive potential only exists between the hydronium oxygen and water oxygen and hydrogen.

One can now write out expressions for the electrostatic, van der Waals, and repulsive interactions for H_{AA} and H_{AR} . In order to calculate the contributions from these interactions to H_{AB} , effective charges and van der Waals parameters are employed. For electrostatic interactions, an effective particle s in EVB complex B as seen by particles in EVB complex A has the charge

$$q_s = \sum_i c_{Bi}^2 q_{sH} + \sum_{j \neq i} c_{Bj}^2 q_{sW} + \sum_{i,j,i \neq j} c_{Bi} c_{Bj} q_{\text{ex},ij} A(r), \quad (8)$$

assuming that particle s is part of a hydronium ion in state i , and part of a water molecule in all the other states. Here, q_{sH} , q_{sW} are the charges of particle s when it is part of the hydronium and water molecules, respectively, and c_{Bi} , c_{Bj} are the EVB coefficients of states i and j for EVB complex B found by diagonalizing the corresponding EVB matrix H_B . The matrix H_B is formed following a similar procedure as was done for H_A , and will be discussed later. The last term of Eq. (8) is the contribution from exchange charges for particle s , with $A(r)$ being the same geometrical scaling function as used in Eq. (2). The exchange charge $q_{\text{ex},ij}$ is nonzero only if the coupling element between state i and state j is not zero, as defined in Ref. 33.

In a CI-like big-matrix approach to the multiproton problem, there will never be any interaction between the off-diagonal exchange charges of different EVB complexes. We opt to take the same approach, so that interactions between off-diagonal charges are intentionally suppressed in SCI-MS-EVB. This is achieved by removing the contribution of the last term of Eq. (8) when calculating the off-diagonal terms in the EVB matrix for EVB complex A .

For van der Waals interactions, we define parameters α , β for effective particle s using the following formulas:

$$\alpha_s = c_{Bi}^2 \alpha_{sH} + \sum_{j \neq i} c_{Bj}^2 \alpha_{sW}, \quad (9)$$

$$\beta_s = c_{Bi}^2 \beta_{sH} + \sum_{j \neq i} c_{Bj}^2 \beta_{sW}, \quad (10)$$

where the same notation as used in Eq. (8) is followed.

The only extra parameters needed in the multiproton case are the repulsive interaction parameters between the hydronium oxygen atoms and between each hydronium oxygen atom with the other hydronium hydrogen atoms. Since the repulsive interactions as described by Eq. (7) are very short range and repulsive interactions between hydronium oxygen and effective particles in the other EVB complexes are negligible compared to all other interactions, a detailed parametrization of these interactions is not important at this point. In this study, no distinction is made between hydronium oxygen and water oxygen and hydronium hydrogen and water

hydrogen in the effective EVB complexes so far as the repulsive potential in Eq. (7) is concerned. This definition makes it unnecessary to determine the repulsion parameters in Eq. (7) for effective particles s since they are always the same for all linear combinations of EVB states.

The same analysis applies to the Hamiltonian for EVB complex B , such that

$$H_B = H_{BB} + H_{BA'} + H_{BR}, \quad (11)$$

with the interactions between EVB complexes A and B replaced by $H_{BA'}$. The prime on A indicates that particles in EVB complex A are replaced by effective particles based on EVB vector c_A .

For the same pair of electrostatic interactions between atoms s in EVB complex A and t in EVB complex B , assuming atom s is part of a hydronium in state i and atom t is part of a hydronium in state i' , we have

$$\begin{aligned} H_{AB'} = & \left(\sum_i c_{Ai}^2 q_{sH} + \sum_{j \neq i} c_{Aj}^2 q_{sW} \right) \cdot \left(\sum_i c_{Bi'}^2 q_{tH} + \sum_{j' \neq i'} c_{Bj'}^2 q_{tW} \right) \\ & + \left(\sum_{i,j,i \neq j} c_{Ai} c_{Aj} q_{s,\text{ex},ij} A(r) \right) \\ & \times \left(\sum_i c_{Bi'}^2 q_{tH} + \sum_{j' \neq i'} c_{Bj'}^2 q_{tW} \right) \\ & + \left(\sum_{i',j',i' \neq j'} c_{Bi'} c_{Bj'} q_{t,\text{ex},i'j'} A(r') \right) \\ & \times \left(\sum_i c_{Ai}^2 q_{sH} + \sum_{j \neq i} c_{Aj}^2 q_{sW} \right) = H_{BA'}. \end{aligned} \quad (12)$$

Thus we can drop the prime and write H_{AB} in place of $H_{AB'}$ and $H_{BA'}$. It is seen that H_{AB} depends on c_{Ai} and c_{Bi} simultaneously. Thus, the solution of the EVB Hamiltonian for EVB complex A depends on the solution of the EVB Hamiltonian for EVB complex B and vice versa. A self-consistent solution may be obtained by solving Eq. (3) with an educated guess for solution of Eq. (11), and then constructing the EVB Hamiltonian in Eq. (11) using the solution of Eq. (3), and then iterate both solutions to convergence. In a MD simulation, the converged EVB vector of the previous time step can be used as an initial guess. For the first MD step, a unit vector that corresponds to the “most likely” classical hydronium solution suffices.

Having obtained the self-consistent solution for the two-proton problem, the total system energy can then be expressed as

$$\begin{aligned} E_{\text{total}} = & \langle a_0 | H_{AA} + H_{AR} | a_0 \rangle + \langle b_0 | H_{BB} + H_{BR} | b_0 \rangle \\ & + \langle a_0 b_0 | H_{AB} | a_0 b_0 \rangle + E_{RR} \\ = & E_{AA} + E_{BB} + E_{AB} + E_{AR} + E_{BR} + E_{RR}, \end{aligned} \quad (13)$$

where $|a_0\rangle$ and $|b_0\rangle$ are the ground-state solutions for EVB complex A and EVB complex B , respectively.

After self-consistent solutions have been found, the Hellmann–Feynman theorem can then be used to calculate the forces. This result can be shown using the following arguments: It is obvious that the Hellmann–Feynman theorem holds if the total energy simultaneously reaches a mini-

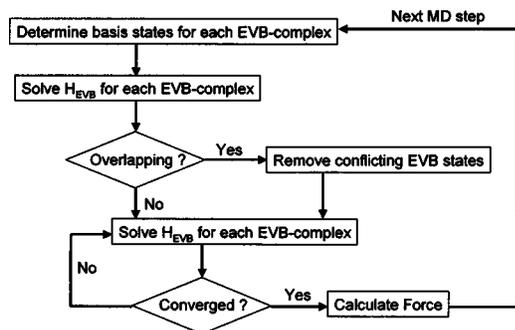


FIG. 2. Flow diagram showing various steps in the SCI-MS-EVB molecular-dynamics algorithm.

imum with respect to all coefficients c_{Ai} and c_{Bi} . The EVB vector c_{Ai} is obtained by optimizing H_A , as described in Eq. (3), and c_{Bi} is obtained by optimizing H_B , as shown in Eq. (11). If EVB vector c_{Ai} is perturbed while satisfying orthonormal constraints, the summation of the first, third, and fourth terms in Eq. (13) will necessarily increase. If EVB vector c_{Bi} is perturbed, then the summation of the second, third, and fifth terms of Eq. (13) has to go up. Thus, we can safely say that the coefficients c_{Ai} and c_{Bi} minimize the total energy E_{total} simultaneously, and the Hellmann–Feynman theorem therefore holds.

The above discussion can be extended to treat systems with more than two EVB complexes. However, the above arguments run into difficulty if one of the water molecules belongs to more than one EVB complex at the same time. In the following discussion we will call this water in the overlapping region of different EVB complexes. The EVB states that utilize this water molecule to form a hydronium will be called overlapping states. Although it is possible to treat overlapping regions explicitly in SCI-MS-EVB, it is worth noting that a water molecule that belongs to two EVB complexes simultaneously will be doubly protonated with the probability $c_{Ai}^2 c_{Bi'}^2$. A state with large values for both c_{Ai}^2 and $c_{Bi'}^2$ will be quite high in energy and should not significantly contribute to the ground EVB state. Thus, to a good approximation, we can simply suppress contributions from overlapping states by only assigning the water to the EVB complex that gives it the largest hydronium character. There are multiple ways to accomplish this; we adopt here the procedure described in the following paragraph.

As shown in Fig. 2, in the SCI-MS-EVB method the first iteration is carried out keeping all the states of each EVB complex. In the case of overlap, the relative weights of hydronium character, as defined by the squares of the corresponding EVB coefficients of the overlapping states, are compared. The water molecule is then assigned to the EVB complex that has the largest weight on this water molecule among all the overlapping EVB complexes, whereas the rest of the overlapping states are removed. Subsequent iterations are then carried out using the resulting overlap-free EVB basis states until the EVB energies and EVB vectors for each EVB complex no longer change. At this point, the EVB vectors converge to the self-consistent solution. In this study, we have used a convergence tolerance where the maximum relative change in the EVB energies as defined in Eqs. (3) and

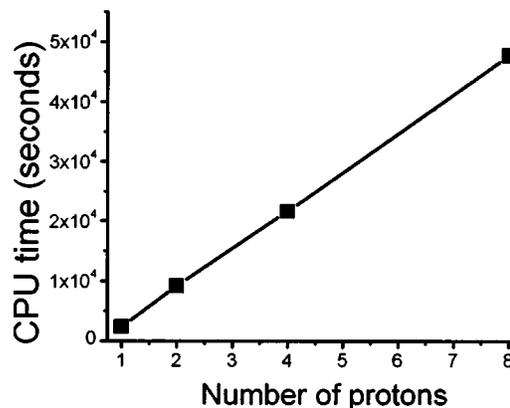


FIG. 3. Scaling of CPU time with respect to the number of excess protons in the system. Calculations were for 2 ps trajectories for a system with 512 water molecules.

(11) of all the EVB complexes between two iterations is within 10^{-6} . Test calculations indicate that EVB vectors have also properly converged when this criterion is used.

In practice, the number of iterations required for convergence depends on the excess proton concentration. For a system as concentrated as 5 M HCl, around five iterations were usually needed for the energies to converge to a relative error of 10^{-6} . For lower concentration systems such as 1 to 2 M HCl, an average of three iterations was observed. The scheme is found to be very stable, since the interactions between hydronium molecules are always repulsive at any physically meaningful geometry.

It should also be noted that force evaluations are not needed during the iterative procedure. The system force and total energy only need to be evaluated once after the converged EVB vectors have been obtained.

The computational cost for an N proton system using the SCI-MS-EVB method scales linearly with respect to the number of excess protons. The parallelization of this method over multiple CPUs can also be implemented efficiently since the solution of the EVB energy and vector of each EVB complex in each iteration is carried out independently. The only communication needed in a parallel calculation is the EVB vectors of the previous iteration. This method has been implemented in the code that has been developed in our group and parallelized over the number of protons. Figures 3 and 4 show the scaling of the method with respect to the number of protons in the system and the number of CPUs used during the simulation, respectively. Linear scaling in both cases is observed.

B. Definition of center of excess charge

In MS-EVB simulations, the position of the center of excess charge (CEC) is usually used as a convenient (continuous) coordinate to track the motion of the protonic defect. In this current study, each EVB complex has its own CEC defined following the definition given in Ref. 33 i.e.,

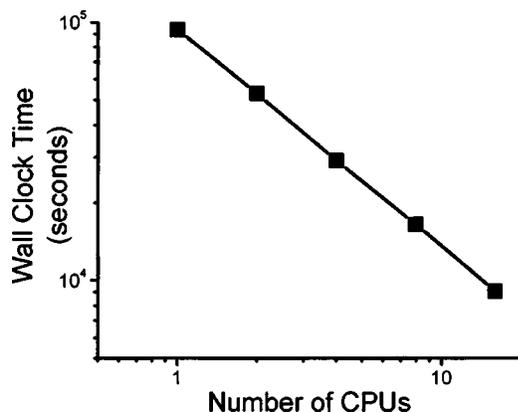


FIG. 4. The wall clock time as a function of the number of CPUs used during the simulation. Calculations were for 2 ps trajectories for a 16 HCl solution with 512 water molecules.

$$\mathbf{r}_{\text{CEC}} = \sum_i^{N_{\text{EVB}}} c_i^2 \mathbf{r}_{\text{CEC}}^i, \quad (14)$$

where $\mathbf{r}_{\text{CEC}}^i$ is the coordinate of the center of charge for the hydronium in state i .

C. Modified definition of asymmetric stretch coordinate q

The MS-EVB2 method as described in Ref. 33 uses the same definition of the asymmetric stretch coordinate q as described by Eq. (22) of Ref. 24. This is given by

$$q = \frac{R_{\text{OO}}}{2} - R_{\text{OH}}, \quad (15)$$

where R_{OH} is defined as the shorter of the two OH bonds formed between the central proton and one of the two oxygen atoms in the H_5O_2^+ . Because of the way this coordinate is defined, when the central proton is away from the OO axis a small cusp in force can be observed when the proton is moved across the OO bisector. Test calculations indicate that the cusp in force does not cause a significant unphysical behavior, but all forces should in principle be continuous. In the SCI-MS-EVB implementation we have therefore decided to remove the cusp in force by defining R_{OH} in Eq. (15) as the distance between the central proton and the center of the OO bond, as done in Ref. 42. All other parameters remain the same as in the MS-EVB2 model. In subsequent discussions we will refer to the modified model as MS-EVB2b and the older version, with the force cusp, as MS-EVB2a.

III. SIMULATIONS

Test simulations were carried out for a 0.44 M HCl solution at 298 K using the MS-EVB2b model. The concentration was obtained in two different ways: (1) by putting one HCl in a cubic periodic box with 128 water molecules and (2) by putting two HCl in a cubic periodic box containing 256 water molecules. The experimental density for HCl of 1.006 g/cm³ was used to determine the box sizes,⁴³ and the Ewald sum was employed in both cases for the long ranged electrostatics. The Cl⁻ anion was modeled as a negative point

charge plus a short-range LJ interaction with parameters $\sigma = 4.400 \text{ \AA}$ and $\epsilon = 0.1000 \text{ kcal/mol}$.^{44,45} Lorentz–Berthelot mixing rules¹ were used to derive the LJ parameters between different atom pairs. In order to prevent the possible divergent attraction between positive point charges on protons and water hydrogen atoms with the Cl⁻ anion (coming from the singularity of $1/r$ Coulombic interaction at extremely short separations), a very short-range LJ potential was added between the Cl⁻ anions and protons, as well as all the hydrogen atoms in the water molecules. The parameters for this interaction were chosen to make sure that the LJ potential between Cl⁻ and the positive charge centers never exceeds 0.1% of all other LJ interactions for all distances above 1 \AA . These parameters were $\sigma = 1.0 \text{ \AA}$ and $\epsilon = 0.0001 \text{ kcal/mol}$. The mass of the most abundant isotope was used for all species.

Both simulation systems were initially equilibrated for 500 ps at constant NVT conditions using a Nosé–Hoover thermostat with a relaxation time of 0.5 ps.⁴⁶ The systems were then propagated at constant NVT for another 100 ps, with configurations saved every 10 ps. A constant NVE simulation was then carried out using each of the saved configurations as initial conditions, with the kinetic energy of each simulation adjusted to ensure that the average temperature of the constant NVE simulation was 298 K. After a short NVE preequilibrium run of 10 ps, the final statistics were accumulated for 6 ns from trajectories for the single HCl system and 3 ns from trajectories for the two HCl system.

IV. RESULTS AND DISCUSSION

Using the Einstein relation,¹ the diffusion constant of the CEC (excess proton) and water molecules was estimated to be $3.2 \pm 0.1 \times 10^{-5}$ and $3.17 \pm 0.03 \times 10^{-5} \text{ cm}^2/\text{ps}$, respectively, when the 256 water simulation box containing two dissociated HCl molecules was employed. The water diffusion constant is in good agreement with the value of $3.1 \pm 0.2 \times 10^{-5} \text{ cm}^2/\text{ps}$ obtained by Day *et al.*³³ using the MS-EVB2a model. The CEC diffusion constant is on the lower side but within the error bar of the $4.0 \pm 0.9 \times 10^{-5} \text{ cm}^2/\text{ps}$ diffusion constant reported by Day *et al.*³³ This slightly lower excess proton diffusion constant may be due to the new definition of the asymmetric stretch coordinate.

When a 128 water simulation box with a single HCl was used, the CEC diffusion constant was estimated to be $3.2 \pm 0.1 \times 10^{-5} \text{ cm}^2/\text{ps}$, which is identical to that estimated using the larger simulation box to within the error bars. The water diffusion constant was estimated to be $2.97 \pm 0.03 \times 10^{-5} \text{ cm}^2/\text{ps}$ when the smaller box is employed. Although a simulation box of this size ($15.697 \text{ \AA} \times 15.697 \text{ \AA} \times 15.697 \text{ \AA}$) has been proven to be large enough to remove finite-size effects for simulations containing one excess proton, presumably the addition of an extra Cl⁻ has increased these effects, thus having a negative influence on water mobility. This interpretation was verified by a separate 2-ns constant NVE simulation on a 128 water one proton system without the Cl⁻ counterion. (The box size was adjusted to

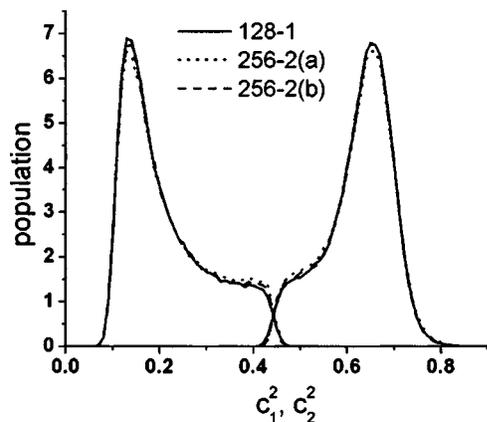


FIG. 5. Relative population of the two largest EVB amplitudes c_1^2 and c_2^2 , where 128-1 indicates the 128 water with 1 HCl system, 256-2 indicates the 256 water with 2 HCl system, and (a) and (b) identify one of the two protons in the 256 water with 2 HCl system.

reflect the density of pure water.) In this case, a water diffusion constant of $3.16 \pm 0.03 \times 10^{-5} \text{ cm}^2/\text{ps}$ was obtained.

Although the calculated proton diffusion constant is not as large as the experimental value of $9.3 \times 10^{-5} \text{ cm}^2/\text{ps}$,⁴⁷ it should be noted that quantum effects have been calculated to increase the proton diffusion by about a factor of two.²⁶ It also seems likely that the MS-EVB potential function could benefit from an overall reparametrization, including the underlying water model used in the potential. These changes in parametrization, however, are not likely to affect the basic mechanistic features of the excess proton diffusion, as described by the MS-EVB model.²¹

In order to determine the relative population of Eigen (H_9O_4^+) versus Zundel (H_5O_2^+) species, the relative population of the largest and second largest MS-EVB amplitudes is reported in Fig. 5. It is seen that at 0.44 M concentration the relative population of Eigen versus Zundel can be obtained reliably by including either one or two excess protons in the periodic box. The agreement between Fig. 5 and previous studies using the MS-EVB2a model is also very good.³³

Determination of site-site radial distribution functions (RDFs) is usually done in MD simulations, since RDFs give important information about liquid state structure and can be compared to experimental measurements. Unfortunately, experimental determination of RDFs for a strong electrolyte solution such as HCl has been a difficult task,⁴⁸ and the accurate determination of the RDFs between atomistic sites has become available only recently due to advances in neutron-diffraction and numerical techniques.⁴⁹⁻⁵¹

For the experimental RDFs of HCl solutions, we refer the reader to a recent publication by Botti *et al.*,⁵² in which the experimental composite partial structure factors (CPSFs) are combined with simulation data using the empirical potential structure refinement (EPSR) methodology to obtain a full set of RDFs at high HCl concentrations (6 M).⁵² In the EPSR procedure, simulations using the simple point-charge/extended (SPC/E) water model are carried out, but the force field is iteratively refined to reproduce experimental measurements. However, in their modeling the excess proton will only participate in a H_3O^+ hydronium state. No Zundel-type H_5O_2^+ protonation state was allowed.

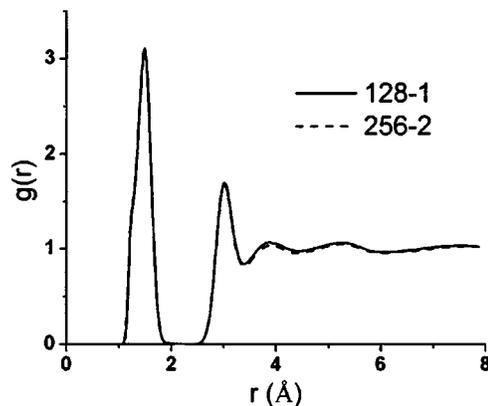


FIG. 6. Hydronium hydrogen (HH) and water oxygen (OW) RDFs. 128-1 indicates the 128 water with 1 HCl system, while 256-2 indicates the 256 water with 2 HCl system.

Figure 6 reports the RDF between hydronium hydrogen (HH) and the oxygen atom in water (OW). In order to be consistent with the hydronium model used in the approach of Botti *et al.*, all three H atoms associated with the EVB state having the largest amplitude are classified as HH. (This is true for all subsequent figures.) Figure 6 is very similar to the corresponding figure in their experimental study. Both figures show the first peak of the RDF, which corresponds to the hydrogen-bonded OW, to be around 1.5 Å. Also, a large depression at 2.2 Å is observed; however, the depression predicted by the MS-EVB2b model appears too wide and deep and the second peak of the RDF curve is pushed somewhat further away compared to the experimental curve. This may be due to an overestimation in the parametrization of the repulsive interactions between the hydronium oxygen and the water oxygen and hydrogen atoms in the MS-EVB2 model. Another possible explanation is due to the restriction of the model used in the EPSR fitting, which only allows for pure hydronium configurations. It should also be noted, however, that the experiments are carried out at more than ten times the HCl concentrations of our simulations. Integration of the RDF indicates that the HCl is fully dissociated since one hydrogen bond is formed for each HH atom.

Figure 7 shows the RDF curve between the oxygen atom in the hydronium (OH) and OW, alongside with the RDF

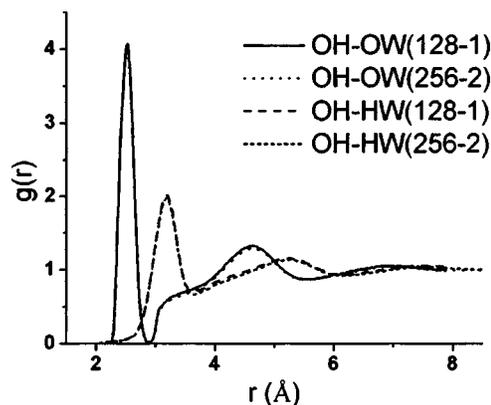


FIG. 7. Hydronium oxygen (OH) and OW RDFs and OH and water hydrogen (HW) RDFs. 128-1 indicates the 128 water with 1 HCl system, while 256-2 indicates the 256 water with 2 HCl system.

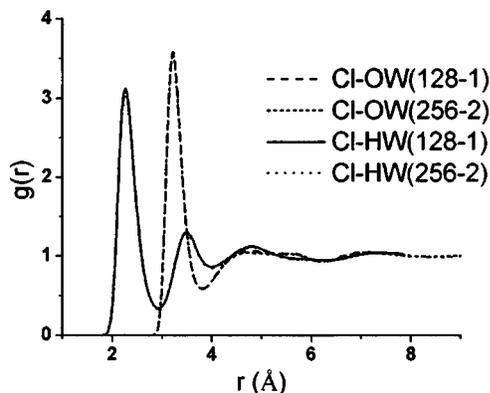


FIG. 8. RDF between Cl^- and hydrogen and oxygen atoms in the water molecules. 128-1 indicates the 128 water with 1 HCl system, while 256-2 indicates the 256 water with 2 HCl system.

between OH and hydrogen atoms in bulk water (HW). The simulation shows the absence of a OH–HW RDF peak at around 2 Å, corresponding to the absence of hydrogen bonding with OH. Immediately after the first peak in the OH–HW RDF around 3.1 Å the onset of a shoulder in the OH–OW RDF is observed. These features agree very well with the experimental measurements after taking into account the large difference in concentrations. It can be seen in this figure as well that the second peak in the OH–OW RDF is slightly pushed away from the origin and the depression in the RDF curve at 2.8 Å is too deep compared with the experimental measurements at the higher HCl concentration.

Figure 8 shows the RDF between Cl^- and OW and HW. When comparing with the experimental curves the first two peaks of the Cl^- –HW RDF are significantly higher and so is the first peak of the Cl^- –OW RDF. This is expected since our simulation concentration, which corresponds to one proton per 128 water molecules, is much lower than the experimental concentration,⁵² which corresponds to about one proton per 9 water molecules. Thus HCl is fully dissociated and Cl^- is better solvated in our simulations.

From the RDF curves reported in Figs. 6–8, it can be seen that there exists a reasonable agreement between the MS-EVB2b simulation results and the experimental measurements, especially considering the large differences in concentrations. The repulsive component between OH and OW and OH and HW may be slightly too strong in the MS-EVB2b model and this will be addressed in a later version of the force field. All of the RDF curves in this study were obtained using both the one HCl and two HCl systems. In Figs. 6–8, good agreement between these two systems was also observed. It therefore seems apparent that there is little correlation between the excess protons up to 0.44 M HCl concentration.

Figure 9 shows the RDFs between Cl^- and the CEC and Cl^- and OH. A common feature of these curves is the existence of two peaks corresponding to the contact ion-pair (CIP) configuration and solvent separated ion-pair (SSIP) configurations.^{44,53–55} However, the RDFs obtained using the one HCl and two HCl systems are rather different. Although in each case the location of the two peaks is the same for both the one HCl and two HCl systems, the RDFs obtained

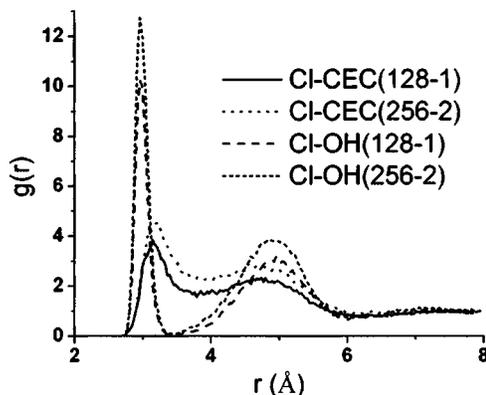


FIG. 9. RDF between Cl^- and the OH and the RDF between the Cl^- and the CEC. 128-1 indicates the 128 water with 1 HCl system, while 256-2 indicates the 256 water with 2 HCl system.

using the two-proton system have significantly larger peak heights. This difference reveals a deficiency of the 128 water plus single HCl system. In a 0.44 M simulation with one proton plus one Cl^- in 128 water molecules, the Cl^- can never have more than one excess proton in either its CIP region or its SSIP region. Thus, the addition of a second HCl, as was done for the 256 water 0.44 M system, significantly changes the height of the peaks in the RDF curves.

The Cl^- –OH RDF is integrated in Fig. 10 to calculate the average number of hydronium in each region. It is found that 0.067 hydronium is in the CIP region when the curve obtained using the single HCl plus 128 water simulation is integrated, but 0.086 hydronium is in the CIP region (0–3.6 Å) when the curve obtained using the two HCl plus 256 water simulation is integrated. Thus, at 0.44 M, around 1% of the Cl^- anions have two hydronium in its CIP region. This also indicates, at this concentration, that Cl^- is 6.5 times more likely to have a single hydronium in its CIP region than having two hydronium in this region. When the same analysis is carried out for the SSIP region (3.6–6.0 Å), it is found that an average 0.30 hydronium is present in the SSIP region for the 0.44 M case of one proton per simulation box (with 128 waters), whereas 0.38 hydronium is found when allowing two protons in the box (with 256 waters). Thus, about 4% of the Cl^- have two protons in their SSIP region at this

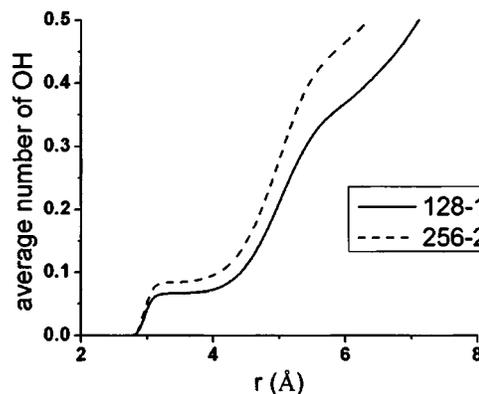


FIG. 10. Average number of hydronium oxygen atoms within a distance r from Cl^- . 128-1 indicates the 128 water with 1 HCl system, 256-2 indicates the 256 water with 2 HCl system.

concentration, and there is an ~ 7.5 ratio between one proton in the SSIP region and two protons in the SSIP region.

Some caution is in order concerning the above discussion, however, since we have yet to simulate the 0.44 M HCl solution with three or more excess protons (though this is currently in progress). However, the effect due to inclusion of more protons is not expected to be large, since the ratio of having two protons versus one proton in the CIP and SSIP regions is already quite small.

The above findings would also indicate that the first and second peaks for the Cl^- -OW, Cl^- -HW, OH-OW, and OH-HW RDFs should also see a small decrease in peak height upon increasing the number of HCl pairs per simulation box. This is indeed observed by carefully examining the peak heights in Figs. 7 and 8. The first two peaks obtained using the two HCl per simulation box are slightly lower in all cases.

Another interesting finding is although the peaks for the SSIP configuration for the Cl^- -CEC RDF and the Cl^- -OH RDF are almost on top of each other, the peaks for the CIP configuration are not, with the CIP peak for the CEC being shifted farther away from Cl^- . Actually, at the location where the OH is most abundant, there is hardly any population corresponding to the CEC. The location of the CEC will be the same as that of OH for Eigen cation configurations, whereas it will shift toward the center of the two O atoms in Zundel configurations and along the bisector of the two OH-HH bonds in H_7O_3^+ -type configurations. The outward shift of the CEC CIP peak indicates a dominance of a more Zundel-type or possibly a H_7O_3^+ configuration. This is not surprising since a Cl^- , being a poor base, is not as good at solvating an excess proton as a water molecule. Once the hydronium is in close contact with the Cl^- as in the CIP configuration, one or more of the water molecules at the other side of the hydronium must contribute more in stabilizing the proton, giving a ground state in which more Zundel-type or H_7O_3^+ character results. This is consistent with experimental findings where it is claimed that H_5O_2^+ is the dominant species at very high HCl concentrations.⁵⁶

In order to validate the above assessment, the distribution of the first and second largest EVB coefficients for both the one HCl 128 water system and a reference system with only one excess proton and 128 water molecules is plotted in Fig. 11. Consistent with our argument, we do observe a decrease in Eigen configurations in the system with the Cl^- present. This is manifested by a lowering in the peak height and heightening of the plateau. This finding also casts some doubt on the experimental Cl^- -OH curve reported in Ref. 52, since only H_3O^+ configurations were included in their model.

It should be possible to calculate Cl^- - Cl^- , OH-OH, and CEC-CEC RDFs based on the trajectory saved from the 256 water plus two HCl simulation. However, the finite-size effect of a limited number of HCl pairs in the same simulation cell is more severe for these RDFs. Simulations including more HCl per simulation box will therefore be required to get statistically reliable curves. Because of this, we have opted not to report this data. Such simulations are in progress and will be reported in subsequent publications.

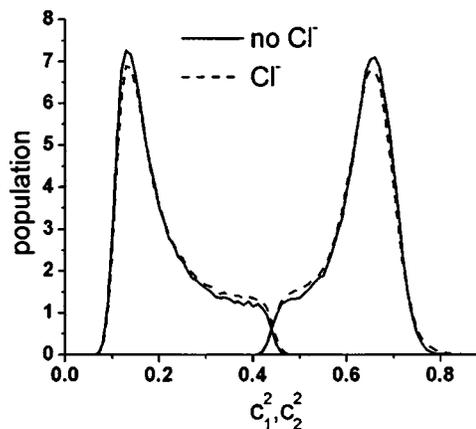


FIG. 11. Relative population of the two largest EVB amplitudes c_1^2 and c_2^2 for the one proton 128 water systems with and without the Cl^- counterion.

V. CONCLUDING REMARKS

In this paper the SCI-MS-EVB method has been presented that enables the treatment of multiple excess protons in aqueous systems. This approach represents a significant and nontrivial generalization of the MS-EVB model. The SCI-MS-EVB method also allows one to treat a multiproton problem with computational cost scaling linearly with respect to the number of excess protons and in a fashion that is highly parallelizable.

In comparing with experimental data, the present simulations indicate that the SCI-MS-EVB method can be used to calculate structural and dynamical properties for multiproton systems with good accuracy. Whereas the calculation of some properties can be done with a single excess proton approach for the concentration studied (0.44 M), the ability to simulate multiple protons in the same box is important for properties such as the Cl^- -OH and OH-OH RDFs. The probability of finding more than one hydronium in the CIP and SSIP regions of the Cl^- was also found to be non-negligible. Zundel cation or possibly H_7O_3^+ -type cation solvation states are found to be especially important for stabilizing the excess proton in the CIP region of a Cl^- anion. Further simulations will be required to study the solvation structures of concentrated acid as a function of pH. The SCI-MS-EVB method will also enable us to study important biological systems in which multiproton interactions may play a key role.

ACKNOWLEDGMENTS

This work was supported in part by the National Science Foundation (Grant No. CHE-0317132) and in part by the U. S. Army Research Laboratory and the U. S. Army Research office under Grant No. DAAD 19-03-1-0121. The authors thank Professor H. B. Schlegel for the helpful discussions concerning the derivation of the forces in SCI-MS-EVB.

¹M. P. Allen and D. J. Tildesley, *Computer Simulation of Liquids* (Clarendon, Oxford, 1987).

²in *Quantum Simulations of Condensed Matter Phenomena*, edited by J. D. Doll and J. E. Gubernatis (World Scientific, Singapore, 1990).

³C. Galli and M. Parrinello, in *Computer Simulations in Material Science*, edited by M. Meyer, and V. Pontikis (Kluwer, Dordrecht, 1991), p. 283.

⁴A. K. Rappé, C. J. Casewit, K. S. Colwell, W. A. Goddard III, and W. M. Skiff, *J. Am. Chem. Soc.* **114**, 10024 (1992).

- ⁵W. D. Cornell, P. Cieplak, C. I. Bayly *et al.*, *J. Am. Chem. Soc.* **117**, 5179 (1995).
- ⁶B. R. Brooks, R. E. Bruccoleri, B. D. Olafson, D. J. States, S. Swaminathan, and M. Karplus, *J. Comput. Chem.* **4**, 187 (1983).
- ⁷A. D. MacKerell, Jr., B. Brooks, C. L. Brooks III, L. Nilsson, B. Roux, Y. Won, and M. Karplus, in *The Encyclopedia of Computational Chemistry*, edited by P. v. R. Schleyer (Wiley, Chichester, NY, 1998), Vol. 1.
- ⁸R. A. Robinson and R. H. Stokes, *Electrolyte Solutions*, 2nd ed. (Butterworths, London, 1959).
- ⁹D. Eisenberg and W. Kauzman, *The Structure and Properties of Water* (Oxford University Press, Oxford, 1969).
- ¹⁰R. P. Bell, *The Proton in Chemistry* (Cornell University Press, Ithaca, NY, 1973).
- ¹¹T. Erdey-Grúz, *Transport Phenomena in Aqueous Solutions* (Hilger, London, 1974).
- ¹²P. W. Atkins, *Physical Chemistry*, 5th ed. (Freeman, New York, 1994).
- ¹³E. Hückel, *Z. Elektrochem.* **34**, 546 (1928).
- ¹⁴J. D. Bernal and R. H. Fowler, *J. Chem. Phys.* **1**, 515 (1933).
- ¹⁵B. E. Conway, in *Modern Aspects of Electrochemistry*, edited by J. O. M. Bockris and B. E. Conway (Butterworths, London, 1964), p. 43.
- ¹⁶M. Eigen, *Angew. Chem., Int. Ed.* **3**, 1 (1964).
- ¹⁷N. Agmon, *Chem. Phys. Lett.* **244**, 456 (1995).
- ¹⁸M. E. Tuckerman, K. Laasonen, M. Sprik, and M. Parrinello, *J. Chem. Phys.* **103**, 150 (1995).
- ¹⁹H. Danneel, *Z. Elektrochem. Angew. Phys. Chem.* **11**, 249 (1905).
- ²⁰C. J. T. de Grotthuss, *Ann. Chim. (Paris)* **58**, 54 (1806).
- ²¹H. Lapid, N. Agmon, M. Peterson, and G. A. Voth, *J. Chem. Phys.* **122**, 014506 (2005).
- ²²U. W. Schmitt and G. A. Voth, *J. Phys. Chem. B* **102**, 5547 (1998).
- ²³U. W. Schmitt and G. A. Voth, *Isr. J. Chem.* **39**, 483 (1999).
- ²⁴U. W. Schmitt and G. A. Voth, *J. Chem. Phys.* **111**, 9361 (1999).
- ²⁵M. Cuma, U. W. Schmitt, and G. A. Voth, *Chem. Phys.* **258**, 187 (2000).
- ²⁶U. W. Schmidt and G. A. Voth, *Chem. Phys. Lett.* **329**, 36 (2000).
- ²⁷T. J. F. Day, U. W. Schmitt, and G. A. Voth, *J. Am. Chem. Soc.* **122**, 12027 (2000).
- ²⁸M. L. Brewer, U. W. Schmitt, and G. A. Voth, *Biophys. J.* **80**, 1691 (2001).
- ²⁹M. Cuma, U. W. Schmitt, and G. A. Voth, *J. Phys. Chem. A* **105**, 2814 (2001).
- ³⁰J. Kim, U. W. Schmitt, J. A. Gruetzmacher, G. A. Voth, and N. F. Scherer, *J. Chem. Phys.* **116**, 737 (2002).
- ³¹A. M. Smondyrev and G. A. Voth, *Biophys. J.* **82**, 1460 (2002).
- ³²A. M. Smondyrev and G. A. Voth, *Biophys. J.* **83**, 1987 (2002).
- ³³T. J. F. Day, A. V. Soudachov, M. Cuma, U. W. Schmitt, and G. A. Voth, *J. Chem. Phys.* **117**, 5839 (2002).
- ³⁴Y. Wu and G. A. Voth, *Biophys. J.* **85**, 864 (2003).
- ³⁵B. Ilan, E. Tajkhorshid, K. Schulten, and G. A. Voth, *Proteins: Struct., Funct., Bioinf.* **55**, 223 (2004).
- ³⁶M. K. Peterson, S. S. Iyengar, T. J. F. Day, and G. A. Voth, *J. Phys. Chem. B* **108**, 14804 (2004).
- ³⁷R. T. Morrison and R. N. Boyd, *Organic Chemistry*, 6th ed. (Prentice-Hall, New Jersey, 1992).
- ³⁸A. Szabo and N. S. Ostlund, *Modern Quantum Chemistry* (Dover, New York, 1996).
- ³⁹J. Aqvist and A. Warshel, *Chem. Rev. (Washington, D.C.)* **93**, 2523 (1993).
- ⁴⁰A. Warshel, *Computer Modelling of Chemical Reactions in Enzymes and Solutions* (Wiley, New York, 1991).
- ⁴¹J. Lobaugh and G. A. Voth, *J. Chem. Phys.* **104**, 2056 (1996).
- ⁴²R. Vuilleumier and D. Borgis, *J. Chem. Phys.* **111**, 4251 (1999).
- ⁴³D. R. Lide, *CRC Handbook of Chemistry and Physics*, 79th ed. (CRC, Boca Raton, FL, 1998).
- ⁴⁴D. E. Smith and L. X. Dang, *J. Chem. Phys.* **100**, 3757 (1993).
- ⁴⁵A. P. Lyubartsev and A. Laaksonen, *J. Phys. Chem.* **100**, 16410 (1996).
- ⁴⁶S. Melchionna, G. Ciccotti, and B. L. Holian, *Mol. Phys.* **78**, 533 (1993).
- ⁴⁷B. Hille, *Ionic Channels in Excitable Membranes*, 2nd ed. (Sinauer Associates Inc., Sunderland, MA, 1992).
- ⁴⁸R. Triolo and A. H. Narten, *J. Chem. Phys.* **63**, 3624 (1975).
- ⁴⁹A. K. Soper, in *Advanced Neutron Sources 1988*, IOP Conf. Proc. No. 97, edited by D. K. Hyer (IOP, Bristol, 1989).
- ⁵⁰A. K. Soper, *Chem. Phys.* **202**, 295 (1996).
- ⁵¹F. Bruni, M. A. Ricci, and A. K. Soper, in *Francesco Paolo Ricci: His Legacy and Future Perspectives of Neutron Scattering*, edited by M. Nardone and M. A. Ricci (Italian Physical Society Conference Proceedings, Bologna, 2001), Vol. 76, p. 37.
- ⁵²A. Botti, F. Bruni, S. Imberti, M. A. Ricci, and A. K. Soper, *J. Chem. Phys.* **121**, 7840 (2004).
- ⁵³S. Winstein, E. Clippinger, A. H. Fainberg, and G. C. Robinson, *J. Am. Chem. Soc.* **76**, 2597 (1954).
- ⁵⁴H. Sadek and R. M. Fuoss, *J. Am. Chem. Soc.* **76**, 5905 (1954).
- ⁵⁵D. E. Irish and M. H. Brooker, in *Advances in Infrared and Raman Spectroscopy*, edited by R. J. H. Clark and R. E. Hester (Heydon, London, 1976), Vol. 2.
- ⁵⁶N. Agmon, *J. Phys. Chem. A* **102**, 192 (1998).