Proton transfer in bulk water by full adaptive QM/MM method: Integration of solute- and solvent-adaptive approaches

Hiroshi C. Watanabe*[†][‡]*, Masayuki Yamada*[§]*, and Yohichi Suzuki*[†]

[†]Quantum Computing Center, Keio University, 3-14-1 Hiyoshi, Kohoku-ku, Yokohama 223-8522, Japan

[‡]PRESTO Japan Science and Technology Agency, 4-18, Honcho, Kawaguchi, Saitama 332-0012, Japan

[§]Graduate School of Information Science and Technology, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8656, Japan

ABSTRACT The quantum mechanical/molecular mechanical (QM/MM) method is a hybrid molecular simulation technique that makes local electronic structures of large systems accessible. It has the strengths of accuracy found in the QM method as well as the strengths of small computational costs found in the MM method. However, it is severe to directly apply the QM/MM method to dynamics of solution systems, particularly to proton transfer. As explained in the Grotthuss mechanism, proton transfer is a structural interconversion between hydronium ion and solvent water molecules. Hence, when the QM/MM method is applied, an adaptive treatment, namely on-the-fly revisions on molecular definitions, is required for both the solute and solvent. Although there have been several solvent-adaptive methods proposed, a full adaptive framework, an approach that also takes into account of adaptation for solutes, still remains untapped. In this paper, we propose a new numerical expression for the coordinate of the excess proton and its control algorithm. Furthermore, we confirmed that this method can stably and accurately simulate proton transfer dynamics in bulk water.

INTRODUCTION

Proton transfer is one of the important phenomena in biology, engineering, and solution chemistry. Although the first proton transfer model was proposed 200 years ago by von Grotthuss, details on the mechanism of proton transfer and solvation structure of hydronium ion (whether it is an Eigen vs. Zundel cation) has long been unclear because experimental observations cannot be straightforwardly interpreted and such cations cannot be uniquely distinguished. In the late nineties and the early noughties, however, developments of computation have facilitated to achieve plausible simulation and advance knowledge on the

mechanism.[1–10] The proton transfer is accompanied by creation/annihilation of covalent bonds between oxygen and hydrogen, which is known as "structural diffusion", and requires quantum mechanical (QM) description that can explicitly count the electronic structures. Therefore, in particular, *ab initio* molecular dynamics (AIMD) played the main role for the progress of proton transfer studies. In AIMD, electronic structure of the entire system is evaluated to obtain potential energy and forces acting on respective atoms.[2,3,7,8,11] On the other hand, AIMD is severely limited in the system size it can handle as well as the duration of the molecular dynamics (MD) simulation due to tremendous computational cost. For instance, the relaxation of the solvation structure including the second solvation shell is supposed to be a rate-limiting process for proton transfer. Therefore, the solvation shell should properly interact with the surrounding environment in proton transfer dynamics. Indeed, a recent study had demonstrated that the radial distribution around the hydroxide ion depends on the system size.[10] However, the extension of the system size brings about a critical explosion of computational cost. Obviously, large molecules such as proteins are also beyond the scope of AIMD, despite the great demands for them in fields such as chemical engineering and biochemistry. Car-Parrinello MD (CPMD),[12] which is one of the most popular AIMD, had introduced fictitious mass for electron. However, anomalously large value has been employed for the fictitious mass to ensure adiabaticity, which can distort hydrogen dynamics as deuterium.[9] It is worth noting that divide-and-conquer (DC) treatment,[13,14] in which the entire system is fragmented into a number of subsystems, reduces the CPU time, for instance, to $O(n^{1.2})$ when used in combination with the density functional tight-binding (DFTB) method[15–17] for homogeneous water systems, where n represents the number of water molecules. However, the DC treatment still requires a massive number of CPUs in accordance with the number of subsystems. In addition, DC treatment can cause discontinuity

when particles diffuse across subsystem borders, although it is assumed to be mitigated to some extent by introduction of buffer zones. As a result, challenges for quantitative investigation still remain in proton transfer mechanism. On the other hand, a hybrid simulation, Quantum mechanics/Molecular mechanics (QM/MM) method, may be an appealing alternative option, because it can reduce computational cost by partially applying QM calculation to the system.[18] Since the computational cost of QM/MM method mainly depends on the size of the local QM region rather than that of the whole system, the QM/MM method has been widely employed in researches for large molecular systems. However, QM/MM method cannot be directly applied to proton transfer dynamics simulation, requiring adaptive treatments to be made for both the solute and solvent.

In conventional QM/MM simulations for solutions, a solute of interest is placed at the center of the QM region so as to be surrounded by QM solvent molecules and molecular definitions are fixed throughout the MD simulation. However, due to free diffusion, the surrounding QM water molecules are replaced by other water molecules (MM). Furthermore, the transferred proton itself is not consistent throughout the MD simulation because of structural diffusion as explained by the Grotthuss mechanism. In such case, the position of the hydronium ion can deviate from the QM center and diffuse across the QM/MM border, leading to the collapse of the MD simulation. Therefore, the definition of the excess proton and the surrounding solvent molecules has to be adaptively updated during the MD simulation so that the hydronium ion is constantly located at the QM center surrounded by QM water molecules.

Several solvent-adaptive QM/MM methods have been proposed to date.[19–38] To understand solvent-adaptive QM/MM methods, it is useful to introduce the concept of QM/MM partitioning, which describes how the entire system is divided into QM and MM regions. In general, adaptive

QM/MM methods are based on multi-partitioning approaches where more than one QM/MM partitions are considered for every MD time step. These partitions share the same QM solute molecule, while they have different numbers and combinations of QM solvent molecules. Potential energies and forces are independently evaluated for respective QM/MM partitions. In the end, the resulting potentials or forces are linearly combined to evaluate effective potential or forces that are used to update the coordinates for the MD simulation. In general, adaptive QM/MM Hamiltonian can be represented as

$$H = \sum_i \frac{p_i^2}{2m_i} + \sum_n^N \sigma^{(n)}(\{r_{i\xi}\})V^{(n)}(\{r_{ij}\}) - \int dq \sum_n^N \frac{\partial \sigma^{(n)}}{\partial q} V^{(n)} \qquad (1)$$

Here, r is the distance between particles where a subscript ξ represents the QM center, while alphabetic subscripts represent particles. $V^{(n)}$ and $\sigma^{(n)}$ are the potential energy and weight function for the nth partitioning. The second term, called the "effective potential", is the weighted sum over potential energies of all partitions. The third term is the "bookkeeping term" introduced to cancel out artificial forces that arise from the derivatives of the weight function.[33,39] As a result, the effective force acting on the ith particle becomes

$$F_i^{\text{eff}} = \sum_n^N \sigma^{(n)} F_i^{(n)} \qquad (2)$$

where $F_i^{(n)}$ represents force acting on the ith particle evaluated for the nth partitioning. Although adaptive QM/MM approaches enable incorporation of quantum chemical effects of solvation into the MD simulation, most approaches suffer from severe artifacts that are termed "temporal" and "spatial" discontinuities. Since we have discussed the discontinuities elsewhere,[40,41] we will briefly introduce them here. Temporal discontinuity is rephrased as the violation of the Hamiltonian conservation caused by discontinuities in the effective potential energy surface.

Some solvent-adaptive QM/MM methods, such as the sorted adaptive partitioning (SAP)[31] and size-consistent multipartitioning (SCMP)[23] methods, are free from temporal discontinuity. On the other hand, spatial discontinuity is manifested as the monotonic drift of the bookkeeping term during the course of the MD simulation. Although some *ad hoc* corrections have been proposed,[25,41] spatial discontinuity is inevitable for any QM/MM method, because it arises from unnatural manipulation of dividing a homogeneous solution into different layers. Therefore, to be fair, static QM/MM method should also be subject to spatial discontinuities, not only adaptive QM/MM. However, spatial discontinuity has not been deeply discussed because it has not been supposed to be a critical factor in static simulation. Note that, compared to multi-size approaches such as SAP[31] and difference-based adaptive solvation (DAS),[33] spatial discontinuity can be rather suppressed by size-consistent treatment, in which the number of the QM solvent is consistent among partitions.[41]

Towards a solute-adaptive method, the first step is to numerically express the position of the excess proton, which is termed "excess proton indicator (EPI)." To date, several groups have addressed the developments and applications of EPI.[42–46]. Previously, Chakrabarti et al propose the expression of the EPI $\boldsymbol{\xi}$ as

$$\boldsymbol{\xi} = \frac{W_i \boldsymbol{r}_i}{\sum_i W_i} \tag{3}$$

where $\boldsymbol{r}_i$ is coordinates of the ith oxygen atom.[44] The weight function for the ith oxygen atom can be written as

$$W_i = \sum_j \frac{1}{1 + \exp\left[\frac{r_{ij} - r_o}{d}\right]} - 2 \tag{4}$$

where r_{ij} is the distance between the ith oxygen and jth hydrogen atoms, while r_o and d are parameters. The weight function W_i defined by eq 4 indicates that hydronium ion is more weighted to estimate EPI in eq 3 than ordinary water molecule whose value is close to zero. Although this indicator has been used as the reaction coordinate in post-processing of MD trajectory for proton transfer in membrane proteins, it cannot be used for an adaptive simulation of bulk water because of its scale-dependency, discontinuity, instability, and computational cost. The problems above mainly arise from the fact that weights for ordinary water molecules distant from the hydronium ion in eq 3 are not necessarily zero. Note that the residual contribution is not negligible and can cause critical error when accumulated. To figure it out, let's suppose that a simulation cell filled with water molecules including a solvated hydronium ion. In the case of a small simulation cell, EPI in eq 3 and 4 may properly work, with EPI corresponding to the position of the hydronium ion throughout the simulation. As the system size increases, however, the indicator would point to the geometric center of the simulation cell regardless of the position of the hydronium ion in reality, because the accumulated residual contributions from ordinary water become comparable to that of the hydronium ion (scale-dependency). In addition, when non-zero weighted water molecules move across the periodic boundary condition, the EPI becomes discontinuous, causing it to violate the Hamiltonian conservation. As a result, temperature unnaturally increases, destabilizing MD simulations (discontinuity). Furthermore, even an ordinary water distant from the hydronium ion can temporarily have a value as large as that of the hydronium ion, when it forms strong hydrogen bonds. This results in extraordinary displacement of the QM center, destabilizing the MD simulation (instability). It should also be noted that the computational cost to evaluate weight N water molecules is almost $2N^2$ considering all possible oxygen-hydrogen pairs and thus, the computational cost for it will

become as large as force evaluation, which is the bottleneck of MD simulations (computational cost).

To avoid such problems, the proton indicator should consist of continuous functions of coordinates of all particles but also be able to detect local structural attributes around the hydronium ion. Although it is obvious that the molecular geometries of solvent water and the solvation structure in the vicinity of the hydronium ion differ from bulk water, it is not trivial whether any index represented by a continuous function (which can identify the position of the hydronium ion without noise and error throughout the MD simulation) exists. Neither is it trivial to control the dynamics of the QM center during MD simulation. As shown in Eq (1), the weight $\sigma^{(n)}$ in solvent-adaptive method should be a function of the distances between particles. Otherwise, the effective force in Eq (2) is not conserved and cannot be derived from eq 1.

In order to achieve an accurate and stable MD simulation, we propose a modified representation for EPI. In addition, we propose a new protocol to control the indicator during the MD simulation, in which we introduce a virtual particle representing the QM center with use of constraint dynamics, the RATTLE method.[47] Lastly, we demonstrate the benchmark simulation for proton transfer in bulk water, in which the Hamiltonian is well conserved, achieving a stable and durable MD simulation.

THEORY AND METHOD

Excess proton indicator. We introduced a virtual site for respective oxygen atoms in hydronium ions and water molecules. Let $r_{i\alpha}$ be defined as the distance between the ith oxygen atom and the

αth hydrogen atom, and throughout this paper, alphabetic and Greek subscripts represent oxygen and hydrogen atoms, respectively. Then, its Gaussian-type score function $\phi'(r)$ is defined as

$$\phi'_{i\alpha}(r_{i\alpha}) = \exp\left[\frac{(r_{i\alpha} - r_0)^2}{\alpha^2}\right] \qquad (5)$$

where α and r_0 are parameters (see Discussion for detail). Next, we supposed a normalized score $\phi_{i\alpha}$ as follows.

$$\phi_{i\alpha} = \frac{\phi'_{i\alpha}}{\sum_\lambda \phi'_{i\lambda}} \qquad (6)$$

The virtual site $\boldsymbol{\eta}_i$ for the ith oxygen atom was defined using hydrogen coordinates $\{\boldsymbol{r}_\lambda\}$ as

$$\boldsymbol{\eta}_i = \sum_\lambda \phi_{i\lambda} \boldsymbol{r}_\lambda \qquad (7)$$

As the next step, a function ψ_i for the ith oxygen is defined by summation of the spline function S_1 of the distance $r_{i\alpha}$ between the ith oxygen and the αth hydrogen,

$$\psi_i = \sum_\alpha S_1(r_{i\alpha}) \qquad (8)$$

Here, S_1 satisfied the following boundary conditions:

$$\begin{array}{ll} S_1(R_{\text{small}}) = 1 & S_1'(R_{\text{small}}) = 0 \\ S_1(R_{\text{large}}) = 0 & S_1'(R_{\text{large}}) = 0 \end{array} \qquad (9)$$

where the spline function $S_1(r_{i\alpha})$ is a bonding score between the ith oxygen and the αth hydrogen, which ranges from zero to one. Note that ψ_i can be regarded as a continuous expression of the number of covalent bonds of the ith oxygen atom (See Figure 1). Here, R_{small} and R_{large} are bonding parameters that satisfy $R_{\text{small}} < R_{\text{large}}$. Note that, although we applied the spline curve here, the function S_1 is arbitrary unless it satisfies eq 9. Next, ψ_i of the ith

oxygen atom was used as an argument for another spline function $S_2(\psi_i)$, which satisfied the following boundary conditions

$$
\begin{aligned}
S_2(2.0) = 0 \quad & S_2{}'(2.0) = 0 \\
S_2(3.0) = 1 \quad & S_2{}'(3.0) = 0
\end{aligned}
\tag{10}
$$

Based on S_2, the weight W_i for the ith oxygen atom was defined as

$$
W_i = \frac{S_2(\psi_i)}{\sum_m S_2(\psi_m)}
\tag{11}
$$

Finally, the excess proton indicator (EPI) $\boldsymbol{\xi}$, which was used as the QM center in the SCMP simulation, was defined as the weighted sum of internally dividing points between the oxygen coordinate $\boldsymbol{r}_i$ and the virtual site $\boldsymbol{\eta}_i$ over all solvent oxygen atoms as

$$
\boldsymbol{\xi} = \sum_i W_i(c\boldsymbol{\eta}_i + (1-c)\boldsymbol{r}_i)
\tag{12}
$$

When the parameter $c = 1$, the EPI is a linear combination of virtual sites $\boldsymbol{\eta}_i$ with weight W_i. In contrast, when $c = 0$, the EPI is a linear combination of water oxygen coordinates $\{\boldsymbol{r}_i\}$, which is indirectly subject to the water hydrogen coordinate through weight W_i. And thus, when the value of c falls in between 0 and 1, the EPI becomes a dividing point between the two positions (Figure 1).

Weight function in size-consistent multipartitioning (SCMP) method. Here, we briefly review the SCMP method. The details of these functions are written in previous literatures. [23,41,48] The weight function $\sigma^{(n)}$ in eq 1 is defined as below for the SCMP method.

$$
\sigma^{(n)}(r_{j\xi}) = \frac{O_{QM}^{(n)}I_{QM}^{(n)}O_{MM}^{(n)}I_{MM}^{(n)}}{\sum_k^N O_{QM}^{(k)}I_{QM}^{(k)}O_{MM}^{(k)}I_{MM}^{(k)}}
\tag{13}
$$

Here, the EPI ξ was adopted as the atom in the QM center. $O_{\mathrm{QM}}^{(n)}$ and $O_{\mathrm{MM}}^{(n)}$ are fade-out functions

for QM and MM solvent molecules in the nth partitioning, respectively, and are defined as

$$O^{(n)} = \prod_{j}^{m} \lambda_j^{(n)}(r_{j\xi}) \qquad (14)$$

where $\lambda_j^{(n)}$ is the progress function of the respective jth QM or MM solvent, which continuously

ranges between zero and one. $I_{\mathrm{QM}}^{(n)}$ and $I_{\mathrm{MM}}^{(n)}$ are fade-in functions for QM and MM solvent

molecules, respectively, and are defined as

$$I^{(n)} = 1 - O^{(n)} \qquad (15)$$

Constraint for the QM center. Using eqs 13-15, the weight of each partitioning was evaluated

based on the distance between the EPI and respective QM and MM solvent water molecules. For

Hamiltonian conservation, the weight has to be a function of a coordinate for any particle. Hence,

we placed a dummy atom at the position corresponding to ξ. The dummy atom does not directly

interact with other particles, but indirectly interacts through constraint. We chose the mass of the

dummy atom to be 1.0×10^{-8} u, which is small enough to well satisfy the constraint condition

and not affect other particles. The velocity Verlet integrator was employed for MD in the present

study, and thus, the RATTLE algorithm[47] was applied to control the constraint. In addition to Eq.

12, the constraint condition in respect to velocity was given as

$$\dot{\xi} = \sum_i \left(\kappa_i(c\dot{\boldsymbol{\eta}}_\iota + (1-c)\dot{\boldsymbol{r}}_\iota) + \frac{d\kappa_i}{dt}\boldsymbol{\eta}_i \right) \qquad (16)$$

The MD algorithm for the ith particle is as follows,

(i) $\quad v'_i = v_i(t) + \dfrac{\Delta t}{2m_i}F_i(t)$

(ii) $\quad r_i(t+\Delta t) = r_i(t) + \Delta t v'_i - \dfrac{\Delta t^2}{2m_i}\dfrac{\partial \xi}{\partial r_i}\lambda_c$

(iii) $\quad v_i\left(t + \dfrac{1}{2}\Delta t\right) = v'_i - \dfrac{\Delta t^2}{2m_i}\dfrac{\partial \xi}{\partial r_i}\lambda_c$

(iv) $\quad v_i(t+\Delta t) = v_i\left(t + \dfrac{1}{2}\Delta t\right) + \dfrac{\Delta t}{2m_i}F_i(t+\Delta t) - \dfrac{\Delta t}{2m_i}\dfrac{\partial \xi}{\partial r_i}\lambda_v$

(17)

where λ_c and λ_v are Lagrange multipliers determined from iteration to satisfy eq 12 and 16, respectively

Adaptive Langevin thermostat. We also carried out Langevin dynamics simulation to maintain constant temperature, where the coupling strength with the thermostat adapts to the QM profile,[23] an index for how much a solvent molecule behaves as a QM molecule. Using the velocity Verlet integrator for Langevin dynamics,[49] the coordinate r and velocity v are propagated as

$$
\begin{aligned}
v\left(t + \frac{\Delta t}{2}\right) &= \left(1 - \gamma\frac{\Delta t}{2}\right)v(t) + \frac{\Delta t}{2m}\{F(r(t+\Delta t)) + R(r(t+\Delta t))\} \\
r(t+\Delta t) &= r(t) + \Delta t\, v\left(t + \frac{\Delta t}{2}\right) \\
v(t+\Delta t) &= \frac{1}{1 + \gamma\frac{\Delta t}{2}}\left(v(t) + \frac{\Delta t}{2m}\{F(r(t)) + R(r(t))\}\right)
\end{aligned}
\qquad (18)
$$

where m and γ are mass and friction coefficient, respectively. F is a deterministic force derived from the potential function V, and R is the Gaussian random force defined as

$$R = \sqrt{\frac{2k_B T \gamma m}{\Delta t}} \zeta \qquad (19)$$

where ζ is a random number that satisfies $\langle \zeta \rangle = 0$ and $\langle \zeta^2 \rangle = 1$. In the limit $\gamma \to 0$, eq (18) reduces to the ordinary velocity Verlet algorithm for Hamiltonian dynamics.

In the SCMP method, the QM profile ω_i for the ith solvent molecule can be represented as

$$\omega_i = \sum_n \delta_i^{(n)} \sigma^{(n)} \qquad (20)$$

where $\delta_i^{(n)} = 1$ if the ith solvent molecule is QM in the nth QM/MM partitioning, while $\delta_i^{(n)} = 0$ if the solvent molecule is MM. As a result, $\omega_i = 1$ when the ith solvent molecule behaves as a pure QM model, which can be rephrased that the ith solvent molecule is defined as QM throughout all weighted partitioning. $\omega_i = 0$ when the ith solvent molecule behaves as a pure MM, which can be rephrased likewise.

In the present study, we associated the QM profile and the friction coefficient of the ith solvent molecule as

$$\gamma_i = (1 - \omega_i)\gamma_0 \qquad (21)$$

where γ_0 is a parameter. Since ω_i changes at every MD time step, correspondingly, γ_i for the respective solvent molecule also changed. This enabled MM solvent molecules to be fully linked to the thermostat and thus, the coupling strength gradually attenuated when the solvent molecule approached the QM center.

Computational details

The SCMP method was implemented in a local version of the GROMACS 5.0.7 package.[50–52] In all SCMP simulations, a total of 80 QM/MM partitioning were considered, where forces and

energy calculation were carried out based on different partitioning. The system consisted of one hydronium ion and 2047 water molecules in a periodic cubic box with a side length of 40.28 Å. Each partitioning has one QM solute hydronium ion and 32 QM solvent water molecules. In the SCMP, the transition parameters $s_{\mathrm{QM}}, t_{\mathrm{QM}}, s_{\mathrm{MM}}$, and t_{QM} were set to be 6.4, 8.4, 4.0 and 6.4 Å, respectively. In the partitioning updating protocol, we allowed updated partitioning to have a degree of order of 75 % for efficiency, as detailed in our previous work.[40] We employed the SPC-Fw water model[53] for the MM water. For the QM part, we employed DFTB3[15,16] implemented in GROMACS, as reported previously,[54] with use of standard 3OB[17] parameter sets. The electrostatic interactions in the MM–MM and QM–MM models were calculated with the particle-mesh Ewald method.[55] Both electrostatic interactions and van der Waals interactions were damped to zero in the range between 8.5 and 9.0 Å. The electrostatic potentials on the QM atoms induced by the charges of the MM atoms were obtained with the smooth particle-mesh Ewald method with a switching function for electrostatic interactions (electrostatic embedding). After equilibration for several picoseconds (ps), all MD simulations were conducted for 100 ps with time step of 0.25 and 0.50 fs for hydronium ion solution and bulk water simulations, respectively. For control of the EPI, we chose $r_0 = 1.3$ Å and $\alpha = 0.129$ Å for eq 5. In addition, we set $R_{\mathrm{small}} = 1.20$ Å and $R_{\mathrm{largre}} = 1.32$ Å for eq 9 and $\nu_{\mathrm{small}} = 2.0$ and $\nu_{\mathrm{large}} = 3.0$ for eq 10. For temperature control, we employed the friction coefficient $\gamma_0 = 100$ ps^{-1} and 10 ps^{-1} for eq 21.

RESULTS

For stable MD simulation and production of an accurate ensemble, the MD simulation needs to conserve the Hamiltonian (total energy) throughout its entire course. As shown in Figure 2a, we evaluated the Hamiltonian with different values of c in eq 12 under microcanonical (NVE) condition in the present simulation for 2.0 ps. The Hamiltonian was well conserved in all simulations over 1.0 ps from the beginning. Based on the fact that proton transfer between the hydronium ion and water molecules was observed several times during the simulation, it can be said that the proton transfer was simulated on a continuous energy surface. Notably, when $c = 0$, the total energy fluctuated around the average value. When $c = 1$, the fluctuation was well suppressed and the simulation seemed to be more stable. However, it violated the conservation of energy after 1.9 ps.

Next, we evaluated the bookkeeping term, which is the third term in eq 1. We found that the drift of the bookkeeping term became more distinct as the parameter c increased. Notably, when $c = 1$, it drifted by 1700 kJ/mol for 1 ps under microcanonical condition, which is about 200 times greater than the shift of the bookkeeping term for pure water or several monoatomic ions solutions. As we have reported in our previous study,[41] the drift of the bookkeeping term is related to spatial discontinuity. This discontinuity arises from the systematic difference of potential energies between QM/MM partitioning whose QM regions are compact and fragmented. Since the diffusion of hydronium ion is much faster than those of water molecules and other monoatomic ions, the deformation of the QM regions occurs sooner, leading to a large error accumulation. The drift of the bookkeeping term results in temperature increase under a microcanonical condition as shown in Figure 2c. If the increasing rate is drastic, the extra energy is locally accumulated around the QM region, destabilizing the MD simulation. Indeed, the present simulation with $c = 1$ was durable only for several picoseconds under a microcanonical

condition and thereafter violated the Hamiltonian conservation due to the extraordinarily fast displacement of the QM center.

As we proposed,[41] the drift of the bookkeeping term can be alleviated by introducing a correction potential U to cancel out the artificial diffusive force as shown in Fig 1b. Here, we employed $U = 1.0$ kJ/mol for both QM and MM solvent oxygen atoms over the distance range defined by transition parameters. Since the correction potential is ad hoc, it requires additional computational cost to find proper condition. As an alternative option to both stabilize the simulation and reproduce plausible proton transfer dynamics, we employed the Langevin thermostat in an adaptive manner so that the coupling strength with solvent molecules gradually changes in accordance with the changes of molecular definition based on its distance from the QM center. To this end, we introduced the QM profile ω_i, as defined in eq 20, which indicates how much a solvent molecule behaves as a QM molecule. In the SCMP method, the QM profile averaged over the MD simulation had smoothly shifted from 1 to 0 as the distance from the QM center increased, indicating that the molecular properties of solvent water had gradually alternated between QM and MM.[23,40] Figure 3 shows the temperature over the course of SCMP simulation with $c = 1$ employing an adaptive Langevin thermostat with the friction coefficient $\gamma_0 = 100$ ps^{-1}. As it can be observed, the adaptive thermostat controls the temperature well, maintaining it at the reference temperature of 300 K, enabling the MD simulation to be durable over hundreds of picoseconds. When $c = 0$ and 0.2, the drifts of the bookkeeping term are more moderate. Thus, a value as small as 10 ps^{-1} is sufficient to stabilize the simulation (See Supporting Information)

It should also be noted that the spatial discontinuity had distorted the dynamics of the solvent molecules located near the QM/MM border rather than the QM center.[41] The thermostat can also alter the dynamics, but under the present adaptive usage, the solute and solvent molecules in the vicinity of the QM center are free from the Langevin thermostat. Therefore, it is plausible to evaluate the dynamical properties from the obtained trajectories of the present simulation as discussed below. As it will be discussed later, the friction coefficient γ_0 does not affect the results of the hydronium ion simulations. Thus, most analyses in the present study are based on the simulation with $c = 1.0$ and $\gamma_0 = 100$ ps^{-1}, unless otherwise stated.

Radial distribution function Figure 4 shows the radial distribution function (RDF) around the oxygen atom O* nearest to the QM center, which should be a part of the hydronium ion. When compared to bulk water simulation, the first peak of O*-O RDF around the hydronium ion shifted from 2.8 Å to 2.6 Å, while the first peak in experiments had been observed at 2.5 Å.[56] DFTB2 simulations show bimodal peaks at around 2.4 Å and 2.8 Å, indicating that a Zundel-type structure was the dominant component.[57] On the other hand, the present O*-O RDF showed a distinct single peak, consistent with DFTB3 simulations.[58] According to the previous AIMD simulation,[2] such result implies that an Eigen-type structure is the major component. The height of the first peak of oxygen atoms of bulk water was 4.1, which was significantly larger than the experimental value of 2.5.[59] The first peak of O*-O RDF was fairly broader than the empirical potential structural refinement (EPSR) of the experimental data.[56] As a result, the coordination number was estimated to be as large as 4.3 by integrating the first solvation shell. As DFTB3, which was employed for the QM region, causes oversolvation and high density for bulk water, such properties seem to be carried over to the hydronium ion simulation. Regarding O*-H RDF, the first peak reflects hydrogen atoms covalently bonded to the hydronium oxygen. While the

bulk water simulation showed its second peak of O*-H RDF at 1.8 Å, it was not seen in the hydronium ion simulations, indicating that the hydronium ion did not accept hydrogen bonds. Note that the previous AIMD simulation showed O*-H RDF of Eigen-type cation had an additional peak at around 1.6 Å[7]. Although an Eigen-type cation was more probable in the present simulation as shown in the next section, we could not find any additional peak in that distance range.

Potential of mean force Figure 5 shows the potential of mean force for proton transfer in comparison with the potential energy in gaseous phase, which are projected on two reaction coordinates, the distance between two oxygen atoms $R_{O^*O'}$ and hydrogen displacement δ. Here, O^* denotes the oxygen atom nearest to the QM center, which is supposed to belong to the hydronium ion. The third nearest hydrogen atom to O^* was selected as H^*, and the oxygen atom nearest to H^* other than O^* was defined as O'. The hydrogen displacement was defined as the difference between the two distances as $\delta = R_{O^*H^*} - R_{O'H^*}$. Note that the EPI parameter c and the friction coefficient γ_0 did not affect the potential mean force (See Supporting Information)

  While the Zundel cation was more stable than an Eigen one in *vacuo*, the balance was inverted in the aqueous phase, which agrees with previous studies.[8,57] Figure 5 gives a clear picture of the proton transfer mechanism in Eigen to Zundel to Eigen sequence, where the energy barrier for the proton transfer disappears as $R_{O^*O'}$ becomes smaller than 2.43 Å. The present SCMP simulation estimated the energy barrier for proton transfer to be about 2.7 kJ/mol, which was in good agreement with DFTB3/3OB simulation.[58] Although DFTB3 leads to the overbinding of OH covalent bonds and underestimation of hydrogen bonds as previously reported,[17,41] the

estimated barrier showed agreement with previous studies of AIMD with DFT using BLYP[7] and HCTH functionals,[60] but was significantly smaller than those of MS-EVB2 and 3 which are estimated to be 8.4 kJ/mol.[61] It is worth noting that the nuclear quantum effect lowered the barrier of proton transfer.[7,9,62,63] Moreover, the energy barrier of around 1 $K_b T$ at room temperature is supposed to disappear by incorporation of nuclear quantum effect, making the topological defect delocalized and transfer to happen at a rate faster than 100 fs.[7,9] In contrast, the present profile showed distinct free energy minimum corresponding to an Eigen structure seen during the resting state of proton transfer. Figure 6 shows the time evolution of proton transfer projected on two reaction coordinates. Here, the proton transfer event was centered around a moment with $\delta = 0$ and the trajectories were averaged over more than 3000 proton transfer events. Note that the proton transfer still appeared as an event within 100 fs despite the free energy barrier. Also note that concerted oscillations were observed along the two reaction coordinates with a period of around 15 fs, which reflected a negative correlation between $R_{O^*O'}$ and δ. Within the framework of classical dynamics simulation, it can be concluded that the Eigen-Zundel interconversion was not the rate-limiting step of the proton transfer, which agrees with previous AIMD studies.[9,64]

Dynamical properties In general, a thermostat can affect dynamical properties. Hence, we employed the adaptive Langevin thermostat. To verify the thermostat influence, we benchmarked the orientational relaxation time for a system of QM water in MM water. Table 1 shows the resulting values obtained by explicit integration of the second rank auto-correlation function of OH bond orientation in water molecules. Note that the relaxation time does not vary according to the friction coefficient of the Langevin thermostat, indicating that the solute

dynamics are retained with adaptive treatment. However, the relaxation time is significantly underestimated due to DFTB3, which leads to fast water diffusion.

Next, we evaluated the mean square displacement (MSD) of the hydronium ion over twelve and eight independent 100ps-trajectories for H_3O^+, where the EPI of the present study is used to estimate the position of the hydronium ion (See Supporting Information). Upon MSD, the diffusion coefficient was evaluated to be 0.61 ± 0.05 $Å^2$/ps, which notably had smaller statistical error when compared to conventional studies.[57] Since the present approach provides a stable and plausible definition for the velocity of the EPI in eq 16, we were also able to evaluate the diffusion coefficient by integration of the velocity autocorrelation function of the EPI, and obtained the value of 0.63 ± 0.09 $Å^2$/ps (see Table 1). Moreover, the agreement of the two analyses indicates an achievement of sufficient sampling. In addition, the obtained values agreed with that in the previous study (0.66 ± 0.20 $Å^2$/ps) based on QM-MD simulation using DFTB3-diag,[57] although it is not directly comparable because the previous study employs a modified parameter for OH repulsive potential.

It is notable that the resulting diffusion coefficients of hydronium ion do not vary according to the friction coefficient. However, the water diffusion coefficient was only in good agreement with the reference DFTB3/3OB simulation when friction coefficient $\gamma_0 =1$ and 10 ps^{-1}, and was distinctively underestimated when $\gamma_0 = 100$ ps^{-1}. These results seem to be inconsistent with orientational relaxation time in Table 1, because they should be correlated. Therefore, we assumed that two factors can affect the solute water dynamics--- the first being formation/deformation of hydrogen bonds of solute water and the other is environmental dynamics that entirely translocates the QM region. The adaptive Langevin thermostat will

affect the latter factor with use of a large γ_0. On the other hand, proton transfer, that is solute dynamics, is faster than environmental water diffusion, suggesting that the diffusion coefficient of hydronium ion is insensitive to the Langevin thermostat.

On the other hand, the present value is smaller by 30% than that by DC-DFTB3-diag. The disagreement between DFTB3-diag and DC-DFTB3-diag was attributed to limited size effect. However, the artifact usually led to underestimation of the diffusion coefficient.[61,65] DFTB3-diag for 128 water system resulted in H_2O diffusion coefficient 0.38 Å^2/ps which was larger than the diffusion coefficient of 0.19 Å^2/ps for DC-DFTB3-diag with 513 water molecules. Taking into account of the agreement in the diffusion coefficient of bulk water between QM-DFTB3/3OB and SCMP-DFTB3/3OB, the derivation may have arisen from either/both the protocol of diffusion coefficient calculation and/or DC treatment. Note that, in the previous study, the hydronium diffusion coefficient was indirectly evaluated by summation of vehicular and Grotthuss diffusion coefficients, where the former corresponds to water diffusion coefficient and the latter was estimated with the proton transfer pitch and rate. In addition, we assumed that there may be a DC shortcoming, such as discontinuity caused by particles crossing the subsystems.

The present diffusion coefficient of the hydronium ion, which was 0.61-0.63 Å^2/ps, was smaller than the experimental value of 0.94 ± 0.01 Å^2/ps,[66] but larger than the values of 0.40 and 0.29 Å^2/ps by MS-EVB2,[43,61] 0.36 Å^2/ps by classical MS-EVB3,[56] 0.50 Å^2/ps by quantum MS-EVB3,[56] and 0.33 Å^2/ps by CPMD with HCTH functional[60] It is notable that most of the simulated diffusion coefficients (including the coefficient in this study) is smaller than experimental values. Unlike previous simulations, limited size effect is less likely in the present simulation because of sufficient system size employed by taking advantage of the QM/MM method. Therefore, the

underestimation is thought to be mainly attributed to either/both shortcoming of DFTB3 and/or missing nuclear quantum effect. As mentioned above, although DFTB3 may have overestimated the proton transfer energy barrier, the Eigen-Zundel interconversion still proceeds within 100 fs, which implies this is less likely to be the rate-limiting step. Previous simulations proposed that the rate-limiting step is the change in the coordination number of the water molecules in the solvation shell.[2,9,64] In addition, water reorientation related to molecular rotation occurred concertedly with formation/deformation of hydrogen bonds.[67] If it is the case, DFTB3 should lead to overestimation of diffusion coefficients. It is known that the underestimation of hydrogen bond energy by DFTB3 results in faster relaxation of the bond orientation of water than that in experiments.[58] Consistently, the diffusion coefficient of DFTB3 water had been significantly overestimated.[57,68] Regardless of the fast diffusion dynamics of DFTB3, the diffusion coefficient of hydronium ion is largely underestimated, which is presumably attributed to the missing nuclear quantum effect. Although nuclear quantum effect can indirectly have competing effects on water diffusion with respect to surrounding water molecules,[69] the effect presumably works to promote proton transfer in total regarding proton transfer.

DISCUSSIONS

To achieve a solute-adaptive QM/MM, it is essential to numerically express the position of the hydronium. The indicator should be capable to distinguish solvation structure specific to the hydronium ion from ordinary bulk water. Intuitively, the number of covalent bonds seems to be such a structural measure, but it cannot be directly used as an EPI because it is discretized. On the other hand, ψ found in eq 8 can be regarded as a continuous expression of the number of covalent bonds of an oxygen atom. To visualize, Figure 7 shows the distribution of ψ_i value of the ith oxygen atoms in eq 8 sampled in course of MD simulations, where the index of the oxygen atom represents its positional order from the QM center. ψ_1 of the nearest neighboring oxygen to the QM center, which is supposed to belong to the hydronium ion as aforementioned, almost took the value of 3.0. Sometimes, it was observed to be below 3.0. In contrast, ψ_3 equaled to 2.0 most of the time and rarely became any larger. The second nearest oxygen to the QM center, which forms the nearest neighboring water to the hydronium ion, showed bimodal distribution in respect to ψ_2. The first majority was at $\psi_2 = 2.0$ which occupied 85 %, while the second majority appeared at 3.0, which amounted to 1.8% (See Supporting Information). The ψ_2 value of 2.0 corresponds to the resting state of proton transfer, in which the hydronium ion takes the Eigen form. During the proton transfer from the hydronium ion to the nearest water molecule, the value of ν_2 gradually increased, which was sometimes accompanied by a decrease in ψ_1. Eventually, the values of ψ_1 and ψ_2 cross over at some point, at which point the indices (according to the displacement of the QM center) also switch. The values of the "original" ψ_1 and ψ_2 continue to decrease and increase, respectively. It is obvious that oxygen atoms farther away from the QM center than the third nearest oxygen atom have a ψ_i value of only 2.0, implying that the EPI was insensitive to ordinary water molecules farther away from the

hydronium ion. This is the most important attribute as an EPI, because otherwise the indicator will momentarily experience large displacement when the ordinary water oxygen has non-zero value of ψ_i. In other words, the hydronium ion and water molecules to be counted for the EPI were localized within the half length of the box size, and thus water displacement across the periodic boundary condition does not cause any discontinuity in the indicator. This also helps drastically reduce the computational cost to evaluate proton indicator because the number of water molecules that need to be considered becomes significantly small.

While EPI in previous studies were a function of oxygen coordinates as defined in eq 3, the EPI in this present study was expressed as a linear combination of dividing points between water oxygen and virtual site coordinates $\boldsymbol{\eta}$ as defined in eq 12. Note that, when $r = r_0$, the Gaussian-type function ϕ' in eq 5 has the maximum value and r_0 was tuned to be larger than ordinary OH covalent bond length. As a result, as shown in eq 7, $\boldsymbol{\eta_i}$ of the ith oxygen atom mainly reflected the hydrogen position that had the longest covalent bond with the ith oxygen atom. If the oxygen coordinates were used in eq 12, that is $c = 0$, the EPI during the resting state of proton transfer (which constitutes most part of the MD simulation) would correspond to the oxygen in the hydronium ion. Hence, the EPI is displaced between two oxygen atoms at an anomalously rapid speed when the proton transfers. However, the EPI with $c > 0$ fluctuated even during the resting state of proton transfer, reflecting the stretch of OH covalent bonds. So, when the proton transferred, it was displaced at a moderate speed that likely reflected the speed of the hydrogen transfer.

In the present study, we achieved Hamiltonian dynamics based on conserved force in a solute-adaptive framework. As aforementioned, the Hamiltonian conservation is one of the most

important criteria that indicates the simulation is free from artifacts arising from the discontinuity on the potential energy surface. Therefore, the existence of Hamiltonian that drives the equation of motion is a critical feature of the MD simulation, regardless of whether stochastic dynamics is employed for its production run. For the conserved force, the weight $\sigma^{(n)}(\{r_{k\xi}\})$ for partitioning in eq 1 should be a function of distances between independent particle coordinates, which means the EPI must represent a coordinate of a particle. To this end, we introduced a dummy atom whose degrees of freedom were also independent variables of the Hamiltonian. Although the dummy atom did not directly interact with other atoms, it indirectly interacted through the RATTLE constraint. Since this artificial treatment does not affect the dynamics, the mass of the dummy atom was set to 1e-8 a.u., sufficiently smaller than other particles. We found that a dummy atom smaller than the one set for our present study did not make any significant difference in the dynamics. However, a dummy atom with a larger mass may slow down the motion of the EPI, and makes the RATTLE iterations unstable, which can lead to the collapse of the MD simulation.

Conclusions

  In this study, we proposed a numerical expression for the excess proton indicator for hydronium ion simulation in bulk water and implemented it in the SCMP code. As a result, we achieved a QM/MM simulation that was both solute- and solvent-adaptive, which we termed as "full adaptive QM/MM method", and successfully demonstrated its stability and efficiency. Based on this new framework, we were able to confirm the total energy conservation of the proton transfer simulation, that is temporal continuity, with Hamiltonian dynamics under microcanonical

condition. However, the bookkeeping term and the temperature constantly drifted due to spatial discontinuity, for which we next conducted a Langevin dynamics simulation by employing the thermostat in an adaptive manner to retain plausible dynamics of the proton transfer. We emphasize that the present computational approach is advantageous to evaluate dynamical properties of proton transfer in bulk water for the following three reasons.
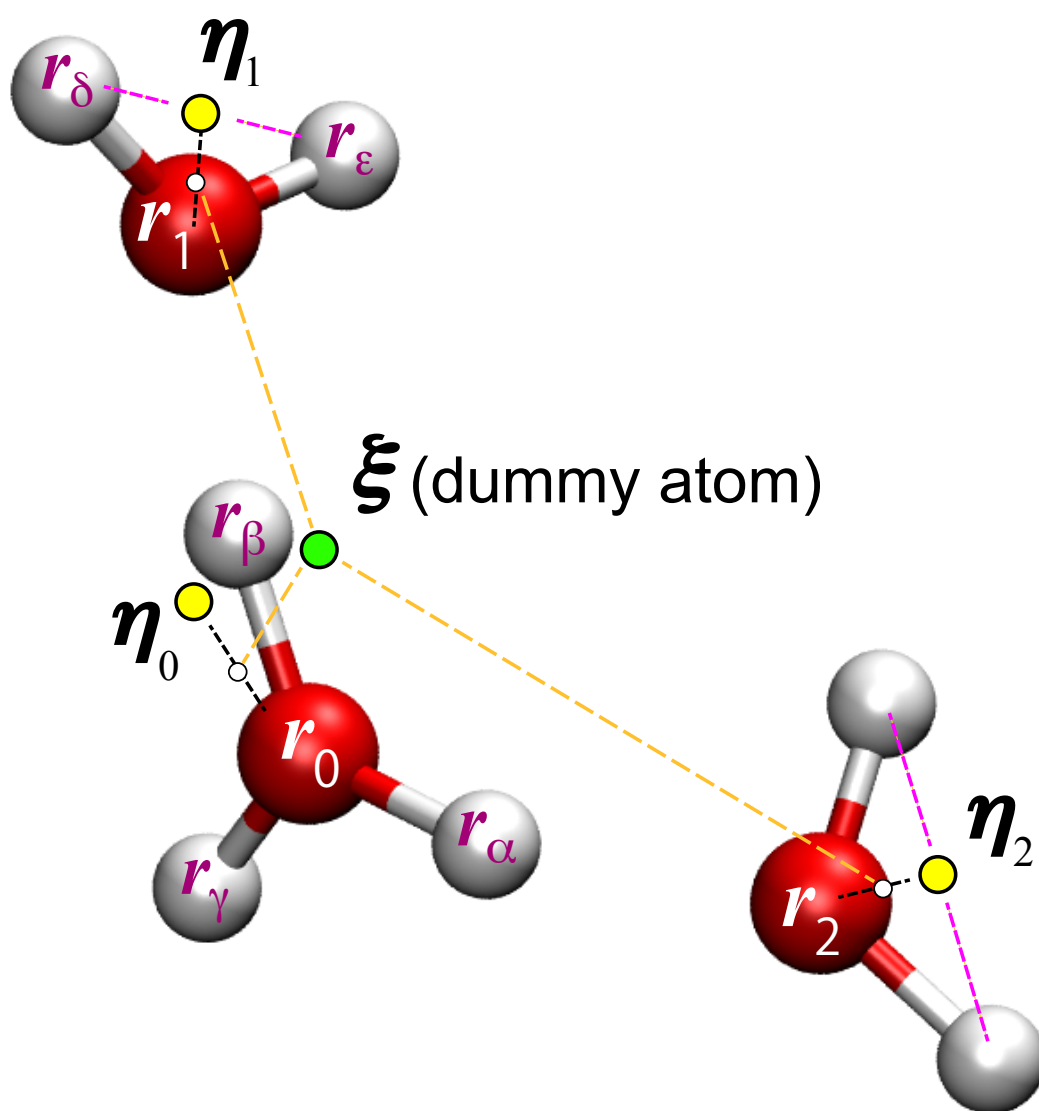
Firstly, the position of excess proton is defined with numerical stability and obtained on-the-fly through MD simulations. This provides direct access to various physical properties, such as diffusion coefficient through MSD or velocity autocorrelation function, that had been indirectly evaluated in previous studies. Moreover, the EPI may be used as a reaction coordinate for enhanced sampling, such as umbrella sampling, by imposing some artificial force on a dummy atom. Notably, the present EPI can be also used for post-MD analysis, as long as the coordinate information is retained. This may be useful to reassess previous AIMD simulations by removing artificial noise.

Secondly, the computational cost required for the present approach based on the QM/MM method is moderate compared to AIMD and thus makes it accessible to longer dynamics trajectory. As a result, statistical errors can be greatly suppressed compared to conventional studies using the AIMD method. The linear shape of the obtained MSD as well as the agreement between diffusion coefficients obtained by MSD and velocity autocorrelation function provide strong grounds that statistical errors were greatly suppressed. We assume that this feature will become more constructive for the analysis of inhomogeneous systems such as the one used in the present study rather than homogeneous systems such as pure water. Since the number of trajectories of interest obtained by a single MD run is remarkably limited, more production runs
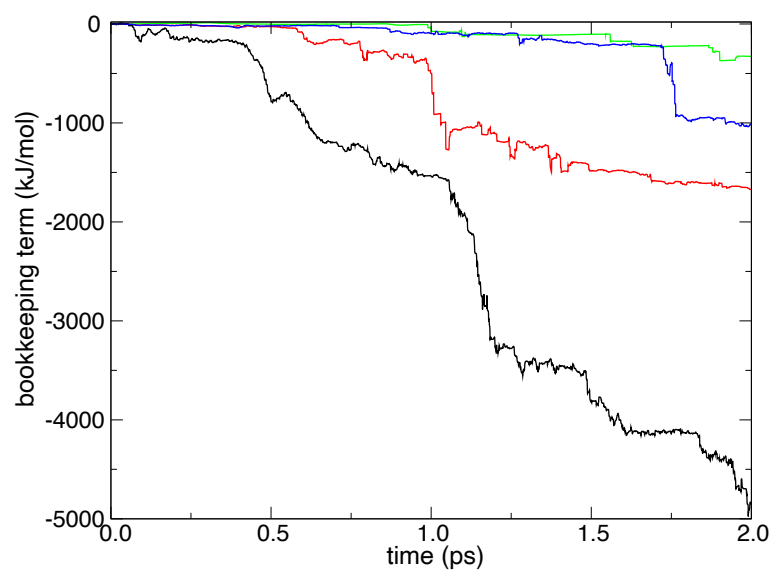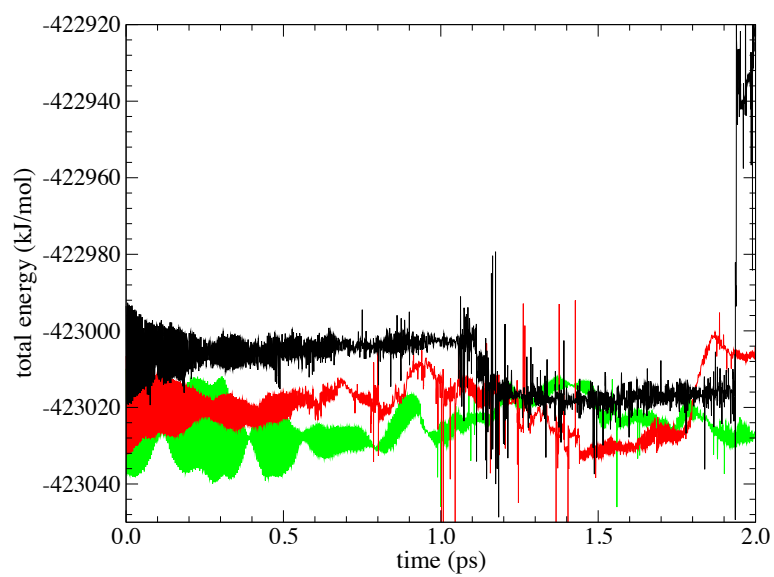
that are longer are required for higher accuracy of simulations for inhomogeneous systems. Hence, computational cost becomes a critical factor.
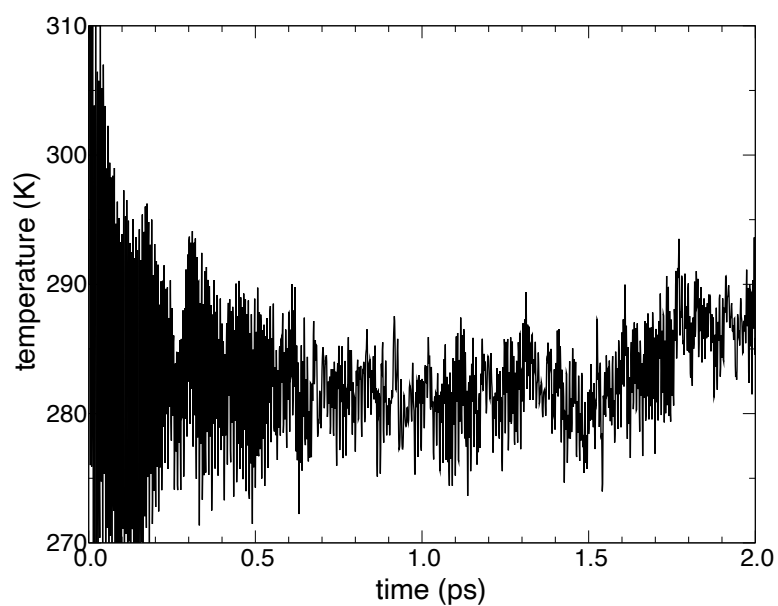
Thirdly, this full adaptive QM/MM method also provides access to gigantic systems that cannot be treated with a full QM method. In the present study, we considered one hydronium ion and 2047 water molecules in a cubic box with side lengths of 39.5 Å, which is beyond the range of application of ordinary AIMD simulations. It is known that diffusion coefficient is subject to limited size effect under periodic boundary condition, resulting in drastic underestimation in previous studies using AIMD simulations. On the other hand, the SCMP method can mitigate the artifacts as we have previously reported.[68] In the present study, therefore, a major part of the deviation of the obtained diffusion coefficient from the experimental value can be attributed to either/both shortcoming of the employed QM model (DFTB3/3OB) or missing nuclear quantum effect. Upon the advantages, the present full adaptive QM/MM method make the hydronium ion simulation accessible with plausible cost and moderate computation time, which had been a long-standing challenge in molecular simulation. By no doubt, this will become an effective tool to advance theoretical analysis for hydronium ion to the next stage.
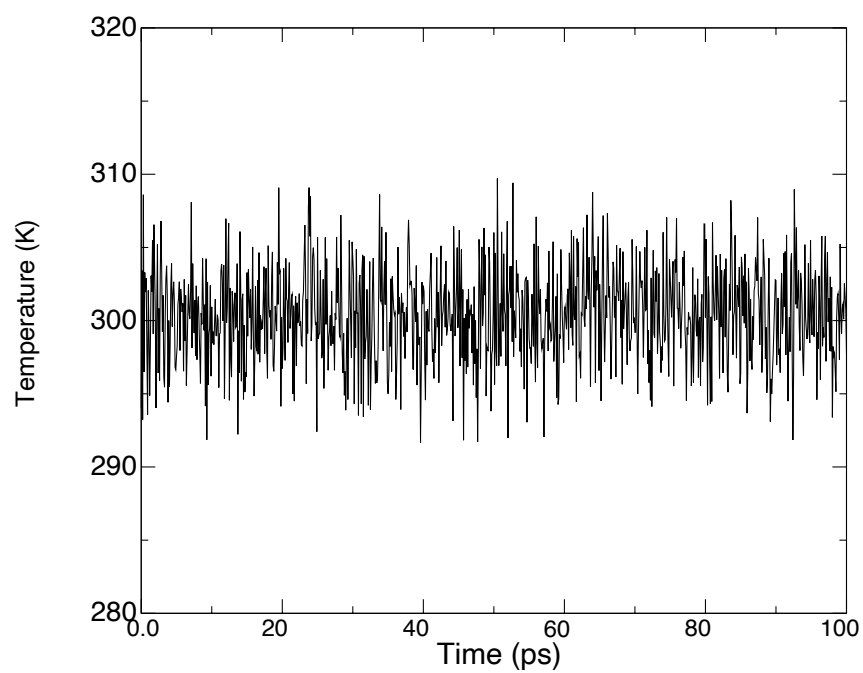
Figures

Figure 1. Schematic illustration for the excess proton indicator $\boldsymbol{\xi}$.
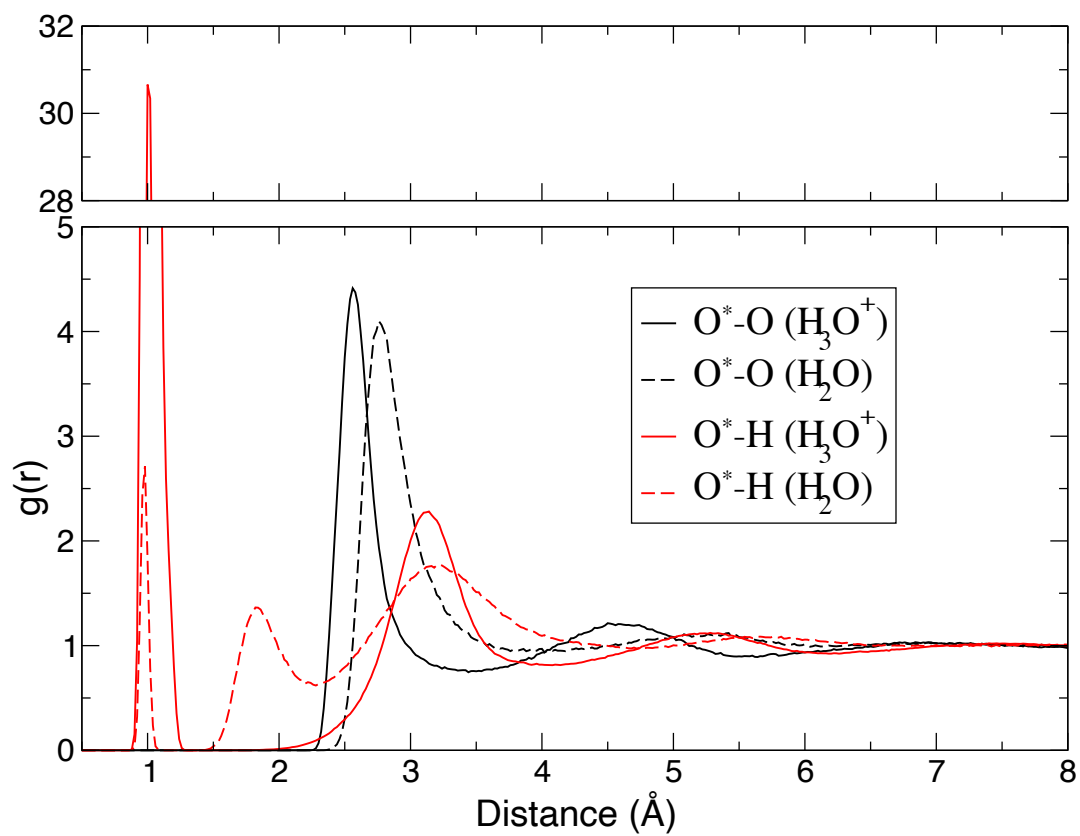
Figure 2. Hamiltonian (total energy) (top), bookkeeping term (middle), and temperature (bottom) in the course of MD simulation time under a microcanonical condition. Black, red, and green lines represent the SCMP simulation with the EPI parameter of $c = 1.0, 0.2$, and 0.0, respectively. The blue line represents the SCMP simulation with $c = 0.2$ and the correction potential $U = 1.0$ kJ/mol.
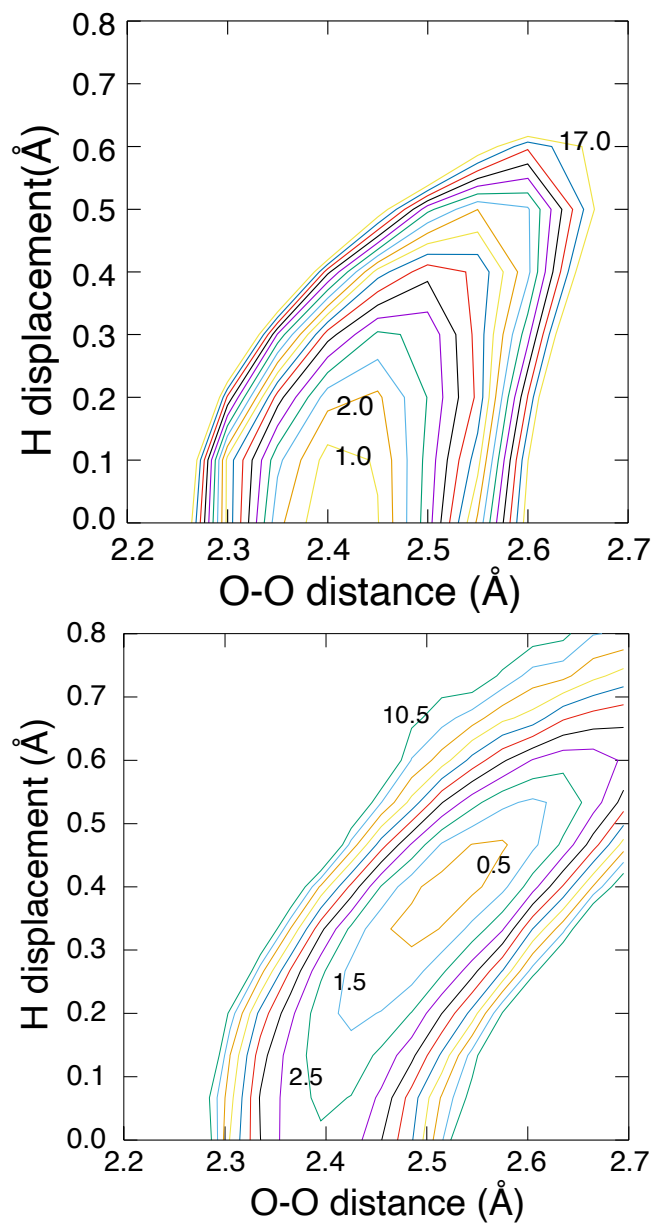
Figure 3. Temperature in the course of SCMP simulation with $c = 1.0$ using an adaptive

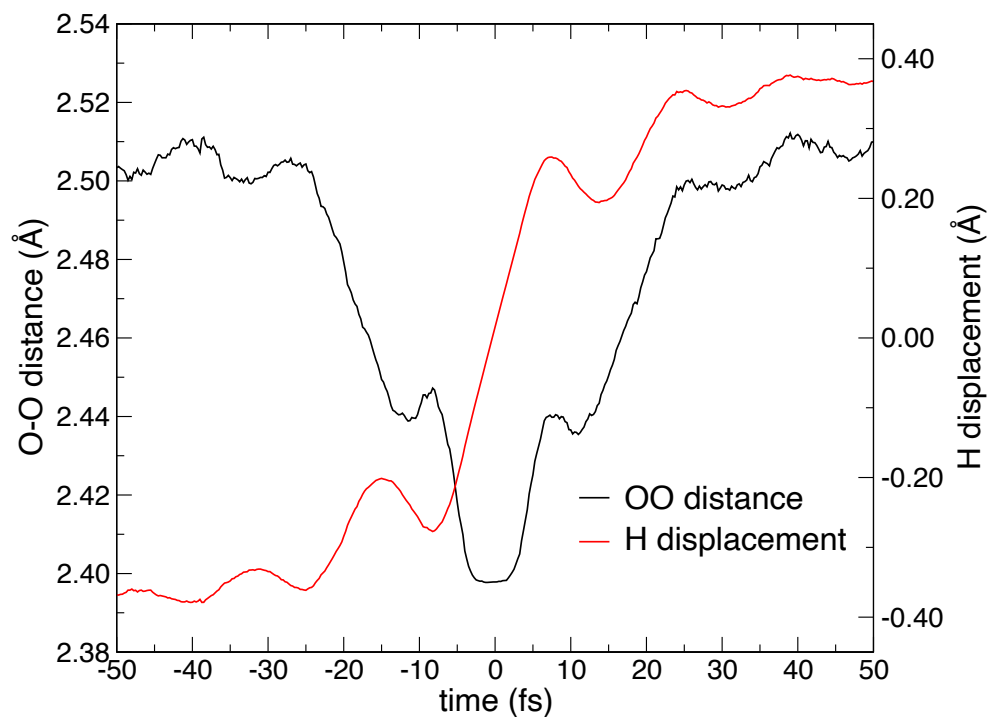Langevin thermostat with a friction coefficient $\gamma_0 = 100$ cm^{-1}

Figure 4. Radial distribution function of oxygen (black) and hydrogen (red) atoms around the oxygen nearest to the QM center, denoted as O*.
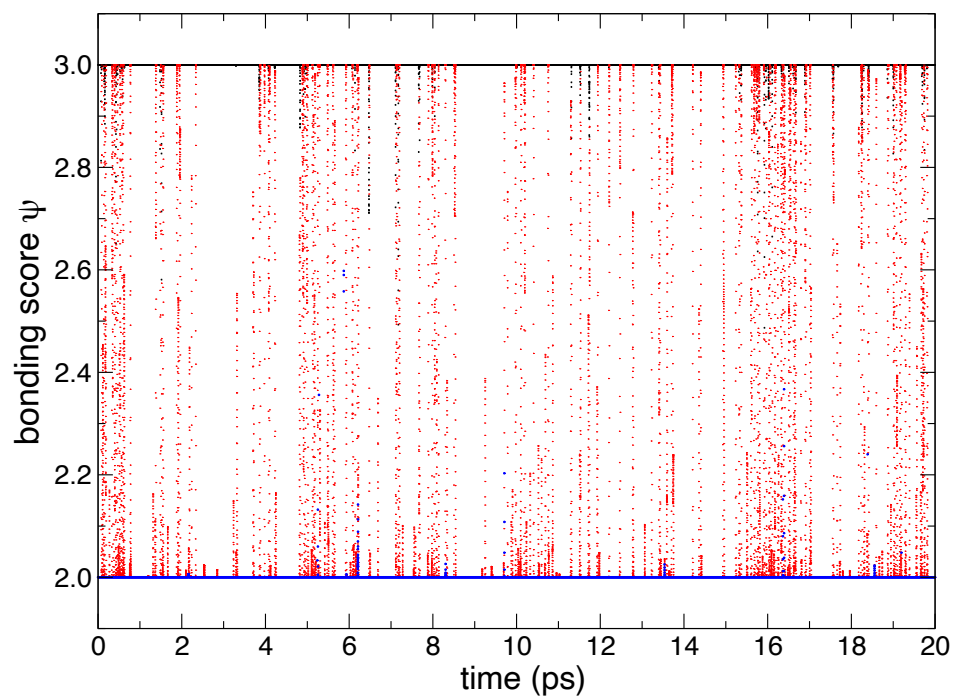
Figure 5. Potential energy in gas phase (top) and potential of mean force in aqueous phase (bottom) of proton transfer. The horizontal axis represents distance between the oxygen atom O* nearest to the QM center and the second nearest oxygen O'. The vertical axis represents the transferred hydrogen displacement $\delta = R_{O'H} - R_{O*H}$.

Figure 6. Time evolution centered on the H-bond switching event. Black line represents the distance between hydronium ion and the nearest water oxygen atoms. Red line represents the displacement of transferred hydrogen atom defined as $\delta = R_{O'H} - R_{O*H}$

Figure 7. Distribution of κ_i values of oxygen in course of the MD simulation. Black, red, and blue dots represent the first, second, and third nearest oxygen to the QM center, respectively.

Table 1. Orientational relaxation time from P2 correlation function (ps) with different friction coefficients

	SCMP+DFTB3/3OBd			DFTB3/3OB	Experiment
γ_0 (ps^{-1})	100	10	1		
	0.53 ± 0.05	0.50 ± 0.08	0.58 ± 0.07	0.7^a	$1.7{\sim}2.6^b$

[a]Goyal et al. (2011), [b]Winkler et al. (2000),[70] Lawrence et al. (2003),[71] Tan et al. (2005),[72] Rezus et al. (2005)[73]

Table 2. Diffusion coefficient ($Å^2$/ps)

γ_0 (ps^{-1})	Experiment	DFTB3 (diag[a])	DC-DFTB3 (diag[b])	SCMP+DFTB3/3OB[d]				
				100		10	1	NVE[c]
				MSD[d]	VAC[e]	MSD[d]	MSD[d]	MSD[d]
H_2O	0.23	0.38 ± 0.03	0.19	0.27 ± 0.03	0.32 ± 0.03	0.48 ± 0.07	0.45 ± 0.03	0.45 ± 0.06[c]
H_3O^+/H_2O	0.94 ± 0.01[f]	0.66 ± 0.20	0.91	0.61 ± 0.05	0.63 ± 0.09	0.62 ± 0.09	0.59 ± 0.15	

[a]MD simulation for 128 solvent water molecules by Goyal et al. (2011). [b]MD simulation for 523 water by Nakai et al. (2016). [c]microcanonical condition. MD simulation for 2,048 bulk water by Watanabe et al. (2017).[68] [d]By linear fitting of MSD. [e]By integration of velocity of autocorrelation function. [f]Weasted et al (1988).[66]

ASSOCIATED CONTENT

(Word Style "TE_Supporting_Information"). **Supporting Information**. A listing of the contents of each file supplied as Supporting Information should be included. For instructions on what should be included in the Supporting Information as well as how to prepare this material for publications, refer to the journal's Instructions for Authors.

The following files are available free of charge.

brief description (file type, i.e., PDF)

AUTHOR INFORMATION

Corresponding Author

*E-mail: hcwatanabe@keio.jp

Notes

Any additional relevant notes should be placed here.

ABBREVIATIONS

QM, quantum mechanics; MD, molecular dynamics; AIMD, *ab initio* molecular dynamics; DC,

divide-and-conquer; CPMD, Car-Parrinello molecular dynamics; QM/MM, quantum

mechanics/molecular mechanics; SCMP, size-consistent multipartitioning; EPI, excess proton

indicator; RDF, radial distribution function; MSD, mean square displacement.

REFERENCES

(1)   Komatsuzaki, T.; Ohmine, I. Energetics of Proton Transfer in Liquid Water. I. Ab Initio Study for Origin of Many-Body Interaction and Potential Energy Surfaces. *Chem. Phys.* **1994**, *180* (2–3), 239–269. https://doi.org/10.1016/0301-0104(93)E0424-T.

(2)   Tuckerman, M.; Laasonen, K.; Sprik, M.; Parrinello, M. Ab Initio Molecular Dynamics Simulation of the Solvation and Transport of H3O+ and OH- Ions in Water. *J. Phys. Chem.* **1995**, *99* (16), 5749–5752. https://doi.org/10.1021/j100016a003.

(3)   Tuckerman, M.; Laasonen, K.; Sprik, M.; Parrinello, M. Ab Initio Molecular Dynamics Simulation of the Solvation and Transport of Hydronium and Hydroxyl Ions in Water. *J. Chem. Phys.* **1995**, *103* (1), 150–161. https://doi.org/10.1063/1.469654.

(4)   Lobaugh, J.; Voth, G. A. The Quantum Dynamics of an Excess Proton in Water. *J. Chem. Phys.* **1996**, *104* (5), 2056–2069. https://doi.org/10.1063/1.470962.

(5)   Vuilleumier, R.; Borgis, D. Molecular Dynamics of an Excess Proton in Water Using a Non-Additive Valence Bond Force Field. *J. Mol. Struct.* **1997**, *436–437*, 555–565. https://doi.org/10.1016/S0022-2860(97)00242-1.

(6)   Schmitt, U. W.; Voth, G. A. Multistate Empirical Valence Bond Model for Proton Transport in Water. *J. Phys. Chem. B* **1998**, *102* (29). https://doi.org/10.1021/jp9818131.

(7)   Marx, D.; Tuckerman, M. E.; Parrinello, M. Solvated Excess Protons in Water: Quantum Effects on the Hydration Structure. *J. Phys. Condens. Matter* **2000**, *12* (8A). https://doi.org/10.1088/0953-8984/12/8A/317.

(8)   Tuckerman, M. E.; Marx, D.; Parrinello, M. The Nature and Transport Mechanism of Hydrated Hydroxide Ions in Aqueous Solution. *Nature* **2002**, *417* (6892), 925–929. https://doi.org/10.1038/nature00797.

(9)   Marx, D.; Chandra, A.; Tuckerman, M. E. Aqueous Basic Solutions: Hydroxide Solvation, Structural Diffusion, and Comparison to the Hydrated Proton. *Chem. Rev.* **2010**, *110* (4), 2174–2216. https://doi.org/10.1021/cr900233f.

(10)   Sakti, A. W.; Nishimura, Y.; Nakai, H. Divide-and-Conquer-Type Density-Functional Tight-Binding Simulations of Hydroxide Ion Diffusion in Bulk Water. *J. Phys. Chem. B* **2017**, *121* (6), 1362–1371. https://doi.org/10.1021/acs.jpcb.6b10659.

(11)   Tuckerman, M. E.; Chandra, A.; Marx, D. Structure and Dynamics of OH - (Aq). *Acc. Chem. Res.* **2006**, *39* (2), 151–158. https://doi.org/10.1021/ar040207n.

(12)   Car, R.; Parrinello, M. Unified Approach for Molecular Dynamics and Density-Functional Theory. *Phys. Rev. Lett.* **1985**, *55* (22), 2471–2474. https://doi.org/10.1103/PhysRevLett.55.2471.

(13)   Yang, W.; Lee, T. S. A Density-Matrix Divide-and-Conquer Approach for Electronic Structure Calculations of Large Molecules. *J. Chem. Phys.* **1995**, *103* (13), 5674–5678. https://doi.org/10.1063/1.470549.

(14) Kobayashi, M.; Kunisada, T.; Akama, T.; Sakura, D.; Nakai, H. Reconsidering an Analytical Gradient Expression within a Divide-and-Conquer Self-Consistent Field Approach: Exact Formula and Its Approximate Treatment. *J. Chem. Phys.* **2011**, *134* (3). https://doi.org/10.1063/1.3524337.

(15) Elstner, M.; Porezag, D.; Jungnickel, G.; Elsner, J.; Haugk, M.; Frauenheim, T.; Suhai, S.; Seifert, G. Self-Consistent-Charge Density-Functional Tight-Binding Method for Simulations of Complex Materials Properties. *Phys. Rev. B* **1998**, *58* (11), 7260–7268. https://doi.org/10.1103/PhysRevB.58.7260.

(16) Gaus, M.; Cui, Q.; Elstner, M. DFTB3: Extension of the Self-Consistent-Charge Density-Functional Tight-Binding Method (SCC-DFTB). *J. Chem. Theory Comput.* **2011**, *7* (4), 931–948. https://doi.org/10.1021/ct100684s.

(17) Gaus, M.; Goez, A.; Elstner, M. Parametrization and Benchmark of DFTB3 for Organic Molecules. *J. Chem. Theory Comput.* **2013**, *9* (1), 338–354. https://doi.org/10.1021/ct300849w.

(18) Warshel, A.; Levitt, M. Theoretical Studies of Enzymic Reactions: Dielectric, Electrostatic and Steric Stabilization of the Carbonium Ion in the Reaction of Lysozyme. *J. Mol. Biol.* **1976**, *103* (2), 227–249. https://doi.org/10.1016/0022-2836(76)90311-9.

(19) Kerdcharoen, T.; Liedl, K. R.; Rode, B. M. A QM/MM Simulation Method Applied to the Solution of Li+ in Liquid Ammonia. *Chem. Phys.* **1996**, *211* (1–3), 313–323. https://doi.org/10.1016/0301-0104(96)00152-8.

(20) Kerdcharoen, T.; Morokuma, K. ONIOM-XS: An Extension of the ONIOM Method for Molecular Simulation in Condensed Phase. *Chem. Phys. Lett.* **2002**, *355* (3–4), 257–262. https://doi.org/10.1016/S0009-2614(02)00210-5.

(21) Shiga, M.; Masia, M. Boundary Based on Exchange Symmetry Theory for Multilevel Simulations. I. Basic Theory. *J. Chem. Phys.* **2013**, *139* (4). https://doi.org/10.1063/1.4816629.

(22) Waller, M. P.; Kumbhar, S.; Yang, J. A Density-Based Adaptive Quantum Mechanical/Molecular Mechanical Method. *ChemPhysChem* **2014**, *15* (15), 3218–3225. https://doi.org/10.1002/cphc.201402105.

(23) Watanabe, H. C.; Kubař, T.; Elstner, M. Size-Consistent Multipartitioning QM/MM: A Stable and Efficient Adaptive QM/MM Method. *J. Chem. Theory Comput.* **2014**, *10* (10), 4242–4252. https://doi.org/10.1021/ct5005593.

(24) Shiga, M.; Masia, M. Quasi-Boundary Based on Exchange Symmetry Theory for Multilevel Simulations. *Mol. Simul.* **2015**, *41* (10–12), 827–831. https://doi.org/10.1080/08927022.2014.938068.

(25) Boereboom, J. M.; Potestio, R.; Donadio, D.; Bulo, R. E. Toward Hamiltonian Adaptive QM/MM: Accurate Solvent Structures Using Many-Body Potentials. *J. Chem. Theory Comput.* **2016**, *12* (8), 3441–3448. https://doi.org/10.1021/acs.jctc.6b00205.

(26) Duster, A. W.; Wang, C. H.; Lin, H. Adaptive QM/MM for Molecular Dynamics Simulations: 5. On the Energy-Conserved Permuted Adaptive-Partitioning Schemes. *Molecules* **2018**, *23* (9). https://doi.org/10.3390/molecules23092170.

(27) Field, M. J. An Algorithm for Adaptive QC/MM Simulations. *J. Chem. Theory Comput.* **2017**, *13* (5), 2342–2351. https://doi.org/10.1021/acs.jctc.7b00099.

(28) Jiang, T.; Simko, S.; Bulo, R. E. Accurate Quantum Mechanics/Molecular Mechanics Simulation of Aqueous Solutions with Tailored Molecular Mechanics Models. *J. Chem. Theory Comput.* **2018**, *14* (8), 3943–3954. https://doi.org/10.1021/acs.jctc.7b01218.

(29) Takahashi, H.; Kambe, H.; Morita, A. Calculation of Solvation Free Energy Utilizing a Constrained QM/MM Approach Combined with a Theory of Solutions. *J. Chem. Phys.* **2019**, *150* (11). https://doi.org/10.1063/1.5089199.

(30) Zeng-hui Yang. On-the-Fly Determination of Active Region Centers in Adaptive-Partitioning QM/MM. *ChemRxiv* **2020**. https://doi.org/10.26434/chemrxiv.12376919.v1.

(31) Heyden, A.; Lin, H.; Truhlar, D. G. Adaptive Partitioning in Combined Quantum Mechanical and Molecular Mechanical Calculations of Potential Energy Functions for Multiscale Simulations. *J. Phys. Chem. B* **2007**, *111* (9), 2231–2241. https://doi.org/10.1021/jp0673617.

(32) Lin, H.; Truhlar, D. G. QM/MM: What Have We Learned, Where Are We, and Where Do We Go from Here? *Theor. Chem. Acc.* **2007**, *117* (2), 185–199. https://doi.org/10.1007/s00214-006-0143-z.

(33) Bulo, R. E.; Ensing, B.; Sikkema, J.; Visscher, L. Toward a Practical Method for Adaptive QM/MM Simulations. *J. Chem. Theory Comput.* **2009**, *5* (9), 2212–2221. https://doi.org/10.1021/ct900148e.

(34) Pezeshki, S.; Lin, H. Adaptive-Partitioning Redistributed Charge and Dipole Schemes for QM/MM Dynamics Simulations: On-the-Fly Relocation of Boundaries That Pass through Covalent Bonds. *J. Chem. Theory Comput.* **2011**, *7* (11), 3625–3634. https://doi.org/10.1021/ct2005209.

(35) Bernstein, N.; Várnai, C.; Solt, I.; Winfield, S. A.; Payne, M. C.; Simon, I.; Fuxreiter, M.; Csányi, G. QM/MM Simulation of Liquid Water with an Adaptive Quantum Region. *Phys. Chem. Chem. Phys.* **2012**, *14* (2), 646–656. https://doi.org/10.1039/c1cp22600b.

(36) Rowley, C. N.; Roux, B. The Solvation Structure of Na + and K + in Liquid Water Determined from High Level Ab Initio Molecular Dynamics Simulations. *J. Chem. Theory Comput.* **2012**, *8* (10), 3526–3535. https://doi.org/10.1021/ct300091w.

(37) Takenaka, N.; Kitamura, Y.; Koyano, Y.; Nagaoka, M. An Improvement in Quantum Mechanical Description of Solute-Solvent Interactions in Condensed Systems via the Number-Adaptive Multiscale Quantum Mechanicalmolecular Mechanical-Molecular Dynamics Method: Application to Zwitterionic Glycine in Aqueous Soluti. *J. Chem. Phys.* **2012**, *137* (2). https://doi.org/10.1063/1.4732307.

(38)   Bulo, R. E.; Michel, C.; Fleurat-Lessard, P.; Sautet, P. Multiscale Modeling of Chemistry in Water: Are We There Yet? *J. Chem. Theory Comput.* **2013**, *9* (12), 5567–5577. https://doi.org/10.1021/ct4005596.

(39)   Ensing, B.; Nielsen, S. O.; Moore, P. B.; Klein, M. L.; Parrinello, M. Energy Conservation in Adaptive Hybrid Atomistic/Coarse-Grain Molecular Dynamics. *J. Chem. Theory Comput.* **2007**, *3* (3), 1100–1105. https://doi.org/10.1021/ct600323n.

(40)   Watanabe, H. C. Improvement of Performance, Stability and Continuity by Modified Size-Consistent Multipartitioning Quantum Mechanical/Molecular Mechanical Method. *Molecules* **2018**, *23* (8). https://doi.org/10.3390/molecules23081882.

(41)   Watanabe, H. C.; Cui, Q. Quantitative Analysis of QM/MM Boundary Artifacts and Correction in Adaptive QM/MM Simulations. *J. Chem. Theory Comput.* **2019**, *15* (7), 3917–3928. https://doi.org/10.1021/acs.jctc.9b00180.

(42)   Pomès, R.; Roux, B. Free Energy Profiles for H+ Conduction along Hydrogen-Bonded Chains of Water Molecules. *Biophys. J.* **1998**, *75* (1), 33–40. https://doi.org/10.1016/S0006-3495(98)77492-2.

(43)   Day, T. J. F.; Soudackov, A. V.; Čuma, M.; Schmitt, U. W.; Voth, G. A. A Second Generation Multistate Empirical Valence Bond Model for Proton Transport in Aqueous Systems. *J. Chem. Phys.* **2002**, *117* (12), 5839–5849. https://doi.org/10.1063/1.1497157.

(44)   Chakrabarti, N.; Tajkhorshid, E.; Roux, B.; Pomès, R. Molecular Basis of Proton Blockage in Aquaporins. *Structure* **2004**, *12* (1), 65–74. https://doi.org/10.1016/j.str.2003.11.017.

(45)   König, P. H.; Ghosh, N.; Hoffmann, M.; Elstner, M.; Tajkhorshid, E.; Frauenheim, T.; Cui, Q. Toward Theoretical Analyis of Long-Range Proton Transfer Kinetics in Biomolecular Pumps. *J. Phys. Chem. A* **2006**, *110* (2), 548–563. https://doi.org/10.1021/jp052328q.

(46)   Duster, A. W.; Lin, H. Tracking Proton Transfer through Titratable Amino Acid Side Chains in Adaptive QM/MM Simulations. *J. Chem. Theory Comput.* **2019**, *15* (11), 5794–5809. https://doi.org/10.1021/acs.jctc.9b00649.

(47)   Andersen, H. C. Rattle: A "Velocity" Version of the Shake Algorithm for Molecular Dynamics Calculations. *J. Comput. Phys.* **1983**, *52* (1), 24–34. https://doi.org/10.1016/0021-9991(83)90014-1.

(48)   Watanabe, H. C.; Banno, M.; Sakurai, M. An Adaptive Quantum Mechanics/Molecular Mechanics Method for the Infrared Spectrum of Water: Incorporation of the Quantum Effect between Solute and Solvent. *Phys. Chem. Chem. Phys.* **2016**, *18* (10), 7318–7333. https://doi.org/10.1039/c5cp07136d.

(49)   Brünger, A.; Brooks, C. L.; Karplus, M. Stochastic Boundary Conditions for Molecular Dynamics Simulations of ST2 Water. *Chem. Phys. Lett.* **1984**, *105* (5), 495–500. https://doi.org/10.1016/0009-2614(84)80098-6.

(50) Bjelkmar, P.; Larsson, P.; Cuendet, M. A.; Hess, B.; Lindahl, E. Implementation of the CHARMM Force Field in GROMACS: Analysis of Protein Stability Effects from Correction Maps, Virtual Interaction Sites, and Water Models. *J. Chem. Theory Comput.* **2010**, *6* (2), 459–466. https://doi.org/10.1021/ct900549r.

(51) Abraham, M. J.; Murtola, T.; Schulz, R.; Páll, S.; Smith, J. C.; Hess, B.; Lindahl, E. GROMACS: High Performance Molecular Simulations through Multi-Level Parallelism from Laptops to Supercomputers. *SoftwareX* **2015**, *1–2*, 19–25. https://doi.org/10.1016/j.softx.2015.06.001.

(52) Hess, B.; Kutzner, C.; van der Spoel, D.; Lindahl, E. GROMACS 4: Algorithms for Highly Efficient, Load-Balanced, and Scalable Molecular Simulation. *J. Chem. Theory Comput.* **2008**, *4* (3), 435–447. https://doi.org/10.1021/ct700301q.

(53) Wu, Y.; Tepper, H. L.; Voth, G. A. Flexible Simple Point-Charge Water Model with Improved Liquid-State Properties. *J. Chem. Phys.* **2006**, *124* (2). https://doi.org/10.1063/1.2136877.

(54) Kubař, T.; Welke, K.; Groenhof, G. New QM/MM Implementation of the DFTB3 Method in the Gromacs Package. *J. Comput. Chem.* **2015**, *36* (26), 1978–1989. https://doi.org/10.1002/jcc.24029.

(55) Darden, T.; York, D.; Pedersen, L. Particle Mesh Ewald: An $N \cdot \log(N)$ Method for Ewald Sums in Large Systems. *J. Chem. Phys.* **1993**, *98* (12), 10089–10092. https://doi.org/10.1063/1.464397.

(56) Botti, A.; Bruni, F.; Ricci, M. A.; Soper, A. K. Eigen versus Zundel Complexes in HCl-Water Mixtures. *J. Chem. Phys.* **2006**, *125* (1). https://doi.org/10.1063/1.2212421.

(57) Goyal, P.; Elstner, M.; Cui, Q. Application of the SCC-DFTB Method to Neutral and Protonated Water Clusters and Bulk Water. *J. Phys. Chem. B* **2011**, *115* (20), 6790–6805. https://doi.org/10.1021/jp202259c.

(58) Goyal, P.; Qian, H. J.; Irle, S.; Lu, X.; Roston, D.; Mori, T.; Elstner, M.; Cui, Q. Molecular Simulation of Water and Hydration Effects in Different Environments: Challenges and Developments for DFTB Based Models. *J. Phys. Chem. B* **2014**, *118* (38), 11007–11027. https://doi.org/10.1021/jp503372v.

(59) Soper, A. K. The Radial Distribution Functions of Water as Derived from Radiation Total Scattering Experiments: Is There Anything We Can Say for Sure? *ISRN Phys. Chem.* **2013**, *2013*, 1–67. https://doi.org/10.1155/2013/279463.

(60) Maupin, C. M.; Aradi, B.; Voth, G. A. The Self-Consistent Charge Density Functional Tight Binding Method Applied to Liquid Water and the Hydrated Excess Proton: Benchmark Simulations. *J. Phys. Chem. B* **2010**, *114* (20), 6922–6931. https://doi.org/10.1021/jp1010555.

(61)   Wu, Y.; Chen, H.; Wang, F.; Paesani, F.; Voth, G. A. An Improved Multistate Empirical Valence Bond Model for Aqueous Proton Solvation and Transport. *J. Phys. Chem. B* **2008**, *112* (2), 467–482. https://doi.org/10.1021/jp076658h.

(62)   Azzouz, H.; Borgis, D. A Quantum Molecular-Dynamics Study of Proton-Transfer Reactions along Asymmetrical H Bonds in Solution. *J. Chem. Phys.* **1993**, *98* (9), 7361–7374. https://doi.org/10.1063/1.464727.

(63)   Kosugi, K.; Nakano, H.; Sato, H. SCC-DFTB-PIMD Method to Evaluate a Multidimensional Quantum Free-Energy Surface for a Proton-Transfer Reaction. *J. Chem. Theory Comput.* **2019**, *15* (9), 4965–4973. https://doi.org/10.1021/acs.jctc.9b00355.

(64)   Andot, K.; Hynes, J. T. HCl Acid Ionization in Water: A Theoretical Molecular Modeling. *J. Mol. Liq.* **1995**, *64* (1–2), 25–37. https://doi.org/10.1016/0167-7322(95)92818-V.

(65)   Yeh, I. C.; Hummer, G. System-Size Dependence of Diffusion Coefficients and Viscosities from Molecular Dynamics Simulations with Periodic Boundary Conditions. *J. Phys. Chem. B* **2004**, *108* (40), 15873–15879. https://doi.org/10.1021/jp0477147.

(66)   Roberts, N. K.; Northey, H. L. Proton and Deuteron Mobility in Normal and Heavy Water Solutions of Electrolytes. *J. Chem. Soc. Faraday Trans. 1 Phys. Chem. Condens. Phases* **1974**, *70*, 253–262. https://doi.org/10.1039/F19747000253.

(67)   Laage, D.; Hynes, J. T. A Molecular Jump Mechanism of Water Reorientation. *Science (80-.).* **2006**, *311* (5762), 832–835. https://doi.org/10.1126/science.1122154.

(68)   Watanabe, H. C.; Kubillus, M.; Kubař, T.; Stach, R.; Mizaikoff, B.; Ishikita, H. Cation Solvation with Quantum Chemical Effects Modeled by a Size-Consistent Multi-Partitioning Quantum Mechanics/Molecular Mechanics Method. *Phys. Chem. Chem. Phys.* **2017**, *19* (27), 17985–17997. https://doi.org/10.1039/c7cp01708a.

(69)   Habershon, S.; Markland, T. E.; Manolopoulos, D. E. Competing Quantum Effects in the Dynamics of a Flexible Water Model. *J. Chem. Phys.* **2009**, *131* (2). https://doi.org/10.1063/1.3167790.

(70)   Winkler, K.; Lindner, J.; Bürsing, H.; Vöhringer, P. Ultrafast Raman-Induced Kerr-Effect of Water: Single Molecule versus Collective Motions. *J. Chem. Phys.* **2000**, *113* (11), 4674–4682. https://doi.org/10.1063/1.1288690.

(71)   Lawrence, C. P.; Skinner, J. L. Vibrational Spectroscopy of HOD in Liquid D2O. III. Spectral Diffusion, and Hydrogen-Bonding and Rotational Dynamics. *J. Chem. Phys.* **2003**, *118* (1), 264–272. https://doi.org/10.1063/1.1525802.

(72)   Tan, H.-S.; Piletic, I. R.; Fayer, M. D. Orientational Dynamics of Water Confined on a Nanometer Length Scale in Reverse Micelles. *J. Chem. Phys.* **2005**, *122* (17), 174501. https://doi.org/10.1063/1.1883605.

(73)   Rezus, Y. L. A.; Bakker, H. J. On the Orientational Relaxation of HDO in Liquid Water. *J. Chem. Phys.* **2005**, *123* (11), 114502. https://doi.org/10.1063/1.2009729.