PCCP

Accepted Manuscript

ROYAL SOCIETY OF CHEMISTRY

www.rsc.org/pccp

An adaptive quantum mechanics/molecular mechanics method for infrared spectrum of water: Incorporation of the quantum effect between solute and solvent

Hiroshi C. Watanabe,[*,†] *Misa Banno,*[§] *and Minoru Sakurai*[§]

[*]corresponding author, e-mail: hwatanabe@protein.rcast.u-tokyo.ac.jp

[†] Research Center for Advanced Science and Technology, The University of Tokyo, 4-6-1 Komaba, Meguro-ku, Tokyo 153-8904 Japan

[§] Center for Biological Resources and Informatics, Tokyo Institute of Technology, B-62 4259 Nagatsuta-cho, Midori-ku, Yokohama 226-8501 Japan

Abstract: Quantum effects in solute–solvent interactions, such as the many-body effect and the dipole-induced dipole, are known to be critical factors influencing the infrared spectra of species in the liquid phase. For accurate spectrum evaluation, the surrounding solvent molecules, in addition to the solute of interest, should be treated with a quantum mechanical method. However, conventional quantum mechanics/molecular mechanics (QM/MM) methods cannot handle free QM solvent molecules during molecular dynamics (MD) simulation because of the diffusion problem. To deal with this problem, we have previously proposed an adaptive QM/MM "size-consistent multipartitioning (SCMP) method." In the present study, as the first application of the SCMP method, we demonstrate the reproduction of the infrared spectrum of liquid-phase water, and evaluate the quantum effect in comparison with conventional QM/MM simulations.

Introduction

Infrared spectroscopy, a nondestructive measurement with high sensitivity, is an essential technique that provides precise information regarding molecular structure and dynamics, and has been successfully adopted for various molecules. Although extraction of 3-dimensional molecular information from 1-dimensional spectrum is not always straightforward, molecular simulation is a practical facilitator for the detailed interpretation of an experimental IR spectrum, because it allows for an atomic level description and replication of dynamics. Note that absorption cross section $\alpha(\omega)$ is connected to dipole moment **M** obtained from molecular simulations as

$$\alpha(\omega) = \frac{4\pi^2}{\hbar cn(\omega)}\omega(1-e^{-\beta\hbar\omega})\sum_i\sum_f \rho_i \left|\langle f|\varepsilon\cdot\mathbf{M}|i\rangle\right|^2 \delta(\omega_{fi}-\omega) \quad (1)$$

where i and f are the state before and after transition, ω is the frequency of light where $\omega_{fi} = \omega_f - \omega_i$, ρ_i is a probability that a system is in the initial state i, $\boldsymbol{\varepsilon}$ is a unit vector along the electric field, $n(\omega)$ is the index of the refraction of the medium, c is the speed of light in vacuo, and $\hbar$ is the Dirac constant.[1, 2] An absorption line shape $I(\omega)$ is then defined as

$$I(\omega) \equiv 3\sum_i\sum_f \rho_i \left|\langle f|\varepsilon\cdot\mathbf{M}|i\rangle\right|^2 \delta(\omega_{fi}-\omega) \quad (2)$$

To obtain the line shape, the two main approaches that have been proposed are normal mode analysis (NMA)[3-5] and Fourier transform of time-correlation functions (FTTCF)[1, 2].

NMA is a static approach, which is applied to a minimized molecular structure. Based on the harmonic approximation, the line shape is then written with a derivative of a dipole moment. Note that in the NMA approach a system should be at the global minimum in the potential surface in order for application of the harmonic approximation. In vacuo, it is relatively straightforward to evaluate the potential, because an isolated molecule does not have complicated intermolecular interaction and the system is moderate in size. Furthermore, identification of the global minimum is usually realized by a simple structural optimization due to the smooth potential surface. In solution, on the other hand, the potential is a function of solvent coordinates as well as those of the solute of interest. Thus, the simulating system size becomes larger, and accordingly, demands more computational effort. Moreover, the intermolecular interaction between solute and solvent must be taken into account by incorporating the

many-body effect, the evaluation of which requires time-consuming computations such as diagonalization or iterative calculation. In addition, intermolecular interactions can cause numerous local minima, which inhibit the identification of the global minimum. Thus, the NMA approach is not practical in liquid systems, and is limited to gaseous systems. In NMA framework, furthermore, neglect of the anharmonic effect causes band peak shifts of 3%–10%.[6, 7] These errors are usually corrected by scaling factors. Such correction is not simple because of the dependency on the frequency range and molecular model. Thus, NMA is based on a priori knowledge, and careful verification is required when it is applied to an unknown system.

The instantaneous NMA (INMA) is an extension of NMA to take into the consideration the solvent effect, which is based on an assumption that the decoupling of solute molecules from the dynamics of solvent molecules. Then, NMA are applied to the solute structure optimized in the frozen solvent cage that is obtained from a molecular dynamics (MD) simulation, as detailed in literatures[8-10]. The INMA has been in practical use for condensed system and large molecule for which the dynamics simulations are not feasible due to considerable computational cost.[11-13] According to Schmitz et al.[14], however, the solute vibrations are decoupled from the dynamics of its environments due to the neglect of the coupling between the kinetic energy operator and eigenstates. Furthermore, in contrast to NMA the interpretation of the INMA spectrum is not straightforward, because it depends on the temporal solvent coordination.

In contrast, FTTCF is a dynamic approach with use of a time series of a dipole moment obtained from a molecular dynamics (MD) trajectory, and the absorption line shape $I(\omega)$ is directly calculated by Fourier transform of the time-correlation function of dipole, as

$$I(\omega) = \frac{1}{2\pi}\int_{-\infty}^{\infty} dt\, e^{-i\omega t}\langle \mathbf{M}(0)\cdot \mathbf{M}(t)\rangle \qquad (3)$$

where $\mathbf{M}(t)$ is a dipole operator and ω is frequency.[1, 2] The absorption cross section, $\alpha(\omega)$, is then written as

$$\alpha(\omega) = \frac{4\pi^2\omega}{3\hbar cn(\omega)}(1-e^{-\beta\hbar\omega})Q(\omega)I(\omega) \qquad (4)$$

where $Q(\omega)$ represents the quantum correction factor for nuclear motion in classical Newtonian dynamics[15]. Note that, in contrast to NMA, the harmonic approximation is not assumed in FTTCF. Moreover, FTTCF can incorporate the solvation effect on

dipole dynamics; e.g., vibration coupling and a dipole-induced dipole, with the use of an appropriate molecular model, as described below. Thus, FTTCF is more advantageous and practical for liquid-phase systems and molecular complexes such as proteins.

According to the Nyquist–Shannon sampling theorem, the spectrum resolution is inversely proportional to the total simulation time. For instance, to reach a resolution of 1 cm^{-1} comparable to that achieved by experiments, it is required to conduct MD simulations of around 30 ps. On the other hand, the time correlation function should be evaluated with sufficient sampling to achieve statistical reliability, although the required simulation time is empirical, and there have been few discussions about effective sampling time.

Note that these sampling problems pose a dilemma that is common to all molecular simulations—the trade-off between model accuracy and statistical accuracy. At present, there are various molecular models available for MD simulations, which are classified into three major groups: quantum mechanics (QM), molecular mechanics (MM), and hybrid QM/MM methods[16].

QM, which counts explicit electronic structure, is an accurate and transferable form of molecular modeling, but demands enormous computational resources. Thus, it has been widely employed to reproduce delicate chemical reactions and physical properties[11, 17, 18], although there is a severe limit on the system size and time scale of QM-MD simulations. The computation of both large molecules such as proteins, and bulk conditions where a solute is surrounded by numerous solvent molecules is beyond the present limits of QM methods with respect to system size. Thus, a bulk condition is usually either substituted with a small cluster consisting of a solute and several solvent molecules in vacuo, or modeled with a small periodic box. However, both can cause artificial boundary effects on the solute of interest. In addition, due to finite simulation time, QM-MD simulations can cover only a small phase space as most spectroscopic studies are based on trajectory at ~ 30 ps.[19-25].

MM models, in contrast, are designed to give high statistical accuracy at the cost of model accuracy. To this end, it adopts an empirical potential for atomic interaction, where the intermolecular interaction is replaced by an effective two-body potential to realize high computational efficiency. As aimed, conventional MM-MD simulation has been successfully applied for global structural sampling of gigantic systems such as

proteins, while it has failed in the reproduction of delicate molecular properties such as molecular vibrations and rotations because of the missing many-body effect and dipole-induced dipole. To deal with this problem, alternative empirical models termed polarizable force fields have been proposed, which can count the many-body effect and induced dipole by the introduction of atomic polarizability and iterative calculations.[20, 26-29] Although the polarizable models take for the computations several times longer than the conventional MM models, they have drastically improved the description of many molecular properties, including the infrared spectrum. However, it should be noted that these polarizable models need a priori parameterization with respect to each molecule in a system of interest, which often requires careful and elaborate procedures.

The third option is the hybrid model QM/MM method, which takes the respective advantages of both QM and MM methods, by dividing the whole system into two parts. The region of interest is treated by the QM method, and the rest is modeled with MM. Since its introduction in 1976[16], the QM/MM method has been successfully used in the investigations of chemical reactions in large systems such as condensed systems and macromolecules. Note that QM/MM has been instrumental in combination in either case of (1) static simulations such as structural optimization and Monte Carlo simulation[11, 17, 30], or (2) dynamics simulations without free QM solvent molecules.[31-35] Thus, it can work fine when the QM solvent molecules are strongly bound to proteins, as demonstrated in the previous studies, in which the water molecules trapped in the internal space of the proteins are included in the QM region.[35, 36]. On the other hand, in the case of dynamics simulations with free QM solvent, careful treatments are required because of the diffusion problems as described below.

Consider a QM/MM MD simulation starting with a certain snapshot structure where the QM region consists of a solute and some solvent molecules in the vicinity of the solute. In the course of the MD simulation, these QM solvents diffuse away from the QM solute, and MM solvents take their place. As a result, the quantum effect on solvation will drop off in the deformed or defragmented QM region, because the solute–solvent interaction will slip into classical mechanics.

A naive solution for this diffusion problem is the introduction of the intermittent update of QM/MM partitioning, such that solvents in the vicinity of a solute of interest are constantly redefined as QM. However, this simple approach would face spatial and temporal discontinuities. Spatial discontinuity arises from different footings in

molecular description. Since QM and MM solvents are no longer equivalent, they cause a spurious boundary between the QM and MM regions, which leads to an artificial solvation structure. Temporal discontinuity manifests as energy or temperature drift, because abrupt switching of molecular definition causes artificial discontinuity in forces.

To deal with these problems, adaptive QM/MM methods have been proposed in the past two decades. As detailed in a previous work,[37] they are classified into three generations. The first generation is designed mainly to alleviate the spatial discontinuity by introducing a smooth switching function, without an energy function (Hamiltonian)[38-40] to be conserved, while the temporal discontinuity remains. The second generation, which is based on size-scaling multipartitioning, challenges the expression of an energy function and improvement of temporal continuity.[41-43] However, these methods suffer from the issue of spatial continuity breakdown as the QM region size increases, while temporal continuity is improved.[44-46]

It is also notable that other approaches that have been proposed, such as FIRES and QBEST, fall outside of the three-generation classification. These are based on a restraint scheme to prevent the exchange between QM and MM solvents.[47, 48] Unlike other adaptive methods, they realize stable and efficient computation. However, they yield spurious dynamics, and spatial discontinuity still seemingly remains.[46]

In the previous study we proposed, as the a third-generation method, the new adaptive QM/MM method of size-consistent multipartitioning (SCMP), which as reported, realizes a stable MD simulation with energy conservation and alleviated spatial discontinuity.[37] Thus, in the present study, we attempt the reproduction of an infrared spectrum as the first application of an SCMP simulation based on the FTTCF approach.

However, the QM/MM method with free QM solvent has another challenge when applied to the reproduction of an infrared spectrum, that of dipole moment description. Since charge fluctuation between solute and solvent plays an important role in the infrared spectrum, solvent molecules surrounding the solute should be also included in the QM region. However, when free solvent molecules diffuse away from the QM solute, the charge transfer effect is consequently not counted, because the solute–solvent interactions are replaced with MM interactions. Furthermore, in a periodic boundary condition, diffused QM solvents can interact with an imaginary QM solute in a neighboring box, which causes severe discontinuity in the time-correlation function of

the dipole, and an artifact in the spectrum. To avoid these problems, in most QM/MM simulations solvent molecules are excluded from the QM region. In this case, however, solute–solvent and solvent–solvent interactions are not different from MM models, which diminish the original advantage of QM/MM methods.

In the present study, we propose a new dipole moment representation to avoid the aforementioned problems, by taking advantage of the SCMP method. As the first application of the SCMP method, we choose liquid-phase water as the benchmark to evaluate the spectrum. Water is one of the most important molecules in infrared spectroscopy, both as a molecule of interest and as a technical fundamental. In addition, liquid-phase water is a special case for which pure QM and MM simulations possess a sampling advantage. Since the solute and solvent are the same molecular species in a uniform system, the number of sampled trajectories for the molecules of interest obtained from a single MD run with pure QM and MM methods is proportional to the system size. In contrast, only a single piece of solute information is available for analysis in the case of the QM/MM method. For these reasons, there have been numerous studies on liquid-phase water using pure QM and MM methods, which are comparable to the present study.

Among these previous studies, we note two remarkable ones that achieved an accurate reproduction of the infrared spectrum of liquid water. First, Del Ben et al. conducted a MD run of 10 ps with a QM description based on Møller-Plesset perturbation theory (MP2) for a 64-water-molecule system. They reproduced the infrared spectrum of liquid water, which revealed the limit of Born-Oppenheimer MD compared with quantum dynamics simulations.[49, 50] Second, polarizable models such as AMOEBA,[51] TTM,[52, 53] CRK,[27] and HHB[29, 54] have also been successfully used to analyze liquid water dynamics. In particular, the Mb-pol model, a polarizable model proposed by Medders et al., can accurately reproduce potential energy surfaces, including higher-level interactions such as correlation energy and charge transfer, by a well-parameterized force field based on coupled-clusters. Note that this model involves a lower computational cost than the pure QM model. This model realizes an extension of MD duration and system size, as Medders et al. demonstrated for a MD run of 27 ps for 216 water molecules in the combination with centroid MD,[29] 100 ps with classical MD simulation.[55]

Regardless of the success with liquid water, these methods are still far from a general

application that can handle a variety of systems because of several lingering problems. As mentioned above, pure QM simulations face severe size limitations, and cannot handle solvation of molecules, even medium-size ones. In addition, when it comes to non-uniform systems, the sampling advantage of pure QM simulations is no longer valid. Therefore, a longer MD simulation time is required for reliable statistics, but such long MD runs are not feasible because of the extremely high computational cost. On the other hand, polarizable models have questionable transferability to cover some systems because of complexity and ambiguity in parameterization, as described above.

In general, a QM/MM method can control the balance between sampling efficiency and model accuracy by selection of the appropriate QM model. In the present study, among the various choices available for the QM method within the SCMP QM/MM framework, we adopt a density functional tight-binding (DFTB) method, which introduces integral approximation to realize a 2-3 order of magnitude faster computation than standard density functional theory (DFT). Although DFTB has limitations in accuracy compared with MP2 and DFT[56-58], various attempts have been systematically made to overcome specific problems.[59-66] For instance, it has been reported that DFTB overestimates the density of bulk water under ambient conditions.[67, 68], and also shown that the minor change in repulsive potential drastically improve the oversolvation properties.[65] Due to the fine balance between accuracy and cost, DFTB has been widely applied to various systems in which higher QM methods are not feasible[30, 32-34, 69-71]. Hence, DFTB can fully exploit the stability of the SCMP MD simulation.

Note that the present simulations are based on Born-Oppenheimer MD, and nuclear quantum effects are neglected. However, according to the previous studies based on centroid MD[29, 72], the nuclear quantum effects can largely affect on infrared spectrum of liquid-phase water, in particular the band peak positions. Thus, the present study focuses on the improvement of the infrared spectra by incorporation of quantum effects in solute-solvent interaction with the SCMP QM/MM method rather than the absolute value of each band peak position.

In this paper, the SCMP method is briefly reviewed in section II, with the introduction of an alternative representation of the dipole moment. In section III, we evaluate the infrared spectra obtained from the SCMP simulations and compared that with the spectra obtained from the conventional QM/MM simulations. In section IV, the conclusions of the study are reviewed.

Materials & Methods

Size-consistent multipartitioning QM/MM

Here, we briefly review the size-consistent multipartitioning QM/MM method, which is detailed in the previous paper.[37] Consider a QM/MM partitioning where the QM layer consists of a solute and q solvent molecules, and the rest of the system is modeled by MM. In the SCMP method, one considers N different QM/MM partitionings such that all partitionings have the same number of QM solvent molecules, i.e., one can write

$$n(S^{(j)}) = q \quad (j = 1,\ldots,N) \tag{5}$$

where $S^{(j)}$ is a QM molecule subset in the j^{th} partitioning.

On the basis of QM/MM force $\mathbf{f}_m^{(n)}$ acting on the m^{th} atom in the n^{th} partitioning, the effective force $\mathbf{F}_m^{\text{eff}}$ in the SCMP method is defined as

$$\mathbf{F}_m^{\text{eff}}(\mathbf{r}) = \sum_n^N \sigma^{(n)}(\mathbf{r}) \cdot \mathbf{f}_m^{(n)}(\mathbf{r}) \tag{6}$$

where $\sigma^{(n)}$ is a normalized weight function of atomic coordinates $\mathbf{r}$ for the n^{th} partitioning. Thus, the effective potential is

$$\begin{aligned} V^{\text{eff}}(\mathbf{r}_{t=T}) &= -\sum_m^M \int \mathbf{F}_m^{\text{eff}}(\mathbf{r}) \cdot d\mathbf{r}_m \\ &= -\sum_n^N \sum_m^M \int \sigma^{(n)}(\mathbf{r})\, \mathbf{f}_m^{(n)}(\mathbf{r}) \cdot d\mathbf{r}_m \\ &= \sum_n^N \sigma^{(n)}(\mathbf{r}) V^{(n)}(\mathbf{r}) - \sum_n^N \sum_m^M \int \nabla_m \sigma^{(n)}(\mathbf{r}) V^{(n)}(\mathbf{r}) d\mathbf{r}_m \\ &= \sum_n^N \sigma^{(n)}(\mathbf{r}) V^{(n)}(\mathbf{r}) - \sum_n^N \sum_m^M \int_{t=0}^T \left[V^{(n)}(\mathbf{r}) \left(\nabla_m \sigma^{(n)} \cdot \frac{\partial \mathbf{r}_m}{\partial t} \right) \right] dt \\ &= \sum_n^N \sigma^{(n)}(\mathbf{r}) V^{(n)}(\mathbf{r}) - \sum_n^N \sum_m^M \sum_{t=0}^T V^{(n)}(\mathbf{r}) \left(\nabla_m \sigma^{(n)} \cdot \mathbf{v}_m \right) \Delta t \end{aligned} \tag{7}$$

where M is the total number of atoms and $V^{(n)}(\mathbf{r})$ is a QM/MM potential energy in the n^{th} partitioning. Note that the second term in eqn (8) corresponds to the book-keeping term introduced by Bulo et al.[41] As described previously, we imposed two conditions on the weight function: (1) $\sigma^{(n)}$ is zero whenever the QM region is fragmented due to the diffusion of the QM solvent, and (2) $\sigma^{(n)}$ is zero whenever the QM region has a compact form. These requirements allow for replacement of a partitioning with a fragmented QM region by one with a compact QM region, without any discontinuity.

To this end, we introduce transition profiles $\Lambda_{\text{out}}^{(n)}$ and $\Lambda_{\text{in}}^{(n)}$ of the i^{th} solvent molecule in the n^{th} partitioning as

$$\Lambda_{\text{out}}^{(n)} = \begin{cases} \lambda_{\text{QM}}^{(n)}(\mathbf{d}_i : s_{\text{QM}}^{\text{out}}, t_{\text{QM}}^{\text{out}}) & \text{if} \quad i \in S^{(n)} \\ \lambda_{\text{MM}}^{(n)}(\mathbf{d}_i : s_{\text{MM}}^{\text{out}}, t_{\text{MM}}^{\text{out}}) & \text{if} \quad i \notin S^{(n)} \end{cases} \tag{8}$$

$$\Lambda_{\text{in}}^{(n)} = \begin{cases} \lambda_{\text{QM}}^{(n)}(\mathbf{d}_i : s_{\text{QM}}^{\text{in}}, t_{\text{QM}}^{\text{in}}) & \text{if} \quad i \in S^{(n)} \\ \lambda_{\text{MM}}^{(n)}(\mathbf{d}_i : s_{\text{MM}}^{\text{in}}, t_{\text{MM}}^{\text{in}}) & \text{if} \quad i \notin S^{(n)} \end{cases} \tag{9}$$

where vector $\mathbf{d}_i$ is a vector from a solute in the QM center to the i^{th} solvent, and λ is a progress function of $\mathbf{d}_i$ that parametrically depends on s^{out}, t^{out}, s^{in}, and t^{in} ($s < t$), and defined as

$$\lambda_{\text{QM}}(\mathbf{d}_i : s_{\text{QM}}, t_{\text{QM}}) = \begin{cases} 1 & \text{if } r_i \le s_{\text{QM}} \\ \dfrac{\left(t_{\text{QM}} - d_i\right)^2 \left(2d_i + t_{\text{QM}} - 3s_{\text{QM}}\right)}{\left(t_{\text{QM}} - s_{\text{QM}}\right)^3} & \text{if } s_{\text{QM}} \le r_i \le t_{\text{QM}} \\ 0 & \text{if } t_{\text{QM}} \le r_i \end{cases} \tag{10}$$

$$\lambda_{\text{MM}}(\mathbf{d}_i : s_{\text{MM}}, t_{\text{MM}}) = \begin{cases} 0 & \text{if } r_i \le s_{\text{MM}} \\ \dfrac{(d_i - s_{\text{MM}})^2 (3t_{\text{MM}} - 2d_i - s_{\text{MM}})}{(t_{\text{MM}} - s_{\text{MM}})^3} & \text{if } s_{\text{MM}} \le r_i \le t_{\text{MM}} \\ 1 & \text{if } t_{\text{MM}} \le r_i \end{cases} \tag{11}$$

Fade-out and fade-in functions, $O(\mathbf{r})$ and $I(\mathbf{r})$ are now introduced as

$$\begin{aligned} O_{\text{QM}}^{(n)}(\mathbf{r}) &= \prod_{i \in S^{(n)}} \Lambda_{\text{out}}^{(n)}(\mathbf{d}_i), \quad I_{\text{QM}}^{(n)}(\mathbf{r}) = 1 - \prod_{i \in S^{(n)}} \Lambda_{\text{in}}^{(n)}(\mathbf{d}_i), \\ O_{\text{MM}}^{(n)}(\mathbf{r}) &= \prod_{i \notin S^{(n)}} \Lambda_{\text{out}}^{(n)}(\mathbf{r}_i), \quad I_{\text{MM}}^{(n)}(\mathbf{r}) = 1 - \prod_{i \notin S^{(n)}} \Lambda_{\text{in}}^{(n)}(\mathbf{d}_i), \end{aligned} \tag{12}$$

The weight function $\sigma^{(n)}(\mathbf{r})$ of the n^{th} partitioning is then defined as

$$\sigma^{(n)}(\mathbf{r}) = \frac{O_{\text{QM}}^{(n)} O_{\text{MM}}^{(n)} I_{\text{QM}}^{(n)} I_{\text{MM}}^{(n)}}{\sum_n^N O_{\text{QM}}^{(n)} O_{\text{MM}}^{(n)} I_{\text{QM}}^{(n)} I_{\text{MM}}^{(n)}} \tag{13}$$

Note that $\sigma^{(n)}(\mathbf{r}) = 0$ when a QM solvent diffuses away beyond $t_{\text{QM}}^{\text{out}}$. Likewise, $\sigma^{(n)}(\mathbf{r}) = 0$ when an MM solvent moves within $s_{\text{MM}}^{\text{in}}$. Thus, $\sigma^{(n)}(\mathbf{r})$ in eqn (14) satisfies the two requirements described above.

Dipole moment

In a pure QM or MM simulation, the total dipole moment of a system is directly used

for the calculation of the infrared spectrum, while it is not straightforward in QM/MM simulations, including the SCMP method, because only the contribution of the QM molecules to the dipole is usually of interest.

Thus, a dipole moment for a QM region may be defined as

$$\hat{\mathbf{M}} = \sum_{i \in S_{\mathrm{QM}}} q_i \mathbf{r}_i \tag{14}$$

where q_i is a Mulliken charge of the i^{th} QM atom at position $\mathbf{r}_i$. QM solvents near the QM center interact with other QM molecules, while those close to the QM/MM boundary interact with MM molecules. Thus, their molecular properties vary depending on the surrounding environment. In eqn (15), however, QM solvents make equal contributions to the dipole moment, regardless of their positions. Thus, an admixture of different signals the in IR spectrum may result, which may be difficult to separate for detailed analysis.

Here, we propose an alternative definition of a local dipole moment in the SCMP simulation, such that the MM solvent contribution and the artifact from the spurious boundary are reduced as much as possible.

$$\hat{\mathbf{M}} = \sum_{n}^{N} \sum_{i} \sigma^{(n)} \delta^{(n)}(i) q_i^{(n)} \mathbf{r}_i \tag{15}$$

where $\sigma^{(n)}$ is the weight of the n^{th} partitioning and

$$\delta^{(n)}(i) = \begin{cases} 1 & \text{if } i \in S^{(n)} \\ 0 & \text{if } i \notin S^{(n)} \end{cases} \tag{16}$$

Note that the sum of Mulliken charges over all QM atoms in a neutral QM region is always zero, i.e.,

$$\sum_{i} \delta^{(n)}(i) q_i^{(n)} = 0 \tag{17}$$

To deduce the local dipole moment defined by eqn (18), the QM profile is a useful index, which we introduced in the previous study.[37] A QM profile $\omega(i)$ of the i^{th} solvent molecule is defined as

$$\omega(i) = \sum_{n}^{N} \sigma^{(n)}(\mathbf{r}) \delta^{(n)}(i) \tag{18}$$

which indicates a ratio between QM and MM properties. Thus, a solvent becomes a perfect QM molecule when $\omega = 1$, and likewise becomes a perfect MM molecule when $\omega = 0$. Since in eqn (19) a dipole moment does not count MM charges, direct

contributions from solvent molecules with $\omega = 0$ are removed. As discussed below, the QM profile is smoothly damped from 1 to 0 as the distance from the QM center become larger. Thus, the representation of the dipole moment enables alleviation of the spurious boundary effect.

Simulation details

For the QM aspects of the SCMP method, we adopted the DFTB method[62, 73] from the various possible choices. This is a semi-empirical QM method, which approximates DFT by introducing parameters for interaction integrals. Since DFTB is characterized by increased computational speed at the cost of accuracy, it can fully exploit the advantage of the SCMP method.

We implemented DFTB3 and the SCMP method in a local version of GROMACS 5.0.2 package.[74-76] The electric potential induced by the charges of MM atoms on the QM atoms was obtained with the smooth particle-mesh Ewald method with a switching function for electrostatic interaction,[77] while van der Waals interactions are damped to zero with a switching function in a range between 9.5 and 10.0 Å.

For QM spectra in vacuo, we adopted both DFTB2 and DFTB3 to evaluate the QM parameter dependency with mio[78-80] and 3ob[63] parameter sets, respectively. Note that in gaseous simulations, the motions of the center of mass were removed to avoid spectrum broadening that was dependent on the initial conditions. As quantum correction factor in eqn (4), we adopt harmonic type correction[15] for classical dynamics as

$$Q(\omega) = \frac{\beta\hbar\omega}{1 - e^{(-\beta\hbar\omega)}} \quad (19)$$

.

For the liquid phase, we carried out both conventional and SCMP QM/MM simulations. In the former simulations, only one water molecule was defined as the QM solute and the rest as the MM solvent. We then tested combinations of DFTB3 with several MM models to evaluate the cross-interaction between the QM and MM models. Of the various MM water models proposed to date, we adopted four major models for the comparison. Two of them, TIP3P[81] and SPC[82], are rigid water models, in which the stretching and bending modes are suppressed by the LINCS algorithm. The others, TIP3P-Fs[83, 84] and SPC-Fw,[85] are extended models that count the intramolecular degrees of freedom by the introduction of stretching and bending parameters.

For the SCMP method, 48 QM/MM partitionings were considered, in which the QM

region of each partitioning consists of one and 31 water molecules as the solute and the solvent, respectively. We then conducted two combinations of DFTB3 with different MM models, TIP3P-Fs and SPC-Fw. A threshold value $\sigma_T = 1.0^{-4}$ was adopted for the partitioning update for the SCMP method. Note that there are several possible options for defining vectors $\mathbf{d}_i$. in eqn (10) and (11). In the present study, we defined $\mathbf{d}_i$ as a vector from an oxygen atom of a water molecule in the QM center to an oxygen atom of the i^{th} water molecule. Transition parameters were set as $s^{in}_{MM} = s^{out}_{MM} = 3.0$ Å, $t^{in}_{MM} = t^{out}_{MM} = s^{in}_{QM} = s^{out}_{QM} = 5.5$ Å, and $t^{in}_{QM} = t^{out}_{QM} = 8.5$ Å.

The MD simulations were conducted over 8.0 ns with a time step of 0.5 fs for the conventional QM/MM methods, and over 800 ps for the SCMP methods. Since the spectrum is in general sensitive to temperature, for the sake of a fair comparison, the Nosé–Hoover thermostat[86, 87] was adopted with a time constant of 500 fs for temperature control at 300 K. Note that, since the thermostat can alter the molecular dynamics, application to the spectrum calculation requires careful treatment. We were able to confirm that temperature control causes little difference in the infrared spectrum, as shown in Supplementary Figure 1.

Results & Discussion

For convenience, conventional QM/MM simulations are denoted as c:QM/MM, and SCMP adaptive QM/MM simulations are denoted as a:QM/MM. In the following sections, the designation "QM/MM" represents both conventional and adaptive simulations, unless otherwise specified.

Water is well known to have three distinctive band peaks: a symmetric stretch (ν_1), an angle bending (ν_2), and an asymmetric stretch (ν_3), which are subject to the surrounding environment. Figure 1 presents gaseous and liquid phase spectra with DFTB3-based simulations. For better evaluation, it is worthy to note the recent study by Goyal et al. that demonstrated the pure DFTB3 spectrum of bulk water with pure DFTB simulations.[65] Note that the whole shape of the a:DFTB3/MM spectra is in agreement with that of the pure DFTB3, although the pure DFTB3 spectrum appears to be lower-frequency shifted with respect to the ν_1 and ν_3 bands. In addition, the comparison indicates the ν_1 and ν_3 bands of the a:DFTB3/MM spectra are more separated showing the distinct peaks, which presumably arises from MM solvent contribution as described below.

Note that the present study is based on classical dynamics, and nuclear quantum effects are excluded that causes shifts of major bands to lower frequency.[72] Thus, although the present simulation does not necessarily yield the exact peak positions, we found several common features in the aqueous spectra that contrast with those of the gas phase spectra: (1) the shifts to a lower frequency of the two OH stretching modes from those in gas phase, (2) the shift to a higher frequency of the bending mode, (3) the broadening of each absorption band, (4) the relative intensity of the peaks, and (5) the continuous far infrared band from 50 to 900 cm^{-1} (see also Supplementary Figure 2). Thus, in the following sections, we focus on each spectral feature in turn, comparing the conventional and SCMP QM/MM simulations.

Dependency of spectra on simulation time

Unlike pure QM simulations, c:QM/MM simulations for the liquid phase generally treat only one solute molecule with the QM method, even if the system contains more than one solute molecule. Thus, QM/MM MD simulations are much faster than pure QM simulations as the system size increases. However, the amount of solute information obtained from a single QM/MM MD run is smaller than that from a pure

QM run in the case of multisolute system, particularly for uniform liquid systems such as pure water, and thus QM/MM simulations require longer MD simulation time for reliable statistics. Nevertheless, there have been few studies regarding the effective MD simulation time required for accurate spectrum evaluation, because there has been no stable QM/MM framework to treat free QM solvent. In the present study, we first evaluate the dependency of the spectrum on simulation time using 800-ps MD trajectories obtained by the SCMP QM/MM method.

As indicated in Figure 2, as the simulation time is extended, the spectrum becomes clearer, particularly with respect to lower-frequency regions where a small shoulder is found at 300 cm^{-1} in the 800-ps simulation, as discussed later. On the other hand, although the positions and half-width at half-maximums (HWHMs) of the major peaks evaluated by Lorentzian fitting almost converge, the high-frequency region around 3500 cm^{-1} still remains slightly rough and blurred, which implies that the spectra are not perfectly converged, even for a simulation time of 800 ps.

Notably, Figure 2 also implies that the positions and HWHMs obtained from a MD run of 30 ps can contain errors larger than 10 cm^{-1} compared with runs at 800 ps, which indicates sufficient sampling as well as model accuracy are critical for an accurate evaluation of the infrared spectrum. It should be noted that statistical error is also critical for pure QM simulations when employed for non-uniform systems, because the sampling advantage of pure simulations is lost. Alternatively, one may evaluate physical properties with multiple independent MD trajectories to gain statistical reliability. It provides rather valuable ensemble, because multiple MD runs possibly realize effectively sampling over different local minima that may not be covered by a single MD run. However, the resulting ensemble with finite simulation time is not necessarily equivalent to that from a single MD run, in particular for flexible molecules with multi-conformations, and as a result the ensemble can rather depend on the initial condition. Therefore, although the multi-MD analysis is suitable for qualitative search for invariant properties among the minima, the extension of a single MD run is still of high interest for quantitative analysis such as free energy evaluation. To this end, the SCMP QM/MM is more advantageous than the pure QM simulation due to computational efficiency.

ν_1 and ν_3 shifts to lower frequency

Compared with the gas phase, the ν_1 and ν_3 bands recorded in the experiments are completely overlapped and their respective positions are difficult to identify (Figure 1). Several research groups have proposed different values for ν_1 and ν_3, but they are consistently shifted by more than 250 cm^{-1} to lower frequency as a result of solvation.[88, 89] In contrast, both the c:QM/MM and a:QM/MM spectra show distinct peaks in the ν_1 and ν_3 bands (Figure 1). This separation enables evaluation of the respective peak shifts and broadenings by Lorentzian fitting.

Note that the absolute band positions obtained from the simulations with DFTB appear to be different from those of experiments as shown in Table 1. The deviations mainly arise from the semi-empirical QM method of DFTB, and thus the errors are also present in the gas water spectrum calculated. (See Table 1). Note that the discrepancy is presumably related to the molecular geometry, which are considerably dependent on the parameter set (see Supplementary Figure 3 and Supplementary Table 1). Since adaptive QM/MM methods are designed for improvement of description of solvation effect, for a proper evaluation we focus mainly on the spectrum difference between the gaseous and liquid phases in the following section rather than their absolute values.

Firstly, Figure 3 indicates that MM solvent models cannot properly reproduce solvation peak shifts of the stretching modes from the gas phase. Secondly, although the ν_1 and ν_3 bands in the c:QM/MM spectra are shifted to lower frequency by around 200 cm^{-1} from those in vacuo, the shifts are not sufficient when compared to experiments reporting shifts of more than 250 cm^{-1}. Finally, ν_1 and ν_3 in the a:QM/MM spectra were additionally shifted to a lower frequency by ca. 50–70 cm^{-1} compared to those of the c:QM/MM spectra, and the total shifts from those in vacuo are about 250 cm^{-1}, in good agreement with experimental values with respect to ν_3, although the ν_1 shift is not sufficient. Therefore, although the absolute positions of the major band peaks by DFTB3 contain large errors, the solvation shifts are well reproduced by the SCMP simulation.

The various theoretical studies have demonstrated that the band shifts principally arise from hydrogen bonding with neighboring molecules, where hydrogen atoms are attracted away from the host oxygen, which consequently attenuates the bond stretching intensity, yielding lower frequency vibration. According to Bouteiller et al., a proton-donating hydrogen bond in a water dimer has ν_1 and ν_3 at frequencies lower by 60 cm^{-1} and 20 cm^{-1}, respectively, while that of ν_3 shift to a lower frequency as the

proton acceptor is small.[90] Indeed, all of the QM/MM simulations indicate that the bonds of liquid phase water molecules extend by ca. 0.02 Å compared to those in gas phase molecules, in agreement with experiments showing an extension by 0.02–0.05 Å (Table 2). The OH bond length and the band positions of the stretching modes exhibit a clear correlation (see Supplementary Figure 4). However, while the MM water models also exhibit the bond extension due to solvation, they do not show the corresponding band shifts. This result evinces that the MM models lack transferability to represent the shift of ν_1 and ν_3 by the bond extension.

For molecular geometry analysis, we made a comparison of the OH bond extensions in isolated water dimers with different hydrogen bonding patterns as shown in Figure 4. Notably, the DFTB3-MM water dimers show a shorter hydrogen bond distance than the DFTB3 pair, and correspondingly larger extension of the covalent bonds of the hydrogen-donating water. The hydrogen bond difference arises from the treatments of non-bonded pairwise interaction as described below. For DFTB3 solute, therefore, MM solvent seemingly yields larger vibration band shifts than DFTB solvent. In condensed phase, however, the DFTB3 solute in the a:QM/MM system has a longer OH bond length than that in the c:QM/MM system (Table 2), which consequently leads to further shifts of the ν_1 and ν_3 peaks to lower frequencies, presumably because of many-body effect and the transferability of DFTB3. Yu et al. demonstrated that DFTB could qualitatively reproduce these asymmetric shifts and confirmed the accumulation of shifts of up to ca. 270 cm^{-1} as the water cluster grew in size[91], which implies the many-body effect is well reproduced by DFTB3. On the other hand, the DFTB3 solute in the c:QM/MM simulations does not show sufficient bond extension compared with the dimer coordinations, because of the missing quantum effect in the solute–solvent.

In the QM/MM framework, the van der Waals interactions are treated with standard force field parameters for both QM and MM atoms. Since the van der Walls parameters of MM water models have been developed assuming bulk water system, the hydrogen bond distance are well reproduced as seen in the QM-MM dimers in Figure 4, although the accuracy of the parameters are questionable under different conditions such as interactions with other molecular species. On the other hand, the repulsive potentials of DFTB are empirically determined by fitting to high-level calculations and experimental values. Although they realizes high performance and transferability, the shortcomings have been also reported, such as underestimation of hydrogen bonding energy and

overbinding of covalent bond as seen in Figure 4.[57, 61, 65, 91] Note that, however, the systematic investigations have clarified the causes of the drawbacks, and manifested possible improvement.[57] Indeed, Goyal et al. demonstrated the drastic improvement of the molecular geometry of water and solvation structure by an adjustment of repulsive potential with a reverse Monte Carlo simulations.[65] Therefore, we conclude that the combination of a:DFTB3/MM outperforms c:DFTB3/MM with respect to the description of the water vibration modes due to incorporation of many body effect, and holds promise for further improvement according to DFTB development.

It is also notable that the c:QM/MM systems with flexible solvents, TIP3P-Fs and SPC-Fw, show additional bond extensions in comparison to the rigid models, TIP3P and SPC (Table 2). Correspondingly, ν_1 was shifted by 10 cm^{-1}, whereas no distinct difference is found with respect to ν_3 (see Figure 5). This indicates that not only the static properties of neighboring solvent but also their dynamic properties affect the solute structure, which is related to the asymmetric shift of ν_1 and ν_3. Since SCMP can treat solvent interactions at the QM level, further frequency shifts in the a:QM/MM spectra compared to those in c:QM/MM come partly from the improved solvent dynamics.

Higher-frequency shift of ν_2

Although Figure 3 indicates that MM/MM systems reproduce the higher-frequency shift of the ν_2 band by solvation, the shift is overestimated at 70 cm^{-1} compared with the experimental values, which report a shift of 19 cm^{-1}. The overestimation is also seen in the c:QM/MM spectra, while the a:QM/MM spectra improve the shift to 30 cm^{-1}, which is slightly larger than the experimental value.

It is reasonable to associate the ν_2 shift with the bending angle. Indeed, we found that MM solutes exhibit a considerable decrease in the bending angle due to solvation (TIP3P-Fs: 104.5°–98.0°, and SPC-Fw: 113.2°–107.8°). Although this is also the case for DFTB3 solutes, which show a decrease of 2° in both c:QM/MM and a:QM/MM systems, the decrease in the angle is much smaller than that in the MM solute. Surprisingly, however, experiments show the bending angle is enlarged by solvation. In other words, the experiments show a positive correlation between the angle and the ν_2 shift (see Supplementary Figure 5), while the simulations show a negative correlation. Note, however, that this is not a fair comparison because there is no overlap of

compared angle ranges between DFTB (108°–110°) and the experiment values (104°–106°).

To analyze these results in more detail, we investigated the cause of the shift in the bending angle. One might presume that this change in the angle arises from the adaption of the solute water to the surrounding cage of water molecules. If this is the case, TIP3P-Fs and SPC-Fw should form cages with different structures, because they take considerably different values of the angle in MM/MM systems (Table 2). In the c:QM/MM system, however, the bending angle of a QM solute water is free from the MM models and takes a specific value of ca. 108.1°, which implies that the bending angle of the QM solute is determined by a solute property rather than the solvent cage.

Next, we evaluated the geometry of a hydrogen-bonded water dimer in vacuo with DFTB3, and found bending angles of 109.83° and 110.44° for the proton-donating and proton-accepting water molecules, respectively, both of which are smaller than that in the monomer (110.50°). Furthermore, when DFTB3 water forms a proton-donating hydrogen bond with MM water (SPC-Fw), the bending angle is 109.33°, which is smaller than those in the DFTB3 water pair (see Figure 4). This result explains the reason for the smaller bending angle in the c:QM/MM systems than in the a:QM/MM systems. The van der Waals interactions in MM models can lead to the overestimation of the band shift of ν_2 by solvation, whereas they are tuned to reproduce the better bond length extensions. On the other hand, since the SCMP method can take into account the solvent electronic structure, it can describe the bending angle more accurately, and will be a promising application in combination with other higher-level QM methods.

Broadening of each absorption band

As shown in Figure 1, an isolated water molecule in vacuo yields a thin line spectrum, where the HWHM is estimated by Lorentzian fitting to be 1–2 cm^{-1} for the three major peaks. In contrast, a water molecule in the liquid phase can have various hydrogen bond patterns. Accordingly, the vibration state varies, resulting in line broadening, where the ν_1 and ν_3 bands form a united broad peak at around 3400 cm^{-1}, and ν_2 band is broadened with an HWHM of 50 cm^{-1}.

Although the MM/MM spectra exhibit broadenings of the respective bands, as shown in Figure 6, they remain far narrower than the experimental widths. In the c:QM/MM spectra, the ν_2 band seems to be moderately broad, while the stretching bands remain

narrow and separated (Figure 5). Interestingly, the flexible MM solvents cause additional broadening in the c:QM/MM spectra when compared to that of the rigid MM solvents. This additional effect is distinct with respect to the ν_1 and ν_2 bands, but is not apparent for the ν_3 band (see Figure 5). Thus, the intramolecular degree of freedom of surrounding solvents inhomogeneously contributes to the vibrational states of the solute molecule.

The a:QM/MM spectra may contain a large error regarding HWHMs due to the blurred spectra, but they clearly show additional line broadening to the major three peaks compared to the c:QM/MM spectra. Although the broadenings of the ν_1 and ν_3 bands may remain insufficient, the band shape becomes closer to that of the experimental spectra by increasing the overlap between the two bands. Therefore, the quantum effect between the solute and the solvent is also significant for the reproduction of realistic line broadening. For the ν_2 band, the two SCMP spectra show MM model dependency, with a moderate width of 53.2 cm^{-1} shown in the TIP3P-Fs spectrum, and an overestimated HWHM of 62.2 cm^{-1} observed in the SPC-Fw spectrum. We assume that this dependency arises mainly through the transition solvent between QM and MM, as described below.

Relative intensities

Incorporation of the quantum effect also improves the relative intensities among the major three bands. Experimentally, the intensity ratio among the major bands is estimated as ν_1 (ν_3):ν_2 = 1.00:0.23, whereas all the c:QM/MM spectra show ν_2 with a larger intensity than those of ν_1 and ν_3 (Figure 5). Although the separation of ν_1 and ν_3 partly accounts for their small intensity, it is unlikely that the experimental ratio would be reached even with the complete overlap. Note that the a:QM/MM spectra incorporating the intermolecular quantum effect shows improved relative intensities, where the ν_1 and ν_3 peaks are higher than those of the c:QM/MM system, in line with experiments (Figure 1 and Supplementary Figure 2). This comparison implies that direct and indirect factors improve the spectrum.

The direct factor is related to mode coupling between the solute and the solvent, where the stretching mode has a positive correlation with its neighbor, while the bending mode is negatively correlated, as reported in the previous study.[21] As a result, the intensities of ν_1 and ν_3 are enhanced, and the ν_2 band is attenuated. These effects are confirmed by the

comparisons among the c:QM/MM spectra, as flexible MM solvents slightly improved the intensities compared to rigid models (Figure 5). Since the intensity is related to the absolute value of the dipole moment, the intensity improvement agrees with the fact that DFTB3 solutes in flexible MM solvents exhibit a dipole moment that is larger by 1% than those in rigid MM solvents (Table 2). Note that the c:QM/MM spectra are evaluated with a dipole moment that consists of only the solute contribution in eqn (14), while the dipole moment for the a:QM/MM system can count the contribution of the surrounding solvent in eqn (15). Thus, the correlation is directly taken into account in the SCMP simulation, leading to better intensity values.

The indirect factor is related to a dipole induced on the solute by the solvent. As shown in Figure 5, the spectra of c:DFTB3/TIP3P-Fs and c:DFTB3/SPC-Fw exhibit two small bands between ν_1 and ν_3. Notably, these peak positions correspond to the stretching modes ν_1 and ν_3 of MM solvent models. As a result, the wells between ν_1 and ν_3 in these spectra appear to be flat compared to those in the rigid solvent simulations (Figure 5). Although in the c:QM/MM simulations a dipole moment does not explicitly count the contribution from MM solvent in eqn (14), as described above, MM solvents induce a dipole moment on a QM solute through its atomic charges, which fluctuates in accord with the MM stretching modes. In the case of the a:QM/MM spectrum, these MM solvent peaks are replaced by those of the QM solvent, which appear at the same positions as those of the QM solute, and lead to the enhancement of the intensities of ν_1 and ν_3. However, note that the ν_1 (ν_3) intensity relative to ν_2 is still smaller than that of experiments, partly because the present spectrum does not have the overtone of ν_2 at around 3300 cm^{-1}, in addition to the separation of ν_1 and ν_3 as described below.

Far infrared region

The liquid water spectrum is known to have a far infrared band characterized by three components at 686, 396, and 200 cm^{-1}, where the first two components originate from librational motion, and the latter is a translational mode accompanied by intermolecular charge fluctuation.[92-94] The SPC-Fw/SPC-Fw spectrum exhibits a broad continuous band of the libration, with a peak at around 580 cm^{-1} and a shoulder at 850 cm^{-1}, while in the TIP3P-Fs/TIP3P-Fs spectrum the librational peak appears at around 550 cm^{-1} and a shoulder at 900 cm^{-1} (see Supplementary Figure 6). Thus, the SPC-Fw model seems to provide a more realistic librational band than the TIP3P-Fs model, which is

reminiscent of the well-reproduced diffusion coefficient of SPC-Fw.[95] On the other hand, in the c:QM/MM spectra (Figure 5) the librational mode with a higher frequency is consistently shifted to ca. 850 cm^{-1} and becomes more distinct, while the other is shifted further to a lower frequency. As a consequence, the librational band forms a plateau that is unlike that in the experimental spectrum. This demonstrates that conventional QM/MM simulation does not necessarily yield better results than MM simulation, in particular with respect to intermolecular interactions between QM and MM molecules.

The a:QM/MM (Figure 6) spectra show a pronounced librational peak at 600 cm^{-1}, while the shoulder at 800 cm^{-1} has a reduced intensity. As a result, although both the peaks are still shifted to a higher frequency, the band shape becomes closer to that observed experimentally.

No band peak is observed for the translational mode at around 200 cm^{-1} in either the MM/MM or c:QM/MM spectra, because the MM solvent model cannot treat intermolecular charge fluctuations. In contrast, although the total molecular charge of the solute averaged over the MD trajectory is zero in the SCMP simulation, the standard deviation of the charge is +0.02 e, which indicates the intermolecular charge fluctuation between the solute and solvents. Correspondingly, the a:QM/MM spectra show the band peak at around 300 cm^{-1}, which is shifted to a higher frequency by 100 cm^{-1} compared to that obtained experimentally.

Thus, we conclude that, although MM solvents can reproduce the solvation effect qualitatively in the far infrared region, the results are limited and unsatisfactory. In the c:QM/MM framework, one of the limitations arises from the fact that the solute–solvent interaction is treated with the MM level to avoid diffusion problems, as mentioned in the first section. The other arises from the dipole moment representation in c:QM/MM that counts only a QM solute contribution. On the other hand, the SCMP method can overcome the limitations and improve the spectra compared to c:QM/MM, although higher QM methods than DFTB3 are required for quantitative reproduction of the peak positions in the far infrared region.

MM contribution

Although the SCMP method largely improves water spectrum, the ν_1 and ν_3 bands still remains separated and lower-frequency shifted compared to the pure DFTB3 spectrum[65]

as mentioned above. Since the band separation is in common with the c:DFTB3/MM spectra, it is attributed to the MM solvent molecules. The attribution is inline with a fact that the band positions of ν_1 and ν_3 in the a:QM/MM spectra are affected by the MM models, TIP3P-Fs and SPC-Fw (see Table 1 and Figure 6).

Figure 7 indicates that the QM profiles of solvent in the SCMP simulation are gradually switched from QM to MM molecules in accordance with their distance from the QM solute. Accordingly, as shown in Figure 8 the radial distribution function (RDF) of the a:DFTB3/SPC-Fw appears to be the admixture of those of the pure DFTB3 and MM models. Since the relation between the solvation structure and the transition parameters was systematically investigated in the previous study[37], we will not detailed it here, but note that the first peak is in good agreement with the pure DFTB3 simulation, while the second peak appears at the distance between those of the pure DFTB3 and SPC-Fw systems. We assume the MM contributions to a:QM/MM spectrum are classified into two types, direct and indirect ones. The direct contribution comes from the fact that the dipole moment represented in eqn (15) counts the contribution from the transition solvents, while the indirect contribution is caused by the MM-like solvation structure and interaction with the transition solvent molecules.

Note that the maximum limit of the indirect contribution should lead to the c:QM/MM spectrum. On the other hand, the understanding of the direct contribution is not straightforward, although eqn (15) indicates each transition solvent contributes to the dipole moment in proportional to its QM profile. In the SCMP method, the effective force acting on a transition solvent molecule is a linear combination of QM and MM forces, while molecular properties such as geometry and oscillator strength may not necessarily be linearly controlled. Indeed, the ν_3 bands of the TIP3P-Fs model in the MM simulations show lower frequency (3675 and 3662 cm^{-1} in the gas and liquid phases, respectively) than those of the SPC-Fw model (3691 and 3686 cm^{-1} in the gas and liquid phases, respectively), as shown in Table 1, while the ν_3 positions are the other way around in SCMP simulations. Therefore, it is difficult to evaluate the effect on the MM model, and further investigations are warranted.

Concluding Remarks

The SCMP method enables a stable and efficient QM/MM-MD simulation without discontinuities. In the present study, we attempted to reproduce the infrared spectrum of water as the first application of the SCMP method. To this end, we also proposed a new representation of the dipole moment based on the SCMP framework to reduce spurious boundary effects and MM solvent effect.

In order to make full use of the advantages of SCMP, we adopted DFTB for the QM method because of its computational efficiency, and recorded 800 ps MD simulations, an outstanding duration for QM/MM simulations with free QM solvent molecules. The long MD simulations reveal that obtaining reliable statistics is a significant limitation for spectrum evaluation, particularly with respect to the far infrared region. Due to its computational efficiency, the SCMP QM/MM simulation will be a practical method for evaluating infrared spectra, particularly for non-uniform systems where there is no sampling advantage of a pure QM and MM simulation. Note that the simulation time for sufficient sampling depends on the system of interest. Thus, even when a pure QM simulation is of interest, the SCMP QM/MM method may become an effective pilot simulation for unknown systems.

On incorporating the quantum effect in solute–solvent interactions, the SCMP QM/MM simulations can yield more realistic spectra than the conventional QM/MM method with respect to the peak shifts, intensities, broadening, and far infrared bands. The DFTB parameters based on each element by counting the electronic structure ensure greater transferability than the MM parameters that are based on each molecule. Nevertheless, differences between the simulated and experimental spectra remain. Since the DFTB parameters were determined with a set of standard molecules to improve the reproduction of various chemical properties, they do not necessarily ensure better results for specific molecules. Improvements in the spectra produced may still be possible with further refinement of the parameters, as demonstrated by a recent study.[65] It should also be noted that the present method contains a large error with regards to the absolute position of the respective band peaks, but the peak shift by solvation can be well reproduced.

Furthermore, although the problem of trade-off between efficiency and accuracy remains, the applications of the SCMP in combination with higher-level QM methods such as DFT and MP2 are also of great interest. Thus, the present study will be a

milestone in the further development and application of adaptive QM/MM methods and infrared spectroscopy.

Table 1. Peak positions (cm^{-1}) of the major three bands estimated with Lorentzian fitting.

State	Method	Solute	Solvent	ν_1	ν_2	ν_3
gas	MM	TIP3P-Fs		3621	1386	3675
		SPC-Fw		3606	1378	3691
	QM	DFTB3		3668	1373	3940
	Experiment [a,b]			3657	1595	3756
liquid	MM/MM	TIP3P-Fs	TIP3P-Fs	3618	1453	3662
		SPC-Fw	SPC-Fw	3618	1454	3686
	c:QM/MM	DFTB3	TIP3P	3461	1453	3745
			SPC	3465	1450	3747
			TIP3P-Fs	3452	1448	3745
			SPC-Fw	3453	1444	3742
	a:QM/MM	DFTB3	TIP3P-Fs	3414	1407	3696
			SCP-Fw	3385	1403	3671
	Experiment [c,d]			3280	1614	3404

[a] Shimanouchi, 1972[96]

[b] Huber et al., 1979[97]

[c] Eisenberg et al. 1969[88]

[d] Bertie et al. 1985[89],

Table 2. Structural properties of a water molecule averaged over the MD trajectory. For the liquid phase, the values are obtained from a "solute" water molecule surrounded by "solvent" water molecules.

State	Method	Solute	Solvent	bond (Å)	angle (degree)	charge [a] (*e*)	dipole [b] (D)
gas	MM	TIP3P		0.9572 [c]	104.52 [c]	− 0.834 [c]	2.310 [c]
		SPC		1.000 [c]	109.47 [c]	− 0.820 [c]	2.274 [c]
		TIP3P-Fs		0.960	104.5	− 0.834 [c]	2.354
		SPC-Fw		1.012	113.2	− 0.820 [c]	2.194
	QM	DFTB2		0.969	107.2	− 0.593	1.635
		DFTB3		0.9552	110.50	− 0.725	1.890
	Experiment			0.957 [d] 0.972 [e]	104.5 [d]		1.855 [f]
liquid	MM/MM	TIP3P-Fs	TIP3P-Fs	0.9785	97.89	− 0.834 [c]	2.568
		SPC-Fw	SPC-Fw	1.0285	107.82	− 0.820 [c]	2.384
	c:QM/MM	DFTB3	TIP3P	0.9762	108.14	− 0.838	2.310
			SPC	0.9762	108.06	− 0.835	2.300
			TIP3P-Fs	0.9768	108.03	− 0.841	2.321
			SPC-Fw	0.9768	108.12	− 0.843	2.323
	a:QM/MM	DFTB3	TIP3P-Fs	0.9778	108.86	− 0.832	
			SPC-Fw	0.9784	108.81	− 0.833	
	Experiment			0.990 [g] 1.01 [h]	106.1 [i]		2.9 [j]

[a] Partial charge at oxygen atom

[b] Dipole is calculated by partial charges on atoms

[c] Fixed values

[d] Hasted, 1972[98]

[e] Cook et al., 1974[99]

[f] Shostak et al., 1991[100]

[g] Soper et al., 2008[101]

[i] Zeidler et al., 2012[102]

[j] Ichikawa et al., 1991[103]

[k] Badyal et al., 2000[104]

Figures

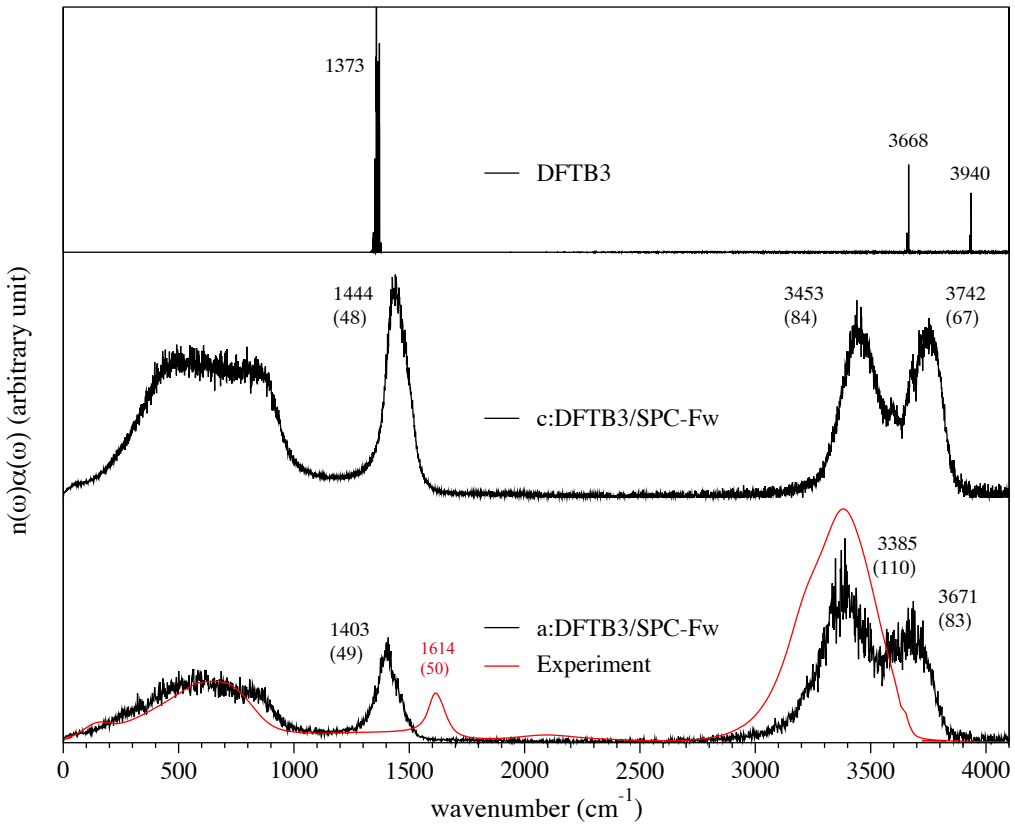

Figure 1. Comparison between the gaseous water spectrum with DFTB3 (top), and liquid water spectra of c:DFTB3/SPC-Fw (middle) and a:DFTB3/SPC-Fw systems (bottom). The experimental spectrum of liquid water is represented by the red line, as a reference. The band peak positions are presented in the figures accompanied by the HWHM values in parentheses. The spectra are represented with a resolution of 2 cm^{-1}.

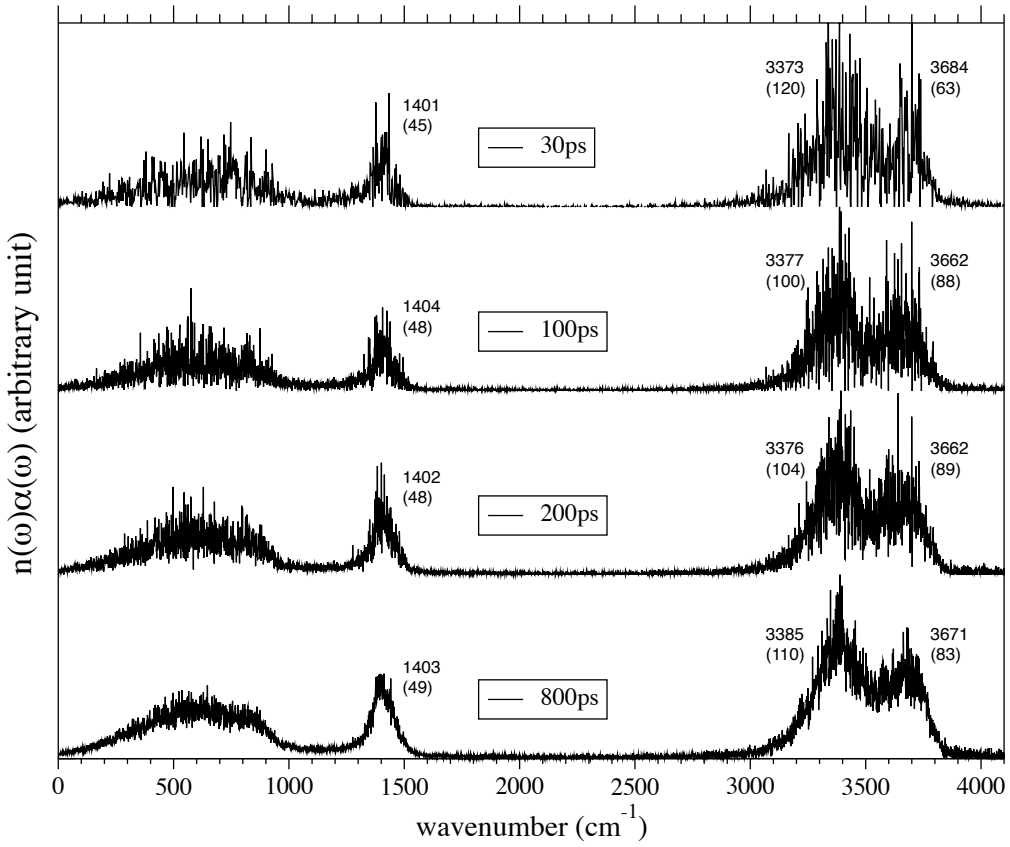

Figure 2. a:DFTB/SPC-Fw spectra evaluated with different simulation durations. The peak positions are presented in the figures accompanied by the HWHM values in parentheses.

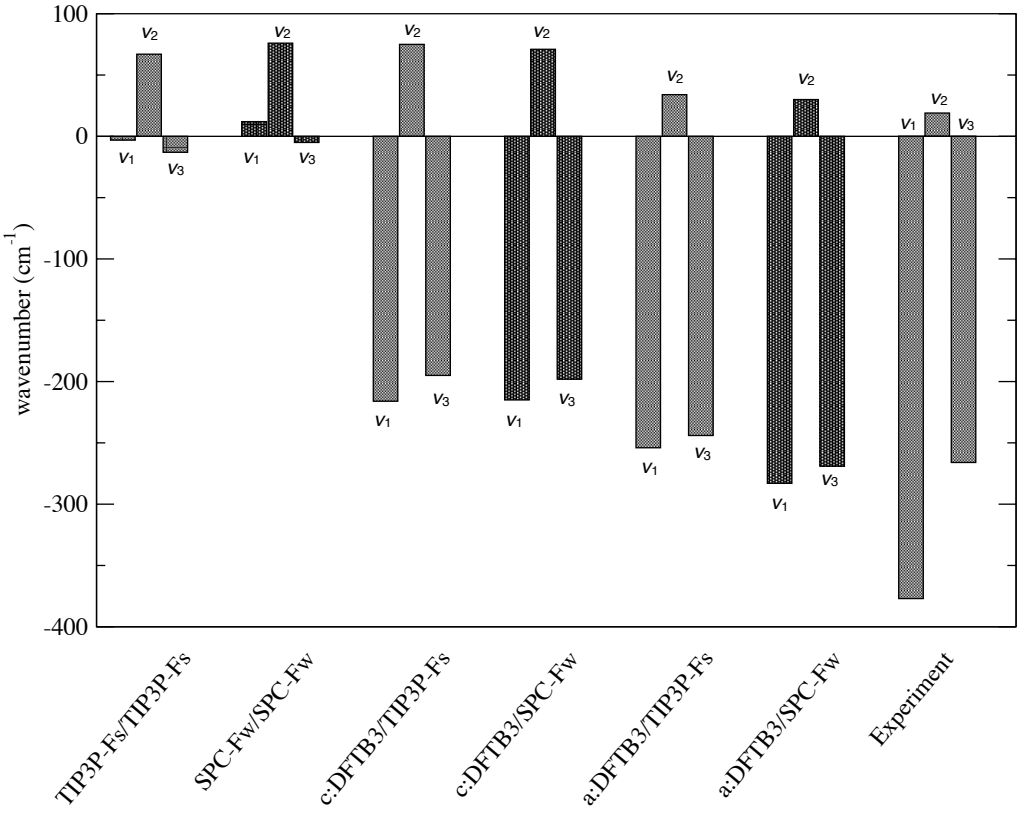

Figure 3. Comparison of the band peak shifts by solvation among different systems, where each reference gas spectrum is evaluated with the identical method applied to the solute in the aqueous calculation. The experimental values are cited from References.[88, 96]

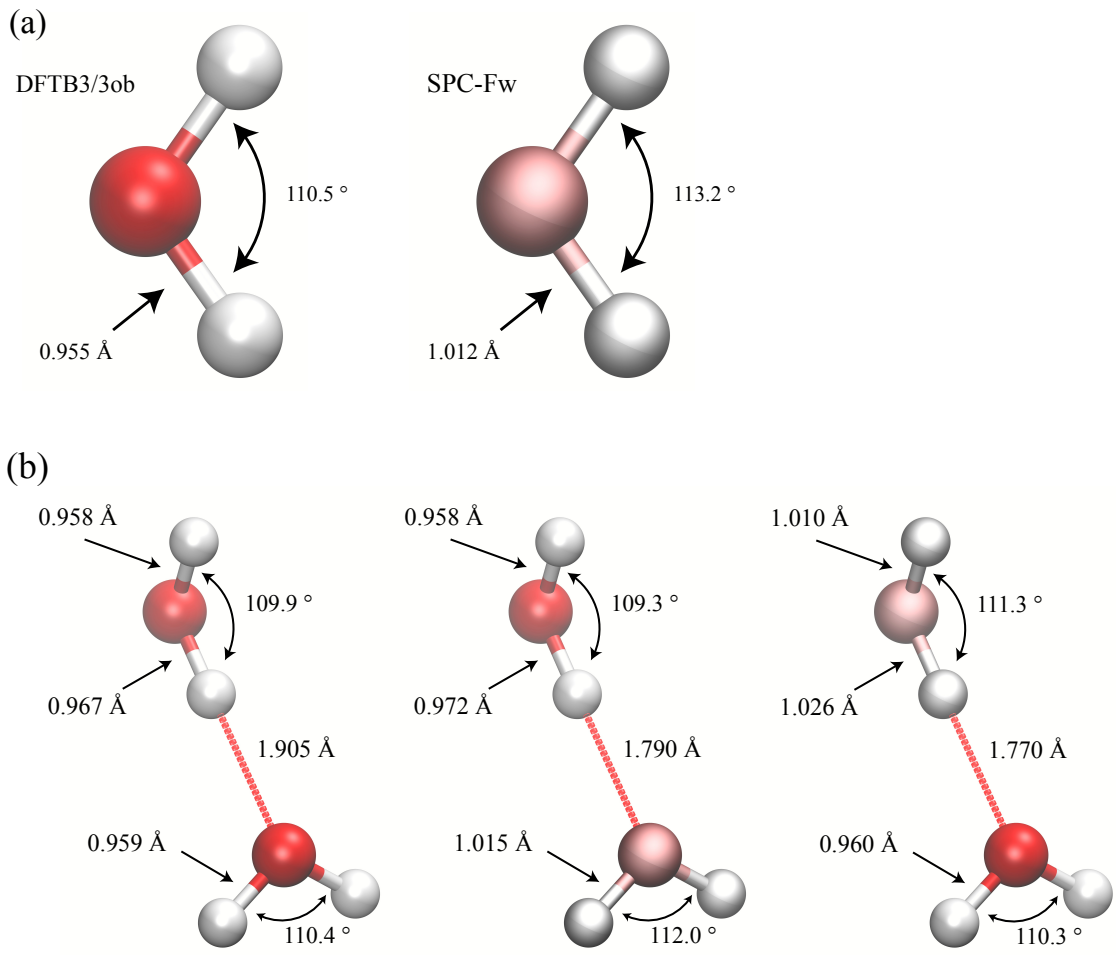

Figure 4. Comparison of water geometries of (a) monomers, and (b) hydrogen-bonded dimers in vacuo. The DFTB3/3ob and SPC-Fw water models are represented as Opaque and Metallic Plastic materials, respectively.

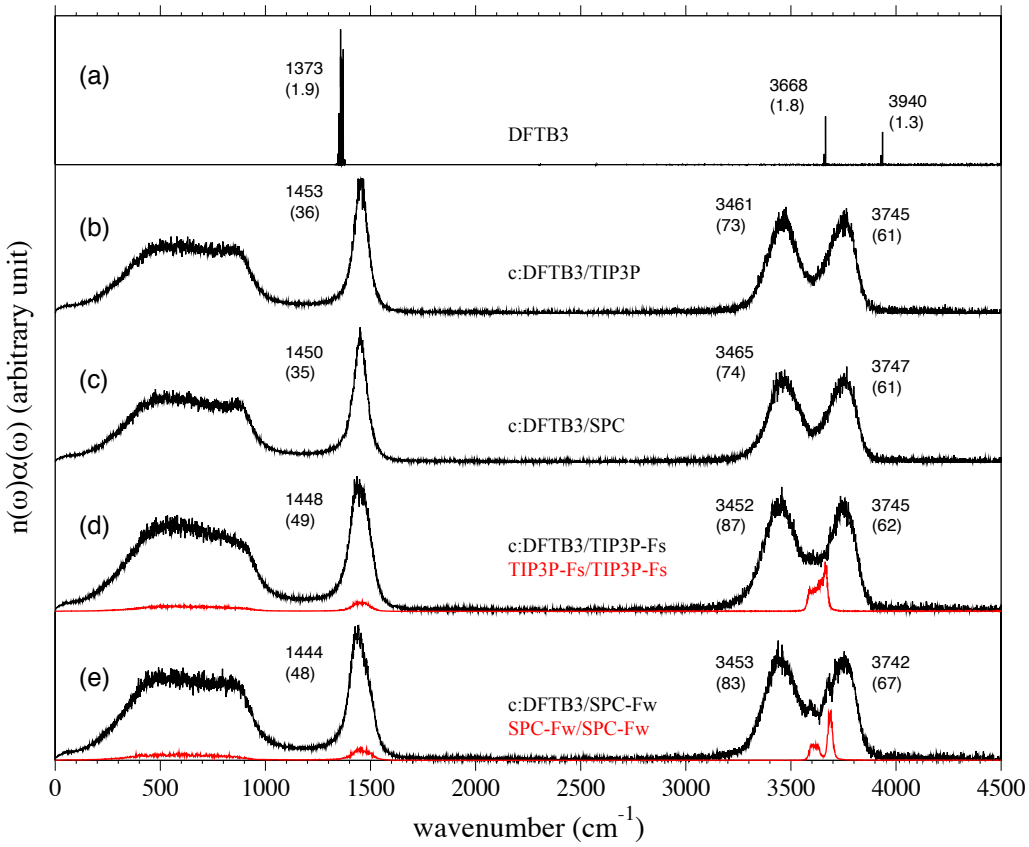

Figure 5. Comparison between (a) the gaseous spectrum with DFTB3 and the aqueous spectra of (b) c:DFTB3/TIP3P, (c) c:DFTB3/SPC, (d) c:DFTB3/TIP3P-Fs, and (e) c:DFTB3/SPC-Fw systems, where the HWHM values of the respective bands are presented in parentheses. As references, TIP3P-Fs/TIP3P-Fs and SPC-Fw/SPC-Fw spectra are represented by red lines. For the sake of clear comparison with QM/MM spectra, MM/MM spectra are scaled down.

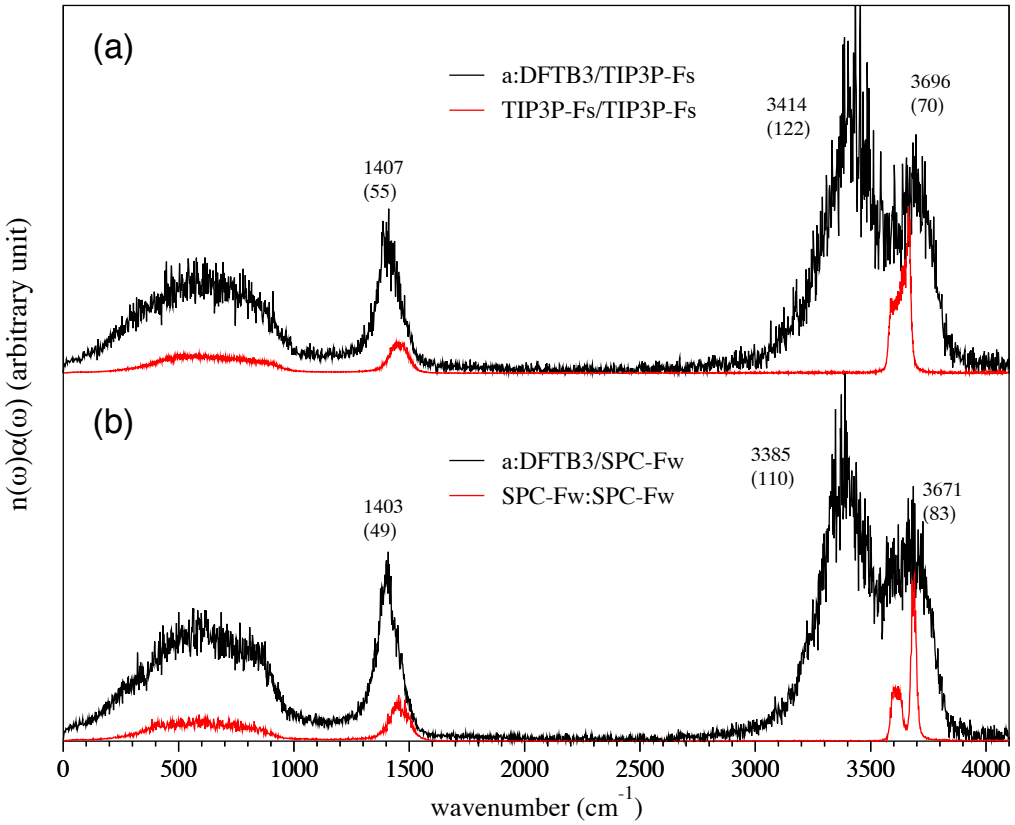

Figure 6. Comparison among aqueous water spectra. (a) The a:DFTB3/TIP3P-Fs system (black), and TIP3P-Fs/TIP3P-Fs system (red) spectra. (b) The a:DFTB3/SPC-Fw system (black) and SPC-Fw/SPC-Fw system (red) spectra. For the sake of clear comparison with QM/MM spectra, MM/MM spectra are scaled down.

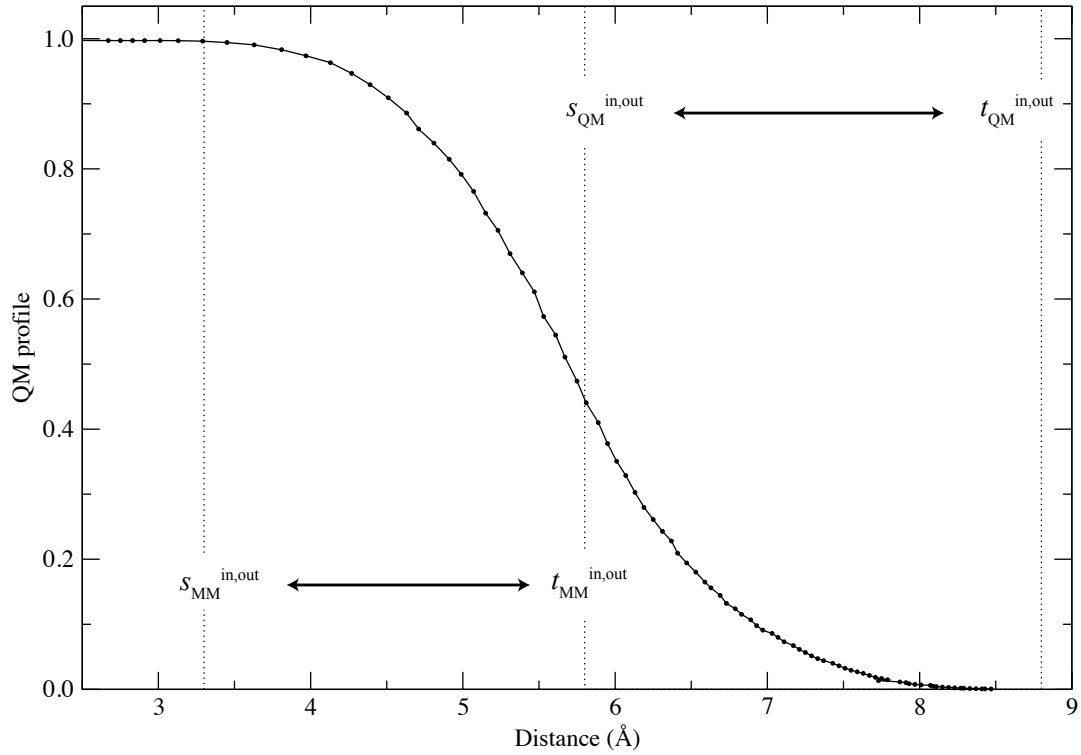

Figure 7. QM profile evaluated from a:DFTB3/SPC-Fw simulation.

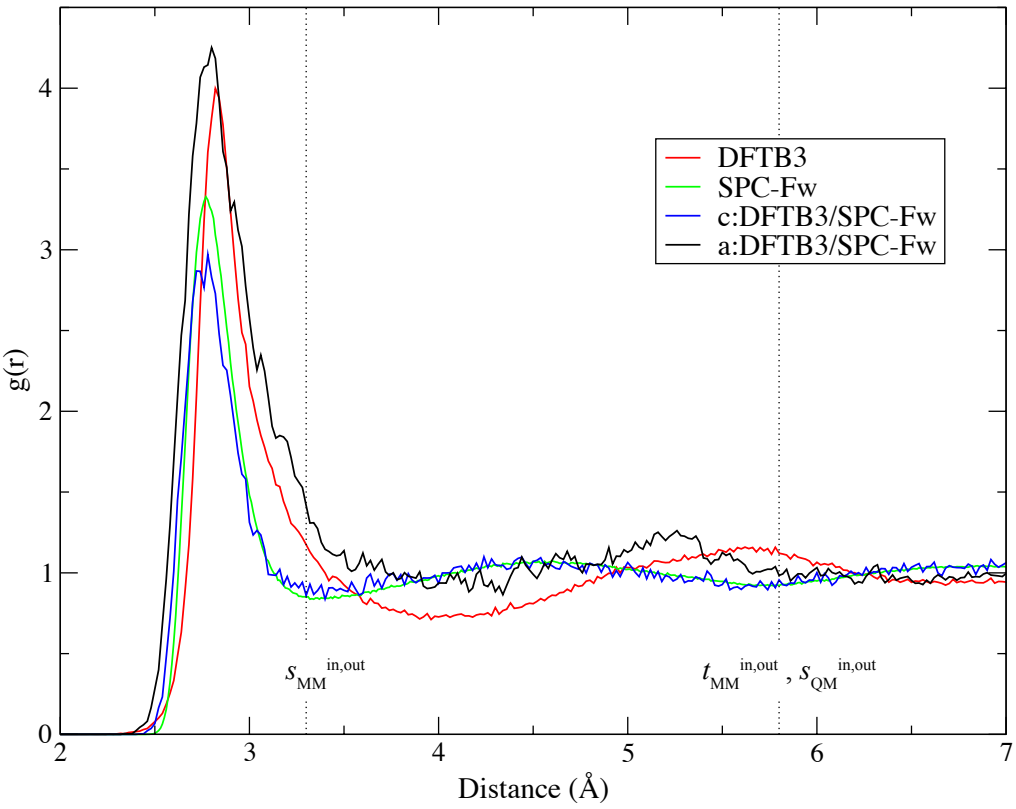

Figure 8. RDF with different approaches: O-O RDF of all oxygen atoms in pure DFTB3 (red), and SPC-Fw water system (green). Oc-O RDF (Oc is central oxyen) of c:DFTB3/SPC-Fw (blue), and a:DFTB3/SPC-Fw (black).

Associated Contents

Supporting Information

Comparison liquid water spectrum between the NVE and NVT simulations, Comparison among MM, c:QM/MM, and a:QM/MM spectra with use of SPC-Fw as MM models, gaseous water spectra in vacuo evaluated with DFTB2 and 3 in NMA and FTTCF approaches, correlation between the OH bond length and the ν_1 frequency, and comparison of far infrared spectra among different simulations. These materials are available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author

*hwatanabe@protein.rcast.u-tokyo.ac.jp

Present Addresses

†Research Center for Advanced Science and Technology, The University of Tokyo, 4-6-1 Komaba, Meguro-ku, Tokyo 153-8904, Japan.

Author Contributions

H.W. designed and coded the size-consistent multipartitioning QM/MM and performed all simulations. The manuscript was written with contributions from all authors. All authors have given approval to the final version of the manuscript.

Funding Sources

Grant-in-aid for JSPS Fellows Grant Number 26-3041 and Research fellowships of the Japan Society for the Promotion of Science to H.C.W.

ACKNOWLEDGMENT

This work was supported by JSPS Grant-in-Aid for JSPS Fellows. The computations were performed at the computer centers of the Tokyo Institute of Technology. The

authors gratefully acknowledge support from Dr. T. Kubař, Dipl. M. Kubillus, and Prof. M. Elstner for helpful discussions about DFTB.

ABBREVIATIONS

NMA, normal mode analysis; INMA, instantaneous normal mode analysis; MD, molecular dynamics; FTTCF, Fourier transform of time-correlation functions; QM, quantum mechanics; MM, molecular mechanics; SCMP, size-consistent multipartitioning; MP2, Møller-Plesset perturbation theory; DFTB, density-functional tight-binding; DFT, density function theory; HWHM, half-width at half-maximum; RDF, Radial distribution function

References

1. R. G. Gordon, *Academic Press*, 1968, **3**, 1.
2. D. A. McQuarrie, *Statistical Mechanics* Harper Collins, New York, 1976.
3. M. Levitt, C. Sander and P. S. Stern, *Int. J. Quantum Chem*, 1983, 181-199.
4. N. Go, T. Noguti and T. Nishikawa, *P Natl Acad Sci-Biol*, 1983, **80**, 3696-3700.
5. B. Brooks and M. Karplus, *P Natl Acad Sci-Biol*, 1983, **80**, 6571-6575.
6. J. P. Merrick, D. Moran and L. Radom, *J. Phys. Chem. A*, 2007, **111**, 11683-11700.
7. B. Temelso and G. C. Shields, *J. Chem. Theory Comput.*, 2011, **7**, 2804-2817.
8. M. Buchner, B. M. Ladanyi and R. M. Stratt, *J. Chem. Phys.*, 1992, **97**, 8522-8535.
9. M. Cho, G. R. Fleming, S. Saito, I. Ohmine and R. M. Stratt, *J. Chem. Phys.*, 1994, **100**, 6672-6683.
10. M. Nonella, G. Mathias and P. Tavan, *J. Phys. Chem. A*, 2003, **107**, 8638-8647.
11. Q. Cui, G. H. Li, J. P. Ma and M. Karplus, *J. Mol. Biol.*, 2004, **340**, 345-372.
12. A. Bastida, M. A. Soler, J. Zuniga, A. Requena, A. Kalstein and S. Fernandez-Alberti, *J. Phys. Chem. A*, 2010, **114**, 11450-11461.
13. A. Bastida, M. A. Soler, J. Zuniga, A. Requena, A. Kalstein and S. Fernandez-Alberti, *J. Chem. Phys.*, 2010, **132**.
14. M. Schmitz and P. Tavan, *J. Chem. Phys.*, 2004, **121**, 12247-12258.
15. R. Ramirez, T. Lopez-Ciudad, P. Kumar and D. Marx, *J. Chem. Phys.*, 2004, **121**, 3973-3983.
16. A. Warshel and M. Levitt, *J. Mol. Biol.*, 1976, **103**, 227-249.
17. Q. Cui and M. Karplus, *J. Chem. Phys.*, 2000, **112**, 1133-1149.
18. Q. Cui and M. Karplus, *J. Phys. Chem. B*, 2000, **104**, 3721-3743.
19. M. S. Lee, F. Baletto, D. G. Kanhere and S. Scandolo, *J. Chem. Phys.*, 2008, **128**, 214506.
20. C. J. Burnham, D. J. Anick, P. K. Mankoo and G. F. Reiter, *J. Chem. Phys.*, 2008, **128**, 154519.
21. W. Chen, M. Sharma, R. Resta, G. Galli and R. Car, *Phys Rev B*, 2008, **77**, 245114.

22. A. Shank, Y. Wang, A. Kaledin, B. J. Braams and J. M. Bowman, *J. Chem. Phys.*, 2009, **130**, 144314.
23. S. Habershon, G. S. Fanourgakis and D. E. Manolopoulos, *J. Chem. Phys.*, 2008, **129**, 074501.
24. W. Kulig and N. Agmon, *Nat Chem*, 2013, **5**, 29-35.
25. Y. Nishimura, Y. P. Lee, S. Irle and H. A. Witek, *J. Chem. Phys.*, 2014, **141**, 094303.
26. A. Warshel, M. Kato and A. V. Pisliakov, *J. Chem. Theory Comput.*, 2007, **3**, 2034-2045.
27. T. Ishiyama and A. Morita, *J. Chem. Phys.*, 2009, **131**, 244714.
28. L. P. Wang, T. Head-Gordon, J. W. Ponder, P. Ren, J. D. Chodera, P. K. Eastman, T. J. Martinez and V. S. Pande, *J. Phys. Chem. B*, 2013, **117**, 9956-9972.
29. G. R. Medders and F. Paesani, *J. Chem. Theory Comput.*, 2015, **11**, 1145-1154.
30. K. Welke, H. C. Watanabe, T. Wolter, M. Gaus and M. Elstner, *Phys. Chem. Chem. Phys.*, 2013, **15**, 6651-6659.
31. M. J. Field, P. A. Bash and M. Karplus, *J. Comput. Chem.*, 1990, **11**, 700-733.
32. P. Phatak, N. Ghosh, H. B. Yu, Q. Cui and M. Elstner, *Proc. Natl. Acad. Sci. U. S. A.*, 2008, **105**, 19672-19677.
33. P. Phatak, J. S. Frahmcke, M. Wanko, M. Hoffmann, P. Strodel, J. C. Smith, S. Suhai, A. N. Bondar and M. Elstner, *J. Am. Chem. Soc.*, 2009, **131**, 7064-7078.
34. J. S. Frahmcke, M. Wanko, P. Phatak, M. A. Mroginski and M. Elstner, *J. Phys. Chem. B*, 2010, **114**, 11338-11352.
35. P. Goyal, N. Ghosh, P. Phatak, M. Clemens, M. Gaus, M. Elstner and Q. Cui, *J. Am. Chem. Soc.*, 2011, **133**, 14981-14997.
36. S. Wolf, E. Freier, Q. Cui and K. Gerwert, *J. Chem. Phys.*, 2014, **141**.
37. H. C. Watanabe, T. Kubar and M. Elstner, *J. Chem. Theory Comput.*, 2014, **10**, 4242-4252.
38. T. Kerdcharoen, K. R. Liedl and B. M. Rode, *Chem. Phys.*, 1996, **211**, 313-323.
39. T. Kerdcharoen and K. Morokuma, *Chem. Phys. Lett.*, 2002, **355**, 257-262.
40. C. F. Schwenk, H. H. Loeffler and B. M. Rode, *J. Am. Chem. Soc.*, 2003, **125**, 1618-1624.
41. R. E. Bulo, B. Ensing, J. Sikkema and L. Visscher, *J. Chem. Theory Comput.*,

2009, **5**, 2212-2221.

42. A. Heyden, H. Lin and D. G. Truhlar, *J. Phys. Chem. B*, 2007, **111**, 2231-2241.
43. N. Takenaka, Y. Kitamura, Y. Koyano and M. Nagaoka, *Chem. Phys. Lett.*, 2012, **524**, 56-61.
44. N. Takenaka, Y. Kitamura, Y. Koyano and M. Nagaoka, *J. Chem. Phys.*, 2012, **137**, 024501.
45. K. Park, A. W. Gotz, R. C. Walker and F. Paesani, *J. Chem. Theory Comput.*, 2012, **8**, 2868-2877.
46. R. E. Bulo, C. Michel, P. Fleurat-Lessard and P. Sautet, *J. Chem. Theory Comput.*, 2013, **9**, 5567-5577.
47. C. N. Rowley and B. Roux, *J. Chem. Theory Comput.*, 2012, **8**, 3526-3535.
48. M. Shiga and M. Masia, *J. Chem. Phys.*, 2013, **139**, 044120.
49. M. Del Ben, O. Schuett, T. Wentz, P. Messmer, J. Hutter and J. VandeVondele, *Comput. Phys. Commun.*, 2015, **187**, 120-129.
50. M. Del Ben, J. Hutter and J. VandeVondele, *J. Chem. Phys.*, 2015, **143**, 054506.
51. P. Y. Ren and J. W. Ponder, *J. Phys. Chem. B*, 2003, **107**, 5933-5947.
52. G. S. Fanourgakis and S. S. Xantheas, *J. Chem. Phys.*, 2008, **128**, 074506.
53. C. J. Burnham, J. C. Li, S. S. Xantheas and M. Leslie, *J. Chem. Phys.*, 1999, **110**, 4566-4581.
54. V. Babin, G. R. Medders and F. Paesani, *J Phys Chem Lett*, 2012, **3**, 3765-3769.
55. G. R. Medders, V. Babin and F. Paesani, *J. Chem. Theory Comput.*, 2014, **10**, 2906-2910.
56. M. Elstner and G. Seifert, *Philos T R Soc A*, 2014, **372**, 20120483.
57. M. Gaus, Q. Cui and M. Elstner, *Wires Comput Mol Sci*, 2014, **4**, 49-61.
58. R. B. Liang, J. M. J. Swanson and G. A. Voth, *J. Chem. Theory Comput.*, 2014, **10**, 451-462.
59. M. Elstner, T. Frauenheim and S. Suhai, *J Mol Struc-Theochem*, 2003, **632**, 29-41.
60. M. Elstner, *J. Phys. Chem. A*, 2007, **111**, 5614-5621.
61. C. M. Maupin, B. Aradi and G. A. Voth, *J. Phys. Chem. B*, 2010, **114**, 6922-6931.
62. M. Gaus, Q. A. Cui and M. Elstner, *J. Chem. Theory Comput.*, 2011, **7**, 931-948.
63. M. Gaus, A. Goez and M. Elstner, *J. Chem. Theory Comput.*, 2013, **9**, 338-354.

64. M. Gaus, X. Y. Lu, M. Elstner and Q. Cui, *J. Chem. Theory Comput.*, 2014, **10**, 1518-1537.
65. P. Goyal, H. J. Qian, S. Irle, X. Y. Lu, D. Roston, T. Mori, M. Elstner and Q. Cui, *J. Phys. Chem. B*, 2014, **118**, 11007-11027.
66. M. Kubillus, T. Kubar, M. Gaus, J. Rezac and M. Elstner, *J. Chem. Theory Comput.*, 2015, **11**, 332-342.
67. P. Goyal, M. Elstner and Q. Cui, *J. Phys. Chem. B*, 2011, **115**, 6790-6805.
68. T. H. Choi, R. Liang, C. M. Maupin and G. A. Voth, *J. Phys. Chem. B*, 2013, **117**, 5165-5179.
69. A. N. Bondar, M. Elstner, S. Suhai, J. C. Smith and S. Fischer, *Structure*, 2004, **12**, 1281-1288.
70. S. Irle, G. S. Zheng, Z. Wang and K. Morokuma, *J. Phys. Chem. B*, 2006, **110**, 14531-14545.
71. P. Goyal, S. Yang and Q. Cui, *Chem Sci*, 2015, **6**, 826-841.
72. S. Fritsch, R. Potestio, D. Donadio and K. Kremer, *J. Chem. Theory Comput.*, 2014, **10**, 816-824.
73. M. Elstner, D. Porezag, G. Jungnickel, J. Elsner, M. Haugk, T. Frauenheim, S. Suhai and G. Seifert, *Phys Rev B*, 1998, **58**, 7260-7268.
74. B. Hess, C. Kutzner, D. van der Spoel and E. Lindahl, *J. Chem. Theory Comput.*, 2008, **4**, 435-447.
75. S. Pronk, S. Pall, R. Schulz, P. Larsson, P. Bjelkmar, R. Apostolov, M. R. Shirts, J. C. Smith, P. M. Kasson, D. van der Spoel, B. Hess and E. Lindahl, *Bioinformatics*, 2013, **29**, 845-854.
76. T. Kubar, K. Welke and G. Groenhof, *J. Comput. Chem.*, 2015, **36**, 1978-1989.
77. T. Darden, D. York and L. Pedersen, *J. Chem. Phys.*, 1993, **98**, 10089-10092.
78. T. Kruger, M. Elstner, P. Schiffels and T. Frauenheim, *J. Chem. Phys.*, 2005, **122**, 114110.
79. K. W. Sattelmeyer, J. Tirado-Rives and W. L. Jorgensen, *J. Phys. Chem. A*, 2006, **110**, 13551-13559.
80. N. Otte, M. Scholten and W. Thiel, *J. Phys. Chem. A*, 2007, **111**, 5751-5755.
81. W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey and M. L. Klein, *J. Chem. Phys.*, 1983, **79**, 926-935.
82. H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren and J. Hermans,

Interaction models for water in relation to protein hydration, D. Reidel Publishing Company, 1981.

83. L. X. Dang and B. M. Pettitt, *J. Phys. Chem.*, 1987, **91**, 3349-3354.
84. U. W. Schmitt and G. A. Voth, *J. Chem. Phys.*, 1999, **111**, 9361-9381.
85. Y. J. Wu, H. L. Tepper and G. A. Voth, *J. Chem. Phys.*, 2006, **124**, 024503.
86. S. Nose, *J. Chem. Phys.*, 1984, **81**, 511-519.
87. W. G. Hoover and B. L. Holian, *Phys. Lett. A*, 1996, **211**, 253-257.
88. D. S. Eisenberg and W. Kauzmann, *The structure and properties of water*, Clarendon Press Oxford, 1969.
89. J. E. Bertie and H. H. Eysel, *Appl. Spectrosc.*, 1985, **39**, 392-401.
90. Y. Bouteiller and J. P. Perchard, *Chem. Phys.*, 2004, **305**, 1-12.
91. H. B. Yu and Q. Cui, *J. Chem. Phys.*, 2007, **127**, 234504.
92. B. Guillot, *J. Chem. Phys.*, 1991, **95**, 1543-1551.
93. H. R. Zelsmann, *J. Mol. Struct.*, 1995, **350**, 95-114.
94. H. Torii, *J. Chem. Theory Comput.*, 2014, **10**, 1219-1227.
95. Y. J. Wu, H. L. Tepper and G. A. Voth, *J. Chem. Phys.*, 2006, **124**, 184507.
96. T. Shimanouchi, *Journal*, 1972.
97. K. P. Huber and G. Herzberg, *Molecular Spectra and Molecular Structure. IV. Constants Of Diatomic Molecules*, Van Nostrand Reinhold Inc., 1979.
98. J. B. Hasted, *Liquid water: Dielectric properties*, Springer New York, 1972.
99. R. L. Cook, F. C. Delucia and Helminge.P, *J. Mol. Spectrosc.*, 1974, **53**, 62-76.
100. S. L. Shostak, W. L. Ebenstein and J. S. Muenter, *J. Chem. Phys.*, 1991, **94**, 5875-5882.
101. A. K. Soper and C. J. Benmore, *Phys. Rev. Lett.*, 2008, **101**, 065502.
102. A. Zeidler, P. S. Salmon, H. E. Fischer, J. C. Neuefeind, J. M. Simonson and T. E. Markland, *J Phys-Condens Mat*, 2012, **24**, 284126.
103. K. Ichikawa, Y. Kameda, T. Yamaguchi, H. Wakita and M. Misawa, *Mol. Phys.*, 1991, **73**, 79-86.
104. Y. S. Badyal, M. L. Saboungi, D. L. Price, S. D. Shastri, D. R. Haeffner and A. K. Soper, *J. Chem. Phys.*, 2000, **112**, 9206-9208.