SCIENTIFIC REP⚙RTS

OPEN

Ab initio molecular dynamics of liquid water using embedded-fragment second-order many-body perturbation theory towards its accurate property prediction

Soohaeng Yoo Willow[1,2,3], Michael A. Salim[1], Kwang S. Kim[2] & So Hirata[1,3]

A direct, simultaneous calculation of properties of a liquid using an *ab initio* electron-correlated theory has long been unthinkable. Here we present structural, dynamical, and response properties of liquid water calculated by *ab initio* molecular dynamics using the embedded-fragment spin-component-scaled second-order many-body perturbation method with the aug-cc-pVDZ basis set. This level of theory is chosen as it accurately and inexpensively reproduces the water dimer potential energy surface from the coupled-cluster singles, doubles, and noniterative triples with the aug-cc-pVQZ basis set, which is nearly exact. The calculated radial distribution function, self-diffusion coefficient, coordinate number, and dipole moment, as well as the infrared and Raman spectra are in excellent agreement with experimental results. The shapes and widths of the OH stretching bands in the infrared and Raman spectra and their isotropic-anisotropic Raman noncoincidence, which reflect the diverse local hydrogen-bond environment, are also reproduced computationally. The simulation also reveals intriguing dynamic features of the environment, which are difficult to probe experimentally, such as a surprisingly large fluctuation in the coordination number and the detailed mechanism by which the hydrogen donating water molecules move across the first and second shells, thereby causing this fluctuation.

Nearly all previous atomistic simulations of liquids have been performed with molecular mechanics (MM) force fields[1–9] or with forces supplied by on-the-fly density-functional theory (DFT) calculations[10–13] or by quantum mechanical/molecular mechanical (QM/MM) calculations[14]. MM is limited to systems that are already well characterized, since force parameterization requires fitting to such *a priori* information. DFT is a great improvement over MM because the former can be directly applied to any liquid systems. Yet, many of DFT functionals lack intermolecular dispersion interactions, despite their importance as a cohesive force in many liquids, or otherwise do not have sufficiently high accuracy. In fact, DFT simulations of bulk water have predicted a considerably elevated melting temperature ($T_m \approx$ 400 K)[15] and thus the existence of its liquid phase in rather unrealistic conditions[12]. Although *ad hoc* dispersion corrections have been shown to improve the description of liquid water[16,17], such improvements are generally nonsystematic, and in these cases DFT falls short of being a predictive theory. One is thus interested in applying systematically improvable theories, such as *ab initio* many-body perturbation (MP)

[1]Department of Chemistry, University of Illinois at Urbana-Champaign, 600 South Mathews Avenue, Urbana, Illinois 61801, USA. [2]Center for Superfunctional Materials, Department of Chemistry, Ulsan National Institute of Science and Technology (UNIST), Ulsan 689-798, Korea. [3]CREST, Japan Science and Technology Agency, 4-1-8 Honcho, Kawaguchi, Saitama 332-0012, Japan. Correspondence and requests for materials should be addressed to K.S.K. (email: kimks@unist.ac.kr) or S.H. (email: sohirata@illinois.edu)

and coupled-cluster (CC) theories, to the simulation of nature's most important liquid, i.e., water, with the eventual goal of reproducing all its observed structural, dynamical, electronic, and spectroscopic features simultaneously, accurately, and from first principles.

An application of *ab initio* MP or CC theory to a bulk liquid has not been possible because of its prohibitive cost. The situation has changed with the advent of the so-called embedded-fragment method[18], which enables a systematic, fast, and parallel-executable application of an *ab initio* theory to large molecular clusters, crystals, and even liquids[19–21]. The method divides a weakly interacting system into overlapping molecular dimers ("fragments"), which are embedded in the electrostatic environment of the whole system, and then applies well-developed molecular theories and software to these fragments to reconstruct an array of whole-system properties, with a computational cost scalable with respect to both system and computer sizes.

Fujita *et al.*[19] reported the first *ab initio* Born–Oppenheimer molecular dynamics (BOMD) simulation of liquid water with the embedded-fragment scheme known as the fragment-molecular-orbital (FMO) method[22]. They employed Hartree–Fock (HF) theory and the 6-31G* basis set as the *ab initio* theory, which does not describe electron correlation and thus lacks the dispersion interaction. The effect of three-body interactions was assessed by Komeiji *et al.*[23] in a water cluster of $(H_2O)_{32}$ and that of electron correlation by Mochizuki *et al.*[24] using second-order many-body perturbation (MP2) theory and the 6-31G* basis set on $(H_2O)_{64}$. Brorsen *et al.*[20,21] repeated FMO-BOMD calculations of liquid water using more rigorous formulas for atomic forces at the HF and DFT level. All of these studies considered only the radial distribution functions and, owing to the smallness of the basis set employed, it is doubtful that the calculated results have much quantitative value. Perhaps the largest-scale electron-correlated calculation of liquid water performed thus far is an isobaric-isothermal Monte Carlo (MC) simulation at the MP2 level using a mixed Gaussian-planewave basis set[25]. Under the ambient conditions, it predicted a reasonable density of 1.02 g/cm³, but did not provide any other meaningful experimental comparison such as detailed structures, diffusion rate, vibrational spectra, or dynamics of water molecule.

In the present work, we perform BOMD simulations of liquid water using on-the-fly atomic forces obtained with thus-far the most accurate electronic structure theory for such simulations, i.e., the *ab initio* spin-component-scaled MP2 (SCS-MP2) method[26] with the aug-cc-pVDZ basis set. The necessary speedup of the *ab initio* calculation is achieved by the embedded-fragment method known as the binary-interaction method (BIM)[27] specifically extended for condensed-phase simulations. The choice of the theoretical level is based on our observation that it accurately reproduces the water-water interaction potentials of the coupled-cluster singles, doubles, and noniterative triples [CCSD(T)] with the aug-cc-pVQZ basis set, which is nearly exact. This theory is, therefore, expected to achieve an equally accurate quantum electronic description of the structural, dynamical, spectroscopic, and electronic properties of liquid water.

We apply this method to the radial distribution function, self-diffusion coefficient, monomer structure and its variation, coordination number and its fluctuation, intershell exchange mechanism, dipole moment and its distribution, as well as infrared (IR) and Raman spectra of liquid water (neat liquid H_2O) under periodic boundary condition. We show that the simulations at the temperature (T) of 250 K and the density (ρ) of 1 g/cm³ predict the aforementioned properties in good agreement with those observed at ambient conditions. The results explain the shapes and widths of the OH stretching bands in the IR and Raman spectra, including the difference between the isotropic and anisotropic Raman components, which is known to reflect the different local hydrogen-bond environments surrounding the OH oscillators. They furthermore reveal intriguing dynamical features of the system, which are difficult to probe experimentally, such as a large fluctuation of the coordination number and the mechanism and time frame in which water molecules move across the first to second shell and thereby cause this fluctuation. They also inform us about the dipole moment and its distribution in the liquid phase as well as its breakdown into the permanent and induced components.

The properties calculated in this work range from the structural to dynamical to spectroscopic to electronic properties, which have hardly been studied simultaneously and with reasonable accuracy by any previously existing simulation methods. Therefore, the present work can provide a useful benchmark against which some controversial issues on collective properties of liquid water can be reexamined.

Results

Potential energy surface. The fidelity of BOMD simulations for bulk systems largely hinges on the reliability of the potential energy surface (PES) obtained from the electronic structure method, and hence verifying the latter for intermolecular interactions is a crucial prerequisite for such simulations. Figure 1 plots the PES's of the water dimer in three geometrical configurations (having C_s, C_i, or C_{2v} symmetry) as a function of the O–O separation (R_{OO}). The binding energy (ΔE) and equilibrium O–O separation ($R_{OO}^{min.}$) of the optimized C_s structure are listed in Supplementary Table S1. Compared with the accurate reference data of MP2/aug-cc-pVQZ or CCSD(T)/aug-cc-pVQZ, the MP2/cc-pVDZ (top panel of Fig. 1) overestimates the intermolecular attractions by as much as 50%, whereas the PES's of MP2/aug-cc-pVDZ (middle panel) and SCS-MP2/aug-cc-pVDZ (bottom panel) are reasonable. This underscores the importance of diffuse basis functions in describing intermolecular forces; namely, basis sets without diffuse basis functions such as 6-31G* are inadequate for BOMD simulations of liquid water and water
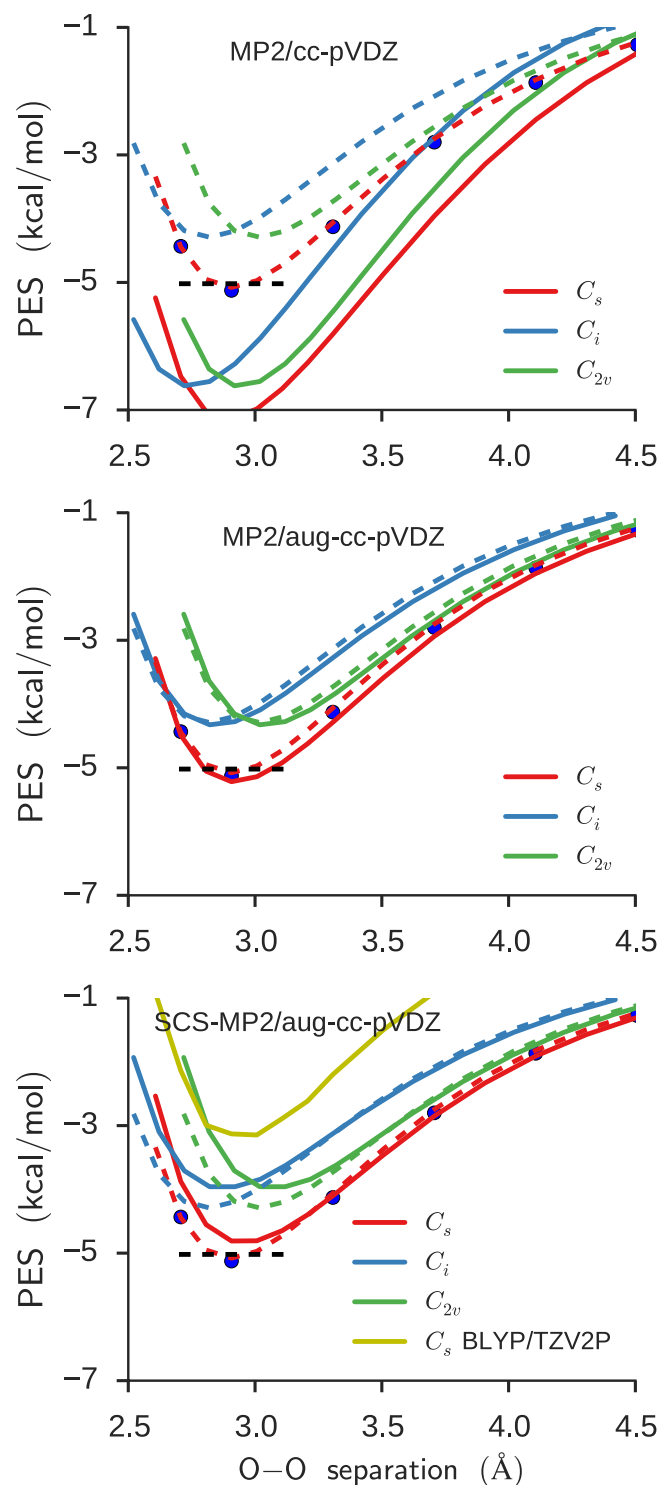
Figure 1. Potential energy surfaces (PES's). The PES's of water dimers in the C_s, C_i, and C_{2v} geometries as a function of the oxygen-oxygen separation (R_{OO}) were obtained at various theoretical levels (solid curves). The dimer geometries were generated by varying R_{OO}, with all other geometrical parameters frozen at their MP2/aug-cc-pVTZ optimized values under the appropriate symmetry constraint. The accurate results from MP2/aug-cc-pVQZ (dashed curves) and CCSD(T)/aug-cc-pVQZ (dots) are superimposed. The binding energy at CCSD(T)/CBS is indicated as dashed black lines.

clusters[19,24]. On the other hand, DFT calculations with the Becke–Lee–Yang–Parr (BLYP) functional and diffuse basis sets tend to underestimate ΔE considerably and are also inadequate (bottom panel and Supplementary Table S1).
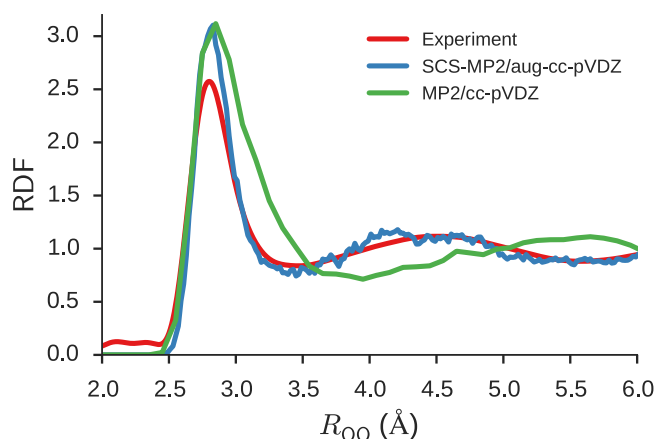
Figure 2. Oxygen-oxygen radial distribution function. The oxygen-oxygen radial distribution functions (RDF's) of liquid water were obtained from the BOMD simulations at $T = 250\,K$ and $\rho = 1\,g/cm^3$ and were compared with the experimental result[34] at $T = 25\,°C$.

SCS-MP2/aug-cc-pVDZ (bottom panel) describes the attractive tail of the PES slightly better than MP2/aug-cc-pVDZ. Its repulsive parts are a little too strong, making the PES's somewhat shallower and equilibrium R_{OO} longer than the reference data. Nonetheless, the value of ΔE predicted by SCS-MP2/aug-cc-pVDZ (4.86 kcal/mol) is closer to the complete-basis-set (CBS) limit of CCSD(T) of 5.02 kcal/mol (dashed lines) than the value obtained at MP2/aug-cc-pVDZ (5.26 kcal/mol), while the MP2-calculated ΔE seems to converge more accurately at the correct limit with a further basis-set extension[28–30].

Note that the calculated values of ΔE and $R_{OO}^{min.}$ in Supplementary Table S1 do not include the zero-point vibrational energy (ZPVE) corrections or nuclear quantum corrections, which have the effect of decreasing ΔE and increasing $R_{OO}^{min.}$ by 0.032 Å[31]. Since such corrections are not at present considered in our BOMD simulations, we find SCS-MP2/aug-cc-pVDZ convenient because its ΔE (4.86 kcal/mol) is smaller and its $R_{OO}^{min.}$ (2.942 Å) is longer than those of CCSD(T)/CBS (5.02 kcal/mol and 2.910 Å) and thus expected to be closer to experimental values that include the ZPVE effect[32,33]. Needless to say, the use of a large basis set such as aug-cc-pVQZ in *ab initio* electron-correlated BOMD simulations takes a prohibitive computational cost and is out of the question.

In this article, therefore, we primarily used SCS-MP2/aug-cc-pVDZ to generate atomic forces. We additionally used MP2/cc-pVDZ as the secondary method to assess the effect of overestimated ΔE and underestimated $R_{OO}^{min.}$ on the simulation results.

Radial distribution function. No *a priori* information exists about the phase diagram of water for the *ab initio* potentials. The first step is, therefore, to identify where the liquid phase boundaries exist. One approach is to determine the melting temperature computationally[15,16]. Another is to vary the temperature such that the BOMD simulation reproduces the experimental liquid structure at ambient conditions[12]. The liquid structure is inferred from the radial distribution function (RDF), which can be observed by neutron or X-ray scattering experiment[34,35] and serves as a fingerprint of local structures and thus phases. We first performed BOMD simulations at $T = 300\,K$, which produced a gas-phase-like RDF, and then reduced the temperature until the calculated RDF coincided with that of the observed result for the liquid phase at ambient conditions.

Figure 2 shows the oxygen-oxygen RDF obtained from the SCS-MP2/aug-cc-pVDZ simulation at $T = 250\,K$. It reproduces the experimental RDF of the liquid phase observed at $T = 25\,°C$ obtained from the X-ray diffraction technique[34]. The calculated position of the first peak is close to the experimental value (~2.8 Å) and the agreement of RDF in the region of $R_{OO} > 3.2$ Å is striking. The slight difference in the calculated peak intensity from the experimental one at $R_{OO} = 2.8$ Å can be diminished by taking into account the nuclear quantum effect, as noted from the previous path-integral molecular dynamics simulation[13]. The average coordination numbers of 4.7 ($R_{OO} \leq 3.4$ Å) or 4.4 ($R_{OO} \leq 3.3$ Å) are in excellent agreement with the experimental values[34]. We therefore conclude that the SCS-MP2/aug-cc-pVDZ simulation at $T = 250\,K$ and $\rho = 1\,g/cm^3$ provides an accurate description of the structural features of liquid water at ambient conditions.

In contrast, the RDF of MP2/cc-pVDZ at $T = 250\,K$ exhibits an excessively broad first peak, which is a manifestation of an over-coordinated local structure with an average coordination number of as large as 8 ($R_{OO} < 3.7$ Å). This, in turn, is due to too strong an attractive intermolecular interaction predicted by MP2/cc-pVDZ, as evidenced by Fig. 1. This first peak is followed by the second peak, which is further displaced at longer R_{OO}, indicating that the second shell is pushed outward by the overcrowded first shell.
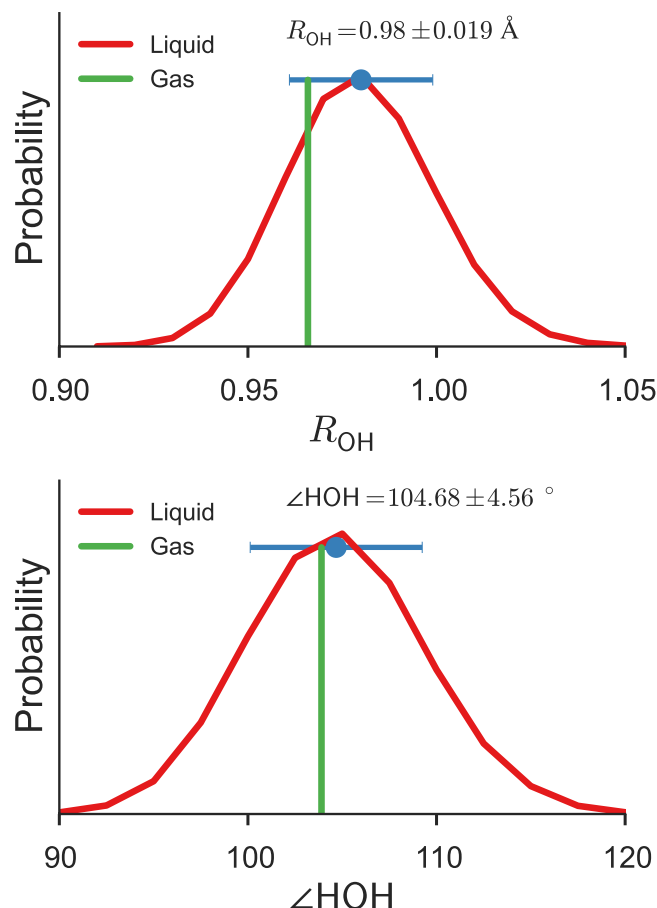
Figure 3. Bond lengths and angles. Distributions of bond lengths (top) and angles (bottom) of water molecules in the liquid phase, compared to the optimized geometry of the water molecule in the gas phase. The blue dots and bars indicate the averages and standard deviations.

Hence, the MP2/cc-pVDZ description of liquid water structure is nonphysical and this is largely due to the smallness of the basis set.

The isobaric-isothermal MC simulations at the MP2 level with a mixed Gaussian-planewave basis set was performed by Del Ben *et al.*[25]. They yielded the liquid density ($\rho = 1.02\,\text{g/cm}^3$) slightly larger than the correct value. This is in qualitative agreement with our simulation, in which the average pressure is found to be -0.6 ± 0.5 GPa at $T = 250\,\text{K}$ and $\rho = 1.0\,\text{g/cm}^3$; the larger calculated liquid density at 1 bar means that a negative pressure is required to maintain it at $1\,\text{g/cm}^3$. The RDF calculated at $T = 295\,\text{K}$ by Del Ben *et al.* was, however, overstructured, suggesting that the liquid is super-cooled, whereas the RDF obtained from our simulation is much closer to experiment. The source of the quantitative difference is hard to pinpoint because there are many approximations involved in both studies. We speculate, however, that the most likely primary source is the different choices of the basis set, in particular, the use of a pseudopotential in the preceding HF calculation in Del Ben *et al.*

Self-diffusion coefficient. Whereas the RDF characterizes time-averaged local structures, the self-diffusion coefficient (D) measures dynamical properties of liquids. Its value is related to the mean square displacement (MSD) by Einstein's diffusion equation. The calculated value of $D \approx 0.27\,\text{Å}^2/\text{ps}$ from our BOMD simulation at $T = 250\,\text{K}$ compares favorably with the experimental value of $0.23\,\text{Å}^2/\text{ps}$ at $T = 25\,°\text{C}$[36]. Hence, we argue that the present simulation properly describes the dynamical properties of liquid water at ambient conditions also.

Water molecular geometry in the liquid phase. Each water molecule undergoes a significant geometry change on going from the gas to liquid phase. The water molecule in the gas phase has the OH bond length (R_{OH}) of $0.966\,\text{Å}$ and HOH angle ($\angle HOH$) of $103.9°$ at the SCS-MP2/aug-cc-pVDZ level, which are in good agreement with the experimental values[37] of $0.957\,\text{Å}$ and $104.5°$. In the liquid phase, the calculated bond lengths and angles form distributions (Fig. 3) with $R_{OH} = 0.980 \pm 0.019\,\text{Å}$ and $\angle HOH = 104.7 \pm 4.6°$. Two experimental investigations of liquid water yielded measurements of ($R_{OH} = 0.970 \pm 0.005\,\text{Å}$; $\angle HOH = 106.1 \pm 0.9°$)[38] and ($R_{OH} = 0.983 \pm 0.008\,\text{Å}$; $\angle HOH = 104.1 \pm 1.9°$)[39].
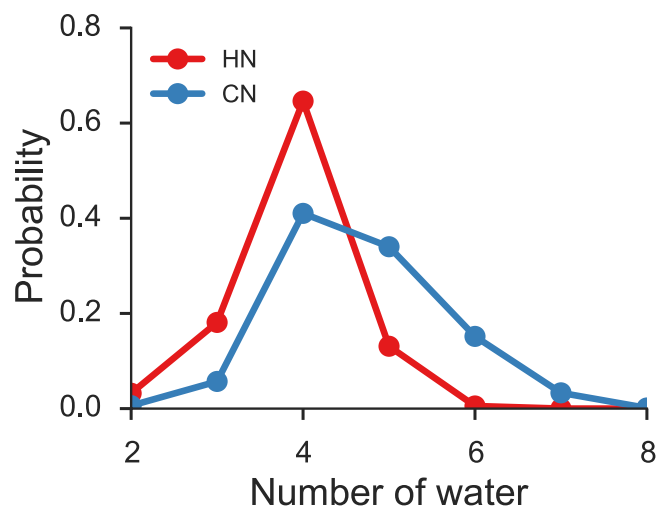
Figure 4. Coordination number and hydrogen-bond number. The distribution of the number (CN) of water molecules in the first shell ($R_{OO} \leq 3.36$ Å) excluding the central molecule and the distribution of the number (HN) of hydrogen bonds with the central molecule (with a hydrogen bond being considered to exist between two water molecules when $R_{OO} \leq 3.36$ Å and $\angle HOO \leq 40°$).
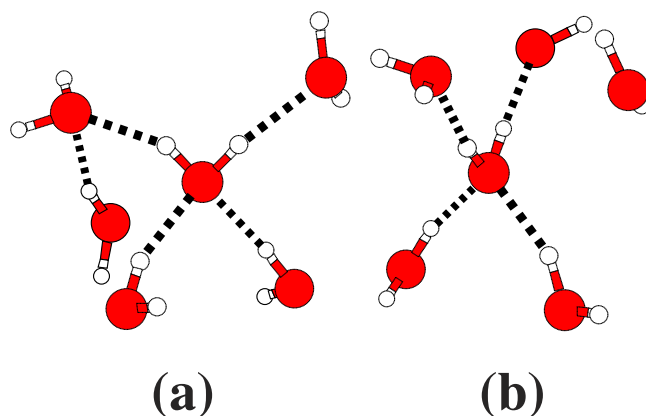

Figure 5. Snapshots of local structures in liquid water. The snapshots of the penta-coordinated water molecule are drawn. It is, on average, hydrogen-bonded to two H-acceptors and two H-donors. The fifth water molecule in the first shell stays non-hydrogen-bonded or weakly hydrogen-bonded. It can be (**a**) near the **D** zone or (**b**) near the **A** zone in the contour maps in Fig. 7.

Remarkably, the BOMD simulation predicts an average OH bond elongation (relative to the gas phase) of 0.014 Å, which is in excellent agreement with the measured elongation of 0.013 Å[38]. It is notable that the bond angle in liquid water ranges from 90° to 120° with a large standard deviation of 4.6°.

Coordination numbers and intershell exchange. Figure 4 plots the distribution of the coordination number, that is, the number of water molecules in the first shell, as well as the distribution of the hydrogen-bond number. According to our simulation, a water molecule has an average coordination number of 4.7 and an average hydrogen-bond number of 3.8. The former value is in good agreement with experiment[34]. In most cases (65%), each molecule is hydrogen-bonded to four neighbors with two serving as a H-donor (hydrogen donating water) and two as a H-acceptor (hydrogen accepting water), while the remaining neighbors in the first shell stay non-hydrogen-bonded or weakly hydrogen-bonded, as shown in Fig. 5.

The coordination number displays a surprisingly large fluctuation (4.7 ± 0.9), ranging from mono-coordination to octa-coordination. Overall, tetra-coordination is the most common (41%), followed by penta-coordination (34%) and hexa-coordination (15%). Tri-coordination (6%) and hepta-coordination (3%) are also significant, while bi-coordination (0.6%), octa-coordination (0.1%), and mono-coordination (0.01%) occur much less frequently. A majority of the penta- and higher-order
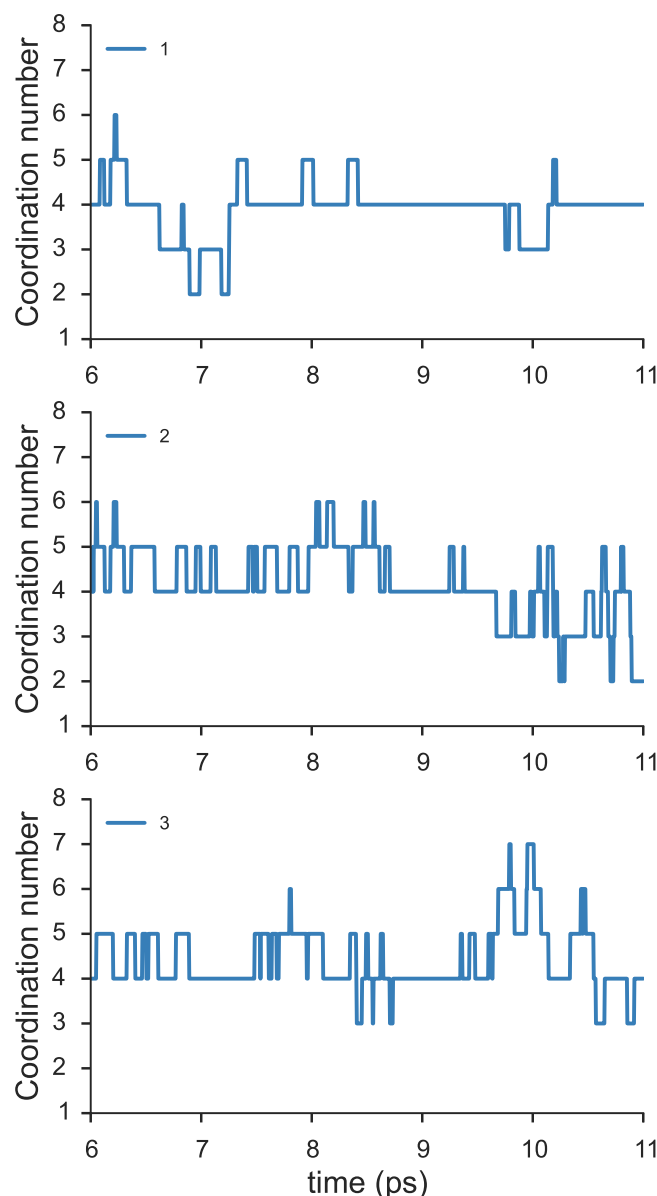
Figure 6. Fluctuation of the coordination number. The number of water molecules in the first shell ($R_{OO} \leq 3.36$ Å) as a function of time is tracked down for three selected water molecules.

coordinations have one or more non-hydrogen-bonded or weakly hydrogen-bonded water molecules in the first shell, judging from the rapid falloff of the hydrogen-bond number after four.

Figure 6 illustrates the time-evolution of the coordination numbers of three randomly selected water molecules. The large and rapid variation in each case may be related to the significant density fluctuations observed even in a nm-sized droplet of liquid water, as discussed in a recent study[40].

Figure 7 shows the (R_{OO}, θ) and (R_{OO}, φ) population contour maps of the oxygen atoms for R_{OO} up to the second-shell RDF peak ($R_{OO} \leq 4.5$ Å) as well as the (θ, φ) population contour map for the first shell ($R_{OO} \leq 3.4$ Å). They reveal that H-acceptors (the **A** zone) are found around $R_{OO} = 2.8$ Å, $\theta = 50°$, and $\varphi = 0°$, while H-donors (the **D** zone) populate around $R_{OO} = 2.8$ Å, $\theta = 120°$, and $\varphi = \pm 90°$. A significant population density is observed at $\theta = 100°$ around $R_{OO} = 3.4$ Å in the (R_{OO}, θ) contour. This is the likely path through which a non-hydrogen-bonded or weakly hydrogen-bonded neighbor in the **D** zone migrates from the first shell to the second shell and vice versa, causing the fluctuation in the coordination number.

Our BOMD trajectory data such as those shown in Fig. 6 provide even more detailed information about key events responsible for the coordination number fluctuation. Let us define a 5-4-5 hetero-exchange as the process by which a penta-coordinated water molecule loses its non-hydrogen-bonded neighbor to become tetra-coordinated, and a *different* water molecule takes its place (hetero-exchange), restoring the penta-coordinated state. This event is frequently observed in our trajectories, lasting an average of *ca.*
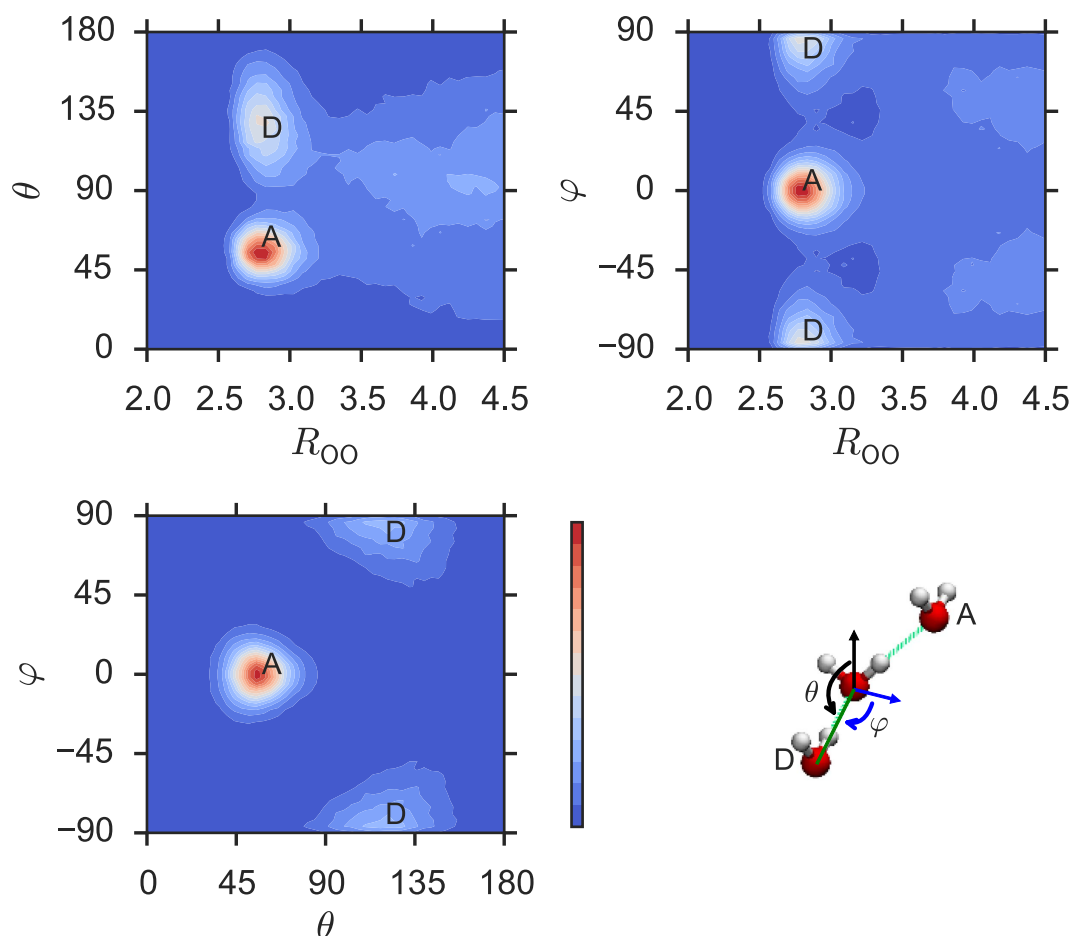
Figure 7. Oxygen-atom population. The (R_{OO}, θ) and (R_{OO}, φ) contour maps of the oxygen atom population within the second-shell RDF peak ($R_{OO} \leq 4.5$ Å) and the (θ, φ) contour map within the first shell ($R_{OO} \leq 3.4$ Å). R_{OO}, θ, and φ are spherical coordinates as defined above. **D** and **A** stand for the hydrogen-donating (H-donor) and accepting (H-acceptor) water molecules, respectively.

100 fs and occurring to a penta-coordinated species approximately every 1.0 ps. Similarly, we use the term the 4-5-4 self-exchange to describe the transient attachment of a non-hydrogen-bonded water molecule to the first shell of a tetra-coordinated water, followed by the detachment of the same molecule. This process occurs, on average, approximately every 1.5 ps with a duration of *ca.* 80 fs. The 5-4-5 self-exchange and 5-6-5 hetero-exchange are also noted at an average frequency of every few ps. The high frequency and long duration of these exchange events involving tetra- and penta-coordinated species are due to their greater stability and thus longer average lifetime.

Dipole moments. The dipole moment of a water molecule in the liquid phase has been under debate for a long time, partly because there is no experimental means of measuring it directly. Accurate theoretical predictions are, therefore, extremely valuable, but previously reported computed values ranged widely from 2.6 to 3.0 D[11,41,42]. An earlier DFT-MD study placed it at 2.66 D[10], but later revised it to 2.95 D[41]. The latter, larger predicted value may be traced to the significantly overestimated polarizability characteristic of DFT. Subsequently, Badyal *et al.*[43] determined the partial charge on the hydrogen atom as $q_H = 0.5 \pm 0.1$ a.u. by X-ray diffraction. On this basis, they suggested the dipole moment of 2.9 ± 0.6 D, which had such a large error bar that it encompassed nearly all computationally suggested values.

The dipole moment of the water molecule in the gas phase calculated by SCS-MP2/aug-cc-pVDZ is 1.878 D, which agrees well with the experimental value of 1.855 D[44]. Figure 8 plots the distributions of the permanent dipole moments $\{\mu_i^0\}$ (obtained with the embedding field turned off) and the total dipole moments $\{\mu_i\}$ (calculated with the embedding field within 9 Å and thus including the induced dipole moments) in the liquid phase. The permanent dipole moment μ_i^0 at SCS-MP2/aug-cc-pVDZ is 1.872 ± 0.07 D, whereas the total dipole moment μ_i is 2.67 ± 0.2 D. Hence, the induced dipole moment is significant and amounts to as much as 0.80 D on average. it could further increase slightly by the nuclear quantum effect[13], but perhaps by not more than 0.05 D.
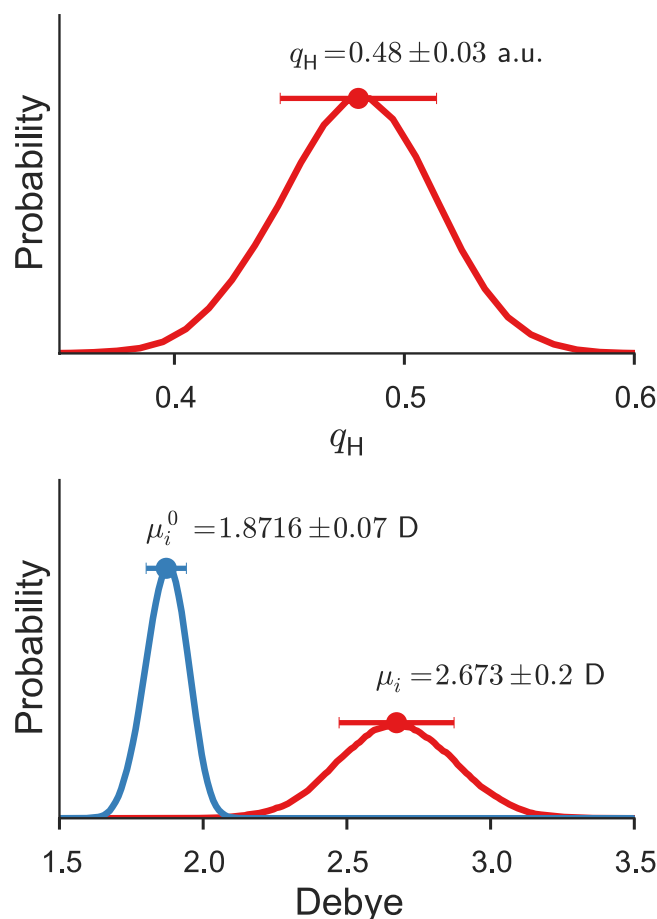
Figure 8. Dipole moments and hydrogen partial charge. The distribution of the partial charges of the hydrogen atoms (top) and the distribution of the permanent (μ_i^0) and total (μ_i) dipole moments (bottom) of the water molecules in the liquid phase. The total dipole moment is obtained with the embedding field within the cut-off distance $R_{EF} = 9\,\text{Å}$, whereas the permanent dipole moment with the field turned off.

Furthermore, the total dipole moment is broadly distributed, showing large fluctuations from 1.8 to 3.5 D, with about 5% exceeding 3.0 D, owing more to the diversity of electrostatic fields than to the monomer geometry fluctuation. The partial charge of the hydrogen atom, $q_H = 0.48 \pm 0.03$ a.u. (Fig. 8), is in good agreement with the mean of the experimental value of Badyal *et al.*, $q_H = 0.5 \pm 0.1$ a.u.[43]. The difference in the total dipole moment between our study and Badyal *et al.* may be traced to the difference in the water structure: the OH bond length proposed[45] by one of the authors of ref. 43 is 1.045 Å and is much greater than our predicted (0.980 Å) or experimental values (0.970–0.983 Å)[38,39]. Our predicted dipole moment ($\mu_i = 2.67$ D) is in excellent agreement with other independent experimental expectation (~2.6 D)[46] and theoretical calculation (~2.65 D)[42]. In short, the dipole moment suggested by Badyal *et al.* (2.9 D) is, in view of this and other accurate theoretical calculations, questionable and likely overestimated.

IR and Raman spectra. The SCS-MP2/aug-cc-pVDZ BOMD simulation predicts the vibrational frequencies of the water molecule in the gas phase at 1676, 3800, and 3900 cm^{-1}. Though BOMD simulations sample the fully anharmonic PES, owing to the classical treatment of vibrations, these frequencies are too high relative to the experimental values (1595, 3657, and 3756 cm^{-1})[37,47], for reasons that are well understood[48]. To correct such deviations rigorously, one should use path-integral molecular dynamics[8,49,50], which can take into account the ZPVE or quantum nuclear effects. A more empirical correction is to simply multiply the calculated frequencies by a scale factor (0.96 for the gas phase and 0.93 for the liquid phase) that can bring the calculated peak positions in better agreement with those observed experimentally.

The IR and Raman spectra of liquid water computed from the SCS-MP2/aug-cc-pVDZ BOMD simulation at $T = 250$ K are compared with the corresponding experimental results at ambient conditions[51–53] in Fig. 9. Unlike most calculated spectra, each band in our simulated spectra is not convoluted with an assumed band shape, but is instead associated with a physically meaningful width that reflects the diversity of local environments in which the vibrations occur.
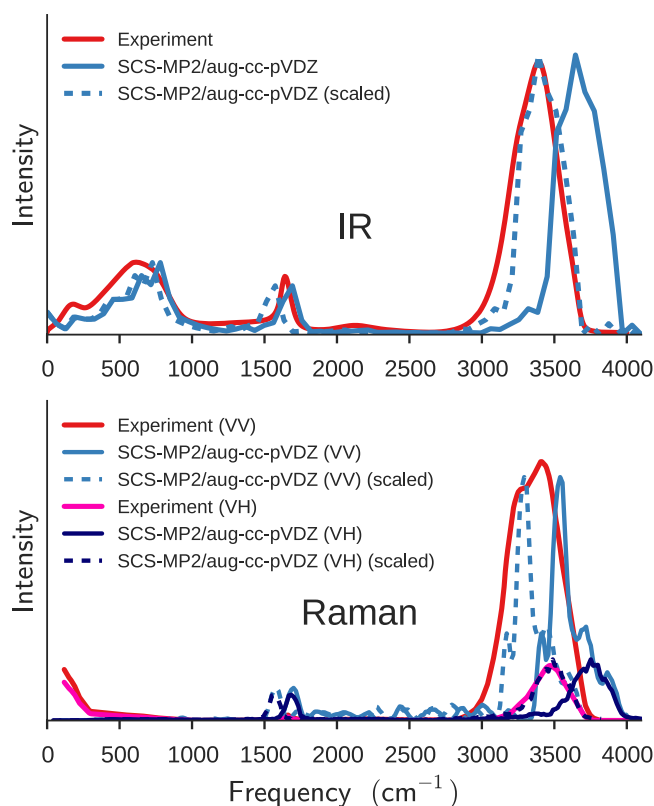
Figure 9. IR and Raman spectra. The IR and Raman spectra of liquid water were calculated with the SCS-MP2/aug-cc-pVDZ simulation at $T = 250\,$K and $\rho = 1\,$g/cm^3. The experimental IR spectrum at $T = 25\,°$C is taken from ref. 52 and the experimental VV and VH Raman spectra at $T = 23\,°$C from ref. 51. The dashed curves are obtained by scaling the frequencies of the calculated spectra by 0.93 to empirically take into account the ZPVE (see the text).

The BOMD simulation reproduces all three major features of the observed IR spectrum (top panel): the broad, intense OH stretching band above 3000 cm^{-1}, the narrower peak due to the bending mode at around 1600 cm^{-1}, and the manifold of intermolecular vibrations below 1000 cm^{-1}.

The OH stretching band in the observed IR spectrum peaks at about 3400 cm^{-1} with a weak shoulder around 3250 cm^{-1} and a full width at half maximum (FWHM) of 375 cm^{-1} [52,53]. Remarkably, the BOMD simulation predicts the OH stretching band with the correct band width (FWHM of about 375 cm^{-1}) and slightly asymmetric band shape, which also agrees with the observed shape, though the calculated peak positions of 3550 and 3700 cm^{-1} are blue-shifted. It is striking that the classical "continuum" picture[48], in which diverse instantaneous hydrogen-bond environments surrounding the OH stretching vibrations give rise to the band width, is borne out by our simulation. Our time-domain computation of the spectra should also naturally account for the motional narrowing of this band, which is estimated to be as much as 30% of the width[54].

The OH stretching band of the observed VV Raman spectrum has bimodal peaks at about 3250 and 3400 cm^{-1} and a FWHM of about 425 cm^{-1}, while the observed VH spectrum shows the OH stretching band around 3460 cm^{-1} with an asymmetric line shape and a FWHM of about 300 cm^{-1} [51]. The calculated VV and VH Raman spectra are also in agreement with the observed spectra. It explains the asymmetric line shape of the OH stretching band in the VH spectrum and the double-peaked nature of the band in the VV spectrum, including the difference in peak positions between the VH and VV components, known as the Raman noncoincidence effect[55]. With the aforementioned frequency scaling by a factor of 0.93, the simulated VV and VH spectra are again in excellent agreement with their experimental counterparts both in peak positions and shapes.

It could be emphasized that our spectral simulation does not involve any adjustable parameter (which is not to say that there is no approximation) in the anharmonic force field, mode-mode coupling, polarizability tensor, etc., unlike previous analyses on the same bands[48]. Note that polarizable model potentials of water[8,56] currently fall short of simultaneously accounting for polarization effects on intermolecular interactions and optical response, such as the Raman effect. Our *ab initio* method can achieve just that, although the peak positions are systematically overestimated because of the classical treatment of vibrations.

Discussion

We have performed *ab initio* BOMD simulations of liquid water at the SCS-MP2 and MP2 levels of theory and computationally reproduced the structural properties (the radial distribution function, monomer geometry and its fluctuation) and structure-driven collective properties (the self-diffusion coefficient, coordination number and its variation, intershell exchange) as well as the electronic and response properties (permanent and induced dipole moments, IR spectra, and isotropic and anisotropic Raman spectra) simultaneously, accurately, and from first principles. The level of theory has been chosen to reproduce the CCSD(T)/aug-cc-pVQZ results of the water dimer, allowing us to study the electronic details of liquid water with unprecedented accuracy.

The calculated properties of liquid water are in good agreement with the experimental results, when the latter are available. The simulated IR and Raman spectra correctly predict the band shapes and widths of the OH stretching bands and the difference between their isotropic and anisotropic Raman components, which are exceedingly difficult with classical MD with parameterized force fields, even if they are polarizable.

For properties that are difficult to probe experimentally, our simulations provide unique insights. The average bond length and angle of the water molecule in the liquid phase are found to increase only minimally, but range from 0.92 to 1.05 Å and from 90° to 120°. The coordination number fluctuates even more greatly from mono-coordination to octa-coordination, though the tetra- (41%) and peta-coordinations (34%) are most common. Such large fluctuation in the coordination number might lead to significant density fluctuations. We have identified 5-4-5 hetero-exchange and 4-5-4 self/hetero-exchange as the key events, occurring approximately every 1–2 ps with an average duration of *ca.* 0.1 ps and causing the large fluctuation in the coordination number. Such exchanges take place as a non-hydrogen-bonded water molecule in the **D** zone migrates between the first and second shells. The total dipole moment ($\mu_i = 2.67 \pm 0.2$ D) of the water molecule in the liquid phase also shows a large fluctuation, ranging from 1.8 to 3.5 D, reflecting more the diversity of the electrostatic field surrounding the water molecule than the geometry fluctuation.

Hence, the present method allows us to assess the validity of uncertain results or conclusions about a whole range of liquid water properties. It can also be exploited to study other unknown collective properties of liquid water and various solvents as well as colligative properties of solute-solvents generally.

Methods

Born-Oppenheimer molecular dynamics. BOMD simulations were carried out in the canonical (*NVT*) ensemble with the velocity Verlet algorithm in conjunction with the Nosé–Hoover chain method[57,58]. A time step of 1 fs was used. An initial liquid structure of $(H_2O)_{32}$ in a cubic cell of side 9.858 Å was obtained from an equilibration run with the TTM3-F force field[8] at $T = 300$ K. Using the same force field, we verified that the 32-water simulation yielded the calculated RDF which was unchanged from the result of a 512-water simulation and was, therefore, adequate. Using SCS-MP2/aug-cc-pVDZ or MP2/cc-pVDZ energies and atomic forces at $T = 250$ K, a 6-ps equilibration run was performed, after which the trajectory was sampled every 4 time steps during a 5-ps production run.

Self-diffusion coefficient. The self-diffusion coefficient (*D*) is related to the MSD by Einstein's diffusion equation:

$$D = \frac{\text{MSD}}{6t} = \frac{1}{6t} \left\langle |\boldsymbol{r}(t) - \boldsymbol{r}(0)|^2 \right\rangle. \tag{1}$$

In turn, MSD was obtained by the calculation of the squared relative displacement of the oxygen atoms as a function of time *t* averaged over all water molecules in the unit cell.

IR and Raman spectra. The IR and Raman spectra were calculated by Fourier transformation of the time-correlation functions of the system's dipole moment and polarizability, respectively. For a direct comparison with the experimental spectra, the harmonic quantum correction was applied to the classically computed spectra[59]. IR intensity $I_{\text{IR}}(\omega)$ at frequency ω then becomes[56]

$$I_{\text{IR}}(\omega) \propto \frac{\omega^2}{k_{\text{B}}T} \int_{-\infty}^{\infty} dt \ e^{-i\omega t} \left\langle \boldsymbol{\mu}(0) \cdot \boldsymbol{\mu}(t) \right\rangle, \tag{2}$$

where $\boldsymbol{\mu}(t)$ is the dipole moment of the system at time *t*, defined as the sum of the dipole moments of monomers, $\{\boldsymbol{\mu}_i\}$, in the unit cell,

$$\boldsymbol{\mu}(t) = \sum_i \boldsymbol{\mu}_i(t). \tag{3}$$

Here, the dipole moment of the *i*th monomer embedded in the electrostatic field, $\boldsymbol{\mu}_i$, is directly computed by the SCS-MP2 method.

The harmonic-quantum-corrected, Bose–Einstein-reduced isotropic and anisotropic Raman spectra were calculated by[56]

$$I_{\text{Raman}}^{\text{iso}}(\omega) \propto \frac{1}{9}\frac{\omega}{k_{\text{B}}T}\int_{-\infty}^{\infty} dt \; e^{-i\omega t}\,\langle\text{Tr}[\boldsymbol{\alpha}(0)]\,\text{Tr}[\boldsymbol{\alpha}(t)]\rangle, \tag{4}$$

$$I_{\text{Raman}}^{\text{aniso}}(\omega) \propto \frac{2}{15}\frac{\omega}{k_{\text{B}}T}\int_{-\infty}^{\infty} dt \; e^{-i\omega t}\,\langle\text{Tr}[\boldsymbol{\beta}(0)\cdot\boldsymbol{\beta}(t)]\rangle, \tag{5}$$

where $\boldsymbol{\alpha}$ is the polarizability tensor of the unit cell and $\boldsymbol{\beta} = \boldsymbol{\alpha} - \text{Tr}[\boldsymbol{\alpha}]\boldsymbol{I}/3$ with $\boldsymbol{I}$ being the identity tensor. Each element of $\boldsymbol{\alpha}$ is the sum of the corresponding elements of the embedded monomer polarizability tensors, $\{\boldsymbol{\alpha}_i\}$. The latter is evaluated as the ratio of the induced dipole moment $\boldsymbol{\mu}_i^{\text{ind.}}$ to the applied electric field $\boldsymbol{E}$, e.g.,

$$(\boldsymbol{\alpha}_i)_{xy} = \frac{(\boldsymbol{\mu}_i^{\text{ind.}})_x}{E_y}. \tag{6}$$

Induced dipole moments are, in turn, evaluated directly by the SCS-MP2 method for an embedded monomer with additional point charges of ∓ 1 a.u. placed at $\boldsymbol{r}_{\text{cm}} \pm 100$ bohr along each Cartesian axis. Here, $\boldsymbol{r}_{\text{cm}}$ denotes the center of mass of the monomer. The parallel-polarized (VV) and perpendicular-polarized (VH) Raman spectra, with each letter denoting the vertical (V) or horizontal (H) polarization direction of the incident and scattered light, are related to the isotropic and anisotropic components by

$$I_{\text{Raman}}^{\text{VV}}(\omega) = I_{\text{Raman}}^{\text{iso}}(\omega) + I_{\text{Raman}}^{\text{aniso}}(\omega), \tag{7}$$

$$I_{\text{Raman}}^{\text{VH}}(\omega) = \frac{3}{4}I_{\text{Raman}}^{\text{aniso}}(\omega). \tag{8}$$

Using the classical MD with the TTM3-F force field[8], we verified that the simulation with a 1-fs time step yields the same IR spectra as the simulation with a 0.2-fs time step. The only exception was the positions of the OH stretching bands (if not their widths and shape), in which a time-step error of up to 50 cm^{-1} (or ~1.5%) was observed.

Energy and gradient evaluation. The potential energy, atomic gradients, and pressure of liquid water in each time step were calculated by the massively parallel implementation of the BIM method[27] using the SCS-MP2/aug-cc-pVDZ or MP2/cc-pVDZ method.

The formalisms of BIM for condensed-phase simulations are documented in the Supplementary Information. Briefly, it divides the liquid into monomers and overlapping dimers embedded in the electrostatic environment of the liquid. The potential energies and gradients of each of the monomers or dimers, treated quantum mechanically (QM), are evaluated with SCS-MP2 or MP2 implemented in a modified NWCHEM[60]. They are used to reconstruct the potential energy, atomic forces, and pressure of the whole liquid under the periodic boundary condition according to Eqs (S1), (S5), and (S6) in the Supplementary Information. The embedding field (EF) consists of atomic point charges that approximate the electrostatic potential of each monomer and are self-consistently determined at the HF level. The calculation thus includes two-body electrostatic, exchange, and dispersion interactions at the SCS-MP2 or MP2 level and three-body and all higher many-body electrostatic interactions at the HF level. The truncation radii of the QM, EF, and long-range (LR) regions, as defined by Eqs (S1)–(S4), are as follows: $R_{\text{QM}} = 7.5$ Å, $R_{\text{EF}} = 9$ Å, and $R_{\text{LR}} = 12$ Å.

References

1. Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W. & Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **79,** 926–935 (1983).
2. Lei, G. C. & Clementi, E. Molecular-dynamics simulation of liquid water with an ab initio flexible water-water interaction potential. *Phys. Rev. A* **33,** 2679–2693 (1986).
3. Berendsen, H. J. C., Grigera, J. R. & Straatsma, T. P. The missing term in effective pair potentials. *J. Phys. Chem.* **91,** 6269–6271 (1987).
4. Fellers, R. S., Leforestier, C., Braly, L. B., Brown, M. G. & Saykally, R. J. Spectroscopic determination of the water pair potential. *Science* **284,** 945–948 (1999).
5. Ren, P. & Ponder, J. W. Temperature and pressure dependence of the AMOEBA water model. *J. Phys. Chem. B* **108,** 13427–13437 (2004).
6. Abascal, J. L. F., Sanz, E., García Fernaández, R. & Vega, C. A potential model for the study of ices and amorphous water: TIP4P/Ice. *J. Chem. Phys.* **122,** 234511 (2005).
7. Bukowski, R., Szalewicz, K., Groenenboom, G. C. & van der Avoird, A. Predictions of the properties of water from first principles. *Science* **315,** 1249–1252 (2007).
8. Fanourgakis, G. S. & Xantheas, S. S. Development of transferable interaction potentials for water. V. Extension of the flexible, polarizable, Thole-type model potential (TTM3-F, v. 3.0) to describe the vibrational spectra of water clusters and liquid water. *J. Chem. Phys.* **128,** 074506 (2008).
9. Wang, L.-P. *et al.* Systematic improvement of a classical molecular model of water. *J. Phys. Chem. B* **117,** 9956–9972 (2013).
10. Laasonen, K., Sprik, M., Parrinello, M. & Car, R. "Ab initio" liquid water. *J. Chem. Phys.* **99,** 9080–9089 (1993).

11. Silvestrelli, P. L., Bernasconi, M. & Parrinello, M. Ab initio infrared spectra of liquid water. *Chem. Phys. Lett.* **277,** 478–482 (1997).

12. Schwegler, E., Grossman, J. C., Gygi, F. & Galli, G. Towards an assessment of the accuracy of density functional theory for first principles simulations of water. II. *J. Chem. Phys.* **121,** 5400 (2004).

13. Morrone, J. A. & Car, R. Nuclear quantum effects in water. *Phys. Rev. Lett.* **101,** 017801 (2008).

14. Georg, H. Electronic properties of water in liquid environments. A sequential QM/MM study using the free energy gradient method. *J. Phys. Chem. B* **116,** 11247–11254 (2012).

15. Yoo, S., Zeng, X. C. & Xantheas, S. S. On the phase diagram of water with density functional theory potentials: The melting temperature of ice Ih with the Perdew-Burke-Ernzerhof and Becke-Lee-Yang-Parr functionals. *J. Chem. Phys.* **130,** 221102 (2009).

16. Yoo, S. & Xantheas, S. S. Communication: The effect of dispersion corrections on the melting temperature of liquid water. *J. Chem. Phys.* **134,** 121105 (2011).

17. Baer, M. D. *et al.* Re-examining the properties of the aqueous vapor-liquid interface using dispersion corrected density functional theory. *J. Chem. Phys.* **135,** 124712 (2011).

18. Gordon, M. S., Fedorov, D. G., Pruitt, S. R. & Slipchenko, L. V. Fragmentation methods: A route to accurate calculations on large systems. *Chem. Rev.* **112,** 632–672 (2012).

19. Fujita, T., Nakano, T. & Tanaka, S. Fragment molecular orbital calculations under periodic boundary condition. *Chem. Phys. Lett.* **506,** 112–116 (2011).

20. Brorsen, K. R., Minezawa, N., Xu, F., Windus, T. L. & Gordon, M. S. Fragment molecular orbital molecular dynamics with the fully analytic energy gradient. *J. Chem. Theory Comput.* **8,** 5008–5012 (2012).

21. Brorsen, K. R., Zahariev, F., Nakata, H., Fedorov, D. G. & Gordon, M. S. Analytic gradient for density functional theory based on the fragment molecular orbital method. *J. Chem. Theory Comput.* **10,** 5297–5307 (2014).

22. Kitaura, K., Ikeo, E., Asada, T., Nakano, T. & Uebayasi, M. Fragment molecular orbital method: an approximate computational method for large molecules. *Chem. Phys. Lett.* **313,** 701–706 (1999).

23. Komeiji, Y., Mochizuki, Y. & Nakano, T. Three-body expansion and generalized dynamic fragmentation improve the fragment molecular orbital-based molecular dynamics (FMO-MD). *Chem. Phys. Lett.* **484,** 380–386 (2010).

24. Mochizuki, Y. *et al.* Fragment molecular orbital-based molecular dynamics (FMO-MD) method with MP2 gradient. *Chem. Phys. Lett.* **504,** 95–99 (2011).

25. Del Ben, M., Schönherr, M., Hutter, J. & VandeVondele, J. Bulk liquid water at ambient temperature and pressure from MP2 theory. *J. Phys. Chem. Lett.* **4,** 3753–3759 (2013).

26. Grimme, S. Improved second-order Møller-Plesset perturbation theory by separate scaling of parallel- and antiparallel-spin pair correlation energies. *J. Chem. Phys.* **118,** 9095 (2003).

27. Hirata, S., Gilliard, K., He, X., Li, J. J. & Sode, O. Ab initio molecular crystal structures, spectra, and phase diagrams. *Acc. Chem. Res.* **47,** 2721–2730 (2014).

28. Kim, J. & Kim, K. S. Structures, binding energies, and spectra of isoenergetic water hexamer clusters: Extensive ab initio studies. *J. Chem. Phys.* **109,** 5886–5895 (1998).

29. Xantheas, S. S., Burnham, C. J. & Harrison, R. J. Development of transferable interaction models for water. II. Accurate energetics of the first few water clusters from first principles. *J. Chem. Phys.* **116,** 1493–1499 (2002).

30. Xantheas, S. S. & Aprà, E. The binding energies of the D_2 and S_4 water octamer isomers: High-level electronic structure and empirical potential results. *J. Chem. Phys.* **120,** 823 (2004).

31. Kim, K. S., Mhin, B. J., Choi, U. & Lee, K. Ab initio studies of the water dimer using large basis sets: The structure and thermodynamic energies. *J. Chem. Phys.* **97,** 6649–6662 (1992).

32. Dyke, T. R. & Muenter, J. S. Microwave spectrum and structure of hydrogen bonded water dimer. *J. Chem. Phys.* **60,** 2929–2930 (1974).

33. Dyke, T. R., Mack, K. M. & Muenter, J. S. The structure of water dimer from molecular beam electric resonance spectroscopy. *J. Chem. Phys.* **66,** 498–510 (1977).

34. Skinner, L. B. *et al.* Benchmark oxygen-oxygen pair-distribution function of ambient water from x-ray diffraction measurements with a wide Q-range. *J. Chem. Phys.* **138,** 074506 (2013).

35. Hura, G., Sorenson, J. M., Glaeser, R. M. & Head-Gordon, T. A high-quality x-ray scattering experiment on liquid water at ambient conditions. *J. Chem. Phys.* **113,** 9140 (2000).

36. Holz, M., Heil, S. R. & Sacco, A. Temperature-dependent self-diffusion coefficients of water and six selected molecular liquids for calibration in accurate ^{1}H NMR PFG measurements. *Phys. Chem. Chem. Phys.* **2,** 4740–4742 (2000).

37. Benedict, W. S., Gailar, N. & Plyler, E. K. Rotationvibration spectra of deuterated water vapor. *J. Chem. Phys.* **24,** 1139–1165 (1956).

38. Ichikawa, K., Kameda, Y., Yamaguchi, T., Wakita, H. & Misawa, M. Neutron-diffraction investigation of the intramolecular structure of a water molecule in the liquid phase at high temperatures. *Mol. Phys.* **73,** 79–86 (1991).

39. Powles, J. G. The structure of the water molecule in liquid water. *Mol. Phys.* **42,** 757–765 (1981).

40. English, N. J. & Tse, J. S. Density fluctuations in liquid water. *Phys. Rev. Lett.* **106,** 037801 (2011).

41. Silvestrelli, P. L. & Parrinello, M. Water molecule dipole in the gas and in the liquid phase. *Phys. Rev. Lett.* **82,** 3308–3311 (1999).

42. Tu, Y. & Laaksonen, A. The electronic properties of water molecules in water clusters and liquid water. *Chem. Phys. Lett.* **329,** 283–288 (2000).

43. Badyal, Y. S. *et al.* Electron distribution in water. *J. Chem. Phys.* **112,** 9206 (2000).

44. Clementi, E. Study of the electronic structure of molecules. II. Wavefunctions for the $NH_3 + HCl \rightarrow NH_4Cl$ reaction. *J. Chem. Phys.* **46,** 3851 (1967).

45. Soper, A. K. The radial distribution functions of water and ice from 220 to 673 K and at pressures up to 400 MPa. *Chem. Phys.* **258,** 121–137 (2000).

46. Gregory, J. K., Clary, D. C., Liu, K., Brown, M. G. & Saykally, R. J. The water dipole moment in water clusters. *Science* **275,** 814–817 (1997).

47. Clough, S. A., Beers, Y., Klein, G. P. & Rothman, L. S. Dipole moment of water from Stark measurements of H_2O, HDO, and D_2O. *J. Chem. Phys.* **59,** 2254–2259 (1973).

48. Auer, B. M. & Skinner, J. L. IR and Raman spectra of liquid water: Theory and interpretation. *J. Chem. Phys.* **128,** 224511 (2008).

49. Paesani, F. & Voth, G. A. The properties of water: Insights from quantum simulations. *J. Phys. Chem. B* **113,** 5702–5719 (2009).

50. Paesani, F., Willow, S. Y., Bakker, H. J. & Xantheas, S. S. Nuclear quantum effects in the reorientation of water. *J. Phys. Chem. Lett.* **1,** 2316–2321 (2010).

51. Scherer, J. R., Go, M. K. & Kint, S. Raman spectra and structure of water from −10 to 90°. *J. Phys. Chem.* **78,** 1304–1313 (1974).

52. Bertie, J. E. & Lan, Z. Infrared intensities of liquids XX: The intensity of the OH stretching band of liquid water revisited, and the best current values of the optical constants of H_2O (l) at 25 °C between 15,000 and 1 cm^{-1}. *Appl. Spectrosc.* **50,** 1047–1057 (1996).

53. Brubach, J.-B., Mermet, A., Filabozzi, A., Gerschel, A. & Roy, P. Signatures of the hydrogen bonding in the infrared bands of water. *J. Chem. Phys.* **122,** 184509 (2005).

54. Eaves, J. D., Tokmakoff, A. & Geissler, P. L. Electric field fluctuations drive vibrational dephasing in water. *J. Phys. Chem. A* **109,** 9424–9436 (2005).
55. Torii, H. Time-domain calculations of the polarized Raman spectra, the transient infrared absorption anisotropy, and the extent of delocalization of the OH stretching mode of liquid water. *J. Phys. Chem. A* **110,** 9469–9477 (2006).
56. Hasegawa, T. & Tanimura, Y. A polarizable water model for intramolecular and intermolecular vibrational spectroscopies. *J. Phys. Chem. B* **115,** 5545–5553 (2011).
57. Nosé, S. A unified formulation of the constant temperature molecular dynamics methods. *J. Chem. Phys.* **81,** 511 (1984).
58. Martyna, G. J., Tuckerman, M. E., Tobias, D. J. & Klein, M. L. Explicit reversible integrators for extended systems dynamics. *Mol. Phys.* **87,** 1117–1157 (1996).
59. Ramírez, R., López-Ciudad, T., Kumar, P. & Marx, D. Quantum corrections to classical time-correlation functions: Hydrogen bonding and anharmonic floppy modes. *J. Chem. Phys.* **121,** 3973–3983 (2004).
60. Valiev, M. *et al.* NWChem: A comprehensive and scalable open-source solution for large scale molecular simulations. *Comput. Phys. Commun.* **181,** 1477–1489 (2010).

Acknowledgments

Author Contributions

The research was designed by S.Y.W., K.S.K. and S.H., the computer program was developed by S.Y.W. and M.S. and the computational simulation was carried out by S.Y.W. The results were analyzed by all authors. The manuscript was co-written and discussed by all authors.

Additional Information

Supplementary information accompanies this paper at http://www.nature.com/srep

Competing financial interests: The authors declare no competing financial interests.

How to cite this article: Willow, S. Y. *et al. Ab initio* molecular dynamics of liquid water using embedded-fragment second-order many-body perturbation theory towards its accurate property prediction. *Sci. Rep.* **5**, 14358; doi: 10.1038/srep14358 (2015).