# Universal tight binding model for chemical reactions in solution and at surfaces. II. Water

A. Y. Lozovoi, T. J. Sheppard, D. L. Pashov, J. J. Kohanoff, and A. T. Paxton

**View Online**          **Export Citation**          **CrossMark**

# Universal tight binding model for chemical reactions in solution and at surfaces. II. Water

A. Y. Lozovoi,[1] T. J. Sheppard,[1] D. L. Pashov,[2] J. J. Kohanoff,[1] and A. T. Paxton[2,a)]

[1]*Atomistic Simulation Centre, School of Mathematics and Physics, Queen's University Belfast, Belfast BT7 1NN, Northern Ireland, United Kingdom*
[2]*Department of Physics, King's College London, Strand, London WC2R 2LS, United Kingdom*

A revised water model intended for use in condensed phase simulations in the framework of the self consistent polarizable ion tight binding theory is constructed. The model is applied to water monomer, dimer, hexamers, ice, and liquid, where it demonstrates good agreement with theoretical results obtained by more accurate methods, such as DFT and CCSD(T), and with experiment. In particular, the temperature dependence of the self diffusion coefficient in liquid water predicted by the model, closely reproduces experimental curves in the temperature interval between 230 K and 350 K. In addition, and in contrast to standard DFT, the model properly orders the relative densities of liquid water and ice. A notable, but inevitable, shortcoming of the model is underestimation of the static dielectric constant by a factor of two. We demonstrate that the description of inter and intramolecular forces embodied in the tight binding approximation in quantum mechanics leads to a number of valuable insights which can be missing from *ab initio* quantum chemistry and classical force fields. These include a discussion of the origin of the enhanced molecular electric dipole moment in the condensed phases, and a detailed explanation for the increase of coordination number in liquid water as a function of temperature and compared with ice—leading to insights into the anomalous expansion on freezing. The theory holds out the prospect of an understanding of the currently unexplained density maximum of water near the freezing point. © *2014 AIP Publishing LLC.* [http://dx.doi.org/10.1063/1.4890343]

## I. INTRODUCTION

This is the second in a series of three papers describing a universal tight binding model for the electronic structure and interatomic forces in condensed phases. In the first,[1] referred to as Paper I in what follows, we presented a scheme for molecular dynamics simulations of organic molecules and here we show how the method is extended for systems in polar solvents.

Of all the polar solvents, we focus here exclusively on water. Water is well known as the universal solvent in biochemistry and the unusual property of its liquid being denser than ice at the melting point is regarded as enabling life to exist on Earth. It has additional features of being able to act as both donor and acceptor of hydrogen bonds, and the hydrogen bonded network found in both the liquid and in ice lies at the heart of its remarkable properties.

Theoretical simulations of water are done nowadays using a variety of methods ranging from *ab initio* quantum chemistry[2] and density functional theory[3] (DFT) to empirical force field models (SPC/E[4] and TIP4P[5] models and their derivatives being most popular during the last couple of decades). The tight binding (TB) approach is the method of choice if the required size of the system (or the length of the simulation in case of molecular dynamics) is beyond the capability of first principles methods, but the effect of electron and proton transfer or bond breaking, not captured in force field models, cannot be ignored.

Earlier it was demonstrated by two of us[6] that the self consistent polarizable ion tight binding (PITB) method[7,8] provides an excellent framework for the description of water. In this method, in addition to on-site energies and bond integrals, atomic species are assigned point multipoles up to a specified angular momentum. These point multipoles are not fixed but rather sought self consistently so that ions respond to varying local electric fields and field gradients during a simulation.

Here we present a revised TB water model to be used within the framework of the self consistent polarizable ion method. The development of the new model is specifically tailored to its application to condensed phases of water, i.e., ice and liquid water. We also correct a few pathological features that were discovered in the previous water model of Ref. 6 (see Secs. II and V).

We require the new water model to be suitable for simulations of organic molecules solvated in water, and also that our model correctly describes interfaces between water and transition metal oxide. In terms of model parameters, such "universality" means in particular that oxygen in water, oxygen in organic molecules, and oxygen in metal oxides must be the same entity, that means described by the same TB parameters. Hence, the development of the water model was linked to the development of the TB models for titanium dioxide and organic molecules comprising C, H, and O atoms. The two latter models are presented in detail elsewhere.[1,9]

---

a)Author to whom correspondence should be addressed. Electronic mail: Tony.Paxton@KCL.ac.uk.

In Paper I[1] we gave a comprehensive account of the self consistent polarizable ion tight binding (PITB) theory. Further details can be found in Refs. 7, 8, 10, and 11. The structure of the present paper is as follows. In Sec. II, as in each paper of the series, we give a description of the particular fitting strategy used. Model parameters are presented in Sec. III and compared to the previous model of Ref. 6. Section IV describes computational details and lists typical conditions and tolerances used in simulations. In Secs. V–VIII we discuss results obtained with the new model: for water monomer and dimer (Sec. V), hexamers (Sec. VI), ice (Sec. VII), and liquid water (Sec. VIII). In the latter section, in particular, we present the temperature dependence of the radial distribution functions (RDF), self diffusion coefficient, and dielectric constant of liquid water. Finally, Sec. IX presents a discussion and the main conclusions of the article.

## II. FITTING STRATEGY

The self consistent PITB model provides a natural framework for the description of water able to capture all essential interactions.[6] Even "intuitive" models (point charge and dipole) are sufficiently good to reproduce various properties of water. Most results obtained with the "genetic dipole model" of Ref. 6 were in good agreement with observations and accurate quantum chemistry calculations. However, we were not entirely happy with three features of the model, namely: (1) the model predicts ice to be denser than liquid water; (2) there appears a shoulder near the first peak of the O–O RDF absent on experimental curves; and (3) the energy separation between oxygen $s$ and $p$ states is insufficient to render correct band structures of metal oxides. Hence, one cannot simulate systems containing water *and* metal oxide without having two distinct oxygen species. Interestingly enough, the point charge model of Ref. 6 did not show any of the above deficiencies, but the point charge model was inferior to dipole models in reproducing the monomer polarizability and long ranged structure of liquid water. That led us to speculate that features (1)–(3) are not unavoidable and can be eliminated in the dipole model. In the present work we achieved exactly that: developed a new dipole model for water free from undesirable features (1)–(3) and thus suitable for simulations of condensed phases of water and water–metal oxide interfaces.

The fitting is achieved in a sequence of stages. In the first stage we focus on quantities that do not depend on the O–O pair potential. For these, we employ the genetic algorithm to fit on site energies, Hubbard $U$'s, crystal field parameter $\Delta$'s, O–H and O–O bond integrals, and O–H pair potential. The targets at this stage are mostly the monomer properties (geometry, vibrational force constants, and the binding energy). We also use dimer angles $\alpha$ and $\beta$ (see Fig. 1(a)) at the equilibrium O–O distance $R_{OO}^{eq} = 2.912$ Å found in CCSD(T) calculations[12] as well as derivatives of the angles with respect to the O–O distance, $d\alpha/dR_{OO}$ and $d\beta/dR_{OO}$, evaluated at $R_{OO}^{eq}$ by CCSD(T) within the present work. The derivatives of the dimer angles were not used in Ref. 6, but since in the present study we do not insist on reproducing $R_{OO}^{eq}$ exactly, it becomes important to have reasonable angles away from $R_{OO}^{eq}$.
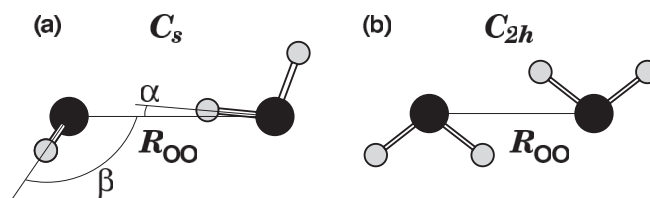


FIG. 1. Two structures of water dimer: (a) H-bonded dimer, point group $C_s$; and (b) planar dimer with antiparallel dipoles, point group $C_{2h}$. The latter has lower energy at short $R_{OO}$,[13] whereas the former provides the global energy minimum (see Fig. 2).

Another new quantity that we added to the objective function was the energy profile corresponding to proton transfer across the dimer. Proton transfer is of direct relevance to hydrogen diffusion in liquid water (the Grotthuss mechanism). Also, it has been discovered in quantum chemical calculations that the water dimer changes its structure at oxygen–oxygen separation shorter than 2.65 Å.[13] This feature was included into the fitting procedure.

The whole set of parameters found in the first stage is then passed to the second stage in which we determine the O–O pair potential. For that, we no longer use the energy vs. O–O distance curve in the water dimer. Instead, we fit the O–O RDF in liquid water directly.

This is a lengthy procedure as every calculation of RDF requires an MD simulation 40–45 ps long. This is the reason why one wants to separate the O–O pair potential from the rest of the model parameters in the first place. The genetic algorithms are not of much help here, instead we generate a few models of the O–O pair potential, obtain the respective RDFs and then "interpolate" between them in order to match the experimental RDF as closely as possible. The pair potential parameters that we thus guess, are used to generate another set of RDFs, and so on. Typically, it takes 3–4 iterations to arrive at a reasonable RDF.

Usually at this stage the density of water is already within 10% of the experimental one. To improve the density further, we take advantage of a novel feature of our model, namely, the fact that the O–O pair potential includes an attractive part decaying as $1/r^6$. This is further discussed in Sec. III. As for the density fitting, we simply fix the repulsive part of the O–O pair potential and change only the attractive part. Such "fine tuning" is possible because density is much more sensitive to the attractive part of the O–O pair potential than RDFs. As before, we create a few models, make NPT molecular dynamics simulations (see Sec. IV below) for each of them to estimate their respective densities, and "interpolate" to obtain the experimental density. This completes the whole cycle of the fitting.

## III. PARAMETERS OF THE MODEL

The complete set of parameters of the new water model is given in Table I. The parameters of the previous TB water model[6] are also shown in the table for comparison. Notation for scaling laws used in Table I ("GSP," "EPL," or "QUAD") is explained in Table II.

There are three major changes in the present model compared to Ref. 6. One is an additional $r^{-6}$ attractive term in

TABLE I. Parameters of the present TB model of water compared to those of the genetic dipole model of Ref. 6. See Ref. 1 for the meaning of parameters and abbreviations. Functions "GSP," "EPL," and "QUAD" denote a particular form of a scaling law and are explained in Table II. All values are given in atomic Rydberg units. Note that parameters of the present model are those listed in Table I of Paper I.[1]

| TB model | Present | | Ref. 6 | |
|---|---|---|---|---|
| | On-site parameters | | | |
| | O | H | O | H |
| $\varepsilon_s$ | –2.1164 | –1 | –1.51 | –1 |
| $\varepsilon_p$ | –1.1492 | | –1.20 | |
| Hubbard $U$ | 1.0775 | 1 | 1.16 | 1.08 |
| $\Delta_{spp}$ | –0.9430 | | –0.924 | |
| $\Delta_{ppd}$ | 0 | | 0 | |
| | Bond integrals, $V_{\ell\ell'm}$, and scaling | | | |
| | O–H | O–O | O–H | O–O |
| Function | GSP | GSP | GSP | GSP |
| $V^0_{ss\sigma}$ | –0.5018 | –0.0150 | –0.348 | –0.080 |
| $V^0_{sp\sigma}$ | | 0.0020 | | 0.050 |
| $V^0_{ps\sigma}$ | –0.4362 | –0.0020 | –0.313 | –0.050 |
| $V^0_{pp\sigma}$ | | 0.0500 | | 0.00012 |
| $V^0_{pp\pi}$ | | –0.0200 | | –0.004 |
| $n_{ss\sigma}$ | 2.0963 | 2 | 1.48 | 2 |
| $n_{sp\sigma}$ | | 2 | | 2 |
| $n_{ps\sigma}$ | 1.5019 | 2 | 1.98 | 2 |
| $n_{pp\sigma}$ | | 3 | | 3 |
| $n_{pp\pi}$ | | 3 | | 3 |
| $n_c$ | 4.0561 | 6 | 4.04 | 4 |
| $r_0$ | 1.8094 | 5.6 | 1.8094 | 5 |
| $r_c$ | 3.7985 | 9.0 | 3.47 | 6.8 |
| | Pair potentials, $\phi$, and scaling | | | |
| | O–H | O–O | O–H | O–O |
| Function | GSP | EPL | GSP | QUAD |
| $\phi^0$ | 0.73669 | | 0.552 | |
| $n$ | 3.3502 | | 3.362 | |
| $n_c$ | 6.3096 | | 6.04 | |
| $r_0$ | 1.8094 | 5.6 | 1.8094 | 5.992 |
| $r_c$ | 3.3550 | | | 3.04 |
| $\phi^0_1$ | | $4.0306 \times 10^{-3}$ | | 0.010 |
| $\phi^0_2$ | | $–2.0265 \times 10^{-3}$ | | 0.647 |
| $m_1$ | | 10 | | |
| $m_2$ | | 6 | | |
| $p_1$ | | 0 | | |
| $p_2$ | | 0 | | |
| | Cut-off distances $[r^{(1)}_{\text{cut}}; r^{(2)}_{\text{cut}}]$ | | | |
| | O–H | O–O | O–H | O–O[a] |
| $r^{(1)}_{\text{cut}}$ | 2.1 | 8 | 2.1 | 5.494 |
| $r^{(2)}_{\text{cut}}$ | 5.5 | 11 | 5.5 | 6.110 |

[a]Augmentative 5th order polynomial applied only to the O–O pair potential.

O–O pair potential. One can think of it as the term providing unaccounted for attraction between water molecules due to dispersion and induction forces. However, here we treat it as just another parameter of the model and use it for fitting rather than pick up some predefined value (as is done in recent "DFT+D" schemes[16]). This is consistent with our empirical tight binding approach in which we search for model parameters that provide better agreement with experiment rather than reproduce DFT results.

The second difference between the present and previous models is the change of the $s - p$ splitting for oxygen. Indeed, in the present model the difference between the oxygen on-site energies $\varepsilon_p$ and $\varepsilon_s$ is about 1 Ry and is close to the value that one would obtain using Hartree–Fock term values[17] ($\varepsilon_s$ = −2.142 Ry and $\varepsilon_p$ = −1.038 Ry). In the previous model, the $s - p$ difference was only 0.3 Ry and led to an incorrect electronic structure in metal oxides.[9]

As discussed in Ref. 6, the reduction of the $s - p$ splitting leads to an increase of the bond angle in a water molecule from 90° (which it would have had in the limit of the large splitting) towards its target value of 104.51° (see Table III). By imposing a lower bound on the $s - p$ splitting, we forced the genetic algorithm to search for other means to achieve the required bond angle, namely, via adjusting the O–H bond integrals and on site Hubbard–U parameters. As we shall see, the resulting model appears to outperform our previous model, which might seem paradoxical given that the new model is derived under additional constraints. The reason for that, in our belief, is that the constraint drove the system into another, more physical basin of the space of parameter values. A small price to be paid is that the model no longer renders correctly the ordering between the symmetric and asymmetric vibrational stretch frequences of the water monomer (see Sec. V).

The third modification to the model is the introduction of explicit cutoffs into all distance dependencies.[18] Given that the present model is intended for simulation of condensed phases, it is essential that at specified distances both pair potentials and bond integrals smoothly turn to zero. Otherwise, long MD runs might be blighted with energy leakage and even become unstable.

In addition to the above three changes, we introduced a short range repulsive H–H pair potential. There are no other bonded interactions between hydrogen atoms in our model since these are considered insignificant. The weak repulsion is added with the sole purpose of avoiding polarization catastrophes. Two H atoms that happen to be close enough during a simulation, may spontaneously pick up opposite charges. That would lead to the pair being pulled toward each other

TABLE II. Explicit form of the scaling laws referred to in Table I. The prefactor $A$ denotes $V^0_{\ell\ell'm}$ in case of bond integrals and $\phi^0$ in case of pair potentials.

| Notation | Function | Explicit form |
|---|---|---|
| GSP | Goodwin–Skinner–Pettifor[14] | $f(r) = A\,(r_0/r)^n \exp\{n[-(r/r_c)^{n_c} + (r_0/r_c)^{n_c}]\}$ |
| EPL | Exponential × power law | $f(r) = \sum_i A_i\,(r_0/r)^{m_i} \exp[-p_i(r - r_0)]$ |
| QUAD | Chadi's quadratic[15] | $f(r) = A_1\epsilon + A_2\epsilon^2$, where $\epsilon = (r - r_0)/r_0$ |

TABLE III. Properties of an isolated water molecule: partial charge on H atom, $\delta$ (in units of the proton charge), equilibrium O–H distance, $R_{OH}$, and HOH angle, $\theta$, vibrational force constants: symmetric stretch, $\nu_1$, bending, $\nu_2$, and asymmetric stretch, $\nu_3$, polarizability, $\alpha_{H_2O}$, dipole moment, $\mu_{H_2O}$, HOMO–LUMO gap, $E_{gap}$, and binding energy $E_b$. Target values are experimental data from the CRC Handbook[20] unless indicated otherwise. $e$ denotes the proton charge.

| | $\delta$ | $R_{OH}$ | $\theta$ | $\nu_1$ | $\nu_2$ | $\nu_3$ | $\alpha_{H_2O}$ | $\mu_{H_2O}$ | $E_{gap}$ | $E_b$ |
|---|---|---|---|---|---|---|---|---|---|---|
| | $(e)$ | (Å) | (deg) | | (Ry/Bohr$^2$) | | (Å$^3$) | (D) | (eV) | (eV) |
| TB (present) | 0.46 | 0.9580 | 104.46 | 1.037 | 0.092 | 0.935 | 1.348 | 1.843 | 11.06 | −11.6 |
| TB (Ref. 6) | 0.47 | 0.9575 | 104.26 | 1.029 | 0.065 | 1.061 | 1.470 | 1.858 | 8.98 | −10.3 |
| Target | | 0.9575 | 104.51 | 1.029[a] | 0.100[a] | 1.062[a] | 1.45 | 1.855 ± 0.005 | 12.31[b] | −12.4[c] |

[a]Reference 21.
[b]Reference 22.
[c]CCSD(T) results, present study. Target value for binding energy $E_b = -0.911$ Ry is the CCSD(T) binding energy of –0.735 Ry plus the spin polarization energy of O atom $E_{sp-pol}$ = −0.176 Ry (see text).

by resulting electrostatic attraction. A repulsive pair potential diverging faster than $1/r$ prevents this from happening (here, we used $1/r^{12}$). It will be straightforward if needed to include molecular hydrogen into the TB model.[19]

## IV. CALCULATION DETAILS

TB calculations are done with our empirical tight binding (TBE) computer code in which the self consistent polarizable ion method is implemented.[23] All the results reported below are obtained using the orthogonal TB method[8, 11] without spin polarization. Both spin polarization and non orthogonality are implemented in TBE and can be included if required. However, since that would increase the computational effort, we try to avoid unnecessary complications unless there is a strong reason to do otherwise.[24, 25]

For MD simulations we employ reversible integrators with Liouville operators.[26] A single Nosé–Hoover thermostat is used for the temperature control, to which the particles and the barostat are coupled with relaxation times of 0.1 ps and 0.4 ps, respectively. To ensure good energy conservation, a small time step of 0.5 fs is chosen in canonical (NVT) and microcanonical (NVE) simulations, whereas in isothermal–isobaric (NPT) runs it is reduced to 0.25 fs. Note that we do not replace hydrogen with deuterium, all the results refer to the light rather than heavy water. A typical MD simulation consists of 20 ps of equilibration followed by >100 ps of production run.

NVT and NPT simulations of liquid water described in Sec. VIII are done at temperatures between 230 K and 350 K with 128 water molecules in a cubic box of the size corresponding to a density of 1 g/cm$^3$ in NVT, or at an external pressure of $P = 0$ in NPT (a few tests at $P = 0.1$ MPa (1 atm.) showed no discernible differences). Ice was calculated in Sec. VII using an orthorhombic cell containing 8 water molecules (24 atoms).

Computations of water monomer, dimer, and hexamers are non periodic, whereas in periodic condensed phase simulations we use a 12×6×6 **k** −point mesh (ice) or only the $\Gamma$-point (liquid water) to sample the electronic Brillouin zone.

Additional quantum chemistry calculations are done at the MP2 and CCSD(T) levels as implemented in the GAMESS--US software package.[27] An augmented, correlation-consistent, valence polarization triple zeta (aug-cc-pVTZ) basis set is used throughout with the basis set superposition error corrected.

## V. WATER MONOMER AND DIMER

Tables III and IV present some properties of the water monomer and dimer as given by our new TB model and by the previous model of Ref. 6. All of these are fitted quantities (with the exception of atomic charges $\delta$ in Table III), for which the last line in the tables lists the target values.

As discussed in Sec. II, we do not put as much emphasis on monomer and dimer as we did in Ref. 6. Therefore, it is not surprising that some of the results worsen compared to the previous model. In particular, the present model fails to reproduce the peculiar ordering of force constants in the monomer $\nu_1 < \nu_3$. In addition, the present model underestimates the equilibrium O–O distance in the water dimer by 0.13 Å, or 4%. Other results represent an improvement over the previous model, such as the bending force constant $\nu_2$, and the dissociation energy of the dimer $E_{diss}$.

The binding energy $E_b$ of the monomer deserves a special comment. The CCSD(T) result, $-0.735$ Ry, is close to the experimental atomization energy at 0 K, $E_{at} = -0.6992$ Ry,[29] so either of these values could be used as a target. However, they both assume that $H_2O$ splits into isolated atoms *in their ground state*, which for O in particular implies the spin polarized state with multiplicity 3. Since in our non spin polarized calculations oxygen is non magnetic, we correct the above $E_b = -0.735$ Ry by the spin polarization energy of the oxygen atom $E_{sp-pol} = -0.176$ Ry and arrive at $-0.911$ Ry as a more suitable target value for $E_b$.

Figs. 2 and 3 refer to additional features included into the fitting: re-orientation of the dimer at short O–O distance

TABLE IV. Properties of the water dimer: equilibrium O–O distance, $R_{OO}$, dimer angles, $\alpha$ and $\beta$ (see Fig. 1(a)) evaluated at $R_{OO}$, and dissociation energy, $E_{diss}$. Target values are CCSD(T) results by Klopper *et al.*[12]

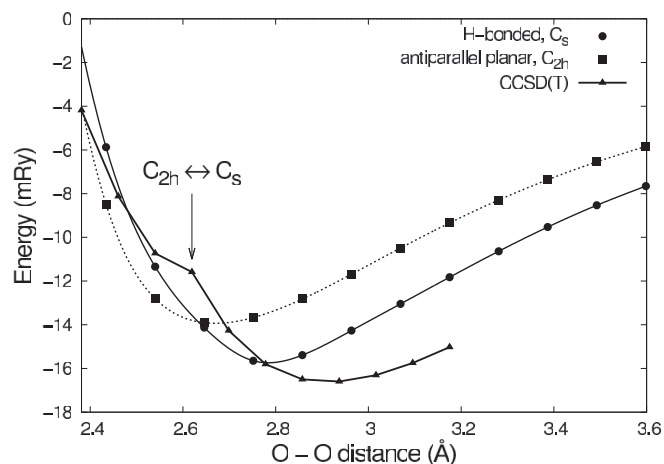| | $R_{OO}$ (Å) | $\alpha$ (deg) | $\beta$ (deg) | $E_{diss}$ (mRy) |
|---|---|---|---|---|
| TB (present) | 2.7851 | 2.3 | 124.9 | −15.8 |
| TB (Ref. 6) | 2.9153 | 3.0 | 113.7 | −15.1 |
| Target | 2.912 ± 0.005 | 5.5 | 124.4 | −16.0 ± 0.2 |

FIG. 2. Binding energy of the water dimer: H-bonded, point group $C_s$ (circles) and planar with antiparallel dipoles, point group $C_{2h}$ (squares). The CCSD(T) curve is shown with triangles, the kink on the curve indicates the transition between $C_s$ and $C_{2h}$ (shown with an arrow). According to Burnham and Xantheas,[13] a non planar version of $C_{2h}$ with point group $C_i$ becomes lowest in energy between 2.5 Å and 2.66 Å. However, we find it unnecessary to go into these fine details for fitting purposes and require only that the $C_{2h}$ ↔ $C_s$ transition is reproduced.

(Fig. 2) and the energy profile arising while the H-bonded proton is being pulled across the dimer (Fig. 3). The TB curves in both cases are compared to the CCSD(T) curves computed in the present study.

A change of the relative orientation of molecules in a water dimer at O–O separation less than 2.66 Å is predicted on the basis of MP2 calculations by Burnham and Xantheas.[13] In our own computations (MP2 relaxation followed by CCSD(T) total energy calculation) we also find that configuration $C_{2h}$, Fig. 1(b), becomes lower in energy than configuration $C_s$, Fig. 1(a), at 2.62 Å. This energy crossover was added to the list of properties to be fitted (see Sec. II) since it might af-
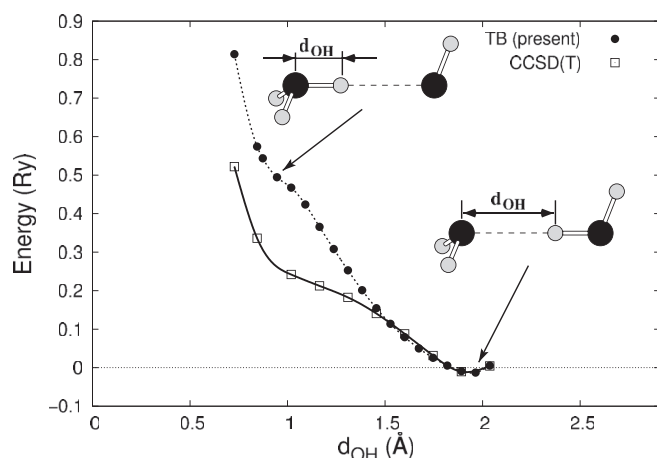


FIG. 3. Binding energy profile corresponding to the transfer of a proton in the H-bonded water dimer as given by the present TB model (circles) and by CCSD(T) (squares). The dimer is in configuration $C_s$ (Fig. 1(a)) at experimental $R_{OO}^{eq} = 2.912$ Å. The oxygen atom of the acceptor molecule is situated at the origin, the oxygen atom of the donor molecule is at the right end of the plot (2.912 Å). The proton moves along the O–O line from the donor molecule toward the acceptor molecule. Positions of the other atoms are fixed. Atomic configurations near the two ends of the curve are shown as two insets.
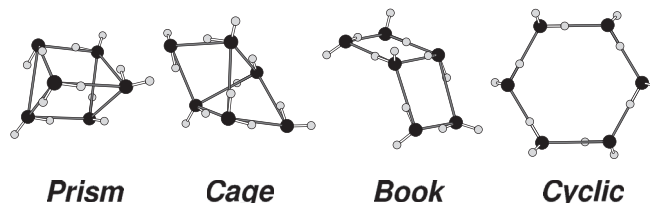


FIG. 4. Atomic configuration of water hexamers discussed in Sec. VI: the prism, the cage, the book, and the planar cyclic hexamer (the ring). The number of hydrogen bonds in these hexamers is 9, 8, 7, and 6, respectively.[28]

fect water–water interaction at short distances and possibly even the shape of the first peak in O–O RDFs. As one can appreciate from Fig. 2, the re-orientation distance is reproduced in the present TB model rather accurately.

The proton transfer curve in Fig. 3 was included in the fitting as this is a way to sample the hydrogen bond profile. The curves in Fig. 3 were obtained by placing the hydrogen atom along the O–O bond at various distances $d_{OH}$, with all other atoms frozen in their equilibrium positions in the $C_s$ dimer, Fig. 1(a). The TB curve closely follows the CCSD(T) curve in the vicinity of the hydrogen equilibrium and up to half way towards the acceptor molecule. From there, the two curves begin to deviate. However, this is not too alarming since the curves appreciably differ only at energies hardly accessible in an ordinary MD run.

The proton transfer curve can also be linked to the water self ionization reaction,

$$2H_2O \rightleftharpoons H_3O^+ + OH^- , \qquad (1)$$

except that the reaction products are not separated to infinity. That is the reason for the absence of a second, local minimum near the high energy end of the curve. The heat of reaction, $Q$, of (1) can be obtained by combining energies of the isolated molecules. The TB result is $Q = 0.929$ Ry (610 kJ/mol $H_2O$) compared to the CCSD(T) result, $Q = 0.719$ Ry (472 kJ/mol $H_2O$). It refers to water self ionization in vacuum at 0 K without any zero point energy correction. Although there is a 30% difference between the TB and CCSD(T) results, what is important is that charged molecules are satisfactorily dealt with in the present model, and that the large positive heat of the reaction (1) is reproduced.

It is also essential that neither of the above curves demonstrates any pathological features at short O–O separation. In the previous model,[6] on the contrary, there was a sudden energy jump in the proton transfer curve as the proton approached the acceptor molecule. It turned out that the three H and one O atoms formed an unphysical molecule, with oxygen charged positively and hydrogens charged negatively. Moreover, we think that the doublet structure of the first peak of the O–O RDF in liquid water originated in this feature. Pushing the O−$p$ orbital on-site energy down relative to the H−$s$ orbitals turns out to be sufficient to cure this problem.

## VI. WATER HEXAMERS

Water hexamers are the natural first objects to test the performance of our model aimed at condensed phase simulations. First, this is because the water–water bonding in liquid

TABLE V. Structure and energies of the four lowest energy water hexamers: prism, cage, book, and the planar cyclic hexamer depicted in Fig. 4. $R_{OO}$ is the distance between adjacent oxygens (in the planar cyclic hexamer, the length of all six O–O bonds are the same), $E_{diss}$ is the energy required to dissociate a hexamer into six water molecules, and $\Delta E_{diss} = E_{diss}$(hexamer) $- E_{diss}$(prism) are the relative energies of each hexamer with respect to the prism hexamer. Target energies are the CCSD(T) results obtained using the aug-cc-pVTZ basis set by Olson *et al.*[30] The target $R_{OO}$ distance in the planar cyclic hexamer is an MP2 result by Santra *et al.*[31]

| Property: | $R_{OO}$ (Å) | $E_{diss}$ (Ry) | $\Delta E_{diss}$ (mRy) | | | |
|---|---|---|---|---|---|---|
| Hexamer: | cyclic | prism | prism | cage | book | cyclic |
| TB (present) | 2.7346 | –0.141 | 0.0 | 1.7 | 7.0 | 13.0 |
| TB (Ref. 6) | 2.6506 | –0.144 | 0.0 | 2.8 | 6.9 | 11.8 |
| Target | 2.7069 | –0.154 | 0.0 | 0.88 | 3.87 | 6.82 |

water and ice is much closer to that in hexamers than in the dimer. Second, water hexamers are perhaps the largest water clusters that have been systematically studied in the literature, including accurate first principles computations.[28,30,31]

In Table V we compare results obtained with our present and previous water models to first principles data. The agreement between all three sets of data is very good. We should note that in both this and the previous model, the hexamers were not included in the fitting and the results in Table V are *predictions*. Importantly, both TB models give the correct energy ordering of the isomers: prism < cage < book < cyclic. These structures are illustrated in Fig. 4. The ordering is commonly believed to arise from the variation in the number of hydrogen bonds: 9 for the prism, 8 for the cage, 7 for the book, and 6 for the cyclic hexamer,[28] although, as noted by Santra *et al.*,[31] there is a certain ambiguity in this argument. Notice that DFT–GGA gives an incorrect ordering.[31]

## VII. ICE

The most common form of ice that we see outdoors in winter or inside an old fridge in the kitchen is hexagonal ice, Ih. Hexagonal ice is a proton disordered phase, but at low temperature (72 K for $H_2O$, 76 K for $D_2O$) the protons order, and ice Ih converts to the ferroelectric orthorhombic phase ice XI which is believed to be the lowest energy modification of ice at ambient pressure and 0 K. According to neutron diffraction experiments the space group of ice XI is $Cmc2_1$.[38,40]

In a recent study Hirsch and Ojamäe[34] considered 16 proton ordered structures that can be arranged in a 24 atom orthorhombic unit cell, including ice XI. It was found that density functional calculations with different types of basis sets and different flavors of generalized gradient approximation (GGA) do favor the ice XI structure, whereas force field models predict a different phase with $P2_12_12_1$ space group symmetry ("phase number 6" in Ref. 34) to be the lowest in energy.

Using our TB model we optimized the unit cell parameters and atomic coordinates of both ice polymorphs, ice XI and the hypothetical phase of Ref. 34. We also found the former to be lower in energy, as with the previous water model of Ref. 6. In Table VI we list the lattice energies and densities of ice obtained with TB along with those found in gradient corrected DFT calculations using GGA (PW91, PBE, BLYP), hybrid functionals (B3LYP, PBE0), meta-GGA (M06–L), and dispersion corrected functionals (BLYP+D). Judging from the table, there is no clear advantage in a particular type of exchange-correlation functional. Perhaps, B3LYP and M06–L provide a somewhat better agreement with experiment. The tight binding results appear in this respect at least as good as the best DFT ones. Besides, in the present TB model, liquid

TABLE VI. Lattice energy, $E_{lat}$, and density, $\rho$, of ice XI at 0 K. The "Hypothetical" phase is "phase number 6" from Ref. 34 which was found to be favored over ice XI in force field models. $\Delta E_{lat}$ is the difference between lattice energies of the hypothetical phase and ice XI. Densities of liquid water are also shown for comparison.

| | Ice XI (space group $Cmc2_1$) | | Hypothetical (space group $P2_12_12_1$) | | Liquid |
|---|---|---|---|---|---|
| | $\rho$ (g/cm³) | $E_{lat}$ (kJ/mol) | $\rho$ (g/cm³) | $\Delta E_{lat}$ (kJ/mol) | $\rho$ (g/cm³) |
| TB (present) | 0.968 | –54.26 | 0.940 | 0.21 | 1.008 |
| TB (Ref. 6) | 0.967 | –51.08 | | 0.67 | 0.926 |
| DFT–GGA (PW91) | 0.995[a] | –68.66[a] | 0.969[a] | 0.96[a] | |
| DFT–GGA (BLYP) | | –55.85[b] | | 0.71[b] | 0.92[c] |
| DFT–GGA (BLYP+D) | | –69.434[e] | | 0.616[e] | 1.07[c] |
| DFT–GGA (PBE) | 0.989[d] | –66.74,[d] –67.901[e] | | 0.571[e] | |
| DFT–GGA (M06–L) | 0.953[d] | –61.04[d] | | | |
| DFT–GGA (B3LYP) | 0.940[d] | –59.29[d] | | | |
| DFT–GGA (PBE0) | 0.981[d] | –63.76[d] | | | |
| Expt. | 0.935[f] | –58.87,[g] –59.25[h] | | >0 | 0.997 |

[a]Reference 34.
[b]Reference 34, structural optimisation only.
[c]Reference 35.
[d]Reference 36.
[e]Reference 37. Results slightly depend on the computer program used. We cite only those obtained with the CP2K code.
[f]Measurements at 5 K.[38]
[g]Experimental estimation at 0 K with zero point energy removed, see Refs. 36 and 39 and references therein.
[h]Estimated using the lattice energy of ice Ih and the enthalpy of the ice Ih → ice XI transition, see Ref. 37.

water is denser than ice, which was an essential requirement of our model (see Sec. II).

## VIII. LIQUID WATER

We move now to the model predictions for liquid water. This is the most critical section of our results as liquid water is where the majority of the model applications are expected to lie. MD simulations reported below were done both in NVT and NPT ensembles. In most cases both yielded similar results. Hence, in this section we present results of NPT simulations unless indicating otherwise.

### A. Radial distribution functions and the "fifth neighbor"

In Fig. 5 we show the radial distribution functions (RDFs) $g(r)$ obtained after NPT simulation at $T = 300$ K and $P = 0$ together with those obtained with the previous model[6] and by neutron diffraction.[32]

The theoretical O–O curve (thin solid line in the top plot of Fig. 5) provides an excellent match to the position of all three maxima of $g_{OO}(r)$ and reasonable match for the position of two minima between them. The first peak is more localized than in experiment but is correctly shaped, which we consider as a significant improvement over the one predicted by our previous model with a shoulder and displaced maximum (dotted line in the same figure). A nicer agreement of the O–O RDF is not just a coincidence, of course, as it was included into the fitting procedure (see Sec. II).

The evolution of oxygen–oxygen RDF with temperature is shown in Fig. 6 together with available experimental RDFs. We observe that $g_{OO}(r)$ at $T = 270$ K, again has a narrower first peak than its experimental counterpart, but starting from the second peak the TB curve remains in perfect agreement with experiment. Increasing the temperature smears the structure of $g_{OO}(r)$, although there is a larger change in going from 270 K to 300 K, than from 300 K to 330 K. The smearing means that the maxima decrease in height and widen; as a result the first minimum moves to larger distances. In addition, the second maximum also slightly moves in the same direction, as has been noted in other simulations (see Ref. 54 and references therein).

One of the characteristic features of the structure of liquid water is a non integer average coordination number $N_c$ of water molecules. If water had a perfect tetrahedral coordination then $N_c$ would be equal to four (as is the case of ice XI). Experiment however predicts a fractional number between 4 and 5. For instance, a recent estimation by Soper and Benmore[41] gives $N_c = 4.67 \pm 0.05$. The reason behind $N_c$ being larger than four is a subject of extensive discussion in the literature, and is often referred to as the "fifth neighbor" problem.[49]

By integrating the RDF up to the first minimum, in our TB model we obtain $N_c = 4.50$, in good agreement with the experimental estimation by Soper and Benmore[41] (see Fig. 7 and Table VII). One might argue that this agreement is simply the consequence of $g_{OO}(r)$ being part of the fitting. We do not think so since our first peak differs in shape from the experiment.
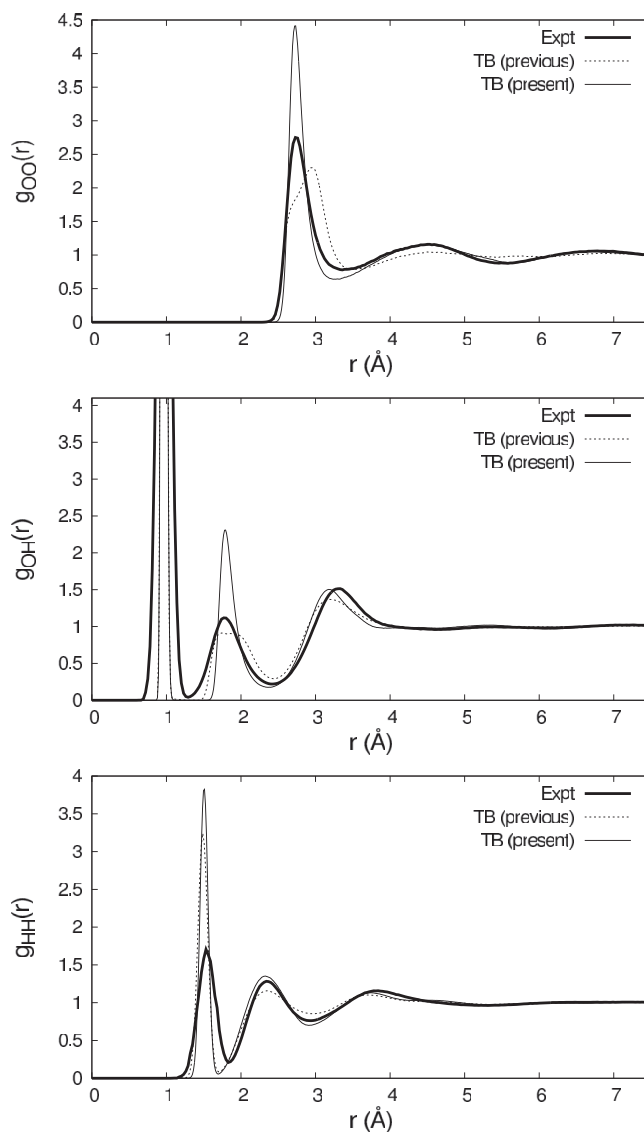
FIG. 5. Radial distribution functions $g(r)$ in liquid water for O–O, O–H, and H–H neighbors obtained with the present TB model (thin line) and with the "genetic" dipole model from Ref. 6 (dotted line). The experimental curves (thick line) are from Refs. 32 and 33 and correspond to $T = 298$ K and $P = 0.1$ MPa. Theoretical curves are the results of an NPT simulation at $T = 300$ K and $P = 0$.

However, we might have a good $N_c$ for a different reason, namely, because the *density* of liquid water is included into the fitting. As a matter of fact, we kept tuning the strength of our "dispersion term" until the density of liquid water became $\approx 1$ g/cm$^3$ (see Sec. II and Table VII), whereas according to Wang *et al.*,[50] it is the dispersion interaction that appears responsible for the coordination number.

Given the weakness of the dispersion forces, it is not surprising that the temperature dependence of $N_c$ is so strong that $N_c$ already approaches that of ice at the left edge of the plot in Fig. 7.

### B. The band gap of liquid water

If we define "band gap" as the energy difference between highest unoccupied and lowest occupied states,[55] averaged
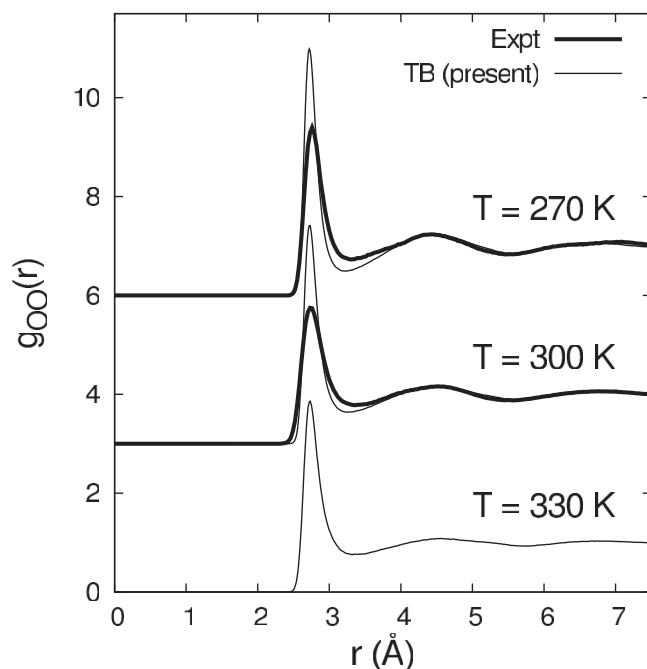
FIG. 6. O–O radial distribution function at different temperatures. TB curves (thin lines) are obtained in NPT simulations at $T = 270$, 300, and 330 K at $P = 0.1$ MPa. Experimental curves (thick lines) are measured at $T = 298$ K, $P = 0.1$ MPa, and $T = 268$ K, $P = 27$ MPa.[32,33] Subsequent sets of curves are shifted vertically by 3 units for clarity. RDFs at standard conditions (the middle set) repeat curves in the top panel of Fig. 5.
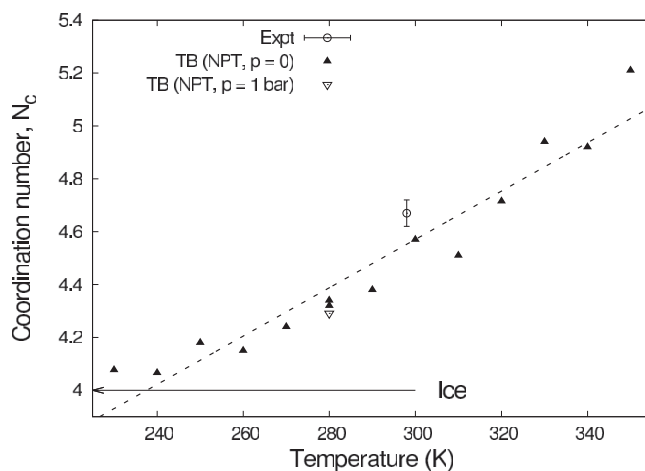


FIG. 7. Temperature dependence of the average O–O coordination number $N_c$ in liquid water: TB results in the NPT ensemble (triangles) and experiment (a circle).[41] The dotted line is a linear fit of $N_c(T)$ and is used to obtain the coordination number at room temperature (see Table VII). The horizontal arrow indicates the coordination number in ice, which is always four apart from amorphous ice and some exotic high pressure phases.[52]

over an NPT MD simulation, then we can examine this quantity as a function of temperature. The result is shown in Fig. 8 with a comparison to an experimental assessment at room temperature. Our predictions may also be compared to *ab initio* zero temperature theoretical calculations[22] which result in a HOMO–LUMO gap of 10.4 eV liquid in water, and 12.31 eV for the gas phase monomer (see Table III). To some

extent the band gap obtained in this way is inherited from the HOMO–LUMO gap of the molecule given in Table III, but since this was not fitted we can regard the result for liquid water as a prediction. However, the relation between HOMO–LUMO gap and an experimental quantity is not at all clear; neither is the one electron picture particularly useful since it does not capture excitonic effects. But what *is* of interest is the temperature dependence: and since this is quite marked, it is valuable to show the result here possibly to provoke further examination of this question. Fig. 8 indicates a large variance in the band gap arising from fluctuations in the atomic positions. It is clear from the data from the monomer and from ice XI that these fluctuations, as expected, are temperature

TABLE VII. Properties of liquid water predicted in our present and previous (Ref. 6) TB models at ambient conditions ($P = 0$, $T = 300$ K): $N_c$ is the average O–O coordination number, $\bar{p}_{tot}$, $\bar{p}_{pc}$ and $\bar{p}_{ind}$ are the average total, point charge, and induced dipole moments of a water molecule, $\epsilon_0$ is the static dielectric constant, and $D_{self}$ is the self diffusion coefficient. TB values for the present model are interpolated into $T = 300$ K using the whole set of data between 260 K and 350 K. Available experimental and DFT–GGA data are also shown for comparison. Diffusion coefficients in parenthesis correspond to $D_2O$ rather than $H_2O$. Values for the dipole moments and dielectric constant in the third line correct erroneous values reported in Ref. 6.

|  | $N_c$ | $\bar{p}_{tot}$ (D) | $\bar{p}_{pc}$ (D) | $\bar{p}_{ind}$ (D) | $\epsilon_0$ | $D_{self}$ ($10^{-5}$ cm²/s) |
|---|---|---|---|---|---|---|
| Expt. | $4.67 \pm 0.05$[a] | $2.95 \pm 0.2$[b] | ... | ... | 78[c] | $2.23 \pm 0.1$[d] (1.9[e]) |
| TB (NVT)[f] | 4.54 | 2.40 | 3.07 | 0.67 | 43.6 | 1.98 |
| TB (NPT)[f] | 4.50 | 2.40 | 3.07 | 0.67 | 43.0 | 2.13 |
| TB (NPT)[g] | 5.0 | 2.55 | 3.28 | 0.74 | 42.7 | 3.0 |
| TB (NPT)[h] |  | 2.51 | 2.51 |  | 58.7 | 3.5 |
| DFT–GGA (PBE) |  | 2.95,[i] 3.09[j] | ... | ... | $67 \pm 6$,[j] 75[k] | (1.6[j]) |
| DFT–GGA (BLYP) |  |  |  |  |  | 0.25,[l] 0.55[m] |
| DFT–GGA (BLYP+D) |  |  |  |  |  | (1.7[n]) |
| DFT–GGA (DRSLL) | 4.90[o] |  |  |  |  | (2.63[o]) |
| DFT–GGA (DCACP) |  |  |  |  |  | ($2.1 \pm 0.23$[p]) |

[a]298 K, Ref. 41.
[b]300 K, Ref. 42.
[c]300 K, Ref. 43.
[d]298 K, Ref. 44.
[e]298 K, Ref. 45.
[f]Present study.
[g]"Genetic model" of Ref. 6.
[h]"Point charge model" of Ref. 6.

[i].318 K, Ref. 3.
[j]330 K, Ref. 46.
[k]Extrapolated to 300 K from data in Ref. 46.
[l].308 K, Ref. 47.
[m]305 K, Ref. 48.
[n]317 K, Ref. 49.
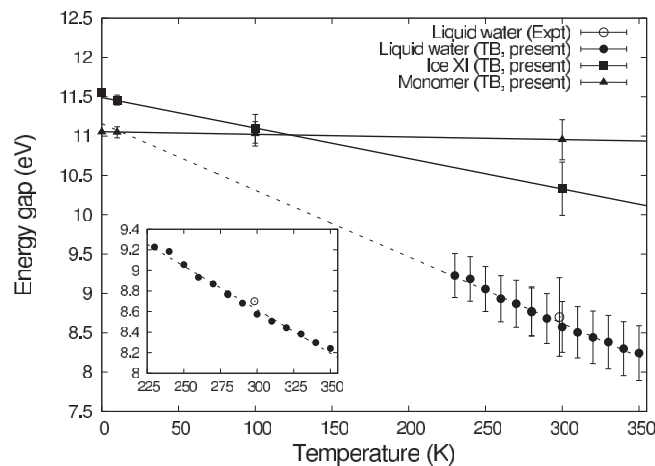[o]300 K, Ref. 50.
[p]325 K, Ref. 51.

FIG. 8. HOMO–LUMO gap in liquid water as a function of temperature. The experimental data point is taken from Ref. 53. The remaining data are taken from our TB simulations; vertical "error bars" show the variance in the energy gap taken over the post-equilibration stage of the simulation. Zero temperature quantum chemistry calculations[22] predict gaps of 10.4 eV liquid in water, and 12.31 eV in the monomer.
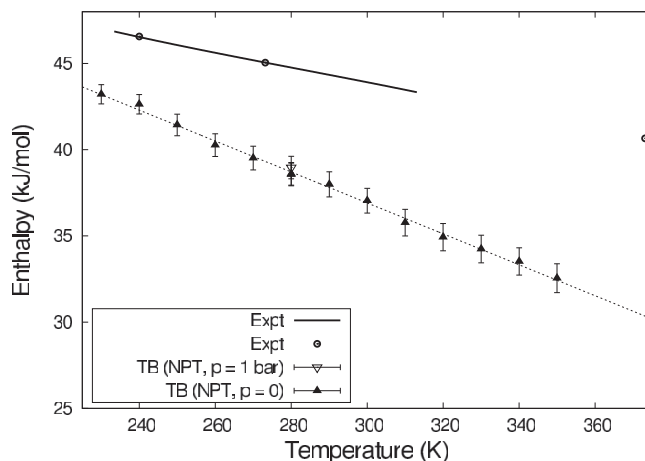


FIG. 9. Enthalpy of vaporisation of liquid water as a function of temperature. The absolute value is underestimated as compared to experiment and discussed in the text. The slope, which is the heat capacity, is overestimated and this is due to the neglect of quantum effects on the low temperature heat capacity.

dependent. On the other hand, the band gap in the monomer is practically temperature independent, while in ice XI the temperature dependence is nearly as large as in the liquid. It is possible that the additional temperature dependence is due to the creation of Urbach tails in the liquid,[53] which are absent in the solid.

### C. Heat of vaporization

The heat of vaporization, $H_{vap}$, is a thermodynamic quantity used to characterize liquid water. It represents the enthalpy difference between liquid and vapor phases and can be extracted from an MD simulation as

$$H_{vap} = -\left(\frac{\langle U \rangle}{N} - U_1\right) + k_B T$$

where $\langle U \rangle$ is the average internal energy of an assembly of $N$ water molecules, $U_1$ is the internal energy of a single water molecule, and $k_B$ is the Boltzmann constant.

The results of our NPT simulations are shown in Fig. 9; interpolated to $T = 298.16$ K these give $H_{vap} = 8.85$ kcal/mol (37.05 kJ/mol or 28.21 mRy/f.u.) which compares favorably with the experimental value of 10.52 kcal/mol.[56] TB models based on the SCC–DFTB method also underestimate $H_{vap}$.[57] Maupin *et al.*[57] argue that the low heat of vaporization in the original SCC–DFTB water model, $4.09 \pm 0.04$ kcal/mol, is a consequence of underestimation of hydrogen bonding in the model and propose an improved model, HBD–SCC–DFTB, with $H_{vap}$ increasing to $7.0 \pm 0.08$ kcal/mol. A similar situation is reported for water simulations employing semi empirical Hamiltonian models, such as AM1 and PM3. The original AM1 and PM3 models underestimate $H_{vap}$, resulting in $7.67 \pm 0.56$ kcal/mol and $4.68 \pm 0.84$ kcal/mol, respectively.[58] Monard *et al.*[58] relate this fact to an insufficiently accurate description of the hydrogen bond, and polarization effects in liquids in particular. They demonstrate that one of the improved models, namely PM3-PIF, results in

$H_{vap} = 9.02 \pm 0.69$ kcal/mol. Classical force field water models, such as SPC and TIP3P, perform much better in this respect and provide $H_{vap}$ within a few percent of experimental value.[59] This is because the room temperature heat of vaporization is usually explicitly included into the fitting set.[60]

### D. The distribution of electric dipoles in liquid water

A peculiar property of water, which must arise from the polarizability of the water molecule, is that the dipole moment of individual molecules *increases* as molecules assemble into first the dimer, then small clusters and finally into bulk phases, liquid and ice.[6,55] In our TB model, the dipole moment of the water monomer is fitted to 1.86 D, and the average dipole moment in the liquid at 300 K is then found to be 2.40 D. The dipole in the liquid is smaller than predicted using DFT, and it is smaller than was claimed in Ref. 6 due to an error in extracting the induced dipole moments from the trajectories. Corrected data using the previous model, and values using our revised model are collected in Table VII and Fig. 10. Our predicted dipole moment in ice XI is 2.59 D; this is consistent with an estimate of 2.60 D made by Coulson and Eisenberg based in classical electrostatics.[61] However, their calculation was based on a disordered proton state—the ordered state would have a larger local moment—and we believe therefore that the TB model fails to capture fully the mutual and collective polarizability in water, and the complete enhancement of the molecular dipole is not achieved.

The physics behind the dipole enhancement in condensed phases is much as explained in Ref. 6 but the details are now properly presented in Fig. 10. First, the dipole *due to point charges alone* is increased from 2.60 D in the monomer to 3.07 D in the liquid; this effect and its temperature dependence is illustrated in Fig. 11. The average dipole moment *induced* on the oxygen anion on account of its atomic polarizability (captured in PITB through the crystal field term) is a little smaller in the liquid than in the gas phase. In the *monomer* the induced dipole points exactly antiparallel to the
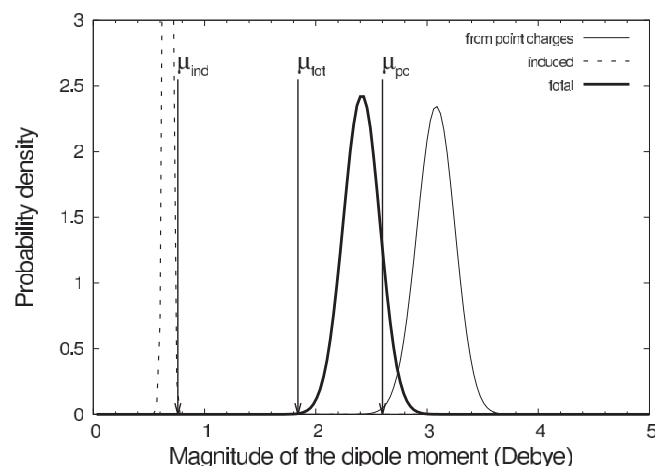
FIG. 10. Distribution of the magnitude of the dipole moment of water molecules at $T = 300$ K. Dipole due to point charges (thin line) and induced dipole of oxygen atoms (dotted line) combine vectorially to form the total dipole moment magnitude (thick line). The induced dipole moment varies little from the monomer value ($\mu_{ind} = 0.76$ D) while the point charge dipole is increased compared to that of the water monomer ($\mu_{pc} = 2.60$ D), and the average total dipole is increased by approximately the same amount (see Sec. VIII D for discussion).

point charge dipole, leading to a maximal cancelation, resulting in the final moment of 1.86 D. (See Fig. 1, Ref. 6.) Conversely in the *liquid*, we find that the cancelation is partly suppressed. The incomplete cancelation then leads to the larger total dipole moment of 2.40 D in the liquid, Fig. 10.

## E. Dielectric constant

The large static dielectric constant, $\epsilon_0$, is a distinguishing feature of liquid water, partly responsible for its acting as a rather universal solvent. We calculate it as

$$\varepsilon_0 = 1 + \frac{4\pi}{3} \frac{\langle \mathbf{P}^2 \rangle}{V k_{\mathrm{B}} T},$$
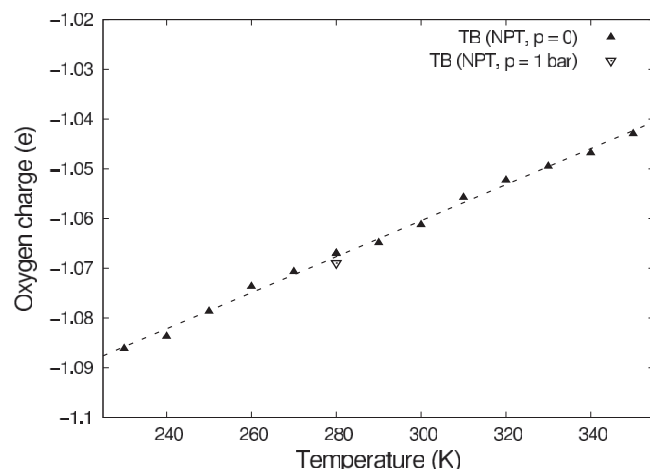


FIG. 11. Average charge on oxygen atoms, $q_O$, in liquid water as a function of temperature. Dotted line gives the linear fit of $q_O(T)$ dependence, and $e$ is the proton charge. For comparison, the oxygen charge in a gas phase water molecule is $-0.92\ e$ (see Table III).
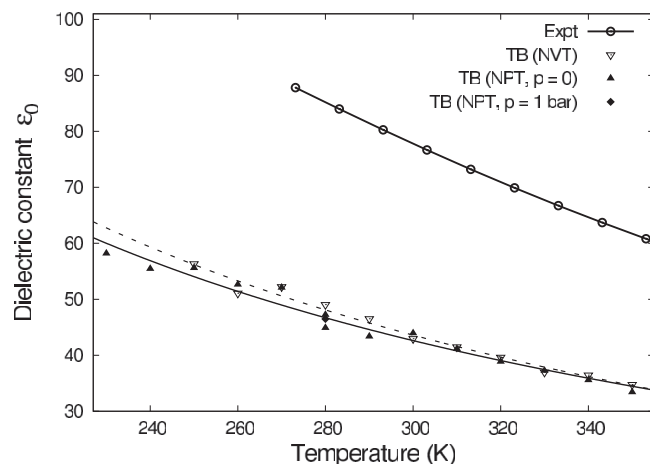


FIG. 12. Temperature dependence of the static dielectric constant $\varepsilon_0$ obtained in both NVT and NPT MD simulations, and comparison to experiment.[62] The lines through the calculated data are inverse temperature fits $f(T) = a + b/T$ to TB points in the interval [230 K, 350 K].

where $\mathbf{P}$ is the dipole moment of a simulation cell having average volume $V$. We employ the method proposed by Sharma *et al.*[46] which relies on the radial integration of the dipole–dipole pair correlation function. The resulting dielectric constant as a function of temperature is shown in Fig. 12 together with its least square fit in the form $\varepsilon_0(T) \simeq a + b/T$. PITB models underestimate the dielectric constant by a factor of about two, while the trend in the temperature dependence is rendered well. This is contrary to the claim made in Ref. 6 in which the dielectric constant was calculated incorrectly due to the same error as was made in extracting the average dipole moment. (Table VII shows corrected data for the genetic dipole model of Ref. 6, whereas there was no error in calculating the point charge model.)

It is not surprising, despite our earlier claim,[6] that the TB model underestimates the polarizability and hence the dielectric constant of water. As other authors have pointed out, a minimal basis set is certain to be incomplete and unable to capture atomic polarization correctly. This is clearly indicated by examining the theory of the Stark effect[63] in which the polarizability of a hydrogen atom is calculated to second order in perturbation theory. It is necessary to extend the sum over both bound and unbound states of the hydrogen atom to obtain a correct answer.[64] Therefore, underestimation of dielectric constant is expected to be an inevitable limitation of tight binding models, and indeed also of localized basis DFT methods.

## F. Self diffusion coefficient

We have used NPT simulations to calculate the self diffusion coefficient, $D_{\mathrm{self}}$, of water. A remark on our choice of the MD ensemble seems required at this point. Being a dynamic rather than static property, $D_{\mathrm{self}}$ must depend on the particular ensemble used in MD simulation. Most diffusion simulations in the literature are done in an NVE ensemble since the alternatives, such as NVT or NPT ensembles, might involve rescaling of the time variable. However, as Tuckerman

points out,[67] it is possible to avoid any time rescaling in both NVT and NPT simulations, which is the way the MD routine is programmed in our TBE code. This opens the possibility to simulate diffusion at specified temperature and pressure, and therefore much better represents experimental conditions. We also verified in an explicit computation that diffusion coefficients obtained in the NVT and NPT ensembles are in agreement with those obtained in a standard NVE ensemble at temperatures between 280 K and 310 K.

Our calculated coefficients of self diffusion are listed in Table VII together with experimental data and DFT results. The agreement between TB and experiment is astonishingly good, much better than the agreement between experiment and DFT–GGA. It is widely recognized that in order to bring DFT–GGA into agreement with experiment one needs to rescale the temperature by about 20%.[48] The situation dramatically improves, however, in the dispersion-corrected GGAs (see the BLYP+D, DRSLL, and DCACP results in Table VII).

The temperature dependence of $D_{self}$ is shown in Fig. 13 together with experimental data. We use an Arrhenius plot of $\log D_{self}$ vs. $1/T$ in which a straight line corresponds to a thermally activated process with neither energy barrier nor attempt frequency being temperature dependent. Experiment suggests that water should be slightly non Arrhenius: indeed, the curve is steeper at low temperatures.[44] Mills[45] provides the following estimates for the diffusion activation energy: $E_a = 20$ kJ/mol for temperatures between $1\,°C$ and $15\,°C$, and $E_a = 18$ kJ/mol in the interval from $15\,°C$ to $45\,°C$. It is encouraging that the TB data points match experiment rather closely in the whole temperature range (see Fig. 13). The scatter of TB data due to the size of the system and the length of MD run is too large to resolve the curvature on the Arrhenius plot. However, if we ignore the curvature and fit a straight line through data points (dotted line in Fig. 13), then we arrive at the activation energy $E_a = 19$ kJ/mol, in excellent agreement with Mills's estimations.
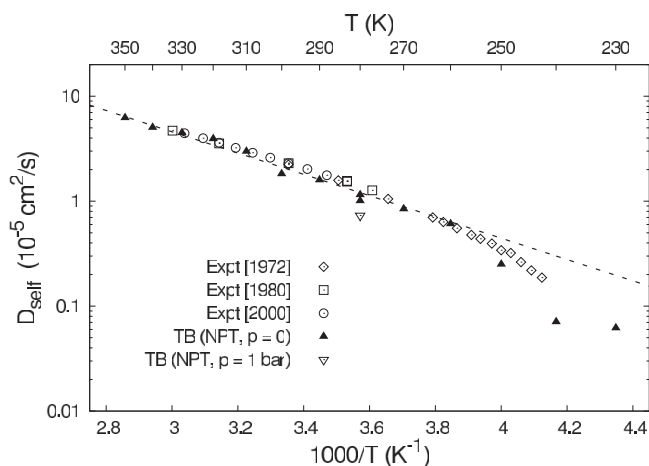


FIG. 13. Arrhenius plot of the self diffusion coefficient $D_{self}$ in liquid water obtained in TB NPT simulations (filled circles). Experimental points are from Ref. 44 (open triangles), Ref. 65 (open squares), and Ref. 66 (open circles). Dotted line represents the Arrhenius fit to TB diffusion data in the temperature interval between 230 K and 350 K. The slope of the line corresponds to diffusion activation energy $E_a = 19$ kJ/mol.
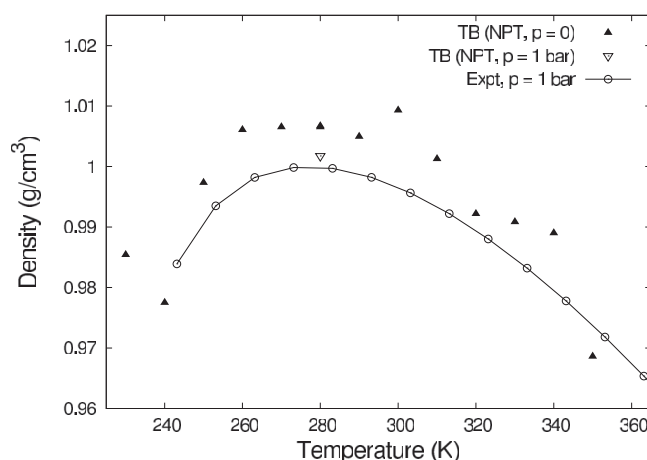


FIG. 14. Temperature dependence of the density of liquid water given by TB (squares) and experiment (circles).[20]

## G. The density maximum of water

One of the most striking and intriguing properties of liquid water is the fact that water density has a maximum at $3.98\,°C$. The reasons behind this phenomenon are still being debated. One of the suggestions[68] is that the maximum is a consequence of the thermodynamic singularity, "the second critical point" of water at temperature $-45\,°C$ which is experimentally unreachable.

In order to test whether our TB model is able to reproduce the density maximum, we plot the average density of liquid water NPT simulations between 230 K and 350 K in Fig. 14. Although the arrangement of the data points might indeed hint at the existence of the maximum, the scatter of the points is too big for this result to be conclusive.

However, due to our current efforts in optimization and efficient parallelization of the TBE program, combined with the progress in computer performance, we expect a sufficiently accurate simulation to become feasible within a year or so. If we find that the density–temperature dependence is indeed rendered properly, then having a highly transferable TB model with controllable parameters might help to resolve the long standing density maximum dispute.

## IX. DISCUSSION AND CONCLUSIONS

We have revised an earlier model for water in order to create a transferable polarizable ion tight binding scheme suitable for organic molecules in gas phase and aqueous solution and also as substrates on oxide surfaces. We have given extensive details here and in Paper I[1] of the fitting procedure, first in order to impress upon the reader the complexity of the problem and second to provide insights into influences upon the properties of the model of individual variations in the parameters. The PITB theory leads to a model that gives a very good account of many measured properties of liquid water, notably the radial distribution functions, heat of evaporation, density, and self diffusivity. In spite of claims made earlier by two of us[6] the PITB fails to account for all of the polarizability in the liquid. This is seen in an underestimation of the collective enhancement of the molecular dipole moment, and in an

underestimation of the dielectric constant by a factor of two. The relative densities of water and ice are properly rendered in the model and there are indications that the density maximum in water is reproduced, although longer simulations in larger systems will be needed to confirm this.

The TB theory, through its construction and parameterization, furnishes us with some insights into the hydrogen bond. Because the model predicts features of water, in particular the binding of hexamers, self diffusivity, and enthalpy of evaporation, we have some confidence that the hydrogen bond is being described faithfully. It is therefore valuable to assess the essential components of the bond as they appear in the TB theory. It is very striking that the O–O bond integrals and repulsive pair potential are the same as we use in titanium dioxide in Paper III[9] and similar to those used earlier for zirconium dioxide.[10] In the oxides the dominant terms are the first neighbor transition metal–oxygen bond integrals and pair potentials; second neighbor O–O bonds can be neglected in a first approximation.[17] It is notable that typical O–O bond lengths in transition metal oxides are comparable to the O–O bond lengths in the water dimer, clusters, and condensed phases. Therefore, it is clear that a contributor to the hydrogen bond is a weak O–O covalent bond. Second, as Coulson points out,[61] the electrostatic attraction arising from the static, polarizable multipole moments of the molecular charge is a significant component, and this is rendered in the TB model through both the self consistent point charges arising from the Hubbard energy and the ionic polarization driven by the crystal field effect. (The latter is equally important and described by the same physics in the bonding in transition metal oxides.[7]) Finally, it is accepted from DFT studies[69] that the London dispersion forces are essential to complete the picture of the hydrogen bond, and these are described in the TB model as an *ad hoc* attractive pair potential, scaling as the inverse sixth power of the bond length, between oxygen atoms.

In Paper III[9] it is shown that the PITB water model can provide insights into the behavior of water molecules on surfaces of titanium dioxide, including a contribution to the question of under which conditions a surface may cause a dissociation of water.

We expect that models of the type presented here, having the freedom to describe charged species, bond breaking, self diffusion and the Grothuss mechanism, in both extended condensed systems and the gas phase, will provide the way forward to large scale studies in electrochemistry among other branches of physical chemistry and chemical physics.

[1]T. J. Sheppard, A. Y. Lozovoi, D. L. Pashov, J. J. Kohanoff, and A. T. Paxton, J. Chem. Phys. **141**, 044503 (2014).

[2]M. Del Ben, M. Schönherr, J. Hutter, and J. VandeVondele, J. Phys. Chem. Lett. **4**, 3753 (2013).

[3]P. L. Silvestrelli and M. Parrinello, Phys. Rev. Lett. **82**, 3308 (1999).

[4]H. J. C. Berendsen, J. R. Grigera, and T. P. Straatsma, J. Phys. Chem. **91**, 6269 (1987).

[5]W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, and M. L. Klein, J. Chem. Phys. **79**, 926 (1983).

[6]A. T. Paxton and J. J. Kohanoff, J. Chem. Phys. **134**, 044130 (2011).

[7]M. W. Finnis, A. T. Paxton, M. Methfessel, and M. van Schilfgaarde, Phys. Rev. Lett. **81**, 5149 (1998).

[8]A. T. Paxton, in *Multiscale Simulation Methods in Molecular Sciences*, NIC Series, Vol. 42, edited by J. Grotendorst, N. Attig, S. Blügel, and D. Marx (Institute for Advanced Simulation, Forschungszentrum Julich, 2009), pp. 145–174. Available on-line at http://webarchiv.fz-juelich.de/nic-series/volume42/volume42.html.

[9]A. Y. Lozovoi, D. L. Pashov, T. J. Sheppard, J. J. Kohanoff, and A. T. Paxton, J. Chem. Phys. **141**, 044505 (2014).

[10]S. Fabris, A. T. Paxton, and M. W. Finnis, Phys. Rev. B **63**, 941011 (2001).

[11]M. W. Finnis, *Interatomic Forces in Condensed Matter* (Oxford University Press, Oxford, UK, 2003).

[12]W. Klopper, J. G. C. M. van Duijneveldt-van de Rijdt, and F. B. van Duijneveldt, Phys. Chem. Chem. Phys. **2**, 2227 (2000).

[13]C. J. Burnham and S. S. Xantheas, J. Chem. Phys. **116**, 1479 (2002).

[14]L. Goodwin, A. J. Skinner, and D. G. Pettifor, Europhys. Lett. **9**, 701 (1989).

[15]D. J. Chadi, Phys. Rev. Lett. **41**, 1062 (1978).

[16]S. Grimme, J. Antony, S. Ehrlich, and H. Krieg, J. Chem. Phys. **132**, 154104 (2010).

[17]W. A. Harrison, *Electronic Structure and the Properties of Solids, the Physics of the Chemical Bond* (W. H. Freeman and Company, San Francisco, 1980).

[18]The cutoff functions used in the present study are fifth order polynomials $P_5(x)$ that produce a smooth step between 1 and 0 with continuous second derivatives:

$$P_5(0) = 1, \qquad P_5'(0) = P_5''(0) = 0,$$

$$P_5(1) = 0, \qquad P_5'(1) = P_5''(1) = 0.$$

Any function $f(r)$ can be smoothly cut to zero between cutoff radii $r_{cut}^{(1)}$ and $r_{cut}^{(2)}$ as $\tilde{f}(r) = f(r)\, P_5[(r - r_{cut}^{(1)})/(r_{cut}^{(2)} - r_{cut}^{(1)})]$. It is straightforward to demonstrate that function $\tilde{f}(r)$ and its first two derivatives coincide with those of $f(r)$ at $r < r_{cut}^{(1)}$, are equal to 0 at $r > r_{cut}^{(2)}$, and are continuously differentiable up to the second order in the whole interval where $f''(r)$ exists and is continuous.

[19]A. J. Skinner and D. G. Pettifor, J. Phys.: Condens. Matter **3**, 2029 (1991).

[20]*CRC Handbook of Chemistry and Physics*, 83rd ed., edited by D. R. Lide (CRC Press, Boca Raton, FL, 2002).

[21]G. Herzberg, *Molecular Spectra and Molecular Structure II: Infrared and Raman Spectra of Polyatomic Molecules* (Krieger Publishing, Malabar, FL, 1991).

[22]P. Cabral do Couto, S. G. Estcio, and B. J. Costa Cabral, J. Chem. Phys. **123**, 054510 (2005).

[23]Empirical Tight Binding computer program (TBE) is a part of the LM electronic structure package (http://www.lmsuite.org/).

[24]A. T. Paxton and M. W. Finnis, Phys. Rev. B **77**, 024428 (2008).

[25]A. T. Paxton and C. Elsässer, Phys. Rev. B **87**, 224110 (2013).

[26]G. J. Martyna, M. E. Tuckerman, D. J. Tobias, and M. L. Klein, Mol. Phys. **87**, 1117 (1996).

[27]P. Piecuch, S. A. Kucharski, K. Kowalski, and M. Musial, Comput. Phys. Commun. **149**, 71 (2002).

[28]S. S. Xantheas, C. J. Burnham, and R. J. Harrison, J. Chem. Phys. **116**, 1493 (2002).

[29]See http://cccbdb.nist.gov for the experimental atomization energy of water monomer.

[30]R. M. Olson, J. L. Bentz, R. A. Kendall, M. W. Schmidt, and M. S. Gordon, J. Chem. Theory Comput. **3**, 1312 (2007).

[31]B. Santra, A. Michaelides, M. Fuchs, A. Tkatchenko, C. Filippi, and M. Scheffler, J. Chem. Phys. **129**, 194111 (2008).

[32]A. K. Soper, Chem. Phys. **258**, 121 (2000).

[33]See www.isis.stfc.ac.uk/groups/disordered-materials/database/ for ISIS disordered materials database.

[34]T. K. Hirsch and L. Ojamäe, J. Phys. Chem. B **108**, 15856 (2004).

[35]Z. Ma, Y. Zhang, and M. E. Tuckerman, J. Chem. Phys. **137**, 044506 (2012).

[36] F. Labat, C. Pouchan, C. Adamo, and G. E. Scuseria, J. Comput. Chem. **32**, 2177 (2011).

[37] Z. Raza, D. Alfè, C. G. Salzmann, J. Klimeš, A. Michaelides, and B. Slater, Phys. Chem. Chem. Phys. **13**, 19788 (2011).

[38] C. M. B. Line and R. W. Whitworth, J. Chem. Phys. **104**, 10008 (1996).

[39] P. J. Feibelman, Science **295**, 99 (2002).

[40] S. M. Jackson, V. M. Nield, R. W. Whitworth, M. Oguro, and C. C. Wilson, J. Phys. Chem. B **101**, 6142 (1997).

[41] A. K. Soper and C. J. Benmore, Phys. Rev. Lett. **101**, 065502 (2008).

[42] A. V. Gubskaya and P. G. Kusalik, J. Chem. Phys. **117**, 5290 (2002).

[43] D. P. Fernández, Y. Mulev, A. R. H. Goodwin, and J. M. H. Levelt Sengers, J. Phys. Chem. Ref. Data **24**, 33 (1995).

[44] K. T. Gillen, D. C. Douglass, and M. J. R. Hoch, J. Chem. Phys. **57**, 5117 (1972).

[45] R. Mills, J. Phys. Chem. **77**, 685 (1973).

[46] M. Sharma, R. Resta, and R. Car, Phys. Rev. Lett. **98**, 247401 (2007).

[47] H.-S. Lee and M. E. Tuckerman, J. Chem. Phys. **126**, 164501 (2007).

[48] M.-V. Fernández-Serra and E. Artacho, J. Chem. Phys. **121**, 11136 (2004).

[49] R. Jonchiere, A. P. Seitsonen, G. Ferlat, A. M. Saitta, and R. Vuilleumier, J. Chem. Phys. **135**, 154503 (2011).

[50] J. Wang, G. Román-Pérez, J. M. Soler, E. Artacho, and M.-V. Fernández-Serra, J. Chem. Phys. **134**, 024516 (2011).

[51] I.-C. Lin, A. P. Seitsonen, M. D. Coutinho-Neto, I. Tavernelli, and U. Rothlisberger, J. Phys. Chem. B **113**, 1127 (2009).

[52] M. Ji, K. Umemoto, C.-Z. Wang, K.-M. Ho, and R. M. Wentzcovitch, Phys. Rev. B **84**, 220105 (2011).

[53] A. Bernas, C. Ferradini, and J.-P. Jay-Gerin, Chem. Phys. **222**, 151 (1997).

[54] C. Pinilla, A. H. Irani, N. Seriani, and S. Scandolo, J. Chem. Phys. **136**, 114511 (2012).

[55] G. Lever, D. J. Cole, N. D. M. Hine, P. D. Haynes, and M. C. Payne, J. Phys.: Conden. Matter **25**, 152101 (2013).

[56] W. Wagner and A. Pruß, J. Phys. Chem. Ref. Data **31**, 387 (2002).

[57] C. M. Maupin, B. Aradi, and G. A. Voth, J. Phys. Chem. B **114**, 6922 (2010).

[58] G. Monard, M. I. Bernal-Uruchurtu, A. van der Vaart, K. M. Merz, Jr., and M. F. Ruiz-López, J. Phys. Chem. A **109**, 3425 (2005).

[59] Y. Wu, H. L. Tepper, and G. A. Voth, J. Chem. Phys. **124**, 024503 (2006).

[60] A. Glättli, X. Daura, and W. F. van Gunsteren, J. Chem. Phys. **116**, 9811 (2002).

[61] C. A. Coulson and D. Eisenberg, Proc. R. Soc. London, Ser. A **291**, 445 (1965).

[62] M. Uematsu and E. U. Frank, J. Phys. Chem. Ref. Data **9**, 1291 (1980).

[63] L. I. Schiff, *Quantum Mechanics*, 3rd ed. (McGraw-Hill, Inc., 1968).

[64] S. Gasiorowicz, *Quantum Physics*, 2nd ed. (John Wiley and Sons, Inc., New York, 1996).

[65] K. R. Harris and L. A. Woolf, J. Chem. Soc., Faraday Trans. 1 **76**, 377 (1980).

[66] M. Holz, S. R. Heil, and A. Sacco, Phys. Chem. Chem. Phys. **2**, 4740 (2000).

[67] M. E. Tuckerman, *Statistical Mechanics: Theory and Molecular Simulation* (Oxford University Press, New York, USA, 2010).

[68] R. J. Speedy and C. A. Angell, J. Chem. Phys. **65**, 851 (1976).

[69] S. Grimme, WIREs Comput. Mol. Sci. **1**, 211 (2011).