LAWRENCE
LIVERMORE
NATIONAL
LABORATORY

LLNL-JRNL-411122

# Isobaric-Isothermal Molecular Dynamics Utilizing Density Functional Theory: An Assessment of the Structure and Density of Water at Near-Ambient Conditions

J. Schmidt, J. VandeVondele, I. W. Kuo, D. Sebastiani, J. I. Siepmann, J. Hutter, C. J. Mundy

March 9, 2009

Journal of Physical Chemistry B

## Disclaimer

# Isobaric-Isothermal Molecular Dynamics Utilizing Density Functional Theory: An Assessment of the Structure and Density of Water at Near-Ambient Conditions

Jochen Schmidt,[1] Joost VandeVondele,[2] I.-F. William Kuo,[3] Daniel Sebastiani,[1] J. Ilja Siepmann,[4] Juerg Hutter,[2] and Christopher J. Mundy[5,*]

[1]*Max-Planck-Institute for Polymer Research, Ackermannweg 10, 55021 Mainz*

[2]*Physical Chemistry Institute, Zurich University, Winterthurerstrasse 190, CH-8057 Zurich, Switzerland*

[3]*Chemical Sciences Division, Lawrence Livermore National Laboratory, P.O. Box 808, Livermore, California 94551*

[4]*Departments of Chemistry and of Chemical Engineering and Materials Science, University of Minnesota, Minneapolis, MN 55455*

[5]*Fundamental and Computational Sciences Directorate, Pacific Northwest National Laboratory, P.O. Box 999, Richland, Washington 99352*

## Abstract

We present herein, a comprehensive density functional theory study towards assessing the accuracy of two popular gradient corrected exchange correlation functionals on the structure and density of liquid water at near ambient conditions in the isothermal-isobaric ensemble. Our results indicate that both PBE and BLYP functionals underpredict the density and overstructure the liquid. Adding the dispersion correction due to Grimme [1, 2] good agreement with the liquid density for both PBE and BLYP is obtaing. Moreover, the addition of the dispersion correction for BLYP yields an oxygen-oxygen radial distribtion function in excellent agreement with experiment. Thus, we conclude that one can obtain a very satisfactory model for water using BLYP and a correction for dispersion.

PACS numbers:

*Electronic address: chris.mundy@pnl.gov

## I. INTRODUCTION

Extended system molecular dynamics (MD) has its origins with the seminal work of Andersen in the derivation of the Hamiltonian equations of motion to sample the isoenthalpic-isobaric ensemble via molecular dynamics where the volume of the simulation cell is treated as a dynamical variable [3]. This pioneering idea of the "extended" system to treat the "bath" degrees of freedom as dynamical variables was quickly extended by Nosé to the canonical (NVT) ensemble [4], and by Parrinello and Rahman who derived a Lagrangian to allow for relaxation of the super cell parameters for simulations of solids under constant stress [5]. Subsequently, it was found that one could formulate the equations of motion for extended systems using a non-Hamiltonian approach leading to generalizations of the Nosé thermostat and provide the equations for the isothermal-isobaric ensemble (NpT) [6]. It should be pointed out, that one technical difference between MD sampling of the NpT ensemble and the equivalent Monte Carlo (MC) approach [7] is that the former requires direct calculation of the virial. For pair potentials, the evaluation of the virial is a straightforward calculation. Although one does not need a virial to sample the NpT ensemble with MC, special techniques are required to sample the intramolecular degrees of freedon for articulated molecules [8, 9]. On the other hand, the calculation of the viral for complicated interaction potentials can be cumbersome and the use of MC has distinct advantages.

Over the past fifteen years, there have been considerable efforts to refine extended system dynamics for both the canonical and NpT ensembles [10–15]. These refinements, in part, have come from a rigorous understanding of the statistical mechanics that arise from the non-Hamiltonian equations of motion[16, 17]. The benefits of being able to sample within the NpT ensemble has proved to be an invaluable tool for assessing the accuracy of interaction potentials. Having the ability to reproduce the experimental densities over a range of temperature is now an integral part of the parameterization of empirical interaction potentials [18, 19]. Despite the availabilty of fast electronic structure codes in conjunction with high performance computing, the application of the NpT dynamics utilizing Kohn-Sham density functional theory (DFT) to represent the interaction potentials is in its infancy. Techniques involving the optimization of lattice parameters for solids at both high and low temperatures are widely used [20–25], but to our knowledge, there have been no applications of NpT dynamics to molecular liquids using DFT interaction potentials. There are many reasons for this. Most early applications of first principles MD (FPMD) simulations in the NpT ensemble have employed plane-wave basis sets.

It is well known that the convergence of the pressure requries a significantly higher basis set cut-off than can be used for standard constant-volume FPMD simulations, thereby leading to a substantial increase in computational expense. Another problem, particularly for liquids, is that the equilibration time for NpT dynamics can be significantly larger as the volume fluctuations are typically low frequency.

We present the first FPMD simulations in the NpT ensemble of liquid water at near-ambient conditions. A previous notable example of overcoming the restriction of constrained volume in FPMD was a simulation of a slab of liquid water consisting of 216 molecules surrounded by a vapor region in NVT ensemble [26]. Recall, that the interfacial geometry is periodic, say in the $xy$ plane, leaving two free interfaces perpendicular to $z$ located at to located at $\pm z_0$. The fact that theses two free interfaces are not confined to be at $\pm z_0$, allows the liquid to expand/contract along the $z$ direction to find its stable density (in principle, in contact with its saturated vapor phase). The density profiles obtained for the DFT aqueous liquid-vapor interface using the BLYP exchange and correlation functionals [27, 28] converged to a bulk value of the specific density of $\approx 0.85$ g/cm$^3$. The aforementioned calculations, although performed in the standard canonical ensemble, provide a glimpse of the transferability to the DFT based interaction potentials to recover the proper liquid density.

The somewhat surprising result of a significant underprediction of the liquid density for water required further investigation of water's properties using a variety of sampling approaches to compute thermodynamic properties. To this end, simulations in the NpT ensemble were carried out utilizing the MC method [29]. In these calculations, a 64 molecule super-cell was utilized. This comprehensive study provided a sensitivity analysis of the calculated density as a function of simulation protocol. Although these calculations were short by the standards of MC sampling for pairwise additive potentials, the trend to lower density at ambient conditions for water utilizing DFT interaction potential is clearly apparent. Recently, McGrath et al. investigated the influence of the exchange-correlation functional, basis set size, density cutoff and even the method of generating the grids, used for numerical integrations, on the specific density of water at ambient condition and on its vapor–liquid coexistence curve [29–31]. In all cases, a liquid density well below 1 g/cm$^3$ was found for the BLYP functional, but the degree of underestimation was smaller for the PBE functional. While that work was performed using the MC approach, a self-consistent check with MD techniques, based on the analytical on-the-fly calculation of the internal pressure is still lacking.

As noted above, a MD simulation in the NpT ensemble relies on the fast computation of the internal pressure. Thus, in this paper we report on a efficient implementation of the virial into the QUICKSTEP module in the CP2K code [32]. Because QUICKSTEP uses standard basis sets for the orbitals, and an auxillary plane wave basis for the electronic density we will also investigate the sensitivity of calculated density to plane wave cut-offs.

## II. THEORETICAL BACKGROUND AND COMPUTATIONAL DETAILS

In the canonical ensemble, the pressure is given by the basic thermodynamic relation

$$P = -\left(\frac{\partial A(N,V,T)}{\partial V}\right)_{N,T} , \tag{1}$$

where $A$ is the Helmholtz free energy, $V$ is the volume of the system, $N$ the number of particles and $T$ the temperature. This can easily be transformed into an ensemble average over the pressure estimator $\Pi$

$$P = \frac{1}{3V}\left\langle \sum_{I=1}^{N}\left(\frac{\boldsymbol{P}_I^2}{m_I} + \boldsymbol{R}_I \cdot \boldsymbol{F}_I\right)\right\rangle = \langle \Pi \rangle \ . \tag{2}$$

Therefore, the evaluation of the internal pressure only requires the momenta $\boldsymbol{P}_I$ of the particles, their positions $\boldsymbol{R}_I$ and the forces $\boldsymbol{F}_I$. In an isobaric-isothermal simulation, the pressure estimator has to be calculated at each step and the result is considered the instantaneous pressure of the system. But in many cases the simple isotropic pressure is not sufficient and therefore a more elaborated, tensorial quantity has to be used. This quantity is the stress tensor which in periodic systems, as investigated in this study, is defined by

$$\Pi_{\alpha\beta} = -\frac{1}{3V}\sum_{\mu=1}^{3}\frac{\partial H}{\partial h_{\alpha\mu}} h^T_{\mu\beta} \ . \tag{3}$$

$H$ is the Hamiltonian of the interacting system and the matrix $\mathbf{h}$ consists of the Bravais lattice vectors $\boldsymbol{a}_1$, $\boldsymbol{a}_2$ and $\boldsymbol{a}_3$, i.e. $\mathbf{h} = [\boldsymbol{a}_1, \boldsymbol{a}_2, \boldsymbol{a}_3]$. The isotropic pressure is then given by the trace of $\mathbf{\Pi}$. In analogy to Eq. (2) the internal stress tensor can be calculated from the momenta and forces of the particles

$$\Pi_{\alpha\beta} = \frac{1}{3V}\sum_{I}\left[F_{I\alpha}\ R_{I\beta} + \frac{1}{m_I}\sum_{\mu i} h^{-1}_{i\alpha} P_{I\mu} P_{Ii} h^T_{\mu\beta}\right] \ . \tag{4}$$

We have implemented the above formula for the stress tensor into the CP2K program package [32], using the Gaussian and Plane Waves (GPW) [33, 34] formulation of DFT. In the GPW

method, the Kohn-Sham energy is given by

$$E[n] = E_T[n] + E_V[n] + E_H[n] + E_{xc}[n] + E_{II} \ , \tag{5}$$

where $E_T[n]$ is the electronic kinetic energy, $E_V[n]$ is the electronic interaction with the ionic cores, $E_H[n]$ is the electronic Hartree energy, $E_{xc}[n]$ is the exchange-correlation energy and $E_{II}$ is the interaction energy of the ionic cores. The electronic density $n$ can be represented by atom centered, contracted Gaussian functions or by plane waves. While the former is used for $E_T[n]$ and $E_V[n]$, the latter is employed for $E_{xc}[n]$ and $E_H[n]$. Therefore, the kinetic and potential energies are calculated purely analytically, while the Hartree and exchange-correlation energies have to be calculated on a numerical grid. This, in turn, results in a simple calculation of the stress tensor following Eq.(4) for $E_T[n]$ and $E_V[n]$, whereas additional terms have to be considered for $E_{xc}[n]$ and $E_H[n]$. These terms arise from the derivative with respect to $h_{\alpha\mu}$ in Eq.(3), because not only the positions of the particles, but also the grid points are affected by a change in $h_{\alpha\mu}$. The grid-dependent contributions have been evaluated following the work of Corso and Resta [35]. A second derivation has been given by Balbas et al. [36], where exclusively sums over grids are used instead of analytical integrals. We will not repeat the whole derivation here, but rather state the results, that we have implemented in our code. The contribution to the stress tensor due to the exchange-correlation energy is given by

$$\begin{aligned}\Pi^{xc}_{\alpha\beta} = &-\frac{1}{3V}\delta_{\alpha\beta}E_{xc}[n] \\ &- \frac{2}{3V}\sum_{\nu\gamma} P^{\nu\gamma} \int d\boldsymbol{r} v_{xc}(\boldsymbol{r})\phi_\nu(R_{I\beta} - r_\beta)\nabla_{I\alpha}\phi_\gamma \\ &- \frac{1}{3V}\int d\boldsymbol{r}\ \frac{\partial e_{xc}}{\partial(\partial_\beta n)}\frac{\partial n}{\partial r_\alpha} \ .\end{aligned} \tag{6}$$

$P^{\nu\gamma}$ is a density matrix element, $v_{xc}$ the exchange-correlation potential, $e_{xc}$ the exchange-correlation energy density, $\phi_\nu$ and $\phi_\gamma$ are basis functions. The first term arises from the scaling of the total volume of the system, while the other terms are linked to the derivative of the integrand. For the Hartree term, we find

$$\begin{aligned}\Pi^{H}_{\alpha\beta} = &-\frac{1}{3V}\delta_{\alpha\beta}E_{H}[n] \\ &- \frac{2}{3V}\sum_{\nu\gamma} P^{\nu\gamma} \int d\boldsymbol{r} v_{H}(\boldsymbol{r})\phi_\nu(R_{I\beta} - r_\beta)\nabla_{I\alpha}\phi_\gamma \\ &- \frac{1}{6V}\int d\boldsymbol{r}\int d\boldsymbol{r}'\ n(\boldsymbol{r})n(\boldsymbol{r}')\frac{(r_\alpha - r'_\alpha)(r'_\beta - r_\beta)}{|\boldsymbol{r} - \boldsymbol{r}'|} \ ,\end{aligned} \tag{7}$$

with the Hartree potential $v_H$. The last part results from the derivative of the energy density $1/|\boldsymbol{r} - \boldsymbol{r}'|$.

All simulations have been performed with the program CP2K [32], utilizing Born-Oppenheimer MD with periodic boundary conditions. Again, the DFT module, called QUICKSTEP, employs the Gaussian and Plane Waves (GPW) method [33, 34]. The equations of motion have been integrated using the algorithm of Martyna, Tobias and Klein [14, 15].

## III. RESULTS AND DISCUSSION

Ten different molecular dynamics simulations of 64 water molecules have been conducted, yielding trajectories between 24 and 48 ps. Four runs have used the BLYP functional [27, 28], the remaining ones the PBE functional [37]. In all cases, the norm-conserving pseudopotentials of Goedecker and co-workers [38] have been applied, the Gaussian basis set was a triple-zeta valence basis set augmented with two sets of d-type or p-type polarization functions (TZV2P). This basis set has been shown to give converged structural and dynamical properties for liquid water at constant volume [39]. The time step was 0.48 fs. Nose-Hoover-thermostats have been applied to all degrees of freedom (so-called "massive" thermostatting), the temperature has been set to 330 K and the time constant to 16.68 fs (corresponding to 2000 $\mathrm{cm}^{-1}$). The external pressure was 1 bar and the time constant of the barostat 300 fs. The first simulation has been started from a geometry with density 1 $\mathrm{g/cm}^3$, arbitrary points from this trajectory have been used as starting geometry for the other runs.

In the context of isobaric MD, the method how to generate the grids used for the numerical integration of the exchange-correlation and the Hartree potential is important, especially for the plane wave cutoff. Since the volume of the supercell is changed in every step of the simulation, the grids have to be regenerated in each step. In principle, two methods are possible. First, the number of grid points can be kept constant, leading to a different cutoff in each step (i.e. the density of the points is variable). Second, the density of the points can be kept constant, yielding a constant cutoff. Both approaches have been used here. If the former one is employed, the grid is generated base on a reference cell with constant size and then rescaled to the instantaneous cell size. In this case, a larger or smaller reference cell, compared to the expected simulation cell in equilibrium, can be used. This will be referred to as LARGEREF or SMALLREF, respectively. The direct NpT dynamics performed without a reference cell is called NOREF. According to the

work of McGrath et al. [29], a density of 0.8 g/cm$^3$ is expected near 300 K, corresponding to a cubic cell of length 13.4 Å. The LARGEREF and SMALLREF runs are using reference cells that correspond to a density of 0.76 and 0.93 g/cm$^3$, respectively. The cutoffs are 600 Ry for LARGEREF and 750 Ry for SMALLREF, which have been chosen to yield a cutoff of 700 Ry at 0.8 g/cm$^3$. The NOREF simulations have been carried out with a cutoff of 600 Ry. In order to assess the effect of the cutoff, a second series of LARGEREF simulations have been conducted with a cutoff of 1200 Ry. Normally, a cutoff of 280 Ry is considered sufficient to obtain a converged energy and forces, but McGrath et al. [29] have shown that for isobaric simulations a higher cutoff may be necessary to yield accurate energy differences as is needed with MC.

An important measure for the quality of an MD simulation is the conserved quantity. In ensembles other than NVE (microcanonical), where the total energy is conserved, this quantity is of a more complex nature and contains the thermostats and barostats. The conserved quantity for the integrator used here is introduced in ref. [15]. The relative change in this quantity for the LARGEREF runs with a cutoff of 600 Ry and 1200 Ry All our simulations yield a very good conservation with a relative change smaller than $1.5{\cdot}10^{-5}$ after 30 ps. Even a MD run in the much simpler NVE-ensemble is not expected to yield better results [34].

The main quantity of interest that can be deduced from an NpT simulation, is the fluctuating density of the system at the given external pressure. We are investigating water at near-ambient conditions, therefore one might expect a density of about 1 g/cm$^3$. However, previous NpT-MC simulations [29] have indicated that BLYP water yields a significantly smaller density, i.e about 0.8 g/cm$^3$. The densities found in our MD simulations are presented in FIG. 1. Although started from different initial geometries (BLYP NOREF was started from a 1g/cm$^3$ geometry, the others from arbitrary snapshots taken from the BLYP NOREF), the simulation time until equilibrium is reached is comparable for all runs. The barostats need at least 10 ps for relaxing the cell parameters. Therefore, such a simulation needs to be significantly longer than for standard FPMD simulations with constant volume. This is because the cell parameters, which are the corresponding dynamic variables in this ensemble, can be subject to low frequency oscillations, that need to be sampled before equilibrium is reached. PBE LARGEREF with 1200 Ry shows a clear negative trend even after almost 50 ps. Thus, the cell cannot be considered completely equilibrated. Still, the four different runs with each functional differ only by about 3%. Thus, no strong connection to the method to generate the grids or other simulation protocal can be found. Therefore, the differences have to be attributed to statistical fluctuations and insufficient

sampling, as alluded to before. McGrath et al. [29] have reported a 10% higher density for a simulation with a cutoff of 280 Ry versus those with 1200 Ry. In the MD runs, we do not find such an effect with cut-offs of 600 and 1200 Ry, presumably because the difference in energy/virial is much larger when increasing the cut-off from 280 to 600 Ry than from 600 to 1200 Ry (see energy–volume curves in Ref. [29]).

The average density from the FPMD simulations with the BLYP functional is 0.75 g/cm$^3$, whereas the PBE functional yields 0.88 g/cm$^3$. This is in line with previous findings in refs. [26, 29, 30]. While the aforementioned MC study with BLYP finds a density of about 0.8 g/cm$^3$ at 300 K (i.e., a 10% lower temperature than used here), a FPMD simulation of a slab of water surrounded by a vapor region, also at 300 K, led to a density of 0.86 g/cm$^3$. We find a density for BLYP at 330 K that is even below these values. PBE leads to a significantly higher density, which has also been reported by others, not only for ambient water, but also at elevated temperatures [31]. The data from our simulations are summarized in TABLE I. The average internal pressure and its RMSD are also reported there. The average pressure has not reached 1 bar in any case, and it shows very large fluctuations of about 3 kbar. This behavior of the pressure estimator is clearly a consequence of the limited system size and again emphasizes the need for very long simulation times in order to reach equilibrium. It should be noted that a consequence of Eq. (3) the internal stress is representing the system's tendency to expand or shrink the cell. Therefore, the pressure can be both positive and negative.

The observation that FPMD simulations with the BLYP and PBE representations of water result in a lower density than found experimentally is most likely correlated to discrepancies in structure and dynamics found in constant-volume FPMD simulations where the density is set to the experimental value [40]. Possible reasons for the deficiencies found are use of incomplete basis-sets [41], the lack of dispersion interactions in most common exchange-correlation functionals [42], or quantum effects [43]. Interestingly, PBE strenghtens the tendency toward overstructuring [44], while the density is closer to the experimental value.

The radial distribution functions (RDF) from our FPMD simulations are shown in Fig. 2. For reference, the experimental result of Hura *et al* is also shown in Fig. 3 [45]. Most commonly, the height and the position of the first peak of the oxygen-oxygen RDF are considered and used as measure for the structure of the liquid. A comprehensive collection of these parameters from varius simulations can be found in ref. [41]. We have summarized our results in TABLE I. The position of the peaks is shifted by about 0.1 Å towards larger values in the case of BLYP. This

could be directly connected to the lower density and reveals the tendency of the BLYP functional to produce weaker hydrogen bonds than that predicted by the PBE functional. In turn, the stronger binding found for the PBE functional leads to a more structured liquid, as expected. We find a height for the first oxygen-oxygen peak, $g_{\rm max}^{\rm OO}$, between 3.12 and 3.24 for the BLYP functional and 3.28 and 3.54 for the PBE functional. When our RDFs are compared to the values for constant-volume FPMD simulations given in ref. [41], our simulations appear to fall within the reported range of 2.9 to 3.6 (BLYP) and 3.2 to 3.8 (PBE). However, it should be noted that the MC simulations for BLYP in the NpT simulations resulted in a somewhat smaller $g_{\rm max}^{\rm OO}$ between 2.6 and 3.0 for the BLYP functional. The corresponding experimental values from neutron diffraction and x-ray scattering measurements are close to 2.8 [45–48]. In the FPMD simulations, the underestimation of the density is accompanied by a locally overstructured liquid.

It should be pointed out that although the present results indicate that there may be severe problems with popular DFT functionals, the successes of DFT cannot be overlooked. Under extreme conditions of pressure and temperature, DFT is able to reproduce the equation of state of water rather well [49]. Fortunately, this cannot be by numerical coincidence. DFT ability to get the correct chemistry and physics in the region of density overlap is likely responsible for the aforementioned successes. This is not entirely different than our understanding of liquid structure from the point of view of perturbation theory, such as the Barker-Henderson (BH) and Weeks-Chandler-Andersen (WCA) theories [50–52]. The salient result of these perturbation theories is that the structure of a simple liquid (at moderate densities) can be solely determined by the repulsive part of the potential. Given that DFT provides an accurate representation of the short-range part of the interaction potential, it is not a surprise that we can produce a reasonable liquid structure at reasonable densities. Moreover, the manifestation of the attractive part of the interaction potential may be important, if not dominant for thermodynamic properties such as liquid-vapor equilibria and pressure (as opposed to structural) for complex molecular liquids

To better test this assumption of the effects on the effects of structure and thermodynamics on the inclusion of dispersion, we have performed simulations using the so-called DFT + Dispersion (DFT-D) method [1, 2]. DFT-D methods have been used by a variety of researchers. Although details of implementation can vary, the salient features of the DFT-D approach are to augment existing DFT exchange and correlation functionals with $C_6 f(r) r^{-6}$. Here $r$ is the internuclear separation, $f(r)$ is a short-range damping function, and $C_6$ is the leading dispersion coefficient. The DFT-D method has been show to be sucessful for approximating the interaction energy of

the water-benzene system [2]. Starting from the final configurations of the simulations from the LARGEREF runs and utilizing a plane-wave cut-off of 600 Ry we continued the runs utilizing the proposed by Grimme [1] and as implemented in CP2K [32]. We choose a cut-off radius of 40.0 Åfor the dispersion interaction which ensured a converged virial and ensured that long-range corrections to the pressure can be neglected. The two runs, denoted BLYP-D and PBE-D were run for a total of at least 25 ps of production. RDFs and thermodynamic data are depicted in Figs. **??** and Table II. As one can see, there is a drammatic improvement in the density and RDF utilizing BLYP-D. Again, for reference, the experimental result of Hura *et al* is also shown in Fig. 3 [45]. The remarkable agreement with the Advanced Light Source experiment (ALS), especially in the second solvation shell, leads us to the conclusion that this simple correction of Grimme to BLYP yeilds a consistent picture of water from the point of view of thermodynamics and structure. It should be pointed out that other corrections to the long-range part of the interaction potential for DFT have also yielded better structure, but to our knowledge have not been benchmarked against thermodynamic properties such as pressure and density [53].

Although the PBE-D yeilds a much improved density, the RDF is still significantly over structured. Given that it is well known that water cluster are known to be underbind and overbind by BLYP and PBE respectively, these results are not a complete surprise. The additional attractive piece to the long-range portion of BLYP and PBE clearly pushes the density in the correct direction. The differences in the intermediate-range portion of the potential are clearly impacted differently by the different parameterizations of the Grimme formulation for BLYP and PBE. As a self-consistent check, simulation for the PBE-D were conducted using the last configuration of the BLYP-D simulations. Identical RDFs were obtained. Thus, from the results of our simulations, it is clear that this cross over region from long to short range plays a significant role in determining the structure for liquid water at near ambient conditions.

## IV. CONCLUSIONS

We present the first isobaric-isothermal MD simulations of near-ambient water from first principles. For this purpose, we implemented analytical expressions for the internal stress tensor into the QUICKSTEP module of the CP2K package [32]. We conducted a series of MD simulations with different exchange-correlation functionals, plane wave cutoffs and grid generation algorithms. While the effect of the plane wave cutoff beyond 600 Ry and the generation of the

| | $\bar{\rho}$ (g/cm$^3$) | $\bar{p}$ (bar) | $\Delta p$ (bar) | $g^{\mathrm{OO}}_{\mathrm{max}}$ | $r^{\mathrm{OO}}_{\mathrm{max}}$ (Å) |
|---|---|---|---|---|---|
| BLYP | | | | | |
| LARGEREF | 0.76 | -30 | 3,025 | 3.18 | 2.83 |
| LARGEREF-2 | 0.74 | 16 | 2,913 | 3.23 | 2.83 |
| SMALLREF | 0.73 | 14 | 2,752 | 3.12 | 2.84 |
| NOREF | 0.78 | 111 | 2,915 | 3.24 | 2.83 |
| PBE | | | | | |
| LARGEREF | 0.87 | 46 | 3,494 | 3.54 | 2.76 |
| LARGEREF-2 | 0.89 | 22 | 3,514 | 3.42 | 2.76 |
| SMALLREF | 0.87 | -57 | 3,563 | 3.28 | 2.76 |
| NOREF | 0.88 | -175 | 3,439 | 3.36 | 2.76 |

TABLE I: Thermodynamic and structural data from the NpT simulations. LARGEREF-2 uses a plane wave cutoff of 1200 Ry. $\rho$ is the density, $\bar{p}$ the average pressure, $\Delta p$ the RMSD of the pressure, $g^{OO}_{max}$ the height of the first peak of the $g^{OO}(r)$ which is located at $r^{OO}_{max}$.

| | $\bar{\rho}$ (g/cm$^3$) | $\bar{p}$ (bar) | $\Delta p$ (bar) | $g^{\mathrm{OO}}_{\mathrm{max}}$ | $r^{\mathrm{OO}}_{\mathrm{max}}$ (Å) |
|---|---|---|---|---|---|
| BLYP-D | | | | | |
| LARGEREF | 0.992 | -0.01 | 3,887 | 2.78 | 2.80 |
| PBE-D | | | | | |
| LARGEREF | 0.944 | -43 | 3,752 | 3.35 | 2.76 |

TABLE II: Thermodynamic and structural data from the NpT simulations utilizing DFT-D. $\rho$ is the density, $\bar{p}$ the average pressure, $\Delta p$ the RMSD of the pressure, $g^{OO}_{max}$ the height of the first peak of the $g^{OO}(r)$ which is located at $r^{OO}_{max}$.

grids do not show a significant influence on the resulting density and structure, the exchange-correlation functionals have a considerable impact. While BLYP drammatically underestimates the liquid density by about 25 %, PBE only yields an underestimation of 12 %. This independent study is in agreement with previous MC simulations of water at ambient conditions [29]. We have placed our findings in the context of other simulations of water based on DFT as well as perturbation theory which predicts that the liquid structure at moderate densities can be described accurately by considering only the repulsive part of the potential. Simulation times ranging from

22-45 ps were reached, but the density was never seen to be completely converged, highlighting the need for extremely long simulations in the NpT ensemble. Additionally, very large systems may be required in order to ameleorate the presence of large fluctuations of the instantaneous pressure, which can reach kbars.

Although the density was below the experimental value for the uncorrected exchange-correlation functionals, the RDFs remain overstructured in all cases. Expectedly, PBE produces a more structured liquid than BLYP. These effects have been found consistently in simulations using other DFT simulation protocol. Using the dispersion corrected exchange-correlation functionals produced a drammatic inprovement in the density for both PBE and BLYP. In addition, the first minima and second maxima of the oxygen-oxygen RDF using the BLYP-D functional leads to a almost perfect agreement with the experimental data from the ALS. As cautioned earlier, longer simulations will have to be conducted with different simulation protocol in order to get a denfinitive number for the liquid density. Despite the good agreement with the experimental density, the RDF for PBE-D still remains overstructured Results for the uncorrected and corrected exchange-correlation functionals point to the fact that the structure both functionals is likely determined by intermediate range interaction. Although other reasons discussed in the literature for the overstructuting of uncorrected exchange-correlation functionals (*e.g.* basis set size, and nuclear quantum effects), our study reveals that the more accurate treatment of the dispersion energy within DFT is required to be able to predict properties of liquids at ambient conditions. Moreover, from a practical point of view, DFT-D is a significant step forward for simulations at constant pressure or with open boundary conditions.

## V. ACKNOWLEDGMENTS

CJM acknowledges fruitful discussions with Phill Geissler regarding connections of this work to perturbation theory. This work was funded by the DFG grant SE 1008/2 (J.S.) and the National Science Foundation (CBET-0756641). CJM is supported by the US Department of Energy's (DOE) Office of Basic Energy Sciences Chemical, Geosciences, and Biosciences division. PNNL is operated by Battelle for the US DOE. A portion of the research was performed using the computing resource NWice located in the EMSL, a national scientific user facility sponsored by the Department of Energy's Office of Biological and Environmental Research and located at

Pacific Northwest National Laboratory. Prepared by LLNL under Contract DE-AC52-07NA27344.

---


[1] S. Grimme, J. Comput. Chem. **25**, 1463 (2004).

[2] U. Zimmerli, M. Parrinello, and P. Koumoutsakos, J. Chem. Phys. **120**, 2693 (2004).

[3] H. Andersen, J. Chem. Phys. **72**, 2384 (1980).

[4] S. Nosé, J. Chem. Phys. **81**, 511 (1984).

[5] M. Parrinello and A. Rahman, Phys. Rev. Lett. **45**, 1196 (1980).

[6] S. Nose and M. Klein, Mol. Phys. **50**, 1055 (1983).

[7] I. McDonald, Mol. Phys. **23**, 41 (1972).

[8] J. Siepmann and D. Frenkel, Mol. Phys. **75**, 59 (1992).

[9] M. Martin and J. Siepmann, J. Phys. Chem. B **103**, 4508 (1999).

[10] W. Hoover, Phys. Rev. A **34**, 2499 (1986).

[11] R. Wentzcovitch, Phys. Rev. B **44**, 2358 (1991).

[12] G. Kalibaeva, M. Ferrario, and G. Ciccotti, Mol. Phys. **101**, 765 (2003).

[13] G. Martyna, D. Tobias, and M. Klein, J. Chem. Phys. **101**, 4177 (1994).

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

[15] M. E. Tuckerman, J. Alejandre, R. López-Rendón, A. L. Jochim, and G. J. Martyna, Journal of Physics A **39**, 5629 (2006).

[16] M. Tuckerman, C. Mundy, and G. Martyna, Europhys. Lett. **45**, 149 (1999).

[17] C. Mundy et al., Nonequilibrium molecular dynamics, in *REVIEWS IN COMPUTATIONAL CHEMISTRY, VOL 14*, volume 14 of *REVIEWS IN COMPUTATIONAL CHEMISTRY*, pages 291–397, 2000.

[18] M. Mahoney and W. Jorgensen, J. Chem. Phys. **112**, 8910 (2000).

[19] B. Chen, J. Xing, and J. Siepmann, J. Phys. Chem. B **104**, 2391 (2000).

[20] K. Thomson, R. Wentzcovitch, and M. Bukowinski, Science **274**, 1880 (1996).

[21] B. Karki, L. Stixrude, and R. Wentzcovitch, Rev. Geophys. **39**, 507 (2001).

[22] M. Gillan, D. Alfe, J. Brodholdt, L. Volcaldo, and G. Price, Rep. Prog. Phys. **69**, 2365 (2006).

[23] R. Martonak, D. Donadio, A. R. Oganov, and M. Parrinello, Nature Materials **5**, 623 (2006).

[24] A. Oganov, R. Martonak, A. Laio, P. Raiteri, and M. Parrinello, Nature **438**, 1142 (2005).

[25] R. Martonak, A. Laio, and M. Parrinello, Phys. Rev. Lett. **90** (2003).

[26] I.-F. W. Kuo and C. J. Mundy, Science **303**, 658 (2004).

[27] A. D. Becke, Phys. Rev. A **38**, 3098 (1988).

[28] C. Lee, W. Yang, and R. G. Parr, Phys. Rev. B **37**, 785 (1988).

[29] M. J. McGrath et al., ChemPhysChem **6**, 1894 (2005).

[30] M. J. McGrath, J. I. Siepmann, I.-F. W. Kuo, and C. J. Mundy, Mol. Phys. **104**, 3619 (2006).

[31] M. J. McGrath et al., J. Phys. Chem. A **110**, 640 (2006).

[32] The CP2K developers group, `http://cp2k.berlios.de/`, 2009.

[33] G. Lippert, J. Hutter, and M. Parrinello, Mol. Phys. **92**, 477 (1997).

[34] J. VandeVondele et al., Computer Physics Communications **167**, 103 (2005).

[35] A. D. Corso and R. Resta, Phys. Rev. B **50**, 4327 (1994).

[36] L. C. Balbás, J. L. Martins, and J. M. Soler, Phys. Rev. B **64**, 165110 (2001).

[37] J. Perdew, K. Burke, and M.Ernzerhof, Phys. Rev. Lett. **77**, 3865 (1996).

[38] S. Goedecker, M. Teter, and J. Hutter, Phys. Rev. B **54**, 1703 (1996).

[39] J. VandeVondele et al., J. Chem. Phys. **122** (2005).

[40] J. C. Grossman, E. Schwegler, E. W. Draeger, F. Gygi, and G. Galli, J. Chem. Phys. **120**, 300 (2003).

[41] H.-S. Lee and M. E. Tuckerman, J. Chem. Phys. **125**, 154507 (2006).

[42] J. VandeVondele et al., J. Chem. Phys. **122**, 014515 (2005).

[43] H. A. Stern and B. J. Berne, J. Chem. Phys. **115**, 7622 (2001).

[44] I.-F. W. Kuo et al., J. Phys. Chem. B **108**, 12990 (2004).

[45] G. Hura, J. M. Sorenson, R. M. Glaeser, and T. Head-Gordon, J. Chem. Phys. **113**, 9140 (2000).

[46] A. K. Soper, F. Bruni, and M. A. Ricci, J. Chem. Phys. **106**, 247 (1997).

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

[48] J. M. Sorenson, G. Hura, R. M. Glaeser, and T. Head-Gordon, J. Chem. Phys. **113**, 9149 (2000).

[49] N. Goldman et al., J. Chem. Phys. (Accepted).

[50] J. Barker and D. Henderson, J. Chem. Phys. **47**, 2856 (1967).

[51] J. Weeks, D. Chandler, and H. Andersen, J. Chem. Phys. **54**, 5237 (1971).

[52] J. P. Hansen and I. R. McDonald, *Theory of Simple Liquids*, Academic Press, New York, 3rd edition, 2006.

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

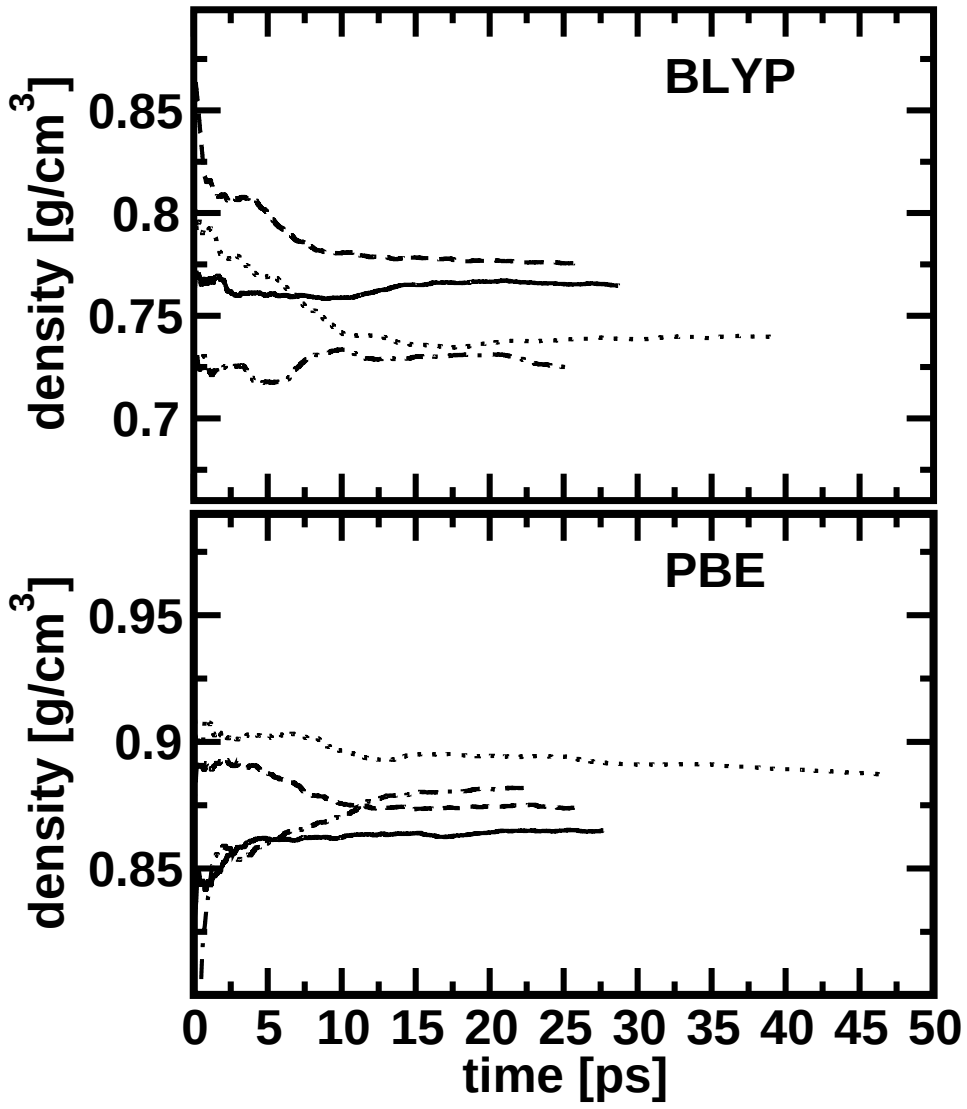


FIG. 1: The density computed from the NpT runs at 330 K and 1 bar. The simulations are: LARGEREF (solid), LARGEREF 1200 Ry (dotted), NOREF (dashed), SMALLREF (dash-dotted). We show a running average, *i.e.* each point represents the average over all previous steps.

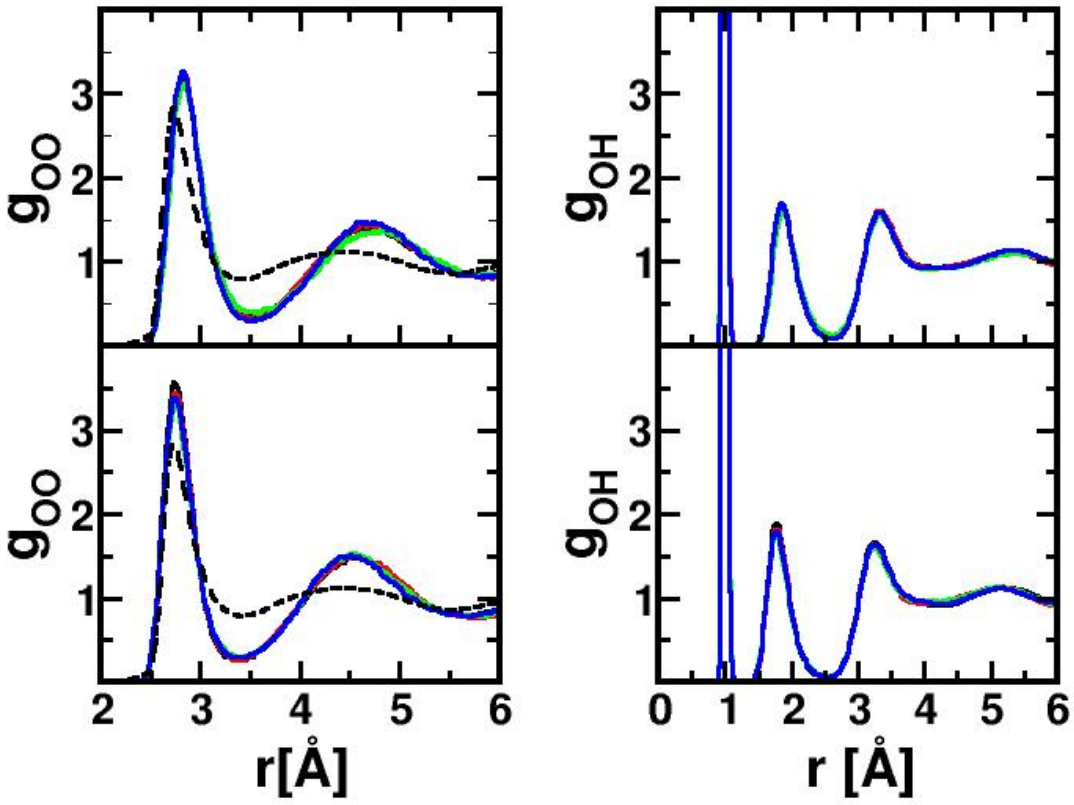


FIG. 2: Radial distribution functions for oxygen-oxygen with BLYP (top left) and PBE (bottom left), and oxygen hydrgen with BLYP (top right) and PBE (bottom right). The experimental oxygen-oxygen radial distribution function was obtained from the ALS and is dipicted with the dashed line [45]. The simulations were conducted at 330K are: LARGEREF (black), LARGEREF 1200 Ry (red), NOREF (green), SMALLREF (blue). A bin size of 0.02Å was used.

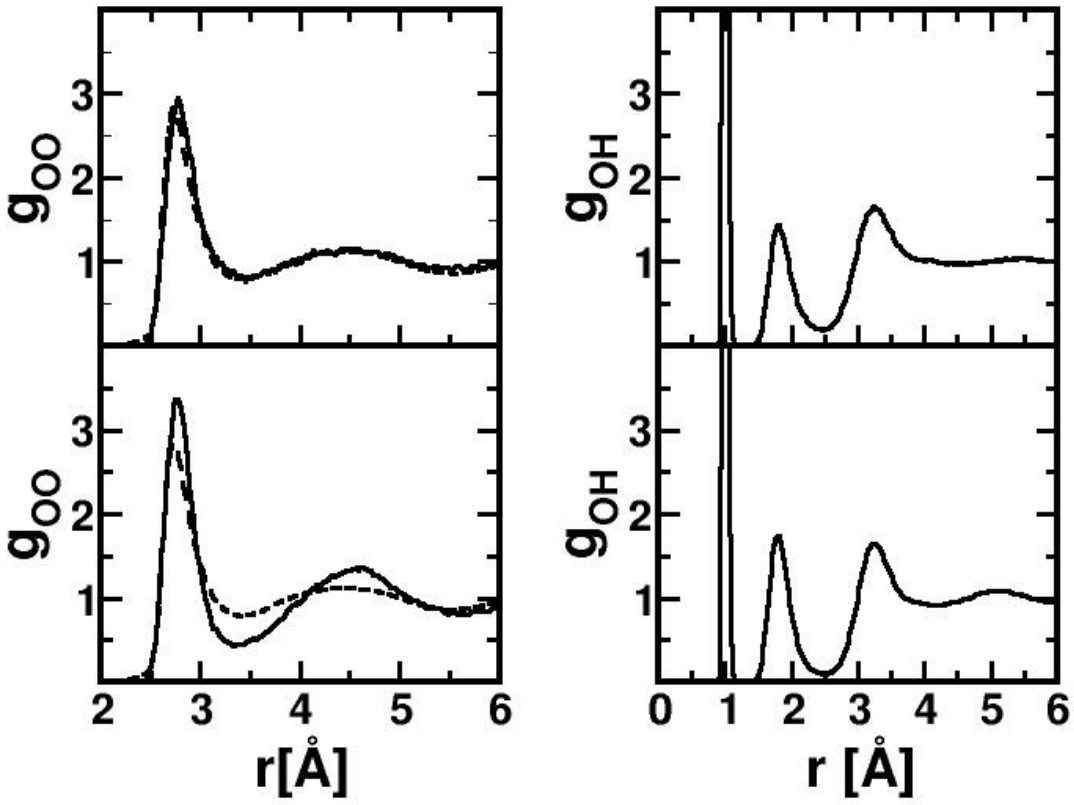


FIG. 3: Radial distribution functions for oxygen-oxygen with BLYP-D (top left) and PBE-D (bottom left) and oxygen-hydrogen with BLYP-D (top right) and PBE-D (bottom right). The experimental oxygen-oxygen radial distribution function was obtained from the ALS and is dipicted with the dashed line [45]. All simulations use the LARGEREF simulation protocol at 330K. A bin size of 0.02Å was used.

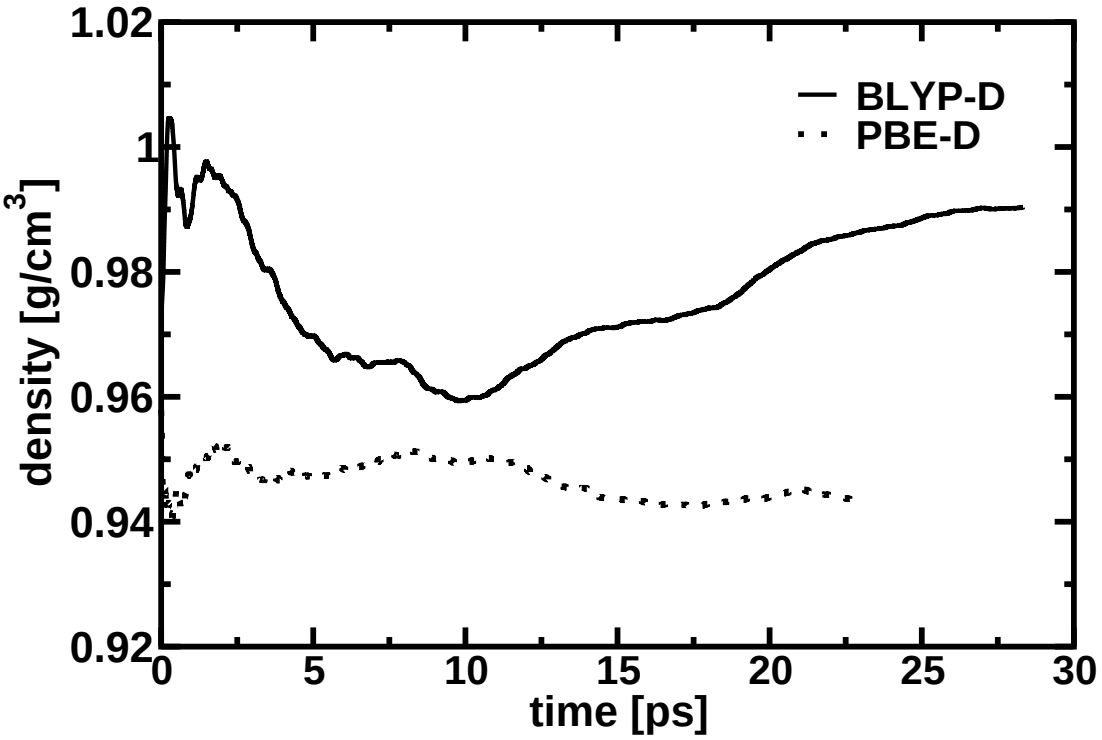


FIG. 4: The density computed from the NpT runs utilizing the dispersion corrected exchange-correlation functionals at 330 K and 1 bar. All simulations use the LARGEREF simulation protocol. We show a running average, *i.e.* each point represents the average over all previous steps.