

Effects of H-bond asymmetry on the electronic properties of liquid water – An AIMD analysis

Imre Bakó^a, János Daru^a, Szilvia Pothoczki^b, László Pusztai^{b,c}, Kersti Hermansson^d

^a*Institute of Organic Chemistry, Research Centre for Natural Sciences, Hungarian Academy of Sciences, Magyar tudósok körútja 2, H-1117 Budapest, Hungary*

^b*Wigner Research Centre for Physics, Hungarian Academy of Sciences, H-1121 Budapest, Konkoly-Thege M. út 29-33., Hungary*

^c*International Research Organization for Advanced Science and Technology (IROAST), Kumamoto University, 2-39-1 Kurokami, Chuo-ku, Kumamoto 860-8555, Japan*

^d*Department of Chemistry-Ångström Laboratory, Uppsala University, Box 538, SE-751 21 Uppsala, Sweden*

Abstract

Three molecular electronic properties (local electronic density of states, molecular dipole moment distribution of the Bader and Wannier types, and net molecular charges) were examined from the collected water trajectories from ab initio molecular dynamics simulations of liquid water at 300 and 350 K using the BLYP-D3 functional. All three molecular electronic properties are found to be sensitive not only to the number of H-bonded neighbours surrounding the molecule but also the nature of the asymmetry of the coordination. That this is the case for the electronic DOS is known from the literature but here we compare the three properties and find that the electronic DOS and the net molecular charge are more affected by the asymmetry than by the number of H-bonded neighbours. The reverse is true for the dipole moment. Moreover, we consistently find that for all three properties, a 3-coordinated water molecule is more perturbed by accepting two H-bonds and donating 1.

Keywords:

Liquid water, ab initio-MD simulations, electronic properties, density of states, molecular dipole moments

Corresponding author:, Email address: bako.imre@ttk.mta.hu (Imre Bakó)

1. Introduction

The electronic energy levels in liquid water, manifested either as the pure electronic DOS spectra from quantum-mechanical calculations or as experimental X-ray photoelectron (XPS), absorption (XAS) and emission (XES) spectra, have been addressed in numerous studies [1-11], some with a focus on the very topic of symmetric versus asymmetric environments. The effect of hydrogen bonding on the lowest excited state associated with the central molecule has been studied on cluster models [10,11]. These calculations provide evidence that the blue shift of excitation energies is mainly determined by acceptor hydrogen bonds, whereas donor hydrogen bonds do not induce significant shifts. This kind of asymmetric nature of the sensitivity to different parts of the hydrogen bond network for the XES and XAS spectra has been addressed by many authors and the resulting implications concerning the local (and not so local) structure of liquid water is still under debate. For example, the group of Pettersson and Nilsson [3,9], based on both elaborate experiments and simulations, have advocated the dominance of the presence of very asymmetrical water molecules in liquid water, while others, such as Kühne and co-workers [12], ascribe the key features in the electronic spectra to normal consequences of the traditional liquid water structure with a fluctuating 4-coordinated H-bond network. In the present paper we will discuss how the basic electronic properties of the water molecules themselves are affected by the H-bond surroundings, without reference to any particular probing technique. We will address whether the perturbation from the H-bond environment and its asymmetry are captured in a consistent way by the molecular dipole moment, the occupied electronic energy levels, and the net molecular charge.

The enhancement of the water molecular dipole moment in liquid water compared to the gas-phase value has also received much attention, as the static and induced dipole moment are two of the most important properties of water molecules in practical applications. Many approaches to estimate the water dipole moment in liquid water and other condensed aqueous media have been attempted, both from experiment and calculation. Computer simulations have shed much light on this problem, both by way of mechanistic non-polarisable or polarisable force fields simulations and by way of electronic structure calculations. The latter, i.e. *ab initio* MD simulations, where the forces are generated from electronic structure calculations, offer a more straightforward possibility to estimate the in-liquid molecular polarisation. Even so, it is *a priori* not clear how a measured macroscopic dipole moment can be subdivided into 'molecular' contributions, in particular in a disordered system like liquid water. Various types of analyses of the electronic wave function to obtain the dipole moment are being practised, yielding e.g. "Wannier-type" dipole moments which involves the analysis of the positions of the nuclei and the calculated Wannier-function centres (WFCs) (which is based on a molecular orbital localization technique) [13,15] or Bader-type dipole moments following a topological analysis of the electron density within the framework of Bader's Quantum Theory of Atoms In Molecules (QTAIM) [14]. The first time WFCs were used for the calculation of the molecular dipole moment in liquid water was the study by Silvestrelli and Parrinello [15] which also gives a comprehensive overview of force-field based approaches, such as those mentioned briefly above. The Bader type analysis of liquid water was first used by Delle Site et al. [16] and Dyer et al. [17]. We make use of both the Wannier and Bader approach in this paper, and discuss both the polarisation and the charge transfer (net dipole moment if the water molecule) for different classes of H-bond surroundings and symmetric and asymmetric environments.

Thus here we have performed *ab initio* molecular dynamics (AIMD) simulations of liquid water, whose results are subjected to detailed analyses in terms of both the atomic and electronic structure. We will first validate the BLYP-D3 MD structures at 300 and 350K by comparison with both X-ray and neutron scattering curves (see references later in the paper). The comparison of the BLYP-D3 liquid water structures with experimental X-ray scattering curves has not been presented before in the literature as far as we are aware. As the comparison turns out favourably it helps to validate that the structures we compute are feasible and worthy of the type of detailed analysis that they will be exposed to.

2. Methods and computational details

DFT-based ab initio molecular dynamics (AIMD) simulations at two temperatures were performed in a periodic setup using the CP2K [18-20] ab initio molecular dynamics code. Details are given below.

2.1 Electronic structure calculations

The Becke-Lee-Young-Parr (BLYP) [21,22] functional was used together with the ‘D3’ dispersion correction, as designed by Grimme et al. [23]. The norm-conserving Goedecker-Teter-Hutter pseudopotentials [24] was used. We employed a triple zeta basis set with double polarization (TZV2P). The plane wave basis sets is controlled by a charge density cut-off in the CP2K program. There has been a long debate in the literature whether the relatively small (approx. 300 Ry) charge density cutoff can provide converged results, especially when applied in the NPT ensemble [25]. Jonchiere and co-workers [26] showed that there is a more important effect on the convergence and numerical problem, namely the smoothing method: smoothing was found to be more important than using a large cutoff. In our simulations at both temperatures, we have chosen to use 300 Ry and the NN50 smoothing method implemented in the CP2K code. During each self-consistent field (SCF) cycle, the wavefunction coefficients were minimized to a threshold of 10^{-7} . In all simulations, the standard deviation of the total energy was about 10^{-4} Hartree.

2.2 Molecular dynamics simulation

The AIMD simulations were performed in the NVT ensemble, at 300 K and 350 K. Periodic boundary conditions have been employed, and in each case, the simulation box was cubic, with edge lengths of 12.43 Å, and contained 64 water molecules; this corresponds to a number density of 0.0333 molecules/Å³. The experimental number densities for liquid D₂O at 300 and 350 K are 0.0332 and 0.0324 molecules/Å³, respectively [27]. For both simulations, the experimental density at 300 K was assumed (for a more detailed reasoning, see below). The time-step was 0.5 fs.

The effect of taking dispersion effects into account is important for liquid water, and applying the Grimme D3 correction leads to a significant improvement of the water structure as has been shown in several publications; see for example Ref. [28]. Moreover, it has been shown that, for water systems of at least 64 molecules or more, the system size in AIMD simulations does not have a large effect on the water structure [28-30]. Some comparisons of the RDFs obtained from simulations using the same functional form (BLYP-D3), but with different system sizes, are presented in the **Supp. Material. Fig. S1**.

As the experimental number density of liquid D₂O water at 350 K is (only little) lower than the value used in our simulation we wanted to estimate the effect decreasing the density slightly, and performed additional (short) simulations at 350 K with a box of 12.55 Å in length, from which we estimated that the error associated with this choice should be less than 1 % in terms of the peak height and the peak is appeared at the same distance in the radial distribution functions (RDFs). Also the dipole moment distributions did not change. These results are presented in **Supp. Mat. Table S1**. In the light of this, and observing that the phase diagram of the BLYP-D3 functional is not known, we decided to assume identical number densities for the two temperatures considered here.

All hydrogen atoms were assigned the mass of deuterium atoms to reduce the influence of quantum effects. More details concerning the technical details of our simulation can be read from our input file to CP2K, listed in **Supp. Mat. Section 2**.

The initial particle configuration was generated from classical MD simulations using the SPC/E [31] interaction potential. Particle configurations were collected for subsequent analyses from 120 ps runs after 40 ps of equilibration (cf. **Fig. S2**).

2.3 Property calculations

The MD snapshots were analysed for a number of properties. The electronic density-of-states, DOS, results were derived by projecting the wave-function (Kohn-Sham orbitals) onto the individual local basis functions of all, or of selected, molecules from the MD snapshots; such projected PDOS graphs are presented. Water molecules with different hydrogen-bond surroundings were analysed in this way. It is worth pointing out that in this paper we will use the term PDOS also for density of states results that refer to a subset of the molecules; others might prefer to use the name localised DOS (LDOS) for these. The Kohn-Sham DOS was calculated in every 500th time-step (250 fs).

The dipole moment of each analysed water molecule was calculated from these data in two ways. One approach was that of Bader et al. (see references in the Introduction), based on the electron densities, using the code of Henkelman et al. [32,33]. Such calculations were performed every 50 fs (100 time steps). The second approach to calculate molecular dipole moments was by means of the centres of the maximally localized Wannier functions (see references in the Introduction using the tools provided within the CP2K code; in both cases some small in-house computer codes streamlined the procedure. This analysis was done every 25 fs (50 time steps) in the trajectories from the production runs.

Molecular net charges were calculated by summing over the atomic charges, derived either from the Bader-type analysis or from Mulliken charges.

Statistical analyses of these resulting molecular dipole moments, electronic PDOS curves and molecular net charges were performed by so-called bootstrapping; this method was first introduced by Efron et al. in 1979 [34]; there are numerous books on the bootstrap theory and its methodology, together with applications [35]. The bootstrap method can be employed to understand the structure and uncertainty of a random variable and also, in other applications, to provide error estimations for existing models. Bootstrap resampling constructs datasets with n points where each point is selected from the full dataset (N ; $n \ll N$), and the averaged value and confidence interval for parameters can be calculated.

In addition, and for comparison, we made some calculations for two optimized water clusters with 20 (dodecahedron) and 30 molecules each using the same electronic structure method. Dipole moments and orbital energy distribution (projected density of states) of the individual water molecules in these clusters were obtained with the same analysis methods as for the liquid water snapshots. The initial cluster geometries were taken from Qian et al. [36].

3. Results

3.1. Validation of the BLYP-D3 simulations: structure, diffusion, and electronic DOS

No direct assessment of the ability of the BLYP-D3 method to reproduce the experimental X-ray-derived total RDF function has been presented in the literature, as far as we are aware. **Figure 1a-b** shows such a comparison between our calculated total X-ray weighted RDF curves (i.e. mostly the O-O partials as they have the largest X-ray diffraction weight) from the 300 K and 350 K simulations, with the total RDF curves from X-ray scattering experiments in the range of 277–307 K and 334–365 K, respectively [37-40]. The figure displays several experimental curves, taken from the literature, for a temperature range rather than for one particular T value, as we do not know the exact melting point of BLYP-D3 water, although recently it has been shown that it lies in the range of 283-325 K [30,41]. In the case of our 300 K simulation, the agreement is close to excellent when compared to all the experimental curves shown in **Fig. 1a**, and the very best agreement is with the 277 K experiment. For the 350 K simulation, the agreement is also excellent in the experimental temperature range shown in **Fig. 1b**. We note that Schmidt et al. [25], using the BLYP-D1 functional in a water simulation at 330 K, found a similar good agreement with X-ray scattering RDF curves from Hura et al. [42].

Next we turn to comparisons with available neutron diffraction results. The O-O, O-H and H-H partial radial distribution functions (RDFs) have been computed from MD simulations near room temperature with the present density functional (BLYP-D3) by several authors in the past decade (see for example Refs [26,30,43,44]). Since software, pseudopotentials, temperature, and systems setups vary between different studies it is not always easy to pinpoint the exact influence originating from the dispersion correction. However, it is fair to say that RDF curves from BLYP-D3 simulations have generally given good agreement with the corresponding curves derived from the neutron diffraction measurements for liquid D₂O at 298 K [45]. Our calculated partial RDFs corresponding to the neutron diffraction results are presented in **Fig. S3** in the **Supp. Mat.** along with a brief discussion of the quality of data.

We also computed total (detailed description on their calculation are given in **Supp. Mat. section 3**) derived from weighting the RDF according to the neutron scattering lengths of H (-0.374 fm), D (0.664 fm) and O (0.583 fm), for D₂O and H₂O, and compared them with the experimentally measured RDFs from neutron scattering measurements at 300 K (**Fig. 2a-b**). The agreement outside 2 Å is good both for the protonated and deuterated water. There is a significant difference at low r values, which mainly arises from the damping factor applied during the analyses of experimental data, and from the fact that the “real” temperature is somewhat lower in our simulation box, as shown above. As far as we are aware such comparisons have not been shown before.

It is here relevant to comment on the consequences of nuclear quantum effects (NQE) for, for example, the radial distribution functions, especially since such effects were not included in our study. The origin and size of the NQE effect on structural and dynamical properties of liquid water have been investigated by several authors; see for example Refs. 46 and 47. Those studies suggest that NQEs on the O—O, O—H, and H—H rdfs are not insignificant but modest. Judging from these studies we conjecture that the NQEs would be small to modest both regarding the total X-ray weighted rdf as displayed in Fig. 1 and the total neutron scattering weighted rdf in **Fig.2**.

The self-diffusion coefficients obtained from our calculations, using the Einstein relation, are 0.16 ± 0.03 at 300 K and 0.31 ± 0.03 Å²/s at 350 K. The error bars were estimated using the block average method. The mean square deviations (MSD) in the Einstein relationship were computed from the relative displacements of the oxygen atoms (which are practically the centres of mass) and averaged over all water molecules, and every 50th configuration of all trajectories.

For comparison, the experimental values have been back-corrected "towards" the actual simulation system size using the procedure from Ref. [26] These modified size-back-corrected self-diffusion coefficients are inversely proportional to the length of the cell and to the viscosity. Since the viscosity is unknown in our calculations, we instead followed the suggestion of [26], where the experimental result (heavy water) is adjusted so that it represents a hypothetical periodic system of 64 water molecules. These values are about 0.18 Å²/s and 0.30 Å²/s at 298 and 330 K (for D₂O), respectively [48], so the agreement between experiment and calculation is satisfactory. The calculated MSDs at 300 K are presented in the **Supp. Mat. Fig. S4**.

Furthermore, we compared experimental and calculated cohesive energies. The cohesive energy of simulated liquid water – calculated as described by McGrath et al. [49] – was found to be -43.4±0.5 and -41.5±0.4 kJ/mol at 300 and 350 K, respectively, which is in a reasonable agreement with the experimentally determined values [50] (-46.34 and -44.2 kJ/mol; in the calculation of experimental cohesive energy an ideal vapor phase is assumed). This finding lends further support the present computational scheme.

Finally, turning to the electronic properties, **Fig. 3a** shows the distribution of the Kohn-Sham eigenvalues at the Γ -point, calculated as an average over 480 snapshots, or 30720 water molecules in total, for each of the two simulations. The difference between the two temperatures is seen to be insignificant. It has been shown previously, the relative positions and features of the occupied states (1b₁, 3a₁, 2b₂, 2a₁) are reasonably well described with only modestly large systems, such as in the

work of Hunt et al [2] and of Prendergast et al. [4] who used 64 molecules and only a single k-point. However, the description of the unoccupied DOS is very sensitive to the size of the simulated system [14].

Indeed, the HOMO-LUMO gap poses some problems in terms of qualitative agreement with experiment. Liquid water can be considered a disordered wide gap insulator; its HOMO-LUMO band gap has been experimentally determined to lie in the range from about 8.7 to 10.5 eV [5,6]. The corresponding band gap from our BLYP-D3 simulation is as low as about 4.5 eV, consistent with a large number of other dispersion-corrected GGA-based AIMD simulations in the literature; see e.g. Refs. [2, 4], who report values in the range 3.6-4.6 eV. It is well known that GGA tends to underestimate the electronic band gap of water, e.g. due to finite size effect, incorrect description of the valence band edge, the so-called derivative discontinuity problem relating to an incomplete treatment of electron transfer and the self-interaction error. A more detailed discussion about the band-gap problem using GGA can be found for example in Refs. 4 and 51-53.

In summary, the pattern of the *peak maxima* in the BLYP-D3 simulated DOS curve (**Fig. 3a**) is in quite good agreement with that from photoemission experiments in, e.g., Refs. [6,54]. The comparison is presented in **Table 1**.

3.2 Effects of H-bond coordination and structural asymmetry

Having confirmed that the experimental liquid structure, at least in terms of the experimentally measurable radial distribution functions (X-ray, neutron), and the overall form of the raw electronic DOS profile (i.e. not adapted to mimic any photoelectronic signal such as that of XES spectroscopy) is quite well described by the BLYP-D3 simulations we will go on to analyse some electronic properties in more detail and whether they correlate with the H-bond pattern and coordination. Such investigations of structure-property relations make good use of computer simulations as they address minute details that are not readily accessible from experiment. Some special focus will be placed on the contributions from the asymmetrically coordinated water molecules components with different number of H-bonded neighbours on the acceptor and donor side, a topic of intense interest in the liquid water literature for some considerable time, as commented on and referenced in the Introduction. First we need to define an H-bond and the coordination classification scheme that we will use.

H-bond coordination classes and networking. In the literature, the existence of a hydrogen bond is usually defined either from energetic or geometric criteria, in the former case by some pair interaction energy cutoff value, in the latter by distance and angle ranges. However, it is worth noting that in the analysis of snapshot structures from MD simulations the AIMD method does not readily provide molecular pair interaction energies. A geometric definition is therefore more straightforward, and is anyway the most common approach in the scientific literature. Different variants of the geometric H-bond definition exist; they all involve the intermolecular $R(O \cdots O)$ or $R(H \cdots O)$ distances between the H-bond donor and acceptor, and usually some H-bond angle; see, e.g., the work of Kumar et al. [55] and references therein.

Here, we define two molecules to be H-bonded if the following two criteria are fulfilled: (i) the intermolecular $R(H \cdots O)$ distance is less than 2.5 Å (first "intermolecular" minimum of the O-H RDF), and (ii) the corresponding $O \cdots O - H$ angle is less than 30°, i.e. if the angle between the intramolecular OH bond and the intermolecular $O \cdots O$ vector is less than 30°.

With these criteria we find an average of 3.6 H-bonds surrounding a water molecule in the 300 K simulations and 3.2 H-bonds at 350 K; see the rightmost column in **Table 2**. As expected, the average number of H-bonds decreases when the temperature is raised. For a given functional and simulation setup, the exact numbers will of course depend on the H-bond criteria, which in our case are rather liberal with respect to the H-bond distance but less so for the H-bond angle. In the

following we will discuss how the surrounding H-bond coordination affects the electronic properties of an in-liquid water molecule.

The snapshots from the simulations were analysed by categorizing the water molecules into different *coordination classes*. In our scheme, which has been used by many others in the literature, we start from a selected "central" molecule and examine how many H-bonds, i.e. H-atoms, it *accepts* from its neighbours, and how many H-bonds it *donates* to its neighbours. These numbers are labelled n_A and n_D , respectively, and the general label for the H-bond coordination around a molecule is thus ' $n_A n_D$ '. For example, 2A:1D denotes a molecule which accepts 2 H-bonds and acts as a donor of 1 H-bond. On the acceptor side, even 3 water molecules may approach, giving 3A:2D. Similar kinds of topological analyses of the hydrogen-bonded network structure in liquid water, as well as the probabilities to find a given number of H-bonded neighbors are available in the literature; a few of all the relevant references are Refs. [56-60].

Our results are shown in **Table 2** and overall agree well with results from the literature. We observe a preference for 4-fold H-bond coordination of the 2D:2A type at both temperatures, with a decrease of the fraction of 4-coordination at 350 K (**Table 2**). Moreover, we find that the two populations with three H-bonds in asymmetric conformations (1A:2D and 2A:1D) are quite abundant and not equal in our simulations. This asymmetry (difference between the two populations) is more pronounced at 300 K. This agrees well with results of DiStasio et al. [61] who found that these two types of coordinations are differently populated. In the following, we present whether and how these different environments entail different electronic properties, namely specifically electronic energy levels, molecular dipole moment and net molecular charge.

Electronic PDOS curves. We have generated the electronic DOS projected onto the basis functions of water molecules belonging to different $n_A n_D$ classes, taken from the MD snapshots. These LDOS curves are shown in **Fig. 3b** and **c** for the 300K and 350 K simulation data, respectively, and the calculated band centers are listed in **Table S2**. Inspection reveals that the order of the "bands" for the different coordination classes is the same for all the valence bands, and for the two independent simulations, supporting the robustness of these results. We actually also find the same order for an optimized water cluster with 30 molecules; see **Figure 3d**.

In summary, the four band centers downshift by 0.3-0.4 eV in going from the least "powerful" H-bond coordination (1A:2D) to the most powerful (3A:2D). The peak positions are seen to be shifted to more negative energy values (i.e. to the left in the figure) when the number of hydrogen bonded neighbors increases. The order in which the "class bands" appears is (least shifted) 1A:2D: \Rightarrow 1A:1D \Rightarrow 2A:1D \Rightarrow 2A:2D \Rightarrow 3A:2D (most shifted). The order follows the number of coordinating H-bonds only if the average of the two 3-coordinated peaks is used. This is shown for the "1b₁" band in **Figs. 4a** and **d**, and looks almost identical for the other bands. The 3-coordinated configurations thus destroy the peak-position-*versus*- n_{HB} correlation, and the 2A:1D case is more affected by the environment than the 1A:2D case.

Molecular dipole moment. Here we have calculated μ_{molecule} in two different ways: from the positions of the nuclei and the Wannier-function centres (WFCs) and from the Bader charges. We – like others – find that the total dipole moment distribution is broad (**Fig. S5**). Our FWHH (full width at half maximum) values are about 0.60-0.65 D for the Wannier and Bader methods alike. The average values of the distributions are rather different, namely about 3.00 D from the Wannier scheme and 2.40 D from the Bader scheme. The published values for the average Wannier-type dipole moment in liquid water are generally reported to lie close to 3.0 D for various AIMD simulations using PBE or BLYP-based functionals in the literature (a selection of references is found in [61-63]). On average the water molecule is thus enhanced by a factor of more than 60% compared to the gas-phase value. This is a large increase, but much smaller than what has been shown to occur in simple ionic crystalline hydrates which can display Wannier-type water dipole moments well over 4 Debye [64-65].

The average molecular dipole moment calculated for MD snapshot populations of water molecules surrounded by 0, 1, 2, 3, and 4 hydrogen bonds is displayed in **Fig. 4b**. Also for this property (as for the PDOS) there is an almost linear dependence on the number of H-bonds. Our Wannier-type dipole moment variation is in excellent agreement with those of DiStasio et al. [61] for the van der Waals-corrected functionals that they tested. Here, we have furthermore made a division of the 3-coordinated class into its main two asymmetric structural motifs, 2A:1D and 1A:2D. The structural asymmetry is indeed translated into a dipole moment asymmetry (**Figs. 4b** and **e**), seen both for the Bader and Wannier-type dipole moments, but the effect is tiny, and much smaller than the effect of changing the number of H-bonded neighbours. The central water molecules in the 2A:1D motifs are on average (only) slightly more polarised than the central water molecules in the 1A:2D motifs. To the best of our knowledge, this has not been presented for the water dipole moment before.

We add a few words about the large difference between the Bader-type and the Wannier-type dipole moments. The AIMD-Bader values are consistently smaller than the AIMD-WFC values, by 0.5-0.6 D. It is clear that the dipole moment of a water molecule cannot be defined unambiguously in condensed matter due to electron delocalization (electron density overlap). This is not least true in H-bonded systems. Actually, the Wannier type dipole moments suffer from the fact that they implicitly include effects of charge transfer *between* molecules, in addition to the polarisation *within* the molecules. Only the latter is what we normally refer to as molecular polarisation. This charge transfer is visible in the Wannier orbitals. It was demonstrated by one of us [66] that the localized Wannier orbitals calculated for a water cluster have small delocalization tails along the hydrogen bonds that are important in determining the resulting dipole moment in the system. This may be part of the explanation for the differences observed between the two types of dipole moment. The physical interpretation of Bader charges was well described in the article by Bader et al. in Ref 14.

Net molecular charge. The distribution of the net charge of a water molecules in liquid water is narrow and close to zero, when calculated either with the Bader or Mulliken schemes (**Fig. S6**). We examined the different H-bond classes and coordination classes, and here, again, we observe considerable asymmetry for the 2A:1D and 1A:2D groups (**Fig. 4c** and **f**), much larger than the difference between the H-bond groups. The asymmetry thus overpowers the effect of changing the H-bond coordination number.

H-bonding is accompanied by electronic charge transfer from the acceptor to the donor molecule, as shown by many studies in the literature (see for example the work of Alfredsson et al. [67]). One would therefore expect 1A:2D to acquire a net negative charge and 2A:1D to become positive. This is, indeed, the case, while a water molecule residing in a 1A:1D or a 2A:2D motif generally remains more neutral, just letting the electronic charge "pass through".

4. Summary and Concluding remarks

For all electronic properties examined here – electronic DOS, molecular dipole moment and molecular net charge - we find a systematic shift as a function of the number of H-bond neighbours. The 3-coordinated systems break the trend for the electronic DOS and the net molecular charge; here the structural asymmetry translates into an electronic asymmetry which is significantly larger than the effect of changing the number of H-bonded neighbours. For the dipole moment, which (also) strongly correlates with the number of surrounding H-bonds, the asymmetry for the two classes of 3-coordinated motifs is consistent with that for the other two properties but here it is overpowered by the effect of the number of H-bonds.

For the PDOS results, on the other hand, the sub-band corresponding to the 2A:1D motifs constitutes the most downshifted group of all in our calculations, even more than the 4- and 5-coordinated configurations, and 1A:2D is the least downshifted, i.e. the least affected by the environment. We note that we obtain the same results for our two independent MD simulations and

the different calculation methods for the dipole moment (Bader, Wannier) or the total molecular charge (Bader, Mulliken) also give consistent results.

It is treacherous to make direct comparisons between results from our "native" DOS bands calculated here and features of X-ray emission (XES) and absorption (XAS) spectra of liquid water, where excited energy levels and elaborate intensity weighting are involved in the process. It may suffice to mention that the 3-coordinated populations and the order between them (most or least affected by the environment) appears to vary between different studies, even for nominally the same type of spectra. Thus calculated XES spectra by Kashtanov et al. [68] show that the peaks of the central water molecules in the 2A:1D population (using our labelling scheme) are more downshifted in energy than the 1A:2D population, in agreement with our results, while the opposite is true for the calculated configuration-resolved spectra of Zhuvtobriukh et al. [9] (see also ample references to the group's previous work in these publications).

Some other molecular properties of water molecules in water-rich environments have also been found to display the asymmetry that we found here, i.e. that the 2A:1D coordination appears to affect the water molecules more than 1A:2D coordination. Thus, Liu and Ojamäe [69] performed quantum-mechanical calculations and analysed the OH stretching vibrations in three large $(\text{H}_2\text{O})_n$ clusters and found that the H-bonded water leg of the central water molecule in the 2A:1D class was overall much more downshifted than those of the 1A:2D class. This is consistent with the results obtained by Ohno et al. for smaller water clusters [70] and with the results obtained for even smaller model clusters by Ojamäe et al. [71].

In this paper, we furthermore showed that the structural properties calculated from simulations using the CP2K/BLYP-D3 approach for liquid D_2O agree well with total and partial RDFs from neutron diffraction experiments of both liquid D_2O and H_2O . Likewise total RDFs functions from X-ray diffraction techniques were mimicked by such calculations and the result was very satisfactory. Via calculations of the self-diffusion coefficients, it could be demonstrated here that our systems were in a (non-supercooled) liquid state; the calculated self-diffusion coefficients were in reasonable agreement with the experimental values, after the latter had been "back-corrected" for MD box size effects.

Acknowledgement

IB acknowledges partial financial support from the Hungarian Scientific Research Fund (at present: National Research, Development and Innovation Office, NKFIH) of Hungary (via grant OTKA K124885 and FK128656 K128136). KH acknowledges grants from the Swedish Research Council (Vetenskapsrådet) and the Swedish strategic e-science research programme eSENCE.. We are welcomed the generous computational support from NIF-HPC cluster. Sz. Pothoczki acknowledges that this project was supported by the János Bolyai Research Scholarship of the Hungarian Academy of Sciences.

References

- [1] J.-H. Guo, Y. Luo, A. Augustsson, J.-E. Rubensson, C. S athe, H.  gren, H. Siegbahn, and J. Nordgren,, *Physical Review Letters* **2002**, 89, 137402- - 4.
- [2] P. Hunt, M. Sprik, R. Vuilleumier, *Chem. Phys. Lett.* **2003**, 376, 68-74.
- [3] Ph. Wernet, D. Nordlund, U. Bergmann, M. Cavalleri, M. Odelius, H. Ogasawara, L.  . N slund, T. K. Hirsch, L. Ojam e, P. Glatzel, L. G. M. Pettersson, A. Nilsson, *Science* **2004**, 304, 995.
- [4] D. Prendergast, J. C. Grossman, G. Galli, *J. Chem. Phys.* **2005**, 123, 014501.
- [5] P. C. do Couto, S. G. Estacio, B. J. C. Cabral, *J. Chem. Phys.* **2005**, 123, 054510.
- [6] B. Winter, E. F. Aziz, U. Hergenhahn, M. Faubel, I. V. Hertel, *J. Chem. Phys.* **2007**, 126, 124504.
- [7] A. Nilsson, L. G. M. Pettersson, *Chemical Physics* **2011**, 389, 1.
- [8] T. Fransson, Y. Harada, N. Kosugi, N.A. Besley, B. Winter, L.G.M. Petersson, A. Nilsson, *Chem. Rev.* **2016**, 116, 7551.
- [9] I. Zhovtobriukh, N. A. Besley, T. Fransson , A. Nilsson, and L. G. M. Pettersson, *J.Chem.Phys.* **2018**, 148, 144507.
- [10] M. Leetmaa, M. P. Ljungberg, A. P. Lyubartsev, A. Nilsson, L. G. M. Pettersson, *Journal of Electron Spectroscopy and Related Phenomena* **2010**, 177, 135.
- [11] L. Liang, P. Rulis, L. Quyang, W. Y. Chin, *Phys. Rev. B* **2011**, 83, 024201.
- [12] R. Z. Khaliullin, T. D. K hne, *Phys. Chem. Chem. Phys.* **2013**, 15, 15746.
- [13] N. Marzari and D. Vanderbilt, *Phys. Rev. B: Condens. Matter Mater. Phys.*, **1997**, 56, 12847.
- [14] R. F. W. Bader and C. F. Matta, *J. Phys. Chem. A* **2004**, 108, 8385.
- [15] P. L. Silvestrelli, M. Parrinello, *J. Chem. Phys.* **1999**, 111, 3572.
- [16] L. Delle Site, A. Alavi, R.M. Lynden-Bell. *Mol. Phys.* **1999**, 96, 1683-1693.
- [17] P. J. Dyer, P. T. Cummings, *J. Chem. Phys.* **2006**, 125, 144519.
- [18] G. Lippert, J. Hutter, and M. Parrinello, *Mol. Phys.* **1997**, 92, 477.
- [19] J. VandeVondele, M. Krack, F. Mohamed, M. Parrinello, T. Chassaing, and J. Hutter, *Comput. Phys. Commun.* **2005**, 167, 103.
- [20] <https://www.cp2k.org>.
- [21] A. D. Becke, *Phys. Rev. A* **1988**, 38, 3098.
- [22] C. Lee, W. Yang, R. G. Parr, *Phys. Rev. B* **1988**, 37, 785.
- [23] S. Grimme, J. Antony, S. Ehrlich, H. Krieg, *J. Chem. Phys.* **2010**, 132, 154104.
- [24] S. Goedecker, M. Teter, J. Hutter, *Phys. Rev. B* **1996**, 54, 1703.
- [25] J. Schmidt, J. VandeVondele, I. F. W. Kuo, D. Sebastiani, J. I. Siepmann, J. Hutter, C. J. Mundy, *J. Phys. Chem. B* **2009**, 113, 11959
- [26] R. Jonchiere, A. P. Seitsonen, G. Ferlat, A. M. Saitta, and R. Vuilleumier, *J. Chem. Phys.* **2011**, 135, 154503.
- [27] P.G. Hill, P.G., R. D.MacMillan, and V. Lee, AECL-7531 (Dec. 1981).
- [28] M. J. Gillan, D. Alf e, and A. Michaelides *J. Chem. Phys.* **2016**, 144, 130901.
- [29] T. D. K hne, M. Krack, and M. Parrinello, *J. Chem. Theory Comput.* **2009**, 5, 235.
- [30] A. P. Seitsonen, T. Bryk, *Phys. Rev. B* **2016**, 94, 184111.
- [31] H. J. C. Berendsen, J. R. Grigera, and T. P. Straatsma. *J. Phys. Chem* **1987**, 91, 6269.
- [32] G. Henkelman, A. Arnaldsson, H. Jonsson, *Comp. Mat. Science* **2006**, 36, 354.
- [33] W. Tang, E. Sanville and G. Henkelman, *J. Phys.: Condens. Matter* **2009**, 21, 084204.
- [34] B. Efron, *Ann. Statist.* **1979**, 7, 1.
- [35] R. Beran. *Stat. Sci.* **2003**, 18, 175

- [36] P. Qian, W. Song, L. Lu, Z. Yang, *Int. J. Quantum Chem.* **2010**, *110*, 1923.
- [37] D. H. Brookes, T. Head-Gordon, *J. Phys. Chem. Lett.* **2015**, *6*, 2938
- [38] A. K. Soper, *J. Phys. Chem. B* **2011**, *115*, 14014.
- [39] L. B. Skinner, C. J. Benmore, J. C. Neufeind, and J. B. Parise, *J. Chem. Phys.* **2014**, *141*, 214507.
- [40] L. B. Skinner, C. Huang, D. Schlesinger, L. G. M. Pettersson, A. Nilsson, C. J. Benmore, *J. Chem. Phys.* **2013**, *138*, 074506.
- [41] T. Morawietz, A. Singraber, C. Dellago, and J. Behler, *PNAS* **2016**, *113*, 836
- [42] G. Hura, J. M. Sorenson, R. M. Glaeser, T. Head-Gordon; *J. Chem. Phys.* **2006**, *113*, 9140
- [43] T. Morawietz, A. Singraber, C. Dellago, and J. Behler, *PNAS* **2016**, *113*, 8368.
- [44] M. J. McGrath, I. F. W. Kuo, J. I. Siepmann, *Phys. Chem. Chem. Phys.* **2011**, *13*, 19943.
- [45] A. K. Soper, *Chem. Phys.* **2000**, *258*, 121 .
- [46] M. Machida, K. Kato, and M. Shiga *J. Chem. Phys.* **2018** *148*, 102324
- [47] M. Ceriotti, W. Fang, P. G. Kusalik, R. H. McKenzie, A. Michaelides, M. A. Morales, and T. E. Markland *Chem. Rev.* **2016**, *116*, 7529
- [48] K. R. Harris, *Phys. Chem. Chem. Phys.*, 2002, *4*, 5841–5845
- [49] M. J. McGrath, I. F. W. Kuo, J. I. Siepmann, *Phys. Chem. Chem. Phys.* **2011**, *13*, 19943.
- [50] K. N. Marsh, Ed., *Recommended Reference Materials for the Realization of Physicochemical Properties*, Blackwell, Oxford, 1987.
- [51] F. Tran and P. Blaha *J. Phys. Chem. A* **2017**, *121*, 3318
- [52] J.P. Perdew, M.; Levy, *Phys. Rev. Lett.* **1983**, *51*, 1884.
- [53] D. Bagayoko *AIP ADVANCES* **2014**, 127104
- [54] B. Winter, E. F. Aziz, U. Hergenhahn, M. Faubel, I. V. Hertel, *J. Chem. Phys.* **2007**, *126*, 124504.
- [55] R. Kumar, J. R. Schmidt, and J. L. Skinner, *J. Chem. Phys.* **2007**, *126*, 204107
- [56] A. Geiger, F. H. Stillinger, and A. Rahman, *J. Chem. Phys.* **1979**, *70*, 4185.
- [57] H. E. Stanley and J. Teixeira, *J. Chem. Phys.* **1980**, *73*, 3404.
- [58] F. Sciortino, and S. L. Fornili, *J. Chem. Phys.* **1989**, *90*, 2786.
- [59] P. Jedlovsky, J. P. Brodholt, F. Bruni and M. A. Ricci, A. K. Soper, R. Vallauri, *J. Chem. Phys.* **1998**, *108*, 8528.
- [60] I. Brovchenko, A. Geiger, A. Oleinikova, *Phys. Chem. Chem. Phys.* **2004**, *6*, 1982.
- [61] R. A. DiStasio Jr., B. Santra, Z. Li, X. Wu, R. Car, *J. Chem. Phys.* **2014**, *141*, 084502.
- [62] F. C. Lightstone, E. Schwegler, R. Q. Hood, F. Gygi and G. Galli, *Chem. Phys. Lett.* **2001**, *343*, 549.
- [63] S. Bogatko, E. Cauet, E. Bylaska, G. Schenter, J. Fulton and J. Weare, *Chem. – Eur. J.* **2013**, *19*, 3047.
- [64] P. D. Mitev, I. Bakó, A. Eriksson, K. Hermansson, *Phys. Chem. Chem. Phys.* **2014**, *16*, 9351.
- [65] A. Sen, P. D. Mitev, A. Eriksson, and K. Hermansson, *International Journal of Quantum Chemistry* **2016**, *116*, 67–80.
- [66] I. Bakó, I. Mayer, *J. Phys. Chem A.* **2016**, *120*, 631.
- [67] M. Alfredsson, K. Hermansson, *J. Chem. Phys.* **1999**, *111*, 1993.
- [68] S. Kashtanov, A. Augustsson, Y. Luo, J.-H. Guo, C. Sâthe, J.-E. Rubensson, H. Siegbahn, J. Nordgren, and H. Ågren, *Phys. Review B* **2004**, *69*, 024201.
- [69] Y. Liu, L. Ojamäe, *J Mol Model* **2014**, *20*, 2281.
- [70] K. Ohno, M. Okimura, N. Akai, Y. Katsumoto, *Phys. Chem. Chem. Phys.* **2005**, *7*, 3005–3014 .
- [71] L. Ojamäe, K. Hermansson, *J. Phys. Chem.* **1994**, *98*, 4271–4282.

Tables

Table 1. The orbital energy difference between the three main valence band peaks in the electronic DOS spectrum calculated from the analysis of the data from our BLYP-D3 AIMD simulation of liquid water at 300 K. " ϵ " is the average peak value of the orbital band in question. Comparison with photoelectron spectroscopy experiments is shown.

	$\epsilon(1b_1)-\epsilon(3a_1)$ (eV)	$\epsilon(1b_1)-\epsilon(1b_2)$ (eV)	$\epsilon(1b_1)-\epsilon(2a_1)$ (eV)
AIMD 300K (this work)	1.9	5.7	18.9
Experiment [6]	2.1	5.6	17.8

Table 2.

Characteristics of hydrogen bonding in liquid water from AIMD simulations using the BLYP-D3 functional at two temperatures. The table gives the fraction of water molecules belonging to a certain class of H-bond coordination, labelled $n_A A: n_D D$, where n_A is the number of H-bonds accepted by a central water molecule and n_D is the number of donated H-bonds; $n_D + n_A = n_{HB}$. The line labelled "Others" accounts for all the very small remaining contributions from very scarcely populated configurations.

A. 300 K							
$n_A A: n_D D$	$n_{ab}=0$	$n_{HB}=1$	$n_{HB}=2$	$n_{HB}=3$	$n_{HB}=4$	$n_{HB}=5$	$\langle n_{HB} \rangle$
1A:1D			6.7%				2
2A:1D				10.4%			3
1A:2D				16.7%			3
2A:2D					57.8%		4
3A:2D						5.4%	5
Others	0.2%	1.1%	0.9%	0.1%	0.6%	0.1%	
All data	0.2	1.1%	7.6%	27.2%	58.4%	5.5%	3.59
B. 350 K							
$n_A A: n_D D$	$n_{HB}=0$	$n_{HB}=1$	$n_{HB}=2$	$n_{HB}=3$	$n_{HB}=4$	$n_{HB}=5$	
1A:1D			13.4%				2
2A:1D				15.0%			3
1A:2D				21.7%			3
2A:2D					37.5%		4
3A:2D						4.7%	5
Others	0.3	3.5%	2.6%	0.1%	1.2%	0.1%	
All data	0.3	3.5%	16.0%	36.7%	38.7%	4.8%	3.23

Figures

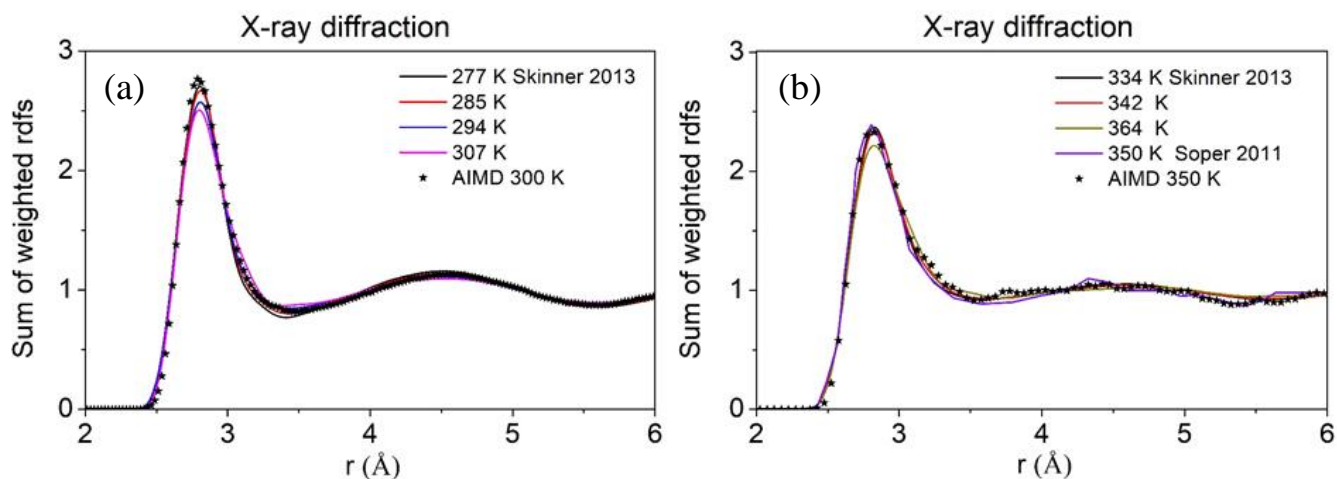


Fig. 1. Total weighted (according to the atomic X-ray scattering factors) radial distribution functions (rdfs) based on trajectories from our BLYP-D3 AIMD simulations and from X-ray diffraction experiments [38-40]. See text for details. (a) 300 K, (b) 350 K.

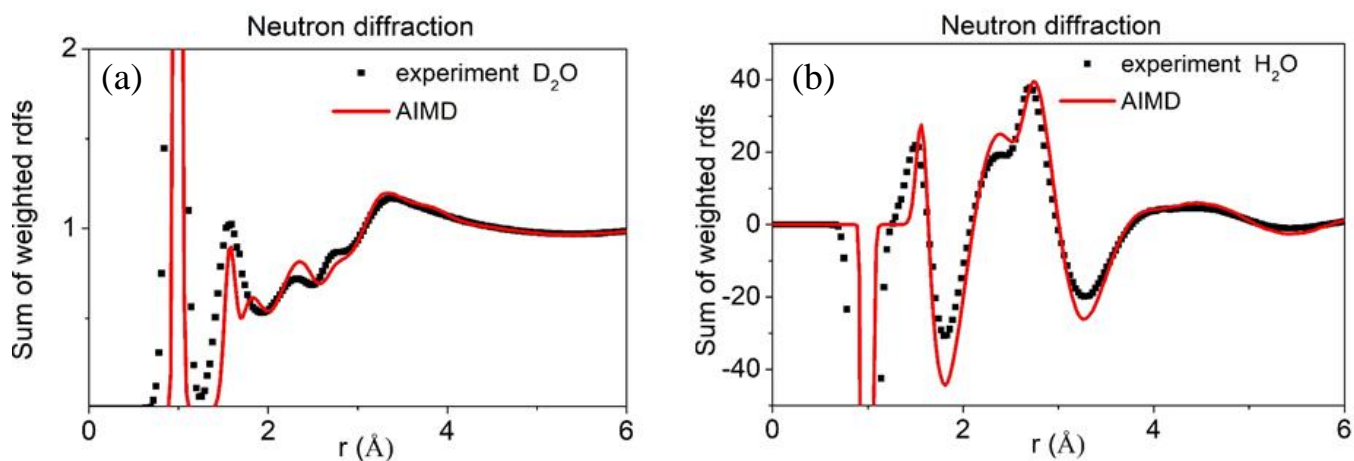


Fig. 2. Total weighted rdfs (based on partial rdfs and the neutron scattering lengths). The partial rdfs are either based on trajectories from our BLYP-D3 AIMD simulations or from neutron diffraction experiments [45], both at 300 K. See text for details. (a) D₂O liquid water, (b) H₂O liquid water. The negative values seen in (b) are due to the negative neutron scattering length of the ¹H isotope.

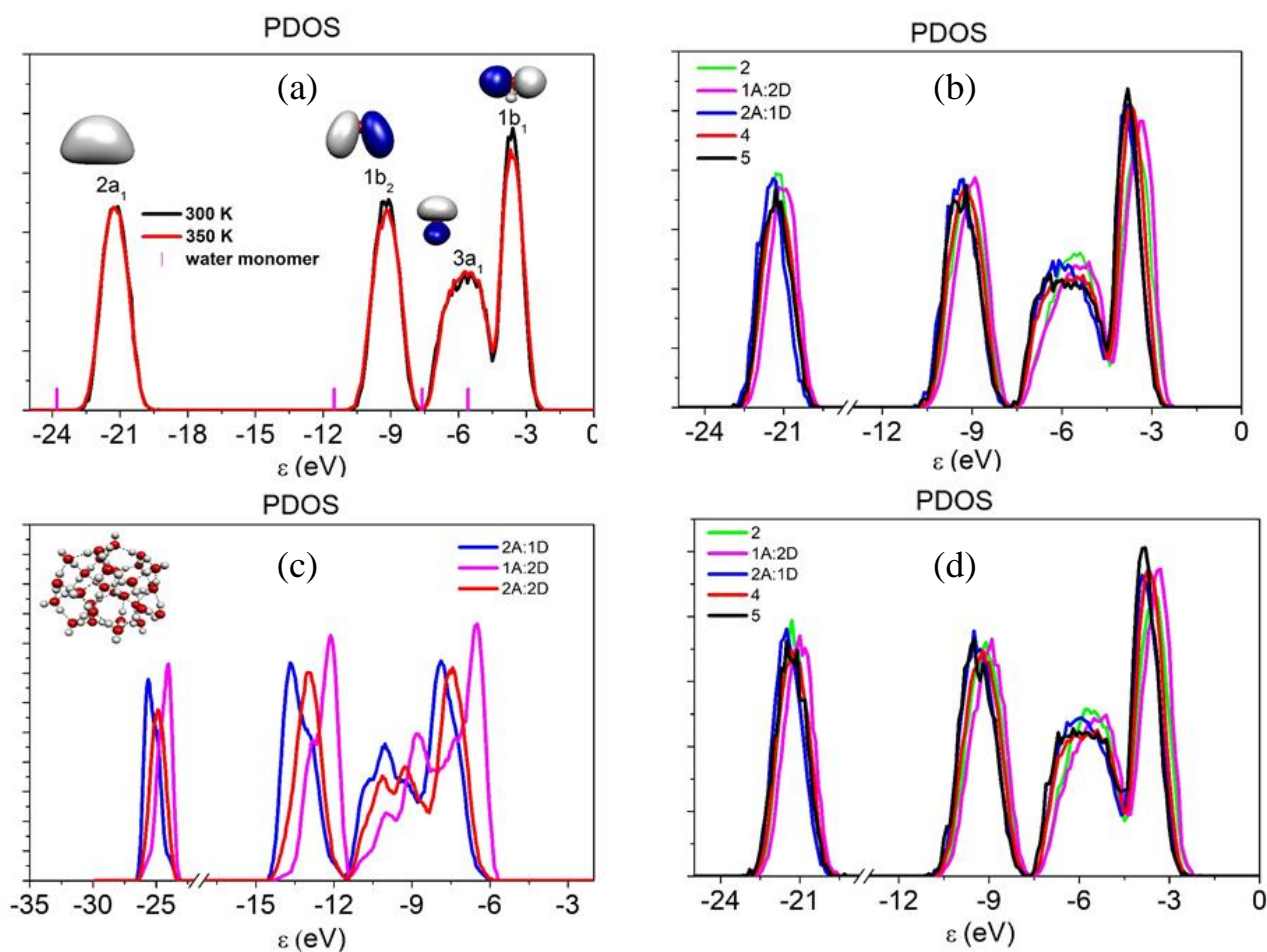


Fig 3. Calculated electronic PDOS graphs from our AIMD simulation of liquid water and from an optimised cluster cluster calculation. The PDOS curves were calculated for 480 snapshots, giving in total 30720 analysed water molecules. No specific alignment was made between the cluster and liquid results so the relative PDOS scales are arbitrary.

- (a) PDOS for all water molecules in the many selected snapshots at 300 and 350 K.
- (b) PDOS calculated separately for the different coordination classes at 300K.
- (c) PDOS calculated separately for the different coordination classes at 350K.
- (d) PDOS calculated separately for three of the coordination classes for an optimized $(\text{H}_2\text{O})_{30}$ cluster. The cluster calculations were performed with periodic calculations and the BLYP-D3 functional (just as the liquid water simulations).

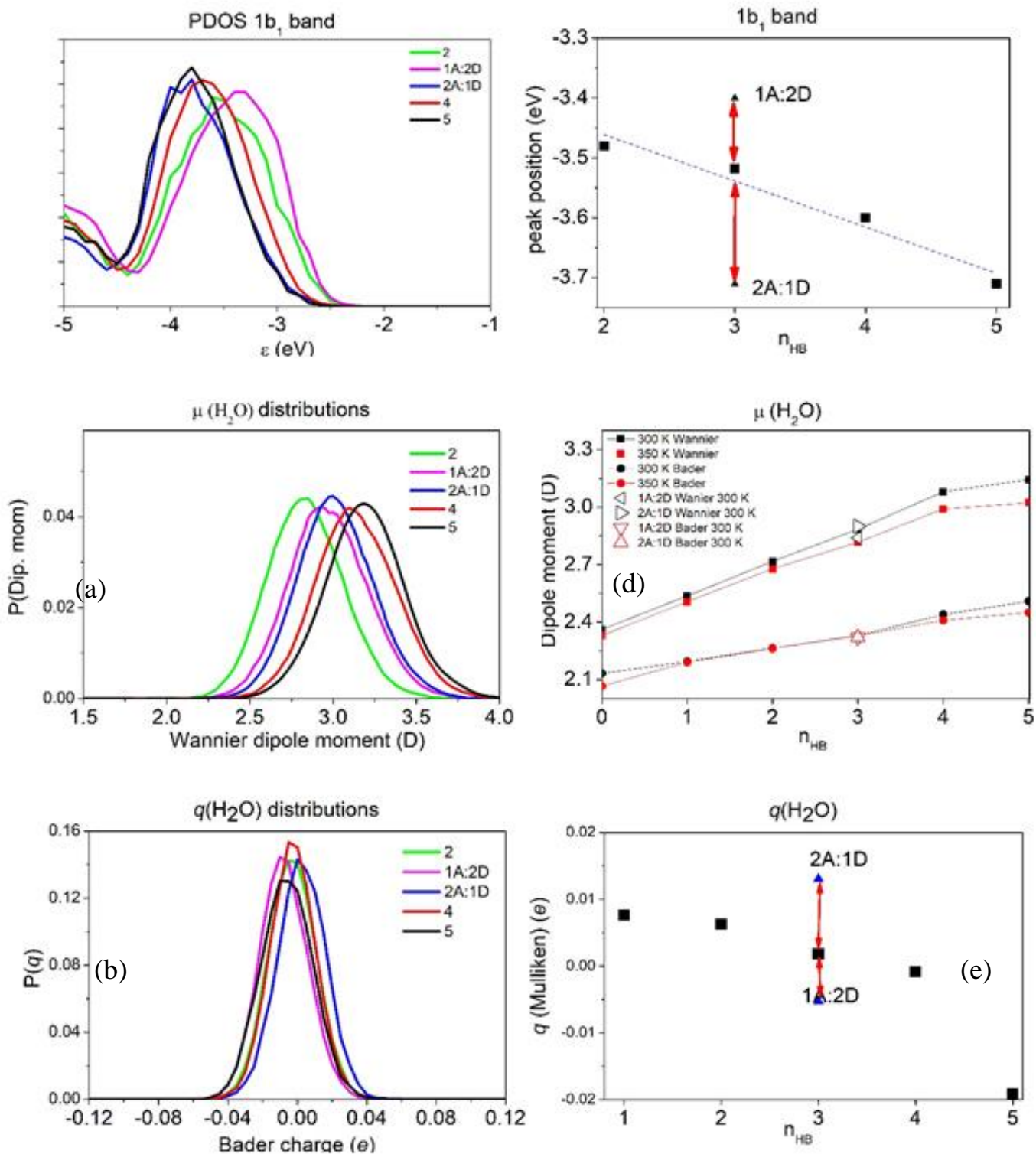


Fig 4. Electronic properties of water molecules in different H-bonding environments analysed from the 300 K AIMD simulations.

- (a) PDOS ($1b_1$ band) for water molecules in the different coordination classes.
- (b) Dipole moment distributions for water molecules in the different coordination classes.
- (c) Net molecular Mulliken charge distributions for water molecules in the different coordination classes.
- (d) Average position of the $1b_1$ peak *versus* n_{HB} and coordination class.
- (e) Average molecular dipole moment *versus* n_{HB} and coordination class.
- (f) Average net molecular Mulliken charge *versus* n_{HB} and coordination class.