

Models

David J. Huggins<sup>a, b, c</sup>

Affiliations:

<sup>a</sup>University of Cambridge, Cambridge Molecular Therapeutics Programme, Hutchison/MRC Research Centre, Hills Road, Cambridge, CB2 0XZ, United Kingdom

<sup>b</sup>University of Cambridge, Department of Chemistry, Lensfield Road, Cambridge, UK CB2 1EW, United Kingdom

<sup>c</sup>University of Cambridge, TCM Group, Cavendish Laboratory, 19 J J Thomson Avenue, Cambridge CB3 0HE, United Kingdom

Corresponding Author: David J. Huggins

Email: djh210@cam.ac.uk

Updated: 20 December 2011

## Abstract

Water is one of simplest molecules in existence, but also one of the most important in biological and engineered systems. However, understanding the structure and dynamics of liquid water remains a major scientific challenge. Molecular dynamics simulations of liquid water were performed using the water models TIP3P-Ewald, TIP4P-2005, TIP5P-Ewald and SWM4-NDP to calculate the radial distribution functions, the relative angular distributions and the excess enthalpies, entropies and free energies. In addition, lower-order approximations to the entropy were considered, identifying the fourth-order approximation as an excellent estimate of the full entropy. The second-order and third-order approximations are approximately 20% larger and smaller than the true entropy respectively. All four models perform very well in predicting the radial distribution functions, with the TIP5P-Ewald model providing the best match to the experimental data. The models also perform well in predicting the excess entropy, enthalpy and free energy of liquid water. The TIP4P-2005 and SWM4-NDP models are more accurate than the TIP3P-Ewald and TIP5P-Ewald models in this respect. However, the relative angular distribution functions of the four water models reveal notable differences. The TIP5P-Ewald model demonstrates an increased preference for water molecules to act both as tetrahedral hydrogen bond donors and acceptors, whereas the SWM4-NDP model demonstrates an increased preference for water molecules to act as planar hydrogen bond acceptors. These differences are not uncovered by analysis of the RDFs or the commonly employed tetrahedral order parameter. However, they are expected to be very important when considering water molecules around solutes and are thus a key consideration in modelling solvent entropy.

## Introduction

At the atomic level, water is one of the simplest molecules in existence. However, understanding the bulk phases of water remains a major scientific challenge and has inspired debate for many years [1-2]. The majority of efforts in modeling liquid water have been directed toward Monte Carlo (MC) and molecular dynamics (MD) simulations. These allow time-averaged properties such as density or heat capacity to be calculated and compared with experiment [3]. MC and MD simulations that employ a forcefield require a set of parameters to describe the structure and properties of water molecules and there has been considerable research into the resulting water models. The simplest models include three sites in total, two hydrogen atoms and an oxygen atom. Such models include the SPC [4], SPC/E [5], TIP3P [6] and TIP3P-Ewald [7] models. A polarisable three site AMOEBA water model has also been developed [8]. Four site models such as TIP4P [9], TIP4P-Ewald [10] and TIP4P-2005 [11] use an extra atom with zero mass to represent a charged site. Five site models have also been developed. The ST2 [12], TIP5P [3] and TIP5P-Ewald [13] models use two lone pairs whereas polarisable models such as SWM4-DP [14] and SWM4-NDP [15] use one massless charged site and one Drude particle [16]. The six site model TIP6P, designed primarily for studying ice, comprises an oxygen atom, two hydrogen atoms, two lone pairs and one massless charged site at the centre of mass [17]. These models differ in their sophistication and also in the time required for simulation. When attempting to model a real system using MC or MD with explicit water molecules, it is thus important to select a water model which provides the correct balance of speed and accuracy. However, whilst the speed is easy to measure, it is very difficult to gauge the accuracy of the different models because experiment has not provided us with a complete picture of the structure and dynamics of liquid water.

The experimental techniques of X-ray scattering [18-19] and neutron diffraction [20-21] provide structural data which can be compared to the predictions of simulation. Application of these methods has provided radial distribution functions (RDFs) for the oxygen-oxygen (O-O), oxygen-hydrogen (O-H), and hydrogen-hydrogen (H-H) distances in liquid water. Previous work has compared some or all of such data with the simulation results for these RDFs for a subset of the water models [22-24]. In this work the RDFs given by the four water models TIP3P-Ewald, TIP4P-2005, TIP5P-Ewald and SWM4-NDP were compared to these three experimental RDFs. These represent three, four and five sites water models in addition to a polarisable model. In addition, these four models were explored in terms of the relative orientations of neighboring water molecules. This property has been considered previously, but has not been completely explored yet it is expected to be very important in understanding liquid water. In particular, orientational correlations have an important effect on the entropy of water, both in bulk water [25-26] and in biological complexes [27-29]. Recent work has shown that the choice of water model affects protein folding [30] and the results of binding free energy calculations [31]. The effect of the water model on orientational correlations in liquid water is thus an area of great importance and one which has not been fully explored.

In this study, orientational correlations in the water models TIP3P-Ewald, TIP4P-2005, TIP5P-Ewald and SWM4-NDP were considered. These models are displayed schematically in Figure 1 and the parameters for the models are presented in Table I.

## Materials and Methods

MD simulations of bulk water were performed using NAMD[32] with the water models TIP3P-Ewald, TIP4P-2005, TIP5P-Ewald and SWM4-NDP. RDFs, orientational correlations and the calculated excess enthalpies, entropies and free energies were then considered.

### *System Setup*

To create the systems, the SOLVATE program[33] version 1.0 from the Max Planck Institute was used to generate a water sphere of radius 50 Å. No ions were included in the simulations. This stage of preparation was undertaken to generate a reasonable water density. This sphere was then cut to a cube of length 25.0 Å. All hydrogen atoms were then deleted from the systems and all the necessary hydrogen atoms, lone pairs and Drude particles were built using the appropriate geometry for each water model. The new water molecules were all oriented with the dipole vector aligned with the x-axis and the hydrogen atoms lying in the xy-plane. This stage of preparation was undertaken to ensure that the geometries of the water molecules were standardized. Each water cube contained 526 water molecules.

### *Equilibration*

All systems were treated using periodic boundary conditions and the electrostatics were modeled using the particle mesh Ewald method [34]. The systems were first subjected to MD equilibration for 100 ps at 300 K in an NPT ensemble and then MD equilibration for 1 ns at 300 K in an NVT ensemble. All systems were brought to equilibrium before continuing the

simulations by verifying that the systems reached a point where the energy fluctuations were stable.

### *Molecular Dynamics*

Production simulations were performed for 40.0 ns at 300 K. All MD simulations were performed using the NAMD program version 2.7b3[35] with the CHARMM27 force field[36-37]. TIP3P-Ewald, TIP4P-2005 and TIP5P-Ewald were simulated using an MD time step of 2.0 fs and SWM4-NDP was simulated using an MD time step of 1.0 fs. Electrostatic interactions were modelled with a uniform dielectric and a dielectric constant of 1.0 throughout the setup and production runs. Van der waals interactions were truncated at 12.0 Å with switching from 8.0 Å. Water molecules were modelled as rigid, employing the SETTLE algorithm [38]. All MD simulations were performed using NAMD compiled for use with CUDA-accelerated GPUs.

### *Energy Evaluations*

The average interaction energy ( $E_{\text{interaction}}$ ) was calculated from 1000 snapshots across the first 8ns of each simulation with one taken every 8 ps. All water molecules in the periodic box were considered, including their interactions with neighboring boxes. To generate the excess enthalpy from the calculated interaction energy, three adjustments are required. The first is a correction to include the potential energy of the Drude particle ( $E_{\text{drude}}$ ) for the SWM4-NDP model.

$$E_{Drude} = 1/2 kd^2 \quad (1)$$

In this equation,  $k$  is the Drude force constant of  $1000 \text{ kcal/mol/\AA}^2$  and  $d$  is the distance of the Drude particle from its reference position at the oxygen atom. This contribution is included in the potential energy term calculated by NAMD. The second adjustment is to include a polarisation correction ( $E_{pol}$ ), which accounts for the energetic penalty incurred due to the increase in the dipole of a water molecule in the liquid phase. This correction is dependent on the calculated liquid phase dipole of each model and has the form [5, 10, 39]:

$$E_{pol} = \frac{1}{2} (\mu_{liquid} - \mu_{gas})^2 / \alpha_{gas} \quad (2)$$

$\mu_{liquid}$  and  $\mu_{gas}$  are the dipole moments in the liquid and the gas phase and  $\alpha_{gas}$  is the mean polarisability of a gas phase water molecule. This correction takes the values of  $1.132 \text{ kcal/mol}$ ,  $0.975 \text{ kcal/mol}$ , and  $0.911 \text{ kcal/mol}$  for the models TIP3P-Ewald, TIP4P-2005, and TIP5P-Ewald respectively. The SWM4-NDP model does not require a polarisation correction as the polarisation penalty  $E_{drude}$  is calculated by equation 1. The third adjustment is to calculate the excess enthalpy from the excess energy. This requires a work term to describe the change in volume between the liquid and the gas. As  $PV$  for the gas is much greater than  $PV$  for the liquid, this is approximated by  $\Delta(PV) = PV_{liquid} - PV_{gas} = -PV_{gas} = -RT$ . The total excess enthalpy per mole ( $H_{excess}$ ) is therefore:

$$H_{excess} = E_{interaction} + E_{drude} + E_{pol} - RT \quad (3)$$

In previous papers, water models were assessed by considering the interaction energy ( $E_{interaction}$ ) or the standard heat of vaporisation ( $-H_{excess}$ ), in some cases corrected by  $E_{pol}$ . This paper focuses on the corrected excess enthalpy ( $H_{excess}$ ) given in Equation 3.

### *Correlation Functions*

The relative positions of each pair of water molecule can be described by the distance between the oxygen atoms  $R$ , and five angles  $\omega^{rel}$ . The five  $\omega^{rel}$  angles are denoted  $\theta_1$ ,  $\theta_2$ ,  $\chi_1$ ,  $\chi_2$ , and  $\varphi$  and are illustrated in Figure 2. The angles  $\theta_1$  and  $\theta_2$  describe the angles between the dipole vectors of each water molecule and the intermolecular axis. The angles  $\chi_1$  and  $\chi_2$  describe the rotation of each water molecule around its dipole vector and are calculated from the angles between the intermolecular axis and vectors normal to the plane of the water molecules. The angle  $\varphi$  describes the rotation around the intermolecular axis [25]. The correlation functions were calculated from the production simulations, using all available water pairs in each case. The correlation functions can be used to calculate the excess entropy ( $S_{excess}$ ):

$$S_{excess} = -\frac{1}{2}k\rho \int [g(R, \omega^{rel}) \ln g(R, \omega^{rel}) - g(R, \omega^{rel}) + 1] dR d\omega^{rel} \quad (4)$$



$k$  is Boltzmann's constant,  $\rho$  is the number density of bulk water, and  $g(R, \omega^{rel})$  is the correlation function. The excess entropy can be separated into translational and orientational contributions to allow the calculation of the excess translational entropy ( $S_{trans}$ ) and the excess orientational entropy ( $S_{orient}$ ):

$$g(R, \omega^{rel}) = g(R)g(\omega^{rel}|R) \quad (5)$$

$$S_{excess} = S_{trans} + S_{orient} \quad (6)$$

$$S_{trans} = -\frac{1}{2}k\rho \int [g(R) \ln g(R) - g(R) + 1]dR \quad (7)$$

The RDFs were calculated as histograms with radial bin sizes of 0.03 Å and were compared to the experimentally determined RDFs [20]. The orientational entropy was computed separately for each radial shell ( $S_{shell}$ ) and then combined to yield the total orientational entropy:

$$S_{shell} = -\frac{1}{\Omega} \int g(\omega^{rel}|R) \ln g(\omega^{rel}|R) d\omega \quad (8)$$

$$S_{orient} = -\frac{1}{2}k\rho \int g(R)S_{shell} dR \quad (9)$$

$\Omega$  is the integral over all five  $\omega$  angles. The orientational pair correlation functions (PCFs) were calculated as histograms in shells of radial bin size 0.1 Å and angular bin size 10°. The correlation function  $g(\omega^{rel}|R)$  can be integrated, taking advantage of the symmetry of the water molecule. The interchangeability of the water molecules allows  $\theta_1$  to be integrated over the range 0 to  $\pi$  and  $\theta_2$  to be integrated from  $\theta_1$  to  $\pi$ . Furthermore, the symmetry of the water molecules allows  $\chi_1$ ,  $\chi_2$  and  $\varphi$  to be integrated over the range 0 to  $\pi$  [25]. To compare relative angular distributions in five dimensions, the RMSD between the probability densities was calculated for each radial shell from each pair of 40ns production simulation:

$$Angular\ RMSD(R) = \sqrt{\sum_{\omega^{rel}} \frac{[g^a(\omega^{rel}|R)dRd\omega^{rel} - g^b(\omega^{rel}|R)dRd\omega^{rel}]^2}{N}} \quad (10)$$

In equation 10,  $g^a(\omega^{rel}|R)$  and  $g^b(\omega^{rel}|R)$  are the angular correlation functions for water models a and b and N is the number of angular bins. The total RMSD between the PCFs quantifies the difference between two distributions across all radial bins and all five sets of angular bins and was also calculated:

$$Total\ RMSD = \sqrt{\sum_R \sum_{\omega^{rel}} \frac{[g^a(R,\omega^{rel})dRd\omega^{rel} - g^b(R,\omega^{rel})dRd\omega^{rel}]^2}{N}} \quad (11)$$

$g^a(\mathbf{R}, \omega^{\text{rel}})$  and  $g^b(\mathbf{R}, \omega^{\text{rel}})$  are the correlation functions for water models a and b and N is the total number of bins.

### *Entropy Approximations*

In addition to the six dimensional PCF calculations, probability density distributions were also calculated for each of the five individual angles  $\theta_1$ ,  $\theta_2$ ,  $\chi_1$ ,  $\chi_2$ , and  $\varphi$  as well as for each pair, triple, and quartet of angles. This facilitates the calculation of lower order approximations to the entropy and allows the lower order distributions to be viewed. The lower order entropy approximations are generated by truncations of the mutual information expansion [40-41]. The first, second, third and fourth order approximations to the entropy were computed, in addition to the F7 approximation that has been employed previously [25].

$$S_{1st\ Order} = \sum_{\binom{5}{1}C} S(\alpha) \quad (12)$$

$$S_{2nd\ Order} = \sum_{\binom{5}{2}C} S(\alpha, \beta) - 3 \sum_{\binom{5}{1}C} S(\alpha) \quad (13)$$

$$S_{3rd\ Order} = \sum_{\binom{5}{3}C} S(\alpha, \beta, \gamma) - 2 \sum_{\binom{5}{2}C} S(\alpha, \beta) + 3 \sum_{\binom{5}{1}C} S(\alpha) \quad (14)$$

$$S_{4th\ Order} = \sum_{\binom{5}{4}C} S(\alpha, \beta, \gamma, \Delta) - \sum_{\binom{5}{3}C} S(\alpha, \beta, \gamma) + \sum_{\binom{5}{2}C} S(\alpha, \beta) - \sum_{\binom{5}{1}C} S(\alpha) \quad (15)$$

$$S_{F7} = S(\theta_1, \theta_2) + S(\chi_1, \chi_2) + S(\theta_1, \chi_2) + S(\theta_2, \chi_1) + S(\varphi) - S(\theta_1) - S(\theta_2) - S(\chi_1) - S(\chi_2) \quad (16)$$

In these equations,  $S(\alpha)$  is the entropy associated with the angle  $\alpha$ ,  $S(\alpha, \beta)$  is the joint entropy associated with the angles  $\alpha$  and  $\beta$ ,  $S(\alpha, \beta, \gamma)$  is the joint entropy associated with the angles  $\alpha$ ,  $\beta$  and  $\gamma$  and  $S(\alpha, \beta, \gamma, \Delta)$  is the joint entropy associated with the angles  $\alpha$ ,  $\beta$ ,  $\gamma$  and  $\Delta$ . The indices  ${}^5C_m$  on each sum represent all combinations of the five angles for a given order  $m$ . All entropies calculated here exclude internal vibrational entropy changes, as the water models are treated as rigid. The excess free energies can be calculated using equation 17.

$$G_{excess} = H_{excess} - TS_{excess} \quad (17)$$

The calculated excess enthalpies, entropies and free energies for the four models can then be compared with the experimental data.

#### *Orientational Order Parameters*

The tetrahedrality of each water model was assessed by calculating an orientational order parameter  $q$  [42-43]:

$$q = 1 - \frac{3}{8} \sum_{j=1}^3 \sum_{k=j+1}^4 (\cos\psi + \frac{1}{3})^2 \quad (18)$$

$\psi$  is the angle formed by the lines joining the oxygen atom of the water molecule to the oxygen atoms of its nearest neighbouring water molecules  $j$  and  $k$ . The four nearest neighbours are considered in each case. This order parameter describes the tetrahedral arrangement of oxygen atoms around an oxygen atom. The value of  $q$  varies between 0 in a random network and 1 in a tetrahedral network. In this case the average value of  $q$  was calculated from 10000 snapshots, considering all waters within 5 Å of the centre of the box. An additional orientational order parameter  $w$  was also calculated:

$$w = 1 - \frac{9}{16} \sum_{m=1}^2 \sum_{n=1}^2 (\cos\lambda + \frac{1}{3})^2 \quad (18)$$

$\lambda$  is the angle formed by the oxygen atom to hydrogen atom bond vector of one water molecule ( $m$ ) and the vector between its oxygen atom and a hydrogen atom of a neighbouring water molecule ( $n$ ). The two nearest neighbouring hydrogen atoms are considered. This order parameter describes the tetrahedral arrangement of hydrogen atoms around an oxygen atom. The value of  $w$  varies between 0 in a random arrangement and 1 in a tetrahedral arrangement. The average value of  $w$  was calculated from 10000 snapshots, considering all waters within 5 Å of the centre of the box.

## Results

The ability of each model to accurately recreate the experimentally determined RDFs was assessed initially. The results for the O-O RDF are shown in Table II. The TIP5P-Ewald model clearly recreates the experimental data most accurately, with the smallest RMSD against the experimental data of 0.035. The agreement is very good, as can also be seen in Figure 3 which shows the plots for each model against experiment. Each 40ns simulation for a given water model was also split into five blocks of 8ns and compared each 8ns block with every other 8ns block to compute the RMSD between the two O-O RDFs. The maximum RMSD between any pair of 8ns simulations for TIP3P-Ewald, TIP4P-2005, TIP5P-Ewald and SWM4-NDP models was 0.0050, 0.0054, 0.0051, and 0.0055 respectively. This is approximately tenfold lower than the RMSD between any different pair of water models. Thus, the differences between the models appear to be significant. It is interesting to note that the SWM4-NDP model shows a marked difference from the experimental data with an RMSD of 0.102, mainly due to the trough after the first peak in the RDF. This was noted in the original paper describing this model [15] and remains a problem. In addition, the TIP4P-2005 model has the first peak in the O-O and O-H RDFs at a higher amplitude than seen in the experimental data. TIP5P also recreates the experimental data for the O-H RDF most accurately, as shown in Table II and Figure 4. However, in this case the RMSD against the experimental data is 0.073 and the model does not fit the experimental data as well as for the O-O RDF. For the H-H RDF, all the models match the experimental data reasonably well, as shown in Table II and Figure 5. The TIP4P-2005 model is most accurate with an RMSD of 0.020 and in this case the TIP5P model is the least accurate with an RMSD of 0.043. All four models show a double peak in the H-H RDF around 4.0 Å that is not present in the experimental data. The two peaks seen for all four models at 3.5 Å and 4.5 Å

can be attributed to two distinct patterns of ordering in the second solvation shell that does not reflect the dynamics of liquid water observed experimentally. In summary, the four models all do reasonably well at recreating the three RDFs of liquid water. The TIP5P-Ewald model is the most accurate, whilst the SWM4-NDP model shows a discrepancy in the O-O RDF. As expected, the more complex water models require greater computational expense. As seen in Table II, the TIP4P-2005, TIP5P-Ewald and SWM4-NDP models run 1.7, 2.6, and 5.2 times slower than the TIP3P-Ewald model.

From a consideration of the three RDFs, the distributions of the relative angles for the four models was then considered. Whilst there is no experimental data for comparison in this case, the models can be compared with one another and the results can be analysed. The correlation functions for the angles  $\varphi$ ,  $\chi_1$ , and  $\theta_1$  within the range of 2.7-2.8 Å for the O-O distance in the four water models are shown in Figures 6, 7 and 8 respectively. The distribution of  $\varphi$  angles is flat for all four models, as noted previously for the TIP4P water model [25] and for other three and four site models [44]. From the simulations in this work, the five site TIP5P-Ewald model and the polarisable SWM4-NDP model are also unstructured along this angle. The  $\chi_1$  angle shows a similar distribution to that reported previously for the TIP4P water model [25] in Figure 7, with a peak at around 90°. However, the TIP5P-Ewald model shows a reduced preference for the 90° angle and an increased preference for the 0° (180°) angle. This increased preference for a 0° angle may reflect an increase in the tendency of the water to act as a hydrogen bond acceptor in a more structured fashion due to the directionality of the lone pair that is not present in the other models. A water model with perfect tetrahedral symmetry and two pairs of massless charge sites with opposite charge should show an equal preference for the 0° and 90° angles, as the hydrogen bond donor and acceptor geometries will be the same. The TIP3P model shows the

least structure and the SWM4-NDP model shows the greatest preference for the  $90^\circ$  angle with probability densities above 3.0. The most striking difference between the angular correlation functions is in the  $\theta_1$  distribution, shown in Figure 8. All four models show a peak at approximately  $50^\circ$  corresponding to the tetrahedral hydrogen bond donor configuration, but the TIP3P-Ewald, TIP4P-2005 and SWM4-NDP models show the second peak corresponding to the hydrogen bond acceptor configuration at  $180^\circ$ . The TIP5P-Ewald model shows this peak at approximately  $130^\circ$ , which reflects tetrahedral hydrogen bond acceptor geometry. Thus, the two peaks in the  $\theta_1$  distribution for TIP5P are consistent with a tetrahedral geometry and the preferred orientation for TIP5P is in good agreement with the calculated global minimum energy configuration for water dimers, which lies at  $\theta_1 \approx 55^\circ$  and  $\theta_2 \approx 125^\circ$  [25, 45]. Conversely, the SWM4-NDP shows a much greater preference for the  $180^\circ$  angle than the other models and a larger probability density at the  $50^\circ$  angle. It is thus acting as a hydrogen bond donor with a strong preference for tetrahedral geometry but as hydrogen bond acceptor with a greater preference for planar geometry rather than tetrahedral geometry. These differences are illustrated more compellingly by considering the distributions of pairs of angles. The distribution for  $\theta_1$  and  $\theta_2$  is shown in Figure 9. The TIP3P-Ewald and TIP4P-2005 models have similar distributions, with peaks of probability density 8.9 and 7.5. The SWM4-NDP model has a much higher peak with a maximum at 13.6. The TIP5P-Ewald model shows two peaks at  $g(\theta_1, \theta_2) \approx 10$  corresponding to hydrogen bond donation and acceptance. The distribution for  $\chi_1$  and  $\chi_2$  is shown in Figure 10 and again the TIP5P-Ewald model differs from the other, showing no peak around  $\chi_1 = 90^\circ$  and  $\chi_2 = 90^\circ$ . The SWM4-NDP model shows the highest peak at this point, with a local maximum in the probability density of  $g(\chi_1, \chi_2) \approx 2.2$ . The greater preference for planar geometry in the SWM4-NDP model is likely due to increased electrostatic interactions when the accepting



water polarises parallel to the donating O-H bond and places its hydrogen atoms at their maximum distance from the donating water. This corresponds to a co-planar arrangement of the donating O-H bond with the accepting water molecule.

In addition to the relative angular distributions, the orientational order parameters  $q$  and  $w$  were calculated for each water model. The value of  $q$  quantifies the overall tetrahedrality of oxygen atoms around each water molecule and the value of  $w$  quantifies the overall tetrahedrality of hydrogen atoms around each water molecule. The values of  $q$  for the TIP3P-Ewald, TIP4P-2005, TIP5P-Ewald and SWM4-NDP models are 0.59, 0.67, 0.64, and 0.63 respectively. These results are in line with previous calculations [43]. Interestingly, TIP4P-2005 is calculated to be the most tetrahedral water model using the orientational order parameter  $q$ , albeit only slightly. This is at odds with the results from considering the relative orientational angles. The values of  $w$  for the TIP3P-Ewald, TIP4P-2005, TIP5P-Ewald and SWM4-NDP models are 0.68, 0.74, 0.78 and 0.69 respectively. Thus TIP5P-Ewald is calculated to be the most tetrahedral water model using the orientational order parameter  $w$ , in line with results from considering the relative orientational angles. There is an important distinction between the tetrahedral geometry of the water network, quantified by the order parameter  $q$ , and the tetrahedral geometry of hydrogen bonding, quantified by the order parameter  $w$ . It is clear from the angular correlation functions and the order parameter  $w$  that the TIP5P model prefers tetrahedral hydrogen bonding geometries. However, whilst this is not surprising, it may not accurately reflect liquid water. Importantly, there is currently no experimental data that is able to assess this issue definitively and it must, for the present, remain as an observation on the models.

After considering the angular correlation functions, the enthalpies, entropies and free energies were calculated for the four models. The experimental values are derived from the experimental

data for  $H_{\text{excess}}$  and  $S_{\text{excess}}$  [46-47]. The results can be seen in Table II. All four models match the experimental excess enthalpy reasonably well, with the TIP3P-Ewald and TIP5P-Ewald models performing slightly less well, underestimating by 1.245 and 1.165 kcal/mol respectively. However, the TIP3P-Ewald and TIP5P-Ewald models provide the best match to the experimental excess entropy and both are within 1 cal/mol/K of the experimentally determined value of -14.054 cal/mol/K, as shown in Table II. The TIP4P-2005 model provides the best estimate of the excess free energy, within 1% of the experimental value, although it achieves this by complementary inaccuracies in the predictions for the excess enthalpy and entropy. The TIP3P-Ewald and TIP5P-Ewald models are the least accurate, 20% and 28% higher than the experimental value. Considering the sum squared differences between the predicted values of  $H_{\text{excess}}$  and  $TS_{\text{excess}}$  and the values from experiment, the SWM4-NDP model is the most accurate, though the TIP4P-2005 model is nearly as good.

The importance of each radial shell in contributing to the orientational entropy can be seen in Figure 11. For all four models, over 70% of the orientational entropy is derived from radial shells up to 3.4 Å and 95% is derived from radial shells up to 5.4 Å. This finding is in good agreement with previous work on the SPC water model [24]. In summary, all four models make reasonable predictions for the excess quantities, with the TIP5P-Ewald performing slightly better and the TIP4P-2005 model performing slightly worse. In addition to calculating the full 6-dimensional entropy, it is interesting to consider the lower order approximations to the entropy, as these require less data and thus shorter simulations to converge. Table III compares the lower order approximations with the true entropy. The first order approximations grossly underestimate the true entropy in all cases, as expected, but the second order approximation is reasonable and is as

close as the F7 approximation to the true entropy. The ability of the fourth order approximation to match the full entropy is most likely due to the flatness of the  $\varphi$  angular distribution.

Whilst analysis of the single and pair relative angular distributions and predictions of the excess enthalpy are revealing, an additional metric was also considered that compares the four models and quantifies the difference between their relative angular distributions. Table IV shows the RMSDs between all pairs of the full five dimensional PCFs across all radial shells for the four models, calculated using equation 11. This RMSD quantifies the similarity of the two five-dimensional relative angular distributions across all radial shells, with 0.0 representing identical distributions. Figure 12 shows the RMSD between the five-dimensional relative angular distributions for all pairs of models in each radial shell from 2.5 to 8.0 Å. The RMSD was calculated for each radial shell using equation 10. Only shells greater than 2.5 Å were considered, as closer shells do not contain sufficient data for accurate comparison. There are relatively few samples for the lowly populated bins between 0 and 2.5 Å and this provides a correspondingly high uncertainty in the predicted relative angular distribution at these distances. However, this is of little consequence in relation to the entropy predictions, as shells below 2.5 Å make only a very small contribution to the entropy due to their very small radial probability densities.

As well as comparing the relative angular distributions in each shell, the total RMSD between pairs of complete six-dimensional distributions was also calculated. For one water model, comparing the first 16ns of the production simulation with the second 16ns allows the convergence of the statistical average of the relative angular distributions to be assessed. This can be seen as the blue, red, green and orange lines for TIP3P-Ewald TIP4P-2005 TIP5P-Ewald, and SWM4-NDP in Figure 12. When considering only one water model, the RMSD is below

0.00045 in all models for all shells. However, when comparing the entire 40ns of the production simulations for different water models, the RMSD is as high as 0.0049, with particular differences below 3.2 Å, in the first solvation shell. This is more than a tenfold increase compared with considering one water model with itself. It is again notable that the TIP5P-Ewald model is clearly different from the other three models, which are more similar to one another. However, all four models have different relative angular distributions. This difference is not entirely clear from the one and two dimensional distributions plotted in Figures 6, 7, 8, 9, and 10. At distances above 6.5 Å, the RMSD between the models converges to 0.00045, representing the uncertainty inherent in the binned sampling. The differences between the models become even more pronounced when multiplied by the radial probability density at the first peak in the RDF between 2.6 Å and 3.1 Å, as for the calculation of the shell entropy in equation 9.

## Discussion

Computational modeling has the potential to offer significant insights into the thermodynamics of liquid water in the bulk and around biological or non-biological solutes. However, the results from this paper suggest that these insights will differ depending on the water model used and this fact may be masked by comparisons with the experimental RDFs. Further experimental data is needed to effectively validate these water models and this in turn is vital in improving the performance of computer simulations. A key result from this paper is that the five site water model TIP5P-Ewald shows notable differences from the other three models in the relative angular distribution, with a marked preference to act as a tetrahedral hydrogen bond donor and acceptor. Importantly, this fact is masked by the order parameter  $q$ , which quantifies the tetrahedral geometry of the water network. However, it is clear from the order parameter  $w$ , which quantifies the tetrahedral geometry of hydrogen bonding, and from the relative angular distributions, particularly the  $\theta_1/\theta_2$  distribution in Figure 9. The TIP3P-Ewald, TIP4P-2005, and SWM4-NDP models do not show a preference to act as a tetrahedral hydrogen bond acceptor and the SWM4-NDP model in particular shows a greater preference to act as a hydrogen bond acceptor with the OH bond from the donating water coplanar with the accepting water. It is important to note that a preference for tetrahedral hydrogen bonding geometry is not the same as a "tetrahedral structure" for water and the proclivity of these models to form chains or rings has not been assessed here [48]. Indeed, recent experimental results suggest that water may be comprised of a mixture of tetrahedral-like and distorted structures [49-50]. These distorted structures have been predicted to include a trigonal arrangement where oxygen acts as a single hydrogen bond acceptor and a trigonal bipyramidal arrangement where oxygen acts as a triple hydrogen bond acceptor [51]. Controversy still exists as to whether these should be considered as

stable states or as transition states between tetrahedral arrangements. Interestingly, the results of this study suggest that trigonal and trigonal bipyramidal arrangements would be more prevalent for some models than others, as the probability densities at  $\theta_1=180$  differ notably, as seen in Figure 8. This effect is even most pronounced for the SWM4-NDP model where  $g(\theta_1=180) \approx 2$  and least pronounced for the TIP5P-Ewald model where  $g(\theta_1=180) \approx 0.5$ . For the TIP3P-Ewald and TIP4P-2005 models,  $g(\theta_1=180) \approx 1$ .

The existence of experimental data on water structure is very useful to evaluate different water models. However, whilst neutron diffraction data is able to validate the predictions of the O-O, O-H, and H-H RDFs for these models, where the TIP5P-Ewald model performs best at recreating the experimental data, there is no experimental data to validate the exact predictions for the relative angular distributions. To comprehensively assess these water models would require new experimental techniques with high accuracy and the ability to consider a short timescale [52]. Failing this, it might be instructive to perform quantum mechanical simulations of liquid water at high levels of accuracy [19, 53-54]. This would allow the different water models to be compared with a more sophisticated model that includes orbital overlap, anisotropic polarisability and dispersion energies. Whilst the water models considered here represent four different classes, in the context of a forcefield it would be interesting to compare polarisable five site models [55], models with anisotropic polarisability [56] and models with two Drude particles per oxygen.

The similarities and differences between the four models is clear from analysis of the one- and two-dimensional angular distributions in Figures 6, 7, 8, 9, and 10. Whilst it is not possible to view the complete five-dimensional relative angular distributions, Figure 12 highlights that the four models have significantly different orientational correlations below  $3.5 \text{ \AA}$ . This is an important issue, as the relative angular distribution in bulk water has been employed as a

substitute for the relative angular distribution around a solute [28], and it is clear from Figure 11 that the region below 3.5 Å is the most important in determining the entropy [24]. There has been significant work on the thermodynamics of water molecules around solutes such as model hydrophobic enclosures [57], buried pockets [58] and the surface of proteins [59]. Furthermore, whilst the  $\phi$  angle is predicted to be unstructured in liquid water, it may be structured in the location of a solute and this should be considered when using data from a simulation of bulk water in calculations of water structure around a solute. This will also affect the accuracy of the lower-order approximations to the entropy. Whilst the fourth-order approximation to the entropy performs very well here, this may not be the case around a solute.

All four models perform reasonably well in predicting the experimental excess entropy of water. The TIP3P-Ewald and TIP5P-Ewald models are more accurate and are both within 5.0% of the experimental excess entropy. The TIP4P-2005 and SWM4-NDP models are within approximately 10% of the experimental excess entropy. Given the very large amounts of sampling required to calculate the full entropy, the lower-order approximations provide a reasonable estimate. Whilst the fourth order approximation is within 5% for all four models, the second order approximation is only 20% larger than the full entropy and requires significantly less data to calculate. The predictions are also good for the excess enthalpy. The differences from the experimental excess enthalpy for the TIP3P-Ewald, TIP4P-2005, TIP5P-Ewald and SWM4-NDP models are 13.4%, 3.9%, 12.5% and 0.9% respectively. The prediction of the experimental excess free energy is particularly good for the TIP4P-2005 model and is within 0.3%. However, this arises from complementary inaccuracies in the predictions for the excess enthalpy and entropy. Overall the TIP4P-2005 and SWM4-NDP models perform well, with the TIP3P-Ewald and TIP5P-Ewald models being less accurate. The reason for this appears to be that these models

were developed without considering the polarization correction from equation 2. If this term is ignored then both models perform significantly better. The choice of water model affects protein folding [30] and binding free energy calculations [31], making the accuracy of the thermodynamics for these models very important.

In summary, the ability of the TIP3P-Ewald, TIP4P-2005, TIP5P-Ewald, and SWM4-NDP water models to reproduce the O-O, O-H and H-H RDFs masks differences in the relative angular distributions of the four models. In particular, the TIP5P-Ewald water molecule acts as a tetrahedral hydrogen bond donor and acceptor in the liquid whereas the SWM4-NDP water molecule acts as a tetrahedral hydrogen bond donor and a co-planar acceptor. Thus, any insights into the hydrogen-bond structure in water taken from a simulation are likely to be dependent on the water model used and may not be true insights. The accuracy of these models should be explored in future work, either experimentally or using quantum mechanical methods. Designing an experiment to probe the relative angular distribution of liquid water would be an invaluable contribution to the field. This will lead to a greater understanding of water and determine whether the added computational expense of more complex water models is necessary. The impact of the water model on correlations around solute molecules must also be explored, to effectively model solvent entropy.



## **Acknowledgements**

Acknowledgements go to Mike Payne for careful reading of the manuscript, Peter Freddolino, Chris Baker, David Payne and Bracken King for helpful discussions, Stuart Rankin for technical help and the Cambridge HPCS for use of the CUDA-accelerated GPUs. All calculations were performed using the Darwin Supercomputer of the University of Cambridge High Performance Computing Service (<http://www.hpc.cam.ac.uk/>) and were funded by the EPSRC under grant EP/F032773/1. Thanks for financial support go to the MRC, Wellcome Trust and EPSRC.

## Figure Legends

### Figure 1 – The four water models used in this study.

The TIP3P-Ewald, TIP4P-2005, TIP5P-Ewald and SWM4-NDP water models. Oxygen atoms are displayed as red spheres, hydrogen atoms are displayed as light grey spheres, lone pairs are displayed as cyan spheres, massless charge sites are displayed as dark blue spheres and Drude particles are displayed as yellow spheres

### Figure 2 – The relative orientations of two water molecules.

The five angles describing the relative orientation of two water molecules. Oxygen atoms are displayed as red spheres and hydrogen atoms are displayed as light grey spheres. The  $\theta$  angles describe the angles between the dipole vectors of each water molecule and the intermolecular axis, the  $\chi$  angles describe the angles between the intermolecular axis and vectors normal to the plane of the water molecules and the  $\varphi$  angle describes the rotation around the intermolecular axis.

### Figure 3 – The O-O RDFs for the four water models.

The oxygen-oxygen radial distribution functions for TIP3P-Ewald (blue), TIP4P-2005 (red), TIP5P-Ewald (green) and SWM4-NDP (orange) compared to the experimental data from neutron diffraction (black).

**Figure 4 – The O-H RDFs for the four water models.**

The oxygen-hydrogen radial distribution functions for TIP3P-Ewald (blue), TIP4P-2005 (red), TIP5P-Ewald (green) and SWM4-NDP (orange) compared to the experimental data from neutron diffraction (black).

**Figure 5 – The H-H RDFs for the four water models.**

The hydrogen-hydrogen radial distribution functions for TIP3P-Ewald (blue), TIP4P-2005 (red), TIP5P-Ewald (green) and SWM4-NDP (orange) compared to the experimental data from neutron diffraction (black).

**Figure 6 – The  $\phi$  Angular Distribution Functions for the Four Models**

The  $\phi$  angular distribution functions from TIP3P-Ewald (blue), TIP4P-2005 (red), TIP5P-Ewald (green) and SWM4-NDP (orange) between 2.7 Å and 2.8 Å.

**Figure 7 – The  $\chi_1$  Angular Distribution Functions for the Four Models**

The  $\chi_1$  angular distribution functions from TIP3P-Ewald (blue), TIP4P-2005 (red), TIP5P-Ewald (green) and SWM4-NDP (orange) between 2.7 Å and 2.8 Å.

**Figure 8 – The  $\theta_1$  Angular Distribution Functions for the Four Models**

The  $\theta_1$  angular distribution functions from TIP3P-Ewald (blue), TIP4P-2005 (red), TIP5P-Ewald (green) and SWM4-NDP (orange) between 2.7 Å and 2.8 Å.

**Figure 9 - Surface plots of the  $\theta_1/\theta_2$  Angular Distribution Functions for the Four Models**

The  $\theta_1/\theta_2$  pair distribution functions for the four models between 2.7 Å and 2.8 Å. The probability density  $g(r)$  is represented by the level of the surface and coloured in bands of height of 1.0.

**Figure 10 – Surface plots of the  $\chi_1/\chi_2$  Angular Distribution Functions for the Four Models**

The  $\chi_1/\chi_2$  pair distribution functions for the four models between 2.7 Å and 2.8 Å. The probability density  $g(r)$  is represented by the level of the surface and coloured in bands of height of 1.0.

**Figure 11 – The orientational entropy of the water models for each radial shell from 2.5 to 8.0 Å.** The two particle entropy from the models TIP3P-Ewald (blue), TIP4P-2005 (red), TIP5P-Ewald (green) and SWM4-NDP (orange) in each shell between 2.5 and 8.0 Å.

**Figure 12 – The RMSD of the orientational correlation function for all pairs of water models.** The angular RMSD between the orientational PCF for each radial shell from pairs of the water models (a) TIP3P-Ewald, (b) TIP4P-2005, (c) TIP5P-Ewald and (d) SWM4-NDP. The RMSD was calculated between 2.5 and 8.0 Å using equation 10. The RMSDs in each shell for the four models are coloured blue, red, green, and orange respectively.



## Tables

**Table I – Parameters for the water models used in this study.**

<b>Model</b>	<b>TIP3P-Ewald</b>	<b>TIP4P-2005</b>	<b>TIP5P-Ewald</b>	<b>SWM4-NDP</b>
<b>qO (e)</b>	-0.83	0.0	0.0	0.0
<b>qH (e)</b>	+0.415	+0.5564	+0.241	+0.55733
<b>qLP (e)</b>	-	-1.1128	-0.241	-1.11466
<b>dO-H (Å)</b>	0.9572	0.9572	0.9572	0.9572
<b>dO-LP (Å)</b>	-		0.7	-
<b>∠ H-O-H (°)</b>	104.52	104.52	104.52	104.52
<b>∠ LP-O-LP (°)</b>	-	-	109.47	-
<b>ε (kcal/mol)</b>	-0.102	-0.185205	-0.178	-0.21094325
<b>R<sub>min</sub> (Å)</b>	1.7892	1.7729	1.737914	1.78692899

Table II- Simulation results for the water models used in this study.

Model	TIP3P- Ewald	TIP4P- 2005	TIP5P- Ewald	SWM4- NDP	Experimental
Runtime (ns/day)	11.8	6.9	4.5	2.3	NA
1 <sup>st</sup> O-O Peak (Å)	2.72	2.78	2.75	2.78	2.75
O-O g(r) Max	2.95	3.28	2.85	3.04	2.75
1 <sup>st</sup> O-O Trough (Å)	3.32	3.38	3.47	3.35	3.38
O-O g(r) Min	0.79	0.75	0.82	0.91	0.78
RMSD in O-O g(r)	0.10	0.09	0.03	0.10	NA
RMSD in O-H g(r)	0.08	0.10	0.07	0.10	NA
RMSD in H-H g(r)	0.03	0.02	0.04	0.02	NA
$S_{\text{trans}}$ (cal/mol/K)	-2.91	-3.48	-3.13	-3.29	-3.01 (b)
$S_{\text{orient}}$ (cal/mol/K)	-10.49	-12.07	-11.65	-12.44	-11.04 (c)
$S_{\text{excess}}$ (cal/mol/K)	-13.40	-15.54	-14.78	-15.73	-14.05 (a)
$H_{\text{excess}}$ (kcal/mol)	-9.27	-10.94	-9.35	-10.43	-10.51 (a)
$G_{\text{excess}}$ (kcal/mol)	-5.28	-6.31	-4.95	-5.74	-6.33 (a)

The RMSD in  $g(r)$  is the root of the sum of the squares of the difference between the experimental and the predicted  $g(r)$  at each radial distance. It represents the overall difference from experiment between the RDF for each model. For the O-H RDF, only the region above 1.29 Å was considered, as the RDFs below this point depend on molecular vibrations and these are all static models. For the same reason, only the region above 1.83 Å was considered for the H-H RDF. (a) The experimental data was taken from Wagner [46] and . (b) The experimental  $S_{\text{trans}}$  was calculated from the experimental RDF from Soper [20]. (c) The experimental  $S_{\text{orient}}$  was calculated from the experimental  $S_{\text{total}}$  and  $S_{\text{trans}}$ .



**Table III - The Orientational Entropies for the Four Models and the Lower Order Approximations.**

<b><math>\Delta S_{\text{excess}}</math> (cal/mol/K)</b>	<b>TIP3P- Ewald</b>	<b>TIP4P- 2005</b>	<b>TIP5P- Ewald</b>	<b>SWM4- NDP</b>	<b>Mean Pairs Per Bin</b>
<b>Full</b>	-10.491	-12.065	-11.648	-12.441	340
<b>4th Order Approximation</b>	-10.214	-11.546	-11.254	-12.054	6135
<b>3rd Order Approximation</b>	-9.088	-9.776	-9.503	-10.564	110430
<b>2nd Order Approximation</b>	-12.568	-14.600	-14.252	-14.418	1987805
<b>F7 Approximation</b>	-9.239	-9.850	-10.130	-10.049	1987805
<b>1st Order Approximation</b>	-3.245	-3.274	-3.106	-3.883	35780525

**Table IV - The RMSD Between the Orientational Pair Correlation Functions Across All The Radial Shells**

<b>RMSD</b>	<b>TIP3P-Ewald</b>	<b>TIP4P-2005</b>	<b>TIP5P-Ewald</b>	<b>SWM4-NDP</b>
<b>TIP3P-Ewald</b>	0.000249			
<b>TIP4P-2005</b>	0.000685	0.000249		
<b>TIP5P-Ewald</b>	0.000897	0.000887	0.000245	
<b>SWM4-NDP</b>	0.000540	0.000512	0.000987	0.000247

## References

- [1] J. D. Smith *et al.*, Science **306**, 851 (2004).
- [2] P. Wernet *et al.*, Science **304**, 995 (2004).
- [3] M. W. Mahoney, and W. L. Jorgensen, The Journal of Chemical Physics **112**, 8910 (2000).
- [4] H. Berendsen *et al.*, Intermolecular forces **11**, 331 (1981).
- [5] H. Berendsen, J. Grigera, and T. Straatsma, J Phys Chem-Us **91**, 6269 (1987).
- [6] W. L. Jorgensen, J. Am. Chem. Soc. **103**, 335 (1981).
- [7] D. J. Price, and C. L. Brooks III, The Journal of Chemical Physics **121**, 10096 (2004).
- [8] P. Ren, and J. W. Ponder, The Journal of Physical Chemistry B **107**, 5933 (2003).
- [9] W. L. Jorgensen, The Journal of Chemical Physics **77**, 4156 (1982).
- [10] H. W. Horn *et al.*, J. Chem. Phys. **120**, 9665 (2004).
- [11] J. L. F. Abascal, and C. Vega, J. Chem. Phys. **123**, 234505 (2005).
- [12] A. Rahman, and F. Stillinger, Journal of the American Chemical Society **95**, 7943 (1973).
- [13] S. W. Rick, J. Chem. Phys. **120**, 6085 (2004).
- [14] G. Lamoureux, A. D. MacKerell Jr, and B. Roux, The Journal of Chemical Physics **119**, 5185 (2003).
- [15] G. Lamoureux *et al.*, Chem. Phys. Lett. **418**, 245 (2006).
- [16] G. Lamoureux, and B. Roux, The Journal of chemical physics **119**, 3025 (2003).
- [17] H. Nada, J. Van der Eerden, and Y. Furukawa, Journal of crystal growth **266**, 297 (2004).
- [18] J. M. Sorenson *et al.*, The Journal of Chemical Physics **113**, 9149 (2000).
- [19] G. Hura *et al.*, Phys. Chem. Chem. Phys. **5**, 1981 (2003).
- [20] A. Soper, Chemical Physics **258**, 121 (2000).
- [21] A. Soper, and M. Phillips, Chemical Physics **107**, 47 (1986).

- [22] W. L. Jorgensen *et al.*, J. Chem. Phys. **79**, 926 (1983).
- [23] C. Vega *et al.*, Phys. Chem. Chem. Phys. **7**, 1450 (2005).
- [24] J. Zielkiewicz, J. Chem. Phys. **123**, 104501 (2005).
- [25] T. Lazaridis, and M. Karplus, J. Chem. Phys. **105**, 4294 (1996).
- [26] T. Lazaridis, J. Phys. Chem. B **102**, 3531 (1998).
- [27] Z. Li, and T. Lazaridis, Abstr. Pap. Am. Chem. Soc. **226**, U442 (2003).
- [28] Z. Li, and T. Lazaridis, J. Phys. Chem. B **110**, 1464 (2006).
- [29] D. J. Huggins, M. Marsh, and M. C. Payne, J. Chem. Theory Comput. **7**, 3514 (2011).
- [30] P. Florová *et al.*, J. Chem. Theory Comput. **6**, 3569 (2010).
- [31] E. Fadda, and R. J. Woods, J. Chem. Theory Comput. **7**, 3391 (2011).
- [32] J. C. Phillips *et al.*, J. Comput. Chem. **26**, 1781 (2005).
- [33] H. Grubmüller, 1996).
- [34] T. Darden, D. York, and L. Pedersen, J. Chem. Phys. **98**, 10089 (1993).
- [35] B. R. Brooks *et al.*, J. Comput. Chem. **4**, 187 (1983).
- [36] A. D. MacKerell *et al.*, J. Phys. Chem. B **102**, 3586 (1998).
- [37] A. D. Mackerell, M. Feig, and C. L. Brooks, J. Comput. Chem. **25**, 1400 (2004).
- [38] S. Miyamoto, and P. A. Kollman, J. Comput. Chem. **13**, 952 (1992).
- [39] R. H. Henchman, J. Chem. Phys. **126**, 064504 (2007).
- [40] H. Matsuda, Physical Review E **62**, 3096 (2000).
- [41] B. J. Killian, J. Y. Kravitz, and M. K. Gilson, The Journal of chemical physics **127**, 024107 (2007).
- [42] P. L. Chau, and A. Hardwick, Mol. Phys. **93**, 511 (1998).
- [43] J. R. Errington, and P. G. Debenedetti, Nature **409**, 318 (2001).
- [44] L. Wang *et al.*, J. Chem. Theory Comput. **5**, 1462 (2009).
- [45] F. H. Stillinger, and A. Rahman, The Journal of Chemical Physics **60**, 1545 (1974).

- [46] W. Wagner, and A. Pruss, J. Phys. Chem. Ref. Data **31**, 387 (1999).
- [47] P. J. Linstrom *et al.*, *NIST Chemistry WebBook* (National Institute of Standards and Technology Gaithersburg, MD, 1997).
- [48] T. Head-Gordon, and M. E. Johnson, **103**, 7973 (2006).
- [49] T. Tokushima *et al.*, Chem. Phys. Lett. **460**, 387 (2008).
- [50] C. Huang *et al.*, Proceedings of the National Academy of Sciences **106**, 15214 (2009).
- [51] R. H. Henchman, and S. J. Irudayam, The Journal of Physical Chemistry B **114**, 16792 (2010).
- [52] K. Ramasesha *et al.*, The Journal of chemical physics **135**, 054509 (2011).
- [53] J. VandeVondele *et al.*, The Journal of Chemical Physics **122**, 014515 (2005).
- [54] J. C. Grossman *et al.*, The Journal of Chemical Physics **120**, 300 (2004).
- [55] H. A. Stern *et al.*, The Journal of Chemical Physics **115**, 2237 (2001).
- [56] E. Harder *et al.*, J. Chem. Theory Comput. **2**, 1587 (2006).
- [57] R. Abel *et al.*, J. Chem. Theory Comput. **6**, 2924 (2010).
- [58] R. Abel *et al.*, J. Am. Chem. Soc. **130**, 2817 (2008).
- [59] D. J. Huggins *et al.*, PLoS Comput. Biol. **6**, (2010).