

Comparison of molecular dynamics simulations of water with neutron and X-ray scattering experiments

Veronika Reich^{1,*}, Arnab Majumdar^{1,**}, Martin Müller^{1,2,3,***}, and Sebastian Busch^{1,****}

¹German Engineering Materials Science Centre (GEMS) at Heinz Maier-Leibnitz Zentrum (MLZ), Helmholtz-Zentrum hereon GmbH, Lichtenbergstr. 1, 85748 Garching, Germany

²Institute of Materials Physics, Helmholtz-Zentrum hereon GmbH, Max-Planck-Str. 1, 21502 Geesthacht, Germany

³Institut für Experimentelle und Angewandte Physik (IEAP), Christian-Albrechts-Universität zu Kiel, Leibnizstr. 19, 24098 Kiel

Abstract. The atomistic structure and dynamics obtained from molecular dynamics (MD) simulations with the example of TIP3P (rigid and flexible) and TIP4P/2005 (rigid) water is compared to neutron and X-ray scattering data at ambient conditions. Neutron and X-ray diffractograms are calculated from the simulations for four isotopic substitutions as well as the incoherent intermediate scattering function for neutrons. The resulting curves are compared to each other and to published experimental data. Differences between simulated and measured intermediate scattering functions are quantified by fitting an analytic model to the computed values. The sensitivity of the scattering curves to the parameters of the MD simulations is demonstrated on the example of two parameters, bond length and angle.

1 Introduction

The structure and dynamics of matter can be examined on an atomic level with various methods. Neutron and X-ray scattering data [1] are generally interpreted by building analytical models that are then refined against the experimental data. With molecular dynamics (MD), simulations reflect the structure and dynamics on the same time and length scales. The simulations depend on a series of input parameters for the interaction potentials, which we propose optimizing with the measured scattering data. In the present contribution, we take the first step toward this goal by using liquid water as an experiment subject.

Water is an ideal candidate because it has been investigated in all possible states; from gaseous over liquid, supercooled, and glassy to the many different phases of ice. Its structure and dynamics can be – and often are – measured on the molecular level with neutron and X-ray scattering [2]. Water, in turn, is also often used to calibrate these instruments [3, 4]. Diffraction measurements do not analyze the energy transfer between probe and sample; instead, the scattered intensity as a function of momentum transfer contains information about the sample structure. Samples can be probed by neutrons, which scatter at the nuclei, or X-rays, which scatter at the electron shells [5, 6]. Spectroscopic measurements analyze not only the momentum but also the energy transfer between probe and sample, which makes it possible to observe the dynamics. For decades, quasielastic neutron scattering (QENS) has been

the method of choice for studying the molecular motions of water [7–10]. Over the years, MD simulations have also been frequently used to study water on the molecular level [11–19], and the rapid advances in computing power have unlocked approaches that were hitherto not feasible.

Although MD simulations have been successfully used to evaluate the nanoscopic (picoseconds and Ångströms) structure and dynamics probed by scattering data [8, 20–24], the interaction potentials used by the simulations are generally optimized to reproduce macroscopic, thermodynamic properties and densities [12, 13], sometimes with the aim of being able to cover a large portion of the phase diagram [13]. A model that is directly optimized to reproduce scattering experiments will likely yield a better description of the nanoscopic properties, albeit probably at the cost of generality. This principle is already employed for the evaluation of diffraction [25–27] and QENS [28, 29] experiments. So far, however, simulations have been refined either against diffraction or QENS measurements; we propose a simultaneous fit of both.

The current work uses two prominent water models: TIP3P [12, 30] and TIP4P/2005 [13]. Both are point models of a water molecule and simplify reality considerably but still depict the behavior of water molecules fairly accurately [15, 31, 32]. Both models can be rigid or flexible in consideration of bond length and angle. We will use TIP3P rigid as a base and compare it to TIP3P flexible as well as the TIP4P/2005 rigid model. All models will be explained in more detail in section 2.2. The experimental data, simulations, and analysis tools are introduced in section 2 with a particular focus on the calculation of scattering patterns described in section 2.5. Different varieties of the simulations are then compared to each other in sec-

*e-mail: veronika.reich@hereon.de

**e-mail: arnab.majumdar@hereon.de

***e-mail: martin.mueller@hereon.de

****e-mail: sebastian.busch@hereon.de

tion 3 as well as to experimental data. Since no perfect agreement between experimental data and simulation was achieved, a parameter optimization scheme is suggested in section 3.4. It demonstrates the sensibility of the scattering curves to MD parameter changes, thereby laying the groundwork for a parameter optimization scheme against scattering data to be presented in future work.

2 Materials and Methods

In this work we compare simulations and experiments as directly as possible: We use diffraction data in reciprocal space and spectroscopic data as intermediate scattering functions rather than derived quantities like the oxygen-oxygen pair correlation function or diffusion coefficient. This allows the scattering patterns to be calculated far more easily from the simulations as opposed to using a Fourier transform to convert the measured data into real space [33, 34].

The complete data analysis workflow has been made available in an accompanying data publication [35].

2.1 Experimental data

Neutron and X-ray diffraction data were taken from a meta analysis by Soper [36] at 298 K. For neutron scattering, diffractograms of H₂O, “null water” (an H₂O/D₂O mixture with an average coherent hydrogen scattering length of zero), HDO, and D₂O are provided. By averaging the coherent hydrogen scattering lengths to zero in null water, only oxygen-oxygen correlations are visible in the resulting diffraction data. For X-ray scattering, “normalization II” as defined by Soper [36] was employed:

$$D(Q) = \frac{I_x(Q)}{\sum_{\alpha} c_{\alpha} f_{\alpha}^2(Q)} - 1.0 \quad (1)$$

Quasielastic incoherent neutron scattering curves were taken from Teixeira et al. [37] at 293 K. The data were measured in energy space up to a maximal energy transfer of $\hbar\omega = 3$ meV. This corresponds very roughly to a shortest probed time of $t_{\min} \approx 1$ ps. All processes that occurred in the MD simulation at shorter correlation times cannot be expected to be visible in this dataset. Rather than taking the data directly, the fit function used to evaluate the data was also utilized to reconstruct the intermediate scattering function

$$I(Q, t) = DWF(Q) \cdot R(Q, t) \cdot T(Q, t) \quad (2)$$

The interpretation of this formula as a decoupling of fast vibrations, rotational motions, and translational motions is, for most of the present contribution, irrelevant – any fit function that describes the measured data with the same accuracy would suffice for the comparison with the simulation. The individual parts of the formula are the Debye-Waller factor due to fast vibrations

$$DWF(Q) = \exp\left[\frac{-Q^2\langle u^2 \rangle}{3}\right], \quad (3)$$

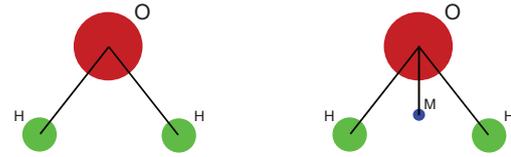


Figure 1. Sketches of TIP3P (left) and TIP4P/2005 (right).

the rotational diffusion

$$R(Q, t) = j_0^2(Qa) + 3j_1^2(Qa) \exp\left[\frac{-t}{3\tau_1}\right] + 5j_2^2(Qa) \exp\left[\frac{-t}{\tau_1}\right] \quad (4)$$

with

$$\tau_1 = \tau_1^0 \cdot \exp\left[\frac{E_A}{k_B T}\right] \quad (5)$$

where E_A is 1.85 kcal/mol, and the translational diffusion

$$T(Q, t) = \exp[-\Gamma(Q)t] \quad (6)$$

with

$$\Gamma(Q) = \frac{DQ^2}{1 + DQ^2\tau_0} \quad (7)$$

where

$$D = \frac{L^2}{6\tau_0}. \quad (8)$$

Numerical values for these quantities are given in Table 2.

2.2 Water models

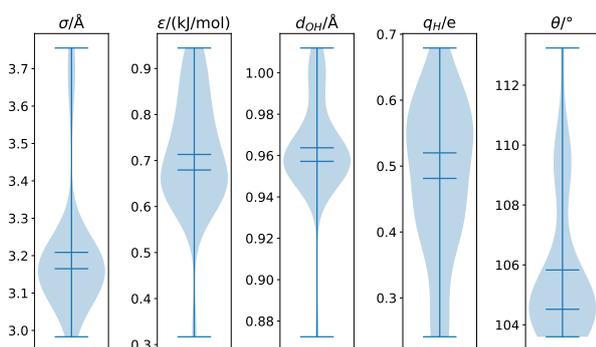
Most water models reduce the complexity of the nearly spherical and polarizable electron cloud of a water molecule [5] to the summation of a small number of interaction sites. How many of these sites are included and which interaction functions are employed varies between models. In this work, two widely used water models, the three-site TIP3P and the four-site TIP4P/2005, were employed. Their most important parameter values are summarized in Table 1 [38].

In the TIP3P model [12], the interaction sites are collocated with the positions of the nuclei, (see Figure 1). The O–H bond length and the H–O–H bond angle can be either rigid or flexible. While TIP3P reproduces some quantities quite well and is frequently used especially in biomolecular simulations, it is known to have deficiencies [30, 31, 39].

The TIP4P/2005 water model [13] is thought to be superior in many aspects [13, 14, 40–43]. Its four interaction sites are located at the three nuclear positions as well as a massless site M, which is located along the bisector of the H–O–H bond angle at 0.1546 Å from the oxygen atom. The negative partial charge is positioned at M rather than the oxygen nucleus, from where the van der Waals interaction is calculated. The TIP4P/2005 water model can be either rigid [13] or flexible [14]. In this contribution, only the rigid model was used.

Table 1. Table of the force field parameters used in the MD simulations. Hydrogen atoms did not participate in Lennard-Jones interactions. The flexible model used harmonic intramolecular potentials.

Parameter	TIP3P rigid	TIP3P flexible	TIP4P/2005 rigid
$r(\text{OH}) / \text{\AA}$	0.9572	0.9572	0.9572
$K r(\text{OH}) / (\text{kcal/mol}/\text{\AA}^2)$	—	450	—
$\angle \text{HOH} / \text{deg}$	104.52	104.52	104.52
$K \angle \text{HOH} / (\text{kcal/mol}/\text{rad}^2)$	—	55	—
$q(\text{H}) / e$	0.415	0.415	0.5564
$\sigma_{\text{O}} / \text{\AA}$	3.188	3.188	3.1589
$\epsilon_{\text{O}} / (\text{kcal/mol})$	0.102	0.102	0.1852

**Figure 2.** Violin plot of the distribution of water model parameters in the literature for a subset of parameters. The shaded area presents an approximated probability distribution to find a model with such a parameter value while the horizontal lines give minimum/maximum, mean, and median of the values used in this analysis.

Water models can either neglect or consider intramolecular vibrations, resulting in either a flexible or rigid model. Which version to use depends on computational cost (with flexible models being more costly) [14, 44], accuracy and artifacts [44], dielectric properties [45], or quantum treatment [46]. As will be illustrated in section 3.2.1, this choice changes the structure and dynamics of the models.

While water models may differ in the number and mathematical function of the interactions, some models with the same principal interactions may differ only in the functions' employed parameters. Figure 2 shows a distribution of a subset of parameter values for many models existing in the literature [11–14, 47–70]. The models used for this overview include differences in parameter values and functional forms. This ambiguity of parameter choice is due to the fact that water models can only reflect a subset of real water properties – and different models reproduce different properties [13].

2.3 MD simulations

The MD simulations were carried out with the program LAMMPS (Large-scale Atomic/Molecular Massively Parallel Simulator) [71]. The initial input files were taken from the literature [12, 30, 38, 40, 72], including the corresponding cutoff distances.

The cubic simulation box contained 8,000 water molecules (24,000 atoms). These molecules were placed in a large starting box that was subsequently compressed to an edge length of 62.086 \AA , resulting in a density of 1.0000 g/cm^3 or 0.1 atoms/ \AA^3 . A series of equilibration steps followed to ensure thermodynamical equilibrium of the simulation. The total setup runs cover a time span of 110 ps.

Once the simulation was stable, a production run with a length of 2 ns was performed; the coordinates were saved every picosecond. Subsequently, a further production run was performed covering 20 ps, in which the coordinates were saved every femtosecond. Both production runs were performed in the *NVT* ensemble; pressure coupling was consciously avoided to prevent seeing atom movements caused by the box resizing algorithm in the intermediate scattering function. All MD simulations in this work were run at 296 K. The pressure in the simulation box was 133 ± 83 bar (TIP3P rigid), -714 ± 170 bar (TIP3P flexible), and 112 ± 91 bar (TIP4P/2005 rigid).

The simulations were run with periodic boundary conditions. Two different versions of the coordinates were saved: 1) A *wrapped* set, where all coordinates were wrapped into the central image of the box. This box always had the correct density and was used for structural calculations. 2) An *unwrapped* set was saved where atoms are not kept in the central image of the periodic box but can travel to neighboring images. Since this version avoided unphysical movements of the atoms across the box length, it was used for the determination of all dynamic quantities.

2.4 Real space analysis

Radial distribution functions (RDFs, $g(r)$) and root mean square displacements (RMSDs) were calculated with the program Visual Molecular Dynamics (VMD) [73]. The RDFs were calculated taking the periodic boundary conditions of the wrapped simulation box into account; the RMSDs were averaged over all atoms and calculated from the unwrapped trajectory. Rotations and translations of the whole simulation box were removed by the algorithm before computing the RMSD, although we did not expect these to occur due to the careful setup of the MD simulations.

2.5 Calculation of scattering patterns

There are several open-source tools available to calculate scattering curves from MD simulations, such as nMOL-DYN [74], LiquidLib [75], or Sassena [76, 77]. We chose Sassena for this work because of its flexibility and scalability.

Sassena can calculate neutron and X-ray single crystal Bragg diffraction, orientationally averaged powder diffraction, coherent and incoherent intermediate scattering functions, or elastic coherent and incoherent structure factors from MD simulations. Parallelization was used as a guiding design principle and was implemented in the used version via MPI and thread parallelism.

In order to increase the calculation speed, we have introduced vectorization to make full use of modern CPUs' wide registers as well as OpenMP shared memory parallelization. Vectorization improves the program speed of each core, thereby benefiting on local computers as well as high performance systems. Hybrid parallelization combining MPI and OpenMP retained the scalability of the code on complex multi-node systems in large clusters.

2.5.1 X-ray and neutron diffraction

X-ray and neutron diffraction data can be reduced to one term due to interference scattering (i. e., without single atom scattering)

$$D(Q) \sim \left\langle \sum_j^N \sum_{k>j}^N b_j^c b_k^c \exp \left[i \vec{Q} \cdot (\vec{R}_j(t_0) - \vec{R}_k(t_0)) \right] \right\rangle_{t_0, \Omega} \quad (9)$$

where Q is the modulus of the scattering vector, N is the number of atoms in the sample, b_l^c is the (Q -independent) coherent neutron scattering length [78] or (Q -dependent) atomic X-ray scattering factor [79] of scatterer l , i is the imaginary number, $\vec{R}_l(t)$ is the position of scatterer l at time t , and t_0 is a starting time. $\langle \dots \rangle_{t_0, \Omega}$ denotes an average over all possible starting times t_0 and selected orientations Ω of the Q -vectors.

This quantity is obtained from the Sassena result $F(Q)$ in the neutron case via

$$D(Q) = \left[\frac{F(Q)}{N} - \sum_i c_i b_i^2 \right] / 100 \quad (10)$$

since the single atom scattering is included in the Sassena result and has to be subtracted; c_i are the concentrations of nuclei with coherent neutron scattering length b_i . The factor 100 is due to a unit conversion from fm^2 to barn. In the case of X-ray scattering, the single atom scattering is also subtracted and the interference scattering is divided by the square of the mean atomic scattering factor [36]:

$$D(Q) = \frac{F(Q)}{N \cdot \sum_i c_i f_i(Q)^2} - 1 \quad (11)$$

where in the present case the atomic scattering factors were calculated from the analytical approximation as a sum of Gaussians [79].

Sassena was used to compute the diffraction patterns from the MD simulations up to $Q = 30 \text{ \AA}^{-1}$ with a spacing between points of 0.05 \AA^{-1} . A Monte Carlo scheme was used to average over the different orientations of Q -vectors; the X-ray diffractograms were averaged over 1,000 random orientations and the neutron diffractograms over 100. We observe that the results were not particularly sensitive to the number of Q -orientations in these disordered systems, in contrast to the situation in crystalline structures.

Different isotopical substitutions were evaluated. $b_O = 5.803 \text{ fm}$ was kept the same through all evaluations but the original value of $b_H = -3.7390 \text{ fm}$ was changed to $b_{null} = 0 \text{ fm}$, $b_{HD} = 1.466 \text{ fm}$, and $b_D = 6.671 \text{ fm}$ for the evaluation of each isotope as described in section 2.1.

Sassena offers the option to subtract the average scattering length from all atoms in the system. While this removes the small-angle scattering caused by the finite simulation box, it distorts the wide-angle diffractogram and was therefore not used for this contribution. Due to the relatively large size of the simulation box, the corresponding small-angle scattering signal was only present at very small Q .

2.5.2 Incoherent quasielastic neutron scattering

The incoherent intermediate scattering function can be calculated as

$$I(Q, t) \sim \left\langle \sum_j^N b_j^i b_j^i \exp \left[i \vec{Q} \cdot (\vec{R}_j(t_0 + t) - \vec{R}_j(t_0)) \right] \right\rangle_{t_0, \Omega} \quad (12)$$

where Q is the magnitude of the scattering vector, t the time, and b_j^i the incoherent scattering length of scatterer j [78]. $\vec{R}_j(t)$ is the position of scatterer j at time t .

Sassena was used to compute the intermediate scattering function at Q -values from 0.1 to 5.0 \AA^{-1} with a step size of 0.1 \AA^{-1} , each averaging over 10 orientations of the Q -vector. The real part of the result was used and normalized to the value at $I(Q, t = 0)$.

2.6 Analytical evaluation of the computed intermediate scattering functions

Albeit not the main point of this contribution, the fit formula commonly used for water [37] given in equation 2 was also used to evaluate the computed intermediate scattering functions. The values of $\sqrt{u^2}$, τ_1^0 , τ_0 , and L were fitted by the method of least squares, the bond length a was changed from the value of 0.98 \AA used in the literature [37] to the bond length used in the simulations, 0.9572 \AA . The fit was carried out in the time range from 1 ps to 2 ns and for Q -values from 0.1 \AA^{-1} to 5.0 \AA^{-1} in steps of 0.1 \AA^{-1} .

The error bars of the fitted parameters were approximated by assuming the Hessian matrix H to be $J^T * J$, where J is the Jacobian matrix returned from the fit routine. H^{-1} was multiplied with the reduced χ^2 and diagonalized. The square roots of the diagonal elements are expressed as error bars in Table 2. It should be noted that

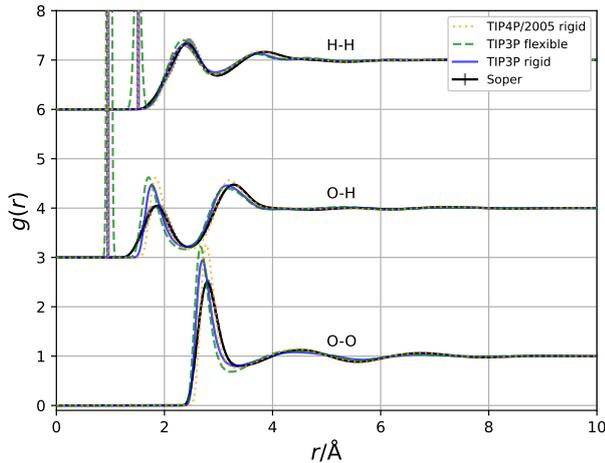


Figure 3. RDFs $g(r)$ of TIP3P rigid, TIP3P flexible, and TIP4P/2005 rigid for O–O, O–H, and H–H correlations. O–H was shifted by an offset of 3 on the y -axis and H–H was shifted by an offset of 6 on the y -axis. The computed RDFs are compared to literature values from [36]

the input data – the evaluated fit function found in the literature [37] – do not have error bars.

3 Results

3.1 Real space analysis

The structure of liquid water in real space can be characterized by the RDFs $g_{OO}(r)$, $g_{OH}(r)$, and $g_{HH}(r)$. The functions computed from the simulations are compared to the RDFs obtained for water at 298 K by Soper [36] in Figure 3. It should be noted that these $g(r)$ should not be referred to as *experimental data* since they are already the result of a computer simulation that reproduced the actual experimental data, which are taken in reciprocal space.

The first sharp peak in the O–H and H–H RDFs is caused by intramolecular correlations and is therefore much sharper than the following intermolecular features. The rigid models exhibit very sharp peaks while the distribution is more smeared out in the flexible model. By and large, there is a good agreement of the intermolecular $g(r)$ with the literature values. The flexible TIP3P model produces slightly more pronounced features in the curves at minimally shorter distances than the rigid model. TIP4P/2005 does not reproduce the literature values substantially better than TIP3P.

The dynamics of a diffusing liquid was extracted from the simulations by the RMSD, averaged over all atoms. The resulting curves are shown in Figure 4 together with the experimental long-time behavior, given by the relation

$$\text{RMSD}(t) = \sqrt{\left\langle \frac{1}{N} \sum_{j=1}^N \left(\vec{R}_j(t_0 + t) - \vec{R}_j(t_0) \right)^2 \right\rangle_{t_0}} \quad (13)$$

where N is the number of atoms and $\vec{r}_k(t)$ the position of atom k at time t . $\langle \dots \rangle_{t_0}$ denotes an average over all starting

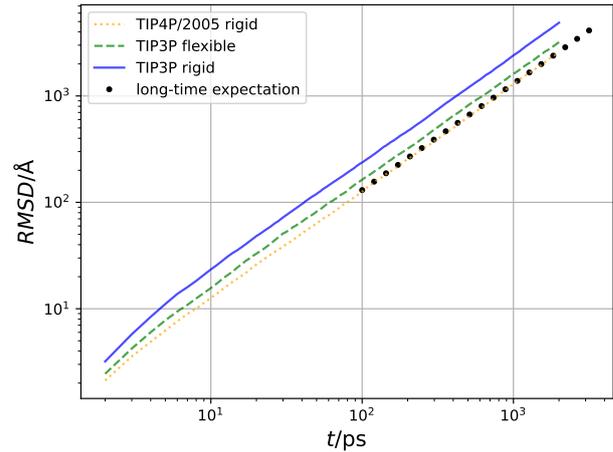


Figure 4. RMSD averaged over all atoms compared to the long-time expectation given by the macroscopic diffusion coefficient of water.

times t_0 . The macroscopic water diffusion coefficient for 296 K was taken to be $2.1745 \cdot 10^{-3} \text{ mm}^2/\text{s}$ [80].

None of the considered models shows a slower long-range diffusive motion than expected from the macroscopic value. The TIP4P/2005 rigid model reproduces the experimentally determined macroscopic diffusion coefficient nearly perfectly. In the case of TIP3P, the diffusion slows when flexibility is introduced to the model – and is in either case further away from the macroscopic value than TIP4P/2005.

3.2 Reciprocal space analysis

Figure 5 shows the computed diffraction patterns and intermediate scattering functions of the three simulated water models together with published experimental neutron and X-ray diffraction data [36] and the published intermediate scattering functions [37].

It should be noted that all curves of the given models were computed from the same MD simulation of H_2O . For diffraction, it is possible in Sasena to vary the probe between X-rays and neutrons and even assign the hydrogen atoms different scattering lengths in the case of neutron scattering, reproducing a measurement series with isotopic substitution. Rather than assigning individual hydrogen atoms random scattering lengths of H and D, all hydrogen atoms were given the same averaged scattering length.

3.2.1 TIP3P rigid versus TIP3P flexible

The O–O correlations that dominate the “null water” neutron diffraction pattern in Figure 5 are hardly affected by the choice between rigid and flexible water molecules. The O–H correlations, particularly visible in the H_2O and D_2O neutron diffraction patterns, show small differences at low Q -values (up to $\sim 6 \text{ \AA}^{-1}$) and a slight shift to smaller Q -values of the features at higher Q when switching from rigid to flexible. The oscillation amplitude of the high- Q features decreases only very slightly when changing from

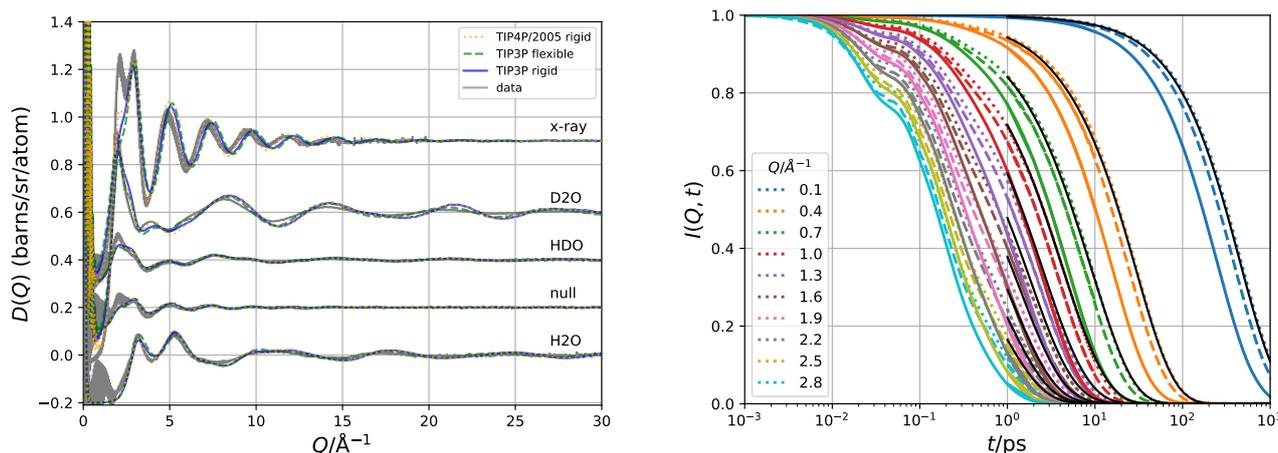


Figure 5. Left: diffractograms for X-rays and neutrons with four different isotopic substitutions for each of the three water models as well as diffraction data from [36] with different offsets for better visibility respectively. Right: intermediate scattering functions, black solid lines refer to intermediate scattering functions from [37], solid colored lines refer to TIP3P rigid, dashed lines to TIP3P flexible, and dotted lines to TIP4P/2005 water model.

a rigid to a flexible model and is in either case much more pronounced than in the data.

The long-range diffusion of the water hydrogens (and therefore the water molecules) is probed by the intermediate scattering functions with small Q and long times, while localized motions dominate the signal at higher Q and short times. The motions slow upon flexibilization – i. e., $I(Q, t)$ decays at longer times compared to the rigid model. This effect is visible across the whole time scale.

3.2.2 TIP3P rigid versus TIP4P/2005 rigid

The diffraction patterns in Figure 5 show a marked difference at the lowest Q values where TIP4P/2005 has a much more pronounced peak around 2 \AA^{-1} than TIP3P. While the neutron H_2O and D_2O diffractograms show very similar O–H correlations between the two models at high Q (i. e., predominantly short distances), the O–O correlations probed by X-ray diffractogram have their oscillations shifted to lower Q in TIP4P/2005.

The intermediate scattering functions of both rigid water models differ strongly. This variation was already indicated in the display of the RMSD curves shown in Figure 4.

3.2.3 Simulations versus scattering experiments

In Figure 5 one can clearly see a discrepancy in all diffractograms between the simulation and experiment at low Q . Particularly striking is the difference for X-ray data, where the first peak in the diffraction pattern is missing altogether for the simulation. This first peak is also not well reproduced for the neutron measurements. It is obvious that the computed diffractograms of TIP4P/2005 align the best with the low Q region of the data and at least the computed curves of the X-ray data have a small feature in this range. At higher Q the oscillations of all three models have roughly the correct amplitude for the X-ray and “null

water” measurements, where O–O correlations dominate, with TIP4P/2005 having the best alignment. On the other hand all are shifted to larger Q than the experimental data, indicating packing that is too close. The H_2O and D_2O curves of both rigid models in particular exhibit oscillations that are too large at high Q , indicating an O–H distance that is too defined.

For small Q , the intermediate scattering function of the flexible TIP3P simulation shows better agreement with the experiment than the rigid one. The intermediate scattering function, dominated by long-range diffusion at low Q , decays faster than the literature values and the same holds for local motions probed at higher Q . TIP4P/2005 seems however to have managed to match the experimental data in the long-range regime better than TIP3P. Still, also TIP4P/2005 deviates significantly from the data at higher Q .

3.3 Interpretation of the simulated intermediate scattering function

The fit formula given in section 2.5.2 was used for a global fit of the computed intermediate scattering functions over the whole Q -range. As shown in Figure 6, the formula is able to reproduce the simulated intermediate scattering functions very well in this time and Q -range. Table 2 compares the original parameters from [37] to the current fit results.

The first component of the fit formula is the Debye-Waller factor where the simulations exhibited a vibrational amplitude $\sqrt{\langle u^2 \rangle}$ that was about 1.3 times as large as the literature value [37]. The rotational motion includes the radius of rotation as a parameter, which was not fit but set to 0.98 \AA in the literature and 0.9572 \AA here, in accordance with the water models’ O–H bond length. The relaxation time of the rotational motion τ_1 was two to three times larger than in the literature, indicating that the rotational motions of the simulation were too slow. The residence

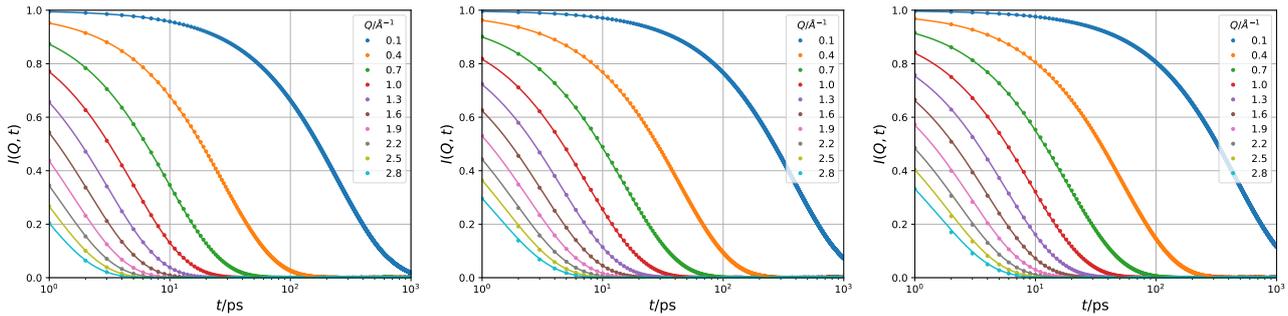


Figure 6. Fit of $I(Q, t)$ computed from the TIP3P rigid (left), TIP3P flexible (middle), and TIP4P/2005 rigid (right) simulations, shown as dots, with an analytic expression from the literature [37]. The best global fit of this formula is shown as drawn-out lines.

Table 2. Parameters from fits to experimental data published in the literature [37] compared to the values obtained from fits to the intermediate scattering patterns computed in this work for both TIP3P and the TIP4P/2005 rigid water models.

Parameter	lit. [37]	TIP3P rigid	TIP3P flexible	TIP4P/2005
$\sqrt{\langle u^2 \rangle} / \text{\AA}$	0.48	0.594 ± 0.001	0.617 ± 0.000	0.618 ± 0.001
τ_1^0 / ps	0.0485	0.124 ± 0.001	0.095 ± 0.000	0.150 ± 0.001
τ_0 / ps	1.25	0.277 ± 0.001	0.623 ± 0.001	0.581 ± 0.002
$L / \text{\AA}$	1.29	0.820 ± 0.001	0.991 ± 0.001	0.864 ± 0.002

time of the jump diffusion was between one half and one fifth in the simulations compared to the literature, meaning that the water molecules are “stuck” for a much shorter time in the simulation than the analysis of the experiment suggests. If they moved, however, their “jump length” L was about 1.5 times smaller in the simulation than in the literature.

3.4 Parameter variation

As a last step, the input parameters of the TIP3P rigid water model were varied manually to investigate their impact on the structure and the dynamics. The result is shown in Figure 7 for two of these parameter changes, O–H bond length and H–O–H bond angle. The purpose of this procedure is to evaluate if and how sensitive the scattering curves react to parameter changes. It will be shown that the scattering curves react with different sensitivities and in different manners to these changes, which we will use in future works to create an optimization algorithm for the simulation input parameters. The success of a coupled change of these two parameters was recently shown in comparison to neutron spectroscopy [10, 81].

The first parameter variation concerns the O–H bond length. In the literature there are many different values for the bond length of water, which sometimes depend on the temperature or even the experimental technique used to measure it. We employed the bond length of 0.9572 Å originally used in TIP3P [12, 82] and 1.00 Å [83] (which is the same as in the SPC water model [47]).

The “null water” and X-ray diffractograms sensitive to O–O correlations were not very sensitive to this parameter variation (see Figure 7). A slight shift of the high- Q oscillations to lower Q as the bond length was increased shows that the molecules grew slightly. The change of the O–H bond length had a slightly larger effect on the H₂O

and D₂O diffractograms, where a change of bond length resulted in a shift of the oscillations at large Q .

In contrast to the rather small changes of the diffraction patterns, the influence of the O–H bond length on the intermediate scattering function shown in Figure 7 is drastic. Prolonging the bond length led to a decrease of dynamics in a time range from a few tens of femtoseconds to nanoseconds. The final decay of the correlations at long times, probably corresponding to the long-range diffusion of the molecules, is more affected than the processes at sub-picosecond times. Comparing the whole intermediate scattering function to the data clearly yields much more information than looking only at the macroscopic diffusion coefficient.

The second example of a parameter change is the bond angle. The original TIP3P value was 104.52° [12, 82], and the comparison simulation was run with an angle of 109.50° (the angle of a tetrahedron), which is the same value used in the SPC water model [47]. The corresponding plots are also shown in Figure 7.

The diffractograms were very insensitive to this parameter change. Only the scattering patterns corresponding to the D₂O sample are slightly affected, since the H–H correlations weigh more than the other isotope combinations. However, it is questionable if this difference would be noticeable when comparing the computed curves to real data.

The influence on the intermediate scattering functions is much smaller than when varying the bond length. An increase of the bond angle leads to an increase in the dynamics. This acceleration primarily affects the long-time processes and leaves fast motions with correlation times below ~1 ps basically unaltered.

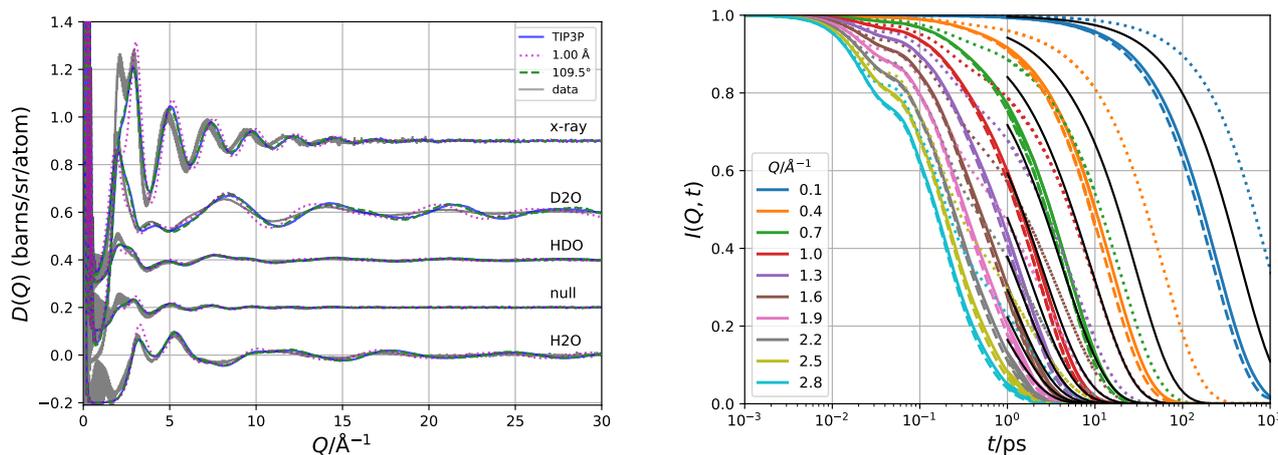


Figure 7. Parameter change in TIP3P rigid: 0.9572 Å O–H bond length versus 1.00 Å O–H bond length, and 104.25° H–O–H angle versus 109.50° H–O–H angle. Left: diffractograms for X-rays and neutrons with four different isotopic substitutions for each of the parameter changes as well as data [36] with different offsets for better visibility respectively. Right: intermediate scattering functions, black solid lines refer to literature data [37], solid colored lines refer to the standard TIP3P with a 104.25° angle and a 0.9572 Å bond length, dotted lines to a 1.00 Å bond length, and dashed lines refer to a 109.50° angle.

4 Discussion

Although no model is perfect, good models simplify complex situations while reproducing relevant observables. Which observables are relevant depends on the purpose of the model. This results in different possible parameter values for MD simulations, depending on if the model’s creator placed more emphasis on reproducing for example the phase diagram or the nanoscopic structure and dynamics.

The experimental data used in this contribution were taken completely from the literature [36, 37]. The QENS experiments were performed at 293 K, the simulations were run at 296 K, and the diffraction data were taken at 298 K. Due to the slightly different temperatures, small variations between the simulation and experiment can be expected; however, because the density of the simulation was constrained and the atoms in the simulation moved faster than in the experiment, we do not expect the effects described in this contribution to be merely due to the temperature. The QENS data were used in the form of the intermediate scattering function reconstructed from the fit formula utilized in the literature [37]. This procedure circumvents problems associated with using the measured data of the scattering function $S(Q, \omega)$ directly, such as cutoff effects. However, we will attempt in the future to perform the comparison between simulation and experiment closer to the actually measured data.

Some parameters in simulations are mostly independent of the the water model chosen, such as cutoff distances or box sizes. For most of these parameters, we chose the consensus values published in the literature [12, 30, 38, 40, 72]. The size of the simulated box was considerably bigger than in many previous studies [8, 12, 13, 15, 31, 33], but simulations of even larger boxes have been published [19]. The use of a fairly large simulation box is important since too small of a box size is

expected to influence both the structure and dynamics in the resulting scattering curves. While the structure in simulations is not expected to be influenced strongly by periodic boundaries, the calculation of $D(Q)$ is: peaks will broaden and small-angle scattering will become visible. For a disordered system like a liquid, the peak broadening effect is assumed to be negligible. The small-angle scattering, caused by the cubic simulation box effectively floating in vacuum, is confined to very small Q -values for a box of this size. For the incoherent intermediate scattering function, problems will arise in the simulation itself [84] while computation of the scattering function from the trajectories should not be hampered. A production run with an NPT ensemble is planned for future work to evaluate the impact of the pressure coupling on the intermediate scattering functions. Barostats regulate the pressure in the simulation box by resizing the box and scaling all atom coordinates accordingly. This change in coordinates is probably negligible but an NVT ensemble was used for this contribution to exclude any possible influences on the intermediate scattering function.

After minimizing influences due to the general setup of the simulations, we could concentrate on comparing different water models. All models do not reproduce the diffraction data over the whole Q -range; the TIP4P/2005 model is at low Q better than the TIP3P variants for all probes and is closer to the X-ray data also at higher Q -values. No significant change between the TIP3P rigid and flexible model was visible in the diffractograms. In particular, the O–H bond length appeared to be too well defined in either case, resulting in oscillations in the diffractograms at high Q with a too high amplitude. Since the bond stiffness was also constrained by other measurements such as Raman spectroscopy, this could be connected to nuclear quantum effects, which were not accounted for in the current simulation [46, 85, 86].

TIP4P/2005 captures the dynamics very well at low Q , which is consistent with the good reproduction of the long-range diffusion coefficient. The long-range mobility of TIP3P increases upon flexibilization. The local dynamics probed at higher Q is not well reproduced by any of the models shown here. Interestingly, the effect of introducing flexibility into TIP3P manifests itself mainly at the picosecond time scale, although the bond vibrations take place on a much shorter time scale.

A more detailed characterization of the simulated intermediate scattering function was done in section 2.6. The formula was able to reproduce the computed points in the covered time and Q -range. However, as seen in Table 2, the fitted values differ from the values in the literature by significant amounts. This disagreement is visible for both water models, highlighting that the simulated microscopic dynamics can deviate significantly even if the macroscopic diffusion coefficient is replicated correctly by the model. A further analysis of the individual values was not undertaken in the current contribution.

The rigid TIP3P model was modified by changing some exemplary parameters individually. Examples of the variation in the O–H bond length and H–O–H angle are shown in this work. We observed that a small variation in bond length can significantly impact the structure and dynamics, whereas a change in bond angle does not. The sensitivity of the computed scattering patterns to simulation parameter adaptations implies that an optimization of the agreement between computed and measured scattering curves could help refine the set of simulation parameter values. It should be noted that a water model consists of a set of parameters, so optimizing only one parameter is not expedient.

5 Summary and outlook

This work established a functional workflow of running MD simulations, calculating X-ray and neutron scattering data, and comparing these curves to already published experimental data. It was found that the scattering curves do not agree completely with the measured data; this disagreement was quantified in the case of dynamics. It was also observed that the scattering curves reacted with different sensitivity to different parameter changes.

Future work will compare the simulated curves with experimental data to use the differences between the two to optimize the set of simulation parameters. A multi-parameter fitting with Monte Carlo methods for this purpose is already in development, and possible machine learning techniques will also be evaluated. The aim is to obtain a procedure that can find a set of parameters for a system to reproduce the real nanoscopic structure and dynamics as probed by neutron and X-ray scattering.

References

[1] D.S. Sivia, *Elementary scattering theory* (Oxford University Press, London, England, 2011)

[2] K. Amann-Winkel, M.C. Bellissent-Funel, L.E. Bove, T. Loerting, A. Nilsson, A. Paciaroni, D. Schlesinger, L. Skinner, *Chemical Reviews* **116**, 7570 (2016)

[3] K. Nishikawa, N. Kitagawa, *Bulletin of the Chemical Society of Japan* **53**, 2804 (1980)

[4] P. Lindner, F. Leclercq, P. Damay, *Physica B: Condensed Matter* **291**, 152 (2000)

[5] A.H. Narten, H.A. Levy, *The Journal of Chemical Physics* **55**, 2263 (1971)

[6] A.H. Narten, M.D. Danford, H.A. Levy, *Discuss. Faraday Soc.* **43**, 97 (1967)

[7] F. Cavatorta, A. Deriu, D.D. Cola, H.D. Middendorf, *Journal of Physics: Condensed Matter* **6**, A113 (1994)

[8] D.D. Cola, A. Deriu, M. Sampoli, A. Torcini, *The Journal of Chemical Physics* **104**, 4223 (1996)

[9] B.J. Niebuur, W. Lohstroh, M.S. Appavou, A. Schulte, C.M. Papadakis, *Macromolecules* **52**, 1942 (2019)

[10] A. Arbe, G.J. Nilsen, J.R. Stewart, F. Alvarez, V.G. Sakai, J. Colmenero, *Physical Review Research* **2** (2020)

[11] F.H. Stillinger, A. Rahman, *The Journal of Chemical Physics* **60**, 1545 (1974)

[12] W.L. Jorgensen, J. Chandrasekhar, J.D. Madura, R.W. Impey, M.L. Klein, *The Journal of Chemical Physics* **79**, 926 (1983)

[13] J.L.F. Abascal, C. Vega, *The Journal of Chemical Physics* **123**, 234505 (2005)

[14] M.A. González, J.L.F. Abascal, *The Journal of Chemical Physics* **135**, 224516 (2011)

[15] J. Zielkiewicz, *The Journal of Chemical Physics* **123**, 104501 (2005)

[16] Y. Ding, A.A. Hassanali, M. Parrinello, *Proceedings of the National Academy of Sciences* **111**, 3310 (2014)

[17] T.A. Ho, A. Striolo, *Molecular Simulation* **40**, 1190 (2014)

[18] A. Henao, S. Busch, E. Guàrdia, J.L. Tamarit, L.C. Pardo, *Physical Chemistry Chemical Physics* **18**, 19420 (2016)

[19] F. Alvarez, A. Arbe, J. Colmenero, *The Journal of Chemical Physics* **155**, 244509 (2021)

[20] G. Smith, W. Paul, M. Monkenbusch, D. Richter, *Chemical Physics* **261**, 61 (2000)

[21] M. Brodeck, F. Alvarez, A. Arbe, F. Juranyi, T. Unruh, O. Holderer, J. Colmenero, D. Richter, *The Journal of Chemical Physics* **130**, 094908 (2009)

[22] J.C. Smith, M. Krishnan, L. Petridis, N. Smolin, in *Dynamics of Soft Matter* (Springer US, 2011), pp. 189–204, https://doi.org/10.1007/978-1-4614-0727-0_7

[23] A. Arbe, F. Alvarez, J. Colmenero, *Polymers* **12**, 3067 (2020)

[24] N. Zec, G. Mangiapia, A.C. Hendry, R. Barker, A. Koutsoubas, H. Frielinghaus, M. Campana, J.L. Ortega-Roldan, S. Busch, J.F. Moulin, *Membranes*

- 11**, 507 (2021)
- [25] R.L. McGreevy, L. Pusztai, *Molecular Simulation* **1**, 359 (1988)
- [26] A.K. Soper, *Molecular Simulation* **38**, 1171 (2012)
- [27] T. Youngs, *Molecular Physics* **117**, 3464 (2019)
- [28] A.J. Markvardsen, R.L. McGreevy (2017)
- [29] J. Armstrong, *Journal of Physics Communications* **6**, 102002 (2022)
- [30] D.J. Price, C.L. Brooks, *The Journal of Chemical Physics* **121**, 10096 (2004)
- [31] P. Mark, L. Nilsson, *The Journal of Physical Chemistry A* **105**, 9954 (2001)
- [32] A. Pérez, A. Rubio, *The Journal of Chemical Physics* **135**, 244505 (2011)
- [33] K.T. Wikfeldt, M. Leetmaa, M.P. Ljungberg, A. Nilsson, L.G.M. Pettersson, *The Journal of Physical Chemistry B* **113**, 6246 (2009)
- [34] M. Rovira-Esteva, N.A. Murugan, L.C. Pardo, S. Busch, J.L. Tamarit, S. Pothoczki, G.J. Cuello, F.J. Bermejo, *Physical Review B* **84** (2011)
- [35] V. Reich, A. Majumdar, M. Müller, S. Busch, *Data publication to comparison of molecular dynamics simulations of water with neutron and x-ray scattering experiments* (2022), <https://dx.doi.org/10.5281/zenodo.7225368>
- [36] A.K. Soper, *ISRN Physical Chemistry* **2013**, 1 (2013)
- [37] J. Teixeira, M.C. Bellissent-Funel, S.H. Chen, A.J. Dianoux, *Physical Review A* **31**, 1913 (1985)
- [38] M. Orsi, *Molecular Physics* **112**, 1566 (2013)
- [39] E.E. Ong, J.L. Liow, *Fluid Phase Equilibria* **481**, 55 (2019)
- [40] C. Vega, J.L.F. Abascal, M.M. Conde, J.L. Aragones, *Faraday Discuss.* **141**, 251 (2009)
- [41] L. Pusztai, O. Pizio, S. Sokolowski, *The Journal of Chemical Physics* **129**, 184103 (2008)
- [42] R.B. Best, J. Mittal, *The Journal of Physical Chemistry B* **114**, 14916 (2010)
- [43] G. Guevara-Carrion, J. Vrabec, H. Hasse, *The Journal of Chemical Physics* **134**, 074508 (2011)
- [44] I.G. Tironi, R.M. Brunne, W.F. van Gunsteren, *Chemical Physics Letters* **250**, 19 (1996)
- [45] D.E. Smith, A.D.J. Haymet, *The Journal of Chemical Physics* **96**, 8450 (1992)
- [46] S. Habershon, T.E. Markland, D.E. Manolopoulos, *The Journal of Chemical Physics* **131**, 024501 (2009)
- [47] H.J.C. Berendsen, J.P.M. Postma, W.F. van Gunsteren, J. Hermans, in *The Jerusalem Symposia on Quantum Chemistry and Biochemistry* (Springer Netherlands, 1981), pp. 331–342, https://doi.org/10.1007/978-94-015-7658-1_21
- [48] H.J.C. Berendsen, J.R. Grigera, T.P. Straatsma, *The Journal of Physical Chemistry* **91**, 6269 (1987)
- [49] M.W. Mahoney, W.L. Jorgensen, *The Journal of Chemical Physics* **112**, 8910 (2000)
- [50] J.R. Grigera, *The Journal of Chemical Physics* **114**, 8064 (2001)
- [51] S.W. Rick, *The Journal of Chemical Physics* **114**, 2276 (2001)
- [52] H.A. Stern, F. Rittner, B.J. Berne, R.A. Friesner, *The Journal of Chemical Physics* **115**, 2237 (2001)
- [53] H.W. Horn, W.C. Swope, J.W. Pitera, J.D. Madura, T.J. Dick, G.L. Hura, T. Head-Gordon, *The Journal of Chemical Physics* **120**, 9665 (2004)
- [54] S.Y. Liem, P.L.A. Popelier, M. Leslie, *International Journal of Quantum Chemistry* **99**, 685 (2004)
- [55] S.W. Rick, *The Journal of Chemical Physics* **120**, 6085 (2004)
- [56] H. Yu, W.F. van Gunsteren, *The Journal of Chemical Physics* **121**, 9549 (2004)
- [57] J.L.F. Abascal, E. Sanz, R.G. Fernández, C. Vega, *The Journal of Chemical Physics* **122**, 234511 (2005)
- [58] P. Paricaud, M. Předota, A.A. Chialvo, P.T. Cummings, *The Journal of Chemical Physics* **122**, 244511 (2005)
- [59] G.S. Fanourgakis, S.S. Xantheas, *The Journal of Physical Chemistry A* **110**, 4100 (2006)
- [60] Y. Wu, H.L. Tepper, G.A. Voth, *The Journal of Chemical Physics* **124**, 024503 (2006)
- [61] P.T. Kiss, A. Baranyai, *The Journal of Chemical Physics* **138**, 204507 (2013)
- [62] L.P. Wang, T. Head-Gordon, J.W. Ponder, P. Ren, J.D. Chodera, P.K. Eastman, T.J. Martinez, V.S. Pande, *The Journal of Physical Chemistry B* **117**, 9956 (2013)
- [63] W. Yu, P.E.M. Lopes, B. Roux, A.D. MacKerell, *The Journal of Chemical Physics* **138**, 034508 (2013)
- [64] S.J. Bachmann, W.F. van Gunsteren, *The Journal of Chemical Physics* **141**, 22D515 (2014)
- [65] S. Izadi, R. Anandakrishnan, A.V. Onufriev, *The Journal of Physical Chemistry Letters* **5**, 3863 (2014)
- [66] R. Qi, L.P. Wang, Q. Wang, V.S. Pande, P. Ren, *The Journal of Chemical Physics* **143**, 014504 (2015)
- [67] R. Fuentes-Azcatl, M.C. Barbosa, *Physica A: Statistical Mechanics and its Applications* **444**, 86 (2016)
- [68] S. Izadi, A.V. Onufriev, *The Journal of Chemical Physics* **145**, 074501 (2016)
- [69] Y. Khalak, B. Baumeier, M. Karttunen, *The Journal of Chemical Physics* **149**, 224507 (2018)
- [70] C.L. Zhao, D.X. Zhao, C.C. Bei, X.N. Meng, S. Li, Z.Z. Yang, *The Journal of Physical Chemistry B* **123**, 4594 (2019)
- [71] A.P. Thompson, H.M. Aktulga, R. Berger, D.S. Bolintineanu, W.M. Brown, P.S. Crozier, P.J. in 't Veld, A. Kohlmeyer, S.G. Moore, T.D. Nguyen et al., *Comput. Phys. Commun.* **271**, 108171 (2022)
- [72] C. Vega, E. de Miguel, *The Journal of Chemical Physics* **126**, 154707 (2007)
- [73] W. Humphrey, A. Dalke, K. Schulten, *Journal of Molecular Graphics* **14**, 33 (1996)
- [74] T. Róg, K. Murzyn, K. Hinsén, G.R. Kneller, *Journal of computational chemistry* **24**, 657 (2003)
- [75] N.P. Walter, A. Jaiswal, Z. Cai, Y. Zhang, *Computer Physics Communications* **228**, 209 (2018)

- [76] B. Lindner, J.C. Smith, *Computer Physics Communications* **183**, 1491 (2012)
- [77] B. Lindner, dissertation, University of Tennessee (2012), https://trace.tennessee.edu/utk_graddiss/1589/
- [78] V.F. Sears, *Neutron News* **3**, 26 (1992)
- [79] P.J. Brown, A.G. Fox, E.N. Maslen, M.A. O'Keefe, B.T.M. Willis, in *International Tables for Crystallography* (International Union of Crystallography, 2006), pp. 554–595, <https://doi.org/10.1107/97809553602060000600>
- [80] O. Dietrich, *Diffusion coefficients of water*, <https://dtrx.de/od/diff/>
- [81] L.P. Wang, T.J. Martinez, V.S. Pande, *The Journal of Physical Chemistry Letters* **5**, 1885 (2014)
- [82] J.B. Hasted, in *The Physics and Physical Chemistry of Water* (Springer New York, 1972), pp. 255–309, https://doi.org/10.1007/978-1-4684-8334-5_7
- [83] A. Zeidler, P.S. Salmon, H.E. Fischer, J.C. Neufeind, J.M. Simonson, T.E. Markland, *Journal of Physics: Condensed Matter* **24**, 284126 (2012)
- [84] I.C. Yeh, G. Hummer, *The Journal of Physical Chemistry B* **108**, 15873 (2004)
- [85] I. Bakó, Á. Madarász, L. Pusztai, *Journal of Molecular Liquids* **325**, 115192 (2021)
- [86] A. Eltareb, G.E. Lopez, N. Giovambattista, *Physical Chemistry Chemical Physics* **23**, 6914 (2021)