# Lawrence Berkeley National Laboratory
## Chemical Sciences

### Title
Water structure as a function of temperature from X-ray scattering experiments and ab initio molecular dynamics

### Permalink
https://escholarship.org/uc/item/86z7z9r1

### Journal
Lawrence Berkeley National Laboratory

### Authors
Hura, Greg
Russo, Daniela
Glaeser, Robert M.
et al.

### Publication Date
2003-03-01

Peer reviewed

eScholarship.org Powered by the California Digital Library
University of California

**Corresponding author:** Prof. Teresa Head-Gordon, Department of Bioengineering, Donner 459A, University of California, Berkeley, CA 94720, USA

**Title: Water Structure as a Function of Temperature from X-ray Scattering Experiments and Ab Initio Molecular Dynamics**

**Greg Hura[1], Daniela Russo[2], Robert M. Glaeser[1,3], Teresa Head-Gordon[1,2]**
*Graduate Group in Biophysics[1], Department of Bioengineering[2], Department of Molecular and Cell Biology[3], University of California, Berkeley CA 94720*

**Matthias Krack and Michele Parrinello**
*Swiss Center for Scientific Computing, Via Cantonale, CH-6928 Manno, Switzerland and ETH Zurich, Physical Chemistry, Hoenggerberg-HCI, CH-8093 Zurich, Switzerland*

---

We present high-quality x-ray scattering experiments on pure water taken over a temperature range of 2°C to 77°C using a synchrotron beam line at the Advanced Light Source (ALS) at Lawrence Berkeley National Laboratory. The ALS x-ray scattering intensities are *qualitatively* different in trend of maximum intensity over this temperature range compared to older x-ray experiments. While the common procedure is to report both the intensity curve and radial distribution function(s), the proper extraction of the real-space pair correlation functions from the experimental scattering is very difficult due to uncertainty introduced in the experimental corrections, the proper weighting of OO, OH, and HH contributions, and numerical problems of Fourier transforming truncated data in Q-space. Instead we consider the direct calculation of x-ray scattering spectra using electron densities derived from density functional theory based on real-space configurations generated with classical water models. The simulation of the experimental intensity is therefore definitive for determining radial distribution functions over a smaller Q-range. We find that the TIP4P, TIP5P and polarizable TIP4P-Pol2 water models, with DFT-LDA densities, show very good agreement with the experimental intensities, and TIP4P-Pol2 in particular shows quantitative agreement over the full temperature range. The resulting radial distribution functions from TIP4P-Pol2 provide the current best benchmarks for real-space water structure over the biologically relevant temperature range studied here.

---

## Introduction

Liquid water structure is characterized by x-ray (or neutron) diffraction that measures experimental intensities as a function of momentum transfer, $Q=4\pi sin(\theta/2)/\lambda$, where $\lambda$ is the wavelength and $\theta$ is the scattering angle with respect to the incident beam. The most recent x-ray data taken at ambient conditions at the Advanced Light Source (ALS) at Lawrence Berkeley National Laboratory exhibited significant differences when compared to the scattering curves of past x-ray experiments[1-3]. We believe the new measurements to be superior since they benefit from the use of a third generation synchrotron source, more accurate intensities using a modern CCD detector, and careful control of error in the applied corrections[3]. However no such high quality x-ray measurements exist yet at other thermodynamic state points. In this study we detail the collection and processing of x-ray diffraction intensity data from pure liquid water throughout the temperature range of 2°C to 77°C at ambient pressure, a temperature range that is most important for water related to biological studies.

The intensity is the true experimental observable in which error-bars are well-defined. However, it is typical practice for water scattering experiments to also report radial distributions in addition to the intensity profile, primarily because it is more convenient and practical to consider water structure in terms of real-space distribution functions [1,4-14]. In actuality the $g_{OO}(r)$'s extracted from the x-ray diffraction measurements are highly-derived functions, not necessarily unique, and there is no way of quantifying the errors introduced in their extraction based on finite $Q$ data[3]. Therefore it is not surprising that the reported oxygen-oxygen radial distribution function, $g_{OO}(r)$, obtained by various x-ray and neutron scattering groups over the last 30 years have been inconsistent[4,10,12]. Recently, independent experimental approaches of x-ray[1] and neutron[14] scattering, using different analyses to extract radial distribution

functions[2,15], independently converged on the same oxygen-oxygen radial distribution function for ambient temperature and pressure as shown in Figure 1.

However extracting real-space pair correlation functions from the intensity remains a primary source of uncertainty that potentially degrades the actual high-quality of the experimental data itself. Currently it is thought that the intensity out to Q=20-25Å$^{-1}$ would be needed to reliably Fourier transform the data to get converged real-space functions. Based on the uncertainties in any approach for extracting real-space functions, such as *Q*-space continuation[7], fitting procedures in *r*-space[2], and reverse Monte Carlo methods[16,17] and its extensions[15], it is important to further test the assumptions made in deriving radial distribution functions based on a smaller range of Q. With the advent of completely *ab initio* simulations of liquid water[18-21], it is now possible to predict the x-ray scattering from the distribution of electron density, $\rho(\mathbf{r})$, generated from the condensed phase simulation trajectory[22] using Eq. (1)

$$I(Q) = \left\langle |F(\mathbf{q})|^2 \right\rangle$$
$$\text{with } F(\mathbf{q}) = \int_0^\infty \rho(\mathbf{r}) \exp(i\mathbf{q}\mathbf{r}) d\mathbf{r} \tag{1}$$

This allows for the straight simulation of the true experimental observable, the intensity, without requiring the contrived and error-prone process of extracting radial distribution functions from intensity data over a finite Q-range. If the simulated intensity curve using Eq. (1) agrees well with experiment (with only moderately large Q-ranges), then we can analyze the corresponding real-space trajectory to ascertain the pair distribution function(s).

In this study we consider the direct calculation of x-ray scattering spectra using ab-initio density functional theory with the LDA functional over the temperature range studied by experiment. The generation of the real-space "snapshots" could come from either a first

principles molecular dynamics calculation[18], that describes the time-evolution of the electronic structure, or a classical molecular dynamics simulation based on an empirical water model. Due to the current expense of simulating the full *ab initio* trajectory, we instead evaluate equation (1) by integrating over positions generated by the classical simulation of various water models.

The combined experimental and theoretical benchmarks are applied to judging the quality of various TIP non-polarizable[23,24] and polarizable[25] water force fields over the range of temperature of 2-77°C that we have examined. We find that the TIP5P non-polarizable water model[24] and the TIP4P-Pol2[25] polarizable model, with DFT-LDA densities, show very good agreement with the experimental intensities, and TIP4P-Pol2 in particular shows quantitative agreement over the full temperature range. The resulting radial distribution functions from TIP4P-Pol2, and for TIP5P for temperatures greater than 25°C, provide the current best benchmarks for real-space water structure over the biologically relevant temperature range studied here. We also examine the relative error introduced by LDA electron densities vs. error in water force field configurations, using an independent analysis method described in [2,3].

## Experimental Methods

*Experimental setup.* The data collection was performed at the Advanced Light Source (ALS) at Lawrence Berkeley National Laboratory on beam line 7.3.3. Doubly distilled and degassed water was used in each experiment, and several data sets with independent fillings of the sample holder were collected at each temperature (2°C, 11°C, 22°C, 33°C, 44°C, 55°C, 66°C, and 77°C). A silicon monochrometer was used to produce an x-ray beam with spot size on the order of 100μm, with 99% of the energy at 12.800 ± 0.001 keV or a wavelength of 0.9686Å. This wavelength provided the best compromise between sufficient flux and maximizing our

accessible Q-range. Data sets were collected with a flat slab water sample using a transmisson geometry, with the sample tilted with respect to the incoming x-ray beam by an angle of $30^{o}$.

A Bruker Charge Coupled Device (CCD) area detector, mounted on a Huber diffractometer and with dimensions of 9.6cm x 9.6cm collected the diffracted x-rays. In order to realize the full range of $0.1\text{Å}^{-1} < Q < 11.1\text{Å}^{-1}$ provided in this study, the data were collected with the detector in three different positions which were pieced together for the final result. The sample to detector positions were determined to 100μm precision while the angle between the x-ray beam and the normal to detector face were characterized on the order of mrads. The geometries at each detector position were determined by collecting PbS powder patterns with the powder placed within the sample holder and fitting the resulting sharp Bragg rings to determine all the geometric parameters. A more detailed description of this procedure can be found in previous work[1]. The three sample to detector positions and collection times are detailed in Table 1. The lowest angles were limited by the beam stop at $0.8832^{o}$ or a $Q$=$0.1\text{Å}^{-1}$, the highest angles were limited by the monochrometer housing and sample holder aperture at $118^{o}$ or at a $Q$ =$11.1\text{Å}^{-1}$.

Our sample holder was based on a similar design of our previously described sample holder [1] with several additions in order to control the temperature of the sample, and is shown in Figure 2 and described in the figure caption. Temperature control of the sample relied on a Peltier device or thermoelectric cooler. A thermocouple placed within the sample chamber, monitored the temperature. An Alpha Omega Instruments Series 800 Controller acted as the thermostat. One face of the Peltier device was placed in thermal contact with the aluminum plate augmented by a zinc oxide thermal grease. In order to achieve temperatures below room temperature a water cooled aluminum block was placed on the other side of the Peltier device as

a heat sink. In order to prevent condensation below the dew point a helium shroud was built around the sample holder. The entrance and exit windows of the helium shroud were made of 1μm thick Mylar purchased from Goodfellow Inc.

*Experimental corrections*. The collected raw intensity data was transformed to a circularly integrated scattering cross-section on a per electron scale versus *Q*. The following divides the corrections into two parts. The first deals with corrections that must be applied generally to all detector positions such as absorption, geometric corrections, and polarization of the radiation. The second set of corrections deals with overlapping data from different panels and thus the contribution of background and sample holder scattering.

Corrections to the intensities due to absorption by air, water, and window material are given by the form

$$\frac{I}{I_0} = \left[ exp\left(\frac{-\mu_\rho t}{\sin\tau}\right) - \exp\left(\frac{-\mu_\rho t}{\nu}\right)\frac{\nu \sin\tau}{\mu_\rho(\sin\tau - \nu)}\right] \qquad (2)$$

where $I_o$ is the intensity if there were no absorption in the sample, $I$ is the measured intensity, $t$ is the thickness, $\tau$ is the angle between the plane of the sample and the incident x-ray beam, and $\nu$ has the form $\nu = cos\theta sin\tau + sin\theta cos\phi cos\tau$. The absorption coefficients, $\mu_\rho$, can be obtained from http://www-cxro.lbl.gov/optical_constants using a tabulated format based on [26]. The absorption coefficient for water is 2.396 ($cm^2/g$) for 1C water at our operating wavelength and varies proportionally with the density for different temperatures. Thickness of samples were determined by measuring the x-ray transmisson thorough the sample using a PIN diode. The error in the thickness was on the order of 0.02cm and did not contribute significantly to the error of the intensity curve.

When the incident radiation is plane-polarized as in our experiment, the in-plane and out-of-plane polarization is treated separately, and the measured intensities must be rescaled by the factor

$$P = P_{in-plane}\left[1-(\sin\phi\,\sin\theta)^2\right]+P_{out-plane}\left[1-(\cos\phi\,\sin\theta)^2\right] \quad (3)$$

where $P_{in\text{-}plane}$ and $P_{out\text{-}plane}$ are the fractional polarizations of the beam in the plane and out of the plane of the circulating electrons, respectively.

The data were also corrected for the $1/r^2$ fall off of intensity, and correction for pixel orientation with respect to the incident radiation. We discovered that detector manufacturers attempt to provide a correction of this sort in their bundled software for data read-out. Corrections to the image collected by the CCD includes distortions introduced by the fiber optic tapers connecting the plate to the chip, a dark current correction that takes into account photoelectrons ejected by thermal motions in the CCDs and any low-level background radiation in the experimental hutch, and finally a flat field correction that provides a correction for variations in pixel sensitivity. This correction involves a calibration measurement that collects a reference image intensity using either a radioactive source or scattering from a fluorescent material at a set distance from the detector. The image is used to generate a scale factor for each pixel such that, when applied to the flat field image, each pixel measures the same intensity. The set distance with which this correction is calibrated varies among manufacturers. If the calibration experiment involves a relatively small source to detector distance, then the geometric corrections ($1/r^2$ fall off, pixel orientation) are implicitly taken into account by the manufacturer in the flat field correction. This is the case for the Bruker detector, so we need to "uncorrect" the collected image for the manufacturers geometry, and reapply the correct geometric corrections for our experiment. This is an insidious problem since these detector corrections are applied

before the “raw image” is available to the experimentalist, and the correction is not documented in the manufacturer literature.

We next turn to the procedure for the subtraction of background that is required in order to match segments of the intensity over the full range of $Q$ in this experiment. Figure 3a shows the fully corrected, filled and empty sample holder data at one temperature, at a point before the three separate positions of the detector have been pieced together. We had intended to use the overlapping regions of the curves to determine an internal scale factor that would place all three segments of the data on the same scale. This is easily done for overlapping the second and third detector positions. However while the overlap between the first and second detector positions for the filled sample is quite good, the overlap between the first and second detector positions for the background is very poor. We attribute this mismatch to additional air-scattering by the x-ray beam once it exits the sample, something that is less of a problem for the filled sample because of its greater absorption cross-section. As a result, we could not simply overlap panels for the filled sample, overlap panels for the empty cell, and subtract the latter from the former to provide a correction for scattering background.

Instead, the background can be subtracted from the first detector position by making use of the prominent kapton window peak at $Q = 0.4\text{Å}^{-1}$ which occurs in a region where there are no sharp features in the liquid water scattering. The factor required for the background subtraction is found when the kapton peak is clearly removed and the intensity in the associated $Q$ region is flat. This factor can be empirically fine-tuned by extrapolating the intensity to the value of the isothermal compressibility at $Q$=0, once the intensity is on an electron units scale. In order to overlap the second panel with the first panel, two scale factors need to be found; a factor related

to the flux over the exposure time ($a_1$) and a scaling for the kapton subtraction ($a_2$). This is summarized in Eq. (4)

$$a_1 I_2(Q) - a_2 I_k(Q) = I_1(Q) \tag{4}$$

where $I_2(Q)$ is the intensity of the filled sample holder in the second detector position, $I_k(Q)$ is the intensity of the background (empty cell), and $I_1(Q)$ is the intensity of the filled sample holder in the first detector position. These two factors are determined by standard least-squares fitting over the overlapping $Q$ range. The third detector position that already overlapped in shape with the second detector position used the same kapton subtraction factor, $a_2$. The resulting curve is shown in Figure 3b.

Once we have the intensity curve over the full $Q$-range, we make a correction for inelastic or Compton scattering as discussed in [1] using a fit to CISD calculations available for water [27,28]. The Krogh-Moe [29], [[30] method was used to normalize the scattering intensity to an absolute scale.

## Theoretical Methods

It is conventional in the analysis of the scattering from molecular liquids to separate the intensity into contributions from individual molecules (self-scattering), and that arising from intermolecular correlations defined by the structure factor, $H(Q)$.

$$I(Q) = \sum_{ij} x_i x_j f_i(Q) f_j(Q) \frac{\sin Qr_{ij}}{Qr_{ij}} + \sum_{i \le j} x_i x_j f_i(Q) f_j(Q) H_{ij}(Q) \tag{5}$$

where the sums are over the atom types present in the sample, $x_i$ is the atomic fraction and $f_i(Q)$ the $Q$-dependent atomic scattering factor[31], for atom type $i$. The structure factor is

$$H_{ij}(Q) = 4\pi\rho\int_0^\infty r^2\,dr\left(g_{ij}(r)-1\right)\frac{\sin Qr}{Qr} \tag{6}$$

where ρ is the atomic density, $r_{ij}$ are the intramolecular distances between atom centers, and $g_{ij}(r)$ is the radial distribution function describing intermolecular correlations between atom types *i* and *j*. In the case of x-rays the assumption is commonly made that the scattering can be represented as arising from independent neutral atoms, each with a spherical electron density distribution. However, it is clear from our previous work that a superposition of spherical electron densities of free atoms centered on the individual atoms of chemically bonded water is not adequately representative of the true charge distribution[2].

We have recently introduced a modification of the atomic scattering factors for liquid water which rescales them properly at low *Q* where chemical bonding effects are known to be significant, while retaining their values at large *Q* where chemical bonding effects on the core electron distribution should be negligible[2]. The modified atomic form factors (MAFF) $f'(Q)$ is of the form

$$f'(Q) = \left[1+(\alpha-1)exp\left(-Q^2/2\delta^2\right)\right]f(Q) \tag{7}$$

where $f(Q)$ is the atomic scattering factor for the isolated atom, α is a scaling factor giving the redistribution of charge, and δ is a parameter to be fit, representing the extent of valence-electron delocalization induced by chemical bonding[2]. For liquid water we choose α to be 2.8D, and the unknown parameter δ was fit by requiring the Debye expression curve to agree with the *ab initio* CI results for gas phase water[28]. Figure 4 shows the difference in a simulated intensity when using the same partial structure factors ($h_{OO}(Q)$, $h_{OH}(Q)$, and $h_{HH}(Q)$) but free atom form factors based on isolated atoms or the MAFFs that take into account chemical bonding. The difference

arises because of chemical bonding effects, which now weight the O-O terms as ~85% of the signal, while O-H terms contribute ~15%, with negligible contribution from H-H correlations.

Explicit quantum treatments of electronic structure in aqueous simulations have been reported, where all have relied on a local density approximation to density functional theory[18-21,32]. The cost of these calculations rises as $N^3$ where N is the number of basis functions or plane waves per atom, with a large coefficient relative to classical simulations that make these approaches computationally demanding, but potentially more predictive. Differences among functionals in the condensed phase have been investigated for liquid water in [19], and commonly used functionals include BLYP[33-35] and PBE[36]. In principle, these approaches provide the current best theoretical means for providing intensity curves directly. These calculations can also be corrected for quantum effects.

As discussed in [22], a direct calculation of the coherent x-ray scattering spectrum of liquid water by means of DFT is feasible when using an implementation of the Gaussian augmented plane wave (GAPW) method[37], called Quickstep, which can perform all-electron ab-initio DFT calculations for periodic systems[38]. The evaluation of an intensity curve from a 10-ps-long CPMD simulation of 64 water molecules with the gradient-corrected BLYP functional and with an average ionic temperature of 318K[20] using the full electronic charge distributions was compared to the ALS2000 experiment[1], which was conducted at room temperature. Overall these results are promising, but are not in themselves definitive because of the short timescale of the simulation and problems of large temperature fluctuations for these small system sizes, where differences of 10-20 degrees in temperature will affect the structural comparisons with room temperature experimental data[2].

The computational cost of dynamics generated with CPMD and the evaluation of the full electronic charge distribution are too prohibitive for the present study. To make these calculations more feasible for the large number of experiments considered here, we have repeated the calculation using the 64-water CPMD trajectory and a *superposition* of DFT molecular densities as an approximation to the full electronic density. Figure 5a shows that this approximation is extremely good for liquid water, and provides significant computational savings for the analysis.

Furthermore, there is no need to employ the computationally more expensive BLYP functional any longer since the charge distribution of an isolated water molecule is already well described by the LDA functional as shown in Figure 5b, where the DFT-BLYP and the DFT-LDA results for an isolated water molecule are compared to recent accurate configuration interaction calculations that included single and double excitations (CISD)[27] which are also in good agreement with coupled cluster calculations (CCSD)[28]. All the Quickstep calculations presented in this work used a double-zeta valence plus p polarization (DZVP) and a triple-zeta plus d polarization basis set (TZVP) for H and O, respectively.

We also provide an alternative examination of classical water force fields by using them as the propagators of nuclear positions, and gathering snapshots of the condensed phase trajectories upon which we evaluate intensities based on molecular densities generated using DFT-LDA and Eq. (1). The classical force fields provide us with a much less expensive way to simulate much longer trajectories and larger system sizes while still simulating the intensity curve to make direct contact to the experimental observable.

We calculated *I(Q)* reported in the Results section using 1-ns-long MD trajectories generated with the classical water models TIP4P[23] and TIP5P[24] for a liquid water model

system with 256 water molecules in a simple cubic box, and we used Ewald sums to treat long-ranged electrostatics. 1,000 atomic configurations equally spaced by 1ps were extracted from each trajectory and the corresponding intensity curves were calculated for each atomic configuration. We also calculated *I(Q)* using Eq. (1) using Monte-Carlo (MC) simulations with the TIP4P-Pol1 and the TIP4P-Pol2 water models[25], which are polarizable versions of the TIP4P model. In total 1,000,000 MC cycles were performed during each MC simulation and after every 1,000 MC cycles an atomic configuration was obtained. Each MC cycle consisted of 500 MC steps to ensure a comparable statistic with respect to the MD simulations.

## Results

Figures 6a-6c show the experimental intensity, with error bars, for representative temperatures of 2°C, 25°C, and 77°C, compared to x-ray scattering data reported by Narten over 30 years ago. The ALS data for these temperatures, and additional temperatures of 10°C, 44°C, 55°C, and 66°C, are provided in tabular form in the appendix. It is apparent that the ALS x-ray scattering intensities are *qualitatively* different in trend with respect to maximum intensity ($I_{max}$) over this temperature range, compared to the older x-ray experiments, and the difference between them falls well outside of the ALS experimental error bars. Given the improvements in experimental design such as using a bright and highly monochromatic source, a well-characterized polarization and Compton scattering correction, and more accurate intensities using a modern Charge Coupled Device (CCD) detector, we are confident that the ALS results are more correct. Considering only the ALS data, the features that most prominently change as a function of increasing temperature include decreasing $I_{max}$, the shifting of the position of the

main diffraction peak to higher $Q$, and the "melting" of the shoulder/peak in the region of 2.5Å$^{-1}$ < $Q$ < 3.5Å$^{-1}$.

It is typical practice for water scattering experiments to report radial distributions in addition to the intensity profile, especially since experimental measures of water structure are very important in validating empirical water force fields that describe water structure in terms of real-space distribution functions. In our previous study, we introduced a procedure that combines real-space functions to find the optimal $g_{OO}(r)$ which best fits the experimental data[2]. While we believe the independent convergence to the same radial distribution function for ambient conditions by several techniques ensures that the reported ambient $g_{OO}(r)$ is sound (as shown later), we break from this tradition here because in actuality the $g_{OO}(r)$'s extracted from the diffraction measurements are highly-derived functions, not necessarily unique, and there is no way of quantifying the errors introduced in their extraction based on finite $Q$ data[3]. Therefore we only report experimental intensity curves in $Q$-space.

However, we can provide some judgment of the quality of various classical force water models using Eq. (1), where DFT-LDA is used to evaluate the electron density of liquid water based on configurations generated with classical dynamics of a series of empirical water models at various state points. If the simulated intensity curve agrees well with experiment than we can analyze the corresponding real-space functions. In this work we judge the quality of various TIP non-polarizable and polarizable water force fields over the range of temperature of 2-77°C that we have examined experimentally.

Figures 7-9 provides a comparison of the experimental intensity data at the state point of 1bar/25°C (Figure 7), 1bar/2-4°C (Figure 8), and 1bar/75°C (Figure 9) with simulated intensities generated with Eq. (1) and DFT-LDA densities, using configurations generated with the non-

polarizable TIP4P and TIP5P models, and the polarizable TIP4P-Pol1 and TIP4P-Pol2 water models. We begin by using the position of the main peak and shape of the shoulder region over the temperature range to qualitatively evaluate the models over the temperature range studied.

In Figure 7 we see that TIP4P reproduces the position of the peak shoulder, but the shoulder is too soft. TIP5P and TIP4P-Pol1 shift the main peak slightly, with slightly better agreement in the shoulder. The best agreement with the experimental intensity is patently obvious for TIP4P-Pol2 model, and shows almost quantitative agreement over the whole curve. For the non-ambient state point at 1bar/2-4°C (Figure 8), close to the temperature of the density maximum of water, the primary feature to note is that the shoulder at room temperature has evolved to a well-defined peak in the experimental intensity. The TIP4P model does not capture the presence of the second peak, and moreover, the main peak height is underestimated. TIP5P and TIP4P-Pol1 show a shifted main peak, but the second peak is well reproduced by both models. TIP4P-Pol2 realizes better agreement with the position of the main peak and very good agreement over the whole intensity curve. At the higher temperature state point of 1bar/75°C (Figure 9), we see a melting of the shoulder/second peak as determined from experiment. All models capture this melting reasonably well, although TIP4P and TIP4P-Pol1 still show a small amount of over-structuring near the shoulder region. TIP5P and TIP4P-Pol2 show similar performance and good agreement with experiment over most of the curve.

Quantitatively, we have defined experimental error bars as low as 1% for some regions of the intensity curve, and an average error of ~2.5% over the range of $0.1\text{Å}^{-1} < Q < 10.0\text{Å}^{-1}$. In Table 2 we report chi-squared values for measuring agreement between simulated and experimental intensities over this $Q$-range, and judge the quality of a model based on not exceeding the 2.5% upper bound in average error. We see from Table 2 that the TIP4P-Pol2

model is quantitatively accurate over the temperature range studied, and therefore we fully endorse its structural properties as being the best among the four models. TIP4P and TIP5P are quantitatively sound for 25°C and above, but falls outside our accepted error bars for the 4°C data, although small changes for the better would not be surprising if the simulation were rerun at 2°C at which the experimental data was collected. TIP4P-Pol1 seems to systematically fail by this quantitative measure throughout the entire temperature range.

Given this level of performance of the simulated intensities using the TIP4P-Pol2 trajectory and DFT-LDA intensities, we can analyze the radial distribution functions generated from the real-space trajectories themselves. Figures 10 shows the oxygen-oxygen radial distribution functions for TIP4P-Pol2 over the temperature range analyzed here. In our previous theoretical analysis, we concluded that peak position values of $g_{OO}(r)$ for 25°C should be reproducible to within 1%, and that a sharper first peak with a *lower bound* in height of 2.6-2.8, best reproduced our ALS2000 experimental data. The current experiments and theoretical models examined here continue to support those results, and further strengthen the finality of the results at room temperature. It is apparent that the trend in the radial distribution functions as a function of temperature is a very structured $g_{OO}(r)$, with sharp peaks and troughs at low temperatures, which melt at higher temperatures. Evidence of tetrahedral bonding at the highest temperatures is found in the still well-defined second peak. The rather good quantitative results between the TIP4P-Pol2 and experimental intensities make these $g_{OO}(r)$*'s* the current best benchmark for real-space water structure.

The direct calculation of x-ray scattering spectra using ab-initio density functional theory based on configurations generated from empirical water force fields provides an alternative way to realize real-space pair correlation functions from the experimental scattering. However, it is

important to quantify the relative error introduced by DFT-LDA electron densities vs. error in water force fields introduced through the water configurations. Another consideration is the considerable computational cost of evaluating the intensity using Eq. (1) for the ~40 known water models over the temperature range studied here.

We can fortunately examine these questions by comparing to an alternative approach of generating intensities using the MAFFs discussed in the Methods Section and Eq. (5). Figure 11 shows a comparison of intensities derived from DFT-LDA and MAFFs, based on nuclear positions generated with the TIP5P water model (and Ewald electrostatics), to see which approximation agrees better with experiment at 25°C. We conclude from Figure 11 that there are small errors in DFT-LDA densities for small $Q$ with respect to experiment, although regions from $3.5\text{Å}^{-1} < Q < 5.0\text{Å}^{-1}$ are better using DFT-LDA. Therefore disagreement seen at low $Q$ in Figures 6-8 may involve the LDA functional, DFT, or lack of nuclear quantum effects. Overall the densities by DFT-LDA are robust, with the primary error residing in the configurations generated using the underlying classical force fields. We should also emphasize that we have no evidence that configurations generated by DFT dynamics would improve upon this situation.

Given the comparable quality of intensities generated with MAFFs vs. DFT-LDA, we can screen potentially worthwhile water models by first considering their simulated intensity curves using MAFFs against experiment, before proceeding to the more expensive DFT-LDA calculations. Table 3 uses the MAFF analysis to generate percentage error differences for the four models studied here, and various other models including CC[39], MCY[40], NCC-vib[41], SPC[42], SPC/E[43], TIP3P[23], TIP-FQ[44], ST2[45], and ST4[46]. The MAFF analysis provides an inexpensive filter to assess the promise of various empirical water force fields for reproducing experiment over the range of temperatures examined here.

## Conclusion

We have presented new x-ray scattering data for pure water over a temperature range from 2-77°C collected at ALS on beam line 7.3.3, which includes the important state point region near the temperature of the density maximum. The ALS x-ray scattering intensities are *qualitatively* different in trend over this temperature range compared to the older x-ray experiments, and falls well outside of the ALS experimental error bars.

The common procedure in the field of water structure is to report both the intensity curve (or structure factor) and an "extraction" of the real-space pair correlation functions from the experimental scattering. While we have a good measure of error in the resulting intensity curve, we have no real definition of error in the resulting real-space functions due to uncertainty introduced in the experimental corrections, the proper weighting of OO, OH, and HH contributions, and numerical problems of Fourier transforming truncated data in Q-space. For the first time we report the direct calculation of x-ray scattering spectra using ab-initio density functional theory. The simulation of the experimental intensity is therefore definitive for determining radial distribution functions over a smaller Q-range.

The *ab initio* simulations are still too expensive to generate adequate statistics for simulating the intensity, and no firm conclusions can yet be drawn as to the adequacy of DFT dynamics. Instead, DFT with the LDA functional was used to evaluate the electron density of liquid water based on configurations generated with classical dynamics of a series of empirical water models at various conditions, including TIP4P, TIP5P, TIP4P-Pol1 and TIP4P-Pol2.

The experimental and theoretical benchmarks of intensity data provide a means for measuring the progress made in the modeling of water structure. The series of TIPnP non-polarizable models have shown continuous improvement over the years, and the most recent

TIP5P model has exhibited excellent agreement with the experiment at room temperature reported in 2000[2,3], and which has been confirmed again here. Many research groups have actively acknowledged this progress by adopting TIP5P in application to the liquid-liquid phase transition in water[47,48] and disseminating the model in molecular mechanics programs[49,50]. We note that quantitative agreement is not adequate at 4°C, although further analysis may be required in regard to treatment of long-ranged electrostatics and simulating with the temperature fixed to the experimental value of 2°C. We do find that the TIP5P model simulated without Ewald, consistent with the Jorgensen protocol[24], does perform better than the model simulated with Ewald treatment of long-ranged electrostatics (see Table 3).

We also focused on the TIP4P-Pol polarizable models developed by Chen and co-workers[25] based on fluctuating charge models[25,44,51,52], which allow for significant shifts of electron density from one atom to surrounding atoms or bonds within the same molecule. The approach adopted in [25] fits to properties of the liquid phase at ambient conditions and to properties of gas-phase monomers and clusters, with the reasoning that fitting to two different thermodynamic states (with respect to density) should also yield a good description for other densities. We found impressive quantitative agreement with experiment over the whole temperature range from 2°C-77°C for the TIP4P-Pol2 model, and it provides the current best benchmark for radial distribution functions over this temperature range.

Given the comparable quality of intensities generated with MAFFs vs. DFT-LDA, we can screen potentially worthwhile water models by first considering their simulated intensity curves using MAFFs against experiment, before proceeding to the more expensive DFT-LDA calculations. In fact, the TIP4P-Pol2 model was chosen for further study with DFT-LDA because of pre-screening with the MAFF approach. While it might be argued that the MAFF analysis

provides an equally good, but less expensive, approach to DFT-LDA (even with the further approximation of molecular superposition) for evaluating simulation intensities, we see these approaches as mutually beneficial. Ultimately first principles approaches will be more generally applicable, although expensive. MAFFs would be preferred based on reduced computational expense and comparable accuracy, but would require a refitting of parameters at greater extremes of temperature and pressure, and require extension to other molecules and heterogeneous solutions.

## Acknowledgment

M. K. thanks J. I. Siepmann for his support and Bin Chen for performing the MC simulations. THG thanks funding from NIH GM65239-01. GH was supported in part by a National Science Foundation training grant under the Graduate group in Biophysics. The Berkeley group thanks the Advanced Light Source at LBNL for beam time and for assistance on beamline 7.3.3.

## Tables

**Table I.** Specification of data collection for three different positions of the detector relative to the sample.

| Position | Collection time (seconds) | Detector center to sample holder distance (cm) | Angle between x-ray beam and detector face normal | Q range ($Å^{-1}$) |
|---|---|---|---|---|
| 1 | 200 | 9.523 | 21.654° | 0.1 – 5.6 |
| 2 | 300 | 8.532 | 53.501° | 3.1 – 8.6 |
| 3 | 400 | 7.717 | 86.893° | 6.4 – 11.1 |

**Table II.** Percentage difference between experimental and simulated intensities using DFT-LDA over the *Q*-range of 0.4-10.0 $Å^{-1}$ at three different temperatures for different empirical water models.

| Temperature (°C) | TIP4P | TIP5P | TIP4P-Pol1 | TIP4P-Pol2 |
|---|---|---|---|---|
| 2 | 3.3% | 4.7% | 5.8% | 1.8% |
| 25 | 1.8% | 2.2% | 4.4% | 1.0% |
| 75 | 2.2% | 2.0% | 5.5% | 1.7% |

**Table III.** Percentage difference between experimental and simulated intensities using MAFFs over the *Q*-range of 0.4-10.0 $Å^{-1}$ at three different temperatures for different empirical water models.

| Water Model | 2°C | 25°C | 77°C |
|---|---|---|---|
| CC | | 5.3% | |
| MCY | | 11.5% | |
| NCC-vib | | 6.2% | |
| SPC | | 2.0% | |
| SPC/E | 2.8% | 1.7% | 1.8% |
| TIP3P | | 4.2% | |
| TIP4P | | 1.0% | |
| TIP5P (no Ewald) | | 3.5% | |
| TIP5P (Ewald) | 3.8% | 5.8% | 1.4% |
| TIP4P-Pol1 | 3.7% | 3.5% | 3.2% |
| TIP4P-Pol2 | 1.2% | 0.9% | 1.7% |
| TIPAB | | 4.3% | |
| ST2 | 8.3% | 7.9% | 6.3% |
| ST4 | | 5.0% | |

## Figure Captions

**Figure 1.** Comparison of $g_{OO}(r)$ derived from r-space fitting using ALS experimental data [2](dashed line) and neutron experimental data and empirical potential structural refinement (black line)[14].

**Figure 2.** *Details of sample holder design.* The sample was held in a cavity between a disc shape transparent Lucite plate and a heat conducting rectangular aluminum plate, and sealed together by a fitted rubber gasket. Each plate had a 3.0mm hole drilled through them for the x-ray beam. The aluminum plate was fitted with inlet and outlet tubes so that water could fill the sample chamber after assembly, and a groove cut into it that allowed x-rays through the hole when the sample holder was tilted by $30^{o}$ with respect to the x-ray beam. The Lucite plate had a 120° cone shaped tapered hole on its exiting face so as not to obstruct scattered x-rays. The 3mm holes were sealed with 38μm thick kapton (polyimide) window on the inward faces of the plates, in contact with the sample. The sample chamber was machined to nominally 500μm thickness but the actual thickness varied depending on the filling and was at times thicker, presumably due to the bowing of the kapton windows.

**Figure 3 (a).** Fully corrected, filled and empty sample holder intensity data before the three separate positions of the detector have been pieced together. **(b)** The total intensity curve for filled and empty sample holder after piecing together.

**Figure 4.** The difference in simulated intensity when using the same partial structure factors ($h_{OO}(Q)$, $h_{OH}(Q)$, and $h_{HH}(Q)$) but free atom form factors based on isolated atoms (black line) or the MAFFs (gray line) that take into account chemical bonding.

**Figure 5. (a)** Coherent x-ray scattering intensities for liquid water calculated with Quickstep from a CPMD simulation of 64 water molecules using the full electronic density (dashed

line) and a superposition of molecular densities (solid line). **(b)** Coherent x-ray scattering intensities of an isolated water molecule calculated with Quickstep using the functionals BLYP (solid line) and LDA (dashed line) compared with the intensities from a CISD calculation (squares).

**Figures 6.** A comparison of ALS experiment (gray line) and older x-ray experiments (black line) **(a)** Narten 4°C, ALS 2°C . **(b)** Narten 50°C, ALS 44°C . **(c)** Narten 75°C , ALS 77°C.

**Figure 7.** Comparison of experimental spectrum at 27°C (black line) with simulated x-ray scattering intensities at 25°C calculated with QUICKSTEP and using trajectory data from a 1-ns-long simulation of 256 water molecules (gray line). **(a)** TIP4P, **(b)** TIP5P, **(c)** TIP4P-Pol1, and **(d)** TIP4P-Pol2.

**Figure 8.** Comparison of experimental spectrum at 2°C (black line) with simulated x-ray scattering intensities at 4°C calculated with QUICKSTEP and using trajectory data from a 1-ns-long simulation of 256 water molecules (gray line). **(a)** TIP4P, **(b)** TIP5P, **(c)** TIP4P-Pol1, and **(d)** TIP4P-Pol2.

**Figure 9.** Comparison of experimental spectrum at 75°C (black line) with simulated x-ray scattering intensities at 77°C calculated with QUICKSTEP and using trajectory data from a 1-ns-long simulation of 256 water molecules (gray line). **(a)** TIP4P, **(b)** TIP5P, **(c)** TIP4P-Pol1, and **(d)** TIP4P-Pol2.

**Figure 10.** Real-space distribution functions from the TIP4P-Pol2 model that best reproduces the experimental intensity for 2°C (black), 25°C (gray), and 77°C (dashed). **(a)** $g_{OO}(r)$, and **(b)** $g_{OH}(r)$.

**Figure 11.** Comparison of DFT-LDA and MAFF's using the same underlying real-space trajectory for reproducing the experimental intensity.

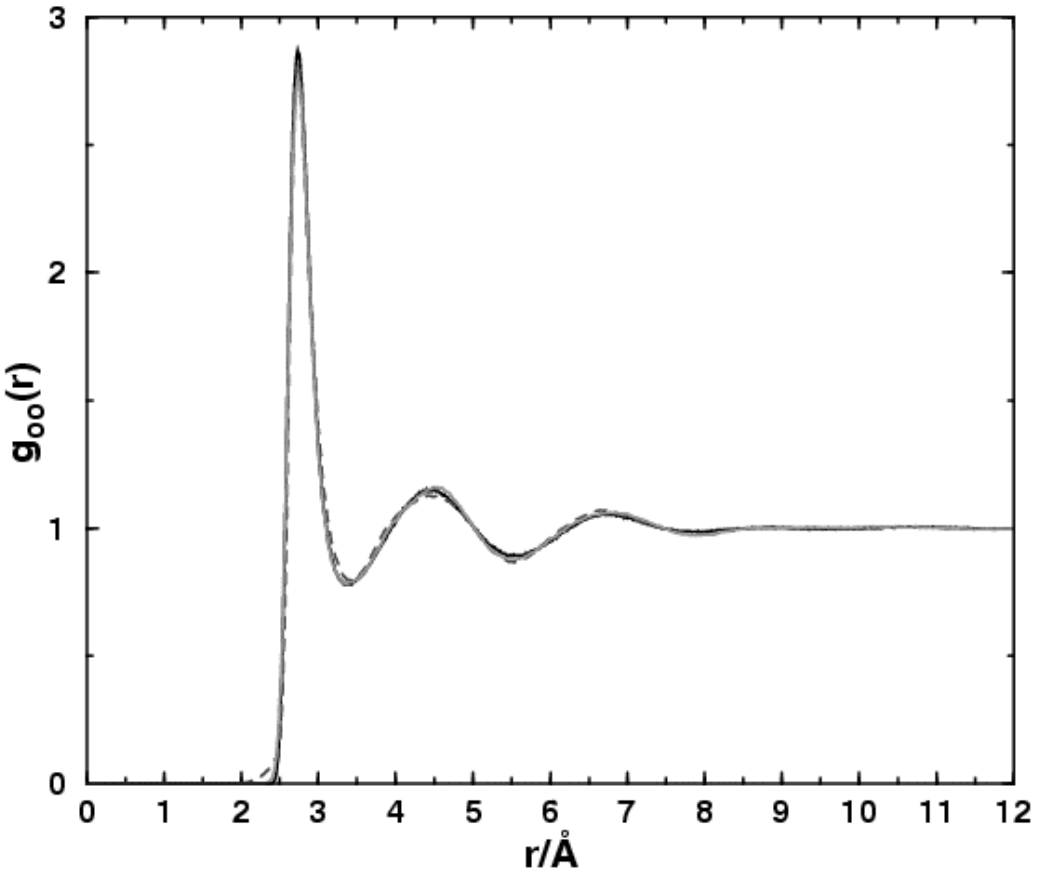



Figure 1. Hura et al.

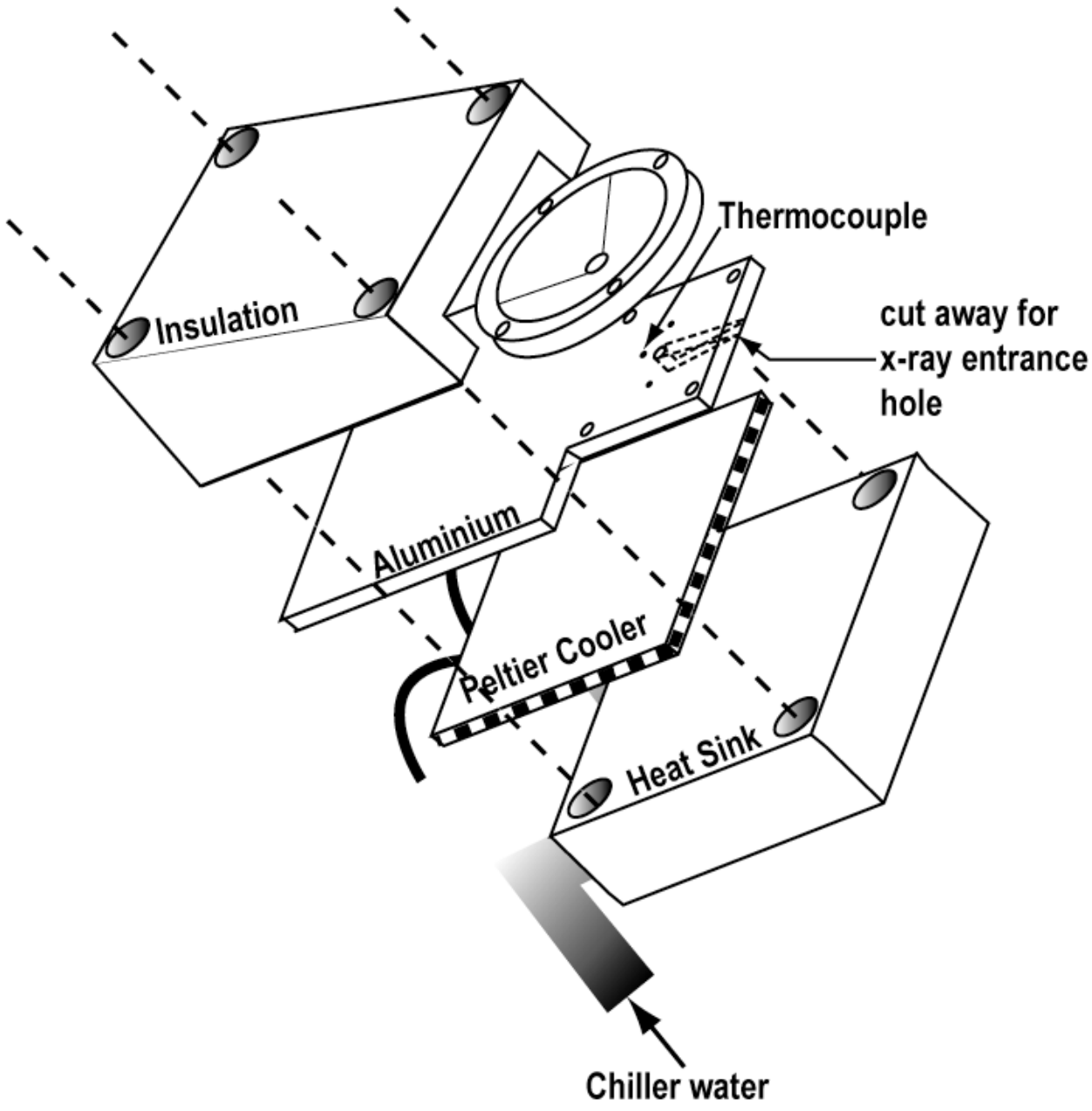



Figure 2. Hura et al.

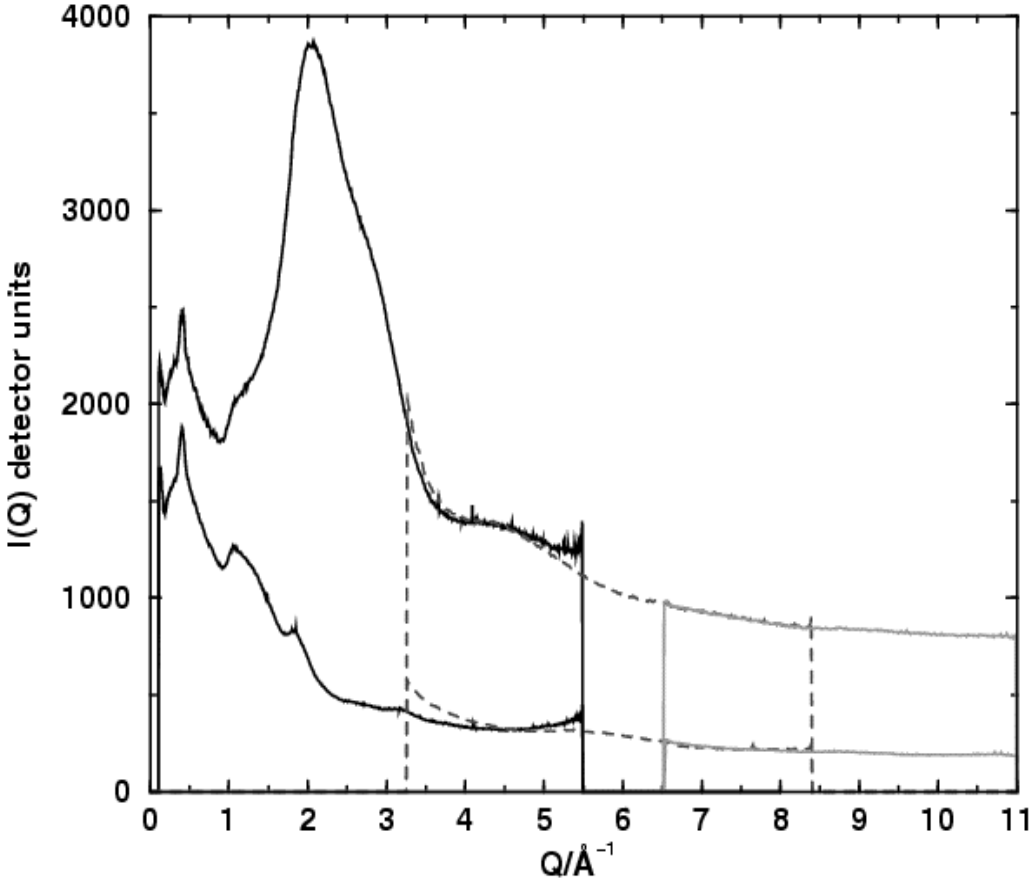



Figure 3a. Hura et al.

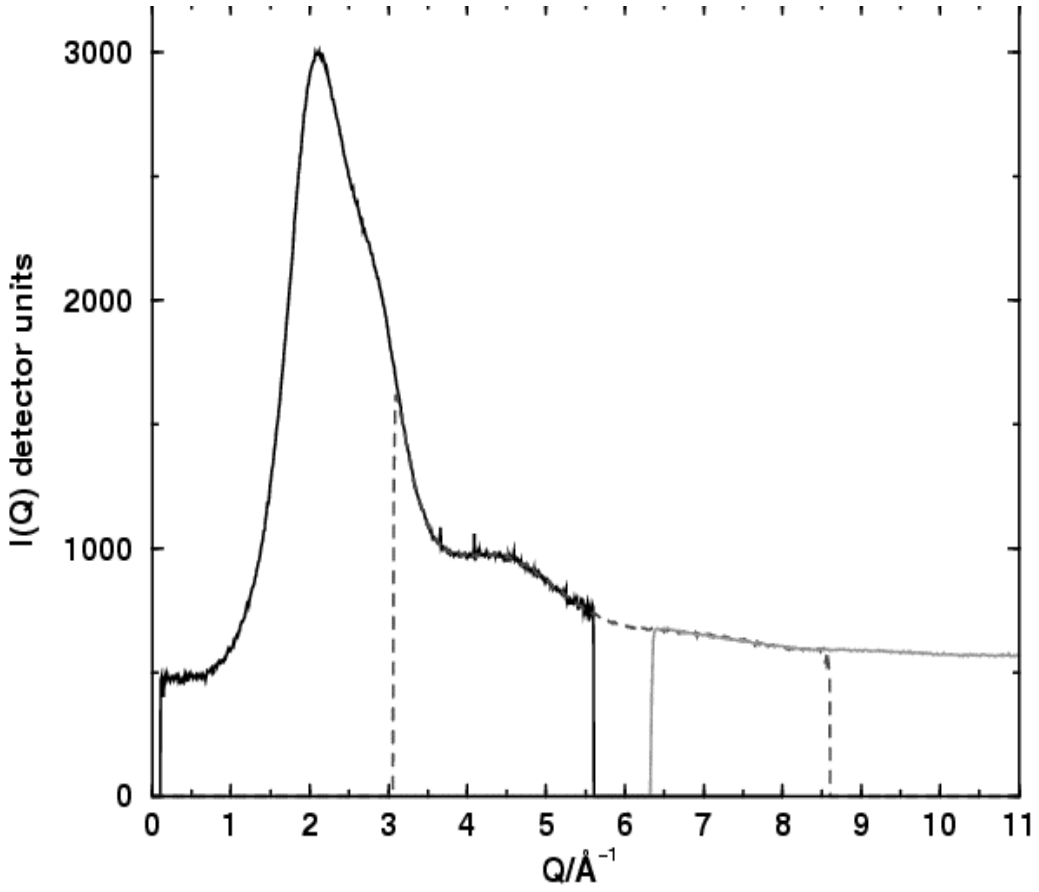



Figure 3b. Hura et al.

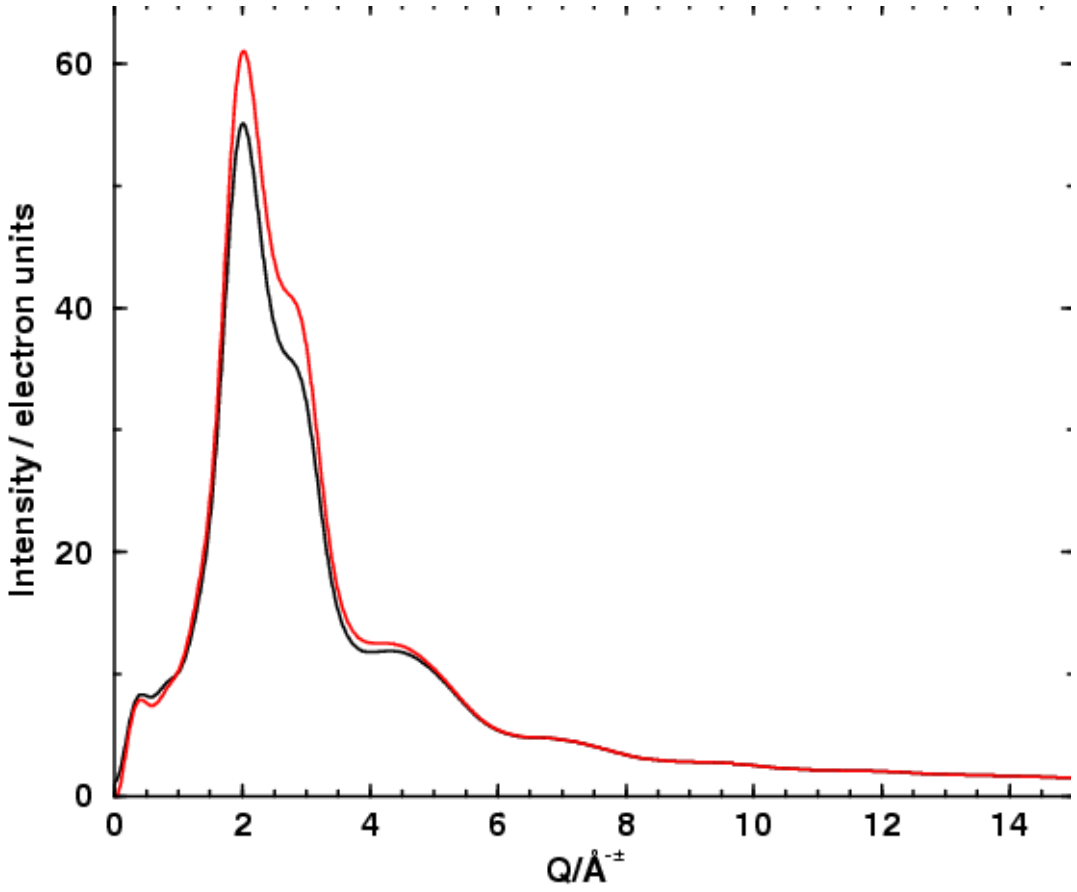



Figure 4. Hura et al.

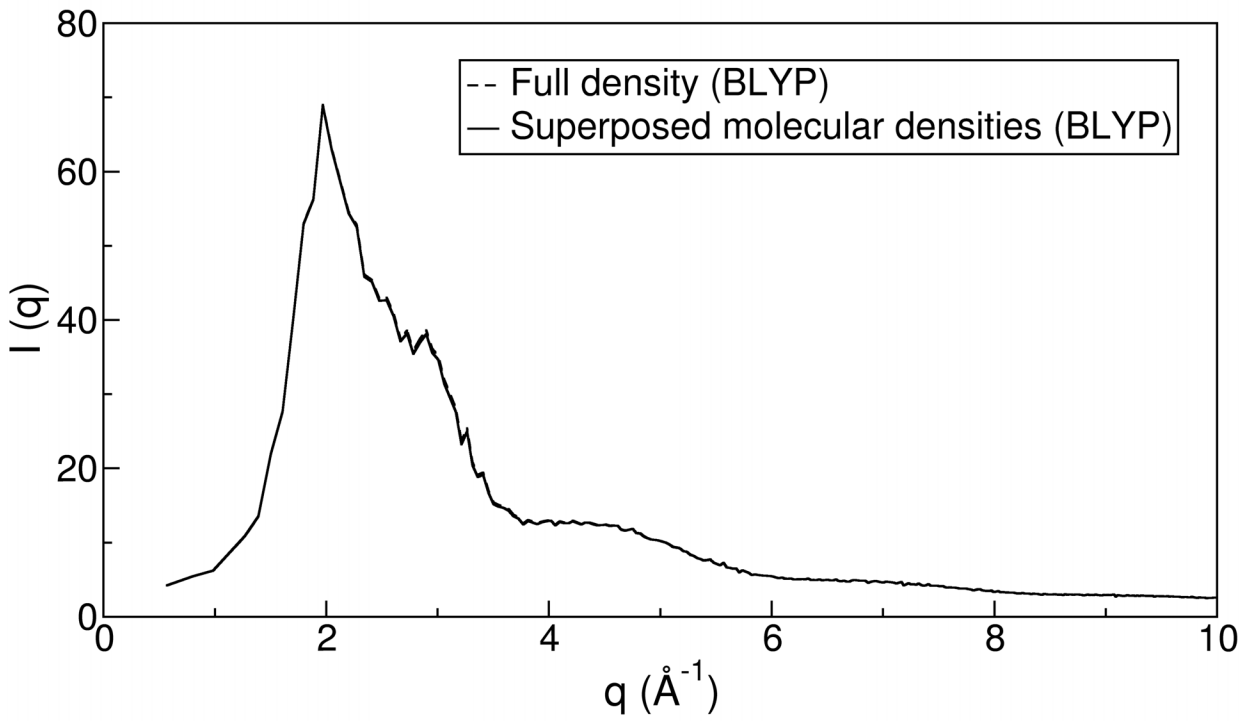



Figure 5a. Hura et al.

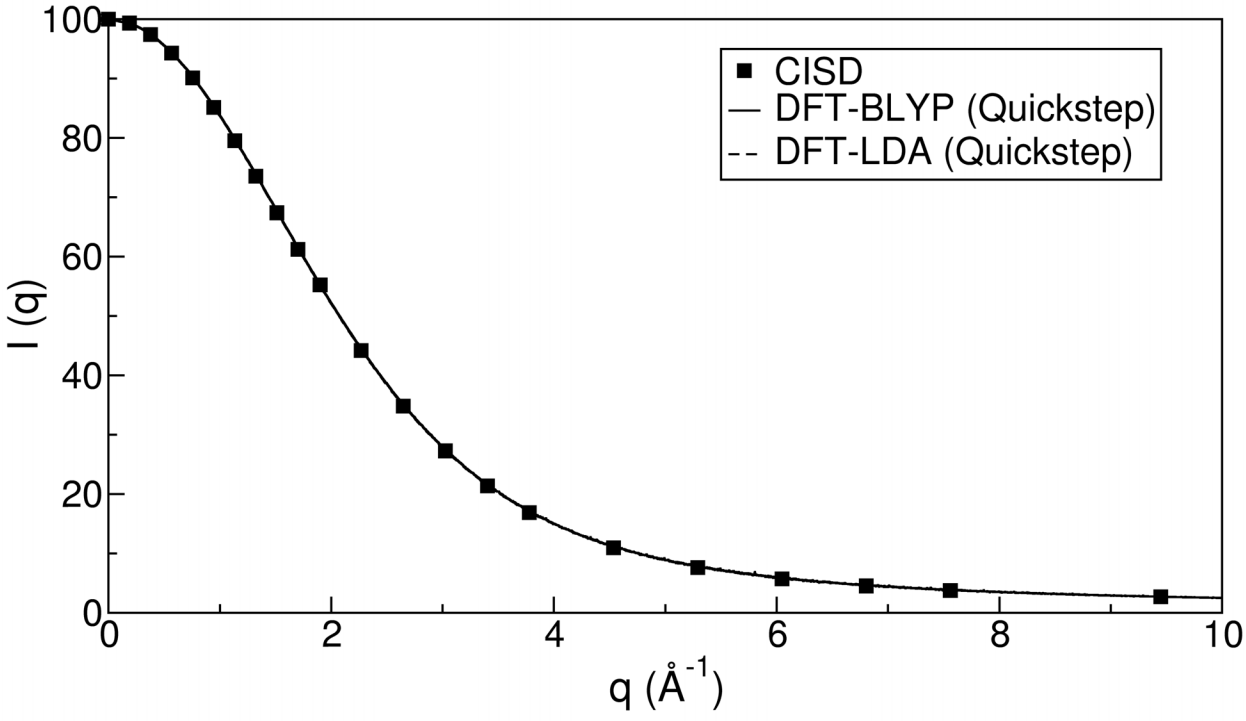



Figure 5b. Hura et al.

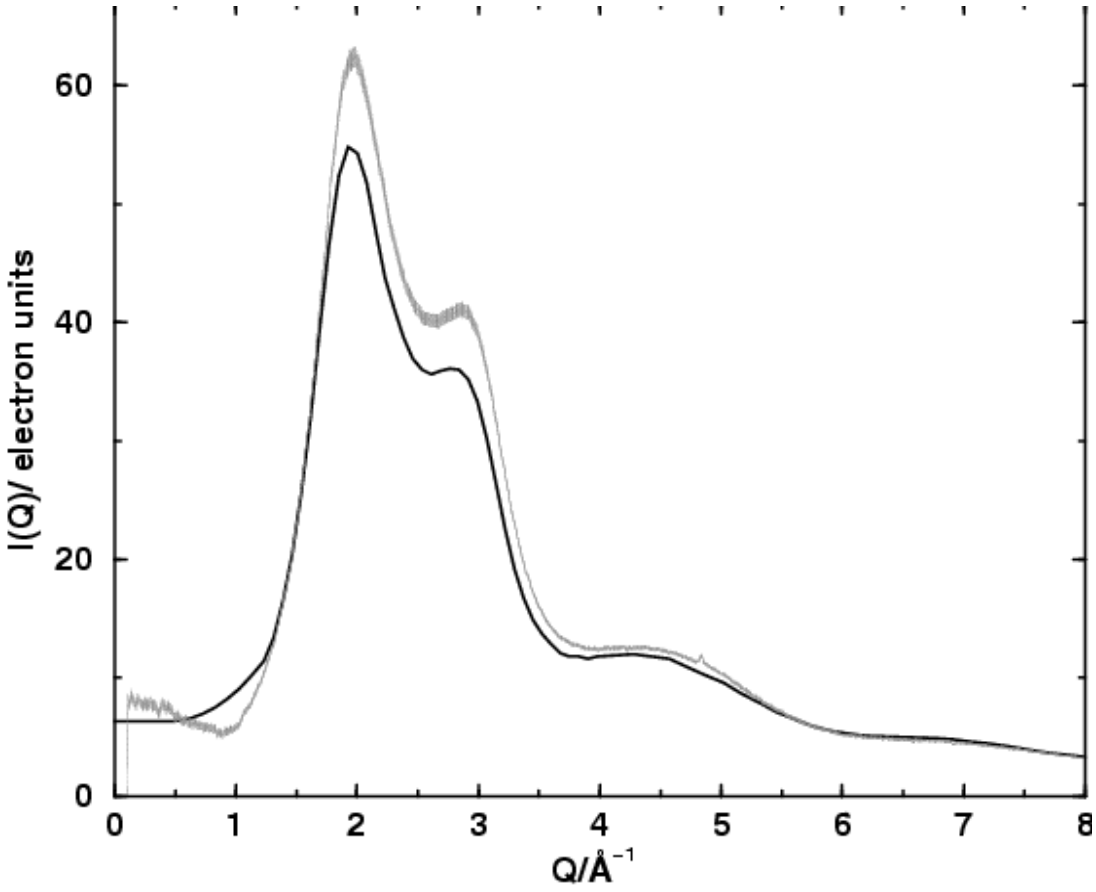



Figure 6a. Hura et al.

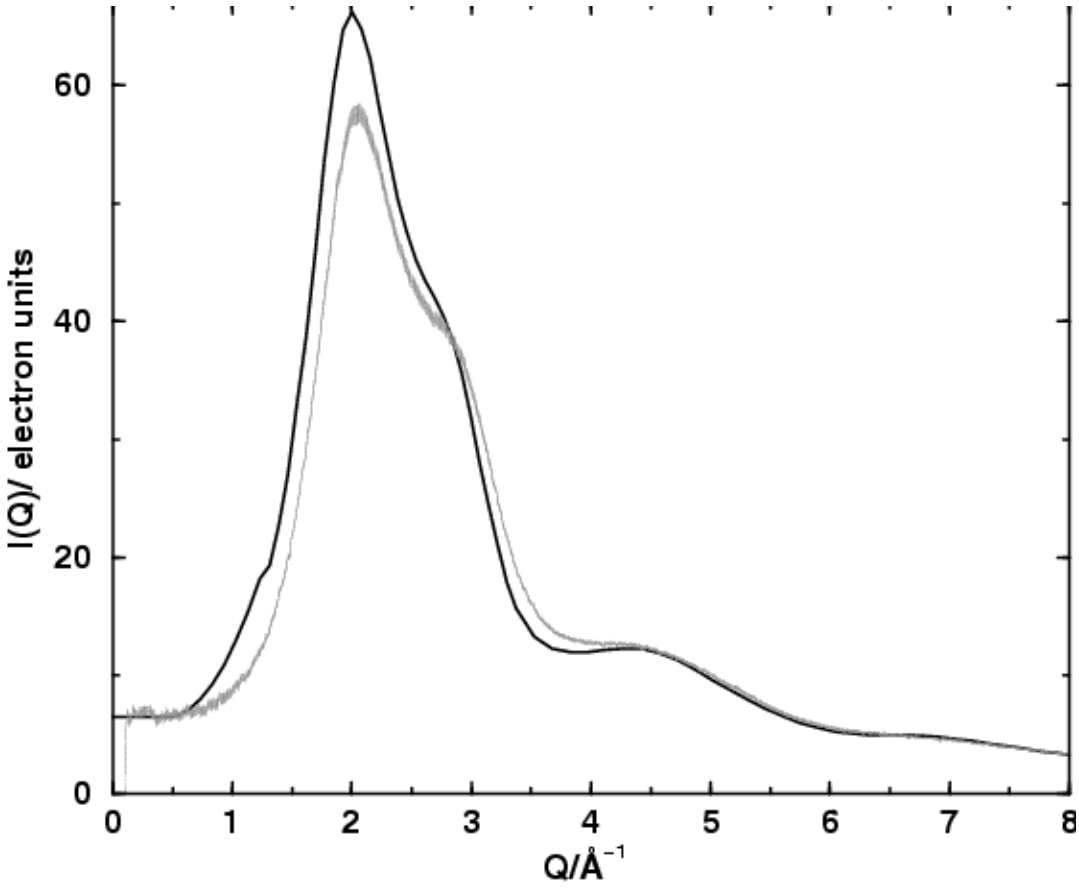



Figure 6b. Hura et al.

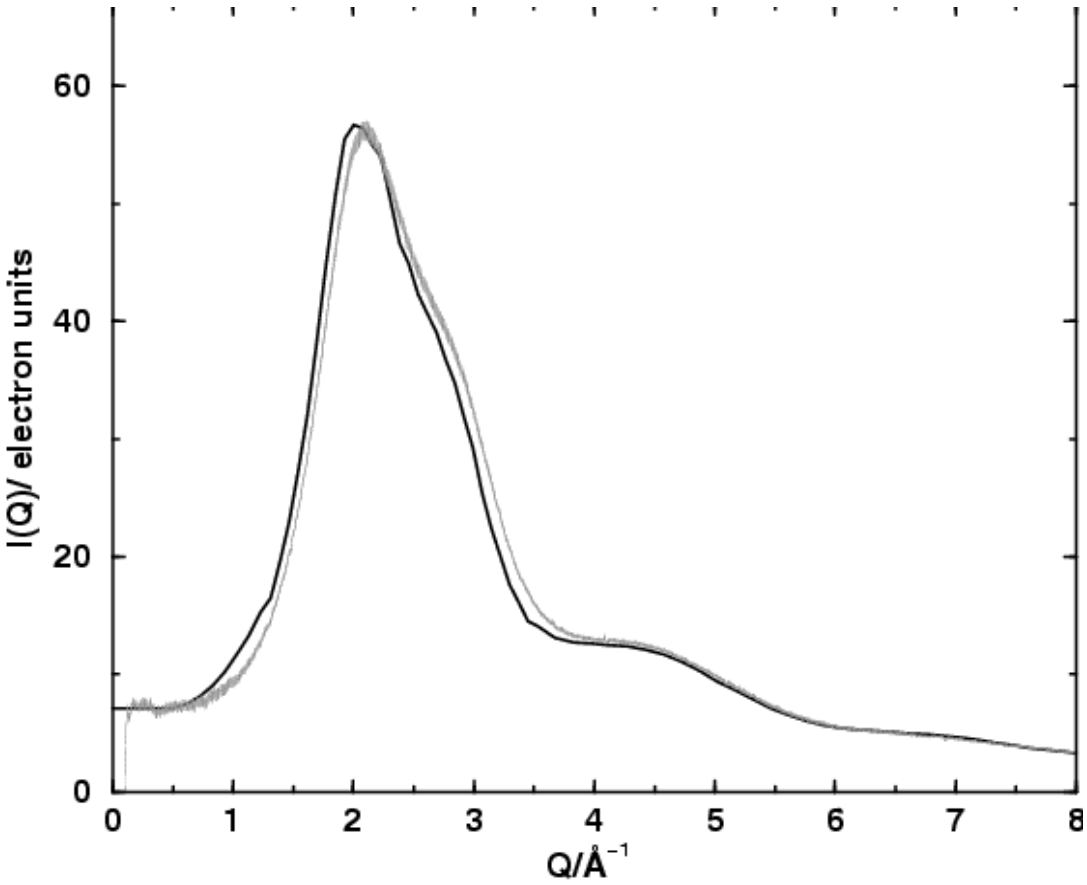



Figure 6c. Hura et al.

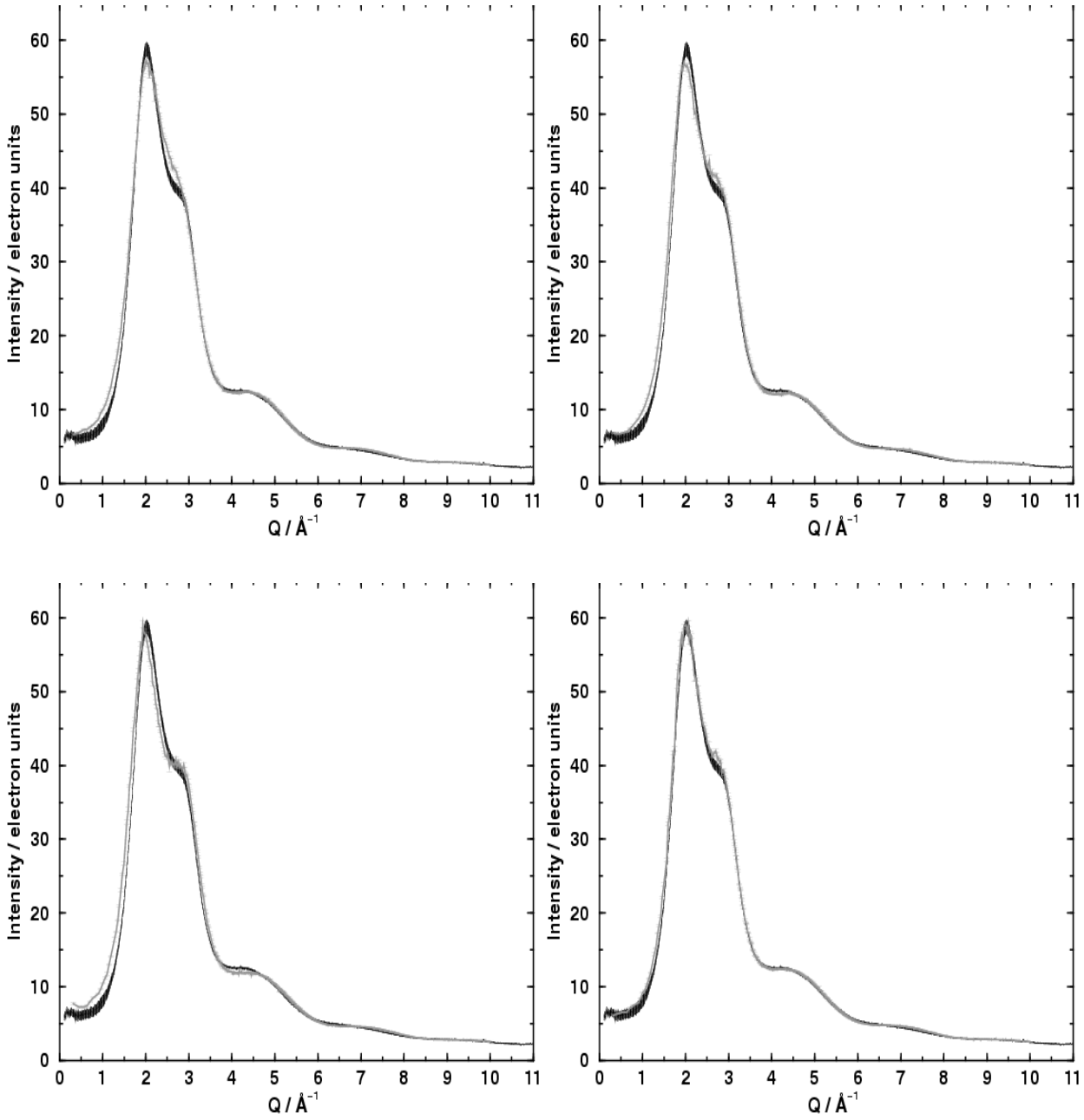



Figure 7. Hura et al.

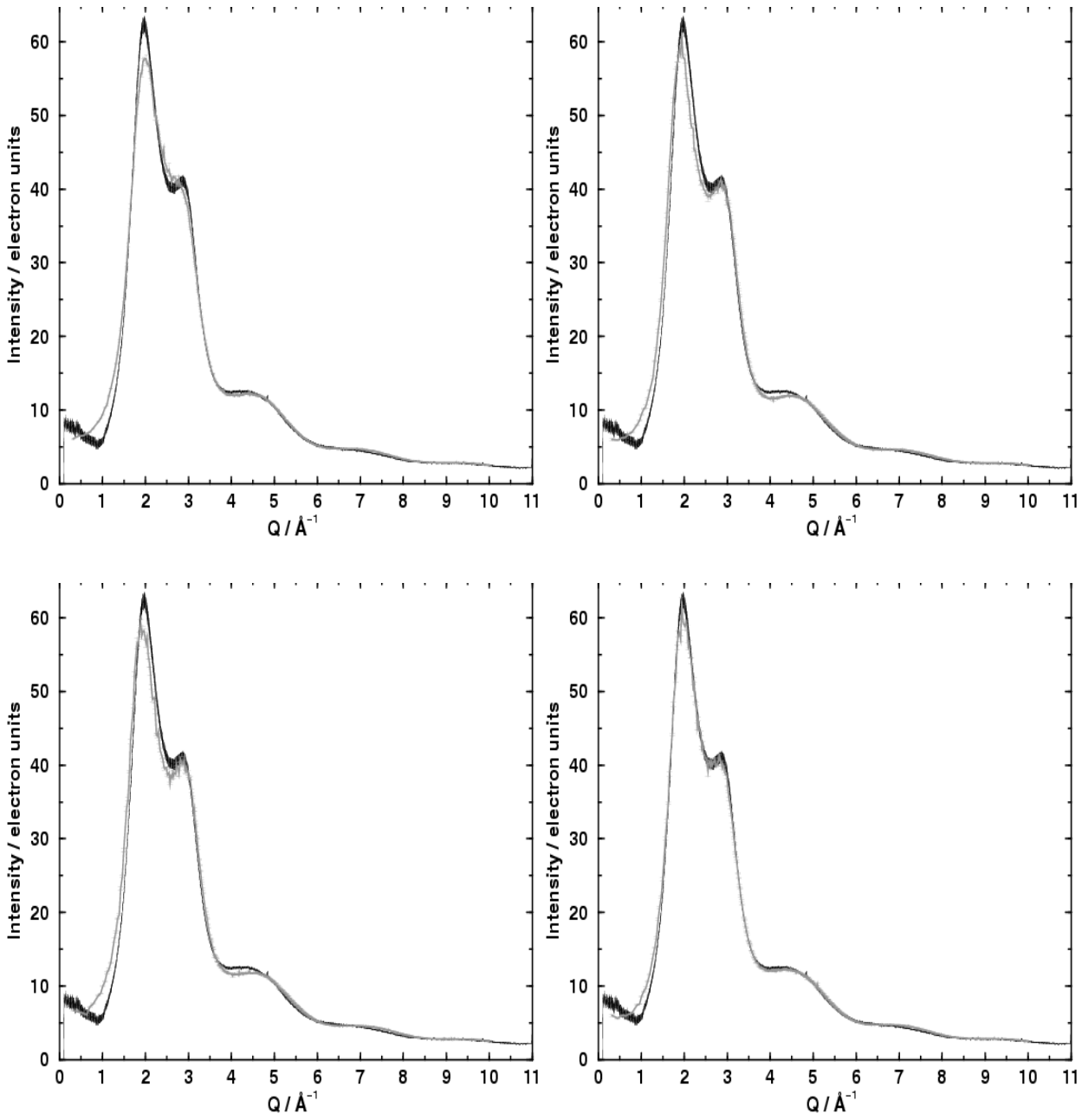



Figure 8. Hura et al.

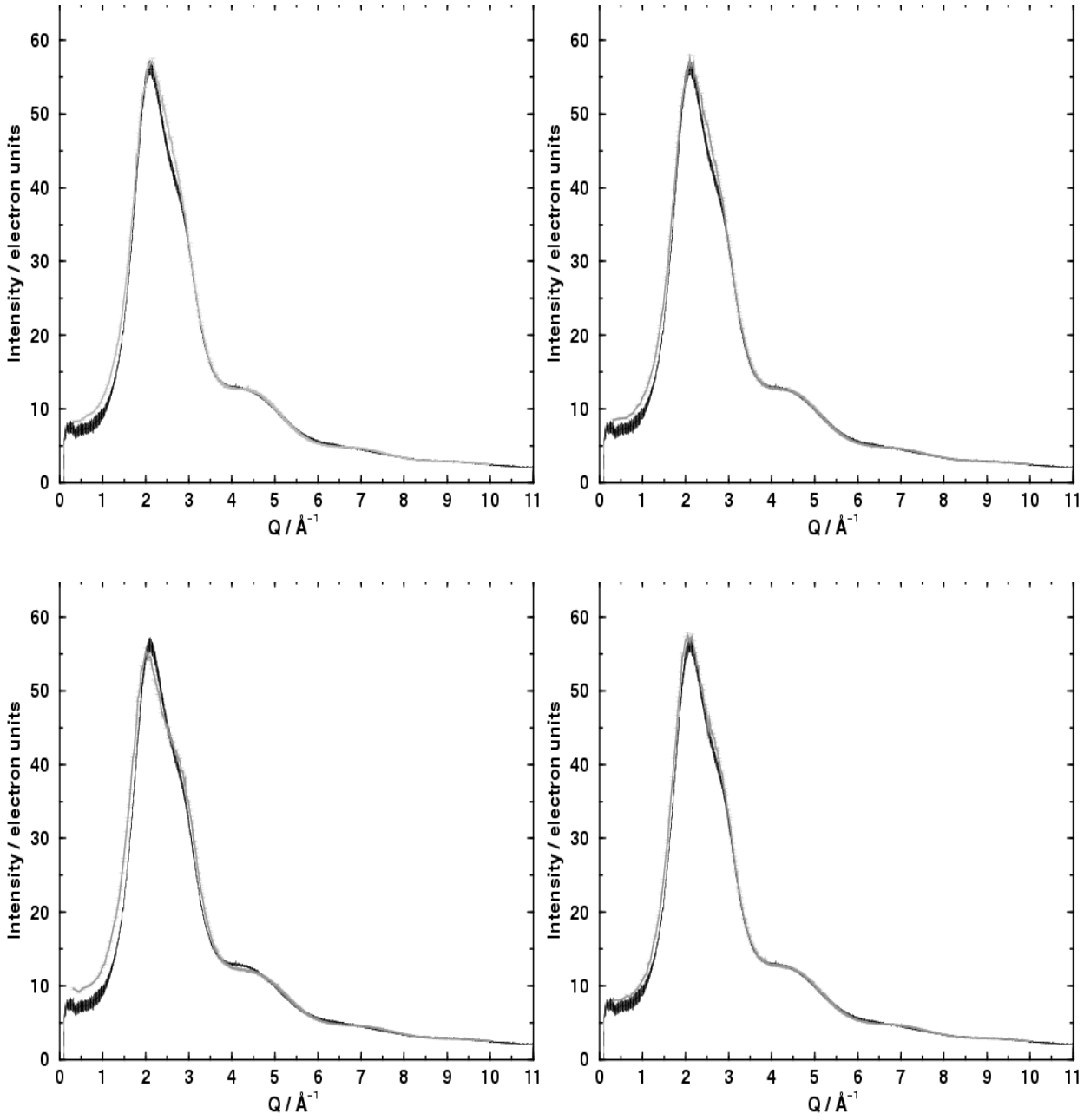



Figure 9. Hura et al.

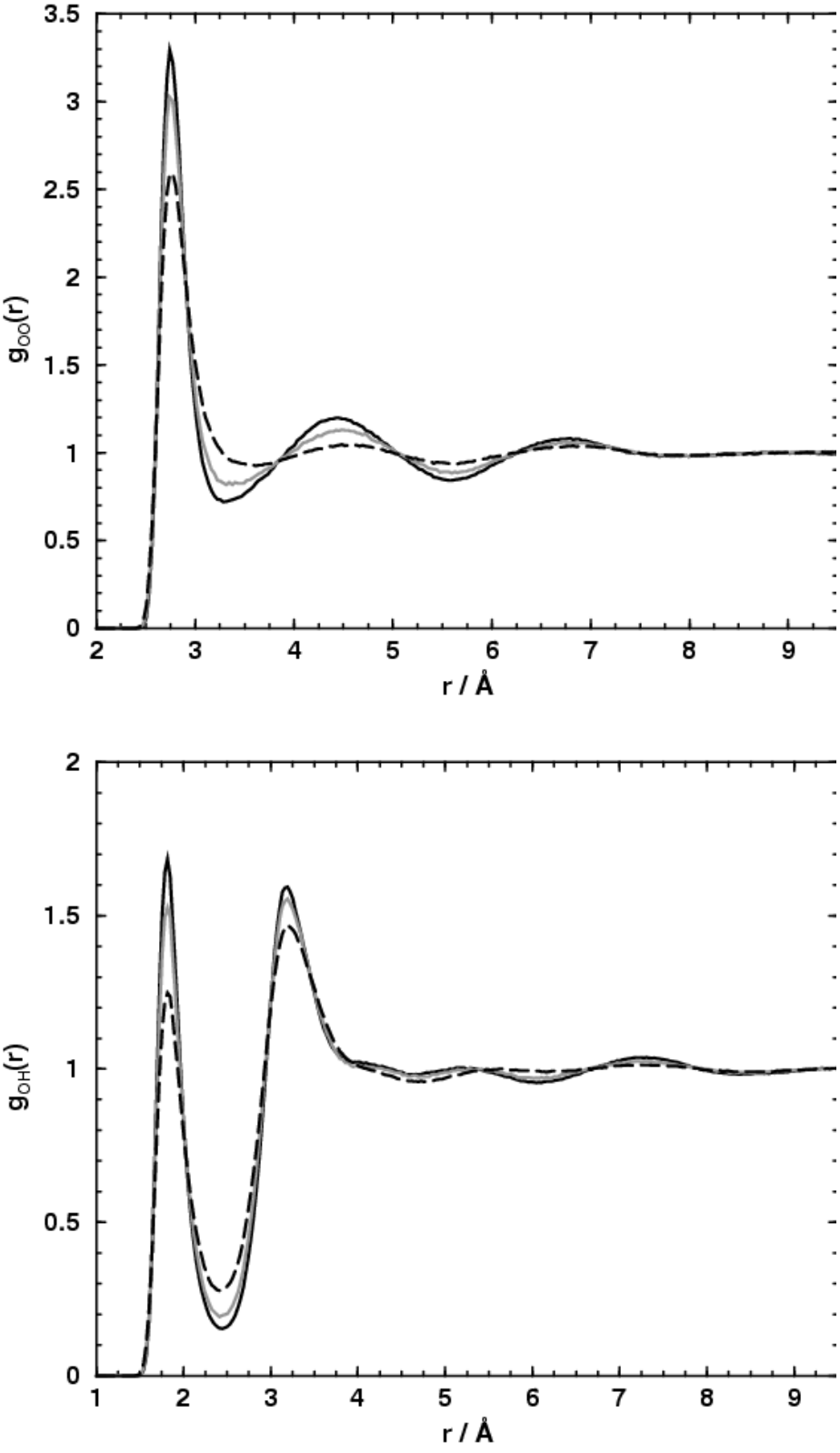



Figure 10. Hura et al.

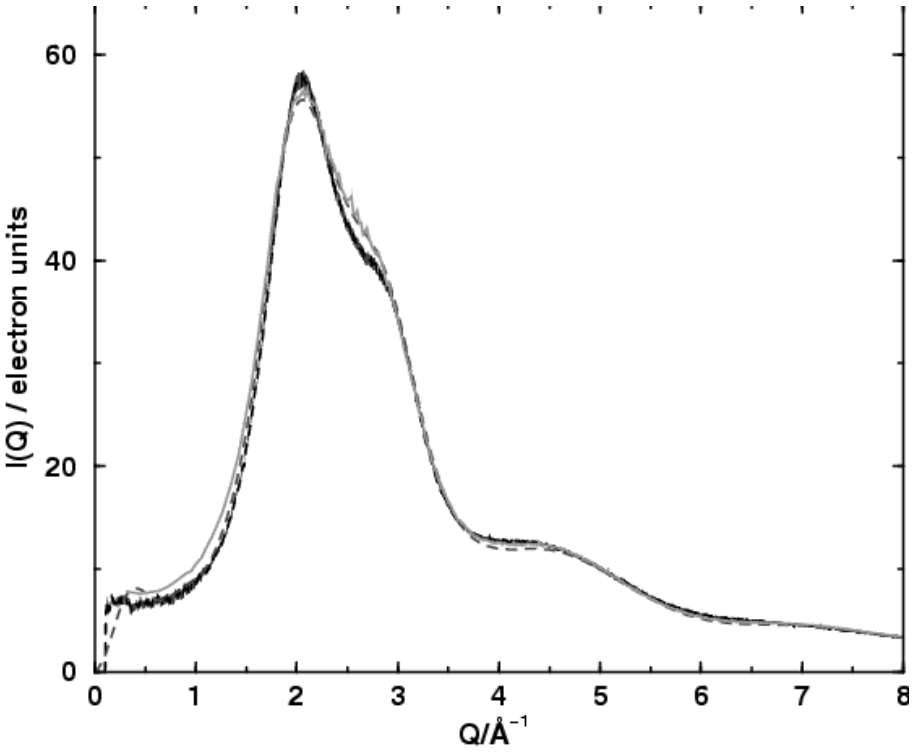



Figure 11. Hura et al.

## References


1. Hura G, Sorenson JM, Glaeser RM, Head-Gordon T: **A high-quality x-ray scattering experiment on liquid water at ambient conditions**. *Journal of Chemical Physics* 2000, **V113**:9140-9148.

2. Sorenson JM, Hura G, Glaeser RM, Head-Gordon T: **What can x-ray scattering tell us about the radial distribution functions of water?** *Journal of Chemical Physics* 2000, **V113**:9149-9161.

3. Head-Gordon T, Hura G: **Water structure from scattering experiments and simulation**. *Chemical Reviews* 2002, **V102**:2651-2669.

4. Narten AH, Levy HA: **Liquid water: molecular correlation functions from X-ray diffraction**. *Journal of Chemical Physics* 1971, **55**:2263-2269.

5. Narten AH, Levy HA: **Liquid Water: Scattering of X-rays**. In *Water A Comprehensive Treatise*. Edited by Franks F: Plenum Press; 1972:311-332. vol 1.]

6. Narten AH, Thiessen WE: **Atom pair distribution functions of liquid water at 25 degrees C from neutron diffraction**. *Science* 1982, **217**:1033-1034.

7. Narten AH, Venkatesh CG, Rice SA: **Diffraction pattern and structure of amorphous solid water at 10 and 77K**. *Journal of Chemical Physics* 1976, **64**:1106-1121.

8. Soper AK, Silver RN: **Hydrogen-hydrogen pair correlation function in liquid water**. *Physical Review Letters* 1982, **49**:471-474.

9. Gorbaty YE, Demianets YN: **An X-ray study of the effect of pressure on the structure of liquid water**. *Molecular Physics* 1985, **55**:571-588.

10. Soper AK, Phillips MG: **A new determination of the structure of water at 25 degrees C**. *Chemical Physics* 1986, **107**:47-60.

11. Dore JC: **Structural studies of water and other hydrogen-bonded liquids by neutron diffraction**. In *Hydrogen Bond Physics International Conference*; *Barga, Italy*: 1991:193-211.

12. Soper AK, Bruni F, Ricci MA: **Site-site pair correlation functions of water from 25 to 400 degrees C: Revised analysis of new and old diffraction data**. *Journal of Chemical Physics* 1997, **V106**:247-254.

13. Dore J: **Structural studies of water in confined geometry by neutron diffraction**. *Chemical Physics* 2000, **V258**:327-347.

14. Soper AK: **The radial distribution functions of water and ice from 220 to 673 K and at pressures up to 400 MPa**. *Chemical Physics* 2000, **V258**:121-137.

15. Soper AK: **Empirical potential Monte Carlo simulation of fluid structure**. *Chemical Physics* 1996, **202**:295-306.

16. McGreevy RL, Puszati L: **Reverse Monte Carlo Simulation: a new technique for the determination of disordered structures**. *Molecular Simulation* 1988, **1**:359.

17. McGreevy RL: **Rmc - Progress, Problems and Prospects**. *Nuclear Instruments & Methods in Physics Research Section a-Accelerators Spectrometers Detectors and Associated Equipment* 1995, **V354**:1-16.

18. Car R, Parrinello M: **Unified approach for molecular dynamics and density-functional theory**. *Physical Review Letters* 1985, **55**:2471-2474.

19. Sprik M, Hutter J, Parrinello M: **Ab Initio Molecular Dynamics Simulation of Liquid Water - Comparison Three Gradient-Corrected Density Functionals**. *Journal of Chemical Physics* 1996, **V105**:1142-1152.

20. Silvestrelli PL, Parrinello M: **Structural, electronic, and bonding properties of liquid water from first principles**. *Journal of Chemical Physics* 1999, **V111**:3572-3580.

21. Schwegler E, Galli G, Gygi F: **Water under pressure**. *Physical Review Letters* 2000, **V84**:2429-2432.

22. Krack M, Gambirasio A, Parrinello M: **Ab initio x-ray scattering of liquid water**. *Journal of Chemical Physics* 2002, **V117**:9409-9412.

23. Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML: **Comparison of simple potential functions for simulating liquid water**. *Journal of Chemical Physics* 1983, **79**:926-935.

24. Mahoney MW, Jorgensen WL: **A five-site model for liquid water and the reproduction of the density anomaly by rigid, nonpolarizable potential functions**. *Journal of Chemical Physics* 2000, **V112**:8910-8922.

25. Chen B, Xing JH, Siepmann JI: **Development of polarizable water force fields for phase equilibrium calculations**. *Journal of Physical Chemistry B* 2000, **V104**:2391-2401.

26. Henke BL, Gullikson EM, Davis JC: **X-Ray Interactions - Photoabsorption, Scattering, Transmission and Reflection at E=50-30,000 Ev, Z=1-92 (Vol 54, Pg 181, 1993)**. *Atomic Data and Nuclear Data Tables* 1993, **V55**:349-349.

27. Wang JH, Tripathi AN, Smith VH: **Chemical Binding and Electron Correlation Effects in X-Ray and High Energy Electron Scattering**. *Journal of Chemical Physics* 1994, **V101**:4842-4854.

28. Watanabe N, Ten-no S, Pal S, Iwata S, Udagawa Y: **Size-extensive calculations of static structure factors from the coupled cluster singles and doubles model**. *Journal of Chemical Physics* 1999, **V111**:827-832.

29. Krogh-Moe J: *Acta Crystallographica* 1956, **9**:951.

30. Norman N: *Acta Crystallographica* 1957, **10**:370.

31. Rahman A, Stillinger FH: **Hydrogen-bond patterns in liquid water**. *Journal of the American Chemical Society* 1973, **95**:7943-7948.

32. Izvekov S, Voth GA: **Car-Parrinello molecular dynamics simulation of liquid water: New results**. *Journal of Chemical Physics* 2002, **V116**:10372-10376.

33. Lee C, Yang W, Parr RG: *Phys. Rev. B* 1988, **37**:785.

34. Becke AD: **Density-functional exchange-energy approximation with correct asymptotic behavior**. *Physical Review A (General Physics)* 1988, **38**:3098-3100.

35. Miehlich B, Savin A, Stoll H, Preuss H: *Chem. Phys. Lett.* 1989, **157**:200.

36. Perdew JP, Burke K, Ernzerhof M: **Generalized Gradient Approximation Made Simple**. *Physical Review Letters* 1996, **V77**:3865-3868.

37. Lippert G, Hutter J, Parrinello M: **The Gaussian and augmented-plane-wave density functional method for ab initio molecular dynamics simulations**. *Theoretical Chemistry Accounts* 1999, **103**:124-140.

38. Krack M, Parrinello M: **All-electron ab-initio molecular dynamics**. *Physical Chemistry Chemical Physics* 2000, **2**:2105-2112.

39. Chialvo AA, Cummings PT: **Engineering a Simple Polarizable Model for the Molecular Simulation of Water Applicable over Wide Ranges of State Conditions**. *Journal of Chemical Physics* 1996, **V105**:8274-8281.

40. Matsuoka O, Clementi E, Yoshimine M: *Journal of Chemical Physics* 1976, **64**:1351.

41. Corongiu G, Clementi E: **Liquid Water with an Abinitio Potential - X-Ray and Neutron Scattering from 238-K to 368-K**. *Journal of Chemical Physics* 1992, **V97**:2030-2038.

42. Berendsen H, Postma J, van Gunsteren WF, Hermans J: **Intermolecular forces : proceedings of the Fourteenth Jerusalem Symposium on Quantum Chemistry and Biochemistry held in Jerusalem, Israel, April 13-16, 1981**. In *Intermolecular forces : proceedings of the Fourteenth Jerusalem Symposium on Quantum Chemistry and Biochemistry*. Edited by Pullman B: D. Reidel

Sold and distributed in the U.S.A. and Canada by Kluwer Boston; 1981:ix, 567. vol 14.]

43. Berendsen HJC, Grigera JR, Straatsma TP: **The missing term in effective pair potentials**. *Journal of Physical Chemistry* 1987, **91**:6269-6271.

44. Liu YP, Kim K, Berne BJ, Friesner RA, Rick SW: **Constructing ab initio force fields for molecular dynamics simulations**. *Journal of Chemical Physics* 1998, **V108**:4739-4755.

45. Rahman A, Stillinger F: *Journal of Chemical Physics* 1971, **55**:3336.

46. Head-Gordon T, Stillinger FH: **An Orientational Perturbation Theory for Pure Liquid Water**. *Journal of Chemical Physics* 1993, **V98**:3313-3327.

47. Yamada M, Mossa S, Stanley HE, Sciortino F: **Interplay between time-temperature transformation and the liquid-liquid phase transition in water - art. no. 195701**. *Physical Review Letters* 2002, **V8819**:U119-U122.

48. Stanley HE, Barbosa MC, Mossa S, Netz PA, Sciortino F, Starr FW, Yamada M: **Statistical physics and liquid water at negative pressures**. *Physica a-Statistical Mechanics and Its Applications* 2002, **V315**:281-289.

49. Lindahl E, Hess B, van der Spoel D: **GROMACS 3.0: a package for molecular simulation and trajectory analysis**. *Journal of Molecular Modeling* 2001, **V7**:306-317.

50. Pearlman DA, Case DA, Caldwell JW, Ross WS, Cheatham TE, Debolt S, Ferguson D, Seibel G, Kollman P: **Amber, a Package of Computer Programs for Applying**

**Molecular Mechanics, Normal Mode Analysis, Molecular Dynamics and Free Energy Calculations to Simulate the Structural and Energetic Properties of Molecules**. *Computer Physics Communications* 1995, **V91**:1-41.

51. Rick SW, Stuart SJ, Berne BJ: **Dynamical Fluctuating Charge Force Fields - Application to Liquid Water**. *Journal of Chemical Physics* 1994, **V101**:6141-6156.

52. Rick SW: **Simulations of ice and liquid water over a range of temperatures using the fluctuating charge model**. *Journal of Chemical Physics* 2001, **V114**:2276-2283.