Diffraction and IR/Raman data do not prove tetrahedral water

Mikael Leetmaa,[1] Kjartan Thor Wikfeldt,[1] Mathias P. Ljungberg,[1] Michael Odelius,[1] Jan Swenson,[2] Anders Nilsson,[1,3] and Lars G. M. Pettersson[1,a)]
[1]*FYSIKUM, Stockholm University, AlbaNova University Center, SE-106 91 Stockholm, Sweden*
[2]*Department of Applied Physics, Chalmers University of Technology, SE-412 96 Göteborg, Sweden*
[3]*Stanford Synchrotron Radiation Laboratory, P.O. Box 20450, Stanford, California 94309, USA*

(Received 21 February 2008; accepted 16 July 2008; published online 25 August 2008)

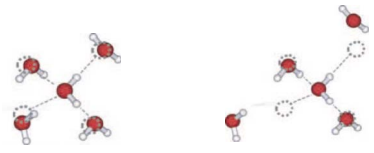
We use the reverse Monte Carlo modeling technique to fit two extreme structure models for water to available x-ray and neutron diffraction data in q space as well as to the electric field distribution as a representation of the OH stretch Raman spectrum of dilue HOD in D_2O; the internal geometries were fitted to a quantum distribution. Forcing the fit to maximize the number of hydrogen (H) bonds results in a tetrahedral model with 74% double H-bond donors (DD) and 21% single donors (SD). Maximizing instead the number of SD species gives 81% SD and 18% DD, while still reproducing the experimental data and losing only 0.7–1.8 kJ/mole interaction energy. By decomposing the simulated Raman spectrum we can relate the models to the observed ultrafast frequency shifts in recent pump-probe measurements. Within the tetrahedral DD structure model the assumed connection between spectrum position and H-bonding indicates ultrafast dynamics in terms of breaking and reforming H bonds while in the strongly distorted model the observed frequency shifts do not necessarily imply H-bond changes. Both pictures are equally valid based on present diffraction and vibrational experimental data. There is thus no strict proof of tetrahedral water based on these data. We also note that the tetrahedral structure model must, to fit diffraction data, be less structured than most models obtained from molecular dynamics simulations. © *2008 American Institute of Physics*. [DOI: 10.1063/1.2968550]

I. INTRODUCTION

There is at present a debate around the structure of liquid water initiated by the x-ray absorption (XAS) study of Wernet *et al.*[1] suggesting that a majority of the molecules in the liquid experience a distorted, asymmetric hydrogen (H) bonding environment with only one strong donating H-bond per molecule (see Fig. 1). This result is at strong variance with the established picture of a continuous distribution of structures around a mainly tetrahedral network of H bonds,[2,3] which, furthermore, are viewed as broken only fleetingly.[4,5] The basis for this traditional picture is a mass of experimental data from x-ray and neutron diffraction studies, infrared (IR) and Raman spectroscopy, NMR studies, recent ultrafast pump-probe two-dimensional (2D) vibrational spectroscopies, as well as molecular dynamics (MD) simulations using all established force fields or *ab initio* quantum mechanical forces. It is thus a bold challenge to the established picture that is suggested by the application of XAS to liquid water by Wernet *et al.*[1]

Nevertheless this study has provoked several critical evaluations of the available experimental data to investigate whether an asymmetrical H-bond model could indeed be accommodated. Soper[6] used the empirical potential structure refinement (EPSR) technique, with starting potentials having either symmetrical or asymmetrical charges on the hydrogens and found equal goodness of fit to x-ray and neutron diffraction data in q space for all models tested.[7] This was criticized by Head-Gordon and Johnson[8,9] who went on to show that artificially changing the charges in MD force fields to give asymmetric hydrogens not too unexpectedly leads to unphysical properties in the simulations.[10] The most asymmetrical structure model obtained from the EPSR fit was furthermore shown to generate a bimodal distribution[11,12] in the electric field (E-field) distribution, which has been proposed to provide a representation of the vibrational OH stretch Raman spectrum.[13,14] Indeed, details in the initial fits by Soper prompted further study using EPSR noting larger uncertainties than previously realized in the experimentally derived pair correlation functions (PCFs) and that softer than expected potentials were required in the EPSR procedure to reproduce the data.[15] However, no reason was found to question the established tetrahedral structure model of water.

In all of these studies assumptions about the forces act-



Tetrahedral Double Donor (DD) Asymmetrical Single Donor (SD)

FIG. 1. (Color online) Schematic illustration of two different classes of water species with different hydrogen bond distortions from the ice tetrahedral structure. Left: the distortions of the hydrogen bonds are relatively small and the tetrahedral symmetry can be approximated as maintained, denoted double donor (DD); the two hydrogen atoms are both involved in donating hydrogen bonds. Right: the distortions are asymmetrical where one donor hydrogen bond is intact and the other is highly distorted, denoted single donor (SD).

a)Author to whom correspondence should be addressed. Electronic mail: lgm@physto.se.

ing on the molecules and the dynamics were included in the analysis of the experimental data. Indeed, in the EPSR approach as much physics as possible is intentionally built in through the starting potential in order to guide the analysis to a physically sound structural solution.[15,16] This, however, introduces an uncertainty since the assumed initial potential and dynamics may introduce a bias, as is evident from Ref. 7, where different final solutions were obtained for different initial potentials. An alternative is represented by reverse Monte Carlo (RMC) modeling,[17] which takes the opposite approach, i.e., relying only on the experimental data (q space for diffraction data) and, if required, introducing specific geometrical constraints to conform to known physics if not enforced by the data; RMC thus allows an analysis *independent of MD force fields* producing a maximally disordered structure consistent with the data and constraints, but not necessarily consistent with any thermodynamical distribution.[18] A strength of the RMC approach is that any observable that can be derived from the atom positions in the simulation box can be included in the fit. Constraints may furthermore be applied to drive the fit toward specific structure models to evaluate their compatibility with the experimental data, but in contrast to the EPSR technique no interaction potential or information on dynamics is obtained from the RMC fitting. An excellent account of the RMC methodology is given in Ref. 19.

Most evaluations of water models are, in terms of structure, based on comparison of radial PCFs with experimental data. However, the diffraction data are obtained in q space and are not directly transformable into r-space PCFs through a simple inversion procedure. The most appropriate evaluation of different structure models is therefore to compare directly with the q-space diffraction data without any attempted inversion into r space. If an acceptable representation of the diffraction data in q space has been obtained for a particular structure model, then the different PCFs can be derived from positions of the atoms in the model. However, there is no guarantee that the structure model represents a unique solution to the diffraction data in q space.

Since the inversion of the data into an average structure or effective interaction potential does not have a unique solution, as we will further underline in the present work for the structure and as previously shown for the potential by Soper,[7] we will here limit ourselves to investigating what bounds the data actually set to the fractions of tetrahedrally and asymmetrically H-bonded molecules in the liquid. We emphasize that the resulting structure models must be considered as neither unique nor necessarily correct.[19] Given that two structure models fit the data equally well within error bars, one cannot claim that one of the models is preferred over the other based on the data included in the fit. In the present case we thus use RMC simply as a tool to test structural motifs against experimental data and illustrate that diffraction and IR/Raman experiments can be interpreted in at least two very different ways.

RMC modeling of diffraction data on liquid water has a history. Jedlovszky *et al.*[20] showed in 1996 that the structure produced by RMC when fitting to PCFs from MD simulations was, in general, more disordered than the original MD structure. In a 1998 paper[21] they went on to study the H-bonding network in liquid water at different temperatures by performing RMC fits to the available experimentally derived (from neutron diffraction) PCFs. The authors illustrated the bad agreement of two MD potential models with the experimentally derived PCFs, a finding that we illustrate with even more clarity in the present work. Finally, Pusztai[22] used RMC to critically evaluate the quality of the then existing data sets, where in particular neutron diffraction data taken on samples with high hydrogen content were found very difficult to reproduce and thus likely to introduce errors in the generated PCFs.

Here we extend these earlier applications of RMC by simultaneously fitting the x-ray diffraction data set of Hura *et al.*[23] together with newly taken, high-quality neutron diffraction data on five isotopic mixtures of D_2O and H_2O (Ref. 15) as well as fitting to the E-field distribution representing the OH stretch vibrational Raman spectrum of dilute HOD in D_2O;[13,14] the resulting overall quality of the fits indicates the consistency of these data sets. In addition to these data sets we constrain the internal geometries to reproduce the zero-point vibrational distribution from simulations of the liquid with quantum effects included through path integral techniques.[24] We then introduce specific constraints on the H-bond geometries driving the fits toward one of two extreme structure models, either maximizing the number of molecules with two donating H bonds (double donor, DD) or maximizing the number of asymmetrically coordinated molecules (single donor, SD); the former is similar to the traditional tetrahedral model, but less structured than models typically obtained from MD simulations, while the latter is similar to the type of average structure originally proposed by Wernet *et al.*[1] We find that these two extreme structure models cannot be distinguished by the established experimental diffraction and IR/Raman data sets.

An important outcome of the present work is that the PCFs corresponding to the measured diffraction data are broader and less well defined than PCFs from MD simulations. This implies a significantly more distorted H-bond network in the liquid, but from the data it cannot be determined whether the distortions are symmetric (DD model) or asymmetric (SD model) or a mixture; this ambiguity is inherent in the scattering data.

It is therefore not possible to claim preference for either model based on the data included in these fits; thus neither diffraction nor IR/Raman can be regarded as proof for either tetrahedral or asymmetrical water structure models. Since the H bond depends on many-body correlations (in our H-bond definition a three-body correlation), our conclusions underline that diffraction data are unable to resolve higher-order correlations in a complex system with nonpairwise additive interactions such as water.[19] Even when IR/Raman, which is sensitive to higher-order correlations, is included in the fit, the combined power of all the included data sets is still insufficient to put strict restrictions on the H-bond populations.

II. METHODS

A. RMC fitting of diffraction data

For all the RMC fitting a locally modified version of the RMC++ code[25] was used. A detailed description of the implementation can be found in the online supplementary material.[26] The neutron diffraction data sets from Soper[15] include the isotopic mixtures 100%D_2O, 75%D_2O +25%H_2O, 50%D_2O+50%H_2O, 25%D_2O+75%H_2O, and 100%H_2O and were measured at 25 °C, while the x-ray diffraction (XD) data set of Hura *et al.*[23] was measured on H_2O at 27 °C. The respective q ranges used in the fits were 0.2–10.8 Å^{-1} for the x-ray data and 0.6–50.0 Å^{-1} for the neutron data [data points below 0.6 Å^{-1} in the neutron diffraction (ND) data sets were omitted in the RMC fits due to larger experimental uncertainties].

The XD and ND patterns are calculated in RMC via the partial structure factors (PSFs) obtained by Fourier transforming the PCFs:

$$A_{\alpha\beta}(q) = 1 + 4\pi\rho \int r^2 (g_{\alpha\beta}(r) - 1) \frac{(\sin(qr))}{qr} dr, \tag{1}$$

where $g_{\alpha\beta}(r)$ is the partial PCF for atoms of types α and β, and ρ is the atomic number density. The total structure factors are then formed as weighted combinations of the PSFs, with the partial weights given in the case of neutrons by the atomic scattering lengths and for x rays by the q-dependent form factors,

$$S_{\mathrm{ND}}(q) = \sum_{\alpha} \sum_{\beta} (2 - \delta_{\alpha\beta}) c_\alpha c_\beta \langle b_\alpha \rangle \langle b_\beta \rangle (A_{\alpha\beta}(q) - 1), \tag{2}$$

$$S_{\mathrm{XD}}(q) = \sum_{\alpha} \sum_{\beta} (2 - \delta_{\alpha\beta}) c_\alpha c_\beta f_\alpha(q) f_\beta(q) \times (A_{\alpha\beta}(q) - 1) / \left(\sum_i c_i f_i^2(q) \right), \tag{3}$$

where c_i is the concentration of atomic species i in the sample, the Kronecker delta avoids double counting, $f_i(q)$ are the atomic form factors, b_i is the atomic scattering lengths and the angular brackets denote the spin and isotopic average. RMC includes also the intramolecular distances in the PSFs, so the intramolecular scattering (excluding the atomic self-scattering) is included in the total structure factors above. The atomic scattering lengths are taken from standard tables.[27] Hydrogen and its isotope deuterium have widely different scattering lengths, a fortunate fact that facilitates the use of ND by isotopic substitution to separate out the PSFs in methods such as RMC.

The atomic form factors were calculated at the required q points by using an analytical five Gaussian fit[28] to the original Dirac–Fock calculation.[27] Since the formation of chemical bonds modifies the electronic structure of the atoms, with an accumulation of charge on the oxygen atom in the case of water, the atomic form factors were modified according to

$$f_i^{\mathrm{mod}}(q) = (1 + \alpha_i e^{-q^2/2\delta_i^2}) f_i(q), \tag{4}$$

as suggested in Ref. 29. The α parameter, which describes the charge transfer, was chosen to correspond to transfer of 0.4 electrons from each hydrogen atom to the oxygen atom, and the δ parameter, describing the diffuseness of the electron cloud, was set to its estimated gas phase value of 2.2.[29] These modified atomic form factors (MAFFs) retain their high-q behavior in accordance with the assumption that the core electrons are unaffected by the chemical bond. A marked improvement is expected from this procedure;[29] from here on we simply denote the MAFFs by $f(q)$.

Before the experimental XD intensity, $I(q)$, could be used in RMC, it had to be calibrated to an absolute scale and normalized to obtain the total structure factor, $S(q)$. This is performed in one step by the following standard procedure:

$$S_{\mathrm{XD}}(q) = \frac{kI(q) - I_{\mathrm{self}}(q)}{I_{\mathrm{norm}}(q)}, \tag{5}$$

with the independent atomic self-scattering given by $I_{\mathrm{self}}(q) = \Sigma c_i f_i^2(q)$ and the normalization factor by $I_{\mathrm{norm}}(q) = I_{\mathrm{self}}(q)$, corresponding to the single-atom scattering normalization procedure and consistent with the normalization used in RMC [Eq. (3)]. The constant k can be calculated either by requiring the high-q region of $S(q)$ to oscillate around zero or by the Norman integral method.[30] The integral method was applied since the q range used for the x-ray data set[8,23] was limited to q=0.2–10.8 Å^{-1}; this gave consistent results and a good fit in RMC. A single q point (q=9.85 Å^{-1}) was eliminated since it deviated significantly from the otherwise continuous curve; the resulting $S(q)$ was smoothed with a five-point smoothing formula to eliminate random noise in the data set.[31] Finally, $S(q)$ was taken smoothly to zero above q=10.0 Å^{-1} by multiplying with a Gaussian to avoid spurious truncation ripples in the $g_{\mathrm{OO}}(r)$ from RMC.

Note that fitting to the extracted $S_{\mathrm{XD}}(q)$ instead of $I(q)$ is merely a way to fit the important intermolecular scattering that is superposed on the dominating intramolecular scattering. One might, through Eq. (5), just as well fit to the experimental intensity $I(q)$, but this would wash away the information contained in the phases and amplitudes of the intermolecular scattering signal. The largest approximation that enters here is associated with the assumed atomic self-scattering, i.e., the MAFFs.

The RMC simulations were performed with a box length of 41.041 Å and 2304 water molecules (giving room temperature density of 0.997 g/cm^3, or equivalently 0.1 atom/Å^3) and using distances up to 20 Å in the Fourier transform of the $g(r)$ functions. The size of the simulation boxes ensured good statistics in the PCFs at an r-space resolution of 0.1 Å.

B. Raman spectrum

The IR/Raman spectrum of HOD in D_2O was modeled by the E-field from the surrounding water molecules on the hydrogen atom projected along the internal OH bond. Fecko *et al.*[14] have suggested a correlation between the projected E-field, using the SPC/E force field, and OH stretch frequen-

cies where the stretch is assumed to be a local uncoupled mode; this computational model has been calibrated against quantum chemical cluster calculations and applied extensively by Skinner and co-workers.[13,32–34]

In the present work the E-field was included as a data set in the RMC simulation. When the simulation was initialized the E-field was computed for each internal O–H bond at a point 0.975 Å from the oxygen and projected along the direction of the selected OH bond. The E-field from the solvated molecule itself was not included. In an earlier work[11] we have found that the calculated E-field converged rapidly with increasing cutoff radius and in the present work a cutoff of 20 Å was used. In the RMC steps the projected E-fields of individual water molecules were updated by adding the difference between the contribution from the moved water molecules before and after the move; in this way a full recalculation of the E-field was avoided which greatly speeded up the procedure. The E-fields were put into a histogram which was compared to a reference histogram obtained from a TIP4P-POL2 simulation[35] to get the χ^2 measure; SPC/E charges were always used to generate the field, however. In Ref. 33 a relation between E-field and IR/Raman spectrum was fitted for structures using SPC/E charges. The SPC/E charges are not supposed to give the real physical electric field but to serve as a representation of the structure; the projected E-field can then be viewed as a collective coordinate used in the *ab initio* frequency map.[13] In this sense the E-field from SPC/E charges can be used to model the IR/Raman spectrum for any structure.

Several weighting parameters, σ, were used in the fitting procedure in different regions of the E-field to be able to better control the fitting. To compensate for the relatively poor statistics in the E-field distribution due to the limited box size, a five point smoothing formula[31] was applied on the calculated E-field before comparing to the reference histogram. This procedure speeded up the fitting notably. In the final analysis of the E-field distributions the five point smoothing formula was not used, but instead a Fourier smoothing was applied that retained the ten lowest Fourier components of the distribution.

C. Generating the two structure models

To generate the two extreme structure models in the RMC fits geometrical cone criteria based on Refs. 1 and 36 were used, i.e., a donating H bond is considered to exist if the accepting oxygen is within a cone with apex on the oxygen of the H-bond donating molecule and centered around the selected O–H group. The cone is expressed as $R_{OO} \leq R_{OO}^{max} - 0.000\,44\theta^2$ where R_{OO} is the oxygen-oxygen distance and θ is the HO–O angle in degrees. Although several different cones were used to generate the two structure models (as is thoroughly described in the online supplementary material where also the generated structures are provided[26]) the final analysis in terms of H bonds and decomposition of the E-field was always done using the cone criterion from Ref. 1 having an $R_{OO}^{max} = 3.3$ Å. The maximally connected DD model of the liquid was created by requiring the program to maximize the number of donating H bonds in addition to the fit to the data; this resulted in an excellent fit for a final structure having 74% DD species complemented by 21% SD and the remainder non-donor (ND) species. For the strongly distorted SD model we instead required the program to fit the experimental data while simultaneously maximizing the number of SD species, i.e., molecules with only one well-defined donating H bond; the resulting distribution was then 81% SD and 18% DD, while still reproducing the experimental data sets. To prevent the asymmetric SD species from being dependent on the particular choice of cone definition the SD species were generated using two cones with $R_{OO}^{max} = 3.1$ and 3.3 Å, respectively. The H-bonded OH of a SD species should thus have an accepting oxygen within the smaller cone while the larger cone around the other OH should not include any accepting oxygen. Note that this imposed asymmetry between the weak and strong H bonds is smaller than previously used[7,8,36] although a more strict definition of broken H bonds is used here compared to Wernet *et al.*[1] See online supplementary material for a more detailed description of the fitting procedure.[26]

Although potential energy functions were not allowed to drive the fit, the consideration of individual contributions to the energy in the final structure, estimated using several force-field models, provided a useful diagnostic of unphysical local situations, which showed up for both structure models as very strongly repulsive interaction energies for some 1% of the 2304 molecules in the unit cell. These were found to correspond to cases where the fit failed to distinguish between intra- and intermolecular O–H correlations, i.e., forming H_2OH-OH moieties with O–O distance below 2.3 Å; constraining the allowed O–O distances to be larger than 2.35 Å eliminated these situations.

To control the internal geometry of the water molecules, distributions of internal O–H distances and internal H–O–H angles were implemented as additional data sets in the fitting. For both RMC runs the internal distances and angles were set to fit Gaussian distributions of approximately the same width as those obtained in path integral quantum simulations by Stern and Berne,[24] thus using a full width at half maximum of 0.15 Å for the distances and 21.6° for the angle distribution, centered at 0.98 Å and 105°, respectively. The distance and angle distributions were fitted independently of each other, i.e., no correlation between the two distributions was taken into account. To prevent too large, unphysical distortions of the internal geometries while still allowing for flexibility in the fitting procedure, smaller σ values (i.e., harder fit) were used on the edges of the distributions than for values closer to the mean geometry. The RMC fits to the internal distributions are shown in the online supplementary material.[26]

III. RESULTS

A. Diffraction

In Fig. 2 we present the resulting fits to the XD and ND data, where the x-ray data have been normalized to $S(q)$ using the modified form factors of Sorenson *et al.*[29] as described above. It is immediately apparent that the quality of the fit with respect to all diffraction data sets is excellent for

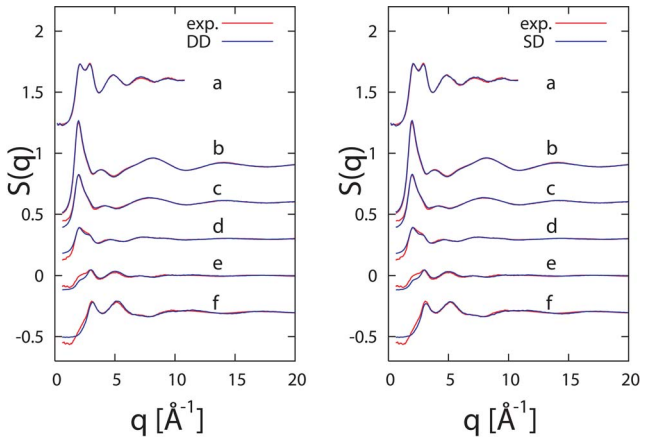


FIG. 2. (Color online) Resulting fits to diffraction data from simultaneous RMC fits to XD and ND data as well as E-field distribution using the DD model (left) and asymmetrical model (right). The diffraction data sets are (a) XD $S(q)$ as derived from the 27 °C data of Ref. 23 and [(b)–(f)] ND from Ref. 15 with (b) 100%D_2O, (c) 75%D_2O and 25%H_2O, (d) 50%D_2O and 50%H_2O, (e) 25%D_2O and 75%H_2O, and (f) 100%H_2O.

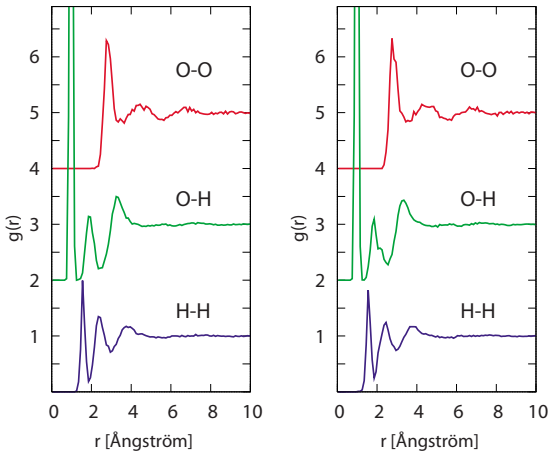


FIG. 3. (Color online) Partial O–O (top), O–H (middle), and H–H (bottom) PCFs for the DD (left) and SD (right) models.

both models in spite of their very different local coordination and resulting H-bond statistics. The simultaneous constraint to maximize the number of double H-bond-donor (DD) species to generate the DD model resulted in 74% DD species complemented by 21% SD and the remainder ND species, as described in Sec. II. Similarly, the asymmetric SD model maximized the number of SD species resulting in completely reversed statistics: 81% SD and 18% DD. In terms of average number of H bonds per molecule the DD model corresponds to 3.41 while the SD structure model has an average of 2.35 H bonds per molecule. Although the absolute values will depend somewhat on the specific H-bond definition (here that of Ref. 1) the large difference in H-bond count applying the same criterion to both models is significant.[26]

We note similar discrepancies in both cases in the representation of the ND data at low q for the samples with high H_2O content and very minor differences in the description of individual data sets; overall the descriptions are, however, equivalent. The two extreme models thus describe the diffraction data equally well and similar to the best fits using the EPSR technique.[15] In terms of diffraction experiments both an asymmetrically distorted structure model with a large fraction of broken H bonds and a symmetrically distorted, maximally coordinated structure model are thus equally valid, as is naturally then also a range of structure models between these two extremes, as well as, e.g., mixtures of the two extreme structure models.

Due to the mix of different data sets and internal and intermolecular geometrical constraints in the fitting procedure with weights (σ values) chosen to enforce stricter fitting in certain parameter regions the absolute measures of the goodness of the fits, i.e., the χ^2 values, carry little meaning. What is instead very clearly seen in Fig. 2 is that all features related to the structural content of the data, i.e., the positions, phases, and amplitudes of the peaks and minima in the diffraction data, are reproduced equally well by both RMC models. Neither of the two structure models can, however, be considered a true, verified average structure model of the liquid; the considered data sets do not give preference to any specific structure model inside the range of ~80%–20% DD and ~20%–80% SD species in terms of H-bond connectivity.

B. The two structure models

The generated structures represent two extreme structure models in terms of H-bond connectivity, which both reproduce the considered experimental data. Although these specific structure models thus should not be viewed as more than representatives of two classes (of probably many) that can reproduce the diffraction and vibrational data it is still important to demonstrate that the various cone definitions used to define them have actually generated two distinctly different structure models.

In Fig. 3 we show the PCFs resulting from the fit with the DD structure model on the left and SD structure model on the right. In this one-dimensional, angularly integrated representation of the structure there are only very minor differences to be seen between the two structure models, as expected since they both reproduce the XD and ND data to the same goodness of fit.

To illustrate the actual differences between the two structure models we show in Fig. 4(a) statistics on the distribution of H–O–O angles, where a value of zero corresponds to a straight, undistorted H bond. We observe for the SD structure model a somewhat bimodal distribution of angles, i.e., compared to the DD structure model both the number of well-defined, nearly straight bonds and angularly very distorted are enhanced in the SD structure model; the distribution of H-bond angles in the DD structure model forms an average of that of the SD structure model. In Fig. 4(b) we show the *difference* in H–O–O angles taken for the two hydrogens on the same molecule where we take the angle involving the nearest oxygen irrespective of whether it is H bonded or not. It is directly obvious that the distortions in the DD structure model are much more symmetrical, i.e., peaking around a difference of zero, than in the SD structure model for which the distribution of angular differences goes to significantly higher angles. In Fig. 4(c), finally, we show the radial asymmetry defined as the difference between the

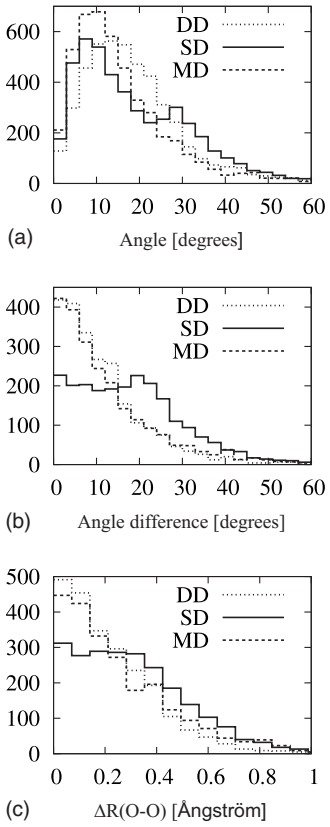


FIG. 4. Histograms characterizing structural parameters for the two different models with DD model (dotted) and SD model (solid line). Comparison is furthermore made against structures from a TIP4P-POL2 MD simulation (Ref. 35) (dashed). [(a), top] Distribution of angles (H–O–O) with 0° corresponding to a straight H bond, [(b), middle] difference between the H–O–O angles of the two hydrogens on the same molecule indicating the degree of angular asymmetry, and [(c), bottom] difference in shortest O–O distance on the donating side of each molecule indicating the degree of radial asymmetry.

shortest O_1-O_2 distances for each molecule, where O_1 is the oxygen in question and O_2 the nearest oxygen to the respective hydrogen. This again illustrates the asymmetry induced in the SD structure model by the constraints applied in the fitting. However, it should be clear that all these differences between the structure models correspond to essentially the same angularly integrated PCFs and can furthermore not be distinguished by the experimental diffraction data or the Raman data as will be discussed below.

Distributions from a room temperature TIP4P-POL2 MD simulation[35] are also shown in Fig. 4 (comparison between MD and diffraction data will be discussed in more detail in Sec. III C). The MD simulation has distributions qualitatively very similar to the DD structure model. While the H-bond distance distribution [Fig. 4(c)] and the asymmetry in H-bond angles [Fig. 4(b)] are nearly identical for the MD and the DD models, the distribution of H-bond angles [Fig. 4(a)] is sharper for the MD structure, indicating a somewhat less distorted, more icelike structure compared to the RMC generated DD structure model.

C. Comparison with MD simulations

Although the DD dominated structure model is meant to represent the traditional tetrahedral, average structure typically obtained from MD simulations the question is how well such simulations actually describe the data. The recent EPSR fit to the diffraction data made by Soper[15] resulted in a significantly softer potential than what is typically used in simulations. Since the O–O pair correlation is directly connected to the XD data it provides a critical comparison of different water models in r space.

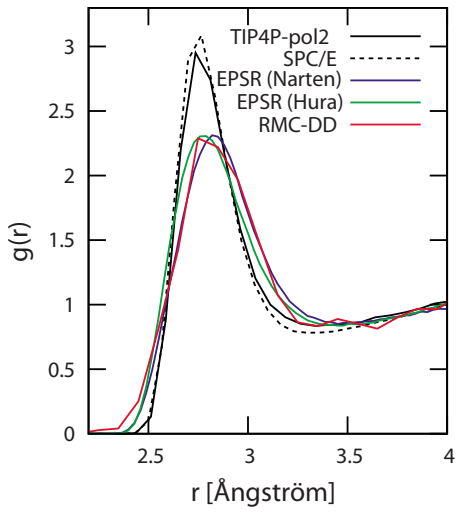


FIG. 5. (Color online) O–O PCFs derived from the RMC fitted DD model and EPSR fits performed by Soper (Ref. 15) of ND data together with either the Hura *et al.* (Ref. 23) or Narten and Levy (Ref. 37) data sets. In the figure we also show a comparison with structure models obtained from MD simulations using the SPC/E and TIP4P-POL2 (Ref. 35) force fields.

We focus in Fig. 5 on the first peak in the O–O PCF, as obtained from the RMC fitted tetrahedral DD structure model, in comparison with two room temperature DD models obtained from MD simulations, SPC/E and TIP4P-POL2.[35] We can directly see that the simulations do not provide an acceptable agreement, the first peak is much higher and sharper and is also shifted toward shorter distances compared to the RMC fit to the experiment. The question is whether this deviation should be considered real or if it could be within the error bars due to the limited q range in the experimental data or due to problems in the present fitting procedure.

To answer this we also show in Fig. 5 the two recent EPSR fitted O–O PCFs of Soper[15] using either data from Hura *et al.*[23] or the much older, but significantly more extended data set obtained by Narten and Levy.[37] The RMC and the EPSR correlation functions[15] both based on the data of Hura *et al.* are in very close agreement in terms of the height and width of the first peak, which indicates that the two independent fitting procedures give consistent results. We furthermore observe that the EPSR fits to the data of Narten and Levy and Hura *et al.* are also consistent although some uncertainties in peak position and shape remain. We conclude that the deviation between MD and the RMC/EPSR fits to the data is real and that the first peak in the O–O correlation from these MD simulations is too high and narrow to describe the experimental diffraction data.

In recent years most simulators have compared their results to either earlier O–O pair correlation functions of Soper[38,39] or as derived by Head-Gordon and co-workers,[29,40]

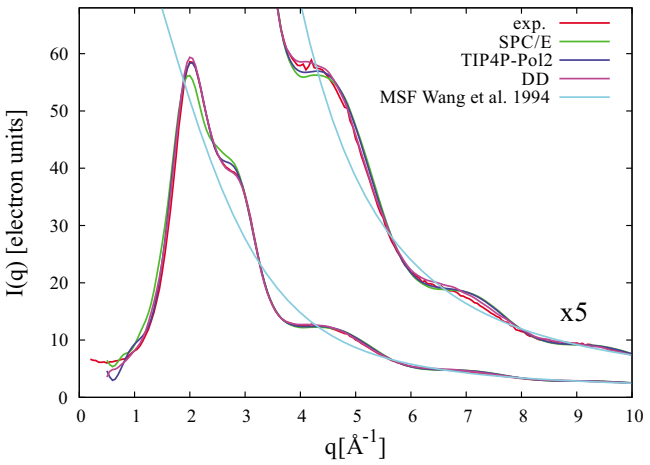


FIG. 6. (Color online) Comparison of calculated total scattering $I(q)$ for SPC/E (green), TIP4P-POL2, (blue) and the RMC derived DD structure model (red) with the experimental $I(q)$ at 27 °C from Hura *et al.* (Ref. 23). The region above 3.5 Å^{-1} has been scaled up by a factor of 5 in the inset to allow a better comparison between measured data and simulations. The molecular scattering factor is also plotted to show its dominant contribution to the scattering intensity.

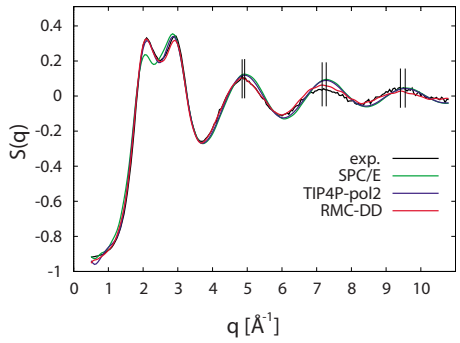


FIG. 7. (Color online) Comparison of the total structure factor $S(q)$ for MD simulations using the SPC/E and TIP4P-POL2 (Ref. 35) force fields and the RMC-DD model with the $S(q)$ derived from the experimental data of Hura *et al.* (Ref. 23). The vertical bars represent the peak position for each oscillation in the experimental and MD obtained $S(q)$.

which indicated a much sharper and higher first peak. Soper, in his latest study,[15] pointed out the relative insensitivity of using only ND to determine the O–O PCF and has therefore combined XD and ND data in the EPSR fitting procedure. Hura *et al.*[23] compared the XD data in terms of the total scattering intensity, $I(q)$, with various MD simulations and concluded that, of the set of investigated MD models, TIP4P-POL2 gives the best agreement with the data. The PCFs were then extracted from this simulation and have subsequently been used as "experimental" reference by many developers of force fields. The $I(q)$ x-ray data are, however, dominated by scattering within the water molecule and therefore differences in the intermolecular scattering between different models become difficult to observe. In Fig. 6 the scattering intensities, $I(q)$, from the RMC-DD model and the SPC/E and TIP4P-POL2 models are compared to experiment. The intermolecular scattering is directly obtained from $S(q)$, whereas the intramolecular scattering is calculated through the Debye scattering equation[30]

$$F^2(q) = \sum_i \sum_j f_i(q) f_j(q) \frac{\sin(qr_{ij})}{qr_{ij}} \exp(-q^2/2\sigma_{ij}), \qquad (6)$$

where $f_i(q)$ are modified atomic form factors[29] with the charge transfer parameter set to 0.4 as before, and the σ_{ij} give the standard deviations of the intramolecular bond distances. The intramolecular parameters were seen to influence the resulting $I(q)$ only very slightly; $r_{OH}=0.97$ Å and $\sigma_{OH}=0.07$ Å from Ref. 41 were used as well as $r_{HH}=1.56$ Å and $\sigma_{HH}=0.13$ Å (the H–H contribution is vanishingly small). In the figure, the RMC-DD structure model is almost impossible to distinguish from the experimental curve, TIP4P-POL2 seems to give a very good fit except for deviations in the shoulder feature around $q=2.8$ Å^{-1}, while SPC/E deviates significantly from the experiment in the region $q=1-3$ Å^{-1}. Note that for $q<1$ Å^{-1} the curves predicted by both the MD and RMC models are unreliable due to truncation effects in the Fourier transform.

This comparison of TIP4P-POL2 with the experimental data is consistent with Ref. 23 where it was concluded that due to the good agreement in $I(q)$, the TIP4P-POL2 model should be considered to provide the best currently available benchmark $g(r)$ functions. However, by scaling up the region above 3.5 Å^{-1} by a factor of 5, as in the inset in Fig. 6, *important discrepancies in terms of both amplitude and phase* between the experimental $I(q)$ and the MD simulations become apparent.

When the large intramolecular scattering contribution (also shown in Fig. 6) is subtracted from the measured data the differences become even more apparent showing that, contrary to what has been claimed,[23] the TIP4P-POL2 force field does not reproduce the diffraction data (see Fig. 7). It should be noted that to generate Fig. 6 the precalculated, model-independent intramolecular scattering was added to the $S(q)$ calculated for the respective model. Here we simply subtract the same large quantity from the experimental data instead; note the different scales in Figs. 6 and 7.

The difference between the $S(q)$ derived from the experimental data and the MD model structures converted into q space becomes apparent at q above 3.5 Å^{-1}. We note a phase shift toward higher q for the MD models with respect to the experimental $S(q)$; this increases with increasing q and thereby directly gives the trend toward shorter distances in the MD O–O pair correlation function. The RMC fitted DD structure model, on the other hand, follows $S(q)$ without any phase shift. The amplitude is also too high for the MD models, which reflects the too sharp first O–O peak in the MD simulations. The mismatch in amplitude between RMC and the experimental data is likely related to problems in the normalization of the experimental data. As can be seen, the RMC fitted curve shows the physically expected damping of the oscillations that, however, the experimental data lacks (the same discrepancy can be seen in published EPSR fits[15]). Note finally that the incorrect intensity ratio for SPC/E in the first double peak in $S(q)$ indicates that the model is too structured and icelike, manifested primarily in the depth of the first minimum in the O–O PCF (see Fig. 5).

The three partial PCFs can be seen as restrictions on

possible structure models. The lower and broader first O–O PCF peak obtained here, and recently by Soper,[15] is in stark contrast to previous diffraction studies where MD potentials have been included in the analysis, as well as to the vast majority of MD simulations. This furthermore indicates that significantly more distorted structures are present in water than what is obtained in MD simulations. Whether these distortions are symmetric or asymmetric is, however, undetermined from the point of view of diffraction data, but it is clear that the soft features of the O–O PCF are consistent with a large range of structure models.

D. E-field distributions (Raman spectra)

Unlike the truly experimental diffraction data the E-field distribution only provides an *approximate* model of the O–H stretch Raman spectrum with uncertainties in terms of a large spread around the average frequency connected to each field strength[14,30,31] as well as the need to obtain the reference distribution from a simulation. In this respect we find the E-field distribution from a TIP4P-POL2 structure, as well as that of the symmetric structure fitted to diffraction data by Soper[7,15] using the EPSR method, to be blueshifted from that obtained from SPC/E, which reflects that the latter potential gives a more icelike structure. Indeed, we have not succeeded to fit the E-field distribution from SPC/E in conjunction with the diffraction data while that from a TIP4P-POL2 structure proved possible albeit time consuming; the MD-based E-field distributions thus to some extent reflect the discrepancy of MD simulations with the diffraction data and caution must be exercised so that the assumed relationship between E-field distribution and Raman spectrum, validated for simulations that do not reproduce the diffraction data, at least does not strongly influence the RMC fitting procedure.

In all cases we have thus focused on the diffraction data to drive the fit, putting less emphasis on the E-field distribution; the resulting distributions still reproduce the target E-field distribution very well and, from a pragmatic viewpoint, the inclusion of the E-field distribution in the fit serves to eliminate some very unphysical situations, e.g., hydrogens from different molecules in close proximity to each other, that are not excluded by the diffraction data alone.[7,11,15]

In Fig. 8 we show the fitted E-field distributions for the two models where the distribution from the TIP4P-POL2 MD simulation of room temperature water[35] was used as reference target; SPC/E charges were used throughout, however. Again we find equivalent fits also to the E-field distribution for both structure models, well within the error bars of the E-field approximation.[13,34,35] This means that even the Raman spectrum of liquid water can be interpreted within the E-field model using either very distorted or maximally tetrahedral structure models, as well as naturally also a range of models between these two extremes.

Two earlier studies[11,12] based on the most asymmetric structure solution of Soper[7] indicated a resulting bimodal distribution of frequencies in contradiction to experiment; in Ref. 12 this led the authors to conclude that *no* asymmetrical structure could fit the data, contrary to what is shown in the present work. It is clear that the flexibility in the three-dimensional H-bond network, together with the one-dimensional, integrated character of the experimental data sets, allows a broad range of structural solutions, not necessarily, however, confined to lie within the range of asymmetrical to symmetrical structures generated here based on H-bonding connectivity criteria.

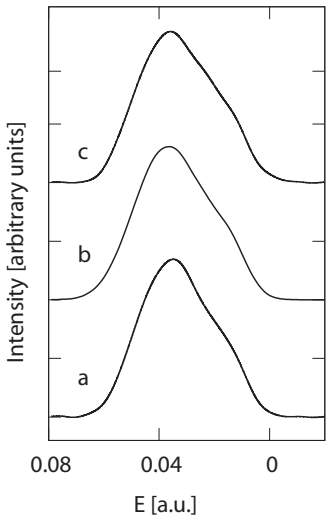


FIG. 8. Resulting E-field distributions from the simultaneous RMC fits to XD and ND data as well as E-field distribution using the (a) DD and (c) asymmetrical models. The reference distribution (b) is taken from a TIP4P-POL2 simulation (Ref. 35).

E. Energetics

Although the RMC procedure does not guarantee a thermodynamically equilibrated structure the very good fit to the diffraction data shows that the resulting structure reproduces the thermodynamically equilibrated distributions of interatomic distances in the liquid; if only pairwise interactions are assumed, the energy is then completely specified through integration of the PCFs with the expression for the potential. It is thus of relevance to estimate the difference in interaction energy between these two extreme structure models for which a too large value clearly cannot be accepted.

We use classical pair potentials [SPC,[42] SPC/E,[43] TIP3P (Ref. 44)] and, in addition, the recent *ab initio* derived, polarizable TTM2.1-F potential of Fanourgakis *et al.*[45] to estimate the difference in interaction energy between the two structure models. Note that no averaging or dynamical sampling is performed. We simply evaluate the interaction energy for the two specific RMC optimized fixed structures using the various potential models; the atom coordinates in the two structures are given in the supplementary material.[26]

Although both structure models also reproduce the quantum distribution of intramolecular geometries, the reference distributions are taken from a simulation[24] and not directly observed experimentally. In order to eliminate uncertainties connected to this we evaluate the energy both for the RMC optimized structures excluding zero-point vibrations, i.e., excluding the intramolecular energy, and also by using the fixed SPC/E geometry for each molecule (R(O–H)=1.0 Å, H–O–H angle 109.47°).

Let us first discuss the tetrahedral RMC DD structure model for which we find interaction energies of −37.3, −40.5, and −40.1 kJ/mol using, respectively, the SPC, SPC/E, and TIP3P flexible potentials; changing to fixed geometry has only minor (0.1–0.2 kJ/mol) effects. For the SPC and SPC/E potential models the MD equilibrated interaction energy has been reported as −41.7 and −46.6 kJ/mol, respectively,[43] while the TIP3P value is −41.0 kJ/mol;[46] the larger negative value for the SPC/E model is due to the correction, −5.22 kJ/mol,[43] to the interaction energy for the induced dipole moment in the model. The RMC DD structure model energies are thus 4.2, 5.7, and 0.9 kJ/mol higher than the various MD values.

Evaluated with the MD potentials the RMC DD structure thus gives energies corresponding to weaker and less well-defined H-bond interactions. The MD potential models have been carefully optimized to generate the thermodynamically correct total energy, but, as shown here and in Ref. 15, based on structures that contain too many short and well-defined H bonds to be consistent with the diffraction data.[15] Integrating the MD energy expressions over the more accurate experimentally derived PCFs obtained here from the RMC fits, thus leads to predicted energies that are too high. This indicates that the MD potentials have been fitted to the thermodynamically correct total energy at the cost of overstructured PCFs. With a heat capacity of 75.4 J/mol K, 5 kJ/mol corresponds roughly to the energy needed to heat room temperature water by 66°. Based on the energy differences (0.9–5.7 kJ/mol) the overstructuring in the MD simulations thus roughly corresponds to cooling the water ~12–76°. However, this estimation is only to put the deviation on a relative thermodynamic scale and we do not imply that MD simulations performed at higher temperature with the potentials under discussion would correctly reproduce the PCFs for the liquid at ambient conditions.

As an estimate of error bars associated with energy estimates using force fields in general, we compare these numbers with the difference between the experimental and simulation-based latent heat of vaporization[47] as reported in the careful and extensive study by Paesani *et al.*[48] using the more advanced polarizable TTM2.1-F force field,[45] which contains some many-body contributions through the polarizability. This force field has not been parametrized against experiment and we can thus take the observed deviation from the experimental value as a reasonable estimate of the uncertainty in the simulations and energy estimates. The quantum simulations for the latent heat of vaporization of H_2O and D_2O are 2.9 and 4.2 kJ/mol, respectively, higher for this potential than the respective experimental values;[48] the above estimates for the tetrahedral RMC DD structure model using the various potentials clearly lie within this range.

The RMC SD structure model with 81% asymmetrically H-bonded water molecules is as expected less energetically stable than the tetrahedral DD structure model with 74% double donors, but by a surprisingly small amount; using flexible SPC, SPC/E, or TIP3P the interaction energy is only 0.7–0.8 kJ/mol higher. This reflects the obviously only minor differences in the partial PCFs since both structure models reproduce the same diffraction data. Using instead the fixed geometry this is increased to 1.3–1.4 kJ/mol, which may be regarded as an upper bound since the SD species are structurally more affected by going to an average fixed internal geometry than are the DD species. Specifically, fixing the angle has a larger probability for the asymmetric species to weaken the single strong donating H bond.

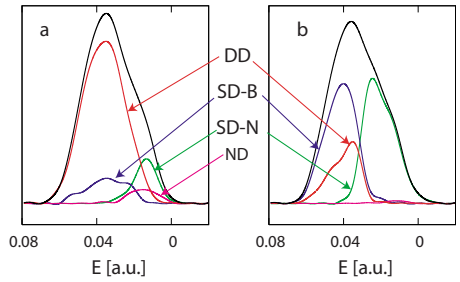


FIG. 9. (Color online) Decomposition of the E-field distributions into contributions from OH groups in doubly H-bonded (DD) species, H-bonded OH in asymmetric SD species (SD-B), and non-H-bonded OH in asymmetric species (SD-N). ND contributions are small in the DD model and negligible in the asymmetrical SD model. (a) Symmetric DD model and (b) asymmetrical SD model.

For the TTM2.1-F potential a similar difference in interaction energy is found with 1.3 and 1.8 kJ/mol for flexible and fixed molecules, respectively. Considering the uncertainties, the small magnitude of the energy difference between the DD and SD structure models and the likely higher entropy of the more distorted structure model in terms of, e.g., softer and thus thermally accessible librational modes we conclude that neither structural solution, here obtained without including any specific interaction model, can be excluded based on these energy estimates.

F. Connection to ultrafast pump-probe studies

Additional experimental information on the H-bond network and, in particular, its dynamics is obtained from state-of-the-art pump-probe ultrafast 2D-vibrational spectroscopies where the time evolution of a specific selected excitation is monitored.[49] The ultrafast spectral shift toward redshifted frequencies when exciting on the blue side of the OH stretch band has been taken as evidence that H bonds are broken only fleetingly[4,5] and that a predominance of asymmetrically coordinated species with broken H bonds could not describe the liquid.[4,5] Based on the data included in the fit we cannot directly address the dynamics, but we can, however, investigate the decomposition of the Raman spectrum in terms of an assumed connection between frequency and structure and thus critically evaluate what an observed change in frequency actually strictly implies in terms of H-bond changes; we find that this is strongly model dependent.

Considering that the shape of the E-field distribution used to represent the Raman OH stretch spectrum can be reproduced both with a structure model that maximizes the number of H bonds and one that maximizes the number of highly H bond broken and distorted species we conclude that also the interpretation of 2D pump-probe ultrafast spectroscopies must depend on which model is preferred. In Fig. 9 we show the decomposition of the E-field into contribu-

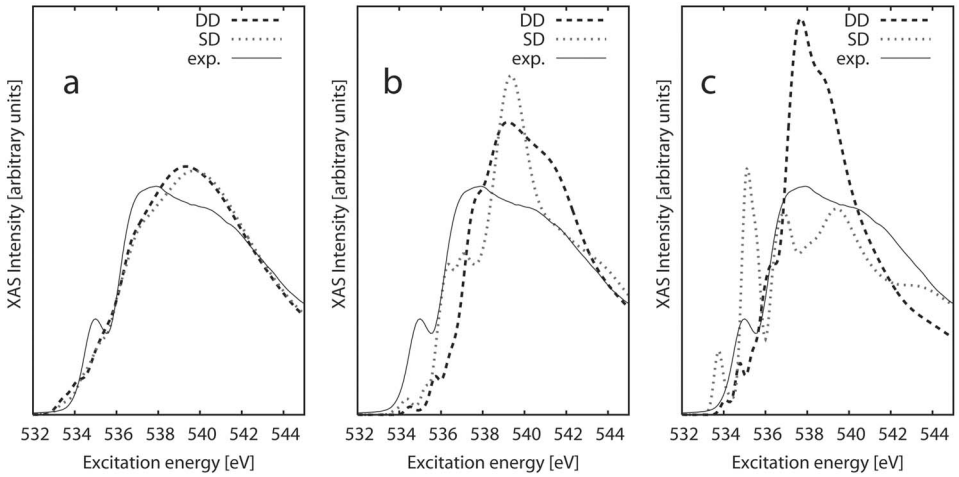


FIG. 10. Computed XAS spectra for the two different models in comparison with experiment from Ref. 1. (a) Total spectra for the DD and SD models summing 200 and 196 individual spectra, respectively. [(b) and (c)] Individual spectra for the same molecule in the respective dump where the original environment has been modified according to the two different RMC fits. The experimental spectrum has been area normalized to the mean integrated (up to 545 eV) cross section of the two computed sum spectra in A.

tions from various species for the two extreme structure models. It is clear that with the very large structural difference between the models the distributions of contributions giving the same total spectrum will be very different. How does that affect the interpretation?

What is observed experimentally is a shift in frequency as function of time after the initial pump step excites, e.g., an isotopically isolated OH stretch vibration in a selected frequency region; the connection with H bonding and structure is, however, given by a model. Considering first the traditional structure model in Fig. 9(a) the very limited range spanned by the non-H-bonded SD species makes it clear that even small frequency changes after exciting in this region must, within this model, be interpreted as a change in the H-bonding network, e.g., a SD species converting to DD, leading to the conclusion that the H-bond dynamics is ultrafast and that broken H bonds exist only transiently.[4] The experimentally, however, equally valid picture in Fig. 9(b) makes a totally different interpretation possible. Here the ultrafast shift in frequency on the non-H-bonded blue side of the spectrum represents the dynamics of the free O–H interacting less specifically with each of several neighbors in a cavity with direction-dependent electrostatic fields affecting the frequency, but without changes in the H-bonding situation. On a slightly longer time scale a jump to a strongly redshifted frequency within the OH stretch band could simply correspond to the non-H-bonded OH in a librationally highly excited SD species trading roles with the H-bonded OH of the same molecule, a process that does not change the H-bonding situation. Thus, H-bond dynamics in terms of converting SD into DD species may not necessarily be ultrafast although frequency changes are; the experimental data do not directly provide this connection.

G. X-ray absorption spectrum

Since the SD structure model intentionally mimics the local asymmetry deduced from the XAS spectrum in Ref. 1 the present investigation will not be complete without also considering the predicted XAS spectra from the respective DD and SD structure models. Before proceeding it should be noted, however, that no RMC fitting was made directly toward the XAS data in the present work. We use the STOBE DFT code[50] in the standard transition potential, half-core hole approach[51] to calculate the XAS spectra for the SD and DD structures using cluster models containing 32 molecules with the same computational parameters as in previous work[11,36] except that the Perdew–Burke–Ernzerhof functional[52] was used. The absolute energy scale is determined for each spectrum contribution through variationally determining the excitation energy of the first core-excited state[11,36,53] taking into account also the influence of the selected functional;[54] the computational procedure is fully described in Ref. 11.

Similar to what was found in Ref. 36 when applying the local cone-based H-bond criterion from Ref. 1 we obtain icelike computed XAS spectra for both the SD and DD structure models of the present work, see Fig. 10(a), where we have summed spectra from 200 (DD) and 196 (SD) randomly selected molecules in the respective dump. The similarity of the resulting spectra is perhaps surprising considering the obvious structural differences shown in Fig. 4 and also indicated by the very different H-bond statistics and could perhaps be thought to represent an insensitivity of XAS to the structure in contradiction to the claim of Wernet *et al.*[1] However, we have already in Ref. 36 shown that the local H-bond criterion of Ref. 1 is insufficient to predict a correct XAS spectrum for water. Instead it was demonstrated that a spectrum similar to experiment could be obtained by modifying the cone criterion to enhance the H-bond-donor asymmetry by introducing a rather large buffer zone between intact and broken H bond.[36] It is, however, difficult to see how a single-component water model could be constructed taking this criterion into account while still reproducing the diffraction data and density. In fact, in the present RMC fits it was deemed infeasible to attempt to enforce a significantly larger asymmetry. However, the earlier interpretation from XAS in terms of two dominating structural motifs in the liquid is supported by new x-ray emission spectroscopy (XES) data,[55] which opens the possibility of rather different structure models, e.g., mixtures of two distinctly different local structures where one component may be very distorted and the other very structured.[55]

An alternative interpretation of the similarity of the computed DD and SD XAS spectra is that the transition potential approach[51] is not sufficiently accurate to correctly predict

XAS spectra as has been claimed by Prendergast and Galli.[56] Earlier applications to gas phase molecules, e.g., pyridine,[53] methanol,[57] water,[58] as well as to condensed phase, e.g., bulk methanol,[57] surface adsorbates such as glycinate on Cu(110),[59] $H_2O/Pt(111)$,[60] $H_2O/Cu(110)$,[61] $H_2O-OH/Pt(111)$ [62] have, however, yielded excellent results in comparison with experiment. In those cases the structure to be computed has been known; a clear prerequisite for the spectrum calculations. The structure is, however, less clear in the case of liquid water.

In Figs. 10(b) and 10(c) we demonstrate the sensitivity of the spectroscopy and the spectrum calculations to structural variations by showing typical individual spectra from the two models. The RMC simulations giving the two different structure models, DD and SD, were started from the same initial MD structure, which allows us to track RMC induced changes around individual molecules. In the two figures we compare computed spectra for the same two randomly selected molecules from the initial structure, but after structural changes according to the two different RMC fits. The resulting spectra show significant differences, but the main observation is that not one of the computed individual spectra bears any resemblance with the final sum. The total spectra, which here in both cases appear rather featureless and icelike, are thus the result of a statistical averaging of rather disparate individual spectrum contributions where certain features are enhanced and other features quenched through the summation.

We have earlier shown that a much larger radial asymmetry could generate a spectrum in reasonable agreement with experiment,[36] but this is likely difficult to combine with diffraction data, at least within a continuum model as used here. Considering the statistical averaging of spectral features illustrated by Fig. 10 we can speculate that the sharp pre-edge and well-defined main edge in the experiment reflect semilocal correlations in the H-bond network that could add statistical weight to these spectral regions. Energetically preferred relative orientations of free OH groups in water clusters have been described, e.g., by McDonald *et al.*[63] but such correlations are difficult to include in RMC without an interaction potential, which could bias the results in an uncontrolled way. It is furthermore not clear what, if any, effect this would induce on the XAS spectrum. We conclude that more work is needed to determine the local and intermediate range structural features or correlations that could, after summing a significant number of rather different-looking spectral contributions, add up to a spectrum consistent with the experiment.

IV. DISCUSSION

As we have shown here, the established techniques for experimental structure determination of liquid water allow a very large variety of structural solutions. This was realized early on and MD simulation models based on known or assumed physics were introduced to complement the data and guide the analysis.[3] These MD models have had a great impact and with improvements in computing resources, potential models, animations, and algorithms they have evolved into standard tools for analysis and interpretation, guiding structural analyses toward distributions around icelike tetrahedral local structures. They are, however, only models that must continuously be tested against experiment.

With the current experimentally derived PCFs it is clear that the potentials used in MD need to be softer in order to describe the data.[15] The question is then if a MD simulation using such softer potentials can be expected to also reproduce other observables such as thermodynamic properties, diffusion, dielectric constant, surface tension, etc. The inclusion of quantum effects in the dynamics has recently been recognized as very important and this could change the picture. The recent path integral simulation of the TTM2.1-F model has lowered the first peak in the O–O pair correlation in comparison to classical dynamics, but it is still too sharp compared to the correlation shown in Fig. 5.[48] The discrepancy with experiment shows that the potential is not yet fully optimized and errors from the approximate path integral treatment of the quantum dynamics must furthermore be considered. It is interesting, however, to note that the ratio of SD to DD species nearly doubled between the classical and quantum dynamics simulations, but the ratio is still far from that reported experimentally by Wernet *et al.*[1] using XAS and also by Tokushima *et al.*[55] using XES. An interesting question is then how this ratio would be in a simulation that can simultaneously describe the diffraction data and also the thermodynamic properties.

In the current study we have approximated the internal localized OH stretch mode Raman spectrum with the E-field distribution as obtained from a TIP4P-POL2 simulation. However, this is not an experimental observable and in this sense we have simply demonstrated that we can obtain similar E-field descriptions for two extreme water structure models that are rather different from the traditional MD models. Corcelli *et al.*[13,33] calibrated the E-field description to OH stretch frequencies using quantum chemical calculations of various local water environments from MD simulations. Although there is a large scatter in this correlation they have generated IR and Raman spectra of various water models that have been compared directly to experimental spectra.[13,33] In the most recent study they show that most of the MD models they have tested are redshifted in comparison to the experimental spectra.[64] In Ref. 64 they use a linear mapping from E-field to frequency, while in Ref. 32 a quadratic mapping is used, which makes the MD models peak at approximately the experimental frequency. Certainly then, MD-like models are not excluded based on the OH stretch Raman spectrum, but given the error bars in the association between E-field and frequency and considering the very limited sampling due to the use of MD structures, E-fields that are considerably blueshifted cannot be excluded. From our failure to obtain a simultaneous fit to the diffraction data and the E-field from SPC/E we, however, conclude that this E-field is too redshifted to be compatible with the diffraction data.

Although recently introduced synchrotron-radiation-based spectroscopy techniques also depend on interpretation,[1,55,56,65,66] they do offer a new approach to the important question of the structure and properties of liquid

water, which seriously challenges the traditional picture.[1,55,66] In the present work we have emphasized the well-known fact that agreement with experiment for a specific model does not imply that this is the only valid way to interpret the data. We have shown that available diffraction and vibrational data can be described equally well within the textbook tetrahedral model, although significantly more distorted than MD models, and the kind of highly distorted SD dominated structure model suggested by Wernet *et al.*[1] However, even though the XAS experiment puts significant and different constraints on the local coordination, the interpretation of the spectrum in terms of asymmetric SD species may still not be unique. The experiments indicate a large fraction of uncoordinated OH groups (pre-edge) but also many OH groups in strongly H-bonding environment (postedge). There are several different possibilities to outline structures where these two OH groups are either on the same molecule or on distinct different species, as well as combinations of these situations.

In order to develop a consistent picture of the structure and dynamics of liquid water encompassing also the new synchrotron radiation experiments a significant amount of new experimental and theory development will be needed. The first step, however, consists in recognizing the nonexclusivity of the popular model of water as a mainly tetrahedrally connected liquid.

ACKNOWLEDGMENTS

We are grateful to A. K. Soper for supplying the neutron data, to T. Head-Gordon for making the XD data available, to J. I. Siepmann for sending the TIP4P-POL2 MD trajectory, to F. Paesani and G. Voth for making available and modifying their TTM2.1-F code used for the energy evaluations, and to H. Bakker, T. Elsaesser, R. Kjellander, S. Kohara, L. Pusztai, and A. Lyubartsev for comments and discussions. This work was supported by the Swedish Foundation for Strategic Research, the Swedish Research Council (VR), and the National Science Foundation (U.S.) under Grant Nos. CHE-0518637 and CHE-0431425. Portions of this research were carried out at the Stanford Synchrotron Radiation Laboratory, a national user facility operated by Stanford University on behalf of the U.S. Department of Energy, Office of Basic Energy Sciences.

[1] Ph. Wernet, D. Nordlund, U. Bergmann, M. Cavalleri, M. Odelius, H. Ogasawara, L.-Å. Näslund, T. K. Hirsch, L. Ojamäe, P. Glatzel, L. G. M. Pettersson, and A. Nilsson, Science **304**, 995 (2004).
[2] B. J. Guillot, J. Mol. Liq. **101**, 219 (2002).
[3] F. H. Stillinger, Science **209**, 451 (1980).
[4] J. D. Eaves, J. J. Loparo, C. J. Fecko, S. T. Roberts, A. Tokmakoff, and P. L. Geissler, Proc. Natl. Acad. Sci. U.S.A. **102**, 13019 (2005).
[5] A. Tokmakoff, Science **317**, 54 (2007).
[6] A. K. Soper, Mol. Phys. **99**, 1503 (2001); Phys. Rev. B **72**, 104204 (2005).
[7] A. K. Soper, J. Phys.: Condens. Matter **17**, S3273 (2005).
[8] T. Head-Gordon and M. E. Johnson, Proc. Natl. Acad. Sci. U.S.A. **103**, 7973 (2006).
[9] T. Head-Gordon and M. E. Johnson, Proc. Natl. Acad. Sci. U.S.A. **103**, 16614 (2007).
[10] T. Head-Gordon and S. W. Rick, Phys. Chem. Chem. Phys. **9**, 83 (2007).
[11] M. Leetmaa, M. Ljungberg, H. Ogasawara, M. Odelius, L.-Å. Näslund, A. Nilsson, and L. G. M. Pettersson, J. Chem. Phys. **125**, 244510 (2006).
[12] J. D. Smith, C. D. Cappa, B. M. Messer, W. S. Drisdell, R. C. Cohen, and R. J. Saykally, J. Phys. Chem. B **110**, 20038 (2006).
[13] S. A. Corcelli and J. L. Skinner, J. Phys. Chem. A **109**, 6154 (2005).
[14] C. J. Fecko, J. D. Eaves, J. J. Loparo, A. Tokmakoff, and P. L. Geissler, Science **301**, 1698 (2003).
[15] A. K. Soper, J. Phys.: Condens. Matter **19**, 335206 (2007).
[16] S. E. McLain, S. Imberti, A. K. Soper, A. Botti, F. Bruni, and M. A. Ricci, Phys. Rev. B **74**, 094201 (2006).
[17] R. L. McGreevy and L. Pusztai, Mol. Simul. **1**, 359 (1988).
[18] G. Toth and A. Baranyai, J. Chem. Phys. **107**, 7402 (1997); Mol. Phys. **97**, 339 (1999).
[19] R. L. McGreevy, J. Phys.: Condens. Matter **13**, R877 (2001).
[20] P. Jedlovszky, I. Bako', G. Palinkas, T. Radnai, and A. K. Soper, J. Chem. Phys. **105**, 245 (1996).
[21] P. Jedlovszky, J. P. Brodholt, F. Bruni, M. A. Ricci, A. K. Soper, and R. Vallauri, J. Chem. Phys. **108**, 8528 (1998).
[22] L. Pusztai, Phys. Rev. B **60**, 11851 (1999).
[23] G. Hura, D. Russo, R. M. Glaeser, T. Head-Gordon, M. Krack, and M. Parrinello, Phys. Chem. Chem. Phys. **5**, 1981 (2003).
[24] H. A. Stern and B. J. Berne, J. Chem. Phys. **115**, 7622 (2001).
[25] G. Evrard and L. Pusztai, J. Phys.: Condens. Matter **17**, S1 (2005).
[26] See EPAPS Document No. E-JCPSA6-129-610833 for details on the RMC fit, fits to internal structure and full coordinate files for the two models. For more information on EPAPS, see http://www.aip.org/pubservs/epaps.html.
[27] *International Tables for Crystallography*, Vol. C, A. J. C. Wilson (Ed.) (Kluwer, Dordrecht, 1992).
[28] D. Waasmaier and A. Kirfel, Acta Crystallogr., Sect. A: Found. Crystallogr. **51**, 416 (1995).
[29] J. M. Sorenson, G. Hura, R. M. Glaeser, and T. Head-Gordon, J. Chem. Phys. **113**, 9149 (2000).
[30] B. E. Warren, *X-Ray Diffraction* (Dover, New York, 1990).
[31] M. P. Allen and D. J. Tildesley, *Computer Simulation of Liquids* (Oxford University Press, New York, 1987).
[32] B. Auer, R. Kumar, J. R. Schmidt, and J. L. Skinner, Proc. Natl. Acad. Sci. U.S.A. **104**, 14215 (2007).
[33] S. A. Corcelli, C. P. Lawrence, and J. L. Skinner, J. Chem. Phys. **120**, 8107 (2004).
[34] R. Kumar, J. R. Schmidt, and J. L. Skinner, J. Chem. Phys. **126**, 204107 (2007).
[35] B. Chen, J. H. Xing, and J. I. Siepmann, J. Phys. Chem. B **104**, 2391 (2000).
[36] M. Odelius, M. Cavalleri, A. Nilsson, and L. G. M. Pettersson, Phys. Rev. B **73**, 024205 (2006).
[37] A. H. Narten and H. A. Levy, J. Chem. Phys. **55**, 2263 (1971).
[38] A. K. Soper, F. Bruni, and M. A. Ricci, J. Chem. Phys. **106**, 247 (1997).
[39] A. K. Soper, Chem. Phys. **258**, 121 (2000).
[40] G. Hura, J. M. Sorenson, R. M. Glaeser, and T. Head-Gordon, J. Chem. Phys. **113**, 9140 (2000).
[41] A. K. Soper and M. G. Phillips, Chem. Phys. **107**, 47 (1986).
[42] K. Toukan and A. Rahman, Phys. Rev. B **31**, 2643 (1985).
[43] H. J. C. Berendsen, J. R. Grigera, and T. P. Straatsma, J. Phys. Chem. **91**, 6269 (1987).
[44] W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, and M. L. Klein, J. Chem. Phys. **79**, 926 (1983).
[45] G. S. Fanourgakis, G. K. Schenter, and S. Xantheas, J. Chem. Phys. **125**, 141102 (2006).
[46] M. W. Mahoney and W. L. Jorgensen, J. Chem. Phys. **112**, 8910 (2000).
[47] The latent heat of vaporization corresponds to the interaction energy plus the difference in intramolecular energy between liquid and gas phase.
[48] F. Paesani, S. Iuchi, and G. A. Voth, J. Chem. Phys. **127**, 074506 (2007).
[49] E. T. J. Nibbering and T. Elsaesser, Chem. Rev. (Washington, D.C.) **104**, 1887 (2004).
[50] K. Hermann, L. G. M. Pettersson, M. E. Casida, C. Daul, A. Goursot, A. Koester, E. Proynov, A. St-Amant, D. R. Salahub, V. Carravetta, A. Duarte, N. Godbout, J. Guan, C. Jamorski, M. Leboeuf, M. Leetmaa, M. Nyberg, L. Pedocchi, F. Sim, L. Triguero, and A. Vela, STOBE-DEMON, DEMON Software, Stockholm, Berlin, 2005.
[51] L. Triguero, L. G. M. Pettersson, and H. Ågren, Phys. Rev. B **58**, 8097 (1998).
[52] J. P. Perdew, K. Burke, and M. Ernzerhof, Phys. Rev. Lett. **77**, 3865 (1996).
[53] C. Kolczewski, R. Püttner, O. Plashkevych, H. Ågren, V. Staemmler, M.

Martins, G. Snell, A. S. Schlachter, M. Sant'Anna, G. Kaindl, and L. G. M. Pettersson, J. Chem. Phys. **115**, 6426 (2001).
[54] O. Takahashi and L. G. M. Pettersson, J. Chem. Phys. **121**, 10339 (2004).
[55] T. Tokushima, Y. Harada, O. Takahashi, Y. Senba, H. Ohashi, L. G. M. Pettersson, A. Nilsson, and S. Shin, Chem. Phys. Lett. **460**, 387 (2008).
[56] D. Prendergast and G. Galli, Phys. Rev. Lett. **96**, 215502 (2006).
[57] K. R. Wilson, B. S. Rude, R. D. Schaller, T. Catalano, R. J. Saykally, M. Cavalleri, A. Nilsson, and L. G. M. Pettersson, J. Phys. Chem. B **109**, 10194 (2005).
[58] M. Cavalleri, M. Odelius, D. Nordlund, A. Nilsson, and L. G. M. Pettersson, Phys. Chem. Chem. Phys. **7**, 2854 (2005).
[59] M. Nyberg, M. Odelius, A. Nilsson, and L. G. M. Pettersson, J. Chem. Phys. **119**, 12577 (2003).
[60] H. Ogasawara, B. Brena, D. Nordlund, M. Nyberg, A. Pelmenschikov, L. G. M. Pettersson, and A. Nilsson, Phys. Rev. Lett. **89**, 276102 (2002).
[61] T. Schiros, S. Haq, H. Ogasawara, O. Takahashi, H. Öström, K. Andersson, L. G. M. Pettersson, A. Hodgson, and A. Nilsson, Chem. Phys. Lett. **429**, 415 (2006).
[62] T. Schiros, L.-Å. Näslund, K. Andersson, J. Gyllenpalm, G. S. Karlberg, M. Odelius, H. Ogasawara, L. G. M. Pettersson, and A. Nilsson, J. Phys. Chem. C **111**, 15003 (2007).
[63] S. McDonald, L. Ojamäe, and S. J. Singer, J. Phys. Chem. A **102**, 2824 (1998).
[64] J. R. Schmidt, S. T. Roberts, J. J. Loparo, A. Tokmakoff, M. D. Fayer, and J. L. Skinner, Chem. Phys. **341**, 143 (2007).
[65] O. Fuchs, M. Zharnikov, L. Weinhardt, M. Blum, M. Weigand, Y. Zubavichus, M. Bär, F. Maier, J. D. Denlinger, C. Heske, M. Grunze, and E. Umbach, Phys. Rev. Lett. **100**, 027801 (2008); L. G. M. Pettersson, T. Tokushima, Y. Harada, O. Takahashi, S. Shin, and A. Nilsson, *ibid.* **100**, 249801 (2008); A. Nilsson, Ph. Wernet, D. Nordlund, U. Bergmann, M. Cavalleri, M. Odelius, H. Ogasawara, L.-Å. Näslund, T. K. Hirsch, L. Ojamäe, P. Glatzel, and L. G. M. Pettersson, Science **308**, 793a (2005); J. D. Smith, C. D. Cappa, K. R. Wilson, B. M. Messer, R. C. Cohen, and R. J. Saykally, *ibid.* **306**, 851 (2004).
[66] S. Myneni, Y. Luo, L.-Å. Näslund, M. Cavalleri, L. Ojamäe, H. Ogasawara, A. Pelmenschikov, Ph. Wernet, P. Väterlein, C. Heske, Z. Hussain, L. G. M. Pettersson, and A. Nilsson, J. Phys.: Condens. Matter **14**, L213 (2002).