The structure of liquid water by polarized neutron diffraction and reverse Monte Carlo modelling

2007 J. Phys.: Condens. Matter 19 335207

(http://iopscience.iop.org/0953-8984/19/33/335207)

IOP PUBLISHING JOURNAL OF PHYSICS: CONDENSED MATTER

J. Phys.: Condens. Matter **19** (2007) 335207 (12pp) doi:10.1088/0953-8984/19/33/335207

The structure of liquid water by polarized neutron diffraction and reverse Monte Carlo modelling

László Temleitner[1], **László Pusztai**[1] **and Werner Schweika**[2]

[1] Neutron Physics Laboratory, Research Institute for Solid State Physics and Optics, Hungarian Academy of Sciences, POB 49, H-1525 Budapest, Hungary
[2] Institut für Festkörperforschung, Forschungszentrum Jülich, D-52425 Jülich, Germany

E-mail: lp@szfki.hu

Received 3 April 2007
Published 4 July 2007
Online at stacks.iop.org/JPhysCM/19/335207

Abstract

The coherent static structure factor of water has been investigated by polarized neutron diffraction. Polarization analysis allows us to separate the huge incoherent scattering background from hydrogen and to obtain high quality data of the coherent scattering from four different mixtures of liquid H_2O and D_2O. The information obtained by the variation of the scattering contrast confines the configurational space of water and is used by the reverse Monte Carlo technique to model the total structure factors. Structural characteristics have been calculated directly from the resulting sets of particle coordinates. Consistency with existing partial pair correlation functions, derived without the application of polarized neutrons, was checked by incorporating them into our reverse Monte Carlo calculations. We also performed Monte Carlo simulations of a hard sphere system, which provides an accurate estimate of the information content of the measured data. It is shown that the present combination of polarized neutron scattering and reverse Monte Carlo structural modelling is a promising approach towards a detailed understanding of the microscopic structure of water.

1. Introduction

Precise knowledge of the microscopic structure of water is of utmost importance for most chemists and molecular biologists—the reason, quite simply, is that life on Earth is based on water and that it takes place essentially in aqueous solutions. For this reason, liquid (and also other forms of) water has(/have) been the subject(s) of a huge number of diffraction studies (see, e.g., [1–6]). Computer simulation investigations using classical (for a recent overview example, see [5]) and quantum mechanical (see, e.g. [7–9]) force fields also abound. An outstanding point to note here is that the development of force fields is—naturally—biased by experimental results; this is yet another reason why reliable diffraction data are indispensable.

Despite the (huge) volume of relevant literature, liquid water is considered to be one of the most notorious puzzles. For instance, very recently, a debate concerning the number of hydrogen (H-) bonds per water molecule has surfaced: Wernet *et al* [10] suggest that this number may be as low as two, whereas (for instance) Head-Gordon *et al* [11] maintain the more traditional view, with four H-bonds per molecule. It is also argued sometimes that a small uncertainty exists regarding even the position of the first intermolecular O–H distance—a crucial distance as it characterizes hydrogen bonding [12]. The reason that these questions are still open is the presence of *hydrogen*.

X-ray diffraction is a less sensitive probe for hydrogen and is essentially determined only by the oxygen–oxygen and oxygen–hydrogen correlations. In neutron diffraction the three distinct correlations of this binary system have a more similar weight: the difficulty is that for separating the three partial contributions, a contrast variation with H/D substitution is required. The H/D substitution, in principle, allows us to derive the most detailed information on the microscopic structure of hydrogeneous (i.e., containing ^{1}H) systems. However, the exceptionally high level of spin incoherence of ^{1}H renders most of the measured diffraction signal from pure H_2O useless ('background') from the structural point of view (see, e.g., [1]). As a consequence of this situation, the structure factor of H_2O, as well as the set of partial pair correlation functions derived, is still under debate [1–3, 12].

To make the situation even worse, in neutron diffraction, the structure factor of H_2O is the one with the highest information content, for the negative coherent scattering length of ^{1}H: as a result, negative peaks would signify preferred O–H distances in the total pair correlation function. At the same time, the H_2O structure factor is by far the most unreliable and therefore even small errors in it have a large impact on the O–H partial pair correlation function.

Recognizing this, and with the aim of correcting for these problems, a non-standard approach, involving reverse Monte Carlo (RMC) modelling [13] of the measured total structure factors (TSFs), was suggested a few years ago [12]. Separation of the partial structure factors (PSFs) and partial radial distribution functions (PRDFs) could be carried out without making use of neutron data from samples with high(er than 33%) ^{1}H content. However, it was obvious that the position of the H-bonding peak in the O–H PRDF changed when data from hydrogeneous samples were included in the modelling. Also, application of coordination constraints made it clear that there are more than one (and in some cases, significantly different) sets of PRDFs that are consistent with the set of TSFs applied during RMC modelling. That is, the current situation cannot be considered as satisfactory: it is clear that reliable neutron total structure factors of water samples with high light water content (ideally, of pure H_2O) would be decisive concerning H-bonding in water.

Spin-incoherence, on the other hand, can be tackled by separating coherent and incoherent parts of the measured diffraction signal; this can be achieved by using polarized neutrons (see, e.g., [14]). Interestingly, potentialities of polarized neutron diffraction have not been exploited in this field; a possible reason for this is that available instruments provide data over only limited momentum transfer ranges, so that traditional evaluation (involving direct Fourier transformation) would not be applicable. On the other hand, as has been shown a couple of times before (see, e.g., [17]), an RMC-based evaluation is less sensitive to the extent of the experimental Q-range. These findings form the basis of the present approach—which is a reverse Monte Carlo based analysis of polarized neutron scattering data.

This paper is concerned with the determination of the structure of pure liquid water, at ambient conditions of pressure and temperature, varying the scattering contrast by varying the ratio of light and heavy water in the samples. First, a brief overview of the polarized neutron diffraction experiments is given. Then computational methods, including reverse Monte Carlo and (reference) hard sphere Monte Carlo modelling, are introduced and some details of the

Table 1. Weighting factors and scattering cross sections for mixtures of D_2O and H_2O. The first four mixtures correspond to the actual samples measured; data for pure H_2O are given for comparison. (Cross section values are given in barns.)

D_2O (%)	H_2O (%)	Coherent partial coefficients H–H	O–H	O–O	σ_{coh} (barn/atom)	$\sigma_{spin-inc}$ (barn/atom)	$\sigma_{isotop-inc}$ (barn/atom)
100	0	0.1978	0.1720		5.138	1.382	0
80	20	0.0937	0.1184		3.177	13.287	1.450
60	40	0.0279	0.0647	0.0374	1.941	24.707	2.175
35.925	64.075	0.00	0.00		1.41	37.811	2.090
0	100	0.0622	−0.0966		2.58	56.058	0

calculations are also provided. The results section contains an analysis of the structure in terms of atom–atom (partial) radial distribution functions, coordination numbers and angle distribution functions. The relative information content of the experiments and the constraints applied is also investigated. Finally, conclusions are put forward.

2. Experiment and data evaluation

Mixtures of H_2O and D_2O, with D_2O contents of 100, 80, 60 and 36% molar ratios have been investigated on the DNS instrument [18] installed at Forschungszentrum Jülich. In order to minimize incoherent background, as well as high absorption and multiple-scattering contributions, the samples were put in double thin-walled Al-containers fabricated specifically for this experiment. (Note that standard vanadium containers would have been disadvantageous, because of the high incoherent scattering cross section of V.) This way, the geometry of the samples was that of a hollow cylinder.

Using the DNS instrument in its diffractometer set-up, scattering intensities have been collected in both spin-flip and non-spin-flip modes. After careful optimization, each sample has been investigated for 4 h at ambient pressure and temperature, using neutrons with a wavelength of 3.3 Å. This way, a momentum transfer range of 0.2–3.4 $Å^{-1}$ could be covered. Before further analyses, the Bragg peaks of aluminium (at about 2.7 and 3.1 $Å^{-1}$) have been cut out of the spectra. The usual corrections for polarization efficiency of the instrument (which is quite high, about 95%, but not ideal) and multiple scattering have also been carried out before data processing. These latter accounted for only a few per cent of the measured signal.

Coherent and incoherent contributions to the total scattering have been separated in the usual manner [14], using the following formulae:

$$I_{coh}(Q) = I^{NSF}(Q) - \tfrac{1}{2} I^{SF}(Q) \tag{1}$$

and

$$I_{incoh}(Q) = \tfrac{3}{2} I^{SF}(Q) \tag{2}$$

where the 'NSF' and 'SF' indices refer to intensities measured in 'non-spin-flip' and 'spin-flip' modes, respectively. This separation, which may be taken as a 'recipe' to be followed in each case, takes place before evaluating the coherent structure factor. To demonstrate the ratio of coherent and incoherent scattering, table 1 shows relevant scattering cross sections for the samples considered in this study. For a more detailed discussion of coherent and incoherent scattering lengths of hydrogen, the reader is referred to [15, 16].

Corrected coherent static structure factors, $S(Q)$—as defined for a molecular unit—are shown in figure 1.

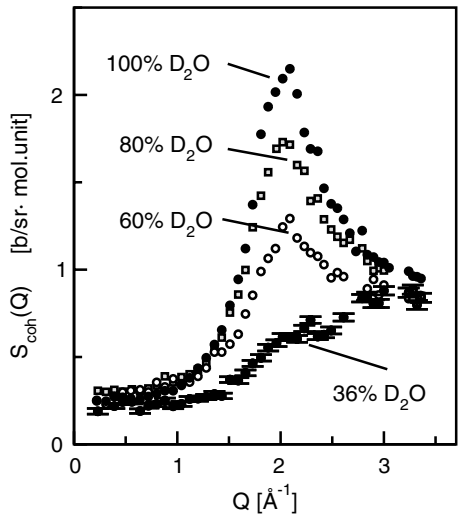

Figure 1. Experimental coherent structure factors of the different isotopic mixtures of pure water.

A further analysis could be (i) a separation into partial scattering $S_{ij}(Q)$ and (ii) a subsequent Fourier transform to obtain distinct spatial correlation functions of H–H, O–H and O–O pairs. One may expect that the present data could be particularly accurate with respect to the hydrogen correlations because of the use of polarization analysis. On the other hand, because of current experimental limitations, the Q-range is smaller compared to previous studies and does not allow for a direct Fourier-transform of the data. This is one but not the only reason why in this study we apply a more general and powerful approach via RMC modelling.

The RMC approach (i) can verify whether the observation is compatible with a possible real structure, (ii) provides pair correlation functions without introducing artefacts from truncation effects, (iii) allows us to search for other structural correlations such as bond angles, which are higher order correlation functions that are more or less constrained by the measured pair correlations, and (iv) allows us to combine the present data with other diffraction data, like x-ray diffraction and/or previous unpolarized neutron diffraction (including a check of their mutual consistency), and/or to introduce further constraints from other experimental observations or theoretical considerations.

3. Reverse Monte Carlo modelling

A general introduction to, and technical details of, reverse Monte Carlo modelling [13] can be found in a recent description of the most up-to-date RMC code, RMC++ [19]. In short, RMC is an inverse method of structural modelling in which atoms are moved around so that structural quantities—in our case, the coherent static structure factors—of the simulation box would be consistent (within experimental errors) with experimental results.

It is instructive to follow how in RMC the volume of the configuration space available for a structural model is being gradually restricted by *constraints*. The simplest one is the density, which, together with particle sizes ('cut-off' distances), defines the packing fraction. For a molecular system, such as liquid water, molecules have to be introduced by using appropriate constraints that keep them together during the entire run. (Note that diffraction data do not contain information on how atoms are connected in molecules, only on the distances; the

molecular structure, therefore, must be provided based on additional—external—knowledge.) Here, the concept of 'fixed neighbour constraints' (FNCs) [19, 20] is applied: that is, in the present case, the 'identity' (serial number in the particle configuration) of the two hydrogen atoms bound permanently to an oxygen atom is preserved. The O–H intramolecular distance is allowed to vary between 0.95 and 1.05 Å, whereas a variation of the intramolecular H–H distance between 1.5 and 1.7 Å is permitted. This way, flexible 'V-shaped' molecules could be defined. The next—and in many cases, final—level of constraints is the introduction of diffraction data.

Simulations were performed in a cubic box with an edge of 53 Å length. The number of molecules was $N = 5000$, so that the atomic number density of 0.1 Å^{-3} was achieved.

The calculations were started from a random arrangement of molecules; this particle configuration was generated by hard sphere Monte Carlo (HSMC) simulation (the HSMC option is built into the RMC++ code). This random (FNC constrained) hard sphere configuration of water molecules also served as a reference system during analyses of the resulting structural models (see below).

In RMC it is possible to further restrict the volume of the configuration space available for a given simulation, by imposing additional geometrical constraints on the arrangements of molecules. For liquid (and also amorphous) water, the application of such coordination constraints proved to be fruitful [12, 21]. The constraints in question were aimed at maximizing the number of hydrogen bonds (H-bonds) in the system—the rationale behind this intention is that, based on many sources of information, a well defined network of H-bonds is generally believed to be an essential constituent of liquid water. (Note, again, that information on the inter-molecular *topology*, that is, on the way hydrogen bonds connect water molecules, is not provided by diffraction experiments, similarly to the case of the intra-molecular connections of atoms.)

Since no experimental (diffraction) data are infinite and infinitely accurate, one cannot expect that any evaluation of these data would provide a uniquely defined structure, even at the level of two-body correlations. This means that there is always a range of structural models that are consistent with a given set of diffraction data (and possibly with additional constraints, too, as in our case). To be able to estimate this *range* of possible structures, particularly in terms of the extent of hydrogen bonding, the essential intermolecular constraint, the one that requires that each O atom should have two H neighbours within the (broadly defined) H-bonding distance range of 1.7–2.0 Å, was introduced, similarly to [12, 21]. These limiting values allow H-bonding distances (as defined by the peak position of the O–H PRDF) proposed previously [2, 3, 12, 22] to be realized, between 1.8 and 1.9 Å. Comparisons between 'constrained' and 'unconstrained' models, as well as between (reference) HSMC and RMC models, will be made, based on the outcome of four calculations. Finally, as a consistency check between our polarized neutron diffraction data and the best known set of PRDFs [22], the O–O, O–H and H–H partials (obtained from the website of the ISIS facility, UK [23]) have also been incorporated into our calculations.

RMC++ uses the following definition for the distinct part of the total coherent scattering functions, $F(Q)$:

$$F(Q) = \sum_i \sum_j c_i c_j \bar{b}_i \bar{b}_j (A_{ij}(Q) - 1) \tag{3}$$

where c_i is the concentration and $\overline{b}_i$ the coherent scattering length for species i in the sample. $A_{ij}(Q)$ are the Faber–Ziman partial structure factors (O–O, O–H and H–H). The transformation between $S_{\text{coh}}(Q)$ (see figure 1) and $F(Q)$ involves only a multiplicative factor and a constant; the RMC++ software carries out this simple transformation automatically.

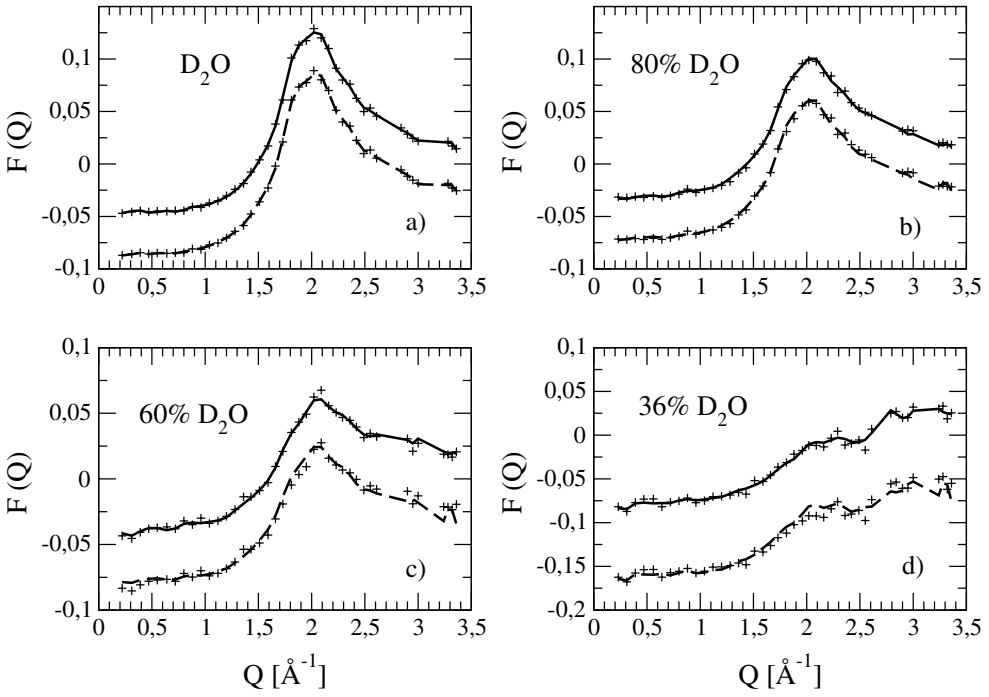

Figure 2. Total coherent scattering functions of the different isotopic mixtures of pure water (symbols), together with reverse Monte Carlo fits (solid lines; shifted by 0.05 (0.07 for 36% D_2O)). Dashed lines correspond to RMC calculations where partial radial distribution functions from [22] were also modelled. (a) 100% D_2O; (b) 80% D_2O; (c) 40% D_2O; (d) 36% D_2O. Note the missing parts around 2.7 and 3.1 $Å^{-1}$. (Symbols represent experimental points and RMC results were calculated at the same Q-values only. Connecting lines are only guides to the eye.)

4. Results and discussion

In figure 2 (upper curves), RMC fits to the experimental total coherent scattering functions of light/heavy water mixtures are shown. All of the four experimental data sets were modelled simultaneously. The quality of the fits did not deteriorate when the 'H-bond' coordination constraint was introduced, indicating that the required local arrangement is also consistent with the measured data (among many others, see below).

The level of agreement with each of the coherent functions $F(Q)$ is excellent, which immediately demonstrates the power of polarized neutron diffraction experiments for liquid water (and, as a matter of fact, for other ^1H containing materials like polymers—see also [24]). Data corrections are straightforward and an unprecedented level of consistency with an extensive set of $S^{coh}(Q)$ of hydrogenated water samples could be achieved by RMC effortlessly, using two (constrained and unconstrained) structural models.

The lower curves in figure 2 correspond to RMC calculations where the four experimental coherent functions $F(Q)$ were modelled together with the three (O–O, O–H and H–H) partial radial distribution functions reported in [22] (that is, seven data sets were considered simultaneously). The (rather high) level of agreement indicates that there is no substantial discrepancy between PRDFs derived on the basis of non-polarized neutron diffraction and the present experimental data. It is not possible to judge at this stage whether the small deviations observed for the coherent scattering functions corresponding to samples with higher hydrogen

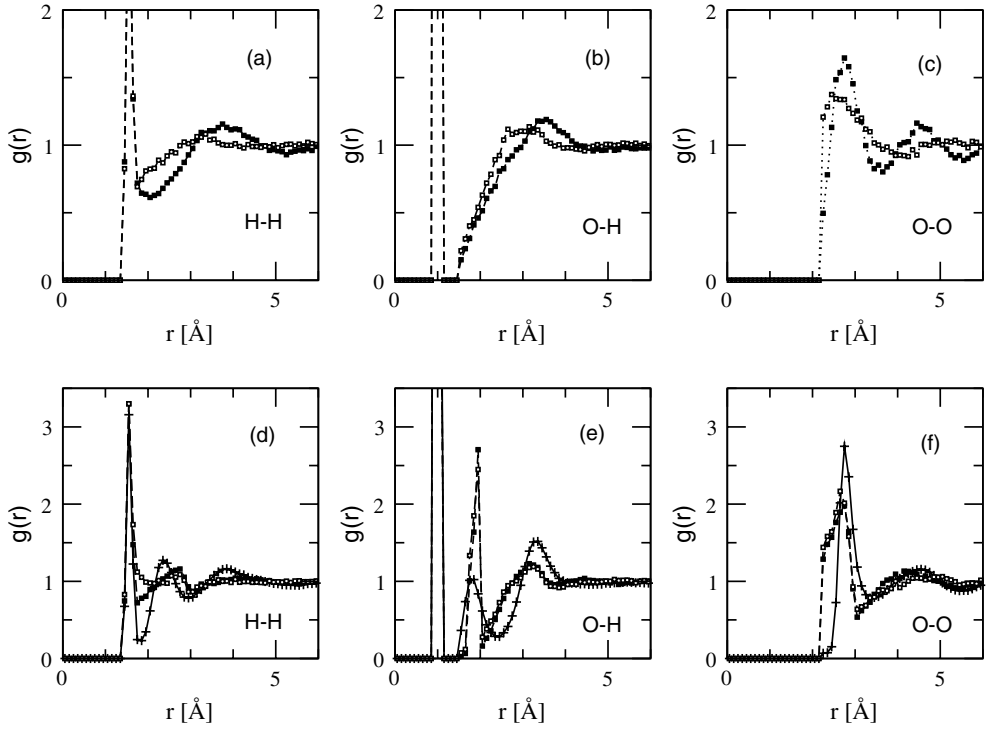

Figure 3. Partial radial distribution functions for each calculation of the present study. Upper half (parts (a)–(c)), unconstrained calculations; lower half (parts (d)–(f)), constrained calculations, as well as results from simultaneous fitting of the PRDFs of [22] (lines with '+' signs). Lines with empty symbols, HSMC; lines with full symbols, RMC.

content (figure 2, parts c and d) are significant. The fact that these discrepancies appear on exactly these functions $F(Q)$ is consistent with common sense expectations: it is the higher hydrogen content that makes experiments (and their evaluation) more difficult.

O–H, H–H and O–O partial pair correlation functions are shown in figure 3 for each calculation. Looking at the unconstrained O–O and H–H partials, it becomes obvious that experimental data do represent important constraints on the volume of the available configuration space, even though the Q-range here is extremely limited. Note that, at present, it is far from trivial to find polarized neutron instruments that would be able to cover a substantially wider scattering vector range—that is, one must find other means for selecting meaningful structures from the (rather wide) choice allowed by the data themselves. This is where the 'H-bond' constraint can play an important role, as is shown in the following paragraphs in detail. Indeed, as exemplified by the O–H and O–O PRDFs (see figure 3, parts (b), (c), (e) and (f)), the 'H-bond' constraint seems to possess a very strong ordering ability—without causing a noticeable deterioration in terms of the quality of the fit to experimental structure factors.

The O–H PRDF needs a longer discussion, partly because it is the most important one (with direct relevance to hydrogen bonding) but also because it is the one that exhibits the most remarkable behaviour of the three. The first thing to note is that on the unconstrained result no sign of the—very much expected—hydrogen bonding can be found. This is partly because of the closeness of the weighting factors of the O–H and H–H partials, which combines

unfavourably with the closeness of several O–H and H–H peak positions between 1.6 and 2.3 Å. One must not, however, underestimate the importance of the lack of high Q information—note that it is exactly at high Q that the peak positions at low r may be resolved (instead of mixing them together, as in the present case). Note, however, that when the coordination constraint is applied then a huge peak appears in this region (between 1.7 and 2.0 Å); the position of the maximum is 1.95 Å. When data are modelled (RMC run), the peak represents a situation where about 80% of the O atoms have perfect H-bonding with two neighbouring H atoms. This finding may also be formulated as the 'H-bonding' constraint re-shaped the topology of the configuration space available for solutions and, within this space, a new (sharp) minimum, corresponding to the O–H intermolecular distance of about 1.95 Å, was found.

It is interesting to note here that when a constrained RMC simulation was started from the final configuration of the unconstrained RMC the goodness-of-fit parameter, χ^2, started to decrease; the H-bonding constraint helps to find regions of the configuration space that correspond to—somewhat, by about 15%—better agreement with the present experimental data. Increasing the number of exactly twofold coordinated O atoms up to a ratio of about 50% was easy and χ^2 was quite sharply decreasing until reaching an 'H-bonded' percentage of 46%. Starting from the other end, from the resulting configuration of the constrained RMC calculation with the H-bonding constraint fulfilled up to about 80% and loosening the constraint, χ^2 also started to decrease. This indicates that most probably the value of 80% is too high for achieving the best possible agreement with our experimental data. A further, more extensive series of constrained RMC calculations is underway, with the aim of finding the most appropriate regions of configuration space corresponding to the present set of diffraction data.

Comparing the constrained and unconstrained simulations and taking into account the findings just described, one can make a statement concerning the range of possible structures that are consistent with the present set of polarized neutron diffraction data: the percentage of oxygen atoms with exactly two H-bonded hydrogen atoms would most probably fall between about 10% and 80%. We note here that the value of 80% was achieved by constrained HSMC simulations; on further, very long, RMC calculations, this value has not changed, even though the requirement for increasing this ratio was still active during the calculation.

It is the upper bound which is the more important: results of computer simulations giving a higher percentage should be taken with some suspicion. Nevertheless, narrowing the above range must be one of the (if not 'the') most important task(s) ahead. Apart from substantially extending the scattering vector range in polarized neutron diffraction experiments (towards which efforts are underway), combination with neutron and x-ray diffraction data on D_2O seems feasible.

One should note that the constrained HSMC and constrained RMC calculations provided nearly identical O–H PRDFs. This indicates that the arguably most essential feature of water structure, the H-bonded network, can be approximated via very simple considerations: the molecular structure and only one ('H-bond') coordination constraint. This finding is even more remarkable if we note that for the H–H and O–O partials noticeable deviations are found between HSMC and RMC results, suggesting that perhaps their behaviour is influenced by more subtle factors.

The position of the H-bonding (O–H) peak at 1.95 Å also deserves a note. First of all, this value is consistent with findings of previous reverse Monte Carlo calculations on liquid water [12] and is slightly higher than the value of 1.8–1.85 Å found on the basis of more conventional isotopic substitution neutron diffraction measurements [2, 3, 22] (see figure 3, part (e)). It is instructive to notice that the value of 1.95 Å appears for the HSMC calculation, as well, even though the (maybe naive) expectation for the hard sphere behaviour would be the exploitation of the lowest allowed value (at 1.7 Å). This indicates that in a 'constrained

random' (hard sphere) arrangement (bonding and closest approach) relations are such that, without further bias (see below), 'bonds' shorter than 1.95 Å do not want to form.

Concerning the O–O PRDF, the influence of the coordination constraint (remember, imposed on the O–H partial!) is enormous, particularly as far as the intensity of the first peak is concerned. The constrained HSMC and RMC results differ visibly beyond 3 Å, which, again, reflects the considerable information content of data taken over even a very limited Q-range.

A further point to note is that at the first minimum, the value of the O–O $g(r)$ is closer to unity than to zero, indicating that coordination shells are not well defined, even in the constrained case. For the unconstrained runs, even the position of the minimum is rather vague. For these reasons, the O–O coordination number is not a very well suited structural characteristic. Just to give an indication that the calculations are meaningful, we note that for the constrained configurations values very close to 4 were found for the O–O (partial) coordination number.

Concerning the H–H PRDF (figure 3, parts (a) and (d)), it is obvious at a glance that the ordering power of the 'H-bond' constraint influences the orientation of water molecules most strongly. This is reflected in the huge change of the H–H PRDF between the 'on' and 'off' statuses of the constraint. It can also be noted here that (considerable) differences between HSMC and RMC results provide further proof that measured data are essential for narrowing the range of possible structures even when the main structural feature of a material (in this case, the H-bond) can be captured via simpler (geometrical) means.

The lower part of figure 3 contains PRDFs obtained by also fitting the partials taken from [22]. As was visible from figure 2, these PRDFs are consistent with our polarized neutron data (at least) nearly within errors, just outside the margin achieved by fitting the coherent functions $F(Q)$ only. Still, differences in terms of the PRDFs are huge, particularly in the region of the H-bonding distance, between 1.5 and 2.5 Å. The first thing to note here is that the experimental data introduced here do allow such a diversity. This, in the light of the discussion above (and of [12]), is not entirely surprising, perhaps apart from one feature, the mean H-bonding distance. For this—extremely important—value, again, only lower and upper bounds can be given on the basis of the present data: the position of the first intermolecular maximum of the O–H PRDF may be between 1.85 and 1.95 Å. To narrow this range, polarized neutron data over a wider Q-range would be necessary—quite obviously, our strong ('H-bonding' and 'PRDF' (from [22])) constraints confine the available configuration space more effectively than the data themselves.

From the configurations, further characteristics of the structure can be calculated. One of the simplest of these is cosine distributions of 'bond angles' (i.e. of angles that can be defined within a given distance range from the centre; for a more detailed description, see, e.g., [25]). In figure 4, first the H–O–H intramolecular angles are displayed. Clearly, the FNCs are able to keep the molecular geometry sensible: the most probable H–O–H angle is very nearly (the canonical) 106° in each case.

The O–H . . . O angles exhibit a clear peak at (or very near) 180° for most calculations; the exception is the unconstrained HSMC run. It is still remarkable that even this distribution shows no resemblance at all to a (common) hard-sphere-like local arrangement. That is, surprisingly, already the density and the molecular geometry (without the coordination constraint!) can capture some characteristics of the structure of liquid water.

O . . . O . . . O angles are characterized within the—not very well defined—first coordination shell. A very weak tendency for the formation of tetrahedral angles ($\cos\Theta = -1/3$) may be spotted—but not for the HSMC results. Also note that in this case the 'H-bond' coordination constraint has not caused sharpening of the (rather flat) maximum. However, the overall picture is rather disordered, as the peak at around 60° ($\cos\Theta = 1/2$) signifies. The

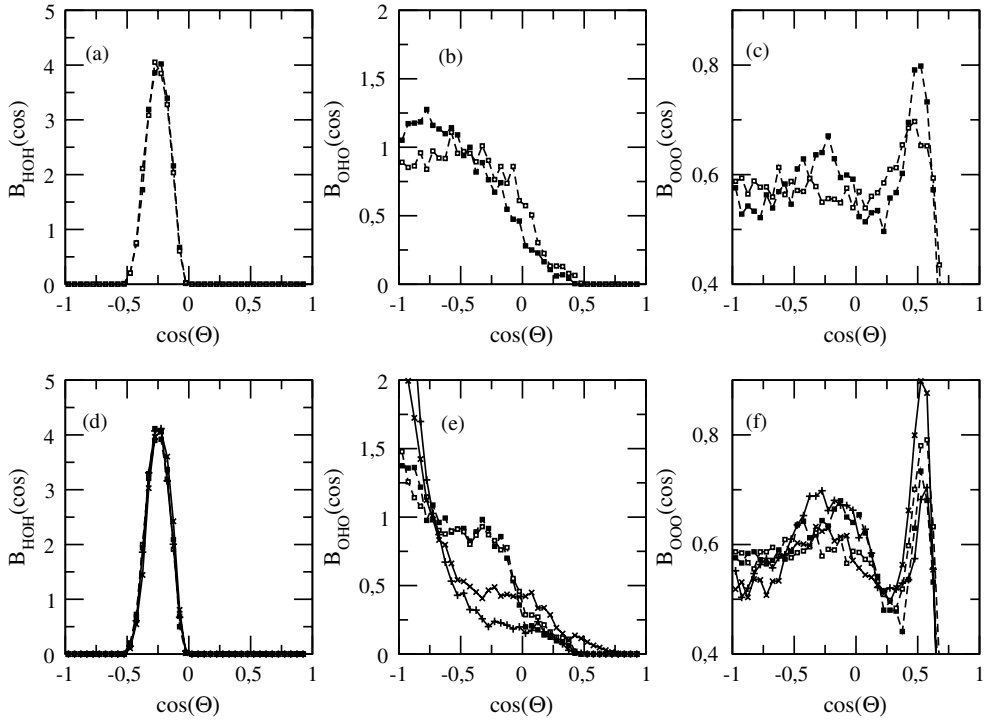

Figure 4. Distribution of the cosines of H–O–H (parts (a) and (d)), O–H . . . O (parts (b) and (e)) and O . . . O . . . O (parts (c) and (f)) angles. Upper half, unconstrained calculations; lower half, constrained calculations, including results from (simultaneous) modelling of (the present coherent functions $F(Q)$ and) PRDFs from [22]. Lines with empty squares, HSMC; lines with full squares, RMC with modelling functions $F(Q)$ only; lines with 'x', RMC with modelling PRDFs of [22]; lines with '+', RMC modelling both functions $F(Q)$ and PRDFs.

importance of these angles is probably exaggerated, and by appropriate angular constraints their occurrence could be lowered. This has not been attempted here, for the reason that no experimental information on the exact number of such angles is available—that is, their presence cannot safely be excluded.

In each part of figure 4, results from modelling the PRDFs of [22] only, as well as from the simultaneous modelling, together with the coherent functions $F(Q)$ measured in this work, are shown. Important observations may be made concerning both the O–H . . . O (parts (b) and (d)) and the O . . . O . . . O (parts (c) and (f)) distributions: in both cases, the inclusion of the present (polarized neutron) data resulted in a better defined structure. In the former the ratio of straight hydrogen bonds increased, whereas in the latter the ratio of (roughly) tetrahedral angles became significantly larger (and in both cases the ratio of irregular angles decreased substantially). This is another indication that the present data, even though strongly restricted in Q, do represent an important constraint on the volume of the configuration space available for water models.

Note also that, in contrast to $g_{\mathrm{OH}}(r)$, the $B(\cos\Theta)$ values of unconstrained RMC calculations do indicate the presence of nearly straight hydrogen bonds; that is, different characteristics reflect the same feature to rather different extents (cf. the enormous peak at 1.95 Å on the O–H PRDF calculated from the constrained runs). This is why as many different aspects of the structure have to be revealed as possible and why structural modelling (like RMC) is(/should be) an essential part of diffraction data analyses.

5. Summary and outlook

The primary aim of the present study was to establish whether polarized neutron diffraction, in combination with reverse Monte Carlo structural modelling, may serve as a standard tool for structural investigations of hydrogeneous (with ^{1}H) materials. To this end, the coherent static structure factors of liquid water samples with high (maximum 64%) light water content have been determined by using the DNS instrument in FZ Jülich (Germany). After a routine (meaning also fast and problem-free) correction procedure, the resulting structure factors could easily be modelled by the RMC method, achieving an unprecedented level of consistency between model and diffraction data from a hydrogeneous material. It has therefore been established that the 'protocol' works well (almost beyond expectations).

Possibilities for further improvements are (a) trying to extend the available scattering vector range, by 'non-standard' use of existing polarized neutron diffraction instruments, (b) the inclusion of scattering data from pure H_2O, which should not be a problem when applying polarization analysis, and which should provide further relevant information particularly on O–H correlations, and (c) combining results from DNS with conventional (neutron and x-ray) diffraction results (over a wider Q-range).

During the interpretation of the present results it soon became apparent that the momentum transfer range of the present experiment would not allow for a more unambiguous distinction between peak positions of O–H and H–H PRDFs at low r than done before. For this reason, a geometrical constraint, for maximizing the number of 'perfectly H-bonded' O atoms, has been used extensively. This coordination constraint required that each oxygen atom had to have exactly two hydrogen neighbours between 1.7 and 2.0 Å, which distance range allows for the formation of both shorter [2, 3, 22] and longer [12] hydrogen bonds. Hard sphere reference systems have been constructed and partial radial distribution functions taken from [22] have been simultaneously modelled. Via these auxiliary means the following conclusions can be drawn concerning the structure of liquid water.

(i) As derived by RMC modelling of the DNS data without and with the 'H-bonding' constraint, the percentage of O atoms with exactly two H atoms in the H-bonding distance range may be between about 10 and 80%, if consistency with the present data is to be preserved. Extending data to be modelled will determine this range to a better accuracy (and in principle optimum values outside of the range suggested here may also prove to be acceptable).

(ii) The H-bonding distance may be between 1.85 (as follows from modelling DNS data together with PRDFs from [22]) and 1.95 Å (suggested by RMC modelling with the 'H-bonding' constraint).

(iii) This distance of 1.95 Å also appears in the HSMC calculations using the 'H-bonding' constraint, suggesting that it is actually geometrical factors (molecular shape; closest approaches) that set this value.

(iv) Although the main feature of the structure, the H-bond, can be described surprisingly well with the help of just one geometrical constraint, for further details of the microscopic structure, like the exact shape of the H–H PRDF or the ratio of straight hydrogen bonds, the present (even though rather restricted) data prove to be indispensable.

Acknowledgments

LP is grateful to the Forschungszentrum Jülich for for providing beamtime and financial support for the experiment on the DNS instrument. Financial support was provided by the Hungarian Basic Research Fund (OTKA), under grants Nos T048580 and T043542.

References

[1] Thiessen W E and Narten A H 1982 *J. Chem. Phys.* **77** 2656
[2] Soper A K 1994 *J. Chem. Phys.* **101** 6888
[3] Soper A K, Bruni F and Ricci M A 1997 *J. Chem. Phys.* **106** 247
[4] Hura G, Sorenson J M, Glaeser R M and Head-Gordon T 2000 *J. Chem. Phys.* **113** 9140
[5] Hura G, Russo D, Glaeser R M, Head-Gordon T, Krack M and Parrinello M 2003 *Phys. Chem. Chem. Phys.* **5** 1981
[6] Dore J C, Garawi M and Bellissent-Funel M-C 2004 *Mol. Phys.* **102** 2015
[7] Sprik M, Hutter J and Parrinello M 1996 *J. Chem. Phys.* **105** 1142
[8] Silvestrelli P L and Parrinello M 1999 *J. Chem. Phys.* **111** 3572
[9] Schwegler E, Galli G and Gygi F 2000 *Phys. Rev. Lett.* **84** 2429
[10] Wernet Ph, Nordlund D, Bergmann U, Cavalleri M, Odelius M, Ogasawara H, Naslund L A, Hirsch T K, Ojamae L, Glatzel P, Nilsson A and Pettersson L G M 2004 *Science* **304** 995
[11] Head-Gordon T and Johnson M E 2006 *Proc. Natl Acad. Sci.* **103** 7973
[12] Pusztai L 1999 *Phys. Rev.* B **60** 11851
[13] McGreevy R L and Pusztai L 1988 *Mol. Simul.* **1** 359
[14] Brückel Th and Schweika W (ed) 2002 *Polarized Neutron Scattering* (*Schriften des Forschungszentrums Jülich, Reihe Materie und Material/Matter and Materials* vol 12) (Jülich: Forschungszentrum Jülich GmbH)
[15] Squires G L 1978 *Thermal Neutron Scattering* (Cambridge: Cambridge University Press)
[16] Brückel Th, Heger G, Richter D and Zorn R (ed) 2003 *Laboratory Course on Neutron Scattering* (*Schriften des Forschungszentrums Jülich, Reihe Materie und Material/Matter and Materials* vol 15) (Jülich: Forschungszentrum Jülich GmbH)
[17] Gereben O and Pusztai L 1995 *Phys. Rev.* B **51** 5768
[18] Schweika W and Böni P 2001 *Physica* B **297** 155
[19] Evrard G and Pusztai L 2005 *J. Phys.: Condens. Matter* **17** S1
[20] Pusztai L and McGreevy R L 1997 *NFL Studsvik Annual Report for 1996* OTH:21, NFL Studsvik University of Uppsala
[21] Pusztai L 2000 *Phys. Rev.* B **61** 28
[22] Soper A K 2000 *Chem. Phys.* **258** 121
[23] http://www.isis.rl.ac.uk/disordered/Database/DBMain.htm
[24] Genix A-C, Arbe A, Alvarez F, Colmenero J, Schweika W and Richter D 2006 *Macromolecules* **39** 3947
[25] McGreevy R L and Pusztai L 1990 *Proc. R. Soc.* A **430** 241