Assessing the level of consistency between diffraction experiments and interaction potentials: A combined molecular dynamics (MD) and Reverse Monte Carlo (RMC) approach

László Pusztai [a,*], Ildikó Harsányi [a], Hector Dominguez [b], Orest Pizio [c]

[a] Research Institute for Solid State Physics and Optics, Hungarian Academy of Sciences, H-1525 Budapest, P.O. Box 49, Hungary
[b] Instituto de Investigaciones en Materiales, Universidad Nacional Autonoma de Mexico (UNAM), Ciudad Universitaria, Circuito Exterior s/n., Coyoacan D. F. 04510, Mexico
[c] Instituto de Quimica, UNAM, Ciudad Universitaria, Circuito Exterior, Coyoacan D. F. 04510, Mexico

ARTICLE INFO

ABSTRACT

A protocol for finding the applicability limits of interaction potential models, used in molecular dynamics (MD) computer simulations, is presented. The essence of the approach is to combine experimental structure factors and radial distribution functions from MD simulations in one single structural model, generated by Reverse Monte Carlo modeling. This way, it becomes possible to tell which parts of the structure are represented by the potential model in question. As an example, we use aqueous rubidium bromide solutions. We show that a great deal of structural information from MD may be consistent with neutron diffraction experiments at lower concentrations.

© 2008 Elsevier B.V. All rights reserved.

1. Introduction

Comparison with experimental data is of primary importance for validating results of computer simulation methods [1] that use interaction potential models. As an essential part of this comparison, structural quantities from molecular dynamics (MD) and Monte Carlo (MC) simulations are calculated and contrasted to their experimental counterparts. The quantity most frequently used for this purpose is the radial distribution function [rdf, $g(r)$] [2]. It should be noted, however, that the rdf is not a genuine experimental result but rather, an interpretation of the primary experimental outcome, the total scattering structure factor (tssf), $F(Q)$; this fact has been demonstrated in Ref. [3].

Since the tssf has a central role in the present study, it may be helpful to mention that $F(Q)$ is defined throughout this contribution according to Ref. [4]:

$$G(r) = \sum_{i,j=1}^{n} c_i c_j \bar{b}_i \bar{b}_j [g_{ij}(r) - 1] \tag{1a}$$

$$F(Q) = \rho_0 \int_0^\infty 4\pi r^2 G(r) \frac{\sin Qr}{Qr} \, dr \tag{1b}$$

In Eqs. (1a) and (1b), c_i and b_i are the molar ratio and the scattering length of species i, $g_{ij}(r)$ are the partial radial distribution functions, $G(r)$ is the total radial distribution function, ρ_0 is the number density and Q is the scattering variable (proportional to the scattering angle); indexes i and j run through nuclear species

of the system. There are, unfortunately, relatively few simulation studies available where the calculated tssf is shown together with the same function from a diffraction experiment; for an old and a recent example, see Refs. [5,6], respectively.

A further, major difficulty lies in the assessment of the quality of the match between experimental and simulated tssf/rdf. Agreement within experimental uncertainties would validate the potential model in question from the structural point of view. Only in the case of such a perfect match it would be justifiable to draw detailed conclusions concerning the partial rdf's and higher order distribution functions, particularly if the number of experimental datasets is lower than the number of partial rdf's. Regrettably, MD and MC simulations hardly ever (if at all) produce an agreement between experimental and calculated tssf's within experimental uncertainties. In the most common cases, where discrepancies between experiment and simulation are apparent, visual inspection (and judgment based upon that) is the tool of assessment [5]. Such assessments are necessarily biased and result in statements like 'our simulation reproduces diffraction data qualitatively/semi-quantitatively'. In the overwhelming majority of cases (like the ones described in Refs. [5,6]), detailed description of structural features is only valid for the model mimicking the real system and *not* for the real system itself. On the other hand, potential models (used for, for instance, simulating electrolyte solutions [1,2,5–7]) certainly capture a good deal of structural features of the corresponding real systems – it is just not possible at present to tell which are the features that are and that are not compatible with available results of diffraction experiments.

The main goal of the present work is to provide a possible tool for a detailed validation of interaction potentials which are applied

* Corresponding author. Fax: +36 13922589.
E-mail address: lp@szfki.hu (L. Pusztai).

with the aim of describing real systems, from the point of view of the microscopic structure. We wish to achieve this goal by combining the experimentally determined total scattering structure factor(s) and partial radial distribution functions from molecular dynamics simulations in one structural model, to be constructed by the Reverse Monte Carlo (RMC) method [8–11]. We require perfect agreement (within experimental uncertainties) with diffraction data and see how well the potential-based partial rdf's can be approached at the same time. In this way, it will be possible to tell if (or which of) the simulated partial radial distribution function(s) is(/are) consistent with the experimental tssf('s), thus shedding light on the strengths/weaknesses of the (most frequently, pairwise) potential model.

For demonstrating the capabilities of the new scheme, the example of aqueous rubidium bromide solutions was chosen. This system is rather comfortable for the purpose since recent neutron diffraction data at two concentrations [12], as well as molecular dynamics simulations at five concentrations [6] are available. Varying the concentration helps to explore the range in which a given potential model may be applied; for this reason, the proposed scheme looks particularly promising for studying solutions (and mixtures in general).

2. The computational method: Reverse Monte Carlo modelling

As details of the Reverse Monte Carlo method have been described in several publications [8–11], here we only provide a brief practical introduction, focusing on the feature that is essential concerning the purpose of this work.

Reverse Monte Carlo is a simple tool for constructing large, three-dimensional structural models that are consistent with the total scattering structure factors (within the estimated level of their errors) obtained from diffraction experiments. Via random movements of particles, the difference between experimental and model total structure factors (calculated similarly to the χ^2-statistics) is minimised. As a result, by the end of the calculation a particle configuration is available that is consistent with the experimental structure factor(s). If the structure is to be analysed further, partial radial distribution functions, as well as other structural characteristics (neighbour distributions, cosine distribution of bond angles) can be calculated from the particle configurations.

For the present purposes the most attractive feature of the RMC method is that it can take any external information (diffraction and EXAFS data, as well as ideas based on other experimental or theoretical sources) that can be calculated directly from the coordinates of the particles. Obviously, partial radial distribution functions from MD simulations fall into this category. In this case, if consistency with all input data is reached then it may be stated that these input data are consistent with each other, as well as with the resulting particle configuration. If, however, some of the input data cannot be approached within their uncertainties then it means that parts of the input data set are not consistent with other pieces of input information. In our case this would mean that some of the input prdf's from MD would not be consistent with the experimental input total scattering structure factor.

In the RMC calculations that compose the basis of the present research one total scattering structure factor from neutron diffraction [12] and three partial radial distribution functions from the MD simulation [6] are applied as input data for each calculation, at both concentrations. Two sets of MD results were selected as input for RMC. The first one contained O–O, Rb–O and Br–O partial rdf's, characterising water–water, cation–water and anion–water correlations, respectively; these correlations are the ones that appear to be the most sought after in structural studies of electrolyte solutions (see, e.g., [2,5]). The second set of input partial rdf's in-

Table 1
Details of the Reverse Monte Carlo calculations performed in the present study

Calculation code	Concentration (M)	Partial rdf's	Number density (Å^{-3})	Number of ion pairs/ water molecules
Run1	2	O–O, Rb–O, Br–O	0.094	128/3392 (1/26.5)
Run2	5	O–O, Rb–O, Br–O	0.090	270/3000 (1/11)
Run3	2	H–H, O–H, Br–H	0.094	128/3392 (1/26.5)
Run4	5	H–H, O–H, Br–H	0.090	270/3000 (1/11)

Note that in all calculations, the corresponding experimental $F(Q)$ was also modelled, so that the number of data sets was 4 in each case.

cluded the H–H, O–H and Br–H functions (hydrogen was in the form of deuterium); these are the functions that contribute the most to the neutron weighted tssf. We thus carried out four calculations altogether; some details of these RMC runs can be found in Table 1. Naturally, in a more detailed investigation, more than one diffraction datasets and different (and more, than three) simulated functions may be selected, depending on the specific questions posed concerning the real system under study.

The input partial rdf's for the present Reverse Monte Carlo study were results of standard (NVE) molecular dynamics simulations, carried out for aqueous rubidium bromide solutions at several concentrations as described in Ref. [6]. A rigid water model, SPCE [13], was applied [14], while the ionic interactions were mimicked by a 'Coulomb-plus–Lennard–Jones' parameter set [7], so that the (pair) potential energy function between the ith and the jth particles took the following general form:
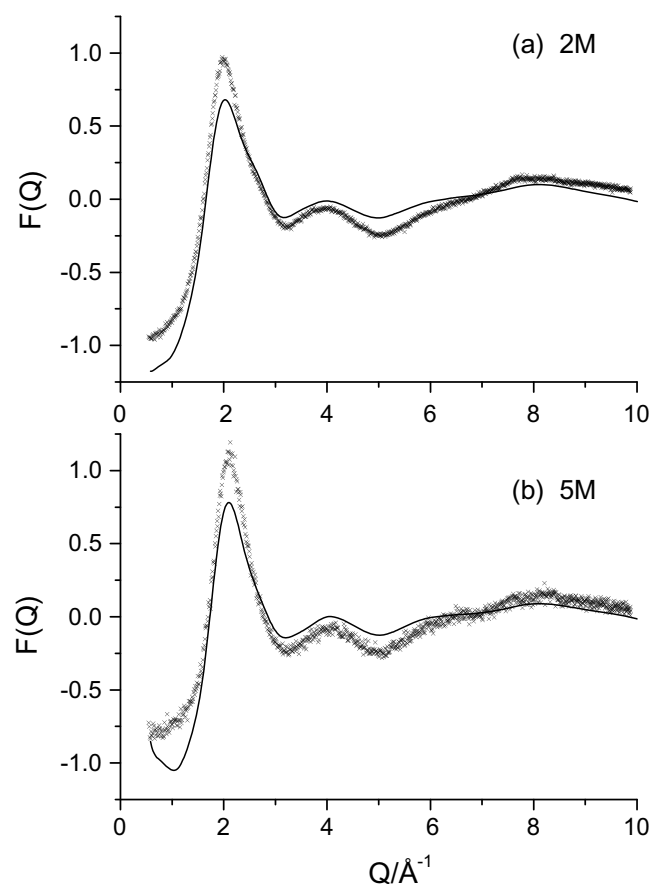

Fig. 1. Experimental (symbols) and MD simulated (solid line; calculated in Ref. [6]) total scattering structure factors of aqueous rubidium bromide solutions. (a) Two molal solution; (b) 5 M solution.

$$V_{ij}(r) = \frac{q_i q_j}{r_{ij}} + \frac{A_i A_j}{r_{ij}^{12}} - \frac{B_i B_j}{r_{ij}^{6}} \qquad (2)$$

In Eq. (2), q_i are charges put on the interaction sites whereas A_i and B_i are (adjustable) Lennard–Jones parameters. The A_i and B_i pair potential parameters for all of the particle pairs were identical to those applied in Ref. [15]. Perhaps it may have been advantageous to make an attempt with a flexible water potential, as well; it was felt, however, that for testing purposes, the – arguably – most popular potential parameters would be sufficient.

The experimental input information, the total scattering structure factors, have been obtained from neutron diffraction experiments conducted for RbBr solutions at two concentrations [2 and 5 molal (to be noted as '2 M' and '5 M', where the ion:water number ratio is 1:13 and 1:5.5, respectively] as described in Ref. [12], using the PSD two-axis neutron diffractometer installed at the Budapest Research Reactor (Hungary) [16].

3. Results and discussions

Fig. 1 provides a direct comparison between simulated [6] and experimental [12] neutron weighted total scattering structure factors. It is clear that although the match is not perfect (i.e., certainly not within experimental uncertainties), the simulated curves do resemble those from neutron diffraction. That is, it is worth trying to find features (partial rdf's in the present case), using the protocol described in the preceding Sections, that are consistent with diffraction data.

Figs. 2 and 3 contrast the cases of 2 and 5 M solutions where the experimental tssf and O–O, Rb–O and Br–O partial radial distribution functions from MD were modelled (calculations Run1 and Run2). The match between experimental and calculated $F(Q)$'s is within the experimental uncertainties for both concentrations, similarly to what happened during the preceeding RMC study [12]. (Note that it was our intention that whatever additional data set is added to the experimental tssf, consistency with experiment was to be maintained.) On the other hand, concerning fits to partial rdf's from MD, the difference between results for the 2 and 5 M salt concentrations is striking: agreement between input (MD) data and RMC is nearly perfect for the 2 M case, whereas for the 5 M solution, none of the three RMC partial rdf's may be called as consistent with the corresponding target function. The conjecture following these observations would be that pair potential parameters [15] applied in the prededing MD study [6] are adequate at lower concentrations whereas in concentrated RbBr solutions, they become inapplicable.
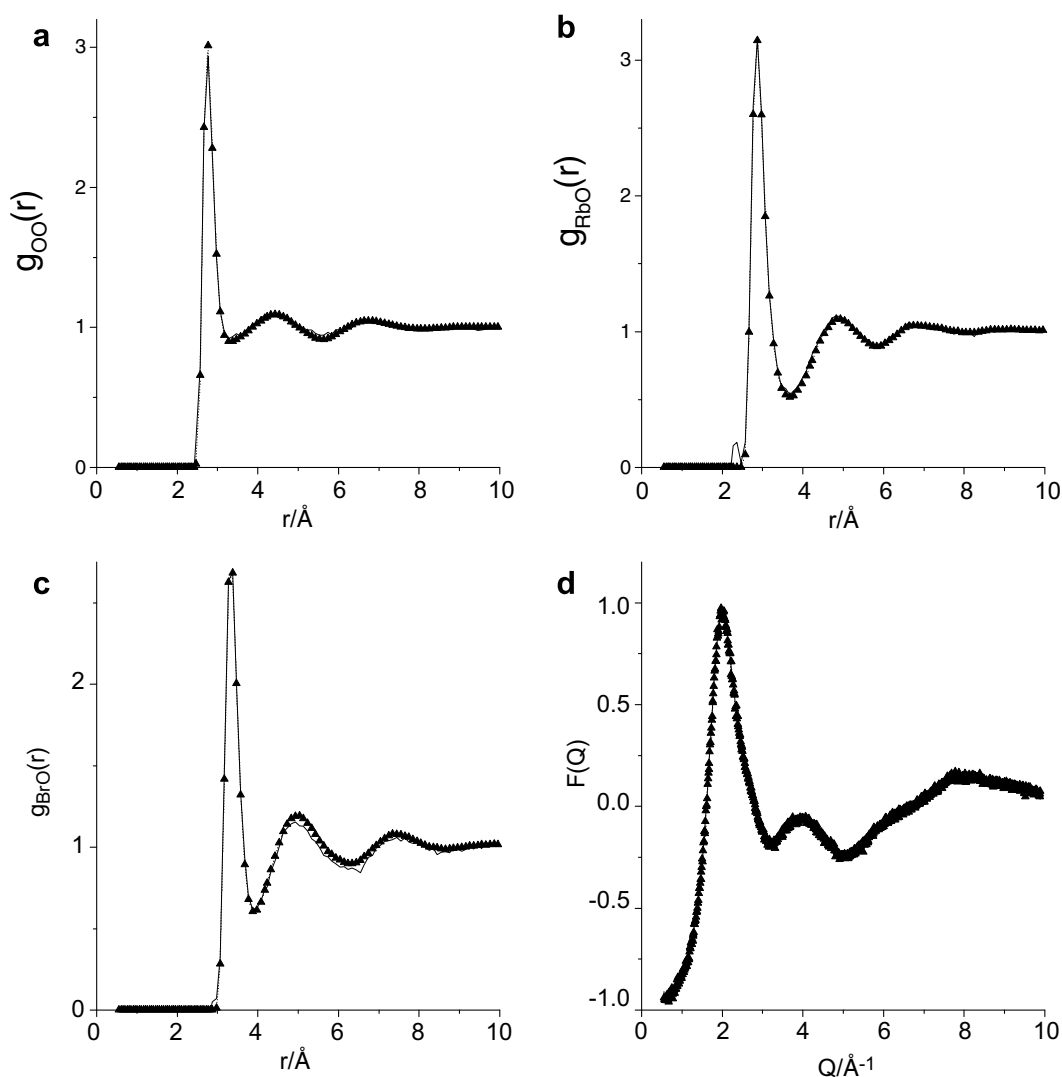

Fig. 2. Results from RMC modelling O–O (a), Rb–O (b) and Br–O (c) partial radial distribution functions from MD simulation, together with neutron diffraction experimental data, for the 2 M solution of RbBr in water. (a–c) Line + symbols: input partial rdf's from MD simulations [6]; solid line: RMC. (d) Line + symbols: experimental total scattering structure factor [13]; solid line: RMC.
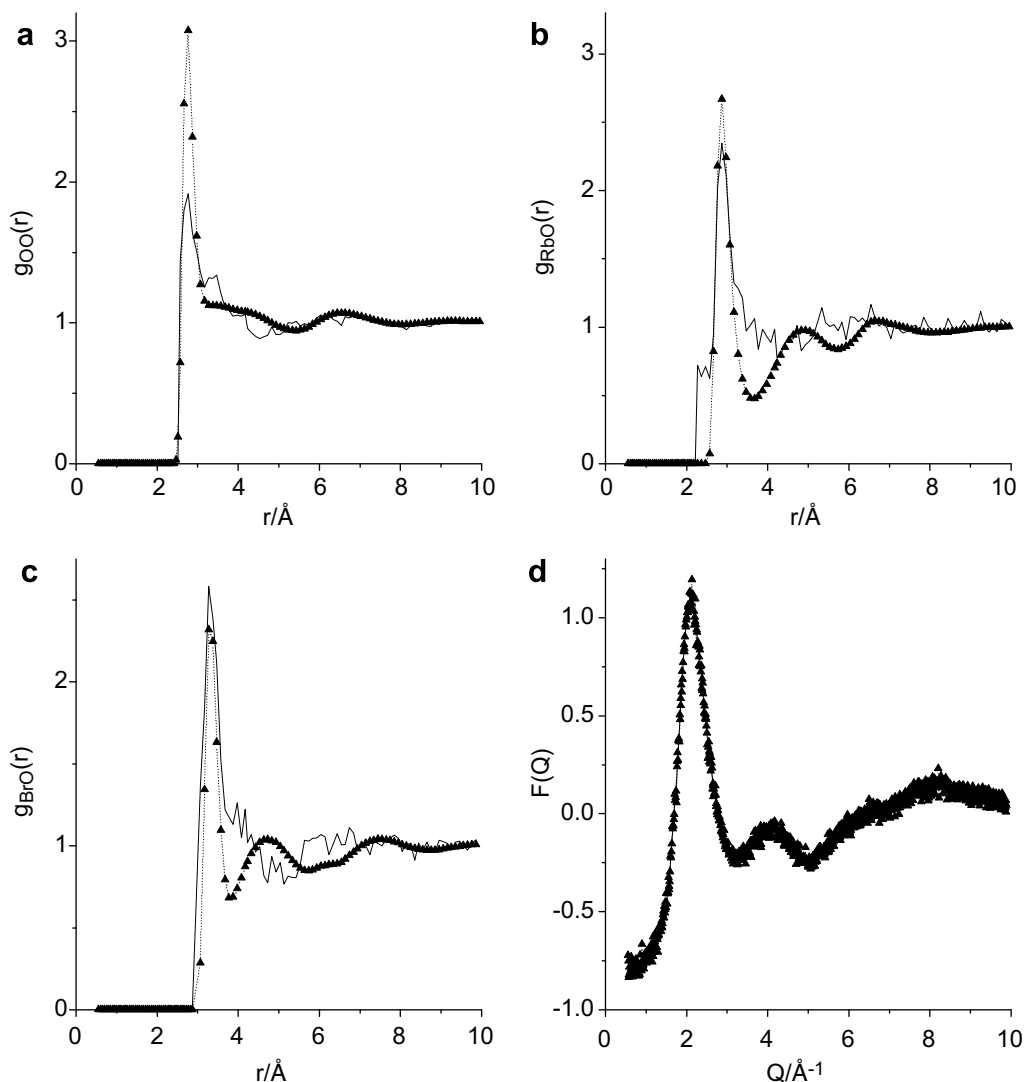
Fig. 3. Results from RMC modelling O–O (a), Rb–O (b) and Br–O (c) partial radial distribution functions from MD simulation, together with neutron diffraction experimental data, for the 5 M solution of RbBr in water. (a–c) Line + symbols: input partial rdf's from MD simulations [6]; solid line: RMC. (d) Line + symbols: experimental total scattering structure factor [13]; solid line: RMC.

In order to check if arriving at this statement was not a consequence of the – from the point of view of scattering properties, arbitrary – selection of partial rdf's used in Run1 and Run2, we have carried out further calculations in which radial distribution functions that have the highest contributions to the tssf were taken from MD (H–H, O–H and Br–H). Figs. 4 and 5 summarise results from these RMC simulations (Run3 and Run4 for the 2 and 5 M cases, respectively). The picture that emerges is very similar to that found previously: while experimental tssf's could be matched within the experimental uncertainties, the new selection of partial rdf's could be fitted only for the less concentrated solution. For both selections of the partial rdf's it is the ionic distributions that pose the larger difficulties whereas water-related rdf's proved to be easier to make consistent with neutron diffraction results.

It now can be stated with confidence that at the lower concentration, where any ion may be surrounded by about 13 water molecules on average, the MD-based prdf's selected for the present study proved to be consistent with neutron diffraction results, since it was possible to generate a structural model (via Reverse Monte Carlo) that agrees with both experimental data and MD simulation results. In this particular case, as well in similar cases,

it can be stated that the selected pieces of the outcome of the molecular dynamics simulation *may be* representative of the corresponding real system and therefore, further interpretation of these pieces is justified (or at least, can be supported strongly).

On the other hand, at the higher concentration, none of the partial rdf's could be made consistent with the corresponding neutron diffraction result (see Figs. 3 and 5); that is, it was not possible to generate a single structure that would bear characteristics of both experiment and simulation. This finding indicates that pair potential parameters that were successful for the lower concentration case are not applicable for 5 M solution (in which less that five water molecules available for any given ion on the average). This finding is in accordance with the general attitude that computer simulations are carried out only for the lower concentration regime of electrolyte solutions (see, e.g., Refs. [2,14]).

It has to be emphasized that the scheme we are proposing is not a potential fitting procedure and therefore, it is not applicable for providing better potential parameters directly. Instead, one might try available potential parameter sets and select the one that provides the highest level of consistency with experiment(s). It should also be stated that even though a perfect consistency between experiment and simulation is achieved, it is no proof that the true
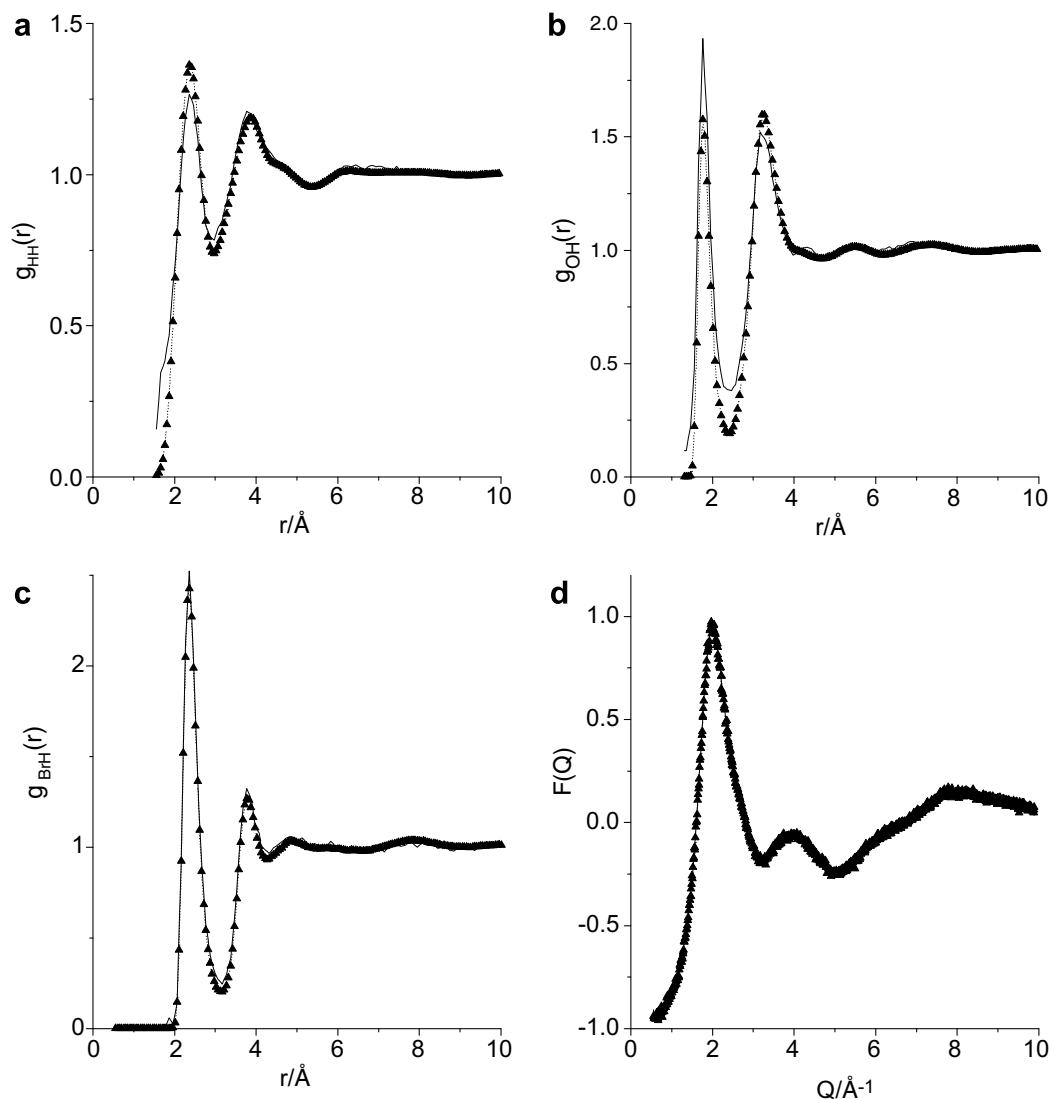
Fig. 4. Results from RMC modelling H–H (a), O–H (b) and Br–H (c) partial radial distribution functions from MD simulation, together with neutron diffraction experimental data, for the 2 M solution of RbBr in water. (a–c): Line + symbols: input partial rdf's from MD simulations [6]; solid line: RMC. (d) Line + symbols: experimental total scattering structure factor [13]; solid line: RMC.

structure is found – such an agreement means that simulation results may be a representation of the real structure.

As a 'side-effect' of the investigations described above, future prospects concerning the applicability of the combination of diffraction data and MD (or MC) computer simulations have become clearer. Present trends point towards an even more increased proportion of studies on complex systems, which almost always contain more than three atomic species. For such multi-component systems, obtaining 'full' structural information (where the number of independent total scattering functions would be equal to the number of partial rdf's) from diffraction (and possibly, from EXAFS) measurements is impossible. In these cases, computer simulations with interaction potentials will gain more emphasis – and the need for validating results from them will be even more important. The scheme suggested here is one possibility for providing a bridge between diffraction experiments and interaction potential models; it is, however, our firm opinion that a tighter connection would be necessary in due course.

As a final thought, it should be noted that not even perfect consistency between diffraction data and simulated partial rdf's can provide a guarantee for that the Hamiltonian in question is capable

of describing the real system considered from every respect. We would like to emphasize the importance of thermodynamic (like the solubility of the salt [17]) and transport (like the self-diffusion constant [2,6]) properties; it would be desirable to attempt to include at least some of these quantities into RMC calculations in the future.

4. Summary and outlook

We have demonstrated the capabilities of a novel combination of molecular dynamics and Reverse Monte Carlo computer simulation methods for a direct, quantitative assessment of the validity of interaction potential models from the structural point of view. The 'merger' of experiment and simulation introduced here (and realized via the construction of large 3D models of the structure) is perhaps the most natural way for spotting details of the interaction models that are (and that are not!) consistent with experimental results.

Our initial results for aqueous rubidium bromide solutions are in accordance with commonsense expectations: at the lower
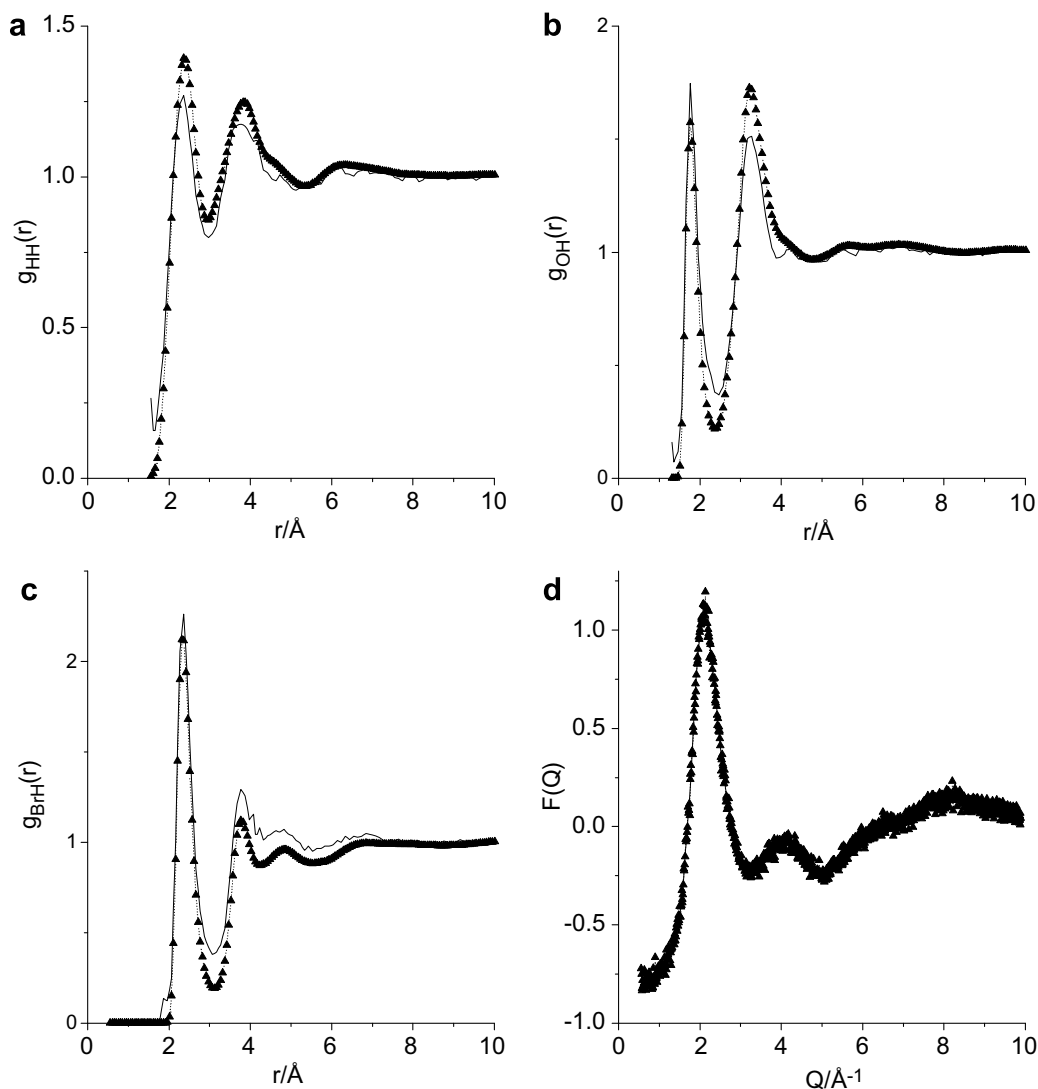
Fig. 5. Results from RMC modelling H–H (a), O–H (b) and Br–H (c) partial radial distribution functions from MD simulation, together with neutron diffraction experimental data, for the 5 M solution of RbBr in water. (a–c): Line + symbols: input partial rdf's from MD simulations [6]; solid line: RMC. (d) Line + symbols: experimental total scattering structure factor [13]; solid line: RMC.

concentration, two sets of three important partial radial distribution functions from the corresponding MD simulation could be made consistent with neutron diffraction data (the best case being the O–O, H–H and Br–H rdf's), whereas at the higher concentration, the entire potential model seems to collapse (the worst being the ionic parts).

The obvious extension of this pilot study is to consider more than one potential models for both water and ionic interactions and more than one sets of diffraction data, as well as to investigate more solutions (like that of rubidium chloride, for which a fairly detailed Reverse Monte Carlo analysis is available [18]). Moreover, in the specific case of aqueous solutions, a similar investigation for pure liquid water, involving rather extensive sets of both diffraction experiments and interaction potential models, should be conducted. Work towards these objectives is underway.

References

[1] M.P. Allen, D.J. Tildesley, Computer Simulation of Liquids, Clarendon Press, Oxford (UK), 1987.
[2] K. Heinzinger, in: C.R. A Catlow, S.C. Parker, M.P. Allen (Eds.), Computer Modelling of Fluids, Polymers and Solids, Kluwer Academic Publishers, Dordrecht, NL , 1990, p. 357.
[3] L. Pusztai, Phys. Rev. B 60 (1999) 11851.
[4] D.A. Keen, J. Appl. Crystallogr. 34 (2001) 172.
[5] M.M. Probst, T. Radnai, K. Heinzinger, P. Bopp, B.M. Rode, J. Phys. Chem. 89 (1985) 753.
[6] I. Harsányi, L. Pusztai, J.-C. Soetens, P.A. Bopp, J. Mol. Liq. 129 (2006) 80.
[7] J. Aqvist, J. Phys. Chem. 94 (1990) 8021.
[8] R.L. McGreevy, L. Pusztai, Mol. Simul. 1 (1988) 359.
[9] R.L. McGreevy, J. Phys.: Condens. Matter 13 (2001) R843.
[10] G. Evrard, L. Pusztai, J. Phys.: Condens. Matter 17 (2005) S1.
[11] O. Gereben, P. Jóvári, L. Temleitner, L. Pusztai, J. Optoelectron. Adv. Mater. 9 (2007) 3021.

[12] I. Harsányi, P. Jóvári, Gy. Mészáros, L. Pusztai, P.A. Bopp, J. Mol. Liq. 131–132 (2007) 60.
[13] H.J.C. Berendsen, J.R. Grigera, T.P. Straatsma, J. Phys. Chem. 91 (1987) 6269.
[14] We note here that in Ref. [6], the TIP4P water potential was said to have been used; we confirm that this statement was an error in Ref. [6].
[15] G. Ferlat, A. San Miguel, J.F. Jal, J.C. Soetens, Ph.A. Bopp, I. Daniel, S. Guillot, J.L. Hazemann, R. Argoud, Phys. Rev. B 63 (2001) 134202.
[16] E. Sváb, Gy. Mészáros, F. Deák, Mater. Sci. Forum 228–231 (1996) 247.
[17] E. Sanz, C. Vega, J. Chem. Phys. 126 (2007) 014507.
[18] I. Harsányi, L. Pusztai, J. Phys.: Condens. Matter 19 (2007) 335208.