Thesis Booklet

Computer modeling investigations concerning the structure of pure water:
consistency between interaction potentials and measured diffraction data

Ph.D. Theses

STEINCZINGER, Zsuzsanna

Secondary school teacher of Mathematics and Chemistry
(Gárdonyi Géza Primary and Secondary School, Érd, Hungary)

Supervisor:
Dr. JÓVÁRI, Pál
(Wigner Research Centre for Physics, Hungarian Academy of Sciences)

Chemistry Doctoral School, Eötvös University, Budapest, Hungary
(director: Prof. INZELT, György)
„Theoretical and physical chemistry, structure of materials" programme
(leader: Prof. SURJÁN, Péter)

2016

I. INTRODUCTION

Water is the most abundant compound on Earth, and the H_2O molecule is the second most frequent in the Universe (following H_2). Living organisms contain 10 to 95 % water; on average, 2/3 of the human body is made of H_2O. Water represents the strongest force that can shape the landscape; it is essential in the photosynthesis, as well as a medium for temperature regulation. Most chemical reactions take place in aqueus solutions. More than half of all the living organisms live in water, which liquid is indipensable for most of activities in life (biological, hygienical, healthcare, transport, production). The most frequent solute is also water, and biological processes take place in water. In order to understand complicated systems, such as those containing proteins, we must understand the behaviour of simpler solutions, and indeed, of pure water itself, in detail.

Beyond its aboundance and practical importance, water is also highly important from the scientific point of view. According to the most widely spread description of the water molecule [Franks 2000], in the „V"-shaped species the O-H distance is 0.958Å and the H-O-H bond angle is 104° 27'. Water in its standard state is a liquid, even though hydrides of the neighbours (in the periodic table) of oxygen are all gases. Another well-known anomaly is the density maximum at 4 °C-on, as well as the cariation of density with temperature. Apart from these, there are more than 60 anomalous properties of water known [Chaplin 2003]; all of these originate from the structure, more precisely, the presence of hydrogen bonds [Baranyai 2011]. Hydrogen bonding is the most important interaction that stabilize the structure of biologically relevant macromolecules. Studying this interaction is the easiest in water, since ammonia and hydrogén-fluride are poisonous gases in their standard state, whereas multicomponent systems, like alcohols, are more complicated [Soper 2013].

For all the above reasons, the material that is most studied by physicists, chemists and biologists is arguably this compound containing two hydrogen and one oxygen atoms, whose atomic/molecular level structure has been investigated with outstanding attention. Despite this fact there are still many uncertainties concerning the liquid structure of water. This statement may be best underlined by mentioning that recently the leading expert of the field, Prof. A. K. Soper has published a review paper with the title „The Radial Distribution Functions of Water as Derived from Radiation Total Scattering Experiments: **Is There Anything We Can Say for Sure?"** [Soper 2013]. It was this situation that inspired us to scrutinize certain details in connection with the structure of water.

The liquid structure of water is the subject of a huge volume of literature; for this reason, we are in possession of many pieces of information. A number of neutron [Soper 1997] and X-ray [Fu 2009, Hart 2005, Skinner 2013] diffraction data sets are available, in the form of the structure factor (SF; $F(Q)$). Computer simulation methods (e.g. molecular dynamics, MD) that work on the basis of

interparticle interactions also provide information on the structure [Baranyai 1995]. The most frequently mentioned structural characteristics, the partial pair correlation functions (PPCF, $g_{ij}(r)$), may be derived on the basis of both diffraction measurements and computer simulations. Ideally, the two routes would lead to a consensual structure that may be called „the structure of water". Unfortunately, this is not the case (see the title of the publication by Soper above), so a quantitative comparison of the various data sets, and revealing the present (or missing) „harmony" („consistency") between them may be instructive for many reasons. This is why, in this dissertation, I have chosen to investigate the question of consistency between $F(Q)$ and $g_{ij}(r)$ functions that come from different sources. Also, I have attempted to characterize the level of consistency quantitatively.

In my PhD dissertation first I have introduced the functions that are available for characterizing the structure of a liquid, the most apparent one being the pair correlation function (PCF). Next, the experimental methods of X-ray and neutron diffraction are mentioned, as well as the resulting structural functions and their relationship with the PCF. Since results from classical computer simulations (molecular dynamics, MD; Monte Carlo, MC) [Allen 1987] are closely related to my thesis work, I have provided a brief overview of them. Special attention has been paid to the Reverse Monte Carlo (RMC) technique [McGreevy 1988], since all my results shown in the dissertation have been obtained via RMC. Finally, calculations performed and results obtained have been described in detail.

II. THE RESEARCH WORK PERFORMED

II.1. Objectives

The motivation for my doctoral activities may be summarized as follows:

I have compared to approches of structure determination of liquid water, diffraction measurements and computer simulations (MC, MD). Considering any of the two approaches, there are quite a few approximations present in the various methods and therefore the final outcome, in our case the partial pair correlation functions, are burdened by uncertainties. Following diffraction experiments, data evaluation is problematic; this is why there is still a permament development activity is going on here. The weak point of computer simulation methods is the interatomic (pair) potential function: this is why for water, new potentials are appearing year by year. Again, in a fortunate situation, the experimental and simulation routes would provide equivalent structural characteristics.

For a thorough characterization of the structure of water it is worth making use of all the knowledge that has been gathered so far. Accordingly, I have exploited knowledge on the geometry of the molecule, on the measured structure factors ($F(Q)$), as well as on $g_{ij}(r)$-s from MD simulations for clarifying issues concerning the liquid structure.

I have investigated (the level of) mutual consistency between measured diffraction data and interatomic potential functions (represented by the the PPCF-s). In order to do so the following have been done:

(1) I've been searching for the interatomic potentials that would be most suitable for reproducing X-ray and neutron diffraction results from pure water. For this, partial $g(r)$-s from MD have been fitted by the Reverse Monte Carlo (RMC) method and from the large RMC particle configurations SF-s measurable via neutron and X-ray diffraction have been computed and compared to measured data. This „protocol" has been applied for finding water potentials that are the most successful from the sructural point of view.

(2) I have tried to contrast, i.e., model together in RMC, partial $g(r)$-s from MD with all possible combinations experimental data. With this approach I wished to find potential functions that would be most suitable for enhancing RMC structural models that originate from direct RMC modeling of diffraction data.

(3) Based on suitable number of independent diffraction experiments (in our case, this number is 3) it is possible to determine the O-O, O-H and H-H PPCF-s via a traditional route (i.e., excluding computer modeling) [Soper 2000]. Partial *g(r)*–s obtained this way carry systematic errors from the measurements (somtimes transforming them into completeley unrecognizable forms). It appeared useful to undetake consistency analyses, similar to what has just been intruduced under point (2), with regard also to such PPCF-s [Soper 2000], using, again, the RMC technique. The question answer here is the how much uncertainties are introduced into the PPCF-s by the process of (traditional) data evaluation.

In order to handle the problems mentioned just above systematically, as well as to make results comparable, I had to perform extensive studies concerning the appropriate values of a number of parameters important during RMC calculations, such as the spacing of the *g(r)* histograms, the precise distance ranges of *g(r)* –s, or the suitable values of the σ parameters. I have considered 4 experimental structure factors and altogether 10 sets of partial pair correlation functions (one set from a traditional evaluation of measured SF-s [Soper 2000] and 9 sets from computer simulations), as well as the consistency of these data sets with each other. I have performed computer experiments by RMC modeling, by which method I have attempted to fit a varying number of data sets simultaneously.

II.2. Methods; the procedure of consistency checks

According to the above objectives, I have carried out the following:

First, I modeled the four (1 from neutron, 3 from X-ray diffraction) measured data sets, one by one, by applying the RMC method. From particle configurations that could reproduce experimental data, I determined and analyzed the PPCF-s. Later, measured SF-s were considered in pairs (1 neutron, together with 1 X-ray structure factor) and PPCF-s were calculated.

Based solely on experimental results it hasn't been possible to create physically entirely meaningful particle configurations; for this reason, the set of input data was complemented by $g_{ij}(r)$-s from MD simulations that applied various interatomic potential functions. The PPCF-s used were compared to each other and then I checked whether they were fittable by RMC.

Following the initial checks, I have included partial *g(r)*-s from computer simulations, in addition to neutron and X-ray data sets, into RMC modeling (single measured sets, as well pairwise combinaitons of them were considered). As a result of these investigations we have learned how well/badly we were able to fit PPCF-s obtained on the basis of interatomic potentials while agreement with experimental SF-s was kept „perfect" (i.e., not (much) worse than it had been when

only SF-s were modeled). Comparison was made by inspecting the functions when displayed, as well as by comparing numerical values charaterizing average differences between corresponding functions (R_W). In order to handle the problems mentioned above systematically, as well as to make numerical results (R_W) comparable, I had to perform extensive studies concerning the appropriate values of a number of parameters important during RMC calculations, such as the spacing of the $g(r)$ histograms, the precise distance ranges of $g(r)$ –s, or the suitable values of the σ parameters. These parameters have been standardize before drawing any further conclusions. Interaction potentials have then been listed in the order of increasing R_W-s, thus facilitating the selection of water potentials that may be matched best with diffraction results.

I have also studied how PPCF-s from computer simulations, that use pair potentials, can reproduce measured diffraction data (more) directly. First, I created RMC particle configurations that fitted PPCF sets, corresponding to various water potentials, perfetly. For these particle sets total structure factors, identical to what had been obtained from diffraction experiments, have been computed (using the RMC software, for convenience). Via comparing measured and computed total $F(Q)$-s it was possible to select water potentials that are the most consistent with experimental data.

Finally, I have also performed the above consistency check for PPCF-s that had been obtained from experimental data via the traditional method [Soper 2000].

III. SUMMARY, THESES

Good quality neutron and X-ray diffraction results from liquid water have been found simultaneously fittable by the RMC method, within errors, even in the presence of PPCF-s that resulted from computer simulations. That is, in every single cases studied it has been possible to create particle configurations that were consistent with one neutron and one X-ray diffraction measurement. The goodness-of-fit deteriorated only marginally when $g_{ij}(r)$-s from MD were also included in the set of input information.

Thesis No. 1: **The neutron diffraction data [Soper 1997] could easily be fitted together with each of the three X–ray SF-s considered in this work. This was valid for the calculations where PPCF-s from MD simulations were also present: that is, no inconsistency at this stage could be detected [Steinczinger 2013].**

X-ray diffraction results from light/heavy water, neutron diffraction data from heavy water and the constraint that reasonably shaped water molecules must be maintained during the calculations are together not sufficient for determining the structure of water, not even at the partial $g(r)$ level. For this reason I have complemented these primary experimental information with $g_{ij}(r)$ functions from MD simulations: applying RMC modeling, I have prepared particle configurations that would fit both experimental data (as a major requirement) and $g_{ij}(r)$-s from MD (as an additional requirement). Based on this „protocol", the ability of the various simulated $g_{ij}(r)$ sets for improving Reverse Monte Carlo models of water may be assessed. The TIP3P water potential [Jorgensen 1983], somewhat unexpectedly, has proven to be rather successful in this respect.

Thesis No. 2: **Partial pair correlation functions from the BK3 [Kiss 2013] and SWM4-DP [Lamoureux 2005] polarizable, and from the TIP3P non-polarizable water potentials could be approached the closest in RMC calculations in which agreement within errors was demanded simultaneously with structure factors from diffraction experiments [Steinczinger 2013].**

Consistency checks between pair potentials and measurements are possible at the level of structure factors. By means of the RMC method, particle configurations that reproduce all $g_{ij}(r)$–s within errors have been generated. Later, from these configurations the total structure factors corresponding to the measured data could be computed. To the best of our knowledge, this simple approach has not yet been applied systematically for characterizing the reliability of interatomic potentials for liquid water.

Thesis No. 3: **Based on differences between experimental SF-s and *F(Q)*-s calculated from RMC particle configurations that were fully consistent with *$g_{ij}(r)$*-s produced by molecular dynamics simulations it can be stated that the SWM4-DP and the BK3 polarizable water potentials proved to be most successful in reproducing measurable diffraction data. I found the ST2 [Stillinger 1974] potential the least successful in this respect.**

Although direct criticism of diffraction measurements cannot be the subject, either the aim, of this dissertation, still, it is worthwhile resording our experience in this matter: following the multilevel series of comparisons various data sets it appears that of the combinations looked at, partial *g(r)*-s from MD simulations may be made the most consistent with the neutron diffraction data of Soper [Soper 1997] and the X-ray structure factor of Skinner [Skinner 2013].

Thesis No. 4: Having scrutinized a number of combinations of input data sets I found that the set of experimental structure factors that would be most suitable for forming the base of future structural models of liquid water could consist of the neutron diffraction data of Soper [Soper 1997] on heavy water an the X-ray diffraction data of Skinner et al. [Skinner 2013].

PPCF-s obtained from isotopic substitution neutron diffraction measurements of liquid water via the traditional method (separation of total structure factors into partials, then Fourier transformation to provide PPCF-s) [Soper 2000] are, by and large, consistent with the original neutron diffraction result on heavy water [Soper 1997], just as they are with X-ray SF-s of Hart [Hart 2005] and [Fu 2009]. Interestingly, however, the smallest deviations were found for the O-O partial *S(Q)* of Skinner and coworkers [Skinner 2013] – this much later work that is based on a different technique shows a closer relation with the „experimental" PPCF-s [Soper 2000] than with the SF [Soper 1997] from which these PPCF-s had been derived from! This fact illuminates the problems that exist to date in connection with diffraction experiments.

Thesis No. 5: Partial pair correlation functions derived via the traditional route [Soper 2000] were found to be in a tighter connection with the X-ray total structure factor of Skinner [Skinner 2013] than with the neutron SF on heavy water [Soper 1997] that was (in part) the basis of the aforementioned „experimental" PPCF-s had been derived from.

In connection with the above topic it was also shown that RMC modeling the „experimental" PPCF-s together with one of the X-ray [Fu 2009] and with the neutron [Soper 1997] structure factor it is still possible to achieve a perfect fit to these SF-s. Meanwhile, fitting the „experimental" O-O

PPCF becomes much harder, in comparison with the case when consistency with only the neutron data is sought for [Steinczinger 2012].

Thesis No.6: **Of the experimental partial pair correlation functions of Soper [Soper 2000], it is $g_{OO}(r)$ that is fittable the least together with measured structure factors from one X-ray and one neutron diffraction experiment [Steinczinger 2012].**

Research papers published in connection with this dissertation

[Steinczinger 2012] Steinczinger, Z., Pusztai, L.(2012) An independent, general method for checking consistency between diffraction data and partial radial distribution functions derived from them: the example of liquid water. *Condensed Matter Physics*, Vol. 15, No 2, 23606: 1–6.

[Steinczinger 2013] Steinczinger, Z., Pusztai, L. (2013) Comparison of the TIP4P-2005, SWM4-DP and BK3 interaction potentials of liquid water with respect to their consistency with neutron and X-ray diffraction data of pure water. *Condensed Matter Physics* Vol. 16, No. 4, 43604:1-8.

Cited references

[Allen 1987] Allen, M. P. Tildesley, D. J. (1987) *Computer Simulation of Liquids.* Oxford: Clarendon Press

[Baranyai 1995] Baranyai, A., Pusztai, L. (1995) Disorder in condensed phases (Rendezetlenség kondenzált fázisokban). In: Csákvári, B., szerk. *A kémia újabb eredményei 80.* Budapest: Akadémai kiadó

[Baranyai 2011] Baranyai, A. (2011) A víz fizikai kémiája. *Magyar Tudomány,* 12. kötet, pp. 1459-1466

[Chaplin 2003] Chaplin, M. (2003).http://www1.lsbu.ac.uk/water/water_anomalies.html (utoljára megtekintve 2016 . 05.15.)

[Franks 2000] Franks F. *Water: A Matrix of Life* (2000) RSC Publishing, Second Edition

[Fu 2009] Fu L., Bienenstock A., Brennan S. (2009) X-ray study of the structure of liquid water. *J. Chem. Phys.,* 131, 234702.

[Hart 2005] Hart, R. T, Benmore, C. J., Neuefeind, J. C., Kohara, S., Tomberli, B., Egelstaff, P.A. (2005) Temperature Dependence of Isotopic Quantum Effects in Water. *Phys. Rev. Lett*. 94, 047801.

[Jorgensen 1983] Jorgensen, W. L. Chandrasekhar, J., Madura, J. D., Impey, R. W., Klein, M. L. (1983) Comparison of simple potential functions for simulating liquid water. *The Journal of Chemical Physics,* 79, pp. 926-935.

[Kiss 2013] Kiss, P. T., Baranyai, A., (2013) A systematic development of a polarizable potential of water. *Journal of Chemical Physics,* 138, 204507.

[Lamoureux 2005] Lamoureux, G., Harder, E., Vorobyov, I. V., Roux, B., MacKerell, A. D. Jr.(2005) A polarizable model of water for molecular dynamics of biomolecules. *Chemical Physics Letters,* 418, pp. 241-245.

[McGreevy 1988] McGreevy, R. L., Pusztai, L. (1988) Reverse Monte Carlo Simulation: A new technique for the determination of disordered structures. *Molecular Simulation.* 1, 359.

[Skinner 2013] Skinner, L. B., Huang, C., Schlesinger, D., Pettersson, L. G., Nilsson, A.(2013) Benchmark oxygen-oxygen pair-distribution function of ambient water from x-ray diffraction measurements with a wide Q-range. *J. Chem. Phys.* 138, 074506 .

[Soper 1997] Soper, A. K., Bruni, F., Ricci, M. A. (1997) Site–site pair correlation functions of water from 25 to 400 °C: Revised analysis of new and old diffraction data. *J. Chem. Phys*. 106, 247. http://dx.doi.org/10.1063/1.473030.

[Soper 2000] Soper, A. K. (2000) The radial distribution functions of water and ice from 220 to 673 K and at pressures up to 400 Mpa. *Chem. Phys*. 258, pp. 121-137. doi:10.10116/S0301-0104(00)00179-8.

[Soper 2013] Soper, A. K. (2013) The Radial Distribution Functions of Water as Derived from Radiation Total Scattering Experiments: Is There Anything We Can Say for Sure? *ISRN Physical Chemistry*; Volume 2013 , Article ID 279463. DOI:10.1155/2013/279463

[Stillinger 1974] Stillinger, F. H., Rahman, A. (1974) Improved simulation of liquid water by molecular dynamics. *J. Chem. Phys*. 60, pp. 1545-1557.