Published in final edited form as:

J Phys Chem B. 2007 November 15; 111(45): 12941–12944. doi:10.1021/jp076501p.

Direct comparisons of experimental and calculated neutron structure factors of pure solvents as a method for force field validation

Jennie L. Thomas^{*}, Douglas J. Tobias^{*}, and Alexander D. MacKerell Jr.[‡]

^{*}*Department of Chemistry, University of California, Irvine, 1102 Natural Sciences 2, University of California, Irvine, California, 92697-2025*

[‡]*Department of Pharmaceutical Sciences, School of Pharmacy, University of Maryland, Baltimore, 20 Penn Street, Baltimore, MD 21201*

Abstract

In the present letter we directly compare neutron structure factors calculated from force field (FF)-based molecular dynamics simulations with experimental structure factors for water, methanol, and tetrahydrofuran (THF). For water, the difference in measured structure factors is more significant than differences between the FFs. It is shown that the inclusion of electronic polarization in the force field improves the agreement with experiment for the more polar methanol while the results are comparable for the additive and polarizable FF models of the less polar THF. The data presented here confirm that comparing calculated scattering profiles from FF based MD simulations to measured neutron structure factors is a promising method for FF validation and development.

Scattering experiments based on isotopic substitution neutron diffraction techniques have been used to measure the structural correlation functions in liquids for many years.^{1,2} Experiments do not measure pair interactions directly, but rather measure the sum of all contributing pair interactions in reciprocal space. The measured structure factors and the resulting difference functions provide a wealth of information about the molecular structure of liquids that may be used to validate and guide the refinement of empirical force fields (FF). The usual approach is to compare radial distribution functions (RDFs) from condensed phase FF simulations to RDFs extracted from scattering data. However, while in principal simple, the inversion of the scattering data to real space suffers from well-documented^{3,4} systematic errors that limit the utility of RDFs in FF validation. In some cases simulations are compared to diffraction data directly. For example, in the case of empirical potential structure refinement (EPSR),^{4,5} the data are used to constrain a simulation of the liquid, and the structure of the model solution is driven towards configurations that are consistent with the measured partial structure factors. In other cases, structure factors calculated from FF-based molecular dynamics (MD) simulations are calculated and compared with measured data.^{6,7} In this letter we directly compare neutron structure factors calculated from additive and polarizable FF-based MD simulations with experimental structure factors for water, methanol, and THF in order to evaluate FF quality. The comparison demonstrates the potential utility of incorporating direct comparisons with neutron diffraction data into well-established all-atom FF development strategies.⁸

Neutron diffraction experiments measure the total structure factor as a function of the magnitude of the scattering vector, k ,

$$k = \frac{4\pi}{\lambda} \sin\theta \quad (1)$$

where λ is the wavelength of the incoming neutrons and 2θ is the scattering angle. The neutron structure factor is defined as

$$S(k) = \frac{1}{N} \sum_{\alpha} \sum_{\beta} b_{\beta} b_{\alpha} \langle \exp[ik \cdot (r_{\alpha} - r_{\beta})] \rangle \quad (2)$$

where b is the coherent neutron scattering length, $\mathbf{r}$ represents the coordinates of each atom in real space, α and β are atom indices, and the sum is over all atoms N .⁹ The neutron structure factor is also the weighted sum of each contributing partial structure factor

$$S(k) = \sum_{\alpha \leq \beta}^N (2 - \delta_{\alpha\beta}) c_{\alpha} c_{\beta} b_{\alpha} b_{\beta} (S_{\alpha\beta} - 1) \quad (3)$$

where c is the respective concentration of each species and b is the coherent neutron scattering length. To avoid double counting the Kronecker delta function is used. The partial structure factors, $S_{\alpha\beta}$, are given by the Fourier transform of each contributing pair correlation function

$$S_{\alpha\beta}(k) = 4\pi\rho \int_0^{\infty} r^2 g_{\alpha\beta}(r) - 1 \frac{\sin(kr)}{kr} dr - 1 \quad (4)$$

where ρ is the density, $g_{\alpha\beta}(r)$ is the radial pair distribution function as a function of r , and k is the magnitude of the scattering vector.² We have calculated the total structure factors using Equation 2 and the partials using Equation 4, by FT of each pair correlation function. The scattering lengths for all atoms are summarized in Table 1. Because the magnitudes of the k vectors supported by the simulation cells occur at unequal intervals the calculated structure factors were grouped into 0.1 Å⁻¹ bins and averaged within the bin before plotting; the same averaging scheme was used for the experimental data for ease of analysis. We compare our calculations to measured structure factors for liquid water,¹⁰ methanol,^{11,12} and THF.¹³ In the case of water two relatively independent experimental data sets were used to evaluate FF quality, referred to as experimental A (new) and B (old). The reader is referred to Ref. 10 for a more complete description of the differences between the two data sets.

MD simulations were performed using four water models: POL3,¹⁴ SPC/E,¹⁵ SWM4-NDP,¹⁶ and TIP3P⁶ in the NPT ensemble from which structure factors were calculated for H₂O and D₂O and compared with measured data. The calculated and measured data points were placed on a 0.1 Å⁻¹ grid and the root mean squared deviations (RMSDs) were calculated using two different experimental data sets (Table 2). The RMSDs were calculated with and without including points in the range of 0.5-1.7 Å⁻¹ due to differences in the experimental structure factors at low k . It is evident that the level of agreement with experiment depends on the water FF, although the RMSD values are more sensitive to the experimental data set used than to differences in the water models. The variation in the experimental results (RMSDs for the two H₂O and D₂O data sets are 0.024 and 0.043, respectively, using data in the range $k=0.5-1.7$ Å⁻¹) indicates that further refinement of these data is needed in order to be useful for future water FF development. In addition, comparison of the experimental and calculated data may be performed qualitatively. For example, the shape of the first peak ($k \sim 2$ Å⁻¹) for the calculated TIP3P D₂O structure (see Supporting Information Figure S2) does not agree as well with the measured data. However, this is not reflected in the overall RMSD, pointing out that both qualitative and quantitative comparisons are useful in determining FF quality.

Calculated structure factors for two models of liquid methanol based on the CHARMM22 additive potential⁸ and a polarizable FF based on a classical Drude oscillator¹⁷ are shown in Figure 1, with the associated RMSD data in Table 3. The following isotopic mixtures were considered: pure CD₃OH, a 1:1 mixture of CD₃OH:CD₃OD, and pure CD₃OD. The polarizable force field (Figure 1b) is in better agreement with the measured data, especially at low k , with the largest difference between the measured and calculated structure factors observed for CD₃OD. The shape of the first peak (near $k=1.8 \text{ \AA}^{-1}$) is also in better agreement with the polarizable model (Figure 1b). The improved agreement is reflected in the RMSD values for the two models (Table 3), which confirms the ability of the present approach to discriminate between FFs.

Analysis of partial structure factors and pair radial distribution functions from the methanol simulations (selected partials and RDFs are shown in Figures 2 and 3) that contribute to the methanol structure factors suggest that the polar atom to nonpolar atom interactions improve upon inclusion of electronic polarizability. The partials shown are for the methanol-D (CD₃OD) isotope. Significant differences between the two models are seen in the D-O (methyl hydrogens to hydroxyl oxygen in 2a), OD-C (hydroxyl proton to methyl carbon in 2b) and C-O (methyl carbon to hydroxyl oxygen in 2c) partials. Such improvements are consistent with the inclusion of polarization, which is anticipated to more accurately treat interactions associated with the induction of local dipoles on the non-polar carbon atoms when adjacent to the polar oxygen. This conclusion is consistent with recent simulations and analysis on the microstructure of liquid methanol¹⁸ that confirms methanol is a strong hydrogen bonding liquid. The experimental partial structure factors for methanol are largely unavailable due to the number (and difficulty) of isotopic substitution experiments required to experimentally measure each partial. However, if such data were available, direct comparison of the calculated and experimental scattering profiles could be used to identify atoms for which optimization of their nonbond parameters (e.g. Lennard-Jones terms or electrostatic parameters) may lead to improvements in the model. In the present methanol calculations it is evident that such improvements have been made via the explicit inclusion of electronic polarization in the FF.

Including the molecular polarizability is less important for the overall calculated structure factor of liquid THF, as shown by the similar RMSD for simulations using additive and polarizable FFs¹⁹ (Table 3). As in the case of methanol, three isotopic mixtures were considered: C₄H₈O, a 1:1 mixture of C₄H₈O:C₄D₈O, and C₄D₈O. Both FF models are in good agreement with the measured structure factors for all isotopic mixtures. The lack of improvement upon going from the additive to polarizable model is proposed to be associated with the nonpolar nature of THF ($\epsilon = 7.5$),²⁰ such that the improved electrostatic representation in the FF has less of an impact compared to methanol ($\epsilon = 33.0$).²⁰

We have directly compared neutron structure factors calculated from FF-based MD simulations with measured structure factors for water, methanol, and THF. In the first case, water, comparison of the calculated and experimental structure factors show that the four models are not identical; however, differences in the experimental structure factors are more significant than differences between force fields. Second, for methanol it is shown that the inclusion of electronic polarizability in the FF improves the model, as evidenced by the significant improvement in the agreement with the experimental structure factors. Third, in the case of THF, both the additive and polarizable models are in good agreement with the measured data for three isotopic mixtures analyzed.

Inclusion of structure factors from scattering experiments in parameterization represents a significant step for the field of FF development. To date, FF optimization has relied primarily on the reproduction of macroscopic physical properties such as the density, specific heat capacity or diffusion coefficient. While important, these measured quantities do not contain

the detailed structural data available from neutron diffraction experiments. The substantial contribution of intramolecular correlations can mask the intermolecular correlations that are our main interest, hence somewhat diminishing the utility of the approach (see Figure S5 of the supporting information and associated discussion). Therefore, comparing data from experiments (and simulations) employing a variety of isotopes, whereby different interactions are highlighted, or in the ideal case comparing difference functions, is important for evaluating force fields. A move towards evaluating FF quality against measured structure factors will result in greater confidence that conclusions derived from the models are correct on both the macroscopic and microscopic level. The results reported here confirm the utility of direct comparison of calculated and experimental neutron structure factors for FF validation, and show that there is a bright future for using such methods in FF optimization.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGMENT

The authors thank Drs. Alan Soper and Daniel Bowron for helpful discussions and providing electronic copies of the measured structure factor data and Drs. Igor Vorobyov and Victor Anisimov for providing simulation data for SWM4 water, THF and methanol, the NSF (Grants #CHE-0431512 and #CHE-0417158) and the NIH (GM50501) for funding.

REFERENCES

1. Enderby JE, North DM, Egelstaf PA. *Phil. Mag* 1966;14:961.
2. Neilson GW, Enderby JE. *J. Phys. Chem* 1996;100:1317.
3. Soper AK. *Chem. Phys* 1986;107:61.
4. Soper AK. *Phys. Rev. B* 2005;72:12.
5. Soper AK. *Mol. Phys* 2001;99:1503.
6. Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML. *J. Chem. Phys* 1983;79:926.
7. Mason PE, Neilson GW, Enderby JE, Saboungi ML, Brady JW. *J. Am. Chem. Soc* 2005;127:10991. [PubMed: 16076206]
8. MacKerell AD, Bashford D, Bellott M, Dunbrack RL, Evanseck JD, Field MJ, Fischer S, Gao J, Guo H, Ha S, Joseph-McCarthy D, Kuchnir L, Kuczera K, Lau FTK, Mattos C, Michnick S, Ngo T, Nguyen DT, Prodhom B, Reiher WE, Roux B, Schlenkrich M, Smith JC, Stote R, Straub J, Watanabe M, Wiorkiewicz-Kuczera J, Yin D, Karplus M. *J. Phys. Chem. B* 1998;102:3586.
9. Squires, GL. *Introduction to the theory of thermal neutron scattering*. Dover Publications; Mineola, NY: 1978.
10. Soper AK. *J. Phys.: Condens. Matter* 2007;19:335206.
11. Yamaguchi T, Hidaka K, Soper AK. *Mol. Phys* 1999;97:603.
12. Yamaguchi T, Hidaka K, Soper AK. *Mol. Phys* 1999;96:1159.
13. Bowron DT, Finney JL, Soper AK. *J. Am. Chem. Soc* 2006;128:5119. [PubMed: 16608347]
14. Caldwell JW, Kollman PA. *J. Phys. Chem* 1995;99:6208.
15. Berendsen HJC, Grigera JR, Straatsma TP. *J. Phys. Chem* 1987;91:6269.
16. Lamoureux G, Harder E, Vorobyov IV, Roux B, MacKerell AD. *Chem. Phys. Lett* 2006;418:245.
17. Anisimov VM, Vorobyov IV, Roux B, MacKerell AD. *J. Chem. Theory Comp.* 2007 In Press.
18. Zoranić L, Sokolić F, Perera A. *J. Chem. Phys* 2007;127:024502. [PubMed: 17640132]
19. Vorobyov I, Anisimov VM, Greene S, Venable RM, Moser A, Pastor RW, MacKerell AD. *J. Chem. Theory Comput* 2007;3:1120.
20. CRC Handbook of Chemistry and Physics. 84th. CRC Press; New York: 2003.

Figure 1. Neutron structure factors for liquid methanol using the CHARMM22 additive force field (a) and using a classical Drude oscillator (b) to treat the polarizability (CD₃OH:CD₃OD and CD₃OH are offset by +1, and +2 respectively).

Figure 2.

Partial structure factors contributing to the difference between the additive and Drude models for deuterated liquid methanol (CD_3OD). The atom labels are as follows: D denotes methyl deuterium, OD denotes hydroxyl deuterium, O denotes oxygen, and C denotes carbon.

Figure 3.

Selected RDFs contributing to the difference between the calculated structure for the additive and Drude models for liquid methanol. The atom labels are as follows: H denotes methyl hydrogen, OH denotes hydroxyl hydrogen, O denotes oxygen, and C denotes carbon.

Table 1

The coherent neutron scattering lengths (b in Equation 2 and Equation 3) used for carbon, oxygen, deuterium, and hydrogen for the neutron structure factor calculations.

Atom	Coherent neutron scattering length (fm)
Carbon	6.646
Oxygen	5.803
Deuterium (^2H)	6.671
Hydrogen (^1H)	-3.739

Table 2
Calculated RMSDs for H₂O and D₂O using data from two different sets of neutron diffraction experiments.

Molecule	Model	Calculated using k points from 0.5-10 Å ⁻¹		Calculated using k points from 1.7-10 Å ⁻¹	
		Exp. A	Exp. B	Exp. A	Exp. B
H ₂ O	POL3	0.024	0.036	0.015	0.020
	SPC/E	0.025	0.024	0.016	0.020
	SWM4	0.026	0.025	0.017	0.022
	TIP3P	0.025	0.025	0.015	0.021
D ₂ O	POL3	0.024	0.033	0.017	0.019
	SPC/E	0.024	0.033	0.017	0.019
	SWM4	0.040	0.029	0.031	0.026
	TIP3P	0.036	0.042	0.036	0.033

Table 3

Calculated RMSDs for methanol and THF. Calculated data based on simulations using the additive CHARMM force field and using an electronic polarizable force field based on a classical Drude oscillator.

Molecule	Model	Isotopic mixture	Calculated using k points from 0.5 - 10 Å ⁻¹	Calculated using k points from 1.7 - 10 Å ⁻¹
Methanol	additive	CD ₃ OH	0.058	0.041
		1:1 mixture	0.069	0.040
		CD ₃ OD	0.080	0.038
	Drude	CD ₃ OH	0.031	0.027
		1:1 mixture	0.033	0.027
		CD ₃ OD	0.032	0.029
THF	additive	C ₄ H ₈ O	0.067	0.060
		1:1 mixture	0.038	0.024
		C ₄ D ₈ O	0.026	0.013
	Drude	C ₄ H ₈ O	0.067	0.060
		1:1 mixture	0.041	0.023
		C ₄ D ₈ O	0.027	0.014