



# Structural Behavior of Aqueous t-Butanol Solutions from Large-Scale Molecular Dynamics Simulations

S. D Overduin, Aurélien Perera, G. N. Patey

## ► To cite this version:

S. D Overduin, Aurélien Perera, G. N. Patey. Structural Behavior of Aqueous t-Butanol Solutions from Large-Scale Molecular Dynamics Simulations. The Journal of Chemical Physics, 2019, 150 (18), pp.184504. 10.1063/1.5097011 . hal-02179317

**HAL Id: hal-02179317**

**<https://hal.sorbonne-universite.fr/hal-02179317>**

Submitted on 10 Jul 2019

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Structural Behavior of Aqueous *t*-Butanol Solutions from Large-Scale Molecular Dynamics Simulations

S. D. Overduin,<sup>1</sup> Aurélien Perera,<sup>2,\*</sup> and G. N. Patey<sup>1,†</sup>

<sup>1</sup>*Department of Chemistry, University of British Columbia,  
Vancouver, British Columbia, Canada V6T 1Z1*

<sup>2</sup>*Laboratoire de Physique Théorique de la Matière Condensée (UMR CNRS 7600),  
Université Pierre et Marie Curie, 4 Place Jussieu, F75252, Paris cedex 05, France*

(Dated: March 13, 2019)

## Abstract

Large-scale molecular dynamics (MD) simulations are reported for aqueous *t*-butanol (TBA) solutions. The CHARMM generalized force field (CGenFF) for TBA is combined with the TIP4P/2005 model for water. Unlike many other common TBA models, the CGenFF model is miscible with water in all proportions at 300 K. The main purpose of this work is to investigate the existence and nature of microheterogeneous structure in aqueous TBA solutions. Our simulations of large systems (128000 and 256000 particles) at TBA mole fractions of 0.06 and 0.1 clearly reveal the existence of long-range correlations ( $> 10$  nm) that show significant variations on long time scales ( $\sim 50$  ns). We associate these long-range, slowly varying correlations with the existence of supramolecular, domain-like structures that consist of TBA-rich and water-rich regions. This structure is always present but continually changing in time, giving rise to long-range, slowly varying pair correlation functions. We find that this behavior appears to have little influence on the single particle dynamics; the diffusion coefficients of both TBA and water molecules lie in the usual liquid state regime, and mean square displacements provide no indication of anomalous diffusion. Using our large system simulations, we are able to reliably calculate small angle x-ray scattering (SAXS) and small angle neutron scattering (SANS) spectra, except at very low wave vector, and the results agree well with recent experiments. However, the present paper shows that simulation of the relatively simple TBA/water system remains challenging. This is particularly true if one wishes to obtain properties such as Kirkwood-Buff factors, or scattering functions at low wave vector, which strongly depend on the long-range behavior of the pair correlations.

## I. INTRODUCTION

The structure of aqueous *t*-butanol solutions (TBA) has been debated for many years.<sup>1-27</sup> Thermodynamic anomalies at low TBA concentration suggest that some aggregation of TBA molecules occurs,<sup>1</sup> and generally model simulations have supported the idea of some sort of TBA aggregation.<sup>13,17,19-21,23,24,28</sup> However, many TBA models commonly used in simulations of aqueous solutions have been found to undergo phase separation in larger systems, casting doubt on the applicability of earlier results to understanding the structure of aqueous TBA.<sup>19,28</sup> For example, the Transferable Potential for Phase Equilibria-United Atom (TraPPE-UA) model,<sup>29</sup> which has proven to be an excellent model of pure TBA,<sup>17</sup> phase separates at 298 K, over the concentration range  $0.1 \leq x_t \leq 0.5$  ( $x_t = N_{TBA}/N$  is the mole fraction of TBA), and possibly at lower and higher concentrations, which we did not consider. The behavior of the TraPPE-UA model is briefly discussed in the Supplementary Material. We note that in a relatively large system (32000 particles), complete phase separation requires up to 50-70 ns, and that the transient finite-size “clusters” appearing along the pathway to phase separation could easily be mistaken for stable aggregates if simulations are not sufficiently long. A further complication is that phase separations observed in simulations of large systems may not occur in the smaller systems often employed, presumably because phase transition is inhibited by the cost of interfacial tension. Thus, employing molecular simulations in an effort to understand TBA-water solutions has proven to be less straightforward than one might have anticipated.

In a recent paper<sup>28</sup> it was shown that the CHARMM generalized force field (CGenFF) for TBA combined with the TIP4P/2005 water model does not demix at any concentration for temperatures below 320 K, in qualitative agreement with experiment. Additionally, at low TBA concentrations ( $x_t \lesssim 0.1$ ) the model agrees well with experimentally measured thermodynamic properties of TBA-water solutions, as well as with diffusion coefficients. In these earlier simulations, which employed up to 32000 particles, long-range correlations that appeared slow to converge were observed, but not explored in detail. In the present paper, we report simulations of very large systems ( $N = 128000$  and  $256000$  particles) of aqueous TBA solutions at two concentrations ( $x_t = 0.06$  and  $0.1$ ). The purpose of this work is to investigate the range and nature of the structural correlations, and how they reflect the microheterogeneous nature of these solutions. A particularly important point

is that all TBA models which we have considered<sup>19,28</sup> show TBA aggregation and species microsegregation, but the CGenFF appears to be the only model where TBA aggregation does not evolve into full macroscopic demixing. Thus, it appears that maintaining a stable phase which exhibits significant microsegregation but avoids demixing is a central problem in modeling solutions of this type.

Our results show that simulations of TBA-water solutions remain challenging. The presence of TBA-rich and water-rich regions leads to small but very long range pair correlations that exceed half the length of the simulation even for 256000 particles ( $\gtrsim 10$  nm). In addition, the pair correlations undergo very slow fluctuations that reflect the slow rearrangement of the microheterogeneous structure. This means that long simulations (hundreds of nanoseconds) of large systems are necessary, and even then the true long-range behavior of the correlation functions remains somewhat uncertain. The implications of this for simulation studies, and, in particular, for the determination of structure factors and Kirkwood-Buff integrals is discussed. We show that over the measured region, our simulation results are in good agreement with SAXS and SANS experiments.<sup>15,23</sup>

The remainder of this paper is organized as follows. The models and simulation details are discussed in Section II, simulation results are described and compared with experiment in Section III, and our conclusions are summarized in Section IV.

## II. MODEL AND SIMULATION METHOD

Molecular dynamics simulations with 8000, 128000, and 256000 molecules were performed using the GROMACS 4.6.2 simulation package.<sup>31</sup> The equations of motion were integrated using a 2 fs time step, and electrostatic interactions were calculated using the particle mesh Ewald method.<sup>32,33</sup> Short-range interactions between all atom pairs were truncated at 0.95 nm. The equilibrium density was obtained by performing simulations at fixed pressure ( $P$ ) and temperature ( $T$ ), using the Berendsen barostat<sup>34</sup> to control the pressure at 1 bar for at least 5 ns. Following equilibration with the Berendsen barostat, constant pressure simulations were performed using the Parrinello-Rahman barostat.<sup>35</sup> In all simulations the temperature was controlled using the Nosé-Hoover thermostat.<sup>36,37</sup> For some systems,  $NVT$  simulations were also performed using the density obtained from  $NPT$  simulations.

Simulations with 8000 particles were used only to determine the excess enthalpy. Excess

enthalpies computed from larger simulations are in good agreement with those obtained with 8000 particles. A summary of the 128000 and 256000 particle simulations performed at 300 K is provided in Table I. *NPT* simulations were all performed at 1 bar, and the *NVT* simulation at  $x_t = 0.06$  at the average *NPT* density ( $\sim 0.97$  gm/cc).

For most simulations presented here, TBA is modelled using the CHARMM generalized forcefield (CGenFF)<sup>38,39</sup> and the TIP4P/2005 model<sup>40</sup> is used for water. We also show results in the Supplementary Material for simulations performed with the Transferable Potential for Phase Equilibria-United Atom (TraPPE-UA)<sup>29</sup> coarse-grained TBA model, employing 32000 particles.

### III. RESULTS

In earlier simulations,<sup>28</sup> it was shown that the CGenFF TBA model combined with TIP4P/2005 water remains mixed at all concentrations for  $T \leq 320$  K, although phase separation possibly occurs at higher temperatures with  $x_t = 0.1$ . The CGenFF model also agrees reasonably well with many experimental properties of aqueous TBA solutions, including the density, isothermal compressibility, and heat capacity, especially at low TBA concentrations.<sup>28</sup> In Fig. 1, we further evaluate the CGenFF model, comparing with experimental results for the excess enthalpy. Overall, the model agrees well with experiment, although there are some differences. For example, experimentally, the excess enthalpy decreases with increasing TBA concentration up to a concentration of  $x_t \approx 0.06$ , while the decrease is more gradual and extends out to  $x_t \approx 0.2$  in the model simulations. Nevertheless, the qualitative features of the experimental curve are reproduced by the model, and the magnitudes of excess enthalpy at most concentrations are in good agreement. Taken together with earlier results,<sup>28</sup> the agreement with experimental excess enthalpies indicates that the CGenFF model provides a reasonably accurate representation of real TBA/water systems.

In previous simulations (using 32000 particles) employing CGenFF TBA with TIP4P/2005 water, long-range correlations that appeared slow to converge were observed.<sup>28</sup> Here we perform long simulations of very large systems (128000 and 256000 particles) to further investigate the structure of aqueous TBA solutions at  $x_t = 0.06$  and  $x_t = 0.1$ . These simulations aim to clarify the nature of structural correlations, including possible long-range

correlations associated with TBA-rich and water-rich regions.

Configurational snapshots from 256000 particle *NPT* simulations of solutions at  $x_t = 0.06$ , and 0.1 are shown in Fig. 2. Snapshots from *NVT* simulations (not shown) exhibit identical features. We note that, while TBA-rich and water-rich regions are evident in the snapshots, one does not see TBA clusters of well-defined size and shape. From simulations with the Lee and van der Vegt model<sup>13</sup> (which Gupta and Patey<sup>19</sup> found demixes at  $x_t = 0.1$  in 64000-particle simulations), Bagchi and coworkers<sup>24</sup> observed that TBA molecules form a spanning cluster in the vicinity of  $x_t = 0.06$ . Noting that most thermodynamic anomalies occur in the region  $x_t = 0.03 - 0.07$ , they suggest a possible relationship between anomalies and the percolation transition. Using the current model (CGenFF TBA with TIP4P/2005 water) we find that the TBA molecules also form a spanning cluster at  $x_t \approx 0.06$ .

#### A. Pair correlation functions

In order to illustrate correlations between the TBA-rich and water-rich regions evident in Fig. 2, in Fig. 3 we plot selected atom-atom pair correlation functions (three each for TBA-TBA, water-water, and TBA-water correlations) for 256000-particle systems ( $x = 0.06$ ) at *NVT* (top) and *NPT* (bottom). The atom labels used here and elsewhere in the paper are CC for the central carbon atom of TBA, OA for the oxygen of TBA, OW for the oxygen of water, and HW for the hydrogen of water. We have estimated finite-size corrections (given in the figure caption), which are further discussed below. Uncorrected results are provided in the supplemental material (Fig. S2). The main panels and insets of Fig. 3 show that the atomistic details present in the correlation functions at short range (see main panel) are completely washed out at long range (see inset). In the *NVT* system, there is no significant structure beyond  $\sim 3$  nm. In contrast, in the *NPT* case, there are weak, decaying correlations that extend to the edge of the simulation cell. These very long-range correlations may be associated with domain-like structure, similar to that previously reported by one of us for aqueous 1-propanol solutions.<sup>41</sup>

Pair correlation functions between the central carbons of TBA [ $h_{CC-CC}(r)$ ] at  $x_t = 0.06$  obtained with 128000 particles are shown for *NVT* and *NPT* simulations in Figs. 4 and 5, respectively. We note that the pair correlation functions converge to  $0 \pm 0.004$  for  $r \gtrsim 2$  nm, however, despite the large system sizes and long simulation times we employ, results obtained

over 50 ns simulation blocks show surprising variation in the behavior of the long-range tails, as well as some variation in the heights of the first peak. The long-range tails apparent in Figs. 4 and 5, though relatively weak, vary significantly in shape. In some simulation blocks the decay appears monotonic, reminiscent of what one might associate with concentration fluctuations, whereas other blocks show structure suggesting correlations between TBA-rich regions. The long-time averages tend to show less structure than some individual 50 ns blocks, largely due to the cancellation of structural features amongst the blocks. As in the 256000 particle case (Fig. 3), under *NPT* conditions the long-range correlations extend to the edge of the simulation cell.

The water-water  $[h_{OW-OW}(r)]$  and TBA-water  $[h_{CC-OW}(r)]$  pair correlation functions show similar variations (Figs. S3 and S4). This is also true of the 256000 particle simulations at  $x_t = 0.1$  (Figs. S5, S6, and S7). The variations we observe in the pair correlation functions indicate the presence of slow structural fluctuations involving TBA-rich and water-rich regions, and, even with our long simulations of large systems, it is difficult to achieve adequate sampling.

Although the level of convergence we are aiming for is high (for example, variations in the long-range tails occur in the third or fourth decimal place), the fact that we are unable to obtain truly converged pair correlation functions in such large systems and long simulations is unexpected given the apparent simplicity of the system (a mixture of small molecules). It is also surprising considering the single-particle dynamics, which show that diffusion of both TBA and water molecules proceeds normally (mean square displacements show no subdiffusive timescale, and thus no indication of multiple diffusive environments) with diffusion constants at  $x_t = 0.06$  of  $0.47 \times 10^{-5} \text{ cm}^2/\text{s}$  and  $1.30 \times 10^{-5} \text{ cm}^2/\text{s}$  for TBA and water, respectively. In other words, over a 50 ns time period a TBA molecule travels on average  $\sim 10 \text{ nm}$  (roughly half the length of the simulation cell in the 256000 particle systems), which is typically considered sufficient for sampling purposes. This indicates that the slow variation of the pair correlation functions is due to the “supramolecular” structure consisting of TBA-rich and water-rich regions, as noted above.

Pair correlation functions from *NVT* and *NPT* simulations are compared in Fig. 6. Despite the large fluctuations present in both *NVT* and *NPT* simulations discussed above, the averages do suggest some ensemble dependence which would not be surprising, given the long correlation range present in these systems. As noted above, the correlations extend

to larger  $r$  in the  $NPT$  case. It is possible that the restricted volume fluctuations in the finite-size  $NVT$  system influences the density and concentration fluctuations, resulting in the formation of slightly different structures. However, we note that the difference between the 128000 and 256000  $NPT$  results (between 3 and 6 nm) is comparable to the differences observed between  $NVT$  and  $NPT$  results at the same  $N$ .

Long-range correlations and inaccuracies in determining pair correlation functions have implications for computing other quantities, including experimental observables such as Kirkwood-Buff (KB) integrals.<sup>42</sup> Because KB integrals are sensitive to short- and long-range structure, they have been used as a test of a model's validity.<sup>13,30</sup> In fact, parameters for some TBA models have been chosen by attempting to reproduce KB integrals.<sup>13</sup> Therefore, it is of interest to investigate how the discrepancies we observe in the pair correlation functions from different simulations might influence the KB integrals.

In order to accurately determine KB integrals, finite-size effects in the pair correlation functions should be properly corrected. In finite-size systems,  $NVT$  pair correlation functions  $h_{ij}(r)$  are expected to decay to  $\sim -S_{ij}(0)/(\sqrt{x_i x_j} N)$  rather than zero at long-range.<sup>43–45</sup> Given the variation we observe in the pair correlation functions at long-range, it is impossible to unambiguously determine the value of the asymptote, although we expect it to be small for such large systems. Instead, we approximate the correction by shifting the  $h_{ij}(r)$  so that they converge to zero at large separations. This has been done for the atom-atom pair correlation functions plotted in Fig. 3, and the small corrections applied are given in the figure caption. This correction procedure is somewhat arbitrary due to the irregular shapes of the long-range  $h_{ij}(r)$  tails, but our main purpose is to determine how sensitive the KB integrals are to small corrections in such large systems. KB integrals at  $x_t = 0.06$  obtained from corrected and uncorrected pair correlation functions as functions of the upper limit of integration  $R$  [ $G_{ij}(R) = 4\pi \int_0^R h_{ij}(r) r^2 dr$ ] are shown in Fig. 7. Note that the value of a KB integral is defined as the  $R \rightarrow \infty$  limit. Rather than converging to constants at large  $R$  ( $R \gtrsim 4$  nm), the uncorrected  $G_{CC-CC}(R)$  and  $G_{OW-OW}(R)$  slope downward, and  $G_{CC-OW}(R)$  slopes upward, due the  $h_{ij}(r)$  converging to non-zero values. Note that this behavior is consistent with the expected (at least in the  $NVT$  ensemble) negative corrections for TBA-TBA and water-water, and the expected positive correction for TBA-water. In most cases, estimating finite-size corrections for the  $h_{ij}(r)$ , as described above, appears to largely resolve the issue resulting in roughly horizontal  $G_{ij}(R)$ . However,



for 256000 particles with *NPT*, there are large differences between the corrected and uncorrected results, and it is unclear if the asymptotic limit has been achieved. This reflects the very long-range correlations for this system, as noted above (Fig. 3).

Fig. 7 illustrates that the KB integrals do not always reach their asymptotic values for  $R < 2 - 4$  nm, and, moreover, there are significant differences between the *NVT* and *NPT* estimates, as well as between *NPT* estimates from different system sizes. Concerning the first point, until recently most simulations were performed with  $N \leq 2000$ , which corresponds to a maximum  $R$  of  $\sim 2$  nm. Additionally, even if large systems are employed, slow fluctuations (discussed above) mean that long simulation times ( $> 100$  ns for 128000-particle systems) are necessary to even approach adequate sampling. With respect to the second point, the ensemble differences observed in the  $h_{ij}(r)$  are apparent in the  $G_{ij}(R)$ , which begin to deviate at  $\sim 2.5$  nm. At large  $R$ ,  $G_{ij}(R)$  from *NVT* and *NPT* simulations with 256000 particles differ by more than a factor of 2. However, we must note that, particularly in the *NPT* case, the correlations extend to the end of the simulation cell, and it is unlikely that true convergence has been achieved. For example, if the correlation functions were to oscillate about zero at still larger distances, this could lead to different estimates of the tail correction, and to different estimates of the KB integrals. Thus, the presence of supramolecular structure, and associated long-range correlations, makes the estimation of KB integrals very uncertain.

## B. Structure factors

Partial structure factors for different atomic pairs are obtained using

$$S_{ij}(k) = \delta_{ij} + 4\pi\rho\sqrt{x_i x_j} \int_0^\infty dr r \frac{\sin(kr)}{k} h_{ij}(r), \quad (1)$$

where  $i$  and  $j$  are atom labels. At low wavenumber, partial structure factors, like the KB integrals, are sensitive to inaccuracies and corrections to the  $h_{ij}(r)$ . The low- $k$  region of the partial structure factors obtained at  $x_t = 0.06$  from *NVT* and *NPT* simulations with 128000 and 256000 particles are shown in Figs. 8 and 9, respectively, both with and without applying an estimate of the finite-size corrections to the pair correlation functions.

If finite-size corrections to the  $h_{ij}(r)$  are omitted, the atom-atom structure factors shown in Figs. 8 and 9 have peaks (minima in the cross terms) in  $S_{ij}(k)$  at  $k \sim 0.7$  nm<sup>-1</sup>,

which could suggest the presence of correlations between domain-like structures. Previously, Perera suggested that aggregation in aqueous alcohols could result in the presence of low- $k$  peaks (prepeaks) in the structure factor.<sup>41</sup> However, for the present TBA/water system, our analysis indicates that the existence or not of a prepeak in the atom-atom structure factors depends very much on details of the long-range behavior of the corresponding pair correlation functions. For the examples shown in Figs. 8 and 9, the prepeaks occurring for the uncorrected results, are removed or much diminished if finite-size corrections estimated as discussed above are applied. The structure factors obtained for  $x_t = 0.1$  are shown in Fig. S9. Given the large fluctuations in the  $h_{ij}(r)$  discussed above, and the arbitrariness involved in estimating finite-size corrections, we cannot with confidence determine the precise behavior of the structure factors at very low  $k$ . We conclude that there are no prepeaks above  $k \sim 1 \text{ nm}^{-1}$ , but below this value our simulations do not allow a firm conclusion, due to imprecise knowledge of the long-range behavior of the pair correlation functions.

In order to compare with experiments,<sup>23</sup> we calculate the total scattering<sup>46</sup>

$$I(k) = \rho \sum_{ij} \sqrt{x_i x_j} \sum_{a_i b_j} f_{a_i}(k) f_{b_j}(k) S_{a_i b_j}^{(M)}(k) , \quad (2)$$

where the outer sum is over molecular species  $i$  and  $j$ , the inner sum is over all atom pairs  $a_i$  and  $b_j$ , and the atomic form factors are given by  $f_{a_i}(k)$  and  $f_{b_j}(k)$ . The structure factor  $S^{(M)}(k)$  is defined as

$$S_{a_i b_j}^{(M)}(k) = w_{a_i b_j}(k) + \rho \sqrt{x_i x_j} \int dr r \frac{\sin(kr)}{k} h_{a_i b_j}(r), \quad (3)$$

and includes intramolecular terms  $w_{a_i b_j}(k) = \sin(kd_{ab})/kd_{ab}$ , where  $d_{ab} = |\mathbf{r}_a - \mathbf{r}_b|$ .

To compare with the experiments of Artola *et al.*,<sup>23</sup> we plot  $I(k)/B(k)$ , where  $B(k) = \sum_i x_i \sum_{a_i} f_{a_i}^2(k)$ . The outer sum in the  $B(k)$  expression is over all molecular species  $i$ , and the inner sum is over all atoms  $a_i$  of species  $i$ . The atomic form factors for both SANS and SAXS are given in Ref. 47. For the SANS scattering, the form factors are independent of  $k$ , as is  $B(k)$ .

Our SANS and SAXS results at  $x_t = 0.1$  are compared with experiment<sup>23</sup> in Figs. 10 and 11, respectively. These results were obtained with the pair correlation functions obtained directly in *NPT* simulations with 256000 particles, with no attempt to correct for finite size. Partial structure factors for TBA-TBA, water-water, and TBA-water are included in

the plots. The sum of these contributions gives the total scattering, and we note massive cancellation amongst the like and cross terms. These cancellations were previously discussed by Perera for other systems.<sup>41</sup> We observe that the MD and experimental curves clearly do not agree in magnitude, and there appears to be some missing scaling factor. However, since the paper of Artola *et al.*<sup>23</sup> does not indicate any scaling, we plot the experimental values as they are given in the paper, and our results as obtained by direct application of Eqs. (2) and (3). Overall, the simulation and experimental curves agree in shape, with the present model giving most of the structural features observed in the experiments. We note that the very small  $k$  behavior determined by our simulations is uncertain for the reasons discussed above, and the experiments do not extend to values of  $k$  low enough to determine the behavior as  $k \rightarrow 0$ . We do note that SANS results reported by Misawa *et al.*<sup>15</sup> do not appear to show the presence of a prepeak at very low  $k$ . Finally, we note that, due to large cancellations amongst the structure factors occurring on in Eq. (2), the absence of a prepeak in the experimentally measured total scattering,  $I(k)$ , does not rule out the possibility of domain-related prepeaks in the atom-atom structure factors  $S_{a_i b_j}^{(M)}(k)$ .

#### IV. SUMMARY AND CONCLUSIONS

We have employed very large scale MD simulations to investigate the structural features of aqueous TBA solutions modeled using the CGenFF model for TBA combined with TIP4P/2005 water. We emphasize that among the generally available TBA models which we have considered, only the CGenFF model is miscible with water in all proportions at 300 K, in agreement with the situation for real aqueous TBA solutions. Other evidence in favor of this model for TBA/water solutions was presented in earlier work,<sup>28</sup> where it was shown that calculated compressibilities, heat capacities, and diffusion coefficients are all in reasonable agreement with experiment. Here we show that this is also true of the excess entropy, and of structural properties as measured by SANS and SAXS experiments.

In the present paper, we demonstrate the existence of exceptionally long range structural correlations that fluctuate very slowly in time. These correlations reflect slow variations in a supramolecular, domain-like structure that appears to be a stable state of the TBA/water system. We note that, despite the presence of the microheterogeneous structure, the single particle dynamics exhibits no unusual features. The diffusion coefficients for both water and

TBA are in the range that is normal for the liquid state, and mean square displacement plots show no indication of anomalous diffusion. Thus, the time scale of single particle motion is apparently decoupled from the much slower time scale associated with fluctuations of the domain-like structure. This rather complex physical picture of TBA/water solutions clearly emerges from large-scale simulations.

The long-range, slowly fluctuating structure also poses problems for the calculation of some mixture properties based on MD simulations. This is particularly true of Kirkwood-Buff factors and of scattering functions at small wave vector, both of which strongly depend on the long-range behavior of the pair correlation functions. Our results show that simulations of very large systems for relatively long times are necessary to obtain estimates of these quantities, and even then significant uncertainties remain. Additionally, we note that even with 256000 particles, there appear to be significant differences between *NVT* and *NPT* simulations, almost certainly due to fluctuations of the microheterogeneous structure.

Finally, we note that a TBA/water mixture is a relatively simple system. Therefore, we expect that the behavior described in this paper, likely exists for other aqueous solutions of small amphiphilic molecules. This clearly has implications for simulation of such solutions, especially if one is interested in phase behavior and/or any property that is dependent on long-range correlations.

### Acknowledgments

The financial support of the Natural Science and Engineering Research Council of Canada is gratefully acknowledged. This research has been enabled by the use of WestGrid and Compute/Calcul Canada computing resources, which are funded in part by the Canada Foundation for Innovation, Alberta Innovation and Science, BC Advanced Education, and the participating research institutions. WestGrid and Compute/Calcul Canada equipment is provided by IBM, Hewlett Packard and SGI.

- 
- \* Electronic address: `aup@lptmc.jussieu.fr`
- † Electronic address: `patey@chem.ubc.ca`
- <sup>1</sup> Y. Koga, *Solution Thermodynamics and its Application to Aqueous Solutions* 2<sup>nd</sup> Ed. (Elsevier: Amsterdam, The Netherlands, 2017) chapters V and VI, and references therein.
- <sup>2</sup> Y. Koga, *J. Phys. Chem.* **100**, 5172 (1996).
- <sup>3</sup> K. Iwasaki and T. Fujiyama, *J. Phys. Chem.* **81**, 1908 (1977); **83**, 463 (1979).
- <sup>4</sup> Y. Koga, *Chem. Phys. Lett* **111**, 176 (1984).
- <sup>5</sup> H. Tanaka and K. Nakanishi, *Fluid Phase Equilib.* **83**, 77 (1993).
- <sup>6</sup> D. T. Bowron, J. L. Finney and A. K. Soper *J. Phys. Chem. B* **102**, 3551 (1998).
- <sup>7</sup> K. Tamura, A. Osaki and Y. Koga, *Phys. Chem. Chem. Phys.* **1**, 121 (1999).
- <sup>8</sup> J. L. Finney, D. T. Bowron and A. K. Soper, *J. Phys. Condens. Matter* **12**, A123 (2000).
- <sup>9</sup> P. G. Kusalik, A. P. Lyubartsev, D. L. Bergman, and A. Laaksonen, *J. Phys. Chem. B* **104**, 9533 (2000).
- <sup>10</sup> J. L. Finney, D. T. Bowron, R. M. Daniel, P. A. Timmins and M.A. Roberts, *Biophys. Chem.* **105**, 391 (2003).
- <sup>11</sup> M. Kiselev, D. Ivlev, Y. Puhovski, and T. Kerdcharoen, *Chem. Phys. Lett.* **379**, 581 (2003).
- <sup>12</sup> A. Fornili, M. Civera, M. Sironi, and S. L. Fornili, *Phys. Chem. Chem. Phys.* **5**, 4905 (2003).
- <sup>13</sup> M. E. Lee and N. F. A. van der Vegt, *J. Chem. Phys.* **122**, 114509 (2005).
- <sup>14</sup> S. Paul and G. N. Patey, *J. Phys. Chem. B* **110**, 10514 (2006).
- <sup>15</sup> M. Misawa, T. Sato, A. Onozuka, K. Maruyama, K. Mori, S. Suzuki and T. Otomo, *J. Appl. Cryst.* **40**, s93 (2007).
- <sup>16</sup> H. A. R. Gazi and R. Biswas, *J. Phys. Chem. A* **115**, 2447 (2011).
- <sup>17</sup> B. Kežić and A. Perera, *J. Chem. Phys.* **137**, 014501 (2012).
- <sup>18</sup> M. D. Hands and L. V. Slipchenko, *J. Phys. Chem. B* **116**, 2775 (2012).
- <sup>19</sup> R. Gupta and G. N. Patey, *J. Chem. Phys.* **137**, 034509 (2012).
- <sup>20</sup> D. Subramanian, J. B. Klauda, J. Leys and M. A. Anisimov, <http://arxiv.org/abs/1308.3676> (2013).
- <sup>21</sup> S. Banerjee and B. Bagchi, *J. Chem. Phys.* **139**, 164301 (2013).
- <sup>22</sup> D. S. Wilcox, B. M. Rankin and D. Ben-Amotz, *Farad. Discuss.* **167**, 177 (2013).

- <sup>23</sup> P. A. Artola, A. Raihane, C. Crauste-Thibierge, D. Merlet, M. Emo, C. Alba-Simionesco, and B. Rousseau, *J. Phys. Chem. B* **117**, 9718 (2013).
- <sup>24</sup> S. Banerjee, J. Furtado, and B. Bagchi, *J. Chem. Phys.* **140**, 194502 (2014).
- <sup>25</sup> K. Zemánková, J. Troncoso, C. A. Cerdeiriña and L. Romaní, *Chem. Phys. Lett.* **640**, 184 (2015).
- <sup>26</sup> B. M. Rankin, D. Ben-Amotz, S. T. van der Post, H. J. Bakker,
- <sup>27</sup> K. Zemánková, J. Troncoso, C. A. Cerdeiriña, L. Romaní and M. A. Anisimov, *Chem. Phys.* **472**, 36 (2016).
- <sup>28</sup> S. D. Overduin and G. N. Patey, *J. Chem. Phys.* **147**, 024503 (2017).
- <sup>29</sup> B. Chen, J. J. Potoff, and J. I. Siepmann, *J. Phys. Chem. B* **105**, 3093 (2001).
- <sup>30</sup> S. Weerasinghe and P. E. Smith, *J. Phys. Chem. B* **107**, 3891 (2003); *J. Chem. Phys.* **118**, 10663 (2003); **119**, 11342 (2003); **121**, 2180 (2004).
- <sup>31</sup> B. Hess, C. Kutzner, D. van Der Spoel, and E. Lindahl, *J. Chem. Theory Comput.* **4**, 435 (2008).
- <sup>32</sup> U. Essmann, L. Perera, M. L. Berkowitz, T. Darden, H. Lee and L. G. Pedersen, *J. Chem. Phys.* **103**, 8577 (1995).
- <sup>33</sup> T. Darden, D. York and L. Pedersen, *J. Chem. Phys.* **98**, 10089 (1993).
- <sup>34</sup> H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, A. DiNola and J. R. Haak, *J. Chem. Phys.* **81**, 3684 (1984).
- <sup>35</sup> F. Trudu, D. Donadio, and M. Parrinello, *Phys. Rev. Lett.* **97**, 105701 (2006).
- <sup>36</sup> S. Nosé, *Mol. Phys.* **52**, 255 (1984).
- <sup>37</sup> W. G. Hoover, *Phys. Rev. A* **31**, 1695 (1985).
- <sup>38</sup> K. Vanommeslaeghe, E. Hatcher, C. Acharya, S. Kundu, S. Zhong, J. Shim, E. Darian, O. Guvench, P. Lopes, I. Vorobyov and A. D. MacKerell Jr., *J. Comput. Chem.* **31**, 671 (2010).
- <sup>39</sup> W. Yu, X. He, K. Vanommeslaeghe, and A. D. MacKerell Jr., *J. Comput. Chem.* **33**, 2451 (2012).
- <sup>40</sup> J. L. F. Abascal and C. Vega, *J. Chem. Phys.* **123**, 234505 (2005).
- <sup>41</sup> A. Perera, *Phys. Chem. Chem. Phys.* **19**, 28275 (2017).
- <sup>42</sup> J. G. Kirkwood and F. P. Buff, *J. Chem. Phys.* **19**, 774 (1951).
- <sup>43</sup> J. J. Salacuse, A. R. Denton and P. A. Egelstaff, *Phys. Rev. E* **53**, 2382 (1996).
- <sup>44</sup> T. L. Hill, *Statistical Mechanics* (McGraw-Hill, New York, 1956).

- <sup>45</sup> J. L. Lebowitz and J. K. Percus, *Phys. Rev.* **122**, 1675 (1961).
- <sup>46</sup> A. Perera and B. Lovrinčević, *Molecular Physics* (2018) doi 10.1080/00268976.2018.1483040
- <sup>47</sup> *International Tables for Crystallography*, ed. E. Prince, International Union of Crystallography, 2006, vol. C.

TABLE I: Summary of simulations performed at 300 K. At  $x_t = 0.1$ , two independent *NPT* simulations were performed at 1 bar. The *NVT* simulations were performed at the average *NPT* densities, which were 0.9680 and 0.9686 gm/cc for  $N = 128000$  and 256000, respectively.

| $x_t$ | $N$    | Ensemble | Time (ns) |
|-------|--------|----------|-----------|
| 0.06  | 128000 | NVT      | 200       |
| 0.06  | 128000 | NPT      | 200       |
| 0.06  | 256000 | NVT      | 100       |
| 0.06  | 256000 | NPT      | 100       |
| 0.10  | 256000 | NPT      | 100       |
| 0.10  | 256000 | NPT      | 100       |



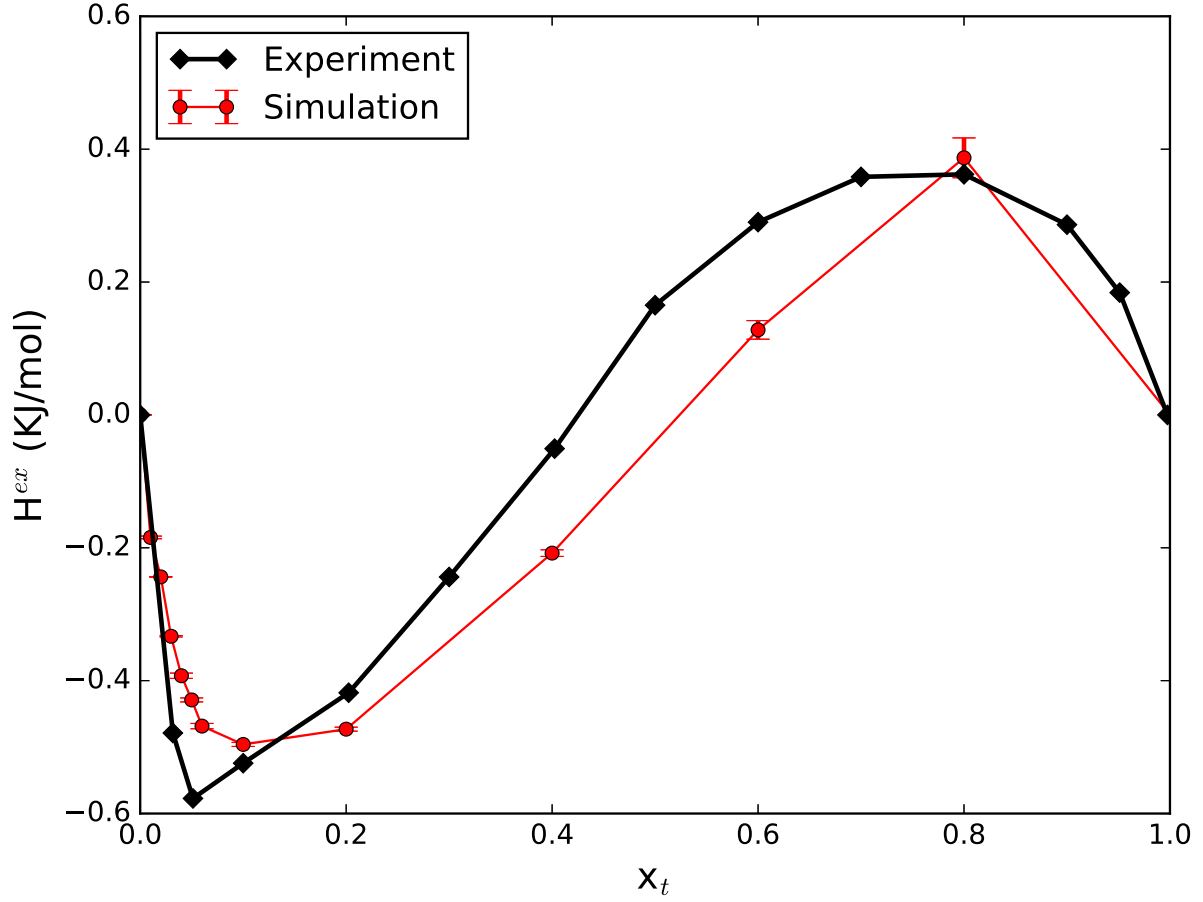


FIG. 1: Excess enthalpy for the CGenFF-TBA/TIP5P2005-water model compared with experiment (Supplementary Material of Ref. 23) at 298 K. The values plotted were obtained in 8000 particle simulations. The errors represent the standard deviation of the average values from four 10 ns simulation blocks.

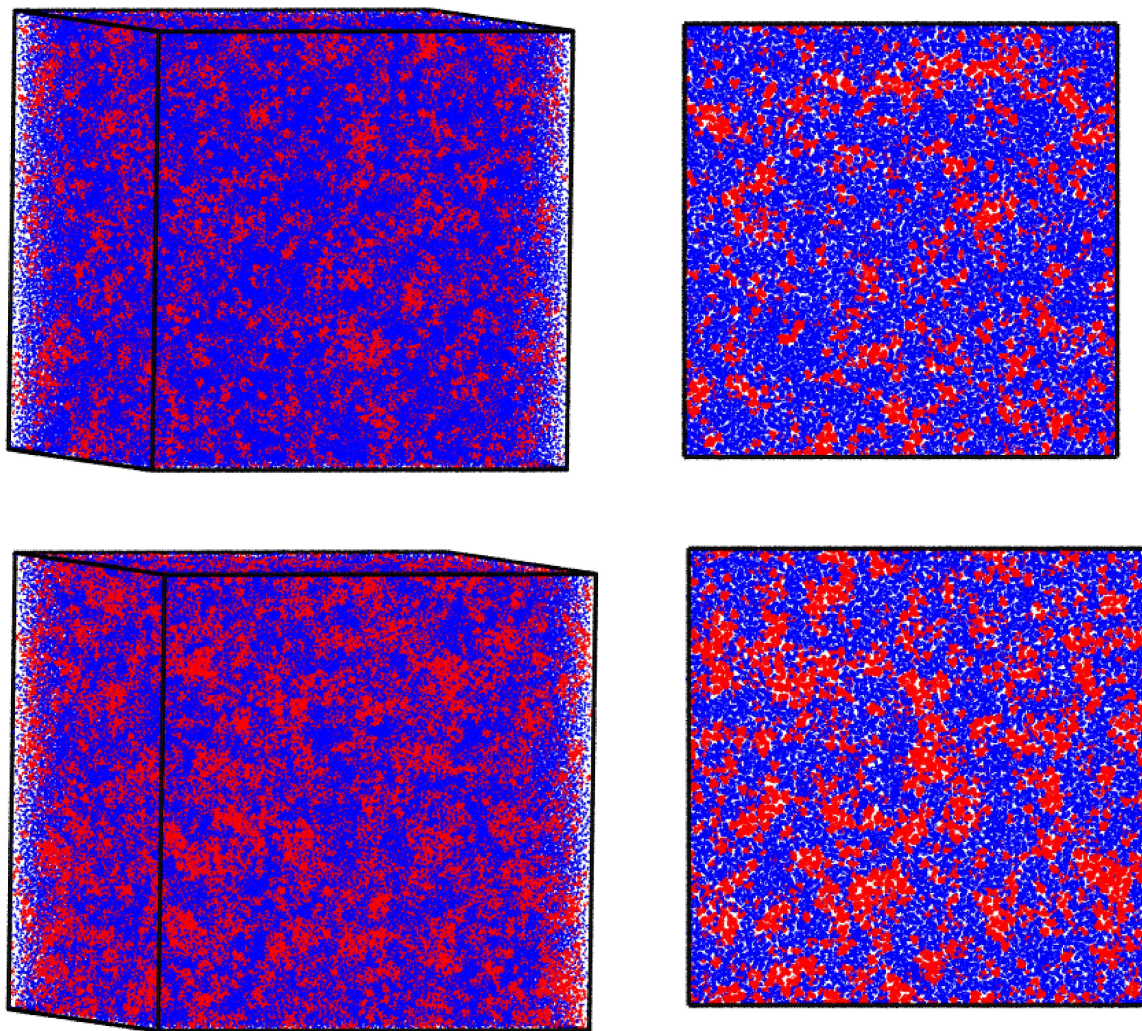


FIG. 2: Snapshots from 256000 particle *NPT* simulations with  $x_t = 0.06$  (top) and  $x = 0.1$  (bottom). The panels on the right are views from cross sections 1 nm thick.

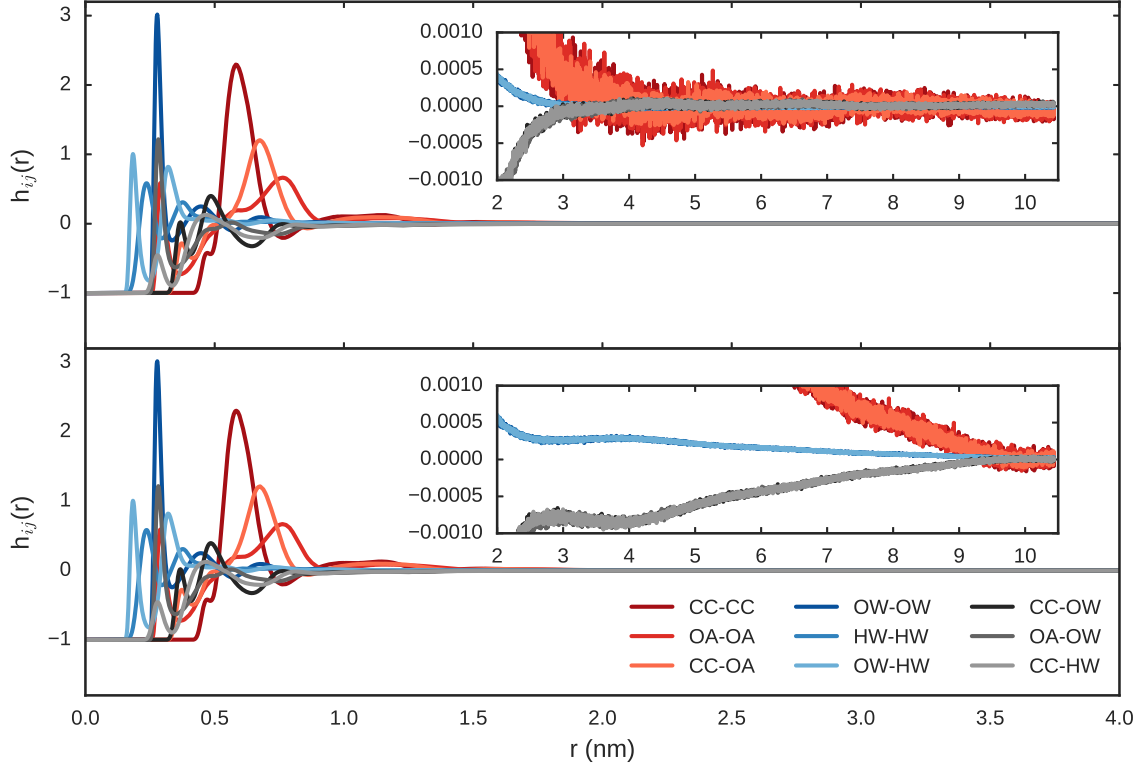


FIG. 3: Illustration of long-range correlations associated with the domain-like structures apparent in Fig. 2. Selected atom-atom correlation functions from 100 ns *NVT* (top) and *NPT* (bottom) simulations with 256000 particles ( $x = 0.06$ ) are shown. The particular atom-atom correlation functions are as indicated in the legend. CC denotes the central carbon of TBA, OA the oxygen of TBA, OW the oxygen of water, and HW the hydrogen of water. To correct for finite-size, as discussed in the text, the TBA-TBA, TBA-water, and water-water *NVT* correlation functions are shifted by  $+0.0001$ ,  $-0.00002$ , and  $+0.000006$ , respectively. Similarly, the TBA-TBA, TBA-water, and water-water *NPT* correlation functions are shifted by  $+0.0008$ ,  $-0.00024$ , and  $+0.0001$ , respectively.

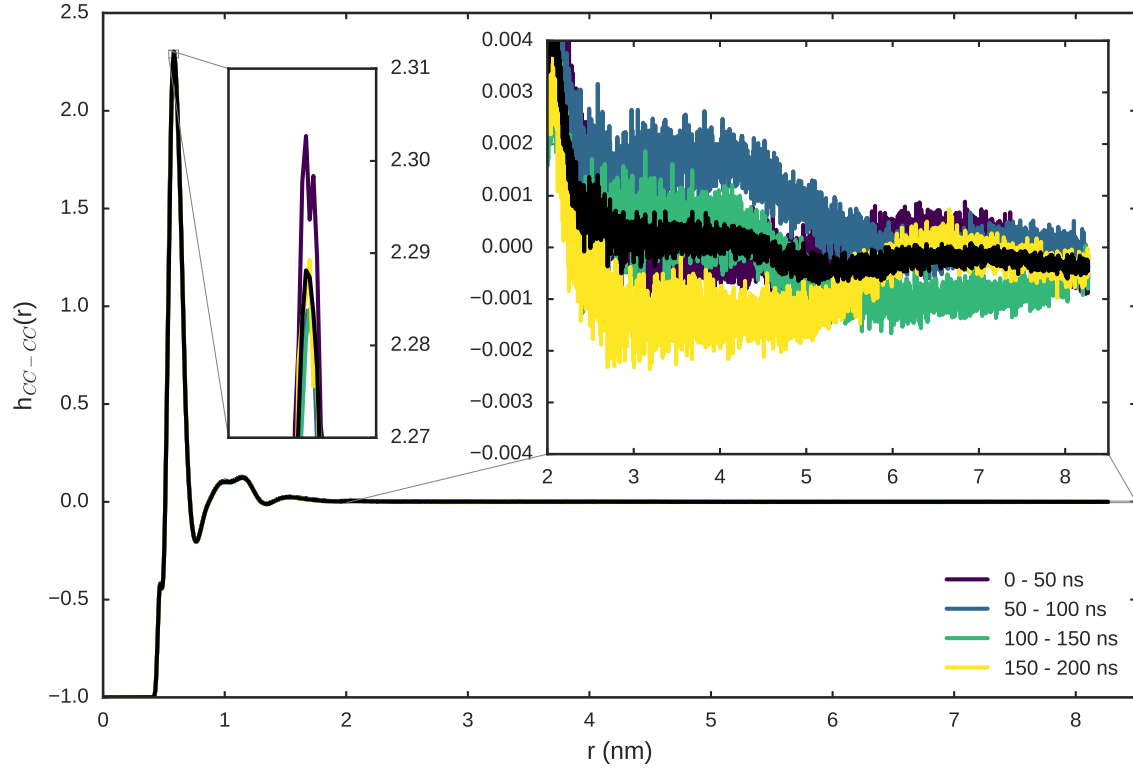


FIG. 4: TBA-TBA pair correlation functions [ $h_{CC-CC}(r) = g_{CC-CC}(r) - 1$ ] at  $x_t = 0.06$  from 50 ns simulations blocks of a 200 ns *NVT* simulation with 128000 particles. The average over the full 200 ns is shown in black.

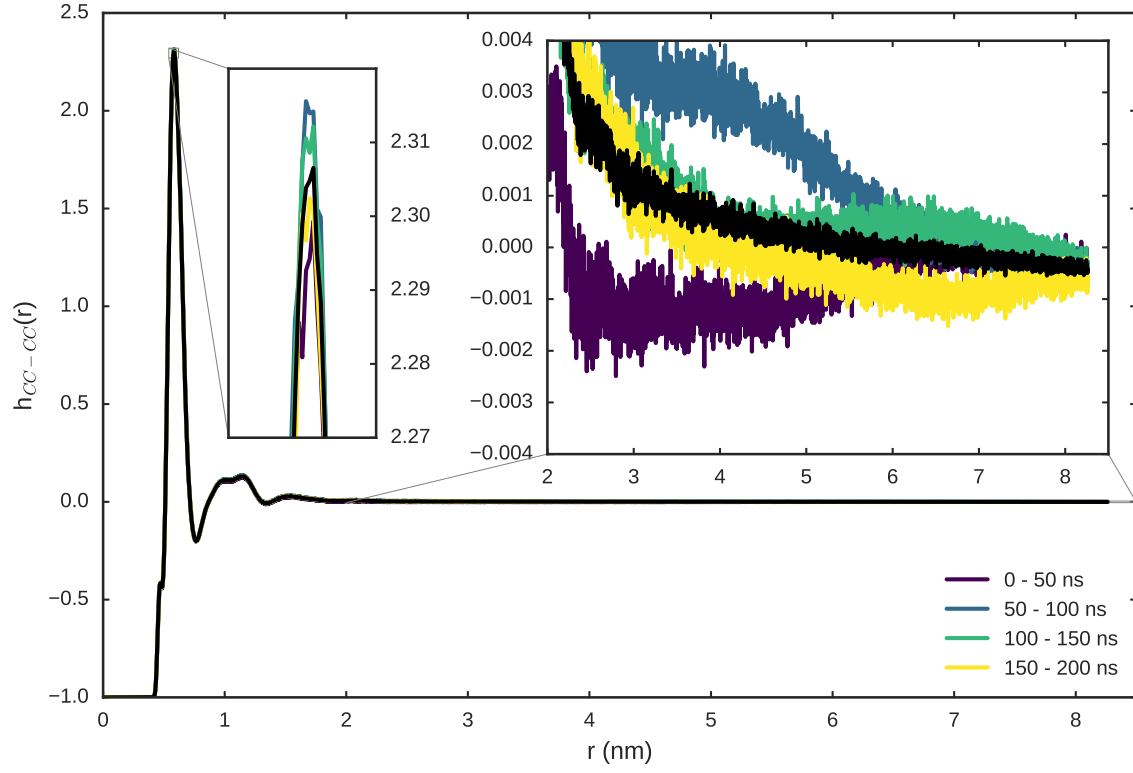


FIG. 5: TBA-TBA pair correlation functions at  $x_t = 0.06$  from 50 ns simulations blocks of a 200 ns *NPT* simulation with 128000 particles. The average over the full 200 ns is shown in black.

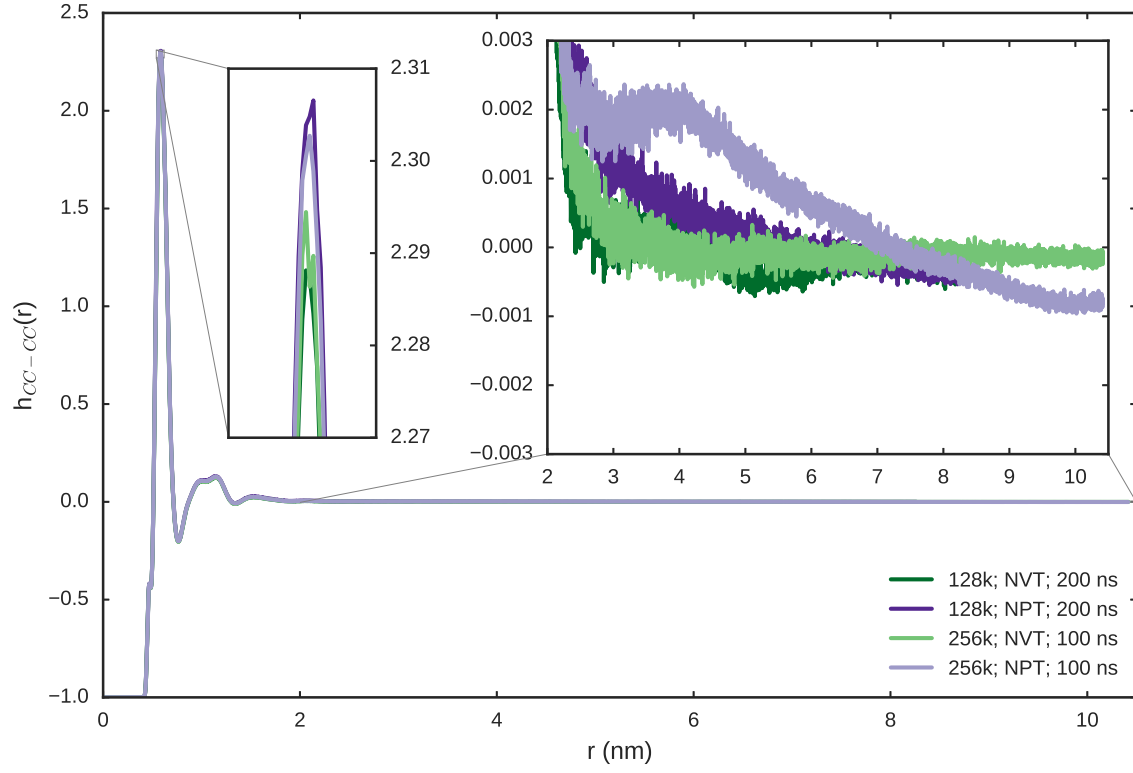


FIG. 6: TBA-TBA pair correlation functions at  $x_t = 0.06$  from *NVT* (green lines) and *NPT* (purple lines) simulations with 128000 (dark lines) or 256000 (light lines) particles. The 128000 and 256000 particle simulations were 200 ns and 100 ns, respectively.

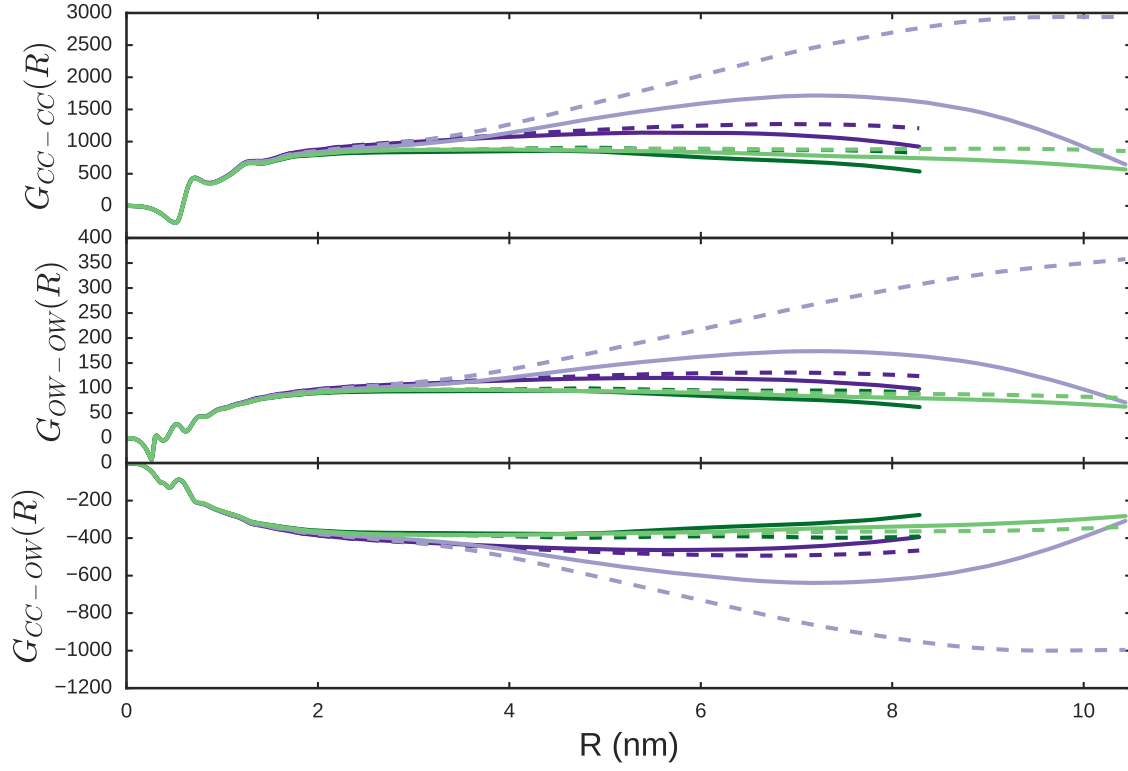


FIG. 7: KB integrals,  $G_{ij}(R)$ , at  $x_t = 0.06$  as functions of  $R$  obtained from *NVT* (green lines) and *NPT* (purple lines) simulations with 128000 (dark lines) or 256000 (light lines) particles. The solid and dashed line represented results without and with estimated finite-size corrections to the pair correlation functions, respectively.

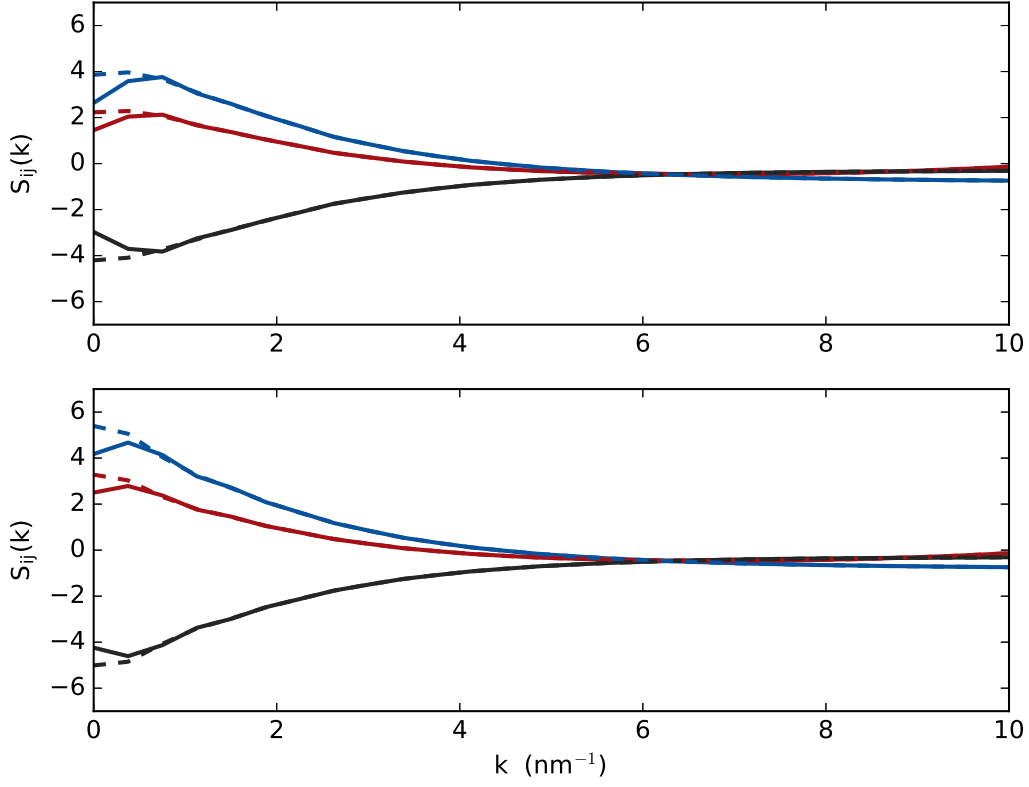


FIG. 8: Partial structure factors at  $x_t = 0.06$ ,  $S_{CC-CC}(k)$  (red),  $S_{CC-OW}(k)$  (black), and  $S_{OW-OW}(k)$  (blue), from 200 ns *NVT* (top) and *NPT* (bottom) simulations with 128000 particles. Results obtained both with and without estimated finite-size corrections to the pair correlations are shown as dashed and solid lines, respectively.



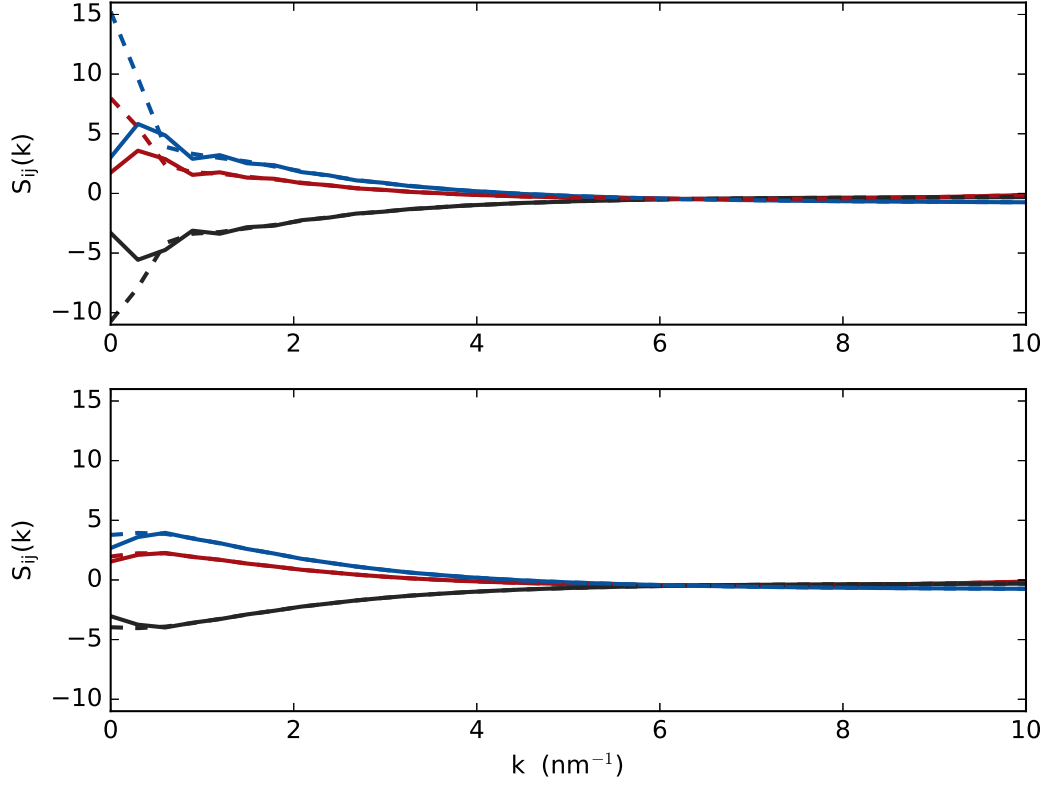


FIG. 9: Partial structure factors at  $x_t = 0.06$ ,  $S_{CC-CC}(k)$  (red),  $S_{CC-OW}(k)$  (black) and  $S_{OW-OW}(k)$  (blue), from 100 ns *NVT* (top) and *NPT* (bottom) simulations with 256000 particles. Results obtained both with and without estimated finite-size corrections to the pair correlation functions are shown as dashed and solid lines, respectively.

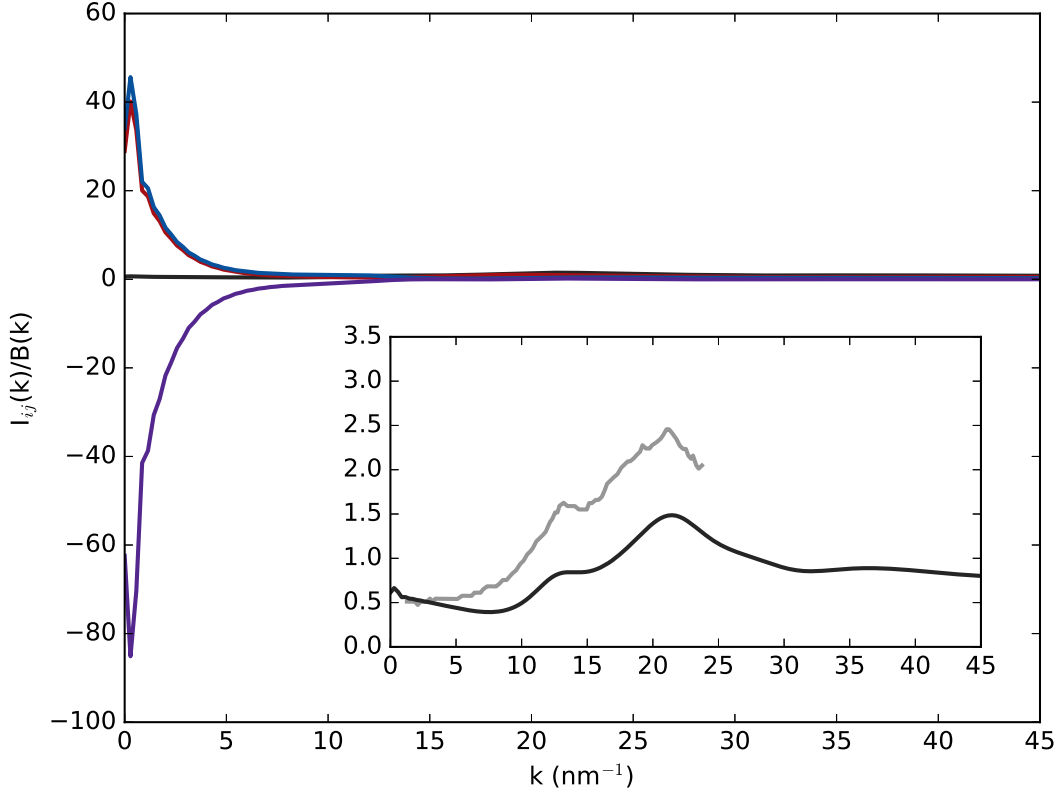


FIG. 10: SANS scattering functions at  $x_t = 0.1$ . The results shown are averages over two independent *NPT* simulations (200 ns in total) employing 256000 molecules. Partial structure factors for TBA-TBA (red), water-water (blue), and TBA-water (purple) are shown. The deuterium scattering length was used for all hydrogen atoms. The total scattering from simulations (black) is compared with experimental results<sup>23</sup> (grey) in the inset.

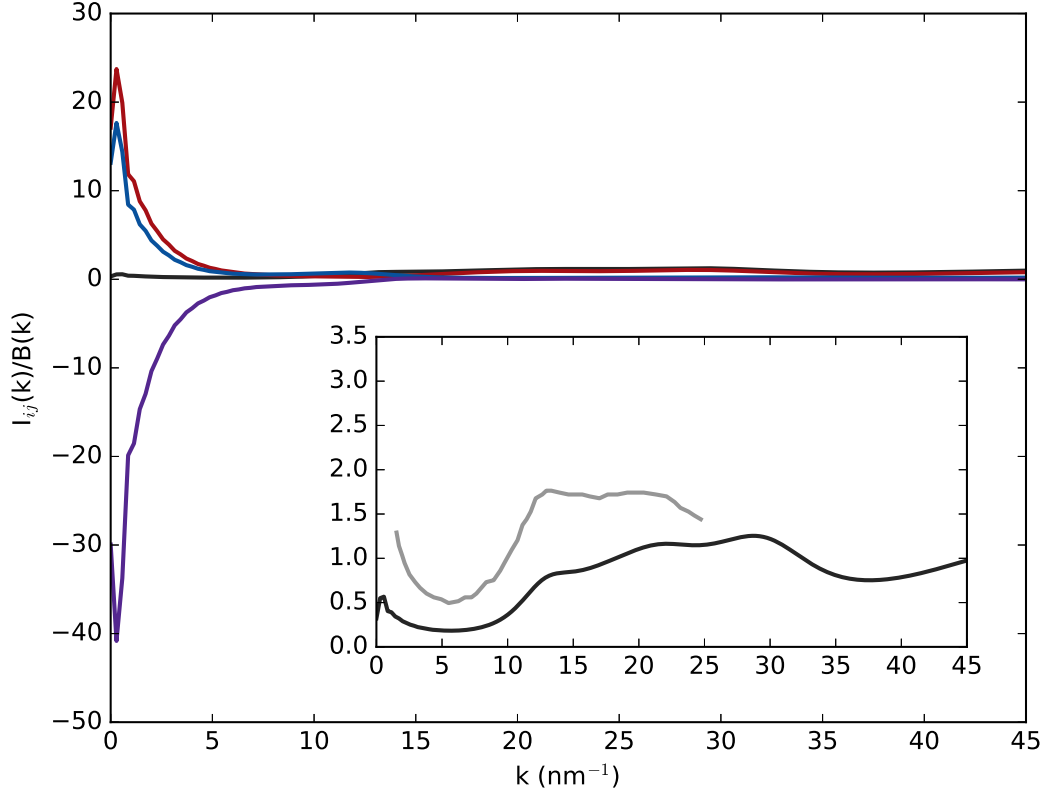


FIG. 11: SAXS scattering functions at  $x_t = 0.1$ . The results shown are averages over two independent *NPT* simulations (200 ns in total) employing 256000 molecules. Partial structure factors for TBA-TBA (red), water-water (blue), and TBA-water (purple) are shown. The deuterium scattering length was used for all hydrogen atoms. The total scattering from simulations (black) is compared with experimental results<sup>23</sup> (grey) in the inset.