Concentration Fluctuation/Microheterogeneity Duality Illustrated with Aqueous 1,4-Dioxane Mixtures

Alena Kolaříková, Aurélien Perera

Concentration fluctuation / micro-heterogeneity duality illustrated with aqueous 1,4-dioxane mixtures

Alena Kolaříková[†‡] and Aurélien Perera[†]

March 9, 2024

[†]Sorbonne Université, Laboratoire de Physique Théorique de la Matière Condensée (UMR CNRS 7600), 4 Place Jussieu, F75252, Paris cedex 05, France.

[‡]Tomas Bata University in Zlín, Faculty of Technology, Department of Physics and Materials Engineering, Nám. T.G. Masaryka 5555, 76001 Zlín, Czech Republic.

Abstract

The structural properties of aqueous 1-4 dioxane mixtures are studied by computer simulations of different water and dioxane force field models, from the perspective of illustrating the link between structural properties at molecular level and measurable properties such as radiation scattering intensities and Kirkwood-Buff integrals (KBI). A strategy to consistently correct the KBI obtained from simulations is proposed, which allows to obtain the genuine KBI corresponding to a given pair of molecular species, in the entire concentration range, and without necessitating excessively large system sizes. The application of this method to the aqueous dioxane mixtures, with all-atom CHARMM dioxane model and 2 water models, namely SPC/E and TIP3P, allows to understand the differences in structure of the corresponding mixtures at molecular level, particularly concerning the role of the water aggregates and its model dependence. This study allows to characterize the dual role played by the concentration fluctuations and the domain segregation, particularly in what concerns the calculated x-ray spectra.

1 Introduction

It is routine nowadays to read papers reporting simulations of complex molecular mixtures, involving mixtures of aqueous and organic solvent, together with ions, co-ions and with much larger molecules such as enzymes, and in the perspective of simulating realistic liquids, such as those involved in biological or pharmaceutical systems[1,2] . The recent interest in the structure of the solvation shell of hyaluronan oligosaccharides in aqueous organic solvent mixtures is among such modern simulation topics.[3–5] In the opposite side of the spectrum in terms of complexity, there are still studies of liquids such

as alcohols or binary aqueous mixtures, with particular focus on their micro-structure. The computer simulations of these latter systems reveal appreciable model dependence, which often focus on minute studies in molecular details. In contrast, the much larger simulations mentioned earlier, rarely focus on the variability that different force field models might introduce. Of course, both types of approach differ in their aims: the first is mostly concerned with the physical properties of the system, while the second is concerned with the question of how well models can describe the local heterogeneity and stability of the mixture. The point of view developed in this paper is that the local heterogeneity plays a capital role in the physics of the system, precisely when different molecular scales are in presence. More precisely, the micro-domains from segregated species could compete in size with larger molecules, and this could affect the distribution of other smaller particles in the system. A particular version of this problem is the influence of the competition of the water solvation shell around a larger object,[6] such as polymer with co-colvent species, the latter which could replace it.[7] For this reason, it appears as fundamental to understand both the role of the local heterogeneity and the how it is influenced by molecular models.

In the present paper, we would like to examine closer binary mixtures of aqueous dioxane, which has been used lately as aqueous organic solvent in the study of more complex mixtures involving hyaluronan and ions.[3–5] In particular, we are interested in why these works choose the TIP3P model for water,[8] rather than more modern water models, such as SPC/E,[9] TIP4P,[10] TIP4P-2005[11] or TIP4P-ew.[12] One explanation might be that the protein community has adapted their investigations on this earlier water model. However, the present study suggests that TIP3P may be a weaker hydrogen bonding water model, hence does not raise the strong micro-segregation problem encountered with these later models. In other words, the TIP3P water model might more "bio-friendly" than recent ones. Yet, as will be demonstrated below, the TIP3P model does not describe well the Hbond structure, that recent model capture. There is clearly a conflict between the necessity of using weak water models when dealing with large scale molecules and the associated heterogeneity, and the necessity of avoiding the strong small scale heterogenity induced by more realistic water models. More specifically, we would like to investigate the role played by the water models in the appearance of micro-heterogeneous/ micro-segregation of water and the organic solvent, which has been the focus of several previous studies,[13–16] sometimes in relation to spurious demixing issues,[17–20] possibly generated to exaggerated concentration fluctuations.[21] These issues are best monitored through the Kirkwood-Buff integrals.[22–25]

The present work was partly motivated by the fact there are only few studies of aqueous 1,4-dioxane from computer simulation perspective,[26,27] and none of these studies mention the fact that water tends to form mostly chain-like aggregates, particularly visible in the simulations at low water concentrations, independently of the models investigated. This latter point is in contrast with the behaviour with most co-solvents, where water tends to form globular aggregates, which are often at the origin the the growth of segregated micro-domains leading to spurious phase separation. The fact that water forms linear aggregates in a particular solvent means that it does not favour the more "spherical/isotropic" tetrahedral connectivity, but rather the chain-like hydrogen bonding with only two water neighbours (not counting the branching parts). This implies a greater miscibility, hence an expected more homogeneous mixing with less anticipated

problems in the evaluation of the KBI. Yet, as will be shown later in this study, these problems are still present, which suggest that water "domain" correlations are intrinsic part of the physical behaviour of this particular liquid at microscopic level. Dioxane is therefore a particularly interesting co-solvent, unlike alcohols for instance.

The outline of the paper is as follows. In Section 2 we remind all the technical details in order to make this paper self consistent, as well as provide a description of the molecular models, related issues and simulation details. Section 3 displays our results while Section 4 provides a discussion of the main observations. Finally, our conclusions are gathered in Section 5.

2 Technical details

2.1 Kirkwood-Buff integrals: theory and experiments

The Kirkwood-Buff Integrals (KBI) are defined as the integrals of the site-site distribution functions $g_{a_i b_j}(r)$, where a_i is an atom of molecular species i and b_j of species j :

$$G_{ij} = \int d\vec{r} \left[g_{a_i b_j}(r) - 1 \right] \tag{1}$$

One can define a corresponding running integral:

$$G_{ij}(r) = 4\pi \int_0^r ds s^2 \left[g_{a_i b_j}(s) - 1 \right] \tag{2}$$

with the KBI given by $G_i = \lim_{r\to\infty} G_{ij}(r)$.

Several important remarks are in order.

(i) The G_{ij} depend only of species index, and not atom indexes (unlike the distribution functions $g_{a_i b_j}(r)$) since the integral is independent of the choice of the integration center in a molecule.

(ii) The distribution functions are defined as the following ensemble average

$$g_{a_i b_j}(r) = \frac{1}{\sqrt{N_i N_j}} \int d\vec{r}' < \sum_{a_i, b_j (i \neq j)} \rho_{a_i}(r') \rho_{b_j}(|\vec{r} - \vec{r}'|) > \tag{3}$$

where $\rho_{c_k}(r) = \delta(\vec{r} - \vec{r}_{c_k})$ is the microscopic density per atom c_k of species k, and N_k is the number of molecules of species k (each atom in a molecular species is counted as unique, hence there are as many atom c_k as the number of molecule of species k). This definition shows that, when species index j and i concern the same species ($i = j$), then the sum excludes the self-contribution. Hence, the limit at large distance r of distribution functions can be written as:[28,29]

$$\lim_{r\to\infty} g_{a_i b_j}(r) = 1 - \frac{\alpha_{ij}\delta_{ij}}{\sqrt{N_i N_j}} \tag{4}$$

where α_{ij} is a coefficient which depends on the statistical ensemble[28,30] and δ_{ij} is the Kronecker symbol. Eq.(4) shows clearly that for identical species, the asymptote of $g_{a_i b_i}(r)$ has an $1/N$ dependence.

(iii) Both equations technically hold in the thermodynamic limit when the total number of particles and the volume are infinite. In finite size systems, hence in conditions met in molecular simulations, the $1/\sqrt{N_i N_j}$ dependence renders the use of Eqs.(1,2) useless, since its integral in Eq.(1) would diverge because of the second constant term in the rhs of Eq.(4). There has been considerable literature on how to evaluate KBI within simulations,[31–34] but this simple and essential fact does not appear to have been considered as the primordial hindering factor. Instead, issues about approximations of the fluctuation related α_{ij} term have been proposed, neglecting the fact that without the exact α_{ij} term, the divergence issue cannot be escaped. Tests conducted on weak microheterogeneous systems with nearly idea KBI, while seamingly support these approximations, may not succed in case of stronger heterogeneity. In Ref.[29] (in the Supplemental Information) we have reviewed the problems met by approaches which ignore the point reminded here. The only way to properly evaluate expressions in Eqs.(1,2) is to restore the infinite number of particle limit $N_i \to \infty$ in Eq.(4), hence killing the size dependent second term. We have proposed to do that by shifting the distribution functions such that they tend to 1 instead of the limit in Eq.(4).

In practice, this is done empirically by numerically estimating the term $\gamma_{ij} = \frac{\alpha_{ij}\delta_{ij}}{\sqrt{N_i N_j}}$, and then applying the correction[29]

$$g_{a_i b_j}^{\text{(corrected)}}(r) = \left(\frac{1}{1-\gamma_{ij}}\right) g_{a_i b_j}(r) \tag{5}$$

We will once more illustrate this point below in the Results section.

In the case of binary mixtures, the KBI can also be evaluated[23,35] from the experimental values of the partial molar volumes $\bar{V}_i$, the isothermal compressibility κ_T and the derivative of the chemical potentials μ_i:

$$D = x_i \left(\frac{\partial \beta \mu_i}{\partial x_i}\right)_{TP} \tag{6}$$

where $\beta = 1/k_B T$ and x_i the mole fraction of species i. Using the following approximations, $\bar{V}_i \approx V_i$ (where V_i is the molar volume of species i), $\kappa_T \approx 0$ (which is justified for dense liquids), and approximating the volume V of the mixture by $V \approx x_1 V_1 + x_2 V_2$, the KBI are given by the expressions:

$$\begin{aligned} G_{12} &= -\frac{V_1 V_2}{V D} \\ G_{11} &= G_{12} + \frac{1}{x_1}\left(\frac{V_2}{D} - V\right) \\ G_{22} &= G_{12} + \frac{1}{x_2}\left(\frac{V_1}{D} - V\right) \end{aligned} \tag{7}$$

Our previous investigations have repeatedly shown[35] that these approximate expressions provide KBI that are nearly as accurate as those obtained through the exact expressions

We note that, from the G_{ij} one can extract three independent expressions for D, which should be the same if the G_{ij} are obtained consistently:

$$D_{12} = -\frac{V_1 V_2}{V G_{12}}$$

$$D_{11} = \frac{V_2}{x_1 (G_{11} - G_{12}) + V} \tag{8}$$

$$D_{22} = \frac{V_1}{x_2 (G_{22} - G_{12}) + V}$$

We apply this methodology to the experimental KBI of the aqueous-dioxane mixture from Ref.[23] We use the convention that species 1 is water (W) and species 2 is dioxane (D). In Fig.1 the original KBI extracted from Ref.[23] are shown in blue for G_{WW}, green for the cross term G_{WD} and red for dioxane G_{DD}. From these three KBI, we extract the three values of D, as shown in corresponding colors in the inset. It is seen that there are some small discrepancies. It is possible to draw a smooth fit of a unique D through these three lines, as shown by the dots. From this unique D, the three KBI are reconstructed using the equations above, and the corresponding curves are shown in cyan, pale green and magenta. The near perfect superposition shows that the fitted D is a good representation of D, and that the approximation in writing the equations for the G_{ij} are correct.

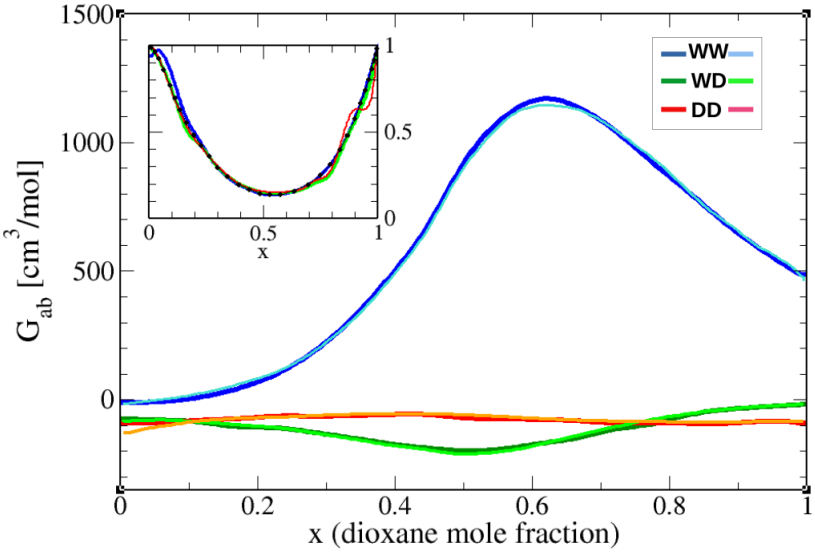

Figure 1: Experimental KBI together with those tested for internal consistency (see text). The inset shows the coefficient D, as obtained by the three routes (lines) and the fitted one (dots).

2.2 Structure factors and radiation scattering intensity

The essential ingredients for the calculation of the radiation scattering are the atom-atom structure $S_{a_i b_j}(k)$ defined as the Fourier transforms of the corresponding atom-atom distribution functions $g_{a_i b_j}(r)$:

$$S_{a_i b_j}(k) = \delta_{ij} + \sqrt{\rho_i \rho_j} \int d\vec{r} \left[g_{a_i b_j}(r) - 1 \right] \exp(i\vec{k}.\vec{r}) \tag{9}$$

where $\rho_m = N_m/V$ is the number density of component m (N_m is the number of molecules of species m and V the system volume). The calculation of the $S_{a_i b_j}(k)$ necessitate properly corrected $g_{a_i b_j}(r)$, for instance using Eq.(5), hence is narrowly related to the calculation of the KBI discussed in the previous sub-section 2.1.

The scattering intensity is calculated through the Debye formula that we have used previously[36]

$$I(k) = r_0^2 \rho \sum_{ij} \sum_{a_i b_j} f_{a_i}(k) f_{b_j}(k) S^{(T)}_{a_i b_j}(k) \tag{10}$$

where $\rho = N/V$ is the number density (with N the total number of particles in the volume V), the $f_{c_n}(k)$ are the form factors of atom c (which are independent of the molecular species n), $r_0 = 2.8179 \times 10^{-13}$cm is the electronic radius, and the total structure factors $S^{(T)}_{a_i b_j}(k)$ are defined in terms of the atom-atom structure factors $S_{a_i b_j}(k)$:

$$S^{(T)}_{a_i b_j}(k) = \delta_{ij} w_{a_i b_i}(k) + S_{a_i b_j}(k) \tag{11}$$

where $w_{a_i b i_i}(k)$ are the intra-molecular atom-atom structure factors, which is obviously a same species property (hence the Kronecker δ_{ij}) in Eq.11). These are directly related to the W-matrix elements in the Reference Interaction Site Model (RISM),[37,38] and are defined for rigid molecule by $w_{a_i b i_i}(k) = j_0(k d_{a_i b_i})$, where $j_0(x) = \sin(x)/x$ is the zeroth-order spherical Bessel function and $d_{a_i b_i}$ is the atom-atom distance between atom a_i and b_i inside the same molecule of species i. However, in the present case, since we consider flexible molecular models, the functions $w_{a_i b_i}(r)$ have been evaluated numerically from the trajectory files by directly sampling of the intramolecular distances and distributions, and then Fourier transformed to obtain the $w_{a_i b_i}(k)$.

2.3 Models and simulation details

There are several examples in past reports of molecular simulations of binary aqueous mixtures which lead to unexpected demixing.[15,17,18,39] Hence, the model dependence of both water and co-solvent is crucial to consider.

In this context, we found that the all atom (AA) CHARMM model of 1,4-dioxane was most appropriate, since it never showed any signs of phase separation, even under long simulations with the most constraining SPC/E model. We studied the united atom(UA) TraPPE model for dioxane under ambient conditions, and found that it had tendency to show liquid-gas phase separation when using $N = 2048$ particles, which it did not show when using $N = 1024$ particles. We note that the website of the TraPPE

model for 1,4-dioxane[40] indicates that this model was tested for $N = 500$ molecules, which might be a reason why it did not show phase separation. In addition, this model is developed precisely to describe liquid-gas phase separation in agreement with experimental binodals. Since critical and sub-critical density fluctuation are most sensitive to finite size conditions, the problems we encountered for $N = 2048$ could result from such considerations. We did not investigate this point any further and restricted ourselves to the AA-CHARMM model.

As mentioned in the Introduction, we consider two water models, namely the SPC/E model and the rather old TIP3P model. The reason for using this latter model is because it is used in other context precisely because it does not pose spurious demixing when used with many other components, including large macromolecules. We are well aware of the existence of suitable TIP4P model families, and tested the TIP4P-2005 model, which was found to give results very close to those with the SPC/E model. However, we did not explore this model any further in the context of the present work.

In the context of classical simulations, the hydrogen bonding interactions are in fact Coulomb charge pairing interaction between the partial charges born by the atomic sites. The oxygen of the SPC/E model has a partial charge of $q_{O_W}^{(\text{SPC/E})} = -0.848e$, that of the TIP3P model is $q_{O_W}^{(\text{TIP3P})} = -0.834e$ and that of the dioxane oxygen atom is $q_O^{(\text{dioxane})} = -0.4$. The hydrogen atoms partial charges are $q_{H_W}^{(\text{SPC/E})} = +0.424e$, $q_{H_W}^{(\text{TIO3P})} = +0.417e$ and $q_H^{(\text{dioxane})} = +0.09$. Finally, the partial charge on the dioxane carbon atom is $q_C^{(\text{dioxane})} = +0.02e$. From these values, it is easy to understand the the water oxygen and hydrogen will generally tend to pair between themselves, leaving aside the weakly charged the dioxane sites. Moreover, this self pairing will be stronger for SPC/E model than for the TIP3P model. All the structural differences discussed in this work are quite simply the result of this simple physics, but made complex through manybody pairing competition.

We have used $N = 2048$ molecules for the base simulations of every mixtures. These systems were studied with 3ns and 6ns trajectories. While this was found sufficient for the TIP3P model, in the case of the SPC/E model, because of need to describe longer ranged correlations with proper statistics, it was found necessary to double the previous size with $N = 16864$ molecules. These larger systems were studied with 20ns long trajectories. For small dioxane mole fractions up to $x = 0.3 - 0.4$, the large size simulations provided good estimates of the RKBI asymptotes. For larger x values $x = 0.6 - 0.9$, the asymptotes of the RKBI could not be stabilized to flat ones, and these continue to meander differently for each run. This is a systematic problem encountered for many other systems, but which is rarely mentioned or discussed in the literature. As mentioned in the sub-section above, the self consistent methodology presented herein helps find the proper KBI value which is found to pass through the meandering of the RKBI curves (see Fig.8-10).

Molecular dynamics simulations of the dioxane-water mixtures were conducted in the Gromacs program package.[41] The initial configurations for all systems were created with Packmol.[42] These initial configurations were first energy minimized and then equilibrated for 3 ns. The simulations were done in the NpT ensemble at T = 300 K and p = 1 bar. The temperature was maintained with the Nosé-Hoover thermostat, whereas

the the Parrinello-Rahman barostat was utilized to keep the pressure constant. The temperature algorithm had a time constant of 0.2 ps and the pressure algorithm was set at 2 ps.

The leap-frog algorithm was used as the integration algorithm, at every time-step of 2 fs. The short-range interactions were calculated within the 1.5 nm cut-off radius. The long-range electrostatics were handled with the partial mesh Ewald method and the constraints with the LINCS algorithm.

3 Results

We present our results in 4 parts. The first one concerns the micro-structure of the aqueous- 1,4-dioxane mixtures as observed through computer simulations, as well its model dependence. Then, in the second part, we analyse how this microstructure is reflected through the various atom-atom distribution functions. In a third part we discuss the reliable obtention of the Kirkwood-Buff integrals through simulation, which are also observable related to the micro-structure of binary mixtures, what is their model dependence and how they match the corresponding experimental KBI. And finally, we analyse how x-ray scattering "experiments" conducted through simulations reflect this micro-structure, in comparison with experimental data.

3.1 Micro-structure of the mixture

A look at snapshots in Fig.2 for both models and three different typical dioxane mole fractions $x = 0.2$, 0.5 and 0.8, from left to right. The most prominent feature is the predominance of chain-like clustering for water molecules, aside the obvious micro-heterogeneity of all these mixtures. Although not very perceptible on these snapshots, there is less clustering in the case of the TIP3P water and the chain clusters are generally smaller than for the SPC/E model. This will be more visible in the correlation functions later in the following section. Neverthless, a closer look at the water clusters reveals that there are also branched chaining, which make this clustering differ from the true chain-like clusters observed for instance in neat alcohols[36] .

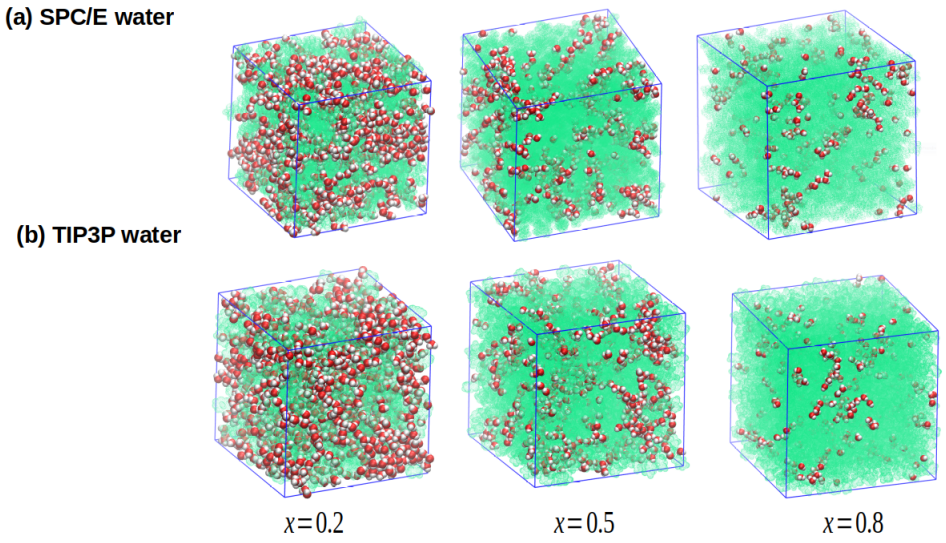

Figure 2: Snapshots of aqueous dioxane for SPC/E water (upper row) and TIP3P water (lower row), and for three dioxane mole fractions of $x = 0.2, 0.5$ and 0.8. Water models are shown with red sphere for the oxygen atom and white sphere for hydrogen atom. Dioxane molecules are shown in "ghost" mode with grass green color.

3.2 Correlation functions

Typical atom-atom distribution functions for the present mixtures are those between the oxygen atoms of each species, which inform us about the local micro-heterogeneity. This latter property is most visibly demonstrated by comparing the pair distribution functions between hydrogen bonding atoms (herein water oxygen atoms) and those between other atoms, as shown below. Additional details on the other atom-atom pair distribution functions are equally shown in section I the Supplemental Material (SM) document.

The water oxygen correlation functions $g_{O_w O_w}(r)$ are shown in Fig.3 for the SPC/E water model and in Fig.4 for the TIP3P model. It is clearly seen that the SPC/E water model has more strongly hydrogen bonded oxygen atoms than the TIP3P model, as can be seen from the magnitude of the first peaks from both insets. In addition, we observe that the range of bonding is larger for the SPC/E model than for the TIP3P model, supporting the fact that water chain clusters are generally longer for the SPC/E model than for the TIP3P model.

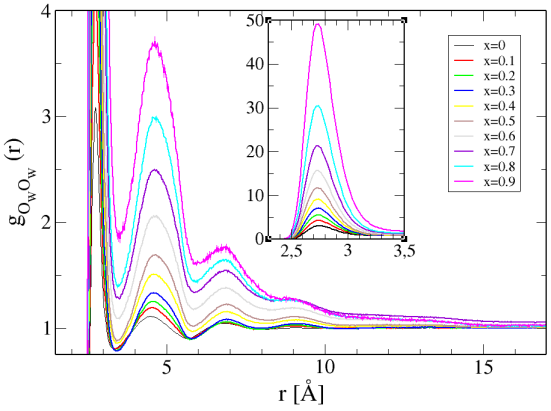

Figure 3: Pair distribution function $g_{O_W O_W}(r)$ between the SPC/e water model oxygen atoms for different dioxane mole fractions x. The inset shows a zoom over the first neighbour peak.

However, the typical chain-cluster feature that is observed in truly chain-like hydrogen bonded oxygen atoms molecular systems, such as neat alcohols[36] or aqueous-DMSO mixtures,[43] which is observed through the depletion correlation of second neighbours and highers, namely the fact that the second and third minimum of $g_{OO}(r)$ are markedly below 1, is not observed in Fig.3 or Fig.4. Instead, the first minimum in $g_{O_W O_W}(r)$ increases steadily with dioxane mole fraction x, even if this is not as fast as in other aqueous mixtures which show micro-heterogeneity, such as aqueous-acetone,[15,20] or aqueous alcohols.[44,45]

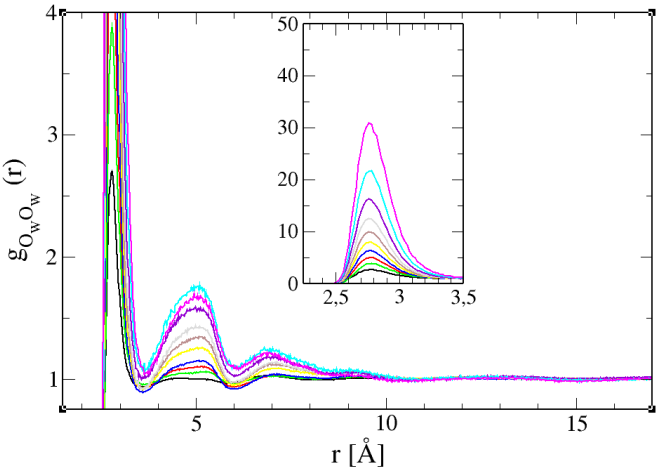

Figure 4: Pair distribution function $g_{O_W O_W}(r)$ between the TIP3P water model oxygen atoms for different dioxane mole fractions x. The inset shows a zoom over the first neighbour peak. The line labels are as in Fig.3

The cross pair distribution function $g_{O_W O}(r)$ between the oxygen atoms of water and dioxane molecules is shown in Fig.5. Both sets of curves look quite similar, unlike previously for the water oxygen-oxygen correlations, which differ quite a bit between the 2 models. This feature indicates that the water interactions and local distribution with dioxane are similar, despite marked differences in water models. This is a direct demonstration of that fact that water aggregation is mostly driven water hydrogen bonding, and that the dioxane molecules act as a solute "bath". This is also a direct consequence of the differences in Coulomb partial charges between the 2 water models and the oxygen atom of the dioxane model. Finally, another remarkable point is the marked minimum at $r \approx 3.5\mathring{A}$, but also at $r \approx 6\mathring{A}$ which witness the depletion correlation between water and dioxane, because of domain segregation between the 2 species. This will induce negative cross KBI $G_{O_W O}$ as the result of integration over the $g_{O_W O}(r)$.

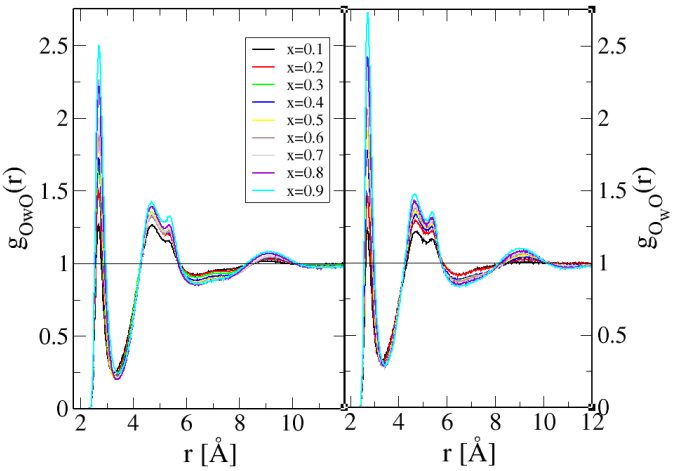

Figure 5: Pair distribution function $g_{O_W O}(r)$ between the oxygen atoms of the water and dioxane molecules are shown for both the SPC/E model (left panel) and TIP3P model (right panel).

The pair distribution functions between the oxygen atoms of the dioxane molecules are shown in Fig.6, comparing between the 2 water models. It can be seen that both sets of curves look very similar, indicating that the water models play little role in the structuring of the dioxane oxygens. In addition, we notice that these curves look much more modest in amplitude than those between the water oxygen atoms seen in Fig.3 and Fig.4, confirming that the micro-heterogeneity is driven by the water molecules.

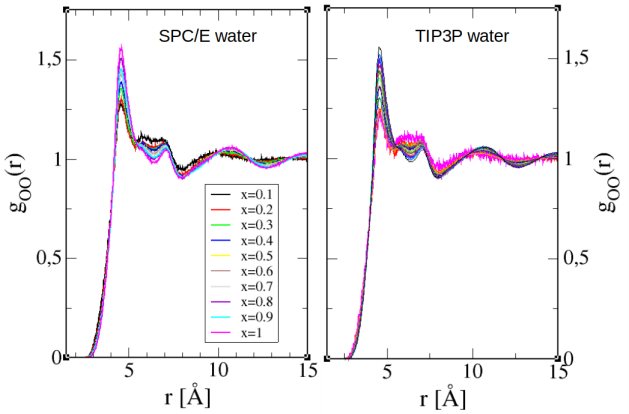

Figure 6: Pair distribution function $g_{OO}(r)$ between the dioxane oxygen atoms for different dioxane mole fractions x.

The pair distribution functions between the carbon atoms of the dioxane molecules for both mixtures are shown in Fig.7. Again, we observe features very similar to those

between the dioxane oxygen atoms, namely that the general features are the same indifferently to water models, and the amplitude of the correlations are rather small.

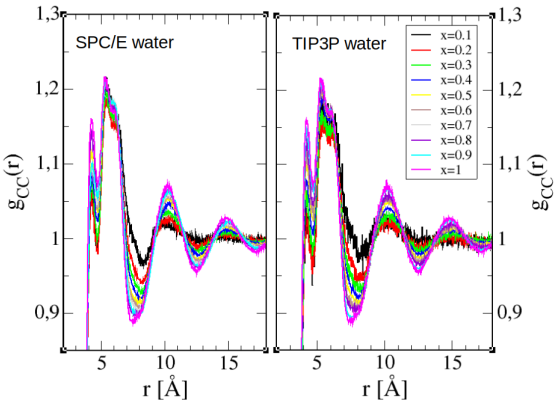

Figure 7: Pair distribution function $g_{CC}(r)$ between the dioxane carbon atoms for different dioxane mole fractions x.

This analysis of the atom-atom pair distribution confirms the micro-heterogeneous status of these mixtures and that it is essentially driven by the clustering of the water molecules. Refer to the SM document for more distribution functions than those displayed here.

3.3 KBI

Before computing the KBI as in Eq.(1) from the pair distributions discussed in the previous sections, we need to address the problem of the asymptotes of these functions. As discussed in our previous works,[29,45] these asymptotes are affected by two problems, that related to finite size discussed in Section 2.1, and the distortion induced by the micro-heterogeneous nature of these mixtures, which affect the correlation functions in the medium to long range part, essentially above $10\mathring{A}$.

These 2 problems are illustrated in Fig.8 for the case of the $x = 0.2$ aqueous-dioxane with SPC/E water. The three RKBI functions, the water-water $G_{WW}(r)$ (in blue), the cross water-dioxane $G_{WD}(r)$ (in green) and the dioxane-dioxane $G_{DD}(r)$(in red) are displayed for different simulation conditions, together with the experimental KBI of Fig.1 and the predicted KBI, both shown in horizontal lines with dotted and dashed lines, respectively. They are seen to be very close, indicating that the KBI of the $x = 0.2$ system are quite well described by this mixture model. This is not always the case, particularly at higher dioxane content.

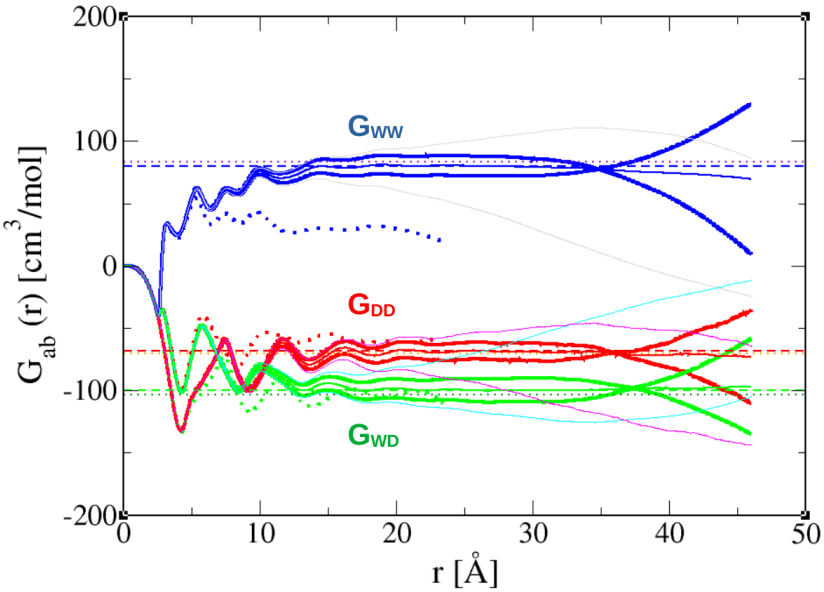

Figure 8: RKBI for the $x = 0.2$ aqueous-dioxane mixture with the SPC/E water model.

For each of the RKBI, 6 different curves are shown. The RKBI for $N = 2048$ system is shown in dotted curves. For the $N = 16864$ system, the original (distorted) curves are shown pale thin lines (gray for $G_{WW}(r)$, cyan for $G_{WD}(r)$ and magenta for $G_{DD}(r)$). The asymptote-corrected curves are shown in their respective colors, and for 2 independent runs. The average of the 2 runs is equally shown in thin curve, and is seen to follow nearly horizontal line, which is used as a guide to evaluate the calculated KBI. In addition, once a first estimate of the 3 KBI is obtained from direct evaluation of the corresponding integrals, these are checked through the procedure of calculating the corresponding D_{ab}, as explained in Section 2.1. In general, it was found that the G_{DD} KBI was the best obtained, because of the more or less flat asymptote of the corresponding RKBI, owing to the lesser long range distortion in the g(r) induced in the distribution of dioxane, which is more homogeneous as compared with that of water. Then, the D_{ab} of component pair with the most distorted asymptote (usually the water-water component D_{WW}) is adjusted to match the remaining two others, by shifting appropriately the first estimation of this KBI. In the case for $x = 0.2$, there was very little adjustment to make since the horizontality of the calculated curves is very apparent in Fig.8. This is not the case for the next example. On this example, we also note that the small system $N = 2048$ does not permit to obtain proper KBI, particularly for G_{WW}. The example also illustrates the importance of the shifting procedure in Eq.(5) which allows to obtain nicely horizontal RKBI from widely initial distorted curves. This step is not properly appreciated in the KBI calculation of the literature,[31,33,34] which the calculations are often applied to systems that are not demanding.

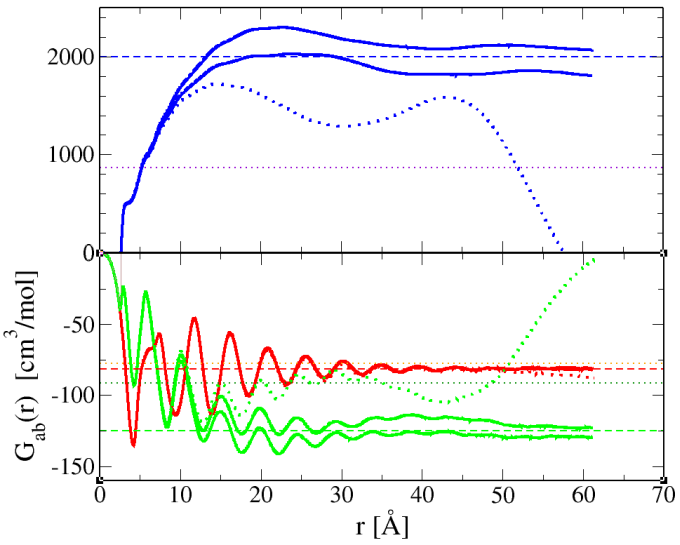

Figure 9: RKBI for the $x = 0.8$ aqueous-dioxane mixture with the SPC/E water model.

Fig.9 shows a similar calculation but for the dioxane content $x = 0.8$ and the same mixture model. This time around, the original distorted KBI are not shown. Instead, the RKBI obtained from a 3ns run of the $N = 16864$ system are shown in dotted lines. The difference in statistics with the two different 20ns runs of the same system size are quite remarkable, indicating the importance of both using large systems and long runs. It can be seen that the long runs still gives different KBI/RKBI estimates, even though with similar and closer trends. This sensitivity of the KRBI is one of the plague of the calculation of KBI from simulation, which is not well documented in other work on the KBI. An important feature concern the long range oscillatory nature of the asymptote of the RKBI curves, particularly visible on the water-water $G_{ww}(r)$ which we attribute to domain oscillations,[45] that is the correlation between the chain-like domains that can be seen in the rightmost snapshots in Fig.2. The difference between the experimental KBI (dotted horizontal lines) and the simulation estimate values (dashed horizontal lines) is quite apparent, indicating that the model is not as appropriate as it was for the $x = 0.2$ case.

For completeness, we show an example for the mixture with the TIP3P model for $x = 0.9$ in Fig.10. In this case, we see that domain oscillations are quite marked, making the estimation of G_{WW} uneasy. In this case, G_{WW} is evaluated through the procedure of Section 2.1 and the predicted value nicely fits in the middle of the domain oscillations of $G_{WW}(r)$. The domain oscillations are also visible for the cross RKBI $G_{WD}(r)$. Another remarkable feature is the closeness of the experimental and predicted KBI values for G_{DD} and G_{WD}, as compared to the large difference for the G_{WW}. This is a direct consequence of the fact that the micro-segregation is driven by water: the distribution of dioxane is insensitive to large variations in micro-heterogeneity of the water clusters.

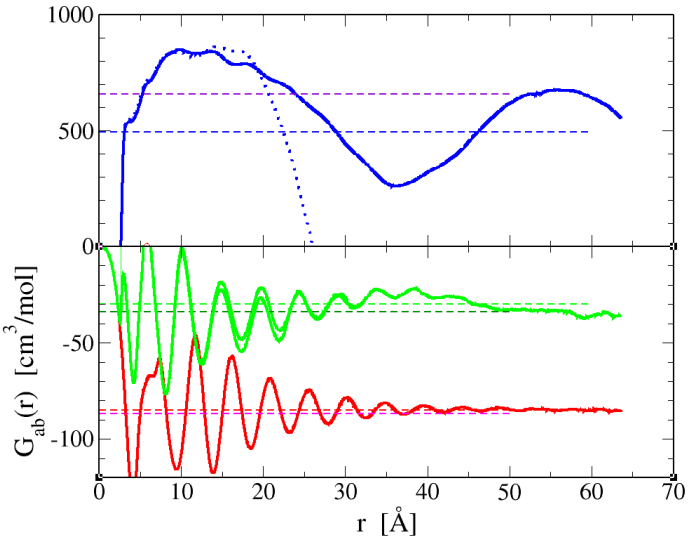

Figure 10: RKBI for the $x = 0.9$ aqueous-dioxane mixture with the TIP3P water model.

From these results one can finally compute a reasonable estimate of the calculated KBI for the two mixture models, as illustrated in Fig.11. From these results, the efficiency of both water models can be appreciated. Both models are relatively good at small concentration, and the SPC/E model has a better range of validity (up to $x \approx 0.5$), while the TIP3P model is off already from $x > 0.25$. On the other hand, for large x values the SPC/E model clearly overestimates the KBI, while the TIP3P model underestimate them. Both models have a maximum/extremum shifted to $x \approx 0.8$ instead of the experimental value $x = 0.6$. The overestimation of the SPC/E is directly related to the overestimation of the water aggregation of this model, noticed for other mixtures.[15,18] In contrast, the TIP3P model is clearly a weaker model of aggregation. This may be one reason why it is preferentially used in calculation of large simulations with complex mixtures involving several components and large molecules. In such cases, the segregative nature of SPC/E may be an obstacle.

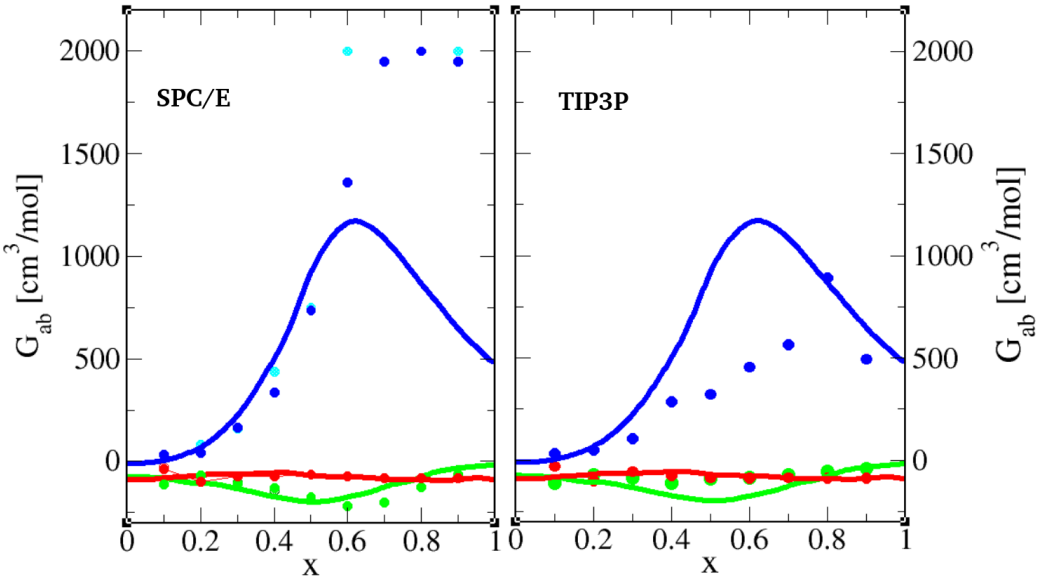

Figure 11: Calculated KBI for the 2 mixture models compared with the experimental results. Simulation date is shown in dots, blue for G_{WW}, green for G_{WD} and red for G_{DD}. The original estimates prior to the correcting method of Section 2.1 are shown in cyan dots for the SPC/E model, where this method is most effective.

3.4 Results for the x-ray scattering

3.4.1 The structure factors

The three structure factors of the oxygen atoms of the mixture model with the SPC/E water model are shown in Fig.12 for typical dioxane concentrations of 0% (pure water), 20%, 50%, 80% and 100% (pure dioxane). The left panel shows the water oxygen structure factor $S_{O_W O_W}(k)$. It is interesting to observe that the shoulder peak of water (curve in orange) at $k \approx 2\mathring{A}^{-1}$, which represents water-water contact[46] , vanishes with dioxane content, while the Hbonding peak at $k \approx 3\mathring{A}^{-1}$[46] remains unchanged even for the 80% dioxane mixture. The other interesting feature is the rise near $k \approx 0$, which reflects both the concentration fluctuations at $k = 0$, and the formation of clusters for $0 < k < 1.5\mathring{A}^{-1}$. We emphasize that the exact $k = 0$ are not representative, because of the large deviations of the RKBI at the end of the half box (as can be seen in the blue curves Fig.8 for instance), directly affect the $k = 0$ value. This problem is the plague of the calculation of structure factors from simulations.[29]

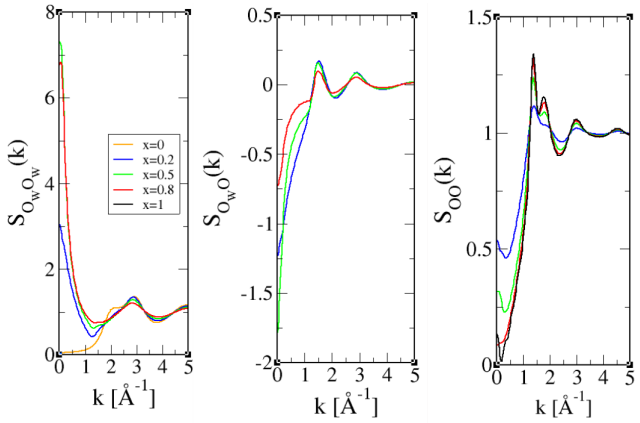

Figure 12: Oxygen-oxygen structure factors for SPC/E water $S_{O_WO_W}(k)$ (left panel), cross water-dioxane $S_{O_wO}(k)$ (middle panel) and dioxane $S_{OO}(k)$ (right panel), for the typical dioxane mole fractions $x = 0.2$ (blue), 0.5 (green) and 0.8 (red). Pure water (orange) and dioxane (black) are equally shown in the left and right panels, respectively.

The $k = 0$ raise is more dramatic for water than for dioxane (right panel) directly witnessing the differences in fluctuations and aggregations discussed in the previous sections. Finally, the negative $k = 0$ part of the cross water-dioxane structure factor $S_{O_WO}(k)$ are a direct consequence of the domain segregation between the two species and the resulting anti-correlations, which were discussed in commenting Fig.5. This point is important since it will contribute to decrease the large positive $k = 0$ contributions to the scattering intensities $I(k)$ from $S_{O_WO_W}(k)$, as will be discussed below. Complementary information on other atom-atom structure factors for both mixture models are displayed in section II of the SI document.

3.4.2 The scattering intensities

The x-ray scattering intensities, as obtained from Eq.(10) are shown in Fig.13 for both sets of mixture models. It can be observed that there are small differences between the 2 water models, but the overall shape is the same for every dioxane concentration. The most remarkable difference is in fact that for pure water (blue curves) between the 2 water models. The very small k-values cannot be trusted because of the intrinsic difficulties in evaluation the KBI integrals mentioned in Section 3.3.

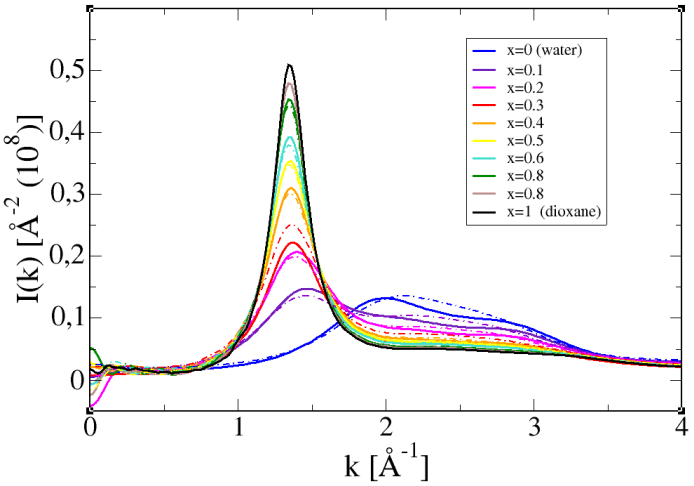

Figure 13: X-ray scattering intensities $I(k)$ for both mixture models and all concentrations. The SPC/E data is shown in thick lines and the TIP3P data in thinner dash-dotted lines.

Fig.14 shows a comparison of the calculated x-ray scattering experiments against the experimental values[47] for the quantity $kI(k)$. This method conveniently kills the $k = 0$ problems and highlights large k differences. It can be seen that both mixtures models describe quite well the overall features of I(k). The main differences are seen for pure water ($x = 0$) where the TIP3P model is seen to totally miss the very specific shoulder peak feature of water at $k = 2\mathring{A}^{-1}$ and $3\mathring{A}^{-1}$, which the SPC/E model reproduces relatively well (minding the amplification of differences due to the multiplying factor k). This is an important flaw of the TIP3P model, which is seen not to reproduce well the pronounced differences of contact between the diameter contact at $\sigma = 3\mathring{A}$ ($k = 2\pi/\sigma = 2\mathring{A}^{-1}$) and the Hbond contact at $r_{HB} = 2\mathring{A}$ ($k = 2\pi/r_{HB} = 3\mathring{A}^{-1}$).

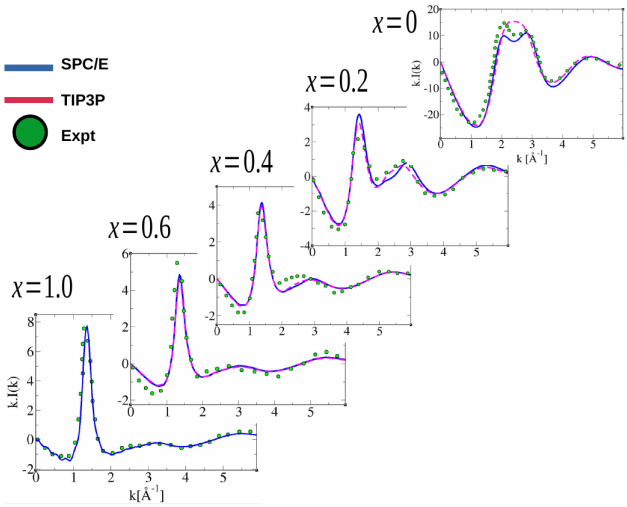

Figure 14: X-ray scattering intensities compared with experimental results from Ref.[47] for various dioxane concentration x. The scanned experimental data is shown in symbols, while the SPC/E and TIP3P mixture data is shown in blue and magenta lines, respectively.

In fact the TIP3P model smears these fundamental differences of the real water model, which are responsible for the typical features of this liquid, such as the hydrophobic effect in particular, which has been initially described[48] as a direct consequence of the existence of microscopic patches of tetrahedrally linked water and disordered water, a model which appear to be still in use.[49,50] This smearing may explain why this model is popular in context where the strong self-heterogeneity of water may come as an obstacle in large scale simulation of complex systems.[3–5]

4 Discussion

In the title we have highlighted the duality between concentration fluctuations and micro-heterogeneity, a topic we have discussed several times previously,[16,45] but which gains a new importance in this study of this particular mixture, for the following reasons. In most of our previous studies of aqueous mixtures, such as aqueous acetone[15,20] or aqueous alcohols,[44] we have encountered spurious demixing issues, which have been equally met by several other authors.[17,39,51] Notable exceptions concern aqueous amides,[52,53] aqueous amines[54] and aqueous-DMSO.[43] In this latter binary mixture, it was noted that water-water correlations show a marked second and third neighbour depleted correlations, which was equally reported previously in Ref.[55] but without comment about this peculiarity. In Ref.,[43] we have traced back this behaviour to water chain-clustering, bearing its very strong similarity with chain-clustering in neat alcohols.[36]

Aqueous dioxane present a similar chain-like clustering, visible in the snapshots

(Fig.2) albeit with no obvious depletion correlations of the first neighbours (Fig.3 and Fig.4). Yet, these mixtures exhibit stable micro-segregation without demixing. From this point of view 1,4-dioxane appears as a very interesting solute to mix with water models in the perspective of studying micro-segregation driven by water.

Fig.15 shows the species partial contributions, $I_{WW}(k)$ in blue, $I_{WD}(k)$ in green and $I_{DD}(k)$ in red, to the total scattering $I(k)$, such that $I(k) = I_{WW}(k) + I_{WD}(k) + I_{DD}(k)$. It can be clearly seen that the magnitudes of the total intensity $I(k)$ does not reflect that of the partial contributions, and in fact the result of a cancellation between their large positive and negative contributions. The signs of such contributions are a direct consequence of the large domain segregation observed in Fig.2 for $x = 0.5$, the cross species negative contribution due to cross domain correlations below 1 as in Fig.5 and mirrored in Fig.12 (middle panel).

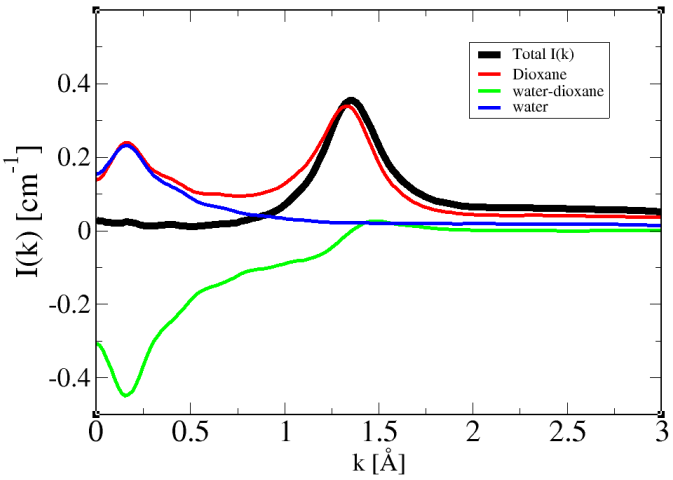

Figure 15: Illustration of partial species contributions to total scattering intensity $I(k)$, for the SPC/E mixture model and dioxane mole fraction $x = 0.5$.

This figure is also a nice illustration of the duality between concentration fluctuations, from $I(k) = 0$[56,57] and the $k \neq 0$ part in the range $0 < k < 0.5\mathring{A}^{-1}$ range, where the prominent positive and negative contributions are observed, and which we attribute to the segregated domain correlation range. From this interpretation, $I(k)$ alone would show only concentration fluctuations $I(0)$ and the domain segregation would be *hidden* to observation through radiation scattering. In particular, the pre-peak like features cancel exactly, to leave no sign of them in the final $I(k)$. For instance, this is not the case in neat alcohols, where the positive and negative pre-peak features leave an observable pre-peak trace in $I(k)$,[36] of for the case of micellar solutions, where the pre-peak in $I(k)$ is equally observed.[58] We conjecture here that the fact that the pre-peak is observable or not is directly related to the underlying *cybotactic order*, in the sense that self-assembled structures have observable lifetimes, such as chain aggregates in alcohols or micelles in micro-emulsions. Further developments into this con-

jecture would necessitate the development of a meso-scale description of such systems, which means no-coarse graining description where the fluctuation aspect of such self-assembled structure would be irremediable lost. Instead, it would be desirable to build a meso-scale description which would preserve the duality of concentration fluctuation and micro-segregation, possibly through field theoretic techniques.[59,60] We conjecture that such approaches would allow a theoretical study of complex systems such as those in Ref.,[3–5] which are currently studied by computer simulation using forgiving water models, which are a way to avoid cumbersome water hydrogen bond network problems.

Finally, there is this issue of which water model is "better". The answer to this question depend on which experimental observable one wants to focus on. From the KBI on Fig.11 we can say that SPC/E is able to well describe the shape of the curves in the high water content region. TIP3P clearly underestimate the KBI in the entire range. So SPC/E "wins". From Figs.13 and 14 for the x-ray scattering intensities, it is less obvious to tell the 2 models apart, except for pure water, where SPC/E clearly wins because it describes the Hbonding better than TIP4P. It is interesting that this feature of SPC/E does not allows to describe the water aggregate distribution in the low water content region. This illustrates indirectly the importance of the description of the micro-heterogeneity itself, which we have advocated since our earlier works.

5 Conclusion

In this work, we have studied the aqueous 1,4-dioxane mixtures under the perspective of how different water models handle the micro-heterogeneity of binary aqueous mixtures, how this property affects the structural properties such as the pair distribution functions, and how it affects the way computer simulations can describe it, in particular through the calculation of the Kirkwood-Buff integrals. The calculation of these latter properties are recurrent considerations in the past the decades. Herein, we have contributed to new ways of consolidating the robustness of the calculations by using a self-consistent method. While this method provides robust KBI estimates, it is still not able to provide proper $S_{ab}(k = 0)$ values in case of strong heterogeneity, as is the case in this study. This is a direct consequence of how the long range part of the micro-heterogeneity is difficult to capture by statistical methods in computer simulation, even in the case of large systems and long runs. This problem has been noticed before,[61] and is related to the inherent difficulties of handling large scale heterogeneity and the associated kinetics. Perhaps the most important message here is the fact that structure of water in mixing conditions is rather long ranged, possibly beyond the $30\mathring{A}$ scale, and this is associated with specific kinetics which are beyond the 20ns range. This appears as very counter intuitive, since spectroscopic analysis show that water decorrelates quite fast near solutes.[62,63] We note that this fast water-solute decorrelation is not incompatible with the existence of long range Hbond water-water correlations that meander through the system. Our study shows that some of these issues remain unsolved, and we have conjectured in the Discussion section 4, that very different methods, such as mesoscale methodologies, may be required to advance further into this topic.

Acknowledgment

A.K. thanks the Erasmus program for funding and the Laboratoire de Physique Theorique de la Matière Condensée for her stay in the course of her PhD intership.

References

[1] Tamar Schlick, Stephanie Portillo-Ledesma, Christopher G. Myers, Lauren Beljak, Justin Chen, Sami Dakhel, Daniel Darling, Sayak Ghosh, Joseph Hall, Mikaeel Jan, Emily Liang, Sera Saju, Mackenzie Vohr, Chris Wu, Yifan Xu, and Eva Xue. Biomolecular modeling and simulation: A prospering multidisciplinary field. *Annual Review of Biophysics*, 50(1):267–301, 2021. PMID: 33606945.

[2] Bruno Rizzuti and Bruno Rizzuti. Molecular simulations of proteins: From simplified physical interactions to complex biological phenomena. *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics*, 1870(3):140757, 2022.

[3] Eva Kutálková, Marek Ingr, Alena Kolaříková, Josef Hrnx□ičík, Roman Witasek, Martina Hermannová, Ondřej Štrympl, and Gloria Huerta-Ángeles. Structure and dynamics of the hyaluronan oligosaccharides and their solvation shell in water: organic mixed solvents. *Carbohydrate Polymers*, 304:120506, 2023.

[4] Alena Kolaříková, Eva Kutálková, Josef Hrnx□ičík, and Marek Ingr. Hyaluronan oligosaccharides form double-helical duplexes in water:1,4-dioxane mixed solvent. *Carbohydrate Polymers*, 326:121632, 2 2024.

[5] Eva Kutálková, Josef Hrnx□ičík, Roman Witasek, Marek Ingr, Gloria Huerta-Ángeles, Martina Hermannová, and Vladimír Velebný. The rate and evenness of the substitutions on hyaluronan grafted by dodecanoic acid influenced by the mixed-solvent composition. *International Journal of Biological Macromolecules*, 189:826–836, 10 2021.

[6] Amish J. Patel, Patrick Varilly, and David Chandler. Fluctuations of water near extended hydrophobic and hydrophilic surfaces. *The Journal of Physical Chemistry B*, 114(4):1632–1637, 2010. PMID: 20058869.

[7] Divya Nayar and Nico F. A. van der Vegt. Cosolvent effects on polymer hydration drive hydrophobic collapse. *The Journal of Physical Chemistry B*, 122(13):3587–3595, 2018. PMID: 29443520.

[8] William L. Jorgensen, Jayaraman Chandrasekhar, Jeffry D. Madura, Roger W. Impey, and Michael L. Klein. Comparison of simple potential functions for simulating liquid water. *The Journal of Chemical Physics*, 79(2):926–935, 07 1983.

[9] H. J. C. Berendsen, J. R. Grigera, and T. P. Straatsma. The missing term in effective pair potentials. *The Journal of Physical Chemistry*, 91(24):6269–6271, 1987.

[10] William L. Jorgensen. Revised TIPS for simulations of liquid water and aqueous solutions. *The Journal of Chemical Physics*, 77(8):4156–4163, 10 1982.

[11] J. L. F. Abascal and C. Vega. A general purpose model for the condensed phases of water: TIP4P/2005. *The Journal of Chemical Physics*, 123(23):234505, 12 2005.

[12] Hans W. Horn, William C. Swope, Jed W. Pitera, Jeffry D. Madura, Thomas J. Dick, Greg L. Hura, and Teresa Head-Gordon. Development of an improved four-site water model for biomolecular simulations: TIP4P-Ew. *The Journal of Chemical Physics*, 120(20):9665–9678, 05 2004.

[13] S. Dixit, J. Crain, W.C.K. Poon, J.L. Finney, and A.K. Soper. Molecular segregation observed in a concentrated alcohol-water solution. *Nature*, 416:829–832, 2002.

[14] Y. Marcus. Preferential solvation in mixed solvents x. completely miscible aqueous co-solvent binary mixtures at 298.15 k. *Monatshefte fuer Chemie*, 132:1387–1411, 2001.

[15] A. Perera and F. Sokolić. Modeling nonionic aqueous solutions: The acetone-water mixture. *The Journal of Chemical Physics*, 121:11272, 2004.

[16] A. Perera. From solutions to molecular emulsions. *Pure and Applied Chemistry*, 88:189, 2016.

[17] Raymond D. Mountain. Microstructure and hydrogen bonding in water - acetonitrile mixtures. *The Journal of Physical Chemistry B*, 114(49):16460–16464, 2010. PMID: 21090638.

[18] Rajappa Chitra and Paul E. Smith. A comparison of the properties of 2,2,2-trifluoroethanol and 2,2,2-trifluoroethanol/water mixtures using different force fields. *The Journal of Chemical Physics*, 115(12):5521–5530, 09 2001.

[19] S.K. Allison, J.P. Fox, R. Hargreaves, and S.P. Bates. Clustering and microimmiscibility in alcohol-water mixtures: Evidence from molecular-dynamics simulations. *Physical Review B*, 71:024201, 2005.

[20] B. Kežić and A. Perera. Revisiting aqueous-acetone mixtures through the concept of molecular emulsions. *The Journal of Chemical Physics*, 137:134502, 2012.

[21] M. Sedlák and D. Rak. On the origin of mesoscale structures in aqueous solutions of tertiary butyl alcohol: The mystery resolved. *The Journal of Physical Chemistry B*, 118(10):2726–2737, 2014. PMID: 24559045.

[22] A. Ben-Naim. Inversion of the kirkwood - buff theory of solutions: Application to the water-ethanol system. *The Journal of Chemical Physics*, 67(11):4884, 1977.

[23] Enrico Matteoli and Luciano Lepori. Solute - solute interactions in water. ii. an analysis through the kirkwood - buff integrals for 14 organic solutes. *The Journal of Chemical Physics*, 80(6):2856–2863, 03 1984.

[24] J.G. Kirkwood and F.P. Buff. The statistical mechanical theory of solutions. i. *The Journal of Chemical Physics*, 19(6):774, 1951.

[25] P. E. Smith. On the kirkwood - buff inversion procedure. *The Journal of Chemical Physics*, 129(12):124509, 2008.

[26] Peter I. Nagy, Gergely Völgyi, and Krisztina Takćcs-Novák. Monte carlo structure simulations for aqueous 1,4-dioxane solutions. *The Journal of Physical Chemistry B*, 112(7):2085–2094, 2008. PMID: 18220380.

[27] I. Bakulin, N. Kondratyuk, A. Lankin, and G. Norman. Properties of aqueous 1,4-dioxane solution via molecular dynamics. *The Journal of Chemical Physics*, 155(15):154501, 10 2021.

[28] J.L. Lebowitz and J.K. Percus. Long-range correlations in a closed system with applications to nonuniform fluids. *Physical Review*, 122(6):1675, 1961.

[29] Aurélien Perera, Martina Požar, and Bernarda Lovrinčević. Camel back shaped Kirkwood - Buff integrals. *The Journal of Chemical Physics*, 156(12):124503, 03 2022.

[30] Michael P. Allen and Dominic J. Tildesley. *Computer Simulation of Liquids*. Oxford University Press, 06 2017.

[31] Peter Krüger, Sondre K. Schnell, Dick Bedeaux, Signe Kjelstrup, Thijs J. H. Vlugt, and Jean-Marc Simon. Kirkwood - buff integrals for finite volumes. *The Journal of Physical Chemistry Letters*, 4(2):235–238, 2013. PMID: 26283427.

[32] Pritam Ganguly and Nico F. A. van der Vegt. Convergence of sampling kirkwood–buff integrals of aqueous solutions with molecular dynamics simulations. *Journal of Chemical Theory and Computation*, 9(3):1347–1355, 2013. PMID: 26587597.

[33] Jasmin Milzetti, Divya Nayar, and Nico F. A. van der Vegt. Convergence of kirkwood - buff integrals of ideal and nonideal aqueous solutions using molecular dynamics simulations. *The Journal of Physical Chemistry B*, 122(21):5515–5526, 2018. PMID: 29342355.

[34] R. Cortes-Huerto, K. Kremer, and R. Potestio. Communication: Kirkwood - Buff integrals in the thermodynamic limit from small-sized molecular dynamics simulations. *The Journal of Chemical Physics*, 145(14):141103, 10 2016.

[35] A. Perera, F. Sokolić, L. Almásy, and Y. Koga. Kirkwood - buff integrals of aqueous alcohol binary mixtures. *The Journal of Chemical Physics*, 124(12):124515, 2006.

[36] Martina Požar, Jennifer Bolle, Christian Sternemann, and Aurélien Perera. On the x-ray scattering pre-peak of linear mono-ols and the related microstructure from computer simulations. *The Journal of Physical Chemistry B*, 124(38):8358–8371, 2020.

[37] J.-P. Hansen and I.R. McDonald. *Theory of Simple Liquids*. Academic Press, Elsevier, Amsterdam, 3rd edition, 2006.

[38] Bernarda Kežić and Aurélien Perera. Towards a more accurate reference interaction site model integral equation theory for molecular liquids. *The Journal of Chemical Physics*, 135(23):234104, 12 2011.

[39] M.E. Lee and N.F.A. van der Vegt. A new force field for atomistic simulations of aqueous tertiary butanol solutions. *The Journal of Chemical Physics*, 122(11):114509, 2005.

[40] Samuel J. Keasler, Sophia M. Charan, Collin D. Wick, Ioannis G. Economou, and J. Ilja Siepmann. Transferable potentials for phase equilibria united atom description of five- and six-membered cyclic alkanes and ethers. *The Journal of Physical Chemistry B*, 116(36):11234–11246, 2012. PMID: 22900670.

[41] S. Pronk, S. Páll, R. Schulz, P. Larsson, P. Bjelkmar, R. Apostolov, M.R. Shirts, J.C. Smith, P.M. Kasson, D. van der Spoel, B. Hess, and E. Lindahl. Gromacs 4.5: a high-throughput and highly parallel open source molecular simulation toolkit. *Bioinformatics*, 29:845–854, 2013.

[42] J.M. Martínez and L. Martínez. Packing optimization for automated generation of complex system's initial configurations for molecular dynamics and docking. *Journal of Computational Chemistry*, 24:819, 2003.

[43] Aurélien Perera and Redha Mazighi. On the nature of the molecular ordering of water in aqueous DMSO mixtures. *The Journal of Chemical Physics*, 143(15):154502, 10 2015.

[44] A. Perera. Charge ordering and scattering pre-peaks in ionic liquids and alcohols. *Physical Chemistry Chemical Physics*, 19:1062, 2017.

[45] A. Perera. Molecular emulsions: from charge order to domain order. *Phys. Chem. Chem. Phys.*, 19:28275–28285, 2017.

[46] Aurélien Perera. On the microscopic structure of liquid water. *Molecular Physics*, 109(20):2433–2441, 2011.

[47] Toshiyuki Takamuku, Atsushi Yamaguchi, Masaaki Tabata, Nobuyuki Nishi, Koji Yoshida, Hisanobu Wakita, and Toshio Yamaguchi. Structure and dynamics of 1,4-dioxane-water binary solutions studied by x-ray diffraction, mass spectrometry, and nmr relaxation. *Journal of Molecular Liquids*, 83(1):163–177, 1999. Progress in the Physical Chemistry of Ionic Liquids A collection of invited papers in honour of Professor M. Chemla.

[48] Philip Ball and John E. Hallsworth. Water structure and chaotropicity: their uses, abuses and biological implications. *Phys. Chem. Chem. Phys.*, 17:8297–8305, 2015.

[49] T. Tokushima, Y. Harada, O. Takahashi, Y. Senba, H. Ohashi, L.G.M. Pettersson, A. Nilsson, and S. Shin. High resolution x-ray emission spectroscopy of liquid water: The observation of two structural motifs. *Chemical Physics Letters*, 460(4):387–400, 2008.

[50] Katrin Amann-Winkel, Marie-Claire Bellissent-Funel, Livia E. Bove, Thomas Loerting, Anders Nilsson, Alessandro Paciaroni, Daniel Schlesinger, and Lawrie Skinner. X-ray and neutron scattering of water. *Chemical Reviews*, 116(13):7570–7589, 2016. PMID: 27195477.

[51] Samantha Weerasinghe and Paul E. Smith. Kirkwood - Buff derived force field for mixtures of acetone and water. *The Journal of Chemical Physics*, 118(23):10663–10670, 05 2003.

[52] Redha Zoranić, Larisa Mazighi, Franjo Sokolić, and Aurélien Perera. On the micro-heterogeneity in neat and aqueous amides: A molecular dynamics study. *The Journal of Physical Chemistry C*, 111(43):15586–15595, 2007.

[53] Redha Zoranić, Larisa Mazighi, Franjo Sokolić, and Aurélien Perera. Concentration fluctuations and microheterogeneity in aqueous amide mixtures. *The Journal of Chemical Physics*, 130(12):124315, 03 2009.

[54] M. Požar and A. Perera. On the micro-heterogeneous structure of neat and aqueous propylamine mixtures: A computer simulation study. *Journal of Molecular Liquids*, 227:210, 2017.

[55] Alenka Luzar and David Chandler. Structure and hydrogen bond dynamics of water - dimethyl sulfoxide mixtures by computer simulations. *The Journal of Chemical Physics*, 98(10):8160–8173, 05 1993.

[56] A. B. Bhatia and D. E. Thornton. Structural aspects of the electrical resistivity of binary alloys. *Phys. Rev. B*, 2:3004–3012, Oct 1970.

[57] Keiko Nishikawa. Simple relationship between the kirkwood - buff parameters and the fluctuations in the particle number and concentration obtained by small-angle x-ray scattering: Application to tert-butyl alcohol and water mixtures. *Chemical Physics Letters*, 132(1):50–54, 1986.

[58] M. Teubner and R. Strey. Origin of the scattering peak in microemulsions. *The Journal of Chemical Physics*, 87(5):3195, 1987.

[59] A Ciach and W T Góźdź. Self-consistent theory for systems with mesoscopic fluctuations. *Journal of Physics: Condensed Matter*, 28(41):414010, aug 2016.

[60] G. R. Andersen and J. C. Wheeler. Theory of lower critical solution points in aqueous mixtures. *The Journal of Chemical Physics*, 69(7):3403–3413, 08 2008.

[61] S. D. Overduin, Aurélien Perera, and G. N. Patey. Structural behavior of aqueous t-butanol solutions from large-scale molecular dynamics simulations. *The Journal of Chemical Physics*, 150(18), May 2019. 184504.

[62] Masanari Nagasaka, Hayato Yuzawa, and Nobuhiro Kosugi. Microheterogeneity in aqueous acetonitrile solution probed by soft x-ray absorption spectroscopy. *The Journal of Physical Chemistry B*, 124(7):1259–1265, 2020. PMID: 31990199.

[63] Andres S. Urbina, Lyudmila V. Slipchenko, and Dor Ben-Amotz. Quantifying the nearly random microheterogeneity of aqueous tert-butyl alcohol solutions using vibrational spectroscopy. *The Journal of Physical Chemistry Letters*, 14(50):11376–11383, 2023. PMID: 38078837.