Micro-heterogeneity versus clustering in binary mixtures of ethanol with water or alkanes

Martina Požar, Bernarda Lovrinčević, Larisa Zoranić, Tomislav Primorać, Franjo Sokolić, Aurélien Perera

Micro-heterogeneity versus clustering in binary mixtures of ethanol with water or alkanes

Martina Požar[1,2], Bernarda Lovrinčević[2], Larisa Zoranić[2],Tomislav Primorać[2] , Franjo Sokolić[2] and Aurélien Perera[1*]

[1]Laboratoire de Physique Théorique de la Matière Condensée (UMR CNRS 7600), Université Pierre et Marie Curie, 4 Place Jussieu, F75252, Paris cedex 05, France.

[2]Department of Physics, Faculty of Sciences, University of Split, Ruđera Boškovića 37, 21000, Split, Croatia.

Abstract

Ethanol is an hydrogen bonding liquid. When mixed in small concentrations with water or alkanes, it forms aggregate structures reminiscent of, respectively, the direct and inverse micellar aggregates found in emulsions, albeit at much smaller sizes. At higher concentrations, micro-heterogeneous mixing with segregated domains is found. We examine how different statistical methods, namely correlation functions, structure factor and cluster distribution analysis, can describe efficiently these morphological changes in these mixtures. In particular, we explain how the neat alcohol pre-peak of the structure factor evolves into the domain pre-peak under mixing conditions, and how this evolution differs whether the co-solvent is water or alkane. This study clearly establishes the heuristic superiority of the correlation function / structure factor analysis to study micro-heterogeneity, since the cluster distribution analysis is insensitive to domain segregation. Correlation functions detect the domains, with a clear structure factor pre-peak signature, while the cluster techniques detect the cluster hierarchy within domains. The main conclusion is that, in micro-segregated mixtures, the

*corresponding author (aup@lptmc.jussieu.fr)

domain structure is a more fundamental statistical entity than the underlying cluster structures. These findings could help a better comparative understanding of radiation scattering experiments, which are sensitive to domains, versus spectroscopy-NMR experiments, which are sensitive to clusters.

Motivation

This work illustrates, perhaps for the first time, the profound difference between clustering and micro-segregation in complex liquids, despite the fact that clustering is at the origin of micro-segregation. This difference allows us to attribute a heuristic importance to some physical observable (structure factor) with respect to others (cluster distribution). Although our work is theoretical, our findings should impact upon the experimental ones, related to each of these observable -radiation scattering versus spectroscopy and NMR, as well as favouring a deeper understanding of molecular association in soft-matter.

1 Introduction

In the statistical analysis of computer simulations, it is important to distinguish between various types of observable[1]. Some observable, such as correlation functions, have a deep meaning from statistical physics, since they can be related to many physical properties of the system through various integrals involving them[3, 4]. However, correlation functions themselves are not physical observable, although some of them can be extracted through scattering experiment[5]. Other observable can be introduced, which provide useful insight about the microscopic state of the system. Hydrogen bond and clusters counting [6], are such examples, which can be recouped by many experimental techniques, such as various spectroscopy techniques[7]. Such observable often appear to be more useful than those related to statistical physics, since they provide finer details on the microscopic structure of the system. An important methodological question is whether or not the introduction of such convenient observable can be compared to the fundamental ones. We illustrate here a case where this question can be answered precisely.

In computer simulations of neat ethanol, the hydroxyl groups are found to form H-bonded chain-like structures[8], which span the entire system. When mixed with alkanes, such as hexane for example, or benzene, these hydrogen bonded structures persist, since they are energetically favourable[9], and induce subsequent local segregation of ethanol from the alkane molecules, at any mixing ratios[10]. Consequently, in alkane environments, ethanol clusters are rather well characterised structures. In contrast, when mixed with water, ethanol can also hydrogen bond with water molecules. Since in all classical force fields representation of the interactions, the value of the partial charges of water are larger than those of the ethanol hydroxyl group, water molecules are found to generally prefer to hydrogen bond with themselves, rather than with ethanol[11, 12, 13]. This competition tends to destroy the chain-like structure of the ethanol clusters, making these rather fuzzy aggregated structures. In ethanol-alkane mixtures, the hydroxyl groups of ethanol are hidden inside the ethanol clusters[10], while in water, these groups are rather dispersed. Following these facts, one could compare the micro-segregation of ethanol in benzene with that of surfactant-in-oil type emulsions, while ethanol in water would have analogies with surfactant-in-water type emulsions.

How does these visually appealing findings translate into observable in physical-chemistry? With the help of computer simulations, we compute structural statistical quantities, such as the pair distribution functions between different atoms, as well as cluster distributions. These calculations help to provide more clear answer to the question of the nature of the cluster structure of an hydrogen bonding molecule in various types of solvent. Indeed, the structure of ethanol in alkanes is more on the cluster side of the description, while in water, these look more like concentration fluctuations. Since the words "cluster" and "concentration fluctuation" mean very different measurable physical characteristics, our analysis should help to understand the relationship between microscopic molecular association and different macroscopic observable which are related to the local distribution of the molecules.

The principal argument of this paper is to show that micro-segregated domains are not reducible to the clusters of which they are made. We reach this conclusion by analyzing the differences between structure factor pre-peaks and direct cluster analysis. The latter analysis cannot tell the difference between concentration fluctuations and segregated domain structures that show up as a pre-peak in atom-atom structure factors. As shown in our previous works[10, 14, 15] , this difference is essential to understand clustering in complex mixtures. In particular, concentration fluctuations are thermodynamic observable through the Kirkwood-Buff integrals[10, 16, 17, 18], while the pre-peak in the atom-atom structure factor is a proof of micro-segregated domains [10, 14, 15]. Interestingly, both type of analysis are more in agreement between them when segregated domains are single clusters, such as the case of low concentration ethanol in alkanes. There differences in methodology have a heuristic significance, which we discuss in the last section of this paper.

The remaining of this paper is organized as follows. In the next section, we recall important theoretical details, describe our simulation protocol and give details about our cluster analysis methodology. We display our findings in the Results

Section shows. Finally, we discuss these findings and present our conclusion in the last part.

2 Theoretical and computational details

Thermodynamic quantities such as the energy or the density do not really reflect the micro-structure of liquids. For this reason, it is preferable to compute the atom-atom radial distribution function. Such functions are defined as fluctuations of the microscopic density of atoms of type a: $\rho_a(\vec{r}) = \sum_i \delta(\vec{r} - \vec{r}_i)$, where $\vec{r}_i$ is the position of any atom of type a. Considering this microscopic quantity as a random variable, one can compute usual statistical quantities, such as the mean density of atom a, $\rho_a =< \rho_a(\vec{r}) >$where the average is taken over a suitable statistical ensemble. The second moment $\rho_{ab}^{(2)}(|\vec{r} - \vec{r}'|) < \rho_a(\vec{r})\rho_b(\vec{r}') >$ is related to the correlation function through $\rho_{ab}^{(2)}(r) = \rho_a \rho_b g_{ab}(r)$. The Kirkwood-Buff theory relates the integrals of the g_{ab} functions -the so-called Kirkwood-Buff integrals (KBI)- to the composition fluctuations $< N_a N_b > - < N_a >< N_b >$, where N_a and $N_b$are the number of atoms of species a and b, respectively. It turns out that these integrals are the simply zero wave vector $k = 0$ values of the corresponding structure factors $S_{ab}(k)$, which are the Fourier transforms of the $g_{ab}(r)$. In a case of multicomponent molecular systems made of molecules instead of atoms, the integrals of the atom-atom $g_{ab}(r)$ are also the KBI of this system, because of the invariance of these integrals with respect to any arbitrary center of mass of the molecules[3]. This is summarised in the following expression

$$S_{ab}(k=0) = \delta_{ab} + \sqrt{\rho_a \rho_b} G_{ab} = \frac{< N_a N_b > - < N_a >< N_b >}{\sqrt{< N_a >< N_b >}} = \varepsilon_{ab}$$

where

$$G_{ab} = \int d\vec{r}\, [g_{ab}(r) - 1]$$

are the KBI, the integrals over the pair distribution functions $g_{ab}(r)$ between atomic sites a and b belong to two related molecular species,

$$S_{ab}(k) = \int d\vec{r} \exp(i\vec{k}.\vec{r})\, [g_{ab}(r) - 1]$$

is the Fourier transform of the correlation functions, and the last term ε_{ab} is related to the thermodynamics

$$\varepsilon_{ab} = \frac{1}{\rho \sqrt{x_a x_b}} \left(\frac{\partial \rho_a}{\partial \beta \mu_b} \right)_{TV\mu_k}$$

through the partial derivatives involving the number density ρ_a of species a, the mole fraction x_a and the chemical potential μ_b ($\beta = 1/k_K T$ is the Boltzmann factor). This expression, however, does not give any indication about the nature of the clustering and domain segregation in the mixtures.

The relationship between concentration fluctuations and micro-hetrogeneous clustering is not very clear, and this remains an important currently unsolved problem in the statistical description of liquids. Since the local segregation of one species with respect to the others indicates an heterogeneity in spatial distribution, it can be mistaken for a concentration fluctuation. Conversely, concentration fluctuations which occur during a critical demixing are clearly not an arrested clustering of either species. Such phase separation process are well understood both from theoretical and computational points of view[20], and their approach is signalled by the growth and divergence of all the partial structure factors exactly at $k = 0$, according to the corresponding diverging growth of concentration fluctuations. Such growth can be unambiguously detected in computer simulations[20, 21]. However, this scenario is not what occurs in micro-segregation. In a series of papers[10, 15, 14, 19] , we have argued that the micro-heterogeneous clustering is a non-zero wave vector fluctuation of the microscopic density, which arises at a specific k-vector which corresponds to mean size of the heterogeneity and should be manifested as a pre-peak in specific atom-atom structure factors $S_{ab}(k)$. This assumption was confirmed by computer simulation on a variety of systems we have studied, both aqueous and non-aqueous. However, the prediction of details of such micro-heterogeneity pre-peak from microscopic details of interactions between various types of molecules remains an open field of investigation[15].

2.1 Simulation details

All the calculations have been conducted using the Gromacs 4.5.5 package[22]. We have used N=16000 particles in order to have a good description of the domain structure. In our previous analysis[11, 12, 13], we used mostly N=2048 particles, which was sufficient to obtain many thermodynamic properties, but clearly insufficient to determine long range domain oscillations in the correlation functions. The initial configurations were all started using the very convenient PackMol code[23].The run lengths for statistics are of few nano-seconds, between 2 and 10, depending of systems, with a time step of 2fs in all cases. We use ambient conditions of T=300K and 1bar atmospheric pressure. Nose-hoover thermostat and Stillinger-Rahman barostat are used, with time constant of 0.1ps. We used the SPC/E [24] force field for water, TraPPe force fields for ethanol [25], and OPLS force field for alkanes [26].

2.2 Cluster analysis details

Cluster analysis depends crucially on the criteria defining how two particles are connected neighbours. Since in a dense liquid, two neighbouring particles can be very close, any criteria describing such situation can be a robust descriptor. This way, one can describe clustering in a simple Lennard-Jones liquid [27, 28]. However, there is a strong difference between such a simple liquid and an associated liquid, as in the case of hydrogen bonding system, where clustering has an element of reality. We have previously studied clustering in pure alcohol and water [29, 30]. We found that the cluster size distribution in neat alcohol show a specific peak at some particular cluster size (broadly around 5-7 particles),whereas water has a cluster size distribution much like a Lennard-Jones system [31, 32], with the maximum occuring for the monomer[33]. In addition, we found that the specific clusters of the alcohols had a precise shape (chain and loop for methanol, globular clusters for tbutanol)[29, 30]. In contrast, water has no such characteristic clusters. Here we compute the same property, but in mixtures.

The cluster is defined as the group of particles where each particle has at least one connection with the neighbor particles. The connectivity criteria can be geometrical constraints, or for example the Hills energetic criteria where particles are consider to be connected if their attractive interaction energy is higher then their relative kinetic energy[28]. Here we used Stilinger distance criteria [34] where the cutoff distance is defined by the first minima of the particle-particle radial distribution function. This way, the interactions between bonded particles are indirectly related to their interactions through the radial distribution function. The cluster size distributions are calculated for the clustering of the like-like sites, using several different statistical approaches. We show the results for the cluster size probability functions:

$$s_n = \frac{\sum_{k=1}^{N_c} s(n,k)}{\sum_{k=1}^{N_c} \sum_{j=1}^{N_{mol}} s(j,k)}$$

where s_n is the probability for the cluster formed of n sites, $s(k,n)$ represents the number of clusters of the size n in the configuration k. Varying the contact distance between neighbouring atoms that are part of a cluste distance around the first minima, shows a relative robustness in the resulting cluster distributions[29, 30]. The cutoff distances defined in this work are $r_c = 3.5$Å between the oxygens of water molecules, $r_c = 3.7$Å between the oxygens of the ethanol molecules, $r_c = 4.5$Å between the methyl groups of the ethanol molecules, and $r_c = 6$Å between the carbon atoms of the benzene molecules.

3 Results

Ethanol-water mixtures were previously studied by computer simulations in our group[11, 12, 13]. There is a major difference in clustering between ethanol and water. Neat ethanol contains specific clusters in the form of chains and loops, much like methanol[29, 30]. In contrast, neat water does not produce any specific clusters[29, 30]. The principal reason seems to be the distribution of partial charges in each molecule. The ethanol has only one hydroxyl group. Therefore, the hydrogen bonds can form chaining patterns ..OH-OH-OH...Despite thermal agitation, small chains can be relatively stable, and conserved through the sample. This is what we observe in simulation of many linear alcohols. In contrast, in water, there are two hydroxyl groups disposed in tetrahedral conformation, that allows branched OH chaining, which is more fragile to thermal agitation because on the increased topological constraints to maintain such a network over large distances. As a result, no robust clustering is observed, despite permanent tetrahedral H-bonding. Recent spectroscopic studies[35] suggest that linear OH clusters exist, but, in our opinion, these clusters are fragilized by permanent competition with potential trimer of quadrumer branching. These intuitive arguments find some support in our recent study of the aqueous-DMSO mixtures[36], where we found that water forms linear clusters in presence of DMSO, and at all concentrations. In contrast, only bulky cluster of water are found in alcohols[29, 30] and solvents such as acetone[31, 32].

3.1 Snapshot analysis

Snapshots represent only one micro-state of the system, and it would be generally unadvised[1] to make any serious conclusions of the general behaviour of any system, based on such single micro-state. However, in the case of micro-heterogenerous mixtures, with at least one associating species, much can be learned from a single micro-state. In fact, this single micro-state is a very good representation of all possible micro-states, since they appear to be simple permutations of the segregation patterns. This is an interesting peculiarity of micro-heterogeneous systems, pertains to a local "symmetry" property, which deserves further scrutiny.

Fig.1 summarizes the findings that we want to report here, namely the morphology of the aqueous-ethanol (upper figures) and alkane-ethanol mixtures (lower figures), each for 3 concentration of ethanol, namely $x_{Eth} = 0.2$(left column), 0.5(middle column) and 0.8(right column). Let us first focus on the upper figures, concerning aqueous-ethanol mixtures. The left-most figure shows the loose domain structure of ethanol molecules in water (shown as semi-transparent dark blue molecules). The oxygen(red) and hydrogen(white) atoms of ethanol are put into evidence, as to better visualise the chain-like clusters. The methyl united atoms are shows as semi-transparent groups. We notice that there are many non-bonded ethanol hydroxyl groups. These groups are in fact bonded to the surrounding water molecule. As a result, despite segregation, the ethanol domains are rather fuzzy. The central figure shows the water molecules, with the ethanol molecules in semi-transparent representation, for $x_{Eth} = 0.5$ and the picture in the right shows a similar representation for $x_{Eth} = 0.8$. We can see that in both pictures, water is segregated in domains, which are also loose, although the hydrogen bonding between the hydroxyl groups is quite apparent. The general picture that emerges from these 3 snapshots, is that both water and ethanol form fuzzy micro-segregated domains. The fuzziness comes from the incomplete self-hydrogen bonding of each species with it own kind. From this observation, we expect that the cluster distributions of will not show any peak at some particular cluster size.

In the lower set of figures we have shown comparative clustering in 3 different alkanes. The lower left figure shows 20% ethanol in hexane, with the hexane molecules shown in semi-transparent, and the ethanol molecules shown with the same convention as in the figure just above. It is seen that the ethanol molecules are segregated from hexane. In addition, we can clearly see the hydroxyl groups within each domain, are bound in chains and loops. In fact, almost all hydroxyl groups are bound into such shape, as will be confirmed below in the cluster analysis. The middle picture shows 50% ethanol in benzene, with a representation of the molecules analogous to the previous snapshot. Once again, we see clearly the segregation in species domains, as well as chain/loop clusters of the hydroxyl groups inside the ethanol domains. The lower right picture shows 80% ethanol in pentane, and this time the ethanol molecules are shown entirely. We again observe a domain segregation by species, and geometric clusters of the hydroxyl groups. In fact, ethanol in this latter system is clustered more or less like in pure ethanol, which is not surprising, and this will be confirmed by the cluster analysis in the next sub-section.

The study of the snapshots shows a profound difference in domain segregation between the aqueous and the alkane mixtures with ethanol, with fuzzy domains in the first and ethanol domain underlying precise geometrical hydroxyl cluster in the second. These differences obviously come from the fact that water offers hydrogen possibilities to the ethanol hydroxyl groups, contrary to alkanes.

3.2 Correlation function analysis

Fig.2 shows the correlation function between the oxygen sites of ethanol, for 3 different concentration of ethanol, while the inset shows the correlations between the oxygen sites of water. The pure liquid correlations are also shown in black. In all cases we observe the strong first peak, which witness the underlying hydrogen bonding between hydroxyl groups, which is at the heart of micro-segregation.

However, micro-segregation is seen in the long range correlation between segregated domain, and not in the short range correlations. In all cases, correlation between the oxygens of water is more important than those between the oxygen of ethanol. We equally observe a feature we have pointed out in other aqueous mixtures[13, 31, 37, 38]: water OO correlations tend to increase at contact with decreasing water concentrations, while solute (here ethanol) correlations at contact tend to decrease with decreasing solute concentrations. This can be seen clearly through the identical trends of the first peaks with same color codes, while they correspond to different concentrations in terms of the concerned species (except of course for pure components shown in black). This remarkable feature is not however specific to water, and is equally seen for any associating molecule mixed with a less associating one. It indicates that the less associating species bonds less and less with itself with the increase of concentration of the more associating species, while the more associating species bond more and more with itself when its concentration decreases.

Fig.3 show correlations between the oxygen sites of ethanol, but in alkanes. The main panel, which shows ethanol oxygen correlations confirms the feature discussed above. Since ethanol is now the associating species, the first peak increases with decreasing ethanol content. The various alkanes correlations - shown in the inset- show the opposite trend, although these correlations concern very different alkanes. These findings, common to Fig2 and Fig3 prove an universal feature in mixtures of associating liquids. We also observe in the main panel the very strong first peak, much stronger than anything in Fig.2. It indicates the stronger hydrogen bonding of the hydroxyl groups of ethanol when in an alkane environment. This is the inverse micelle effect that we have mentioned in the Introduction, supporting the energetically favoured association of the hydroxyl sites.

The contrast of short range association between hydroxyl groups of ethanol in water and ethanol in alkanes will be reconfirmed below in the cluster distribution study. However, the micro-segregated domains affect the medium and long range correlations, and this is better analysed by looking at the structure factors.

3.3 Structure factor analysis

Fig.4 shows the structure factors for the correlations shown in Fig.2, with the same color conventions. In addition, the oxygen-oxygen structure factors of the pure components are shown in black. As noted before[39], pure water has a main peak about $k \approx 2\text{Å}^{-1}$,corresponding to the water diameter $\sigma_W \approx 3\text{Å}$, and a shoulder-peak at $k \approx 3\text{Å}^{-1}$ corresponding to the hydrogen bonding distance $r_{HB} \approx 2\text{Å}$. Pure ethanol has only one main peak around $k \approx 2.8\text{Å}^{-1}$, which corresponds more to an hydrogen bonding distance $r \approx 2\text{Å}$, as well as a pre-peak around $k \approx 0.8\text{Å}^{-1}$, which corresponds to the chain and ring clusters[13], similar to those observed in the snapshots in the previous section 3.1. In other words, in contrast to water, ethanol is entirely structured by the hydrogen bonding, since both peaks are related to this interaction. So, these two hydrogen bonding associating liquids have a very different micro-structure, a fact that we recognized in earlier works[29, 30] to be equally shared by other alcohols such as methanol and tbutanol.

In mixing conditions, by monitoring the behaviour of these peaks, we can account for changes in the micro-structure, with respect to pure fluid state. The structure factors in Fig.4 show remarkable microscopic changes.

Let us focus first in the water structure in the inset. As ethanol is added, the main peak at $k \approx 2\text{Å}^{-1}$ changes little until ethanol mole fraction 80% (red) where is nearly disappears. The Hbond peak at $k \approx 3\text{Å}^{-1}$ diminishes more clearly. From these facts, we can conclude that water is less and less hydrogen bonded when ethanol concentration increases. Fig.4 shows another remarkable feature: an intense pre-peak growing at $k \approx 0.4 - 0.2\text{Å}^{-1}$, which corresponds to water domain sizes of $d \approx 12 - 30\text{Å}$. These numbers match roughly the domains seen in the uppers snapshots seen in Fig.2. These pre-peak witness the water-solute domain segregation under mixing. In order to see this clearly, it is necessary to use N=16000 particles instead of N=2048 as we did previously[11, 13]. A remarkable feature is that these pre-peak are maximal at lower ethanol concentrations (20% and 50%) - witnessing the large water segregated domains that we observe in the snapshots, but diminish as this concentration increases (80%) as the water domains become smaller. Gathering all the peak informations, we see that the small water segregated domains (at large ethanol concentrations) have less hydrogen bonded water molecules than in pure water. This picture confirms the fuzzy water cluster picture that we have found from the snapshot analysis.

Turning now to the ethanol structure factor in the main panel, we see that the Hbond peak at $k \approx 2.8\text{Å}^{-1}$ diminishes very strongly with water content increase, while the cluster pre-peak at $k \approx 0.8\text{Å}^{-1}$ diminishes and shifts to higher k-values. The overall picture is that of less hydrogen bonded ethanol molecules at lower concentrations, with an apparent diminution of cluster sizes. Much like water, ethanol also develops a domain pre-peak around $k \approx 0.1 - 0.2\text{Å}^{-1}$, which corresponds to the segregated domains complementary to those of water. These domains are seen to grow, as the population of the cluster peak diminishes. This implies that smaller Hbonded clusters populate the large ethanol segregated domains, suggesting a fuzzyness of these domains. But it is also an indication that there is more to ethanol domain segregation than just ethanol self hydrogen bonding. Indeed, since ethanol molecules are less Hbonded at low concentrations, and yet they are gathered into a growing pre-peak, it means that these ethanol molecules are grouped through their interaction with water, and not by their own self Hbonding. This is a direct manifestation of the so-called hydrophobic effect[40, 41], of which we see here an interesting microscopic insight, through the hydroxyl groups of the solute, while this is usually described in terms of the hydrophobic groups of the solutes [41].

Fig.5 shows the oxygen-oxygen structure factors for ethanol in alkanes, as well as the carbon-carbon structure factors of the alkanes in the inset. Again, the pure ethanol structure factor is shown in black. We note that the ethanol Hbond peak at $k \approx 2.8\text{Å}^{-1}$ is not affected by mixing with alkanes, contrarily to what happened with water. It is a direct indication of the robustness of ethanol Hbonded clusters -as opposed their fuzziness in water. Now, however, with the increase of the alkane concentration, we see a phenomenon different than in water. We see that it is the hydroxyl group cluster pre-peak, at $k \approx 0.8\text{Å}^{-1}$, that moves, with the addition of alkanes, into a domain pre-peak at smaller k-values $k \approx 0.1 - 0.15\text{Å}^{-1}$. This is a remarkable result, since it confirms the visual information that we gathered through the snapshots in Fig.1: the ethanol domains are essentially made of hydroxyl group clusters, larger than those found in pure ethanol. We note that the alkane structure factors in the inset have a main peak around $k \approx 1.4\text{Å}^{-1}$, which corresponds the the diameter on the carbon atoms in various force field models $\sigma_C \approx 4\text{Å}$. Despite large differences in the various alkane molecules, the structure factors look nearly the same around this value of k. We note that the increase of the domain segregation leads to an increase of these structure factors but only at $k = 0$. In other words, these liquids witness concentration fluctuations instead of segregated domains, unlike the ethanol molecules. This asymmetry of the solvent behaviour between water and alkanes under the same ethanol insertion is remarkable. It confirms the picture of simple and complex disorder which we previously introduced[10, 14].

3.4 Cluster distributions

We turn now towards the cluster distribution. Perhaps the most important challenge in this study is to see if it can confirm the micro-heterogeneous structure of mixtures involving associating molecules.

Fig.6 shows the cluster distribution of water oxygen atoms in aqueous-ethanol, for different concentrations of ethanol. Since these aqueous mixtures are micro-segregated, we expect to see this in the cluster distribution. We note that these curves present no specific peak -ie- the probability distribution of a cluster of smaller size is always greater than that of a larger size, which is a trivially expected behaviour for simply disordered liquids. Indeed, the first inset shows the cluster probability in a Lennard-Jones type mixture, which is strikingly similar to that of water in water-ethanol. This latter mixture is in fact a one-liquid carbon-tetrachloride (we have used the OPLS model [42]), which is artificicially treated as a mixture by simple labelling of molecules. We considered the central carbon atom for computer cluster distribution of this system. The second inset shows the probability distribution of clusters of the pentane carbon atom in ethanol-pentane mixtures, for different ethanol concentrations, which are again trivial clsuter distributions. All these curves in Fig.6 show an additional common property: for a given size, the cluster probability at lower concentrations are always larger than that at larger concentrations. This property is also a trivial effect of random mixing at different concentrations. From these curves in Fig.6, we learn that there are almost no differences in these various distributions, which is very counter-intuitive, particularly after having noticed the strong micro-segregation in aqueous-ethanol mixtures in previous sections.

Fig.7 shows a comparison of the cluster distribution of ethanol oxygen atoms in pentane (main panel) and water (inset), for different concentrations of ethanol, including pure ethanol(shown in black). We note that the pure ethanol cluster peak (around 6-7 oxygen atoms) in pentene, increases with the decreasing ethanol concentrations, which confirms the clustering trend observed through the pre-peak analysis of the structure factors in Fig.5. The inset, however, shows only the trivial clustering, as seen in the previous Fig.6, despite the micro-segregation present in aqueous-ethanol. From the difference in clustering of ethanol, that we have observed in the previous sub-sections, we see that the cluster distribution is only able to detect clusters that are not fuzzy. By extension, we could say that cluster analysis is more performant for surfactant in oil, rather than surfactant in water.

Fig.8 shows a comparison of the clustering of the methyl group in ethanol-pentane mixtures, and aqueous-ethanol (inset). Since these methyl groups are randomized in pure ethanol, we do not expect to see any specific-peak, which is indeed confirmed for pure ethanol. However, since there is strong clustering of ethanol at small concentrations in benzene, we expect to see a some signs of specific clustering, which is absent from these plots: they look very similar for ethanol in benzene and in water, despite obvious differences.

The obvious, and almost counter-intuitive conclusion of this sub-section is that direct cluster calculation is not generally able to detect micro-heterogeneous distribution of molecules. To be more precise, it detects all clusters, but there seems to be more to micro-segregation that just clustering. This is why only the correlation function analysis, and particularly the structure factor analysis, can account for micro-segregation properly. The fact that these latter observable have a sound theoretical and statistical basis, is certainly in favour of these methods, as opposed to cluster detection, which is empirical and cannot be related to any quantity in statistical physics of the disordered liquids.

4 Discussion and Conclusion

The principal idea behind simple and complex disorder in liquids is the fact that all liquids -being disordered systems, are characterised by the same order parameter, namely the number density[4], but the description of complexity requires a new type of order parameter. Indeed, the H-bond interaction is not a Landau type order parameter since it is related to a pair interaction. Landau type order parameters are, by definition[4], related to external fields and corresponding 1-body functions. In that sense, it is not possible to describe the local order produced by the H-bond induced clustering through a classical Landau-type order parameter description. On the other hand, it is clear that a proper statistical description of the local order produced by the H-bonding interaction is required, if one wishes to describe complexity emerged from the hydrophobic interaction, for example. One way around this problem is to consider that specific fluctuations related to the H-bonding can be conveniently averaged into the concentration fluctuations. This is the route taken by the KBI formalism[16, 17, 18], and also field theoretic variants[40, 41]. These routes can explain only the part that concentration fluctuations contribute to the complex local order produced by the H-bonding. In particular, such approaches ignore the presence of a non-zero pre-peak in the structure factor. As shown here and in our previous works[10, 14, 37], this pre-peak witnesses the specificity of the clustering over concentration fluctuations.

The present study reveals a non-intuitive finding since direct cluster analysis is not able to reveal micro-segregation. This is very surprising since micro-segregation can be interpreted as a form of clustering. The only possible explanation, is that cluster analysis can only detect the clusters within the domains, but cannot detect the domains themselves, when these are made of groups of disjoint clusters. This is the case of the fuzzy domains in ethanol-water, but not the case of ethanol in benzene, where the base of the domain is made of underlying ethanol OH group clusters. Both scenarios were confirmed through the analysis of snapshots and structure factors. This explanation shows that the cluster study of mixtures with fuzzy domain structure is deceitful since it predict distributions indistinguishable from that found in a Lennard-Jones mixture. Although this result is the correct, it does not give any information on the micro-segregation of these systems.

This difference in information about the morphology of complex mixtures, as given by structure factors and cluster distribution, has a direct impact in the corresponding experiments, which are radiation scattering methods -which detect domain pre-peaks, and NMR, infrared and mass spectrometry -which detect clusters. Our study shows that these two different sets of techniques, may not detect the same type of aggregation of molecules. This important point deserves further scrutiny.

The asymmetry of the prediction of the cluster structure in aqueous mixtures and alkane-alcohol mixtures can be connected to the direct and inverse micelle structure, when extrapolated to the binary emulsions, such as water-surfactant and oil-surfactant. Inverse micelles in a oil-surfactant system, consist of dense core of hydroxyl groups, which are bound by energetical restraints. In a way, such micelles are energetically simple to obtain, and do not require any intervention of the surrounding oily solvent. On the other hand, direct micelles do require the solvent (water) to cooperate in order to shy away the oily parts of the surfactant inside a micellar core. Such micelles require more coordination at molecular level than the formers. In view of this, it is not surprizing that ethanol clustering in alkanes gives a specific clustering in alkanes, as opposed to ethanol in water.

From an heuristic point of view, the fact that the correlation function formalism of liquid state theory has a sounder statistical and theoretical basis than the direct cluster distribution analysis, supports the findings of the present work. It confirms that meaningful studies of the micro-segregation in complex liquid mixtures should be investigated through statistical physics of liquids.

Acknowledgements

This work has been partially supported by the Croatian Science Foundation under the project 4514 "Multi-scale description of meso-scale domain formation and destruction". M. Požar thanks the French Embassy in Croatia for financial support through "bourse du Gouvernement Français".

References

[1] Allen M. P and Tildesley D. J. *Computer Simulation of Liquids* (Oxford, 1987)

[2] Binder K. and Heermann D. W, *Monte Carlo Simulation in Statistical Physics: An Introduction* (Springer, Berlin, Heidelberg, 1986)

[3] Hansen J. P and McDonald I. R, *Theory of Simple Liquids* (Academic, London, 1986)

[4] Chaikin P. M and Lubensky T. C, *Principles of condensed matter physics*, (Cambridge University Press 1995).

[5] Fisher H. E. , Barnes A. C. and Salmon P. S. (2006) Neutron and X-ray diffraction studies of liquids and glasses, Rep Prog Phys 69:233–299.

[6] Rappaport D. C. (1983) Hydrogen bonds in Water, Molecular Physics 50(5) :1151

[7] Guo J.-H. et al. (2003) Molecular Struture in Alcohol-Water Mixtures, Phys. Rev. Lett. 91(15): 157401-1

[8] Benmore, C. J. and Loh, Y. L. (2000). The structure of liquid ethanol: A neutron diffraction and molecular dynamics study. J. Chem. Phys., 112: 5877

[9] Murdoch K. M., Ferris T. D., Wright J. C. and Farrar T. C. (2002) Infrared Spectroscopy of ethanol clusters in ethanol-hexane binary solutions. J. Chem. Phys. 116: 5717

[10] Požar M. et al. (2015) Simple and complex disorder in binary mixtures with benzene as common solvent, Phys. Chem. Chem. Phys. 17: 9885

[11] Mijaković M. et al. (2011) Ethanol-water mixtures: ultrasonics, Brillouin scattering and molecular dynamics, J. Mol. Liq. 164: 66

[12] Asenbaum A. et al.(2012) Structural changes ethanol-water mixtures: Ultrasonics, Brillouin scattering and molecular dynamics studies, Vibrational Spectroscopy 60: 102

[13] Mijaković M., Polok K. D., Kežić B., Sokolić F. , Perera A. and Zoranić L. (2014) A comparison of force fields for ethanol-water mixtures, Molecular Simulation, J. Mol. Sim 42: 699

[14] Perera A. (2016) From Solutions to Molecular Emulsions, Pure Appl. Chem. 88: 189

[15] Perera A. in "Fluctuation Theory of Solutions: Applications in Chemistry, Chemical Engineering, and Biophysics", Ed. P. E. Smith, J. P. O'Connell and E. Matteoli, CRC Press Taylor and Francis (2012) , ISBN 9781439899229

[16] Kirkwood J. G. and Buff F. (1950) The Statistical Mechanical Theory of Solutions, J. Chem. Phys. 19: 774

[17] Matteoli E. and Lepori J. (1984) Solute-solute interactions in water. II. An analysis through the Kirkwood-Buff integrals for 14 organic solutes, J. Chem. Phys. 80: 2856

[18] Ben-Naim A., (1977) Inversion of the Kirkwood-Buff theory of solutions: Application to the water-ethanol system, J. Chem. Phys. 67: 4884.

[19] Perera A., Kežić B., Sokolić F. and Zoranić L. in "Molecular Dynamics" (Vol 2), Ed. L. Wang (InTech, Rijeka, 2012)

[20] K. Binder, Kinetics of Phase separation in Stochastic Non linear systems, Vol 8, Springer Series in Synergetics, Ed. L. Arnold and E. Lefever 1980.

[21] J. Zauschn, P. Virnau, K. Binder, J. Horbach and R. L. Vink (2009), Statics and dynamics of colloid-polymer mixtures near their critical point of phase separation: A computer simulation study of a continuous Asakura–Oosawa model, J. Chem. Phys. 130: 064906

[22] van der Spoel. D., Lindahl E., Hess B., Groenhof G., A. E. Mark and Berendsen H. J. C. (2005) GROMACS: fast, flexible and free, J.Comp. Chem. 26: 1701

[23] Martínez L., Andrade R., Birgin E. G., Martínez J. M.(2009) Packmol: A package for building initial configurations for molecular dynamics, Journal of Computational Chemistry, 30(13):2157.

[24] Berendsen J. C., Postma J. P. M., Von Gusteren W. F. and Hermans J. , in *Intermolecular Forces*, edited by B. Pullman (Reidel, Dordrecht, 1981)

[25] Chen B., Potoff J.- J.and Siepmann J. I.(2001) Monte Carlo calculations for alcohols and their mixtures with alkanes. Transferable potentials for phase equilibria. 5. United-atom description of primary, secondary, and tertiary alcohols, J. Phys. Chem 105(15): 3093.

[26] Jorgensen W. L., Madura J. D. and Swenson C. J., J. Am. Chem. Soc. 106, 6638 (1984); *ibid*. Jorgensen W. L. , J. Phys. Chem. **90**, 1276 (1986)

[27] Jonsson H., Andersen H. C. (1988) Icosahedral Ordering in the Lennard-Jones Crystal and Glass, Phys. Rev. Lett. 60: 2295

[28] Pugnaloni L. A. and Vericat F.(2002) New criteria for cluster identifaction in continuum systems, J. Chem. Phys. 116: 1097.

[29] Perera A., Sokolic F. and Zoranic L. (2007) Microstructure of neat alcohols, Phys. Rev. E75:060502-(R)

[30] Zoranic L., Sokolic F. and Perera A. (2007) Microstructure of neat alcohols: A Molecular Dynamics study, J. Chem. Phys. 12**7**: 024502.

[31] Perera A. and Sokolic F. (2004) Modeling nonionic aqueous solutions: the acetone-water mixture, J. Chem. Phys. 121(22): 11272

[32] Kežić B. and Perera A. (2012) Revisiting aqueous-acetone mixtures through the concept of molecular emulsions , J. Chem. Phys. 137: 134502

[33] A. Geiger A., F. H. Stillinger F. H. , A. Rahman A. (1979) Aspects of the Percolation Process for Hydrogen-Bond Networks in Water, J. Chem.Phys. 70: 4185

[34] Stillinger F. H. (1963) Rigorous Basis of the Frenkel-Band Theory of Association Equilibrium, J. Chem. Phys. 38: 1486

[35] Wernet Ph. et al. (2004) The structure of the first coordination shell in liquid water, Science 104: 995

[36] Perera A. and Mazighi R. (2015) Simple and complex forms of disorder in ionic liquids, J. Chem. Phys. 143: 154502

[37] Kežić B. and Perera A. (2012) Aqueous tert-butanol mixtures: a molecular-emulsion, J. Chem. Phys. 137: 014501

[38] Perera A. and Kežić B. (2013) Fluctuation and micro-heterogeneity in mixtures of complex liquids, Faraday Discuss. 167: 145

[39] Perera A. (2011) On the microscopic structure of liquid water, Mol. Phys. 109: 2433

[40] Lum K. Chandler D. and Weeks J. D. (1999) Hydrophobicity at Small and Large Length Scales, J. Phys. Chem B 103:4570

[41] Chandler D. (2005) Interfaces and the driving force of hydrophobic assembly , Nature (London) 437: 640

[42] Duffy A. M., Severance D. L. and Jorgensen W. L. (1992) Solvent Effects on the Barrier to Isomerization for a Tertiary Amide from ab Initio and Monte Carlo Calculations , J. Am. Chem. Soc. 114:7535

Figure captions

- Fig.1 - Selected snapshots of aqueous-ethanol (top figures) and alkane-ethanol (lower figures) mixtures. Figures on right correspond to 20% ethanol, in the middle for 50% ethanol and on the right for 80% ethanol. See the text for details on color conventions for different molecules.

- Fig.2 - Oxygen-oxygen correlation function in ethanol-water mixtures. Main panel for ethanol, inset for water. Blue curves for 20% ethanol, green for 50% ethanol and red for 80% ethanol. The pure component is shown in black. These color conventions are preserved in all subsequent figures.

- Fig.3 - Site-site correlations in ethanol-alkane mixtures. Main panel: ethanol oxygen-oxygen correlation function. Inset: methyl-methyl correlation (blue for pentane, green for benzene and red for hexane). Color convention according to ethanol mole fraction as in Fig.2.

- Fig.4 -Structure factors for the correlation functions shown in Fig.2, with same conventions. Structure factor of neat liquids shown in black.

- Fig.5 - Structure factors for the correlation functions shown in Fig.3, with same conventions.

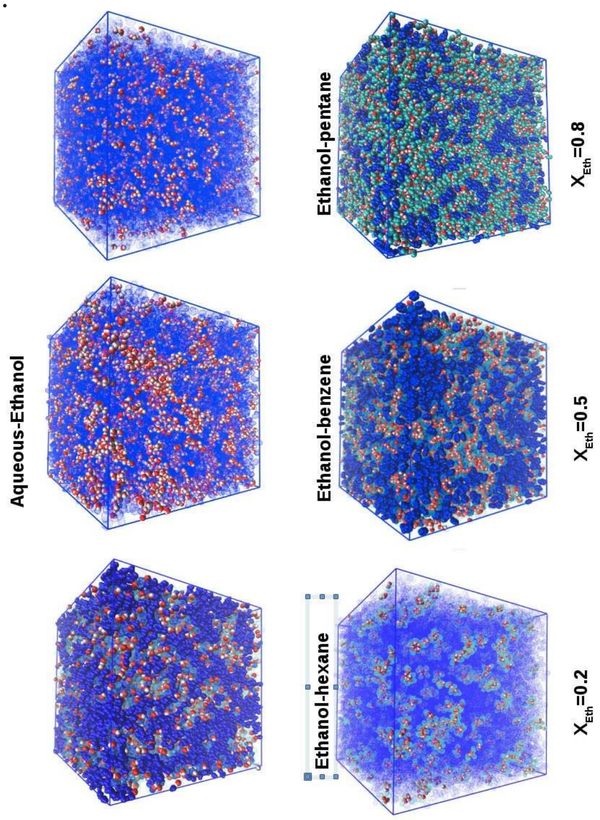
- Fig.6 - Cluster distribution functions. Main panel, for water oxygen atoms in aqueous ethanol mixtures (color conventions according to ethanol mole fraction as in Fig.2). Top inset: cluster distributions in a binary Lennard-Jones type mixture (see text). Lower inset: cluster distribution for the pentane central carbon atom in ethanol-pentane mixtures.

- Fig.7 - Cluster distribution functions. Main panel, for ethanol oxygen atoms in ethanol-pentane mixtures. Inset, for ethanol oxygens in aqueous mixtures (color conventions according to ethanol mole fraction as in Fig.2)

- Fig.8 - clusters distribution functions for the ethanol methyl group. Main panel, for ethanol-pentane mixtures. Inset, for aqueous-ethanol mixtures (color conventions according to ethanol mole fraction as in Fig.2).

Fig.1 - Selected snapshots of aqueous-ethanol (top figures) and alkane-ethanol (lower figures) mixtures. Figures on right correspond to 20% ethanol, in the middle for 50% ethanol and on the right for 80% ethanol. See the text for details on color conventions for different molecules.

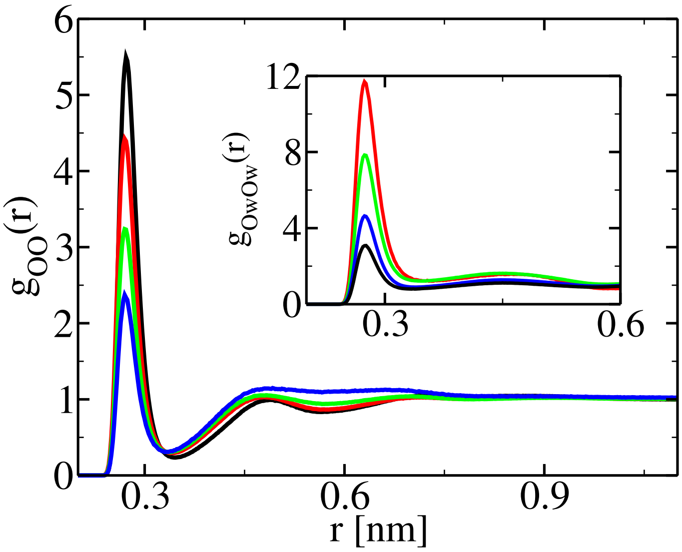
.



.

Fig.2 - Oxygen-oxygen correlation function in ethanol-water mixtures. Main panel for ethanol, inset for water. Blue curves for 20% ethanol, green for 50% ethanol and red for 80% ethanol. The pure component is shown in black. These color conventions are preserved in all subsequent figures.

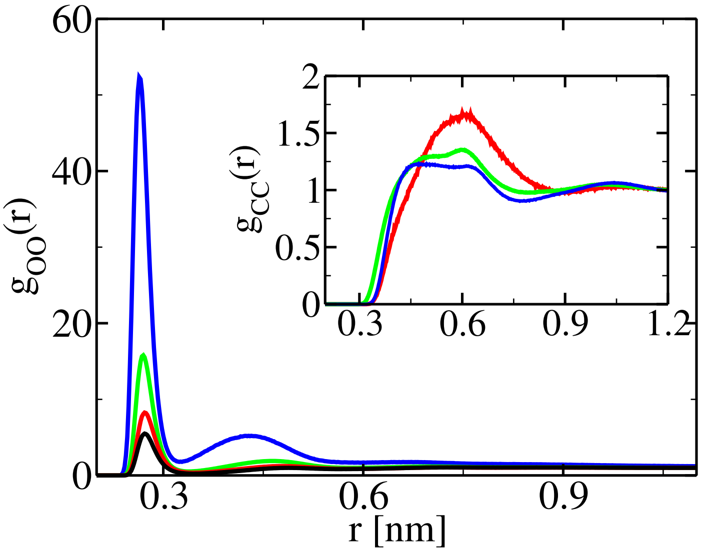
.



.

Fig.3 - Site-site correlations in ethanol-alkane mixtures. Main panel: ethanol oxygen-oxygen correlation function. Inset: methyl-methyl correlation (blue for pentane, green for benzene and red for hexane). Color convention according to ethanol mole fraction as in Fig.2.

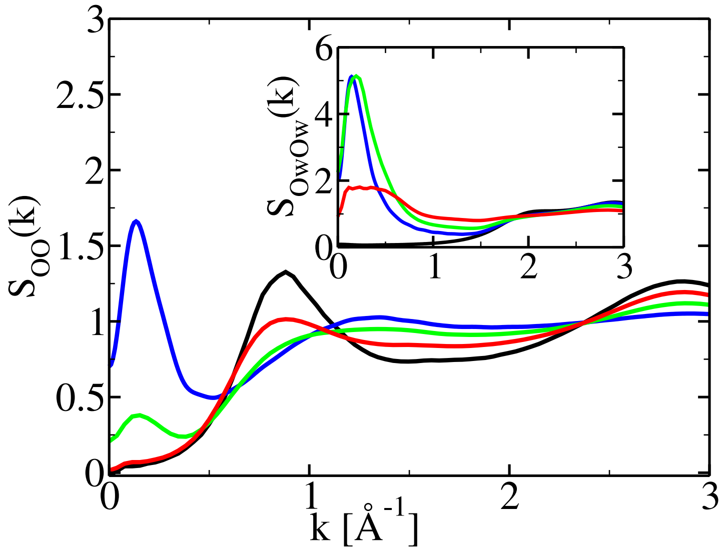
.



.

Fig.4 - -Structure factors for the correlation functions shown in Fig.2, with same conventions. Structure factor of neat liquids shown in black.

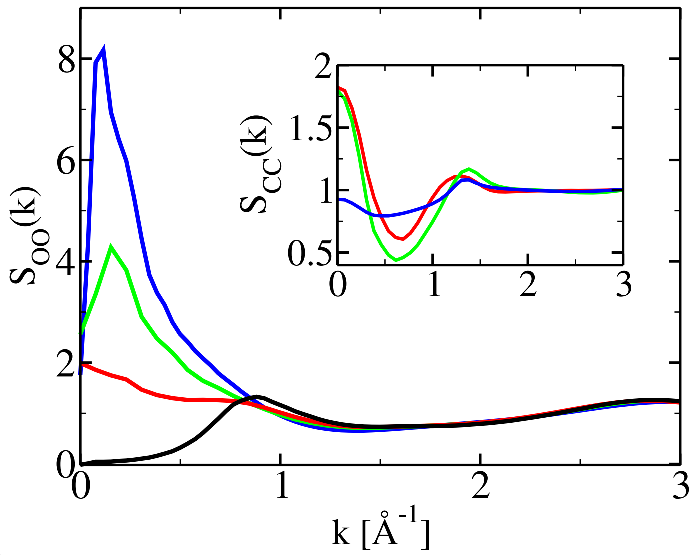
.



.

Fig.5 - Structure factors for the correlation functions shown in Fig.3, with same conventions.

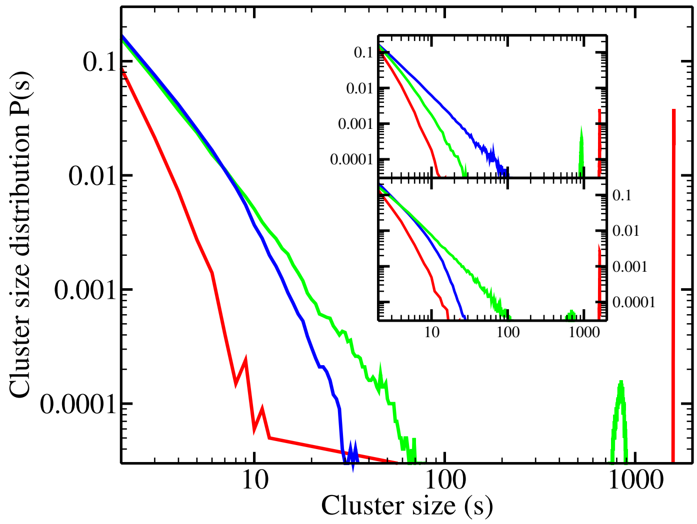
.



.

Fig.6 - Cluster distribution functions. Main panel, for water oxygen atoms in aqueous ethanol mixtures (color conventions according to ethanol mole fraction as in Fig.2). Top inset: cluster distributions in a binary Lennard-Jones type mixture (see text). Lower inset: cluster distribution for the pentane central carbon atom in ethanol-pentane mixtures.

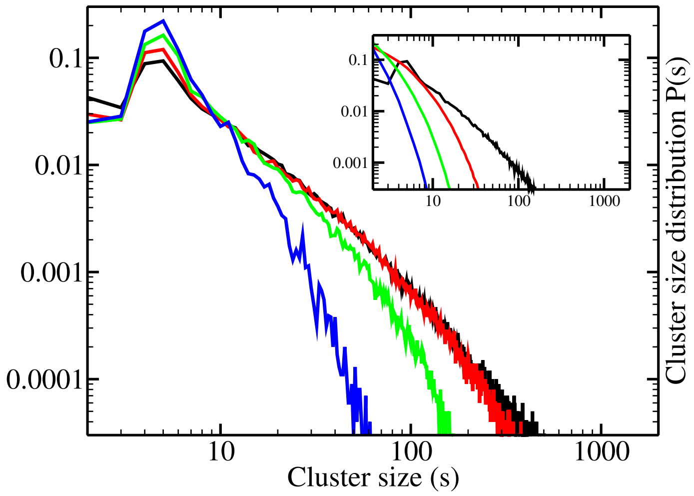
.



.

Fig.7 - Cluster distribution functions. Main panel, for ethanol oxygen atoms in ethanol-pentane mixtures. Inset, for ethanol oxygens in aqueous mixtures (color conventions according to ethanol mole fraction as in Fig.2)

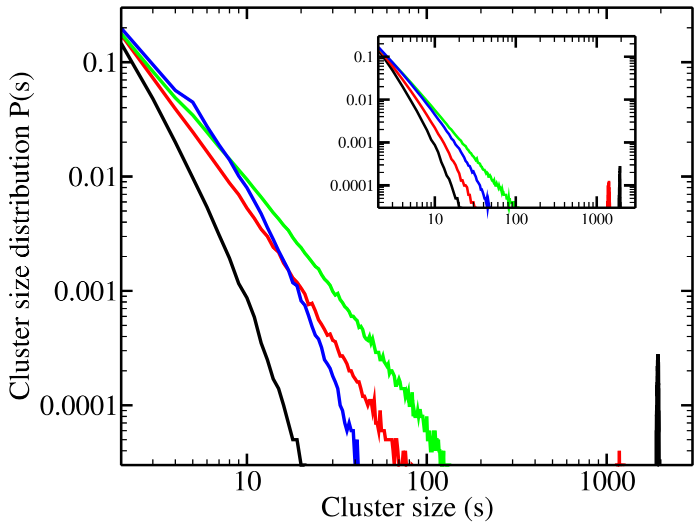
.



.

Fig.8 - clusters distribution functions for the ethanol methyl group. Main panel, for ethanol-pentane mixtures. Inset, for aqueous-ethanol mixtures (color conventions according to ethanol mole fraction as in Fig.2).

.



.