Microheterogeneity in Ethanol-Water Mixture

Mijaković, Marijana

University of Split
Faculty of Science
PhD Study of Biophysics

Marijana Mijaković

Microheterogeneity in Ethanol-Water Mixture

Doctoral thesis

Split, 2012

TEMELJNA DOKUMENTACIJSKA KARTICA

Sveučilište u Splitu
Prirodoslovno - matematički fakultet

Doktorski rad

MIKROHETEROGENOST U MJEŠAVINI ETANOL - VODA

Marijana Mijaković

Rad je izrađen na: Prirodoslovno - matematičkom fakultetu u Splitu

Sažetak

Metodom molekularne dinamike analizirana je mikroskopska struktura tekuće mješavina etanola i vode pri sobnim uvjetima, u cijelom rasponu koncentracija. Definirana su tri strukturna režima miješanja od kojih svaki predstavlja različitu realizaciju koja ima svojstvo mikroheterogenosti, odnosno lokalnog nemiješanja komponenti. Pri malim koncentracijama etanola zbog hidrofobnog efekta vode distribucija hidrofobnih dijelova molekula etanola nije homogena kao u čistom etanolu. Dodavanjem etanola, na otprilike 0.15 molarnog udjela, etanol i voda počinju graditi bi-kontinuiranu isprepletenu mrežu u cijelom sustavu. Drugi prijelaz je definiran na otprilike 0.65 molarnog udjela etanola gdje voda više ne može formirati kontinuiranu strukturu mreže vodikovih veza preko cijelog sustava. U prvom redu ovi režimi su potvrđeni analizom radijalne distribucijske funkcije koja se odnosi na klastere, što predstavlja novi pristup u analizi strukture molekularnih otopina. Osim toga provedena je detaljna analiza različitih modela za etanol i testirana je njihova sposobnost reproduciranja promjene strukture s promjenom molarnog udjela komponenti. Testirani su modeli parametrizirani na termodinamičkim svojstvima čistih tekućina i jedan model parametriziran na aktivitetu otopljene tvari u vodenoj otopini.

Broj stranica: 160
Broj slika: 84
Broj tablica: 11
Broj literaturnih navoda: 140
Broj priloga: 2
Jezik izvornika: Engleski
Rad je pohranjen u: Nacionalnoj sveučilišnoj knjižnici u Zagrebu, Sveučilišnoj knjižnici u Splitu, Knjižnici prirodoslovno – matematičkog fakulteta u Splitu, Knjižnici Nastavnog zavoda za javno zdravstvo SDŽ

Ključne riječi: etanol, mikroheterogenost, mikrostruktura, mješavina etanola i vode, modeli etanola, molekularna dinamika, vodena otopina etanola.

Mentori: Dr.sc. Franjo Sokolić, redoviti profesor
Dr. sc. Larisa Zoranić, docent

Ocjenjivači: 1. Dr.sc. Željko Crljen, redoviti profesor
2. Dr. sc. Ivica Aviani, izvanredni profesor
3. Dr. sc. Nives Štambuk-Giljanović, redoviti profesor

Rad prihvaćen: 24. 10. 2012.

BASIC DOCUMENTATION CARD

University of Split
Faculty of science

Ph.D. thesis

MICROHETEROGENEITY IN ETHANOL – WATER MIXTURE

Marijana Mijaković

Thesis performed at Faculty of Science in Split

Abstract

Ethanol-water liquid mixture was analyzed by the method of Molecular Dynamics, over the whole composition range to elucidate its micro structure. A recently introduced concept of microheterogeneity was used to describe mixture's behaviour. It was discovered that there are at least three structural regimes with respect to composition range in aqueous ethanol at ambient conditions. Up to approximately 0.15 mole fraction of ethanol, hydrophobic ethanol's sites distribution declines from homogenous. From 0.15 to approximately 0.65 mole fraction of ethanol, ethanol and water form bi-continuous intertwined microheterogenous mixture of segregated domains, and above 0.65 mole fraction of ethanol water network loses its connectivity over the whole system. Results from the novel approach in cluster analysis, performed for the first time on this kind of system, namely the calculation of cluster radial distribution function, were especially important for supporting the picture of three structural regimes. This is the first time that such detailed overall picture of the microscopic level structure of this liquid mixture is given. Besides, in-detail analysis of different ethanol models in their ability to reproduce this structural change with mole fraction is presented; models that are parametrized on pure substance thermodynamical properties, as well as one model parametrized on the solute activity in the aqueous mixture.

Number of pages: 160
Number of figures: 84
Number of tables: 11
Number of references: 140
Number of appendices: 2
Original in: English
Thesis deposited in National and University Library in Zagreb, University Library in Split, Library of Faculty of Science in Split, Library of Institute for Public Health of Split's-Dalmatian County

Keywords: aqueous solutions, ethanol, ethanol models, ethanol-water mixture, microheterogeneity, microstructure, molecular dynamics.

Supervisors: Dr.sc. Franjo Sokolić, full professor
Dr. sc. Larisa Zoranić, assistant professor

Reviewers: 1. Dr.sc. Željko Crljen, full professor
2. Dr. sc. Ivica Aviani, associated professor
3. Dr. sc. Nives Štambuk-Giljanović, full professor

Thesis accepted: 24 October 2012

Acknowledgements

First of all I would like to acknowledge SRCE, University Computing Centre at University of Zagreb, as all the calculations were done there, on Isabella cluster.
My thanks also goes to the Institute of Public Health of Split's-Dalmatian County, that have financed first two years of my Ph.D. study.

Of all the people who have contributed to this work, first place is for my mentor Franjo Sokolic. Then I would like to thank my jury members; Zeljko Crljen, Ivica Aviani and Nives Stambuk-Giljanovic and all the teachers from my Ph.D. study, primarily study leader Davor Juretic. Big thanks to my group members; Aurelien Perera, Bernarda Kezic and Kamil Polok, and to my co-authors on two papers; Augustinus Asenbaum, Christian Pruner and Emmerich Wilhelm. I thank all my study colleagues and co-workers from the Institute of Public Health. For all the administrative help I thank Irena Bitunjac.
For all the support and encouragement I thank all my friends and my family. Finally I thank my friend and my co-mentor Larisa Zoranic because she believed in me, even in dark times when I did not.

I dedicate this work to Leo. Some may think he is not even a real person, but I know better.

Contents

List of Tables

List of Figures

List of Illustrations

List of Symbols and Abbrevations

χ mole fraction

MD Molecular Dynamics

LJ Lennard-Jones

LB Lorentz Berthelot

IUPAC International Union of Pure and Applied Chemistry

OPLS Optimized Intermolecular Potential Functions for Liquid Simulations

TIPS Transferable Intermolecular Potential Functions

NpT isothermal-isobaric ensemble

SPC Single Point Charge model for water

TraPPE Transferable Potentials for Phase Equillibria

UA united atoms

KBFF force field based on Kirkwood-Buff theory

KB Kirkwood-Buff

SPC/E Single Point Charge Extended model for water

RDF radial distribution function

NMR nuclear magnetic resonance

NVT canonical ensemble

E ethanol

W water

AA all atoms

TIP4P Transfferable Intermolecular Potential 4 Points

TFE triflourethanol

WS Weerasinghe Smith model

TBA tertiary butanol

GROMOS GROningen MOlecular Simulation computer program package

AMBER Assisted Model Building with Energy Refinement program package

CHARMM Chemistry at HARvard Macromolecular Mechanics: name of a widely used set of force fields for MD

N number of particles

b.a. bigger angle

s.a. smaller angle

σ parameter for the radius in the Lennard-Jones potential

ε parameter for the energy well depth in the Lennard-Jones potential

CPU central processing unit

FFT Fast Fourier transform

-hrm and harm angle restraint potentials in DLPOLY

NVE microcanonical ensemble

τ_β coupling constant in thermostat

g(r) radial distribution function

3ps three peaks of water oxygen radial distribution function

CN coordination number

μVT grand canonical ensemble

KBI Kirkwood-Buff integral

rKBI running Kirkwood-Buff integral

cRDF cluster radial distribution function

1 Introduction

1.1 Aims of this research

The problem dealt with in this work is in the core of the behaviour of the many bio-systems on molecular level. It is the problem of the behaviour of the water solutions of amphiphilic molecules, where ethanol molecules are taken as one of the smallest and the least complex amphiphilic molecules. Only recently it has been understood that these kinds of binary mixtures, when mixing two hydrogen bonded liquids such as water and ethanol, posses microheterogeneous structure [16], and very big contribution to this understanding came from the previous work on microheterogeneous liquid binary mixtures of the group candidate is working with [17], [18]. A multidisciplinary approach is important when working on this kind of problems, as physical chemistry measurements are essential to be compared with the physic's calculations and simulation results. This type of problems also has significant informatics and mathematical principles part. The aim of this research is to examine the crossover behaviour of the ethanol water mixture that seems to be in between that of the simple mixtures of disordered liquids and more complex amphiphilic systems that form micelles. This crossover behaviour is not yet fully understood. By comparing previously known experimental data - excess enthalpy, heat capacity, compressibility and azeotropy, to our measurement of the speed of sound, and Molecular Dynamics simulation results on thermodynamical properties and, more important structural properties, it appears that non-monotonic changes of macroscopic properties with mole fraction change, take place in this liquid mixture. It appears that there are two microscopic heterogeneous structure changes, first at $\chi_1 \approx 0.15$ mole fraction of ethanol, and the other at $\chi_2 \approx 0.65$ mole fraction of ethanol. This is in accordance with the idea of three different mixing schemes in aqueous alcohols, presented in 1996 by Tanaka et al. [19]. First change is the change of the rigidification of the water hydrogen bonded network and clustering of the ethanol molecules. Below χ_1 ethanol molecules are shielding their hydrophobic sites together, in small clusters. The increased amount of

this small ethanol clusters within water network is rigidify it, in a way that it becomes less compressible. Hydrogen bonded water network is flexible, but when small clusters of ethanol are distributed within this network, these areas with ethanol inside lose their compressibility, so this areas occupied with ethanol molecules diminish the overall water network's compressibility. When mole fraction of ethanol reaches approximately χ_1 value, the structural change occurs that overturns this effect of rigidifying of the water network; hydrophobic ethanol parts are not shielded as effectively as before and ethanol and water start to form bi-continuous microsegregated phase. The second change approximately at χ_2 corresponds to the change in the water network, it starts to break into smaller fragments. This change is less marked in the various thermodynamical properties, but it is still obvious, especially from the MD data.

Molecular Dynamics simulations stand together with the experimental techniques in investigations of the molecular as well as more complex biological systems. With an increase of the computer processing power, larger space and time scales become accessible, and it is a shift in the direction of getting more realistic simulation data. However, the foundation of the simulations are force field models, so it is necessary to critically address their reliability under different conditions and transferability to mixed systems.

Generally, classical force fields are effective two body potentials, comprising the Lennard-Jones (LJ) potential for the van der Waals interaction and the electrostatic interaction of partial charges that models hydrogen bonding. Also in a standard procedure, potential models for mixtures are built from the pure system force fields where cross interactions are calculated in a standard way: LJ part using geometrical or arithmetical rules (so called Lorentz-Berthelot rules, LB rules); and Coulombic part by superposition. However, the transferability of the pure system models to a mixed system is not granted. As recent results from simulations of aqueous organic compounds mixtures (aside from ethanol-water simulations) show, some excess quantities as excess enthalpy, and dynamic properties as diffusion coefficient, which are the most sensitive to mixing properties, are difficult to reproduce correctly. In this curse, it is reasonable to ask if models parametrized on pure liquids

are expected to be successful enough in reproducing excess quantities of the mixture. The other aim of this work is to test different force fields in order to give an answer if all these models are reliable enough in the reproduction of the microscopic structure and its change with mole fraction of aqueous ethanol mixture.[1]

1.2 Previous investigations

1.2.1 Alcohols

Alcohol is a chemical compound with a hydroxyl group(-OH) bound to the saturated carbon atom (-C). Simple acyclic, monohydric (having one -OH group) alcohols have the general formula: $C_nH_{2n+1}OH$. The first few acyclic monohydric alcohols are listed here:

Chemical Formula	IUPAC Name
CH_3OH	Methanol
C_2H_5OH	Ethanol
C_3H_7OH	Isopropyl alcohol
C_4H_9OH	Butyl alcohol (Butanol)
$C_5H_{11}OH$	Pentanol

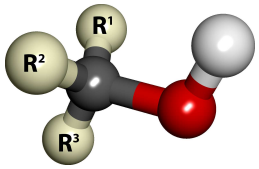


Illustration 1: General model of alcohol molecule

Alcohols are amphiphilic molecules. On one side they have polar hydroxyl group, and on the other a hydrophobic tail. The length of this tail determines the degree of the alcohol's solubility in water. Only the first few alcohols are miscible with water (up to butanol). The rest of them have a hydrophobic

[1] All results concerning testing of different ethanol model are from article in preparation[20]

tail too big to allow total miscibility with water, even though the hydroxyl part of the molecule is polar and tends to form H-bonds with water molecules.

1.2.2 Ethanol, general facts

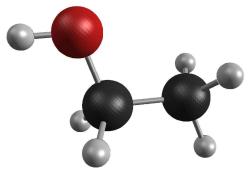


Illustration 2: Model of ethanol molecule

Ethanol is a clear, colourless liquid in the temperature range from 159 K (-114 C) to 351 K (78 C) at atmospheric pressure. Its molecular formula is C_2H_5OH, empirical formula C_2H_6O and molecular weight 46.068 g/mol. Ethanol is miscible with ether, acetone, benzene, acetic acid, many organic solvents and water. It is a one H-bond donor and one H-bond acceptor. Ethanol-water mixtures have less volume than the sum of their individual volumes at the given mole fractions. The reaction of mixing ethanol and water is exothermic. Mixtures of ethanol and water form an azeotrope at about 89 mole fraction of ethanol at normal pressure and T = 351 K temperature.

Ethanol and methanol are the only two linear alcohols that crystallize, bigger alcohols form glasses. Solid ethanol is composed of linear hydrogen-bonded chains of molecules, with molecules arranged in an alternating sequence. For ethanol, only one crystalline phase is found at atmospheric pressure. The pairs of molecules along the chain are linked together in *trans* and *gauche* conformations, with the carboxyl group directed away from the centre of each pair. These pairs alternate along the chain. On the contrary, the high-pressure structure of ethanol crystal is represented by chains consisting only of molecules in *trans* conformation and they are linked in each chain so that their carboxyl groups are coplanar and aligned in the same direction

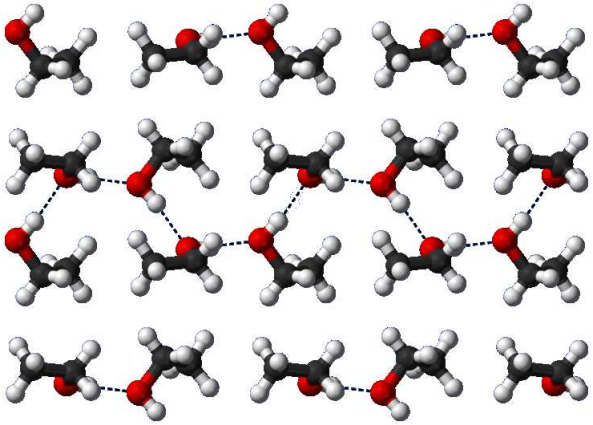


Illustration 3: Structure of solid ethanol (Ball-and-stick model of part of the crystal structure of ethanol at 87K (186C). X-ray crystallographic data from [14]

along the b axis of the cell [21]. In the vapour phase separate molecules are of the average size 0.4 nm [22].

1.2.3 Models of liquid ethanol

For determining the structure of liquid ethanol (and many more molecular systems) computer simulations present the opportunity to obtain detailed insight into the structure and dynamics of the system in question. The key issue for computer simulations is determination of the model for potential functions that properly describe inter-atomic interactions (potentials) in the modeled systems.

Various models of liquid ethanol based on Coulomb potential of point charges and Lennard-Jones potential are present in the literature, rigid model that neglects internal degrees of freedom in the molecule, and flexible models that comprise the flexibility of dihedral angle, flexibility of angles between H-O-CH_2 and O-CH_2-CH_3 bonds in ethanol molecule, and some that take into account the elongation of the bonds. There are models of the united-atoms type, that reduce contribution of the group of atoms to one site, such as CH_2 and CH_3, and others, all-atoms type, take into account every atom

separately. polarizable models explicitly take into account the polarizability of the molecules. Some of the models are parametrized by fitting the liquid-vapour coexistence curve, and others by fitting to the experimental Kirkwood-Buff integrals.

The geometry, bond lengths and angles, of the ethanol molecule was probed through various experimental techniques; microwave spectroscopy [23], neutron diffraction [24], [25], [26], x-ray diffraction [27], electron momentum spectroscopy [28].

Here is a short overview of the most commonly used models for liquid ethanol, in historical order as they appeared in the literature. Five of these models are used in this work (details in section 2), and results obtained by some of those not used here are discussed later in this work.

Based on the molecular geometry given in [23], Jorgensen developed rigid united-atoms model for liquid ethanol [29]. One model commonly used for liquid ethanol, OPLS, Optimized Intermolecular Potential Functions for Liquid Simulations, developed by Jorgensen [30], has been derived by directly fitting experimental thermodynamic and structural data of pure organic liquids, liquid water, and aqueous solutions of organic molecules and ions representative of peptide constituents. It was developed as an extension on Jorgensen model for pure liquid hydrocarbons [31], which was done by reparametrization of the TIPS (Transferable Intermolecular Potential Functions) for water, alkanes, alcohols and ethers [32]. For liquid alcohols TIPS and OPLS potentials have the same form, but differ slightly in the value of their parameters. In his work Jorgensen [30] did the Monte Carlo simulation of the thermodynamic and structural properties of liquid ethanol at fixed atmospheric pressure (p= 1 atm) and fixed temperature T=298 K in the isothermal-isobaric (NpT) ensemble. The system consisted of 128 molecules in the cubic cell. Jorgensen confirmed that ethanol molecule in a liquid state has a *trans* conformer and two mirror image *gauche* conformers. This result is in accordance with ab initio calculations. Hydrogen bonds were confirmed to be nearly constant at $2.7-2.8\mathring{A}$. Computed heat of vaporization of 9.99 kcal/mol was found to be in good agreement with experimental value of 10.11 kcal/mol from [33]. The density of 0.748 g/cm^3 was also found to be

in good agreement with experimental data of 0.7851 g/cm^3 from [33]. Heat capacities and isothermal compressibility were also found to be in agreement with the experimental data. The results on the location and area of the first peak for the O-O radial distribution function were in excellent agreement with the X-ray diffraction data [34, 35, 27].

All-atoms potential with the flexible dihedral angle was developed by Muller-Plathe [36], aimed to be used with SPC water [37] for simulation of polymer membranes.

In 1996 Jorgensen [38] developed all-atom type potential with stretching, bending and internal rotation degrees of freedom. Cornell potential for organic molecules [39] (AMBER) is based on the same approach as Jorgensen's all-atom OPLS, but they differ in values of parameters.

Chen, Jeffrey and Siepmann [40], a group working on the development of the transferable potentials for phase equilibria force fields, proposed new TraPPE and TraPPE-UA (united atoms) force fields for alcohols, because it was found [41] that Jorgensen's OPLS was not transferable to the longer alcohols and to elevated temperatures. The authors claimed that TraPPE-UA force field had satisfactory performance for the prediction of the thermophysical properties. Structural analysis of the liquid alcohols showed that for all neat alcohols (methanol, ethanol, propan-1-ol, propan-2-ol, butan-2-ol, 2-methylpropan-2-ol, pentan-1-ol, pentane-1,5-diol and octan-1-ol) on average two hydrogen bonds are formed per hydroxyl group, regardless the difference in the cluster size distributions.

In 2005 Serbanovic at al. [42] used the flexible OPLS-UA force field from [43] in a Monte Carlo computer simulation that predicted vapour-liquid equilibrium of few binary systems with propane + ethanol as one of them. Flexible OPLS-UA force field differs from Jorgensen's semi-flexible force field in the added flexibility of the angle between bonds [44]. The model was found to be in good agreement with the experimental data.

In the year 2005 Schnabel et al. [45] proposed rigid united-atom model that uses three nuclei off-center LennardJones united atoms for the methyl, methylene and hydroxyl group. Geometry was obtained by quantum chemistry calculations and model was developed to predict Henry's law constants for

various solutes in ethanol as solvent. Model was parametrized to give accurate density and vapour pressure of pure ethanol.
In the year 2012 Jiao [46] developed force field for liquid ethanol based on Kirkwood-Buff (KB) theory of solutions, KBFF, because many existing force fields perform poorly in their ability to reproduce the experimental KB integrals of the alcohol-water solutions. This force field was developed to be used with the Simple Point Charge/Extended (SPC/E) water model.
There are few polarizable models for ethanol: Gao [47] developed a polarizable intermolecular potential by defining atomic dipoles, [48] included charge carrying Drude particles in their molecular model and [49] used an extended Hamiltonian approach to allow molecules to respond to the environment and developed flexible polarizable model for ethanol. In this model intramolecular potential varies with the field and atomic charges also fluctuate with the field. At any instant of time each molecule in the bulk has its own set of charges and intramolecular potentials. This model predicts longer chains of H-bonded molecules, as H-atom in this model is allowed to adapt possition in response to the field, and in that way optimize the H-bonded structures.

1.2.4 Simulations of pure ethanol

Here are presented two simulations of pure ethanol, to give an insight at various properties that are attainable by Molecular Dynamics simulations.
Saitz, Padro and Guardia [50] carried out molecular dynamics simulation of liquid ethanol at four thermodynamic states ranging from T=137 K to T=348 K. They used OPLS potentials, and reported thermodynamic, structural and transport properties for liquid ethanol. Heat of vaporization was found to be in overall agreement with the experimental data, only 1% higher than the experimental value, and in good agreement with the data from other simulations. They calculated the percentage of *gauche* conformers in liquid at 298K to be 50% and concluded that more precise experimental data, and more detailed simulation are needed to decide on precise molecular conformation of liquid ethanol. Radial distribution functions (RDF) data

that best describe the structure of liquids, was found to be in good agreement with the earlier computer simulations. They confirmed that liquid becomes more structured as the temperature decreases. The changes of temperature do not affect the positions of the maxima and minima, but only their height and depth. They compared the RDF data with the data for the liquid methanol and proposed that these two liquids have the similar structure, predominantly governed by chains of H-bonded molecules. At T=298K, authors found that 80% of ethanol molecules have two H-bonds, 14% have 1 H-bond, 5% have 3 H-bonds and 1% are not H-bonded. Branching of the H-bonded chains was reported to be nearly constant with the increase in temperature,, and with decreasing temperature chains become longer. Authors conclude that at room temperature there are very few hexamer closed chains in the liquid ethanol, in the contrast with Sarkar and Joarder [51] who suggested that closed hexamer chains are in majority at room temperature. Self diffusion coefficients for the centre of mass were calculated and they show a decrease with the temperature decrement, what they suggest to be in good agreement with the formation of the longer H-bonded chains at lower temperatures. Diffusion coefficients were reported to be in good agreement with the data from NMR experiments, although somewhat higher. Another comprehensive study on the structure, and also the clustering of neat liquid ethanol is the one by Benmore and Loh [26], using neutron scattering in pure ethanol at room conditions, and Molecular dynamics simulation. The authors measured directly 10 structure factors, out of 21 possible partial structure factors that exist for pure ethanol. The data have been used to obtain the details of the conformation of the ethanol molecule in the liquid state. The intermolecular structure was compared to the molecular dynamic simulation done with four-site and nine-site Jorgensen potential for liquid ethanol. The authors obtained six hydrogen-hydrogen partial structure factors, and four composite structure factors. Molecular dynamic simulations were done with 125 semi-flexible molecules in a cubic cell, in the NVT ensemble. The Nose-Hoover thermostat was used to control the temperature at 300K and density at 0.7873 g/cm^3. Simulation runs were 100 ps long. These were still small systems and very short simulation runs. It was

found that the position of the first peak of the radial distribution function for O-O sites agrees better with the experimental data for the four-site model, and the peak position of the H1-H1 sites has better agreement in the nine-site model, but in general they were found to be in good agreement with the experimental data concerning the positions of the first peaks in radial distribution functions. On coordination numbers authors claimed that hydroxyl hydrogen is surrounded by 2.0 ± 0.2 hydroxyl atoms at 3.0 $\mathring{A}$, and oxygen is surrounded by 0.95 ± 0.3 hydroxyl atoms at 2.1 $\mathring{A}$ and 2.0 ± 0.2 oxygen atoms at 3.0 $\mathring{A}$. Authors concluded that experimental and simulation data are inconsistent with the presence of the closed hexamer clusters, but consistent with the existence of the H-bonded winding chains of molecules, and that the U-shaped chain agreement is very good.

Ethanol is a H-bonding liquid and as such it is an associated liquid. It has the structure of H-bonded clusters. Up to now there has been an ongoing discussion about the exact form of clustering in neat ethanol. Different cluster sizes were proposed and it is still not clear if these clusters are linear or cyclic in nature [33, 35, 51]. In 1999 Ludwig, Weinhold and Farrar [52] presented a quantum cluster equilibrium theory for liquid ethanol. The authors claimed that at room temperature neat liquid ethanol consists of approximately equal parts of monomer, cyclic tetramer, and cyclic pemntamer clusters.

1.2.5 Simulations of ethanol-water systems

Data from simulations presented here, on enthalpy, excess enthalpy and diffusion coefficients, will be later compared to the results from this work.

Wensink et al. [53] studied binary mixtures of alcohols (methanol, ethanol and 1-propanol) with water in order to compute excess properties of mixing, in particular the relation between mobility and viscosity. Most of the simulated properties were found to be in good agreement with the experimental data. Enthalpy of mixing and excess density were particularly well reproduced, even though excess density is underestimated, and simultaneously excess

enthalpy (enthalpy of mixing) is overestimated. The authors attributed this to the fact that parameters in the force field may not be well suited to the mixtures, even though they give excellent results for the neat compounds. The authors proposed their work to serve as a benchmark for further improvements on the models. Self diffusion coefficients were computed from the mean square displacement using Einstein's relation (details in subsection 3.7). Overestimated values were obtained for the diffusion coefficients, but with correct trends. The maximum in the excess viscosities obtained from this simulation was found to be shifted to the higher value of alcohol concentration than in the experimental data and the excess viscosity was found to be underestimated, whereas the pure components were reproduced rather well. Authors adopted the view that the Stokes - Einstein's relation, that relates diffusion with viscosity holds at the molecular level, supported by the [54, 55], so that allowed them to compute the effective hydrodynamic radius of the particles in the mixture. Experimental radii have the minimum at the low ethanol concentration, which is reproduced well by the simulation (minimum at $\chi = 0.2$ mass fraction, $\chi = 0.09$ mole fraction ethanol) and authors concluded that there is no evidence for the diffusion of the larger entities, 'collective diffusion' in any of the mixtures. At the end the authors concluded that the OPLS model for water and short alcohols gives qualitatively correct answers for properties like energy, density and viscosity, but that the exact micro structure of this mixtures is still not fully understood.

One more extensive study on ethanol water mixtures was done by Noskov, Lamoureux and Roux [48] focusing mainly on the number of hydrogen bonds in the system, namely elucidation of the structure of the hydrophobic hydration in ethanol water mixture. The second aim of the work was to develop a new polarizable force field model for ethanol. The system under investigation consisted of 250 molecules. Simulations were performed in the NpT ensemble, at room temperature and pressure (T = 300 K, p = 1 atm). The calculated enthalpy of vaporization for neat ethanol of 10.19 kJ/mol, is in good agreement with the experimental data authors showed of 10.15 kJ/mol [56]. The authors plotted the radial distribution function for $O^{ethanol}$-$O^{ethanol}$ and $O^{ethanol}$-$H^{ethanol}$ sites compared with the radial distribution functions from the

non-polarizable model [30] and found positions of the first peak to be in good agreement with the both, [30] and the experimental data, although somewhat shifted to the smaller radius than the non-polarizable model predicts. The model was able to reproduce the positions of the minima in the self diffusion coefficients for water at ethanol concentration 30% and for ethanol at ethanol concentration of 20%. For pure liquid water, authors report an average of 3.03 H-bonds per water molecule in the neat water system, with the $(d_{HO} < 2.4\mathring{A})/150^{\circ}$ geometric definition of the H-bond, and 1.65 H-bonds per ethanol molecule in the neat ethanol system. From positive excesses for the W/W and E/W numbers of H-bonds and the negative excesses for the W/E and E/E numbers of H-bonds authors concluded that water is an overall better solvent than ethanol: at any ethanol fraction, both water and ethanol are preferably solvated by water. When solvated in neat ethanol, a water molecule loses 14% of its H-bonds, while on the contrary, when ethanol molecule is solvated in water, it gains more H-bonds, going from 1.65 to 2.23, as the concentration of ethanol goes from 100% to 0%. The maximum water H-bonding excess was found to be at 15% ethanol concentration. The authors presented the analysis of the fraction of the water hydrogen bonded clusters, ethanol hydrogen bonded clusters and all hydrogen bonded clusters with the respect to the cluster size. They concluded that the ethanol component does not percolate at any concentration, while they claimed that water percolates for the concentrations of ethanol below 30%. As the indication of the existence of the percolation the authors took the existence of the clusters with the size close to the whole system size. The authors claimed that there is an optimal number of water molecules to solvate each ethanol molecule at 20% ethanol concentration. Finally, the net effect from the entire first hydration shell is a reduction rather than an excess of water H-bonds, and the dominant contribution arises from the structuring of water in the second hydration shell of ethanol. The authors also claimed that the presence of water clustering exists from zero to very high ethanol concentrations in this mixture.

Zhang et al. [57] in 2006 provided the study of the self diffusion coefficients and mutual diffusion coefficients over the whole range of the concentrations

in ethanol - water mixture, done by molecular dynamics simulation. The authors used the OPLS-AA (optimized potentials for liquid simulations - all atoms) force field for ethanol and the SPC (simple point charge) for water. The system consisted of 250 molecules in a cubic cell. Simulations were performed in NVT ensemble at the temperature of T=298 K. The equilibration lasted for 600 ps, and the run that was used to collect the necessary statistics lasted for 100 ps. In a binary mixture there are two kinds of diffusion coefficients; self diffusion coefficient and mutual diffusion coefficient. Self diffusion coefficient is described with the velocity autocorrelation function of the molecules of each species. The authors found the minimum in the self diffusion coefficient of ethanol to be at the 30% mole fraction of ethanol, and sharp decrease in the self diffusion coefficient of water up to 30% mole fraction of ethanol. For bigger ethanol concentrations the authors found self diffusion coefficient of water to remain constant. The mutual diffusion coefficient describes the mobile ability of species A diffusing into species B. It has a kinematic part and a thermodynamic part. Mutual diffusion coefficients were found to be larger than zero due to non-ideality of the ethanol-water mixture.

Zhang and Yang [58] reported another study on structure and diffusion properties of ethanol water mixture at 298.15 K and atmospheric pressure by molecular dynamic simulation. A simple 'rigid molecule' model was used for ethanol, and TIP4P for water. The total number of molecules was 500, and the simulations were done in NVT and NpT ensembles. The validity of the rigid model was verified by obtaining the correct enthalpy of vaporization for neat water and neat ethanol, in good agreement with [53]. The absolute value of the excess of heat of mixing was underestimated at low ethanol concentrations, and in good agreement with the experimental data above $\chi = 0.90$ mole fraction of ethanol. The authors presented O-O and O-H radial distribution functions, and concluded that water-water correlation is enhanced in the mixture compared to that in pure water, while ethanol-ethanol H-bonding structure is gradually broken as the ethanol concentration increases. The strong interaction between ethanol and water molecules leads to the enhancement of the correlation as the ethanol mole

fraction increases. Self diffusion coefficients were computed to be larger than the experimental data, but this is usually the case with the self diffusion coefficients from simulations. However, they were found to be lower than the data from reference simulation [53]. The self diffusion coefficients calculated from the velocity autocorrelation function were slightly better than those computed from mean square displacement. The mutual diffusion coefficients obtained from the MD simulation were in fair agreement with the experimental data in the ethanol - water mixture.

1.2.6 Microheterogeneity

Microheterogeneity and the microheterogenous structure change along three regimes with the mole fraction change is one of the key features of ethanol-water mixture that is analyzed in this work. Neat ethanol and neat water are both associated liquids, because they are both hydrogen bonding liquids, and each of them has its own microstructure, which are not the same. Water has a dense, flexible, 3-dimensional H-bonded network structure, unlike all other liquids. When ice melts to water, it should be expected that the H-bonded ice structure is destroyed while liquefying. That is not so. When melting, the H-bonded network is not disrupted in most part, it just starts to have fast librations of the individual constituent H_2O molecules. Experimental results show that the rotational angles of circa 35 degrees are attained in approximately 700 ps time [59]. These hindered rotations (librations) induce great distortions in the H-bonded network of liquid water and they are at the origin of water's fluidity and flexibility [60]. Even though, the structure of this familiar and abundant liquid still poses an unanswered question [61]. Ethanol, on the other hand, has an yet unresolved clustered microstructure. Microstructures of ethanol and water are different from the neat Lennard-Jones liquids, that have dense packing of atoms with the coordination number of 12, because H-bonds in associating liquids require a certain number of neighbours in the vicinity of every molecule. Microstructure of these liquids results from these two requirements, on one side the dense packing structure

of liquids, and on the other constraints to the structure demanding from the H-bonds. When these liquids are mixed, a microheterogeneous structure emerges as a consequence of the competition between their microstructures, because of the numerous possibilities of local order. Each liquid prefers its own microstructure, so they locally separate. This micro-separation is called microheterogeneity. An important point is that these liquids mix well, as if micro-separation helps them to stay together[2]. Another important point is that this is happening in thermodynamical equilibrium. Usually formation of segregated domains is seen close to the phase transition point.

Three significant papers published in 2007 from A. Perera, F. Sokolic and L. Zoranic [17], L. Zoranic, F. Sokolic and A. Perera [18] and [62] for the first time introduced the clear terminology in the topics of associated liquids, addressing the microheterogeneity of the neat alcohols (methanol and tert-butanol) as the microstructure of neat liquids, while the segregation of the small domains of the two components in their binary mixtures with water was called microheterogeneity. In the [18] authors reported the feature of the microstructure found in neat methanol at room temperature and pressure, by the molecular dynamic simulation, to be in agreement with the experimental findings, namely the prepeak in the structure factor, corresponding to the distance of about 6 Å. This feature corresponds to the first time noticed bump in the probability of H-bonded clusters of oxygen-oxygen sites at the cluster size of 5, for neat methanol. Authors claimed that this indicates that about five oxygen atoms cluster preferentially, through H-bonding mechanism, in accord with the experimentally known fact that methanol molecules tend to form chains with rich topology [63]. The non-H-bonding sites were reported to have structure factor typical of an atomic liquid, and the H-bonding sites to have a prepeak that corresponds to a 7.5 Å distance. In order to confirm that this prepeak is entirely due to the local organization coming from H-bonds, tert-butanol system simualtion was done, under the same conditions, but without partial charges, and the prepeak has disappeared.

[2]On this point one clarification of the terminology is useful, as it is not consistent in the literature. In this work microheterogeneity is the name reserved for the micro-separation of the components in the mixture. Structure of the neat associated liquids will be referred as micro-structure.

Authors also showed that oxygen-oxygen cluster size probability function has a pronounced peak at the cluster size around 4, while the probability of finding monomers was smaller than that. The striking feature reported in this work is that neat alcohols show structure strongly reminiscent of that in microemulsion after the disorder to order phase transition, but in the alcohols it is within the disordered phase. The viewpoint introduced in this work is that associated liquids can be viewed dually as constituted of molecules interacting through strong directional forces, and at the same time as the mixture of the microclustered molecular domains.

1.2.7 Problems with excess enthalpy and diffusion coefficients

In two works 2011 Chitra and Smith [64], [65] performed molecular dynamic simulations on 2,2,2-trifluoroethanol (TFE) - water mixture with a variety of standard models for TFE. None of the tested models reproduced excess enthalpy and diffusion coefficients of mixture with satisfactory accuracy. They all give overestimated self-association of solute.
Molecular dynamic simulation of acetone-water mixture by Perera and Sokolic in 2004 [66] for OPLS acetone model produced results for excess enthalpy that were not in good agreement with experimental data, while Weerashinge Smith (WS) acetone model [67] results were in much better agreement with experimental data. However, WS model wasn't successful in reproducing internal energies with satisfactory agreement with experimental data, so it seems that the choice of the model is always a trade-off, different models reproduce better different properties of mixtures.
Lee and van der Vegt [68] in 2005 reported results for simulations of aqueous tertiary butanol (TBA) solutions with model they developed by targeting experimental Kirkwood-Buff integrals with reparametrization of distribution and magnitude of partial atomic charges of GROMOS model, because results they obtain with GROMOS [69] and OPLS [30] force fields didn't yield satisfactory results, again reporting big disagreements in excess enthalpies of mixture.

Excess enthalpy for aqueous methanol mixtures simulation, with ethanol modeled by OPLS [30] force field were reported to be too high and not in the best agreement with experimental data, 2006 by Gonzalez-Salgado and Nezbeda [70] and 2010 by Perera et al. [71].
Kang and Smith [72] reported excess enthalpies for aqueous solution of N-methylacetamide (NMA) modeled with CHARMM [73] force field to be in disagreement with experimental data, while data obtained by their new model parametrized on mixture's Kirkwood-Buff integrals were in much better agreement.
In 2009 Zoranic et al. [74] reported excess enthalpies for aqueous solutions of three amides modeled by force filed from [75] and [76] that were too high for all three amides.
In 2010 Dai et al. [77] reported molecular dynamic simulation data of various binary mixtures of organic molecules. Simulations were done using OPLS [38]force fields for organic molecules. The general conclusion was that excess enthalpies for these mixtures and force fields are higher than experimental values.
Few possible routes for the improvement of force fields for mixtures exist in the literature. One are recently emerging new force fields, that include explicit polarizability in the models. Also, the issues of the combination rules are critically discussed [45], [70]. Yet another route is to propose models for mixtures that are parametrized on the mixture properties, such as the solute activity change with solute concentration. As discussed in [78], it is not clear if models parametrized on data related to chemical potentials along the temperature change are successful in reproducing Kirkwood-Buff integrals on fixed temperature along the composition range. It is indeed true that such model (TraPPE) has problems in reproducing it, as confirmed in this work.
In our recent publication, using standard models, highly overestimated experimental excess enthalpies for ethanol-water mixture were obtained[6].

1.2.8 Conclusion on previous investigations

As it can be seen, early work was concentrated at the finding of the proper models for real systems, and later on the agreement of the various calculated properties with the experimental data. Up to now much has been said on the diffusion coefficients and local hydrogen bonding of this system. There were some attempts on deciphering the structure of the mixture, but not a single one to give the global view of the structural changes of this mixture with the molar fraction change. All the studies were concentrated mainly on giving the description of the local structure around H-bonding atoms, and the behaviour and structure of this mixture is governed by an emergent phenomenon - microheterogeneity. There is obviously a lot more to it that can be found and explained.

There are lots of evidence in the literature, listed in the previous section, that classical non-polarizable models for alcohols do not reproduce some of the mixture's properties, as excess enthalpy and diffusion coefficients.

2 System and methods

2.1 System

2.1.1 System size

The system under investigation consisted of a fixed number of N=2048 of particles. This number was found to be enough to get the density correlations correct [18] - [74]. The exact number of molecules for each mole fraction is listed in the Table 1

Table 1: Number of molecules for each mole fraction

Ethanol mole fraction	Number of Ethanol molecules	Number of Water molecules
0.00	0	2048
0.05	102	1946
0.08	163	1885
0.10	204	1844
0.12	245	1803
0.15	307	1741
0.18	368	1680
0.20	409	1639
0.22	450	1598
0.25	512	1536
0.30	614	1434
0.40	819	1229
0.50	1024	1024
0.60	1228	820
0.70	1433	615
0.80	1638	410
0.90	1843	205
1.00	2048	0

2.1.2 Ethanol models

Torsional motion about C-O bond puts ethanol molecule in two mirror imaged *gauche* and one *trans* conformer. It can be clearly seen from the rotational energy function in Figure 1 from [30]. Trans conformer has H-atom positioned at 180°, ant two gauche conformers at 60° i -60°.

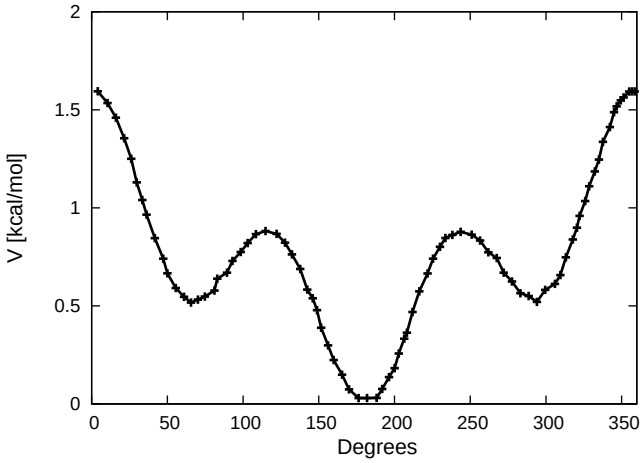


Figure 1: Rotational energy as the function of dihedral angle for ethanol

Five force field potential models for ethanol were used, all five of non-polarizable united-atom type, and all four consisting of Coulomb and Lennard-Jones interaction. In all of these models ethanol has four sites: H, O, CH_2 and CH_3, numbered 1,2,3 and 4, respectively. The first one is semi flexible Optimized Potential for Liquid Simulations, OPLS, from [30]. It is somewhere referenced as OPLS-UA (united atoms) to point out the fact that some of the atoms are united in one site and then parameters are attributed to the site as a whole interaction point. Second (it is called here: fully flexible OPLS with bigger angle (b.a.)) and third (it is called here: fully flexible OPLS with smaller angle (s.a.) models were fully flexible OPLS, from [43] and [42]. These two models differ from the semi flexible model in a way that while semi flexible model implements only the flexibility of the dihedral angle inside the all four sites of the molecule, (Illustration 4), the fully flexible model implements also the flexibility of two angles between neighbouring

chemical bonds in the molecule, namely the angle between H-O and O-CH_2, and the angle between O-CH_2 and CH_2-CH_3 chemical bonds, (Illustration 5). The fourth model is Transferable Potential for Phase Equilibria United Atoms, TraPPE-UA, from [40]. It is fully flexible, but differs in the values of the parameters from the fully flexible OPLS-UA potential. The last model is Kirkwood-Buff force filed, KBFF, from [46].

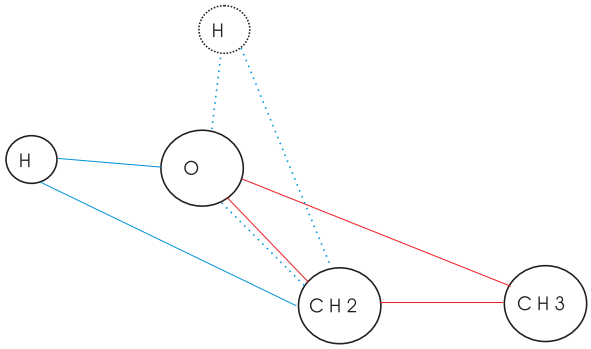


Illustration 4: Semi-flexible model of ethanol molecule

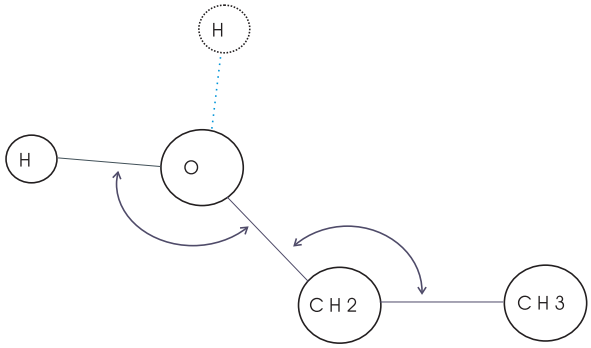


Illustration 5: Fully-flexible model of ethanol molecule

It is important to mention here that the parametrization of all included models was performed on systems with the size which is smaller than the system in this work.
The general form of the used force fields is as follows:

$$E(r^N) = E_{bonds} + E_{angles} + E_{dihidral} + E_{nonbonded} \tag{2.1}$$

with:

$$E_{bonds} = \sum_{bonds} K_r(r - r_0)^2 \tag{2.2}$$

$$E_{angles} = \sum_{angles} k_\theta(\theta - \theta_0)^2 \tag{2.3}$$

$$E_{dihidral} = \frac{A_1}{2}(1 + cos\phi) + \frac{A_2}{2}(1 - cos2\phi) + \frac{A_3}{2}(1 + cos3\phi) \tag{2.4}$$

$$E_{nonbonded} = \sum_{i>j} \frac{q_i q_j}{4\pi\epsilon_0 r_{ij}} + \sum_{i>j} 4\epsilon_0((\frac{\sigma}{r_{ij}})^{12} - (\frac{\sigma}{r_{ij}})^6) \tag{2.5}$$

Instead of elastic potential for the bonds, rigid bonds were used in all models. Semi-flexible OPLS does not have E^{angles}, but uses six length constraints. Interactions between non-bonded atoms are presented by pairwise additive Coulomb and Lennard-Jones potential. van der Waals force is modeled by Lennard-Jones potential. Cross-site parameters for Lennard-Jones potential (Illustration 6) were calculated by the following Lorentz-Berthelot rules [79]:

$$\sigma_{ij} = \frac{\sigma_{ii} + \sigma_{jj}}{2} \tag{2.6}$$

$$\epsilon_{ij} = \sqrt{\epsilon_{ii}\epsilon_{jj}} \tag{2.7}$$

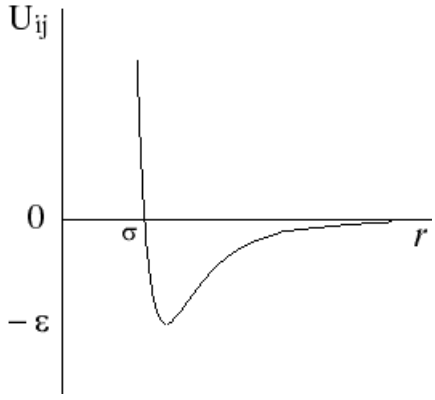


Illustration 6: Shape of the Lennard-Jones potential

The list of used parameters for all four ethanol force fields is listed in Table 2

Table 2: Force field parameters for five used ethanol models

	Semi flexible OPLS	Fully flexible OPLS b.a.	Fully flexible OPLS s.a.	TraPPE	KBFF
Sites masses [g/mol] and charges [e]					
H m	1	1	1	1	1
O m	16	16	16	16	16
CH_2 m	14	14	14	14	14
CH_3 m	15	15	15	15	15
H q	0.4350	0.4350	0.4350	0.4350	0.5200
O q	-0.7000	-0.7000	-0.7000	-0.7000	-0.8200
CH_2 q	0.2650	0.2650	0.2650	0.2650	0.3000
CH_3 q	0.0000	0.0000	0.0000	0.0000	0.0000
Length constraints between sites [Å]					
H-O	0.945	0.945	0.945	0.945	0.945
O-CH_2	1.430	1.430	1.430	1.430	1.430
CH_2-CH_3	1.530	1.530	1.530	1.540	1.530
H-CH_2	1.948				
O-CH_3	2.3954				
Angles potential parameters [kJ/mol], [degrees]					
H-O-CH_2		460.00	460.00	460.00	450.00
k_θ					
O-CH_2-CH_3		420.00	420.00	420.00	520.00
k_θ					
H-O-CH_2		108.50	108.50	108.50	108.50
θ_0					
O-CH_2-CH_3		109.47	108	109.47	109.50
θ_0					
Dihedrals potential parameters [kJ/mol]					
A_1	3.4900	3.4900	3.4900	3.4900	3.4900
A_2	-0.4860	-0.4860	-0.4860	-0.4860	-0.4860

Table 2: Force field parameters for five used ethanol models

	Semi flexible OPLS	Fully flexible OPLS b.a.	Fully flexible OPLS s.a.	TraPPE	KBFF
A_3	3.1275	3.1275	3.1275	3.1275	3.1275
van der Waals potential parameters, ϵ [kJ/mol], σ [Å]					
H ϵ	0.00000	0.00000	0.00000	0.00000	0.08800
O ϵ	0.71131	0.71131	0.71131	0.77325	0.65060
CH_2 ϵ	0.49396	0.49396	0.49396	0.38247	0.41050
CH_3 ϵ	0.86612	0.86612	0.86612	0.81482	0.86720
H σ	0.00000	0.00000	0.00000	0.00000	1.58000
O σ	3.07100	3.07100	3.07100	3.02000	3.19200
CH_2 σ	3.90500	3.90500	3.90500	3.95000	4.07000
CH_3 σ	3.77500	3.77500	3.77500	3.75000	3.74800

TraPPE model differs from OPLS in the way that it has slightly longer bond length between CH_2 and CH_3 sites. The charge of the sites is the same, while Lennard-Jones parameters are different; oxygen site is smaller and has deeper ϵ than OPLS, allowing H-bonded sites to come closer, while CH_2 site is bigger with shallower potential well (smaller ϵ). KBFF model differs significantly from OPLSes and TraPPE in charges on sites and in Lennard-Jones parameters.

2.1.3 Water models

B. Guillot [80] listed 46 distinct models for liquid water, each of them in accordance with a subset of water properties, but none of them successful in reproducing the complete set of thermodynamic, structural, anomalous properties of water. A detailed review of water models is available in [15]. In this work Single Point Charge/Extended, SPC/E, model [81] was used. It is a simple, three-site, effective rigid pair potential model, composed of

Lennard-Jones and Coulombic terms [82], that does not use a lot of CPU time and is reasonably good in reproducing structural properties, internal energy, density and diffusivity of liquid water under ambient conditions [83]. Another model; Transferable Intermolecular potential with 4 Points, TIP4P, from [84], was also used in few cases as a test of dependance of some results on water model. The list of used parameters for SPC/E model is in Table 3 and for TIP4P in Table11. Illustration 7 illustrates water models. SPC/E is 'a' type, while TIP4P is 'c' type.

Table 3: Force field parameters for SPC/E water model

	Ow	Hw	Hw
Mass [g/mol]	16	1	1
Charge [e]	-0.8476	0.4238	0.4238
ϵ[kJ/mol]	0.65036	0.00000	0.00000
σ[Å]	3.16500	0	0
Length constraints between sites [Å]			
Ow-Hw1	1		
Ow-Hw2	1		
Hw1-Hw2	1.633		

Table 4: Force field parameters for TIP4P water model

	Ow	Hw	Hw	Mw
Mass [g/mol]	16	1	1	0
Charge [e]	0.0000	0.5200	0.5200	-1.0400
ϵ[kJ/mol]	0.64870	0.00000	0.00000	0.00000
σ[Å]	3.15365	0.00000	0.00000	0.00000
Length constraints between sites [Å]				
Ow-Hw1	0.9572			
Ow-Hw2	0.9572			
Ow-Mw	0.1500			

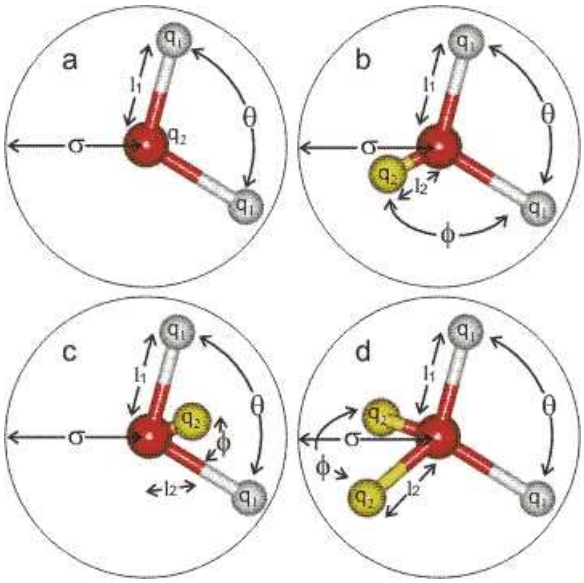


Illustration 7: Different water models, SPC/E is 'a' type [15]

2.2 Methods

2.2.1 DL POLY 2 and DL POLY Classic 1.4 software

The method of investigation was Molecular Dynamics computer simulation. The calculations were conducted on the Isabella cluster at SRCE, by DL POLY 2.14 [85] and DL POLY Classic 1.4 [86], [87] packages. With the chosen values of size parameters, the investigated system falls within permitted limits of both packages. This was an opportunity to test both versions on the same set of data, to ensure their compatibility, as in future research bigger systems are going to be needed. The DL POLY 2 uses kbar as a pressure unit, while DL POLY Classic uses katm as a pressure unit. Both packages use Leapfrog Verlet as a default integration algorithm. Four different systems were tested, test file number 10 from the DL POLY 2.14 package, Lennard-Jonnes liquid, SPC/E water and OPLS Fully flexible s.a., third model from this work. The difference in integration algorithm was tested on all four test systems, and the difference in kbar vs katm was tested on ethanol system. There was found to be no difference at all due to the integration algorithm used, and only insignificant difference in the value of the pressure due to the difference in the default pressure units. In the NpT ensemble, pressure is the quantity that fluctuates the most. It was found that all four test systems have slightly different calculated absolute values of energies and volumes from different DL POLY versions. In the molar values, these differences will be reasonably small to allow the conclusion that both versions are compatible. However, it is not clear what is the root of noticed differences in the results. The decrement in the speed of the calculations with the increase in the number of computer cores was exponential, and the biggest gain was found to be for the ethanol system, smaller for the water system and the smallest for the Lennard-Jones liquid.

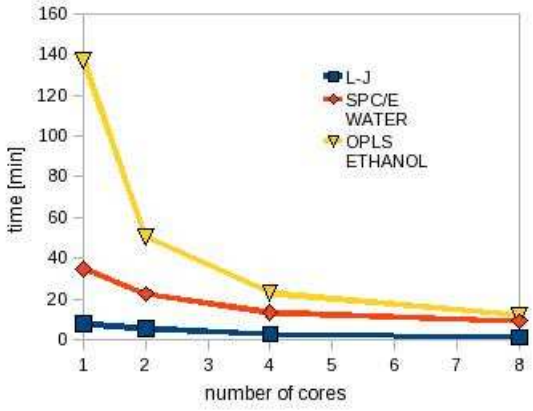


Figure 2: Cpu time for 500 steps runs with the respect to the number of computer cores

Test results are listed in Table 5.
In the case of ethanol it can be seen that the absolute value of the van der Waals energy differs less between two different versions, 2 and Classic, than between two builds of the same version of the application, 2.14 and 2.16. The same applies for the total configurational energy. The corresponding molar configurational energies are 36.9586 kJ/mol for DL POLY 2.14, 36.9214 kJ/mol for DL POLY 2.16 and 36.9648 kJ/mol for DL POLY Classic.

Table 5: Results from the compatibility testing of the DL POLY 2 and Classic ver. Energies are expressed in kJ, and volumes in $Å^3$.

1. TEST10 from DLPOLY2.14 package

DLPOLY version	temp_tot	eng_cfg	eng_vdw	eng_cou	eng_bnd	eng_ang	eng_dih	volume	press
2.14 test file	584.45	-9261.1	1258.8	-14379	47.67	3189	622.01	50225	3.121
2.14 on SRCE	584.45	-9261.1	1258.8	-14379	47.67	3189	622.01	50225	3.121
2.16 on laptop	584.45	-9261.1	1258.8	-14379	47.67	3189	622.01	50225	3.122
Classic 8cores	596.15	-9391.9	1264.8	-14547	42.28	3232.1	615.8	50225	2.927
Classic 4cores	596.15	-9391.9	1264.8	-14547	42.28	3232.1	615.8	50225	2.927
Classic 2cores	596.15	-9391.9	1264.8	-14547	42.28	3232.1	615.8	50225	2.927
Classic 8cores leapfrog	596.15	-9391.9	1264.8	-14547	42.28	3232.1	615.8	50225	2.927

2. Lennard-Jones liquid; 500 steps; 4000 particles

DLPOLY version	temp_tot	eng_vdw	volume	press
2.14 on SRCE	100.06	-22284	679890	-0.002473
2.16 on laptop	100.06	-22284	679890	-0.002473
Classic 8cores	99.99	-22258	680070	0.001090
Classic 4cores	99.99	-22258	680070	0.001090
Classic 2cores	99.99	-22258	680070	0.001090
Classic 8cores leapfrog	99.99	-22258	680070	0.001090

3. SPC/E WATER; 500 steps; 2048 molecules

DLPOLY version	temp_tot	eng_cfg	eng_vdw	eng_cou	volume	press
2.14 on SRCE	300.23	-95560	18045	-113600	61256	0.013
2.16 on laptop	300.27	-95560	18049	-113610	61256	0.017
Classic 8cores	299.81	-95499	17978	-113480	61250	-0.014
Classic 4cores	299.81	-95499	17978	-113480	61250	-0.014
Classic 2cores	299.81	-95499	17978	-113480	61250	-0.014
Classic 8cores leapfrog	299.81	-95499	17978	-113480	61250	-0.014

4. ETHANOL 500 steps; 500 steps; 2048 molecules

DLPOLY version	temp_tot	eng_cfg	eng_vdw	eng_cou	eng_ang	eng_dih	volume	press
2.14 on SRCE	299.94	-75691	-25009	-60296	5454.1	4160.5	195530	-0.009
2.16 on laptop	300.06	-75615	-24989	-60254	5467.1	4160.7	195850	0.023
Classic 8cores	300.17	-75704	-24999	-60347	5434.3	4207.1	195960	0.007
Classic 4cores	300.17	-75704	-24999	-60347	5434.3	4207.1	195960	0.007
Classic 2cores	300.17	-75704	-24999	-60347	5434.3	4207.1	195960	0.007
Classic 8cores leapfrog	300.17	-75704	-24999	-60347	5434.3	4207.1	195960	0.007

The difference between 2 and Classic versions is 0.005 kJ/mol, and it is less than the difference in some of the reported experimental enthalpy of vaporization values as from 42.30 kJ/mol in [33] to 42.47 kJ/mol in [56] of 0.17 kJ/mol.

OPLS fully flexible models were calculated on DL POLY 2.14 version and OPLS semi flexible, TraPPE, and KBFF model on DL POLY Classic version of the application.

2.2.2 Molecular Dynamic simulation

The essence of Molecular Dynamic simulation is simply stated: it numerically solves classical equations of motion for N-body system [88]. That was accomplished for the first time by Adler and Wainwright in 1957 [89] -1959 [90] for the system of hard spheres. In that case, particles move at constant velocity between perfectly elastic collisions, and it is possible to solve dynamic problem without making any approximations, within the limits imposed by machine accuracy [79]. Several years later Rahman for the first time solved equations of motion for the Lennard-Jones system [91]. After that, computer simulations developed rapidly. The first step in a computer simulation is when a real system is reduced to a mathematical model, and then this model is solved by the use of the computer. When using computer simulation, one must be aware that usage of computer to compute the behaviour of the many-body model does not absolutely guarantee that the computed behaviour is representative of that model, and much less that the model mimics reality. Result from a simulation is at the first place the test of the underlying model, used in a computer simulation. Eventually, if a model is a good one, simulation results can be compared with the experimental data, and assist experimentalist in the interpretation of the new results. Nevertheless, computer simulations are one of the most important tools in the study of liquid systems [92]. It provides a direct route from microscopic details of a system (masses of atoms, interactions between them, molecular geometry) to macroscopic properties of experimental interest [79].
Molecular dynamics is a numerical realization of the system's unfolding trajectory in phase space. The unfolding is governed by the classical equations of motion. For the equilibrium MD, the system is confined to the hypersurface of constant Hamiltonian in phase space, and if the system is isolated, the Hamiltonian is the total energy. Once the MD simulation has generated a phase space trajectory, it serves as raw data for obtaining time averages of the properties. This time averages distinguish MD from other form of simulation, Monte Carlo, that computes ensemble averages, and from the formal statistical mechanics. In order to provide reliable time averages, a

MD simulation must generate trajectories in a way that the sampling of the constant-Hamiltonian surface is apparently random. The phase space trajectory in MD is deterministic, and not random, because successive points are determined by solving classical equations of motion. The way to make this deterministic trajectory apparently random is by the mean of molecular interactions. After only a few interactions positions and velocities are essentially unrelated to their earlier states: the correlation time is short, so consequently, after a short time Maxwell distribution of velocities develop and the time averages for properties can be computed.

2.2.3 Treatment of the boundary conditions

MD is typically applied to the systems consisting of a few thousands of molecules. In order to mimic the real bulk liquid, periodic boundary conditions are used, to overcome the surface effects that would be dominant in such small systems. The cubic box is replicated all over the space to form an infinite lattice. As a particle moves in the primary box, its images move in the surrounding boxes in exactly the same way. If a particle leaves the central box, one of its images will enter it from the opposite wall of the box. A two-dimensional illustration of the periodic boundary conditions is shown in the Illustration 8.

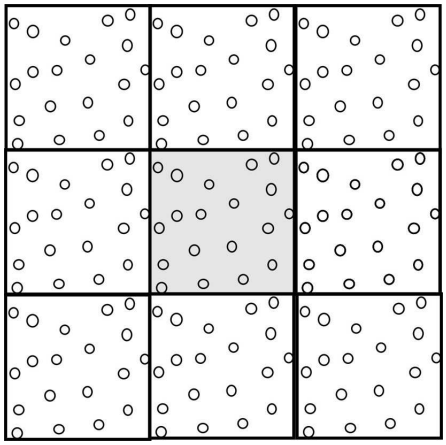


Illustration 8: 2D periodic boundary conditions

2.2.4 Treatment of the long-range forces: Ewald summation

A long ranged force is often defined as one in which the spatial interaction falls off no faster than r^{-d} , where d is the dimensionality of the system. The Coulomb interaction between charged sites falls inevitably in this category, as it has dependence r^{-1}. This force poses a serious problem to the simulator [79], as it's range if far greater than the half of the simulation cell, where the other forces are truncated. This interaction was treated with Ewald sum method [93]-[94] in OPLS models and with its improved variant, Smoothed Particle Mesh Ewald (SPME) [95], in TraPPE model. In short, Ewald sum procedure goes as follows: original charge distribution of point charges described by delta functions is screened by superimposing an equal in magnitude and opposite in charge Gaussian charge distribution over every point charge. The form of the Gaussian distribution is:

$$\rho_G(\vec{r}) = \frac{1}{(2\pi)^{3/2}\sigma^3} e^{\frac{-|\vec{r}|^{-2}}{2}\sigma^{-2}} \tag{2.8}$$

where σ is the standard deviation of the Gaussian distribution. Parameter used in the DL POLY is $\alpha = 1/(\sqrt{2}\sigma)$. Limit of this distribution when $\sigma \to 0$ is the delta function of a point charge. This extra distribution acts as an ionic atmosphere, screening the interactions between neighbouring charges. The screened interactions are now short ranged, and the total potential is calculated by summing over all of the sites in the central cell and their images in the real space. The potential field of Gaussian distribution can be obtained by solving Poisson's equation:

$$\nabla^2\phi_\sigma(\vec{r}) = -\frac{\rho_G(\vec{r})}{\epsilon_0} \tag{2.9}$$

and the total Coulomb interaction energy is

$$U_{Coulomb} = \frac{1}{2}\sum_{i=1}^{N} q_i\phi_{[i]}(\vec{r_i}) \tag{2.10}$$

where $\phi_{[i]}(\vec{r_i})$ is the potential field generated by all of the sites plus their images, excluding site i.

One more canceling Gaussian distribution is then added to cancel the first one, exactly the same, but opposite in charge to the first. This canceling distribution is summed up in the reciprocal space. The complete Ewald sum requires an additional correction, known as the self energy correction, which arises from a Gaussian that is acting on its own site, and that is constant. Ewalds method, therefore, replaces a potentially infinite sum in real space by two finite sums: one in real space and one in reciprocal space; and the self energy correction. There is one more necessary correction for the molecular systems, as the intramolecular Coulomb interactions must be excluded from the sums. In a bit simplified form, total Coulomb interaction energy can be written as [96]:

$$\begin{aligned} U_{Coulomb} &= U^{shortr.} + U^{longr.} - U^{self} \\ &= \frac{1}{4\pi\epsilon_0}\frac{1}{2}\sum_{\vec{n}}\sum_{i=1}^{N}\sum_{j(j\neq i)=1}^{N}\frac{q_1 q_j}{|\vec{r_i} - \vec{r_j} + n\vec{L}|} erfc(\frac{|\vec{r_i} - \vec{r_j} + n\vec{L}|}{\sqrt{2}\sigma}) \\ &+ \frac{1}{2V\epsilon_0}\sum_{\vec{k}\neq 0}(e^{-\sigma^2 k^2/2}|S(\vec{k})|^2) \\ &- \frac{1}{4\pi\epsilon_0}\frac{1}{\sqrt{2\pi}\sigma}\sum_{i=1}^{N} q_i^2 \end{aligned} \tag{2.11}$$

The summation for U^{shortr} is short ranged in real space, truncated by the erfc function and the summation for U^{longr} is short ranged in reciprocal space, truncated by the exponential factor. In practice the convergence of the Ewald sum is controlled by three variables: the real space cutoff rcut, the convergence parameter α and the largest reciprocal space vector $\vec{k_{max}}$ used in the reciprocal space sum. DL POLY option that ensures automatic calculation of the and $\vec{k_{max}}$ parameters, with the precision set to 106. Smooth particle mesh Ewald method's main difference from the Ewald sum is in its treatment of the reciprocal space terms. By means of an interpolation procedure involving basis spline functions[3], the sum in the reciprocal space

[3]Basis spline (B-spline) is a spline function that has minimal support with respect

is represented on a three dimensional rectangular grid. In this form the Fast Fourier Transform (FFT) may be used to perform the primary mathematical operation, which is a 3D convolution. While computing Fourier's transform onf N points using standard way is an algorithm of order of N^2 the FFT is of the order of N(logN) [97].The efficiency of these algorithm greatly reduces the time cost of the performing the reciprocal space sum.

2.2.5 MD algorithm

With all the potentials provided, the MD simulation application starts to solve classical equation of motion $F = ma$, as $F = -gradU$, with the initial conditions provided in the input configurational file. Basically, there are two repetitive steps in MD calculation; first the calculation of total force on all of the particles at given time t, and second the advancement of the coordinates of all particles at time t + time step [98]. The Leapfrog Verlet integration algorithm was used. As previously shown, Leapfrog and Velocity Verlet produced exactly the the same results on DL POLY Clssic, as expected, because MD should give equal results regardless of the way of obtaining them.

In the Leapfrog algorithm, velocity is half-step ahead of positions and forces (there the name came from). First velocity at half time step $(t + \Delta t)$ is calculated:

$$\vec{v}(t + \frac{1}{2}\Delta t) = \vec{v}(t - \frac{1}{2}\Delta t) + \vec{a}(t)\Delta t \tag{2.12}$$

to a given degree, smoothness, and domain partition. Spline is a sufficiently smooth piecewise-polynomial function. In interpolating problems, spline interpolation is often referred to as polinomial interpolation. A piecewise-defined function (also called a piecewise function) is a function whose definition changes depending on the value of the independent variable. A fundamental theorem states that every spline function of a given degree, smoothness, and domain partition, can be uniquely represented as a linear combination of B-splines of that same degree and smoothness, and over that same partition.

and then the position advances for the full time step:

$$\vec{r}(t+\Delta t) = \vec{r}(t) + \vec{v}(t+\frac{1}{2}\Delta t)\Delta t \tag{2.13}$$

During this step the current velocities are calculated, as they are needed so that energy at time t can be calculated:

$$\vec{v}(t) = \frac{1}{2}(\vec{v}(t+\frac{1}{2}\Delta t) + \vec{v}(t-\frac{1}{2}\Delta t)) \tag{2.14}$$

Velocity Verlet algorithm goes as follows:
It stores values of position, velocity and acceleration at time t.
1. positions at $(t+\Delta t)$ are calculated.

$$\vec{r}(t+\Delta t) = \vec{r}(t) + \vec{v}(t)\Delta t + \frac{\vec{a}(t)}{2}(\Delta t)^2 \tag{2.15}$$

2. velocities at $(t+1/2\Delta t)$ mid step are calculated

$$\vec{v}(t+\frac{1}{2}\Delta t) = \vec{v}(t) + \frac{\vec{a}(t)}{2}(\Delta t) \tag{2.16}$$

3. forces and accelerations at $(t+\Delta t)$ are computed
4. velocity move completed

$$\vec{v}(t+\Delta t) = \vec{v}(t+\frac{1}{2}\Delta t) + \frac{1}{2}\vec{a}(t+\Delta t)\Delta t \tag{2.17}$$

For the treatment of the bond length constraints SHAKE algorithm was used with the leapfrog Verlet integration algorithm. SHAKE algorithm goes as follows[99]:

1. For the treatment of the bond length constraints SHAKE algorithm was used with the leapfrog Verlet integration algorithm. SHAKE algorithm goes as follows[99]:atoms in the system are moved using the Leapfrog Verlet algorithm, assuming an absence of rigid bonds (constraint forces). (This is stage one of the SHAKE algorithm.)

2. The deviation in each bond length is used to calculate the corresponding

constraint force that (retrospectively) corrects the bond length.

3. After the correction has been applied to all bonds, every bond length is checked. If the largest deviation found exceeds the desired tolerance, the correction calculation is repeated.

4. 1.Steps 2 and 3 are repeated until all bond lengths satisfy the convergence criterion (this iteration constitutes stage 2 of the SHAKE algorithm).

2.2.6 Ensemble and settings of general MD parameters

All of the simulations were done in the NpT ensemble, as it corresponds to the realistic experimental conditions of the liquid in an open tube at atmospheric temperature and pressure. Brendsen thermostat and barostat were used to fix the pressure at the atmospheric value of 1 atm, and the temperature at T=300 K, with thermostat and barostat relaxation times of 0.1 ps and 0.5 ps respectively. The timestep for each evaluation of the equations of motion was t=2 fs. The timestep must be chosen in a way for it to be smaller than the timestep of the dynamics of the system. Librations of the water molecule are at ps scale [59]. The stretching of the H-bond is on the picosecond scale [20], too. It means that it was safe to use the 2 fs timestep for the simulation. Data for the site-site correlation functions were gathered every 20 steps (40 fs). Short range interactions (van der Waals) cutoff was set to 15 Å. The width of the border to be used in the Verlet neighbour list construction was set to 15 Å for all the mole fractions. The Verlet neighbour list is updated whenever two particles move more than half of the width of the border from their previous positions at the last update of the Verlet list.
In all the fully flexible models valence angle potential, E_{angle}, was evaluated as an angle restraint potential (it is denoted by: '-hrm' in DL POLY), in which the angle subtended by a triplet of atoms, maintained around some preset value θ_0 , is handled as a special case of angle potentials. As a consequence angle restraints may be applied only between atoms in the same molecule. Unlike with application of pure angle potentials (denoted by: 'harm' in DL POLY), electrostatic and van der Waals interactions between the pair of

atoms are still evaluated when distance restraints are applied [99].

2.2.7 Thermostat and barostat

In order to conserve the temperature and pressure at the atmospheric conditions (300 K, 1 atm), Berendsen thermostat (weak-coupling method) and barostat [100] were used. With the introduction of the heat bath and thermostat, the system no longer samples the microcannonical ensemble (NVE), as the microcannonical ensemble does not correspond to the conditions under which experiments are taken out. Only the Nose-Hoover algorithm (extended system method) [101] generates trajectories in the canonical (NVT) ensemble, while the other thermostats will produce properties that typically differ from canonical averages by the order of 1/N. In the Berendsen algorithm the instantaneous temperature of the system is scaled at each step by scaling the velocities of the particles in the system by $\chi(t) = \sqrt{1 + \frac{\Delta t}{\tau_\beta}(\frac{T_0}{T(t)} - 1)}$ as the instantaneous temperature is $T(t) = \frac{2E_kin}{K_b N_f}$, where N_f is the number of degrees of freedom in the system, $N_f = 3N - N_{constraints} - 3$ [102]. This scaling is the last step in the Leapfrog algorithm, after the full step velocities are obtained. It is done in a few iterations. τ_β is used as an empirical parameter to adjust the strength of the coupling. Its value should be chosen in a appropriate range. Too large value (loose coupling) may cause a systematic temperature drift. In the limit $\tau_\beta \to \infty$, the Berendsen thermostat is inactive leading to the MD equation of motion, which samples a microcanonical ensemble. On the other hand, a too small value (tight coupling) will cause unrealistically low temperature fluctuations. In the lowest limit it will sample canonic (NVT) or Gibss (NpT) ensemble. All the values in the middle does not sample any proper ensemble. In this work value of 0.1 ps was chosen, as it is usually done in the MD simulations of liquids.

2.2.8 Runs lengths

Equilibration and production runs lengths, for each ethanol model used, are listed in the Table 6

Table 6: Equilibration and production runs lengths for each ethanol model

Ethanol mole fraction	Equilibration length [ps]	Production length [ps]
1. Semi flexible OPLS		
∀ mole fractions	6*1000	1000
2. Fully flexible OPLS (big angle)		
0.05	6*128	128
0.08	6*128	128
0.10	6*128	128
0.12	6*128	128
0.15	6*128	128
0.18	7*128	128
0.20	7*128	128
0.30	6*128	128
0.40	6*128	128
0.50	4*128	128
0.60	6*128	128
0.70	5*128	128
0.80	5*128	128
0.90	4*128	128
1.00	5*128	400
3. Fully flexible OPLS (small angle)		
∀ mole fractions	1000	1000
4. TraPPE		
∀ mole fractions	1000	1000
5. KBFF		
∀ mole fractions	1000	1000

3 Results and discussion

3.1 Thermodynamical results

3.1.1 Test for equilibrium

Before reporting any of the results, the test was taken to check if the systems really reached an equilibrium state before the production run was set on. To confirm this, the configurational energies on Figure 3- 4 and the volumes on Figure 5- 6 of the system were plotted against time, over the duration of the production runs. It is obvious that the values of the configurational energies and the volumes at all concentrations show no drift. This confirms that the system has reached an equilibrium. The results for the TraPPE model are shown, but the test was done for all the models, and it confirmed that all of them reached equilibrium.

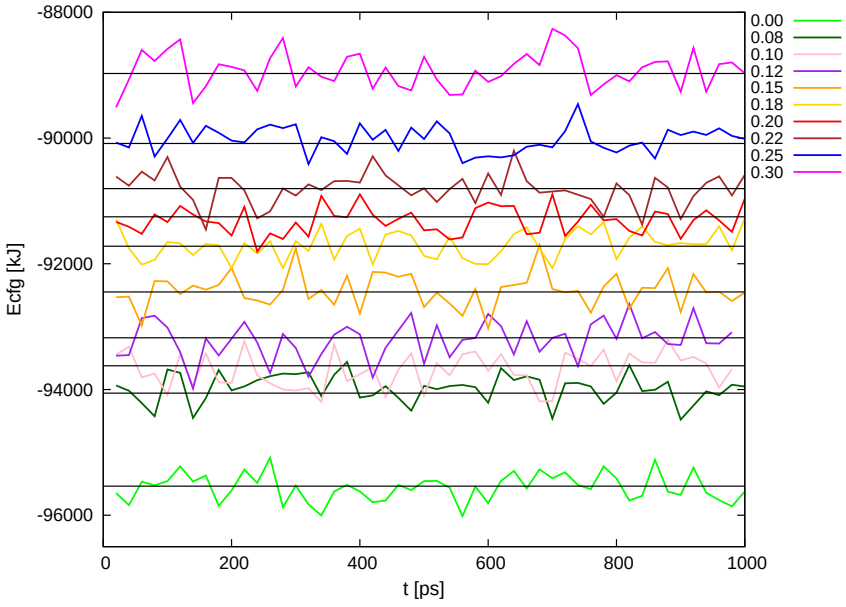


Figure 3: Configurational energies during the production run for the TraPPE-SPC/E system, in the range from 0.0 to 0.3 mole fractions of ethanol

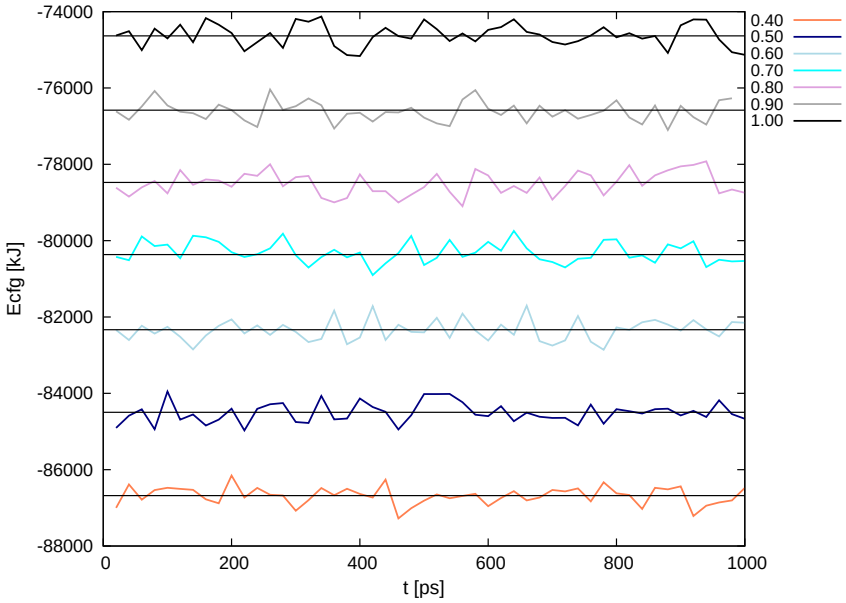


Figure 4: Configurational energies during the production run for the TraPPE-SPC/E system, in the range from 0.4 to 1.0 mole fractions of ethanol

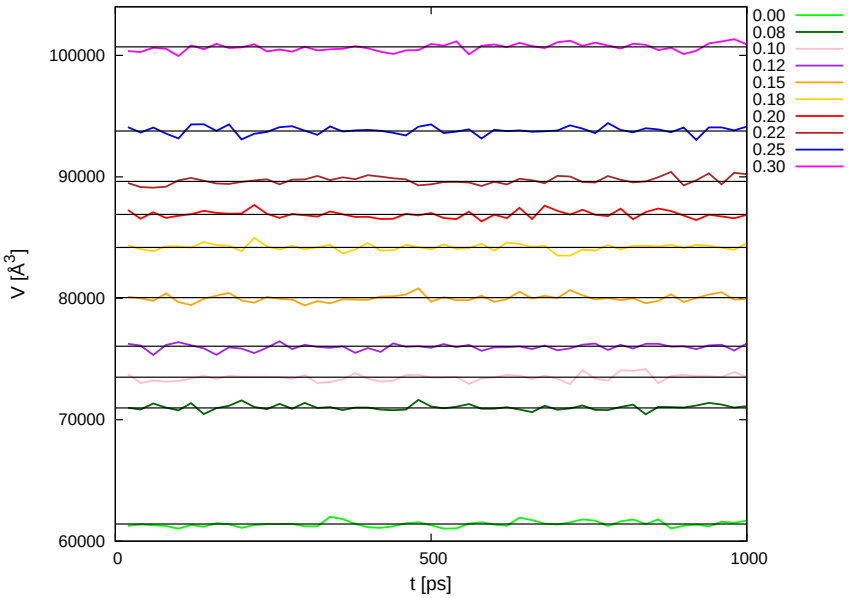


Figure 5: System volumes during the production run for the TraPPE-SPC/E system, in the range from 0.0 to 0.3 mole fractions of ethanol

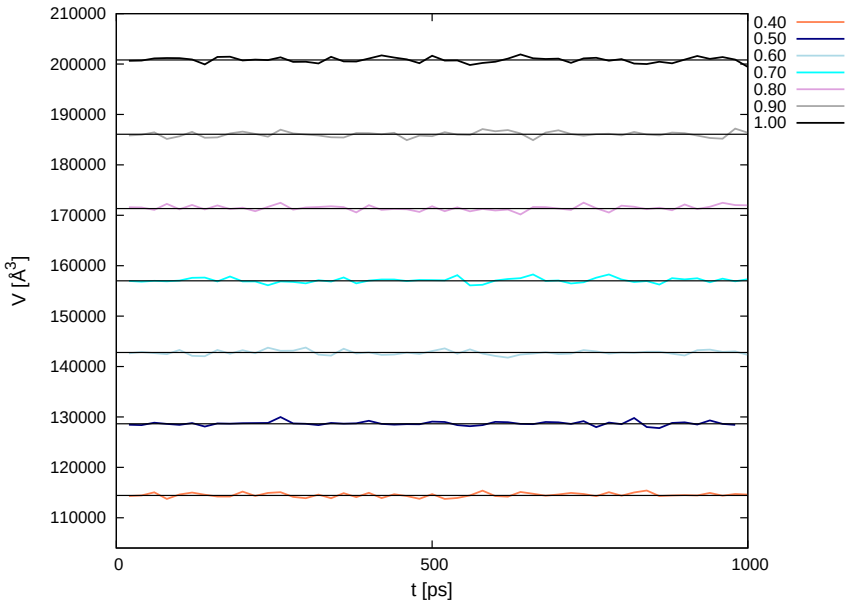


Figure 6: System volumes during the production run for the TraPPE-SPC/E system, in the range from 0.4 to 1.0 mole fractions of ethanol

3.1.2 Enthalpy of vaporization

Molar enthalpy of vaporization, ΔH_{vap}, is the heat of vaporization, the energy required to transform one mole of the liquid to vapour phase at a given (often atmospheric) pressure [98][30]. For the neat component, it tests the accuracy of the used model, when compared with the experimental value.

$$\Delta H_{vap} = H_{vapour} - H_{liquid} \tag{3.1}$$

$$\Delta H_{vap} = U_{vapour} + pV_{vapour} - U_{liquid} - pV_{liquid} \tag{3.2}$$

With U =internal energy and $pV_{vapour} = RT >> pV_{liquid}$, for the ideal gas limit.

$$\Delta H_{vap} = U_{vapour} + RT - U_{liquid} \tag{3.3}$$

With $U = U_{kin} + U_{pot}$, and $U_{kin_{vapour}} = U_{kin_{liquid}}$ at $T = const.$

$$\Delta H_{vap} = U_{pot_{vapour}} + RT - U_{pot_{liquid}} \tag{3.4}$$

Potential energy has an intramolecular E_{intra} part coming from the torsion and angle flexibility of the potential and intermolecular part E_i , (bonded and nonbonded part), so:[50]

$$\Delta H_{vap} = E_{intra_{vapour}} - E_{intra_{liquid}} - E_{i_{liquid}} + RT - (H^0 - H) \qquad (3.5)$$

Where $(H^0 - H)$ part comes from the difference of the enthalpy between real and ideal gas, as the nonbonded part in vapour = 0 for ideal gas [30][50]. From [30] data for the pure ethanol it is justified to set $E_{intra_{vapour}} - E_{intra_{liquid}} \approx 0$, as Jorgensen calculated $E_{intra_{vapour}} = 2.077\ kJ/mol$ and $E_{intra_{liquid}} = 2.035\ kJ/mol$. $RT = 2.494 kJ/mol$ and $(H^0 - H) = 0.251 kJ/mol$ was taken from [30].

Experimental data and references for the enthalpy of vaporization for pure ethanol are shown in Table 7

Table 7: Experimental enthalpy of vaporization for pure ethanol, in kJ/mol

Reference	ΔH_{vap}
[56]	42.47 kJ/mol at $T = 298K$
[103]	42.3 kJ/mol at $T = 298.15K$
[33]	42.3 kJ/mol at $T = 298.15K$
[104]	42.297 kJ/mol at $T = 300K$

Table 8: Molar enthalpy of vaporization for pure ethanol from other simulations, in kJ/mol

Model	reference	system size	duration	ΔH_{vap}
Jorgensen's semi flexible OPLS at 298K	[30]	128 molecules	MC $2x10^6$ steps	41.83
Jorgensen's semi flexible OPLS at 298K	[50]	125 molecules	MD 250 ps	42.80
TraPPE at 300K	[40]	300 molecules	MC 50000 steps	43.3
Jorgensen's semi flexible OPLS at 298K	[53]	391 molecules	2.2 ns	41.43
fCINTRA polarizable flexible at 298K	[49]	200 molecules	MD	45.77
PIPF at 298K	[47]			42.17
polarizable with Durde particle at 300K	[48]	250 molecules	6 ns	42.63
polarizable OPLS at 298K	[105]	216 molecules	MD 100 ps	40.70

Results for the enthalpy of vaporization for pure ethanol, for different models from this work are given in the Table 9 :

Table 9: Enthalpy of vaporization for pure ethanol from this work, in kJ/mol

Model	$-E_i$	ΔH_{vap}
Semi flexible OPLS	41.3	43.5 ± 0.1
Fully flexible OPLS b.a.	41.9	44.2 ± 0.1
Fully flexible OPLS s.a.	41.5	43.7 ± 0.1
TraPPE	41.1	43.3 ± 0.1
KBFF	47.5	49.8 ± 0.1

The standard deviation σ of 0.1 kJ/mol in Table 9 was calculated in a way that RT and $(H^0 - H)$ were taken as constants, and the standard deviation of enthalpy of vaporization was evaluated from:

$$\sigma = \sqrt{(\sigma_{config})^2 + (\sigma_{angle})^2 + (\sigma_{dihedral})^2} \tag{3.6}$$

It was assumed that $E_{intra_{vapour}} - E_{intra_{liquid}} \approx 0$, because this contribution is the order of magnitude smaller than the standard deviation of 0.1 kJ/mol. The TraPPE model gives exactly the same value as the reference [40], from which the parameters for the model were taken, namely the Monte Carlo simulations in the Gibbs (canonical, NVT) and grand-canonical (μVT) ensembles. OPLS models slightly overestimate enthalpy of vaporization for pure ethanol. On the other hand, the KBFF model largely overestimates it. When looking at the data from Table 8 it is curious that simulations with smaller systems and for shorter runs give better results and there are no references in literature of systems this big, and runs this long.[4]

Experimental enthalpy of vaporization for water at 298 K is (43.99 ± 0.1) kJ/mol from [106]. Enthalpy of vaporization from the simulation was calculated as [107]:

$$\Delta H_{vap} = -E_{i_{liquid}} + RT + Q \tag{3.7}$$

[4]to the knowledge of the author

where $Q = -0.23 kJ/mol$ is the quantum correction for the vibrational energy of the water molecule in vapour and liquid phase, and corrections due to intermolecular interactions in the liquid [107]. Enthalpy of vaporization for water from this work is in Table 10 and is in excellent agreement with the experimental result[5].

Table 10: Molar enthalpy of vaporization for pure water, in kJ/mol

Model	$E_{i_{liquid}}$	$E_{i_{liquid}} + 5kJ/mol$	ΔH_{vap}
SPC/E	-46.6	-41.6	43.9 ± 0.1
TIP4P	-41.68		43.95 ± 0.09

Value of approximately 5 kJ/mol is the correction for the self-polarization energy in the liquid, that needed to be included when working with the SPC/E model for water [81][80].
The TIP4P model [84] was also calculated, as a test.
Out of various simulated models of water, here are just a few results for the enthalpy of vaporization of water (or internal energy): Van der Spoel [108] reports $E_{i_{liquid}}$ of -47.2 and -46.4 kJ/mol at 301 K, and Mark and Nilsson [82] of-45.4 kJ/mol at 298 K. Kiss et al. report it of -41.5 kJ/mol, with the correction for the self-polarization effects added. Glattli et al. [107] report the enthalpy of vaporization of 43.6 kJ/mol at 301 K.

3.1.3 Density

The density of liquid ethanol under ambient conditions is 0.789 g/cm^3 [1]. In this work, OPLS, TraPPE and KBFF models give density of ethanol of $0.8 \pm 0.1 g/cm^3$, and density of water of $1 \pm 0.07 g/cm^3$ from SPC/E model,

[5] $E_{i_{liquid}}$=-46.644±0.1 from DL POLY 2 package, and -46.649±0.1 from DL POLY 4 package, confirming that the difference in the packages is inconsequential.

and $1 \pm 0.06 g/cm^3$ from TIP4P. Density over the whole composition range is plotted on Figure 7.

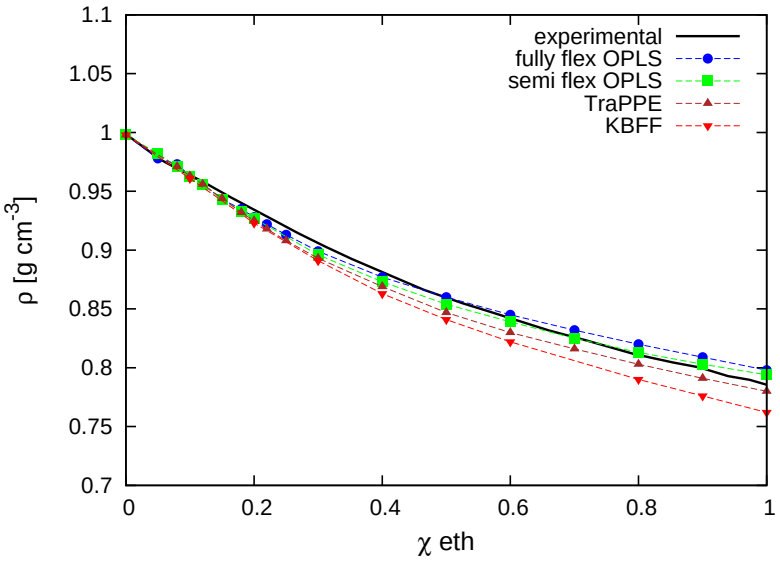


Figure 7: Densities from $\chi = 0$ (left) to the $\chi = 1$ mole fraction ethanol from this work. Line for experimental data from [1].

This is in good agreement with experimental data, even though OPLS models slightly overestimate the value for pure ethanol, and TraPEE model slightly underestimates it. KBFF model has a rising trend of underestimating densities for higher alcohol concentrations.

3.1.4 Excess volume

If two liquids would mix ideally, the volume of the mixture would be simply the sum of volumes of its constituents. That is the case when mixing Lennard-Jones atoms of the same size. Excess volume is the deviation from this ideal mixing, defined as:

$$V_{excess} = V_{mixture} - (\chi_A V_A + \chi_B V_B) \tag{3.8}$$

where $\chi_A + \chi_B = 1$, and A and B are two different components in the mixture. When mixing Lennard-Jones atoms of different size, excess volume is negative, and bigger when difference in size of atoms is bigger, as shown on Figure 8:

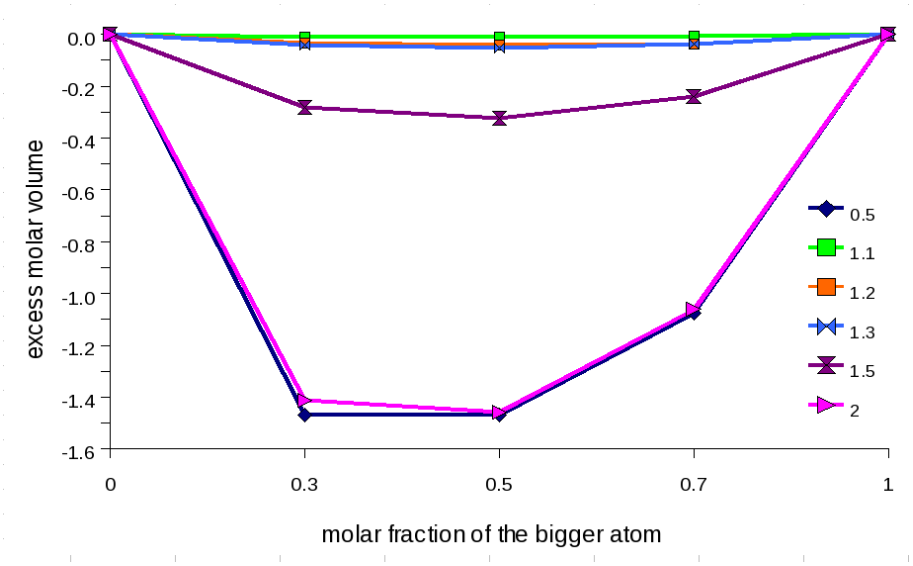


Figure 8: Excess volume for Lennard-Jones atoms, different lines for different size proportions, from [2]

Excess volumes from this work is plotted in Figure 9, compared to the experimental data. Experimental line on Figure 9 resembles the shape of the excess volume of Lennard-Jones mixtures on Figure 8, with the minimum shifting to the left from the $\chi = 0.50$. Excess volume has good reproducibility for all models. The best one, that follows the shape and size (even if it is slightly too negative), is the KBFF model. Other models have smaller excess value than the experimental data.

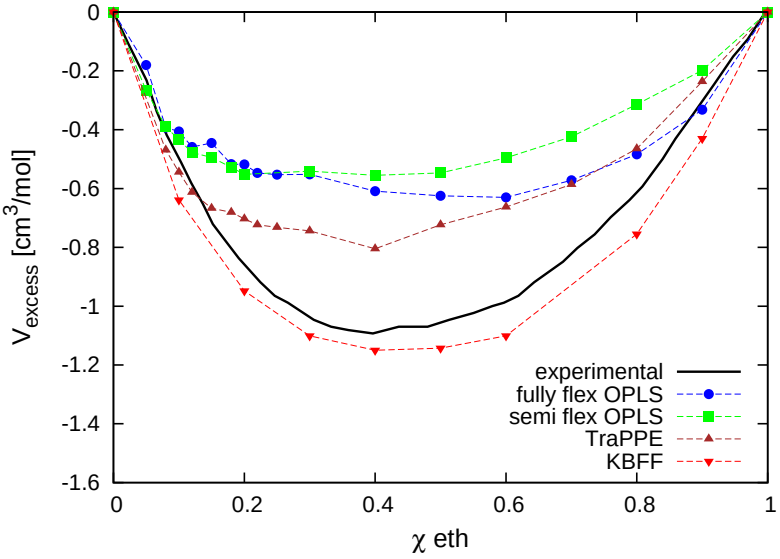


Figure 9: Excess volume, experimental data from [1] line

3.1.5 Excess enthalpy

When mixing two liquids, excess enthalpy, or the heat of mixing, is defined as the difference of the mixture's real enthalpy and the expected enthalpy of the ideal mixture. An ideal mixture has zero excess enthalpy, meaning that the intake of energy when breaking bonds in the liquid upon mixing and the release of energy when new bonds are formed are the same.

$$E_{excess} = E_{i_{mixture}} - (\chi_A E_{i_A} + \chi_W E_{i_W}) \tag{3.9}$$

On Figure 10 excess configurational energies are shown for binary mixtures of Lennard-Jones liquids with atoms of different size. The bigger difference in constituent's size means bigger excess energy in these simple mixtures. The shape of the excess energy is symmetric, it has a minimum at $\chi = 0.50$ mole fraction.

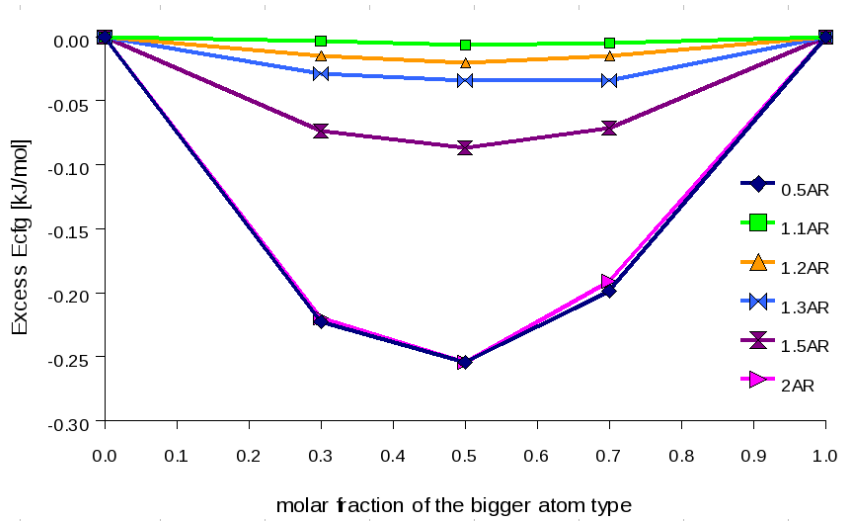


Figure 10: Excess of negative configurational energies for Lennard-Jones atoms, different lines for different size proportions, from [2]

In the case of the ethanol-water solution, to our knowledge, excess enthalpy simulation data were reported only five times, which is peculiar considering the vast number of simulation studies. On Figure 11 we show experimental values [1] and simulation data from: Muller-Plathe [36], 1996., with SPC water [37], 2003. Wensink et al. [53] all-atoms OPLS and TIP4P water [84], 2005 Zhang and Yang [58] rigid ethanol and TIP4P; 2011. Guevara-Carrion et al. [109] rigid united-atoms ethanol [45] with TIP4P/2005 [110] and SPC/E [81], and data from this groups work [6] on flexible OPLS and SPC/E water model.

The experimental curve for ethanol-water under ambient condition is negative, which is an indication that mixing is energetically favorable, however, it also has highly nontrivial mole fraction dependence. Minimum is shifted to the mole fraction of $\chi = 0.15$ that indicates the difference of this hydrogen bonded mixture to the simple Lennard-Jones mixture, where minimum is at the mole fraction of 0.50. Negative excess enthalpy can be partially attributed to the size effects, but the shift of the minimum to smaller mole fractions reveals the structural changes that take place in this system. This feature will be commented more in section 3.8, where experimental data are discussed more. The wide range of different behaviour of simulation data emphasizes

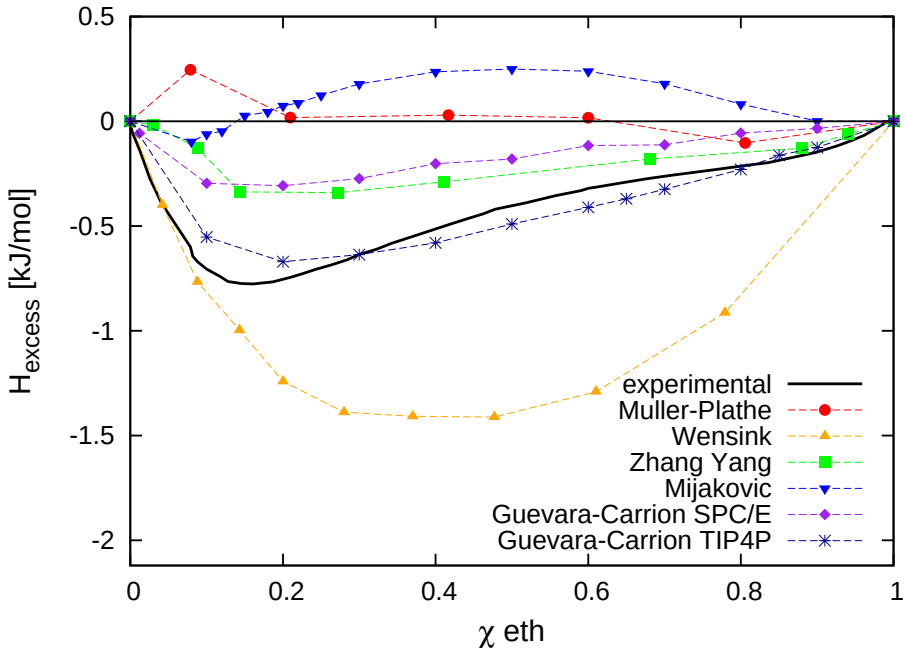


Figure 11: Excess enthalpies from previous works

the difficulty in reproducing the experimental results. Simulation data can be evaluated if they can reproduce negative excess, or if they can follow the overall trend (thought the shape of the curve or extremes and inflection points). If the negative excess is not reproduced, but the overall trend is, it can be argued that these models nevertheless capture, to some extent, mixing behaviour of the real system. On the other hand, if the trend is missed, even thought values are negative, that might be indication that fine restructuring along the composition range is not reproduced.

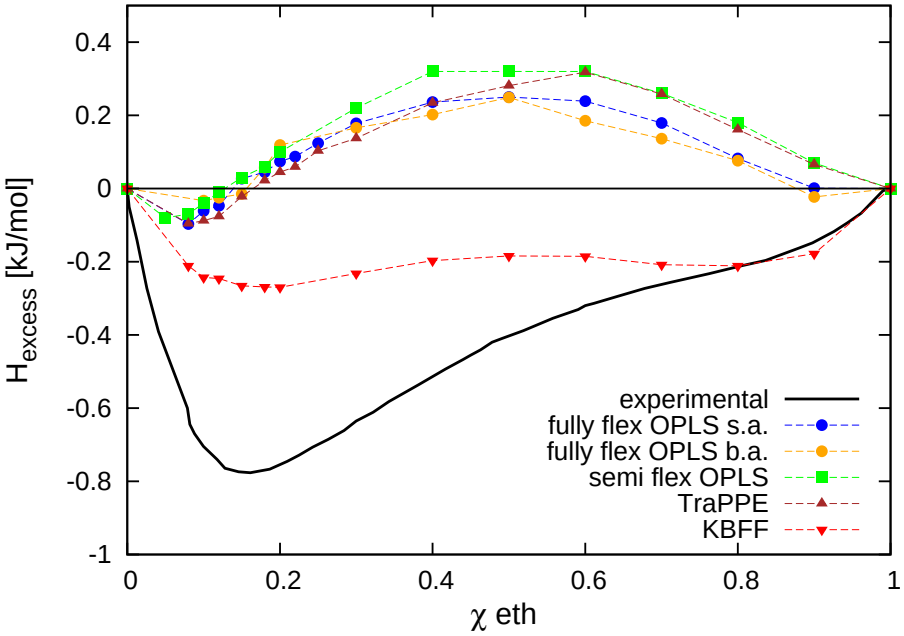


Figure 12: Excess enthalpies [kJ/mol] vs. mole fraction of ethanol from this work. Experimental data from [3]

In Figure 12, excess enthalpies for all the listed models from this work are shown, compared to the experimental data. Different ethanol models from this work all give the same problem with the excess enthalpy: it is evidently too high. KBFF model data are in much better agreement with experimental line (it is negative), than OPLS and TraPPE models, but excess enthalpy is still too high. It is not clear if the fact that KBFF model's line is still too high, is the consequence of the model, or of insufficient length of production runs. In this work, production runs lasted only 2 ns, while in the reference article for KBFF model simulation, [46], production runs were 30 ns long.
At first glance, data, apart from KBFF model, is far away from capturing behavior of ethanol-water mixture. However, they follow a certain trend, which is traceable to the one in the experimental curve. Namely, simulation plots can be divided into three regions. The first one, up to $0.3 - 0.4$ alcohol mole fraction, which includes a minimum and point of change of the sign; central part; and then from the change of the slope around $0.5 - 0.6$, the third region. This division is similar to the one that can be used to discuss

experimental curve. It is not an unique explanation, but we will use this connection to explore the subject further. The remaining question is: Why is the KBFF model able to reproduce excess enthalpies correctly, and other models fail in it? And the second one is: Are models that missed negative excess, suitable to represent real system properties?

Following the differences between partial charges, in Figure 13 we show separately the excess of vdW and Coulomb energy. All models show approximately symmetric shape of the curves. The central part of the Coulombic excess inclines towards the positive values, showing that excess results in a net repulsion. The vdW excess is similar for all models, it is negative, showing that vdW interactions enhance mixing. TraPPE model has the highest Coulombic and the lowest vdW, with the Coulombic maximum shifted towards higher alcohol concentrations. This corresponds to the maximum of excess enthalpy at $\chi = 0.60$ mole fraction of ethanol for TraPPE model.

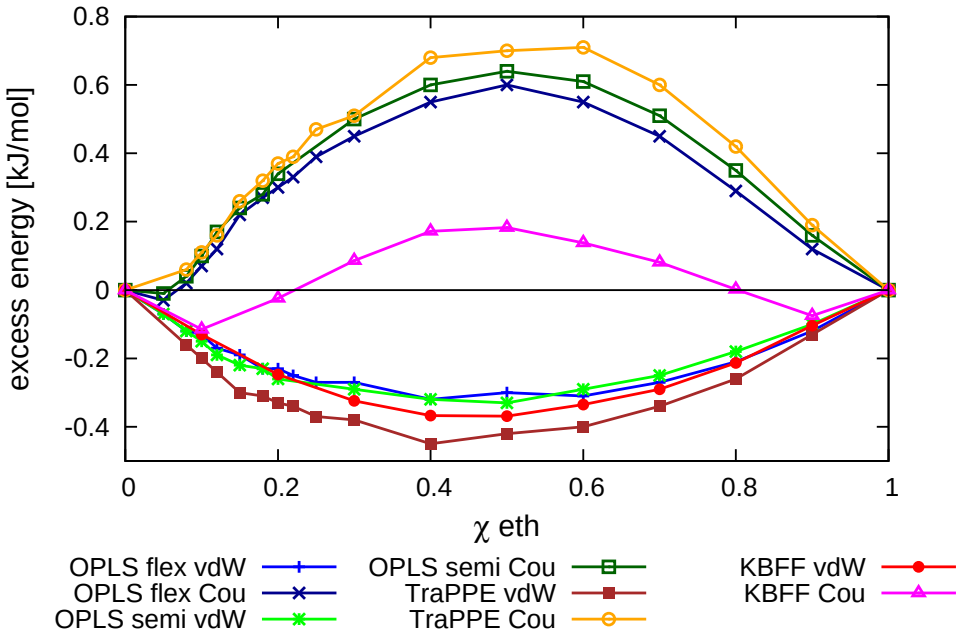


Figure 13: Excess van der Waals and Coulomb energy

3.1.6 Comparison of SPC/E and TIP4P results

A study by van der Spoel et al. [108] confirmed that SPC/E water model gives the best agreement with experimental results for bulk water, but suggests that it is not so relevant in studies of solutes in water. Tieleman and Berendsen [111] reported that the SPC/E has dubious results in studies of solvated biological membranes. Neither SPC [37] nor SPC/E models preform good with biological membranes, but they reported that the usage of SPC/E leads to sharper interface and lower area per head group, all in disagreement with experimental data. Kiss and Baranyai [83] in 2011 stated that the many-body structure of real liquid water is more similar to models created by TIP4P, than to models created by SPC/E or TIP3P. Along these lines, the mixture of the TIP4P [84] water model and the semi flexible OPLS ethanol model was used to simulate the mixture at three mole fractions of ethanol; $\chi = 0.20, \chi = 0.50, \chi = 0.80$, in order to investigate if the excess quantities would be better reproduced with this water model.
The enthalpy of water for the TIP4P model was calculated to be at $(43.95 \pm 0.09) kJ/mol$ and that result is similar to $(43.9 \pm 0.1) kJ/mol$ obtained with the SPC/E model. Results in Table 11 show that there is no significant difference in excess enthalpy, or density, when compared with results for the system with SPC/E water and the same ethanol model (semi flexible OPLS), under exactly the same conditions.

Table 11: Excess enthalpy and density for mixture with different types of water models

χ ethanol	TIP4P ρ $[g/cm^3]$	SPC/E ρ $[g/cm^3]$	TIP4P ΔH [kJ/mol]	SPC/E ΔH [kJ/mol]
0.20	0.918	0.927	0.069	0.096
0.50	0.851	0.854	0.301	0.318
0.80	0.813	0.813	0.144	0.177

3.2 Radial distribution functions

3.2.1 Theoretical introduction

The pair correlation function g(r) expressed explicitly as a function of the distance r, in a homogenous, isotropic liquid, is often referred as radial distribution function, RDF [112]. This function plays a central role in the theory of liquids, because it may be obtained from the simulations or experimentally from the x-ray and neutron diffraction experimental data, and from solving Ornstein-Zernicke integral equation. It provides a direct insight into the microstructure of the liquid on an average level, as in a liquid there are no persistent structures as the crystal lattice, but instead there are persistent structural correlations. Finally, many thermodynamical properties can be expressed in terms of RDF, such as isothermal compressibility, configurational part of the internal energy, structure factor [113][114]. In an isotropic homogenous liquid RDF is defined as the ratio between the pair density of the system and the homogenous one-body number density ρ. A detailed derivation is provided in [112][113][114]. Here is shortened derivation.

Functions of the fundamental importance in the theory of liquids are equilibrium and time dependent distribution functions. For system of N particles and volume V, state of the system at any moment is completely specified by N position vectors $\vec{r^N} = \vec{r_1}, \vec{r_2}..\vec{r_N}$ and N momenta $\vec{p^N} = \vec{p_1}, \vec{p_2}..\vec{p_N}$.

Values of these 6N variables define a point in 6N-dimensional phase space of the system. Hamiltonian of the system is:

$H(\vec{r^N}, \vec{p^N}) = K_N(\vec{p^N}) + V_N(\vec{r^N})$

where $K_N(\vec{p^N})$ is sum of the kinetic energies of N particles, and $V_N(\vec{r^N})$ is inter-particle potential energy.

The aim of equilibrium statistical mechanic is to calculate observable properties of the system as an average over trajectories in real space (Boltzman), or average over an ensemble of systems (Gibbs). In Gibbs formulation the distribution of phase-space points is described by phase-space probability density:

$f^{[N]}(\vec{r^N}, \vec{p^N}, t)$

It is the probability that at time t system is in a microscopic state inside the element of phase-space $d\vec{r}^N d\vec{p}^N$, so its integral over all phase-space is equal to 1.

$$\iint_{\Delta\Omega} f^{[N]}(\vec{r}^N, \vec{p}^N) d\vec{r}^N d\vec{p}^N = 1 \tag{3.10}$$

Given a complete knowledge of phase-space probability density it would be possible to calculate the average value of any function of $\vec{r}$ and $\vec{p}$. Liouville's theorem:

$$\frac{df^{[N]}}{dt} = 0 \tag{3.11}$$

states that probability density does not change with time.

If we are interested in a subset of the system, say n particles of the whole system of N particles, the redundant information can be eliminated by integrating over the coordinates and momenta of the rest N-n particles. In that way the reduced phase-space probability density, or reduced phase-space distribution function is obtained:

$$f^{(n)}(\vec{r}^n, \vec{p}^n, t) = \frac{N!}{(N-n)!} \iint_{\Delta\Omega} f^{[N]}(\vec{r}^N, \vec{p}^N) d\vec{r}^{(N-n)} d\vec{p}^{(N-n)} \tag{3.12}$$

When reduced phase-space distribution function is integrated it gives the probability of finding n particles in the element of the reduced phase-space $d\vec{r}^n d\vec{p}^n$, irrespective of positions and momenta of the rest N-n particles of the system.

With Hamiltonian of the system $H = K_N + V_N$, where K_N is a sum of independent terms, for a system of fixed N, V and T, reduced phase-space equilibrium probability distribution function can be written as:

$$f_0^{(n)}(\vec{r}^n, \vec{p}^n, t) = \rho_N^{(n)}(\vec{r}^n) f_M^{(n)}(\vec{p}^n) \tag{3.13}$$

where

$$f_M^{(n)}(\vec{p}^n) = \frac{1}{(2\pi m k_B T)^{\frac{3n}{2}}} e^{-\beta \sum_{i=1}^{n} \frac{|p_i|^2}{2m}} \tag{3.14}$$

is the product of n independent Maxwell distributions, and the equilibrium n-particle density is:

$$\rho_N^{(n)}(\vec{r}^n) = \frac{N!}{(N-n)!}\frac{1}{Z_N}\int e^{-\beta V_N} d\vec{r}^{(N-n)} \tag{3.15}$$

where Z_N is the configurational integral

$$Z_N = \int_{V^N} e^{-\beta V_N} d\vec{r}^N \tag{3.16}$$

The n-particle density yields the probability of finding n particles of the system with coordinates in the volume element $d\vec{r}^n$, irrespective of positions of all the rest of N-n particles, and of all the momenta. Particle densities and closely related equilibrium particle distribution functions, $g_N^{(n)}(\vec{r}^n)$, provide a complete description of the structure of a fluid, while knowledge of the low-order particle distribution functions, in particular of the pair density $\rho_N^{(2)}(\vec{r}_1, \vec{r}_2)$, is often sufficient to calculate the equation of state and other thermodynamic properties of the system.
n-particle distribution function $g_N^{(n)}(\vec{r}^n)$ is defined as:

$$g_N^{(n)}(\vec{r}^n) = \frac{\rho_N^{(n)}(\vec{r}_1..\vec{r}_n)}{\prod_{i=1}^n \rho_N^{(1)}(\vec{r}_i)} \tag{3.17}$$

For homogenous system particle distribution function measures the extent to which the structure of a fluid deviates from the complete randomness.

$$\rho^n g_N^{(n)} = \rho_N^{(n)}(\vec{r}^n) \tag{3.18}$$

$$g_N^{(n)} = \frac{\rho_N^{(n)}(\vec{r}^n)}{\rho^n} \tag{3.19}$$

In the homogenous, isotropic system pair distribution function $g_N^{(2)}(\vec{r}_1, \vec{r}_2)$ is a function of a separation $r_{12} = |\vec{r}_1 - \vec{r}_2|$, and is called radial distribution function, RDF or g(r).

$$g_N^{(2)} = \frac{\rho_N^{(2)}(\vec{r}_1, \vec{r}_2)}{\rho^2} \equiv g(r) \equiv RDF \tag{3.20}$$

Single-particle density of a uniform fluid is equal to the overall number density, ρ.
Particle densities are expressible in form of delta functions, so distribution function is

$$g(r) = \frac{1}{\rho N} \langle \sum_{i}^{N} \sum_{j \neq i} \delta(\vec{r} - \vec{r}_{ij}) \rangle \tag{3.21}$$

with angle brackets meaning the ensemble average.
If $\rho_N^{(2)}(\vec{r}, \vec{r_1})$ and $\rho_N^{(2)}(\vec{r}, \vec{r_2})$ are independent, there is no correlation between them and the RDF equals 1, which is exactly the case when $r \to \infty$, where this is understood as the r being large enough compared with the size of the molecule, but still within the open system. At the small distances around one molecule in the liquid, the positions of the other molecules surrounding it are correlated to its position, and the RDF measures this correlation. For the closed system the limit of the RDF, when $r \to \infty$ is not 1, but $1 - 1/N$ dependent [114], because in the system of N particles, placing one of them at the fixed position changes the density from N/V to $(N-1)/V$. Simulation, even in case of the N-constant ensemble, represent pseudo infinite system, where global density fluctuation are achieved with appropriate counting formula and with the use of the periodic boundary condition [88]. In NPT ensemble any chosen sub-volume can exchange particles with the rest of the cell, and that implicates that the N-constant property is lost. However, when the cutoff radius reaches the size close to the half of the cell, the reservoir of the particles becomes smaller (in the case of the R equal to half cell size, the reservoir consists only of particles in the cell angles). Therefore, fluctuations of the particle number are not good anymore. This affects the behavior of the RDFs tail in such a way that it does not reach the limit of the open system, which is 1. The asymptotic form of RDF for a finite-N system is usually described in terms of the 1/N correction.

In a closed ideal gas system, where the particles are not interacting, this is exactly:

$$\lim_{r\to\infty} g(r) = 1 - 1/N \tag{3.22}$$

However, in a closed system of the interacting particles there is additional contribution to the closure correlation, apart from the change in the density because one particle that is fixed is missing. It is due to the change in the density because the fixed particle interacts with its surroundings [112], the result proven in a general way by Lebowitz and Percus [115], and first found by Ornstein and Zernike [116].

$$\lim_{r\to\infty} g_{i,j}(1,2) = 1 - \frac{1}{N}\frac{\partial\rho_i}{\partial\beta\mu_j} \tag{3.23}$$

Whether the system is affected with N-constant limit, or it has an issue with the long-range fluctuations depends on the system size. The probability of the fluctuation of the size L depends on the probability that we will find particle outside the cutoff radius $w = (V(sphere)/V(box)) * N_paticle$, and for the small number of particles is negligible. Therefore, system with the small number of particles is close to the closed system, and the correction of the N-constant ensemble should be applied. Larger the system sizes and the number of particles, RDFs tails will be less affected.

It was checked [117] for many systems that if the asymptote for the ideal gas was taken, it was never a serious problem, except in a case when Kirkwood-Buff integral was computed. An empirical way for correcting the asymptote was explained in detail in [118]. It consists of shifting the incorrect asymptote value a_{ij} to 1 with the help of a switch function $S_{ij}(r)$:

$$g_{ij}^{corrected}(r) = g_{ij}(r)[1 + (1-a)S_{ij}(r)] \tag{3.24}$$

with

$$S_{ij}(r) = 0.5(1 + \tanh((r - R_{ij})/\kappa_{ij}) \tag{3.25}$$

where $R_{ij} = \sigma_{a_i} + \sigma_{b_j}$ and $\kappa_{ij} = 1\mathring{A}$. the same effect is achieved by simply adding or subtracting a constant from radial distribution function from R_{ij} on.

Corrected and uncorrected RDF for the O-Ow correlation at mole fraction of 0.12 ethanol are shown on Figures 14 - 15. The size of the correction is seen to be very small, a constant of only 0.0006 had to be subtracted from the RDF data, to get the correct asymptote.

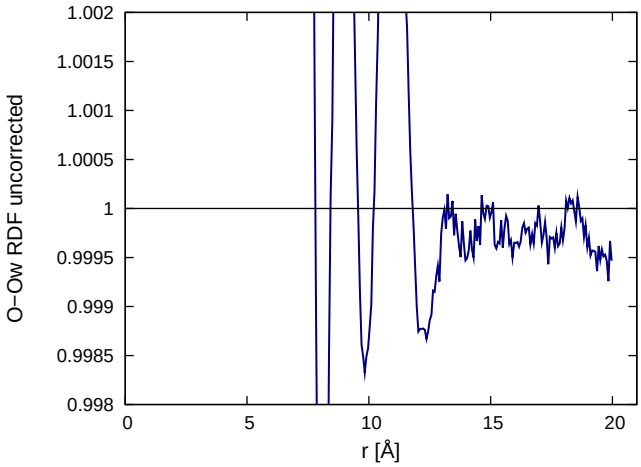


Figure 14: Uncorrected tail of the O-Ow sites RDF for 0.12 mole fraction of ethanol

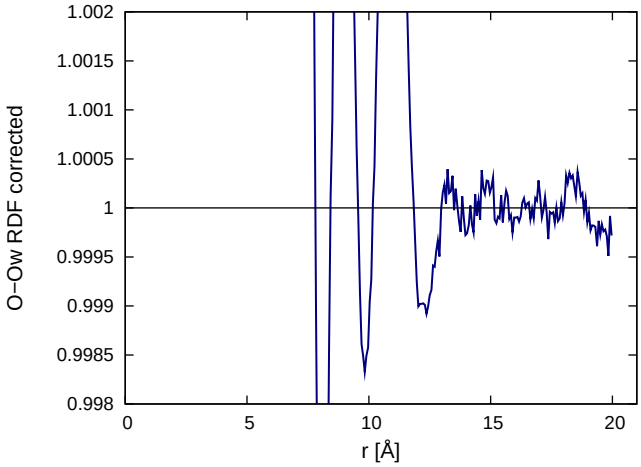


Figure 15: Corrected tail of the O-Ow sites RDF for 0.12 mole fraction of ethanol

3.2.2 Pure systems RDF results

Before venturing into radial distribution functions analysis of the ethanol-water system, it is instructive to show the plot of the radial distribution function of the neat Lennard-Jones system, namely liquid argon, at Figure 16.

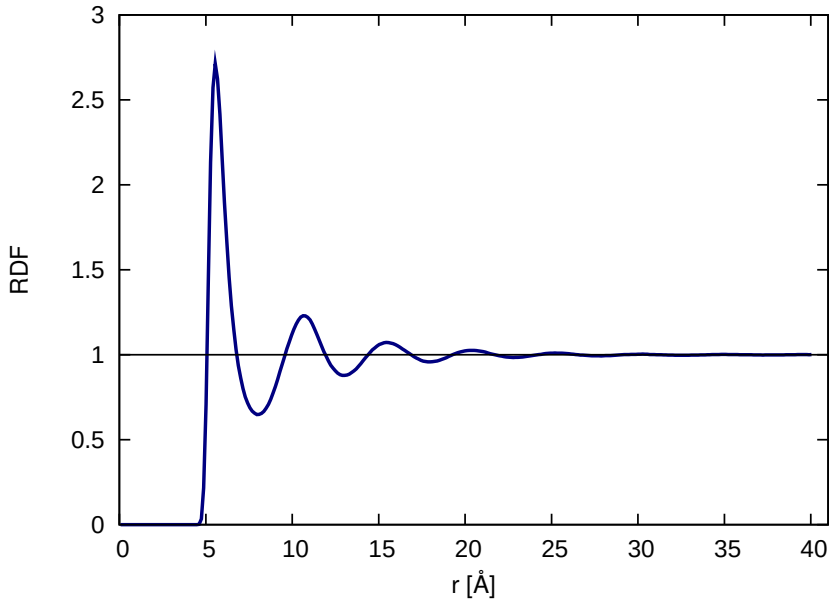


Figure 16: Radial distribution function of liquid argon from [2], system size is 864 atoms, at T=100K

The typical long-range oscillatory pattern is the 'signature' of dense Lennard-Jones systems. The period of oscillations is approximately the length of the atoms separation. The first maximum, is at the minimum of the L-J potential function. Below this r, radial distribution function quickly vanishes to zero, due to the strong repulsive forces that insures particles don't collapse one into another. Above first maximum, RDF develops oscillatory behaviour that reflects the fact that in Lennard-Jones systems particles tend to pack in concentric and nearly equidistant spheres around the central spherical particle.
Figure 17 shows radial distribution function for argon-like system with atoms of the same sizes, for different mole fractions. From this plot it is visible that

change in mole fractions does not affect the height or the shape of RDF.

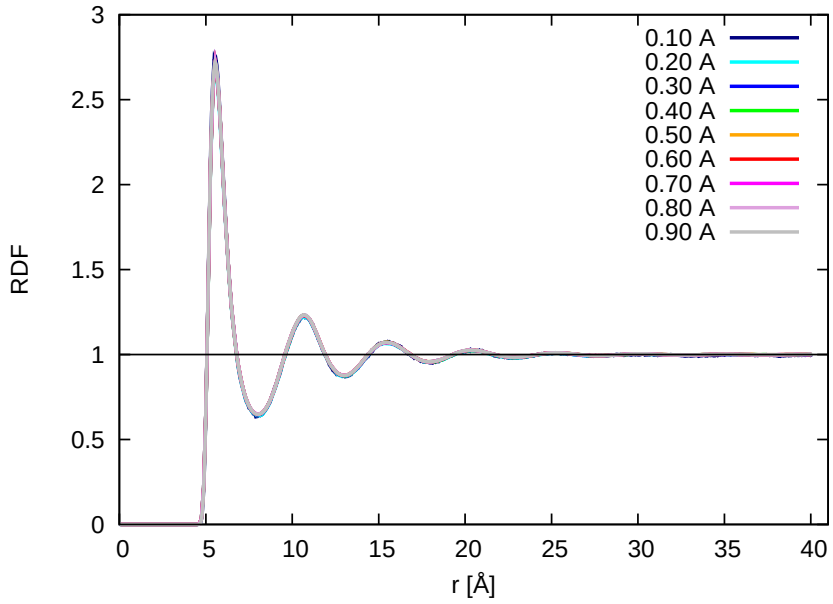


Figure 17: Radial distribution function of liquid argon-like system for different mole fractions from [2]

If RDF's height rises with increased mole fraction it means that local density rises faster than overall homogenous number density of the system. When height of the RDF drops with increased mole fraction, it would mean that increased density does not mean more neighbours, or increased local density, because at the beginning there was already a 'saturation' in local density. And the trivial case is when increased global density of the system increases local density in the same proportion, so RDF remains the same, as on Figure 17.

At Figure 18 are given the radial distribution functions of ethanol sites in pure ethanol, for the fully flexible OPLS model. CH_2 and CH_3 sites have the behaviour that resembles the behaviour of the Lennard-Jones liquid; a long-range oscillatory behaviour, with exponential decay of the amplitude of oscillations. First peaks differ from liquid argon's first peak because this sites

are not single spherical atoms, but are incorporated into the whole ethanol molecule, and so are influenced by the behaviour of the rest of the molecule.

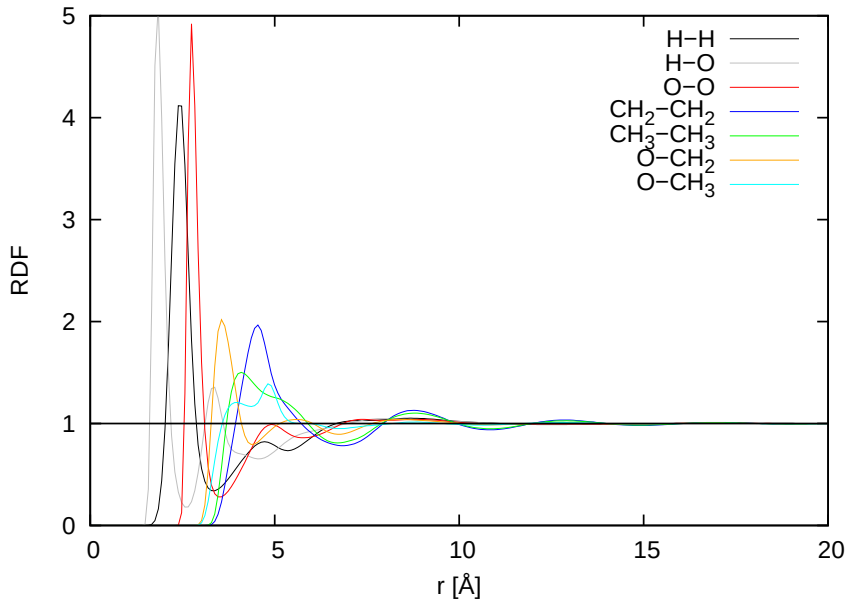


Figure 18: Radial distribution functions of ethanol sites in pure ethanol

Oxygen's first peak is narrower and higher than CH_2 and CH_3 peaks, suggesting stronger and more directed correlation. Broadness of the CH_2 and CH_3 sites suggests only optimal space packing of molecules, without any specific bonding.

Radial distribution functions of water sites in neat water are on Figure 19-20. Hydrogen bonding sites have the same pattern in both liquids; narrow and high first peak that comes from the strong and directional hydrogen bond.

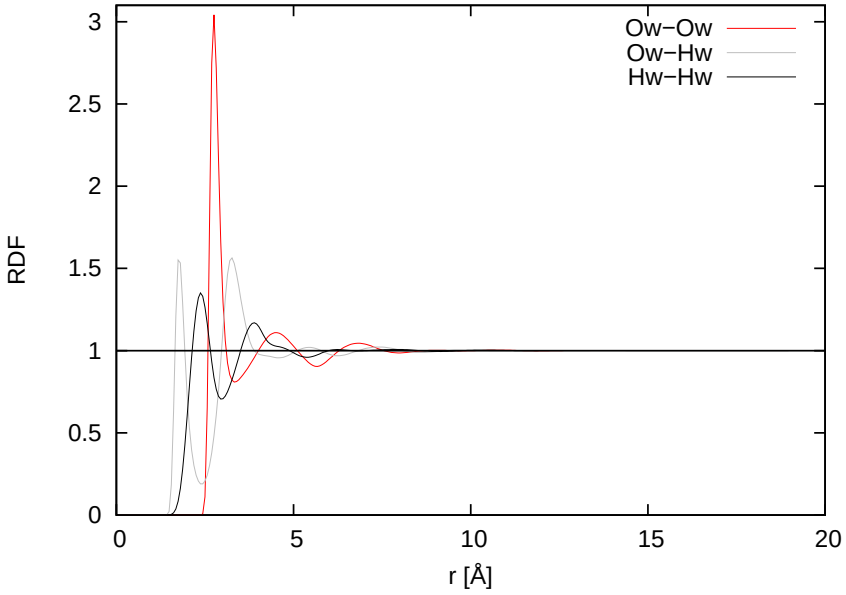


Figure 19: Radial distribution functions of water sites in pure water

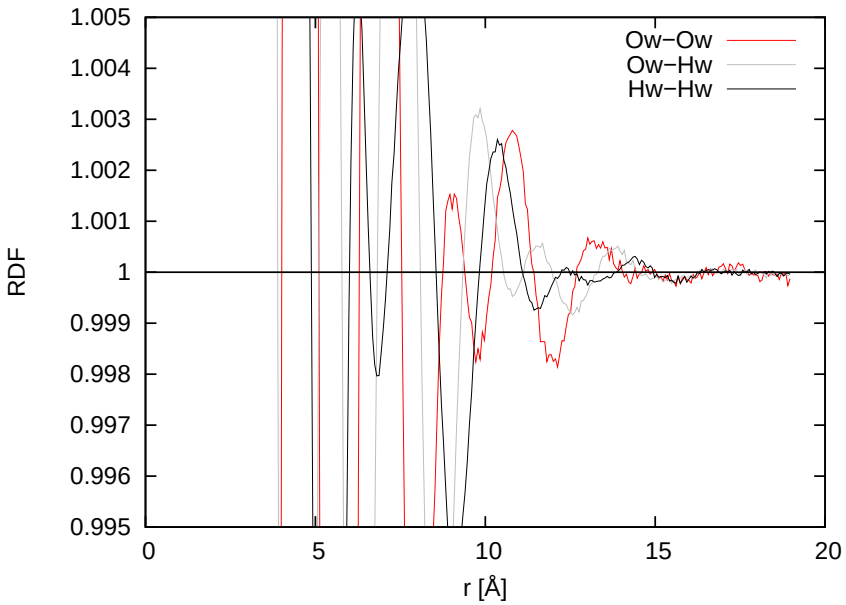


Figure 20: Radial distribution functions of water sites in pure water, zoomed

On Figure 19 so called '3ps' - three peaks shape of water's oxygen's RDF, reported in literature [61], is visible, that seems to disappear beyond 9Å.

However, if we take a look at Figure 20, the oscillatory pattern is still visible up to 15Å, but with much smaller amplitude. The difference in the decay pattern of the hydrogen bonding sites and CH_2 and CH_3 sites is obvious, while the last have exponential decay, the first have step decay, that is reflecting the step from the firs peak of the Ow-Ow correlation to the second. Two peaks of the same height for the correlation of Hw sites around Ow site come from the hydrogens belonging to the same water molecule. First peak is at the 1.93Å, the length of the hydrogen bond between oxygen and hydrogen, and the second peak is approximately 1.63Å away, the distance of two hydrogen atoms in one water molecule for the SPC/E water model. The feature of these two peaks of the same size is then repeated again, on a smaller scale.

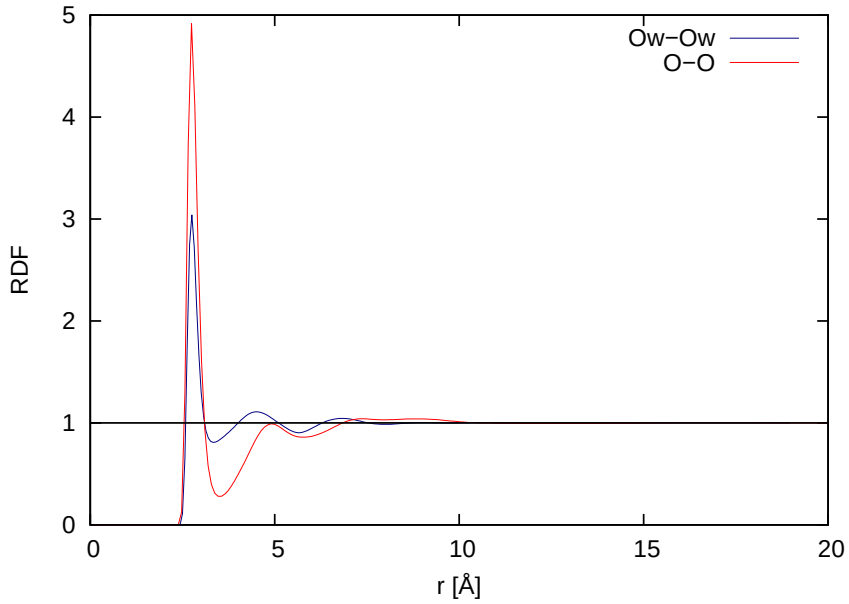


Figure 21: Radial distribution functions of oxygen sites in pure ethanol and pure water

On Figure 21 are RDFs for oxygen sites in pure ethanol and pure water. When compared to water, ethanol's first peak starts to rise a bit earlier, because of the slightly smaller Lennard-Jones σ radius of the oxygen site in ethanol. First peaks are on the same distance for both liquids, while second and third are shifted to bigger radii in ethanol, due to the fact that

ethanol molecules are bigger, and as the radial distribution function is an averaged description of the system, this shift is expected. The first peak in ethanol is higher, even though in ethanol the number of first neighbours for oxygen sites is smaller than in water ($\approx$ 2 in ethanol, $\approx$ 4 in water) because microscopic one particle density of the system, ρ, that comes into denominator of the radial distribution function's definition, is smaller for ethanol, than for water. This should be viewed as the density in the whole system of oxygen sites when they are part of the bigger molecule, to their density when they are part of the smaller molecule.

3.2.3 RDF results for mixtures

Radial distributions functions of various sites in mixtures are going to be examined in this section, to see if any proof for the proposed three structural regimes of ethanol-water mixture can be found in their behaviour. Figures here are from the full flexible OPLS model data, other model's figures are provided in the Appendix.
On Figure 22 - 23 are radial distribution functions of CH_2 sites, at different mole fractions of ethanol.

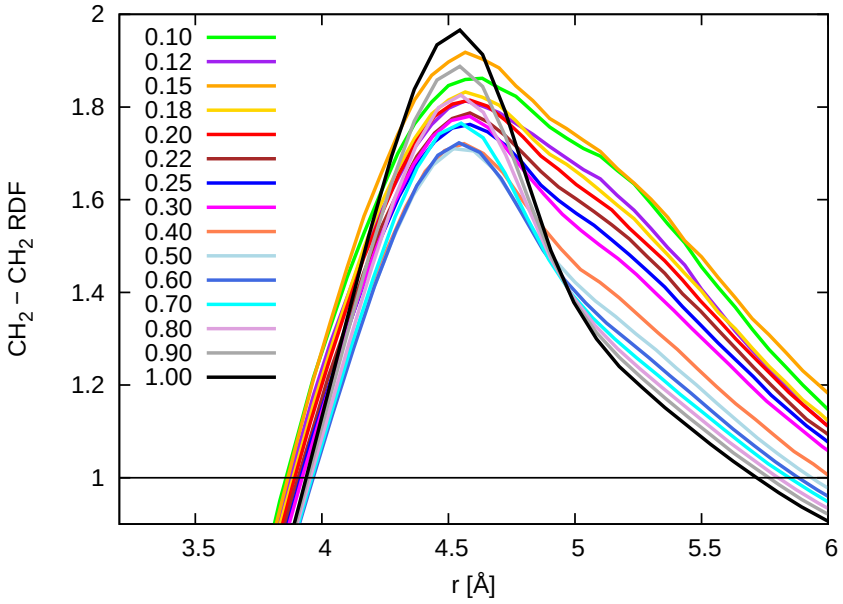


Figure 22: First peak of radial distribution functions of CH_2 sites, at different ethanol mole fraction

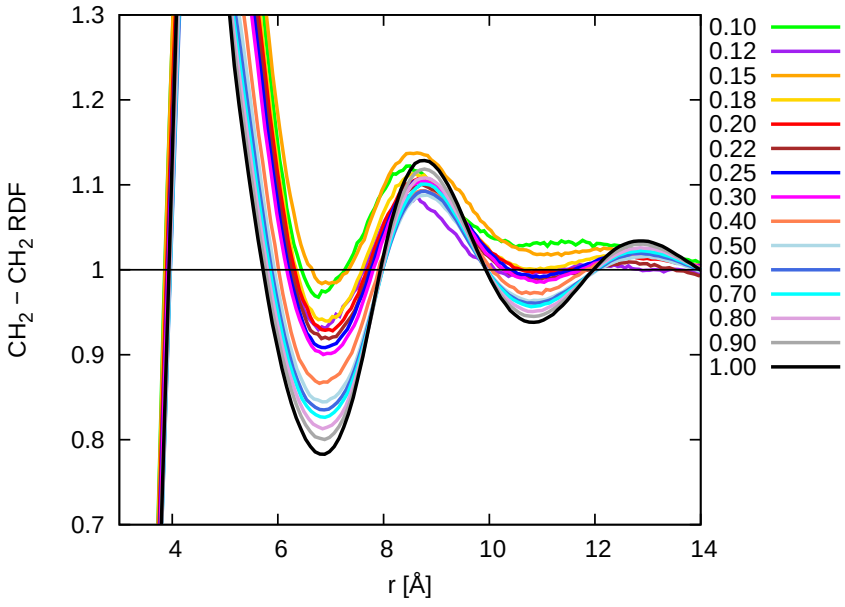


Figure 23: Second and third peak of radial distribution functions of CH_2 sites, at different ethanol mole fraction

On Figure 22 is first peak of the radial distribution function. For 0.15 mole fraction of ethanol it has the highest value, except the one for the pure ethanol. It is a fact in favour of the picture that ethanol-water system has three regimes of mixing, and that important, and in many properties evident structural changes occurs at approximately 0.15 and 0.65 mole fractions of ethanol. On Figure 23 the same characteristic can bee seen for the second peak of the RDF. While first peak is the highest for 0.15 mole fraction, it is the lowest for 0.60 mole fraction of ethanol. The strongest correlation between CH_2 sites, that is approximately centre of mass for ethanol molecule, at 0.15, suggests the biggest number of small ethanol clusters dispersed in water network. From pure ethanol, correlation decreases with the decrement of the ethanol mole fraction, until 0.40-0.60, then it rises until 0.15, and then decreases again. This non-trivial behaviour is reflecting structural changes in the mixture, and the fact that CH_2 site is on one side pulled by the oxygen sites that form H-bonds, and on the other side by CH_3 sites that are subjected to the effect of exclusion from the charged surroundings. Going from pure ethanol to smaller ethanol concentrations, intensity of the first peak goes down, as the ethanol is less abundant in the system, as expected, approximately until the point where percolation of water network is established. From that point, in the middle, bi-continuous phase, correlations increase with the rarification of the ethanol molecules. This is suggesting there is a microheterogenous structure of the system, where the properties of the neat ethanol are lost. Below the point of 0.15 mole fraction of ethanol, correlations drop down again, as further rarification of ethanol molecules leads to the smaller number of ethanol clusters in the mixture. The similar behaviour of the first peak's maximum of CH_2 sites is seen for TraPPE and semi-flexible OPLS model in the Appendix, even though TraPPE model gives the highest correlations for 0.18 mole fraction, and semi flexible OPLS has the same height for 0.15 and 0.18 mole fractions of ethanol. KBFF model marks the first change at 0.12-0.15 (plot in the Appendix), and second change at 0.30, from where correlations rise up to the ones in the pure ethanol. It is interesting to notice the difference between relative ratios of high concentrations correlations (0.60-1.0) to low ones; there is much more

evident difference between them in the KBFF model, high ones are much higher than low ones. For OPLSes and TraPPE model they are all in the same range.

The CH_3 sites that feel no Coulomb interaction, but are subjected to the effect of exclusion from the charged environment, have first peak of the radial distribution function increasing from pure ethanol to the 0.15 mole fraction of ethanol emphasizing the same structural change of merging of the small ethanol domains into biggest micro-segregated domains at 0.15, as the CH_2 sites, Figure 24. When going from pure ethanol to 0.15 mole fraction of ethanol, the first peak also has a very subtle shift to the left, to the shorter distances, indicating the most dense packing of these sites at 0.15, in average. Semi flexible OPLS model gives the highest first peak for CH_3 correlations at 0.15 too, while TraPPE gives it for 0.18, and KBFF at 0.12, as can be seen in the Appendix. Second peaks are also the highest for the 0.15 mole fraction of ethanol, Figure 25. The CH_3 sites have the smallest height of the first peak for the pure ethanol, as a consequence of the hydrophobic effect, that tends to keep CH_3 sites together. The rarification of the sites decreases one particle homogenous system density, that is in the denominator of the RDF function, and two particle correlation stays approximately constant and this results in the rising of the RDF's peak height.

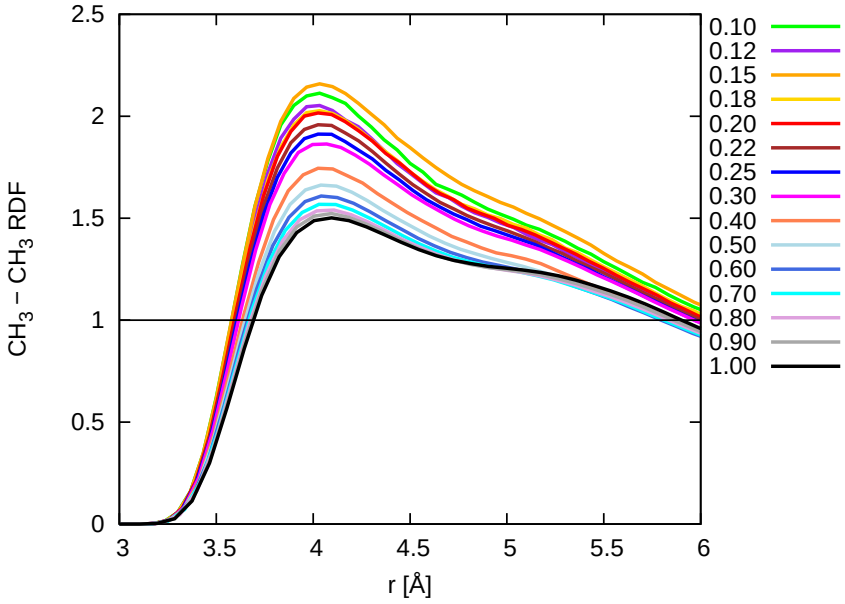


Figure 24: First peak of radial distribution functions of CH_3 sites, for different ethanol mole fractions

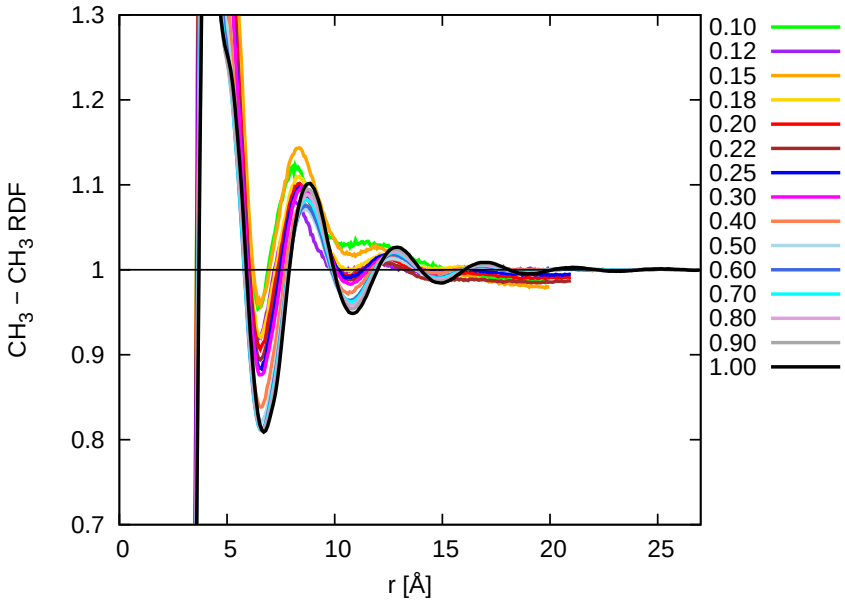


Figure 25: Second and third peak of radial distribution functions of CH_3 sites, for different ethanol mole fractions

Radial distribution function of ethanol oxygen sites is shown on Figure 26, and its first peak on inset. The height of the first peak decreases with decreasing of the ethanol mole fraction, suggesting that oxygen from ethanol preferentially bonds to water's oxygen via H-bond, if possible, as the rarification of O sites in the mixture decreases the height of the first peak. behaviour of this site's radial distribution function does not mark proposed structural changes.

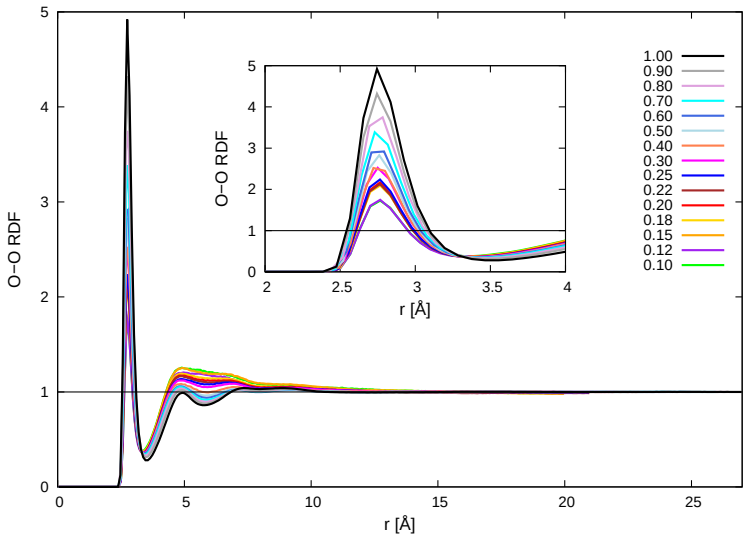


Figure 26: Radial distribution functions of O sites, first peak on inset, for different ethanol mole fractions

Water oxygen sites, Ow, increase radial distribution function's first peak as water's molar fraction decreases from pure water, inset in Figure 27. This suggests that there is an interaction that keeps sites of the same kind together. Of course, it is the H-bond of water oxygen sites, meaning that when water is rarifyed in the system, water oxygen sites stay bonded one to each other, rather than form bonds with ethanol oxygen sites, and the majority of Ow sites is always bonded via H bonds, no matter the mole fraction. Water oxygen sites radial distribution function is on Figure 27, with the three peaks shape visible for all the mole fractions.

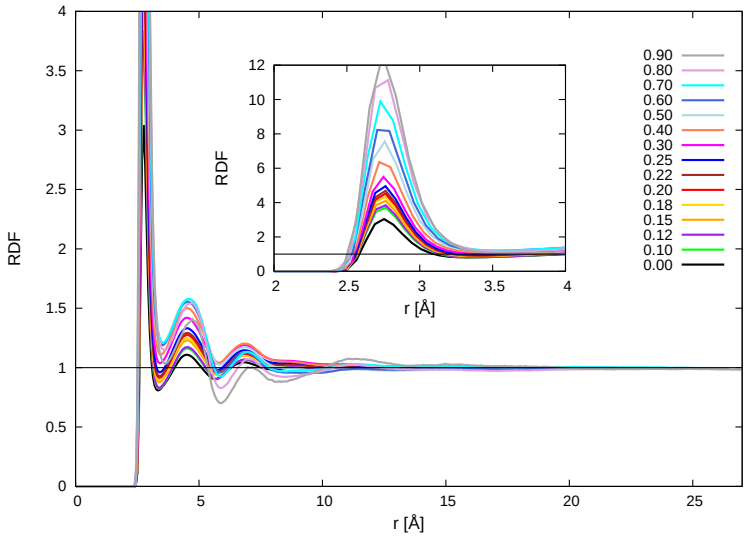


Figure 27: Radial distribution functions of Ow sites, first peak on inset, for different ethanol mole fractions

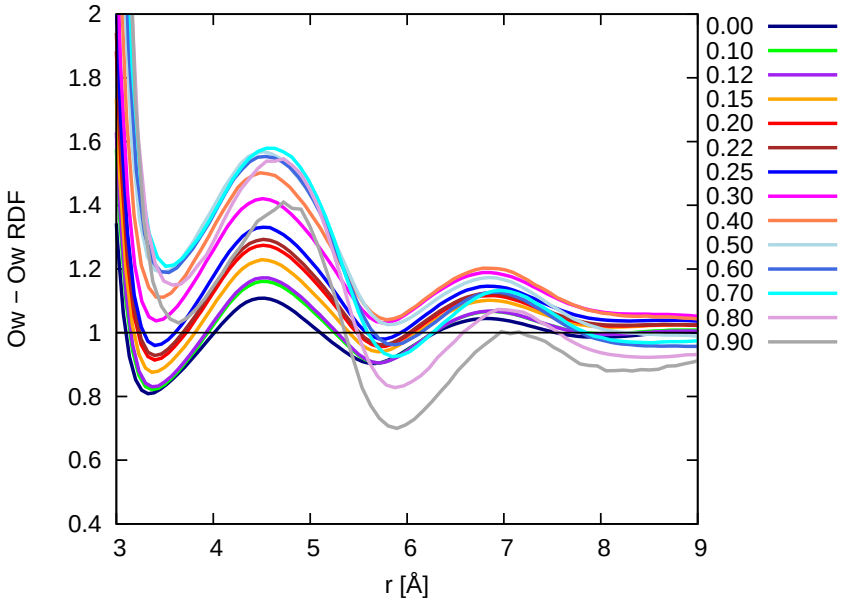


Figure 28: Second and third peak of radial distribution functions of Ow sites, for different ethanol mole fraction

On Figure 28 is one of the facts in favour of the here proposed behaviour of the ethanol-water's microstructure with composition, namely the shift in

the second peak of Ow-Ow radial distribution function for the mole fractions above 0.60. This shift at 0.60 corresponds to the proposed mole fraction where water network is not percolated any more, and becomes thorn apart in smaller fragments. This shift in the position of the second peak of water oxygen sites is confirmed in simulations with all of the models in this work, and it is the most pronounced for KBFF model, figures for the rest of the models are in the Appendix.

One more fact that supports it is the behaviour of the correlation of the CH_3-Ow sites on Figure 29. The fact that CH_3 sites come more in the contact with water at mole fractions above 0.60 is seen as the increased second peak in the CH_3-Ow radial distribution function. The group of curves that represent functions for the mole fractions above 0.60 have obvious difference in behaviour at the second peak, and even at the third peak, from the rest of the RDFs.

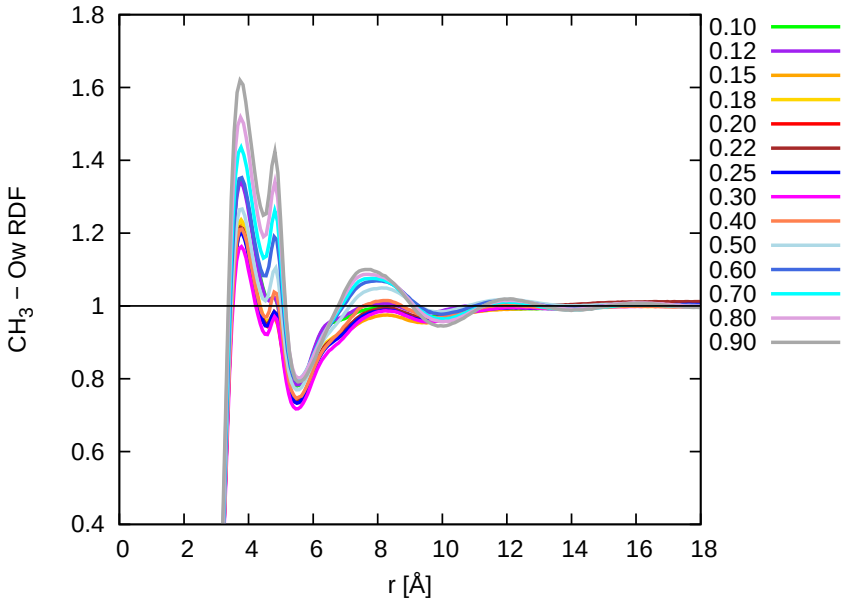


Figure 29: Radial distribution functions for CH_3-Ow sites, for different ethanol mole fraction

It is interesting to look at the radial distribution function of CH_3 sites

around ethanol oxygen O, Figure 30. The central peak in the first peak has minimum at 0.60 mole fraction, and maximum at 0.15, just the two mole fractions where proposed structural change occurs. From 0.0 to 0.15 mole fraction the height rises, as more ethanol is added, and it is the highest for 0.15, where ethanol's hydrophobic sites distribution differs from the randomness the most. Then addition of more ethanol induces the breaking of the confinement of the hydrophobic sites in water network, and O-CH_3 first peak goes down, because ethanols had 'the most dense' packing at 0.15. The peak decreases until 0.60 - the same as CH_3-CH_3, but CH_3-CH_3 is decreasing from 0.15 to 100, and O-CH_3 from 0.15 to 0.60. After 0.60 mole fraction, until 1.00 peak rises again, so now it behaves as ethanol oxygens at this mole fractions, more and more of ethanol, and no specific force between O and CH_3 sites.

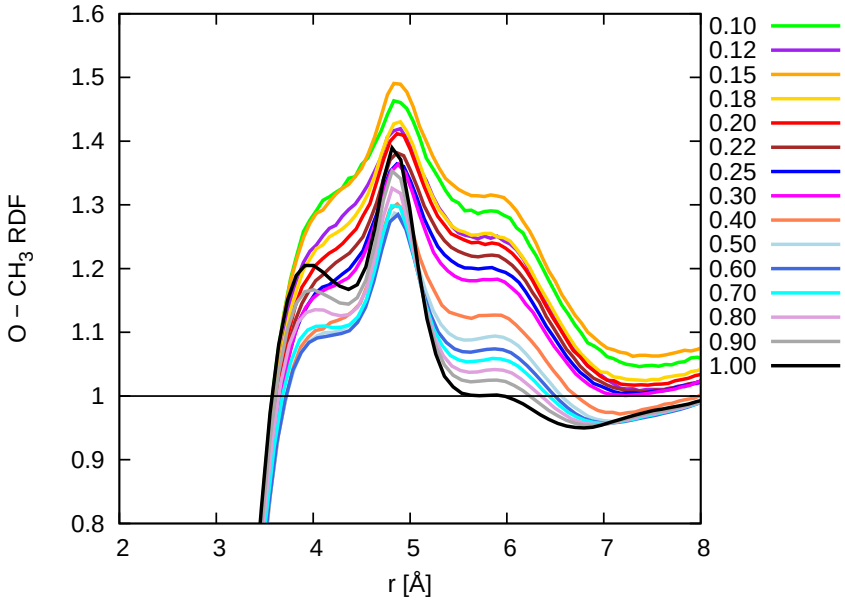


Figure 30: Radial distribution functions for O-CH_3 sites, for different ethanol mole fraction

Summary of the behaviour of first peak heights is piloted in Figure 31. CH_2 and CH_3 sites clearly indicate three regions along the mole fraction change in this mixture. In the first region to the left in Figure 31 increase of

the ethanol mole fraction increases local density of all ethanol sites more than their global density. This is interpreted in a way that ethanol molecules are clustered together in this region. In the middle region only ethanol oxygen sites increase their local density more than global density, while there is overturn in the behaviour of CH_2 and CH_3 sites, at the left border of this region there is the highest ratio of their local to global density. Another overturn happens again at the beginning of third, right region. In the right region CH_2 sites now follow the behaviour of oxygen ethanol sites, meaning that adding more ethanol increases their local density more than their global density, while CH_3 sites correlations are still dropping, meaning they are released from their imprisonment that water network forced on them in the first mole fraction region.

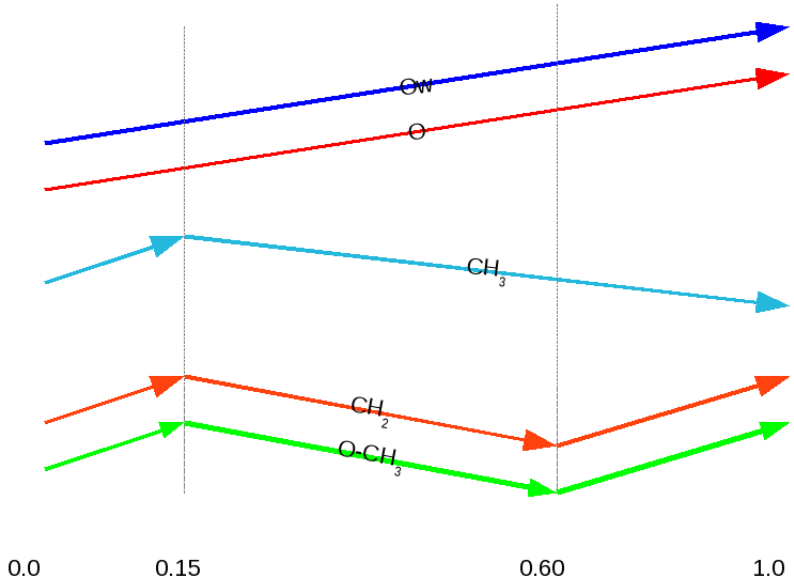


Figure 31: Summary of the behaviour of RDF's first peak heights along the mole fraction range from pure water on the left

3.2.4 Comparison of results for different ethanol models

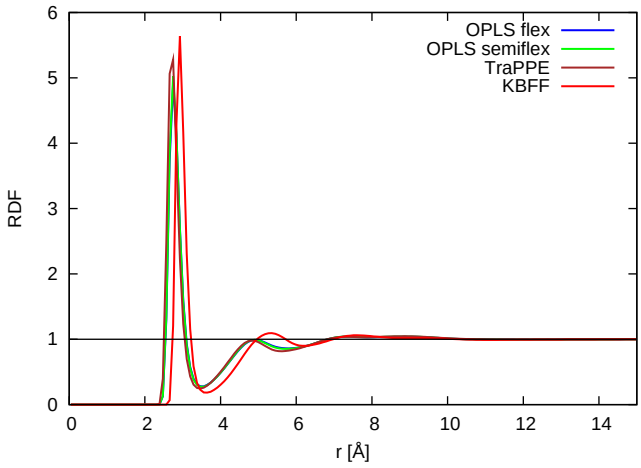


Figure 32: Radial distribution functions of O-O sites in pure ethanol, for various ethanol models

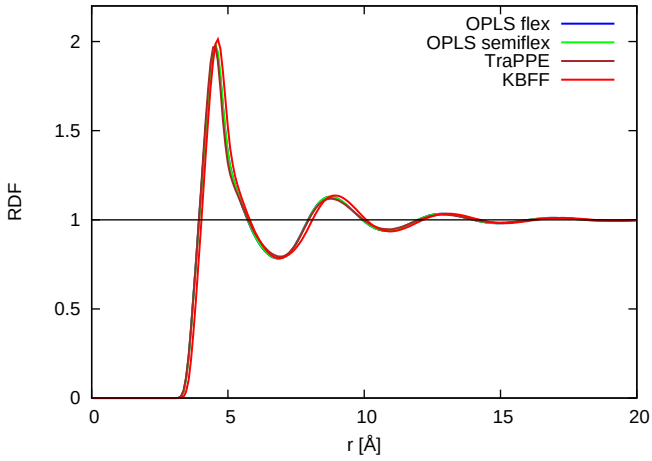


Figure 33: Radial distribution functions of CH_2-CH_2 sites in pure ethanol, for various ethanol models

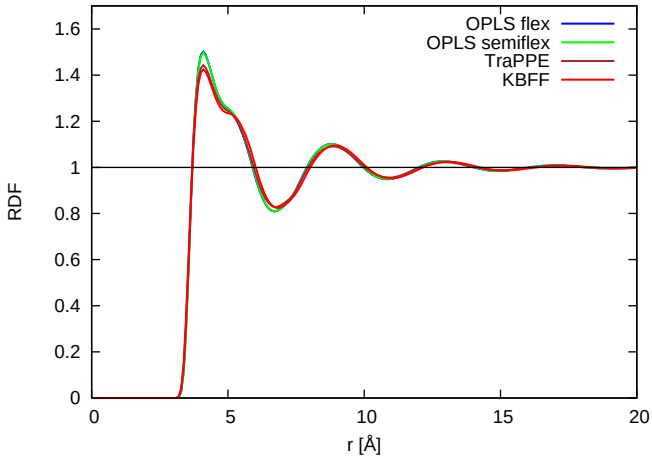


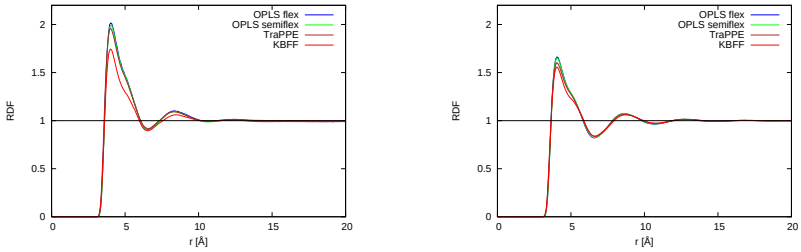
Figure 34: Radial distribution functions of CH_3-CH_3 sites in pure ethanol, for various ethanol models

Figure 32-34 present RDFs of various ethanol models for pure ethanol. Fully flexible OPLS and semi flexible OPLS model give the same radial distribution functions for all three sites. This is expected, as both models have the same parameters. TraPPE model, having slightly different parameters, gives slightly different radial distribution functions, but in general their features do not differ much. The positions of minima and maxima are at the same distances, and small differences can be seen only at the depth of minima and the height of peaks. KBFF model has all hydrogen-bonded correlation shifted outwards, mainly due to the non-zero Lennard -Jones parameter for radius of the H-atom (see section 2.1). The O-O correlations are slightly larger due to the stronger charges, which is then compassed by the smaller CH_3 correlations.

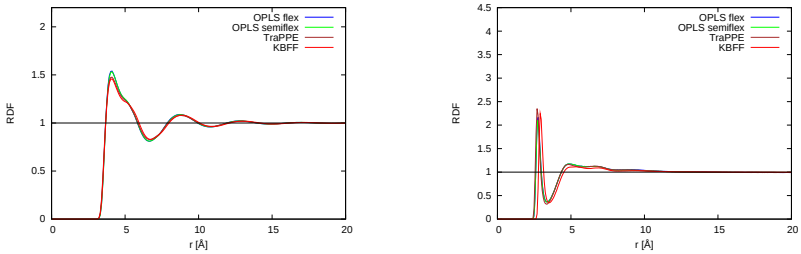
Differences in mixture's RDFs for various ethanol models are plotted on Figure 35. KBFF model shows noticeable difference with respect to other models, and the main one is that it increases mixing, which is seen as increasing of the cross-correlations. Nevertheless, all models follow general trends of the concentration dependence, which can be summed up in three points:

a) correlations for O-O alcohol sites increase with alcohol concentration, as

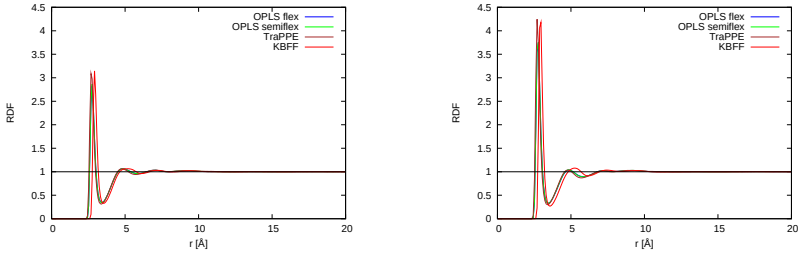
adding more alcohol increases H-bonding between hydroxyl group;
b) water structure is more enhanced with alcohol mole fraction;
c) hydrophobic correlations slightly decrease with rarefying of the alcohol.



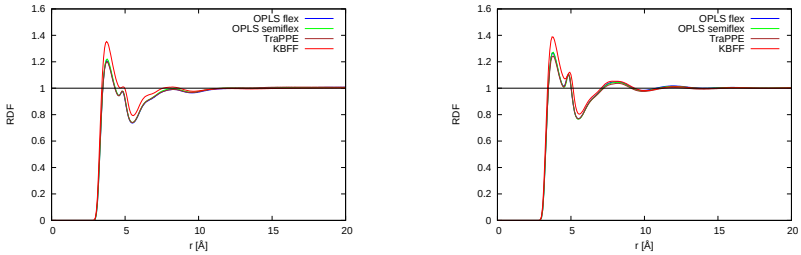
(a) $\chi = 0.20$ CH_3-CH_3 **(b)** $\chi = 0.50$ CH_3-CH_3



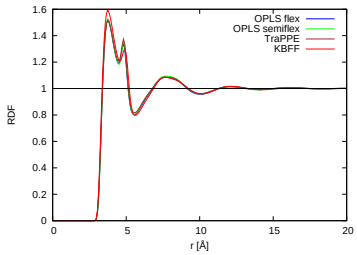
(c) $\chi = 0.80$ CH_3-CH_3 **(d)** $\chi = 0.20$ O-O



(e) $\chi = 0.50$ O-O **(f)** $\chi = 0.80$ O-O



(g) $\chi = 0.20$ CH_3-Ow **(h)** $\chi = 0.50$ CH_3-Ow



(i) $\chi = 0.80$ CH_3-Ow

Figure 35: Radial distribution functions for various concentrations, sites and ethanol models

3.2.5 Comparison of SPC/E and TIP4P results

As previously said, three mole fractions; $\chi = 0.20, \chi = 0.50, \chi = 0.80$ were simulated, to test if TIP4P model for water would yield better results of excess enthalpy for the mixed system. Figure 36 presents Ow-Ow radial distribution functions in pure water at ambient conditions for SPC/E and TIP4P models compared with experimental x-ray results from [119] and [120]. Ow-Hw and Hw-Hw are plotted on inset. Both models are in fairly good agreement with the experimental data, as concluded in [120]. The plot obtained here for Ow-Ow RDFs is exactly the same as in reference [120].

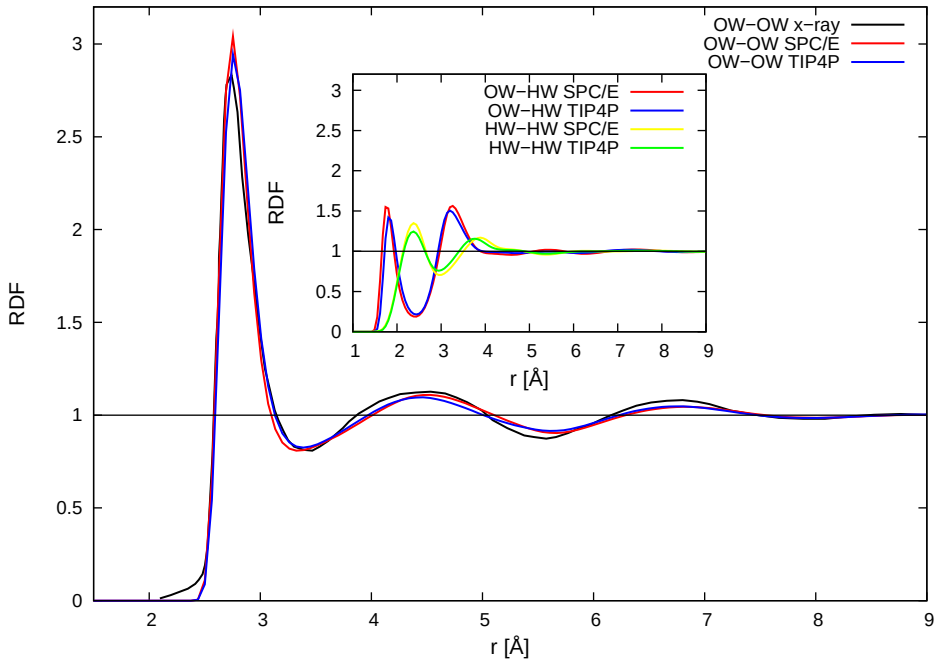
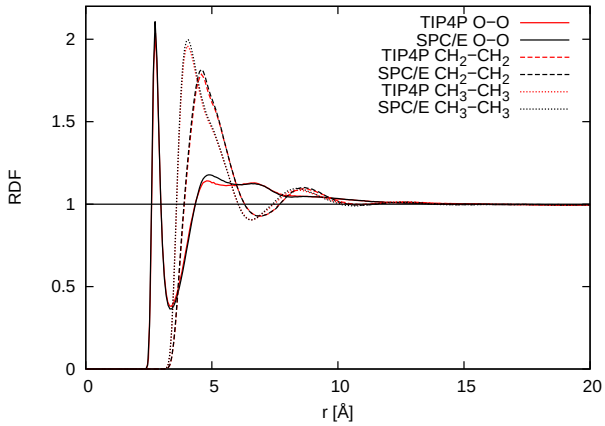
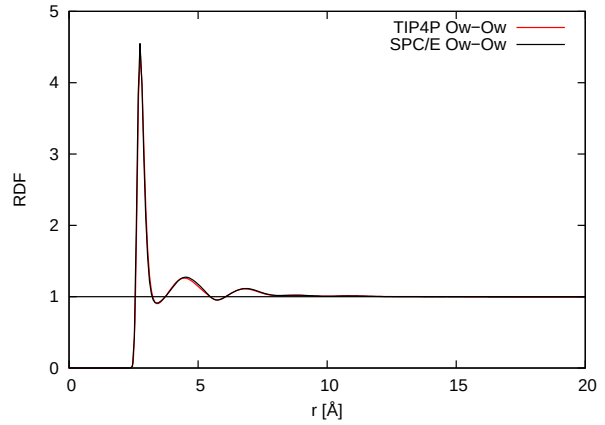


Figure 36: Radial distribution functions in pure water, comparison of SPC/E and TIP4P models for water

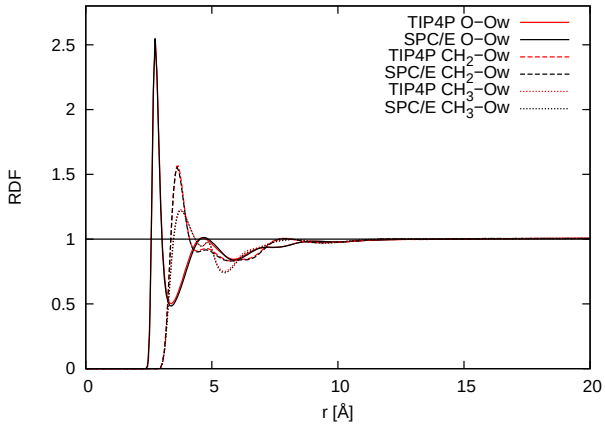
On Figure 37 are radial distribution functions for both models of water with semi flexible OPLS model for ethanol at $\chi = 0.20$ mole fraction of ethanol. Obviously there are no significant differences in radial distribution functions. SPC/E model gives slightly higher first peak in mixture, as well as in pure water.



(a) Ethanol



(b) Water



(c) Ethanol-water

Figure 37: Radial distribution functions at $\chi = 0.20$ mole fraction of ethanol, with TIP4P and SPC/E water models

3.3 Coordination number

Coordination number is the number of neighbouring sites in the first shell around some central site. It is defined as:

$$CN = 4\pi\rho \sum r^2 g(r) \Delta r \tag{3.26}$$

where summation goes from $r = 0$ to $r = r_{1.minimum}$. ρ is number density and $g(r)$ is radial distribution function, RDF. Coordination numbers for various sites from ethanol - water mixture for semi flexible OPLS, as well as coordination numbers for the system of binary mixture of Lennard-Jones atoms of the same size is shown on Figure 38

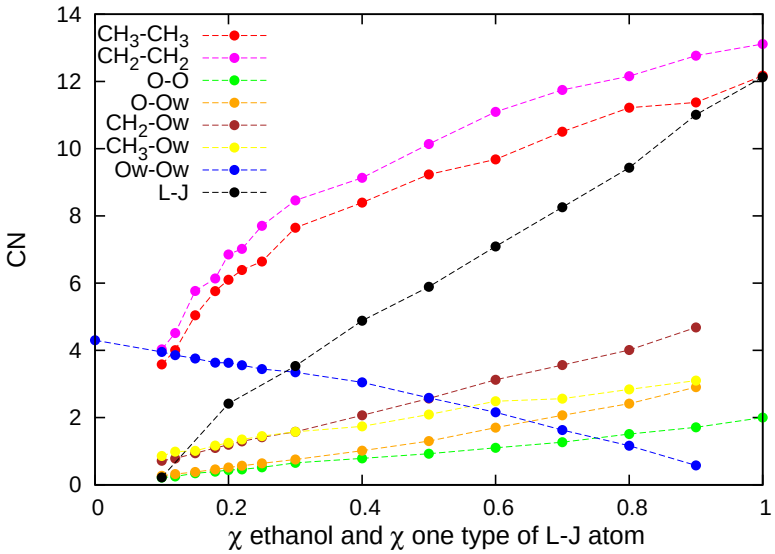


Figure 38: Coordination numbers

Coordination number for ethanol oxygen around ethanol oxygen from neutron diffraction study by Benmore and Loh [26] was determined to be 2.0 ± 0.2 by $3\mathring{A}$ distance. From MD study by Noskov, Lamoureux, Roux [48] it was calculated to be 1.94.

In this work ethanol oxygen number was calculated to be 2, as seen on Figure 38, what is in good agreement with the mentioned results. The change

in ethanol oxygen coordination number is almost linear with the change of mole fraction.

Coordination numbers of CH_3-CH_3[6] and CH_2-CH_2 sites sharply rise until $\chi = 0.15$ mole fraction of ethanol, where first proposed structural change in this mixture takes place. This rise is very different from the monotonous rise of the Lennard-Jones coordination number, even though in the pure ethanol this numbers are very close to the value of 12.1 of the pure Lennard-Jones liquid. Behaviour of CH_2-CH_2 and CH_3-CH_3 coordination numbers is evidence for ethanol's more pronounced shielding of the methyl groups at mole fractions below $\chi = 0.15$. Marked change at this mole fraction reveals the change in the structure of the system at the level of the first neighbours distances. O-Ow[7] coordination number is higher than O-O, reflecting the fact that ethanol prefers to bond to water, than to ethanol, as expected, as a consequence of the stronger charge at Ow site.

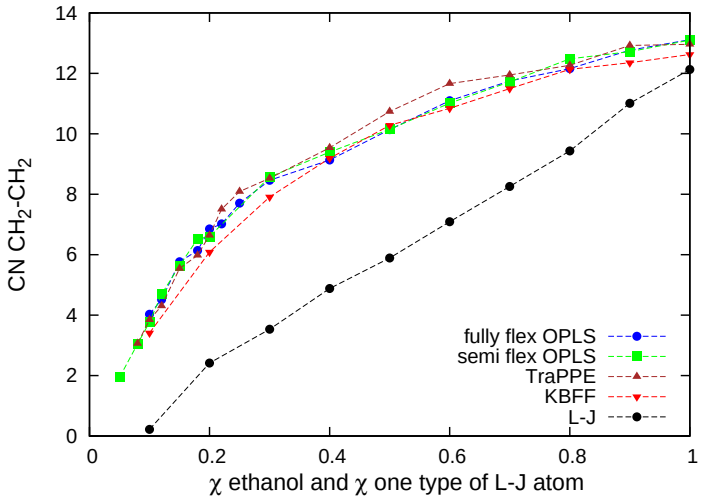
Water's coordination number for SPC/E model was calculated by Wu, Tepper, Voth [122] to be 4.34 and authors report experimental result of 4.26. Bagchi 2012 [123] reports coordination number of water to be 5 at 300 K for TIP5P model. In this work water oxygen - water oxygen coordination number was calculated to be 4.3, that is in good agreement with the previous simulation results.

[6] CH_3-CH_3 meaning: CH_3 site around CH_3 site coordination number

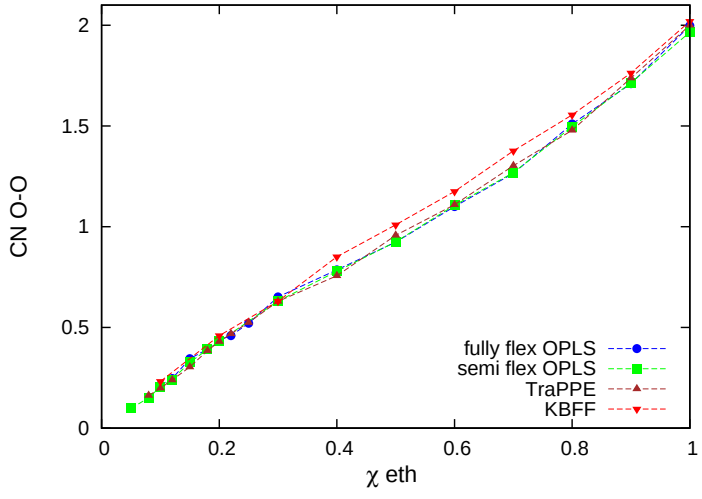
[7] O-Ow meaning ethanol oxygen around water oxygen. It is not to be confused with Ow-O that would mean water oxygen around ethanol oxygen, and that is a different thing, as can be seen from [121]

3.3.1 Comparison of different ethanol models results

On Figures 39-40 are coordination numbers for different ethanol models.

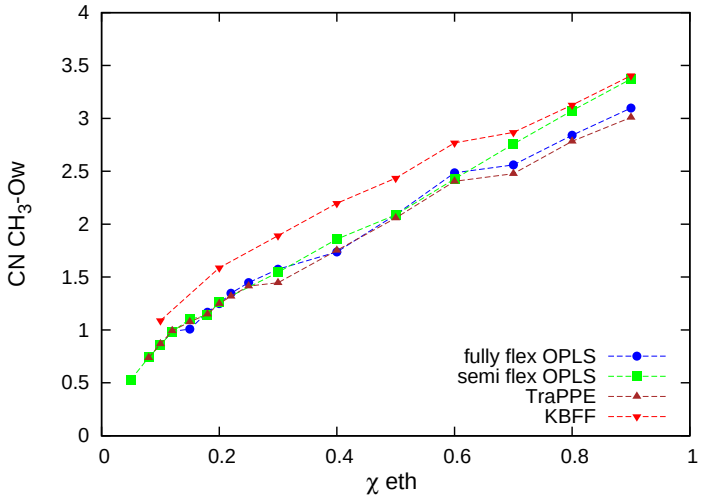


(a) CH_2-CH_2

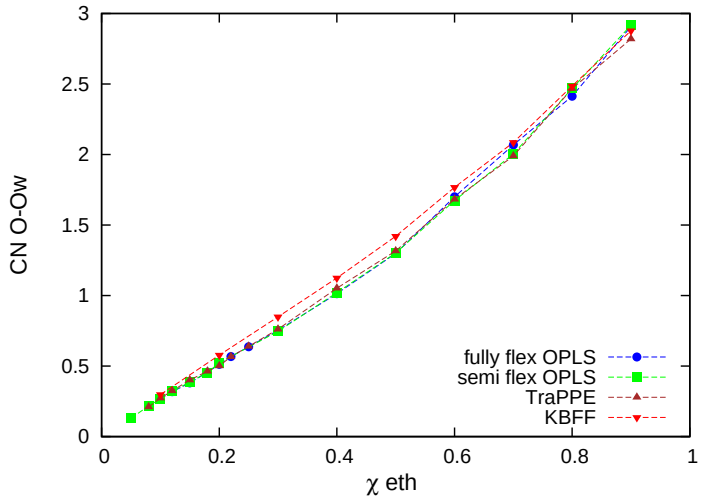


(b) O-O

Figure 39: Ethanol-ethanol coordination number for different ethanol models



(a) CH_3-Ow



(b) O-Ow

Figure 40: Ethanol-water coordination numbers for different ethanol models

In the region around 60 % TraPPE model has largest CN for hydrophobic sites. This corresponds to the shift of the maximum for TRAPPE excess enthalpies (see Figure 12). On the low alcohol mole fraction up to approximately 40 % OPLS and TraPPE models have larger number of first neighbors for hydrophobic sites then KBFF. Trying to interpret this in terms of Coulomb contribution, segregation of hydrophobic sites leads to enhancement of the repulsive part of the interaction, since it brings together positive charges from

CH_2 closer. Even thought the CH_2 charges are small, comparing to other partial charges, the nett effect of CH_2-repulsion is traceable through positive enhancement of the excess energies. This effect is obvious in the highest TraPPE excess Coulomb energy (see Figure 13), especially at $\chi = 0.6$, as it can be interpreted as a consequence of the highest CH_2-CH_2 CN for this model.

For the higher mole fractions models differ in the the hydrogen-bonded sites behaviour. Excess Coulomb energy in this concentration region is positive for all models, but the effect is much less for KBFF (see Figure 13). In this region O-O and O-Ow CN are bigger for the KBFF models, showing that these sites are also more bonded, resulting in the lower overall energy.

The most noticeable difference between the models is in the cross-correlations between hydrophobic sites and water, where KBFFs coordination numbers for CH_n-Ow (n=2,3) cross-correlations for all concentrations are larger. The cross-CN for O-Ow correlations differ mostly in the middle region where there is approximately equal number of water and alcohol. In the bases of this are the stronger charges on ethanol hydroxyl group in KBFF model, that are more "eager" to form bonds either with molecule of water or another molecule of ethanol. Therefore the structure of the interface is more open, allowing more contact, also between CH_n and Ow sites. KBFF model therefore acts against micro segregation, and favors more homogeneous structures. However, this is a subtle difference, and it does not imply that there is no micro segregation in mixtures modeled by KBFF model. Microsegregation exists in all these systems, modeled by any of these models, and it is clearly obvious from the shape of CH_2-CH_2 CN behaviour and its difference from the Lennard-Jones CN.

3.4 Kirkwood-Buff integrals

3.4.1 Theoretical introduction

Kirkwood-Buff theory was first published in 1951. [124], where Kirkwood and Buff derived some new relationships between thermodynamic quantities and radial distribution functions in two-component system in the open μ, V, T ensemble. It is considered to be the most general and the most powerful theory of solutions according to Ben-Naim [112]. It provides direct relationship between thermodynamical quantities such as compressibility, partial molar volumes and derivatives of chemical potential in terms of Kirkwood-Buff integrals (KBI), that are the measure of concentration fluctuation in a system. KBI gives measure of the tendency of a molecule j to concentrate around a central molecule i [125]. KBI is defined as:

$$G_{ij_{\mu VT}} = \int_0^\infty 4\pi[g_{ij}(r) - 1]r^2 dr \tag{3.27}$$

where i and j are different species, and g_{ij} is the corresponding radial distribution function. The main result of Kirkwood-Buff theory is the fact that from integrals of the radial distribution function one may calculate thermodynamical properties of the system. With the definition of two auxiliary quantities:

$$\eta = \rho_A + \rho_B + \rho_A\rho_B(G_{AA} + G_{BB} - 2G_{AB}) \tag{3.28}$$

$$\xi = 1 + \rho_A G_{AA} + \rho_B G_{BB} + \rho_A\rho_B(G_{AA}G_{BB} - G_{AB}^2) \tag{3.29}$$

with A and B being two species, and ρ number density, it is possible to express thermodynamic quantities in terms of molecular quantities, KBI:

$$\kappa_T = \frac{\xi}{kT\eta} \tag{3.30}$$

$$\overline{V}_A = \frac{1 + \rho_B(G_{BB} - G_{AB})}{\eta} \tag{3.31}$$

$$\overline{V}_B = \frac{1 + \rho_A(G_{AA} - G_{AB})}{\eta} \tag{3.32}$$

$$\mu_{AA} = \frac{\rho_B kT}{\rho_A V \eta} \tag{3.33}$$

$$\mu_{BB} = \frac{\rho_A kT}{\rho_B V \eta} \tag{3.34}$$

$$\mu_{AB} = \mu_{BA} = -\frac{kT}{V\eta} \tag{3.35}$$

where κ_T is compressibility, $\overline{V}_i$ partial molar volume of the species i, and $\mu_{ij} = (\partial\mu_i/\partial N_j)_{N_i,p,T}$ derivatives of the chemical potential. In 1978 Ben-Naim [126] published the inversion of Kirkwood-Buff theory, so G_{ij} were extracted from measurable thermodynamic quantities (formal derivation can be found in [112]):

$$G_{AB} = kT\kappa_T - \rho\overline{V}_A\overline{V}_B/D \tag{3.36}$$

$$G_{AA} = kT\kappa_T - \frac{1}{\rho_A} + \frac{\rho_B\overline{V}_B^2\rho}{\rho_A D} \tag{3.37}$$

$$G_{BB} = kT\kappa_T - \frac{1}{\rho_B} + \frac{\rho_A\overline{V}_A^2\rho}{\rho_B D} \tag{3.38}$$

where $\rho = \rho_A + \rho_B$, and $D = \frac{\chi_A}{kT}(\frac{\partial\mu_A}{\partial\chi_A})_{p,T}$ is a term related to concentration fluctuations [127].

3.4.2 Calculation and measurement of the KBI

KBI is strictly property of the μ, V, T ensemble, however it can be calculated from the simulation data in NpT ensemble from running KBI (rKBI), under a few approximations: that simulation data can represent an open system, that KBI at infinity is equal to rKBI up to certain range R that is taken under the assumption that beyond this radius the system has already reached the homogeneity, so there are no long-range correlations in the system.

$$G_{ij_{\mu VT}} = \int_0^\infty 4\pi[g_{ij}(r) - 1]r^2 dr \approx G_{ij_{NpT}} = \int_0^R 4\pi[g_{ij}(r) - 1]r^2 dr \tag{3.39}$$

When calculating KBI from simulation in NpT ensemble, the correction to the tail of the RDF must be applied, as explained when radial distribution

function is discussed. However, sometimes even with this correction, it is impossible to get RDF to oscillate around any horizontal asymptote. This may be attributed to the small system size in simulations of microheterogenous systems, where the integral must be obviously cut at some radius (as necessary approximation in the NpT ensemble), so microheterogeneous structure may not have enough space to develop fully. The problem with alcohol-water systems is, as stated in introduction, that they are between micelle-forming systems, that have more or less defined size of microsegregated domains, and random systems. Systems of alcohols and water have microsegregated domains, as it appears, of no specific shape, so it is very hard to determine their size and consequently the correlation length. In order to properly evaluate KBI cutoff radius must be larger than the correlation length in the system. With ethanol-water type of microheterogenous system it is not always easy to evaluate KBI from simulation data.

When evaluating KBI from simulations it is suggested in the literature [64] that a cutoff distance (that is used as the integral's upper limit) equal to the range over which the intermolecular forces dominate the distribution of the particles, is a good approximation. Experimental estimates suggest that the radius at which one molecule influence another extends over several molecular diameters, although this is somewhat dependent on the density [128]. In [64] radius of 10 Å was taken. It is reasonably to take this cutoff radius in a way that the rest of the volume is sufficiently big reservoir of particles, to mimic infinitely big system. But this does not solve the problem of possible appearance of big microsegregated domain in the system, so that the correlation length becomes bigger than the chosen cutoff radius.

It should be also noted that there are some difficulties in obtaining accurate KBI values from the available thermodynamic data [125]. Quantities used for calculation of KBI are extremely sensitive to experimental precision and the accuracy of the of fitting of vapour-liquid equilibrium or activity data. A review of different values of KBI obtained by different authors on the same systems is provided in [125]. Another method for obtaining KBI values is from small-angle x-ray scattering, as Nishikawa and Ilijima did 1993. [129] for ethanol water system, and neutron scattering experiments.

3.4.3 KBI results

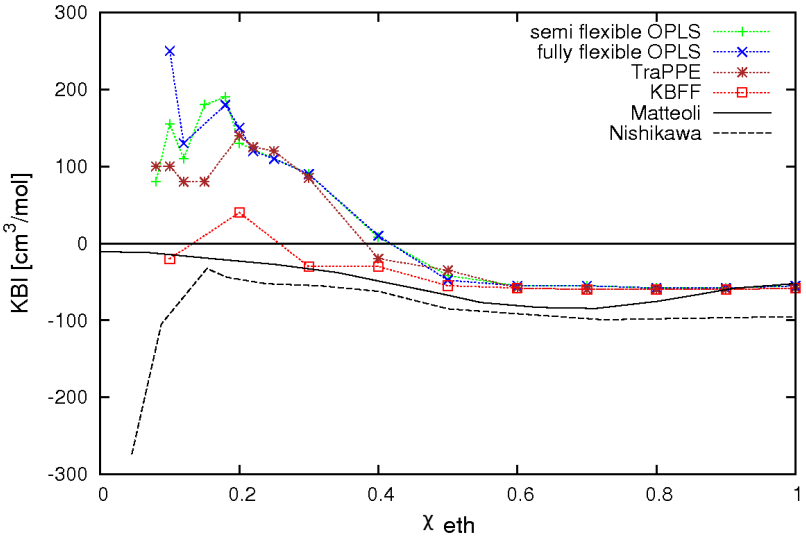
Kirkwood-Buff integrals for various models used in this work are shown on Figure 41. Results are compared to the experimental data from thermodynamic measurement [125] and from small-angle x-ray scattering [129]. First it is obvious that experimental data are not in good agreement with each other. G_{EE} has very different behaviour in the region $\chi \leq 0.2$. Here experimental data from [129] show steep rise, that is very different from thermodynamic measurement. KBFF model is in good agreement with [125], while OPLS and TraPPE models overestimate ethanol-ethanol KBI in the region $\chi \leq 0.2$. This fact can be interpreted in a way that small changes in ethanol model's parameters, from KBFF to TraPPE, or OPLS, have big effect on ethanol-ethanol KBI in the water rich region. It can be understood in the light of the proposed structural regime of small ethanol clusters in water network, as this structure is assumed to be highly frustrated, and thus easily removed from balance if the parameters of the model are guessed slightly wrong. At lower concentrations the difference G_{EE} and G_{EW} indicates that ethanol has higher preference to bond to ethanol, than to water[8]. From $\chi > 0.4$ ethanol shows very slight preference to water. This is understandable, as in the water rich region hydrophobic effect is stronger. From $\chi > 0.4$ on, all the models are in fairly good agreement with experimental data for G_{EE}. It can be concluded that OPLS and TraPPE models overestimate concentration fluctuation for ethanol-ethanol, and underestimate concentration fluctuation for ethanol-water, in the region $\chi \leq 0.4$, while KBFF model is in much better agreement with experimental data.

Water-water concentration fluctuations are systematically overestimated in region from $\chi \geq 0.15$ to $\chi \leq 0.4$ by OPLS and TraPPE models. The conclusion can be drawn that OPLS and TraPPE models enhance microsegregation at small mole fractions in this system. From the behaviour of experimental and simulated difference between G_{WW} and G_{EW} it can be concluded that water is preferentially solvated by water at all mole fractions

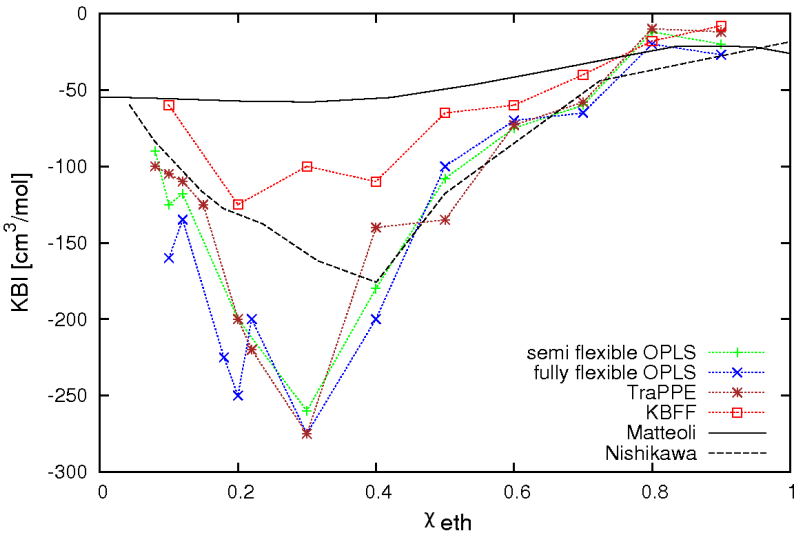
[8] If G_{EE} is higher than G_{EW} it means that ethanol is preferentially solvated by ethanol, than water, and vice versa [112]

in this system. In the proposed middle bi-continuous region water-water KBI has maximum value.
In the ethanol rich region $\chi > 0.65$ ethanol and water have slight preference for water. The proposed structural change at this mole fraction occurs as the result of the breaking up of the percolated water network, that has no big effect on concentration fluctuations, and thus it is not expected to be noted in the behaviour of the KBIs.

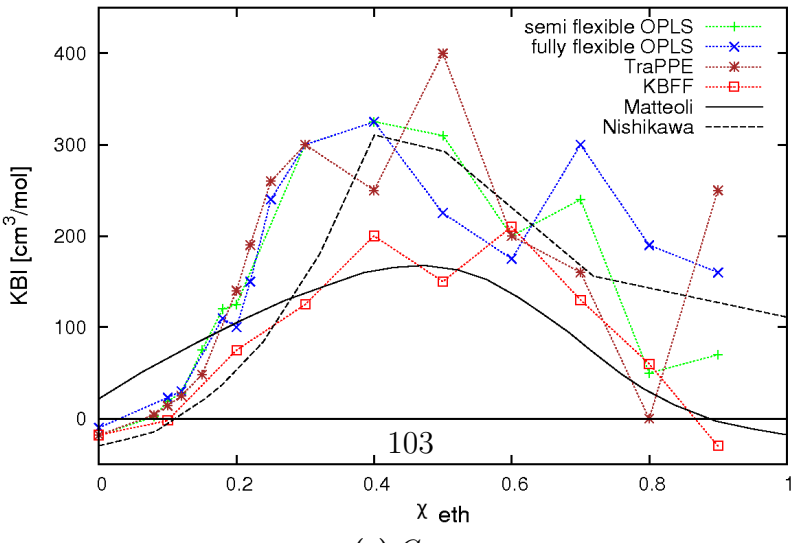
It is worthwhile to explain why KBI calculation using KBFF models gives better results. The reason is that the force filed is parametrized to give correct KBI. However, this is achieved by the reproducing more homogeneous behaviour of RDFs tail, which is then translated in the correct behaviour of rKBI, and calculation of KBI. This model promotes mixing, and therefore it makes more homogeneous system, which is obvious from the fact that constancy of tail is reached for the system size used in the simulation. Other models, are more segregated, and therefore the stability of the RDF is affected, which then shows up as larger KBI. This is not an indication that other models are wrong, they just have stronger segregation, which affects fluctuation and the RDF s tail. The main difference is not at the large alcohol concentrations, as it would be expected, since models differ in ethanol force field, but at small alcohol concentrations. An explanation can be given, considering the differences between models: OPLS and TRAPPE have weaker charges, and they are 'less competitive' to join the hydrogen-bonded network, which becomes even more evident when there is less of them. In the region of small alcohol concentrations, many properties show extremes, which is an indication that in this region the system is more frustrated, and therefore more sensitive to the small changes.



(a) G_{EE}



(b) G_{EW}



(c) G_{WW}

Figure 41: Kirkwood-Buff integrals

3.5 Cluster analysis

One way of accessing microscopical segregation in liquids is via clustering analysis. Hill's theory [130] provides the statistical mechanics formalism to describe clustering in equilibrium classical system. Following the Hill's definition; two particles belong to a same cluster if they are connected through a path of bonded pairs of particles. There are different definitions of a bonded pair. Hill's definition of a bonded pair states that two particles are bonded if their relative kinetic energy is less than pair's negative potential energy. The Stillinger criterion which is a geometrical one, states that a pair is bonded if two particles in a given configuration are separated by distance less than some predefined distance d [131]. Vericat and Pugnaloni included dynamical criterion in definitions of physical and chemical clusters [132]. In this work Stillinger definition is used, as the prime goal is to elucidate static structural properties. Imposed geometrical criterion does not take into account the interaction between particles, the particles are considered bonded if their relative distance is less than some predefined value. However, the distinction will be noticed between bonded pairs of hydrophobic sites (site clusters) and bonded pairs of H-bonding particles (interaction clusters).
Following the definition of Coniglio et al [133] the cluster pair correlation function $g_\gamma(\vec{r_1}, \vec{r_2})$, cRDF, is the joint probability density of finding two particles that belong to the same cluster of kind γ at positions $\vec{r_1}$ and $\vec{r_2}$ respectively (definition taken from [132]).
The probability $p(n)$ of finding a cluster of size n is defined as:

$$p(n) = \frac{\sum_k s(k,n)}{\sum_{n,k} s(k,n)} \tag{3.40}$$

where $s(k,n)$ is the number of clusters of size n in the configuration k. Average cluster size is defined as:

$$n_{average} = n \sum_n p(n) \tag{3.41}$$

On Figure 42[9] it is visible that for cluster size from 4 to 10, the biggest probability is for the $\chi = 0.18$ mole fraction. It refines the picture of the structural transformation at approximately this mole fraction, where there is the largest number of small hydrophobic site clusters, and not ethanol clusters.

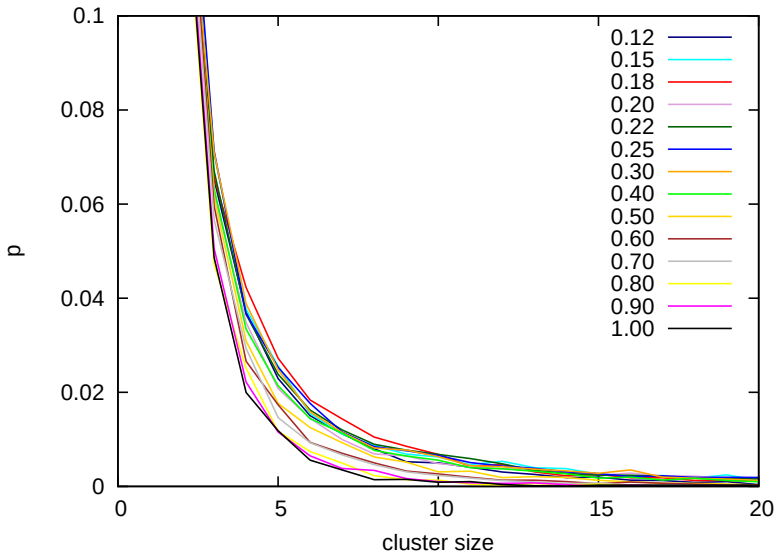


Figure 42: Cluster size probability function for CH_3 sites clusters

[9]All Figures in this section are from [134]

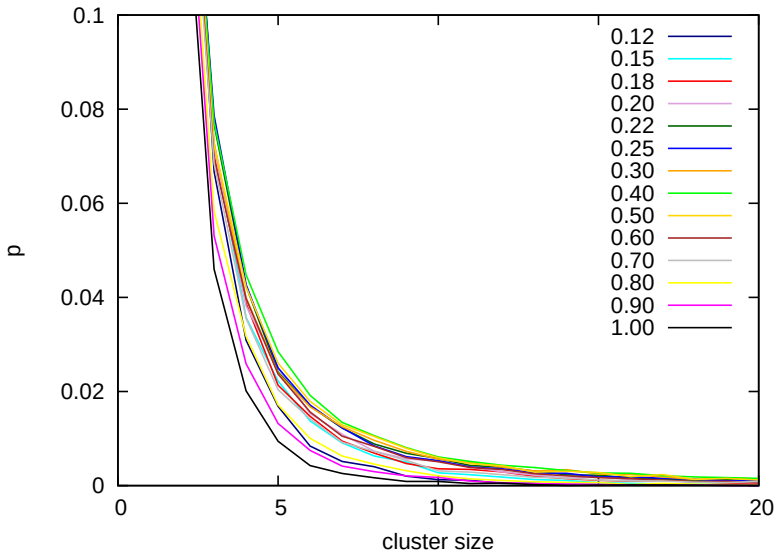


Figure 43: Cluster size probability function for CH_2 sites clusters

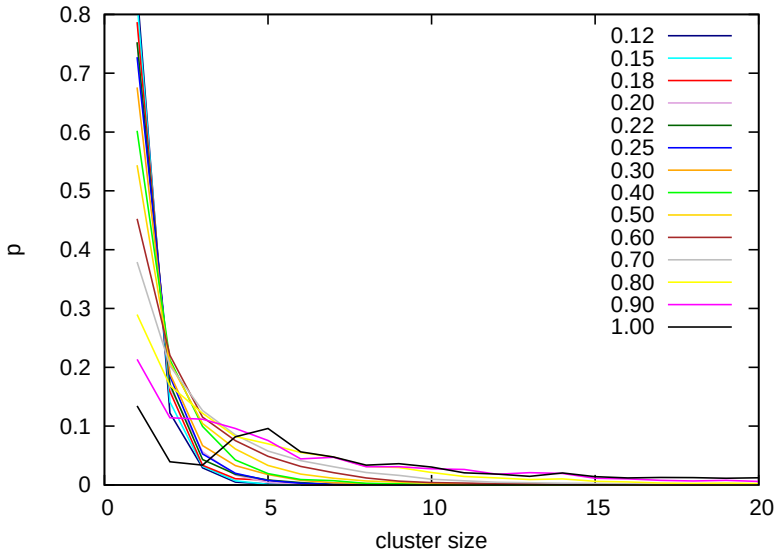


Figure 44: Cluster size probability function for O sites clusters

O clusters show preferential clustering at the size of approximately 5 for pure ethanol, and that structuring is somewhat preserved at $\chi = 0.90$ mole fraction.

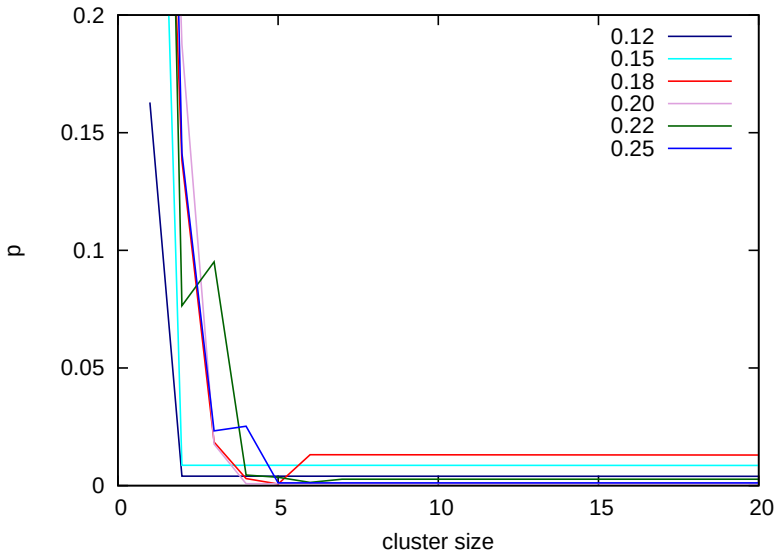


Figure 45: Cluster size probability function for Ow sites clusters, small concentrations

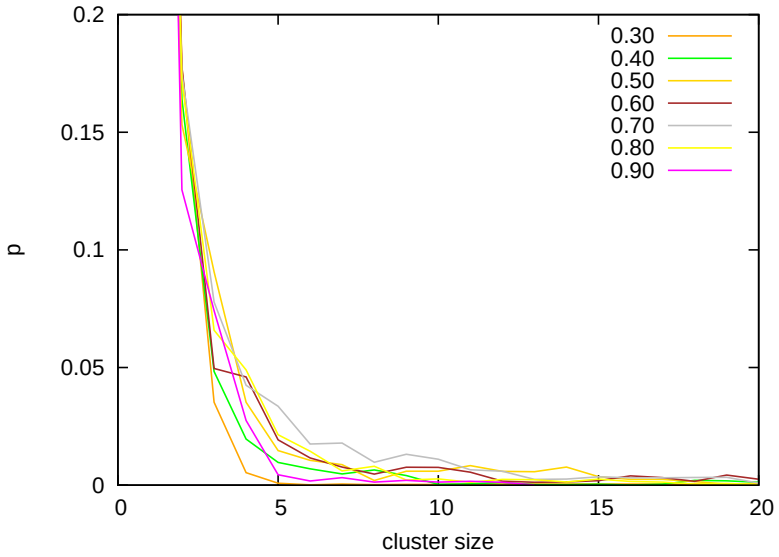


Figure 46: Cluster size probability function for Ow sites clusters, higher concentrations

The biggest and the average clusters are plotted on Figure 47. Average cluster size is a statistical measure defined by the whole distribution, while

the biggest cluster is one of the fluctuations of the cluster size. When the system is close to the percolation of some specific sites, site's clusters of the size that spans the system appear. Therefore, the size of the biggest cluster may be the sign of approaching the regime change. CH_3 average clusters show two regimes; one exponential up to $\chi = 0.30$, and linear above. Biggest CH_3 cluster has linear behavior, steeper line up to $\chi = 0.40$, and closer to the line $\chi \times 2048$ above. CH_2 biggest clusters behave almost exactly as the CH_3 average, exponential rise up to $\chi = 0.30$, and linear above. Average CH_2 clusters have exponential rise up to $\chi = 0.60$, and linear above. The oxygen ethanol clusters stay small in size over the whole mole fraction range. The exception is when approaching pure system, they show an increase, the biggest, and the average as a consequence, meaning that in the mixture there is negligible probability that ethanol O-sites would form a percolated network, as ethanol tends to preserve it's structure of the pure liquid, clustering O sites in clusters of approximately 5. Water cluster behaviour shows that water slightly changes its network's connectivity at $\chi = 0.30$, but the most dramatic change occurs at $\chi = 0.60$. Above $\chi = 0.60$ percolation of the water network is evidently broken as there are only small clusters present in the system. The change in the size of the average cluster is very rapid, shown by the angle that line forms at this mole fraction.

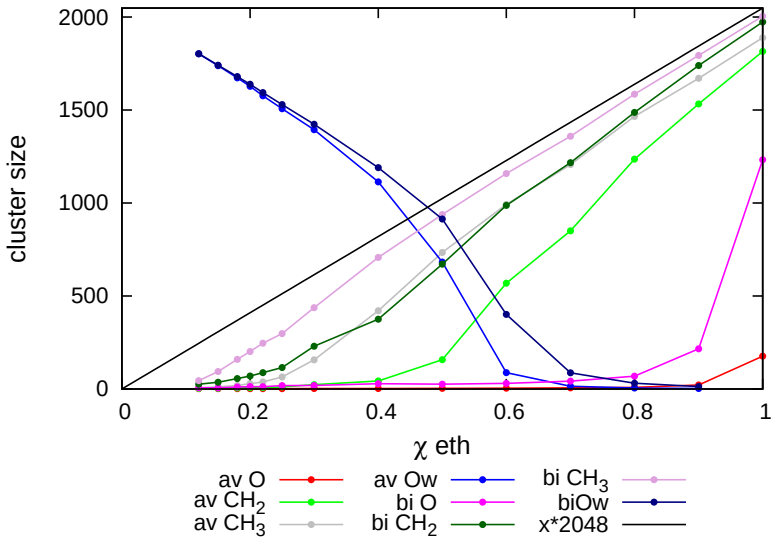


Figure 47: The biggest and average clusters

Cluster radial distribution functions are plotted on Figure 48 to 51[10]. As the total number of sites in the system, N, is used as the normalization factor, and if all of the N particles are not bonded in clusters, the cRDF oscillates below the line $y = 1$. On plots some of the functions have a sharp end of the first peak, which is a consequence of the used cutoff distance for the cluster definition. Long range behaviour of the cRDFs has three main types. First one is when the cRDF oscillates around a horizontal line, meaning that almost all sites are clustered, and sites form a homogenous distribution. Second behaviour is when the cRDF tail spans the half cell size, but it has an exponential decay. This means that while large clusters appear as a fluctuation, smaller clusters are predominant in the system. And the last type of behaviour is when exponential decay of the tail leads to zero before reaching the half cell radius, meaning that big, percolated clusters do not appear at all in the system, not even as a fluctuation.

[10]These plots are for the flexible OPLS model for ethanol. Other models are plotted in the Appendix, showing the same characteristics.

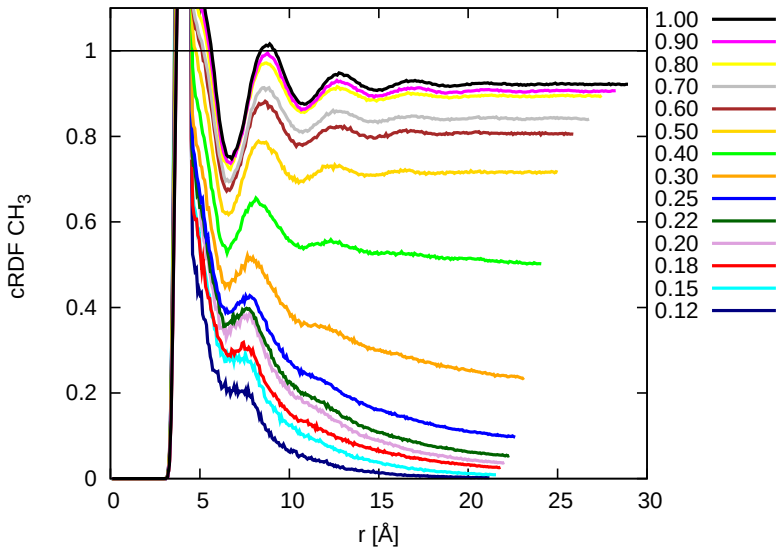


Figure 48: Cluster radial distribution function for CH_3 sites

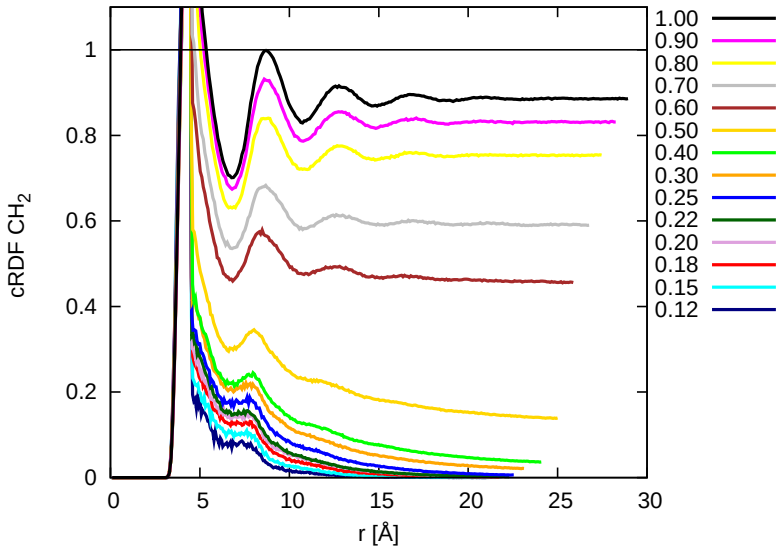


Figure 49: Cluster radial distribution function for CH_2 sites

Hydrophobic sites for high alcohol concentrations show horizontal cRDF's tails, therefore the sites are homogeneously organized. This random organization is lost for CH_2 sites at $\chi = 0.60$, and for CH_3 at $\chi = 0.40$ (or $\chi = 0.30$, depending on ethanol model). The cRDF for oxygen ethanol clusters show

that O sites have tendency to form smaller clusters, and their cRDF's tail show second type of behaviour.

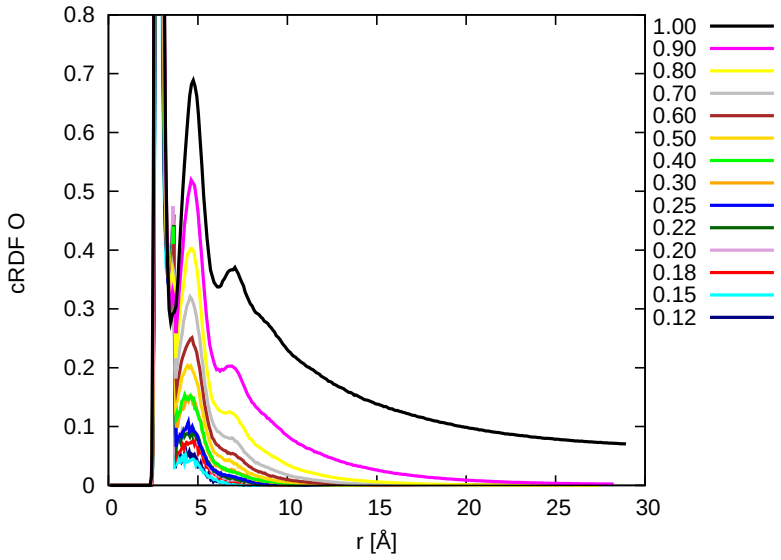


Figure 50: Cluster radial distribution function for O sites

In the range below $\chi = 0.60$ ethanol, water cRDFs are horizontal, meaning that water is percolated, its network spanning the whole system. At $\chi = 0.60$ mole fraction of ethanol the system approaches the structural change and water cRDF tail has an exponential decay, as the big percolated clusters become rarefied. Above $\chi = 0.60$ water network is broken is smaller fragments, and cRDF decays to zero before reaching the end of the simulation cell.

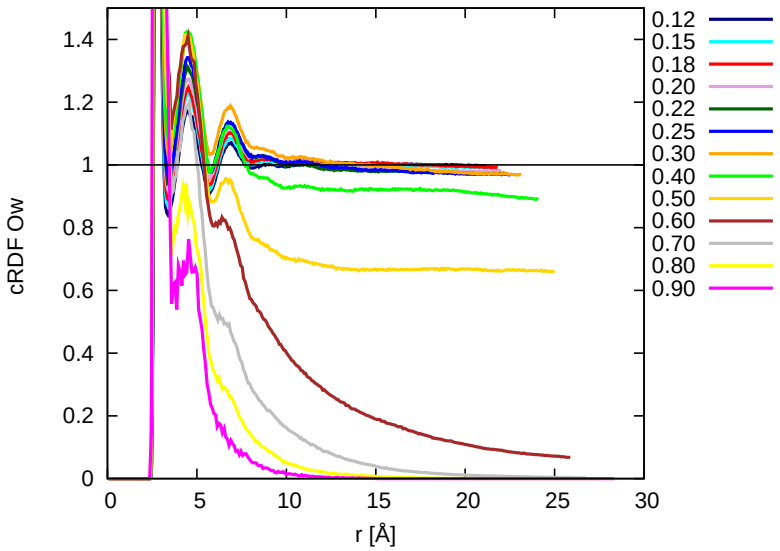


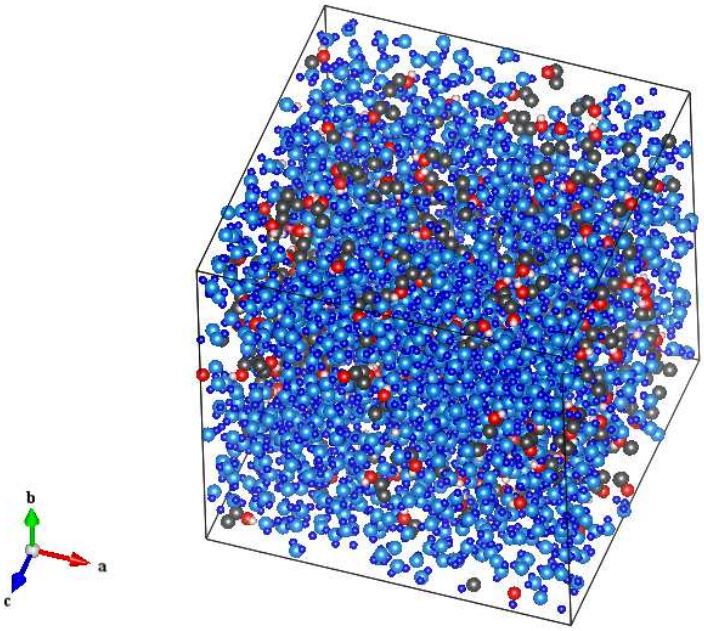
Figure 51: Cluster radial distribution function for Ow sites

Summing up all the information from clusters it can be concluded that the change at approximately $\chi = 0.15$ mole fraction of ethanol is due to the change of the hydrophobic sites organization, as they are the most clustered at this concentration, and from this concentration on they manage to achieve better degree of randomness. Second change is due to the water reorganization, namely the breaking up of the percolated water network into smaller domains. All of the tested models show the same behaviour, with just small variations, so it can be concluded that described properties are not model sensitive.

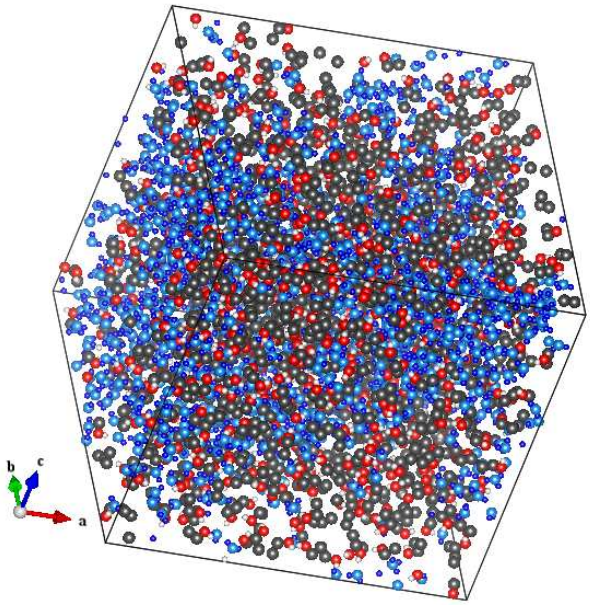
3.6 Snapshots

Snapshot, the instantenous configuration of the system, (all of them in this work are produced by VESTA 3 application [135]), is not statistically averaged feature.. However, it is reasonable to assume that one state randomly chosen will show any particular global structural organization that is seen in the averaged properties of the system. On Figure 52 system at two mole fractions is shown[11]:

[11] On all the Figures in this section plotted sites are 1/3 of their real size, for the reason of visibility

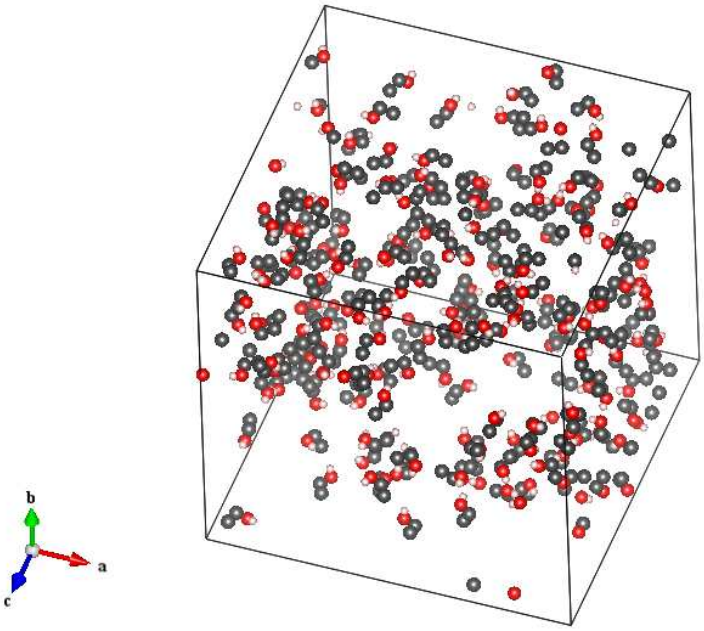


(a) $\chi = 0.10$

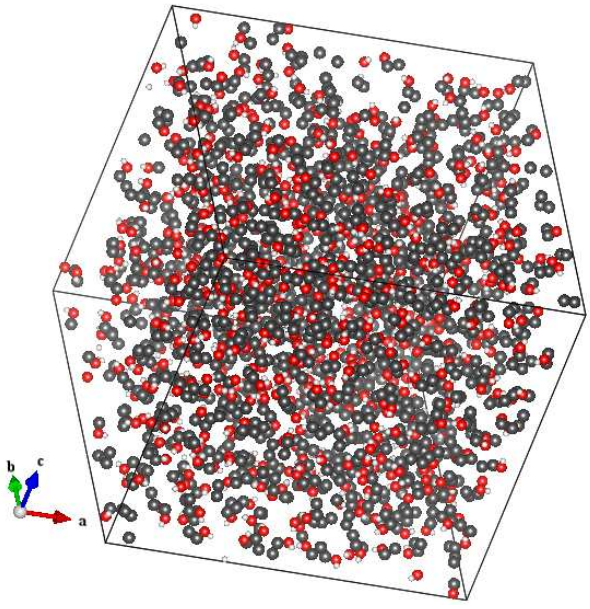


(b) $\chi = 0.40$

Figure 52: Snapshots of ethanol(red-black)-water(blue) for two different mole fractions of ethanol



(a) $\chi = 0.10$

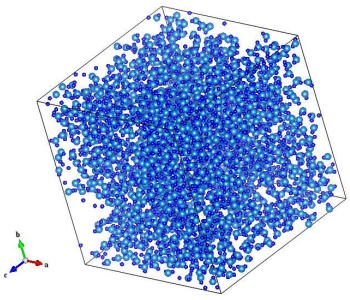


(b) $\chi = 0.40$

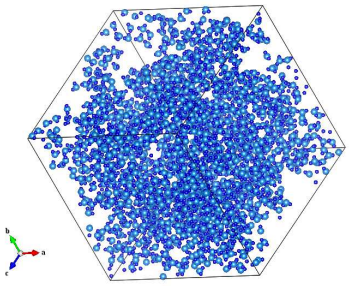
Figure 53: Snapshots of ethanol molecules for two different mole fractions of ethanol

In the Figure 52 (a) the small ethanol clusters are seen in the water, with

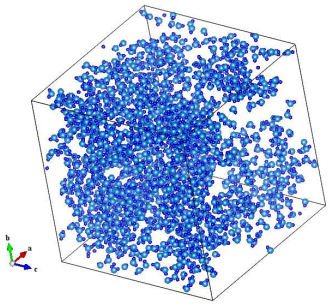
some single ethanol molecules still present in the system, while in the region of bi-continuous microheterogeneous mixture at Figure 52 (b) these clusters are merged into bigger ethanol associations. On Figure 53 there are only ethanol molecules at the same mole fractions, and small ethanol clusters and single molecules are easily identified at 53 (a). In order to track down the change in the water network at approximately $\chi = 0.65$ mole fraction, it is instructive to look at water molecules along the changing mole fraction on Figure 54.



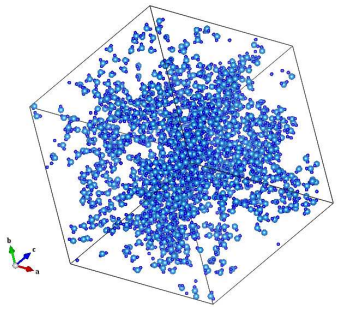
(a) $\chi = 0.10$



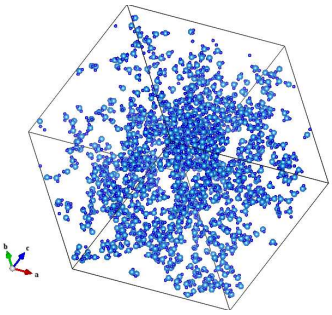
(b) $\chi = 0.20$



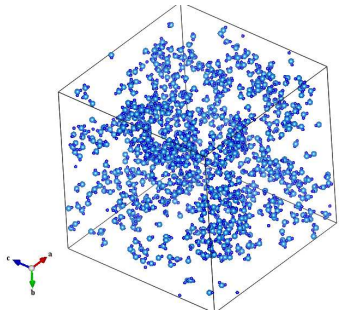
(c) $\chi = 0.30$



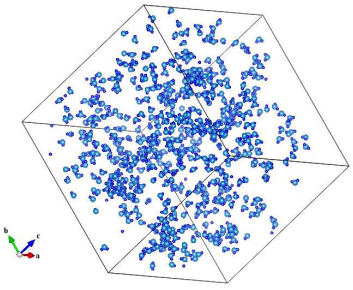
(d) $\chi = 0.40$



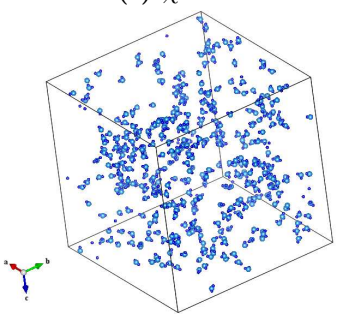
(e) $\chi = 0.50$



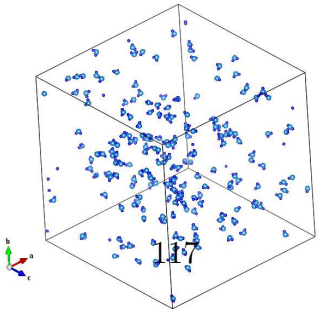
(f) $\chi = 0.60$



(g) $\chi = 0.70$



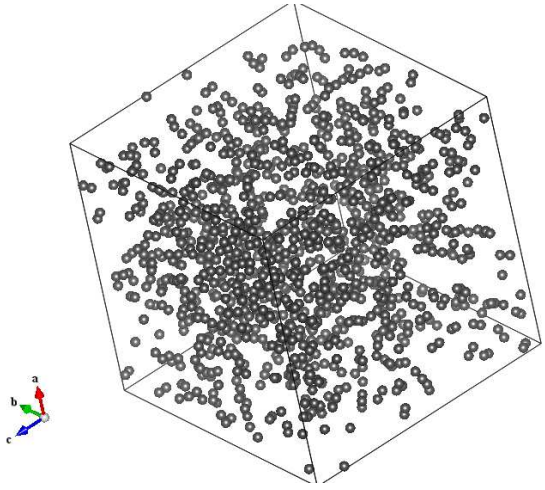
(h) $\chi = 0.80$



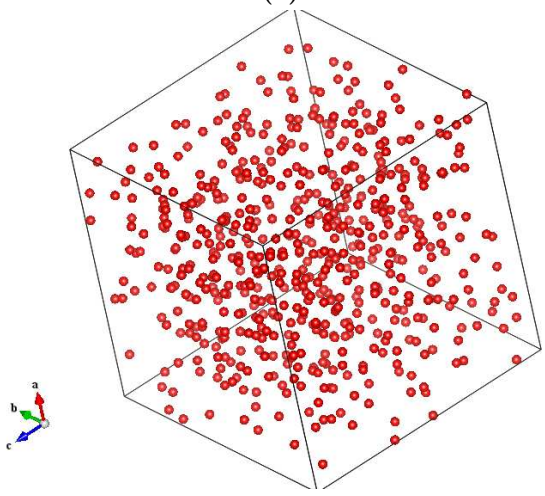
(i) $\chi = 0.90$

Figure 54: Snapshots of water molecules for nine different mole fractions of ethanol

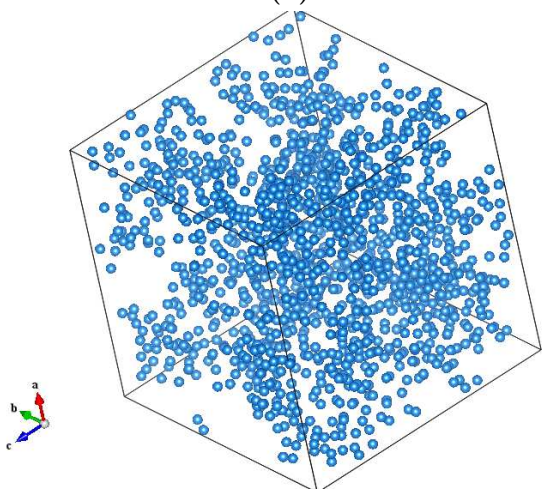
At Figure 54 (a) water seems to fill the whole volume of the cell uniformly. From Figure 54 (b) to Figure 54 (f) thinning of the water network is visible, but the percolation seems to hold over all of this mole fraction range.
Microheterogeneous quality of the water network is clearly visible in this mole fraction range. Microheterogeneous areas of water (or holes in water network) do not have any kind of definable geometry, but regions with water and regions without water seem to form shapeless pattern of interlaced areas. This is the reason why it is not expected to observe a prepeak in the structure factor, because there is no specific size that can be attributed to these interlaced microsegregated domains. From Figure 54 (g) on, breaking of the water network becomes visible, and especially fragments of two molecules that correspond to the shift to the right of the Ow-Ow RDF second peak, mentioned before.
When looking at ethanol oxygen sites and ethanol CH sites the mole fraction of $\chi = 0.30$ in comparison with the water oxygen sites at $\chi = 0.70$, at Figure 55, similarities between Ow and CH sites are more obvious, both having bigger empty areas in their space distribution, while ethanol oxygens microstructure can be thought of as microheterogeneity of a smaller scale, having finer granularity than the other two sites structure.



(a) CH



(b) O



(c) Ow

Figure 55: Snapshots of ethanol sites at $\chi = 0.30$ and water oxygen site at$\chi = 0.70$ mole fraction of ethanol

This is expected, as ethanol oxygens are distributed on the interface toward water, while CH sites tend to be shielded further away from water.

3.7 Diffusion

Self diffusion coefficients are dynamical properties of mixtures that describe translational mobility of molecules in the mixture relative to the similar molecules. Self diffusion coefficients can be calculated from molecular dynamics simulation by using Green-Kubo relation (velocity autocorrelation function):

$$D_i = \frac{1}{3}\int_0^\infty \langle v_i(t_0) v_i(t_0+t)\rangle dt \tag{3.42}$$

or Einstein relation (mean square displacement):

$$D_i = \frac{1}{6}\lim_{t\to\infty}\frac{d}{dt}\langle [r_i(t_0) - r_i(t_0+t)]^2\rangle \tag{3.43}$$

[79], [136].

Self diffusion coefficients calculations for different models from the literature show that this feature is not easily reproducible. Wensink et al [53] calculated self diffusion coefficients for mixture of all-atoms OPLS and TIP4P models, using Einstein relation. Noskov et al [48] reported self diffusion coefficients for polarizable ethanol model. Self diffusion coefficients for the mixture of OPLS-AA and SPC water were reported by Zhang et all [57], calculated by Green-Kubo method, and Zhang and Yang [58] reported them calculated from MD simulation of rigid ethanol model and TIP4P water by both, Green.Kubo and Einstein relation. Guevara-Carrion et all [109] reported selfdiffusion coefficients calculated by Green-Kubo relation. All of these results for water are presented on Figure 56, together with experimental results from [4] and for ethanol on Figure 57, together with experimental results from [5]. All models except polarizable [48] and rigid [109] with TIP4P/2005 water overestimate diffusion coefficients for both, water and ethanol. Polarizable model underestimates it, while results for the rigid model from [109] are in excellent agreement with experimental data for ethanol, and slightly underestimated for water, but better than polarizable model.

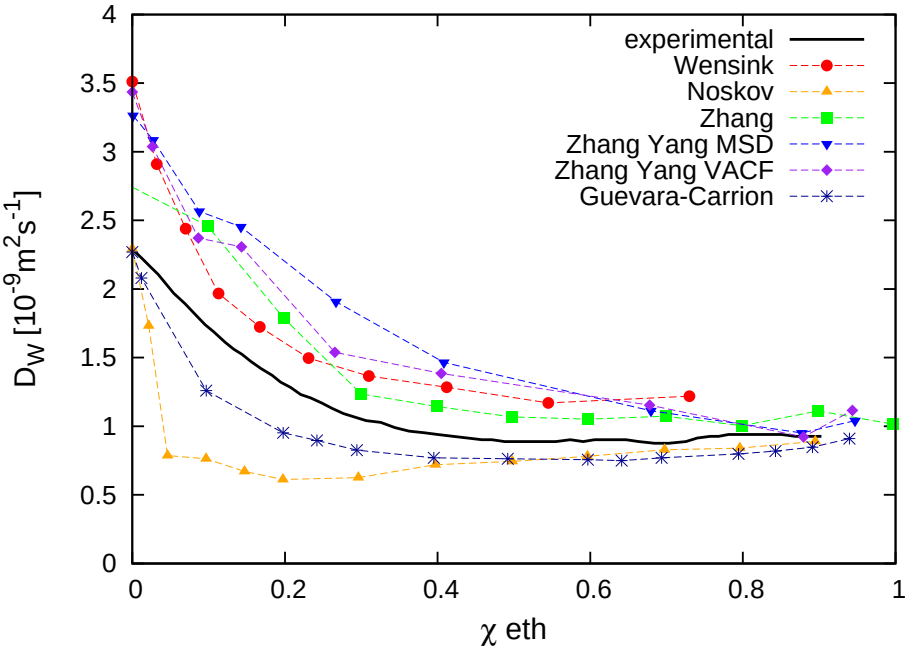


Figure 56: Self diffusion coefficient for water, experimental data from [4].

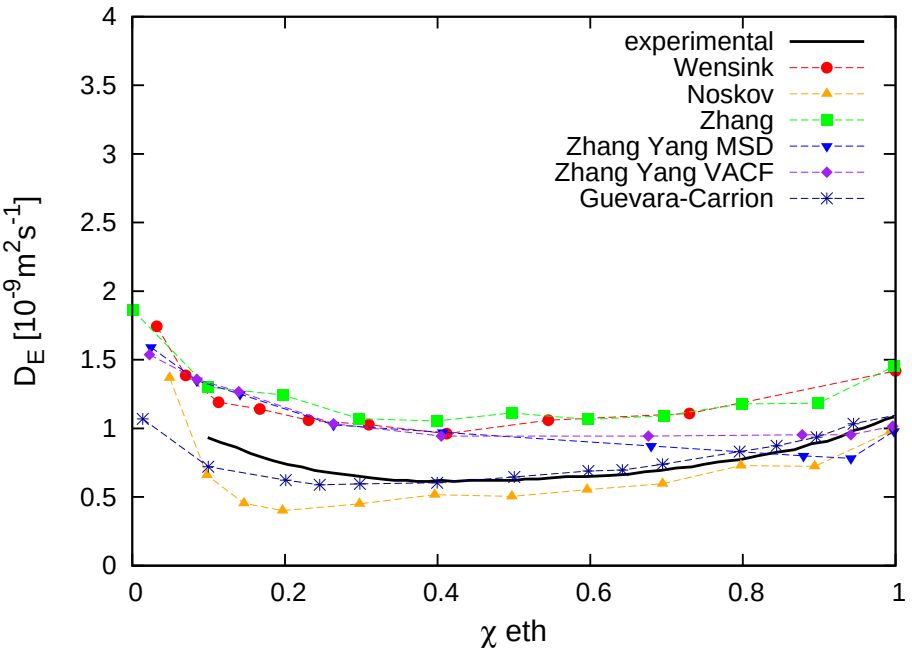


Figure 57: Self diffusion coefficient for ethanol, experimental data from [5].

Figures 58-59 apresent results for different ethanol models. All the data follow the experimental line, even though they are all slightly too high. The best results are from KBFF model, as expected. This can be easily explained with the fact that excess Coulomb energy is more positive for OPLS and TraPPE models, meaning enhanced repulsive interaction, resulting in slightly faster system, while KBFF model gives slightly slower system.

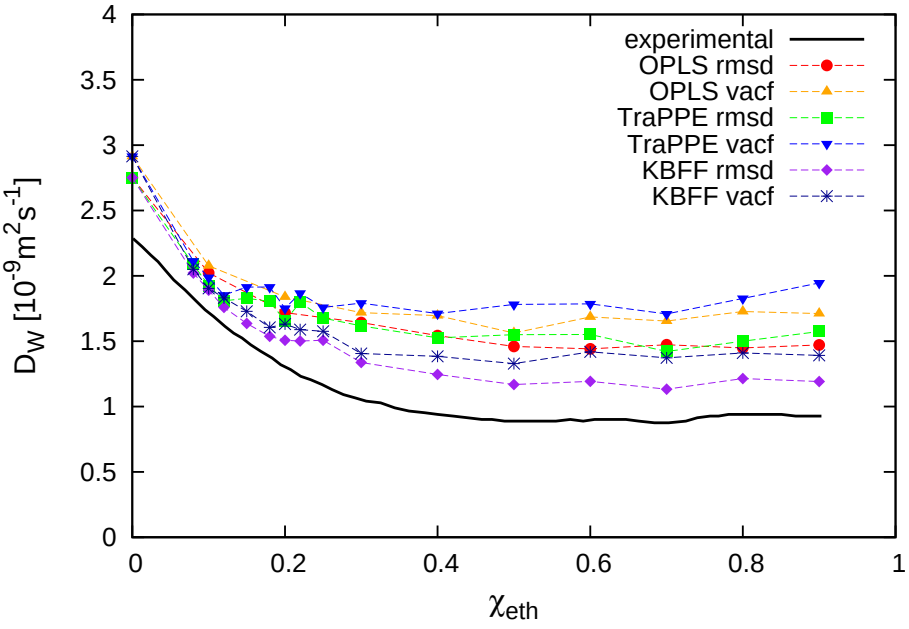


Figure 58: Self diffusion coefficient for water.

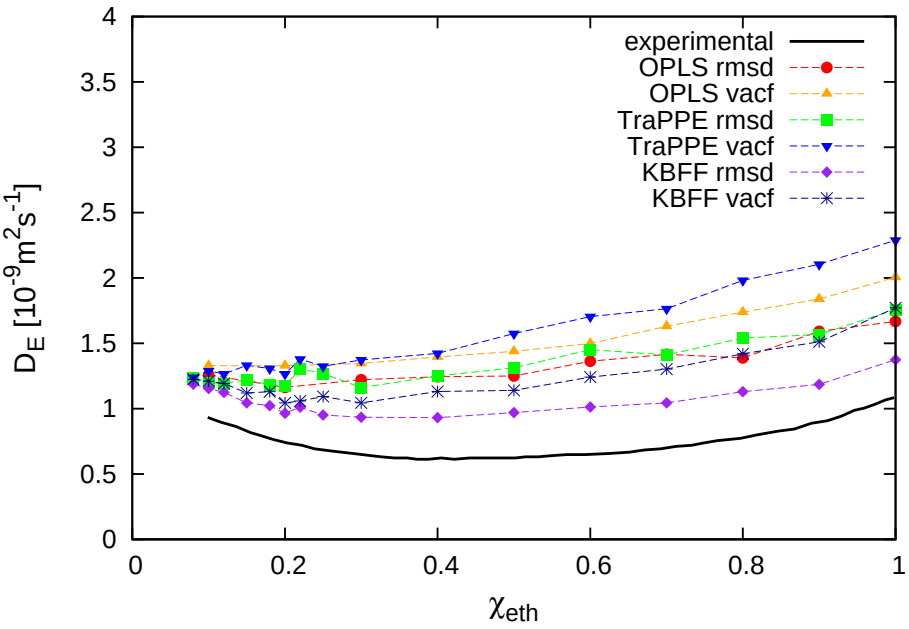


Figure 59: Self diffusion coefficient for ethanol.

3.8 Additional experimental results

Experimental results of the behaviour of various properties in the aqueous ethanol are known to undergo non-trivial changes with the change of the ethanol-water concentration. In this work a few of them are highlighted: speed of sound, excess enthalpy, heath capacity, compressibility and azeotropy. All of the listed properties exhibit the behaviour that has similarities with three regimes separated by two mole fraction (namely $\chi_1 \approx 0.15$ and $\chi_2 \approx 0.65$ mole fraction of ethanol) where the transition from one regime to another occurs in this mixture, that is proposed as the hypothesis of this work. It is justified to ask if that changes can be tracked in the structural changes that are rewived it the context of the acquired MD data.

3.8.1 Speed of sound

Recent highly accurate results from the measurement of ultrasound and hypersound speed have been obtained from the group's collaboration work [12] [6], [137]. Ultrasound is defined as a sound wave with frequency from $2 * 10^4$ to 10^9 Hz. Higher portion of the frequency spectrum, from 10^9 to 10^13 Hz is defined as hypersound.

On Figure 60 is the variation of ultrasonic sound speed in ethanol-water mixture, at different mole fractions, and at different temperatures, from [6]. In the range from 0℃ to 70℃ sound speed in pure water and pure ethanol exhibits very different behaviour. In the case of pure ethanol it is decreasing linearly as the temperature increases. This is expected behaviour, as the increased thermal disorder is expected to lower the speed at which the sound propagates through the media. In water sound speed increases in a nonlinear fashion at this temperature range. This fact is one of the many known anomalies of liquid water [138], [139]. In this whole temperature range speed of sound is greater in pure water than in pure ethanol. Both of these liquids are hydrogen bonded, but they have different qualities of hydrogen-bonded structures. Water kind of hydrogen-bonded association is more rigid and

[12] with Austrian group from University of Salzburg: A. Asenbaum, C. Pruner and E. Wilhelm

more uniform, so the sound propagates faster in water, than in ethanol. Ethanol has hydrophobic part that prevents it from forming the same kind of association as water.

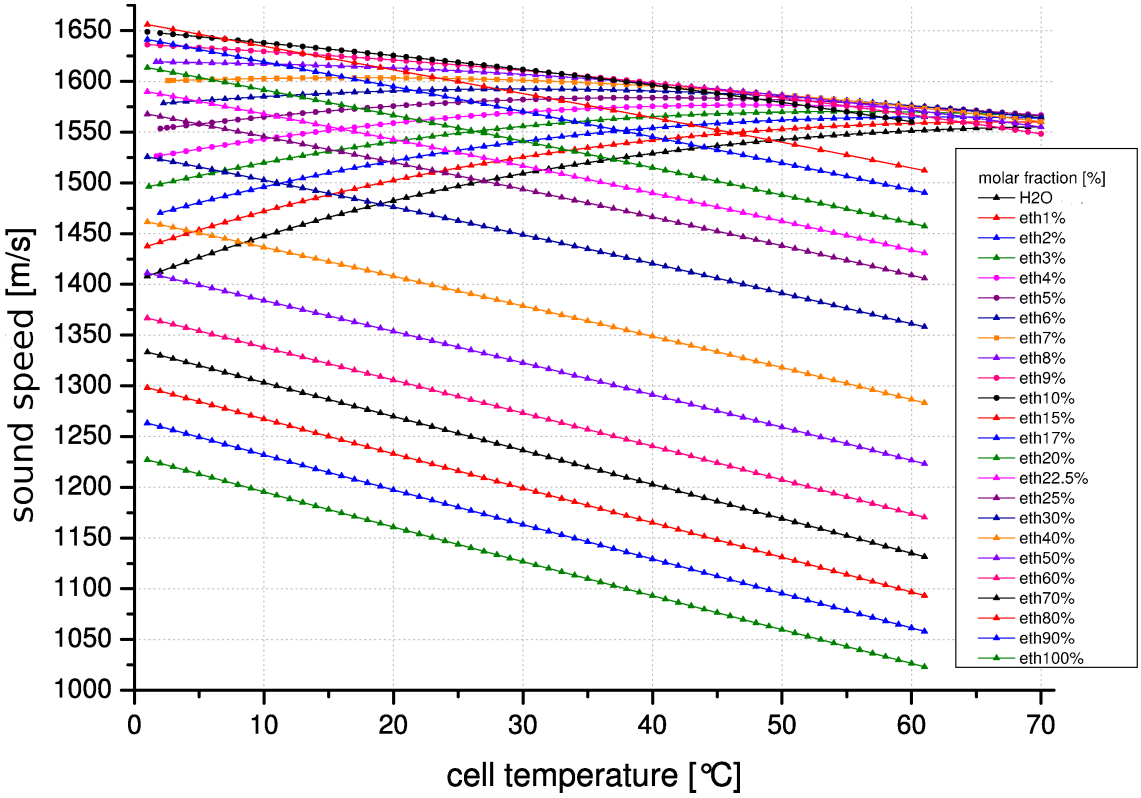


Figure 60: Speed of ultrasonic sound in ethanol-water system, as function of temperature and molar fraction, from [6]

When ethanol and water are mixed, the speed of sound goes through a maximum, that manifests itself at the mole fraction of approximately $\chi_1 \approx 0.15$ ethanol, at ambient conditions, as seen in Figure 61. For higher temperatures this maximum is slightly shifted to the smaller concentrations of ethanol. In the first region below $\chi_1 \approx 0.15$ addition of ethanol to water is not pushing the sound speed down towards ethanol values, but on the contrary, it enhances the water's network structure in a way that sound propagates even faster in this mixture than in pure water. This picture implies that in this region ethanol is clustered in small domains incorporated into flexible water network. In the middle region sound speed starts to drop

down toward pure ethanol value, as ethanol domains begin to grow bigger in the water network. After $\chi_2 \approx 0.65$ water network is not percolated any more and speed of sound starts to drop linearly.

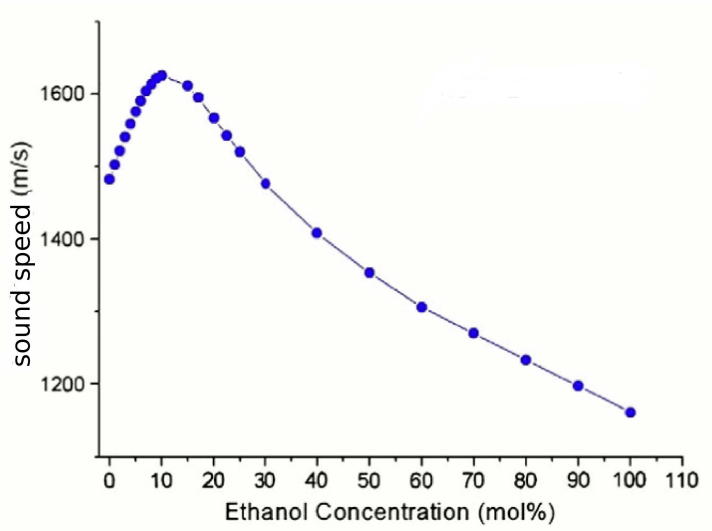


Figure 61: Speed of ultrasonic sound under ambient conditions as function of the ethanol mole fraction, from [6]

Another interesting feature of the sound speed in this mixture is dispersion effect (Figure 62).

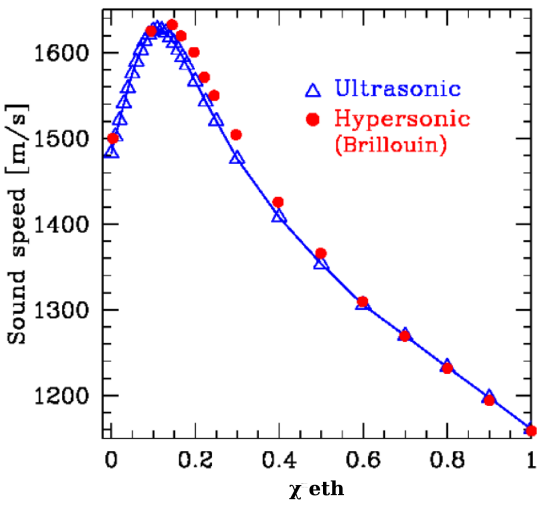


Figure 62: Ultrasonic (triangles) and hypersonic (dots) sound speed, from [7]

Dispersion effect is the difference in ultrasound and hypersound speeds, that can be seen in the region up to $\chi_2 \approx 0.65$ [140]. The higher frequency hypersound speed is higher than ultrasonic sound speed. It is interesting to note that dispersion effect is the biggest around first proposed structural change at $\chi_1 \approx 0.15$, indicating that this effect is somehow related to it. Dispersion effect vanishes above $\chi_2 \approx 0.65$. To reconcile all of the data from the experimental sound speed measurements, the increment of the sound speed at small ethanol mole fractions, meaning the rigidifying of the mixture, and the high frequency dispersion effect that suggests the presence of inhomogeneities, it is reasonable to conclude that in the region from 0 to χ_1 water network is rigidyfied by adding ethanol molecules, that cluster themselves in small clusters. As more ethanol is added to the mixture, the more of this clusters is formed, subtracting areas accessible to the water network's flexibility, as this regions occupied with ethanol molecules can't be more squeezed, thus acting as hard spots in the flexible water network.
When negative excess sound speed is compared to the previously discussed excess enthalpy (see Section 3.1.5) at Figure 63 it can be seen that both lines have minimum at approximately the same mole fraction of ethanol, $\chi_1 \approx 0.15$, and an inflection point at $\chi_2 \approx 0.65$.

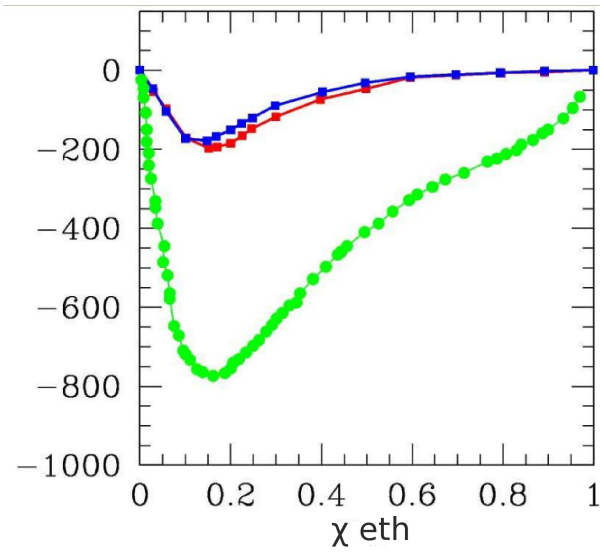


Figure 63: Experimental excess enthalpy ftom [8] as green circles (in J/mol) and negative excess sound speed from [6] as squares (blue for ultrasonic speed and red for hypersonic speed (in m/sec))

3.8.2 Response functions

Isothermal compressibility, κ_T, is a measure of the relative change in volume of a system due to the change in pressure, for an isothermal (dT=0) process. Isentropic compressibility, κ_S measures the change in volume of a system due to the change in pressure during a reversible adiabatic (dS=0) process. Adiabatic process is a process in which a system does not exchange heat with its surroundings.

$$\kappa_T = -\frac{1}{V}(\frac{dV}{dp})_T \tag{3.44}$$

$$\kappa_S = -\frac{1}{V}(\frac{dV}{dp})_S \tag{3.45}$$

Data from [9] on isothermal compressibility of the ethanol-water mixture at 25℃ shows the minimum in the compressibility at $\chi_1 \approx 0.15$ (Figure 64) that supports the previous conclusion, as the minimum in the excess compressibility would confirm the forming of the highest number of incompressible regions in the water network. Excess in isentropic compressibility from [10] (Figure 65)

also has a minimum at approximately the same mole fraction. At $\chi_2 \approx 0.65$ isothermal compressibility has an inflection point, while excess isentropic compressibility has a maximum that falls at the higher mole fraction.

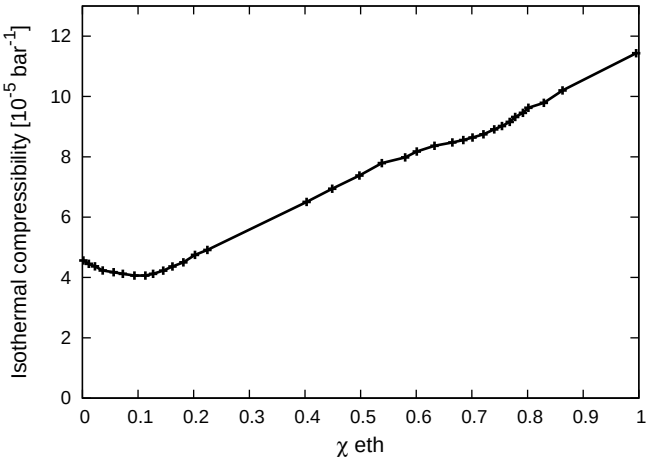


Figure 64: Isothermal compressibility of the ethanol-water mixture, from [9]

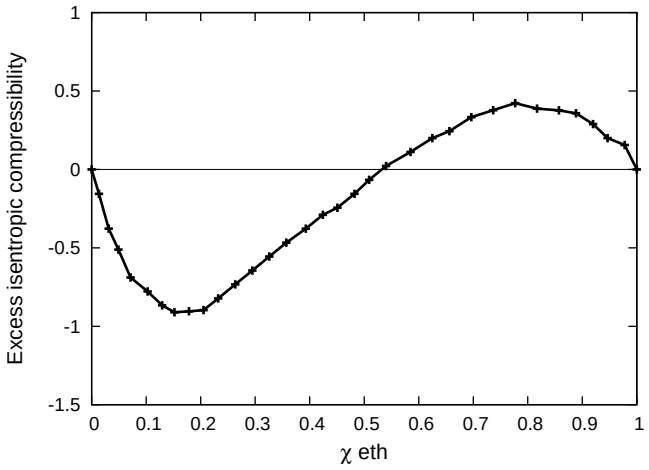


Figure 65: Variation in excess isentropic compressibility with ethanol mole fraction, from [10]

with V=volume, T=temperature, p=pressure and S=entropy of a system. Constant pressure heat capacity is amount of heat that needs to be added to

the system to increase the temperature for a given unit, at constant pressure. Molar heat capacity is heat capacity per mole of substance.

$$C_p = (\frac{dQ}{dT})_p \tag{3.46}$$

with Q=heat. Excess molar heat capacity (Figure 66) also marks both proposed significant mole fractions, it has maximum at $\chi_1 \approx 0.15$ and inflexion point at $\chi_2 \approx 0.65$.

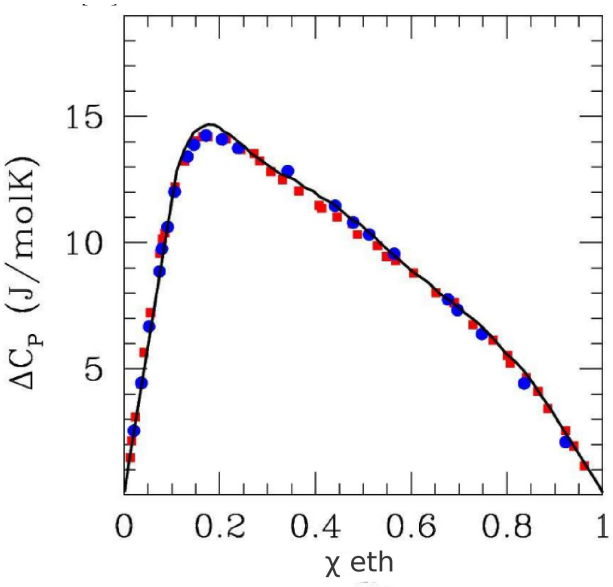


Figure 66: Excess heat capacity. Line from [8], squares from [11] and dots from [12]

3.8.3 Azeotropy

Mole fraction of ethanol in liquid aqueous ethanol vs. mole fraction of ethanol in vapour phase is shown on Figure 67. The curve is called azeotropy curve, while the point where it crosses x=y line is called azeotropy point. Azeotropy line shows maximum at $\chi_1 \approx 0.15$ and inflexion point at $\chi_2 \approx 0.65$ mole fraction of ethanol in liquid phase.

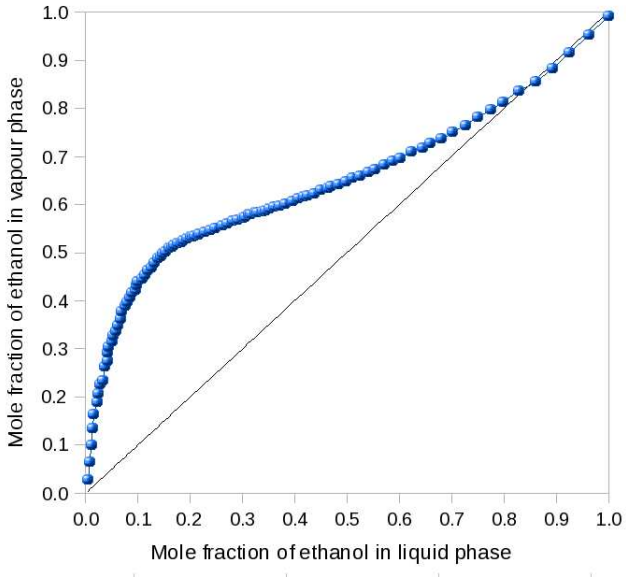


Figure 67: Vapour - liquid equilibrium diagram for ethanol - water. Data from [13]

Maximum in the mole fraction of ethanol molecules in the vapour phase at χ_1 points to the structural organization of the liquid phase in the way that ethanol molecules are the easiest to evaporate from the liquid at this mole fraction, supporting the picture of the highest number of small ethanol clusters embedded into water network. After that point, bigger associations of ethanol are starting to form, thus diminishing the ratio of the ethanol to water molecules that are evaporating. Change in convexity at χ_2 supports the breaking up of the water network into smaller fragments, which are now easier to evaporate water molecules at the higher rate.

3.8.4 Summary of additional experimental data

It is interesting that such different physical properties as sound speed, excess compressibility, excess enthalpy, excess heat capacity (Figure 66) and azeotropy, when plotted against mole fraction of ethanol in the ethanol - water liquid mixture, all show non-trivial behaviour with two approximately the same significant points. This implies that the same changes in micro-structure

of the system are responsible for all of them. Proposed picture of three structural regimes is in accordance with the observed data. However, all the extrema in the data presented in this section do not fall exactly at the same mole fractions. It is still justified to claim that the same structural change that takes place along the change of the ethanol's mole fraction is responsible for all observed features, as this features are all different things, heat capacity is enthalpic effect and speed of sound has more to do with the geometry of the structure, so it is not expected for all of them to register the change in the structure at the same, sharply defined point.

4 Conclusion

The Ethanol-water mixture was discussed in this work by the means of simulation and experimental data. Simulation data on static structural properties was found to identify three structural regimes; first up to $\chi_1 \approx 0.15$ mole fraction of ethanol, second between $\chi_1 \approx 0.15$ and $\chi_2 \approx 0.65$, and third above $\chi_2 \approx 0.65$. This was in agreement with two mole fractions in this mixture where various experimental properties exhibit extrema or inflexion points, namely: sound of speed, excess enthalpy, excess heath capacity, compressibility and azeotropy curve. The aim of this work was to achieve the understanding of the microscopic structure picture, and to connect it with the changes observed in experimental properties. Behind this non-trivial behaviour of experimental properties can be the recently introduced new concept of microheterogeneity of aqueous mixtures. Microheterogeneity is defined as the local imiscibility of species in a mixture, that appears homogenous at macro scale. It is important to point out that microheterogeneity is the property of the system in equilibrium. Formation of domains is common in system that is out of the equilibrium, for example when system is close to phase separation. Microheterogeneity should be distinguished from the concentration fluctuation, as concentration fluctuation is statistical property that represents the fluctuation in number of particles of a given species, when looked through a window at various realizations of the system. It is the variation in number of particles of one species in a given ΔV in various microstates of the system. KBI measures concentration fluctuation in a system. Even the microscopically homogenous mixture posses concentration fluctuation, because the system is not frozen, it goes through various realizations in time. When a system has microsegregated domains, they will also look like concentration fluctuation, but they will have a permanency to them. In the micelle forming systems these domains look like large particles. In the aqueous mixtures microsegregated domains are of not such a specific shape. Even further, in this work it was shown how the microheterogenous structure changes it's form with the change of mole fractions of mixture's constituents.

First indications of the existence of microheterogeneity were the data on experimental entropy that was too small, so this fact points out to the existence of some order in the system. Only in 2002 Soper et al. [16] published a paper about microheterogeneity in methanol-water mixture, so it became 'official'. In that paper microheterogeneity was discussed via radial distribution functions obtained from scattering experiments. Since then there were lots papers dealing with microheterogeneity, for example Zoranic et al. [62]
discussed microheterogeneity in aqueous amides, but it was never truly separated from concentration fluctuations, as it is not clear how to extract microheterogeneity from RDF. In microsegregated systems microheterogeneity is the separation of species observed at snapshots, and concentration fluctuation is variation in time of this microsegregated domains. So it is related to variations in clusters, which are not well defined [113]. In this work the tool for separating microheterogeneity from concentration fluctuations is introduced for the first time. Namely it is pair connectedness function (cRDF), that uses clusters, as they are the statical property of the structure, to identify microsegregated domains. With its help it was concluded that microheterogeneity is responsible for the existence of the three structural regimes in this mixture, and for the first time here is given microscopic picture that describes these three different microheterogenous structures.
Molecular dynamic simulation is one of the tools that allows accessing microstructural properties of liquids. Though, it has some difficulties, one being the size of the simulated systems. For systems of associated liquids, as water and alcohol, it is not clear how big are the microsegregated domains in the mixture, as they appear to be of no specific size and shape. For instance; the prevalent picture of pure water is that of the flexible network that spans the whole system. In this mixture up to $\chi = 0.65$ mole fraction of ethanol water seems to span the whole system with its percolated network, so it is hard to speak of any specific size of its domains/clusters. As water preferentially bonds to water, and 'tries' to keep it's H-bonded network undisrupted, it can be identified as a 'driving force' in aqueous alcohols systems.

Another difficulty is the time of the simulations, due to the unknown dynamic of the microheterogeneity. The example is tail of the radial distribution function that shows some changes even after 4 ns statistics, which is a relatively long time for the MD simulation of 2048 molecules.

Yet the most important question is the choice of the force fields for the simulation. Classical force fields parametrized on pure components, as OPLS and TraPPE, seem to overestimate microsegregation in this system, even though they give better pure component's properties. The KBFF force field parametrized on mixture's Kirkwood-Buff integrals gives interface between domains of apparently more realistic size. However, all the force fields gave the similar structural and clustering properties. The differences can be seen at coordination numbers, radial distribution functions, and the most obvious at the value of Kirkwood-Buff integrals. The energetic contribution differs the most in the excess Coulomb energy and consequently in the excess enthalpy of the system. However all of the force fields in this work were able to track restructuring of the mixture at approximately the same mole fractions that are in line with the mole fractions where previously mentioned experimental properties exhibit extrema or inflexion points.

The short description of three regimes structures is presented next, along with the listed facts from simulation and experimental data that supports the given picture.

In the first region, from pure water up to $\chi \approx 0.15$, water pushes away CH_3 sites, which is a known hydrophobic effect. This effect can be tracked to the properties of the CH_3 site-site clustering. Probability for the CH_3 clusters of the sizes from 4 to 10 is rising up to $\chi \approx 0.15 - 0.18$, where it is the highest, then it drops down. The CH_3 average cluster sizes are very small and their concentration dependance has non-linear shape. The cRDF's first peak for CH_3 sites is the highest at $\chi = 0.15$ indicating thy are the most clustered at this mole fraction. Exponential decay to zero of CH_n sites cRDFs indicates that these sites are grouped in smaller clusters. The CH_2 and CH_3 coordination numbers exhibit a steep rise in this region, steeper than for the Lennard-Jones liquid, reflecting the fact that hydrophobic sites are distributed in an inhomogeneous way. First peak height of the correlations

of the CH_3-CH_3 and CH_2-CH_2 sites also rises in this region. The behaviour of the experimental data also marks the $\chi \approx 0.15$. Speed of sound rises from the value of the pure water up to $\chi \approx 0.15$ mole fraction of ethanol. Compressibility has a drop in this region. This drop occurs because system is more rigid due to the fact that hydrophobic effect has pushed ethanol molecules in small less compressible regions. Mole fraction of ethanol in the vapour phase to mole fraction of ethanol in the liquid phase ratio also rises up to a maximum value at $\chi \approx 0.15$ mole fraction of ethanol, indicating that ethanol molecules are the easiest to evaporate. Minimum in excess enthalpy and maximum in excess heat capacity both occur at this mole fraction, resulting from the added effect of the negative vdW excess and the minimum in the excess Coulomb energy. The minimum in the excess Coulomb energy means that the opposite charges have the best positions relative to each other: all water molecules forming percolated network, supported with the maximum number of H-bonded ethanol oxygenes, with all CH_2 sites away from positive charges, and CH_3 sites shielded from water. With the adding of more ethanol entropy maximization tendency pulls hydrophobic parts of ethanol molecule and overcomes energetic effects, resulting in formation of the bi-continuous microsegregated ethanol-water phase in the middle region. Middle region, from $\chi \approx 0.15$ to $\chi \approx 0.65$, is an intertwined ethanol water phase. With the rise of the ethanol mole fraction, the first peaks of RDF for hydrophobic sites go down, indicating their more homogenous placement than in the first region. Water network is thinning, as it can be seen from the size of the average water cluster. Up to $\chi = 0.60$ big system-size clusters appear as a fluctuation, the tail of the water's cRDF has an exponential decay, but it still spans the whole system size. The size of the biggest water cluster has a large drop from $\chi = 0.50$ to $\chi = 0.60$ mole fraction of ethanol, marking the second border between regimes. Speed of sound decreases in this region as the water network becomes thinner, excess enthalpy goes to the less negative values, as the positioning of the charged sites becomes disrupted by the growing hydrophobic parts of the mixture. The ratio of mole fraction of ethanol in the vapour phase to liquid phase becomes smaller, as the thinning of the water network acts in favour of releasing more and more

water molecules from liquid. At approximately $\chi_1 \approx 0.65$ a border between regions can be identified, that corresponds to the breaking of the connectivity of the water network. Azeotropy curve has an inflection here that suggests that there is a change in favour of easier release of water molecules from the liquid. Excess enthalpy and heath capacity also both have inflection point roughly around this mole fraction.
In the third region cRDFs for Ow sites have exponential decay and their tails go to zero before reaching the end of the cell, indicating that connectivity of the water's network is broken. The same thing can be observed from the biggest and average cluster sizes. So it is pictured as a region with smaller fragments of thorn up water network immersed in ethanol, that is visually confirmed with snapshots of the water. The same can be seen from the average size of the Ow clusters. In the Ow-Ow radial distribution functions there is another confirmation for the proposed picture. The positions of the second peaks change with the concentration, reflecting the longer distance of the second neighbour. In the previous regimes, due to the percolation of the water, these peaks remain at the same position. In third region because of the broken water's network there are now patches of water domains so their edges contribute to the shifting of the second peak.

As a final conclusion it can be said that here is presented a new picture that identifies different types of microheterogeneity in aqueous ethanol system. Concept of microheterogeneity is important for all aqueous solutions, and consequently for bio-systems, too. Microheterogeneity, as being a local property, does not have yet an adequate theoretical description, so here introduced pair connectedness function that allows accessing it has an important role. In the future work it is planned to use it at ionic liquids, various aqueous alcohols, as tert-butanol, tro-component mixtures and so on. About the question of ethanol models posed in this work it can be concluded that all the models were able to identify trends in the behaviour along mole fractions described here. Evolution of force fields towards ones sensitive to the environment, such as polarizable force fields, may be the best choice to overcome differences and problems presented in this work.

5 References

[1] E. E. Washbrun, *International Critical Tables of Numerical Data, Physics, Chemistry and Technology.* 2003.

[2] M. Mijakovic, "Teorija dvokomponentnih Lennard-Jones sustava," Master's thesis, Sveuciliste u Splitu, 2007.

[3] V. P. Belousov and I. L. Makarova *Vestn. Leningr. Univ. Fiz. Khim.*, vol. 4, p. 101, 1970.

[4] W. Price, H. Ide, and Y. Arata *J. Phys. Chem. A*, vol. 107, p. 4784, 2003.

[5] A. J. Easteal and L. A. Woolf *J. Phys. Chem.*, vol. 89, p. 1066, 1985.

[6] M. Mijakovic, B. Kezic, L. Zoranic, F. Sokolic, A. Asenbaum, C. Pruner, E. Wilhelm, and A. Perera *J. Mol. Liq.*, vol. 164, p. 66, 2011.

[7] B. Kezic, M. Mijakovic, L. Zoranic, F. Sokolic, A. Asenbaum, C. Pruner, E. Wilhelm, and A. Perera *in preparation.*

[8] J. B. Ott, C. E. Stouffer, G. V. Cornett, B. F. Woodfield, C. Guan-quan, and J. J. Christensen *J. Chem. Therm.*, vol. 19, p. 337, 1987.

[9] Y. Tanaka, T. Yamamoto, Y. Satomi, H. Kubota, and T. Makita *Review of Physical Chemistry of Japan*, vol. 47, p. 12, 1977.

[10] J. A. Hockey *J. Phys. Chem.*, vol. 74, p. 2570, 1970.

[11] J. P. E. Grolier and E. Wilhelm *Fluid Phase Equilibria*, vol. 6, p. 283, 1981.

[12] G. C. Benson, P. J. DArcy, and O. Kiyohara *J. Solution Chem.*, vol. 9, p. 931, 1980.

[13] A. Wakisaka, K. Matsuura, M. Uranaga, T. Sekimoto, and M. Takahashi *J. Mol. Liq.*, vol. 160, p. 103, 2011.

[14] P. G. Jonsson *Acta Cryst. B*, vol. 32, p. 232, 1976.

[15] www.lsbu.ac.uk/water/models.html, May 2012.

[16] S. Dixit, J. Crain, W. C. K. Poon, J. L. Finney, and A. K. Soper *Nature*, vol. 416, p. 829, 2002.

[17] A. Perera, F. Sokoli, and L. Zoranic *Phys. Rev.E*, vol. 75, p. 060502(R), 2007.

[18] L. Zoranic, F. Sokolic, and A. Perera *J. Chem. Phys.*, vol. 127, p. 024502, 2007.

[19] S. H. Tanaka, H. I. Yoshihara, A. W.-C. Ho, F. W. Lau, P. Westh, and Y. Koga *Can. J. Chem*, vol. 74, p. 713, 1996.

[20] L. Zoranic, K. Polok, F. Sokolic, and M. Mijakovic *in preparation.*

[21] D. R. Allan and S. J. Clark *Phys. Rev. B*, vol. 60, p. 6328, 1999.

[22] B. Himmel and T. Gerber *Journal of noncrystalline solids*, vol. 159, p. 235, 1993.

[23] M. D. Harmony, V. W. Laurie, R. L. Kuczkowski, R. H. Schwendeman, D. A. Ramsay, F. J. Lovas, W. J. Lafferty, and A. G. Maki *Phys. Chem. Ref. Data*, vol. 8, p. 619, 1979.

[24] D. G. Montague, I. P. Gibson, and J. C. Dore *Mol. Phys.*, vol. 47, p. 1405, 1982.

[25] Y. Tanaka, N. Ohtomo, and K. Arakawa *Bull. Chem. Soc. Jpn.*, vol. 57, p. 2569, 1984.

[26] C. J. Benmore and Y. L. Loh *J. Chem. Phys.*, vol. 112, p. 13, 2000.

[27] A. H. Narten and A. Habenschuss *J. Chem. Phys*, vol. 80, p. 3387, 1984.

[28] C. et al *Journal of Physics B: At.Mol. Opt. Phys.*, vol. 41, p. 175103, 2008.

[29] W. L. Jorgensen *J. Am. Chem. Soc.*, vol. 103, p. 345, 1981.

[30] W. L. Jorgensen *J. Phys. Chem.*, vol. 90, p. 1276, 1986.

[31] W. L. Jorgensen, J. D. Madura, and C. J. Swenson *J. Am. Chem. Soc.*, vol. 106, p. 6638, 1984.

[32] W. L. Jorgensen *J. Am. Chem. Soc.*, vol. 103, pp. 335,341,345, 1981.

[33] R. C. Wilhoit and B. J. Zwolinski *J. Phys. Chem. Ref. Data*, vol. Supl. 2, 1973.

[34] D. L. Wertz and R. K. Kruh *J. Chem. Phys*, vol. 47, p. 388, 1967.

[35] M. Magini, G. Paschina, and G. Piccaluga *J. Chem. Phys*, vol. 77, p. 2051, 1982.

[36] F. Muller-Plathe *Mol. Sim.*, vol. 18, p. 133, 1996.

[37] E. B. Pullman, *Intermolecular Forces*. 1981.

[38] W. Jorgensen, D. Maxwell, and J. Tirado-Rives *J. Am. Chem. Soc.*, vol. 118, p. 11225, 1996.

[39] W. Cornell, P. Cieplak, C. I. Bayly, I. R. Gould, K. M. Merz, D. M. Ferguson, D. C. Spellmeyer, T. Fox, J. W. Caldwell, and P. A. Kollman *J. Am. Chem. Soc.*, vol. 117, p. 5179, 1995.

[40] B. Chen, J. Potoff, and J. Siepmann *J. Phys. Chem B*, vol. 105, p. 3093, 2001.

[41] M. E. van Leeuwen *Mol. Phys.*, vol. 87, p. 87, 1996.

[42] S. P. Serbanovic, M. L. Mijajlovic, I. R. Radovic, B. D. Djordjevic, M. L. Kijevcanin, E. M. Djordjevic, and A. Z. Tasic *J. Serb. Chem. Soc.*, vol. 70, p. 527, 2005.

[43] W. L. Jorgensen and J. Tirado-Rives *J. Am. Chem. Soc*, vol. 110, p. 1657, 1988.

[44] P. van der Ploeg and A. Berendsen *J. Chem. Phys.*, vol. 76, p. 3271, 1970.

[45] T.Schnabel, J. Vrabec, and H.Hasse *Fluid Phase Equilibria*, vol. 233, p. 134, 2005.

[46] Y. Jiao, *The development of accurate force fields for protein simulation.* PhD thesis, Kansas state university, 2012.

[47] J. Gao, D. Habibollazadeh, and L. Shao *J. Phys. Chem.*, vol. 99, p. 16460, 1995.

[48] S. Y. Noskov, G. Lamoureux, and B. Roux *J. Phys. Chem. B*, vol. 109, p. 6705, 2005.

[49] S. Wang and N. M. Cann *J. Chem. Phys.*, vol. 126, p. 214502, 2007.

[50] L. Saitz, J. A. Padro, and E. Guardia *J. Phys. Chem. B*, vol. 101, p. 78, 1997.

[51] S. Sarkar and R. N. Joarder *J. Chem. Phys*, vol. 100, p. 5118, 1994.

[52] R. Ludwig, F. Weinhold, and T. C. Farrar *Mol. Phys.*, vol. 97, p. 465, 1999.

[53] E. J. W. Wensink, A. C. Hoffmann, P. J. van Mareen, and D. van der Spoel *J. Chem. Phys.*, vol. 119, no. 14, p. 7308, 2003.

[54] R. Wasler, B. Hess, A. E. Mark, and W. F. van Gunsteren *Chem. Phys. Lett.*, vol. 334, p. 337, 2001.

[55] K. Rah and B. C. Eu *Phys. Rev. E*, vol. 60, p. 4105, 1999.

[56] B. D. Smith and R. Srivastava, *Thermodynamic data for pure compounds*. Amsterdam: Elsevier, 1986.

[57] L. Zhang, Q. Wang, Y. C. Liu, and L. Z. Zhang *J. Chem. Phys.*, vol. 125, p. 104502, 2006.

[58] C. Zhang and X. Yang *Fluid Phase Equilibria*, vol. 231, p. 1, 2005.

[59] S. Bratos, J. C. Leicknam, S. Pommeret, and G. Gallot *Journal of Molecular Structure*, vol. 708, p. 197, 2004.

[60] Y. Marechal, *The hydrogen bond and the water molecule*. Amsterdam, Oxford: Elsevier, 2007.

[61] A. Perera *Mol. Phys., to appear*.

[62] L. Zoranic, R. Mazighi, F. Sokolic, and A. Perera *J. Phys. Chem. C*, vol. 111, p. 15586, 2007.

[63] S. K. Allison, J. P. Fox, R. Hargreaves, and S. P. Bates *Phys. Rev. B*, vol. 71, p. 024201, 2005.

[64] R. Chitra and P. E. Smith *J. Chem. Phys.*, vol. 114, p. 426, 2001.

[65] R. Chitra and P. E. Smith *J. Chem. Phys.*, vol. 115, p. 5521, 2001.

[66] A. Perera and F. Sokolic *J. Chem. Phys.*, vol. 121, p. 11272, 2004.

[67] S. Weerasinghe and P. E. Smith *J. Chem. Phys.*, vol. 118, p. 10663, 2003.

[68] M. E. Lee and N. F. van der Vegt *J. Chem. Phys.*, vol. 122, p. 114509, 2005.

[69] R. Wasler, A. E. Mark, W. F. van Gunsteren, M. Lauterbach, and G. Wipff *J. Chem. Phys.*, vol. 112, p. 10450, 2000.

[70] D. Gonzalez-Salgado and I. Nezbeda *Fluid Phase Equilibria*, vol. 240, p. 161, 2006.

[71] A. Perera, L. Zoranic, F. Sokolic, and R. Mazighi *J. Mol. Liq.*, vol. 159, p. 52, 2010.

[72] M. Kang and P. E. Smith *Journal of computational chemistry*, vol. 27, p. 1477, 2006.

[73] A. D. M. et al *J. Phys. Chem. B*, vol. 102, p. 3586, 1998.

[74] L. Zoranic, R. Mazighi, F. Sokolic, and A. Perera *J. Chem. Phys.*, vol. 130, p. 124315, 2009.

[75] J. M. Cordeiro *Int. J. Quantum Chem.*, vol. 65, p. 709, 1997.

[76] M. A. M. Cordeiro, W. P. Santana, R. Cusinato, and J. M. Cordeiro *THEOCHEM*, vol. 759, p. 159, 2006.

[77] J. Dai, X. Li, L. Zhao, and H. Sun *Fluid Phase Equilibria*, vol. 289, p. 156, 2010.

[78] E. A. Ploetz, N. Bentenitis, and P. E. Smith *Fluid Phase Equilibria*, vol. 290, p. 43, 2010.

[79] M. P. Allen and D. J. Tildeslay, *Computer simulation of liquids*. Oxford: Oxford University Press, 1987.

[80] B. Guillot *J. Mol. Liq.*, vol. 101, p. 219, 2002.

[81] H. J. C. Berendsen, J. R. Grigera, and T. P. Straatsma *J. Phys. Chem.*, vol. 91, p. 6269, 1987.

[82] P. Mark and L. Nilsson *J. Phys. Chem. A*, vol. 105, p. 9954, 2001.

[83] P. Kiss and A. Baranyai *J. Chem. Phys*, vol. 134, p. 054106, 2011.

[84] W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, and M. L. Klein *J. Chem. Phys*, vol. 79, p. 926, 1983.

[85] W. Smith, T. R. Forester, and I. T. Todorov, "DL POLY 2 package from Daresbury Laboratory, United Kingdom," 2004.

[86] http://www.ccp5.ac.uk/DL POLY CLASSIC/, April 2012.

[87] W. Smith, T. R. Forester, and I. T. Todorov, *The DL POLY Classic User Manual*. UK: Daresbury Laboratories, 2010.

[88] J. M. Haile, *Molecular Dynamics Simulation*. New York: Wiley interscience publication, 1992.

[89] B. J. Adler and T. E. Wainwright *J. Chem. Phys.*, vol. 27, p. 1208, 1957.

[90] B. J. Adler and T. E. Wainwright *J. Chem. Phys.*, vol. 31, p. 459, 1959.

[91] A. Rahman *JPhys. Rev. A*, vol. 136, p. 405, 1964.

[92] L. Zoranic, *A Molecular Dnamics Study of Microheterogenities in Aqueous Alcohol Solutions*. PhD thesis, Faculty of Science, University of Zagreb, 2008.

[93] P. Ewald *Ann. Phys.*, vol. 64, p. 253, 1921.

[94] S. W. de Leeuw, J. W. Perram, and E. R. Smith *Proc. Roy. Soc. Lond. A*, vol. 373, p. 57, 1980.

[95] U. Essmann, L. Perera, M. L. Berkowitz, T. Darden, H. Lee, and L. G. Pedersen *J. Chem. Phys.*, vol. 103, p. 8577, 1995.

[96] H.Lee and W. Cai, *Ewald Summation for Coulomb Interactions ia a Periodic Supercell*. Stanford University, CA: Department of Mechanical Engineering, 2009.

[97] T. Darden, D. York, and L. Pedersen *J. Chem. Phys.*, vol. 98, p. 12, 1993.

[98] F. Sokolic, *Studies of liquid and gas sulphur dioxide by computer simulation: Towards a good effective intermolecular potential*. PhD thesis, Institute Rudjer Boskovic, 1985.

[99] I. T. Todorov and W. Smith, *DL POLY 4 manual*. Warrington: Daresbury Laboratories, 2010.

[100] H. J. C. Berendsen, J. P. M. Postma, A. DiNola, and J. R. Haak *J. Chem. Phys.*, vol. 81, p. 3684, 1984.

[101] W. G. Hoover *Phys. Rev. A*, vol. 31, p. 1695, 1985.

[102] P. Hnenberger *Adv. Polymer. Sci.*, vol. 173, p. 105, 2005.

[103] M. S. Dionisio, J. J. M. Ramos, and R. M. Goncalves *Can. J. Chem.*, vol. 68, p. 1937, 1990.

[104] www.ddbst.com, June 2012.

[105] M. A. Gonzales, E. Enciso, F. J. Bermejo, and M. Bee *J. Chem. Phys.*, vol. 110, p. 8045, 1999.

[106] E. K. N. Marsh, *Recommended Reference Materials for the Realizatio of Physicochemical Properties*. 1987.

[107] A. Glattli, X. Daura, and W. F. van Gunsteren *J. Chem. Phys*, vol. 116, p. 9811, 2002.

[108] D. van der Spoel, P. J. van Maaren, and H. J. C. Berendsen *J. Chem. Phys*, vol. 108, p. 10220, 1998.

[109] G. Guevara-Carrion, J. Vrabec, and H. Hasse *J. Chem. Phys*, vol. 134, p. 074508, 2011.

[110] J. L. F. Abascal and C. Vega *J. Chem. Phys*, vol. 123, p. 234505, 2005.

[111] D. P. Tieleman and H. J. C. Berendsen *J. Chem. Phys*, vol. 105, p. 4871, 1996.

[112] A. Ben-Naim, *Molecular theory of solutions*. Oxford University Press, 2006.

[113] A. Perera, B. Kezic, F. Sokolic, and L. Zoranic, *in Molecular Dynamics (Vol 2)*. InTech, 2012.

[114] J. P. Hansen and I. R. McDonald, *Theory of simple liquids*. Academic Press, Elsevier, 2006.

[115] J. L. Lebowitz and J. K. Percus *Phys. Rev.*, vol. 122, p. 1675, 1961.

[116] L. S. Ornstein and F. Zernike *Proc. Amsterdam Acad. Sci.*, vol. 17, p. 793, 1914.

[117] L. Zoranic, F. Sokolic, and A. Perera *J. Mol. Liq.*, vol. 136, p. 199, 2007.

[118] A. Perera, L. Zoranic, F. Sokolic, and R. Mazighi *J. Mol. Liq.*, vol. 159, p. 52, 2011.

[119] G. Hura, J. M. Sorenson, R. M. Glaeser, and T. Head-Gordon *J. Chem. Phys.*, vol. 113, p. 9140, 2000.

[120] J. M. Sorenson, G. Hura, R. M. Glaeser, and T. Head-Gordon *J. Chem. Phys.*, vol. 113, p. 9149, 2000.

[121] I. Juurinen, "Water-ethanol mixtures by molecular dynamics and x-ray Compton scattering," Master's thesis, Helsingin yliopisto, 2009.

[122] Y. Wu, H. L. Tepper, and G. A. Voth *J. Chem. Phys.*, vol. 124, p. 024503, 2006.

[123] B. Bagchi *Chemical Physics Letters*, vol. 529, p. 1, 2012.

[124] J. G. Kirkwood and F. P. Buff *J. Chem. Phys.*, vol. 19, p. 7, 1951.

[125] E. Matteoli and L. Lepori *J. Chem. Phys.*, vol. 80, p. 2856, 1984.

[126] A. Ben-Naim *J. Phys. Chem.*, vol. 82, p. 792, 1978.

[127] A. Perera, R. Mazighi, and B. Kezic *J. Chem. Phys.*, vol. 136, p. 174516, 2012.

[128] E. E. Matteoli and G. A. Mansoori, *Advances in Thermodynamics. Fluctuation Theory of Mixtures.* 1990.

[129] K. Nishikawa and T. Iijima *J. Phys. Chem.*, vol. 97, p. 10824, 1993.

[130] T. L. Hill *J. Chem. Phys.*, vol. 23, p. 617, 1955.

[131] F. H. Stillinger *J. Chem. Phys.*, vol. 38, p. 1486, 1963.

[132] L. A. Pugnaloni and F. Vericat *J. Chem. Phys.*, vol. 116, p. 1097, 2002.

[133] A. Coniglio, U. de Angelis, A. Forlani, and G. Lauro *J. Phys. A*, vol. 10, p. 219, 1977.

[134] L. Zoranic, F. Sokolic, and M. Mijakovic *in preparation.*

[135] K. Momma and F. Izumi, "Vesta 3 for three-dimensional visualization of crystal, volumetric and morphology data," *J. Appl. Crystallogr.*, vol. 44, p. 1272, 2011.

[136] J. M. Stoker and R. L. Rowley *J. Chem. Phys.*, vol. 91, p. 3670, 1989.

[137] A. Asenbaum, C. Pruner, E. Wilhelm, M. Mijakovic, L. Zoranic, F. Sokolic, B. Kezic, and A. Perera *Vibrational Spectroscopy*, vol. 60, p. 102, 2012.

[138] P. Ball *Nature*, vol. 452, p. 291, 2008.

[139] K. Stokely, M. G. Mazzaa, H. E. Stanley, and G. Franzese *Proc.Nat. Acad. Sci.*, vol. 107, p. 1301, 2010.

[140] E. Wilhelm and A. Asenbaum, *in Heat Capacities: Liquids, Solutions and Vapours*. The Royal Society of Chemistry, 2010.

6 List of publications and CV

- M. Mijakovic, B. Kezic, L. Zoranic, F. Sokolic, A. Asenbaum, C. Pruner, E. Wilhelm, and A. Perera *J. Mol. Liq.*, vol. 164, p. 66, 2011
- A. Asenbaum, C. Pruner, E. Wilhelm, M. Mijakovic, L. Zoranic, F. Sokolic, B. Kezic, and A. Perera *Vibrational Spectroscopy*, vol. 60, p. 102, 2012
- B. Kezic, M. Mijakovic, L. Zoranic, F. Sokolic, A. Asenbaum, C. Pruner, E. Wilhelm, and A. Perera *in preparation*
- L. Zoranic, F. Sokolic, and M. Mijakovic *in preparation*
- L. Zoranic, K. Polok, F. Sokolic, and M. Mijakovic *in preparation*

In the year 2005 graduated in IT Tehnology at *Odjel za strucne studije Sveucilista u Splitu*. In the year 2007 graduated in physics at Faculty of Science, University of Split. Started doctoral study in biophysics at Faculty of Science, University of Split, in 2008. Works at the Institute of public health, Split's - dalmatian county.

Abstract

Ethanol-water liquid mixture was analyzed by the method of Molecular Dynamics, over the whole composition range to elucidate its micro structure. A recently introduced concept of microheterogeneity was used to describe mixture's behaviour. The microsegregation of species, while on the macroscopic level the liquid appears homogenous, noticed particularly in aqueous mixtures, is addressed as microheterogeneity. It was discovered that there are at least three structural regimes with respect to composition range in aqueous ethanol at ambient conditions. Up to approximately 0.15 mole fraction of ethanol, hydrophobic ethanol's sites distribution declines from homogenous. From 0.15 to approximately 0.65 mole fraction of ethanol, ethanol and water form bi-continuous intertwined microheterogenous mixture of segregated domains, and above 0.65 mole fraction of ethanol water network loses its connectivity over the whole system. Broken parts of the water's network that do not span the whole system are found in the mixture. To prove it the results from MD simulation are listed: RDF, coordination numbers, analysis of clustering, Kirkwood-Buff integrals, snapshots, and diffusion coefficients, all supporting the given picture, as well as results from the sound speed measurement, and previously known experimental data on excess enthalpy, isothermal compressibility, azeotropy and heat capacity. Especially results from the novel approach in cluster analysis, performed for the first time on this kind of system, were important for supporting the picture of three structural regimes. This is the first time that such detailed overall picture of the microscopic level structure of this liquid mixture is given. Besides, in-detail analysis of different ethanol models in their ability to reproduce this structural change with mole fraction is presented; OPLS and TraPPE models that are parametrized on pure substance thermodynamical properties, as well as KBFF model parametrized on the solute activity in the aqueous mixture. As some recent results from simulations of aqueous organic compounds mixtures show, some excess quantities as excess enthalpy, and dynamic properties as diffusion coefficient, which are the most sensitive to mixing properties, are difficult to reproduce correctly using force fields parametrized on pure components. The OPLS, TraPPE and KBFF models were tested at excess properties and diffusion coefficient, and KBFF model was found to perform better in reproducing them. And more important, all three models were found to reproduce properties that confirm the existence of three structural regimes in the mixture.

Sažetak

Metodom molekularne dinamike analizirana je tekuća mješavina etanola i vode u cijelom rasponu koncentracija, da bi se bolje razumjela njena mikrostruktura. Za opis ponašanja mješavine upotrijebljen je novi koncept mikroheterogenosti. Mikroheterogenost je pojava mikroseparacije komponenti u mješavini koja je na makroskopskom nivou homogena. Ova pojava je posebno opažena u vodenim otopinama. Pronađeno je da u tekućoj mješavini etanola i vode, na sobnim uvjetima, postoje najmanje tri strukturna režima u odnosu na molarni udio konstituenata. Prvi je do otprilike 0.15 molarnog udjela etanola, gdje distribucija hidrofobnih *site*-ova ima otklon od homogene distribucije. Od 0.14 do otprilike 0.65 molarnog udjela etanola, etanol i voda formiraju bi-kontinuiranu isprepletenu mikroheterogenu mješavinu mikrosepariranih domena. Iznad 0.65 molarnog udjela etanola mreža molekula vode je pokidana i više se ne proteše kontinuirano cijelim sustavom. Kao osnova za danu sliku izloženi su rezultati simulacije molekularne dinamike: RDF, koordinacijski broj, analiza klastera, Kirkwood-Buff integrali, *snapshot*-ovi i difuzijski koeficijent. Osim toga navedeni su i eksperimentalni rezultati mjerenja brzine zvuka koji podržavaju ovu sliku, kao i neki otprije poznati rezultati za eksces entalpije, kompresibilnost, azeotropiju i toplinski kapacitet. Posebno su vrijedni rezultati analize klastera, zbog upotrebe nove metode, prvi put primijenjene na ovoj vrsti sustava. Ovo je i prvi put da je dana ovakva detaljna slika mikrostruktute tekuće mješavine etanola i vode. Osim toga provedena je detaljna analiza različitih modela za etanol i testirana je njihova sposobnost reproduciranja promjene strukture s promjenom molarnog udjela komponenti. Testirani su OPLS i TraPPE modeli parametrizirani na termodinamičkim svojstvima čistih tekućina i KBFF model parametriziran na aktivitetu otopljene tvari u vodenoj otopini. Prema nekim rezultatima simulacija za vodene otopine organskih molekula, publiciranim u posljednje vrijeme, modeli parametrizirani na čistim tekućinama ne reproduciraju dobro osobine mješavina koje su posebno osjetljive na *interface* miješanog sustava, kao što su eksces entalpije i difuzijski koeficijent. U ovom radu pokazano je da ih KBFF model bolje uspijeva reproducirati od ostala dva modela. Sva tri modela bila su uspješna u reproduciranju tri strukturna režima u promatranoj mješavini.

Appendices

A Radial distribution functions

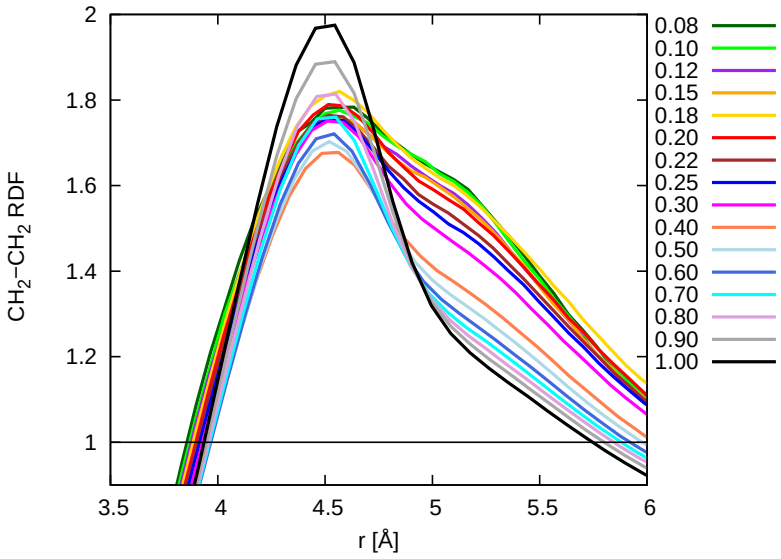


Figure 68: Radial distribution functions of CH_2 sites, first peak, TraPPE model

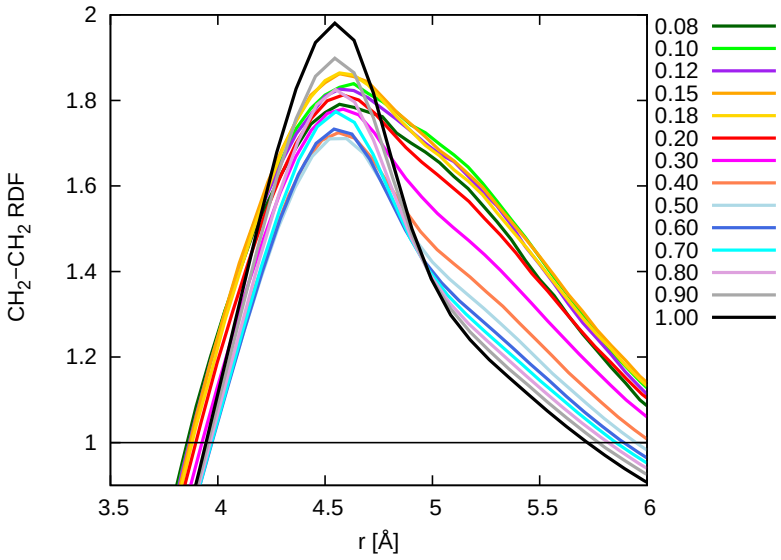


Figure 69: Radial distribution functions of CH_2 sites, first peak, semi flexible OPLS model

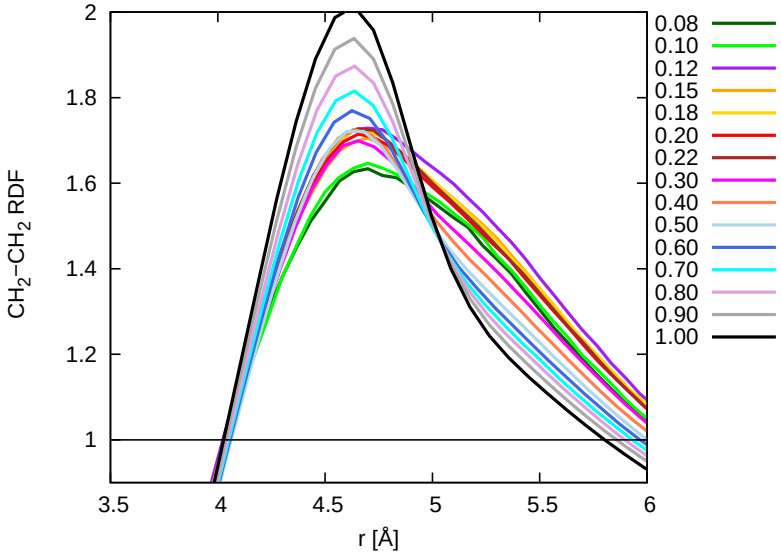


Figure 70: Radial distribution functions of CH_2 sites, first peak, KBFF model

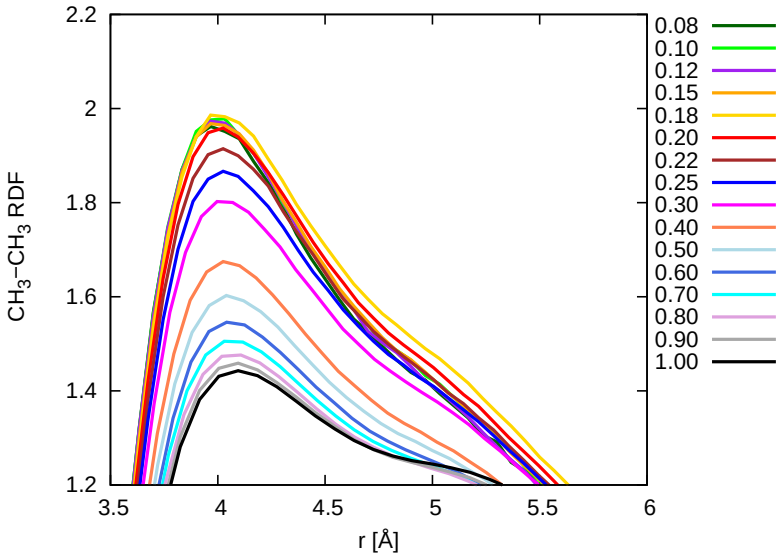


Figure 71: Radial distribution functions of CH_3 sites, first peak, TraPPE model

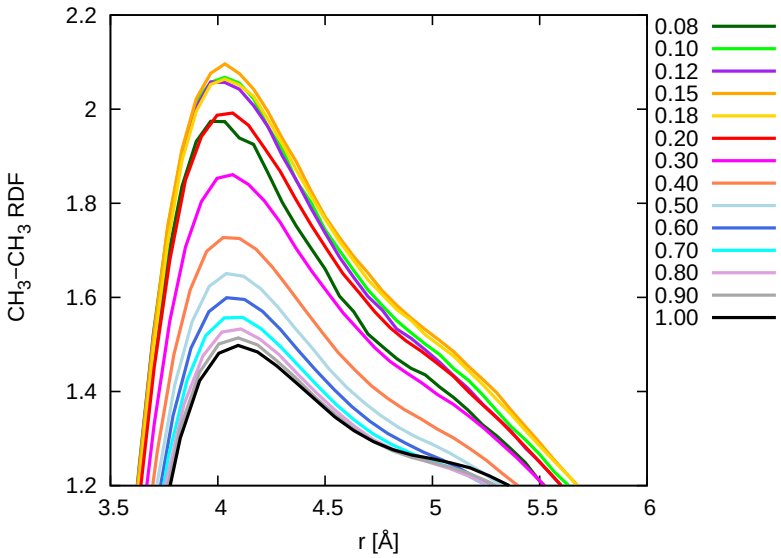


Figure 72: Radial distribution functions of CH_3 sites, first peak, semi flexible OPLS model

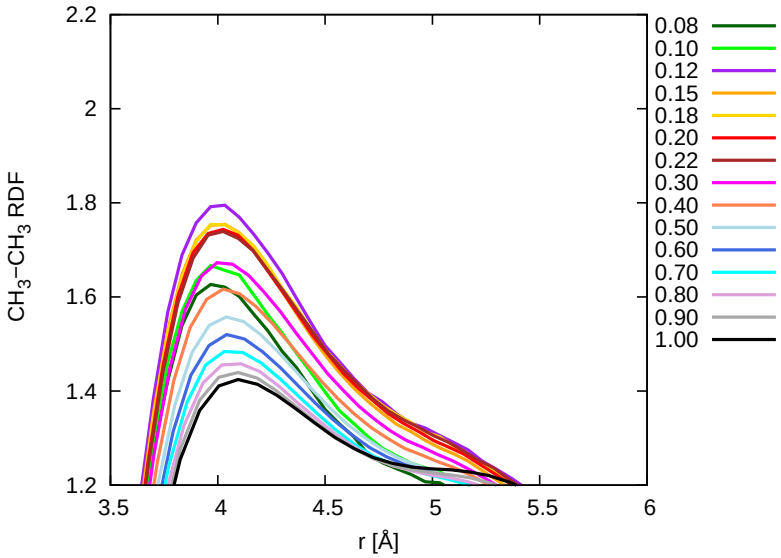


Figure 73: Radial distribution functions of CH_3 sites, first peak, KBFF model

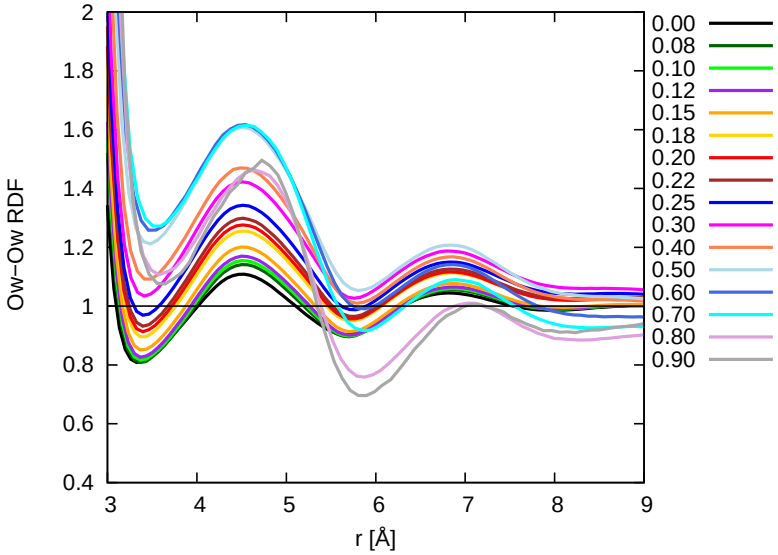


Figure 74: Radial distribution functions of Ow sites, shift of the second peak, TraPPE model

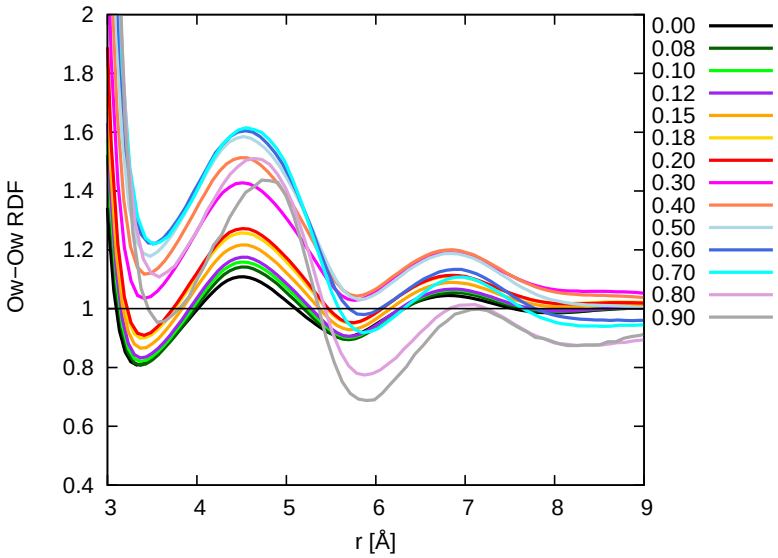


Figure 75: Radial distribution functions of Ow sites, shift of the second peak, semi flexible OPLS model

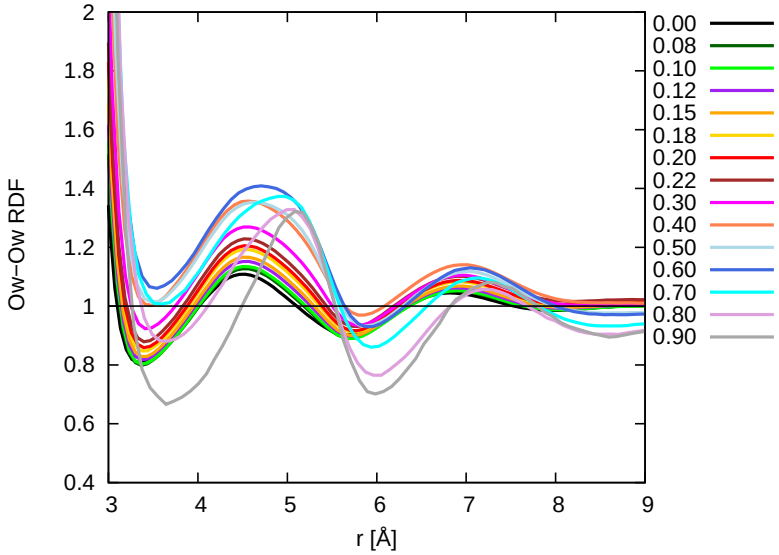


Figure 76: Radial distribution functions of Ow sites, shift of the second peak, KBFF model

B Cluster radial distribution functions

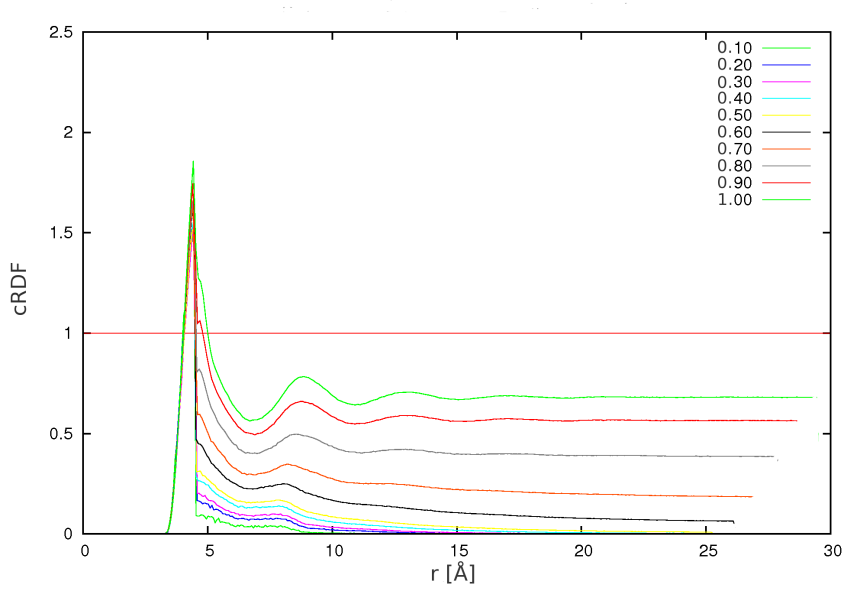


Figure 77: Cluster radial distribution functions, CH_2, KBFF model

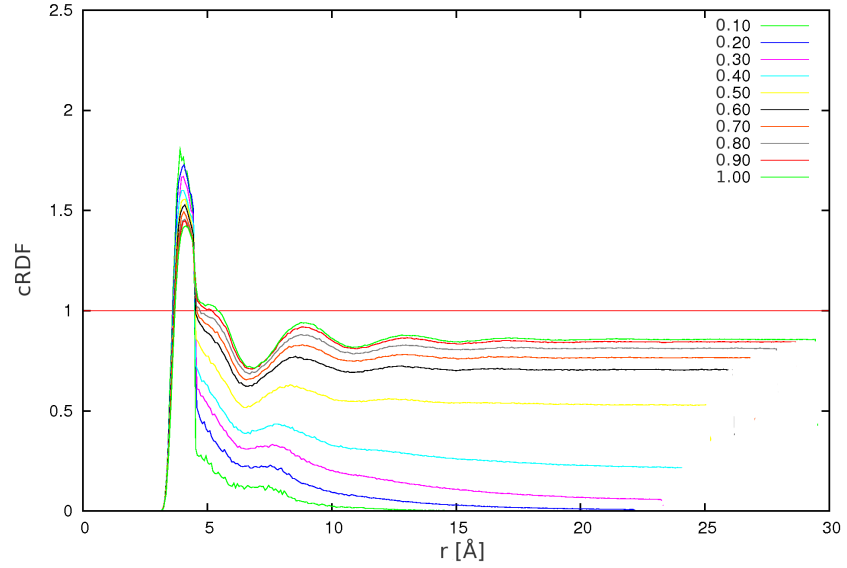


Figure 78: Cluster radial distribution functions, CH_3, KBFF model

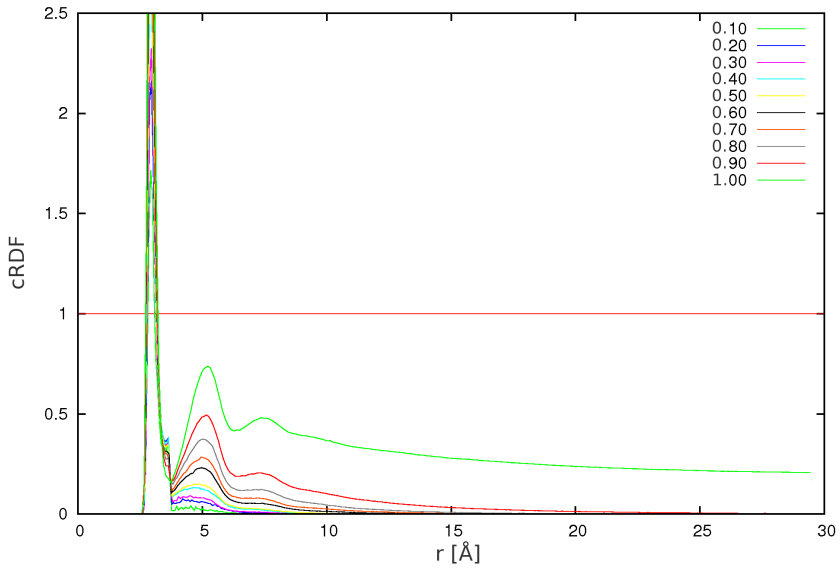


Figure 79: Cluster radial distribution functions, O, KBFF model

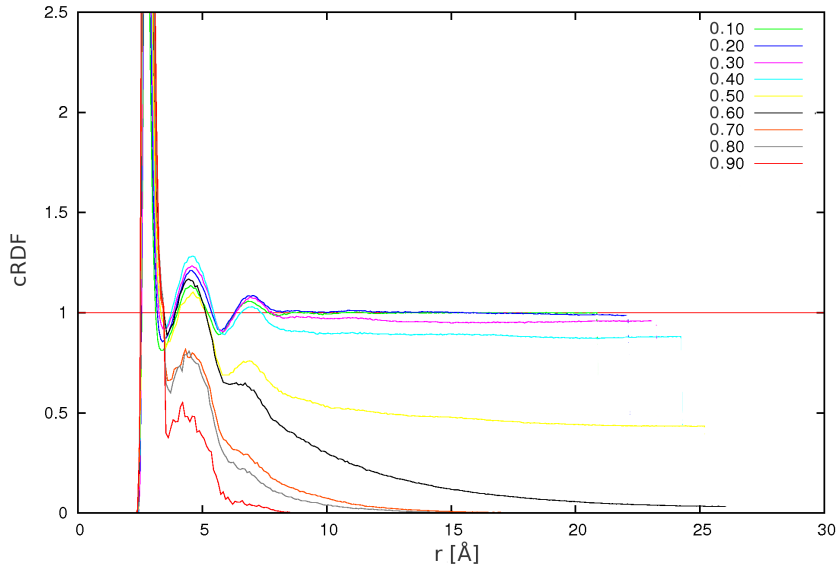


Figure 80: Cluster radial distribution functions, Ow, KBFF model

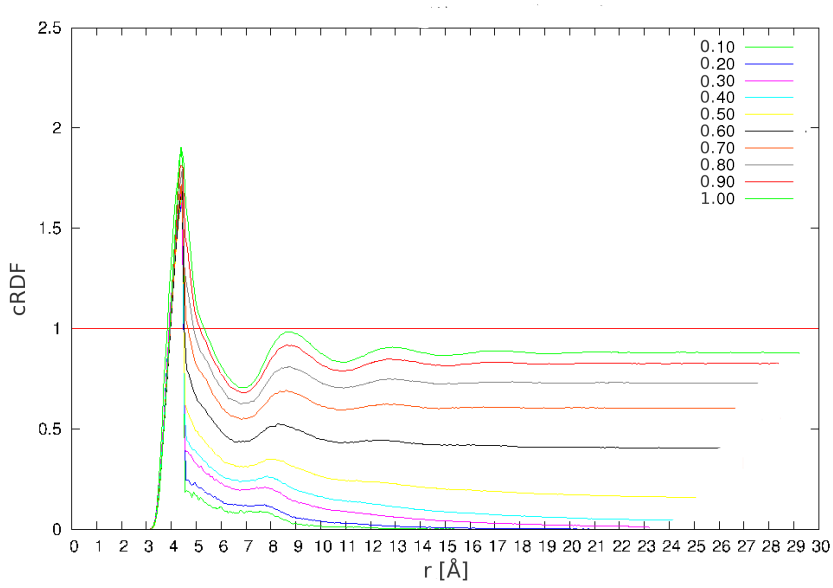


Figure 81: Cluster radial distribution functions, CH_2, TraPPE model

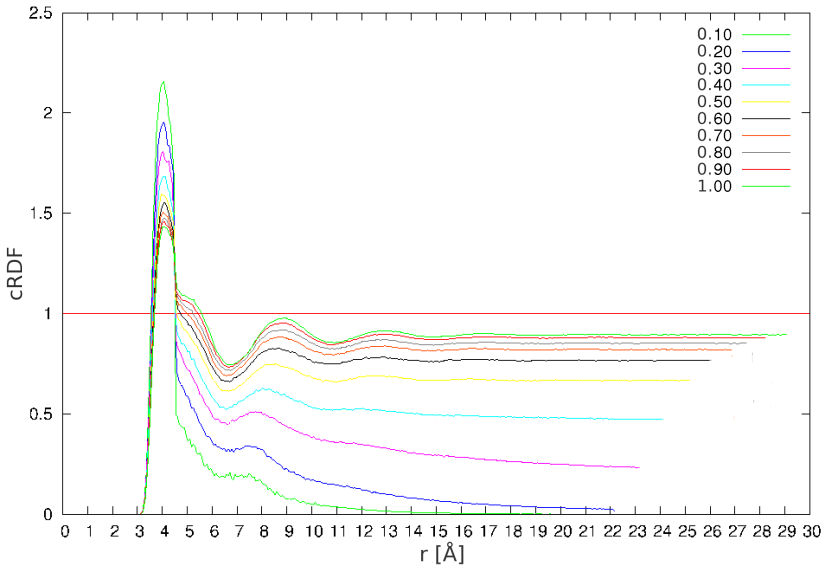


Figure 82: Cluster radial distribution functions, CH_3, TraPPE model

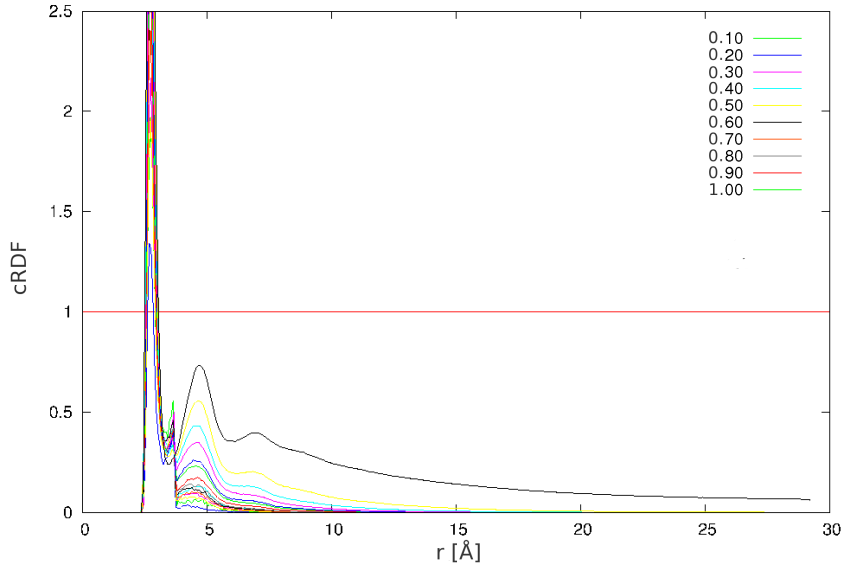


Figure 83: Cluster radial distribution functions, O, TraPPE model

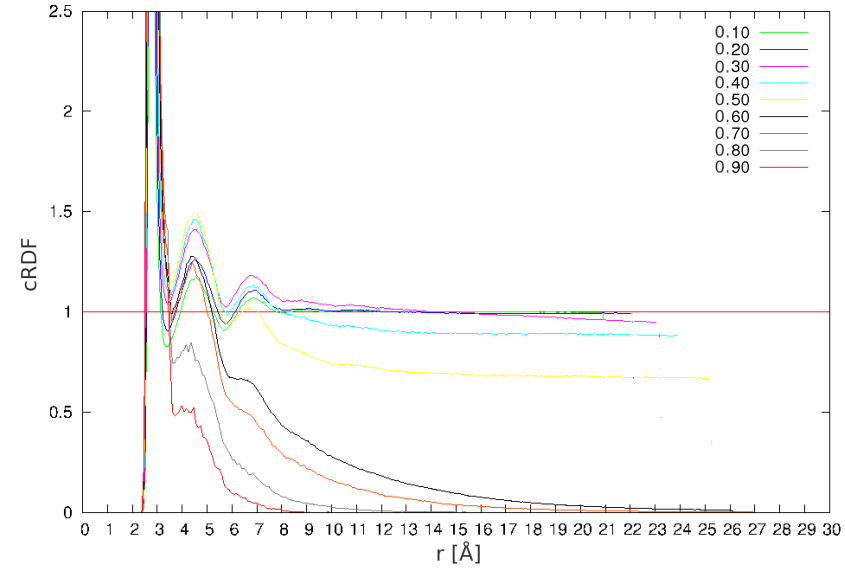


Figure 84: Cluster radial distribution functions, Ow, TraPPE model