Molecular Dynamics Simulations of Ethanol and Ethanol-Water Mixtures.

P.J. Berryman

Submitted for the Degree of
Doctor of Philosophy

University of Surrey

Theory and Advanced Computation Group
School of Electronics and Physical Sciences
University of Surrey
Guildford, Surrey GU2 7XH, U.K.

December 2006

Summary

Molecular Dynamics simulations of ethanol and ethanol-water mixtures are undertaken and compared to experiment. We calculated bond lengths and vibrational frequencies of pure ethanol in the liquid and vapour phases at NTP, pure ethanol liquid as a function of pressure, and ethanol-water mixtures as a function of concentration. The vibrational frequencies show good agreement with experimental results.[9]

We modelled ethanol using a modified version of the Cornell field,[31] and water using a flexible form of the TIP3P potential.[50] The liquid is subjected to hydrostatic pressure in the range -1 to $15kbar$ at room temperature.

We compare the results against the predictions of the solvation pressure model (SPM). The model states that solvated particles experience a pressure equal to the cohesive energy density (CED) of the liquid. Support of the model is noted for some bonds in pure ethanol. Ethanol-water mixtures show no support for the model. We attempt to establish the reasons for the failure of the model in these cases.

This work shows supporting evidence for the bi-percolating nature of alcohol-water mixtures.[59,60] Mixtures of ethanol and water are seen to mix poorly. The ethanol tends to clump together into clusters of neighbouring molecules.

We also see evidence of hydrophobic hydration of ethanol in ethanol-water mixtures. The water molecules tend not to interact with the hydrophobic head of the ethanol molecules, being easily drawn away by other hydrogen bonding interactions. Thus, rather than being compressed, water simply surrounds the ethanol molecule in a clathrate-like configuration.

The model is not able to predict these hydrophobic hydration effects causing a reduced pressure effect on the solute. Further work is required to determine whether water can exert a solvation pressure on molecules with which it mixes more thoroughly.

Key words: Molecular Dynamics, Solvation Pressure, Ethanol, Water, Bi-Percolating

Email: p.berryman@surrey.ac.uk

WWW: http://www.eps.surrey.ac.uk/

Acknowledgements

I would like to thank my Supervisor Dr Faux for his help and guidance, and Dr Everett for useful discussions and advice.

Contents

Chapter 1

Introduction

1.1 Background

A large portion of current scientific interest concerns the behaviour of complex molecules in response to their environment. The formation of proteins, for example, is an important unsolved problem in biology. Proteins are formed in long straight configurations, which quickly take their required shape in a process known as folding. Incorrect folding is widely regarded as the cause of a variety of prion diseases, such as BSE in cattle, scrapie in sheep and Alzheimer's Disease.[1,2]

Proteins are usually surrounded by a largely water based mixture. The folding of the protein is thought to be the result of the hydrophobic effect. This results in hydrophobic portions of the protein burying themselves in the centre of the folding protein, where they are surrounded by the hydrophilic portions. Many groups have explored the effects of hydrophobicity, known as "hydrophobic collapse" on the folding protein.[3–5]

The folded state of the protein may be altered by the pH and ion concentration of its surroundings, as well as to thermodynamic effects such as pressure and temperature. In addition to this, the solvent which surrounds complex molecules may also have an important physical effect. The solvent is speculated to exert a pressure, known as solvation pressure, on the protein. This pressure acts in addition to any external pressure felt by the protein molecule. The effect has been observed in carbon nanotubes,[6–8]

starch grains[9] and solvent mixtures.[10,11]

There are, of course, other models aiming to describe the effect of solvent-solute interactions. Versions of the Perturbed Hard Fluid Model[13] have been used to attempt to describe such interaction. Hutchinson et al derived an equation for the total force $\mathbf{F}_{g-l}$ on the solute bond resulting from its interaction with the solvent (at 1atm).[14]

The focus of studies based on the hard fluid model has been on the van der Waals interaction, specifically the competition between long ranged attractive forces, and short ranged repulsive forces. This model uses a hard sphere representation of the repulsive van der Waals component, relying on solvent compressibility data and the van der Waals radius to determine the strength of the repulsive interaction. Similarly, long range attractive interactions are assumed in the simplest case to be a linear function of density. Additional parameters are used to attempt to mimic the nonlinear behaviour of certain bonds as a function of density, notably the C-H and O-H bonds.

Using this approach, it has been possible for various groups to achieve good agreement with corresponding experimental analyses.[13–15] However, these results are highly dependent on fitting parameters designed specifically for the purposes of recreating experimental work. The solvent-solvent interaction is a notable omission from the analysis, due to the attempts of the model to be a simple recreation of the experimental picture.

The SPM is used as a complementary approach to the understanding of shifts in bond length and associated shifts in vibrational frequency. Instead of using functions to predict attractive and repulsive forces, the concept of a generalised 'solvation pressure' is applied to the interpretation of results of full MD simulations and experimental data. The strength of the solvation pressure model (SPM) lies in its simplicity. It states that the magnitude of the pressure applied to a solute can be obtained simply from the Cohesive Energy Density (CED) of the mixture. The aim of this project is to begin to establish the extent to which the SPM can be applied. The SPM, and its possible limitations, are explained in detail in Section 1.2.

1.2 The Solvation Pressure Model

The Solvation Pressure Model (SPM) is a theory which states that solvents exert a pressure on a solute placed within them. This pressure is thought to be equivalent to an external hydrostatic pressure. The solute is squeezed by the solvent, affecting properties of the molecule, such as the bond lengths and vibrational frequencies.

Some experimental evidence for the model has already been presented in previous papers.[9,10,12] In these studies, different solvents have been used to alter the external pressure of gelation of starch grains. Molecular Dynamic and experimental studies of chloroform were undertaken by Hubel et al.[10]

The magnitude of the solvation pressure is speculated to be equal to the Cohesive Energy Density (CED) of the liquid. The CED is the heat of vapourisation of a molecule per unit volume:

$$CED = \frac{\Delta U_{vap}}{V} \tag{1.1}$$

where ΔU_{vap} is the energy change on vapourisation, and V is the volume of the liquid. The CED of a liquid represents the potential energy of a condensed material compared with that of an ideal vapour at the same temperature. This potential energy is known as the internal energy, and is a negative quantity. The CED is therefore defined as a positive quantity despite representing a negative internal energy.

In their book, Hildebrand and Scott designated the heat of vapourisation as the CED and its square root as the solubility parameter.[16] They used the parameter in an attempt to understand solubility of liquids.

Note that in experimental studies, it is convenient to measure the CED in terms of the enthalpy of the liquid:

$$\Delta H_{vap} = \Delta U_{vap} + PV_m \tag{1.2}$$

Rearranging, and assuming the ideal gas law PV = RT, we obtain a definition of the CED based on the enthalpy of the liquid:

$$CED = \frac{\Delta H_{vap} - RT}{V} \tag{1.3}$$

This form of the CED is seen in the literature,[17] and is convinient because the enthalpy can be easily measured experimentally. However, there is no need for us to make these assumptions in our calculations, since we are able to measure the intermolecular energy of the system directly.

The CED is closely related to the "internal pressure" of a molecule, which is a measure of the change in internal energy of one mole of a solute as it undergoes a small isothermal expansion.

$$\text{Internal Pressure} = \left(\frac{\partial U}{\partial V}\right)_T \tag{1.4}$$

The internal pressure of a molecule is closely related to CED. In non-polar molecules, the internal pressure is approximately equal to the CED of the molecule. However, for polar liquids including water, the internal pressure of the liquid tends to be lower than the CED. This is because a small expansion of the liquid is not sufficient to break the strong bonds, resulting in a small change in energy. The most important contributions to the internal energy come from interactions which vary most rapidly near equilibrium molecular separation. These include dispersion and dipole-dipole interactions. The internal pressure therefore offers a less complete picture of the intermolecular interactions affecting a liquid.

$$\left(\frac{\partial U}{\partial V}\right)_T \approx \frac{\Delta U}{V} \tag{1.5}$$

The CED can be expressed in units of pressure:

$$1 Jm^{-3} \equiv 1 Nm^{-2} \equiv 10^{-5} bar \tag{1.6}$$

This makes the CED appear especially relevant to studies of pressure and represents the energy required to release a single molecule of a liquid as a vapour. The CED is a measure of how strongly a molecule is attracted to its neighbours. It is largest where the

attraction is strongest, such as in water, which forms large numbers of strong hydrogen bonds, for which the CED is $\sim 22.5kbar$, and smallest in weakly interaction liquids such as chloroform ($3kbar$).

These strong intermolecular bonds mean that molecules such as water are pulled very tightly together. It is easy to imagine the forces that may be exerted by these strong cohesive interactions. This makes it a very significant effect, equal to around $7kbar$ for ethanol, and $\sim 22.5kbar$ in water. Note that the 'solute' can be the same type of molecule as the solvent, and so the model also applies to pure liquids.

Thus the solvation pressure is thought to be ubiquitous, applying to all liquids. However, it may not always be the dominant effect in a given mixture. Additional interactions between solvent and solute may dominate, making it appear that the solvation pressure is not influencing these bonds.

Hydrogen bonding in ethanol-water mixtures is a good example of this. Hydrogen bonding between water and ethanol will tend to lengthen the OH bond of ethanol. Because of the strength of this interaction, it will tend to dominate over any pressure effects. Such interactions, which hide the effects of the SPM are known as "masking interactions".[10] In general, a masking interaction is the name given to solvent-solute interactions that prevent a bond acting as a pressure gauge.

However, not all hydrogen bonding interactions can be classified as masking interactions. The large CED of water is cause by hydrogen bonding interactions between water molecules. These interactions would clearly contribute heavily to the solvation pressure placed on the ethanol solute, putting the whole ethanol molecule under pressure and reducing all bond lengths. These solvent-solvent interactions would not be classified as a masking interaction.

Note that these masking interactions do not imply that a solvation pressure is not felt by these bonds. They simply hide the effects of the SPM due to the presence of larger forces. Masking interactions have been noted in previous solvation pressure studies of ethanol,[9] and of chloroform.[10]

In studies of chloroform,[10] the C-H bond length was found to shorten more than expected on transition from the gas to liquid phase. This was deemed to be the result

of masking interactions caused by the comparatively small size of the H atom. Previous study of ethanol also noted masking interactions, including those caused by the hydrogen bonding of the O-H group.

In this project, we look for evidence of these pressures predicted to be exerted by solvents. This is achieved through Molecular Dynamics (MD) studies of ethanol vapour, ethanol liquid under pressure and mixtures of ethanol and water, coupled with the results of experimental Raman spectroscopic analysis performed by Dunstan et al.[9] Ethanol is chosen for several reasons: excellent experimental data is available for bulk properties of ethanol, in addition to the Raman spectroscopic data; ethanol contains a number of the bond types found in proteins and so serves as a prototype for such a system. Also, the CED of ethanol means that it can exert a solvation pressure on chloroform, as well as feeling the solvation pressure effects of water.

1.2.1 Development of the SPM

The SPM developed from the study of Raman spectra from two distinct sources. The first of these sources was Dunstan and coworkers' studies of carbon nanotubes.[6–8] In this series of articles, the group explored the potential of using carbon nanotubes as sensors of stress and strain at the molecular level. They showed that the frequencies of Raman spectroscopic modes of the nanotubes shifted as a function of external pressure. It was noted in passing that the frequency was also dependent on the solvent of the mixture.

The second source was the research of Dixit and coworkers in their spectroscopic analysis of the structure of mixtures of methanol and water in varying concentrations.[11] The main focus of the work was the structure of methanol molecules in mixtures of methanol and water. They noted the formation of chains of methanol molecules, and studies how the length of these chains varied with concentration. They noted that when mixed with small amounts of water, the behaviour of the methanol was as if it were being squeezed closer together without altering its overall structure.

Dixit et al tentatively cited the nanotube research of Dunstan and coworkers as a related phenomenon, and coined the phrase 'Solvation Pressure' in order to describe

it. Further investigation into this effect was suggested. This work was undertaken by Dunstan and coworkers, who produced the most comprehensive work on the subject to date.[9] Both experimental and computational techniques were used. The paper reported some supporting evidence for the SPM.

Dunstan et al applied hydrostatic pressure to starch grains in aqueous solution using a Diamond Anvil Cell.[19] The hydrostatic pressure, P_h applied to the mixture was increased until gelation of the starch molecules occurred. The pressure required to induce gelation of the starch molecule is known as the 'gelation pressure'. They varied the solvation pressure, P_s, applied to the starch grains by altering the concentration, and hence CED, of the ethanol-water mixture. The relationship between concentration and solvation pressure is described in more detail in Section 4.1. Gelation pressure was seen to increase as a function of decreasing solvation pressure. Plotting the gelation pressure against P_s resulted in straight line of slope -1. It followed that:

Gelation Pressure $= P_h + P_s$

and hence that hydrostatic and solvation pressures have an identical effect the starch. This was claimed as the first verifiable example of a protein being placed under "negative effective pressure".

The paper also compared the Raman spectra of pure ethanol under hydrostatic pressure with those of ethanol-water and ethanol-chloroform mixtures. According to the theory, the high CED of water should act to increase the pressure applied to ethanol. Conversely, the low CED of chloroform should reduce the pressure felt by ethanol. Plots of vibrational frequency as a function of solvation pressure should follow the trends of pure ethanol as a function of pressure. Chloroform data should continue along the same lines at lower pressures. The experiment showed that, in some cases, the predictions of the SPM were very accurate. However, for some bond types, the trends did not follow the predictions of the model.

Finally, the paper introduced initial MD studies of ethanol. This included attempts to model the decrease in bond length as ethanol vapour is solvated. This was achieved with a similar level of success - some bond lengths fitted the predictions of the model, while others showed unexpected behaviour.

Further work was clearly required in this area. The work showed that the modelling of the situation was possible. However, no attempt had yet been made to recreate the experimental Raman frequency data directly. Neither had any attempt been made to utilise the advantages of performing a MD study of the system in order to better understand the processes behind the observed effects.

The aims of this work are outlined in the following sections. Complementary work from Queen Mary - University of London aims to provide understanding of solvation pressure in liquids with a low CED, achieved through studies of chloroform.

1.3 Current Work and Implications

Ethanol is used throughout this project, since it is a simple organic molecule containing some of the important bond types seen in proteins. Also, it is a good example of a molecule with an 'average' CED, and hence average solvation pressure. A CED of $6.76kbar$, places ethanol between low CED liquids, such as chloroform,[10] and the high CED of water. Therefore, ethanol is able to exert a solvation pressure on chloroform, as well as feeling the solvation pressure effect of water.

Chapter 3 describes the study of pure ethanol in the liquid and vapour phase. The effects of the SPM are felt by liquids, but not by vapour. We therefore test the effects of the SPM through comparison of bond lengths and vibrational frequencies in ethanol liquid and vapour.

Ethanol-water mixtures are studied in Chapter 5. The CED of the mixture varies with the concentration of the components. We can calculate the effective CED and hence effective solvation pressure of the mixture using linear interpolation. In this way, the solvation pressure felt by an ethanol molecule in an ethanol-water mixture can be altered within the range $6.76kbar - 22.5kbar$.

We use MD simulations to attempt to explain the pressure dependence of experimental Raman frequencies. Analysis of structure and forces should offer a unique insight into the workings of the SPM.

The aim of this project is to establish the range of validity of the SPM. If the model is proven, the implications are likely to extend to much of wet chemistry and biology. The large pressures predicted by the model would have a large impact on a large variety of systems. In particular we hope to investigate the implications of the model in the field of protein folding. The results of this study will help to determine the feasibility of applying the SPM in this more complex field.

1.4 Raman Spectroscopy

Raman spectroscopy is a technique used in chemistry to study vibrational, rotational and other low frequency modes in a system. It relies on inelastic, or Raman scattering of monochromatic light, usually from a laser. Absorption or emission of phonons by the laser light result in the energy of the laser phonons being shifted up or down. The shift in energy of the laser phonons gives information about the system.

The Raman effect occurs when light interacts with the electron cloud of bonds within a molecule. The light incident on a molecule is either scattered or absorbed. Most of the scattered light is scattered elastically, meaning that it does not change frequency. This is known as Rayleigh Scattering. However, a small fraction of the light is scattered inelastically, which means that its frequency is altered by the scattering process. The difference in energy caused by the scattering of the light is proportional to the vibrational energy of the molecules involved in the scattering. This process of energy exchange in scattering molecules is known as the Raman effect.

A change in polarisability of the electron cloud is required in order for a given mode to exibit the Raman effect. The size of the polarisability change determines the intensity of the peak, whilst the Raman shift determines the vibrational modes of the molecule. Vibrational modes which do not change the polarisability are known as "Raman inactive modes", and do not feature in the Raman spectrum.

A similar complementary spectrum can be obtained using infrared spectroscopy. Infrared spectroscopy works by studying the frequencies of infrared light which are absorbed by a sample molecule. The level of absorption of the infrared light is measured as a function of frequency. A given molecule will absorb only the frequencies that match vibrational modes of the molecule.

As with Raman spectroscopy, not all vibrational modes are present in the resulting spectrum. Only those vibrations which cause a change in the dipole moment of the molecule are absorbed by the molecule. This means that some modes are "IR inactive" and not visible in the resulting spectrum. Raman and Infrared spectra are complementary, since modes which are Raman inactive may be active in the IR spectrum and

vice versa. Producing both specra therefore provides the most information about the vibrational modes present in the system.

Spectroscopy is useful in chemistry, since the vibrational information obtained is specific to the bond which caused the scattering. Analysing the the frequencies at which scattering occurs provides a method for identifying the molecules present in a given sample. In this study, we focus on Raman spectroscopy, since we are working with specific modes all of which are Raman active, and to complement existing experimental work.[9]

The basic components of a Raman spectrometer include a laser source and monochromator, collection optics to gather the scattered light, and a detection system. The required laser frequency depends on the type of material being studied. The collection optics consist of simple lenses in order to focus the light onto a spectrograph system and a laser rejection filter, which is used to filter out elastically scattered light. A CCD camera or photomultiplier tube detects the scattered light, producing a plot of intensity as a function of frequency. Figure 1.1 shows the set-up of a basic Raman spectrometer.

In the study of solvation pressure, the Raman frequencies are important because the frequency of light emitted by a molecule is related to the bond length of a given bond. The Raman frequency of a bond is propertional to the square root of the force constant.

$$v = \frac{1}{2\pi c}\sqrt{\frac{k}{m}} \tag{1.7}$$

The force constant can be calculated as the second derivative of the potential function.

$$k = \frac{\partial^2 U(r)}{\partial (r)^2} \tag{1.8}$$

As a given bond is compressed by an external force, its Raman frequency will tend to increase. The Raman frequency of a bond therefore gives us a measure of the pressure it is experiencing.

When a liquid is squeezed by an external hydrostatic pressure, molecules are pushed closer together and the bonds within the molecules are compressed. In turn, the bonds

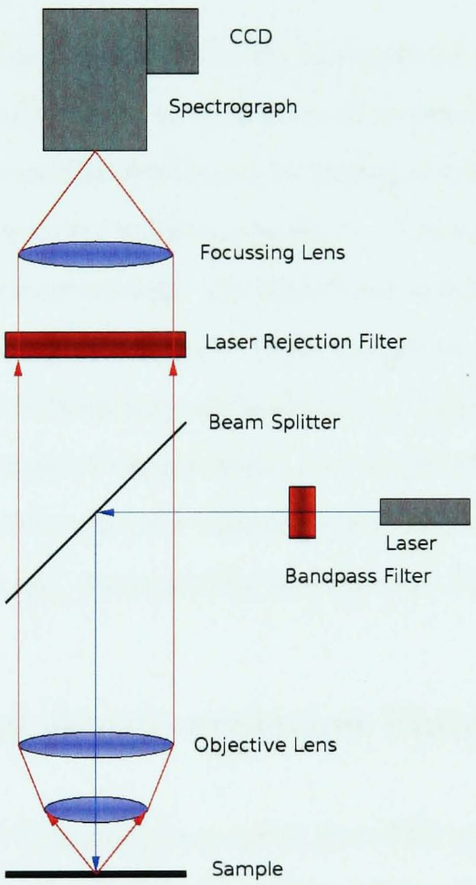


Figure 1.1: Schematic diagram of a simple Raman spectrometer.

vibrate faster, causing the Raman frequency to increase. Similarly, systems of increasing solvation pressure should induce identical frequency shifts in the solute.

Thus the Raman frequency can be used as an experimental tool to monitor the reaction of a molecule to changes in pressure. This makes it an ideal tool for detecting possible shifts in frequency as a function of solvation pressure. The Raman frequencies are recorded as a function of external hydrostatic pressure for a molecule of interest, in this case ethanol.

Knowing the shift in Raman frequency as a function of increasing hydrostatic pressure, we can look for similar shifts as a function of increasing solvation pressure. We increase the solvation pressure felt by ethanol by mixing it with water, which has a higher CED, and hence exerts a larger solvation pressure. The solvation pressure will vary linearly as a function of concentration: the higher the concentration of water, the higher the solvation pressure felt by ethanol. If the solvation pressure effect is equivalent to an hydrostatic pressure, frequency shifts should be identical in each case. We can therefore use Raman spectra as an experimental test of the effects of the SPM in mixtures. The method used to determine the effective CED of a mixture and relationship between solvation pressure and concentration are explored fully in Chapter 4.

1.5 Velocity Autocorrelation Function

We can calculate vibrational frequencies from MD simulations, using a "Velocity Autocorrelation Function" (VACF). This allows direct comparison of solvation pressure in MD and experiment.

Autocorrelation is a method by which time-domain signals, functions or series of values are analysed. It is used frequently in a wide variety of situations in order to find repeating patterns in signals, the presence of a periodic signal which has been buried by noise, or identifying the fundamental frequency of a system which does not contain the frequency itself, but implies it through harmonic frequencies.

The VACF is used in molecular dynamics, and is important because it reveals the underlying nature of the dynamic processes of the simulation. At some time t, for

example the start of the simulation, the velocity components of all atoms in each ethanol molecule in the simulation are recorded

$$\mathbf{v}_i(t_0) = \begin{pmatrix} \mathbf{v}_x(t_0) \\ \mathbf{v}_y(t_0) \\ \mathbf{v}_y(t_0) \end{pmatrix}_i \tag{1.9}$$

for every atom i in the system. The first point on the VACF is obtained by averaging the scalar products $\mathbf{v}_i.\mathbf{v}_i$ for all atoms in the system:

$$C_v(t=0) = \frac{1}{N}\sum_{i=1}^{N}(\mathbf{v}_i(t=t_0).\mathbf{v}_i(t=t_0)) \tag{1.10}$$

At the next time-step of the simulation $t = t_0 + \Delta t$:

$$\mathbf{v}_i(t_0+\Delta t) = \begin{pmatrix} \mathbf{v}_x(t_0+\Delta t) \\ \mathbf{v}_y(t_0+\Delta t) \\ \mathbf{v}_z(t_0+\Delta t) \end{pmatrix}_i \tag{1.11}$$

From this it is possible to calculate the second point on the VACF:

$$C_v(t=\Delta t) = \frac{1}{N}\sum_{i=1}^{N}(\mathbf{v}_i(t_0).\mathbf{v}_i(t_0+\Delta t)) \tag{1.12}$$

We can continue this process until the end of the simulation, giving:

$$C_v(t=n\Delta t) = \frac{1}{N}\sum_{i=1}^{N}(\mathbf{v}_i(t_0).\mathbf{v}_i(t_0+n\Delta t)) \tag{1.13}$$

the results being plotted as a function of time.

If there was no interaction between molecules in the system, the result of this plot would be a straight line, since the atoms in the system would all retain their initial velocity for all time.

If the interaction between particles is weak but not negligible, the magnitude and direction of the velocities are seen to change gradually over time. The scalar product

$\mathbf{v}(t_0).\mathbf{v}(t_0 + n\Delta t)$ decreases on average as the velocity changes. In this case, the VACF is a simple exponential decay, revealing the presence of weak forces slowly decorrelating the velocities. This result is typical of a gas.

In a liquid or solid, the forces are strong. Atoms tend to seek out locations where the attractive and repulsive forces are most balanced. In solids, where positions of atoms are mostly fixed, the atoms tend to oscillate around fixed points. Consequently, the VACF function oscillates strongly between positive and negative values. The oscillations decay in time due to the effects of perturbative forces acting to disrupt the oscillations. The VACF therefore resembles a damped harmonic function.

In a liquid, the oscillations are less defined, since the atoms have no defined positions. Therefore in liquids, the VACF resembles a strongly damped harmonic function.

Taking the Fourier transform of the VACF gives us the vibrational density of states (VDOS) of the system. This reveals the dominant vibrational frequencies of the system. Frequencies present in Raman spectrum are also present in the VDOS function. However, the VDOS may also contain any Raman inactive modes present in the sample. Since we are comparing the positions of known modes, these additional modes will not distract from the results.

Therefore the VACF method allows a direct comparison between the molecular dynamics simulations and experimental data, and a test of the SPM which can be applied in both situations. In some cases, the VDOS function may show additional peaks in frequency compared experimental Raman spectrum. Such peaks reveal the positions of "Raman inactive modes", and occur when a vibration causes no change in the polarisability of the molecule, meaning that the vibration is not present in the Raman spectrum.

Chapter 2

Molecular Dynamics

2.1 Brief introduction to Molecular Dynamics

Molecular Dynamics is the process by which atomic positions, velocities and forces are calculated as a function of time using the laws of classical mechanics, notably Newton's second law:

$$\mathbf{F}_i = m_i \mathbf{a}_i \tag{2.1}$$

for each atom i in a system of N atoms, where $\mathbf{F}$ is the force on the particle, m is its mass and $\mathbf{a}$ is the acceleration experienced by the particle.

Through integration, it is then possible to obtain a trajectory that describes the positions velocities and acceleration of the particles as a function of time. Since there is no analytical method of solving exactly the equations of motion for an ensemble of atoms, a numerical approach is used. Integration algorithms are used in MD simulations to solve the equations numerically. These algorithms make the assumption that the positions, velocities and forces can be approximated by Taylor series expansions.

$$\mathbf{r}(t+\delta t) = \mathbf{r}(t) + \mathbf{v}(t)\delta t + \frac{1}{2}\mathbf{a}(t)\delta t^2 + ... \quad (2.2)$$

$$\mathbf{v}(t+\delta t) = \mathbf{v}(t) + \mathbf{a}(t)\delta t + \frac{1}{2}\mathbf{b}(t)\delta t^2 + ... \quad (2.3)$$

$$\mathbf{a}(t+\delta t) = \mathbf{a}(t) + \mathbf{b}(t)\delta t... \quad (2.4)$$

There are several integration algorithms in current use. In these algorithms there is a trade off between accuracy and computational complexity.

In this work, we used the Verlet Leapfrog (VL) Algorithm. In this algorithm, velocities are first calculated at $t + \frac{1}{2}\delta t$. These are used to calculate the positions at $t + \delta t$.

$$r(t+\delta t) = r(t) + v(t+\frac{1}{2}\delta t)\delta t \quad (2.5)$$

$$v(t+\frac{1}{2}\delta t) = v(t-\frac{1}{2}\delta t) + a(t)\delta t \quad (2.6)$$

In the VL algorithm, velocities and position are not calculated at the same time at any point in the simulation. Calculations of velocity and position 'leapfrog' over each other over the course of a simulation. The advantage of VL is that, unlike in simpler algorithms, velocities are calculated explicitly. The disadvantage is that they are not calculated at the same time as the positions.

It is necessary to use the velocities calculated at times of $t/2$ in order to estimate the velocity of atoms at time t. This is achieved by taking a simple average of the velocity immediately before and after the time t.

$$\mathbf{v}(t) = \frac{1}{2}[\mathbf{v}(t+\frac{1}{2}\delta t) + \mathbf{v}(t-\frac{1}{2}\delta t)] \quad (2.7)$$

Inter-atomic potentials are used to determine the forces on the atoms. In this case, we use classical potentials, which means that they depend only on the relative positions of atoms, and not on any quantum properties of the atoms.

For example, bonds connecting the individual atoms in ethanol are modelled as spring potentials. The force applied by the bond on the two atoms is a function of the distance

between the atoms. There is an equilibrium distance at which the force applied by the bond is zero. At distances closer than equilibrium, the force is repulsive, acting to push the atoms apart. At distances greater than equilibrium, the force acts to bring the atoms closer together.

Properties of the simulation can be calculated using statistical mechanics. For example, the temperature of the simulation can be calculated from the velocities of the atoms

$$T = \frac{\sum_{i=1}^{N} m_i \mathbf{v}_i^2(t)}{k_B f} \tag{2.8}$$

where i is the atom label, k_B is Boltzmann's constant and f is the number of degrees of freedom of the system.

The VL Algorithm generates trajectories in the microcanonical (NVE) ensemble. An ensemble is a statistical mechanical concept which is fundamental to the workings of MD simulations. Imagine a given molecular simulation is replicated many times, each 'copy' having the same physical properties, such as the number of molecules, density and temperature, but achieved through different atomic positions and velocities. Such a collection of copies is known as an ensemble. NVE refers to the parameters which are kept constant during the simulation: Number of Molecules, Velocity and total Energy.

Because of the way that the ensemble is constructed, at any given time, the instantaneous value of the bulk properties may differ from the general value. The true value of a property is calculated as the average of all possible replicas. This is known as the 'ensemble average'.

In the course of a simulation, the atoms are involved in constant dynamic motion. Positions and velocities change as a function of time. Each time-step results in a different 'configuration' of the molecules of the system. Each configuration generates new instantaneous values for the bulk properties of the system, such as the pressure and temperature. An ensemble average must be calculated in order to obtain the true thermodynamic value of the variable. In MD this is achieved by performing the average of all the configurations generated by the simulations. This makes the assumption that

the average of the simulated configurations is equivalent to the ensemble average. This is known as the ergodic hypothesis,[20] and is yet to be rigorously mathematically proven, but holds true assuming good statistical data for systems in equilibrium.

The microcanonical ensemble generates trajectories based on configurations with a common fixed energy E. The energy is fixed by the assumption that the system is isolated. This means that no energy is allowed to enter or leave the system. By definition, isolating the system makes the NVE ensemble the simplest of the ensembles through which trajectories can be calculated.

By placing the system in a heat sink, and allowing exchange of energy between the system and an external source, one can perform simulations at constant temperature. The trajectories generated by the simulation are based on a canonical (NVT) ensemble. The ensemble keeps the following parameters constant: number of molecules, volume and temperature. This is achieved using the Hoover method, in which the equations of motion are modified to include a frictional correction to the acceleration of the atoms in the system.

$$\frac{d\mathbf{v}(t)}{dt} = \frac{\mathbf{f}(t)}{m} - \chi(t)\mathbf{v}(t) \tag{2.9}$$

where χ is the friction coefficient given by:

$$\frac{d\chi(t)}{dt} = \frac{1}{\tau_T^2}\left(\frac{\mathcal{T}}{T_{\text{ext}}} - 1\right) \tag{2.10}$$

The temperature of the simulation is maintained close to a given target value. Fluctuations are permitted, the magnitude of which are controlled using a relaxation constant. This control over the system temperature makes the ensemble useful for equilibrating the system at a given temperature. However, due to the frictional terms introduced into the equations of motion, the velocities are scaled by the algorithm over the course of the simulation. It is therefore not suitable for the data acquisition phase of the simulation, since velocities are required for use in the VACF.

During the course of this work we also make use of the isobaric-isothermal (NPT) ensemble. The number of molecules, pressure and temperature are kept constant during an NPT simulation. The equations of motion are modified through the use of a heat sink, and by allowing the volume of the system to be controlled by a barostat. The barostat forces changes in volume by scaling the lengths of all the bonds in the system.

The NPT ensemble is vital in these simulations in order to place ethanol molecules under external pressures, and also to achieve standard pressure where required. Again, it is not suitable for data acquisition, since properties we wish to measure, notably the bond length, are altered by the barostat. Hence, all simulations use the NVE ensemble for data acquisition.

2.2 DLPoly

We performed Molecular Dynamics simulations using the DLPoly package from Daresbury Labs.[21] DLPoly has a solid reputation as a basis for a wide variety of atomic and molecular simulations. The software has been used successfully within the physics department at the University of Surrey for many years.

Simulated molecules were contained in a regular periodic cube with sides of ~2nm. We used 100 molecules in all pure ethanol simulations. In ethanol-water simulations, we used 326 water molecules and between 1 and 80 ethanol molecules.

Periodic boundary conditions mean that the simulated volume is used to approximate a small section of a larger volume of liquid. On contact with one of the 'walls' of the simulated box, atoms do not rebound as one might expect. Instead they disappear from the simulated volume, as if leaving a region of interest within a larger simulation. The molecule reappears at the opposite boundary, simulating an identical molecule entering the region of interest. The process is illustrated in Figure 2.1.

Because of the use of periodic boundary conditions, we use the Ewald summation technique[22] for calculating electrostatic interactions. The Ewald sum is an efficient method of calculating long range forces in a periodic system.

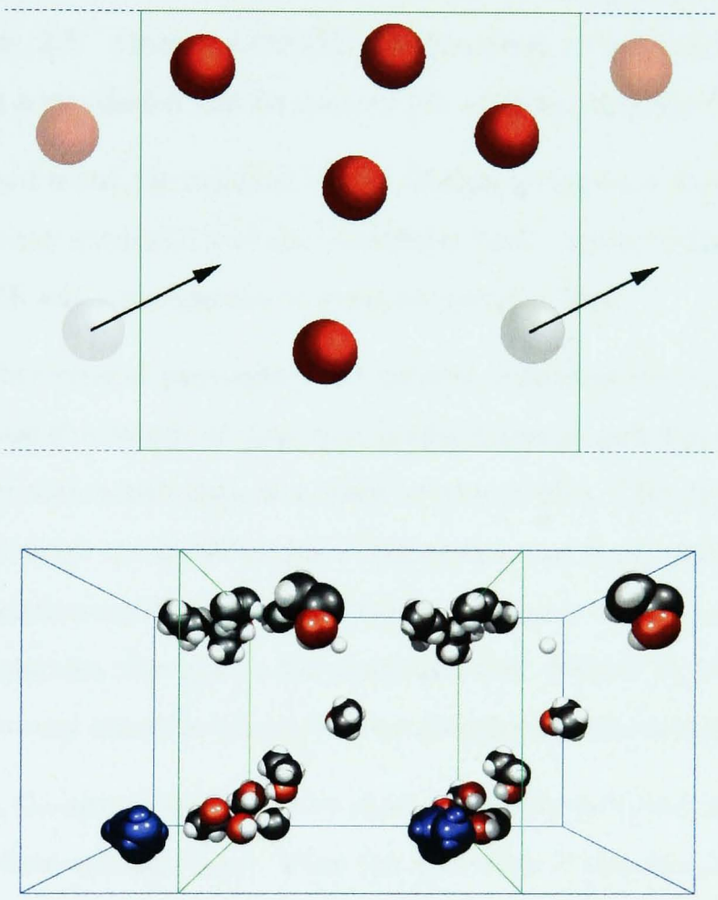


Figure 2.1: Periodic Boundary conditions in DLPoly. In each figure, the simulated box is shown in green. Surrounding 'boxes' simulated by the use of the periodic boundary conditions are shown in blue. A molecule travelling towards the right hand edge of the box is shown as a solid blue circle. When it reaches the edge it will be reappear one box length to the left, with an identical velocity. The effect is as if one molecule leaves a region of interest, whilst another simultaneously enters it.

2.2.1 Input Files

DLPoly requires three input files in order to begin a simulation. The first of these is the CONFIG file. This file details the simulation cell vectors and positions of the molecules at the start of a simulation. The methods used to generate a CONFIG file are described in Section 2.5. Once a CONFIG file has been created, final positions of molecules following a simulation can be used as the starting point for further simulations.

The second input file required by the DLPoly program is the CONTROL file. This file defines many parameters of the simulation itself. These include the choice of ensemble, along with any temperature or pressure specifications.

Many other critical parameters are defined, including the number of steps in the simulation and the length of time that passes between each time-step. We also define the frequency with which data is written to output files. Clearly simulations must be long enough to allow the equilibration of the system and good statistical data to be acquired. The simulation must run for some minimum length of time in order to allow movement of the molecules throughout the simulated box. Failure to allow this movement of the molecules may result in incomplete equilibration of the simulation.

However, the simulations must be short enough so as to not require too much processor time, or data storage space. Thus the time-step of the simulation, the amount of time which passes between each iteration of the molecules, is a crucial parameter.

The time-step of the simulation must be smaller than the oscillation period of the fastest mode in the system. Experimental data[9] shows that ethanol has a maximum vibrational frequency of $2928cm^{-1}$. This corresponds to an oscillation period of $11fs$. It is good practice to choose a time-step at least several orders of magnitude lower than this limit.[22]

Also, if the time-step is too long, the system can become unstable, as the assumption that the acceleration of the particle is constant during a single time-step breaks down at large time-steps. Conversely, if the time-step is too short, the total duration of the simulation is not long enough to gain any useful insight into the interactions.

This was especially evident in previous work.[9] In this case the hydrogen bonding tail

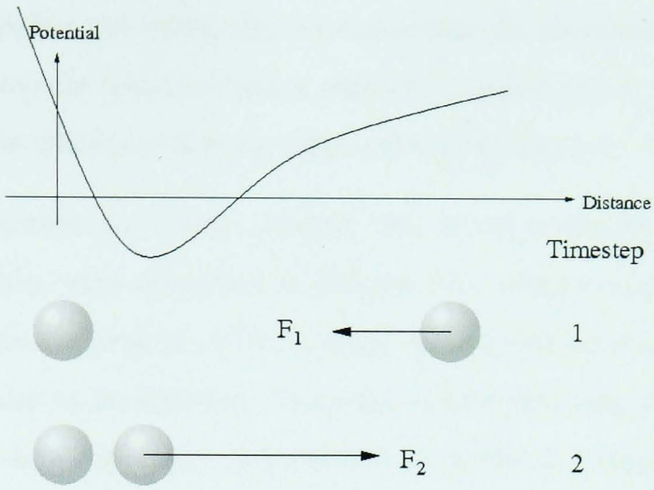


Figure 2.2: The interaction between two particles in one dimension, to illustrate the importance of choosing a time-step carefully.

of the ethanol molecule was especially sensitive to the time-step of the simulation. This was partly due to problems with the molecular model, however it served to illustrate the importance of a choosing a good time-step for the simulation. Based on this information, the time-step chosen for this work is $0.2fs$.

The process by which instabilities develop in simulations with a large time-step is illustrated in Figure 2.2. In this example, the time-step is set too high. The equations of motion assume that the force is constant for one entire time-step. The particle shown on the left moves with a constant velocity v_1 for the duration of the time-step. Because there is no intermediate calculation of the force applied to the particle, it may approach a second particle, shown on the right, without experiencing a repulsive force. This repulsive force should cause the molecule to be slowed and stopped gradually, and prevent the molecules approaching too closely.

Instead, when the forces are recalculated at time-step 2, the molecules are too close together, and hence the repulsive force is high. This causes the particle to be moved even further during the next time-step, increasing the scale of the problem. This 'rattling' of the system continues until eventually the molecules break apart.

Also specified in the CONTROL file, are the cutoff radii for long range parameters such as the van der Waals forces. These parameters are set so that processor time is not wasted calculating large numbers of extremely small forces between separated

atoms. The higher the value, the more accurate the simulation, however, the accuracy gains are extremely small beyond a certain threshold radius. By carefully setting this parameter, the efficiency of computational simulations can be greatly improved.

Another important parameter defined here is the ensemble used for the simulation. These ensembles were described in Section 2.1. During equilibration NPT and NVT ensembles were required in order to allow correct temperature, pressure and structure of the molecules to be achieved. Once the system was equilibrated, an NVE ensemble was used for data collection. This choice of ensemble is accompanied by the relevant target temperatures and pressures.

The FIELD file contains information about the molecules in the system. This includes charges and masses of atoms, the functional form of the bonds and angles, and the parameters used with these functions to calculate the bond strength as a function of distance. If the required functional form is not implemented in the DLPoly code, it is possible to provide a table of the energy of the bond at a range of distances. This feature makes the DLPoly program very flexible in the types of simulations it can perform.

This information is generally obtained through the choice of one of a large number of 'force fields', which are use a variety of simple rules in order to define molecules for a specific purpose. The choice of molecular model, and hence the details used to complete the FIELD file, are discussed in some detail in the following sections.

2.2.2 Output files

DLPoly outputs the simulation data to a small number of files. Some of the details of the contents of these files, and frequency of output, is specified in the CONTROL file.

The OUTPUT file contains a general overview of the simulation. It contains a summary of the input files, as well as brief summaries of the progress of the simulations. All of the information in the OUTPUT file can be found elsewhere. It is intended to be a printed record of the simulation, rather than analysed.

The most important output file is the HISTORY file. This file can contain positions, velocities and forces applied to the molecule at regular intervals. In order to keep

the file size to a minimum, we did not collect force data, but collected positional and velocity data every 25 time-steps. Most of the analysis in this work is performed using data from the HISTORY file.

The STATIS file is another important file for analysis of the simulation. It contains instantaneous values of variables such as the pressure, temperature and energies. STATIS contains all of the information required in order to check that the mixture is equilibrated and proceeding under the right conditions. In this work, the file is updated every 100 time-steps, which is adequate for the purposes of checking the trends of data.

The REVCON file contains the final positions velocities and forces of the atoms in the simulation. Its contents can be copied to a the CONFIG file if a simulation is to be continued.

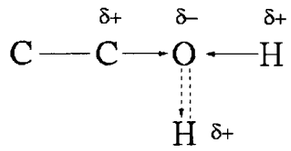
2.3 Choice of Molecular Model

The requirements of this project place strong constraints on the molecular models that are suitable for use in the modelling of ethanol. Firstly, the model must describe all atoms explicitly, since bond lengths must be measured and vibrational properties explored as a function of pressure. For this reason, atoms must be unconstrained in movement along the length of the bond and free to express angular twisting and rocking motions. This rules out simplifying potentials, such as most variants of OPLS,[23] which use single points to represent groups of molecules such as the ethyl and methyl group. It is also inappropriate to use force fields which use rigid geometries. This includes models such as ECEPP[24] and JUMNA,[25] which both make use of fixed bond lengths.

It is also desirable that the model should allow the modelling of full proteins, as well as simple organic molecules, since the project will be extended in the future to include larger molecules. Therefore, models such as MM4,[26] which are tailored specifically for smaller molecules were excluded from consideration.

Finally, in order to keep the simulations simple, it was decided that use of polarisable force fields would be an unnecessary addition to the model. Based on previous experience of modelling ethanol within the department[9] and elsewhere,[27] it was decided

Figure 2.3: In a polar model of ethanol, the charge cloud of an oxygen atom is deformed by hydrogen bonding, reducing the strength of the carbon-oxygen bond. This allows the bond to be further stretched by the hydrogen bonding interaction. This will not be possible in our fixed charge system, removing a potential masking interaction.



that the use of non-polarisable models allow satisfactory modelling of the structure and properties of ethanol. Using polarisable force fields would unnecessarily increase the processor time required to complete simulations, and offer little benefit to the model reproducing the properties required.

One of the potential effects of excluding polarisability of the ethanol molecule is that some hydrogen bonding effects may be weakened. In a polarisable model of ethanol, the charge cloud of an atom is would be strongly deformed by hydrogen bonding. Figure 2.3 shows one such example of the effects of polarisability on atoms involved in hydrogen bonding. In this example, the oxygen of ethanol is hydrogen bonded to the hydrogen atom of a neighbouring molecule. The positive charge of the neighbouring hydrogen atom tends to draw electrons away from the oxygen atom. This in turn reduces the strength of the carbon-oxygen bond, causing a lengthening of the bond and strengthening of the hydrogen bond.

Since the SPM makes no attempt to predict the effects of hydrogen bonding between solvent and solute, such an interaction would likely constitute a "masking interaction". Therefore, in may be beneficial to remove this effect, in order to better test for the solvation pressure effect. If we see evidence for the SPM in this case, it would then be interesting to repeat the simulations using a polarisable force field, providing a useful insight into the nature of the masking interactions.

Several force fields were considered carefully for suitability for this project. These were CHARMM,[28] OPLS-aa,[29] GROMOS,[30] and Cornell.[31]

Of these CHARMM was found to have a history of errors in reproducing the bulk properties of hydrocarbons. In particular, modelling of butane using CHARMM92, reveals

as much as 63% error in density.[32] This error has since been reduced significantly, but at the expense of the simplicity and consistency of the model.

There is evidence to suggest that the GROMOS force field gives a poor approximation of the interaction between methyl groups and water. This is due to the non-standard treatment of water in interactions which do not involve hydrogen bonding. This has been shown to result in an overly hydrophilic nature of the methyl group.[33,34]

The OPLS-aa force field on the other hand, provides good fitting of properties of hydrocarbons. However this is achieved by altering the van der Waals well depths and the radius of hydrogen. This results in good reproduction of density and heat of vapourisation data for larger molecules under normal strain. However, this is at the expense of its accuracy of simulations under stress. This may be detrimental in high pressure simulations, such as the pure ethanol simulations described in Chapter 3, in which ethanol is placed under pressures of up to $15kbar$.

We chose the Cornell model because of its strong reputation in simulations of both small organic molecules, including ethanol[27] and larger proteins including DNA.[35,36] The work of Cheatham et al[27] reports very accurate representation of the ethanol molecule, in particular important bulk properties such as the density and heat of vapourisation were found to be represented accurately by the model. This is achieved without over-complicating the model, as is the case with some alternatives.

Therefore, the Cornell model appears to offer the best overall model for the requirements of this project. The details of the initial model used are shown in Section 2.4. This initial Cornell model is then tested for accuracy of reproduction of the bulk properties of ethanol, and parameters refined accordingly.

Note that in spite of the points made above, it is likely that all of the models described about could have been suitable for this study. The choices made simply reflect the process used in order to select a model for this study. It may be possible to argue a case for all four models, and utilise each successfully.

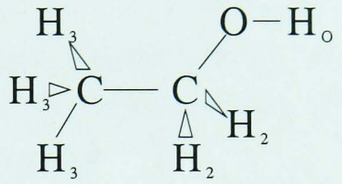


Figure 2.4: The structure of ethanol. The labels on the atoms refer to the naming convention used in this work

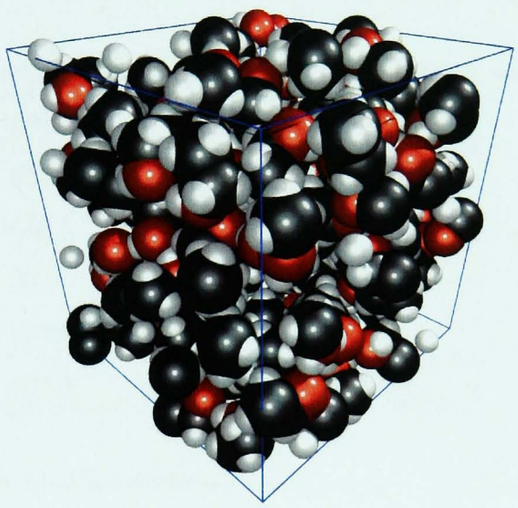


Figure 2.5: A box of 100 ethanol molecules from the simulation. This and all simulation images are generated using custom molecular visualisation software, designed specifically for this project. The software is a front-end to the Persistence of Vision ray-tracer, POV-Ray.[37]

2.4 The Cornell Model of Ethanol

The following is a description of the Cornell Model of ethanol. Unless otherwise stated, equations and parameters are taken from the work of W. D. Cornell et al.[31]

The structure of an ethanol molecule is shown in Figure 2.4. The labels on the atoms show the naming convention used in this work. Figure 2.5 shows a simulated box of 100 ethanol molecules.

In the the Cornell model, all bonds are simple harmonic potentials given by:

$$U(r) = \frac{1}{2}k(r - r_0)^2 \tag{2.11}$$

where k is the spring constant, r is the distance between the bonded atoms, and r_0 is

Table 2.1: Potentials used with equation 2.11 to define chemical bonds. The parameters and equation defining the C-H bond are described in Section 2.6.

Bond type	k ($kcal\ mol^{-1}\ \AA^{-2}$)	r_0 ($\AA$)
C-C	400.0	1.526
C-O	400.0	1.410
O-H_O	600.0	0.960

Table 2.2: Potentials used with equation 2.12 in defining bond angles

Angle Type	k ($kcal\ mol^{-1}\ radian^{-2}$)	θ_0 (deg)
HĈH	70.0	109.5
HĈC	100.0	109.5
CĈH	70.0	109.5
CÔH	110.0	108.5
HĈO	100.0	109.5

the equilibrium distance between bonded atoms.

Three body angular potentials are also harmonic, as given by:

$$U(r) = \frac{1}{2}k(\theta - \theta_0)^2 \tag{2.12}$$

where θ is the angle formed by two bonds, and θ_0 is the equilibrium angle.

Finally dihedral four body potentials:

$$U(r) = A[1 + \cos(m\phi - \delta)] \tag{2.13}$$

where ϕ is the dihedral angle.

Intermolecular forces are modelled with Lennard Jones (12-6) potentials:

$$U(r) = \left(\frac{A}{r^{12}}\right) - \left(\frac{B}{r^6}\right) \tag{2.14}$$

and Coulombic two body terms, whose charges are given by the RESP model[38]. We used the values provided by Cheatham et al[27] for the charges on ethanol. Using these values gives us a set of charges known to produce reliable results in simulations.

Table 2.3: Dihedral Potentials used with equation 2.13 to define skeletal molecular structure

Dihedral	A ($kcal\ mol^{-1}$)	δ (deg)	m
HCCH	0.16	0.00	3.00
HCCO	0.16	0.00	3.00
CCOH	0.17	0.00	3.00
HCOH	0.17	0.00	3.00

Table 2.4: The van der Waals parameters defining the interactions between atoms in different ethanol molecules. H_3 refers to the three hydrogen atoms attached to the carbon atom at the non-polar head of the molecule. H_2 refers to the two hydrogen molecules attached to the central carbon atom, and H_O refers to the hydrogen atom at the polar tail of the molecule.

Lennard Jones Pair	A ($kcal\ \AA^{12}\ mol^{-1}$)	B ($kcal\ \AA^{6}\ mol^{-1}$)
C-H_2	2.29E+04	1.05E+02
C-C	1.04E+06	6.76E+02
C-H_O	0.00E+00	0.00E+00
C-O	7.91E+05	6.93E+02
C-H_3	9.61E+04	1.26E+02
H_2-O	4.63E+04	1.03E+02
H_3-O	6.77E+04	1.24E+02
O-O	5.82E+05	7.00E+02
H_2-H_2	3.26E+03	1.43E+01
H_3-H_3	7.52E+03	2.17E+01
H_2-H_3	4.99E+03	1.77E+01
O-H_O	0.00E+00	0.00E+00
H_O-H_O	0.00E+00	0.00E+00
H_2-H_O	0.00E+00	0.00E+00
H_3-H_O	0.00E+00	0.00E+00

Table 2.5: The charges give by the RESP model, as provided by Cheatham et al.[27]

Atom Type	Charge (u)
C_1	-0.0990
H_3	0.0345
C_2	0.3318
H_2	-0.0294
O	-0.6718
H_O	0.4143

2.5 Equilibrating and Running a Simulation

The starting configurations were generated using a modified version of the rndrot program provided with DLPoly. Rndrot takes two input files containing the coordinates of single molecules of ethanol and water. The program applies random translation and rotational matrices to the molecules in order to generate a starting configuration for the molecules.

In previous work, it was found that poor starting configurations of the molecules could easily lead to the breakup of the ethanol molecule.[9] This was caused by molecules being placed too close together resulting in high initial forces, and hence large trajectory changes between time-steps. There are therefore rudimentary checks to ensure that two molecules are not placed too close together, in order to keep the forces applied to the molecules as low as possible. An atom is relocated if it is placed within $2\mathring{A}$ of its nearest neighbour. RDF data shows this to be the minimum distance separating two atoms in an equilibrated system. If a molecule is closer than this to its nearest neighbour, it is rejected.

No further effort is made to generate a realistic starting configuration, this is left for correction during the equilibration phase of the simulation. The program first places all of the required ethanol molecules in the simulated box. Once this is completed, water molecules are placed around the existing ethanol molecules. Arranging molecules in this order is optimal because the water molecules are smaller, and hence are more easily able to fill gaps.

Initially, rndrot was set up to calculate the required box size for a given simulation, based on the expected density of the liquid mixture. However, since the molecules were placed randomly within the box rather than generating a realistic configuration of molecules, it was found that a large number of molecules were out of place, causing disturbances in the system that required a long equilibration period to settle out.

Therefore it was found to be of benefit to the equilibration process to place the molecule in a much larger box, with sides of around $\sim 40\mathring{A}$, and to reduce the size of the box using an NPT ensemble as the first stage of the equilibration process. Equilibrating in this manner meant that when the target box size is reached, the atoms have assumed a far more natural configuration, and the disturbances in the system are less pronounced and require less time to relax.

The system was equilibrated by performing a zero Kelvin simulation of 10^5 time-steps of $0.2fs$ duration. The temperature was then raised in $50K$ intervals for a further 10^5 time-steps, until standard temperature of $298K$ was reached.

The pressure of the system was raised as required by allowing the cube to shrink during an NPT run during a short run of 10^3 time-steps with a target constant pressure of $30kbar$. This target pressure was not reached due to the short timescale of these runs, the result being regular small increases in pressure. At each pressure, the system was equilibrated using an NVT ensemble to ensure good realignment of the molecules before any data acquisition. As with initial equilibration, this consisted of 10^5 time-steps of $0.2fs$.

Data acquisition was then performed using an NVE ensemble. It was necessary to use this ensemble for data acquisition, as both NVT and NPT types of simulation interfere with properties we wished to measure. For example in an NPT simulation bond lengths are scaled in proportion to the change in box size. Data were collected for $20ps$, again using 10^5 time-steps of $0.2fs$ duration. Atomic positions and velocities were recorded every 25 time-steps, while thermodynamic properties were output every 100 time-steps.

In order to obtain vibrational data, we obtained a VACF from the acquired velocity data and performed a Fourier transform to obtain data as a function of frequency. This process is described fully in Section 1.4.

To produce a Raman spectrum we must take the Fourier transform of the VACF. In a Fourier transform, the maximum frequency is determined by the time that passes between successive data points. We chose time-step and data collection intervals carefully in order to obtain a frequency range comparable with experimental results. The smaller the time-step, the greater the maximum frequency of the theoretical Raman spectrum.

In this case, it is preferable to obtain the same range of frequencies as was determined experimentally in previous work.[9] Dunstan and co-workers examined Raman frequencies in the range $300 - 3500cm^{-1}$. A time-step of $0.2fs$, with data collected every 25 time-steps, corresponds to a maximum frequency of $3333cm^{-1}$. This defines the maximum time-step that can be used.

In addition, Dunstan and co-workers reported a spectral resolution of "better than $1cm^{-1}$".[9] In the Fourier transform, longer simulations result in corresponding improvements to the spectral resolution. It will not be possible to achieve the same resolution, as it would require the collection of significantly more data, but it is an important consideration when choosing these parameters.

2.6 Testing and Refining the Cornell Model

At this stage, we tested the model by calculating the equilibrium density, CED and compressibility of ethanol liquid at NTP. We also compared the Raman spectrum data, and RDFs with their experimental equivalents. The bulk properties of the ethanol model were all found to be in good agreement with experimentally obtained values. However, there were severe problems with Raman spectrum obtained from the MD simulations. It was found that the default Cornell model was unable to reproduce the high frequency ($\sim 3000cm^{-1}$) Raman modes of ethanol liquid. These modes are severely underestimated by the Cornell model, appearing instead at around $\sim 1200cm^{-1}$.

These high frequency Raman modes of ethanol liquid are caused by the vibration of the C-H bonds (C-H_3 and C-H_2). It seems the simple harmonic term describing the C-H bond in the Cornell model is not sufficient to describe these high frequency terms.

However, these modes had been previously modelled successfully using the DREIDING model,[39] which used more complex Morse potentials to model these interactions.

In order to provide better agreement with the experimentally obtained Raman spectrum of ethanol, and hence provide the opportunity for meaningful comparison between the two, it was decided that the C-H bond should be modified. In previous papers, the DREIDING model had produced very good fits to the Raman data in this case, it was decided to utilise the C-H bond parameters of this potential. The DREIDING potential uses a more realistic Morse function.

The Morse potential is defined by the equation:

$$U(r) = E_0([1 - \exp(-k(r - r_0))]^2 - 1) \tag{2.15}$$

We used parameters given as in the DREIDING model, which gave good reproduction of these bonds:

$E_0 = 70 kcal mol^{-1}, r_0 = 1.09\dot{A}, k = 2.24\dot{A}^{-1}$

The model was re-tested as before, except that existing positional data were used as the starting point of the simulation, rather than generating new configurations. The Raman spectrum of the simulation showed much improved correlation with the experimental data. As expected, the DREIDING parameters used for the C-H bond produces far better agreement with the experimental Raman spectrum. As described in the Chapter 3, this does not seem to be to the detriment of other properties of the liquid.

Figure 2.6 compares the morse potential describing the C-H bond and the harmonic potential it replaces. The harmonic potential gave an unusually low vibrational frequency. The figure shows that the gradient of the harmonic potential is notably lower, especially close to the equilibrium value of $1.09\mathring{A}$. This is the likely reason for the inaccuracy of the vibrational frequencies created by the potential. This inaccuracy could be corrected by increasing the spring constant of the bond. However, we felt that it was preferable to replace the potential using values from an established model.

In comparing the morse and harmonic potentials, we raise the question of whether or not the harmonic potential are sufficiently accurate to convey correct responses to

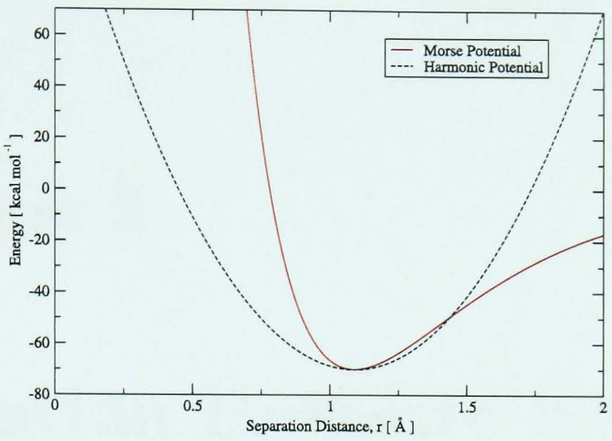


Figure 2.6: The harmonic potential of the Cornell model was replaced by the morse potential of the DREIDING model. This was due to unusually low vibrational frequency of the C-H bond of the Cornell model. This figure compares the two potentials.

pressure. The strength of the harmonic potentials varies as a function of x^2, and hence the force varies linearly with distance. If a pressure is exerted on the molecule, it clearly affects the length of the bond. In a harmonic system, we would expect this to result in a linear change in displacement of the atom, and hence a linear change in bond length.

However, under the influence of the range of pressures we will be using in this simulation, we still remain very close to the bottom of the potential well. It is therefore possible for the harmonic potential to closely approximate the bond potential in the required range. This suggests that the component of the change in bond length caused by a pressure is likely to be linear for all potentials.

Also, the system is significantly more complex than bond interactions alone can convey. Angular and dihedral potentials as well as coulombic interactions also significantly affect the displacement of the atoms. Coulombic interactions in particular allow a significant variation in bond length as a function of pressure. The complete potential describing a given atom allows many ways for a bond to exibit a full range of responses to increasing pressure. Therefore, harmonic potentials are a valid approximation for pressurised systems in the range studied in this work.

Chapter 3

Solvation Pressure in Pure Ethanol

In this chapter we present the results from pure ethanol simulations. Pure ethanol liquid was simulated using 100 ethanol molecules contained within a periodic cube at Normal Temperature and Pressure (NTP). Ethanol vapour is simulated as a single ethanol molecule contained within the same cube, also at NTP. The molecules were initially arranged using equilibrated configurations from the work of Dunstan et al,[9] ensuring the best possible starting separation and molecular structure.

We begin by describing the processes used in order to test the quality of the Cornell model of ethanol. This is shown in Section 3.1. There we present bulk properties of ethanol liquid, and compare these with experimental values. We compare theoretical Radial Distribution Functions (RDFs) of ethanol with experimentally determined peaks.

Section 3.2 describes the process of determining bond lengths of ethanol vapour and liquid under hydrostatic pressure. We demonstrate the change in bond length on transition from vapour to liquid. The SPM describes the magnitude of the change in bond length in many cases.

Section 3.3 shows the vibrational frequency analysis of the pure ethanol data. We describe the method used to calculate vibrational frequency data, and apply it to the

Table 3.1: some bulk properties shown as a check of the quality of the model

Property	Model Value	Experimental value
Density (gcm^{-3})	0.81 ± 0.01	0.789[40]
Compressibility (GPa^{-1})	1.14 ± 0.05	1.12[41]
Cohesive Energy Density ($kbar$)	6.52 ± 0.01	6.76[40]

ethanol simulations. The shift in vibrational frequency on transitions from vapour to liquid is shown to be consistent with the SPM in most cases.

3.1 Testing the Molecular Model

Table 3.1 shows some bulk properties of ethanol calculated as a test of the quality of the model. The CED was calculated using equation 1.1 using a value of ΔU obtained by subtracting the internal energy of the vapour from the average internal energy for the liquid at NTP. The value was calculated to be $6.52kbar$, which is 3.6% lower than the experimental value.[40]

Density was calculated using the standard equation:

$$D = \frac{M}{V} \tag{3.1}$$

where M is the mass of the ethanol liquid. The density was calculated by deducing the mass of 100 ethanol molecules from their molecular masses, and dividing into the volume occupied by ethanol under a pressure of $0kbar$. As shown in table 3.1, this value was calculated to be slightly high in the model, showing a 2.6% deviation from the experimental value.[40]

Compressibility was calculated using

$$C = \frac{1}{V}\frac{\partial V}{\partial P}\bigg|_{P=0} \tag{3.2}$$

and found to be 1.8% higher than expected experimentally, but showing agreement within the limits of the error of the calculation.

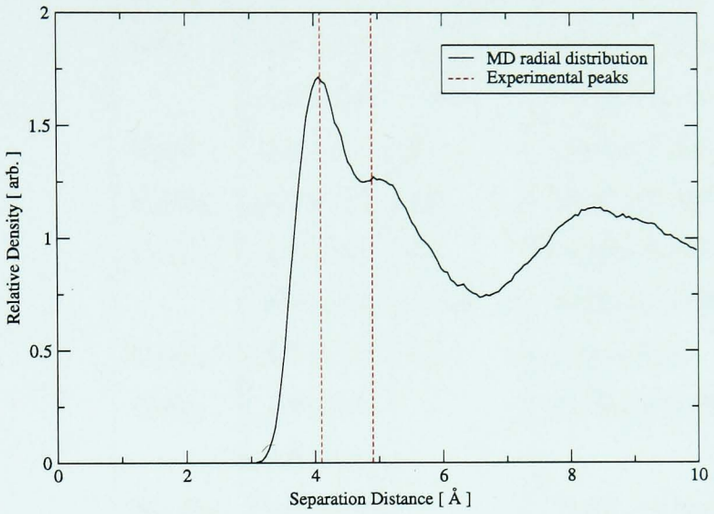
These results represent very good reproduction of bulk properties and show that the Cornell model offers a realistic representation of the molecule.

Radial Distribution Functions (RDFs) also provide a good test of the quality of the ethanol molecule in comparison with experimentally obtained values. The RDF of a system describes how, on average, the atoms are radially packed around each other. They provide a useful method of describing the average structure of disordered systems such as liquids. We calculated RDFs using simple geometry, and compared curve peaks with experimental values. We also calculated the dependence of the vibrational frequency on pressure, giving us a measure of molecule separation as a function of pressure.

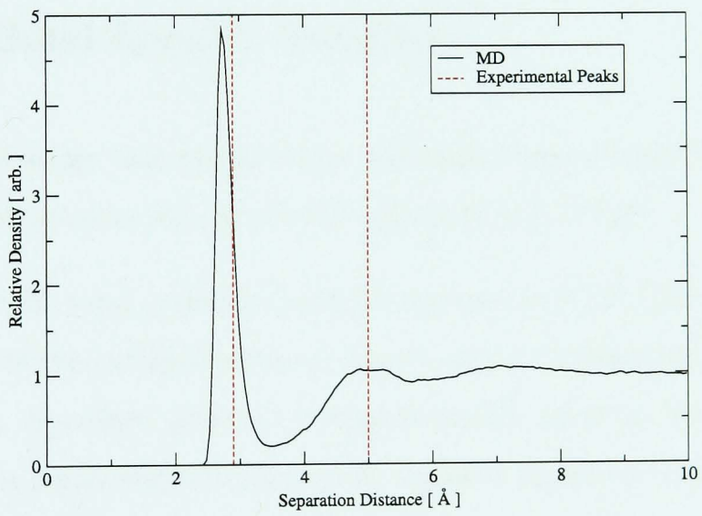
The results of this process for all pairs of atoms are shown in Table 3.2. Generally there is an excellent relationship between the predictions of the model, and the values obtained experimentally[42]. The only bond which seems to be inaccurate in this respect is the O-O distance, which is slightly shorter than experimental x-ray analysis predicts. It is most likely that the harmonic function defining the interaction is slightly too strong in the Cornell field. However, on the whole the data provides good verification of the quality of the MD simulations used in this work.

It is interesting to note how the peak of the RDFs of different atom pairs respond to increasing pressure. As shown in Table 3.2, atoms involved in hydrogen bonding show RDF peaks that are almost independent of pressure. This is not unexpected, due to the high Coulombic charges on the O and H atoms which are able to 'lock' the atoms in their place and resist the effects of increasing pressure. Those at the non-polar head of the molecule, by contrast, are squeezed a factor of ten times closer together over the same range of external pressure.

For this reason it is reasonable to assume that any solvation pressure felt by the O-H_O on the ethanol molecule is masked by a far stronger hydrogen bonding interaction. Studying the RDF data in this way provides a good indication of any bonds that might display interactions acting to mask the solvation pressure effect.



(a) C-C bond



(b) O-O bond

Figure 3.1: Selected RDF data for ethanol. On the left is the distribution of distances from the carbon atom at the head of the molecule to the same atom on another molecule. The data agrees strongly with experimental peaks, shown as dotted lines.[42] Experimental data shows the peaks of the X-ray diffraction taken from the total RDF function. On the right is a similar distribution of oxygen to oxygen distances, showing that the model perhaps slightly overestimates the strength of hydrogen bonding interactions.

Bond	Bond Length MD ($\AA$)	Exp ($\AA$)	Dependence ($\AA kbar^{-1}$)
O-O	2.8 ± 0.1	2.9	0.002 ± 0.001
	5.0 ± 0.3	~ 5.0	0.002 ± 0.001
C_2-O	3.5 ± 0.1	3.7	0.002 ± 0.001
C_2-C_2	4.4 ± 0.1	4.3	0.020 ± 0.005
C_1-C_1	4.1 ± 0.1	4.3	0.020 ± 0.003
	4.9 ± 0.2	~ 5.0	0.020 ± 0.003
C_1-C_2	4.7 ± 0.3	5.0	0.08 ± 0.01
O-H_O	1.8 ± 0.1		0.0035 ± 0.0008
	3.5 ± 0.2		
H_O-H_O	2.5 ± 0.1		0.003 ± 0.001
	5.0 ± 0.1		

Table 3.2: RDF data of the various bonds of ethanol, and the shift in the RDF peak as a function of pressure. Experimental data taken from the work of Jorgenson et al.[42]

3.2 Bond Length Analysis

The SPM states that the solvation pressure of a pure liquid is equal to its CED. Therefore liquid ethanol feels a solvation pressure of $6.76 kbar$.

Consider the total pressure applied to ethanol at NTP. The total pressure is defined as the sum of the external pressure, known as the hydrostatic pressure, and the solvation pressure. Standard pressure is approximately equal to $1 bar$. Under normal circumstances, a liquid will not experience negative pressures or positive strains. Therefore, the minimum total pressure that can be experienced by liquid ethanol is $6.76 kbar$.

By definition, a vapour has a CED of zero and therefore would not feel a solvation pressure. Consequently, at NTP, the total pressure felt by ethanol vapour will be zero. So, if the SPM is correct, the total pressure felt by ethanol vapour will be exactly $6.76 kbar$ lower than that of liquid ethanol at standard pressure.

It is possible to test for this change in the total pressure experienced by the ethanol liquid and ethanol vapour. Two main tests were used to look for this change in total

pressure, and hence for evidence of the solvation pressure. We looked for changes in bond length and vibrational frequency as a function of solvation pressure.

3.2.1 Method

We placed liquid ethanol under positive hydrostatic pressure. We calculated the length of the bonds as a function of increasing pressure. In general, the bond lengths of a liquid can be used as a pressure gauge, since most bonds shrink in reaction to increasing pressure. The trend of the data should show that bond length is inversely proportional to hydrostatic pressure.

We also measured the bond lengths of ethanol vapour. Assuming that a solvation pressure acts on ethanol liquid, their bond lengths should be correspondingly shortened. Using knowledge of the expected size of the solvation pressure, we can plot the vapour data on the same axes.

We plotted the bond lengths of liquid ethanol and ethanol vapour on a scale representing the total pressure applied to the molecules, including solvation pressure. On such a scale, ethanol vapour would appear at zero pressure. Liquid ethanol would appear at pressures equal to or greater than its solvation pressure. However, it is convenient that the zero of the pressure axis should refer to the liquid at NTP. We therefore adopt a convention consistent with previous papers.[9,10]

We use excess pressure P_e, to denote total pressure relative to ethanol at NTP, such that:

$$P_t \text{ (Total Pressure)} = P_h \text{ (Hydrostatic Pressure)} + P_s \text{ (Solvation Pressure)} \quad (3.3)$$

$$P_e = P_t - P_s \text{ (ethanol)} \quad (3.4)$$

$$P_e = P_h + P_s \text{ (liquid)} - P_s \text{ (ethanol)} \quad (3.5)$$

Where P_s (ethanol) is the solvation pressure felt by pure ethanol liquid at NTP. Hence for pure ethanol liquid:

$P_e = P_h$

and for ethanol vapour, which does not feel the solvation pressure:

$P_e = -6.76kbar.$

When plotted in this manner, to reflect the lack of solvation pressure relative to pure liquid ethanol, the vapour data should follow the trends of the liquid data. Extrapolating the best fit straight line through the ethanol liquid to a pressure of $-6.76kbar$, should reveal a bond length equal to that of ethanol vapour. If this is the case, it will represent good initial evidence for the SPM.

3.2.2 Results

Figure 3.3a shows the C-H_3 C-H_2 and C-C bond lengths of ethanol as a function of pressure. Error bars on the vapour data represent standard error ($S_e = \frac{\sigma}{\sqrt{N}}$) of 1000 separate vapour data runs. These vapour runs were performed and recorded automatically, using a Linux BASH script to automate the generation of a starting configuration, running of the simulation, and collection of the relevant data. The errors in the liquid data are approximately equal to the size of the symbol in both directions. The straight line is a guide to the eye, representing a linear regression analysis of the liquid data only.

Figure 3.3a clearly shows that for these bonds, ethanol vapour feels a total pressure $6.76kbar$ lower than pure ethanol at NTP. The best fit straight line passes through the error bars of the vapour data in each case, offering excellent support for the SPM. The total pressure felt by these bonds at NTP is equal to the CED of the liquid. Thus, the SPM describes the bond shortening most effectively.

Figure 3.3b shows the length of the C-O and O-H_O bonds, again as a function of pressure. Neither bond appears to give a fit to the SPM. This means that either the solvation pressure does not affect these bonds, or that the solvation pressure effect is hidden by masking interactions.

Consider first the O-H_O interaction. The equilibrium bond length, as described in Table 2.4, is $0.96\mathring{A}$. As expected, this is the position of the vapour data point. On transition to the liquid phase, we see a marked increase in the length of the bond to $\sim 0.99\mathring{A}$.

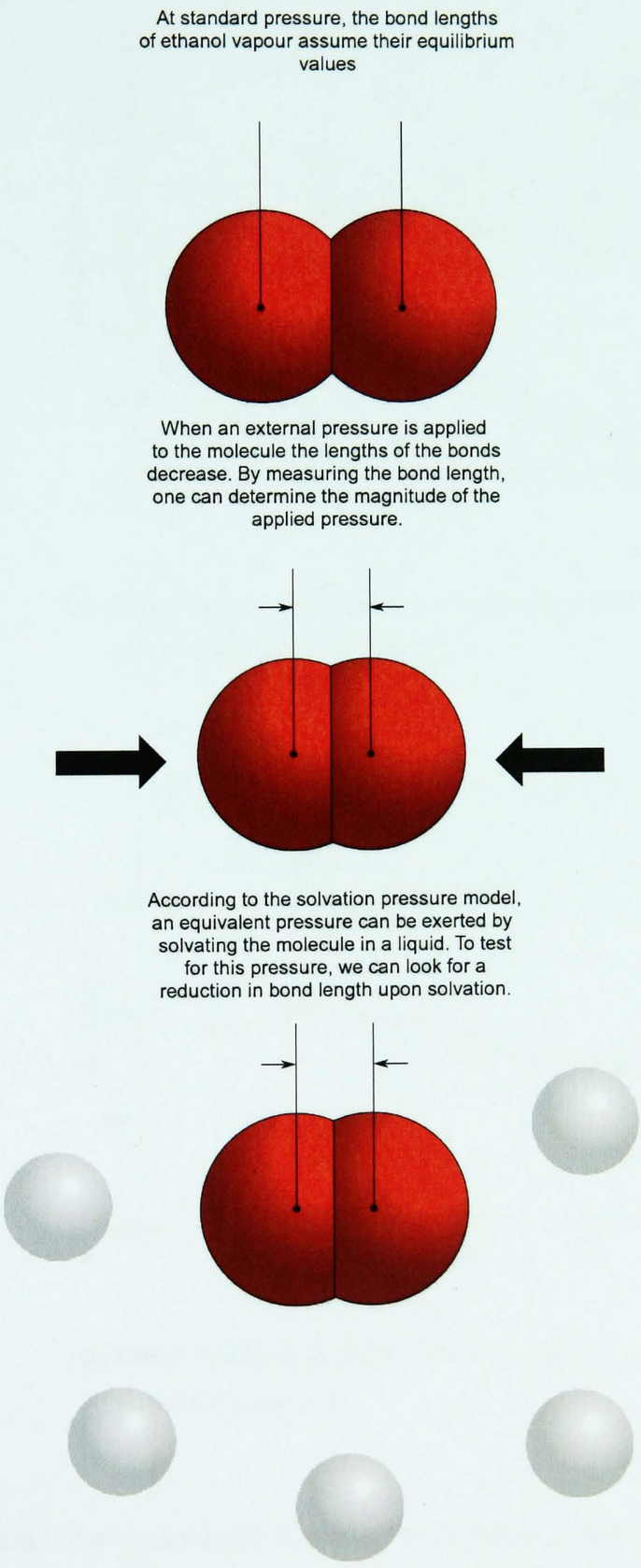
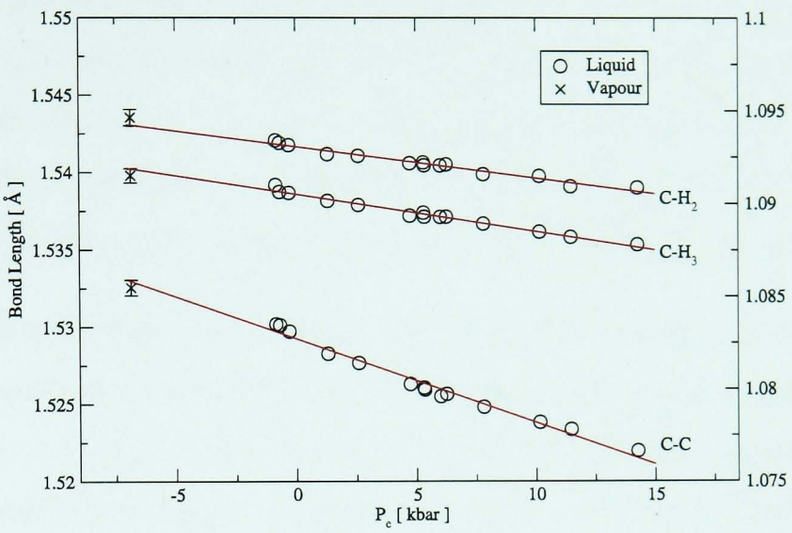
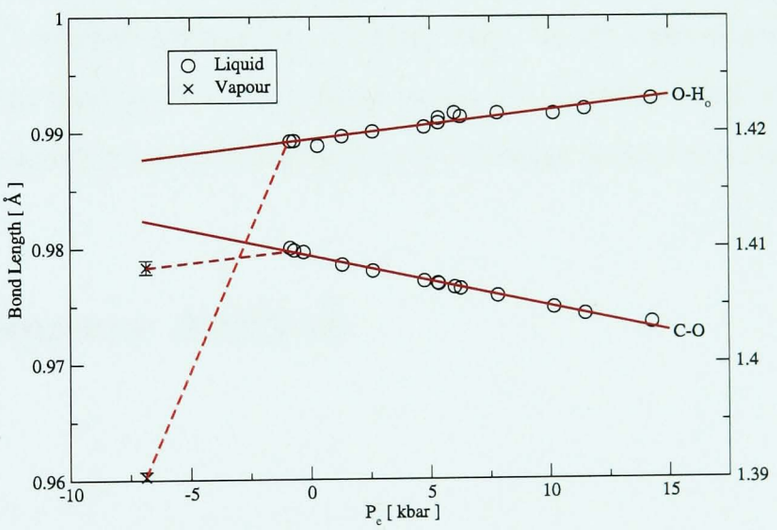


Figure 3.2: Using bond lengths to test for solvation pressure in pure liquids.



(a) Three bonds show behaviour consistent with the SPM: C-C (left axis), C-H_2, and C-H_3 (right axis).



(b) Other bonds show more complex behaviour: O-H (left axis), and C-O (right axis)

Figure 3.3: The bond lengths of ethanol calculated as a function of pressure, with the vapour data offset by the value of the CED, as predicted by the SPM. Three bonds which show strong agreement to the model are shown in Graph 3.3a, while those showing more complex behaviour are shown in 3.3b. The straight lines show linear regression of the liquid data as a guide to the eye.

The length of the bond is significantly lengthened by the OH...O and H...OH hydrogen bonding interactions. Indeed, the strength of this interaction is such that in this case, the bond length increases with increasing hydrostatic pressure. This phenomenon is well known, having been explored for a variety of hydrogen bonding liquids,[43,44] and observed in the initial SPM studies of ethanol.[9] Therefore, this bond will not be used to to test the SPM, since the bond does not contract as a function of hydrostatic pressure.

The case of the C-O bond is more subtle. Again, within the defined error, the vapour data point marks the equilibrium length of the bond. In this case, in the liquid phase, the bond does appear to provide a standard indicator of pressure. Table 2.4 again shows that this bond is stretched by hydrogen bonding interactions. The level of stretching is greatly reduced in comparison with the O-H_O bond. However, it is still significant enough to mask the effects of the solvation pressure.

It is clear that the predictions of the solvation pressure cannot be applied in the case of the O-H_O and C-O bonds. The ubiquitous effects of the solvation pressure are masked by the strong hydrogen bonding interaction. It seems that, in the case of predicting shifts in bond length from the liquid to the vapour phase, the model is most suited to non-polar bonds. These are the bonds for which fewest masking interactions exist.

3.3 Raman Frequency Analysis

3.3.1 Method

The second test of the SPM is to consider the Raman vibrational frequencies of ethanol as a function of pressure. These should be related to the bond lengths of the atoms, and hence should be expected to follow the patterns seen in the bond length analysis. However, measuring the Raman frequencies has the added advantage of allowing direct comparison with experimental data. This will be important in later chapters in the study of mixtures of water and ethanol, since an explanation for the unexpected behaviour of the trends in Raman data in this case was a motivating factor in performing these simulations.

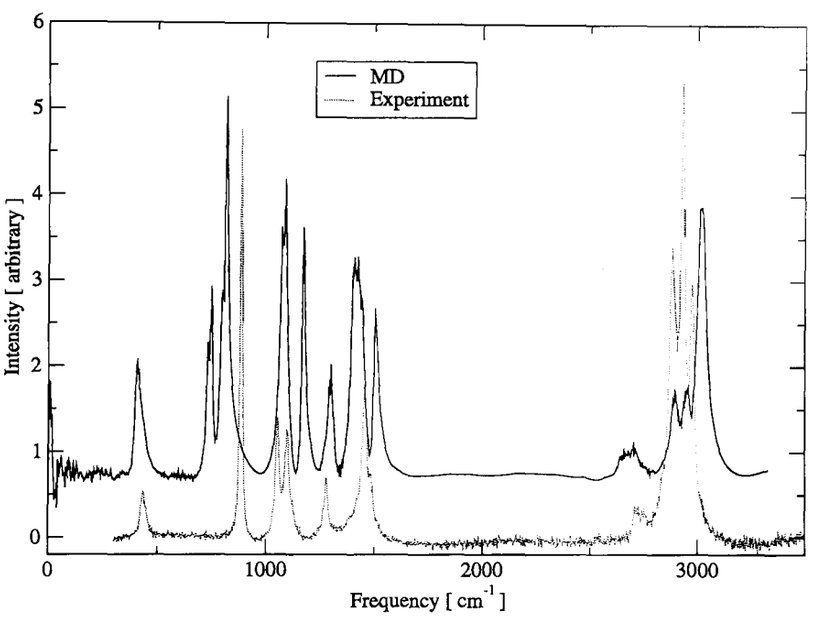


Figure 3.4: The vibrational density of states spectrum of ethanol, obtained theoretically by the Fourier transform of the autocorrelation function, is compared with an experimental Raman spectrum.[9] The MD data is offset vertically by 0.75 units for clarity.

The vibrational density of states spectra were obtained using the method described in Section 1.5. Where applicable, determination of the position of the peak and other data related to the frequencies and frequency range of a peak was determined using the peak fitting software fityk.[45] The program allows automated and manual addition of Lorentzian peaks to the Raman data, and optimisation of the fit using genetic algorithms. Peak frequencies, frequency ranges and the associated errors can the be established, and output to text files for further analysis.

3.3.2 Results

Figure 3.4 compares the Raman frequency of ethanol at NTP we obtained from our MD simulations with the experimental results of Dunstan et al.[9] The MD simulation results are offset vertically for clarity. The broad features of the experimental Raman spectrum are reproduced most satisfactorily. The peaks in intensity are accurate to within around $\sim 100cm^{-1}$ of the corresponding experimental peak. This again displays the quality of the ethanol model, and means that we can provide meaningful comparison with experimental results. .

Some vibrational modes of ethanol as a function of pressure are shown in Figure 3.5. The graph is set up in the same way as the bond length analysis, with the Raman frequencies of ethanol calculated as a function of pressure P_e, as described in Equation 3.3.

The vapour data again represents the average of 1000 separate vapour simulations. Despite the large number of vapour simulations performed the resulting vibrational spectra is slightly noisy due to the fact that there is only one bond to analyse per time-step. The straight line is a guide to the eye, again representing linear regression of the liquid data only. The errors on the vapour data are larger than for the bond length analysis due to the nature of the vapour vibrational spectrum, which results in slightly noisy peaks, and therefore an inherent error in fitting a set of Lorentzian peaks. However, these errors are not large enough to distract from the general trends in data, providing an uncertainty of around $\pm 3cm^{-1}$. In order to assist with the identification of modes, it was possible to compute the autocorrelation function using only atoms of

Label	Raman Mode	Peak Position	
		MD (cm^{-1}	Exp (cm^{-1})
CCOs	CCO Symmetric Stretch	815	883
CCOa	CCO Antisymmetric stretch	1082	1051
MO_2	CO Stretch + CH_3 Rock + in plane COH Deformation	1173	1096
MO_1	CH_2 Twist + in plane COH Deformation	1296	1276
CH_3d	CH_3 Deformation	1420	1453
CH_2	CH_2 Symmetric Stretch	2896	2881
CH_3s	CH_3 Stretch	2948	2928

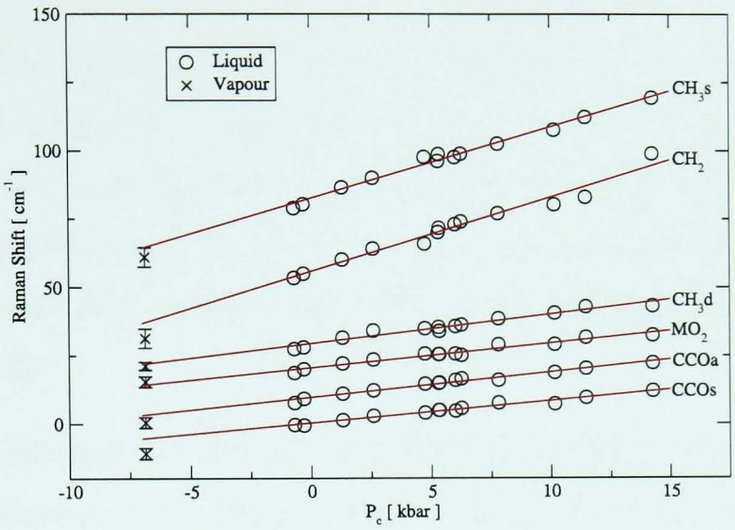
Table 3.3: Major Raman modes of ethanol compared with experimental results of van Uden[9]. We adopt the same naming conventions used in their work.

interest, for example to place the CCO modes, one can perform the Fourier transform using only those atoms.

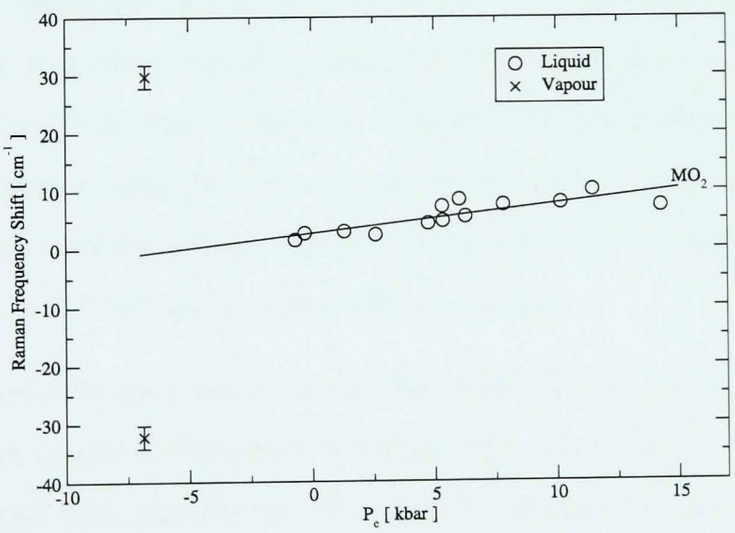
Six of the seven main frequency modes studied seem to provide especially strong support for the SPM. These are the CH_3 deformation, CH_3 Rock + CO stretch + in plane COH deformation, CCO symmetric and antisymmetric stretch, CH_2 Symmetric Stretch and CH_3 Stretch. They are shown in Figure 3.5a. The extent to which this data fits the SPM is quite striking, and provides extremely strong evidence for the model.

The CH_2 twist and in plane COH deformation mode, shown in Figure 3.5b, is a slightly different case. The peak appears to split in the vapour phase, producing two distinct peaks in the vapour Raman spectra where only one appeared in the liquid phase. The average of these two peaks gives a point which fits the SPM. Clearly a complex interaction occurs causing the two modes to combine in the liquid phase.

It would appear that the COH deformation modes occur at a higher frequency in the vapour phase than in the liquid. This is consistent with the bond length data, which shows the O-H_O bond to be significantly stretched by hydrogen bonding in the liquid phase.



(a) Most Raman modes show good agreement with the SPM. In this figure, individual modes are offset vertically by the following amounts: CCOs, 0; CCOa, 10; MO_2, 20; CH3d, 30; CH_2, 50; CH3s, 80 cm^{-1}.



(b) Others show more complex behaviour

Figure 3.5: Raman peak frequency shifts of ethanol as a function of pressure. The vertical axis represents the change in frequency of a mode compared with its value at NTP. Data in good agreement with the SPM is shown in 3.5a, whilst more complex behaviour is shown in 3.5b. The straight lines show linear regression of the liquid data as a guide to the eye.

3.4 Discussion of Pure Ethanol Results

It is clear that the modified Cornell field offers a very good approximation of the relevant properties of ethanol. Raman spectra are reproduced extremely well at a variety of pressures, without compromising important bulk properties, such as density and compressibility. In addition to this, the model enables very good reproduction of the relevant RDF peaks, showing that the structure of the liquid ethanol is reproduced effectively.

In this simple case, with an accurate ethanol model, strong evidence is seen for the SPM. However, the effect appears to be weaker than, or masked by, the effects of hydrogen bonding.

The non-polar bonds of the system appear to react to the solvation pressure of ethanol liquid. The extent to which the bonds are compressed on condensing of the molecule from the vapour to liquid phase is very well predicted by the model. The reasons for the success of the SPM in this case will be explored further in Chapter 7, as further analysis techniques are introduced in order to understand the results of more complex systems. It is likely that this result can be generalised in order to predict the results for similar bonds in other molecules. This includes hydrocarbons, and should also include most bonds in hydrophobic amino acids in proteins, such as Valine and Phenylalanine. Although tests should be performed to determine the response of untested bonds, such as the C=O bond and bonds involving nitrogen.

The highly charged polar tail of the molecule does not follow the predictions of the solvation pressure. Hydrogen bonding interactions stretch these bonds to lengths often far greater than equilibrium. It should be possible to predict whether bonds are likely to be affected by such interactions. It may also prove possible to estimate the change in bond length for these bonds, although further analysis would be needed on a far wider variety of molecules.

In Chapter 4, ethanol is mixed with water, and hence hydrogen bonding will be a large factor. The success of the project in this case requires that the hydrogen bonding in water does not mask the appearance of the solvation pressure in the non polar region

of ethanol. It is unlikely that the solvation pressure will be detected in bond involving the alcohol group of ethanol, since hydrogen bonding is also likely to prove a greater factor in this case. The aim is to explain the unusual shifts in Raman spectroscopic data seen by Dunstan and coworkers during studies of these mixtures.[9]

The established techniques of obtaining spectroscopic and bond data as a function of pressure will be applied to this new situation. The introduction of water will also allow the possibility of obtaining ethanol-water mixtures of widely differing CED, which will provide a strong test of the SPM. Several established methods, as well as some innovative techniques will be utilised in order to explain the results.

Chapter 4

Ethanol-Water Mixtures

4.1 Introduction

The SPM predicts that the pressure applied to a solute by its solvent is equal to the CED of the mixture. In order to test the extent to which high pressure solvation effects act on solutes, we performed simulations on mixtures of ethanol and water. Due to its highly polar nature, water has a very large CED of $\sim 22.5kbar$.[46] Ethanol has a CED of $6.76kbar$. In mixtures we assume the CED varies linearly with concentration, as shown in Figure 4.1. This is consistent with the behaviour documented by John Burke.[47]

Therefore, if the SPM is correct, by varying the concentration, an ethanol molecule in a mixture, can experience solvation pressures in the range $6.76 - 22.5kbar$. Using equation 3.3, this is equivalent to excess pressures, P_e, of $0 - 17kbar$. This is similar to the range of hydrostatic pressures applied to pure ethanol in Chapter 3. Identical hydrostatic and solvation pressures should cause identical shifts in bond length. Thus, if the SPM is correct, we can reproduce the pressure responses of ethanol shown in Figures 3.3a and 3.5 by varying the concentration.

We measured bond length and vibrational frequency as a function of pressure P_e, for varying mixtures of ethanol and water at NTP. In this case, the pressure is caused entirely by the solvation pressure of the mixture. We compared the results to those of

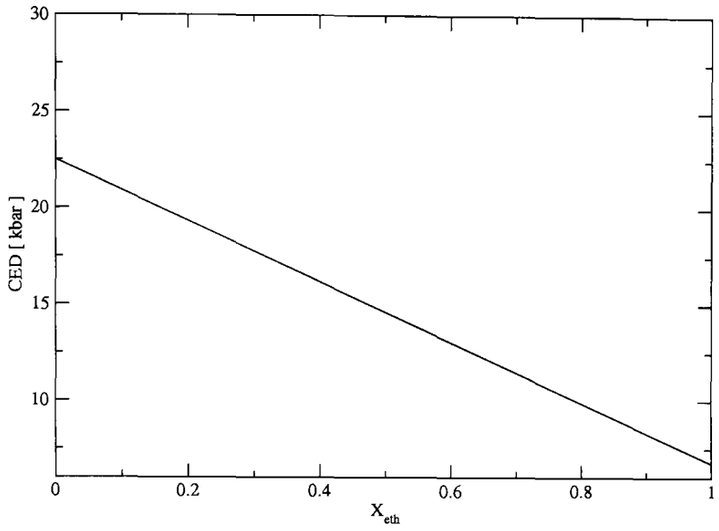


Figure 4.1: The CED of ethanol-water mixtures as a function of solution concentration.

ethanol under hydrostatic pressure, as shown in Chapter 3.

The structure of liquid water is shown in Figure 4.2.

We simulated nine different concentrations of ethanol and water, and calculated bond lengths and Raman frequencies. The CED of a mixture was calculated using linear interpolation:

$$CED_{mix} = \frac{\Delta U_{mix}}{V_{mix}} \tag{4.1}$$

$$= \frac{V_1 CED_1 + V_2 CED_2}{V_1 + V_2}. \tag{4.2}$$

where CED_{mix} is the CED of the mixture, V is the volume of the component and ΔU_{vap} is the energy change on vapourisation. CED_{mix} is equivalent to the solvation pressure of the mixture:

$$P_s \equiv CED_{mix}. \tag{4.3}$$

This can be used with Equation 3.3 to determine the excess pressure applied by water. Water has a high CED, and hence a high solvation pressure This means it will act to

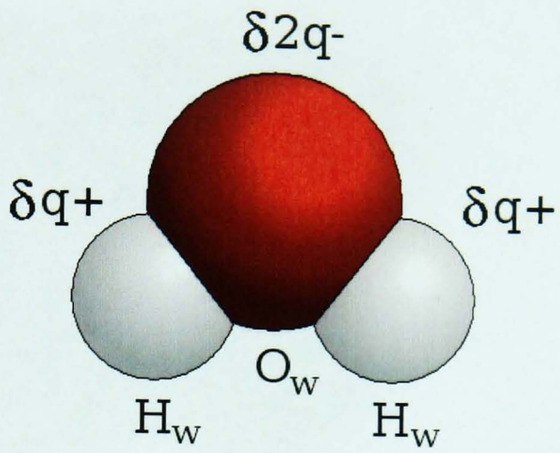


Figure 4.2: The structure of liquid water.

increase the pressure felt by ethanol. The excess pressure, P_e, applied to the ethanol will vary as a function of concentration in the range 0 - $17kbar$.

We decided initially to keep the amount of water in the simulations constant, in order to maintain a high solvation pressure. Therefore, 300 water molecules were used throughout our simulations, and ethanol added 10 molecules at a time, up to a maximum of 80 molecules. Figure 4.3 shows a mixture of 300 water molecules and 40 ethanol molecules.

4.2 Choosing a Model for Water

A recent review of the progress towards the successful simulation of liquid water noted 46 distinct models.[48] This statistic hints at the complexity of modelling liquid water, and the challenge faced when choosing an appropriate model.

In order to break down the search, it is possible to place the water models in two categories. The first category contains simple models. These have three sites representing the three atoms of water and the related charges. This category consists of four main models, the Simple Point Charge (SPC) Model,[49] and the TIP3P potential,[50] in flexible and rigid variants.

The second category contains massless charge (MC) models. MC models separate the

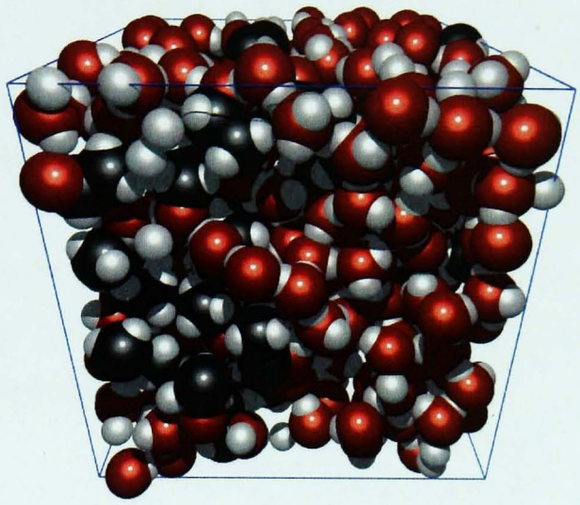


Figure 4.3: A mixture of 300 water molecules and 40 ethanol molecules following equilibration.

charge on the oxygen atom from its location in physical space. Three sites represent the atoms, and additional sites of zero mass represent the charge of the oxygen atom. The category can be divided into two sub-categories: those which have one additional site for the charge on the oxygen, and those which allocate two additional sites. Four point water models range from the simpler TIP4P model,[50] to more complex models, such as SWFLEX-AI.[51] Similarly, five point models range from the simple TIP5P model,[50] to models such as TTM2-F.[52]

DLPoly 2.14 does not include native support for massless charges required for MC models. Implementing a MC model would therefore be a significant undertaking. Since this is an initial study of the SPM in mixtures, it was decided that the use of MC models should be reserved for the case that simpler models prove unsuitable.

Therefore in the first instance, there were four main options. These were to choose between SPC and TIP3P water, and then to further choose between flexible or rigid models.

The two models are similar in both parameters and bulk properties. Several papers on the subject report similar properties of the two models.[53–56]

It was therefore decided that the rigid TIP3P model should be used, since the parameters of the Cornell field detail its use explicitly, suggesting that it is most compatible

Atom	Atomic Mass (amu)	Charge (e)
O_W	15.9994	-0.834
H_W	1.0080	0.417

Table 4.1: The atomic masses and charges of a water molecule, as defined by the TIP3P model.

with the ethanol model derived in Section 2.4. Also, there is again evidence from Cheatham et al[27] to suggest that the TIP3P model works successfully with a system of ethanol molecules.

4.3 The TIP3P Model of Water

Simulations were first performed on pure water, to obtain some bulk properties of the molecule in order to test its quality and suitability. The model used was based on the TIP3P potential,[50] which consists of three atoms of fixed charge. Atomic masses and charges on the molecule are shown in Table 4.1. The molecule is held rigid, with constant bond length and angles. The O_W-H_W bond length is 0.9572$\AA$, and the angle $H_W\hat{O}_WH_W$ is 104.52 deg.

The van der Waals parameters for water-water and water-ethanol interactions are shown in Table 4.2. The values are taken from the work of Cornell et al.[31]

4.4 Testing and Refining the TIP3P Model

We used the same tests as for liquid ethanol; we calculated bulk properties of water, and compared RDF data with experiment. It is of course impossible to produce informative Raman data, due to the rigid nature of the model.

The water produced very good results for density and CED. However, there were concerns over the unusually low compressibility of the molecules. The likely cause of these significant compressibility errors was the inflexible nature of the water molecule, as well as the slight overestimation of the density and CED. Therefore, the model was

Lennard Jones Pair	A ($kcal\ \AA^{12}\ mol^{-1}$)	B ($kcal\ \AA^{6}\ mol^{-1}$)
O_W-H_W	5.82E+05	5.95E+02
H_W-H_W	0.00E+00	0.00E+00
H_W-O_W	0.00E+00	0.00E+00
O_W-C	7.86E+05	6.37E+02
O_W-H_3	6.92E+04	1.16E+02
O_W-H_2	4.76E+04	9.64E+01
O_W-O	5.83E+05	6.46E+02
O_W-H_O	0.00E+00	0.00E+00
H_W-C	0.00E+00	0.00E+00
H_W-H_3	0.00E+00	0.00E+00
H_W-H_2	0.00E+00	0.00E+00
H_W-O	0.00E+00	0.00E+00
H_W-H_O	0.00E+00	0.00E+00

Table 4.2: The van der Waals parameters defining the interactions between atoms in different water and ethanol molecules. H_3 refers to the three hydrogen atoms attached to the carbon atom at the polar head of the ethanol molecule. H_2 refers to the two hydrogen molecules attached to the central carbon atom, and H_O refers to the hydrogen atom at the polar tail of ethanol.

Property	Model Value	Experimental
Density (gcm^{-3})	0.98 ± 0.01	1.00[40]
Cohesive Energy Density ($kbar$)	24.2 ± 0.02	22.4[46]
Compressibility (GPa^{-1})	0.27 ± 0.9	0.46[57]

Table 4.3: The properties of rigid TIP3P water, determined to assess the quality of the model, and its suitability for these simulations. The model produces an excellent density and CED, but the compressibility is low.

Bond type	k ($kcal\ mol^{-1}\ \AA^{-2}$)	r_0 ($\AA$)
O_W-H_W	600.0	0.957

Table 4.4: Potential used with equation 2.11 to define chemical bonds in water.

Angle Type	k ($kcal\ mol^{-1}$ radian^{-2})	θ_0 (deg)
$H_W\hat{O}_WH_W$	100.0	104.5

Table 4.5: Potentials used with equation 2.12 in defining bond angles

revised to include bond vibration and bending parameters, as provided by the Cornell Potentials.[31]

The O_W-H_W bond was allowed to vibrate with simple harmonic motion, using Equation 2.11 and the parameters shown in Table 4.4. We also allowed simple harmonic motion of the angle of the water molecule, using Equation 2.12 with parameters shown in Table 4.5.

The details of the bulk properties of flexible TIP3P water are shown in Table 4.6. Despite changes to the model, the compressibility of the water is underestimated by 25%. However, it is significantly improved compared with the rigid TIP3P model.

The density of the water corresponds very well with the accepted value. This is true of most water molecules, which tend to reproduce density data accurately at NTP.

The CED is slightly overestimated in the simulations, 9% higher than expected. This is possibly further evidence of slight overestimation of the hydrogen bonding potentials in the model. However, this value compares very favourably with other water models. For example, the SPC model, which was also assessed for suitability, significantly underestimated the CED. The calculated value in that case was only $12kbar$, 47% lower than expected, meaning it was unsuitable for this study.

It is preferable for this quantity to be overestimated, since it should result in an overestimation of the solvation pressure. This will make any evidence of the SPM more visible.

It was also necessary to test the quality of mixtures of ethanol and water. We simulated mixtures of ethanol and water with concentrations, X_{eth}, in the range 0.03 - 0.28. The

Property	Model Value	Experimental
Density gcm^{-3}	1.02 ± 0.01	1.00[40]
Cohesive Energy Density kbar	24.5 ± 0.02	22.4[46]
Compressibility GPa^{-1}	0.34 ± 0.09	0.46[57]

Table 4.6: The properties of water, determined to assess the quality of the TIP3P model, and its suitability for our simulations. Note the excellent density and CED of the water.

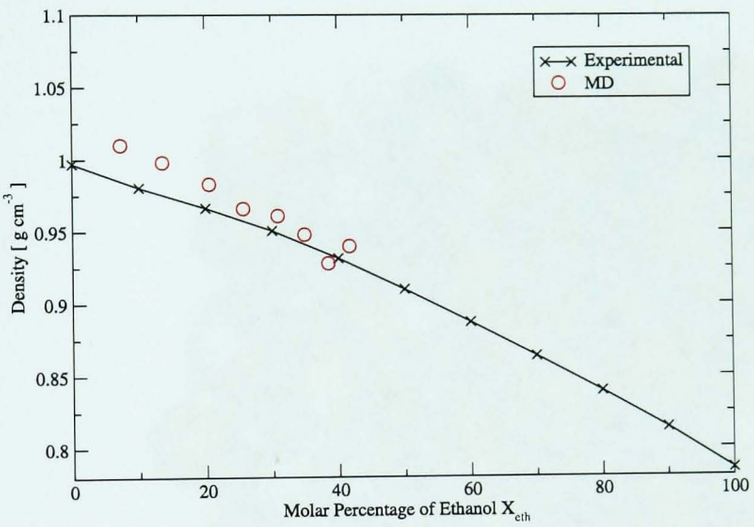


Figure 4.4: Densities of the various mixtures of ethanol and water compared with experimentally obtained values[40] as a test of the validity of the TIP3P model. The experimental values are just slightly lower in most cases.

change in density of the mixture as a function of concentration was measured and compared with experiment. This is shown in Figure 4.4. The density of the mixture remains slightly high as the concentration of ethanol increases, but followed the trend of the experimental density data[40] satisfactorily.

Figure 4.5 shows the structure of ethanol-water mixtures. The structure of the mixture shows that water forms hydrogen bonds with other water molecules as well as with ethanol molecules.

Overall, properties of water and ethanol-water mixtures are reproduced accurately. It was decided that the model is suitable for these simulations.

Figure 4.5: A mixture of ethanol and water. Hydrogen bonding can be seen between water molecules, and between ethanol and water.

The investigation into the SPM in ethanol-water mixtures follows in Chapter 5. Then in Chapters 6 - 9, we use several techniques to attempt to determine the underlying reasons behind these results.

Chapter 5

Solvation Pressure in Ethanol-Water Mixtures

To test SPM in ethanol-water mixtures, we performed bond length and vibrational frequency analyses, as in Chapter 3. The solvation pressure experienced by an ethanol molecule ought to be a linear interpolation of the solvation pressures of ethanol and water, according to equation 4.2.

Mixtures in the concentration range $X_{eth} = 0.03$ - 0.28 were chosen for this study. By focussing on water-rich systems, we study the largest solvation pressure effects possible with these liquids. Therefore solvation pressure effects should be easiest to detect. The simulations were performed at NTP in each case.

The process of achieving zero pressure in a given simulation was not straightforward. It is tempting to allow the simulation to settle at zero pressure during an NPT simulation. But equilibrating such a system with an NVT run, results in a lowering of the system pressure. This is because the system does not equilibrate fully during NPT runs. Following equilibration with a NVT simulation, the average pressure of such a system reduces to as little as $-1.5kbar$. Therefore, it is necessary to allow the NPT simulation to settle at a slightly higher pressure than is required, and attempt to predict the corresponding drop in pressure during the NVT run.

For these reasons, it was normal to repeat the combination of NPT and NVT simula-

tions several times in order to locate the point at which the system would equilibrate at zero pressure. As more simulations are performed, it became possible to predict the box size required for a simulation to equilibrate close to zero pressure. We also noted trends in size of the pressure drop during an NVT run, allowing the equilibration process to become more efficient over time. In addition to this, the pressure during a NVE run is seen to fluctuate naturally during the course of the simulation. Typically this fluctuation is around $\pm 0.2 kbar$.

These difficulties in precisely reaching zero pressure in a given simulation lead to an overall error in the pressure of a simulation of around $\pm 0.5 kbar$.

5.1 Radial Distribution Functions

RDF data for the hydrogen molecules at the head of the molecules bonding to the oxygen in water is shown in Figure 5.1 for various mix concentrations. The head of the ethanol molecule is not strongly attracted to water. The plateau shows that there is a relatively weak interaction between the head of the ethanol molecules and the water. There is no favoured configuration of water molecules approaching the head of an ethanol molecule, and hence a random distribution relative to the H_3 atoms in ethanol.

By contrast, the hydrogen bonding tail of ethanol is strongly attracted to water. This is of course expected, due to the large opposing charges on these atoms. This is shown in Figure 5.2 for high and low ethanol concentrations.

Water molecules are far less attracted to the C-H_3 group in ethanol than they are to other surrounding water molecules. This process is known as 'Hydrophobic Hydration'.[58] Consider a polar molecule, in this case water, interacting with a non-polar molecule, such as the head of ethanol. The polar molecule tends to be easily drawn away by strong interactions with other charged molecules. The greater interaction strength of the hydrogen bonding water-water interaction defines the lowest energy configuration of the water, and hence negates any equilibrium configuration which may otherwise exist between the water and ethanol.

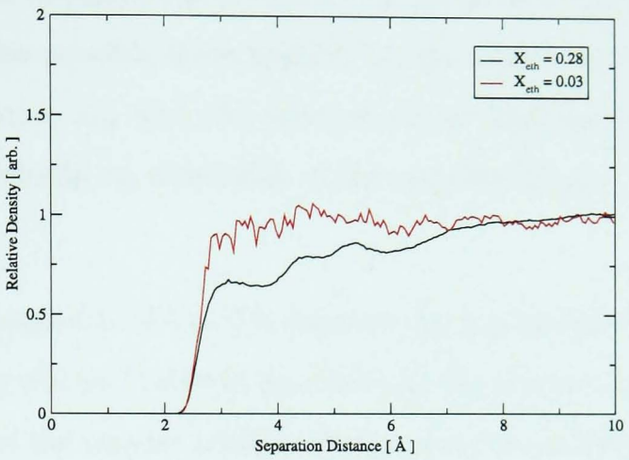


Figure 5.1: The RDFs showing the structure of oxygen atoms in water surrounding the hydrogen in CH_3. The plateau suggests that there is no energetically favourable configuration adopted. X_{eth} is a measure of the concentration of the mixture, representing the mole fraction of ethanol. The number of moles of ethanol is divided by the total number of moles in the system. By definition, the sum of the mole fractions X_{eth} and X_{wat} is zero.

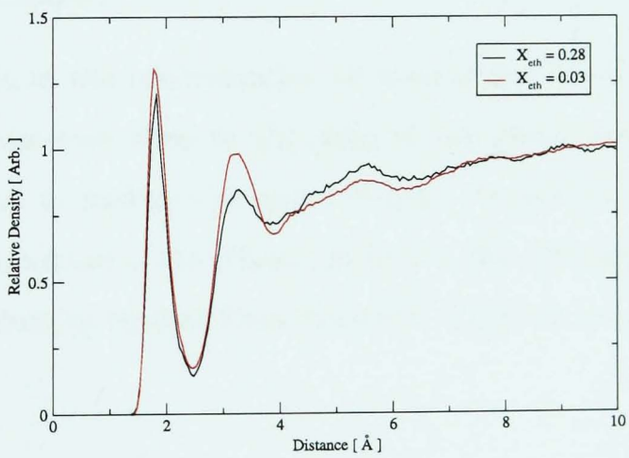


Figure 5.2: In contrast to Figure 5.1, the interaction between water and the tail of the molecule is very strong, showing a definite interaction similar to that of tail-tail interactions of pure ethanol. Shown is the oxygen atom in ethanol bonded to the hydrogen in water. Hydrogen bonding causes water molecules to be separated from ethanol at strongly defined equilibrium positions. The overall structure is similar to that of the hydrogen bonding in pure ethanol.

Given then that there is no equilibrium configuration of the head of the ethanol and the water, and that the water molecule has symmetrical charge and mass distributions, it is possible to speculate the possible force applied by the water. A water molecule has one negative charge centre, -2q, and two positive charge distributions of half the size, +q, positioned symmetrically on either side of the negative charge. This is shown in Figure 4.2.

Since the positioning with respect to ethanol is random, for a given probability p that the large negative charge -2q will be in closest proximity to the ethanol, there is a twice the likelihood, 2p, that one of the smaller positive charges q will be in closest proximity. Therefore on average there would be a total of 2pq of attractive force balanced by a total repulsion of 2pq, and thus zero net force.

These Coulomb interactions represent the bulk of the force applied by water. The Cornell potentials do not define a van der Waals potential to the H_W atom, and the van der Waals forces applied by the O_W atom are around 60 times smaller than the Coulomb interactions. This would explain the lack of response to solvation pressure. The attractive and repulsive interactive forces affecting the ethanol molecule will be discussed in more detail in Chapter 7.

It should also be noted that as the concentration of ethanol molecules increases, the chance of finding a water molecule close to the head of the ethanol molecule is significantly reduced, compared to finding one near the tail. As well as indicating the hydrophobic and hydrophilic nature of the ethanol molecule, this also suggests the possibility that the ethanol molecules tend to form 'clusters'. This will be discussed fully in Chapter 6.

5.2 Bond Length Analysis

In this section, we present bond lengths of ethanol as a function of the pressure applied by the mixture. We perform all simulations at NTP, and hence no hydrostatic pressure is applied to the mixtures. The results are compared to those of pure ethanol under hydrostatic pressure from Figure 3.3. Bond lengths of ethanol in ethanol-water mixtures

are plotted as a function of excess pressure P_e. The solvation pressure of the mixture, P_s, is calculated using Equation 4.3, and used with Equation 3.3 to determine a value for P_e. In all calculations, we used values for CEDs as calculated from the simulations. The data do not appear to support the SPM, showing mostly no frequency or bond length shifts as a function of increasing solvation pressure.

For example, for liquid water at NTP:

$$\begin{aligned} P_e &= P_h + P_s - P_s \text{ (ethanol)} \\ P_e &= 0 + 22.5 - 6.76 \\ P_e &\sim 17kbar \end{aligned}$$

Hence, at NTP, water can be used to increase the pressure applied to liquid ethanol by up to $17kbar$.

We will first consider the C-H_3 bond, shown in Figure 5.3. The bond length increases slightly as a function of increasing solvation pressure. This is the opposite trend to that predicted by the SPM.

It is possible that this is the result of the water molecules not approaching the head of the ethanol molecule closely. Hydrophobic hydration interactions mean that they are easily drawn away by other water molecules, with which it bonds more strongly. This effect is seen in the RDF data earlier in this chapter, and later in the analysis of structural density (Chapter 8.) The effect is also seen subjectively in molecular visualisations, Figure 5.4.

The C-H_2 bond (Figure 5.6b) is an interesting case, showing mostly no change in bond length with increasing solvation pressure. However, as the ethanol concentration reduces to very low levels, the bond length appears to decrease sharply. There is little chance of this being an anomalous point, since if this were the case it would almost certainly be present in at least some of the other bonds. In addition to this, any anomalies should be reduced or removed by the repetition of the simulations.

Also, by determining the nearest neighbours of a molecule, it is possible to calculate the bond lengths of isolated ethanol molecules in higher concentration mixtures. This

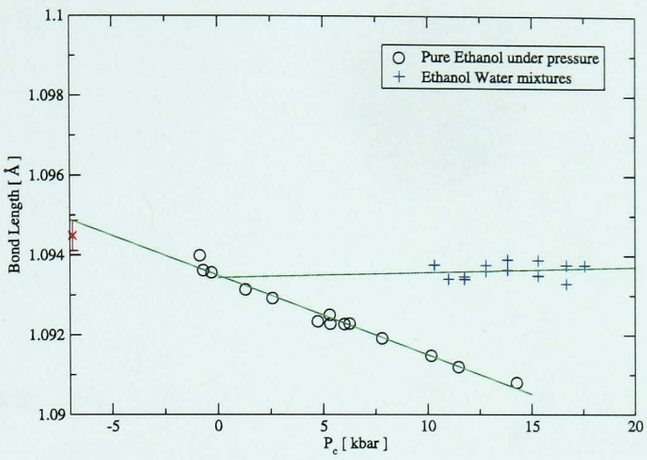


Figure 5.3: The response of the C-H_3 bond to increasing pressure P_e. In ethanol-water mixtures, this pressure is caused by the solvation pressure, P_s, of mixtures of different concentration as calculated using Equation 4.3. Each simulation is performed at NTP, and hence no hydrostatic pressure is applied to the system. The trend of the data is opposite to that predicted by the solvation pressure theorem, showing a slight increase in bond length as solvation pressure increases.

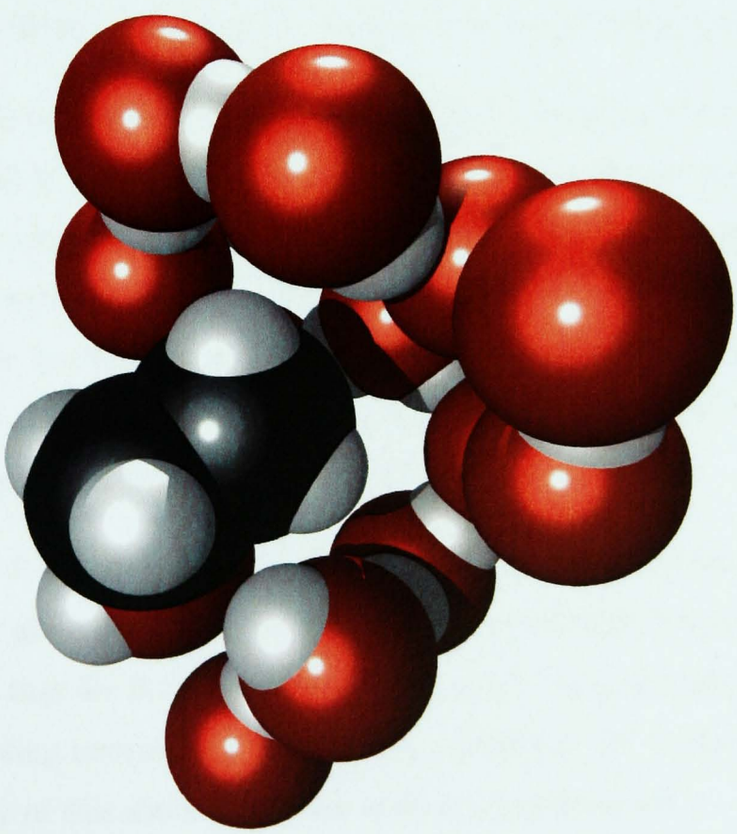


Figure 5.4: Water molecules do not closely approach the C-H_3 group of ethanol. Atoms are shown using standard CPK colours: carbon, black; hydrogen, white; oxygen, red.

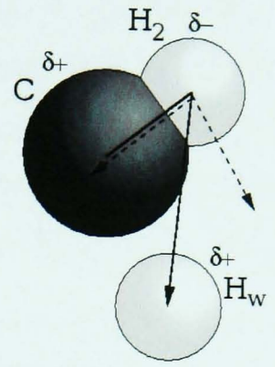


Figure 5.5: The shortening of the C-H_2 bond by hydrogen in water. This interaction is seen most frequently in a mixture of 300 water molecules and a single ethanol molecule. The shortening force can be caused by the hydrogen bonding interaction H_w ... O.

process is described fully in Section 6.2. Such isolated molecules also tend to show this shortening of the bond compared with groups of ethanol molecules.

It is important to understand why a system with only a single ethanol molecule would have this shortened bond length, since it is far closer to displaying behaviour consistent with the SPM than any other bond in this section of the study.

The C-H_2 bond would be most receptive to any change in the level of hydrogen bonding at the tail of the molecule, since it is in closest proximity. Indeed, the size of the water molecule means that it can potentially deform the C-H_2 bond directly, when hydrogen bonding with the tail of an ethanol molecule. The molecule is not large enough to affect any other bond in the same way. Thus it is possible to speculate that an increase in the level of hydrogen bonding between ethanol and water could be responsible for a reduction in the bond length of the C-H_2 bond.

Chapter 7 will show that for this data point, the lengthening forces of oxygen atoms in water are more evenly balanced by the shortening forces of hydrogen atoms. This suggests that for this system with just a single ethanol molecule, there are more hydrogen bonding interactions, since these interaction act to shorten the C-H_2 bond. The geometry of this shortening force is shown in Figure 5.5.

The level of hydrogen bonding decreases as a function of ethanol concentration. Even at low ethanol concentrations, ethanol molecules can begin to 'protect' themselves from

the hydrogen bonding interactions of water. It will be shown in Chapter 6, that even at low concentrations of ethanol, the molecules tend to seek each other out, preferring to have at least one ethanol molecule as a nearest neighbour. It will be further shown in Chapter 8 that neighbouring ethanol molecules in low concentration mixtures show no strong tendency to form hydrogen bonds, although some are present in a given mixture.

It is therefore the ability of a single ethanol molecule to be completely surrounded by water molecules that allows the shortening of the C-H_2 bond to occur. Due to Hydrophobic Hydration, the C-H_3 molecule is not affected in the same way, because water molecules do not approach the non-polar part of the molecule closely. As expected, the shortening relies on the proximity of the bond to the polar tail of the molecule.

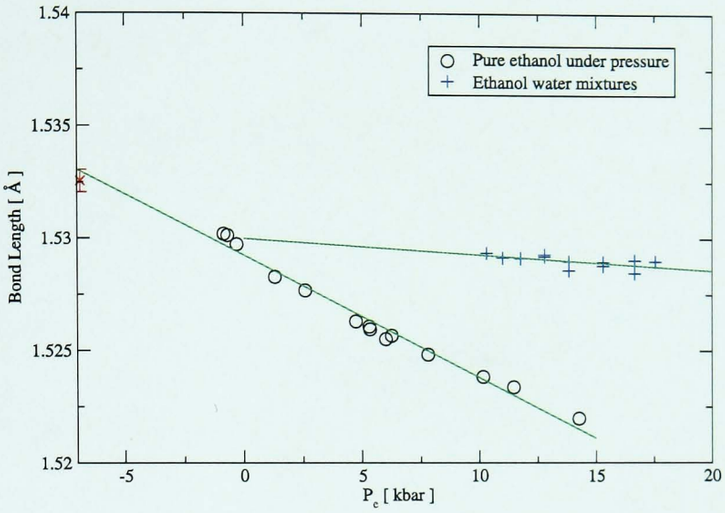
The C-C bond is a very similar case, as shown in Figure 5.6a. There is no change in bond length as a function of increasing solvation pressure.

The O-H_O bond (Figure 5.7) in mixtures shows a similar trend to that of the pure ethanol. However this was not a bond which followed the trends of the SPM, due to the masking interactions of the strong hydrogen bonds. Therefore, the fact that the two sets of data are in good agreement cannot be attributed to the effects of the solvation pressure. This result is most likely due to hydrogen bonds between ethanol molecules being simply replaced by similar hydrogen bonding interactions between ethanol and water.

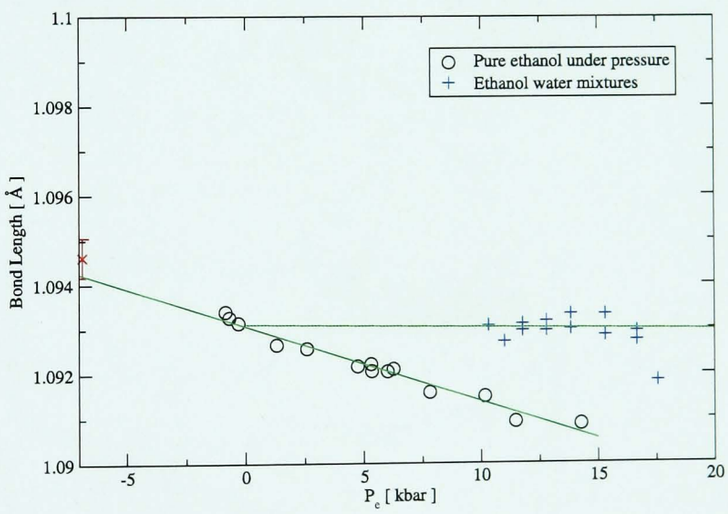
The C-O bond (Figure 5.8) shows a strong increase in bond length with increasing solvation pressure. This is because the C-O bond is lengthened by hydrogen bonding interactions. This is shown in Figure 5.9. The oxygen atom in ethanol is attracted by surrounding hydrogen atoms in the O-H_O group of other ethanol molecules or water.

The equilibrium positions of these interactions are such that the angle C-O...H is generally obtuse. This means that the interaction usually acts to pull the oxygen atom away from the carbon, causing a lengthening of the bond.

As the concentration of water increases, the total number of hydrogen bonding interactions per ethanol molecule increases, since water produces a large number of hydrogen bonding interactions. Therefore, as the concentration of water increases, the average bond length of the C-O bond increases.



(a) C-C



(b) C-H_2

Figure 5.6: If the SPM is correct in stating that the CED can be directly related to a real pressure applied to solutes, the bond lengths for ethanol mixtures should follow the same trends as the data for ethanol under hydrostatic pressure. Clearly, the results in (a) and (b) do not indicate a relationship between mixture concentration (and hence 'solvation pressure') and bond length.

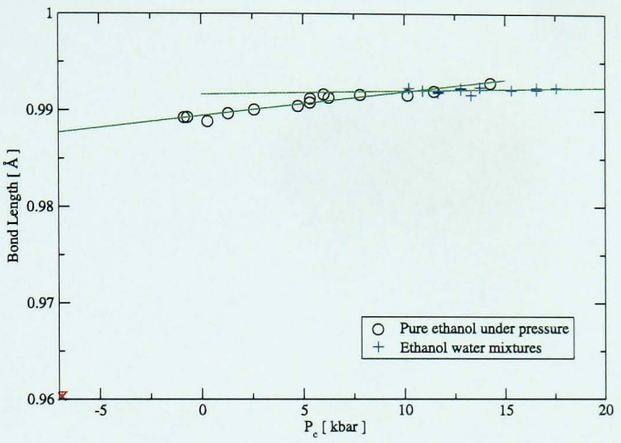


Figure 5.7: The O-H_O bond shows very similar trends to that of the pure ethanol simulation. Although this is a requirement of the SPM, it cannot apply in this case, as the dominant interaction of the bond is through hydrogen bonding. Replacing the hydrogen bonding of ethanol with that of water results in similar variations in bond length.

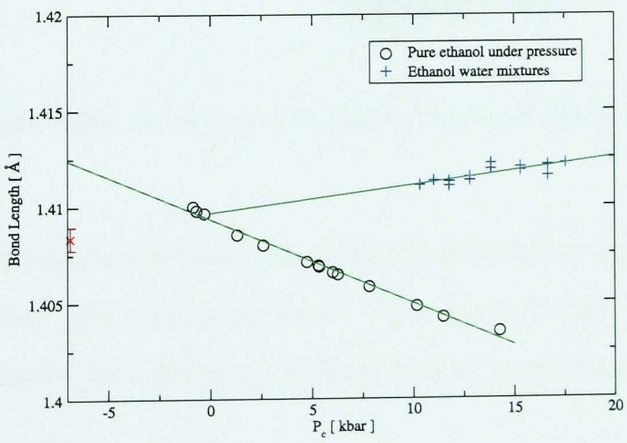


Figure 5.8: The C-O bond shows increasing bond length as a function of solvation pressure. Interestingly, the data can be extrapolated to include the vapour point from the ethanol vapour simulations.

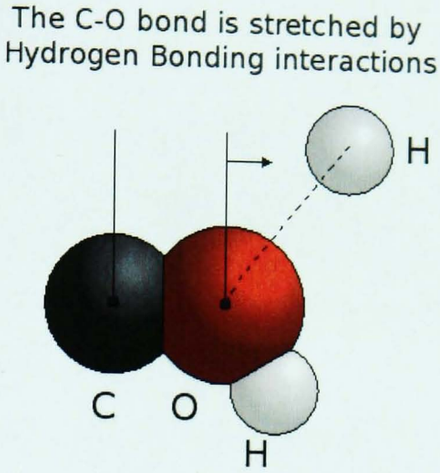


Figure 5.9: Demonstration of the increase in bond length of the C-O bond caused by hydrogen bonding. The positioning of the atoms represents a typical structural arrangement of the atoms, although the O...H distance has been exaggerated for clarity.

Figure 5.10: An example of the Lengthening of the C-O bond from the simulations.

In ethanol vapour the hydrogen bonding interaction are not present, and so the bond does not feel this stretching force. The bond length of the C-O bond in vapour is therefore lower than for pure ethanol at zero pressure, and all mixtures.

Interestingly, despite this not being a solvation pressure interaction, the ethanol-water bond length data can be extrapolated to include the ethanol vapour data point. It seems that the C-O bond may be an 'intermediate' case; applying a solvation pressure alters the length of the bond, but identical hydrostatic and solvation pressures do not cause equivalent shifts in bond length.

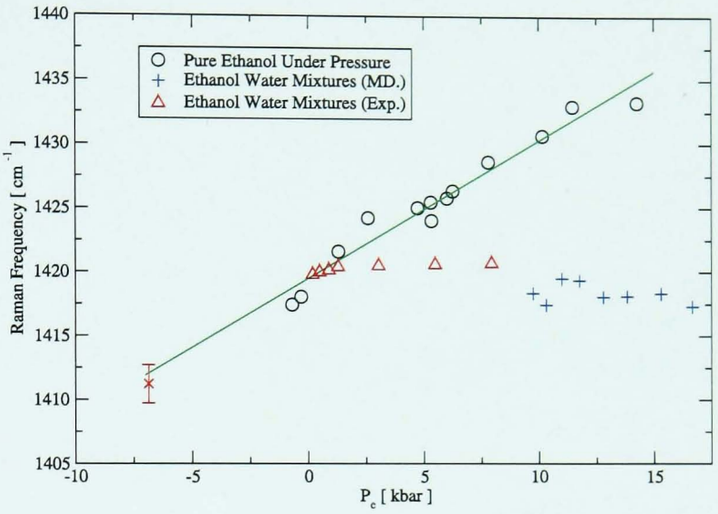
5.3 Raman Frequencies

The vibrational frequencies are analysed by the method described in Chapter 1. The results confirm the trends shown in the bond length analysis in the preceding section. In most cases the results show vibrational frequency independent of solvation pressure. This is consistent with the bond lengths shown in the preceding section, since a change in frequency is dependent on a change in bond length.

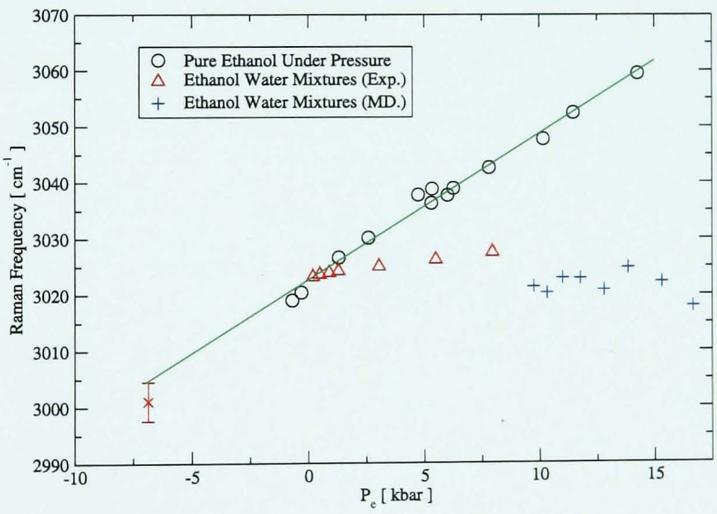
Although the vibrational frequencies shown in Figure 5.11 do not follow the trends of pure ethanol under pressure, they are in good agreement with the shifts for ethanol-water mixtures as obtained by experiment.[9] Frequency shifts caused by the solvation pressure effects of water are far smaller than predicted by the SPM. The agreement with experimental values is quite striking, and offers further validation of the quality of the model used in this work.

As shown, the offset of the MD results compared with the experimental data[9] is less than $10cm^{-1}$. Considering the simplifications introduced in modelling the system, this is an excellent result. Many bonds are modelled using the most simplistic functions possible, and as described in the Chapter 4, the water model is one of the simplest available. Considering these factors, we had anticipated far higher discrepancies in frequency orders of magnitude greater than was achieved.

Since the results are in such good agreement with experiment, there is a good opportunity to gain insight into both sets of results. In order to better understand these results, it is necessary to observe the interaction between the water and ethanol molecules. This is done quantitatively using mathematical methods such as calculating structural density functions, and visually using molecular visualisation software. Over the following chapters, several methods will be used in order to gain insight into the reasons for the lack of solvation pressure in ethanol water mixtures.



(a) CH_3 deformation



(b) CH_3 stretch

Figure 5.11: Vibrational frequency data as a function of excess pressure. We compare our MD results for ethanol-water mixtures with the experimental results of Dunstan et al,[9] and with our results for pure ethanol under hydrostatic pressure. Again there is no perceivable shift in frequency as the concentration of water increases. However, the data shows reasonable correlation with previous experimental studies,[9]

Chapter 6

Clustering of Ethanol in Water

6.1 Evidence of Clustering of Ethanol in Water

As shown in Chapter 5, ethanol-water mixtures do not follow the predictions of the SPM. The following chapters will describe the processes undertaken to explore some of the possible explanations. An obvious starting point for such an analysis is to test the extent of the mixing of the two molecules.

It is possible that ethanol and water are not mixing fully. This possibility was first introduced in Figure 5.1 in the analysis of RDF data. It was noted that the likelihood of finding a water molecule in close proximity to the head of the ethanol molecule significantly decreases as a function of increasing ethanol concentration.

This may simply be a natural result of the increasing ethanol concentration, displaying no greater reduction in proximity than a random mixture of the two liquids. However, it may indicate a tendency for ethanol molecules to form extended structures, known as clusters, in regions containing less water than would be expected in a random arrangement of ethanol and water.

In addition to these observations, this clustering, or 'bi-percolating' behaviour of alcohols has been observed in studies of methanol.[59,60] They studied the segregation of methanol and water using experimental and MD methods. Methanol was modelled using a three-site model, making use of a single methyl atom to represent the head of

the molecule. Methanol concentrations in the range 0.27 to 0.70 were studied. The group witnessed clustering of the methanol molecules. In each case, clusters involving a large percentage of the methanol molecules in the system were observed.

The tails of ethanol molecules are likely to be found at the boundary between ethanol and water. Due to the hydrophobic hydration principle, the heads of the ethanol molecules would therefore tend to clump together in a region protected from the water molecules.

In the same way as oil on the surface of water, it is possible that to an extent, the non-polar heads of the ethanol molecules clump together because they are not attracted to water. This is a similar situation to the folding of proteins. Proteins are made up of chains of amino acids. These amino acids fall into two categories: hydrophobic and hydrophilic. When a protein folds, the hydrophobic portions of the chain are drawn to the centre of the protein structure and surrounded by their hydrophilic counterparts, since these are attracted to water.

In the studies of methanol it was found that although water and methanol 'mix' in the traditional sense of each having an equal chance of occupying a given space, they do in fact remain largely separate.[59] These separations are not total - it is possible, for example, for a cluster to contain a few water molecules.

If the ethanol molecules are indeed forming large clusters, it might provide an explanation for the observed independence of bond length and Raman frequencies as a function of CED of the mixture, since a significant number of ethanol molecules may be largely shielded from the effects of the surrounding water.

6.2 Bond Length as a Function of Cluster Size

6.2.1 Method

In order to test the level of clustering in the mixtures, we used the criterion of Dougan et al.[59] Two molecules are considered to be 'clustered' if the separation of oxygen atoms in ethanol falls within the first peak of the RDF of pure ethanol. Such molecules would

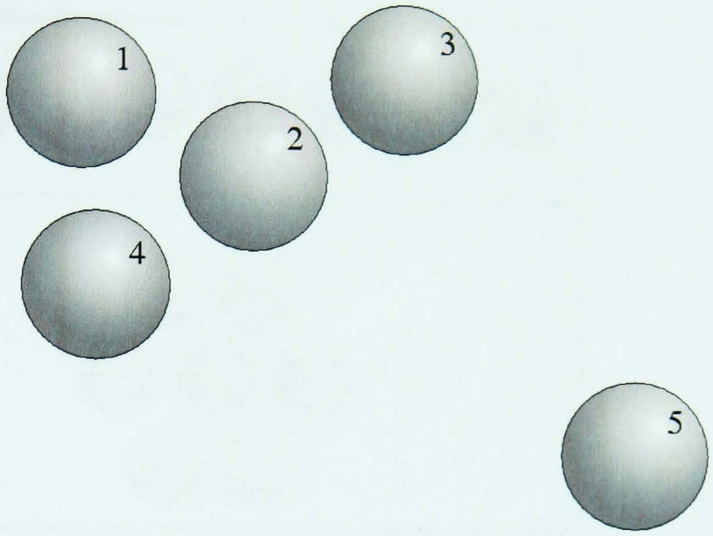


Figure 6.1: An example arrangement of 5 ethanol molecules, shown as grey spheres. Molecules 1-4 are in a cluster, whereas molecule 5 is separate.

therefore be a hydrogen bonded pair, and therefore neighbours. Using simple logic, it is then possible to build up a chain of all the ethanol molecules that are connected and in turn possible to calculate the size of ethanol clusters, and the number of individual clusters in the molecule as a function of time.

Using the RDF of ethanol, the cut-off radius for two molecules to be defined as neighbours was set at $r < 2.4\text{Å}$. Calculating the distance between each pair of ethanol molecules allows a list of clustered pairs to be generated.

For example, take simple example case of a system of 5 molecules. An example arrangement of the molecules is shown as grey spheres in Figure 6.1. The molecules 1-4 are clustered, molecule 5 is isolated. Using this example, we can explain the logic used to determine the size of the cluster.

By analysing each pair of molecules for distance, we establish the following neighbours:

1 - 2

1 - 4

2 - 3

2 - 4

By grouping all the neighbours seen by atom 1, we start to build a chain of neighbours:

1 - 2 - 4

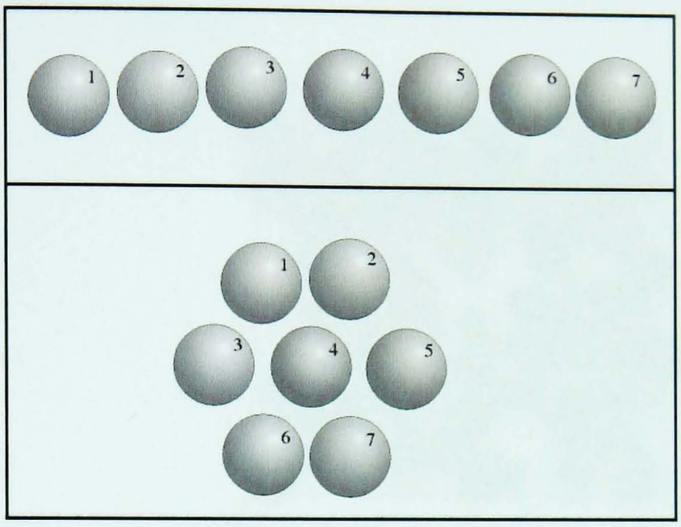


Figure 6.2: Two hypothetical arrangements of ethanol molecules in 2D. The first example clearly offers very little protection from surrounding water molecules. By contrast, in the second example the atoms show much greater protection from water molecules. Despite this, each would be assigned a cluster size of seven molecules.

2 - 3

2 - 4

repeating the process for each atom we can determine the full list of atoms connected by a chain of neighbours:

1 - 2 - 4 - 3

and hence that there are four molecules in the cluster.

Is should be noted that this definition of a cluster does not distinguish the tightness of a cluster. Figure 6.2 shows two example clusters that would not be distinguished by the algorithm. This means that a given cluster size does not guarantee a corresponding level of protection from water molecules.

In practice, the differences are far more subtle. The criterion allows for only a very small percentage of water molecules to be contained within a cluster. This fact can easily be verified visually using 3d visualisation software. We marked water molecules and ethanol molecules not involved in clustering using a blue colouring, and observing the numbers of water molecule contained within a given cluster. An example of such a visualisation is shown in Figure 6.3.

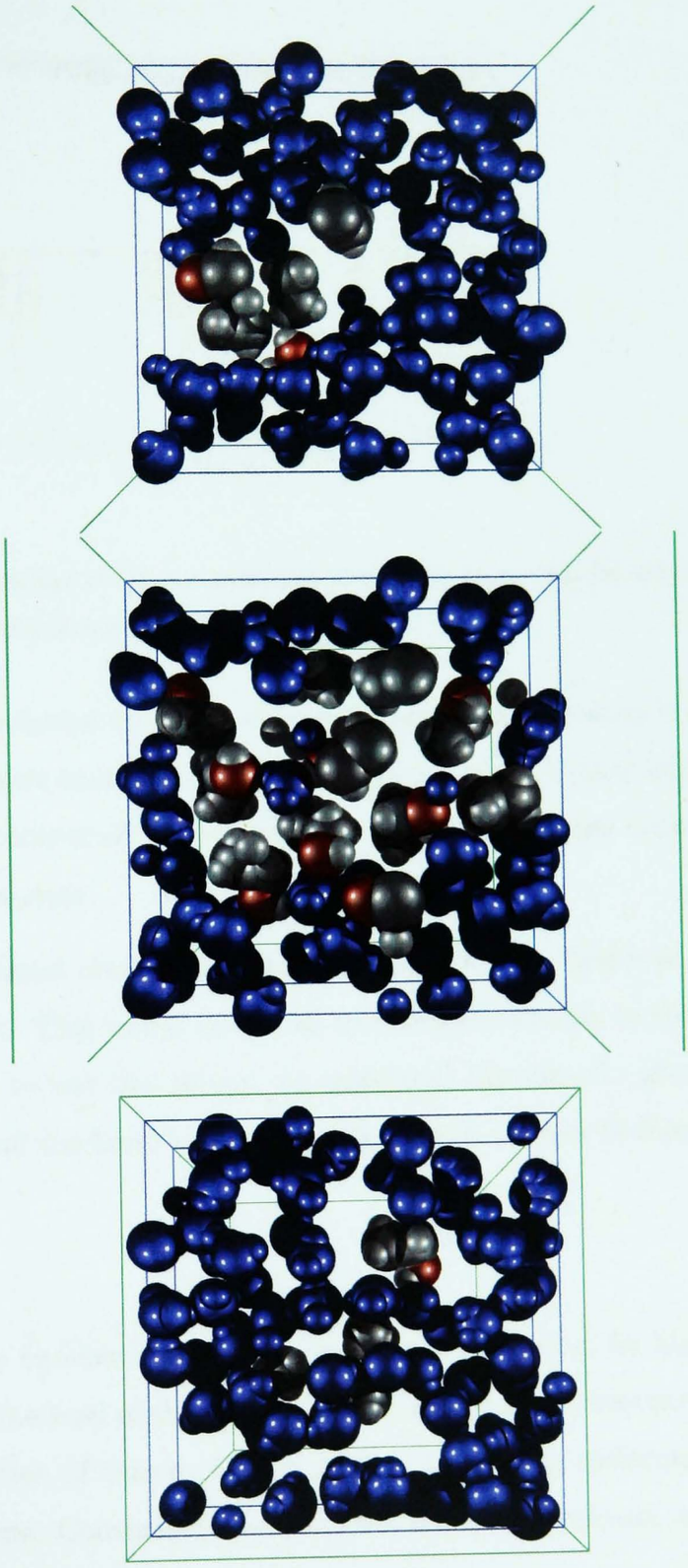


Figure 6.3: Sections through a box of 30 ethanol molecules in water, showing the extent of ethanol clustering. Shown using standard colours is a cluster of 26 ethanol molecules. Water molecules and ethanol that is not part of the cluster are shown in blue.

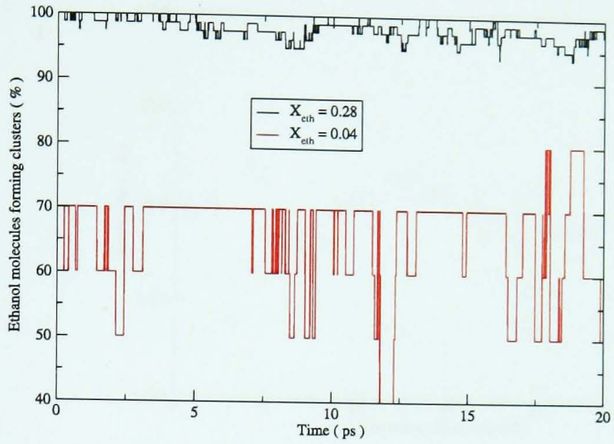


Figure 6.4: The high percentage of ethanol molecules involved in clustering, for low and high concentrations of ethanol molecules.

The level of clustering of ethanol molecules in various mix concentrations is shown in Figure 6.4. The figure shows that the size of the largest ethanol cluster accounts for at least 70% of the total number of ethanol molecules. This percentage increases with increasing ethanol concentration.

The shielding effect of ethanol clusters may be responsible for the failure of the SPM in ethanol-water mixtures. This would create an environment similar to that of pure ethanol at NTP. In order to test this theory, we calculated the size of a given cluster, then took the average of all the bond lengths of molecules in clusters of that size.

6.2.2 Results

Figure 6.4 shows the large number of molecules involved in clustering, for high and low ethanol concentrations. The level of shielding offered to atoms at the centre of a cluster should increase as a function of cluster size. In a large cluster, the molecules may feel no solvation pressure effects. Conversely, small clusters should be exposed, and feel the full effects of the SPM.

Figure 6.5 shows that this is not the case. The figure shows the average bond length as a function of cluster size in a simulation of 30 ethanol and 300 water molecules. As cluster size increases, the bond length remains around $\sim 1.094\mathring{A}$, allowing for random

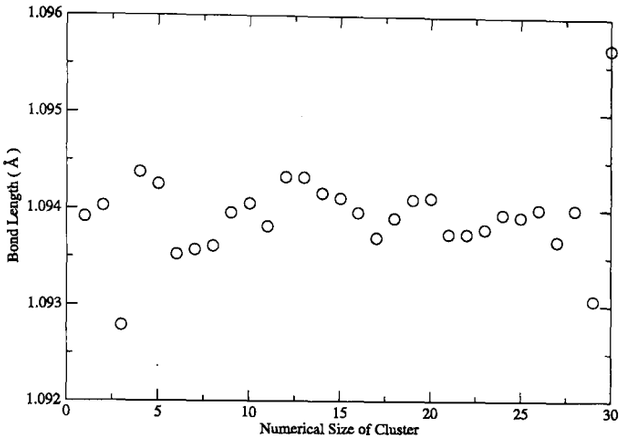


Figure 6.5: The average bond length of the C-H_3 bond as a function of cluster size in a system of 30 ethanol molecules and 300 water molecules. There is seems to be no trend in the bond length as a function of cluster size.

fluctuations due to the large range of bond lengths. It would appear that there is no strong trend relating the bond length to the size of a given cluster.

It should be reiterated that the definition of a cluster may allow water to be contained within its structure. Such water molecules can be seen inside the centre slice of the ethanol cluster shown in Figure 6.3. Before ruling out clustering as a potential factor resulting in the failure of the SPM, we must attempt to implement a more robust method with which to measure shielding effects. This method will be described in the following section.

6.3 Bond Length as a function of Water Proximity

6.3.1 Method

It is possible that simply using the size of a cluster to quantify the shielding effect may be too general an approach. Water molecules can be contained within a cluster, and there is no parameter defining the tightness of a cluster. In other words, nothing in the definition of a cluster implies a given level of protection from surrounding water molecules.

Calculating the distance to the nearest water molecule, d_{sep}, offers a more direct and robust method of assessing the impact of clustering on the SPM. By calculating the average bond lengths of ethanol as a function of d_{sep}, it should be apparent whether ethanol can be protected from water through clustering. This eliminates the possibility of different clusters offering varying protection from water.

This method also solves a potential problem with defining a given cluster as offering uniform protection to all molecules. Within a cluster, atoms at the edge of the molecule will not be protected from the solvation pressure effects of water. Using this method, such atoms will be correctly distinguished.

We grouped together, or 'binned' ethanol molecules within a small range of distance to their nearest water molecule. We calculated the mean bond length of molecules within each group. The average bond length as a function of d_{sep} can then be established.

If clustering is responsible for masking the effects of the SPM, solvation pressure should decrease with increasing distance d_{sep}. Thus, bond length should increase as a function of d_{sep}.

If success is noted in this case, it may be possible to refine the definition of a cluster in order to better describe the level of protection offered. For example it may be possible to count the number of neighbouring pairs. Since the molecules in a long stringy cluster will have less neighbours, this could be used to quantify the resulting level of protection.

6.3.2 Results

The results of this process are shown in Figure 6.6 for the C-H_3 bond. The other non polar bonds, C-H_2 and C-C, follow the same trends. The data becomes noisy at the extremes of proximity to water, due to the small sample size. Although the data spreads at these extremes of proximity, the best fit through this data is a straight line.

The data clearly shows that there is no shortening of the bond when it is in close proximity to the nearest water molecule. These is no difference in bond length at the extremes of separation from water.

The data shows convincingly no effect of the closeness of a water molecule on the bond length of the C-H_3 bond. The gradient of the data shows that the bond length changes by $< 5 \times 10^{-3} \mathring{A}$ per angstrom of separation. This is true of all bonds at the head of the ethanol molecule which do not partake in hydrogen bonding interactions.

By contrast, the hydrogen bonding O-H_O group of ethanol shows an expected reaction to the presence of water. Figure 6.7 shows the relationship between proximity to water and the length of the O-H_O bond in ethanol. Since this bond is seen to be stretched by hydrogen bonding effects, it follows that the bond gets longer as water molecules approach more closely.

The gradient of the data shows that in this case, on average the O-H_O bond is shortened by $\sim 0.03 \mathring{A}$ for each additional Angstrom of distance to the nearest water molecule. The C-O bond is similarly shortened, with a gradient of $\sim 0.03 \mathring{A}$. These values are many times greater than for the non-polar bonds.

Clustering is therefore not responsible for the lack of solvation pressure felt by atoms at the head of the ethanol molecule. Further analysis is required in order to explain why the bond lengths at the head of the molecule might appear unaffected by the proximity of water molecules.

In order to do this, we will first attempt to calculate the forces applied by molecules on a given ethanol bond. We will sort the forces into those which act to lengthen and bond and those which act to compress it. By doing so we hope gain insight into how water molecules affect the bonds of ethanol, and establish why bonds are unaffected by the presence of water molecules. This will be the focus of Chapter 7.

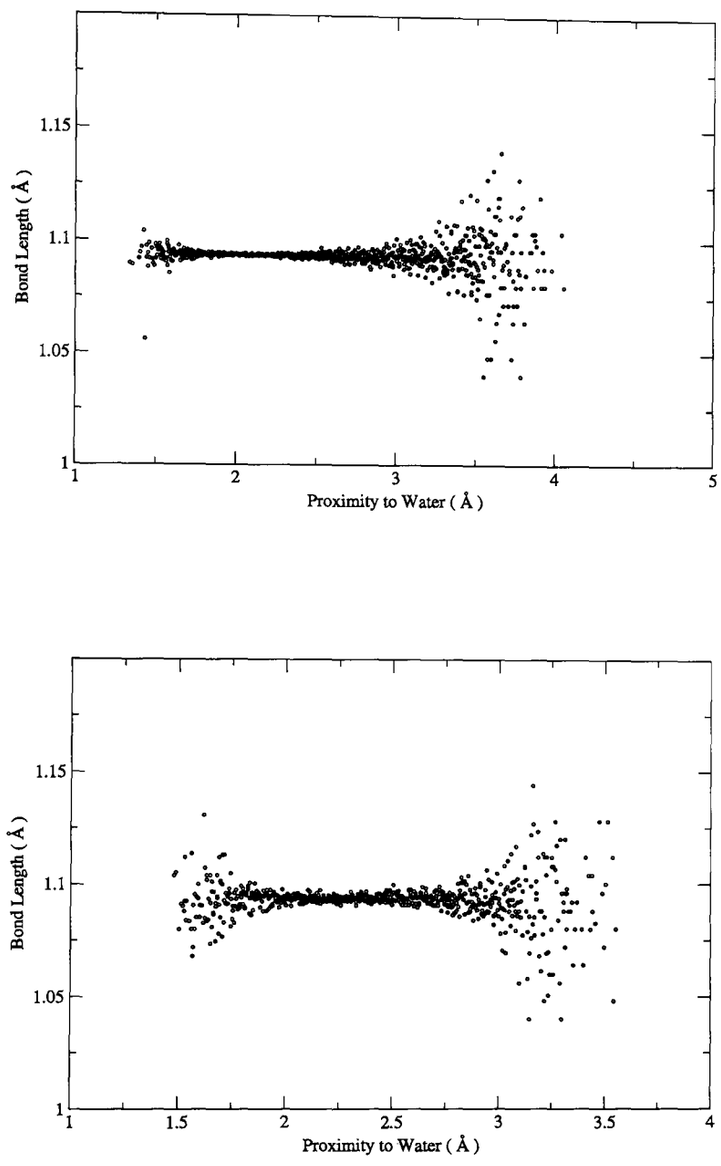


Figure 6.6: Bond lengths of the ethanol C-H_3 bond as a function of increasing distance to the nearest water molecules is shown for high (left) and low (right) concentrations of ethanol.

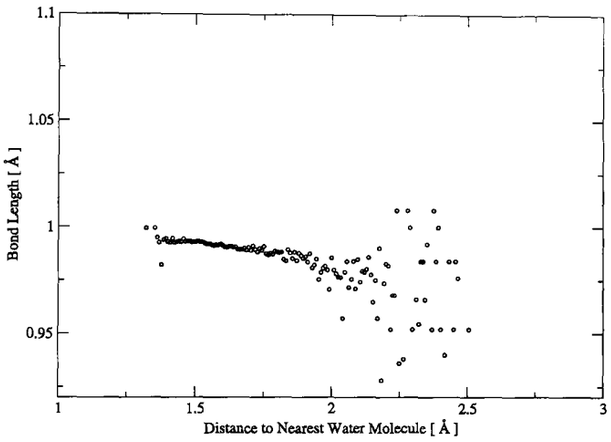


Figure 6.7: O-H_O bond length as a function of increasing distance to the nearest water molecule. Since hydrogen bonding tends to stretch this bond, the length of the bond tends to decrease as water molecules become more distant.

Chapter 7

Analysis of Forces

7.1 Lengthening and Shortening Forces

Chapter 6 offered clustering of ethanol as the first possible explanation for the absence of a solvation pressure effect in ethanol-water mixtures. However, the techniques described suggested that, although extensive clustering is taking place in the mixtures, it does not appear to be responsible for the failure of the SPM. Further analysis is required to explain why the bond lengths of non-polar ethanol bonds might appear unaffected by the SPM. It is therefore a logical step to attempt to calculate and understand the intermolecular forces defining the interaction between ethanol and water.

Figure 7.1 shows a Coulombic interaction F between a hydrogen atom from the methyl group in ethanol and the oxygen in water. The slight positive charge on the hydrogen atom is attracted to the slight negative charge on the oxygen in the water. Because the angle CHO is obtuse, the attractive force will act to pull the hydrogen atom away from the carbon atom, as shown. Using simple vectors, it is therefore possible to calculate the component of the force which acts along the length of the bond. We use the symbol F_b to denote the magnitude of this component of the force.

Figure 7.2 shows the opposite case, in which the C-H_3 bond is shortened by a similar force acting at an acute angle with respect to the bond. The bond is compressed by a force $F\cos\theta$ by the oxygen atom, whose position acts to pull the hydrogen atom closer

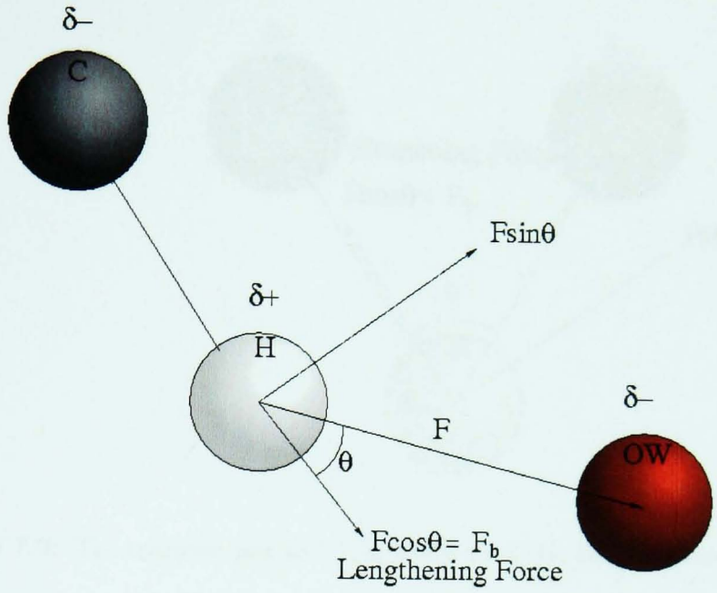


Figure 7.1: An example Coulombic interaction between O_w and H_3. In this case, the C-H_3 bond is lengthened by a force $F_b = F\cos\theta$.

to the carbon atom to which it is bonded.

In this manner it is possible to calculate the contributing force components acting on a bond from different atom types which cause lengthening and shortening of a specific bond. We calculated the average contribution of a given type of atom to the lengthening and shortening of a given bond. Since DLPoly does not provide the level of force statistics required for such calculation, it is necessary to calculate the force components acting on the bond using the positional data in the HISTORY file provided by DLPoly. We calculated lengthening and shortening force contributions every 25 time-steps.

7.2 Force Equations

The equations used to define the intermolecular forces are defined using equations equivalent to those used by the DLPoly program, as detailed below.

The van der Waals component of force is calculated using the standard equation:

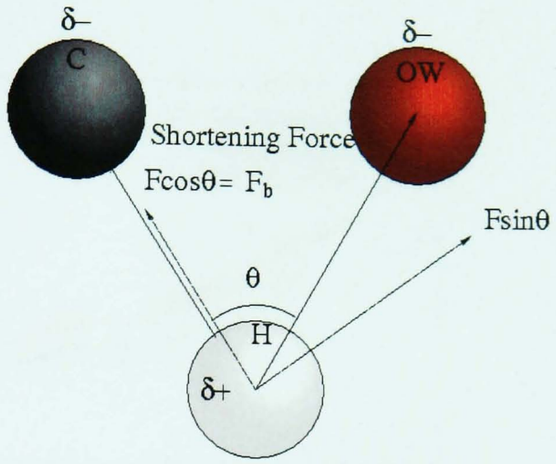


Figure 7.2: The opposite case to 7.1, in which the C-H_3 bond is shortened by a force $F_b = F\cos\theta$.

$$U(r) = 4\epsilon\left(\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^{6}\right) \tag{7.1}$$

$$\underline{f} = -r_{ij}\left(\frac{\delta U(r)}{\delta r}\right)\frac{1}{\underline{r}_{ij}} \tag{7.2}$$

$$\underline{f} = -r_{ij}\left(24\epsilon\left(\frac{2\sigma^{12}}{r^{13}} + \frac{\sigma^{6}}{r^{7}}\right)\right)\frac{1}{\underline{r}_{ij}} \tag{7.3}$$

where σ is the Lennard Jones length parameter, ϵ is the Lennard Jones well depth, and r is the distance between the two interacting atoms.

The long range Coulomb potential is calculated using the distance dependent dielectric method:

$$U(r) = \frac{1}{4\pi\epsilon_0\epsilon(r)}\frac{q_i q_j}{r} \tag{7.4}$$

$$\underline{f} = \frac{\delta U(r)}{\delta r} \tag{7.5}$$

$$\underline{f} = \frac{1}{2\pi\epsilon_0\epsilon}\frac{q_i q_j}{r^4}\underline{r} \tag{7.6}$$

where ϵ_0 is the dielectric constant, $\epsilon(r)$ is the distance dependent dielectric, and q_i and q_j are the charges on the atoms.

For the ethanol to behave in accordance with the SPM, the shortening forces must increase relative to lengthening forces as a function of water concentration. Since the

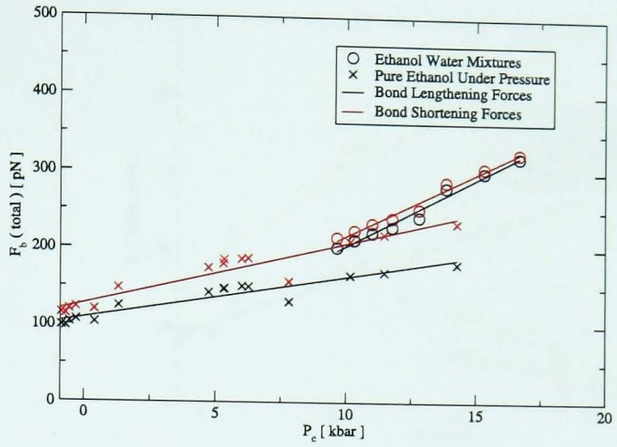


Figure 7.3: The force F_b applied to a C-H_3 bond by van der Waals and Coulomb interactions with surrounding molecules. Forces are divided into those acting to lengthen the bond and those acting to shorten the bond, and are averaged for all bonds of each type in the molecule.

Cornell potentials model hydrogen bonding as an entirely Coulombic interaction, we expect the bulk of this force to be due to large Coulombic interactions.

7.3 Force Analysis: Pure Ethanol

The total intermolecular forces, both lengthening and shortening, affecting the C-H_3 bond in ethanol, as a function of water concentration is shown in Figure 7.3. The intermolecular forces in pure ethanol under hydrostatic pressure are also shown on the same graph for comparison.

Consider first the pure ethanol results. At low pressures, the overall force components acting to lengthen and shorten the bonds are fairly similar. Shortening forces felt by an average C-H_3 bond are just slightly higher than the corresponding lengthening force.

As the pressure increases, ethanol molecules are forced closer together. The graph shows that the shortening forces increase relative to lengthening forces. This is mostly due to an increase in repulsive van der Waals forces acting on the molecules as they are squeezed together. In most cases, this causes the hydrogen atoms of the methyl group

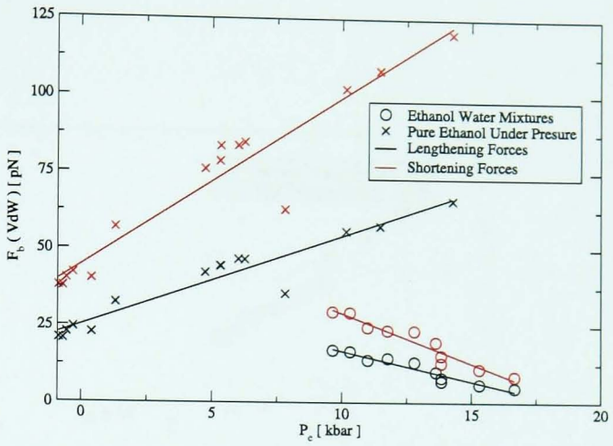


Figure 7.4: The van der Waals force, F_b, acting to lengthen (black) or shorten (red) a C-H_3 bond. Linear regression is applied to the data as a guide to the eye.

to be pushed towards the carbon atom, shortening the bond.

The lengthening and shortening interactions of the van der Waals force is shown in Figure 7.4. As the hydrostatic pressure applied to pure ethanol increases, the van der Waals forces become more repulsive, which tends to shorten the C-H_3 bond.

The net force due to the Coulomb interaction remains constant as a function of hydrostatic pressure. Figure 7.5 shows that lengthening and shortening forces increase largely in parallel with each other in this case. Whereas the van der Waals force is always attractive at large distances and repulsive at short distances, the Coulomb force is always either attractive or repulsive at all distances, depending on the charges on the atoms involved. Therefore, assuming the overall structure of pure ethanol liquid is not greatly altered by increasing pressure, it follows that on average the overall Coulomb force felt by a bond will be unchanged. Larger lengthening forces will be counteracted by larger shortening forces.

By contrast the van der Waals force tends to become more repulsive in all cases where two atoms of any kind are brought together. This causes the total force, shown in Figure 7.6, to be a shortening force.

Notice the intercept of the graph in Figure 7.6. This is the point at which the bonds are neither shortened or lengthened. In the case of pure ethanol, extrapolating the line

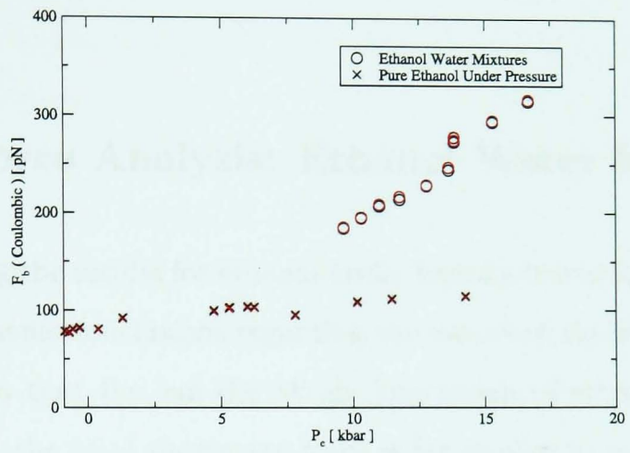


Figure 7.5: The Coulombic force, F_b (Coulombic), acting to lengthen (black) or shorten (red) the C-H_3 bond in ethanol.

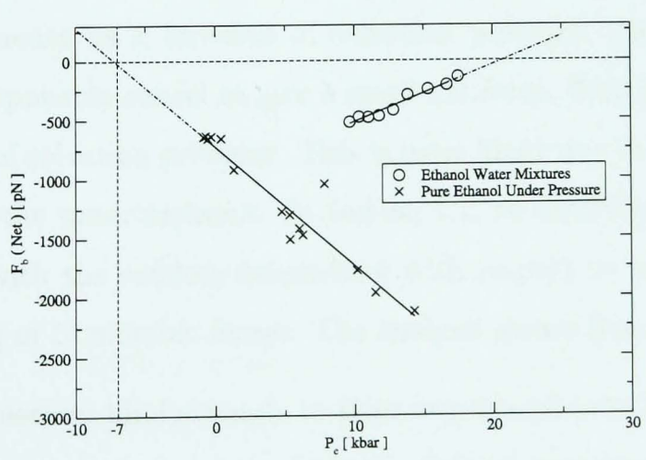


Figure 7.6: The net force, F_b (Net), applied to the C-H_3 bond. This is calculated as the difference between the bond lengthening and bond shortening forces.

of best fit, places the intercept of the data at around $\sim -7kbar$. This value of course represents the position of ethanol vapour on the pressure scale. Therefore the solvation pressure felt by ethanol on transfer from vapour to liquid is confirmed by the force analysis.

7.4 Force Analysis: Ethanol Water Mixtures

Comparing the results for ethanol under high hydrostatic pressure and the mixtures, we can draw some conclusions regarding the nature of the solvation pressure. From Figure 7.4, we see that the van der Waals interaction of ethanol in water is very weak. In particular, the total shortening force is far smaller in mixtures than for ethanol under pressure. This means that the tight binding of water molecules is not acting to squeeze molecules closer together.

It seems that the CED of the water in not sufficient to push the ethanol molecules closer together. This means that increased concentration of water do not induce increased van der Waals forces in ethanol.

The water also does not apply a large net Coulombic force to the head of the ethanol molecule. This is shown in Figure 7.5. Both the lengthening and shortening Coulomb forces increase as a function of solvation pressure. However, as with pure ethanol, these components cancel to give a small net force. This net force remains constant as a function of solvation pressure. This is most likely due to the symmetrical charge distribution of the water molecule. In Section 5.1, we described how this charge distribution, coupled with the random orientation with respect to non polar bonds, may result in cancelling of Coulombic forces. The analysis shown here shows support for this theory.

Water molecules bind strongly to their own kind due to hydrogen bonding. This highly self-attractive nature causes a lack of a definite structure with respect to the non-polar head of ethanol. A lack of structure results in a full range of Coulombic forces being applied, which often cancel to give a small net force.

This gives a clear explanation for the lack of solvation pressure in ethanol water mixtures. The method of force analysis gives a very convenient and convincing method by

which to test the SPM. Future studies should employ these methods as the main test of possible effects on solutes due to solvation pressure.

In Figure 7.6, observe the highest pressure data point in the ethanol-water mixture data. This point represents a mixture of a single ethanol molecule in water, such that the only force contributions affecting the C-H_3 bond are from water molecules. In this case, for the SPM to hold true, the force applied to the single ethanol molecule must fall on the best fit line through pure ethanol data. This would place the point at $\sim -60pN$. Instead, the graph shows that the force applied by water to a single ethanol molecule is in fact lower than in any other system tested. Hence, the force applied to an average ethanol molecule is seen to increase as a function of X_{eth}.

7.5 Transition in Ethanol-Water Mixtures

The force analysis of ethanol-water mixtures shown in Section 7.4 show evidence for a transition in the forces applied to the C-H_3 bond. This transition occurs within the concentration range $0.10 < X_{eth} < 0.14$. This equates to a pressure of $12.8kbar < P_e < 13.8kbar$.

The transition can been seen in the Coulomb, van der Waals and total forces (Figures 7.3 and 7.4 - 7.6). Figure 7.6 with the regression analysis applied separately to simulations on either side of the transition is shown in Figure 7.7. This regression shows that the force applied to the C-H_3 bond jumps suddenly from 7 to $10pN$. A similar transition is seen for the C-C and C-H_2 bonds. However, no transition is seen for the polar bonds C-O and O-H_O.

The reason for this transition is not yet understood. We will see further evidence for a transition in Chapter 8 and will explore the possibility of a transition fully in Chapter 9.

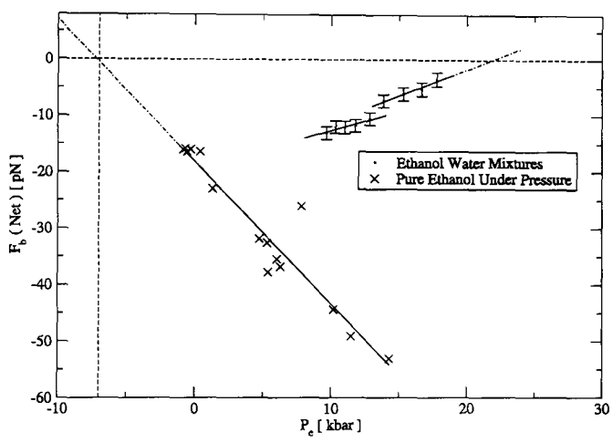


Figure 7.7: The total force component along the length of the C-H_3 bond. This is calculated as the difference between the bond lengthening and bond shortening forces.

Chapter 8

Structural Density Functions

RDFs provide only a limited amount of information about the structure of a system. Using RDFs for a given pair of atoms, such as the O and H_O in ethanol, we can establish the most common distances separating atoms of these types. Using this information we can establish some information about the structure surrounding a typical molecule. For example, the hydrogen bonding interaction is characterised by a tall sharp peak at around $2\mathring{A}$. However, it is often impossible to establish a true picture of the average structure surrounding a typical molecule.

The force ánalysis described in Chapter 7 suggests that there is a transition in the structure of ethanol water mixtures in the concentration range 0.1 - 0.15. It is speculated that this is the result of a change in the structure of the clusters of ethanol molecules. It is possible that at some point the clusters of ethanol molecules switch from loose strings of connected molecules to a more coherent ball. Adopting such a structure allows more water molecules to be expelled from the cluster. This would cause a shift in the force acting on an average molecule, since the nearest neighbours of an atom affect it most significantly.

Molecular visualisations are able to provide some supporting evidence for this transition, as shown in Figures 8.1 and 8.2. However, such evidence is somewhat subjective and limited by the fact that an individual picture can only represent one time-step of a simulation, and one's ability to accurately represent three dimensional structure in a two dimensional medium.

Figure 8.1: A sample cluster in a simulation of 30 ethanol molecules and 300 water molecules. The water molecules have been removed for clarity.

Figure 8.2: A contrasting cluster from a simulation of 40 ethanol molecules and 300 water molecules. Note the more globular nature of the cluster.

Therefore it is important to have a method by which to accurately determine the average local structure surrounding a molecule. The aim is to establish quantitatively the type of molecule surrounding ethanol, the bonds that are interacting and their equilibrium angles.

8.1 Calculating the Structural Density

The Structural Density Function as introduced by Laaksonen et al[61] proved to be a useful tool for presenting this information. A given bond in a molecule is oriented such that the length of the bond is positioned along the z axis, with the centre of the bond at the origin. In the case of ethanol, the central C-C is generally used as the bond of interest, since it represents the centre of the molecule, and placing it at the origin allows the greatest control over positions of the other atoms. The C-C bond is translated so that it is centred at the origin. Rotational matrices are then applied in order to to fix the position of the ethanol atoms:

rotation around the x axis:

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & \sin\theta \\ 0 & -\sin\theta & \cos\theta \end{pmatrix}$$

rotation around the y axis:

$$\begin{pmatrix} \cos\theta & 0 & \sin\theta \\ 0 & 1 & 0 \\ -\sin\theta & 0 & \cos\theta \end{pmatrix}$$

rotation around the z axis:

$$\begin{pmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

If necessary, the molecule is then re-centred such that the centre of the C-C bond is at the origin. The final positions of the skeleton of the ethanol atom is shown in Figure 8.3.

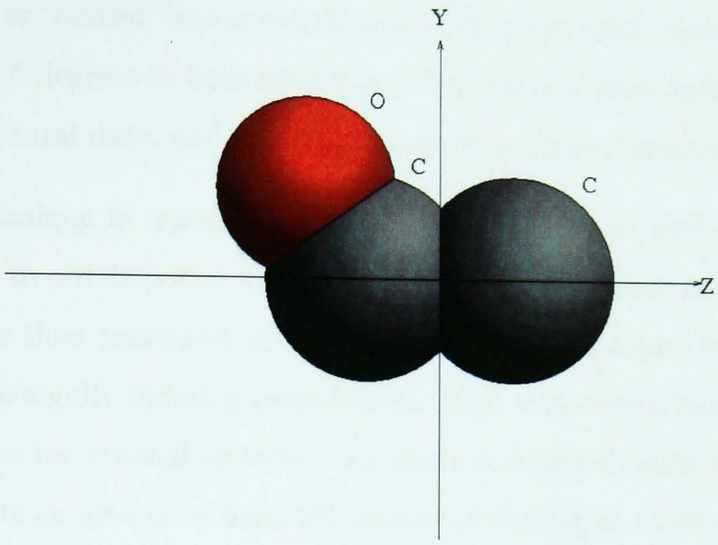


Figure 8.3: The positions of the carbon and oxygen atoms after rotation. The x coordinate is zero for each atom.

The positions of surrounding atoms of interest are analysed in order to determine trends in the angles of interaction. The hydrogen bonding O-H bonds of ethanol and water are of most interest, as they give a clear indication of the types, abundances and positions of molecules surrounding an atom. In order to define the structure immediately surrounding a given molecule, the analysis is based on molecules that fall within the first peak of the RDF. This relates to the few atoms immediately neighbouring the ethanol molecule.

Therefore, using the configuration shown, water molecules within a distance of $4\mathring{A}$ from the C-C bond must be considered. The co-ordinates of atoms of interest are transformed and rotated along with the ethanol molecule. In order to determine the position of the atom with respect to the ethanol molecule, the position of the atom is expressed in spherical polar co-ordinates.

$$r = \sqrt{x^2 + y^2 + z^2} \tag{8.1}$$

$$\theta = \arctan\left(\frac{y}{x}\right) \tag{8.2}$$

$$\phi = \arctan\left(\frac{\sqrt{x^2 + y^2}}{z}\right) \tag{8.3}$$

The spherical co-ordinates θ and ϕ of the atom with respect to the origin are then

indexed or 'binned' into a two dimensional array. Each point in the array represents a range of 5 degrees in each direction. This allows a good balance between resolution of the structural data, and ensuring a large population range of points.

This technique is repeated for each ethanol molecule, and each surrounding atom of interest, to obtain a two dimensional map of the number of atoms at each angle. The data was then processed using the visualisation package OpenDX.[62] We used a ColorMap to signify differing probabilities of an atom being located at a given angle with respect to the ethanol molecule. An angle is coloured using the appropriate colour if it represents an area of at least 2% chance of finding an atom at that point, otherwise it it left transparent. This gave us a two dimensional contour plot of probability data.

We converted the angles from the polar co-ordinates back in to Cartesian x y z values, using a fixed radius of $r = 4\mathring{A}$.

$$x = r \sin\phi \cos\theta \tag{8.4}$$

$$y = r \sin\phi \cos\theta \tag{8.5}$$

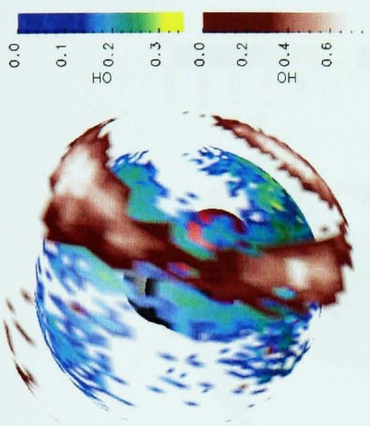
$$z = r \cos\phi \tag{8.6}$$

Using these co-ordinates allowed us to map the data back onto a sphere, using the Compute function. We used Glyphs to represent the positions of the ethanol molecule.

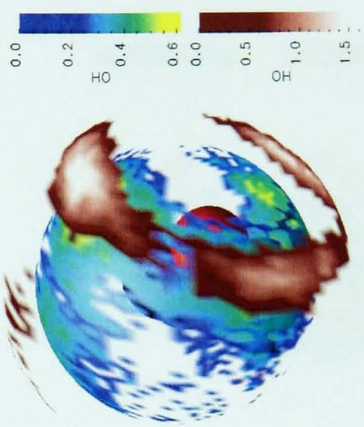
The results of the process are shown in the following sections.

8.2 Structural Densities: Ethanol-Water

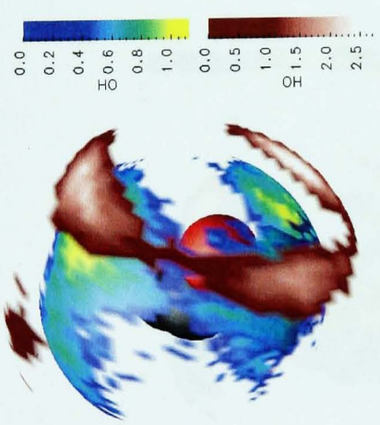
Figure 8.4 shows the structure of water molecules surrounding ethanol. Despite slightly noisy data, the first two figures, Figure 8.4a and 8.4b, show five main sites of interaction between ethanol and water. These five sites are shown most clearly as ethanol concentration increases further in Figure 8.4c. The structural density has a line of symmetry along the plane of the C-C bond in ethanol. Therefore, the red region representing a region often populated by oxygen in water, shown in the bottom left of the figure, has



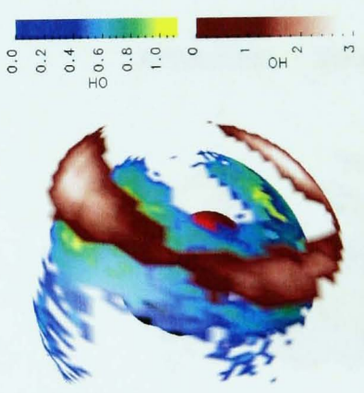
(a) 10 Ethanol Molecules. Mole fraction 0.03



(b) 20 Ethanol Molecules. Mole fraction 0.06

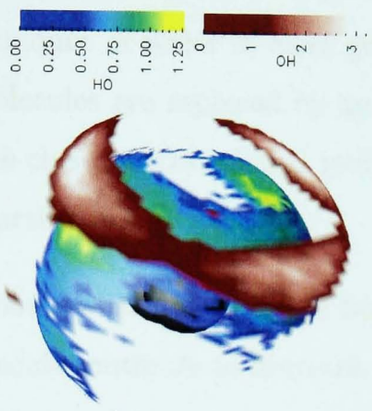


(c) 30 Ethanol Molecules. Mole fraction 0.10

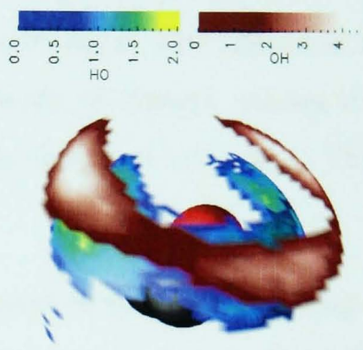


(d) 40 Ethanol Molecules. Mole fraction 0.14

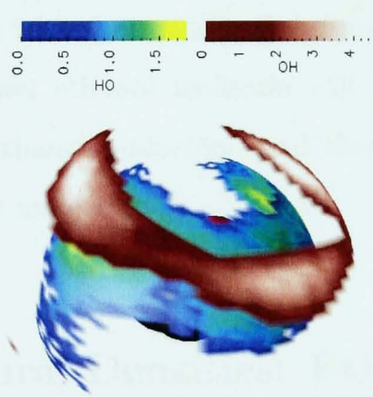
Figure 8.4: Structural Densities of water surrounding ethanol molecules at varying concentrations.



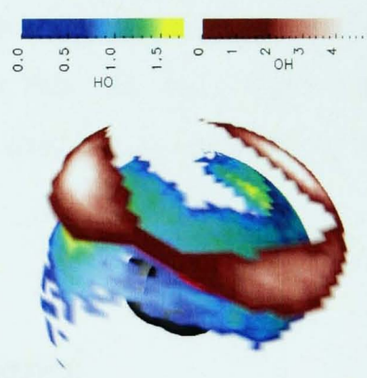
(e) 50 Ethanol Molecules. Mole fraction 0.17



(f) 60 Ethanol Molecules. Mole fraction 0.21



(g) 70 Ethanol Molecules. Mole fraction 0.24



(h) 80 Ethanol Molecules. Mole fraction 0.28

Figure 8.4: Structural Densities of water surrounding ethanol molecules at varying concentrations. (cont.)

a corresponding site to the right of the molecule. However, in the picture shown this region is obscured by other structural information.

As the concentration of ethanol increases again to a fraction of 0.14, as shown in Figure 8.4d the fourth and fifth hydrogen bonding sites seem to disappear, suggesting a possible reduction in the amount of water in close proximity to the ethanol. The likelihood is that the water molecules are replaced by an increased amount of ethanol, raising the possibility that the clustering of ethanol molecules becomes tighter at this point. This will be explored further in Section 8.4.

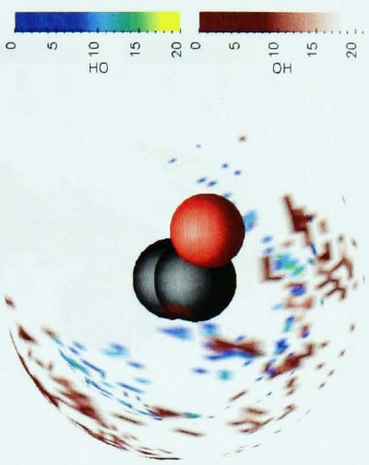
As the concentration of ethanol increases further, the amount of water surrounding a given ethanol molecule continues to decrease. This of course can again be accounted for by the reduction in the amount of water relative to ethanol, but also suggests further increase in the amount of clustering and the tightness of the cluster.

It should be noted that a structural density function depicting, for example, three sites of interaction between ethanol and water, does not necessarily mean that each ethanol molecule interacts with three water molecules. The average number of water molecules surrounding a given ethanol molecule will be lower than this. In many cases sites are shared with ethanol molecules, and therefore not all sites can be expected to be occupied by water molecules at all times.

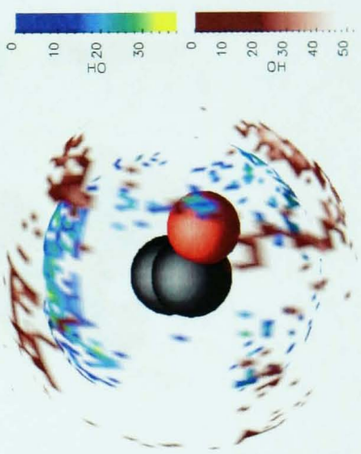
8.3 Structural Densities: Ethanol-Ethanol

Figure 8.5 shows the Structural Density functions representing the positions of the O-H_O group of ethanol with respect to ethanol. The data for the lower concentrations of ethanol (Mole fraction < 0.1), shown in Figures 8.5a and 8.5b are noisy. This is due to the low number of ethanol molecules in close enough proximity for data to be collected. Therefore these data are not useful for analysis.

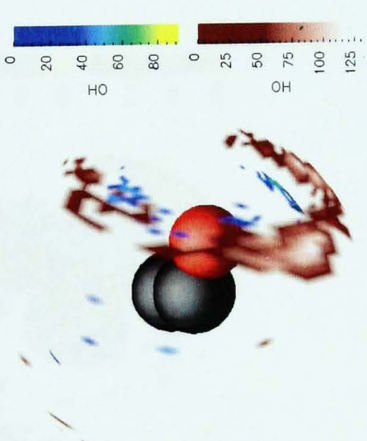
Beginning at 8.5c, at a mole fraction of 0.1, when the ethanol concentration has increased enough to provide useful data, one begins to see the characteristic band of three 'lobes' of oxygen atoms. These lobes exist surrounding the oxygen atom in the ethanol



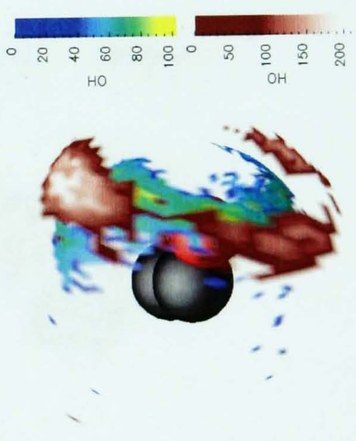
(a) 10 Ethanol Molecules. Mole fraction 0.03



(b) 20 Ethanol Molecules. Mole fraction 0.06

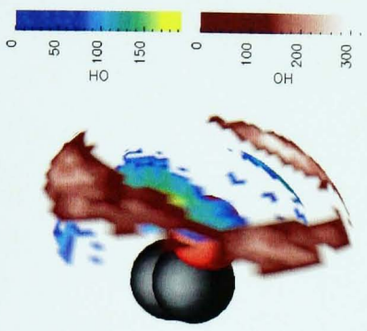


(c) 30 Ethanol Molecules. Mole fraction 0.10

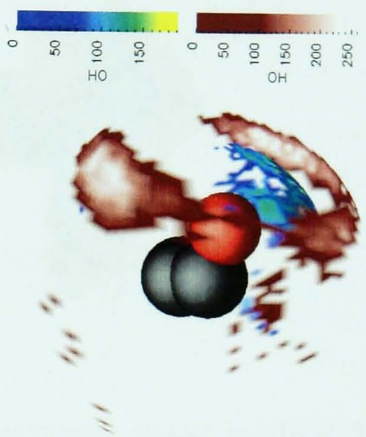


(d) 40 Ethanol Molecules. Mole fraction 0.14

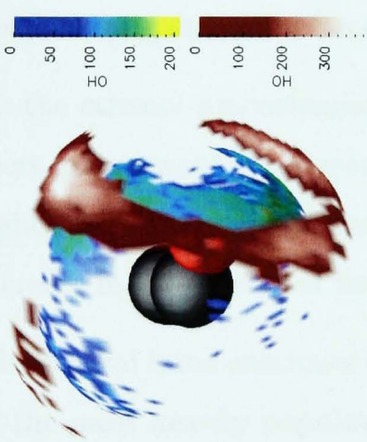
Figure 8.5: Structural Density Functions showing the position of the polar tail of ethanol that surround a given ethanol molecule.



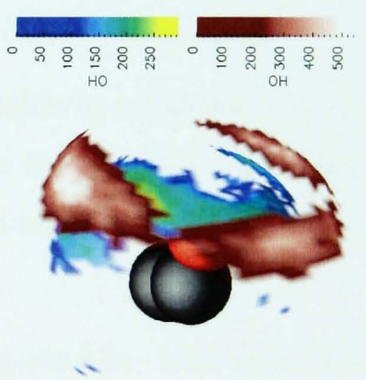
(e) 50 Ethanol Molecules. Mole fraction 0.17



(f) 60 Ethanol Molecules. Mole fraction 0.21



(g) 70 Ethanol Molecules. Mole fraction 0.24



(h) 80 Ethanol Molecules. Mole fraction 0.28

Figure 8.5: Structural Density Functions showing the position of the O-H_O group of ethanol immediately surrounding a given ethanol molecule.

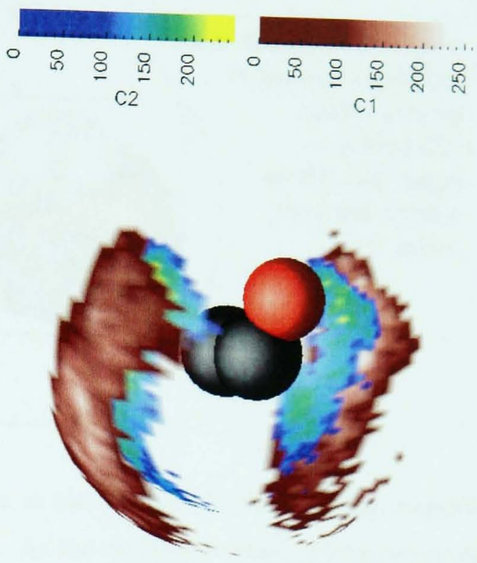


Figure 8.6: The structure of carbon atoms surrounding ethanol molecules. The radius is set at $4.2\mathring{A}$ so as to include atoms the first peak of the radial distribution function. The structure shows that carbon atoms appear most attracted to the sides of the ethanol.

molecule, and are caused by bonding between the oxygen and hydrogen molecules in neighbouring ethanol molecules.

As the ethanol concentration increases further to a fraction of 0.14, the lobes become more pronounced, as shown in 8.5d. On its own this cannot be said to suggest a tightening of the ethanol clustering, since it may also be attributed to continued gradual increase in the quality of the data.

The general trend continues throughout Figures 8.5e - 8.5h, showing increasing intensity of the most heavily populated regions as ethanol concentration increases. Additional peaks in the structure at the fourth and fifth bonding sites, as seen in the pure water data, begin to appear at a mole fraction of ~ 0.2, but quickly disappear again.

Shown in Figure 8.6 is the arrangement of carbon atoms surrounding an ethanol molecule. Those atoms which fall within the first peak of the RDF (Figure 3.1) are included in the structural analysis. The carbon atoms seem to be attracted to the centre of the ethanol molecule, attracted primarily to the carbon atoms.

It is likely that this is simply due to the rectangular nature of the ethanol molecule. Taking molecules within a certain radius of the molecule clearly results in a circle surround-

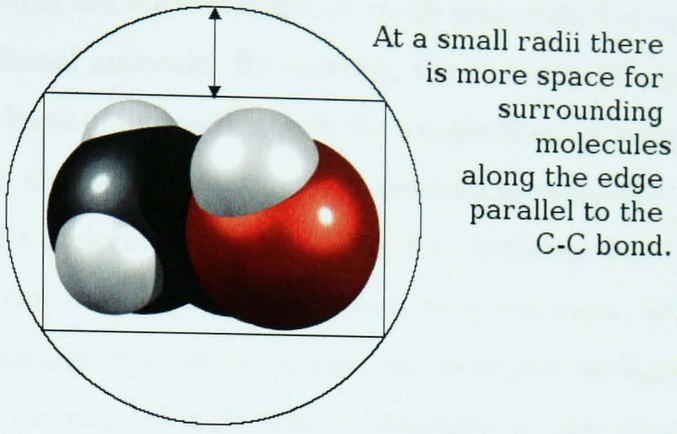


Figure 8.7: At a given radius there is more space for surrounding molecules along the edge parallel to the C-C bond. As the radius increases, this increased space soon becomes negligible. This explains the structure shown in Figure 8.6. It also explains the double peak in the Radial Distribution shown in Figure 3.1.

ing the molecule, rather than a given distance from the edge of the molecule. Therefore, the first space which becomes available for occupation by surrounding molecules is along longest edge of the ethanol molecule, parallel to the C-C bond. This is shown in Figure 8.7. Thus, we witness the pattern seen in Figure 8.6.

As the radius of the circle increases, this increased space soon becomes negligible. Increasing the radius by an extra angstrom allows atoms carbon atoms to surround the molecule on all sides. Therefore, there is a second peak in the RDF at this point.

Repeating the analysis with a cut-off radius of $\sim 5\mathring{A}$, encompassing the second peak of the RDF, shows this more random arrangement of the molecules. The molecule is surrounded fairly evenly at all angles by neighbouring carbon atoms. There is still a slight preference towards the areas shown in Figure 8.6, but this is less pronounced.

8.4 Analysis

Figures 8.4c, 8.4d, 8.5c and 8.5d show the Structural Densities of ethanol water mixtures of concentration 0.1 and 0.14, for which a change in structure was predicted in Chapter 7. Focus first on the system of 30 ethanol molecules (Mole Fraction 0.1), who's structural densities are shown in Figures 8.4c and 8.5c. The water structure shows wa-

ter surrounding the ethanol molecule on all sides with five main locations of interaction with the ethanol molecule. By contrast, the structure of ethanol molecules shows only a very thin band of interaction with the oxygen atom in surrounding ethanol molecules. The picture shown by the structural densities is consistent with that depicted by the snapshot of a molecule shown in Figure 8.1. Although the ethanol molecules are 'clustered' and can be said to be less mixed with the water than you would expect from a random arrangement of molecules, the extended 'stringy' nature of the clustering means that the ethanol molecules are generally in close proximity to water.

Compare this with the system of 40 ethanol molecules (Mole Fraction 0.14), shown in Figures 8.4d and 8.5d. The structure of surrounding water molecules still shows a reasonably high level of interaction with the ethanol. However, the lack of the two lower interaction sites seen in the system of 30 ethanol molecules, suggests that the interaction with water is more inhibited by the presence of surrounding ethanol molecules. This corresponds with a clear increase in the level of interaction with other ethanol molecules. Again, this corresponds with the picture of an ethanol cluster shown in Figure 8.2, which appears to show a tightly packed cluster of ethanol molecules, more consistent with the picture of a cluster as a ball of liquid, protecting its contents from the effects of the water.

Similarly, the structural density of ethanol at a molar volume of 0.21 clearly shows the secondary minima hinted at in the evaluation of forces. This is shown in Figure 8.5f. The area of active hydrogen bonding with water is seen to be reduced when compared with lower ethanol concentrations.

The structural density analysis shown here provides further evidence for a transition in ethanol-water data.

Chapter 9

Transition in Structure of Ethanol-Water Mixtures

In much of the data presented in this document, there is a perceived shift in the properties of ethanol water mixtures in the concentration range 0.1 - 0.15. This is in good agreement with the work of Kiselev et al,[63] who witness corresponding transitions in methanol water mixtures.

The number of hydrogen bonds per water molecule shows such a transition in methanol. The Kiselev paper shows a clear peak in number of hydrogen bonds in methanol as a function of concentration.

In order to recreate this graph, we used identical criteria, as described below, to categorise hydrogen bonds in our ethanol water simulations. The criterion chosen is a geometric, as opposed to an energetic, consideration. To be counted as a hydrogen bond, the oxygen-hydrogen distance must be less than 2.5$\AA$ and also the H..O angle must be greater than 165°, as angles shallower than this cannot be associated with hydrogen bonding.

Using these criteria, we calculated the total number of hydrogen bonds in each mixture. This number was split into contributions of water-water, water-ethanol, ethanol-water and ethanol-ethanol interactions. These contributions were then divided in order to obtain average numbers of hydrogen bonds per molecule.

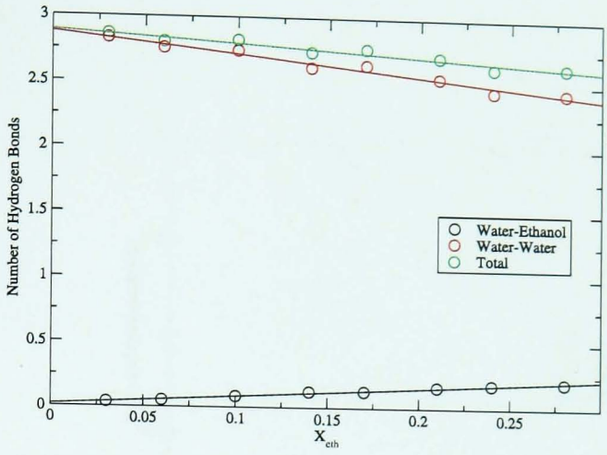


Figure 9.1: A graph showing the water-water and water-ethanol hydrogen bonds per molecule as a function of concentration. The graph resembles that of the ethanol water studies of Noskov et al,[58] rather than the methanol-water studies of Kiselev et al[63] which show a definite transition in structure as a function of concentration.

Shown in Figures 9.1 and 9.3 is the average number of hydrogen bonds per water molecule as a function of ethanol concentration. Unexpectedly, there seems to be no evidence of a transition in this case. Instead, the graph seems to resemble that of Noskov et al[58] in their study of ethanol water mixtures, which shows no evidence for a transition.

It is not clear why this hydrogen bonding analysis should show no transition, when it has been such a feature of other aspects of this study. It is especially unusual in light of the related transition which appears to have been discovered in the work of Kiselev et al.[63] To begin to explain this difference, we must look at the differences in the simulations.

The ethanol model of Noskov et al offers a useful point of comparison, since the simulations were also of ethanol, but using a polarisable force field. The results of the hydrogen bonding in the Noskov paper do not show evidence of a transition. Of course, the most obvious difference is that the Kiselev simulations were of methanol. It is likely that this could account for at least some of the disparity in the results.

It is easy to imagine that any transition would be greatest in the smaller alcohols.

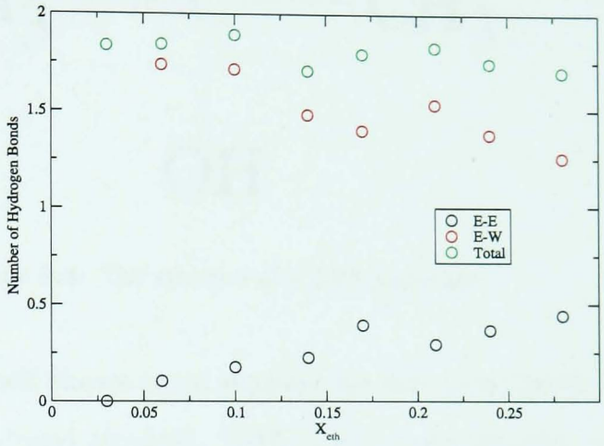


Figure 9.2: The ethanol-ethanol and ethanol-water hydrogen bonds per molecule as a function of concentration. Again, no transition in the number of hydrogen bonds can be observed.

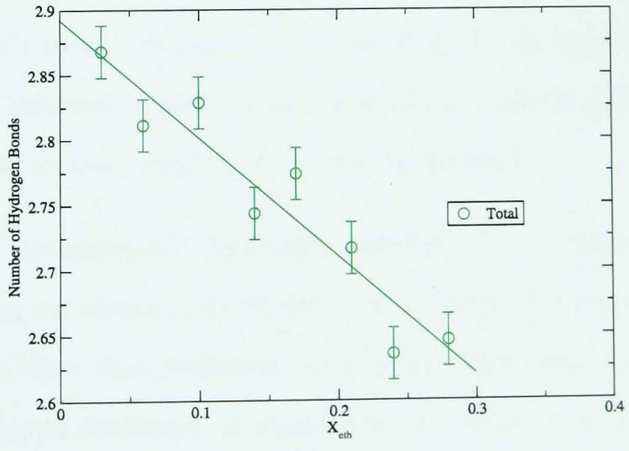


Figure 9.3: A close up of the average total number of hydrogen bonds per water molecule.

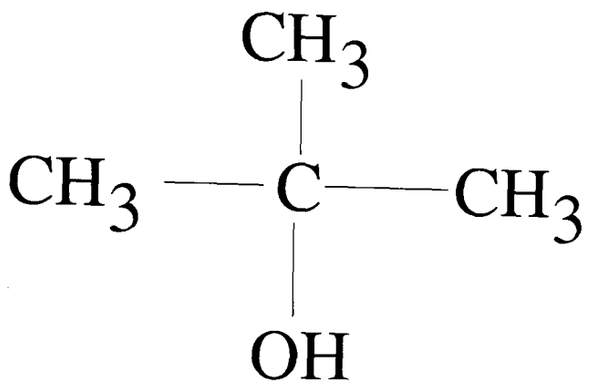


Figure 9.4: The struture of a TBA molecule.

Indeed the Kiselev paper itself shows some support for this hypothesis due to the concurrent study of TBA (tert-butyl alcohol). TBA is a tertiary alcohol, whose structure is shown in Figure 9.4. This larger molecule displays far less of a peak in the level of hydrogen bonding.

Another important difference is the different type of force field used in each simulation. In contrast to the all-atom simulations in this work, and more advance polarisable field of Noskov et al,[58] the Kiselev paper uses a three site potential model to represent the six atoms of methanol, choosing a pseudo-atom to represent the methyl group.

It is possible that the simplification of the Kiselev simulation allows further enhancement of the structural changes which add to the appearance of a transition, sharpening the peak in the level of hydrogen bonding. It follows that if the transition becomes less defined as a function of increasing alcohol size, the use of a single pseudo-atom the methyl group would increase in the prominence of the transition.

It would be interesting to compare the hydrogen bonding in a molecular dynamic simulation of methanol using an all-atom force field. It is likely that such a simulation would result in smaller peak than that presented by Kiselev. However, due to the small size of the head of the methanol molecule, it would almost certainly still be present.

It should be noted that the lack of an apparent transition in the level of hydrogen bonding does not conflict with the view of the structure of ethanol as depicted in the structural density functions in Chapter 8. The number of 'lobes' in the structure surrounding a given molecule does not necessarily correspond to the number of that

particular atom surrounding an average molecule. For example, a structure displaying three lobes may indicate that there are three sites of interaction, each occupied one third of the time. The result would be a structure with three lobes, but an average of one hydrogen bond per molecule.

Chapter 10

Conclusions

10.1 Solvation Pressure

In this study, we have seen some clear evidence for the SPM in pure ethanol simulations. The bond lengths and Raman frequencies of ethanol vapour follow the trends of the SPM for all non-polar bonds. However, those bonds involved in hydrogen bonding interactions showed no evidence of solvation pressure interactions. In this case, the ubiquitous solvation pressure effects are masked by strong hydrogen bonding interactions.

Also, there is no evidence for the SPM in mixtures of ethanol and water. Mixtures of varying quantities of ethanol and water have showed little or no variation in bond length or frequency as a function of solvation pressure.

It is likely that the SPM is only applicable in certain idealised situations. For example a liquid such as methane, which has no strong Coulombic interactions, could be expected to follow the trends of the SPM outlined here. It is likely that all of the bonds in liquid methane are compressed by the solvation pressure.

However, molecules with high partial charges on some atoms, including hydrogen bonding in water, can not be expected to follow such trends. Large Coulombic interactions between ethanol and water have a great impact on structure and properties of a liquid, but do not produce a compressive force conducive to pressure. This is because the

effects of large bond-shortening forces are counterbalanced by the effects of large bond-lengthening forces. This is the case in both pure ethanol and the mixtures. Therefore, in this case, a solvation pressure would require the water molecules to physically squeeze the ethanol molecules together in order to generate repulsive van der Waals forces. This does not happen, because of hydrophobic hydration effects, which mean that the water molecule does not interact strongly with the non-polar head of the ethanol molecule.

This revised condition on the solvation pressure effect limits the potential magnitude of the largest solvation pressures attainable in a mixture, since high Coulombic forces are largely responsible for high CEDs. It also reduces the possible implications of the SPM in some of the potential implications of the model. This is not only because of the reduced theoretical maxima of the solvation pressure effect, but also due to the abundance of hydrogen bonding liquids such as water.

10.2 Implications for Protein Folding

Consider the example of protein folding, cited as a potential implication of the SPM, as in Chapter 1 and the preceding work by Hubel et al.[9] The protein would typically be surrounded by a mixture consisting largely of water. Any fluctuation in solvation pressure would typically represent a reduction in this high average solvation pressure, and theoretically produce a resulting misfolding of the protein. However, if water does not produce a solvation pressure effect on the protein, logically this situation will not arise.

Work is already underway to conduct related research using molecules which do not partake in strong hydrogen bonding. Although the results presented here have highlighted the limitations of the SPM, it would still be useful to produce further test cases involving sample portions of larger proteins to monitor their interactions and the possibility of solvation pressure interactions. It would be useful to build up a picture of which parts of proteins are affected by the solvation pressure, since it cannot be assumed from this study whether, for example, a carboxyl group or nitrogen atom would be similarly unaffected by the solvation pressure of water.

The lack of solvation pressure felt by ethanol in water raises the interesting possibility that water protects molecules from the solvation pressure effect. Consider the C-H_3 bond of ethanol. Using the force analysis shown in Figure 7.6, we can compare the force applied to this bond by water and by other weakly charged atoms. The graph clearly shows that the overall force applied to this bond by water is lower than that applied by ethanol. As has been discussed, this is due to the large Coulomb forces tending to cancel each other out, whereas the van der Waals force tends to become dominated by repulsive forces since this component of the force can be far larger.

Consider a system containing proteins surrounded mostly by water, but also in the presence of other non polar molecules. The sudden random approach of one of the small number of non polar molecules or parts of a molecule, can cause a sudden jump in the overall force applied to the protein. If this occurs during the folding of the protein it may be enough to disturb the process.

It is well known that protein misfolding can be caused by the proximity of other proteins.[65] However, this is not the result of a solvation pressure interaction. Two proteins become entwined when an amino acid links to an amino acid on a second protein, rather than linking with one of its own chain. If a smaller non polar molecule contaminates the folding protein, this is the likely mechanism by which it would interfere with the folding process.

10.3 Other Important Results

During this study it has also been possible to test some other modern theories related to ethanol-water mixtures. It has been possible to extend the 'bi-percolating' model seen by Soper et al[59,60] in methanol-water mixture to the case of these ethanol water mixtures. This is an important success for the bi-percolating model of alcohol water molecules, because it represents not only the extension of the model to include larger alcohols, but also represents the reproduction of the success using a more complicated all-atom model. The original paper used a simplified three site model for the methyl group in methanol, leaving open the possibility that the simplicity of the model revealed an immiscibility which would otherwise not be present. However, these results provide

good support for the results obtained by the Edinburgh group, and serve to allay any doubt over the suitability of their model.

The transition seen in the structure of ethanol water mixtures is also a subject that would benefit from further investigation. It is known from the Kisilev Paper[63] that there appears to be a transition at a similar concentration in methanol. It is possible, therefore that a series of simulations on alcohols of increasing size would give more clues as to the cause of the transition. Measurement of the trends in the position of the transition as well as further structural visual analyses would potentially allow further insight into its causes and implications.

Bibliography

[1] **J. Bradbury** Chaperones: Keeping a Close Eye on Protein Folding *Lancet* 361 (9364) (2003) 1194-5.

[2] **T. Chaudhuri and S. Paul** Protein-Misfolding Diseases and Chaperone-Based Therapeutic Approaches *FEBS J* 273 (7) (2006): 1331

[3] **D. M. Huang and D. Chandler** Temperature and Length Scale Dependence of Hydrophobic Effects and their Possible Implications for Protein Folding *PNAS* 97(2000) 8324-27

[4] **P. R. ten Wolde and D. Chandler** Drying-Induced Hydrophobic Polymer Collapse *PNAS* 99(2002) 6539-43

[5] **R. Zhou, X. H. Claudio, J. Margulis and B. J. Berne** Hydrophobic Collapse in Multidomain Protein Folding *Science* 305(2004) 1605-09

[6] **J. R. Wood, M. D. Frogley, E. R. Meurs, A. D. Prins, T. Peijs, D. J. Dunstan, and H. D. Wagner** Mechanical Response of Carbon Nanotubes under Molecular and Macroscopic Pressures *J. Phys. Chem. B*, 103(1999): 10388-92

[7] **J R Wood, M D Frogley, J. R. Wood, A. D. Prins, D. J. Dunstan, H. D. Wagner** Identity of Molecular and Macroscopic Pressure on Carbon Nanotubes *High Pressure Res*, 18(2000): 153

[8] **J. R. Wood, M. D. Frogley, Q. Zhao, E. R. Meurs, A. D. Prins, T. Peijs, D. J. Dunstan and H. D. Wagner** Carbon Nanotubes: From Molecular to Macroscopic Sensors *Phys. Rev. B*, 62(2000): 7571-5

[9] **N. W. A. van Uden, H. Hubel, D A Faux, A C Tanczos, B Howlin and D J Dunstan** Solvation Pressure as Real Pressure: I. Ethanol and Starch Under Negative Pressure *J. Phys.: Condens. Matter*, 15(2003): 1577-1584,

[10] **H. Hubel, D. Faux, R. B. Jones and D. J. Dunstan** Solvation Pressure in Chloroform *J. Chem. Phys.* 124(2006): 204506

[11] **S. Dixit, W. C. K. Poon and J. Crain** Hydration of Methanol in Aqueous Solutions: a Raman Spectroscopic Study *J. Phys.: Condens Matter* 12(2000): L323-8

[12] **N. W. A. van Uden, H. Hubel, D. A. Faux, D. J. Dunstan, C. A. Royer** Negative Effective Pressures in Liquid Mixtures *High Press. Res.* 24(2003) 205-9

[13] **Y. Melendez-Pagan and D. Ben-Amotz** Intermolecular Forces and Bond Length Changes in High Pressure Fluids. Vibrational Spectroscopic Measurement and Generalized Perturbed Hard Fluid Analysis *J. Phys. Chem. B* 104(2000) 7858-66

[14] **E. J. Hutchinson and D. Ben-Amotz** Molecular Force Measurement in Liquids and Solids Using Vibrational Spectroscopy *J Phys. Chem. B* 102(1998), 3354-6

[15] **M. R. Zakin and D. R. Herschbach** Density Dependence of Attractive Forces for Hydrogen Stretching Vibrations of Molecules in Compressed Liquids *J. Chem. Phys.* 89(1988), 2380

[16] **J. H. Hildebrand and R. L. Scott** Solubility of Nonelectrolytes 3rd Edition *Rheinhold Publishing Company, New York* 1949

[17] **A. M. F. Barton** Handbook of Solubility and Other Cohesion Parameters *CRC Press* 2000

[18] **D. N. Brems, S. M. Plaisted, J. J. Dougherty and T. F. Holzman** The Kinetics of Bovine Growth Hormone Folding are Consistent with a Framework Model *J Biol Chem.* 262(6)(1987) 2590-6

[19] **D.J. Dunstan, and I.L. Spain** The Technology of Diamond Anvil High-Pressure Cells *Journal of Physics E: Scientific Instruments* 22(1989) 913-33

[20] **G. D. Birkhoff** Proof of the ergodic theorem *Proc. Nat. Acad. Sci.* 17(1931), 656-660

[21] **W Smith and T R Forester** A General-Purpose Parallel Molecular Dynamics Simulation Package *J. Molec. Graphics* 14(1996): 136

[22] **M. P. Allen and D. J. Tildesley** Computer Simulation of Liquids *Oxford: Clarendon Press* 1989

[23] **W. L. Jorgensen and J. Tirado-Rives** The OPLS Force Field for Proteins. Energy Minimizations for Crystals of Cyclic Peptides and Crambin *J. Am. Chem. Soc.* 110(1988), 1657-66.

[24] **I. K. Roterman, M. H. Lambert, K. D. Gibson and H. A. Scheraga** A Comparison of the CHARMM, AMBER and ECEPP Potentials for Peptides. II. Phi-Psi Maps for N-Acetyl Alanine N'-Methyl Amide: Comparisons, Contrasts and Simple Experimental Tests. *J. Biomol. Struct. Dyn.* 7(1989) 421-53

[25] **R. Lavery and B. Hartmann** Modelling DNA Conformational Mechanics *Biophys. Chem.* 50(1994), 33-45.

[26] **N. L. Allinger, K. Chen and L. J.-H. Lii** An Improved Force Field (MM4) for Saturated Hydrocarbons. *J. Comput. Chem.* 17(1996) 642.

[27] **T. E. Cheatham, M. F. Crowley T. Fox, and P. A. Kollman** A Molecular Level Picture of the Stabilization of A-DNA in Mixed Ethanol-Water Solutions *PNAS* 94(1997) 9626-30

[28] **B. Brooks, R. Bruccoleri, B. D. Olafson, D. J. States, S. Swaminathan and M. Karplus** A Program for Macromolecular Energy, Minimization, and Dynamics Calculations *J. Comput. Chem.* 4(1983) 187-217

[29] **W. L. Jorgensen, D. S. Maxwell and J. Tirado-Rives** Development and Testing of the OPLS All-Atom Force Field on Conformational Energetics and Properties of Organic Liquids *J. Am. Chem. Soc.* 118(1996) 11225-36

[30] **J. Hermans, H. J. C. Berendsen, W. F. van Gunsteren and J. P. M. Postma** A Consistent Empirical Potential for Water-Protein Interactions *Biopolymers* 23(1984) 1

[31] **W. D. Cornell, P. Cieplak, C. I. Bayly, I. R. Gould, K. M. Merz, Jr., A. M. Ferguson, D. C. Spellmeyer, T. Fox, J. W. Caldwell, and P. A. Kollman** A Second Generation Force Field for the Simulation of Proteins Nucleic Acids and Organic Molecules *J. Am. Chem. Soc*, 117(1995): 5179–5197

[32] **G. Kaminski, E. M. Duffy, T. Matsui and W. L. Jorgensen** Free Energies of Hydration and Pure Liquid Properties of Hydrocarbons from the OPLS All-Atom Model *J. Phys. Chem* 98(1994) 13077-82

[33] **A. E. Mark, S. van Helden, P. E. Smith, Lambert H. M. Jamsen and W. F. van Gunstereno** Convergence Properties of Free Energy: alpha-Cyclodextrin Complexities as a Case Study *J. Am. Chem. Soc.* 116(1994) 6293-302

[34] **J. Aqvist, C. Medina and J. E. Samuelsson** A New Method for Predicting Binding Affinity in Computer Aided Drug Design *Protein Eng.* 3(1994) 385-91

[35] **T. E. Cheatham III and P. A. Kollman** Observation of the A-DNA to B-DNA Transition During Unrestrained Molecular Dynamics in Aqueous Solution *J Mol Biol.* 259(1996) 434-4

[36] **Y. Duan, W. Patricia, M. Crowley and J. M. Rosenberg** Molecular Dynamics Simulation Study of DNA Dodecamer d(CGCGAATTCGCG) in Solution: Conformation and Hydration *J Mol Biol.* 272(1997) 553-4

[37] POV-Ray, The Persistance Of Vision Raytracer. *http://www.povray.org*

[38] **C. I. Bayly, P. Cieplak W. D. Cornell and P. A. Kollman** A Well-Behaved Electrostatic Potential Based Method Using Charge Restraints For Deriving Atomic Charges *J. Phys. Chem.* 97(1993): 10269-80

[39] **S L Mayo, B D Olafson and W. A. Goddard III** DREIDING: a Generic Force Field for Molecular Simulations *J. Phys. Chem.* 94(1990) 8897-990

[40] **D. R. Lide** Handbook of Chemistry and Physics *CRC* 1999

[41] **W. E. Forsythe** Smithsonian Physical Tables (9th Revised Edition) *Knovel* 2003

[42] **W L Jorgensen** Quantum and Statistical Mechanical Studies of Liquids. 12. Simulation of Liquid Ethanol Including Internal Rotation *J. Am. Chem. Soc.* 103(1981): 345-50

[43] **J. H. Nguyen, M. B. Kruger and R. Jeanloz** Compression and Pressure-Induced Amorphization of $Co(OH)_2$ Characterized by Infrared Vibrational Spectroscopy *Phys. Rev. B* 49(1993), 3734-8

[44] **R Ludwig** The Effect of Hydrogen Bonding on the Thermodynamic and Spectroscopic Properties of Molecular Clusters and Liquids *Phys. Chem. Chem. Phys.* 4(2002), 5481-7

[45] Fityk - General-Purpose Nonlinear Curve Fitting and Data Analysis *http://www.unipress.waw.pl/fityk/*

[46] **C-H. Gu, H. Li, R. B. Gandhi and K. Raghavan** Grouping Solvents by Statistical Analysis of Solvent Property Parameters: Implication to Polymorph Screening. *Int. J. Pharmaceutics* 283(2004) 117-25

[47] **J. Burke in C. Jensen (ed.)** Solubility Parameters: Theory and Application *AIC Book and Paper Group Annual* 3(1984): 13-58

[48] **B.Guillot** A Reappraisal of what we have Learnt During Three Decades of Computer Simulations on Water *J. Mol. Liquids* 101(2002) 219-260

[49] **H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren and J. Hermans in B. Pullman (ed.)** Intermolecular Forces *Reidel, Dordrecht* 1981 p331.

[50] **W L Jorgensen, J Chandrasekhar, JD Madura, RW Impey, and ML Klein** Comparison of Simple Potential Functions for Simulating Liquid Water *J. Chem. Phys.* 79, 926-935

[51] **P. J. van Maaren and D. van der Spoel** Molecular Dynamics of Water with Novel Shell-Model Potentials *J. Phys. Chem. B* 105(2001) 2618-26

[52] **G. S. Fanourgakis and S. S. Xantheas** The Flexible, Polarizable, Thole-Type Interaction Potential for Water (TTM2-F) Revisited *J. Phys. Chem. A* 110(2006) 4100-6

[53] **M. W. Mahoney and W. L. Jorgensen** A Five-Site Model for Liquid Water and the Reproduction of the Density Anomaly by Rigid, Nonpolarizable Potential Functions *J. Chem. Phys.* 112(2000) 8910-22

[54] **K. Kiyohara, K. E. Gubbins and A. Z. Panagiotopoulos** Phase Coexistence Properties of Polarizable Water Models *Molecular Phys.* 94(1998) 803-8

[55] **D. van der Spoel, P. J. van Maaren and H. J. C. Berendsen** A Systematic Study of Water Models for Molecular Simulation: Derivation of Water Models Optimized for use with a Reaction Field *J. Chem. Phys.* 108(1998) 10220-30

[56] **P. G. Kusalik and I. M. Svishchev** The Spatial Structure in Liquid Water *Science* 265(1994) 1219-21

[57] **M. N. Rodnikova, T. M. Val'kovskaya, V. N. Kartez and D. B. Kayumova** About Elasticity of Spatial H-bond Network in Liquids *J. Mol. Liquids* 106(2003) 219-2

[58] **S. Y. Noskov, G. Lamoureux and B. Roux** Molecular Dynamics Study of Hydration in Ethanol-Water Mixtures Using a Polarizable Force Field *J. Phys. Chem. B* 109(2005) 6705-13

[59] **L. Dougan, S. P. Bates, R. Hargreaves, J. P. Fox, J. Crain, J. L. Finney, V. Reat and A. K. Soper** Methanol-Water Solutions: A Bi-Percolating Liquid Mixture *J. Chem. Phys.* 121(2004), 6456

[60] **S. Dixit, J. Crain, W. C. K. Poon, J. L. Finney and A. K. Soper** Molecular Segregation Observed in a Concentrated Alcohol Water Solution *Nature* 416(2002), 829

[61] **A. Laaksonen, P. Kusalik and I. M. Svishchev** Three Dimensional Structure in Water Methanol Mixtures *J. Phys. Chem.* 101(1997), 5910-8

[62] OpenDX Data Explorer. *http://opendx.org*

[63] **M. Kiselev and D. Ivlev** The Study of Hydrophobicity in Water-Methanol and Water-Tert-Butanol Mixtures *J. Molec. Liquids* 110(2004) 193-9

[64] **A. H. Narten, M. D. Danford and H. A. Levy** X-Ray Diffraction Study of Liquid Water in the Temperature Range $4 - 200$ deg C, *Faraday Discuss.* 43(1967), 97-107

[65] Mechanism of Protein Misfolding Captured in Computer Simulation *http://www.sandia.gov/media/fold.htm* Retrieved 1/08/2006