Structures in LKα14 dissolved in binary mixtures of water and ethanol

Dissertation

zur Erlangung des Doktorgrades

der Naturwissenschaften (Dr. rer. nat.)

der Fakultät für Chemie und Pharmazie

der Universität Regensburg

vorgelegt von

Ragnheiður Guðbrandsdóttir

aus

Reykjavík, Island

im Jahr 2022

Abstract

Proteins and peptides can influence the local structure of a solvent around the solute, while the solvent influences the conformation of the protein or peptide. Both of these effects can be observed by investigating the model peptide LKα14 in different mixtures of water and ethanol. Molecular dynamics simulations of these systems show how the peptide has different secondary structure depending on the solvent composition, while the presence of LKα14 can induce a separation of the fully miscible solvents, water and ethanol, into a water-rich and an ethanol-rich domains.

The main properties of LKα14 in different water/ethanol mixtures analysed in this work are:

- The secondary structure of LKα14 in the bulk and at the surface of different water/ethanol mixtures. LKα14 forms an α-helix at the surface of all solvent compositions and in the bulk of any solvent composition containing ethanol but not in the bulk of pure water.

- The adsorption of LKα14 to the surface of the different water/ethanol mixtures and the orientation of the peptides at the surface. In pure water and at low ethanol concentration LKα14 adsorbs to the surface. At the surface the α-helix lies parallel to the surface with the leucine residues orientated towards the interface and the lysine residues orientated towards the bulk.

- The solvation shell of LKα14. Depending on the solvent composition there is excess solvation of LKα14 by water or ethanol. The ideal ethanol mole fraction in the solvation shell is around 0.4. For all solvent compositions the ethanol mole fraction of the solvation shell is closer to 0.4 than the ethanol mole fraction of the bulk is. The water molecules and ethanol molecules in the solvation shell are also spatially separated, with ethanol close to the leucine residues and water close to the lysine residues.

- Aggregates formed by LKα14. In pure water and at low ethanol concentration LKα14 peptides form oligomers while at higher ethanol concentrations a microemulsion-like structures are formed, with water and ethanol corresponding to the two solvents and LKα14 to the surfactant. At lower ethanol concentration the aggregates percolate similar to a bicontinuous microemulsion.

Zusammenfassung

In Lösung befindliche Proteine und Peptide können die lokale Struktur des Solvenz in ihrer Umgebung beeinflussen. Umgekehrt kann das Lösungsmittel die Konformation des Proteins bzw. des Peptides beeinflussen. Bei der Untersuchung des Modellpeptides LKα14 in verschieden Mischungen von Wasser und Ethanol können beide Effekte beobachtet werden. Molekulardynamik-Simulationen dieser Systeme zeigen, dass die Sekundärstruktur des Peptides von der Lösungsmittelkomposition abhängt, während das Vorhandensein von LKα14 zu einer Trennung der völlig mischbaren Lösungsmittel Wasser und Ethanol in eine wasserreiche bzw. eine ethanolreiche Domäne führt.

Die Haupteigenschaften von LKα14 in Wasser/Ethanol-Mischungen, die in dieser Arbeit analysiert werden sind:

- Die Sekundärstruktur von LKα14 im Bulk und an der Oberfläche unterschiedlicher Wasser/Ethanol-Mischungen: LKα14 formt eine α-Helix an der Oberfläche aller Lösungsmittelkompositionen und im Bulk jeder Lösungsmittelkomposition, die Ethanol beinhaltet — nicht jedoch im Bulk reinen Wassers.

- Die Adsorption von LKα14 an die Oberfläche unterschiedlicher Wasser/Ethanol-Mischungen und die Orientierung der Peptide an der Oberfläche: In reinem Wasser und bei niedriger Ethanolkonzentration adsorbiert LKα14 an die Oberfläche. An der Oberfläche liegen die α-Helices parallel zur Oberfläche, die Leucinseitenketten zeigen zur Oberfläche und die Lysinseitenketten zum Bulk.

- Die Solvathülle von LKα14: Je nach Lösungsmittelkomposition gibt es Exzesssolvatation von Wasser oder Ethanol am Peptid. Der bevorzugte Molenbruch von Ethanol in der ersten Solvathülle ist ungefähr 0.4. Der Molenbruch von Ethanol in der ersten Solvathülle ist in jeder Lösungsmittelkomposition näher an 0.4, als der Molenbruch von Ethanol im Bulk ist. Die Wassermoleküle und die Ethanolmoleküle in der Lösungshülle sind auch räumlich getrennt — die Ethanolmoleküle näher an die Leucinseitenketten und die Wassermoleküle näher an die Lysinseitenketten.

- Aggregate von LKα14: In reinem Wasser und bei niedriger Ethanolkonzentration formen LKα14-Peptide Oligomere. Bei höheren Ethanolkonzentrationen bilden sich mikroemulsionsähnliche Strukturen, in denen das Wasser und das Ethanol den zwei Lösungsmitteln entsprechen und die LKα14-Peptide den Tensiden. Bei niedrigeren Ethanolkonzentrationen perkolieren die Aggregate, ähnlich wie in einer bikontinuierliche Mikroemulsion.

Contents

List of Figures

Part I.

Introduction

1. Introduction

Proteins are the building blocks of all living organisms. Proteins fulfil many roles in biology, ranging from structural proteins like collagen, which is found in connective tissues such as skin or ligaments or keratin, which is a key component of hair, nails, horns and feathers, to enzymes, which can catalyse chemical reactions. For them to function it is necessary that they possess the correct conformation. This functional conformation is called the native state. The native state is defined both by the secondary structure (local conformation of the peptide at each amino acid) and the tertiary structure (global three dimensional structure of the amino acid chain). In contrast a protein without well defined tertiary or secondary structure is considered unfolded.

In the last two decades another type of folding has increasingly been studied: Proteins can misfold and form insoluble aggregates. This is known to induce different diseases [1–10] – most of them neurological[6–10]. One prominent example is the formation of fibrils by denaturation and aggregation of the protein fragment amyloid-β. It has been linked to Alzheimer's disease[3–7]. Further examples are prion diseases like Creutzfeldt–Jakob disease or scrapie[1–3,7–9]. In both cases the aggregation involves a change of conformation to an insoluble β-sheet[3,5–8].

Since aggregation and precipitation of proteins has been linked to so many diseases, understanding the thermodynamics of protein conformations and the underlying interactions between protein and the surrounding solvent is important. Thermodynamically protein conformations are only weakly stable and the delicate balance between different protein conformations is sensitive to changes in the solvent composition[11–14]. For example the presence of ions[15,16], osmolytes[17,18] (for example urea or sucrose) or denaturants can bias the folding equilibrium. Both ions and osmolytes can, depending on the specific molecules, both denature and stabilise proteins while denaturants, as the name suggests, denature them. In nature these competing effects are exploited as a way of controlling protein folding[19]. Whenever mixed solvents are present the shift of equilibria can be described thermodynamically by preferential solvation. Since the interactions between the solute and the different components of the solvent are of different strength, aggregation of a specific solvent component in the solvation shell of the solute and corresponding depletion of the bulk can occur[20–22]. Focusing on the solvent, this implies that it reacts sensitively to the presence and conformation of the protein. The scenario is then that in the system of protein and solvent both parts influence the structure of one another, but are at the same time subject to alterations in their own structure. That is, the conformation of a protein is influenced by the solvent in which it is dissolved and the protein conformation in turn influences the behaviour of the solvent. This can lead to different solubility and aggregation behaviour of proteins in these different solvents.

One class of solvents which are interesting is colloidal systems: microemulsions and

micellar solutions[23–29]. Since such systems consist of two distinct pseudo-phases (one polar and one apolar) they can be a good way to bring molecules which dissolve in different environments close to each other[28,29]. This in turn can be used to catalyse reactions[28,29] while in other cases the micelle itself or its surface acts as a catalyst[28–31]. Micellar or microemulsion systems are also an option of dissolving water-insoluble or poorly water-soluble molecules[23,29]: The molecules are dissolved inside the micelles which are themselves dissolved in the water, a fact which is considered of interest for drug delivery[32–35] and has been used for cleaning[23,36] and extraction of solutes[37,38]. Micellar solutions have also been used to extract proteins both from water solution[38–40] so it is not unreasonable to suggest they might be capable of dissolving protein aggregates which are water insoluble.

On the application side, the presence of surfactants is often undesired. Many surfactants are toxic[41,42] and most are petrol-based[42,43]. An alternative is the use of simpler solvents. Hydrotropes, which are short amphiphilic molecules which do not form micelles, can also facilitate solubility of non-polar compounds[44–50]. Some examples of hydrotropes are caffeine[47], nicotinamide[49], urea[49,50], sodium xylenesulfonate (SXS)[44,48] and ethanol[47]. One mechanism of hydrotrope action is hydrotrope aggregation in the presence of solutes[48–52]. The solutes are preferentially solved by the hydrotrope in comparison to water which leads to aggregation of the hydrotrope around the solute[49–51]. According to this mechanism, the solvation shell of a protein should be very sensitive to the protein folding state, if one of the conformations possesses large hydrophobic parts at the surface and other conformations do not.

A very simple system, in which such a behaviour can be studied, is a short peptide which folds into one distinct secondary structure when dissolved in a binary mixture of water and a hydrotrope, e.g. ethanol. Binary mixtures of water and ethanol are a good starting point from where it is possible to go on to explore more complicated solvent mixtures, such as water, ethanol and octanol, which are known to form microemulsion like structures[53–55] with octanol in the role of oil and ethanol as the surfactant. As is the case with microemulsions the exact structures of such systems depend heavily on the composition of the system[56,57].

In this work we consider a small model peptide, LKα14. It was designed to form α-helices at the surface of water[58] and adsorbs to polar/apolar surfaces[58–60]. It consist of 14 amino acid residues, half of which are hydrophilic and half of which are hydrophobic. When folded into an α-helix the hydrophobic side chains form a continuous hydrophobic surface on one side of the peptide. Thus it is very well suited for the intended study.

In this work the described system LKα14 in water/ethanol is studied using large scale molecular dynamics (MD) computer simulations. MD simulation with explicit solvent is a well suited method for the study of delicate solvation effects in which the structure of the solvent is important and where different implicit solvation model can therefore not be applied. The peptide was studied in water/ethanol mixtures of different compositions, ranging from pure water to pure ethanol. Both bulk solutions and interfaces were studied.

The following work consist of 11 chapters organised into three main parts. The first part contains the theoretical background and details the methods used. Chapter 2 has

a general discussion of preferential solvation and distribution functions. Chapter 3 discusses protein structure and the method used to determine the secondary structure of LKα14 in this work. Chapter 4 discusses LKα14 and related model peptides. Chapter 5 describes the method used to calculate the orientation of a helix axis from points along the axis. Chapter 6 describes micelles and microemulsions and details how the clusters observed in this work were defined. Chapter 7 contains the details of the performed MD simulations and the force fields used. The second part contains the results of the performed simulations. In chapter 8 the secondary structure of the peptide both in bulk and at the surface is examined. In chapter 9 the adsorption behaviour and the orientation of the peptide is examined. Chapter 10 details preferential solvation of LKα14 in the different solvent mixtures and the structure of the solvation shell. Finally in chapter11 clusters formed by the peptides and water molecules are examined. The third and final part consist of the final remarks.

Part II.

Theory and methods

2. Solvation

When a solute is dissolved in a solvent it influences the structure of the solvent. Such influence includes induced orientation of polar solvent molecules around charged solutes[61] and large solute molecules occupying volume which the solvent otherwise would have occupied[62]. In solvents, such as water, which form networks of hydrogen bonds, these networks also differ around the solute molecule. Those differences depend on whether or not the solute participates in hydrogen bonding[61,63].

The solvation shell of a solute is the interface of the solvent around the solute. Solvent molecules tend to form layers around the solute molecule, with all solvent molecules in a layer at similar distance from the solute. This is, for example, the case for ions in water[61]. There the water molecules of the innermost layer — the first solvation shell — are oriented with one of the hydrogen atoms pointing towards the ion if the ion is positive and with the oxygen atoms pointing towards the ion if it is negative[61]. The water molecules of the next layer, the second solvation shell, then form hydrogen bonds with the water molecules in the first solvation shell[61]. The number of well defined solvation shells depends on the solute and the solvent[61,64].

2.1. Preferential solvation

In a binary solvent the mole fraction of the different solvent components in the solvation shells around the solute is not necessarily equal to the mole fraction in the bulk. If the solute has a stronger affinity with one of the components, the mole fraction of that components is higher in the solvation shells. This difference is called preferential solvation[20,22].

2.2. Radial distribution functions

In a simulation where the coordinates of every atom are known, radial distribution functions (RDFs) are a good way to characterise the solvation shells — both the number of well defined solvation shells and the distance of the different solvation shells from the solute. The RDF between particles of type a and b, where $a \neq b$, is defined as

$$g_{ab}(r) = \frac{\langle p_b(r) \rangle}{\langle p_b \rangle_0} \tag{2.1}$$

where $\langle p_b(r) \rangle$ is the particle density of b at distance r from particles a and $\langle p_b \rangle_0$ is the average particle density of b in the whole volume[65,66]. RDFs have the property that

$$\lim_{r \to 0} g_{ab}(r) = 0 \tag{2.2}$$

and

$$\lim_{r \to \infty} g_{ab}(r) = 1. \tag{2.3}$$

Equation 2.2 corresponds to the fact that no two particle occupy the same point in space and equation 2.3 the fact that the influence of particle a on the distribution of b decreases with distance[65].

In a simulation $\langle p_b \rangle_{\max}$, the average particle density of b in all spheres of radius $r_{\max}$, is used instead[66]. Then equation 2.1 becomes

$$g_{ab}(r) = \frac{\langle p_b(r) \rangle}{\langle p_b \rangle_{\max}}. \tag{2.4}$$

If the solute is a and the solvent b the RDF has a peak for each solvation shell. For large non-spherical solutes, like peptides, there are more than one way to calculate r. Which way is chosen influences how easily the RDF can be interpreted. The simplest way is to calculate r from a single point of the solute (such as the centre of mass) to a single point of each solvent molecule. This is not a good option for solute molecules which depart significantly from a spherical geometry. The simulations used to analyse the solvation shells started from a helical conformation of the peptide and were not long enough for the secondary structure to change. A 14 amino acid long peptide in a helical conformation is decidedly not spherical and this simple method of calculating r is therefore not suitable.

2.3. Spatial distribution functions

Spatial distribution functions (SDFs) are similar to RDFs but include not just the distance to a reference but also the direction from that reference[67,68]. That is $g_{ab}(x, y, z)$ depends on the spatial position in relation to the reference molecule. Therefore the SDFs yield additional information about the 3D distribution of particles around the reference compared to RDFs.

When calculating SDFs with GROMACS the first step is to centre the simulation box around the molecule the SDFs will be calculated from[69]. Then each frame is rotated and translated such that the reference molecule is fitted to its position in the first frame or a reference structure[69]. When calculating SDFs with LKα14 as reference only the backbone of the peptide was used for the fit as the side chains are more flexible. When the trajectory has been prepared in this way, the simulation box is divided into bins in all three dimension and the particle density in each bin calculated. By averaging over time the spatial distribution function around the peptide is obtained[69].

3. Protein structure

The structure of proteins is considered at four different levels: primary structure, secondary structure, tertiary structure and quaternary structure. The primary structure is the sequence of amino acids in the protein. The secondary structure is the local structural elements (α-helix, β-sheet) while the tertiary structure describes how these elements make up the global structure of the protein. The quaternary structure describes how different polypeptide chains arrange together to form large proteins.

For small peptides such as LKα14 only the primary and secondary structures have to be considered since they are too small to have a tertiary structure. Peptides can form ogliomers and clusters, but since they do not have protein function they are not considered quaternary structure.

In this chapter the secondary structure of LKα14 in the bulk and at the surface of different mixtures of water and ethanol will be considered. Since LKα14 is a peptide, not a protein, in the theoretical discussion in 3.1 "peptide" will be used even when discussing facts about secondary structure which also hold for proteins.

3.1. Secondary structure elements

When determining the secondary structure only the conformation of the backbone is relevant. The side chains influence which secondary structure the peptide assumes, but are not directly relevant for determining it.

Along the backbone of a peptide there are three different bonds (excluding the bonds with the side chains and oxygen and hydrogen):

- The peptide bond between different amino acids.
- The bond between the nitrogen atom and the α-carbon atom of the same amino acid.
- The bond between the α-carbon atom and the carbon atom of the CO group.

The peptide bond is planar since it has considerable double bond character[70–72], but rotations around the other two bonds are possible. The torsion angle around the C_α−C bond is denoted as ψ and the torsion angle around the C_α−N bond as ϕ. Both are defined as 180° when all atoms of the backbone lie in the same plane as the peptide bond and grow with a clockwise rotation as seen from the C_α atom. The secondary structure at each residue of a peptide is determined both by the torsion angles and hydrogen bonds between different residues of the peptide. Only hydrogen bonds between the NH and CO

groups of the backbone are considered, since (as mentioned earlier) only the backbone is relevant for determining the secondary structure.

The two most common secondary structure elements are the α-helix and the β-sheet[71,72]. Other secondary structure elements include 3_{10}-helices[73,74], π-helices[73,74], β-hairpin[75] and turns[74,76]. Parts of the peptide chain which are not part of any well defined secondary structure element are considered random coil.

3.1.1. α-helix

The α-helix was first described in 1951 by Pauling, Corey and Branson[70,77]. In an α-helix the hydrogen atom of the NH group of residue i forms a hydrogen bond with the oxygen atom of the CO group of residue $i + 4$ (see figure 3.1). This leads to a helix with 3.6 residues per turn around the helix or a turn of 100° around the helix per amino acid. An ideal α-helix has a translation of 0.15 nm along the helical axis for each amino acid, which corresponds to a pitch of 0.54 nm[71]. Almost all α-helices in proteins are right handed, although short left handed α-helices exist[72,78].

3.1.2. β-sheet

In an β-sheet the backbone of the peptide is extended and two or more extended peptide chains are connected with hydrogen bonds. In peptide chains in the middle of a β-sheet which consist of more than two chains, each residue is connected with two hydrogen bonds to one neighbouring chain while the neighbouring residues form hydrogen bonds with the other neighbouring chain. Two neighbouring chains can either be parallel or anti-parallel[71,72,75]. The peptide chains which partake in an anti-parallel β-sheet can be right next to each other in the overall peptide chain with a turn consisting of two residues between them or strands from different regions of the peptide[75]. In a parallel β-sheet the strands in contact always originate from different regions of the peptide, since a longer connecting chain is needed in order for the amino acids to line up[79].

3.1.3. Other secondary structure elements

As mentioned previously in this chapter other secondary structure elements include β-hairpins, turns, π-helices and 3_{10}-helices. β-hairpins are small β-sheets which consist of two extended peptide strands which are anti-parallel and connected with a turn[75]. A turn is a region where the direction of the peptide chain changes sharply over a few residues, such as two residues connecting the anti-parallel strands in a β-sheet or a region with the same curvature of the backbone as a helix but so short that it does not reach a whole circle[75,76]. Both π-helices and 3_{10}-helices are similar to α-helices but with the hydrogen bond between residues i and $i + 5$ and residues i and $i + 3$ respectively[73,80].

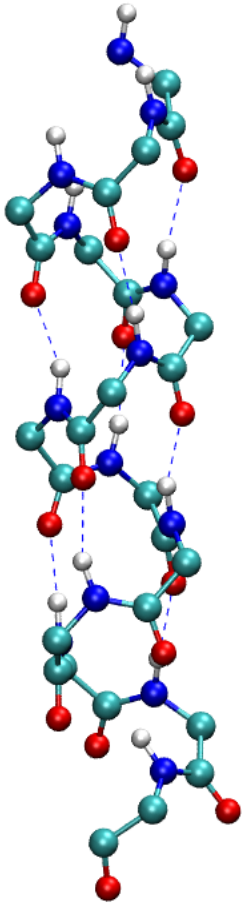


Figure 3.1.: Side view of the backbone of an α-helix. For all residues which are part of the helix there is a hydrogen bond between the CO group of residue i and the hydrogen of the NH group of residue $i + 4$.

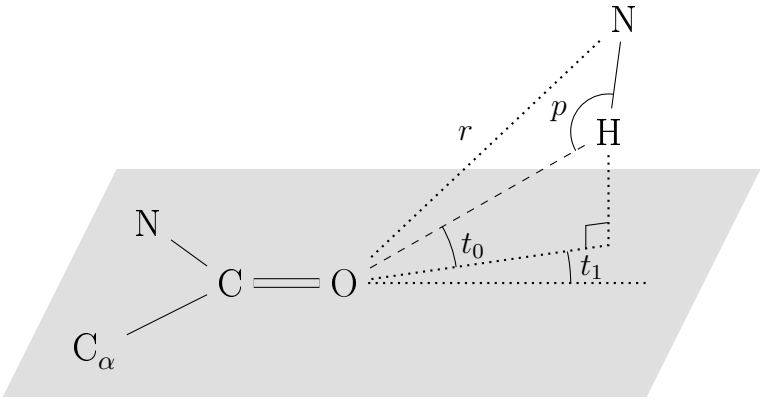


Figure 3.2.: Geometry of a hydrogen bond between the CO group of one amino acid and the NH group of another amino acid in a protein or peptide.

3.2. Stride

The program used to analyse secondary structures in this work is STRIDE[74,81]. The program classifies each residue according to secondary structure[82]. To classify each residue STRIDE calculates both the backbone geometry and the hydrogen bonding pattern. The parameters used to assign an α-helix or extended conformation (β-sheet) to a residue are fitted to assignments by crystallography. For rarely occurring structural elements less data exists and therefore the assignment criteria from theory are used[74]. STRIDE distinguishes between 7 different secondary structures: α-helix, extended conformation, isolated bridge, π-helix, 3_{10}-helix, turn and random coil[82].

For the the recognition of α-helixes and β-sheets STRIDE uses the empirical energy function in eq. 3.1 to calculate the hydrogen bond energy E_{hb} between the CO group of one amino acid and the NH group of an other.

$$E_{\mathrm{hb}} = E_r E_t E_p \tag{3.1}$$

where E_r is the distance term and E_t and E_p describe the direction properties of the hydrogen bond, E_t in relation to the CO group and E_p in relation to the NH group.

$$E_r = \frac{C}{r^8} + \frac{D}{r^6} \tag{3.2}$$

where r is the distance between the two atoms, $C = -3E_m r_m^8$ kcalÅ^8/mol, $D = -4E_m r_m^6$ kcalÅ^6/mol and E_m and r_m are the optimal hydrogen bond energy and length respectively. E_p has the form

$$E_p = \cos^2(p) \tag{3.3}$$

where p is the angle between the hydrogen bond and the NH bond (see figure 3.2) and

E_t has the form

$$E_t = \begin{cases} (0.9 + 0.1\sin(2t_1))\cos(t_0), & 0 < t_1 \leq 90^\circ \\ K_1\left(K_2 - \cos^2(t_1)\right)^3\cos(t_0), & 90^\circ < t_1 \leq 110^\circ \\ 0, & 110^\circ < t_1 \end{cases} \tag{3.4}$$

where t_0 is the angle between the line of the CO bond and the projection of the hydrogen bond on the plane of the CO group, t_1 is the angle between the hydrogen bond and the projection of the hydrogen bond on the plane of the CO group (see figure 3.2) and K_1 and K_2 are empirical constants[74].

The first step in the STRIDE algorithm is checking for α-helixes. A minimal α-helix for STRIDE includes two consecutive hydrogen bonds between residues i and $i+4$ such that

$$E_{\text{hb}}^{i,i+4}\left(1 + W_1^\alpha + W_2^\alpha\frac{P_i^\alpha + P_{i+4}^\alpha}{2}\right) < T_1^\alpha \tag{3.5}$$

where P_n^α is the empirical probability of an α-helix residue having the torsional angles of residue n and W_1^α, W_2^α and T_1^α are empirical weights. If such a minimal α-helix is detected (the condition in eq. 3.5 fulfilled between residues k and $k+4$ and between residues $k+1$ and $k+5$), residues $k+1$, $k+2$, $k+3$ and $k+4$ are considered part of the α-helix. Residues k and $k+5$ are also considered part of the α-helix if $P_k^\alpha < T_2^\alpha$ and $P_{k+5}^\alpha < T_3^\alpha$ respectively, where T_2^α and T_3^α are empirical weights. Overlapping α-helices are merged into a single helix[74].

The next step is checking for β-sheets. A minimal β-sheet in STRIDE involves two consecutive hydrogen bonded β-bridges where each β-bridge involves two hydrogen bonds. Both hydrogen bonds in a parallel β-bridge have to satisfy

$$E_{\text{hb}}\left(1 + W_1^\beta + W_2^\beta\frac{P_a^\beta + P_b^\beta}{2}\right) < T_{\text{p}}^\beta$$

for it to be recognised as a β-bridge (P_a^β and P_b^β are the empirical probabilities of a residue which is a part of a parallel β-sheet having the torsional angles of residues a and b respectively and W_1^β, W_2^β and T_{p}^β are empirical weights). Similarly both hydrogen bonds in an anti-parallel β-bridge have to satisfy

$$E_{\text{hb}}\left(1 + W_1^\beta + W_2^\beta C\right) < T_{\text{ap}}^\beta$$

for it to be recognised as a β-bridge ($C = \frac{P_a^\beta + P_b^\beta}{2}$, $C = P_a^\beta$ if residue b only participates in one hydrogen bond in the β-sheet or $C = P_b^\beta$ if residue a only participates in one hydrogen bond) and T_{ap}^β is an empirical weight. Adjacent β-bridges are merged into a single β-sheet if there are no more than four residues between the bridges on one strand and no more than one residue between them on the other strand. The intervening residues are also considered part of the β-sheet and all residues in the β-sheets are assigned an extended

state as secondary structure. In this step isolated β-bridges which are too far away from other β-bridges to merge with them are also found[74].

Finally STRIDE checks for turns and 3_{10} and π-helices. Here a simple check for hydrogen bond presence is used[74]. If a hydrogen bond is found between residue i and residue $i+3$, $i+4$ or $i+5$ both residues and all intermediate residues are considered part of a turn. If there are two consecutive turns, defined by hydrogen bonds between residues i and $i+3$ and residues $i+1$ and $i+4$ respectively, both turns are merged into a 3_{10}-helix. Two consecutive turns defined by hydrogen bonds between residues i and $i+5$ and residues $i+1$ and $i+6$ respectively similarly define a π-helix. Overlapping helices of the same type are merged into longer helices[74,80].

If after all three checks a residue could be assigned multiple secondary structures, the priority ranking of the different secondary structures is: α-helix, extended conformation, isolated bridge, π-helix, 3_{10}-helix and turn. If a residue does not qualify for any other secondary structure it is assigned random coil.

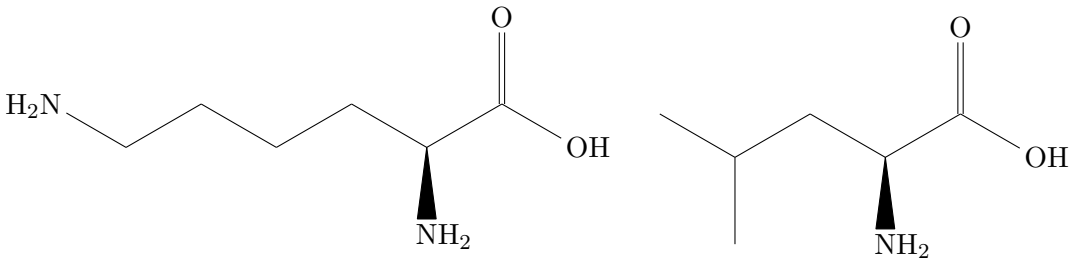
4. The peptide

LKα14 was one of three model peptides designed by DeGrado and Lear in 1985[58]. The three peptides all consist of the same amino acids: leucine (Leu, L) and lysine (Lys, K) in the same proportions. The primary structure of the three peptides is LKLKLKL (LKβ7), LKKLLKL (LKα7) and LKKLLKLLKKLLKL (LKα14). As such only the different primary structure and not any properties inherent to an amino acid can lead to different secondary structures of the peptides[58]. All of the three peptides have periodic primary structures, but the length of the repeat units is different.

DeGrado and Lear showed that these three peptides have different properties in aqueous solutions. In the bulk of an aqueous solution LKα7 is in a random coil conformation at all concentrations. The other two are only in a random coil conformation at low concentration when there is no aggregation. At higher concentration LKβ7 takes on an extended conformation and the peptides form β-sheets with each other. These aggregates precipitate out of the solution. LKα14 forms tetramers with the peptides in an α-helical conformation. Both LKα14 and LKβ7 adsorb to the air/water interface, LKα14 forming α-helices and LKβ7 forming β-sheets, while LKα7 shows only marginal surface activity[58]. Other studies have confirmed the surface activity of both LKα14 and LKβ15 (primary structure LKLKLKLKLKLKLKL) at water/apolar interfaces as well as their conformation when adsorbed[59,60,83–85].

As already mentioned the only properties which differ between the different model peptides are the length and the periodicity. As can be seen in figure 4.1 leucine is an amino acid with an apolar side chain while lysine has a polar side chain. This means that the side chain of leucine is hydrophobic while the side chain of lysine is hydrophilic. All of the model peptides are therefore amphiphilic. Depending on the periodicity of the primary structure different secondary structures result in a spatial separation of the hydrophilic and hydrophobic side chains. For LKβ7 and LKβ15 an extended conformation results in the lysine and leucine side chains being on opposite sides of the extended peptide strand while for LKα7 and LKα14 an α-helical conformation means that a plane parallel to the helix axis divides the two types of residues (see figure 4.2). But as shown by DeGrado and Lear LKα7 is too short to form a stable α-helix[58].

In this work only LKα14 and its behaviour is investigated – both at the surface and in the bulk of water/ethanol mixtures. At the surface the adsorption, orientation and conformation is examined while in the bulk the conformation and aggregation is studied.



(a) Lysine (b) Leucine

Figure 4.1.: The two amino acids which the three model peptides designed by DeGrado and Lear consist of. Lysine has an amino group at the end of the side chain, while the side chain of leucine is aliphatic. Therefore the side chain of lysine is polar and hydrophilic while the side chain of leucine is apolar and hydrophobic.

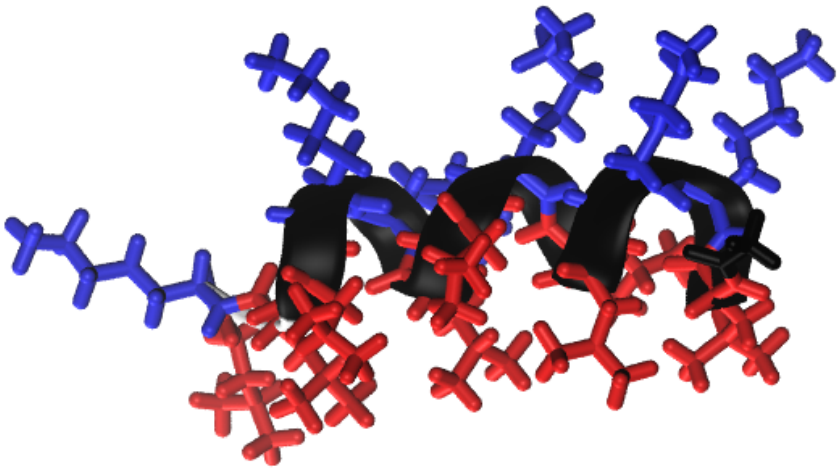


Figure 4.2.: LKα14 in α-helical conformation.Leucine residues are shown in red and lysine residues in blue while the backbone is black. It can clearly be seen how the different amino acids form different sides of the helix.

5. Orientation of helices

In order to determine the angle between the axes of helices and the surface it is necessary to determine which amino acids of the peptides are part of the α-helices and then to find the axes through these helices. As noted before (chapter 3.2) STRIDE[82] is used to determine the secondary structure of the peptides and as such to find which amino acid form part of the α-helices. A method developed by Peter C. Kahn[86,87] is then used to find the axis through each α-helix. In this method points along the axis are determined[86] through which a line is then fitted[87]. Finally the angle between each axis and the plane of the surface is calculated.

5.1. Points along the axis

When simulating an α-helix with an all-atom force field, the information one has about the helix itself is the position of each atom along the backbone. These atoms can be considered to form three helices, one consisting of the α-carbons of each amino acid and the others the nitrogen atoms of the amino groups and the carbon atoms of the carboxyl groups respectively. These helices have the same axis but different radii and run slightly translated from each other. Each of these helices therefore needs to be considered separately when finding the points along the axis (see figure 5.1). Figure 5.2 shows the vectors used and how they relate to the helix.

Take any two amino acids on the helix other than the first and the last one. Let $\boldsymbol{n}_1$ and $\boldsymbol{n}_2$ be the vectors from the origin to the N atoms of the amino acids. Then from the N atom that defined $\boldsymbol{n}_1$ construct a vector $\boldsymbol{a}_1$ to the N atom of the amino acid before it on the helix and a vector $\boldsymbol{b}_1$ to the N atom of the next amino acid along the helix. Define vectors $\boldsymbol{a}_2$ and $\boldsymbol{b}_2$ similarly from the N atom that defined $\boldsymbol{n}_2$. The bisector of the angle defined by either $\boldsymbol{a}_1$ and $\boldsymbol{b}_1$ or $\boldsymbol{a}_2$ and $\boldsymbol{b}_2$ is perpendicular to the helix axis and will intersect it[86]. The sum of $\boldsymbol{a}_1$ and $\boldsymbol{b}_1$ is a vector with the same direction as the bisector of the angle. Let $\boldsymbol{v}_1$ be a unit vector of that same direction and $\boldsymbol{v}_2$ a unit vector similarly obtained from $\boldsymbol{a}_2$ and $\boldsymbol{b}_2$.

Both $\boldsymbol{v}_1$ and $\boldsymbol{v}_2$ are perpendicular to the helix axis and therefore $\boldsymbol{v}_1 \times \boldsymbol{v}_2$ has the same direction as the axis. Let $\boldsymbol{h}$ be a unit vector with this direction and $\boldsymbol{h}_1$ and $\boldsymbol{h}_2$ vectors from the origin to the intersection of the axis and the extension of $\boldsymbol{v}_1$ and $\boldsymbol{v}_2$ respectively. Let r be the radius of the helix and d the distance between the end points of $\boldsymbol{h}_1$ and $\boldsymbol{h}_2$ on the axis. Then

$$\boldsymbol{n}_1 + r\boldsymbol{v}_1 = \boldsymbol{h}_1, \tag{5.1}$$

$$\boldsymbol{n}_2 + r\boldsymbol{v}_2 = \boldsymbol{h}_2 \tag{5.2}$$

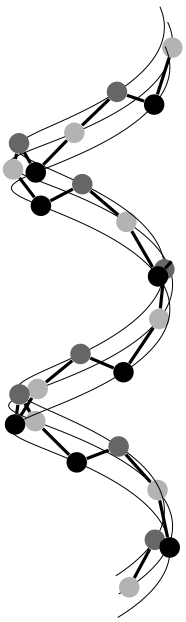
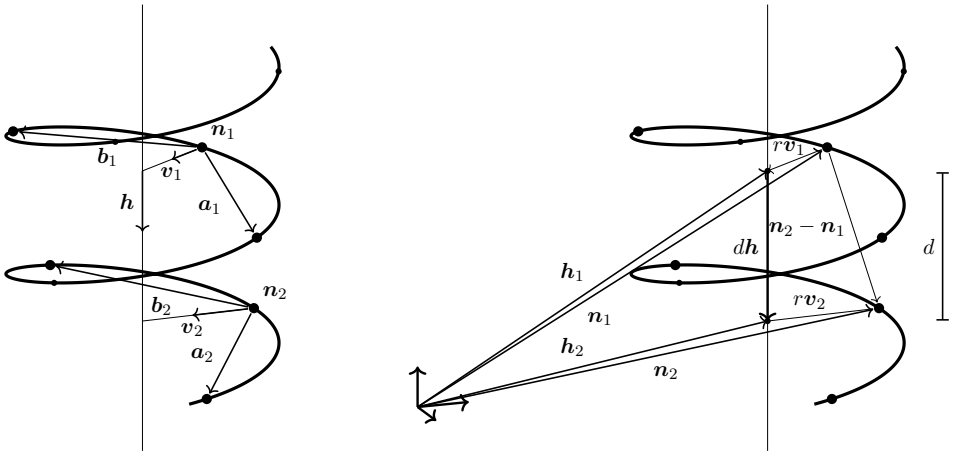


Figure 5.1.: A scematic drawing of an α-helix showing the three partial helices used for the analysis.



(a) (b)

Figure 5.2.: The vectors used to find the axis.

and

$$\boldsymbol{h}_1 + d\boldsymbol{h} = \boldsymbol{h}_2. \tag{5.3}$$

Equations 5.2 and 5.3 show that

$$\boldsymbol{h}_1 + d\boldsymbol{h} = \boldsymbol{n}_2 + r\boldsymbol{v}_2. \tag{5.4}$$

Substituting $\boldsymbol{h}_1$ from equation 5.1 into 5.4 yields

$$r\boldsymbol{v}_1 + d\boldsymbol{h} = \boldsymbol{n}_2 - \boldsymbol{n}_1 + r\boldsymbol{v}_2. \tag{5.5}$$

From equation 5.5 it follows that

$$|r\boldsymbol{v}_1 + d\boldsymbol{h}|^2 = |\boldsymbol{n}_2 - \boldsymbol{n}_1 + r\boldsymbol{v}_2|^2. \tag{5.6}$$

Both $\boldsymbol{v}_1$ and $\boldsymbol{v}_2$ are unit vectors and $\boldsymbol{v}_1 \cdot d\boldsymbol{h} = 0$ since these vectors are perpendicular to each other[86]. Solving equation 5.6 for r therefore yields

$$r = \frac{|d\boldsymbol{h}|^2 - |\boldsymbol{n}_2 - \boldsymbol{n}_1|^2}{2\,|(\boldsymbol{n}_2 - \boldsymbol{n}_1) \cdot \boldsymbol{v}_2|}. \tag{5.7}$$

$\boldsymbol{v}_1$ and $\boldsymbol{v}_2$ are both perpendicular to the axis and therefore lie in parallel planes[86]. This means that the projection of $(\boldsymbol{n}_2 - \boldsymbol{n}_1)$ onto the axis equals d or

$$d = (\boldsymbol{n}_2 - \boldsymbol{n}_1) \cdot \boldsymbol{h}. \tag{5.8}$$

By substituting d from equation 5.8 into equation 5.7 and then substituting the resulting r into equations 5.1 and 5.2

$$\boldsymbol{h}_1 = \boldsymbol{n}_1 + \frac{|((\boldsymbol{n}_2 - \boldsymbol{n}_1 \cdot \boldsymbol{h})\,\boldsymbol{h})|^2 - |\boldsymbol{n}_2 - \boldsymbol{n}_1|^2}{2\,|(\boldsymbol{n}_2 - \boldsymbol{n}_1) \cdot \boldsymbol{v}_2|}\boldsymbol{v}_1$$

and a comparable equation for $\boldsymbol{h}_2$ are obtained[86].

By taking the N atoms of the second and third amino acids of the helix to define $\boldsymbol{n}_1$ and $\boldsymbol{n}_2$ and so on to the other end, one obtains two points on the axis for each pair of amino acids. By doing the same for both the α-carbons and the carbon atoms of the carboxyl groups a further four points are obtained for each pair of amino acids.

5.2. Least squares line

The first step in fitting a line through the points is finding the centroid of the points and translating all the points such that the centroid of the points lies on the origin. This means the fitted line will go through the origin and that there exists a set of two rotations which will map one of the axis to the regression line through the points[87]. Choosing to rotate first by θ around the x-axis and then by ϕ around the z-axis gives the rotation

matrix

$$\boldsymbol{R} = \begin{pmatrix} \cos(\phi) & \cos(\theta)\sin(\phi) & \sin(\theta)\sin(\phi) \\ -\sin(\phi) & \cos(\theta)\cos(\theta) & \sin(\theta)\cos(\phi) \\ 0 & -\sin(\theta) & \cos(\theta) \end{pmatrix}. \tag{5.9}$$

Taking a point p_i' that before the transformation laid on the z-axis and applying the rotation matrix from equation 5.9 to it gives

$$\begin{pmatrix} x_i' \\ y_i' \\ z_i' \end{pmatrix} = \boldsymbol{R} \begin{pmatrix} 0 \\ 0 \\ l_i \end{pmatrix}$$

where l_i is the distance from the origin to p_i' and x_i', y_i' and z_i' are the coordinates of the new point. By choosing the right θ and ϕ the resulting point lies on the fitted line. The distance between this point and a corresponding data point $p_i = (x_i,\ y_i,\ z_i)$ is then

$$d_i = \sqrt{(x_i - x_i')^2 + (y_i - y_i')^2 + (z_i - z_i')^2}$$

which leads to

$$d_i^2 = \left(x_i - R_1 \begin{pmatrix} 0 \\ 0 \\ l_i \end{pmatrix} \right)^2 + \left(y_i - R_2 \begin{pmatrix} 0 \\ 0 \\ l_i \end{pmatrix} \right)^2 + \left(z_i - R_3 \begin{pmatrix} 0 \\ 0 \\ l_i \end{pmatrix} \right)^2$$

where R_1, R_2 and R_3 are the first, second and third rows of $\boldsymbol{R}$ respectively.

The sum of the squared deviations is then $S = S_x + S_y + S_z$ where

$$S_x = \sum x_i^2 - 2\left(\sum x_i l_i\right)\sin(\theta)\sin(\phi) + \left(\sum l_i^2\right)\sin^2(\theta)\sin^2(\phi),$$

$$S_y = \sum y_i^2 - 2\left(\sum y_i l_i\right)\sin(\theta)\cos(\phi) + \left(\sum l_i^2\right)\sin^2(\theta)\cos^2(\phi)$$

and

$$S_z = \sum z_i^2 - 2\left(\sum z_i l_i\right)\cos(\theta) + \left(\sum l_i^2\right)\cos^2(\theta).$$

Minimising S gives the least square fit[87]. This is done by differentiating S with respect to ϕ and θ, which gives

$$\frac{\mathrm{d}S}{\mathrm{d}\phi} = -2\sin(\theta)\cos(\phi)\sum x_i l_i + 2\sin(\theta)\sin(\phi)\sum y_i l_i \tag{5.10}$$

and

$$\frac{\mathrm{d}S}{\mathrm{d}\theta} = -2\cos(\theta)\sin(\phi)\sum x_i l_i - 2\cos(\theta)\cos(\phi)\sum y_i l_i + 2\sin(\theta)\sum z_i l_i \tag{5.11}$$

respectively, and setting the results equal zero. Then equation 5.10 becomes

$$\cos(\phi)\sum x_i l_i = \sin(\phi)\sum y_i l_i$$
$$\tan(\phi) = \frac{\sum x_i l_i}{\sum y_i l_i} \tag{5.12}$$

and equation 5.11 becomes

$$\sin(\theta)\sum z_i l_i = \cos(\theta)\left(\sin(\phi)\sum x_i l_i + \cos(\phi)\sum y_i l_i\right)$$
$$\tan(\theta) = \frac{\sin(\phi)\sum x_i l_i + \cos(\phi)\sum y_i l_i}{\sum z_i l_i}. \tag{5.13}$$

ϕ and θ can then be calculated from equations 5.12 and 5.13. The direction cosines (α_x, α_y and α_z) of the fitted line can be calculated by taking the unit vector along the z-axis and applying the rotation matrix in equation 5.9 on it[87]. The direction cosines and the original centroid of the points are enough to define the fitted line. It is possible to calculate the direction cosines of the fitted line directly from the sums without explicitly calculating ϕ and θ using equations 5.14, 5.15 and 5.16.

$$\alpha_x = \frac{\sum x_i l_i}{\sqrt{\left(\sum x_i l_i\right)^2 + \left(\sum y_i l_i\right)^2 + \left(\sum z_i l_i\right)^2}} \tag{5.14}$$

$$\alpha_y = \frac{\sum y_i l_i}{\sqrt{\left(\sum x_i l_i\right)^2 + \left(\sum y_i l_i\right)^2 + \left(\sum z_i l_i\right)^2}} \tag{5.15}$$

$$\alpha_z = \frac{\sum z_i l_i}{\sqrt{\left(\sum x_i l_i\right)^2 + \left(\sum y_i l_i\right)^2 + \left(\sum z_i l_i\right)^2}} \tag{5.16}$$

How equations 5.14, 5.15 and 5.16 are derived is shown in appendix B.

5.3. Calculating the angle

Since the surface of the solution lies in the xy-plane, only α_z is relevant to calculating the angle between the helix and the surface (ϕ_h). The absolute value of α_z is used which means the resulting angle will be between 0° and 90°. α_z is the cosine of the angle between the helix and the z-axis which means that

$$\phi_h = 90° - \cos^{-1}(\alpha_z).$$

The possible orientations of a line in space span a sphere. Restricting this to positive angles it is only necessary to consider a semi-sphere. When averaging the angle it is on the other hand necessary to correct for the fact that in a uniform distribution of the helix-axis, not all ϕ_h are equally likely. Therefore the average of ϕ_h is weighted

with $\cos(\phi_h)$

$$\bar{\phi}_h = \frac{\sum \phi_h \cos(\phi_h)}{\sum \cos(\phi_h)}. \tag{5.17}$$

6. Micelles and microemulsions

6.1. Surfactants

Surfactant is a word formed from the term "surface active agent"[23,88]. Surfactants are molecules that adsorb to surfaces and lower the surface tension[23,25]. Typical surfactants are amphiphilic molecules with a hydrophilic headgroup and a hydrophobic tail[23,25]. The hydrophobic tail often consist of hydrocarbon chains or aromatic rings. Both polar and ionic headgroups are common[24,88].

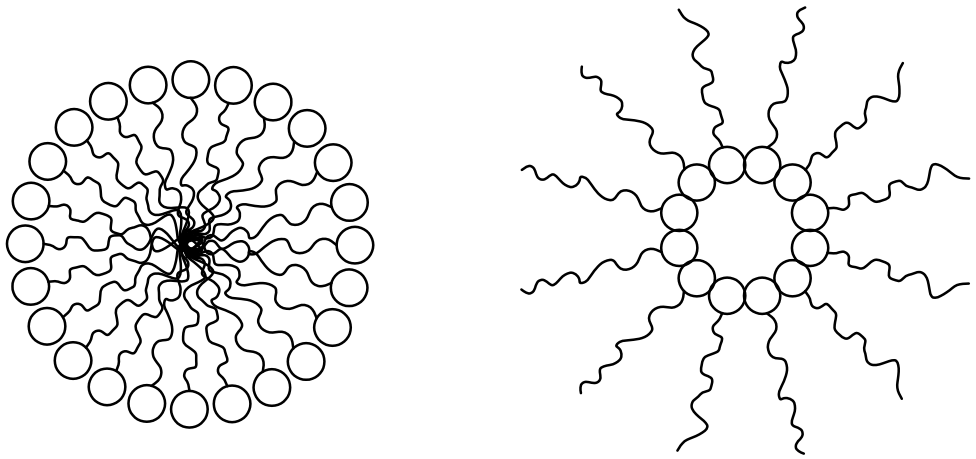
The amphiphilic nature of surfactants influences the solvation behaviour of the molecule in both polar and apolar solvents[24]. The hydrophobic tail is better soluble in apolar solvents and as such facilitates solvation of the surfactant in them while the hydrophilic headgroup is better soluble in polar solvents and facilitates solvation there[23,24]. At the same time both parts of the molecule decrease solubility in their non-prefered solvent type[24]. At polar/apolar surfaces both parts of the molecule can be in their preferred environment simultaneously[23,24,89].

6.2. Micelles

Surfactants tend to form aggregates in aqueous solutions where the hydrophobic tails cluster together and form a hydrophobic core with the hydrophilic headgroups pointing out towards the water (see figure 6.1a)[23–25,90]. Similar structures form in apolar solvents, with the headgroups clustered in the centre while the hydrophobic tails face the solvent[23,25,90]. Those are called inverted micelles (see figure 6.1b)[23,90]. Inverted micelles are typically smaller than normal micelles[24].

Surfactants which form micelles have a characteristic minimum surfactant concentration needed for micelles to form[23,27,89,90]. This concentration is called the critical micelle concentration (CMC). The CMC is associated with a break in the relationship between the surfactant concentration and many properties of the solution, such as the surface tension, the osmotic pressure and the equivalent conductivity.[23–25,89]

The simplest form of a micelle is a spherical micelle. The size of a spherical micelle is more dependent on the surfactant which forms it than on the concentration of said surfactant in the solution[89,90]. The shape and size of the headgroup in relation to the length of the tail influences the curvature of the micelle and so the size of the micelle[27,90]. In a normal micelle the headgroups shield the apolar tails from the polar solvent[23,24]. As such a surfactant with a large headgroup and a short tail will form micelles with fewer surfactant molecules than the micelles a surfactant with a longer tail or a smaller headgroup forms[90]. In an inverted micelle on the other hand long or branched tails



(a) A normal micelle found in water.

(b) An inverted micelle found in oil.

Figure 6.1.: A schematic representations of micelles. The circles are the hydrophilic headgroups. Images by Guðbrandur Guðmundsson.

and small headgroups make it easier to shield the headgroups from the apolar solvent[90]. Therefore surfactants differ both in the size of the micelles they form and in whether they form normal or inverted micelles more easily.

For some surfactants, shapes other than spheres are possible and at high surfactant concentration the micelles can grow[89]. Such micelles have different shapes depending on the shape of the surfactant molecule and the curvature which that shape induces in a surfactant layer[27,90]. Micelles with high curvature tend to grow into cylinders while lower curvature leads to a bilayer[24]. In a bilayer the surfactant molecules form two layers with the part of the molecule which is better soluble pointing out towards the solvent and the less soluble part pointing towards the other surfactant layer[89,90].

The packing parameter of a surfactant is a way of calculating which types of micelles a given surfactant will form from the headgroup and the tail[90]. The packing parameter is defined as

$$PAK = \frac{v}{al} \tag{6.1}$$

where v is the volume and l the length of the surfactant tail and a is the surface area the headgroup fills on the surface of the micelle[26,91]. The preferred micelle structure for surfactants with packing parameter smaller than $\frac{1}{3}$ is spherical. With increasing $\frac{v}{al}$, $\frac{1}{3} < \frac{v}{al} < \frac{1}{2}$, this preferred structure changes to elongated, cylindrical or rod-like, and for $\frac{1}{2} < \frac{v}{al} < 1$ it changes to various interconnected structures. At $\frac{v}{al} \approx 1$ the preferred structure is an extended bilayer and surfactants with even larger packing parameters prefer inverted structures[90,91].

Adding a so-called co-surfactant to the solution can change the shape and the size of

the aggregates that form[89]. Co-surfactants are typically short chain alcohols[25,89]. Like surfactants, co-surfactants have a more polar and a less polar part. Since the size and the shape of the co-surfactant is different from the surfactant, integrating a co-surfactant into the micelle changes the curvature of the surface of the micelle[25].

Micelles, both normal and inverted, can be swollen, with either oil inside a micelle or water inside an inverted micelle[23,90]. This is especially common in inverted micelles[25].

Extended bilayers and liquid crystals

At very high surfactant concentration the micelles are arranged in a regular 3D pattern over large distances in the solution[25,89]. Such orderings are called liquid crystals or mesophases[25]. Liquid crystals come in different forms[89]. Some examples are cylindrical micelles that are arranged parallel to each others[90], lamellar phases where multiple layers of surfactant molecules are stacked on top of each other[89] and hexagonal phases where cylindrical rods are arranged in a hexagonal pattern.[25,89]

A surfactant can also aggregate into vesicles which are composed of a (spherical) bilayer of surfactant molecules and water inside the centre[23,90]. Such vehicles can also join together in a regular structure to form even more different liquid crystal structures[89,90]. Another example of a bilayer is the lipid bilayer of a cell membrane[89,90].

6.3. Microemulsions

Microemulsions are thermodynamically stable and as such form spontaneously upon simple mixing of the components[23,89]. A microemulsion consists of two immiscible solvents, water and oil, and at least one surfactant and/or co-surfactant[23,27]. There are three different types of microemulsions: water in oil, oil in water and bicontinuous[89]. In all three there are water-rich and oil-rich domains in the solution and surfactant molecules at the interface between these domains[27,89]. At the macroscopic scale microemulsions are homogeneous and transparent[23].

Microemulsions only form when the surface tension of the interface between the oil and the water is so low that it is almost nonexistent[23] (below about 10^{-2} mN/m)[27]. Microemulsion droplets are approximately between 1 nm and 100 nm in diameter[92]. The maximum possible amount of dispersed phase inside a droplet depends on the elasticity of the surfactant layer[27]. When the elastic modulus of the surfactant layer is comparable to kT the system can transition smoothly from a water in oil microemulsion through a bicontinuous one to an oil in water microemulsion[27] while for a larger elastic modulus the transition from a water in oil microemulsion to an oil in water microemulsion goes through a lamellar phase[27].

The most common shape of microemulsion droplets is spherical, similar to a swollen micelle, only larger[27,89]. Other shapes such as cylindrical rods are also possible[27]. In bicontinuous microemulsions both domains are continuous but different microemulsions differ in how branched the domains are[90]. The interplay between the concentrations of the different components of the solution and the shape of the surfactant is the deciding

factor in what form a microemulsion takes.[23,89]

Regular structures such as lamellar phases with oil and water in-between the different layers of surfactant molecules or spherical droplets on lattice points with more or less rod-like connections between them correspond to swollen liquid crystals[25]. Lamellar or hexagonal phases are anisotropic in contrast to microemulsions which are isotropic, even when bicontinuous[25].

Methods to characterise microemulsions include measurements of different scattering behaviour, electrical conductivity and viscosity[23]. Scattering measurements can be used to determine the size of the microemulsion droplets or the dimensions of the different domains (periodicity) in a bicontinuous microemulsion[23,93]. Changes in the viscosity can indicate a transition to or from a bicontinuous structure since the interconnected structure of a bicontinuous system leads to a higher viscosity[89]. It is possible to distinguish between a water in oil and an oil in water system by measuring the electrical conductivity since water is a far better conductor than oil[23]. This method distinguishes better between a water in oil microemulsion and a bicontinuous one than between a bicontinuous microemulsion and an oil in water one[23].

6.4. LKα14 in water/ethanol mixtures

Water and ethanol are miscible in all concentrations. Ethanol is however an amphiphilic molecule which can solvate various organic molecules which are insoluble in water[94]. It has been shown that mixtures of water, ethanol and some molecules less hydrophilic than ethanol (for example octanol[53,55] or eugenol[95]) form microemulsion-like systems at the right compositions. In these systems ethanol acts similar to surfactants in that it is present in elevated concentration at the interface between the water-rich and the waterless domains but in contrast to surfactants it is also present in considerable concentration throughout both domains[54].

As mentioned in chapter 4, LKα14 forms oligomers in water. These oligomers tend to be tetramers[58,83]. In the simulations presented in this work such oligomers were seen when the ethanol concentration was low and in pure ethanol. In mixtures with high ethanol concentration larger clusters were formed. These clusters involved both the peptides and water. The solution forms ethanol-rich and water-rich domains with the peptides surrounding the water-rich domain. The water-rich domain has nearly no ethanol at all but water is present in the ethanol-rich domain. (A more detailed description of these clusters is found in chapter 11.)

6.4.1. Defining the clusters

The amount of water trapped inside the clusters is an important characteristic of them and it is therefore necessary to be able to calculate the number of water molecules in the clusters. The clusters analysed in chapter 11 are defined by the absence of ethanol and the presence of at least one peptide. The following method was used to assign the molecules to the clusters.

First the box is overlayed with an equally spaced three dimensional grid, with a spacing of 0.05 nm. For each point the adjacent points in x, y and z direction, as well as the next points on the diagonals are considered neighbours. This means that each grid point has 26 neighbours. No weighing of the different neighbours is used.

At each grid point the distance to the nearest atom centre in an ethanol molecule is calculated. Points which are further than 0.25 nm away from an ethanol molecule are considered as potential cluster points. Then the potential cluster points are sorted into potential clusters, where neighbouring points need to belong to the same cluster.

The next step is to find the closest grid point for each peptide atom. If that grid point is part of a potential cluster the molecule as a whole is counted as part of that cluster. Should the molecule be counted as part of more than one cluster those clusters are merged. The same is then repeated for each water molecule. All potential clusters which do not contain any peptide molecules are then discarded. The result is a list of clusters and for each cluster a list of both grid points and molecules which belong to that cluster.

6.4.2. Analysing the clusters

In addition to the number of clusters and the number of molecules in the clusters, two further properties of the clusters were analysed: Whether the lysine or the leucine side of the peptides is oriented toward the ethanol-rich domain and in how many directions the clusters percolate.

For each cluster all grid points are divided into surface grid points and inner grid points depending on how many of their neighbours belong to the same cluster. If 25 or fewer of the neighbouring grid points also belong to the same cluster a grid point is considered to be a surface grid point. For the surface grid points the distance to the nearest atom which is part of a lysine residue is calculated. The same is done for the distance to the nearest atom which is part of a leucine residue. Then the number of surface grid points closer to a lysine residue than a leucine residue is counted.

For each cluster a path which traverses the whole length of the box and connects to itself across the periodic boundary is sought in all three dimensions. Then the number of dimensions in which such a path can be found is counted.

7. MD

In molecular dynamic (MD) simulations classical mechanics are used to simulate the behaviour of atoms and molecules in time. The forces acting on the atoms are calculated using force fields describing the interactions between the different atoms and from there the trajectory of each atom is obtained.

7.1. Force fields

The force fields used for the simulations in this work are Amber ff03w[96] for the peptide, GAFF[97] for the ethanol and TIP4P/2005[98] for the water.

7.1.1. Amber

Amber is a family of force fields for proteins[99,100]. The Amber ff94 force field was developed in 1994 by Cornell et al[101]. The charges of the atoms were fitted to quantum mechanical calculations in the gas phase, while bond lengths, angles and force constants for both were fitted to reproduce experimental vibrational frequencies[101]. Amber ff94 has been found to result in over stabilisation of α-helices[99,100,102,103].

In 2003 Duan et al. developed an improved version of the Amber ff94 force field called Amber ff03[104]. The difference between ff94 and ff03 is in the charges and the torsion parameters of the protein backbone which were fitted to new quantum mechanical calculations using an continuous solvent model instead of the gas phase[104]. Amber ff03 has been found to over stabilise α-helices[99,105] but also to under stabilise α-helices in very helical peptides[102].

The exact Amber version used in this work is Amber ff03w which was developed in 2010 by Best and Mittal by tuning the torsion parameters of the protein backbone for use with the TIP4P/2005 water model[96]. Amber ff03w describes helix formation better than ff03[96] but still slightly over stabilises α-helices at low temperatures (such as 300 K)[106].

7.1.2. GAFF

GAFF stands for General Amber force field[97]. It was developed in 2004 by Wang et al to be compatible with Amber protein force fields and includes parameters for organic molecules[97]. The force field parameters used for ethanol in this work were generated from GAFF with AnteChamber[97] and converted to GROMACS compatible form with ACPYPE[107].

7.1.3. TIP4P/2005

TIP4P/2005 is a water model developed in 2005 by Abascal and Vega[98]. It has rigid bonds and angle. There are positive charges at the hydrogen atoms but the corresponding negative charge is located at an interaction site along the bisector of the angle[98]. TIP4P/2005 has been shown to reproduce the phase diagram of water quite well qualitatively though both the melting and the boiling points are shifted[98]. It has also been shown to reproduce many thermodynamic properties of water (such as surface tension, density, self-diffusion coefficient and shear viscosity) quite well.

7.2. Simulations details

All simulations were performed with the GROMACS suite of programs (version 5 and later)[108,109]. All nonpolar hydrogen atoms were described as virtual sites and all remaining bond lengths were constrained using the LINCS algorithm[110]. The long-range electrostatic interactions were calculated using smooth particle-mesh Ewald (PME) summation[111] with a grid spacing of 0.12 nm and fourth order interpolation. The Lennard-Jones interactions were smoothly cut off at a distance of 1.0 nm.

In all simulations the LYS residues were charged and chloride ions used for charge neutralisation.

7.2.1. Simulations containing an air/liquid interface

In the simulations containing an air/liquid interface, a 10 nm × 10 nm × 10 nm box of solution containing 20 peptides was equilibrated at 300 K and 1 bar using a stochastic velocity-rescaling thermostat[112] and a Parrinello-Rahman barostat[113]. The equilibrated box was then inserted into the middle of a $10\,\mathrm{nm} \times 10\,\mathrm{nm} \times 30\,\mathrm{nm}$ or $10\,\mathrm{nm} \times 10\,\mathrm{nm} \times 40\,\mathrm{nm}$ simulation box with vacuum on either side. This is a reasonable approximation of an air/liquid interface.

The simulation of this larger simulation box was performed at 300 K using a stochastic velocity-rescaling thermostat[112]. The time step of the simulation was 2.5 fs. The PME calculation was done using a correction for slab geometry[114]. Twelve different compositions were simulated and for each composition two different simulations were run, one where the starting conformation of the peptides was an α-helix and one where it was random coil. The simulations starting from an α-helix were run for 900 ns and the simulations starting from random coil for 450 ns. These simulations were used for the analysis of the behaviour of the peptide at the surface.

To compare the surface tension with and without peptides the peptides were removed from the simulation boxes from the simulations starting from an α-helical conformation. The resulting boxes were then simulated for 45 ns using the same conditions as for the simulations with the peptides.

7.2.2. Bulk simulations with one peptide

Two sets of bulk simulation of a single peptide in a $5\text{nm} \times 5\text{nm} \times 5\text{nm}$ simulation box were performed. Both were performed at 300 K and 1 bar using a stochastic velocity-rescaling thermostat[112] and a Parrinello-Rahman barostat[113] both with a time constant of 1 ps. The time step was 5 fs in all cases.

In the first set of single peptide bulk simulations 11 different solvent compositions were used and the simulation time was 900 ns. All peptides started from an α-helical conformation. This set of simulations was used for the analysis of the solvation shells.

In the second set of single peptide bulk simulations, 12 different solvent compositions were used and the simulation time was 10 μs. Two simulations were performed for each solvent composition, one starting from an α-helical conformation of the peptides and one starting from a random coil conformation. This set of simulations was used for the analysis of the secondary structure in bulk.

7.2.3. Bulk simulations with 20 peptides

Bulk simulations with 20 peptides in a$10\,\text{nm} \times 10\,\text{nm} \times 10\,\text{nm}$ simulation box were also performed. As with the bulk simulations with a single peptide they were performed at 300 K and 1 bar using a stochastic velocity-rescaling thermostat[112] and a Parrinello-Rahman barostat[113] both with a time constant of 1 ps. The time step was 5 fs, the simulation time 1000 ns and 16 different solvent compositions were used. All peptides started from an α-helical conformation. These simulations were used for the analysis of the clusters formed by the peptides.

Part III.

Results

8. Secondary structure

The first property of the peptide that was analysed was the dependency the secondary structure on the composition of the solvent mixture. Since it is known that in water LKα14 assumes different secondary structures in bulk and at the surface[58,83] this was done both for simulations of bulk solution and simulations with a solution/air interface.

8.1. Bulk

The analyses of the secondary structure in bulk was done from simulations of one single peptide in bulk. Two simulations were performed for each composition of the solvent mixture. To prepare for the first one an α-helical structure was simulated for 0.6 ns at 900 K. The resulting structure was then simulated for 10 μs at 300 K. To ascertain that the resulting secondary structure had in fact relaxed into the thermal equilibrium a second simulation was then run, also for 10 μs at 300 K but starting from a random coil conformation for all $x_{\mathrm{Eth}} > 0$ and an α-helical conformation for $x_{\mathrm{Eth}} = 0$. The secondary structure was considered to be converged when both simulations reliably showed a distribution of secondary structures consistent with each other. The secondary structure is stable enough that only after 5 μs could this be considered true for all x_{Eth}. Therefore the last 5 μs of each simulation were combined for the analyses while the first 5 μs were discarded.

Figure 8.1 shows the secondary structure as function of simulation time for each amino acid in the simulations with $x_{\mathrm{Eth}} = 0$, $x_{\mathrm{Eth}} = 0.6$ and $x_{\mathrm{Eth}} = 0.8$. As can be seen in figure 8.2 and is discussed further down, random coil is the dominant secondary structure of LKα14 in water. But figure 8.1b shows that when the simulation starts from an α-helical conformation the α-helix is stable for about 1 μs. Both figure 8.1a and 8.1b also show periods of time where parts of the peptide form α-helices which last for multiple μs. Figures 8.1c, 8.1d, 8.1e and 8.1f show that in $x_{\mathrm{Eth}} = 0.6$ and $x_{\mathrm{Eth}} = 0.8$ α-helix is very clearly the dominant secondary structure and that once formed an α-helix does not denature in these solutions. But at the same time figures 8.1c and 8.1e show that it can take more than 2 μs for the α-helix to form when starting from an random coil conformation. It can also be seen from figures 8.1c, 8.1d, 8.1e and 8.1f that the two amino acids closest to the N-terminus do not form part of the α-helix. Such plots for the other compositions can be seen in appendix B.

Taken together, how long the α-helices last in pure water and how long they take to form in the mixtures, shows how stable secondary structures can be even when they do not represent the global minimum. This is in accordance with literature as finding the true conformation of proteins from simulations is a considerable challenge[115–119]. The

time scale of the changes here should also be kept in mind in part 8.2.

Figure 8.2 shows the probabilities of different secondary structures in the different solvent compositions averaged over all amino acids in the peptide. It shows clearly that only at $x_{\mathrm{Eth}} = 0$ random coil is the dominant secondary structure. At $x_{\mathrm{Eth}} = 0.033$ the probability of an α-helix is 0.73 and at higher x_{Eth} the probability of an α-helix is even higher. The only composition of the solvent which leads to a significant amount of any secondary structure other than either α-helix or random coil is pure water where the probability of other secondary structures is 0.15. Figures 8.1a and 8.1b show that this is mainly turn.

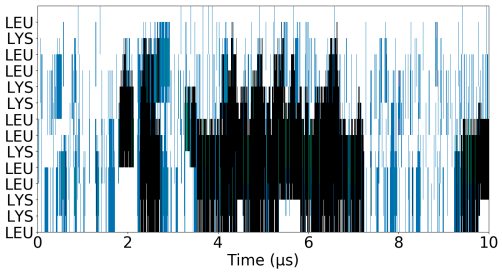
It can be seen from figure 8.2 that the probability of random coil lies between 0.1 and 0.2 for $x_{\mathrm{Eth}} \geq 0.1$ and the probability of α-helix between 0.8 and 0.85 for the same. There is a slight trend towards more α-helix and less random coil the more ethanol there is in the solvent. As mentioned before figure 8.1 shows that the two amino acids closest to the N-terminus do not form part of the helix. As can be seen in appendix B, this holds for other solvent compositions as well. Therefore for $x_{\mathrm{Eth}} \geq 0.033$ the probability of random coil seen in figure 8.2 does not stem from the α-helix denaturing and forming again but rather from the fact that not all amino acids of the peptide are part of the α-helix. In pure water on the other hand, the probability of α-helix does stem from an α-helix spanning at least 5 amino acids forming and disappearing again. This is clear both from figures 8.1a and 8.1b and the fact that, as seen from the definition of an α-helix in section 3.1, the shortest possible α-helix is 5 amino acids long.

8.2. Surface

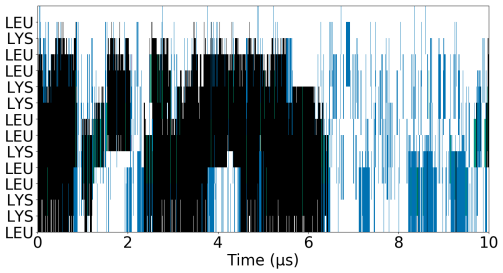
Given that LKα14 was designed to form an α-helix at air/water interfaces[58] and has been shown to do so[58,83], it is of interest whether it shows the same behaviour in simulations. Just as for the secondary structure in bulk, two simulations were made for each solvent composition. In both simulations twenty peptides were simulated in a solvent box with a solvent/air interface. While in one simulation of each composition the peptides started from an α-helical conformation, in the other the peptides all started from the same random coil conformation. The simulations starting from an α-helical conformation were 900 ns long, but the ones starting from a random coil conformation were 450 ns. The first 100 ns of each simulation were considered equilibration and ignored in the secondary structure analyses.

The secondary structure was analysed both for the surface and the bulk in order to determine if there is a difference in secondary structure in these two domains. The surface was determined to be 1.5 nm thick (see section 9.1) and the rest of the solvent box was therefore considered bulk. In contrast to the simulations of single peptide in bulk, the trajectories resulting from the two different starting conformations were analysed separately.

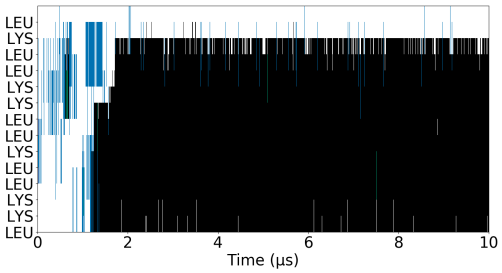
Comparison of figures 8.3 a, b and c on one hand and figures 8.3 d, e and f on the other shows a very clear difference in the secondary structure depending on the starting conformation. This is unsurprising given that when simulating a single peptide in bulk



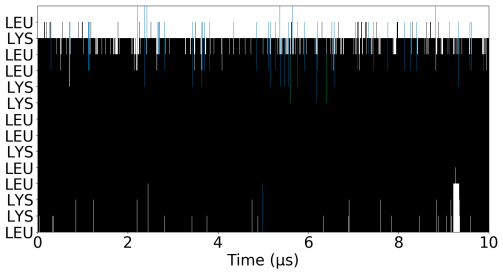
(a) Mole fraction ethanol $x_{\mathrm{Eth}} = 0$ starting from a random coil configuration.



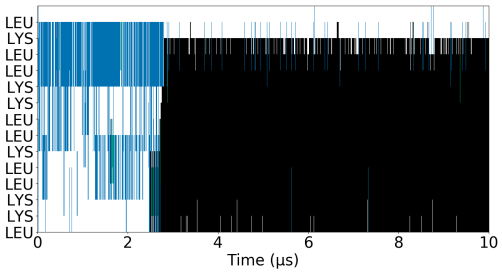
(b) Mole fraction ethanol $x_{\mathrm{Eth}} = 0$ starting from an α-helical configuration.



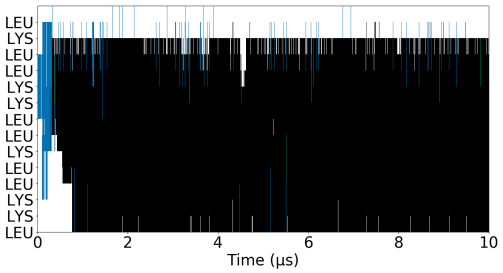
(c) Mole fraction $x_{\mathrm{Eth}} = 0.6$ starting from a random coil configuration.



(d) Mole fraction ethanol $x_{\mathrm{Eth}} = 0.6$ starting from an α-helical configuration.



(e) Mole fraction ethanol $x_{\mathrm{Eth}} = 0.8$ starting from a random coil configuration.



(f) Mole fraction ethanol $x_{\mathrm{Eth}} = 0.8$ starting from an α-helical configuration.

Figure 8.1.: Secondary structure of LKα14 as a function of simulation time for each amino acid. Black is α-helix, white random coil, blue turn, green 3_{10}-helix and pink isolated bridge.

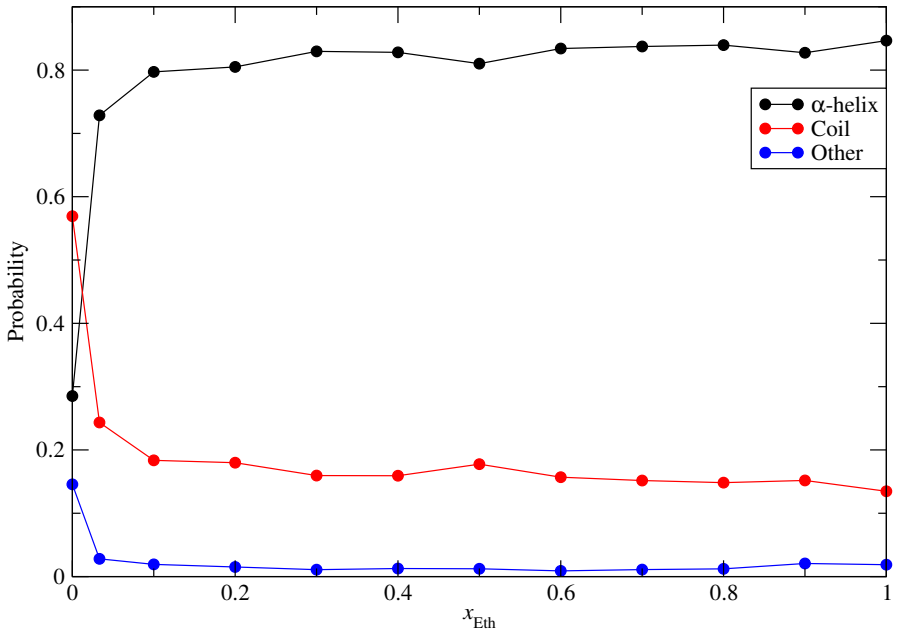


Figure 8.2.: Secondary structure of a single LKα14 in bulk as a function of the ethanol mole fraction x_{Eth}. All probabilities were averaged over two simulations, one starting from an α-helical conformation and the other from a random coil. The probabilities are also averaged over all amino acids in the peptide. The lines serve only to guide the eye.

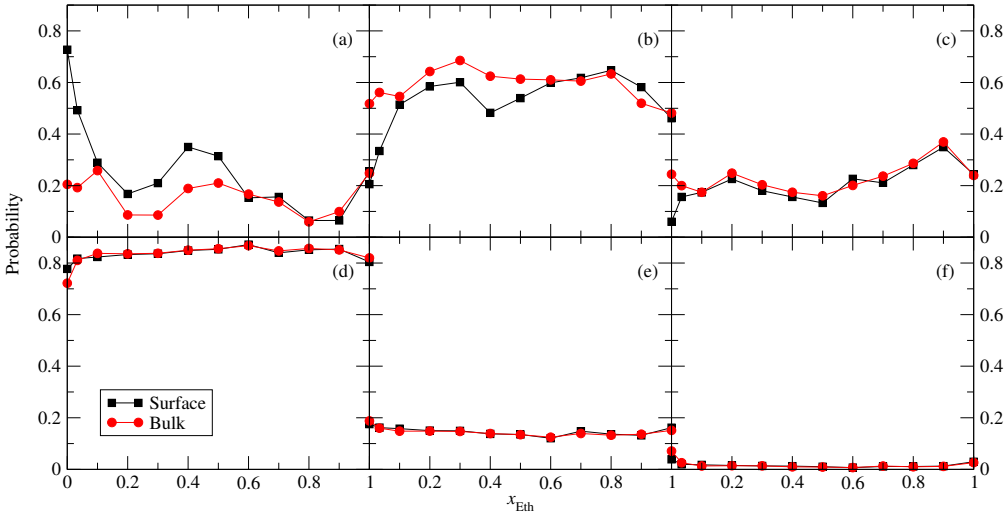


Figure 8.3.: The probability of different secondary structures as function of the ethanol mole fraction, x_{Eth}, at the surface and the in the bulk in a simulation with 20 peptides and solvent/air interface. (a), (b) and (c) started from an random coil conformation, (d), (e) and (f) from an α-helical structure. (a) and (d) show the probability of α-helix, (b) and (e) the probability of random coil and (c) and (f) show the probability of turn. The lines serve only to guide the eye.

the secondary structure could not reliably be considered converged for several μs and these simulations were shorter than 1 μs. Therefore these results can not be considered converged, but they can still give insight into the different folding properties of the peptides at the surface and in the bulk.

As seen in figures 8.3 d, e and f there is only a very slight difference in the secondary structure at the surface and in the bulk when starting from an α-helical conformation of the peptides. Furthermore this difference is only seen in pure water. In all other solvent compositions there is no significant difference seen in secondary structure between the surface and the bulk. This difference between pure water and the other solvent compositions agrees with the fact that as seen in 8.1 pure water is the only composition where random coil is more common in the bulk than an α-helix. For all compositions, except pure water, the probability of an amino acid being part of an α-helix is greater than 0.8 both in bulk and at the surface, the probability of random coil is around 0.15 and the probability of turn is less than 0.03. For pure water the probability of α-helix is 0.78 at the surface and 0.72 in the bulk, the probability of random coil is 0.18 and 0.17 and the probability of turn 0.04 and 0.07 respectively.

Figure 8.3 d shows that the α-helices present at the beginning of the simulations are preserved for the entire simulation, whether the peptides are adsorbed to the surface or not, except for the slight decrease in α-helices prevalence in the bulk of pure water. Since LKα14 was designed to form α-helices at water/air interfaces it is not surprising that the α-helices do not denature at the surface and the secondary structures seen in 8.1 would fit with the α-helices also being stable in the bulk of these simulations (again except for the simulation in pure water).

Figures 8.3 a, b and c show that when starting from a random coil configuration there is a significant difference between the conformation of the peptides at the surface and in the bulk when the ethanol concentration is low. This difference is largest in pure water where the probability of an α-helix is 0.73 at the surface and 0.20 in the bulk. The largest difference in prevalence of both random coil and turn is also in pure water. For $x_{\mathrm{Eth}} \geq 0.6$ there is less difference in the prevalence of the different secondary structures between the surface and the bulk.

Figure 8.3 a shows a general trend toward a declining probability of an α-helix at the surface when x_{Eth} grows while the probability in the bulk shows lesser dependency on x_{Eth}. Figure 8.3 b shows the similar behaviour for the probability of a random coil, though here the probability at the surface rises with growing x_{Eth}. The probability of turn on the other hand shows no significant difference between the surface and the bulk, as can be seen in figure 8.3 c. There is though a weak trend towards higher prevalence of turn with higher x_{Eth}. All these trends though are not very well converged. This is directly linked with how short these simulations were.

8.2.1. Development of the secondary structure with time

It is striking how similar the probabilities of the different secondary structures are at the surface of pure water when starting from the different conformations of the peptides. As mentioned earlier, at the surface of pure water the probability of an α-helix is 0.78, of

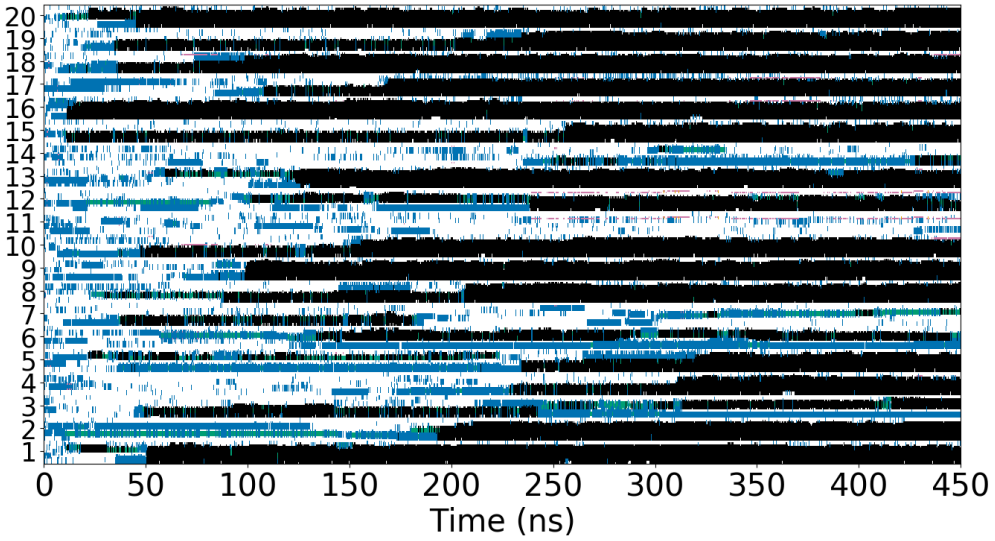
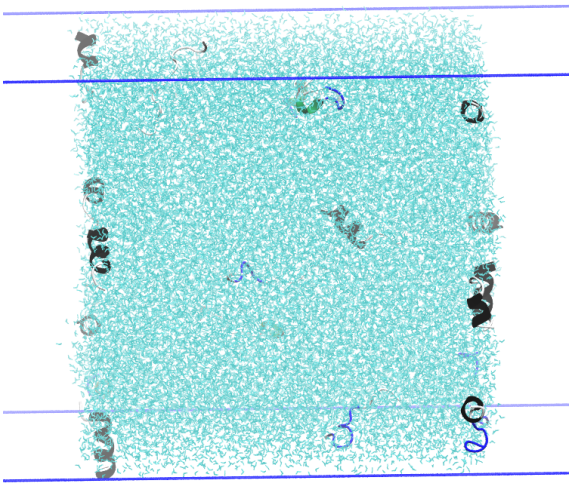


Figure 8.4.: The secondary structure of the peptide in pure water as a function of time starting from a random coil conformation. 20 peptides in a simulation with a solvent/air interface, the numbers on the side correspond to individual peptides. Black is α-helix, white random coil, blue turn, green 3_{10}-helix and pink isolated bridge.

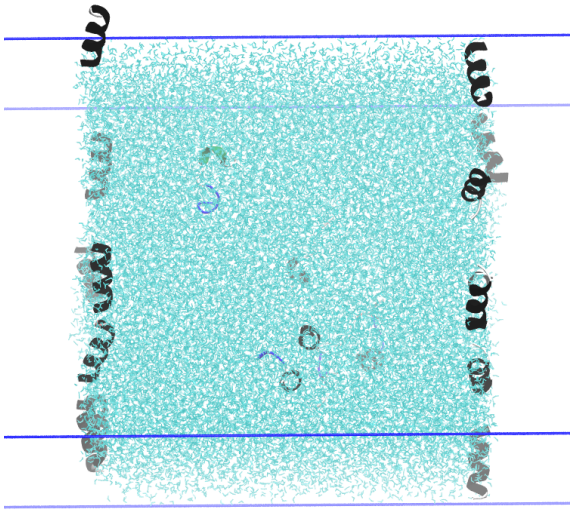
random coil 0.18 and of turn 0.04 when starting from an α-helical conformation. When starting from a random coil conformation the probabilities are 0.73, 0.20 and 0.05 respectively. This is the only combination of solvent composition and spatial domain of the solution where both starting conformations result in similar probabilities, in all other cases there is a stark difference. In all solvent compositions the probability of an α-helix is under 0.5 when starting from a random coil configuration and over 0.8 when starting from an α-helical one, while in the bulk of pure water an α-helical starting conformation leads to a probability of an α-helix of 0.72 in contrast to the probability of 0.20 which results from a random coil starting conformation.

Having 20 peptides in the simulation rather than just one (as in section 8.1) should shorten the simulation time needed to converge the secondary structure prevalence. But since the peptides show different behaviour in the different domains (in the bulk and at the surface) and are not equally distributed between these domains (see chapter 9.2) the extra peptides do not shorten the convergence time equally in both. How great the effect is at the surface and in the bulk respectively also depends on the solvent composition since the adsorption of the peptides to the surface depends on the solvent composition(see chapter 9.2).

Figure 8.4 shows how the secondary structure of the the peptides changes with time in pure water when starting from a random coil conformation. After about 50 ns three peptides have completely folded into α-helices and a further six peptides have a large α-helical part. After about 100 ns five peptides have completely folded into α-helices and



(a) Early in the simulation.



(b) At the end of the simulation.

Figure 8.5.: Snapshots of the simulation in pure water. The blue lines show the boundaries of the simulation box. The water is cyan. The peptides are coloured according to secondary structure, α-helices are black, 3_{10}-helices green, turn blue and random coil white. The peptides at the surface are far more α-helical than the peptides in the bulk.

a further six peptides have a large α-helical part. Figure 8.5a shows a snapshot of the simulation after 100 ns. It can be seen that of the peptides in the bulk only one has any α-helical structure while most of the peptides at the surface are at least partly α-helical. Figure 8.5b shows that after 450 ns all the peptides at the surface are either completely or mostly α-helical while the peptides in the bulk have more diverse secondary structures.

Together figures 8.4 and 8.5 show that LKα14 folds quite quickly into an α-helix at the surface of water. Comparing figure 8.4 to section 8.1 where the secondary structure was only considered reliably converged after 5 μs shows that LKα14 forms an α-helix faster at the surface of water than in the bulk of any of the solvent compositions simulated in this work. As can be seen from figure 8.4 α-helices generally tend to form faster at the surface than in bulk for $x_{\mathrm{Eth}} \leq 0.5$ (see also appendix B.2).

8.3. Discussion

It is known from experiments that LKα14 forms α-helices at an air/water interfaces[58,83] and in oligomers formed at high concentrations[58] but not in dilute bulk[58]. It has also been shown experimentally that in the bulk of pure ethanol the preferred secondary structure of LKα14 is α-helix[120] and that in the bulk of mixtures of water and ethanol the shift from predominantly random coil to predominantly α-helix takes place continuously over the range from $x_{\mathrm{Eth}} = 0$ to $x_{\mathrm{Eth}} = 0.24$[120].

The results from the simulations agree quite well with these experiments though the ethanol mole fraction at which the shift from predominantly random coil to predominantly α-helix takes place is lower in the simulations. As mentioned in chapter 7.1.1 the Amber force field which is used for the peptide in the simulations is known to stabilise α-helices slightly too much[106] which might be part of the reason why the change in secondary structure takes place at a lower x_{Eth}.

Ethanol is know to stabilise α-helices[14,121] though it can also denature the tertinary structure of proteins[13]. It can be concluded, both from the simulations in this chapter and the experiments[120] already mentioned, that LKα14 is no exception to this stabilising effect of ethanol. On the other hand figure 8.3 shows that LKα14 forms an α-helix significantly faster at a water/air interface than either in the bulk or at the surface of pure ethanol. A flat apolar/polar interface is optimal for inducing formation of an α-helix when said α-helix is divided in a hydrophilic and a hydrophobic part along the axis of the helix (see chapters 4 and 9.4). The stabilisation of the α-helix in pure ethanol is of different origin and not depended on the primary structure of the peptide in the same way. The slower formation of the α-helices also indicates that ethanol might not be as strongly stabilising for an α-helix in LKα14 as a water/air interface is. LKα14 does not adsorb to the surface of pure ethanol (see chapter 9.2). The presence of an ethanol/air interface does therefore not influence the secondary structure and the formation time of an α-helix to a significant degree.

When both water and ethanol are present in the solvent mixture the two solvent can spatially separate to form a solvation shell around the peptide which is not uniformly hydrophilic (see chapter 10.2.1). This stabilises the α-helix in a similar way to the water/air

interface, though judging from figure 8.3 not to the same extent. The trend towards a higher α-helix prevalence with higher ethanol mole fraction seen in figure 8.2 indicates that this effect is smaller than the general stabilising effect of ethanol on α-helices. As mentioned in appendix B.1 the presence of both water and ethanol in the solvent might facilitate a faster completion of the folding into an α-helix once the formation of a helix has started. For some organic solvents proteins have been shown to denature in mixtures of organic solvent and water but not in the neat organic solvent, where the proteins are kinetically prevented from denaturing[122]. While here it is the ethanol which stabilises the α-helices, similar effects might influence the folding time, though the spatial division of the solvent into a hydrophilic and a less hydrophilic region presumably also plays a role.

9. Adsorption and orientation at the surface

LKα14 peptides have been shown, both in experiments and theoretical studies to adsorb onto air/water interfaces[58,84]. Once adsorbed to the interface they have been shown to assume α-helical conformation and an orientation parallel to the interface[58,83,84,123]. The same has been shown at the interfaces of water and other hydrophobic substances[123,124]. It has also been shown that at those same interfaces the leucine side chains are oriented away from the water, while the lysine side chains are oriented towards it[60,123,124].

The results discussed in this chapter are all from the simulations with the interface where the starting conformation of the peptides was α-helical. The first 100 ns were discarded as equilibration and only the next 350 ns were used, as the adsorption and the orientation of the peptides at the surface converges far more quickly than the secondary structure. To analyse the behaviour of the peptides in different mixtures of water and ethanol, the partial densities of all components were evaluated (9.1) as well as both the orientation of the helix-axes (9.3) and of the side chains (9.4). Since the lysine residues are charged while the leucine residues are not the dipole moment was used to determine the orientation of the side chains.

For all comparisons of the surface and the bulk in this chapter the surface was defined as 1.5 nm thick. This definition is based on the partial density of the peptides in pure water and in solutions with low x_{Eth}. As will be shown in this chapter, it is not possible to arrive at a clear definition of the surface in pure ethanol or the solutions with high x_{Eth}. Therefore the definition based on the behaviour of the peptides in pure water and in solutions with low x_{Eth} was used for all solutions.

9.1. Defining the surface

Figure 9.1 shows the partial densities for all components along the z-coordinate for solutions with $x_{\mathrm{Eth}} = 0$, $x_{\mathrm{Eth}} = 0.033$ and $x_{\mathrm{Eth}} = 1$. The plots in figure 9.1 should be symmetrical since there is nothing that in theory makes one interface different from the other. In the time frame of the simulation this is not reached because of the strong adsorption at low x_{Eth} and formation of clusters at high x_{Eth}. In all analyses of the surface region an average over both interfaces was used.

As can be seen in figure 9.1 there are sharp peaks in the partial density of the peptide at the interface for $x_{\mathrm{Eth}} = 0$ and $x_{\mathrm{Eth}} = 0.033$, while the partial density of the peptide rises more slowly and a bit further from the edge in pure ethanol. In both pure water and the solution with $x_{\mathrm{Eth}} = 0.033$ there is a region between the surface and the bulk

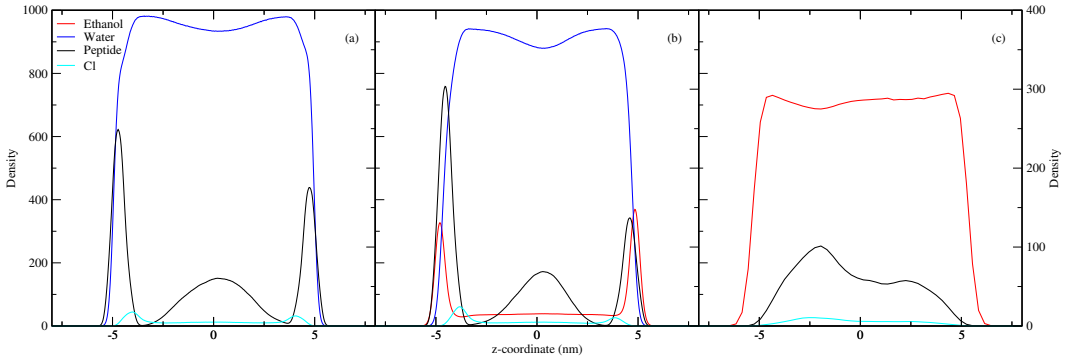


Figure 9.1.: Partial densities of all components along the z-coordinate for (a) $x_{\mathrm{Eth}} = 0$, (b) $x_{\mathrm{Eth}} = 0.033$ and (c) $x_{\mathrm{Eth}} = 1$. Ethanol and water use the scale on the left, the peptide and the ions the scale on the right.

which is depleted of peptides. In pure water the partial density of both water and the peptide starts rising 5.5 nm from the box centre. 5 nm from the centre of the box can be considered the start of the liquid phase, which fits the intention for the simulation. The region between 3 nm and 4 nm from the centre of the box is depleted of peptides. The solution with $x_{\mathrm{Eth}} = 0.033$ shows similar behaviour, though here the depleted region is between 3.5 nm and 4.5 nm from the centre of the box and the bulk region therefore narrower. This seems to be facilitated by adsorption of ethanol to the surface. From both solutions taken together it can be concluded that 1.5 nm is a good estimate of the surface thickness.

In figure 9.1c it can be seen that in pure ethanol the liquid phase starts around 5.3 nm from the centre of the box while the partial density of ethanol starts to rise 6 nm from the centre of the box. The partial density of the peptide starts to rise 5 nm from the centre of the box and reaches the bulk value at about 4 nm from the centre of the box. Right at the edge the peptide partial density is low but it can not be discerned from figure 9.1c whether the bulk should be considered to begin 5 nm or 4 nm from the centre of the box.

When looking at the partial densities of the other solutions (shown in appendix C) it can be seen that the more ethanol there is in the solution the more difficult it is to define where the surface region starts. Therefore the definition that the surface is 1.5 nm thick, based on the partial densities shown in figures 9.1a and 9.1b was used for all solutions.

9.2. Adsorption

Figure 9.2 shows the probability of finding the peptides at the surface as well as in the bulk as a function of x_{Eth}. It can be seen that in pure water the peptide adsorbs to the surface which agrees with previously reported results[58,84].

It can also be seen in figure 9.2 that even at $x_{\mathrm{Eth}} = 0$ the probability of finding the peptides at the surface is lower than of finding them in the bulk. Comparing this to

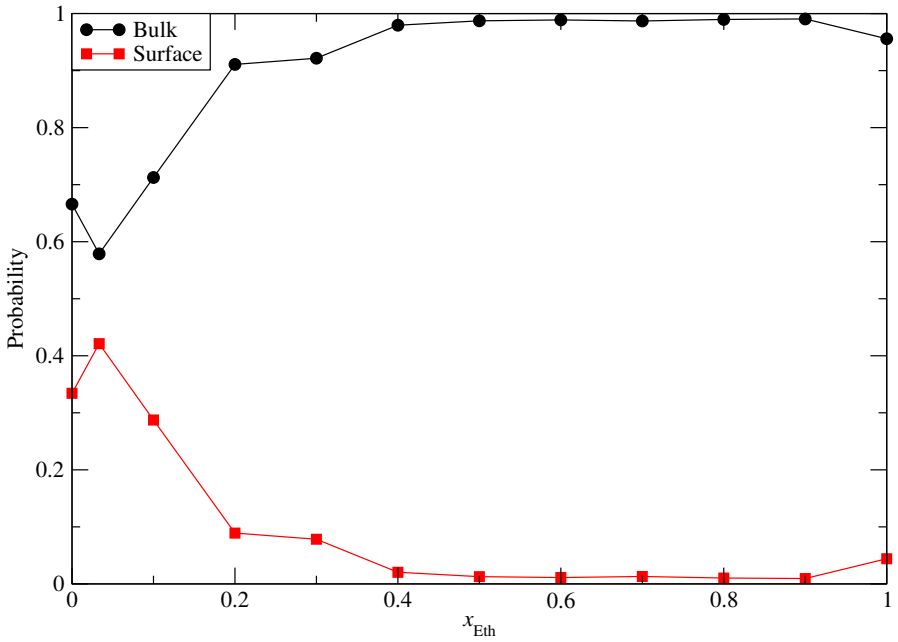


Figure 9.2.: The probability of finding the peptides at the surface and in bulk as a function of x_{Eth}. The black circles show the probability of finding a peptide at the surface and the red squares the probability in the bulk. The lines serve only to guide the eye.

figure 9.1 shows that the concentration is higher at the surface. This discrepancy can be explained by the fact that the bulk is a larger volume than the surface region and is as such not an evidence against adsorption.

Figure 9.2 further shows that the absorption is strongest at $x_{\mathrm{Eth}} = 0.033$ and that from $x_{\mathrm{Eth}} = 0.033$ to $x_{\mathrm{Eth}} = 0.2$ there is a sharp decline in the probability of finding the peptides at the surface. For $x_{\mathrm{Eth}} = 0.2$ and $x_{\mathrm{Eth}} = 0.3$ the probability is around 0.1 though slightly lower for $x_{\mathrm{Eth}} = 0.3$. For x_{Eth} from 0.4 to 0.9 on the other hand all the peptides are in the bulk and the probability of finding a peptide at the surface is 0.02 or lower. The slightly higher prevalence of the peptides at the surface in pure ethanol than in solutions with x_{Eth} between 0.4 and 0.9 is most likely due to a mixture of two things: less presence of clusters in pure ethanol which leads to a more uniform distribution of the peptides in the solution at each point in time (see chapter 11) and the preferential solvation of the peptide by a mixture of water and ethanol (see chapter 10). Similarly the stronger adsorption at $x_{\mathrm{Eth}} = 0.033$ than in pure water is most likely due to the fact that, as can be seen in figure 9.1b), the ethanol also adsorbs to the surface and the adsorbed peptides lie between the adsorbed ethanol and the bulk. At higher x_{Eth} this effect has less influence as there is more ethanol in the bulk.

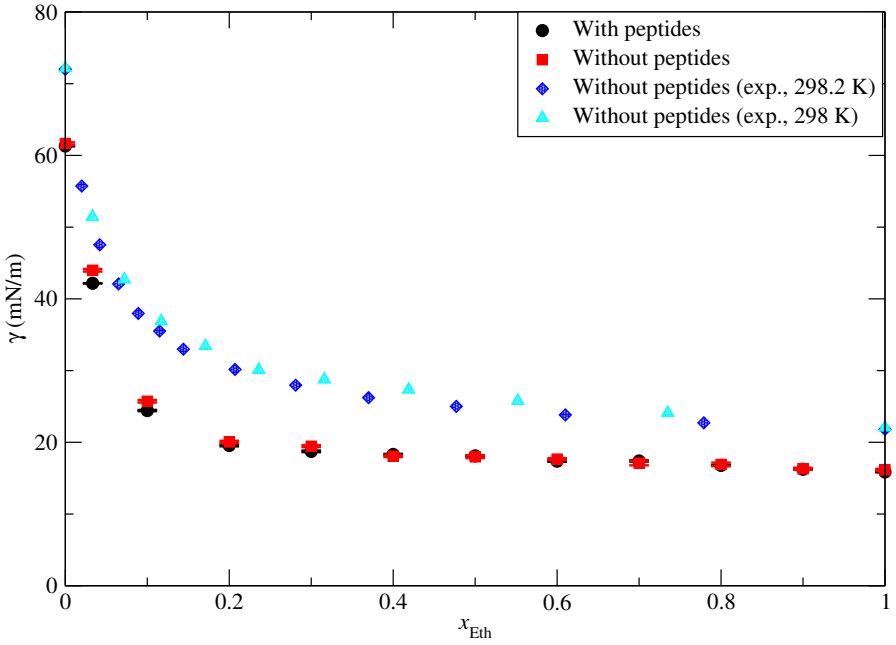


Figure 9.3.: The surface tension, γ, as a function of x_{Eth}. The black circles show the surface tension of a solution with peptides and the red squares the surface tension of a solution without peptides. The blue diamonds[125] and the cyan triangles[126] are two sets of experimental data for surface tension of water/ethanol mixtures at 298 K.

9.2.1. Surface tension

The surface tension of the solution was calculated both with and without peptides. As the surface tension converges faster than the position of the peptides, the simulations used for the calculation of the surface tension without peptides were only 45 ns long and only the first 10 ns were considered equilibration.

Figure 9.3 shows the surface tension of the solution as a function of x_{Eth}. Only two solvent compositions lead to a noticeable difference in surface tension depending on whether or not the peptides are present: $x_{\mathrm{Eth}} = 0.033$ and $x_{\mathrm{Eth}} = 0.1$. For $x_{\mathrm{Eth}} = 0.033$ the difference is 1.8 ± 0.2 mN/m and for $x_{\mathrm{Eth}} = 0.1$ it is 1.3 ± 0.3 mN/m. In pure water the presence of the peptides does also lower the surface tension but only by 0.3 ± 0.2 mN/m which barely counts as significant. For $x_{\mathrm{Eth}} = 0.2$ and $x_{\mathrm{Eth}} = 0.3$ the difference is 0.5 ± 0.3 mN/m and 0.7 ± 0.2 mN/m respectively. For higher x_{Eth} the uncertainty in the surface tension difference is larger than the value itself. Comparing these differences with figure 9.2 shows that when the peptides adsorb to the surface the surface tension is lowered, though neither the correlation nor the effect itself is very strong. The difference in surface tension depending on whether LKα14 is present in the solution has been measured to be larger in pure water (around 20 mN/m) and at low x_{Eth} (around 15 mN/m at $x_{\mathrm{Eth}} = 0.1$) than was calculated in these simulations[120].

Figure 9.3 also shows that the surface tension, of the pure solvent mixture, calculated from the simulations is always lower than the surface tension measured in experiments[125,126]. The surface tension calculated here for pure water is 61.7 ± 0.2 mN/m, which is lower than expected surface tension of tip4p2005 at 300 K (69.3 ± 0.9 mN/m)[127], which it self is slightly lower than the experimental value of 71.73 mN/m[127]. For ethanol the difference from the experimental value is smaller than for water but still significant. Qualitatively the steep descent in the surface tension at low x_{Eth} followed by a slower descent at $x_{\mathrm{Eth}} > 0.2$ agrees well with the experimental values, though in the simulations the slope of the curve is greater at low x_{Eth} and smaller at high x_{Eth} than in the experiments.

9.3. Orientation of helix-axis

As mentioned in chapter 8.2 in the simulations with the interface the simulation time was not long enough for the secondary structure to significantly change during the simulation. When starting from a helical conformation of the peptides, this conformation was therefore kept.

The angle between the helix axis and the surface plane (ϕ_{h}) was analysed for all simulations which were performed with an interface using the method described in chapter 5. When comparing the bulk and the surface region the definition of the surface region from chapter 9.1 was used. As will be shown here the orientation of the helices supports this definition.

The formation of clusters at higher mole fractures of ethanol, mentioned in chapter 9.1 and further discussed in chapter 11, complicated the analyses for those cases, especially

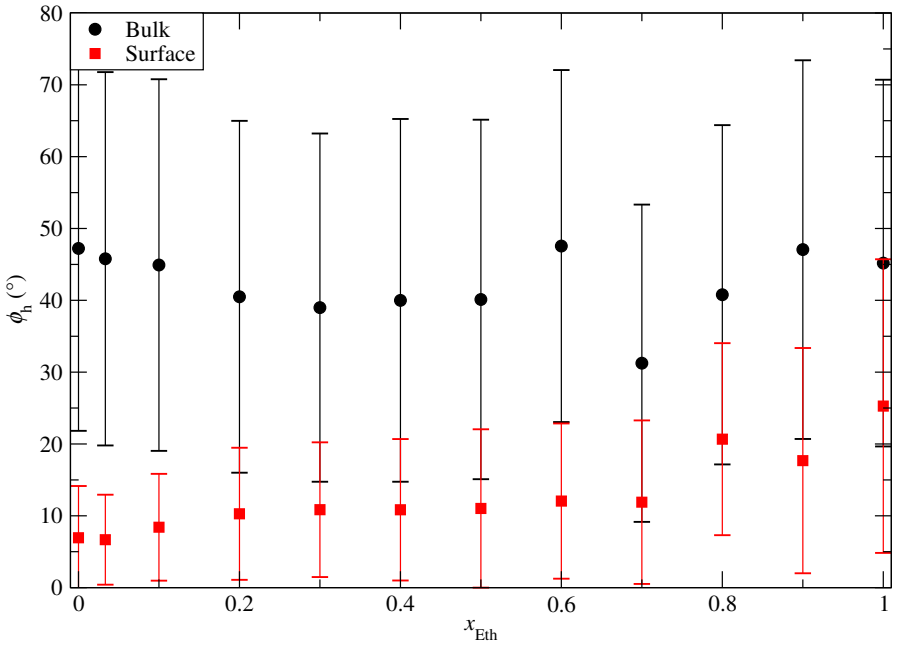


Figure 9.4.: The average angle, $\bar{\phi}_h$, between the surface plane and the helix as a function of the mole fraction of ethanol. The red squares are at the surface and the black circles in the bulk.

for the population of the peptides in the bulk.

As can be seen in figure 9.4, there is a clear difference in the orientation of the peptide at the surface or in the bulk. This difference is greater at lower x_{Eth}. In pure water, $\bar{\phi}_h$ is 6.9° which agrees well with previously reported experimental value of 8°[123]. With more ethanol in the solvent, $\bar{\phi}_h$ in the surface region grows to 10° (at $x_{Eth} = 0.2$). At $x_{Eth} = 0.8$ and higher $\bar{\phi}_h$ is over 15°. The variation in ϕ_h in the surface region also grows with higher x_{Eth} which shows that the difference is in how strongly the parallel orientation of the axis is preferred and not that a different ϕ_h is preferred in the different solutions.

In the bulk, $\bar{\phi}_h$ is around 45° for x_{Eth} from 0 to 0.1 and around 40° for x_{Eth} from 0.2 to 0.5, always with high variance. At higher x_{Eth} the behaviour is more erratic due to the presence of clusters (see chapter 11) but the variance is still high. The difference between x_{Eth} between 0 and 0.1 and x_{Eth} between 0.2 and 0.5 stems from the fact that at $x_{Eth} = 0.2$ and higher the surface region is not as well defined as at lower x_{Eth} and in contrast to lower x_{Eth} there exist no region between the bulk and the surface region which is depleted of peptides (see chapter 9.1). $\bar{\phi}_h$ of around 45° with high variance is consistent with a random distribution of ϕ_h. Due to the symmetry of the bulk, a random distribution of ϕ_h is to be expected.

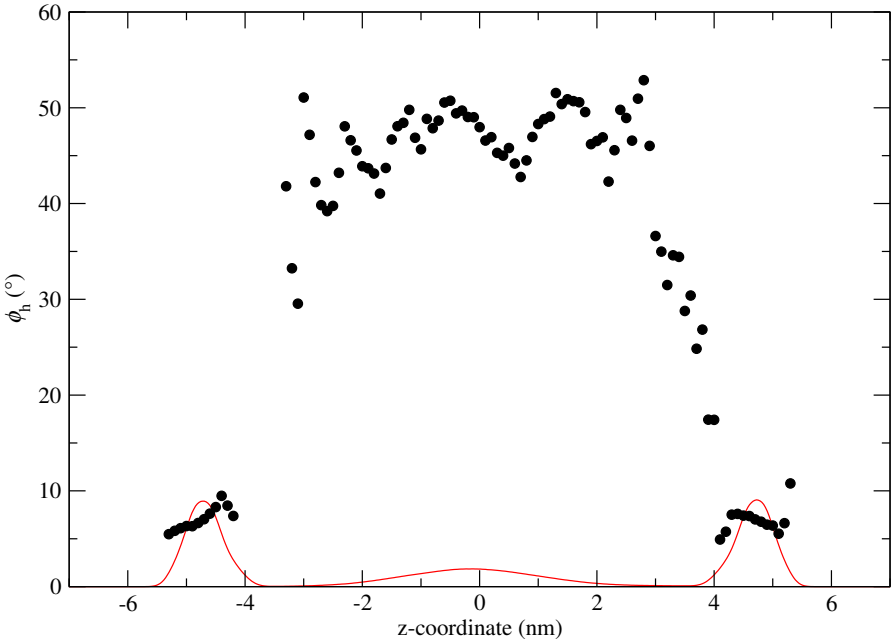


Figure 9.5.: The average angle between the surface plane and the helix, $\bar{\phi}_h$, along the z-axis of the simulation box in pure water, $x_{Eth} = 0$. The red line shows the distribution of the peptides in the box and is there as a reference.

9.3.1. Closer look at pure water

In pure water there is a clear, quite sharp boundary between the surface region and the bulk. This can be seen in figure 9.5 where $\bar{\phi}_h$ is between 5° and 8° at more than 4 nm from the centre of the simulation box, and over 30° (mostly between 40° and 50°) less than 3 nm from the centre. Figure 9.5 also shows that the surface region where population of the peptides is high is also the region where $\bar{\phi}_h$ is low. This supports the definition of the surface region as 1.5 nm thick (from 4 nm from the centre to 5.5 nm from the centre).

Figure 9.6 shows the distribution of ϕ_h at the surface and in the bulk. In the bulk all ϕ_h are equally likely. This is in agreement with the value and high variance of $\bar{\phi}_h$ at $x_{Eth} = 0$ in figure 9.4. At the surface, ϕ_h of less than 2.5° are 10 times as common as they are in the bulk while ϕ_h of more than 25° do not occur. This shows that the helix-axes lie parallel to the surface.

Figures 9.4, 9.5 and 9.6 show that in pure water the peptides are oriented so that the helix-axis is parallel to the surface at the surface and have a random orientation of the helix-axis in bulk.

9.3.2. Closer look at pure ethanol

In pure ethanol there is no clear boundary between the surface region and the bulk where the orientation of the peptide changes. As mentioned before (chapters 9.1 and 9.2) the

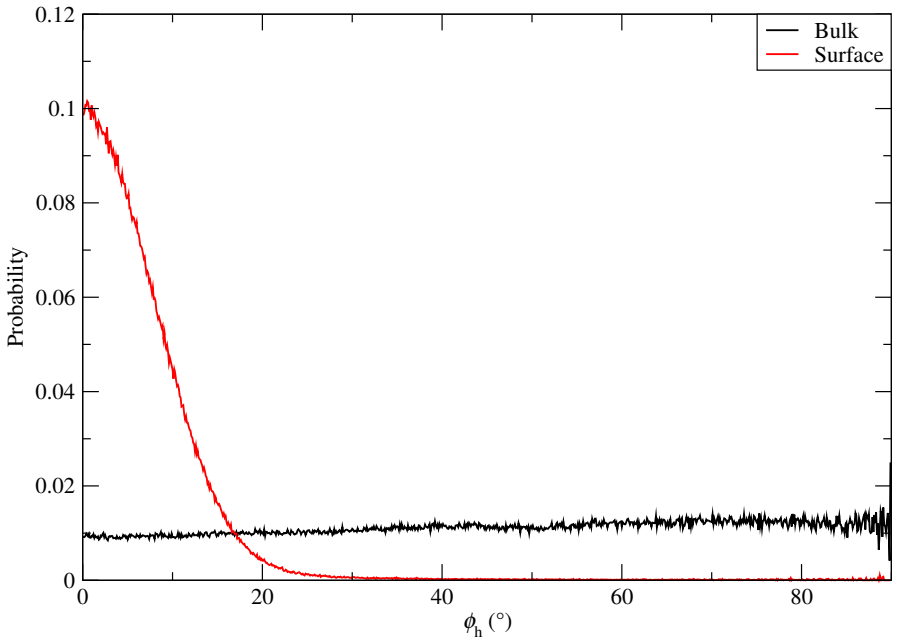


Figure 9.6.: Probability of different ϕ_h in pure water. The red line shows the distribution at the surface and the black line the distribution in the bulk. The reason for the higher noise at larger angles is the correction from equation 5.17.

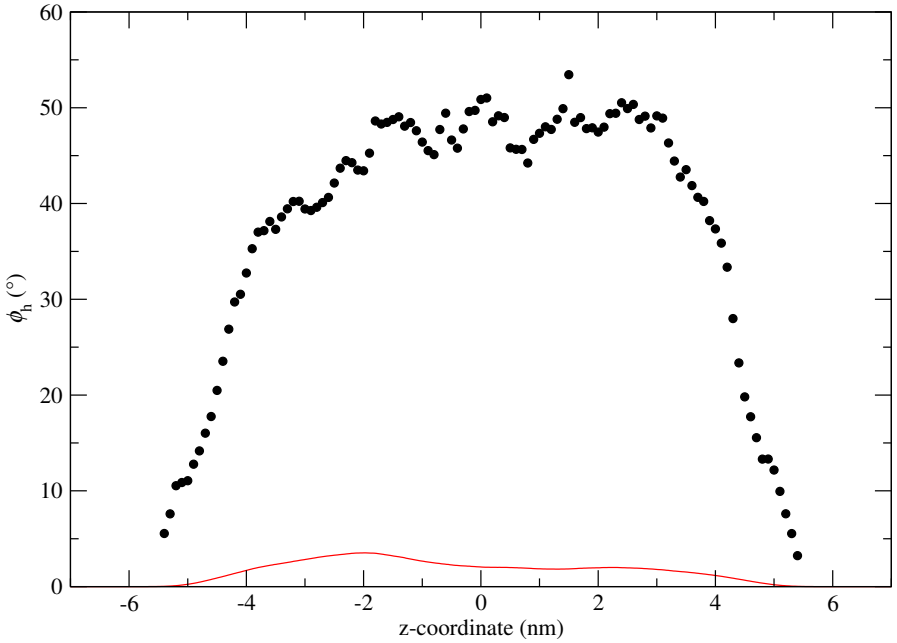


Figure 9.7.: The average angle between the surface plane and the helix, $\bar{\phi}_h$, along the z-axis of the simulation box, $x_{Eth} = 1.0$. The red line shows the distribution of the peptides in the box and is there as a reference.

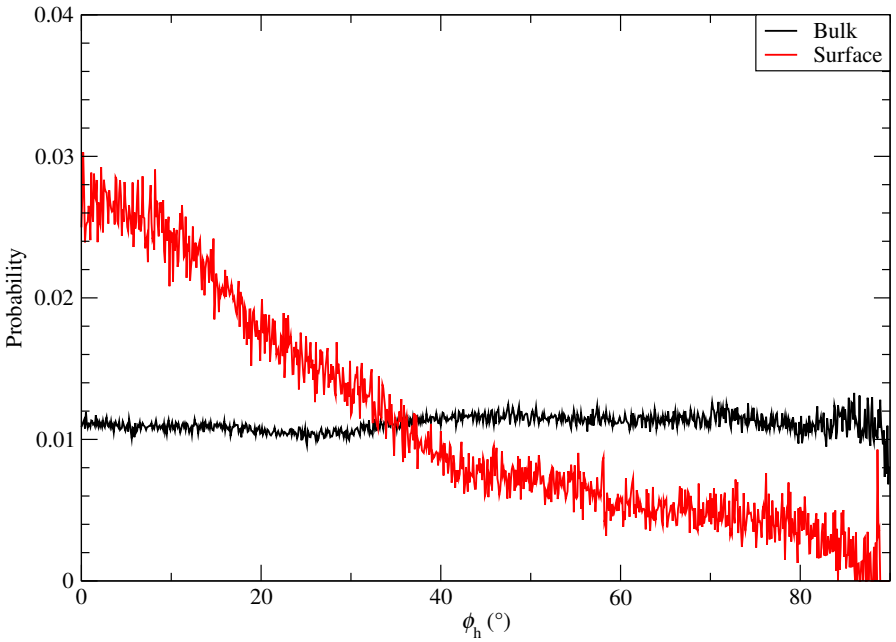


Figure 9.8.: Probability of different ϕ_h in pure ethanol. The red line shows the distribution at the surface and the black line the distribution in the bulk. The reason for the higher noise at larger angles is the correction from equation 5.17.

peptides do not adsorb to the surface in pure ethanol and this can be seen in the lack of a well defined surface region in figure 9.5 which shows $\bar{\phi}_h$ along the z-axis of the simulation box. Right at the interface, more than 5 nm from the box centre, $\bar{\phi}_h$ is smaller than 10° and there is a gradual change in $\bar{\phi}_h$ to $\bar{\phi}_h$ larger than 40° over 2 nm. Figure 9.8 shows the distribution of ϕ_h at the surface and in the bulk and shows well that though there is a difference in the distribution in the two regions it is an order of magnitude smaller than in pure water. Since there are very few peptides in the area defined as the surface, there is also more noise in the distribution there. In the bulk all ϕ_h are close to equally likely, while at the surface small ϕ_h are slightly more likely than large ϕ_h. These distributions of ϕ_h agree well with the values and high variance of $\bar{\phi}_h$ seen in figure 9.4 both at the surface and in the bulk.

The reason $\bar{\phi}_h$ is smaller right at interface is that the peptide does not go into the gas phase, neither fully nor in part. This means that for all peptides which have other orientations than parallel to the interface the centre of mass is not right at the interface and therefore only peptides which are parallel can be encountered right at the interface. There are enough peptides right at the interface for this to influence the distribution but not for small angles to dominate the distribution of ϕ_h in figure 9.8 the same way it does in pure water (see figure 9.6).

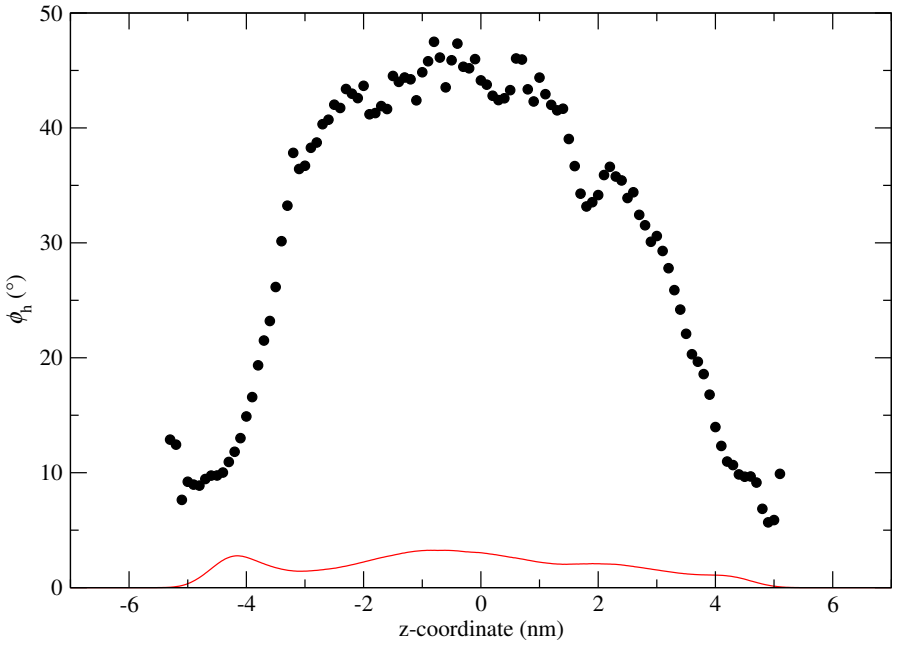


Figure 9.9.: The average angle between the surface plane and the helix, $\bar{\phi}_h$, along the z-axis of the simulation box, $x_{Eth} = 0.3$. The red line shows the distribution of the peptides in the box and is there as a reference.

9.3.3. Closer look at $x_{Eth} = 0.3$

In the solution with $x_{Eth} = 0.3$ the peptide show a behaviour between the behaviour in the two extremes of pure water and pure ethanol, as can be seen in figure 9.9. The surface region is not as clearly defined nor quite as broad as in pure water, but in contrast to pure ethanol there is a distinct surface region where $\bar{\phi}_h$ is around 10° of between 1 nm and 1.5 nm, further than 4 nm from the centre of the simulation box. Figure 9.9 also shows that just as in pure water the region where $\bar{\phi}_h$ is around 10° is the region where there is a population of peptides adsorbed at the surface. This therefore further supports the conclusion that peptides which are adsorbed at the surface are oriented in such way that the helix-axis is parallel to the interface.

Figure 9.10 shows the distribution of ϕ_h at the surface and in the bulk. Here the distribution in the bulk shows a slight preference for smaller ϕ_h, while at the surface $\phi_h < 10°$ are common and $\phi_h > 45°$ do not occur. Comparing figures 9.10 and 9.6 shows that small ϕ_h are less likely when $x_{Eth} = 0.3$ than in pure water and the fall off to zero probability occurs at higher ϕ_h in the solution with $x_{Eth} = 0.3$ than in pure water.

The distributions in figures 9.9 and 9.10 show that in the solution with $x_{Eth} = 0.3$ there is a distinct population of peptides adsorbed at the surface in contrast to pure ethanol. At the same time this surface population is not as separated from the bulk population in space as in pure water and this leads to the distributions of ϕ_h at the surface and in the bulk being influenced by each other.

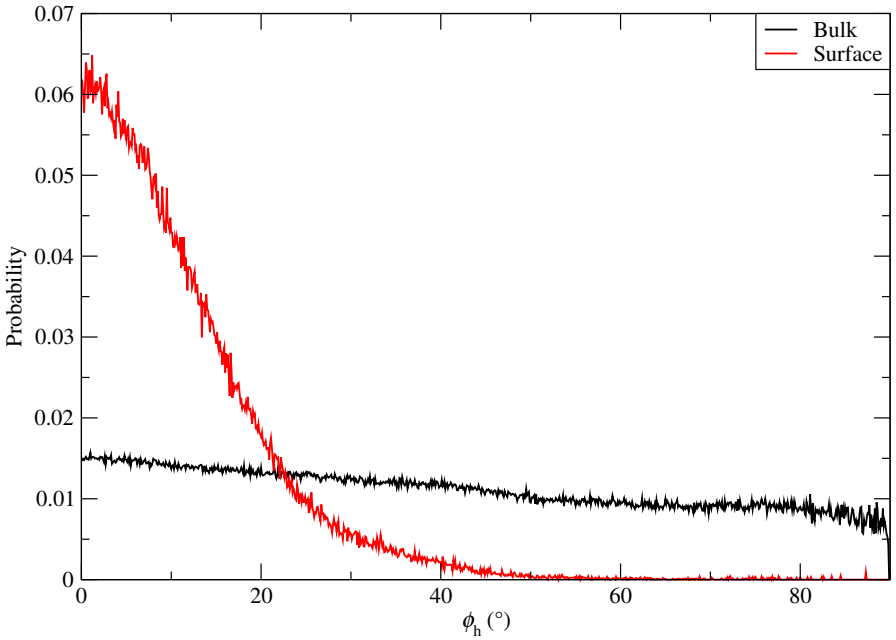


Figure 9.10.: Probability of different ϕ_h in a water/ethanol mixture with $x_{Eth} = 0.3$. The red line shows the distribution at the surface and the black line the distribution in the bulk. The reason for the higher noise at larger angles is the correction from equation 5.17.

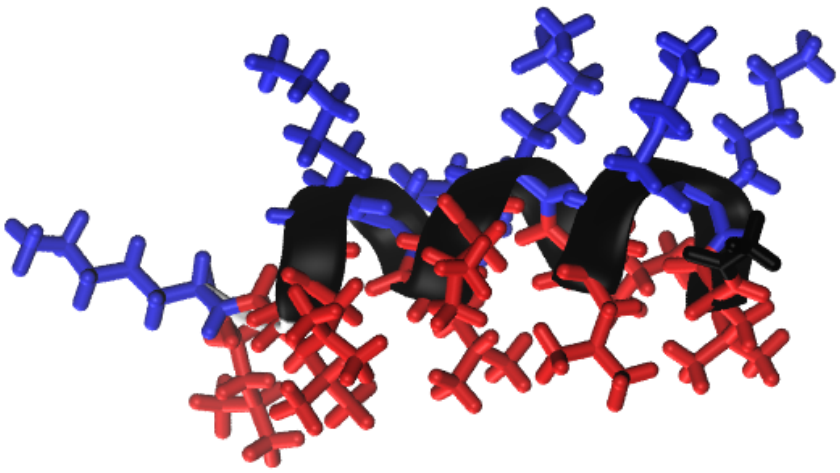


Figure 9.11.: LKα14 in α-helical conformation. Leucine residues are shown in red and lysine residues in blue while the backbone is black. It can clearly be seen how the different amino acids form different sides of the helix.

9.4. Orientation of dipole moment

The LKα14 peptides in the simulations are charged, the amino groups at the side chains of all six lysine residues are protonated and the COO-terminus of the peptide is deprotonated. Since the N-terminus are caped with an acetyl group and therefore not charged, this means that each peptide has a charge of +5. When the conformation of the peptides is α-helical, the lysine residues are on one side of the helix and the leucine residues on the other side as can be seen in figure 9.11. Since the lysine residues are charged and the leucine residues are not, it is easy to come to the conclusion that the orientation of the dipole moment of the peptide should give information about the orientation of the peptides. As will be shown in this chapter the dipole moment of the peptide does contain information about the orientation, but it is necessary to consider the backbone and the side chains separately.

9.4.1. Dipole moment of the whole peptide

The angle between the dipole moment of the peptide and the surface plane was analysed for all simulations which were performed with an interface. The surface region was defined the same way as in chapter 9.1 and earlier in this chapter, as 1.5 nm from the interface. When calculating the average angle between the surface and the dipole moment of the whole peptide, $\bar{\phi}_{\text{whole}}$, the correction from equation 5.17 was used, for the same reason as for $\bar{\phi}_{\text{h}}$.

As can be seen from figure 9.12 there is no clear difference in the value of $\bar{\phi}_{\text{whole}}$ in the bulk and at the surface for $x_{\text{Eth}} < 0.6$. At higher x_{Eth} clusters make the equilibration

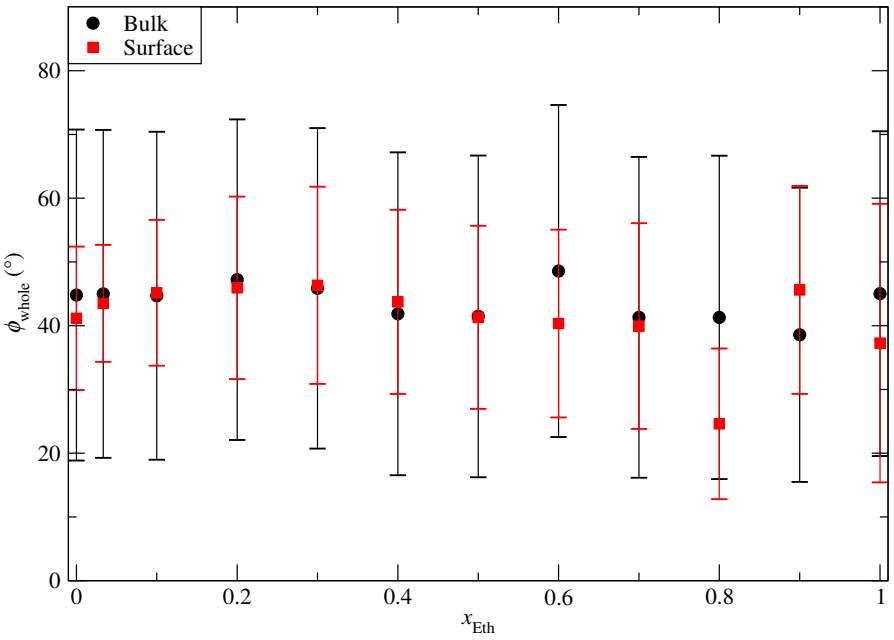


Figure 9.12.: The average angle between the surface and the dipole moment of the whole peptide, $\bar{\phi}_{\text{whole}}$, as a function of the mole fraction of ethanol, x_{Eth}. The red squares are at the surface and the black circles in the bulk.

more difficult and therefore the patterns break down. When $x_{\text{Eth}} < 0.6$ $\bar{\phi}_{\text{whole}}$ is between 40° and 50° both at the surface and in the bulk, which would seem to suggest a random distribution in both regions. But when looking closer at the variance it becomes clear that it is significantly smaller at the surface in pure water. This does suggest a difference in the distribution of ϕ_{whole} in the two regions. The variance in ϕ_{whole} at the surface grows with higher x_{Eth} until in pure ethanol there is little difference in the variance between the surface and the bulk. This agrees well with the results shown in figures 9.7 and 9.8, which show no distinct surface population in pure ethanol. Figure 9.13 shows the distribution of ϕ_{whole} both at the surface and in the bulk in pure water and in pure ethanol. There the difference in the distribution of ϕ_{whole} in the different domains in water can be clearly seen. At the surface there is a broad peak at $\phi_{\text{whole}} = 38.5°$ while in the bulk all angles are equally likely, as they also are in the bulk of pure ethanol. At the surface of pure ethanol $\phi_{\text{whole}} > 60°$ is less common than smaller ϕ_{whole} but the difference between the surface and the bulk is far smaller than in water.

9.4.2. Dipole moment of the backbone and of the side chains

There are two different parts of the peptides which are charged, the carboxyl acid terminus and the amino groups at the ends of the lysine side chains. In figure 9.11 it can be seen that when the peptide is in an α-helixcal conformation a vector pointing from the general area of the leucine residues to the general area of the lysine residues is perpendic-

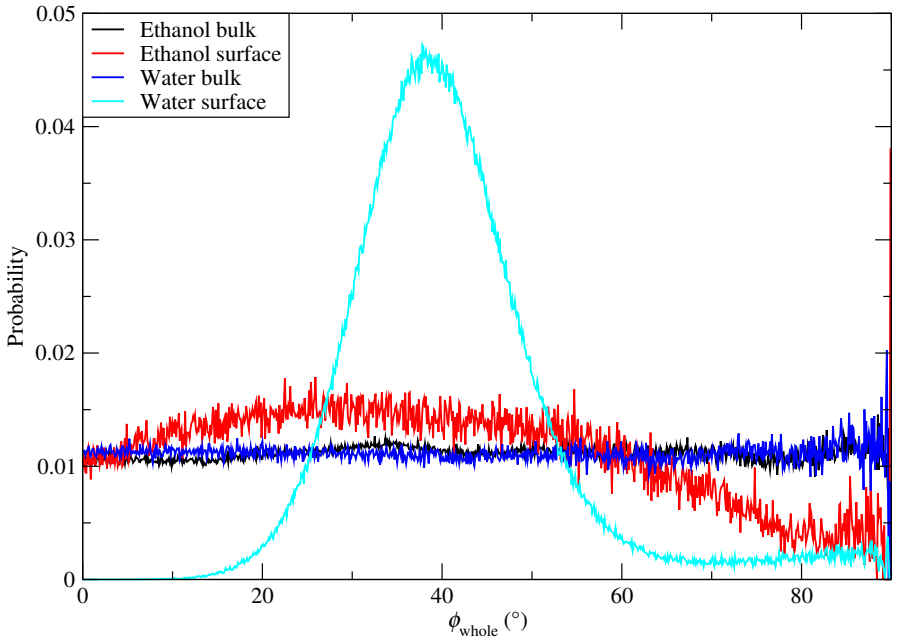


Figure 9.13.: Probability of different ϕ_{whole} in pure water and in pure ethanol. The black and the red line shows the distribution in pure ethanol, in the bulk and at the surface respectively, the blue line the distribution in the bulk of pure water and the cyan line the distribution at the surface of pure water. The reason for the higher noise at larger angles is the correction from equation 5.17.

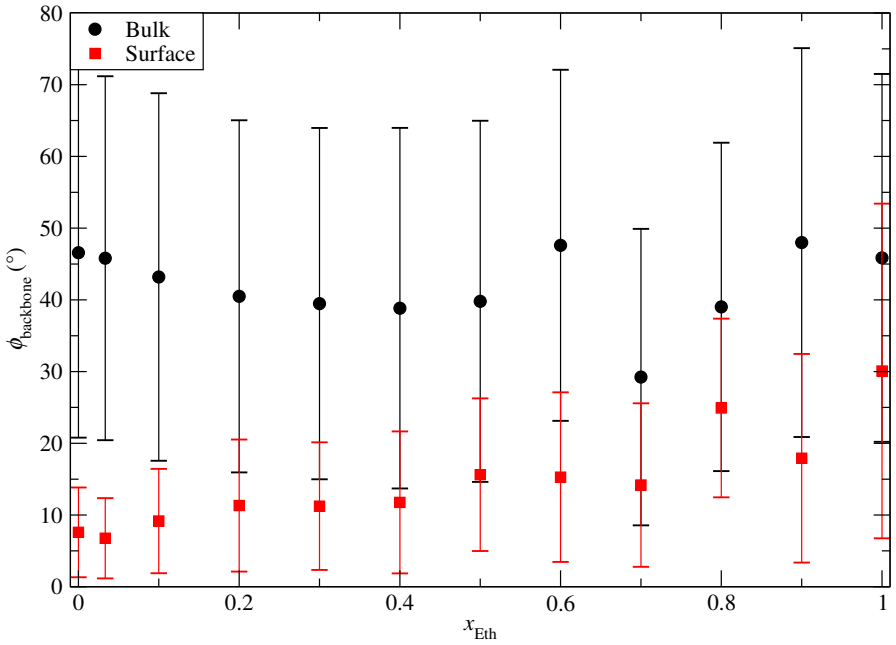


Figure 9.14.: The average angle between the surface and the dipole moment of the peptide backbone, $\bar{\phi}_{\text{backbone}}$, as a function of the mole fraction of ethanol, x_{Eth}. The red squares are at the surface and the black circles in the bulk.

ular to the helix axis. Therefore it is reasonable to conclude that looking at the backbone alone would give similar results to the helix axis and that the dipole moment of only the side chains might be a better measure of the orientation of the residues than the dipole moment of the whole peptide. Figure 9.14 shows the average angle between the interface and the dipole moment of the backbone, $\bar{\phi}_{\text{backbone}}$, at different x_{Eth}. Comparing it to figure 9.4 shows that the behaviour of $\bar{\phi}_{\text{h}}$ and $\bar{\phi}_{\text{backbone}}$ is very similar. This suggest that the dipole moment of the backbone is along the helix axis.

The average angle between the interface and the dipole moment of the side chains, $\bar{\phi}_{\text{side}}$, on the other hand is between 75° and 80° at the surface for $x_{\text{Eth}} < 0.6$ as can be seen in figure 9.15. This is nearly perpendicular to the interface. Since as previously mentioned the dipole moment of the side chains can be considered to line up with a vector perpendicular to a plane dividing the peptide between the lysine and leucine residues, this means that there is a strong tendency for this dividing plane to be parallel to the interface when $x_{\text{Eth}} < 0.6$. In this orientation the lysine residues are oriented towards the bulk while the leucine residues are oriented towards the interface. That this is the case in pure water agrees well with previously reported experimental results[123,124].

Figure 9.16 shows the distribution of ϕ_{whole}, ϕ_{backbone} and ϕ_{side} both in the bulk and at the surface for $x_{\text{Eth}} = 0$, $x_{\text{Eth}} = 0.3$ and $x_{\text{Eth}} = 1$ and figure 9.17 shows $\bar{\phi}_{\text{whole}}$, $\bar{\phi}_{\text{backbone}}$ and $\bar{\phi}_{\text{side}}$ along the z-coordinate in the same solutions. Figures 9.16a and 9.17a show how in pure water the behaviour of ϕ_{whole} is right between the behaviour of ϕ_{backbone}

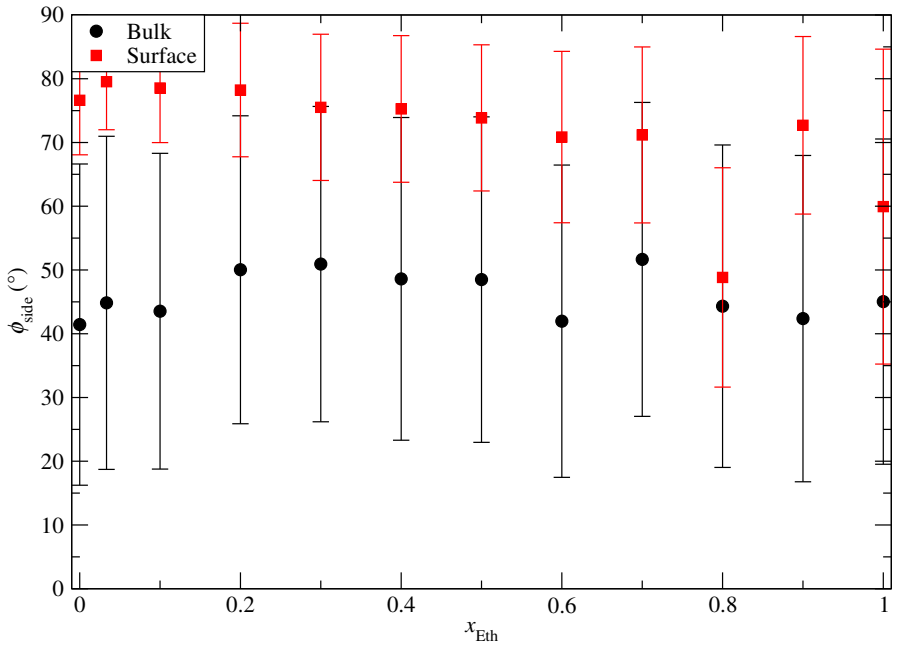


Figure 9.15.: The average angle between the surface and the dipole moment of the side chains of the peptide, $\bar{\phi}_{\text{side}}$, as a function of the mole fraction of ethanol, x_{Eth}. The red squares are at the surface and the black circles in the bulk.

and ϕ_{side}. This is not surprising since the dipole moment of the whole peptide is the sum of the dipole moments of the backbone and of the side chains. Figures 9.16c and 9.17c show that in ethanol the difference between the surface and the bulk is not large. This is consistent with the results in 9.3 for the helix axis. At last figures 9.16b and 9.17b show how the behaviour of ϕ_{whole}, ϕ_{backbone} and ϕ_{side} at the surface of a solution with $x_{\text{Eth}} = 0.3$ lies between their behaviour in pure water and in pure ethanol. In the bulk the same influence from the surface can be seen as can be seen in figure 9.10 for the helix axis.

Comparing figures 9.17a, 9.17b and 9.17c also shows how the sharp divide between the bulk and the surface in pure water becomes more undefined and gradual with lower x_{Eth}.

9.5. Discussion

LKα14 has been shown to adsorb to water/air interfaces both in experiments[58,120,123] and in MD simulations[84]. α-helical LKα14 peptides have also been found to lie parallel to interfaces when adsorbed to them, both water/air interfaces[123] and water/apolar interfaces in general[47,124]. The results presented in sections 9.2 and 9.3 agree well with these previous studies. LKα14 is shown to adsorb to the water/air interface and the helix axis to be parallel to the surface.

At an apolar/polar interface LKα14 α-helices have been shown to orient the hydropho-

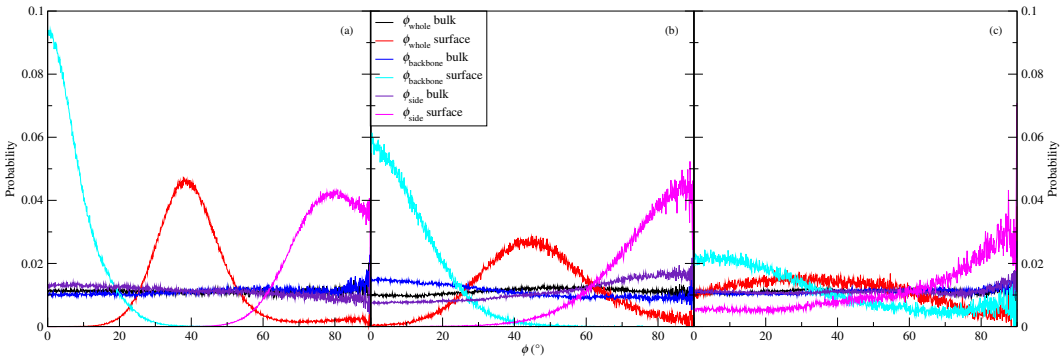


Figure 9.16.: Distribution of the angle between the surface and the dipole moment of the whole peptide, ϕ_{whole}, the backbone, ϕ_{backbone}, and the side chains, ϕ_{side}, in solutions with (a) $x_{\text{Eth}} = 0$, (b) $x_{\text{Eth}} = 0.3$ and (c) $x_{\text{Eth}} = 1$. The colours are the same for all three solutions. The reason for the higher noise at larger angles is the correction from equation 5.17.

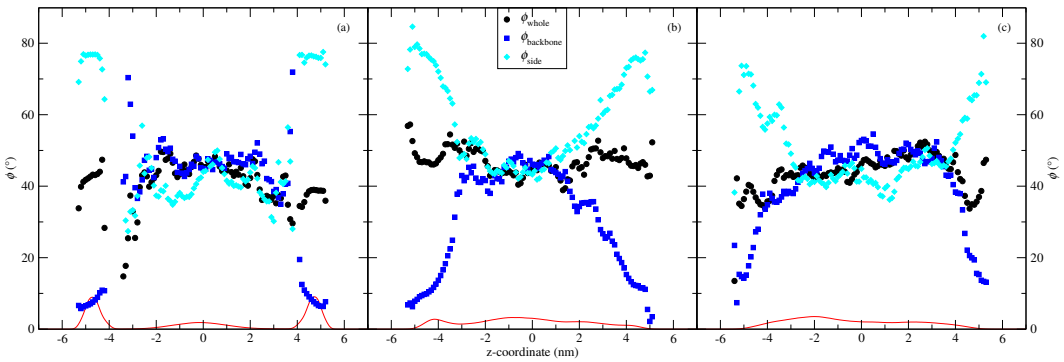


Figure 9.17.: Distribution of ϕ_{whole} (black circles), ϕ_{backbone} (blue squares) and ϕ_{side} (cyan diamonds) along the z-axis in solutions with (a) $x_{\text{Eth}} = 0$, (b) $x_{\text{Eth}} = 0.3$ and (c) $x_{\text{Eth}} = 1$. The red lines show the distribution of the peptides in the box and are there as a reference.

bic leucine residues towards the apolar side and the hydrophilic lysine residues towards the polar side[47,83,84]. This orientation in conjunction with the amphiphilic nature of the α-helical conformation of LKα14 was proposed as the reason for the adsorption and the conformation of LKα14 by DeGrado and Lear, when they designed it, though they did not measure the orientation of the peptides at the surface[58]. The results in section 9.4 also agree well with these previous studies, the analysis of the dipole of the side chains of the peptides shows that the leucine residues are oriented towards the air and the lysine residues are oriented towards the bulk of the solution.

The simulations presented in this chapter show that with higher ethanol mole fraction in the solution the adsorption becomes weaker until at $x_{\mathrm{Eth}} \geq 0.4$ no adsorption is observed. This has also been shown experimentally by Maltseva et al[120]. The orientation at the surface also becomes less uniform with higher ethanol mole fraction in the solution. At higher x_{Eth} there is also a decrease in how sharp the divide between the orientation in the bulk and at the surface is. Both the lack of a sharp divide between the surface and the bulk, and the lack of a uniform orientation at the surface fit with the fact that at high x_{Eth} no adsorption is observed.

For x_{Eth} between 0.4 and 0.9 and especially between 0.6 and 0.9, the patterns in orientation are less strong. This is most likely due to the clusters discussed in chapter 11. The formation of clusters means it takes longer time to sample the whole phase space of the system and therefore longer simulations might lead to a clearer picture of what happens at these x_{Eth}.

The surface tension results in section 9.2.1 also support the conclusion that at low x_{Eth} the peptides adsorb to the surface but at high x_{Eth} they don't. The surface tension changes are in qualitative agreement with experimental results by Maltseva et al.[120] though the changes measured by them are an order of magnitude larger than the changes observed here.

10. Solvation shells

10.1. Determining the thickness of the solvation shells

To determine the thickness of the solvation shells around the peptides, r was defined as the distance from each atom in the solvent molecules to the closest atom in the peptide. This gave a distribution function from which it was possible to determine the thickness of the first solvation shell around the peptide (see figure 10.1). With this method of calculation of r it is difficult to calculate the true normalised $g(r)$ since calculating the volume needed for calculations of $\langle p_b(r) \rangle$ and $\langle p_b \rangle_{\max}$ in equation 2.1 is complicated. The distribution functions in figure 10.1 are therefore calculated using particle counts instead of true particle densities.

Figure 10.1 shows that the main peak for the first solvation shell has a slightly different position depending on whether the RDFs is calculated between the water and the peptide or the ethanol and the peptide. Different concentration of the different solvents on the other hand only change the absolute and relative heights of the peaks but not their positions. Looking at the water RDFs the first solvation shell can be estimated to be from 0.15 nm to 0.3 nm from the surface of the peptide while the estimate from the ethanol RDF is from 0.2 nm to 0.4 nm. The difference in thickness is caused by the difference in the size of water and ethanol molecules. For further analyses of the composition of the solvation shells a thickness of 0.3 nm was used.

10.2. Composition of solvation shells

The composition of the first, second and third solvation shells was calculated for solutions with x_{Eth} from 0 to 1. The simulations used for the analyses were 900 ns long and consisted of a single peptide in an α-helical conformation in a $5 \times 5 \times 5$ nm box of solution. A solvent molecule was considered part of the first solvation shell if any of its atoms was within 0.3 nm of any peptide-atom (in accordance with 10.1) and part of the second or third solvation shells respectively if any of its atoms was within 0.6 nm or 0.9 nm of any peptide-atom and it was not part of an closer solvation shell. That means each solvent molecule was always counted as part of the inner most solvation shell it qualified for and no other.

Figure 10.2 shows x_{Eth} for the first, second and third solvation shells as well as in the bulk as a function of $x_{\mathrm{Eth},0}$, the mole fraction of ethanol in the whole simulation box. There is a slight difference between the third solvation shell and the bulk, which is largest for $x_{\mathrm{Eth},0} = 0.5$ and $x_{\mathrm{Eth},0} = 0.6$. x_{Eth} in the second solvation shell is lower than in the bulk for $x_{\mathrm{Eth},0}$ between 0.3 and 0.9. The difference is largest for $x_{\mathrm{Eth},0} = 0.7$. As is to

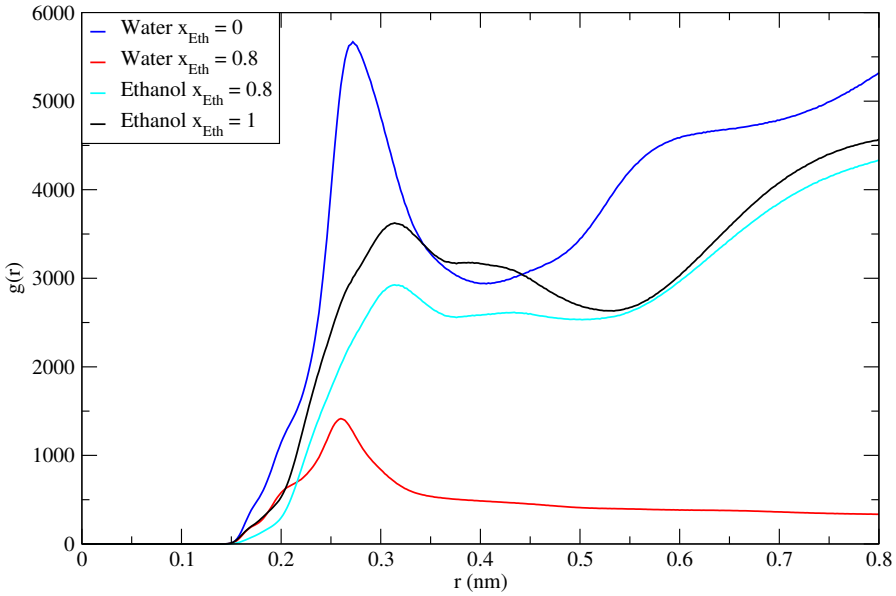


Figure 10.1.: RDFs for water and ethanol calculated from the surface of the peptide. The RDFs are not normalised. Therefore the peak heights are not comparable with each other, but the peak position are. The blue and the red line show water in solution with $x_{\mathrm{Eth}} = 0$ and $x_{\mathrm{Eth}} = 0.8$ respectively and the cyan and the black line show ethanol in solution with $x_{\mathrm{Eth}} = 0.8$ and $x_{\mathrm{Eth}} = 1$ respectively. The positions of the peaks for both ethanol RDFs are the same though both the absolute and the relative heights are different. The same is true of the water RDFs in the region where $r < 0.4$ nm.

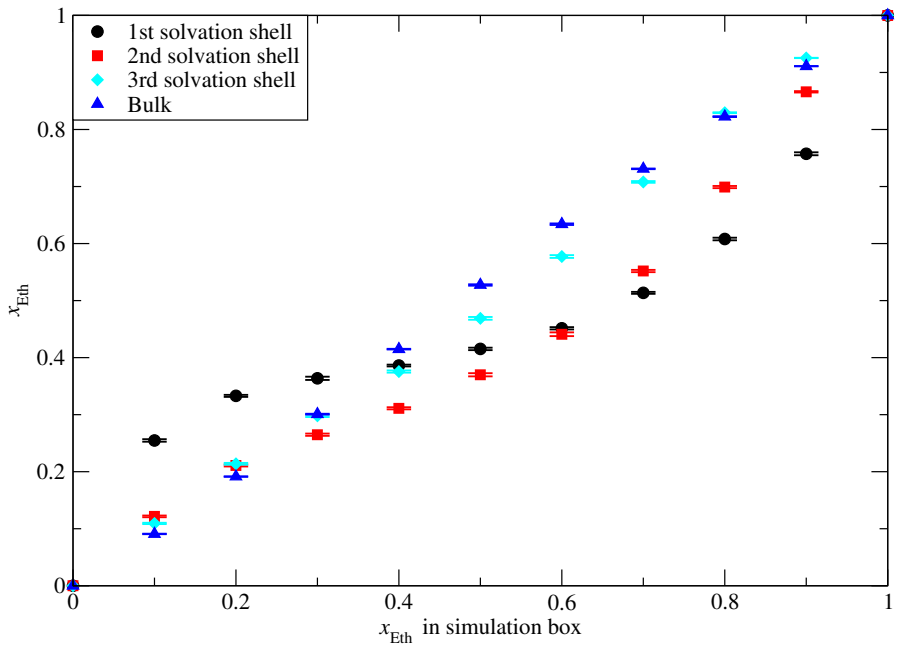


Figure 10.2.: The ethanol mole fraction, x_{Eth}, in the different solvation shells as a function of x_{Eth} in the simulation box. The values for the bulk are not the same as for the whole simulation box, as the molecules in the three solvation shells are not part of the bulk. There is little difference between the third solvation shell and the bulk but the composition of both the first and the second solvation shell deviate from the composition of the bulk.

be expected, the composition of the first solvation shell shows the largest deviation from the composition of the bulk. For small $x_{\mathrm{Eth},0}$, between 0.1 and 0.3, x_{Eth} is larger in the first solvation shell than in the bulk, while for $x_{\mathrm{Eth},0}$ from 0.5 to 0.9 it is lower in the first solvation shell than in the bulk. When $x_{\mathrm{Eth},0} = 0.4$ the mole fraction of ethanol in the first solvation shell is 0.39 and as such the composition of the first solvation shell is very close to $x_{\mathrm{Eth},0}$.

Figure 10.2 shows that x_{Eth} in the first solvation shell is between 0.35 and 0.45 for $x_{\mathrm{Eth},0}$ from 0.3 to 0.6. For $x_{\mathrm{Eth},0}$ between 0.1 and 0.7 the difference between the composition of the first solvation shell and the bulk gets larger the further $x_{\mathrm{Eth},0}$ is from 0.4. For $x_{\mathrm{Eth},0}$ higher than 0.7 the difference between the first solvation shell and the bulk decreases with higher $x_{\mathrm{Eth},0}$. This shows that the preferred composition of the first solvation shell of monomeric LKα14 in a binary water-ethanol solution is around 0.4.

10.2.1. Spatial distribution functions

As shown in chapter 8 the helical conformation of LKα14 is far from possessing spherical symmetry. Not only is it an α-helix and therefore inherently closer to a cylinder than a sphere, but the two types of amino acids in LKα14 form opposite sides of the helix. Since the side chain of lysine is polar and that of leucine is not, it is possible to divide the peptide into a polar half-cylinder and an apolar half-cylinder (see chapter 9.4). Therefore it is reasonable to expect that the distribution of the solute molecules around these two half-cylinders is not the same.

Figure 10.3 shows isosurfaces of the spatial distribution functions of water and ethanol around LKα14 in a solution with $x_{\mathrm{Eth},0} = 0.4$. It shows that the water in the first solvation shell is found around the polar lysine residues and the ethanol is found around the apolar leucine residues. The asymmetric form of the SDFs is due to one end of the α-helix being less rigid and the solvent molecules at that end therefore having less defined structure. This difference in distribution between water and ethanol molecules explains the preference for x_{Eth} of the first solvation shell to be around 0.4. This is the x_{Eth} which results when the lysine residues are surrendered by water and the leucine residues by ethanol. The fact that ethanol is a slightly larger molecule than water, the exact amino acid sequence of LKα14 and the size of the accessible surface of both lysine and leucine all interact to result in the mole fraction of ethanol and water in the first solvation shell not being the same even though each solvent surrounds one half of the peptide.

10.3. Discussion

As mentioned briefly in section 6.4 water and ethanol are miscible in all concentrations. Both experiments[128] and simulations[129,130] have shown that binary mixtures of water and ethanol are not completely homogeneous on the molecular scale. The scale of this heterogeneity is smaller than the scale of the separation into a clear water rich and a clear ethanol rich domain in the solvation shell of LKα14 seen in this work. Thus the presence of LKα14 influences the structure of the solvent mixture. Excess solvation by

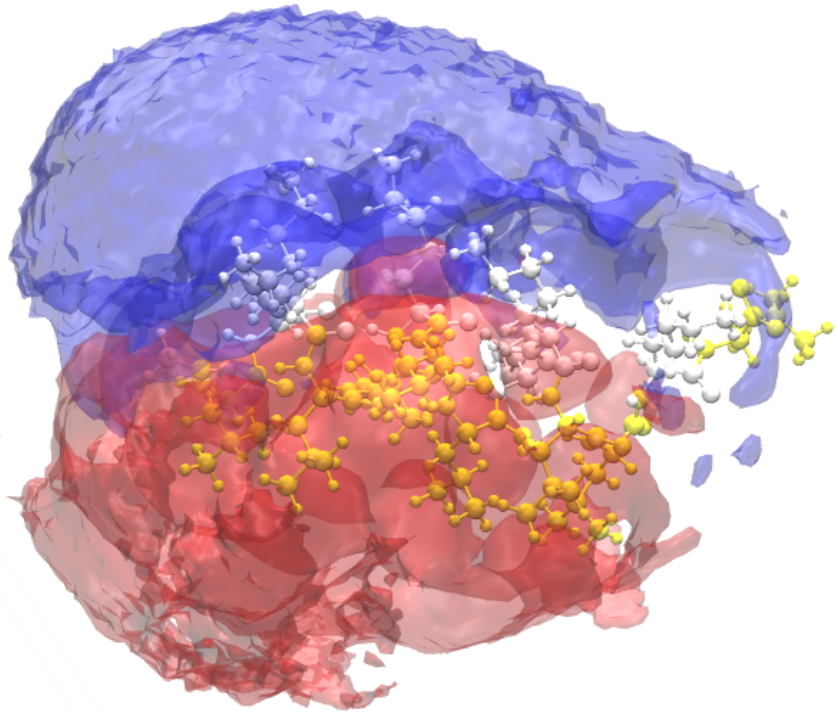


Figure 10.3.: Isosurfaces of the spatial distribution functions of water and ethanol around LKα14 in a solution with $x_{\mathrm{Eth},0} = 0.4$. The SDF of water is shown in blue and the SDF of ethanol in red. Lysine residues are shown in white and leucine residues in yellow.

either water or ethanol has been reported before for some polypeptides[131,132] and the protein lysozyme[21] but the effect is quite large for LKα14 and the change in which solvent is in excess is unusual. As shown in chapter 11.2 the influence of LKα14 on the structure of the solvent can also reach further than the first or second solvation shell of the peptide in more concentrated solutions of LKα14.

The different hydrophobicity/hydrophilicity of different amino acids is an important factor in protein folding[119,133–135]. In correctly folded proteins the more hydrophobic amino acids are in the centre of the protein while the more hydrophilic amino acids are on the surface of the protein where they are in contact with water in aqueous solutions[133,135]. A peptide of the size of LKα14 is not capable of folding in such a way, since its amino acid sequence is simply not long enough. Given the spatial separation of the hydrophilic and hydrophobic amino acids in an α-helical LKα14 peptide it can compensate for this. In the presence of a water/air interface it adsorbs to the surface (see chapter 9) and in more concentrated aqueous solutions it forms oligomers (see chapter 11). In both cases the result is similar to a folded protein, the hydrophilic amino acids are in contact with the water and the hydrophobic amino acids with either other hydrophobic amino acids or air (which is hydrophobic).

The clear spatial separation of α-helical LKα14 into a hydrophilic and hydrophobic parts is reminiscent of surfactants. The main difference is that in LKα14 the separation is along the helix axis while typical surfactant has a polar head group and a long apolar tail[23]. The adsorption of LKα14 to the surface in aqueous solutions (see chapter 9) is similar to the behaviour of surfactants. Similarly the the spatial separation of ethanol and water in the solvation shell of the α-helix is reminiscent of how surfactants orient with the hydrophilic part towards the water and the hydrophobic part towards the oil in microemulsions. In microemulsion-like systems formed by water, ethanol and for example octanol, the ethanol acts similar to a surfactant[53,55] and ethanol forms clearly defined clusters in binary mixtures with apolar solvents[129,130]. It is therefore surprising how here the ethanol takes on the role of the hydrophobic solvent.

11. Clusters

11.1. Oligomers

As mentioned in chapter 4 LKα14 is known to form oligomers, mainly tetramers, in water[58,83]. In order to investigate how the size of such oligomers depends on the composition of the solvent the aggregate size was calculated.

The simulations used for the analysis were 1000 ns long and consisted of 20 peptides in a $10 \times 10 \times 10$ nm box of solution. The cutoff distance used to determine which peptides were part of the same aggregate was 0.35 nm.

Figure 11.1 shows both the maximum aggregate size and the average aggregate size. The graph can be divided into four regions. From $x_{\mathrm{Eth}} = 0$ to $x_{\mathrm{Eth}} = 0.2$ the maximum aggregate size is around 5 peptides and the average aggregate size between 2-3 peptides with small variance in both variables. From $x_{\mathrm{Eth}} = 0.6$ to $x_{\mathrm{Eth}} = 0.9$ the maximum aggregate size is over 17 peptides and the average aggregate size is over 14 peptides with a noticeably larger variance in both variables than at low x_{Eth}. Both the maximum and average aggregate size have a maximum at $x_{\mathrm{Eth}} = 0.75$. From $x_{\mathrm{Eth}} = 0.3$ to $x_{\mathrm{Eth}} = 0.55$ there is a gradual transition between those two regions in the size of the aggregates. The transition in the variance on the other hand is not continuous. In pure ethanol the aggregates are smaller than in mixtures with high x_{Eth} with maximum size of 11 ± 3 peptides and average size of 7 ± 2 peptides.

Figure 11.2 shows the size distribution of the aggregates. The mixtures in the first region in figure 11.1 are shown in black, the mixtures in the second region in red, the third region in cyan and pure ethanol in blue. Figure 11.2 supports the division into the four regions on the basis of figure 11.1, with the exception of $x_{\mathrm{Eth}} = 0.2$ which from figure 11.2 agrees better with the transition region than the low x_{Eth} region.

When interpreting figure 11.2 it is important to note that it shows the number of aggregates of each size, not the number of peptides in aggregates of each size. If the number of dimers is the same as the number of monomers two times as many peptides are part of the dimers than are monomeric.

The solution with $x_{\mathrm{Eth}} = 0.2$ is the one where monomers are the most common, with 5.4 out of 20 peptides being monomers. At low x_{Eth} dimers, trimers and tetramers more numerous than monomers. There is a sharp difference in the frequency of these smaller oligomers and pentamers, hexamers and heptamers which are significantly less common. Larger aggregates are very uncommon and none at all are larger than 13 peptides. In the transition region (which here includes $x_{\mathrm{Eth}} = 0.2$) monomers are more common than oligomers. Then the frequency of each aggregate size decreases exponentially with increasing aggregate size, though for $x_{\mathrm{Eth}} = 0.5$ and $x_{\mathrm{Eth}} = 0.55$ the frequency increases

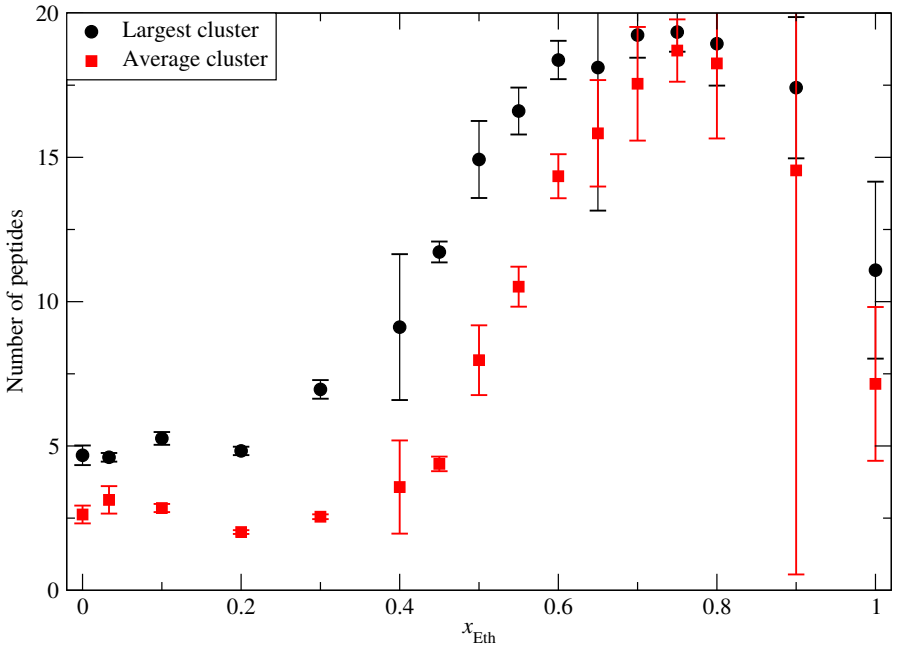


Figure 11.1.: The maximum and average size of aggregates, with aggregates defined as peptides which are within 0.35 nm of each other.

again for aggregate sizes of 16 peptides or more. At high x_{Eth} the most common aggregate size is 20 peptides. Monomers are also more common than oligomers, with two exceptions. At $x_{\mathrm{Eth}} = 0.9$ dimers are more common than monomers and for $x_{\mathrm{Eth}} = 0.7$ monomers and pentamers have the same frequency. In pure ethanol monomers are the most numerous but the largest amount of peptides are part of octamers (5.3 on average). There are five more aggregate sizes which contain on average significantly more peptides than there are monomers: trimers (2.4 peptides), heptamers (1.2 peptides), nonamers (1.4 peptides) and aggregates of size 17 (1.5 peptides) and 20 (1.8 peptides).

11.1.1. Oligomers in water and at low x_{Eth}

Figure 11.3 shows two typical snapshots from the simulation of the peptides in pure water. It shows only the peptides and not the solvent in order to better see the oligomers. It can clearly be seen that in all oligomers the leucine residues are oriented towards the other peptides in the oligomer while the lysine residues are oriented out towards the water. As can be seen by comparing figures 11.3a and 11.3b the tetramers in both have the same conformation of the four peptides: the axis of two and two peptides are parallel and the axis of the two pairs of parallel peptides are perpendicular to each other (see bottom figure 11.3a or top right corner of figure 11.3b). From the right perspective the tetramers look similar to the symbol #. The trimers come in two different conformations: three parallel peptides (see for example centre of figure 11.3b) or two parallel and one perpendicular peptide similar to a tetramer with one peptide missing (for example in the top left corner

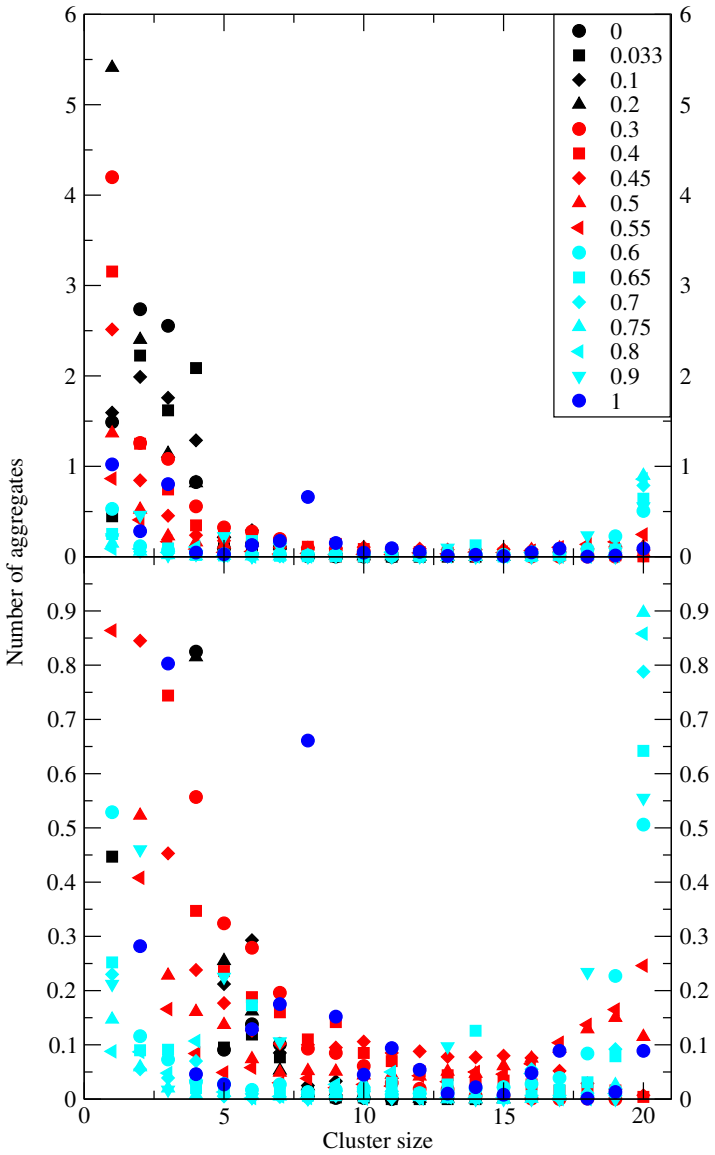
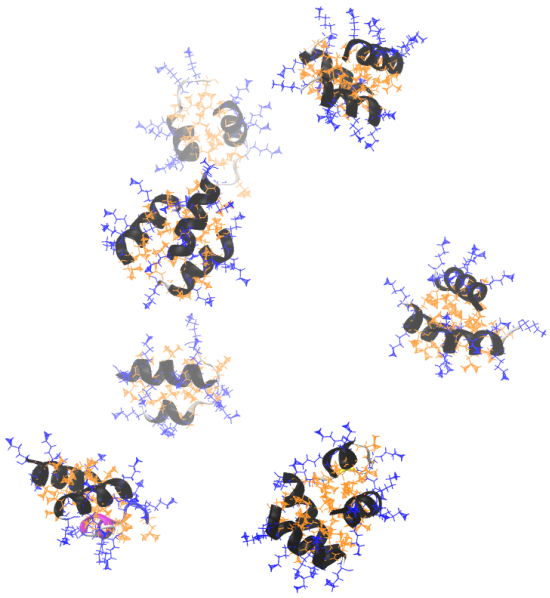
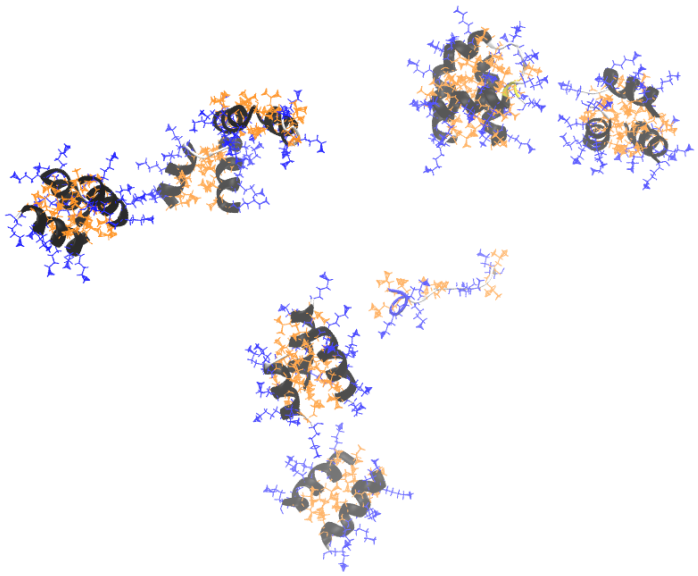


Figure 11.2.: The number of aggregates of each aggregate size, with aggregates defined as peptides which are within 0.35 nm of each other. The legend shows the mole fraction of ethanol, x_{Eth}, of each simulation. The lower figure shows a zoomed in view of the less common aggregate sizes. The colours correspond to different regions of the graph in figure 11.1. If all 20 peptides are in a single aggregate in all frames of the trajectory the number of aggregates of size 20 is 1. If all 20 peptides are monomers the number of aggregates of size 1 is 20.

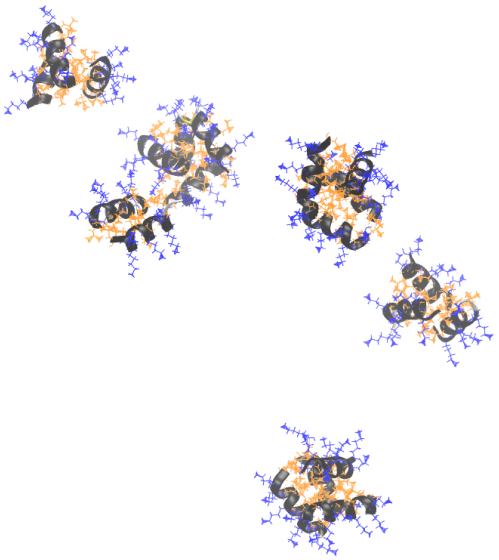


(a) Two dimers, four trimers and a tetramer.

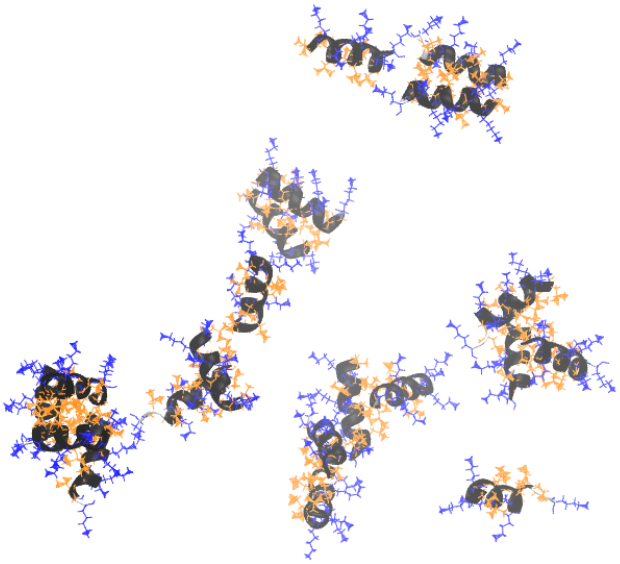


(b) One monomer, three dimers, three trimers and a tetramer.

Figure 11.3.: Two snapshots of the peptides in a simulation with mole fraction ethanol $x_{\mathrm{Eth}} = 0$. The backbone is black in α-helix, blue in turn, purple in 3_{10}-helix and white in random coil. The lysine side chains are blue and the leucine side chains er orange.



(a) $x_{\mathrm{Eth}} = 0.033$. One dimer, two trimers and three tetramers.



(b) $x_{\mathrm{Eth}} = 0.1$. One monomer, one dimer, one trimer and one tetramer are clearly defined. Additionally two cases which are each either a trimer or a dimer and a monomer and one case which is either a tetramer or two dimers.

Figure 11.4.: Snapshots of the peptides in simulations with mole fraction ethanol $x_{\mathrm{Eth}} = 0.033$ and $x_{\mathrm{Eth}} = 0.1$. The backbone is black in α-helix, blue in turn, purple in 3_{10}-helix and white in random coil. The lysine side chains are blue and the leucine side chains er orange.

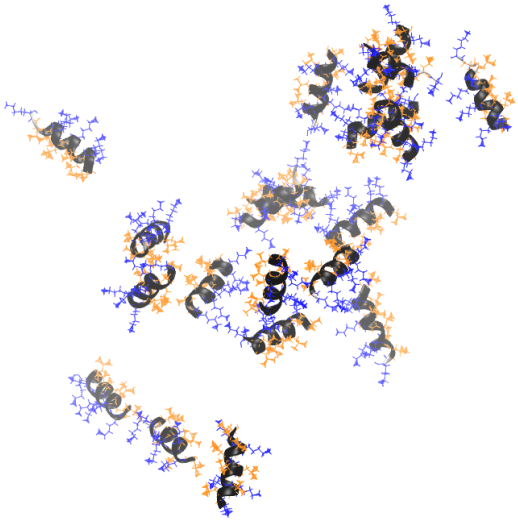


Figure 11.5.: Snapshot of the peptides in a simulation with mole fraction ethanol $x_{\mathrm{Eth}} = 0.2$. It is difficult to categorise or count the oligomers by eye. The backbone is black in α-helix, blue in turn, purple in 3_{10}-helix and white in random coil. The lysine side chains are blue and the leucine side chains er orange.

of figure 11.3b). The second trimer conformation and the tetramer conformation are in good agreement with the conformations found in earlier simulations[83].

In experiments LKα14 forms tetramers not trimers in water[58]. In the simulation both dimers and trimers were more common than tetramers. In the simulation the oligomers are also quite stable over the whole simulation time and did not exchange peptides or change size except for one case of a monomer and a dimer merging to form a trimer. This suggests that the size distribution of the oligomers is not converged though the conformation of them and the number of monomers is. Therefore the discrepancy between the simulation and experiments is not significant and might disappear if the simulation was extended.

Figure 11.4 similarly shows typical snapshots from simulations with $x_{\mathrm{Eth}} = 0.033$ and $x_{\mathrm{Eth}} = 0.1$ and figure 11.5 one from a simulation with $x_{\mathrm{Eth}} = 0.2$. As can be seen in figure 11.4a the conformations of the oligomers are the same when $x_{\mathrm{Eth}} = 0.033$ as in pure water. When $x_{\mathrm{Eth}} = 0.1$ oligomers are less well defined as can be seen in figure 11.4b. Especially noticeable when comparing figure 11.4b with figures 11.3 and 11.4a is that when $x_{\mathrm{Eth}} = 0.1$ more leucine residues are oriented towards the solvent. The distribution of the peptides in figure 11.5 seems chaotic in comparison with even figure 11.4b. There are aggregates but they have far less well defined inside and outside than in solvents with lower x_{Eth}. When considering only the solvents with low x_{Eth} it would be reasonable to conclude that higher x_{Eth} leads to less well defined and fewer aggregates and more monomers. As seen in figure 11.2 there is in fact a maximum in the prevalence of monomers when $x_{\mathrm{Eth}} = 0.2$.

The size distribution of the oligomers in the solutions with $x_{\mathrm{Eth}} = 0.033$ and $x_{\mathrm{Eth}} = 0.1$ is most likely not converged, similarly to the size distribution in pure water. Since there are more exchanges of the peptides between oligomers with higher x_{Eth} the size distribution of $x_{\mathrm{Eth}} = 0.1$ especially is probably closer to the true size distribution than the distribution in pure water is. In the simulation with $x_{\mathrm{Eth}} = 0.2$ the size distribution is very likely converged since there are more monomers floating around and more exchanges of peptides between different oligomers. As in pure water the conformation of the peptides in each oligomer and the number of monomers are converged.

11.1.2. Pure ethanol

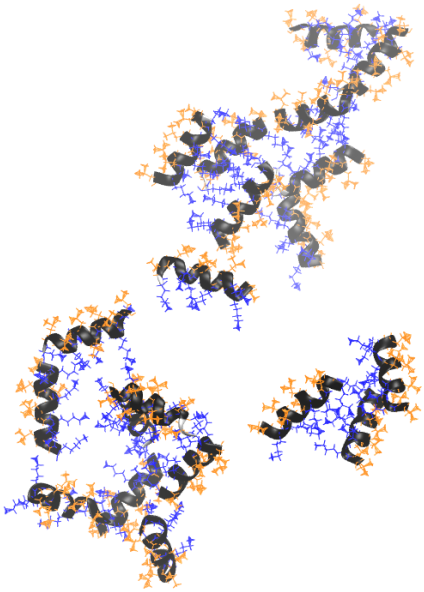
Figure 11.6 shows two snapshots from the simulation of the peptides in pure ethanol. In contrast to the oligomers in solvents with low x_{Eth} where the leucine residues are oriented inwards towards the other peptides of the oligomer in the aggregates in pure ethanol the lysine residues are oriented inwards and the leucine residues out towards the ethanol. In the larger aggregates in pure ethanol there are also some leucine residues which are oriented toward the leucine residues of other peptides as can be seen in figure 11.6b but the lysine residues of the same peptides are also oriented toward the lysine residues of other peptides. The aggregates in pure ethanol are not as well defined and do not have as clear preferred conformation as the ones in pure water but there are fewer peptides with both the lysine and the leucine side chains exposed to the solvent than when $x_{\mathrm{Eth}} = 0.2$.

There is also tendency for the aggregates to merge with simulation time, as can be seen by comparing figures 11.6a and 11.6b and from figure 11.7 which shows the number of peptides in the largest aggregate as a function of time. This suggest that 1000 ns are not enough to quantitatively converge the size of the aggregates in pure ethanol though qualitatively it can be said that large aggregates or even a single aggregate containing all 20 peptides are preferred.

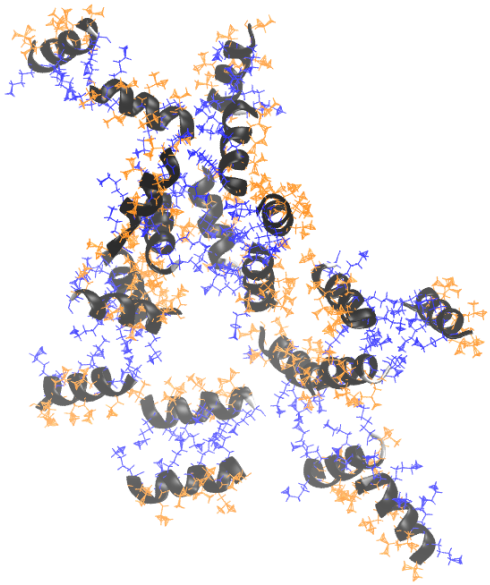
11.2. Clusters of peptides and water

In the solvent compositions not detailed in sections 11.1.1 or 11.1.2 the peptides form clusters with water molecules. As seen in chapter 10.2.1 the water and the ethanol in the solvation shell of single LKα14 α-helix in bulk are spatially separated. In these clusters the water part of the solvation shells is partly shared between different peptides. Depending on the solvent composition additional water molecules which are not part of the solvation shell of the peptides can also be part of the cluster.

Figure 11.8 shows two such clusters in a simulation with $x_{\mathrm{Eth}} = 0.7$. The clusters consist of water molecules surrounded by peptides. The more hydrophilic lysine side chains are oriented towards the water while the more hydrophobic leucine side chains are oriented out towards the rest of the solution. Figure 11.8a shows that the clusters can connect to themselves over the periodic boundaries of the simulation box. Comparing figures 11.8a and 11.8b shows that in the same solvent composition the clusters can take on different shapes.



(a) Early in the trajectory.



(b) Late in the trajectory.

Figure 11.6.: Two snapshots of the peptides in a simulation with mole fraction ethanol $x_{\mathrm{Eth}} = 1$. The backbone is black in α-helix, blue in turn, purple in 3_{10}-helix and white in random coil. The lysine side chains are blue and the leucine side chains er orange.

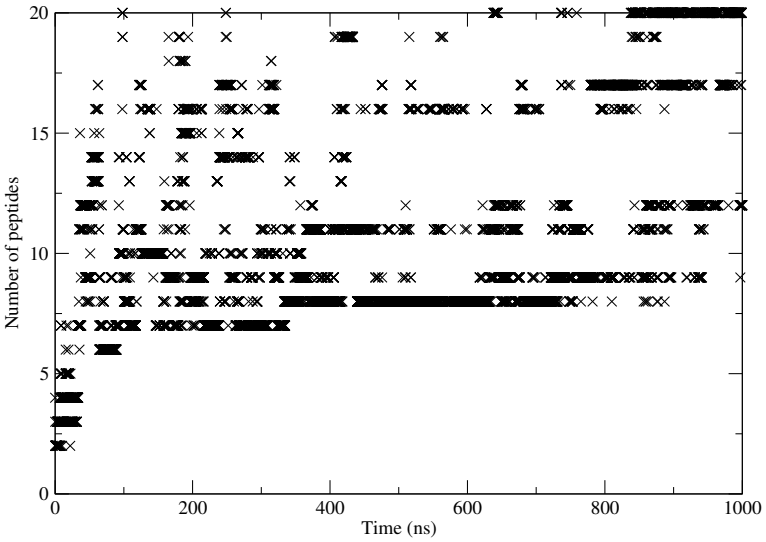


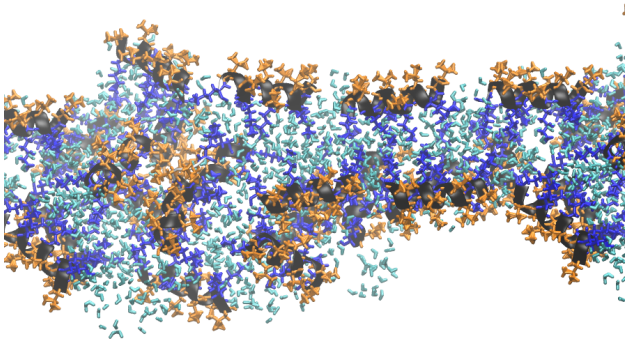
Figure 11.7.: The number of peptides in the largest cluster as function of time for pure ethanol.

Figure 11.9 shows two snapshot of simulation boxes showing all molecules in the simulation, one with $x_{\mathrm{Eth}} = 0.4$ and one with $x_{\mathrm{Eth}} = 0.7$. Both show well how two different domains have formed in the solutions, one water-rich inside the cluster and one ethanol-rich outside of it. Comparing figures 11.9a and 11.9b shows that there is a greater difference in the composition of the ethanol-rich domain than the water-rich domain. In both cases the peptides lie parallel to the interface between the the two domains with the lysine residues oriented in towards the water-rich domain and the leucine residues out towards the ethanol-rich domain.

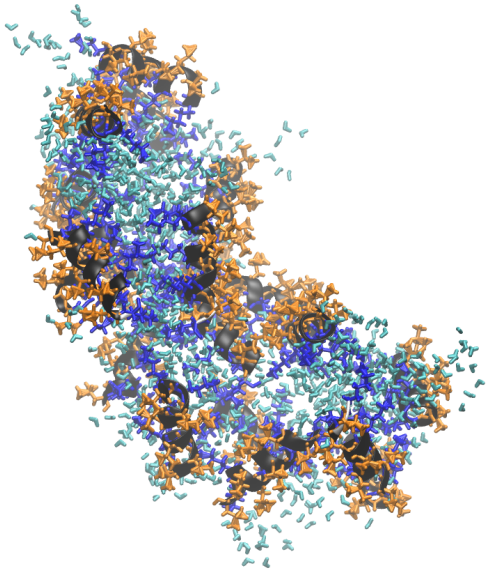
11.2.1. Size

Section 6.4.1 details how the clusters where defined in the analysis. The cluster analysis was performed for the solutions with $x_{\mathrm{Eth}} = 0.3$ to $x_{\mathrm{Eth}} = 0.9$. In pure ethanol it is not possible for clusters of peptides and water to form since there is no water. For compositions with low x_{Eth} it is impossible to distinguish between water molecules in a cluster and out of it. As detailed in section 11.1.1 the aggregate at low x_{Eth} are also of a different type: simple oligomers instead of clusters involving both peptides and water. Clusters involving both peptides and ethanol are not formed at any x_{Eth}.

Figure 11.10 shows snapshots of the simulation with $x_{\mathrm{Eth}} = 0.3$, figure 11.10a shows peptides only while figure 11.10b shows both peptides and water molecules. Even though

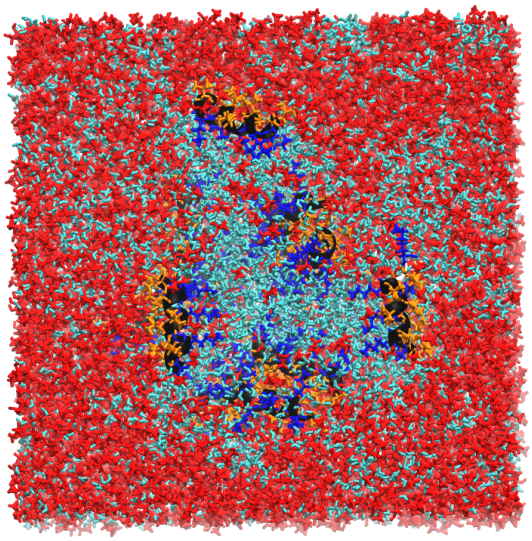


(a) A cluster that connects to itself over the periodic boundary. The snapshot shows a $3 \times 1 \times 1$ lattice of the simulation box.

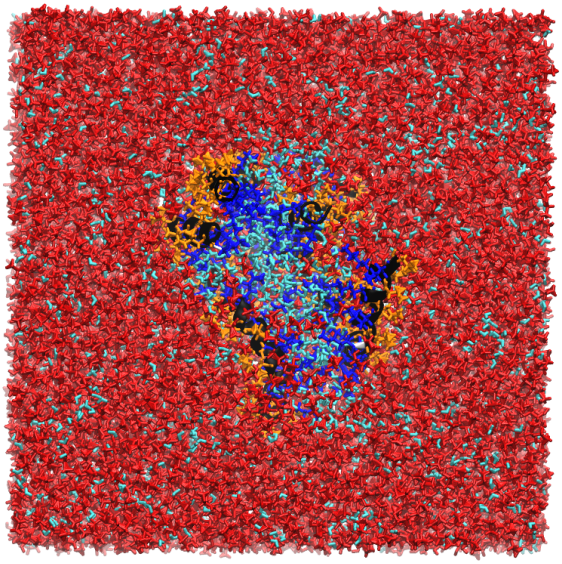


(b) A cluster which does not connect to itself over the periodic boundary.

Figure 11.8.: Two snapshots of clusters of peptides and water molecules in a simulation with mole fraction ethanol $x_{\mathrm{Eth}} = 0.7$. For water molecules to be shown they have to be within 0.2 nm of the peptides or other water molecules which are shown. The lysine side chains are shown in blue and the leucine side chains in orange, water molecules are shown in cyan.

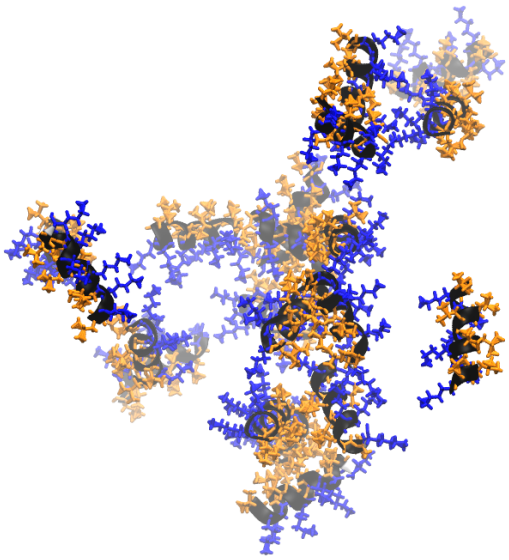


(a) $x_{\mathrm{Eth}} = 0.4$.

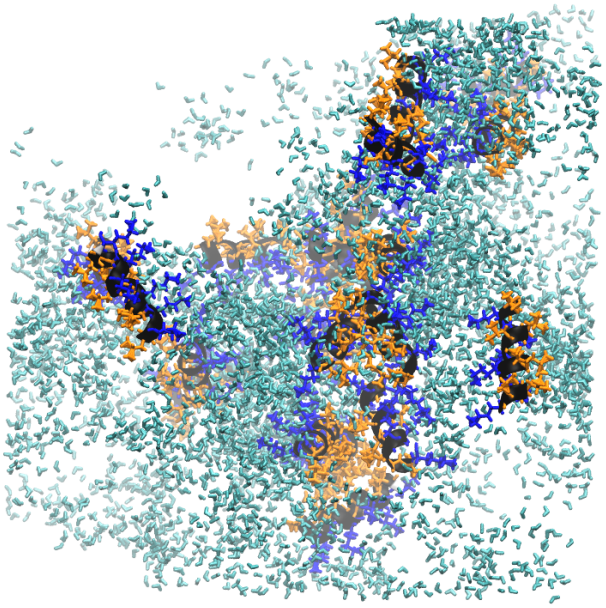


(b) $x_{\mathrm{Eth}} = 0.7$

Figure 11.9.: Snapshots of the simulations with mole fraction ethanol $x_{\mathrm{Eth}} = 0.4$ and $x_{\mathrm{Eth}} = 0.7$. The lysine side chains are shown in blue and the leucine side chains in orange, water molecules are shown in cyan and ethanol molecules in red.

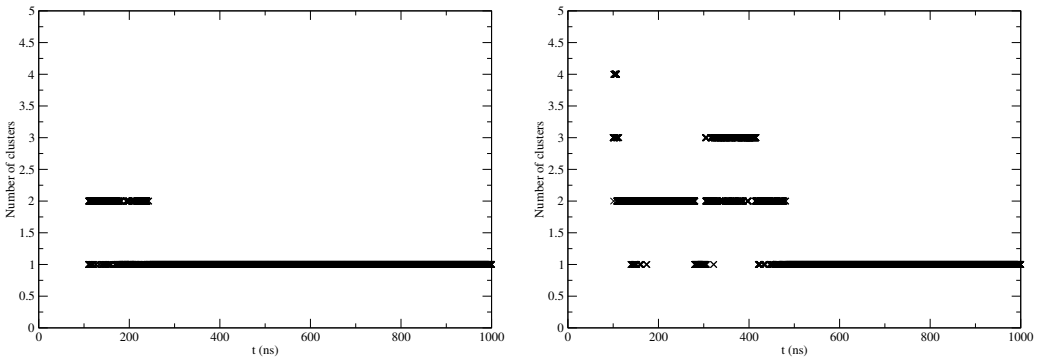


(a) Only peptides.



(b) Peptides and water molecules.

Figure 11.10.: A snapshots of the peptides and water molecules close to them in a simulation with mole fraction ethanol $x_{\mathrm{Eth}} = 0.3$. For water molecules to be shown they have to be within 0.2 nm of the peptides or other water molecules which are shown. The lysine side chains are shown in blue and the leucine side chains in orange, water molecules are shown in cyan.



(a) $x_{\mathrm{Eth}} = 0.65$. (b) $x_{\mathrm{Eth}} = 0.9$

Figure 11.11.: Number of clusters found as a function of simulation time.

the clusters analysis was performed on the trajectory of the simulation with $x_{\mathrm{Eth}} = 0.3$ it is clear from figure 11.10a that some of the aggregates found in section 11.1 for this solvent composition have configurations of peptides which are more reminiscent of the oligomers in figure 11.3 than the clusters in figure 11.8. The analysis method described in section 6.4.1 is guarantied to find one single cluster in a solution with very little ethanol no matter whether a cluster is present or not (since it depends on absence of ethanol as a part of the definition of a cluster). From figure 11.10 this seems to be the case for $x_{\mathrm{Eth}} = 0.3$ though the results from the analysis are included in the subsequent discussions.

The simulations the cluster analysis was performed on were 1000 ns long. The first 100 ns were not used for the cluster analysis since at the start of the simulations the peptides were randomly distributed in the box. For solutions with $x_{\mathrm{Eth}} = 0.55$ to $x_{\mathrm{Eth}} = 0.7$, two simulations were performed from the same starting configuration for reasons detailed further down (section 11.2.3). For 10 out of 15 trajectories which this analysis was performed on, not a single frame was found with more than one cluster. Of the rest one trajectory had a single frame with two clusters and in another trajectory two frames had two clusters. In three trajectories a significant amount of frames have more than one cluster. Those are the trajectories with $x_{\mathrm{Eth}} = 0.8$ and $x_{\mathrm{Eth}} = 0.9$ and one of the two trajectories with $x_{\mathrm{Eth}} = 0.65$ (average number of clusters 1.02, 1.47 and 1.09 respectively). In all three cases the size of the clusters was not converged after 100 ns. Figure 11.11 shows the number of clusters as a function of time for $x_{\mathrm{Eth}} = 0.65$ and $x_{\mathrm{Eth}} = 0.9$. It can clearly be seen that the number of clusters converges to a single cluster in both trajectories. Comparing figures 11.11a and 11.11b also shows that while it took 250 ns for the number of clusters to converge in this simulation with $x_{\mathrm{Eth}} = 0.65$ it took 500 ns when $x_{\mathrm{Eth}} = 0.9$ and only when $x_{\mathrm{Eth}} = 0.9$ are there any frames where more than 2 clusters were found after 100 ns. The case when $x_{\mathrm{Eth}} = 0.8$ is the same as when $x_{\mathrm{Eth}} = 0.6$ except it only takes 140 ns for the number of clusters to converge to 1.

As mentioned in section 11.1.2, 1000 ns were not enough to converge the size of the

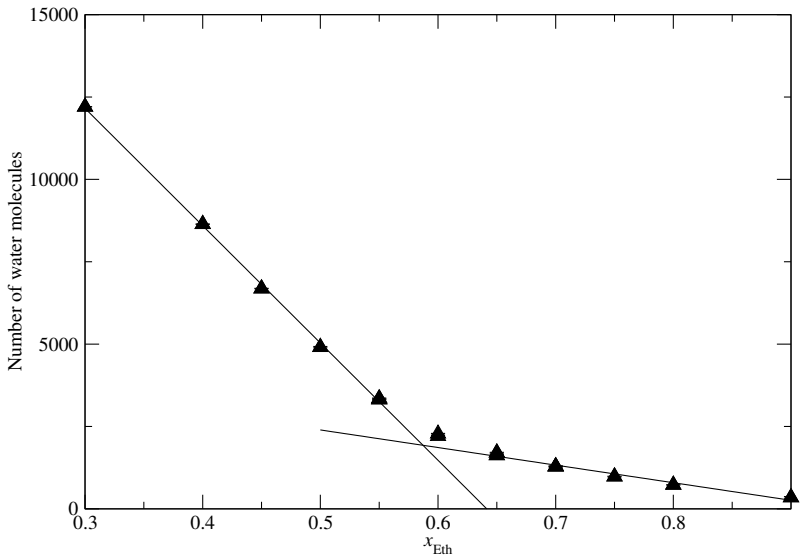


Figure 11.12.: The average number of water molecules in the largest cluster as a function of x_{Eth}. The lines show fits through the points from $x_{\mathrm{Eth}} = 0.3$ to $x_{\mathrm{Eth}} = 0.55$ and from $x_{\mathrm{Eth}} = 0.65$ to $x_{\mathrm{Eth}} = 0.9$.

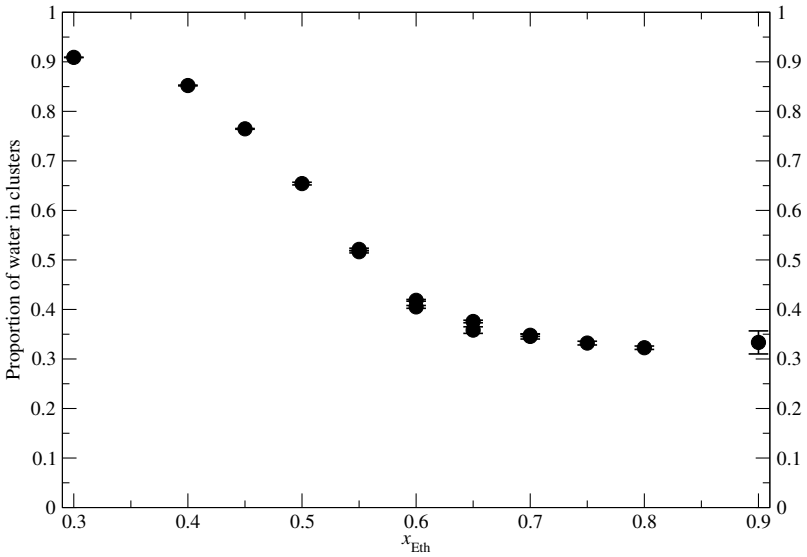


Figure 11.13.: The proportion of water molecules in the largest cluster out of all water molecules in the simulation box as a function of x_{Eth}.

aggregates in pure ethanol. This and the fact that significantly longer time is needed to converge the size of the clusters when $x_{\mathrm{Eth}} = 0.9$ than in any other mixture suggests that the presence of water facilitates faster formation of clusters than when only ethanol is present.

When only a single cluster was found all 20 peptides present in the simulation were part of that cluster. This was the case for all the solvent compositions.

Figure 11.12 shows the number of water molecules in the largest cluster as a function of x_{Eth}. Two distinct regions can be seen in the figure. From $x_{\mathrm{Eth}} = 0.3$ to $x_{\mathrm{Eth}} = 0.55$ and from $x_{\mathrm{Eth}} = 0.65$ to $x_{\mathrm{Eth}} = 0.9$ the number of water molecules in the largest cluster falls linearly with increasing x_{Eth}, though as the can be seen in figure 11.12 the slopes of the lines are different in the two regions. The number of water molecules is different in the different solvent compositions. Therefore figure 11.13 shows which proportion of all the water molecules in the simulation box are in the largest cluster. The same two regions can be seen in figure 11.13 as in figure11.12. For $x_{\mathrm{Eth}} > 0.6$, around a third of the water molecules in the simulation box are in the cluster while for the lower x_{Eth} a larger part of the available water molecules is recognised as being part of the cluster. For $x_{\mathrm{Eth}} = 0.3$ to $x_{\mathrm{Eth}} = 0.6$ the lower x_{Eth} is, the larger is the proportion of the water molecules in the simulation box which are part of the cluster.

11.2.2. Orientation of the peptides

As can be seen in all snapshots of the clusters the peptides tend to lie in the interface between the water-rich and the ethanol-rich domains with the lysine side chains oriented inwards towards the water molecules inside the clusters and the leucine side chains oriented out towards the ethanol-rich domain. In the same trajectories as in sections 11.2.1 and 11.2.3 the orientation of the peptides was analysed.

Section 6.4.1 details how the simulation box was overlayed with grid points and these grid points associated with the clusters. Grid points which are part of a cluster and have fewer than 26 neighbours which are also part of the same cluster are considered on the surface of the cluster. Figure 11.14 shows the proportion of surface points of the largest cluster which are closer to a lysine residue than a leucine residue as a function of x_{Eth}. There is no solvent composition where more than 3.1 % of all surface points of the largest cluster are closer to a lysine residue than a leucine residue. This means that almost all of the surface of the clusters is closer to a leucine residue than a lysine residue, which in turn means that the peptides are oriented with the leucine residues facing out towards the ethanol-rich domain.

Figure 11.14 shows a clear difference in prevalence of surface points which are closer to a lysine residue than a leucine residue between solvent compositions with $x_{\mathrm{Eth}} \geq 0.65$ and with $x_{\mathrm{Eth}} \leq 0.6$. This is the same break point as seen in figures 11.12 and 11.13 in section 11.2.1. When $x_{\mathrm{Eth}} \geq 0.65$ the proportion of surface points which are closer to a lysine residue is 0.16 % or smaller. In these compositions the surface of the clusters is small and closely packed with peptides as seen in figure 11.15a while at lower x_{Eth} the peptides are further apart from each other and more of the interface between the ethanol-rich domain of the solution and the water-rich inside of the cluster is a direct

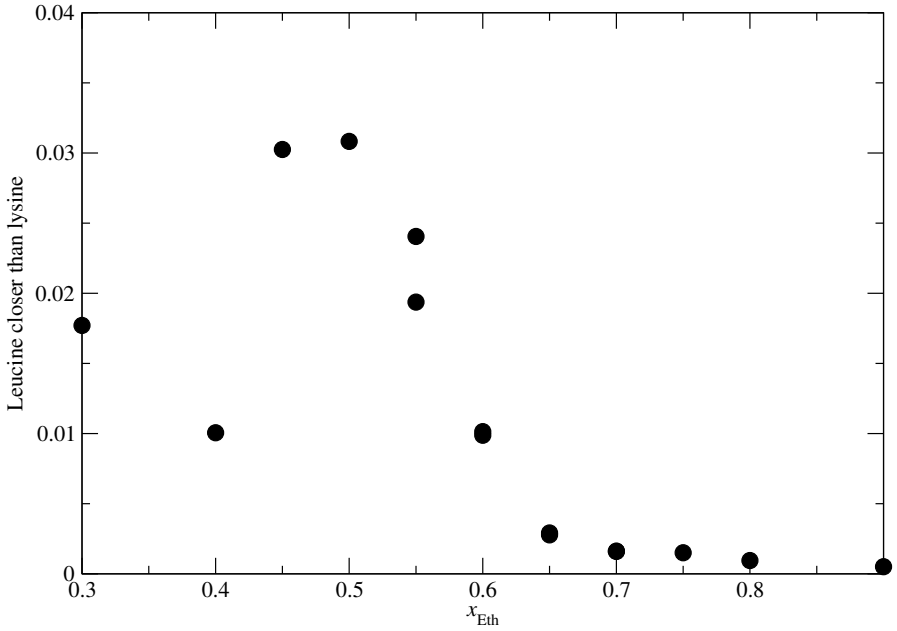
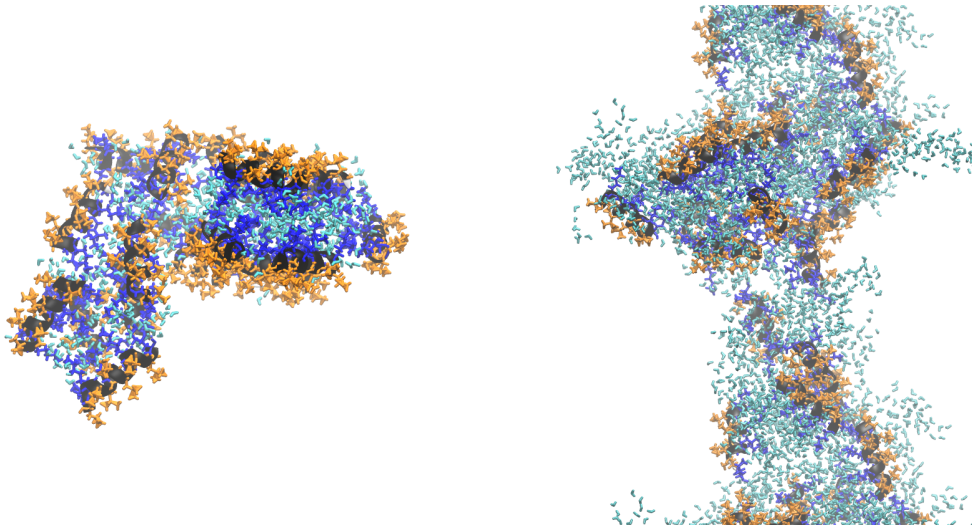


Figure 11.14.: The proportion of surface points of the largest cluster which are closer to a lysine residue than a leucine residue.

interface between these two domains without peptides in between (see figure 11.15b). This difference can also be seen by comparing figures 11.9a and 11.9b. This explains why at lower x_{Eth} slightly more surface points are closer to a lysine residue than a leucine residue. If the surface of the cluster is curved (partly concave) a point on the surface without a peptide can be quite close to a lysine residue located at a different part of the surface (the surface of the clusters is often concave, see figures 11.8, 11.9a, 11.16 and 11.15).

Another aspect of the conformation of the peptides in clusters at high x_{Eth} which can be clearly seen in figures 11.8, 11.9b and 11.15a is that the peptides form two layers with water in between similar to bilayers formed by surfactants (though far less extensive than surfactant bilayer formations tend to be). This is less noticeable at lower x_{Eth} as can be seen in figures 11.9a, 11.16 and 11.15b.

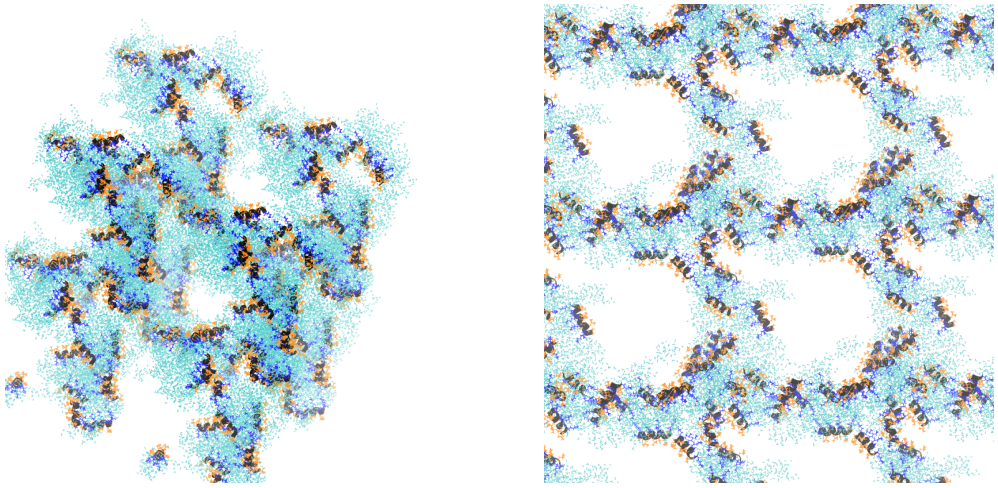
The reason for the difference in both how much of the surface of the clusters is covered by peptides and the tendency to form bilayers-like structures seems to be the fact that more water molecules are part of the clusters at low x_{Eth} (both more water molecules in the system and a higher proportion of them in the clusters). This means that the peptides can be further apart from each other and still be part of the same cluster which means the conformation of the peptides with regards to each other is more flexible (compare figures 11.9a and 11.9b). This also explains why more smaller aggregates where found between $x_{\mathrm{Eth}} = 0.3$ and $x_{\mathrm{Eth}} = 0.6$ when only the peptides were considered, while only one large cluster was found when the water was considered as well.



(a) A cluster in a solution with $x_{\mathrm{Eth}} = 0.8$. The cluster is clearly defined and the surface closely packed with peptides.

(b) A cluster in a solution with $x_{\mathrm{Eth}} = 0.45$. More of the surface is exposed to the rest of the solution without peptides in between and which water molecules should be counted as part of the cluster is a not as easily defined. As the cluster connects with itself across the periodic boundary, additional simulation boxes were added to the edge of the figure in the direction of the connection.

Figure 11.15.: Two snapshots of clusters of peptide and water clusters in simulations with mole fraction ethanol $x_{\mathrm{Eth}} = 0.8$ and $x_{\mathrm{Eth}} = 0.5$. For water molecules to be shown they have to be within 0.2 nm of the peptides or other water molecules which are shown, this distance is not calculated over the periodic boundary. The lysine side chains are shown in blue and the leucine side chains in orange, water molecules are shown in cyan.



(a) A cluster which connects to itself over the periodic boundary in all three cardinal directions, $x_{\mathrm{Eth}} = 0.4$. The snapshot shows a $2 \times 2 \times 2$ lattice of the simulation box. All three directions show different thickness of connections.

(b) A cluster which connects to itself over the periodic boundary in two directions, $x_{\mathrm{Eth}} = 0.45$. The snapshot shows a 3×3 lattice of the simulation box, in the two directions in which the cluster connects to itself. Both directions show similarly wide connections.

Figure 11.16.: Two snapshots of clusters of peptide and water clusters in simulations with mole fraction ethanol $x_{\mathrm{Eth}} = 0.4$ and $x_{\mathrm{Eth}} = 0.45$. For water molecules to be shown they have to be within 0.2 nm of the peptides or other water molecules which are shown, this distance is not calculated over the periodic boundary. The lysine side chains are shown in blue and the leucine side chains in orange, water molecules are shown in cyan.

11.2.3. Form

As mentioned in section 11.2 the clusters can connect to themselves over the periodic boundary conditions. The clusters can percolate this way in one, two or three directions or not at all. Figure 11.16 shows examples of percolation in two and three directions in solutions with $x_{\mathrm{Eth}} = 0.45$ and $x_{\mathrm{Eth}} = 0.4$ respectively. Examples of percolation in one direction and no percolation at all in a solution with $x_{\mathrm{Eth}} = 0.7$ can be seen in figure 11.8.

In order to quantify the percolation the trajectories of section 11.2.1 were analysed. Figure 11.17 shows the proportion of clusters which percolate in one, two or three directions or not at all as function of x_{Eth}. For four solvent compositions $x_{\mathrm{Eth}} = 0.55$, $x_{\mathrm{Eth}} = 0.6$, $x_{\mathrm{Eth}} = 0.65$ and $x_{\mathrm{Eth}} = 0.7$ three points are shown in figure 11.17. For these solvent compositions two simulations were performed starting from the same equilibrated simulation box. Figure 11.17 shows the results of both, as well as the average values of

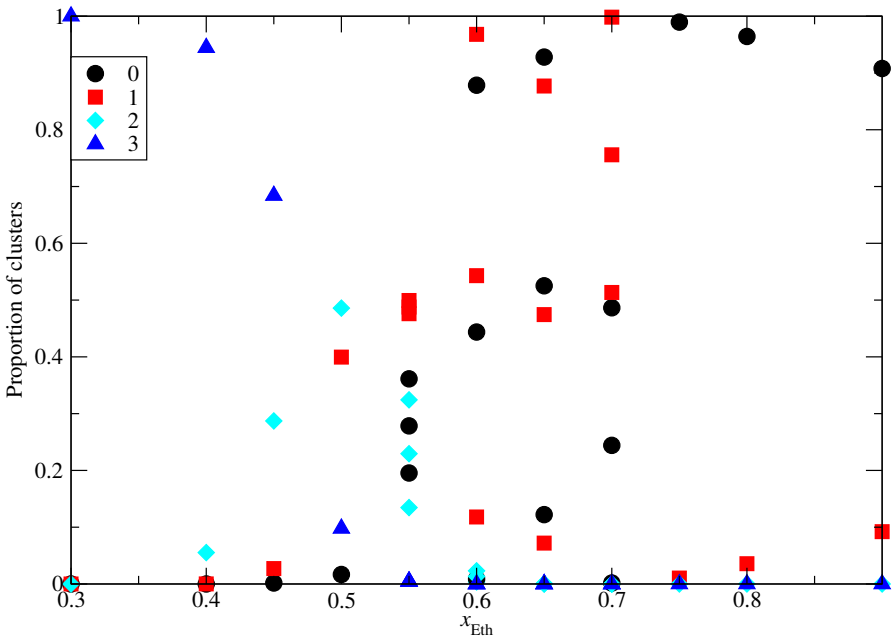
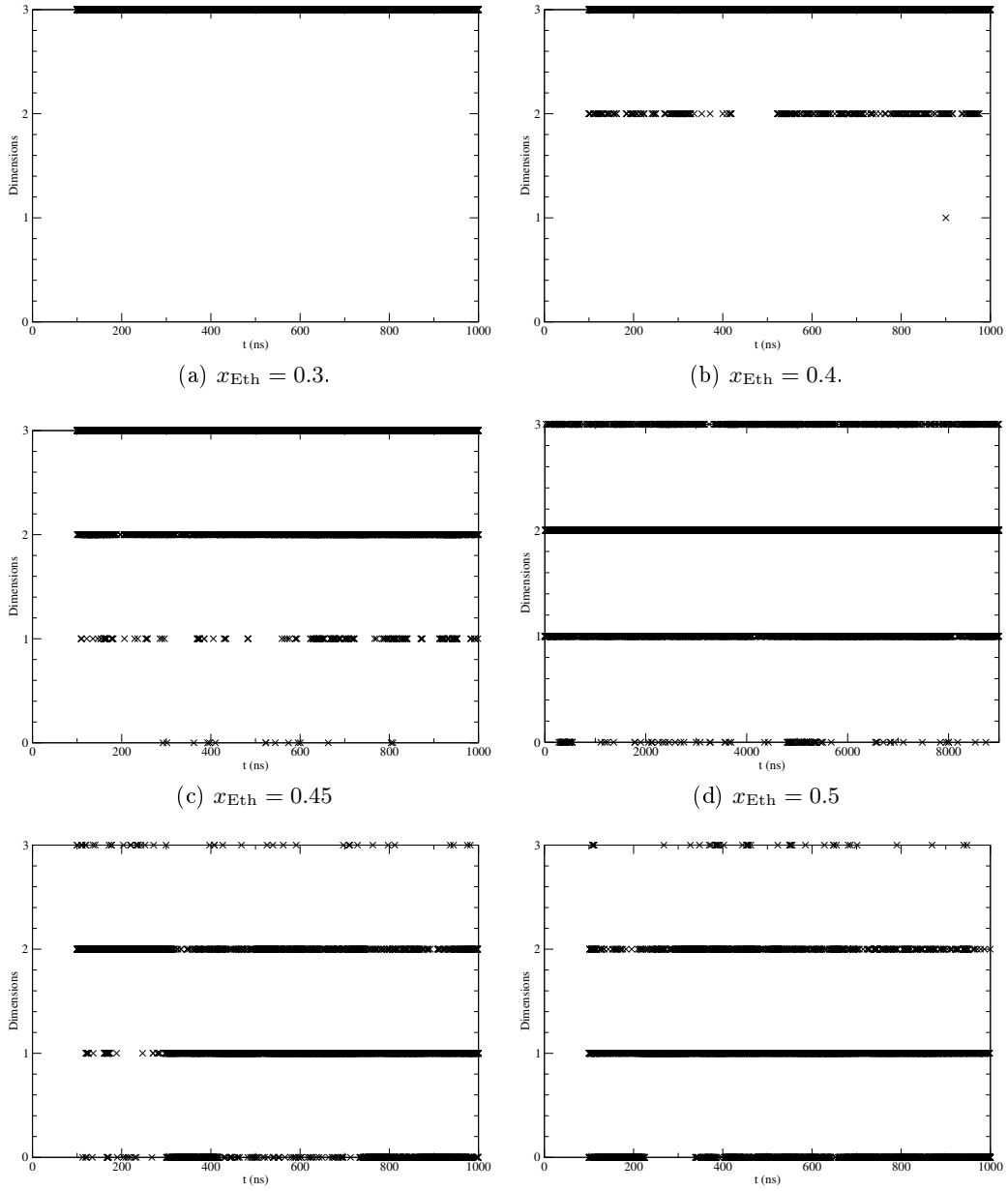


Figure 11.17.: The proportion of clusters which percolate in 1, 2, or 3 dimensions as well as the proportion of clusters which does not percolate at all (0 in the legend) as a function of the ethanol mole fraction, x_{Eth}.

the two simulations. As can be seen for both $x_{\mathrm{Eth}} = 0.6$ and $x_{\mathrm{Eth}} = 0.65$ in one simulation over 85 % of all clusters percolate in one direction while in the other over 85 % of all clusters do not percolate at all. For both $x_{\mathrm{Eth}} = 0.55$ and $x_{\mathrm{Eth}} = 0.7$ there are also clear differences between the two simulations but not as extreme as for $x_{\mathrm{Eth}} = 0.6$ and $x_{\mathrm{Eth}} = 0.65$. The reason the second simulations were performed is that after the first simulation the results were that for $x_{\mathrm{Eth}} = 0.55$ the clusters most commonly percolated in one direction, for $x_{\mathrm{Eth}} = 0.6$ no percolation was the most common, for $x_{\mathrm{Eth}} = 0.65$ percolation in one direction was again the most common and both percolation in one direction and no percolation were about as likely for $x_{\mathrm{Eth}} = 0.7$. As can be seen in figure 11.17 the second simulations changed this radically. In many or most of the trajectories analysed the percolation of the clusters can not be said to be quantitatively converged and statistical analysis in order to estimate error bars for figure 11.17 would not result in any useful information. Qualitative conclusions about the percolation of the clusters can be drawn even from simulations where the percolation has not quantitatively converged and will be described in the following text.

Figures 11.18, 11.19 and 11.20 show in how many dimensions the clusters percolate as a function of time for all trajectories in which the clusters were analysed. The only solvent composition which does not fluctuate between different numbers of dimensions at all is $x_{\mathrm{Eth}} = 0.3$ as seen in figure 11.18a. As discussed in section 11.2.1 it is not quite clear whether this composition should be considered when discussing the clusters; if it



(a) $x_{\mathrm{Eth}} = 0.3$.

(b) $x_{\mathrm{Eth}} = 0.4$.

(c) $x_{\mathrm{Eth}} = 0.45$

(d) $x_{\mathrm{Eth}} = 0.5$

(e) $x_{\mathrm{Eth}} = 0.55$.

(f) $x_{\mathrm{Eth}} = 0.55$.

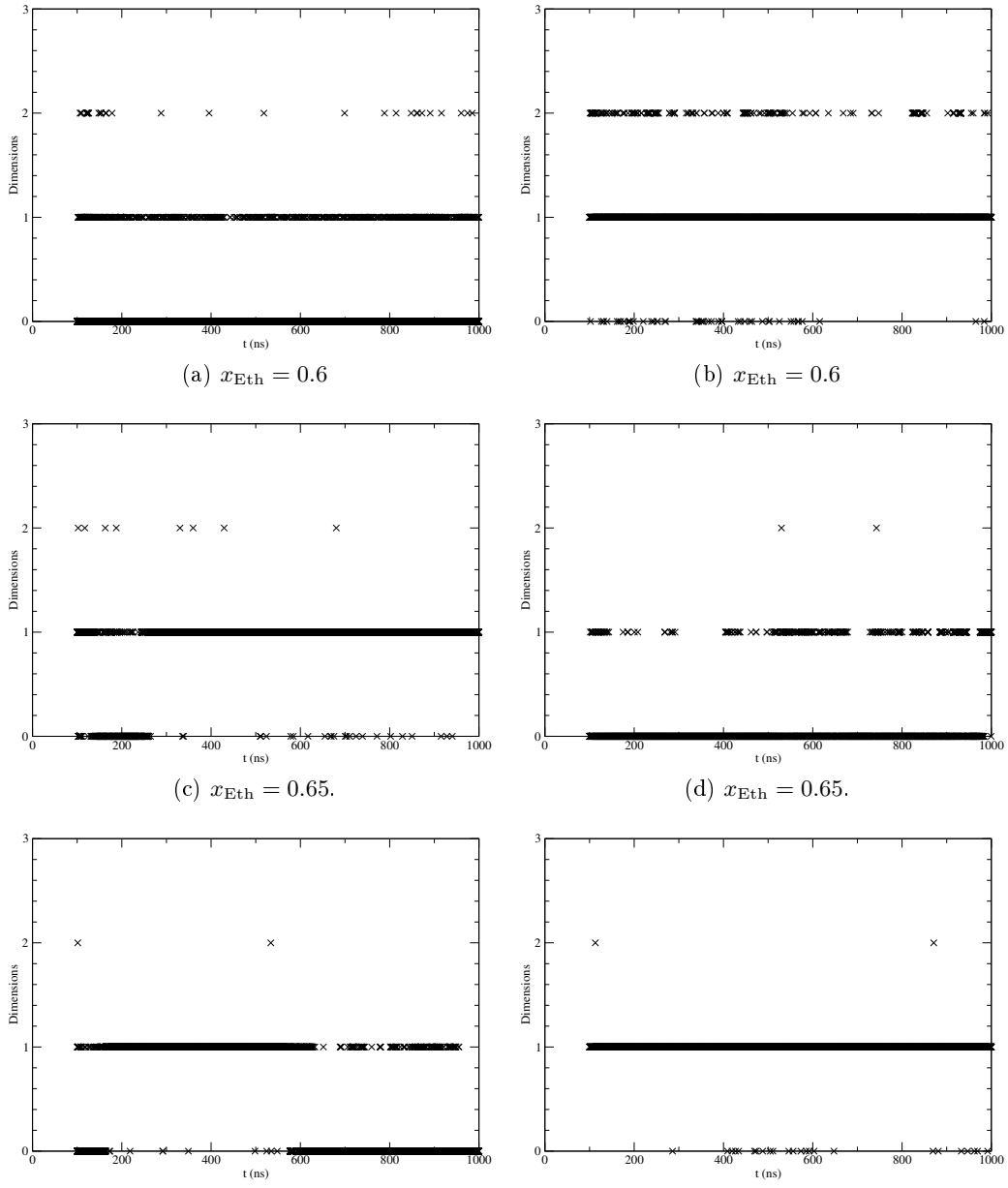
Figure 11.18.: Number of dimensions the cluster percolates in as a function of simulation time.

is, it is clear that here the solution forms a bicontinuous system.

For solvent compositions with x_{Eth} from 0.4 to 0.55 the sampling of the different number of dimensions in which the clusters percolate seems to be quite good judging from figure 11.18, though comparison with figures 11.19c and 11.19d reveals that this is not a guaranty of convergence. Of those solvent compositions only $x_{\mathrm{Eth}} = 0.55$ was simulated twice. Comparing figures 11.18e and 11.18f reveals no clear difference between the two simulations and figure 11.17 also shows similar results for both simulations of this solvent composition.

From $x_{\mathrm{Eth}} = 0.4$ to $x_{\mathrm{Eth}} = 0.55$ there is a trend from a bicontinuous system in which the clusters percolate in all three dimensions through a system in which the clusters form a two dimensional grid to a system of either endless cylindrical clusters or discrete cylindrical clusters. At $x_{\mathrm{Eth}} = 0.4$ the bicontinuous system dominates. The same applies to $x_{\mathrm{Eth}} = 0.45$ though here a two dimensional grid appears about 30 % of the time (figure 11.16b shows an example of such a grid). When $x_{\mathrm{Eth}} = 0.5$ the two dimensional grid and percolation in only one dimension are both quite common (40 - 50 % of the time) while percolation in all three dimensions only appears about 10 % of the time. In all three of those solvent compositions the amount of time when the clusters do not percolate at all is not significant (when $x_{\mathrm{Eth}} = 0.4$ the clusters always percolate in at least one direction). At $x_{\mathrm{Eth}} = 0.55$ percolation in one dimension appears about 50 % of the time while percolation in two dimensions and no percolation at all both account for about 25 % of the trajectory and percolation in all three dimensions does not account for a significant fraction of the trajectory.

Figure 11.19 shows the number of dimensions the clusters percolate in as a function of time for the rest of the solvent compositions which were simulated twice ($x_{\mathrm{Eth}} = 0.6$, $x_{\mathrm{Eth}} = 0.65$ and $x_{\mathrm{Eth}} = 0.7$). In contrast to $x_{\mathrm{Eth}} = 0.55$ figures 11.19c and 11.19d show a clear difference between the two simulations. The same difference can also be seen between the different values at $x_{\mathrm{Eth}} = 0.6$ in figure 11.17. In the first simulation (figure 11.19a) no percolation at all dominates, though percolation in one dimension also appears about 10 % of the time (evenly distributed throughout the simulation). In the second simulation on the other hand percolation in one dimension dominates while no percolation at all does not appear for any significant amount of time. This means that when averaging both simulations there is not a great difference between the proportions of clusters which percolate in one dimension and which do no percolate at all. In both simulations percolation in two dimensions also appears, though only for about 2 % of the time. Figures 11.19c and 11.19d also show a similar difference between the two simulations when $x_{\mathrm{Eth}} = 0.65$. In contrast to when $x_{\mathrm{Eth}} = 0.6$ figure 11.19c also shows a change in the distribution after about 250 ns in the first simulation. Before this time the distribution is similar to the distribution in the second simulation seen in figure 11.19d with no percolation at all dominating, while after this time percolation in one dimension dominates to an even greater degree than no percolation at all dominates in the second simulation. Figure 11.17 reveals that in each case the more common configuration appears around 90 % of the time. For $x_{\mathrm{Eth}} = 0.7$ the situation is again similar in that two distinct distributions can be seen. One in which percolation in one dimension dominates quite



(a) $x_{\mathrm{Eth}} = 0.6$ (b) $x_{\mathrm{Eth}} = 0.6$

(c) $x_{\mathrm{Eth}} = 0.65$. (d) $x_{\mathrm{Eth}} = 0.65$.

(e) $x_{\mathrm{Eth}} = 0.7$ (f) $x_{\mathrm{Eth}} = 0.7$

Figure 11.19.: Number of dimensions the cluster percolates in as a function of simulation time.

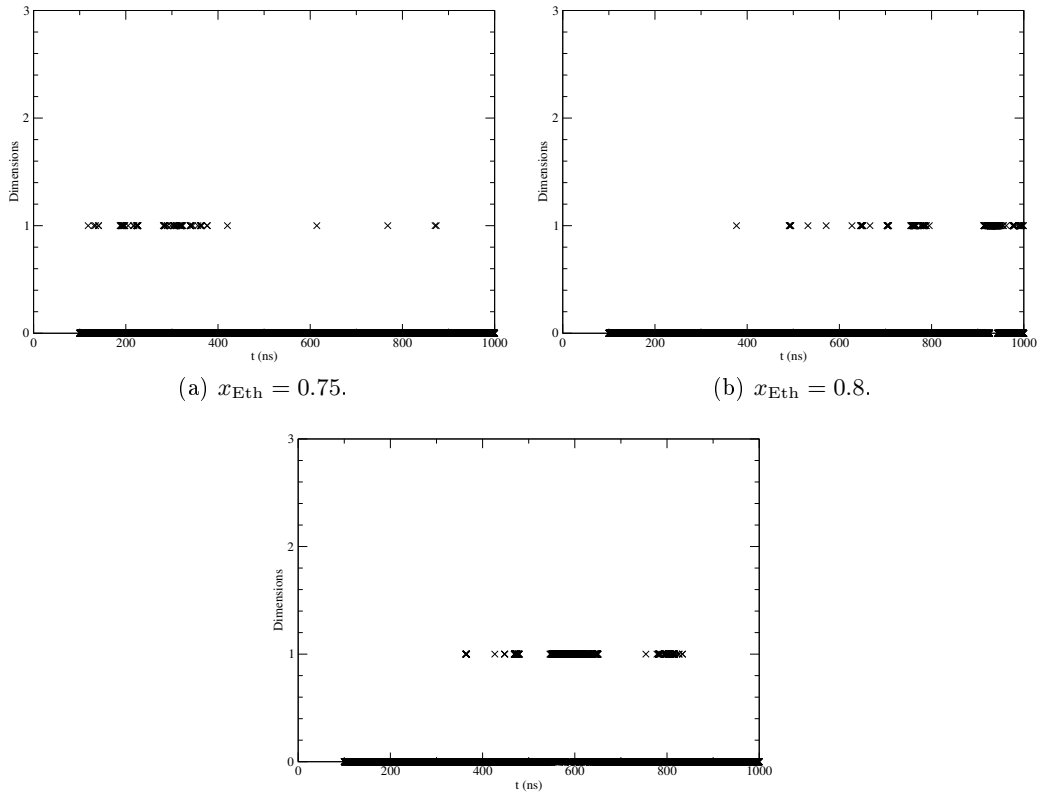
strongly and one in which no percolation at all dominates but a significant amount of time the cluster percolates in one dimension. The second simulation is in the first distribution the whole simulation time (figure 11.19f) while the first simulation changes between the two distributions two times: at about 160 ns it changes from the second distribution to the first one and at about 580 - 620 ns it changes back, snapping between the two distributions a few times while doing so.

All three solvent compositions thus seem able to be in two distinct states associated with two distinct distribution of number of dimensions in which the clusters percolate. In all cases, the change between these two states is so infrequent that two times 1000 ns is not enough to quantitatively converge the distribution of them (but the distribution of the number of dimension in each state is converged, which allows us to draw some conclusions about the states themselves). In each case percolation in one dimension dominates one state and no percolation at all dominates the other. For the discussion of these two states the one dominated by percolation in one dimension will be called state (I) and the state dominated by no percolation at all will be called state (0).

When $x_{\mathrm{Eth}} = 0.6$ state (I) more often shows percolation in two dimensions than in none at all, while state (0) shows a significant amount of percolation in one dimension. When $x_{\mathrm{Eth}} = 0.65$ both states seem more strongly dominated by their respective defining percolation patterns, though this difference to $x_{\mathrm{Eth}} = 0.6$ is more visible for state (0) than for state (I). Here state (I) also clearly shows less percolation in two dimensions than when $x_{\mathrm{Eth}} = 0.6$ (compare figures 11.19b and 11.19c). Both states seem quite similar when $x_{\mathrm{Eth}} = 0.7$ as when $x_{\mathrm{Eth}} = 0.65$. The state (I) is more strongly dominated by percolation in one dimension than when $x_{\mathrm{Eth}} = 0.6$ and state (0) seems more dominated by no percolation at all than when $x_{\mathrm{Eth}} = 0.6$ as well.

Figure 11.20 shows the number of dimensions the clusters percolate in as a function of time for the solvent compositions with $x_{\mathrm{Eth}} = 0.75$, $x_{\mathrm{Eth}} = 0.8$ and $x_{\mathrm{Eth}} = 0.9$. These solvent compositions were only simulated once each. All three trajectories are dominated by no percolation at all. When comparing figure 11.20 with states (0) in figure 11.19 it can be seen that the trajectories in figure 11.20 are more strongly dominated by no percolation at all than states (0) of the solvent compositions in figure 11.19 are. The second simulation of $x_{\mathrm{Eth}} = 0.6$ was in state (0) for the whole simulation time. Therefore the absence of as state comparable to state (I) in figure 11.20 does not prove that such a state does not exist for these solvent compositions. But its absence from all three simulations and the very strong domination of no percolation at all does suggest that if it exist at all for these solvent compositions, it is less easily accessible than at lower x_{Eth}.

The general trend for the percolation of the clusters in all solvent compositions is a change from a bicontinuous system at low x_{Eth} through a system percolating in two dimensions and then in only one dimension to a non-percolating system at high x_{Eth}. The number of percolation dimensions seems to converge better at low x_{Eth} where higher number of percolation dimensions is found than when percolation in one dimension or no percolation at all dominates (though the convergence seems better again at very high x_{Eth} where no percolation at all dominates very strongly). The change from percolation in more than one dimension being common to percolation in only one or no dimensions



(a) $x_{\mathrm{Eth}} = 0.75$.

(b) $x_{\mathrm{Eth}} = 0.8$.

(c) $x_{\mathrm{Eth}} = 0.9$

Figure 11.20.: Number of dimensions the cluster percolates in as a function of simulation time.

at all dominating coincides with change between the two regions seen in figures 11.12, 11.13 and 11.14 in sections 11.2.1 and 11.2.2.

11.3. Discussion

The size, form and convergence time of aggregates formed by LKα14 as well as the orientation of the peptides in these aggregates depends on the composition of the solvent. In pure water and at low x_{Eth} the peptides form oligomers with the lysine residues oriented out toward the solvent while at high x_{Eth} and in pure ethanol larger aggregates are formed in which the lysine residues are oriented in towards the centre of the aggregate. At high x_{Eth} water is part of the aggregates. This is of course not possible in pure ethanol.

The clusters formed at higher x_{Eth} are reminiscent of microemulsions (specifically bicontinuous or water in oil microemulsions) or — at very high x_{Eth} — inverse micelles. Giving that the packing parameter of surfactants is one of the main methods of classifying micelles and predicting their curvature and form it would therefore seem reasonable to calculate the packing parameter of LKα14. This is not possible since it is difficult to define what a, v and l should be in equation 6.1 for LKα14. The equation expects a typical surfactant which has a comparatively bulky head group, which is in contact with the water, and long hydrocarbon tail. In such a surfactant the spatial boundary between the hydrophilic and hydrophobic parts of the molecule is perpendicular to the longest dimension of the molecule (if the tail would be stretched out)[23,90]. In an α-helical LKα14 in contrast, the spatial boundary between the hydrophilic and hydrophobic parts of the molecule is parallel to the helix axis and the helix axis is the longest dimension of the molecule. The α-helix is also rather inflexible and rod-like while the hydrocarbon chain of a typical surfactant is flexible. Therefore a typical surfactant lies perpendicular to the surface of the micelle / microemulsion droplet[23,90] while, as detailed in section 11.2.2, the α-helix lies parallel to the surface of the cluster. This means that it is for example meaningless to talk about the length of the hydrophobic tail in a LKα14 α-helix, the helix does not have a hydrophobic tail but rather a hydrophobic half. As such no packing parameters were calculated for LKα14. The systems equivalence to the critical micelle concentration was not calculated either, since that would have increased the already large simulation effort considerably.

Like in microemulsions[23,25,27,89,90] the structure of the clusters depends on the composition of the system as a whole. At $x_{\mathrm{Eth}} > 0.6$ about a third of all water molecules in the simulation box are part of the clusters. At $0.3 < x_{\mathrm{Eth}} \leq 0.6$ a larger proportion of the water molecules in the box than a third are part of the clusters and the proportion of water molecules in the clusters decreases with increasing x_{Eth}. At $x_{\mathrm{Eth}} \geq 0.6$ the clusters do not percolate in more than 1 direction for any significant amount of time and the peptides were close enough to each other that a cluster analysis involving only the peptides and no water molecules found the average number of peptides in the largest cluster to be over 17. At $0.3 < x_{\mathrm{Eth}} < 0.6$ percolation in more dimensions is more common (at $0.3 < x_{\mathrm{Eth}} \leq 0.45$ percolation in all three dimensions is the most common) and the peptides are on average further apart from each other so that a cluster analysis

involving only the peptides finds smaller clusters. This indicates a phase change, from a system similar to a bicontinuous microemulsion to a system that resembles a swollen inverse micelle.

The rod-like nature of a LKα14 α-helix also influences the form of the clusters. The rod-like structure of the clusters seen in figures 11.8 and 11.15a, both in discrete clusters and in ones that connect over the periodic boundary, can easily be explained by the curvature of the surface of the cluster be direction depended, it is more curved in the direction perpendicular to the axis of the helices than the one parallel to the axis. A large curvature parallel to the helix's axis would mean that either the α-helix would have to be bent or only a part of the peptide would in fact lie in the surface of the cluster, while a curvature parallel to the helix's axis can be achieved by the peptides being oriented at an angle to each other. The orientation seen at high x_{Eth}, as shown in figures 11.8 and 11.15a, is in fact peptides on two sides of the cluster, with the lysine side chains oriented towards the peptide on the other side, reminiscent of a very small bilayer. The bilayer structure is less pronounced at lower x_{Eth} (see figures 11.15b and 11.16). Looking along the cluster the peptides form more of a circle around the water-rich domain at lower x_{Eth} and an ellipsis or two flat surfaces on either side of the water-rich domain at higher x_{Eth} (compare figures 11.9a and 11.9b).

Adding ethanol to water has been shown to influence the critical micelle concentration of surfactants[136,137]. With added ethanol the critical micelle concentration is increased, except at very low x_{Eth}[136–139]. None of these studies reports inverted micelles or either solvent component participating in the micelles in any way[136–139]. Ethanol has also been shown to act similar to surfactants in some microemulsion-like systems[53–55]. In contrast, here the amphiphility of the α-helices separates the fully miscible ethanol and water into two distinct domains and the ethanol takes on the role of oil when comparing the system to microemulsions or micelles (though there still is a significant amount of water molecules in the ethanol-rich domain). In the simulations of one peptide in bulk presented in chapter 8 no such separation of the ethanol and the water took place. It can therefore be concluded that the presence of the peptides is necessary for the separation.

Part IV.

Conclusion

12. Conclusion

In this work LKα14 was simulated in different mixtures of water and ethanol, both with and without an air/solution interface. The secondary structure of the peptides was calculated in bulk and at the surface. The adsorption to the surface and the orientation of α-helical LKα14 at the surface was calculated, both the tilt of the α-helices and which amino acids are oriented towards the bulk and which towards the air. The excess solvation of LKα14 by either water or ethanol was determined as well as the spatial separation of water and ethanol in the solvation shell. Finally the aggregation behaviour of LKα14 in different solvent compositions was examined. The orientation of LKα14 in the different aggregates was examined as well as the size of the aggregates. For the larger clusters formed by both LKα14 and water the percolation behaviour was also investigated.

The simulations presented in this work show that the different solvation interactions between the two distinct solvent molecules and the two distinct amino acids of which LKα14 consists strongly influence all aspects of the behaviour of the peptide as well as the microscopic structure of solvent. This effect is particularly strong given that when the peptide forms an α-helix the two distinct amino acids are spatially separated making the α-helix amphiphilic (see chapter 4).

The increased stability of the α-helix in the bulk of ethanol/water mixtures or pure ethanol compared to the bulk of pure water (shown both here and in experiments[120]) is in agreement with experiments and simulations on other peptides and proteins which have shown that the stability of α-helices is increased upon addition of alcohols[14,121,140]. But as detailed in chapter 8 LKα14 forms even more stable α-helices at the surface of pure water than it does either in the bulk or at the surface in any water/ethanol mixture or pure ethanol. It is therefore reasonable to conclude that the stability of α-helix in water/ethanol mixtures results from two separate effects, the general stabilisation of α-helices by ethanol and the spatial separation of the water and ethanol molecules in the solvation shells of the α-helix discussed in chapter 10.

LKα14 has been shown to adsorb to polar/apolar interfaces both in experiments[58,120,123] and simulations[84]. It has also been shown that when adsorbed to an interface the axis of α-helical LKα14 peptides is parallel to the interface[47,123,124] and the leucine and lysine residues are oriented towards the apolar and the polar side of the interface respectively[47,83,84]. The spatial separation of the solvation shell of the α-helix discussed in chapter 10 means that in a water/ethanol mixture the environment around the peptide is similar to an interface. LKα14 adsorbs less strongly to the solvent/air interface in a water/ethanol mixture than in pure water even at low ethanol concentration. It also does not adsorb to the surface in pure ethanol but as discussed in chapter 9 the effect is not straightforward and the adsorption pattern points to a water/ethanol mixture being the best solvent for LKα14.

Examining the solvation shells of the α-helix strongly supports this conclusion as there is a preferential solvation of the peptide by either water or ethanol depending on the composition of the solvent and a clear spatial separation of the water and ethanol molecules in the solvation shells. In addition the orientation of the peptides in relation to the interface between the water-rich domain and the ethanol-rich domain in the clusters discussed in chapter 11.2 is the same as the orientation LKα14 is known to assume at a polar/apolar interface (with the ethanol-rich domain as the apolar side). The preferential solvation of each side of the peptide by either water or ethanol thus induces a separation of the solvent into these water- and ethanol-rich domains even though water and ethanol are fully miscible in all proportions when the peptides are not present.

In contrast to the adsorption behaviour and the secondary structure of peptides, the clusters and their structure have not been confirmed in experiments. Such experiments to confirm or disprove the formation of these clusters would be needed. If the findings in this work are confirmed, the presence of the observed aggregates might lead to several promising applications. The observed structures share typical properties of microemulsions: two liquids separated by a layer of surfactants (here the peptides)[23,25,27,89,90].

Microemulsions consist of two immiscible solvents and the surfactants[23,25,27,89,90]. The microemulision-like systems where ethanol acts as the surfactant also consist of two immiscible solvents in addition to the ethanol[53–56]. Micelles and microemulsions have been used to extract molecules from one solvent to another and to separate molecules with different solubilities[37,38]. Given that water and ethanol have different solvation properties (though they are more similar than between water and a typical organic solvent)[94] a microemulsion-like system where the two solvents are water and ethanol instead of water and oil might be interesting in the context of separating molecules whose solvation properties are not very dissimilar. Such a system might also be interesting as a solvent for proteins and peptides since ethanol and water are known to influence the conformation (and therefore the solubility in either solvent) in different ways[13,14,121].

As mentioned in the introduction investigating a solute in a binary water/ethanol mixture can be a good starting point for an investigation of the same solute in a more complicated three component mixture such as water/ethanol/octanol. Given that according to simulations LKα14 induces a separation of water and ethanol in the binary mixture it would be interesting to see what effect adding LKα14 to a ternary water/ethanol/octanol mixture would have. In the right composition a ternary water/ethanol/octanol mixture forms a microemulsion-like system with the ethanol acting as the surfactant[53,57]. LKα14 might adsorb to the surface of the microemulsion droplets. It might also form separate clusters in the water-rich domain (which does have a significant amount of ethanol[57]). The presence of octanol might influence the secondary structure or the adsorption behaviour of the peptide and the presence of the peptide might influence the concentrations of the different solvent needed for the formation of the different microemulsion-like structures.

Part V.

Appendix

A. Trigonometry

At the end of 5.2 we have the rotations matrix

$$\boldsymbol{R} = \begin{pmatrix} \cos(\phi) & \cos(\theta)\sin(\phi) & \sin(\theta)\sin(\phi) \\ -\sin(\phi) & \cos(\theta)\cos(\theta) & \sin(\theta)\cos(\phi) \\ 0 & -\sin(\theta) & \cos(\theta) \end{pmatrix} \tag{A.1}$$

and apply it to the unit vector along the z-axis $\boldsymbol{u}_z = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$. This gives a vector $\boldsymbol{\alpha}$ whose coordinates are the direction cosines of the fitted line from 5.2.

$$\boldsymbol{\alpha} = \begin{pmatrix} \cos(\phi) & \cos(\theta)\sin(\phi) & \sin(\theta)\sin(\phi) \\ -\sin(\phi) & \cos(\theta)\cos(\theta) & \sin(\theta)\cos(\phi) \\ 0 & -\sin(\theta) & \cos(\theta) \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} \sin(\theta)\sin(\phi) \\ \sin(\theta)\cos(\phi) \\ \cos(\theta) \end{pmatrix} \tag{A.2}$$

From 5.2 we also have that

$$\tan(\phi) = \frac{\sum x_i l_i}{\sum y_i l_i} \tag{A.3}$$

and

$$\tan(\theta) = \frac{\sin(\phi)\sum x_i l_i + \cos(\phi)\sum y_i l_i}{\sum z_i l_i}. \tag{A.4}$$

With equations A.3 and A.4 it is possible to calculate $\cos(\phi)$, $\sin(\phi)$, $\cos(\theta)$ and $\sin(\theta)$. It will be assumed that $\sum x_i l_i$, $\sum y_i l_i$ and $\sum z_i l_i$ are positive since the angle in 5 is only defined between 0° and 90° and it is therefore possible to take the absolute values.

A.1. Calculating $\cos(\phi)$ and $\sin(\phi)$

Equation A.3 gives

$$\phi = \tan^{-1}\left(\frac{\sum x_i l_i}{\sum y_i l_i}\right). \tag{A.5}$$

Inserting this into $\cos(\phi)$ yields

$$\begin{aligned}\cos(\phi) &= \cos\left(\tan^{-1}\left(\frac{\sum x_i l_i}{\sum y_i l_i}\right)\right) \\ &= \frac{1}{\sqrt{\left(\frac{\sum x_i l_i}{\sum y_i l_i}\right)^2 + 1}} \\ &= \frac{1}{\sqrt{\frac{(\sum x_i l_i)^2 + (\sum y_i l_i)^2}{(\sum y_i l_i)^2}}} \\ &= \frac{\sum y_i l_i}{\sqrt{(\sum x_i l_i)^2 + (\sum y_i l_i)^2}}. \end{aligned} \tag{A.6}$$

Substituting ϕ from equation A.5 into $\sin(\phi)$ yields

$$\begin{aligned}\sin(\phi) &= \sin\left(\tan^{-1}\left(\frac{\sum x_i l_i}{\sum y_i l_i}\right)\right) \\ &= \frac{\frac{\sum x_i l_i}{\sum y_i l_i}}{\sqrt{\left(\frac{\sum x_i l_i}{\sum y_i l_i}\right)^2 + 1}}\end{aligned}$$

which simplifies to

$$\sin(\phi) = \frac{\sum x_i l_i}{\sqrt{(\sum x_i l_i)^2 + (\sum y_i l_i)^2}} \tag{A.7}$$

in an analogue way to equation A.6.

A.2. Calculating $\cos(\theta)$ and $\sin(\theta)$

Equation A.4 gives

$$\theta = \tan^{-1}\left(\frac{\sin(\phi)\sum x_i l_i + \cos(\phi)\sum y_i l_i}{\sum z_i l_i}\right). \tag{A.8}$$

Inserting the results from equations A.6 and A.7 into equation A.8 gives

$$\begin{aligned}\theta &= \tan^{-1}\left(\frac{\frac{\sum x_i l_i}{\sqrt{(\sum x_i l_i)^2+(\sum y_i l_i)^2}}\sum x_i l_i + \frac{\sum y_i l_i}{\sqrt{(\sum x_i l_i)^2+(\sum y_i l_i)^2}}\sum y_i l_i}{\sum z_i l_i}\right)\\ &= \tan^{-1}\left(\frac{\left(\frac{(\sum x_i l_i)^2+(\sum y_i l_i)^2}{\sqrt{(\sum x_i l_i)^2+(\sum y_i l_i)^2}}\right)}{\sum z_i l_i}\right)\\ &= \tan^{-1}\left(\frac{\sqrt{(\sum x_i l_i)^2+(\sum y_i l_i)^2}}{\sum z_i l_i}\right). \end{aligned} \tag{A.9}$$

Substituting this into $\cos(\theta)$ yields

$$\begin{aligned}\cos(\theta) &= \cos\left(\tan^{-1}\left(\frac{\sqrt{(\sum x_i l_i)^2+(\sum y_i l_i)^2}}{\sum z_i l_i}\right)\right)\\ &= \frac{1}{\sqrt{\left(\frac{\sqrt{(\sum x_i l_i)^2+(\sum y_i l_i)^2}}{\sum z_i l_i}\right)^2+1}}\\ &= \frac{1}{\sqrt{\frac{(\sum x_i l_i)^2+(\sum y_i l_i)^2}{(\sum z_i l_i)^2}+1}}\\ &= \frac{1}{\sqrt{\frac{(\sum x_i l_i)^2+(\sum y_i l_i)^2+(\sum z_i l_i)^2}{(\sum z_i l_i)^2}}}\\ &= \frac{\sum z_i l_i}{\sqrt{(\sum x_i l_i)^2+(\sum y_i l_i)^2+(\sum z_i l_i)^2}}. \end{aligned} \tag{A.10}$$

Substituting θ from equation A.9 into $\sin(\theta)$ yields

$$\begin{aligned}\sin(\theta) &= \sin\left(\tan^{-1}\left(\frac{\sqrt{(\sum x_i l_i)^2+(\sum y_i l_i)^2}}{\sum z_i l_i}\right)\right)\\ &= \frac{\frac{\sqrt{(\sum x_i l_i)^2+(\sum y_i l_i)^2}}{\sum z_i l_i}}{\sqrt{\left(\frac{\sqrt{(\sum x_i l_i)^2+(\sum y_i l_i)^2}}{\sum z_i l_i}\right)^2+1}}\end{aligned}$$

which simplifies to

$$\begin{aligned}\sin(\theta) &= \frac{(\sum z_i l_i)\,\frac{\sqrt{(\sum x_i l_i)^2+(\sum y_i l_i)^2}}{(\sum z_i l_i)}}{\sqrt{(\sum x_i l_i)^2+(\sum y_i l_i)^2+(\sum z_i l_i)^2}} \\ &= \frac{\sqrt{(\sum x_i l_i)^2+(\sum y_i l_i)^2}}{\sqrt{(\sum x_i l_i)^2+(\sum y_i l_i)^2+(\sum z_i l_i)^2}}.\end{aligned} \tag{A.11}$$

in an analogue way to equation A.10.

A.3. Calculating α

By inserting the results from equations A.7 and A.11 into the x-coordinate of the vector in equation A.2

$$\begin{aligned}\alpha_x &= \frac{\sqrt{(\sum x_i l_i)^2+(\sum y_i l_i)^2}}{\sqrt{(\sum x_i l_i)^2+(\sum y_i l_i)^2+(\sum z_i l_i)^2}} \cdot \frac{\sum x_i l_i}{\sqrt{(\sum x_i l_i)^2+(\sum y_i l_i)^2}} \\ &= \frac{\sum x_i l_i}{\sqrt{(\sum x_i l_i)^2+(\sum y_i l_i)^2+(\sum z_i l_i)^2}}\end{aligned}$$

is obtained. Similarly

$$\begin{aligned}\alpha_y &= \frac{\sqrt{(\sum x_i l_i)^2+(\sum y_i l_i)^2}}{\sqrt{(\sum x_i l_i)^2+(\sum y_i l_i)^2+(\sum z_i l_i)^2}} \cdot \frac{\sum y_i l_i}{\sqrt{(\sum x_i l_i)^2+(\sum y_i l_i)^2}} \\ &= \frac{\sum y_i l_i}{\sqrt{(\sum x_i l_i)^2+(\sum y_i l_i)^2+(\sum z_i l_i)^2}}\end{aligned}$$

is obtained by inserting the results from equations A.6 and A.11 into the y-coordinate of the vector in equation A.2. Finally by inserting the result from equation A.10 into the z-coordinate of α

$$\alpha_z = \frac{\sum z_i l_i}{\sqrt{(\sum x_i l_i)^2+(\sum y_i l_i)^2+(\sum z_i l_i)^2}} \tag{A.12}$$

is obtained.

B. Secondary structure

For the secondary structure analysis in chapter 8 the secondary structure of the peptide or peptides was calculated every 5 ps. This appendix contains plots of the secondary structure as function of time for all simulations for which the secondary structure was analysed.

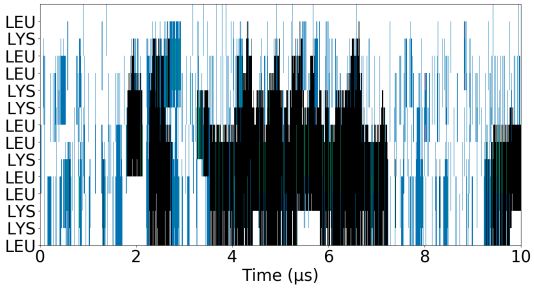
B.1. Bulk

As explained in section 8.1 two simulations were performed for each composition, each with one single LKα14 peptide in bulk. For the first one an α-helical structure was simulated for 0.6 ns at 900 K. The resulting simulation box was then equilibrated at 300 K using the Berendsen barostat. The equilibrated simulation box was then simulated for 10 μs using the Parrinello-Rahman barostat. The starting conformation of the peptide in the other simulation was either random coil or an α-helix depending on whether the peptide showed more tendency towards α-helical or random coil secondary structure. The simulations with $x_{\mathrm{Eth}} = 0.033$ are an exception, there no heating was involved and the two starting conformations were simply an α-helical one and a random coil one.

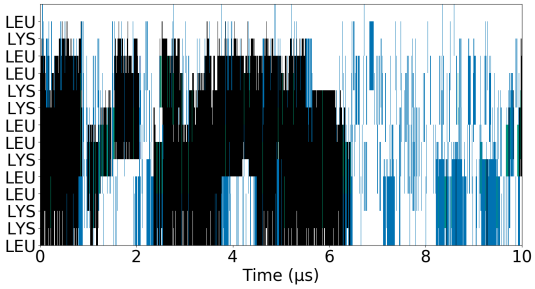
The first intention was to only do one simulation for each composition. Given that an α-helix is a relatively stable structure the system was heated so that other conformations could be reached. As can be seen from figures B.3b, B.3d and B.3f heating the system was not always enough to denature the α-helix. As can be seen from figures B.1, B.2, B.3 and B.4 there was no clear pattern to be seen after 1 μs other than that an α-helix which was present at the start of the simulation would stay (or in the case with $x_{\mathrm{Eth}} = 0.9$ oscillate between α-helix and turn mostly because it was barely long enough to be an α-helix).

Therefore simultaneously a second simulation was started for each composition and the first simulations were continued. The route chosen was to start each of the new simulations from an random coil conformation. This conformation was taken from the simulation with $x_{\mathrm{Eth}} = 0$ and therefore not suitable for a second simulation of that composition. As such the second simulation of the composition with $x_{\mathrm{Eth}} = 0$ was started from an α-helical conformation of the peptide. As can be seen from figure B.2 in the simulation of the compositions with $x_{\mathrm{Eth}} = 0.2$ and $x_{\mathrm{Eth}} = 0.3$ an random coil structure survived longer in the simulations starting from the heated structure than in the simulations starting from the random coil structure. Since the α-helices did not disappear after they had formed a third simulation starting from an α-helical structure was not considered necessary for these systems.

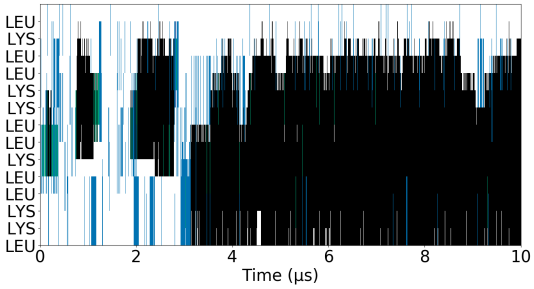
Comparing figures B.1, B.2, B.3 and B.4 shows that the only simulations where an



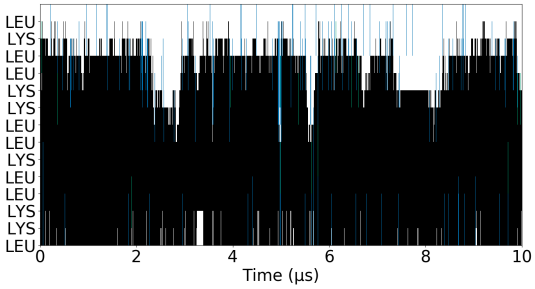
(a) Mole fraction ethanol $x_{\mathrm{Eth}} = 0$ starting from the heated conformation.



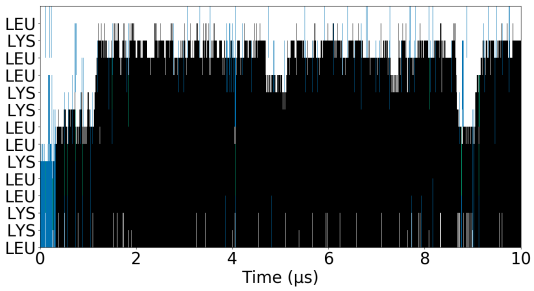
(b) Mole fraction ethanol $x_{\mathrm{Eth}} = 0$ starting from an α-helical configuration.



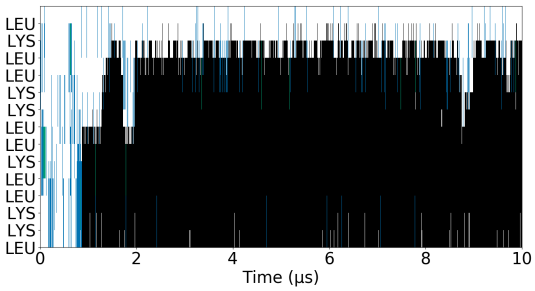
(c) Mole fraction ethanol $x_{\mathrm{Eth}} = 0.033\bar{3}$ starting from a random coil configuration.



(d) Mole fraction ethanol $x_{\mathrm{Eth}} = 0.033$ starting from an α-helical configuration.

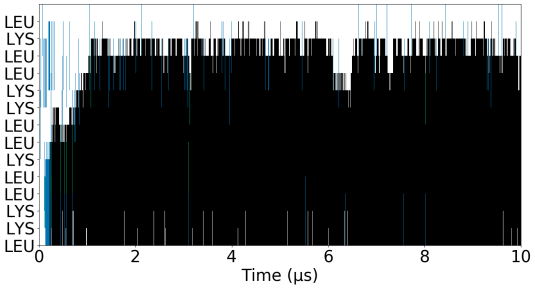


(e) Mole fraction ethanol $x_{\mathrm{Eth}} = 0.1$ starting from a random coil configuration.

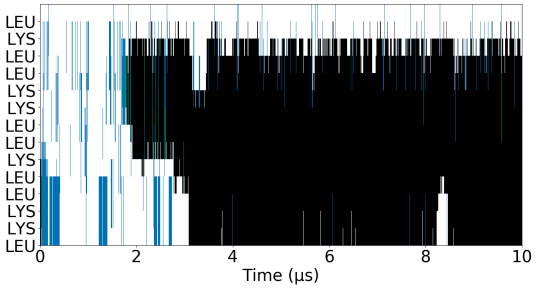


(f) Mole fraction ethanol $x_{\mathrm{Eth}} = 0.1$ starting from the heated configuration.

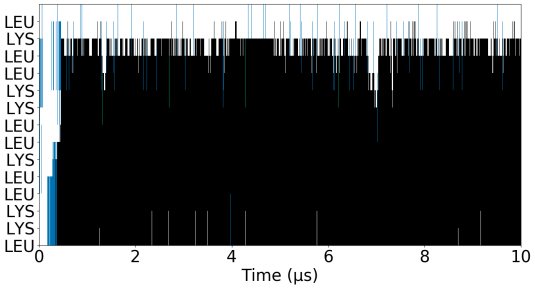
Figure B.1.: Secondary structure of LKα14 as a function of simulation time for each amino acid. Black is α-helix, white random coil, blue turn, green 3_{10}-helix and pink isolated bridge.



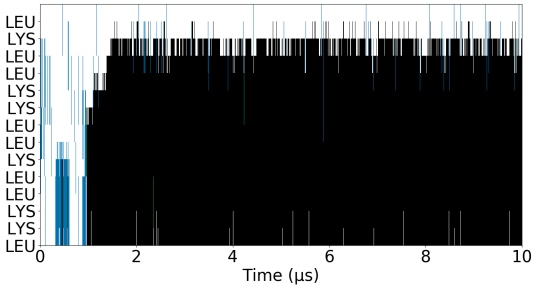
(a) Mole fraction ethanol $x_{\mathrm{Eth}} = 0.2$ starting from a random coil configuration.



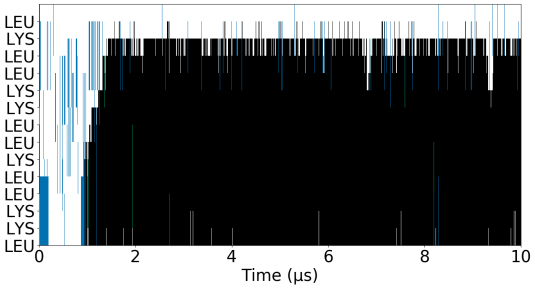
(b) Mole fraction ethanol $x_{\mathrm{Eth}} = 0.2$ starting from the heated configuration.



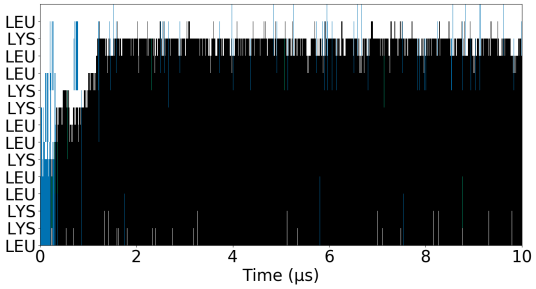
(c) Mole fraction ethanol $x_{\mathrm{Eth}} = 0.3$ starting from a random coil configuration.



(d) Mole fraction ethanol $x_{\mathrm{Eth}} = 0.3$ starting from the heated configuration.

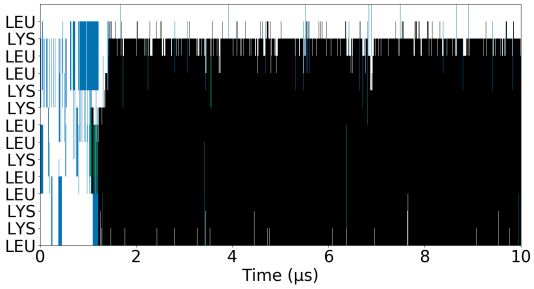


(e) Mole fraction ethanol $x_{\mathrm{Eth}} = 0.4$ starting from a random coil configuration.

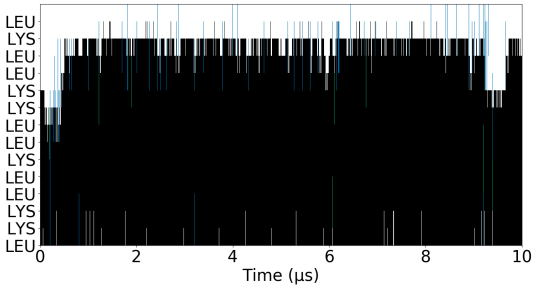


(f) Mole fraction ethanol $x_{\mathrm{Eth}} = 0.4$ starting from the heated configuration.

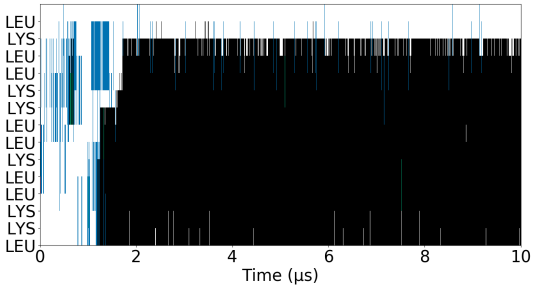
Figure B.2.: Secondary structure of LKα14 as a function of simulation time for each amino acid. Black is α-helix, white random coil, blue turn and green 3_{10}-helix.



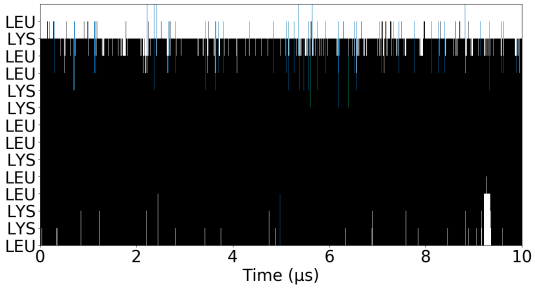
(a) Mole fraction ethanol $x_{\mathrm{Eth}} = 0.5$ starting from a random coil configuration.



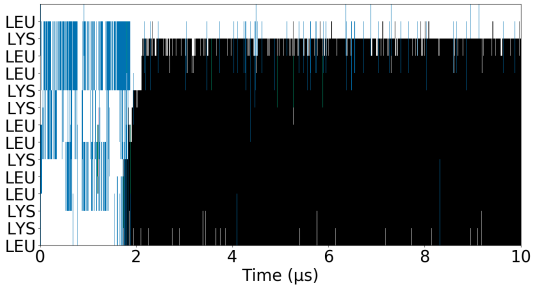
(b) Mole fraction ethanol $x_{\mathrm{Eth}} = 0.5$ starting from the heated configuration.



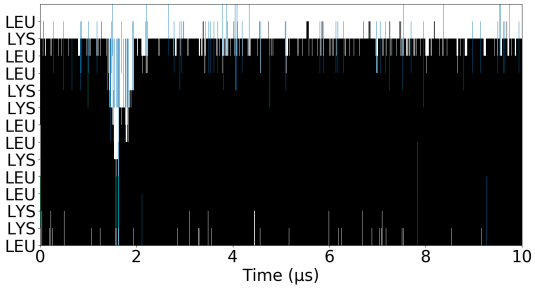
(c) Mole fraction ethanol $x_{\mathrm{Eth}} = 0.6$ starting from a random coil configuration.



(d) Mole fraction ethanol $x_{\mathrm{Eth}} = 0.6$ starting from the heated configuration.

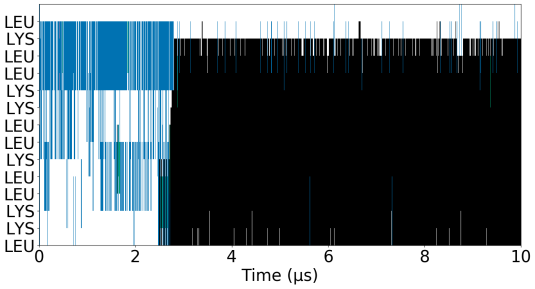


(e) Mole fraction ethanol $x_{\mathrm{Eth}} = 0.7$ starting from a random coil configuration.

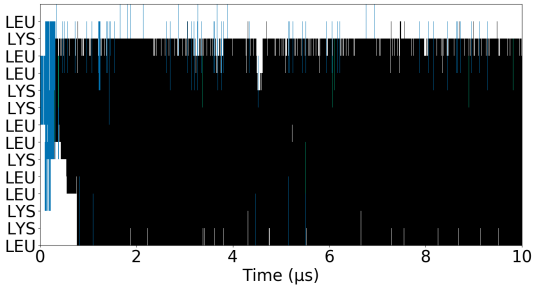


(f) Mole fraction ethanol $x_{\mathrm{Eth}} = 0.7$ starting from the heated configuration.

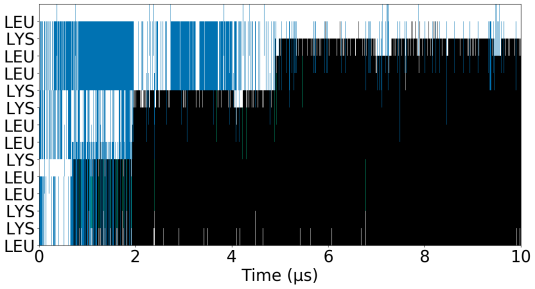
Figure B.3.: Secondary structure of LKα14 as a function of simulation time for each amino acid. Black is α-helix, white random coil, blue turn and green 3_{10}-helix.



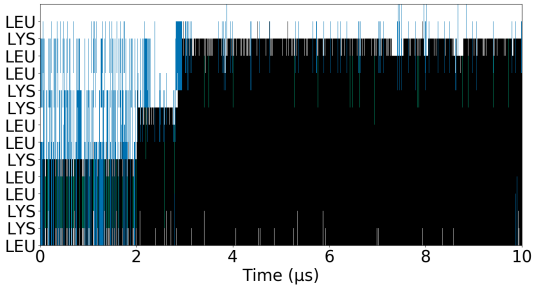
(a) Mole fraction ethanol $x_{\mathrm{Eth}} = 0.8$ starting from a random coil configuration.



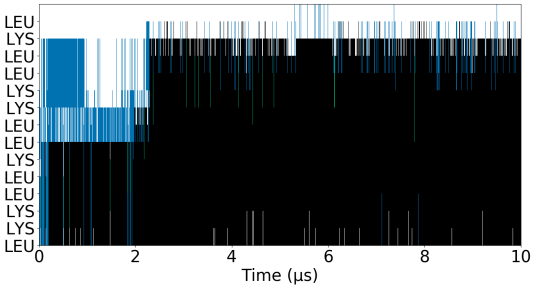
(b) Mole fraction ethanol $x_{\mathrm{Eth}} = 0.8$ starting from the heated configuration.



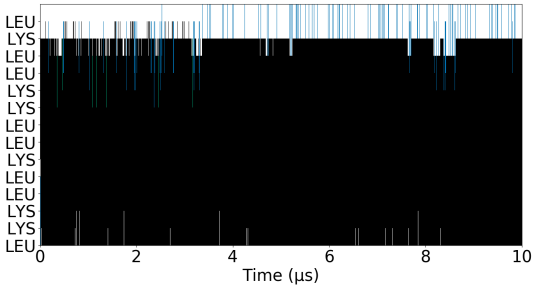
(c) Mole fraction ethanol $x_{\mathrm{Eth}} = 0.9$ starting from a random coil configuration.



(d) Mole fraction ethanol $x_{\mathrm{Eth}} = 0.9$ starting from the heated configuration.



(e) Mole fraction ethanol $x_{\mathrm{Eth}} = 1$ starting from a random coil configuration.



(f) Mole fraction ethanol $x_{\mathrm{Eth}} = 1$ starting from the heated configuration.

Figure B.4.: Secondary structure of LKα14 as a function of simulation time for each amino acid. Black is α-helix, white random coil, blue turn and green 3_{10}-helix.

α-helix completely disappears are both simulations with $x_{\mathrm{Eth}} = 0$ and the simulation with $x_{\mathrm{Eth}} = 0.033$ starting from an random coil simulations (where it then soon forms again and stays for the rest of the simulation). For other compositions there is always at least a turn left. This happens for three compositions with $x_{\mathrm{Eth}} > 0.1$. There is the case of the α-helix becoming shorter and then the last part turning into turn for a short while before the whole α-helix forms again when $x_{\mathrm{Eth}} = 0.7$ and shorter α-helices at the start of simulations turning into turns when $x_{\mathrm{Eth}} = 0.9$ or $x_{\mathrm{Eth}} = 1$. There is also more fluctuation in the length of the α-helix for $x_{\mathrm{Eth}} = 0.033$ and $x_{\mathrm{Eth}} = 0.1$ than when $x_{\mathrm{Eth}} > 0.1$. This shows that the α-helix is quite stable in all conformations other than $x_{\mathrm{Eth}} = 0$.

There is no clear pattern in how long it takes the α-helix to form but for $0.3 \leq x_{\mathrm{Eth}} \leq 0.8$ an α-helix that does form tends to be quick to assume its full length while this takes longer time for $x_{\mathrm{Eth}} \geq 0.9$ or $x_{\mathrm{Eth}} \leq 0.2$. Therefore the presence of both water and ethanol seems to promote the peptide folding into a full length α-helix when a part of it has already formed an α-helix.

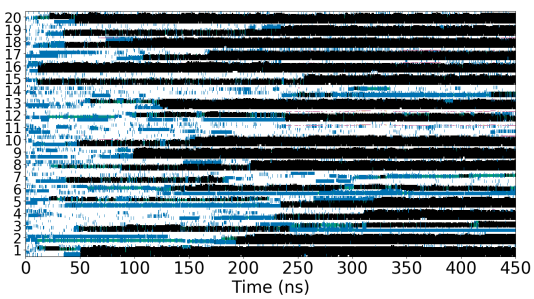
Figures B.1, B.2, B.3 and B.4 also show that the full length of the α-helix is 12 of the 14 amino acids of the peptide. The two amino acids at the N-terminus of the peptide are not part of the α-helix. As mentioned earlier there is more fluctuation in the length of the α-helix when $x_{\mathrm{Eth}} = 0.033$ or $x_{\mathrm{Eth}} = 0.1$ than when $x_{\mathrm{Eth}} > 0.1$. This fluctuation involves the α-helix shortening from the N-terminus end far more often than from the C-terminus end of the helix. Even at higher x_{Eth} where fluctuations in helix length are less common fluctuations are more common at the N-terminus end of the α-helix than at the C-terminus end.

B.2. Surface

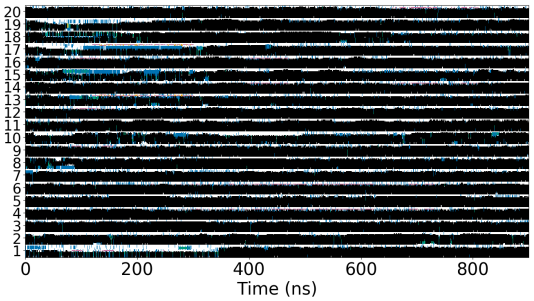
The difference between the secondary structure at the surface and in the bulk of the water/ethanol mixtures was examined by performing two simulations for each composition. Each had 20 peptides and two solution/air interfaces. For each composition, in one simulation the peptides started from an α-helical conformation and in the other simulation they started from the same random coil conformation as was used for the random coil simulations in B.1. The simulations starting from an α-helical conformation were 900 ns while the simulations starting from a random coil conformation were 450 ns. As explained in section 8.2 this was not enough to converge the secondary structures.

Figures B.5, B.6, B.7 and B.8 clearly show how not all amino acids in the peptides are part of the α-helix. The white strip signifying the random coil secondary structure of the amino acids closest to the N-terminus allow the different peptides to be distinguished in the graphs showing the results from the simulations when starting from an α-helical conformation. In the graphs of the results from the simulations starting from a random coil conformation it is only possible to distinguish the different peptides when they do form α-helices or by comparison with the corresponding graph from a simulation starting from α-helical conformation.

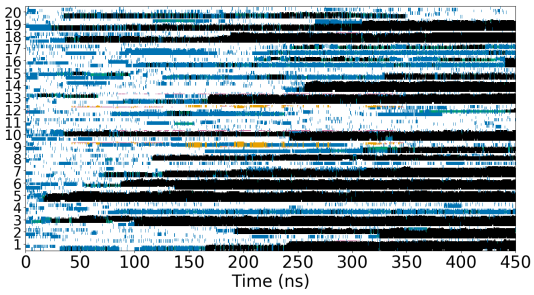
Comparing figures B.5, B.6, B.7 and B.8 shows that even if the simulations starting



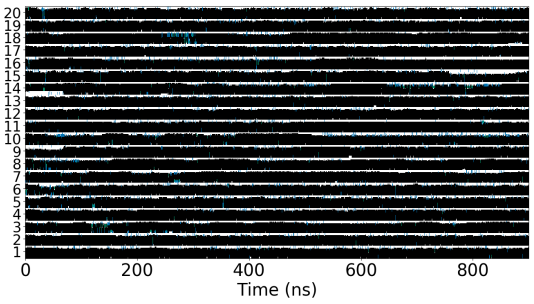
(a) Mole fraction ethanol $x_{\mathrm{Eth}} = 0$ starting from a random coil conformation.



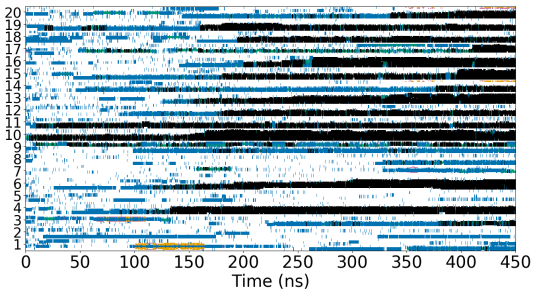
(b) Mole fraction ethanol $x_{\mathrm{Eth}} = 0$ starting from an α-helical configuration.



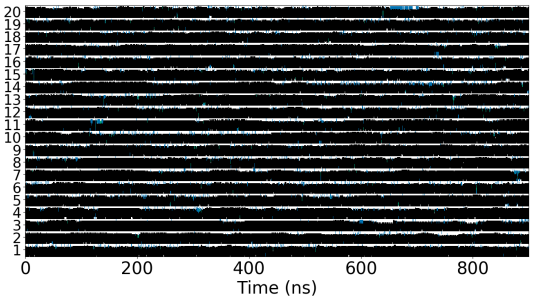
(c) Mole fraction ethanol $x_{\mathrm{Eth}} = 0.033$ starting from a random coil conformation.



(d) Mole fraction ethanol $x_{\mathrm{Eth}} = 0.033$ starting from an α-helical configuration.

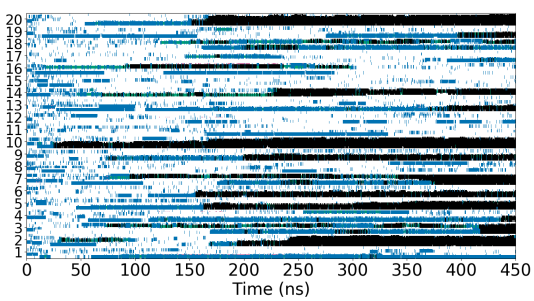


(e) Mole fraction ethanol $x_{\mathrm{Eth}} = 0.1$ starting from a random coil conformation.

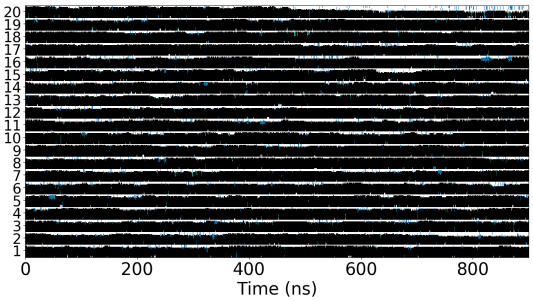


(f) Mole fraction ethanol $x_{\mathrm{Eth}} = 0.1$ starting from an α-helical configuration.

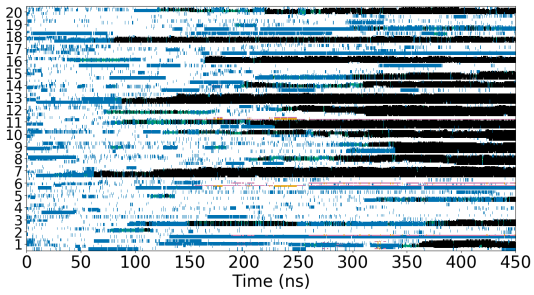
Figure B.5.: Secondary structure of LKα14 as a function of simulation time for each amino acid of all peptides. The order of the amino acids in each peptide is the same as in the single peptides in bulk earlier and the numbers on the side refer to the individual peptides. Black is α-helix, white random coil, blue turn, green 3_{10}-helix, pink isolated bridge and orange extended conformation.



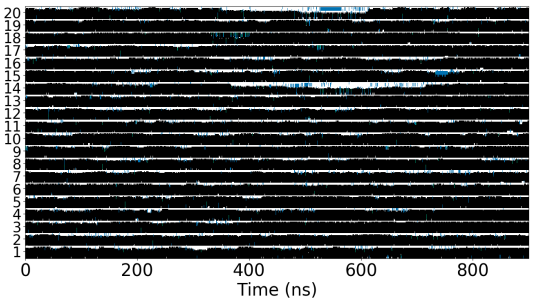
(a) Mole fraction ethanol $x_{\mathrm{Eth}} = 0.2$ starting from a random coil configuration.



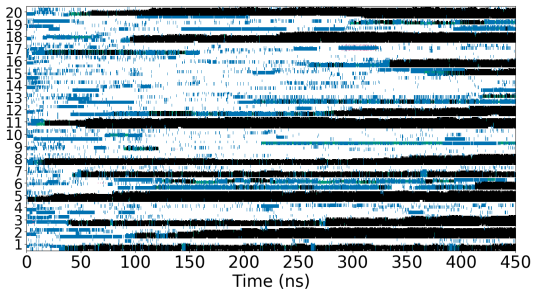
(b) Mole fraction ethanol $x_{\mathrm{Eth}} = 0.2$ starting from an α-helical configuration.



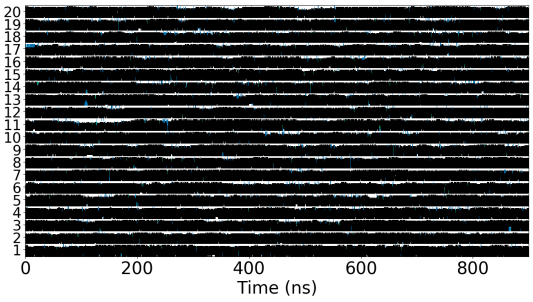
(c) Mole fraction ethanol $x_{\mathrm{Eth}} = 0.3$ starting from a random coil configuration.



(d) Mole fraction ethanol $x_{\mathrm{Eth}} = 0.3$ starting from an α-helical configuration.

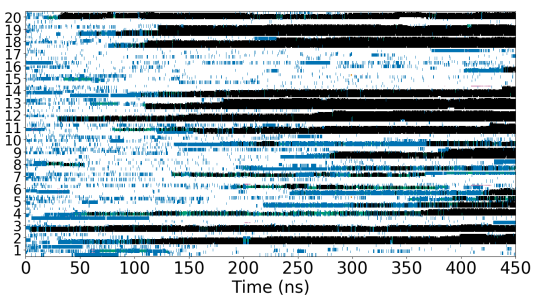


(e) Mole fraction ethanol $x_{\mathrm{Eth}} = 0.4$ starting from a random coil configuration.

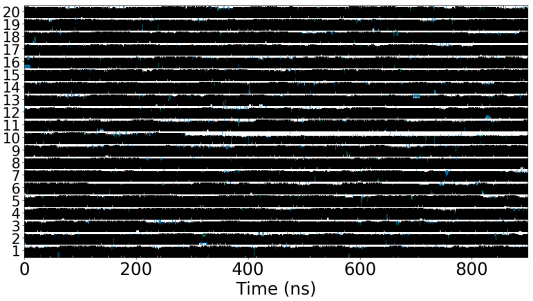


(f) Mole fraction ethanol $x_{\mathrm{Eth}} = 0.4$ starting from an α-helical configuration.

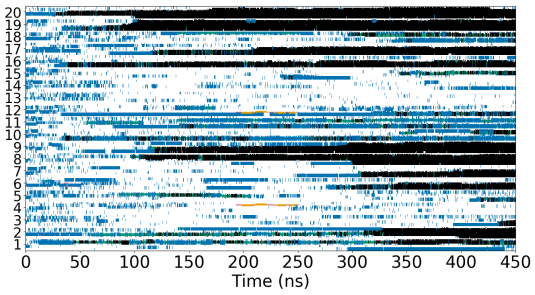
Figure B.6.: Secondary structure of LKα14 as a function of simulation time for each amino acid. The order of the amino acids in each peptide is the same as in the single peptides in bulk earlier and the numbers on the side refer to the individual peptides. Black is α-helix, white random coil, blue turn, green 3_{10}-helix, pink isolated bridge and orange extended conformation.



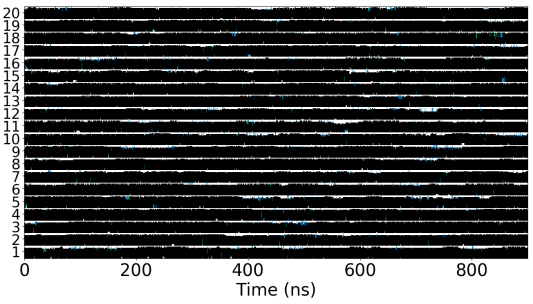
(a) Mole fraction ethanol $x_{\mathrm{Eth}} = 0.5$ starting from a random coil configuration.



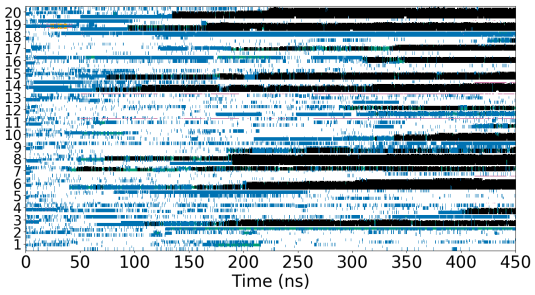
(b) Mole fraction ethanol $x_{\mathrm{Eth}} = 0.5$ starting from an α-helical configuration.



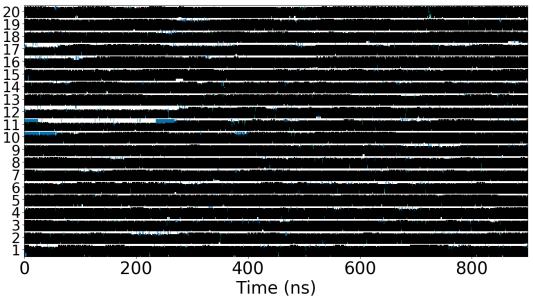
(c) Mole fraction ethanol $x_{\mathrm{Eth}} = 0.6$ starting from a random coil configuration.



(d) Mole fraction ethanol $x_{\mathrm{Eth}} = 0.6$ starting from an α-helical configuration.

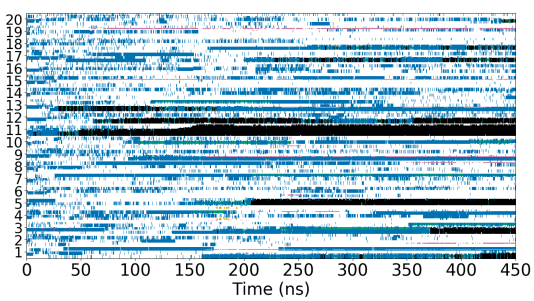


(e) Mole fraction ethanol $x_{\mathrm{Eth}} = 0.7$ starting from a random coil configuration.

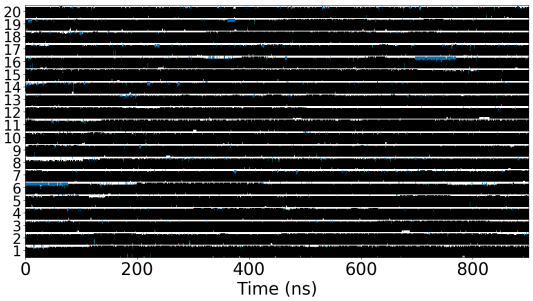


(f) Mole fraction ethanol $x_{\mathrm{Eth}} = 0.7$ starting from an α-helical configuration.

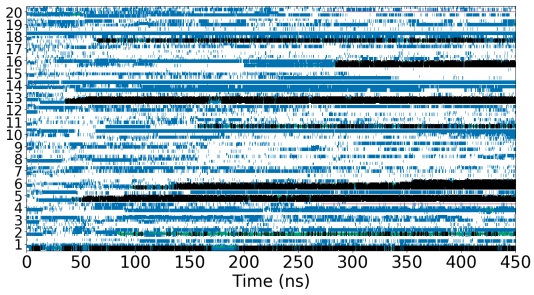
Figure B.7.: Secondary structure of LKα14 as a function of simulation time for each amino acid.The order of the amino acids in each peptide is the same as in the single peptides in bulk earlier and the numbers on the side refer to the individual peptides. Black is α-helix, white random coil, blue turn, green 3_{10}-helix, pink isolated bridge and orange extended conformation.



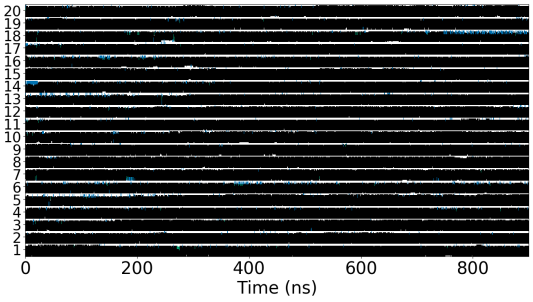
(a) Mole fraction ethanol $x_{\mathrm{Eth}} = 0.8$ starting from a random coil configuration.



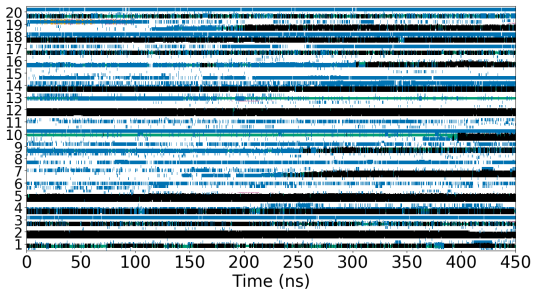
(b) Mole fraction ethanol $x_{\mathrm{Eth}} = 0.8$ starting from an α-helical configuration.



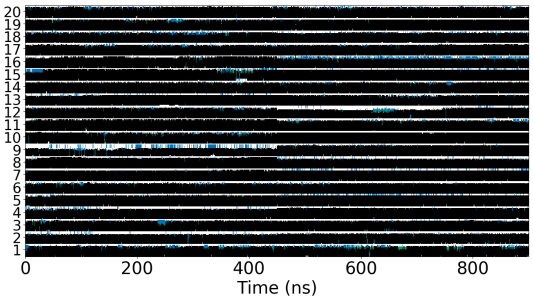
(c) Mole fraction ethanol $x_{\mathrm{Eth}} = 0.9$ starting from a random coil configuration.



(d) Mole fraction ethanol $x_{\mathrm{Eth}} = 0.9$ starting from an α-helical configuration.



(e) Mole fraction ethanol $x_{\mathrm{Eth}} = 1$ starting from a random coil configuration.



(f) Mole fraction ethanol $x_{\mathrm{Eth}} = 1$ starting from an α-helical configuration.

Figure B.8.: Secondary structure of LKα14 as a function of simulation time for each amino acid.The order of the amino acids in each peptide is the same as in the single peptides in bulk earlier and the numbers on the side refer to the individual peptides. Black is α-helix, white random coil, blue turn, green 3_{10}-helix, pink isolated bridge and orange extended conformation.

from an α-helical conformation were two times as long as the ones starting from a random coil conformation more changes can be seen in the secondary structure over time in the simulations starting from a random coil conformation. There is no visible difference between the secondary structure graphs for the different simulations starting from α-helical conformations, though the analysis in section 8.2 reveals a slight difference between pure water and the other compositions (i.e. a slightly lower prevalence of α-helix in the bulk). But even though random coil is the preferred conformation of the peptide in the bulk of pure water, figure B.5b shows that this slight difference must stem from the α-helices being on average shorter in the bulk rather than from any α-helix fully denaturing within the time-frame of the simulation.

The simulations starting from random coil configurations have more interesting results. Figure B.5a shows that in pure water 16 of the 20 peptides form α-helices within the time-frame of the simulation (14 of these are full length α-helices) and 8 peptides start forming α-helices within 100 ns of the start of the simulation. As shown in section 8.2 the majority of these α-helices is at the surface. Since LKα14 was designed to form α-helices at water/air interfaces[58] and has been shown to do so[58,83] it is not surprising that the peptide form α-helices at the surface in the simulations. But comparison of figure B.5a with figures B.1, B.2, B.3 and B.4 shows that the α-helices form more quickly at the surface of pure water than in the bulk of solvent compositions where the α-helix is quite stable.

Comparison of figures B.5, B.6, B.7 and B.8 also shows a tendency towards more α-helix formation upon decreasing ethanol content of the solvent. Especially the solvent compositions where section 9.2 shows that α-helices adsorb to the surface lead to many α-helices formed within the time-frame of the simulation, 15 for $x_{\mathrm{Eth}} = 0.033\overline{3}$ and 14 for $x_{\mathrm{Eth}} = 0.1$ (11 and 6 full length ones respectively). The α-helices in those two solvent compositions form slower than in pure water. In contrast only 5 of the 20 peptides fold into an α-helix within the time-frame of the simulations when $x_{\mathrm{Eth}} = 0.8$ or $x_{\mathrm{Eth}} = 0.9$ (only 1 full length α-helix forms in each). The trend of more α-helices forming with less x_{Eth} is not linear, with more α-helices formed when $x_{\mathrm{Eth}} = 0.4$ (13, 8 full length ones) than when $x_{\mathrm{Eth}} = 0.2$ (13, but only 5 full length ones).

The different times needed for formation of α-helices in different solvent compositions in the simulations both with and without a surface seems to suggest that presence of both water and ethanol induces quicker formation of α-helices. The water/air surface induces an even quicker α-helix formation.

Figures B.5, B.6, B.7 and B.8 all show, that when starting from a random coil conformation, α-helices that have formed tend to not disappear again. This fact and the stark contrast between the secondary structures depending on the starting conformation of the peptides support the conclusion from section 8.2 that 450 or 900 ns long simulations are not long enough to converge the secondary structure even when 20 peptides are used instead of 1.

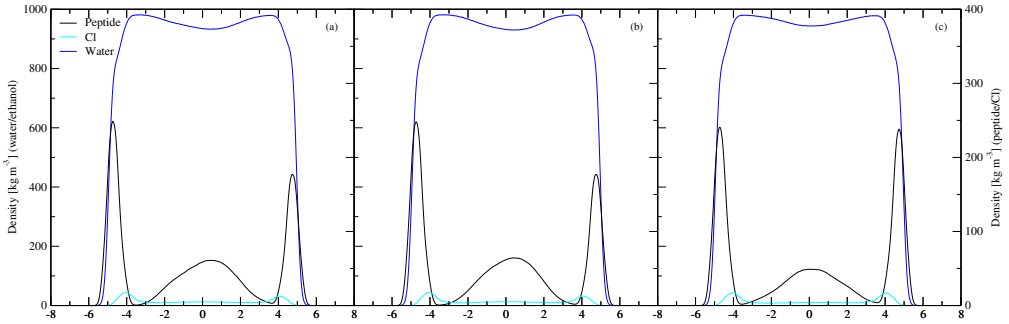
C. Partial densities

The partial densities along the z-coordinate of all the different components in all simulations with an interface are shown in figures C.1, C.2, C.3 and C.4. For the simulations where the peptides start from an α-helical conformation the partial densities were calculated both from the first 450 ns of the simulation and after 900 ns. In all cases the first 100 ns were discarded as equilibration.

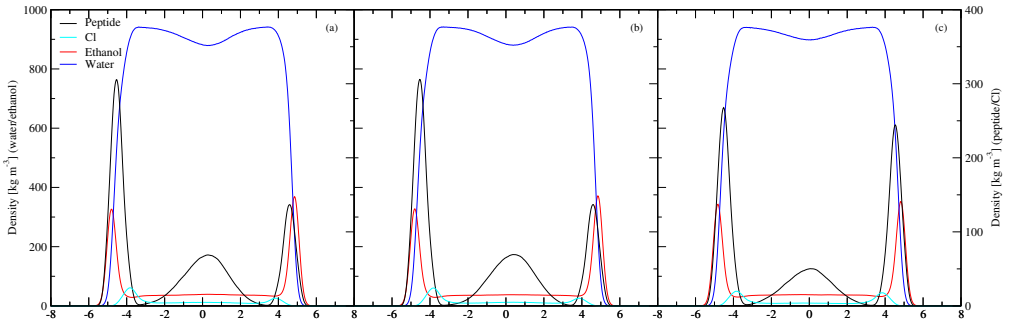
Figure C.1 shows clearly that for low x_{Eth} the peptides adsorb to the surface no matter the starting conformation and that there is no significant difference between the results after 450 ns and 900 ns. When ethanol is present it also adsorbs to the surface. This has been shown to also be the case in experiments of ethanol/water mixtures without any solutes[141,142]. Comparing figures C.1a, C.1b and C.1c reveals no clear difference in the adsorption between these three x_{Eth}though as seen in chapter 9.2 a closer analysis reveals a slight difference in the adsorption.

Figures C.2 and C.3 reveal an interesting dynamic where a longer simulation time results not in more symmetrical curves for the partial densities but rather in more loop sided ones. Many of them have a two peak structure close to one surface (for example $x_{\mathrm{Eth}} = 0.4$, $x_{\mathrm{Eth}} = 0.5$ or $x_{\mathrm{Eth}} = 0.7$) with a peak in the water density between the two peaks in the peptide density. In these solvent compositions the peptides form clusters with the water molecules (see chapter 11). These clusters as a whole adsorb to the surface. There they orient is such a way that about half the peptides in the cluster are close to the surface, then the water molecules form a layer and then the second peptide peak represents the peptide on the other side of the cluster. The position of the clusters in the solution converges slower than the position of monomers or oligomers.

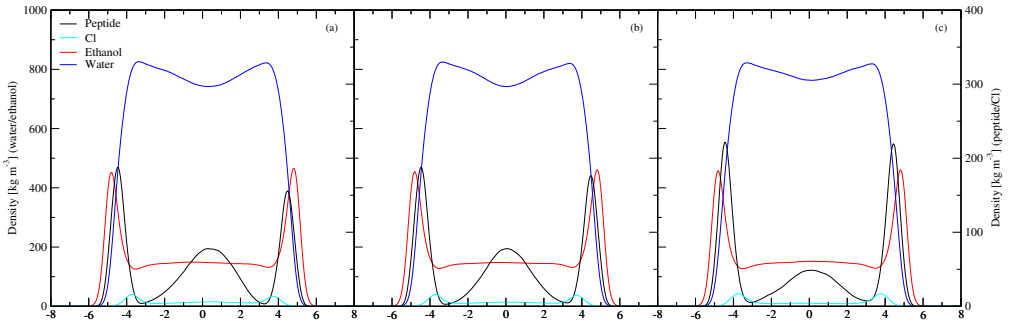
As seen in figure C.4 there is no adsorption of the peptides at high x_{Eth} and for $x_{\mathrm{Eth}} = 0.9$ and $x_{\mathrm{Eth}} = 1$ a longer simulation leads to a more symmetrical curves for the partial densities as would be expected from a better converged simulation.



(a) $x_{\mathrm{Eth}} = 0$

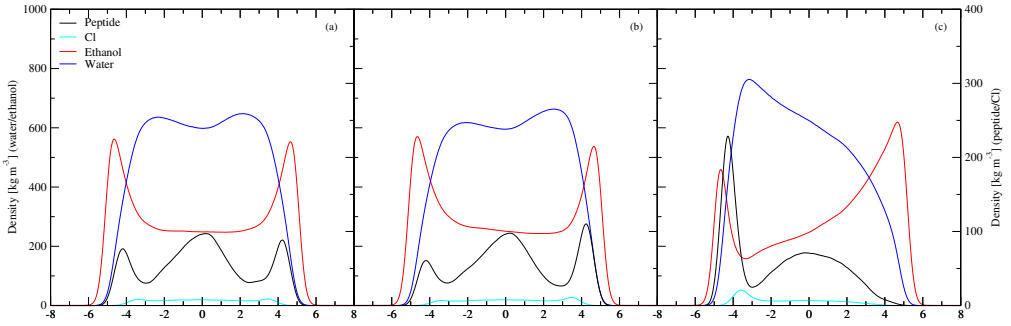


(b) $x_{\mathrm{Eth}} = 0.033$

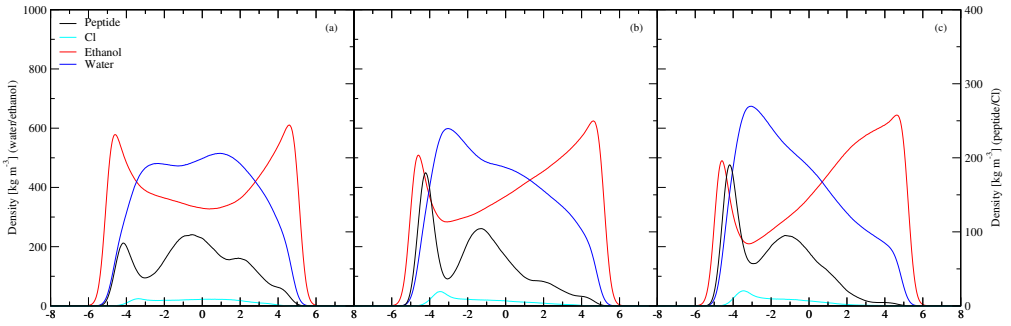


(c) $x_{\mathrm{Eth}} = 0.1$

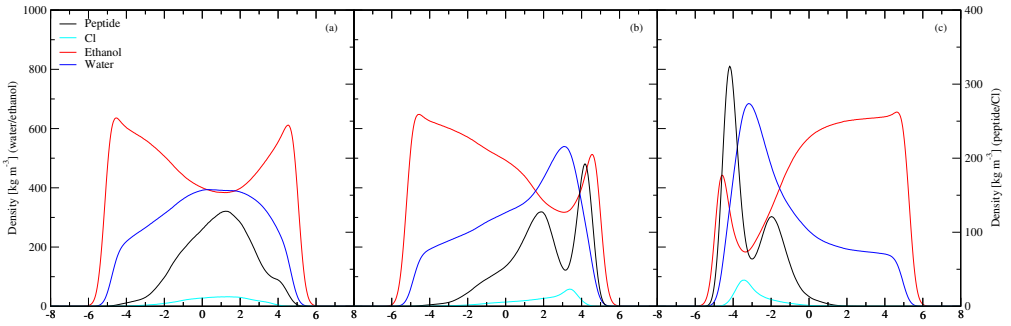
Figure C.1.: Partial densities of each component along the z-coordinate with $x_{\mathrm{Eth}} = 0$, $x_{\mathrm{Eth}} = 0.033$ and $x_{\mathrm{Eth}} = 0.1$. In each subfigure (a) is a 450 ns simulation and (b) a 900 ns simulation, both starting from an α-helical peptide conformation, while (c) is a 450 ns simulation starting from a random coil peptide configuration. Water uses the scale on the left, the peptides and the ions the scale on the right.



(a) $x_{\mathrm{Eth}} = 0.2$

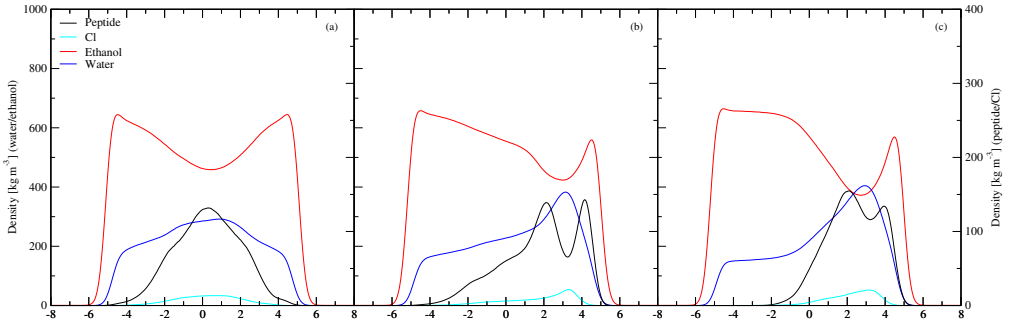


(b) $x_{\mathrm{Eth}} = 0.3$

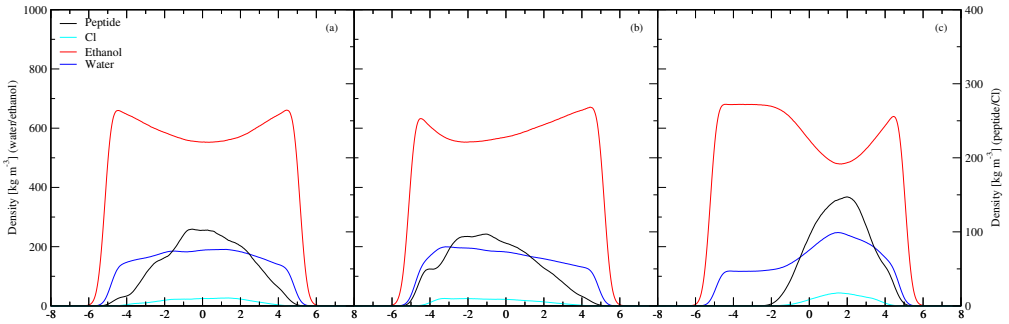


(c) $x_{\mathrm{Eth}} = 0.4$

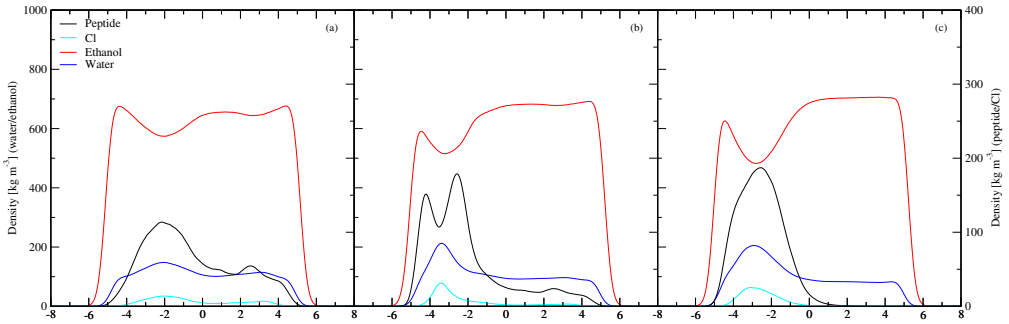
Figure C.2.: Partial densities of each component along the z-coordinate with $x_{\mathrm{Eth}} = 0.2$, $x_{\mathrm{Eth}} = 0.3$ and $x_{\mathrm{Eth}} = 0.4$. In each subfigure (a) is a 450 ns simulation and (b) a 900 ns simulation, both starting from an α-helical peptide conformation, while (c) is a 450 ns simulation starting from a random coil peptide configuration. Water uses the scale on the left, the peptides and the ions the scale on the right.



(a) $x_{\mathrm{Eth}} = 0.5$

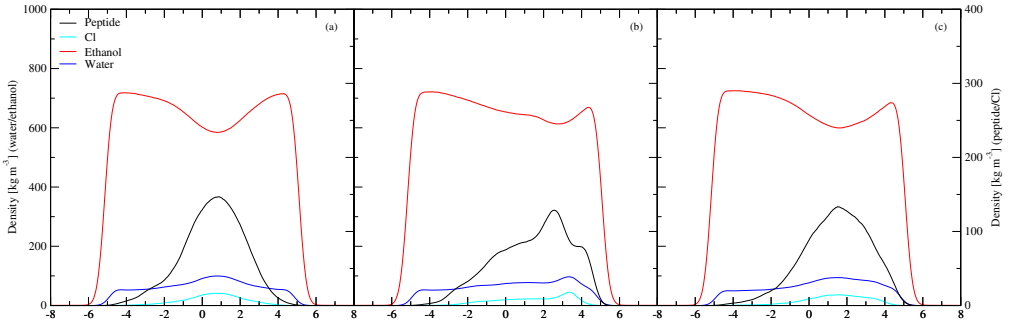


(b) $x_{\mathrm{Eth}} = 0.6$

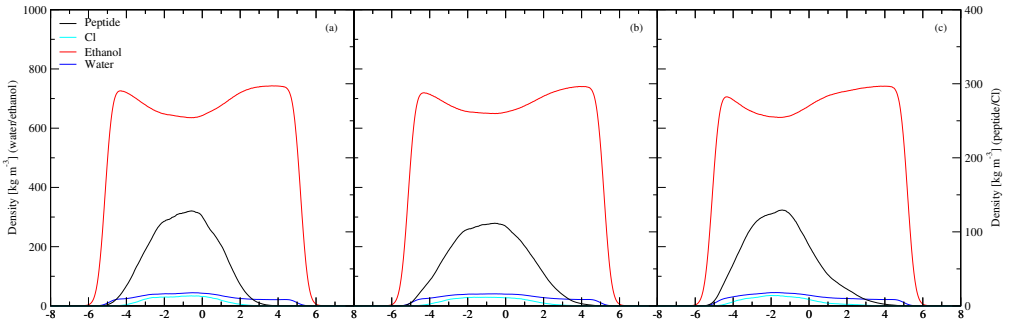


(c) $x_{\mathrm{Eth}} = 0.7$

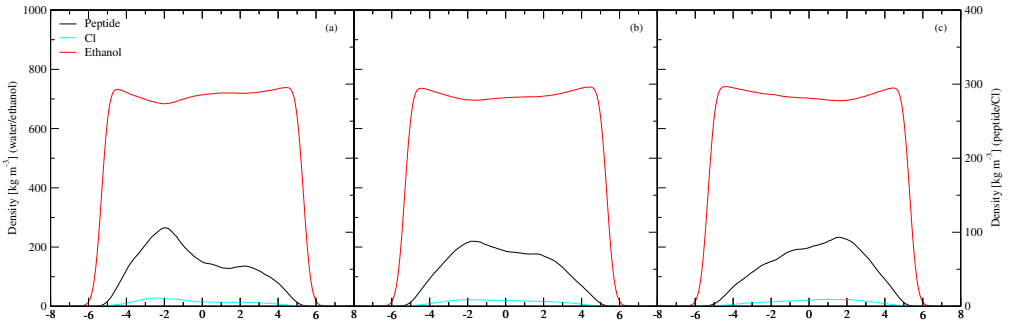
Figure C.3.: Partial densities of each component along the z-coordinate with $x_{\mathrm{Eth}} = 0.5$,$x_{\mathrm{Eth}} = 0.6$ and $x_{\mathrm{Eth}} = 0.7$. In each subfigure (a) is a 450 ns simulation and (b) a 900 ns simulation, both starting from an α-helical peptide conformation, while (c) is a 450 ns simulation starting from a random coil peptide configuration. Water uses the scale on the left, the peptides and the ions the scale on the right.



(a) $x_{\mathrm{Eth}} = 0.8$



(b) $x_{\mathrm{Eth}} = 0.9$



(c) $x_{\mathrm{Eth}} = 1$

Figure C.4.: Partial densities of each component along the z-coordinate with $x_{\mathrm{Eth}} = 0.8$,$x_{\mathrm{Eth}} = 0.9$ and $x_{\mathrm{Eth}} = 1$. In each subfigure (a) is a 450 ns simulation and (b) a 900 ns simulation, both starting from an α-helical peptide conformation, while (c) is a 450 ns simulation starting from a random coil peptide configuration. Water uses the scale on the left, the peptides and the ions the scale on the right.

Bibliography

[1] Adriano Aguzzi and Anna M. Calella. Prions: protein aggregation and infectious diseases. *Physiological reviews*, 89(4):1105–1152, 2009.

[2] Stanley B. Prusiner. Prions. *Proceedings of the National Academy of Sciences*, 95 (23):13363–13383, 1998.

[3] Adriano Aguzzi and Tracy O'Connor. Protein aggregation diseases: pathogenicity and therapeutic perspectives. *Nature reviews Drug discovery*, 9(3):237–248, 2010.

[4] Robin Roychaudhuri, Mingfeng Yang, Minako M. Hoshi, and David B. Teplow. Amyloid β-protein assembly and alzheimer disease. *Journal of Biological Chemistry*, 284(8):4749–4753, 2009.

[5] Sam Gandy et al. The role of cerebral amyloid β accumulation in common forms of alzheimer disease. *The Journal of clinical investigation*, 115(5):1121–1129, 2005.

[6] Frank M. LaFerla, Kim N. Green, and Salvatore Oddo. Intracellular amyloid-β in alzheimer's disease. *Nature Reviews Neuroscience*, 8(7):499–509, 2007.

[7] Christopher A. Ross and Michelle A. Poirier. Protein aggregation and neurodegenerative disease. *Nature medicine*, 10(7):S10–S17, 2004.

[8] Massimo Stefani and Christopher M. Dobson. Protein aggregation and aggregate toxicity: new insights into protein folding, misfolding diseases and biological evolution. *Journal of molecular medicine*, 81(11):678–699, 2003.

[9] Stanley B. Prusiner. Neurodegenerative diseases and prions. *New England Journal of Medicine*, 344(20):1516–1526, 2001.

[10] Akira Kakizuka. Protein precipitation: a common etiology in neurodegenerative disorders? *Trends in Genetics*, 14(10):396–402, 1998.

[11] Serge N. Timasheff. Protein-solvent interactions and protein conformation. *Accounts of Chemical Research*, 3(2):62–68, 1970.

[12] George Nemethy, William J. Peer, and Harold A. Scheraga. Effect of protein-solvent interactions on protein conformation. *Annual review of biophysics and bioengineering*, 10(1):459–497, 1981.

[13] Tsutomu Arakawa, John F. Carpenter, Yoshiko A. Kita, and John H. Crowe. The basis for toxicity of certain cryoprotectants: a hypothesis. *Cryobiology*, 27(4):401–415, 1990.

[14] Masahiro Kinoshita, Yuko Okamoto, and Fumio Hirata. Peptide conformations in alcohol and water: analyses by the reference interaction site model theory. *Journal of the American Chemical Society*, 122(12):2773–2779, 2000.

[15] Werner Kunz, Pierandrea Lo Nostro, and Barry W. Ninham. The present state of affairs with hofmeister effects. *Current opinion in colloid & interface science*, 9 (1-2):1–18, 2004.

[16] Shahar Sukenik, Liel Sapir, Regina Gilman-Politi, and Daniel Harries. Diversity in the mechanisms of cosolute action on biomolecular processes. *Faraday discussions*, 160:225–237, 2013.

[17] Shahar Sukenik, Liel Sapir, and Daniel Harries. Osmolyte induced changes in peptide conformational ensemble correlate with slower amyloid aggregation: a coarse-grained simulation study. *Journal of Chemical Theory and Computation*, 11(12): 5918–5928, 2015.

[18] Liel Sapir and Daniel Harries. Wisdom of the crowd. *Bunsen-Magazin*, 19:152–162, 2017.

[19] George N. Somero. Protons, osmolytes, and fitness of internal milieu for protein function. *American Journal of Physiology-Regulatory, Integrative and Comparative Physiology*, 251(2):R197–R213, 1986.

[20] Gayani N. Pallewela and Paul E. Smith. Preferential solvation in binary and ternary mixtures. *The Journal of Physical Chemistry B*, 119(51):15706–15717, 2015.

[21] Maria G. Ortore, Paolo Mariani, Flavio Carsughi, Stefania Cinelli, Giuseppe Onori, José Teixeira, and Francesco Spinozzi. Preferential solvation of lysozyme in water/ethanol mixtures. *The Journal of chemical physics*, 135(24):12B620, 2011.

[22] Yizhak Marcus. Preferential solvation in mixed solvents. In Paul E. Smith, Enrico Matteoli, and John P. O'Connell, editors, *Flutcuation Theory of Solutions*, pages 65–92. CRC Press, 2013.

[23] Robert J. Hunter. *Foundations of colloid science*. Oxford University Press, Walton Street, Oxford OX26DP, UK, 1989. reprint with corrections in 1993.

[24] Håkan Wennerström and Björn Lindman. Micelles. physical chemistry of surfactant association. *Physics Reports*, 52(1):1–86, 1979.

[25] Yves Chevalier and Thomas Zemb. The structure of micelles and microemulsions. *Reports on Progress in Physics*, 53(3):279, 1990.

[26] Jacob N. Israelachvili, D. John Mitchell, and Barry W. Ninham. Theory of self-assembly of hydrocarbon amphiphiles into micelles and bilayers. *Journal of the Chemical Society, Faraday Transactions 2: Molecular and Chemical Physics*, 72: 1525–1568, 1976.

[27] Dominique Langevin. Micelles and microemulsions. *Annual Review of Physical Chemistry*, 43(1):341–369, 1992.

[28] Laurence S. Romsted, Clifford A. Bunton, and Jihu Yao. Micellar catalysis, a useful misnomer. *Current opinion in colloid & interface science*, 2(6):622–628, 1997.

[29] Giorgio La Sorella, Giorgio Strukul, and Alessandro Scarso. Recent advances in catalysis in micellar media. *Green Chemistry*, 17(2):644–683, 2015.

[30] Vassiliki Papadimitriou, Theodore G. Sotiroudis, and Aristotelis Xenakis. Olive oil microemulsions: enzymatic activities and structural characteristics. *Langmuir*, 23 (4):2071–2077, 2007.

[31] Haralambos Stamatis, Aristotelis Xenakis, and Fragiskos N. Kolisis. Bioorganic reactions in microemulsions: the case of lipases. *Biotechnology advances*, 17(4-5): 293–318, 1999.

[32] Glen S. Kwon and Teruo Okano. Polymeric micelles as new drug carriers. *Advanced drug delivery reviews*, 21(2):107–116, 1996.

[33] Duy Toan Pham, Athittaya Chokamonsirikun, Vipasiri Phattaravorakarn, and Waree Tiyaboonchai. Polymeric micelles for pulmonary drug delivery: a comprehensive review. *Journal of Materials Science*, 56(3):2016–2036, 2021.

[34] Sushant S. Kulthe, Yogesh M. Choudhari, Nazma N. Inamdar, and Vishnukant Mourya. Polymeric micelles: authoritative aspects for drug delivery. *Designed Monomers and Polymers*, 15(5):465–521, 2012.

[35] Vladimir P. Torchilin. Targeted polymeric micelles for delivery of poorly soluble drugs. *Cellular and Molecular Life Sciences CMLS*, 61(19):2549–2559, 2004.

[36] David Chelazzi, Rodorico Giorgi, and Piero Baglioni. Microemulsions, micelles, and functional gels: how colloids and soft matter preserve works of art. *Angewandte Chemie International Edition*, 57(25):7296–7303, 2018.

[37] Constantine D. Stalikas. Micelle-mediated extraction as a tool for separation and preconcentration in metal analysis. *TrAC Trends in Analytical Chemistry*, 21(5): 343–355, 2002.

[38] Hirofumi Tani, Tamio Kamidate, and Hiroto Watanabe. Micelle-mediated extraction. *journal of Chromatography A*, 780(1-2):229–241, 1997.

[39] Javier A. Gomez del Rio and Douglas G. Hayes. Protein extraction by winsor-iii microemulsion systems. *Biotechnology progress*, 27(4):1091–1100, 2011.

[40] Priscila G. Mazzola, Andre M. Lopes, Francislene A. Hasmann, Angela F. Jozala, Thereza C.V. Penna, Perola O. Magalhaes, Carlota O. Rangel-Yagui, and Adalberto Pessoa Jr. Liquid–liquid extraction of biomolecules: an overview and update

of the main techniques. *Journal of Chemical Technology & Biotechnology: International Research in Process, Environmental & Clean Technology*, 83(2):143–157, 2008.

[41] Sharrel Rebello, Aju K. Asok, Sathish Mundayoor, and M.S. Jisha. Surfactants: toxicity, remediation and green surfactants. *Environmental chemistry letters*, 12 (2):275–287, 2014.

[42] Sourav De, Susanta Malik, Aniruddha Ghosh, Rumpa Saha, and Bidyut Saha. A review on natural surfactants. *RSC advances*, 5(81):65757–65767, 2015.

[43] Norah Alwadani and Pedram Fatehi. Synthetic and lignin-based surfactants: Challenges and opportunities. *Carbon Resources Conversion*, 1(2):126–138, 2018.

[44] Pierre Bauduin, Audrey Renoncourt, Andreas Kopf, Didier Touraud, and Werner Kunz. Unified concept of solubilization in water by hydrotropes and cosolvents. *Langmuir*, 21(15):6769–6775, 2005.

[45] Marios Hopkins Hatzopoulos, Julian Eastoe, Peter J. Dowding, Sarah E. Rogers, Richard Heenan, and Robert Dyer. Are hydrotropes distinct from surfactants? *Langmuir*, 27(20):12346–12353, 2011.

[46] Chirravuri V. Subbarao, I.P. Kalyan Chakravarthy, A.V.S.L. Sai Bharadwaj, and Kommuri M.M. Prasad. Functions of hydrotropes in solutions. *Chemical engineering & technology*, 35(2):225–237, 2012.

[47] Sandra Roy, Tsuki L. Naka, and Dennis K. Hore. Enhanced understanding of amphipathic peptide adsorbed structure by modeling of the nonlinear vibrational response. *The Journal of Physical Chemistry C*, 117(47):24955–24966, 2013.

[48] Werner Kunz, Krister Holmberg, and Thomas Zemb. Hydrotropes. *Current Opinion in Colloid & Interface Science*, 22:99–107, 2016.

[49] Jonathan J. Booth, Muhiadin Omar, Steven Abbott, and Seishi Shimizu. Hydrotrope accumulation around the drug: the driving force for solubilization and minimum hydrotrope concentration for nicotinamide and urea. *Physical Chemistry Chemical Physics*, 17(12):8028–8037, 2015.

[50] Jonathan J. Booth, Steven Abbott, and Seishi Shimizu. Mechanism of hydrophobic drug solubilization by small molecule hydrotropes. *The Journal of Physical Chemistry B*, 116(51):14915–14921, 2012.

[51] Rabindranath Paul, Krishna Gopal Chattaraj, and Sandip Paul. Role of hydrotropes in sparingly soluble drug solubilization: Insight from a molecular dynamics simulation and experimental perspectives. *Langmuir*, 37(16):4745–4762, 2021.

[52] Seishi Shimizu and Nobuyuki Matubayasi. The origin of cooperative solubilisation by hydrotropes. *Physical Chemistry Chemical Physics*, 18(36):25621–25628, 2016.

[53] Sebastian Schöttl and Dominik Horinek. Salt effects in surfactant-free microemulsions. *The Journal of chemical physics*, 148(22):222818, 2018.

[54] Sylvain Prévost, Sebastian Krickl, Stjepan Marčelja, Werner Kunz, Thomas Zemb, and Isabelle Grillo. Spontaneous ouzo emulsions coexist with pre-ouzo ultraflexible microemulsions. *Langmuir*, 37(13):3817–3827, 2021.

[55] Michael L. Klossek, Didier Touraud, Thomas Zemb, and Werner Kunz. Structure and solubility in surfactant-free microemulsions. *ChemPhysChem*, 13(18):4116–4119, 2012. doi: https://doi.org/10.1002/cphc.201200667.

[56] Tobias Lopian, Sebastian Schöttl, Sylvain Prevost, Stéphane Pellet-Rostaing, Dominik Horinek, Werner Kunz, and Thomas Zemb. Morphologies observed in ultraflexible microemulsions with and without the presence of a strong acid. *ACS central science*, 2(7):467–475, 2016.

[57] Sebastian Schöttl, Tobias Lopian, Sylvain Prévost, Didier Touraud, Isabelle Grillo, Olivier Diat, Thomas Zemb, and Dominik Horinek. Combined molecular dynamics (md) and small angle scattering (sas) analysis of organization on a nanometer-scale in ternary solvent solutions containing a hydrotrope. *Journal of colloid and interface science*, 540:623–633, 2019.

[58] William F. DeGrado and James D. Lear. Induction of peptide conformation at apolar water interfaces. 1. a study with model peptides of defined hydrophobic periodicity. *Journal of the American Chemical Society*, 107(25):7684–7689, 1985.

[59] Tobias Weidner, Nicholas F. Breen, Kun Li, Gary P. Drobny, and David G. Castner. Sum frequency generation and solid-state nmr study of the structure, orientation, and dynamics of polystyrene-adsorbed peptides. *Proceedings of the National Academy of Sciences*, 107(30):13288–13293, 2010.

[60] Tobias Weidner, Julia S. Apte, Lara J. Gamble, and David G. Castner. Probing the orientation and conformation of α-helix and β-strand model peptides on self-assembled monolayers using sum frequency generation and nexafs spectroscopy. *Langmuir*, 26(5):3433–3440, 2010.

[61] Dan L. Bergman, Alexander P. Lyubartsev, and Aatto Laaksonen. Topological and spatial aspects of the hydration of solutes of extreme solvation entropy. *Physical Review E*, 60(4):4482, 1999.

[62] Gerhard Hummer, Shekhar Garde, Angel E Garcia, Michael E. Paulaitis, and Lawrence R. Pratt. Hydrophobic effects on a molecular scale. *The Journal of Physical Chemistry B*, 102(51):10469–10482, 1998.

[63] David Chandler. Hydrophobicity: Two faces of water. *Nature*, 417(6888):491–491, 2002.

[64] James S. Peerless, G. Hunter Bowers, Albert L. Kwansa, and Yaroslava G. Yingling. Fullerenes in aromatic solvents: correlation between solvation-shell structure, solvate formation, and solubility. *The Journal of Physical Chemistry B*, 119(49): 15344–15352, 2015.

[65] John G. Kirkwood and Elizabeth Monroe Boggs. The radial distribution function in liquids. *The Journal of Chemical Physics*, 10(6):394–402, 1942.

[66] Erik Lindahl, Mark J. Abraham, Berk Hess, and David van der Spoel. Gromacs documentation, release 2020.5. https://doi.org/10.5281/zenodo.4420784, januar 2021. 5.10.4 Radial distribution functions.

[67] Peter G. Kusalik and Igor M. Svishchev. The spatial structure in liquid water. *Science*, 265(5176):1219–1221, 1994.

[68] Igor M. Svishchev and Peter G. Kusalik. Structure in liquid water: A study of spatial distribution functions. *The Journal of chemical physics*, 99(4):3049–3058, 1993.

[69] Erik Lindahl, Mark J. Abraham, Berk Hess, and David van der Spoel. Gromacs documentation, release 2020.5. https://doi.org/10.5281/zenodo.4420784, januar 2021. 3.6.85 gmx spatial.

[70] Linus Pauling, Robert B. Corey, and Herman R. Branson. The structure of proteins: two hydrogen-bonded helical configurations of the polypeptide chain. *Proceedings of the National Academy of Sciences*, 37(4):205–211, 1951.

[71] Jeremy M. Berg, John L. Tymoczko, Gregory J. Gatto jr., and Lubert Stryer. *Stryer Biochemie*, chapter 2. Zusammensetzung und Struktur der Proteine, pages 31–77. Springer-Verlag GmbH Deutschland, Heidelberger Platz 3, 14197 Berlin, Germany, 8 edition, 2018.

[72] Donald Voet, Judith G. Voet, and Charlotte W. Pratt. *Lehrbuch der Biochemie*, chapter 6. Dreidimensionale Struktur von Proteinen, pages 162–219. WILEY-VCH Verlag GmbH & Co. KGaA, Boschstraße 12, 69469 Weinheim, Germany, 3 edition, 2019.

[73] David J. Barlow and Janet M. Thornton. Helix geometry in proteins. *Journal of molecular biology*, 201(3):601–619, 1988.

[74] Dmitrij Frishman and Patrick Argos. Knowledge-based protein secondary structure assignment. *Proteins: Structure, Function, and Bioinformatics*, 23(4):566–579, 1995.

[75] Catherine K. Smith and Lynne Regan. Construction and design of β-sheets. *Accounts of chemical research*, 30(4):153–161, 1997.

[76] Frederic M. Richards and Craig E. Kundrot. Identification of structural motifs from protein coordinate data: secondary structure and first-level supersecondary structure. *Proteins: Structure, Function, and Bioinformatics*, 3(2):71–84, 1988.

[77] David Eisenberg. The discovery of the α-helix and β-sheet, the principal structural features of proteins. *Proceedings of the National Academy of Sciences*, 100(20): 11207–11210, 2003.

[78] Jack D. Dunitz. Pauling's left-handed α-helix. *Angewandte Chemie International Edition*, 40(22):4167–4173, 2001.

[79] John D. Fisk and Samuel H. Gellman. A parallel β-sheet model system that folds in water. *Journal of the American Chemical Society*, 123(2):343–344, 2001.

[80] Wolfgang Kabsch and Christian Sander. Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers: Original Research on Biomolecules*, 22(12):2577–2637, 1983.

[81] Matthias Heinig and Dmitrij Frishman. Stride: a web server for secondary structure assignment from known atomic coordinates of proteins. *Nucleic acids research*, 32 (suppl_2):W500–W502, 2004.

[82] Dmitrij Frishman and Patrick Argos. Stride: Protein secondary structure assignment from atomic coordinates. *Proteins*, 23:455–479, 1995.

[83] Cahit Dalgicdir and Mehmet Sayar. Conformation and aggregation of lkα14 peptide in bulk water and at the air/water interface. *The Journal of Physical Chemistry B*, 119(49):15164–15175, 2015.

[84] Cahit Dalgicdir, Christoph Globisch, Christine Peter, and Mehmet Sayar. Tipping the scale from disorder to alpha-helix: folding of amphiphilic peptides in the presence of macroscopic and molecular interfaces. *PLoS Comput Biol*, 11(8):e1004328, 2015.

[85] Michael A. Donovan, Helmut Lutz, Yeneneh Y. Yimer, Jim Pfaendtner, Mischa Bonn, and Tobias Weidner. Lk peptide side chain dynamics at interfaces are independent of secondary structure. *Physical Chemistry Chemical Physics*, 19(42): 28507–28511, 2017.

[86] Peter C. Kahn. Defining the axis of a helix. *Computers & Chemistry*, 13(3): 185–189, 1989.

[87] Peter C. Kahn. Simple methods for computing the least squares line in three dimensions. *Computers & chemistry*, 13(3):191–195, 1989.

[88] Gustavo Kume, Manlio Gallotti, and George Nunes. Review on anionic/cationic surfactant mixtures. *Journal of Surfactants and Detergents*, 11(1):1–11, 2008.

[89] D. Fennell Evans and Håkan Wennerström. *The colloidal domain*. Wiley-VCH, 111 River Street, Hoboken, NJ 07030, USA, 2 edition, 1999.

[90] Jacob N. Israelachvili. *Intermolecular and surface forces*. Elsevier Inc., 225 Wyman Street, Waltham, MA 02451, USA, 3 edition, 2011.

[91] Ramanathan Nagarajan. Molecular packing parameter and surfactant self-assembly: the neglected role of the surfactant tail. *Langmuir*, 18(1):31–38, 2002.

[92] Stanislaw Slomkowski, José V. Alemán, Robert G. Gilbert, Michael Hess, Kazuyuki Horie, Richard G. Jones, Przemyslaw Kubisa, Ingrid Meisel, Werner Mormann, Stanisław Penczek, et al. Terminology of polymers and polymerization processes in dispersed systems (iupac recommendations 2011). *Pure and Applied Chemistry*, 83(12):2229–2259, 2011.

[93] Max Teubner and Reinhard Strey. Origin of the scattering peak in microemulsions. *The Journal of Chemical Physics*, 87(5):3195–3200, 1987.

[94] CRC Handbook. *CRC Handbook of Chemistry and Physics, 49th Edition*. The Chemical Rubber Co., 18901 Cranwood Parkway, Cleveland, Ohio, 44128, 49 edition, 2007.

[95] Mariano Cáceres, Eduardo Guzmán, Agustín Alvarez-Costa, Francisco Ortega, Ramón G. Rubio, Carlos Coviella, Pablo L. Santo Orihuela, Claudia V. Vassena, and Alejandro Lucia. Surfactantless emulsions containing eugenol for imidacloprid solubilization: Physicochemical characterization and toxicity against insecticide-resistant cimex lectularius. *Molecules*, 25(10):2290, 2020.

[96] Robert B. Best and Jeetain Mittal. Protein simulations with an optimized water model: cooperative helix formation and temperature-induced unfolded state collapse. *The journal of physical chemistry B*, 114(46):14916–14923, 2010.

[97] Junmei Wang, Romain M. Wolf, James W. Caldwell, Peter A. Kollman, and David A. Case. Development and testing of a general amber force field. *Journal of computational chemistry*, 25(9):1157–1174, 2004.

[98] Jose L.F. Abascal and Carlos Vega. A general purpose model for the condensed phases of water: Tip4p/2005. *The Journal of chemical physics*, 123(23):234505, 2005.

[99] Viktor Hornak, Robert Abel, Asim Okur, Bentley Strockbine, Adrian Roitberg, and Carlos Simmerling. Comparison of multiple amber force fields and development of improved protein backbone parameters. *Proteins: Structure, Function, and Bioinformatics*, 65(3):712–725, 2006.

[100] Sandrasegaram Gnanakaran and Angel E. García. Helix-coil transition of alanine peptides in water: Force field dependence on the folded and unfolded structures. *Proteins: Structure, Function, and Bioinformatics*, 59(4):773–782, 2005.

[101] Wendy D. Cornell, Piotr Cieplak, Christopher I. Bayly, Ian R. Gould, Kenneth M. Merz, David M. Ferguson, David C. Spellmeyer, Thomas Fox, James W. Caldwell, and Peter A. Kollman. A second generation force field for the simulation of proteins, nucleic acids, and organic molecules. *Journal of the American Chemical Society*, 117(19):5179–5197, 1995.

[102] Erik J. Thompson, Allison J. DePaul, Sarav S. Patel, and Eric J. Sorin. Evaluating molecular mechanical potentials for helical peptides and proteins. *PLoS One*, 5(4): e10056, 2010.

[103] Takao Yoda, Yuji Sugita, and Yuko Okamoto. Secondary-structure preferences of force fields for proteins evaluated by generalized-ensemble simulations. *Chemical physics*, 307(2-3):269–283, 2004.

[104] Yong Duan, Chun Wu, Shibasish Chowdhury, Mathew C. Lee, Guoming Xiong, Wei Zhang, Rong Yang, Piotr Cieplak, Ray Luo, Taisung Lee, et al. A point-charge force field for molecular mechanics simulations of proteins based on condensed-phase quantum mechanical calculations. *Journal of computational chemistry*, 24 (16):1999–2012, 2003.

[105] Kresten Lindorff-Larsen, Paul Maragakis, Stefano Piana, Michael P. Eastwood, Ron O. Dror, and David E. Shaw. Systematic validation of protein force fields against experimental data. *PloS one*, 7(2):e32131, 2012.

[106] Robert B. Best, Wenwei Zheng, and Jeetain Mittal. Balanced protein–water interactions improve properties of disordered proteins and non-specific protein association. *Journal of chemical theory and computation*, 10(11):5113–5124, 2014.

[107] Alan W. Sousa Da Silva and Wim F. Vranken. Acpype-antechamber python parser interface. *BMC research notes*, 5(1):1–8, 2012.

[108] Herman J.C. Berendsen, David van der Spoel, and Rudi van Drunen. Gromacs: a message-passing parallel molecular dynamics implementation. *Computer physics communications*, 91(1-3):43–56, 1995.

[109] Sander Pronk, Szilárd Páll, Roland Schulz, Per Larsson, Pär Bjelkmar, Rossen Apostolov, Michael R. Shirts, Jeremy C. Smith, Peter M. Kasson, David van der Spoel, et al. Gromacs 4.5: a high-throughput and highly parallel open source molecular simulation toolkit. *Bioinformatics*, 29(7):845–854, 2013.

[110] Berk Hess. P-lincs: A parallel linear constraint solver for molecular simulation. *Journal of chemical theory and computation*, 4(1):116–122, 2008.

[111] Ulrich Essmann, Lalith Perera, Max L. Berkowitz, Tom Darden, Hsing Lee, and Lee G. Pedersen. A smooth particle mesh ewald method. *The Journal of chemical physics*, 103(19):8577–8593, 1995.

[112] Giovanni Bussi, Davide Donadio, and Michele Parrinello. Canonical sampling through velocity rescaling. *The Journal of chemical physics*, 126(1):014101, 2007.

[113] Michele Parrinello and Aneesur Rahman. Polymorphic transitions in single crystals: A new molecular dynamics method. *Journal of Applied physics*, 52(12):7182–7190, 1981.

[114] In-Chul Yeh and Max L. Berkowitz. Ewald summation for systems with slab geometry. *The Journal of chemical physics*, 111(7):3155–3162, 1999.

[115] Thomas J. Lane, Diwakar Shukla, Kyle A. Beauchamp, and Vijay S. Pande. To milliseconds and beyond: challenges in the simulation of protein folding. *Current opinion in structural biology*, 23(1):58–65, 2013.

[116] Young Min Rhee, Eric J. Sorin, Guha Jayachandran, Erik Lindahl, and Vijay S. Pande. Simulations of the role of water in the protein-folding mechanism. *Proceedings of the National Academy of Sciences*, 101(17):6456–6461, 2004.

[117] Christopher D. Snow, Eric J. Sorin, Young Min Rhee, and Vijay S. Pande. How well can simulation predict protein folding kinetics and thermodynamics? *Annu. Rev. Biophys. Biomol. Struct.*, 34:43–69, 2005.

[118] Robert B. Best. Atomistic molecular simulations of protein folding. *Current opinion in structural biology*, 22(1):52–61, 2012.

[119] Ken A. Dill and Justin L. MacCallum. The protein-folding problem, 50 years on. *science*, 338(6110):1042–1046, 2012.

[120] Daria Maltseva, Ragnheidur Gudbrandsdottir, Gönül Kizilsavas, Dominik Horinek, and Grazia Gonella. Location and conformation of the lkα14 peptide in water/ethanol mixtures. *Langmuir*, 2020.

[121] Nami Hirota, Kazuko Mizuno, and Yuji Goto. Group additive contributions to the alcohol-induced α-helix formation of melittin: implication for the mechanism of the alcohol effects on proteins. *Journal of molecular biology*, 275(2):365–378, 1998.

[122] Kai Griebenow and Alexander M. Klibanov. On protein denaturation in aqueous-organic mixtures but not in pure organic solvents. *Journal of the American Chemical Society*, 118(47):11695–11700, 1996.

[123] Lars Schmüser, Steven Roeters, Helmut Lutz, Sander Woutersen, Mischa Bonn, and Tobias Weidner. Determination of absolute orientation of protein α-helices at interfaces using phase-resolved sum frequency generation spectroscopy. *The Journal of Physical Chemistry Letters*, 8(13):3101–3105, 2017.

[124] Michael Deighan and Jim Pfaendtner. Exhaustively sampling peptide adsorption with metadynamics. *Langmuir*, 29(25):7999–8009, 2013.

[125] Gonzalo Vazquez, Estrella Alvarez, and Jose M. Navaza. Surface tension of alcohol water+ water from 20 to 50. degree. c. *Journal of chemical and engineering data*, 40(3):611–614, 1995.

[126] Ibrahim S. Khattab, Farzana Bandarkar, Mohammad A. Abolghassemi Fakhree, and Abolghasem Jouyban. Density, viscosity, and surface tension of water+ ethanol mixtures from 293 to 323k. *Korean Journal of Chemical Engineering*, 29(6):812–817, 2012.

[127] Carlos Vega and Enrique De Miguel. Surface tension of the most popular models of water by using the test-area simulation method. *The Journal of chemical physics*, 126(15):154707, 2007.

[128] Akihiro Wakisaka and Takahiro Ohki. Phase separation of water–alcohol binary mixtures induced by the microheterogeneity. *Faraday Discussions*, 129:231–245, 2005.

[129] Martina Požar, Bernarda Lovrinčević, Larisa Zoranić, Tomislav Primorać, Franjo Sokolić, and Aurélien Perera. Micro-heterogeneity versus clustering in binary mixtures of ethanol with water or alkanes. *Physical Chemistry Chemical Physics*, 18 (34):23971–23979, 2016.

[130] Ivo Jukić, Martina Požar, and Bernarda Lovrinčević. Comparative analysis of ethanol dynamics in aqueous and non-aqueous solutions. *Physical Chemistry Chemical Physics*, 22(41):23856–23868, 2020.

[131] Marco Fioroni, M. Dolores Diaz, Klaus Burger, and Stefan Berger. Solvation phenomena of a tetrapeptide in water/trifluoroethanol and water/ethanol mixtures: a diffusion nmr, intermolecular noe, and molecular dynamics study. *Journal of the American Chemical Society*, 124(26):7737–7744, 2002.

[132] Shuting Zhang, Brian Andrews, Reinhard Schweitzer-Stenner, and Brigita Urbanc. Intrinsic conformational dynamics of alanine in water/ethanol mixtures: An experiment-driven molecular dynamics study. *The Journal of Physical Chemistry B*, 124(51):11600–11616, 2020.

[133] Ken A. Dill. Dominant forces in protein folding. *Biochemistry*, 29(31):7133–7155, 1990.

[134] Emily E. Meyer, Kenneth J. Rosenberg, and Jacob Israelachvili. Recent progress in understanding hydrophobic interactions. *Proceedings of the National Academy of Sciences*, 103(43):15739–15746, 2006.

[135] David Eisenberg and Andrew D. McLachlan. Solvation energy in protein folding and binding. *Nature*, 319(6050):199–203, 1986.

[136] Antonio Cipiciani, Giuseppe Onori, and Gianfranco Savelli. Structural properties of water-ethanol mixtures: a correlation with the formation of micellar aggregates. *Chemical physics letters*, 143(5):505–509, 1988.

[137] Katarzyna Łudzik, Kinga Kustrzepa, Henryk Piekarski, and Małgorzata Jóźwiak. Application of conductance study to analyze micellization behavior of cationic gemini surfactants in water and water–ethanol solvent mixed media. *Journal of Chemical & Engineering Data*, 61(3):1047–1053, 2016.

[138] Halide Akbaş and Çiğdem Batıgöç. Micellization of dodecylpyridinium chloride in water-ethanol solutions. *Colloid journal*, 70(2):127–133, 2008.

[139] Yoshikiyo Moroi, Hiroyuki Takagi, Shigemi Nagadome, Yutaka Hirata, and Gohsuke Sugihara. Micelle formation and phase behavior of dodecylammonium heptafluorobutyrate in a water-ethanol mixture. *Journal of colloid and interface science*, 149(1):252–255, 1992.

[140] Daizo Hamada and Yuji Goto. The equilibrium intermediate of β-lactoglobulin with non-native α-helical structure, 1997.

[141] Johannes Kirschner, Anderson H.A. Gomes, Ricardo R.T. Marinho, Olle Björneholm, Hans Ågren, Vincenzo Carravetta, Niklas Ottosson, Arnaldo Naves de Brito, and Huib J. Bakker. The molecular structure of the surface of water–ethanol mixtures. *Physical Chemistry Chemical Physics*, 23(19):11568–11578, 2021.

[142] Zhi X. Li, Jian R. Lu, Dmitrii A. Styrkas, Robert K. Thomas, Adrian R. Rennie, and Jeffrey Penfold. The structure of the surface of ethanol/water mixtures. *Molecular Physics*, 80(4):925–939, 1993.

Acknowledgements

In the end I would like to thank all those who supported me in accomplishing this work.

First and foremost, my supervisor Prof. Dr. Dominik Horinek who always had an open ear for questions and always took his time to discuss my work in depth.

I also want to thank Prof. Dr. Emanuel Schneck for agreeing to review this thesis and his travels to Regensburg for the examination.

Furthermore I want to thank Prof. Dr. Werner Kunz for being an examiner and Prof. Dr. Manfred Scheer for chairing my doctoral defence.

I want to thank Dr. Grazia Gonella and Daria Maltseva at the Max Planck Institute for Polymer Research, for collaboration.

I am grateful to Philipp Dullinger, Tuan-Minh Do and Þröstur Guðmundsson for proof-reading this thesis and to my father Guðbrandur Guðmundsson for providing two graphics of micelles.

I also want to thank all the people I have shared office and lunch with over the last years (Philipp, Minh, Chris, Eva and Anna among others) for interesting and fun conversations about our respective research as well as diverse others themes.

Last but not least I want to thank all my friends and relations for their support, especially my parents and my sister.

Kærar þakkir öllsömul