

## **Library Declaration and Deposit Agreement**

### **1. STUDENT DETAILS**

Alaina Emanuella O. Emmanuel  
0721368(902)

### **2. THESIS DEPOSIT**

- 2.1 I understand that under my registration at the University, I am required to deposit my thesis with the University in BOTH hard copy and in digital format. The digital version should normally be saved as a single pdf file.
- 2.2 The hard copy will be housed in the University Library. The digital version will be deposited in the University's Institutional Repository (WRAP). Unless otherwise indicated (see 2.3 below) this will be made openly accessible on the Internet and will be supplied to the British Library to be made available online via its Electronic Theses Online Service (EThOS) service. [At present, theses submitted for a Masters degree by Research (MA, MSc, LLM, MS or MMedSci) are not being deposited in WRAP and not being made available via EThOS. This may change in future.]
- 2.3 In exceptional circumstances, the Chair of the Board of Graduate Studies may grant permission for an embargo to be placed on public access to the hard copy thesis for a limited period. It is also possible to apply separately for an embargo on the digital version. (Further information is available in the Guide to Examinations for Higher Degrees by Research.)
- 2.4 (a) Hard Copy I hereby deposit a hard copy of my thesis in the University Library to be made publicly available to readers immediately.  
I agree that my thesis may be photocopied.
- (b) Digital Copy I hereby deposit a digital copy of my thesis to be held in WRAP and made available via EThOS.  
My thesis can be made publicly available online.

### **3. GRANTING OF NON-EXCLUSIVE RIGHTS**

Whether I deposit my Work personally or through an assistant or other agent, I agree to the following: Rights granted to the University of Warwick and the British Library and the user of the thesis through this agreement are non-exclusive. I retain all rights in the thesis in its present version or future versions. I agree that the institutional repository administrators and the British Library or their agents may, without changing content, digitise and migrate the thesis to any medium or format for the purpose of future preservation and accessibility.

#### 4. DECLARATIONS

(a) I DECLARE THAT:

- I am the author and owner of the copyright in the thesis and/or I have the authority of the authors and owners of the copyright in the thesis to make this agreement. Reproduction of any part of this thesis for teaching or in academic or other forms of publication is subject to the normal limitations on the use of copyrighted materials and to the proper and full acknowledgement of its source.
- The digital version of the thesis I am supplying is the same version as the final, hardbound copy submitted in completion of my degree, once any minor corrections have been completed.
- I have exercised reasonable care to ensure that the thesis is original, and does not to the best of my knowledge break any UK law or other Intellectual Property Right, or contain any confidential material.
- I understand that, through the medium of the Internet, files will be available to automated agents, and may be searched and copied by, for example, text mining and plagiarism detection software.

(b) IF I HAVE AGREED (in Section 2 above) TO MAKE MY THESIS PUBLICLY AVAILABLE DIGITALLY, I ALSO DECLARE THAT:

- I grant the University of Warwick and the British Library a licence to make available on the Internet the thesis in digitised format through the Institutional Repository and through the British Library via the EThOS service.
- If my thesis does include any substantial subsidiary material owned by third-party copyright holders, I have sought and obtained permission to include it in any version of my thesis available in digital format and that this permission encompasses the rights that I have granted to the University of Warwick and to the British Library.

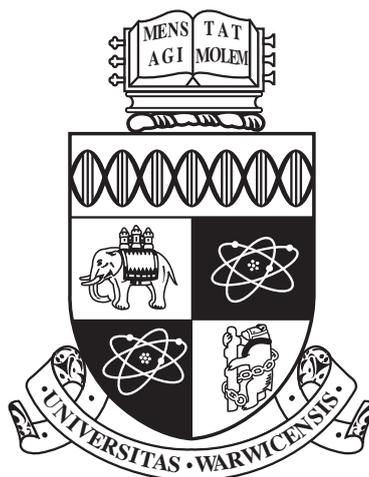
#### 5. LEGAL INFRINGEMENTS

I understand that neither the University of Warwick nor the British Library have any obligation to take legal action on behalf of myself, or other rights holders, in the event of infringement of intellectual property rights, breach of contract or of any other right, in the thesis.

---

*Please sign this agreement and return it to the Graduate School Office when you submit your thesis.*

Student's signature: .....Date: .....



**Molecular Simulation of Ice Growth Inhibition by  
Biomimetic Antifreeze Macromolecules**

by

**Alaina Emanuella O. Emmanuel**

**Thesis**

Submitted to the University of Warwick

for the degree of

**Doctor of Philosophy**

.....

December 2015

THE UNIVERSITY OF  
**WARWICK**

# Contents

<b>Acknowledgments</b>	<b>ii</b>
<b>Declarations</b>	<b>iii</b>
<b>Abstract</b>	<b>iv</b>
<b>Chapter 1 Introduction</b>	<b>1</b>
1.1 Statement of the Problem . . . . .	1
1.2 Background . . . . .	3
1.2.1 Ice structure, crystallization and growth . . . . .	3
1.2.2 Three Macroscopic Antifreeze Properties . . . . .	7
1.2.3 Essential Polymer physics . . . . .	10
1.2.4 AFGPs . . . . .	10
1.3 Key experimental studies . . . . .	12
1.3.1 Previous experiments on the mechanisms of AFGPs . . . . .	12
1.3.2 Small molecules e.g. Sugars: C linked and O linked . . . . .	17
1.3.3 Polymer antifreezes . . . . .	19
1.4 Key simulation studies . . . . .	22
1.4.1 Simulations of small antifreeze molecules . . . . .	22
1.4.2 Simulations of antifreeze proteins or glycoproteins . . . . .	25
1.4.3 Polymer simulations . . . . .	30
<b>Chapter 2 Computer Simulation Methodology</b>	<b>40</b>
2.1 Statistical mechanics . . . . .	40
2.2 Forcefields . . . . .	41
2.2.1 Bonded interactions . . . . .	41
2.2.2 Non-bonded interactions . . . . .	42
2.2.3 Water forcefields: 3-site and 4-site models . . . . .	45
2.3 Potential energy minimization . . . . .	46

2.4	Molecular dynamics . . . . .	47
2.4.1	Time integration algorithms . . . . .	48
2.4.2	Constant temperature MD . . . . .	49
2.4.3	Constant pressure MD . . . . .	51
2.4.4	Bond constraints . . . . .	52
2.4.5	Periodic boundary conditions . . . . .	52
<b>Chapter 3 Mixing and validation of OPLS-AA and TIP4P/Ice force-</b>		
<b>fields for simulations of ice growth in the presence of additives</b>		<b>54</b>
3.1	Introduction . . . . .	54
3.2	Methods . . . . .	57
3.2.1	Forcefield parameters . . . . .	57
3.2.2	Simulation parameters . . . . .	57
3.3	Analysis methods . . . . .	61
3.3.1	Hydrogen Bonding . . . . .	61
3.3.2	Diffusion coefficients . . . . .	61
3.3.3	Orientational Order Parameter . . . . .	62
3.4	Results . . . . .	62
3.4.1	All water models . . . . .	62
3.4.2	Effects of temperatures on TIP4P and TIP4P/Ice . . . . .	66
3.4.3	The effects of dilute salt concentrations on TIP4P/Ice and TIP4P . . . . .	71
3.4.4	The effect of ethanol concentrations on TIP4P/Ice and TIP4P	81
3.4.5	The effect of cut-offs on TIP4P and TIP4P/Ice . . . . .	89
3.4.6	Melting and freezing . . . . .	93
3.5	Conclusion . . . . .	96
<b>Chapter 4 MD simulations of antifreeze active and antifreeze inactive</b>		
<b>polymers in solution</b>		<b>99</b>
4.1	Introduction . . . . .	99
4.2	Methods . . . . .	100
4.2.1	Forcefield parameters . . . . .	100
4.2.2	Simulation parameters . . . . .	100
4.2.3	Initial configuration and equilibration . . . . .	101
4.3	Results . . . . .	103
4.3.1	Macromolecules shape and size . . . . .	103
4.3.2	Water ordering or disordering properties . . . . .	105
4.3.3	Lattice matching properties . . . . .	111

4.3.4	Solvent accessibility . . . . .	114
4.3.5	Hydrogen bonding . . . . .	118
4.3.6	Polymer flexibility and movement . . . . .	122
4.4	Discussion . . . . .	124
4.5	Conclusion . . . . .	127

**Chapter 5 The effect of single poly(vinyl alcohol) and poly(ethylene glycol) polymers on ice crystal growth rates 129**

5.1	Introduction . . . . .	129
5.2	Methods . . . . .	130
5.2.1	Forcefield parameters . . . . .	130
5.2.2	Simulation parameters . . . . .	130
5.2.3	Initial configuration and Equilibration . . . . .	131
5.3	Results . . . . .	132
5.3.1	Total energy profiles . . . . .	132
5.3.2	Growth/melt kinetics . . . . .	133
5.3.3	Evidence of superheating . . . . .	141
5.3.4	Polymer residence regions . . . . .	142
5.3.5	Ice polymorphism . . . . .	142
5.3.6	Ice growth around polymers . . . . .	146
5.4	Discussion . . . . .	156
5.5	Conclusion . . . . .	158

# Acknowledgments

# Declarations

I confirm that my thesis has been prepared in accordance with the University's guidelines on the presentation of a research thesis. This thesis is submitted to the University of Warwick in support of my application for the degree of Doctor of Philosophy. It has been composed by myself, except where acknowledged in the text, and has not been submitted in any previous application for any degree.

Some work in this thesis have been submitted for publishing by myself. These are:

**Paper 1** Title1

**Paper 2** Title2

**Paper 3** Title3

**Paper 4** Title4

# Abstract

# Chapter 1

## Introduction

### 1.1 Statement of the Problem

The growth of ice crystals within tissues is a serious threat to the cryo-preservation of organs, food and crops in cold climates. In the UK alone, the waiting list for organ transplantation currently exceeds 7000[1]. It is important that no organ is wasted. Inspired by antifreeze proteins and glycoproteins (AF(G)Ps)– proteins which permit the survival of organisms in sub-freezing climates – researchers are now seeking better and cheaper synthetic methods of preventing the growth of ice crystals or reducing them into manageable sizes. [Mention the internal and external mechanical damage, dehydration, pockets of high salinity] Challenge and Gaps: Our collaborators recently demonstrated the remarkable ability of low concentrations of polyvinyl alcohol (PVA) to preserve mammalian red blood cells after freeze-thawing. However, the molecular mechanisms of the antifreeze actions of PVA remain unknown. Role of Simulation: Here we employ molecular dynamics computer simulations to address this challenge. This will enable us to correlate the structural properties of this polymer with different antifreeze characteristics i.e. thermal hysteresis, dynamic ice shaping or reduced ice recrystallization, and facilitate the design of new, improved antifreeze polymers for specific control over crystal size and morphology. What we have done: Our initial work has focused on PVA, and its antifreeze inactive isomer, polyethylene glycol (PEG). We have simulated various chain lengths of PVA, PEG and poly-L-hydroxyproline in solution and at the ice/water interface using the TIP4P/Ice water model and the OPLS-AA forcefield for the polymers. One hypothesis is that the antifreeze molecules structure liquid water either in an ice-like configuration, which could promote binding to a growing ice crystal, or in a disordered configuration that could inhibit ice crystal growth. Our simulations

indicate that PVA does not have significant ordering or disordering effect on surrounding water molecules in solution. On the other hand the O-O separation of PVA in solution is 0.46 nm (compared to 0.31 nm and 0.39 nm in PEG), which corresponds to the O-O separation distance on the prism plane of ice, suggesting that direct binding to ice may play an important role in the mechanism. We are currently investigating the interactions of PVA and PEG with an ice/water interface and how this affects the rate of ice crystal growth from two-phase coexistence studies. Ultimately the development of these polymers may prove as better alternatives to conventional cryoprotectants like dimethyl sulfoxide (DMSO) and glycerol. Novelty or nonsense Thus far, simulations of PVA, PEG and poly L-hydroxyproline (HYP) in ice/water systems have not previously been explored with molecular dynamics simulation, and so this work presents a novel contribution to the growing antifreeze simulation community.

OR:

**Big picture:** The growth of ice crystals within tissues is a serious threat to the cryo-preservation of organs, food and crops in cold climates. In the UK alone, the waiting list for organ transplantation currently exceeds 7000[1], therefore creating a huge emphasis on the protecting the limited organs available. Inspired by antifreeze proteins and glycoproteins (AF(G)Ps)– proteins which permit the survival of organisms in sub-freezing climates – researchers are now seeking better and cheaper synthetic methods of preventing the growth of ice crystals or reducing them into manageable sizes. [Mention the internal and external mechanical damage, dehydration, pockets of high salinity] **Challenge and Gaps:** Our collaborators recently demonstrated the remarkable ability of low concentrations of polyvinyl alcohol (PVA) to preserve mammalian red blood cells after freeze-thawing. Unfortunately however the molecular mechanisms of the antifreeze actions of PVA remain unknown. [Mention why and what is specifically challenging about identifying the molecular mechanism. This is on order to add depth to this section and help to explain the role of the simulation in the next part.] **Role of Simulation:** By using molecular dynamics computer simulations we can begin to address this challenge by enabling us to correlate the structural properties of this polymer with different antifreeze characteristics i.e. thermal hysteresis, dynamic ice shaping or reduced ice recrystallization, and facilitate the design of new, improved antifreeze polymers for specific control over crystal size and morphology. **What we have done:** Our initial work has focused on PVA and its antifreeze inactive isomer, polyethylene glycol (PEG). We have simulated various chain lengths of PVA, PEG and poly-L-hydroxyproline in solution and at the ice/water interface using the TIP4P/Ice water

model and the OPLS-AA forcefield for the polymers.

One hypothesis is that the antifreeze molecules structure liquid water in either:

- 1) an ice-like configuration - which could promote binding to a growing ice crystal,;
- 2) in a disordered configuration that could inhibit ice crystal growth.

Our simulations indicate that PVA does not have significant ordering or disordering effect on surrounding water molecules in solution. Interestingly the O-O separation of PVA in solution is 0.46 nm (compared to 0.31 nm and 0.39 nm in PEG), which corresponds to the O-O separation distance on the prism plane of ice, suggesting that direct binding to ice may play an important role in the mechanism.

Our report focuses on these interactions of PVA and PEG with an ice/water interface and in turn, how this affects the rate of ice crystal growth from two-phase coexistence studies. Ultimately the development of these polymers may prove to be better alternatives than conventional cryoprotectants such as dimethyl sulfoxide (DMSO) and glycerol.

**Novelty:** Thus far, simulations of PVA, PEG and poly L-hydroxyproline (HYP) in ice/water systems have not previously been explored with molecular dynamics simulation, and so this work presents a novel contribution to the growing antifreeze simulation community. Give a thesis overview Outline of subsequent Chapters

## 1.2 Background

### 1.2.1 Ice structure, crystallization and growth

The most common ice form on the earth is hexagonal ice (1h), which is stable at 273 K and 1 atm. Although eighteen other crystalline phases exist depending on the temperature and pressure, 1h can be distinguished by its hexagonal symmetry. It possesses a regular structure of water molecules where a single oxygen atom is bonded to four hydrogen; two of which are covalently bonded and the others remain weakly hydrogen bonded. Consequently, four equidistant water molecules surround each water molecule and so the ice lattice structure is held firmly in place.

Many of the physical processes and interactions we observe in both water and ice are predominantly due to the extensive hydrogen-bonding network and their bond angles. For instance, the characteristic trait that ice with a density of  $0.9167 \text{ g cm}^3$  - is less dense than liquid water ( $0.9998 \text{ g cm}^3$ ) at 273.15 K. In fact it is well known that near tetrahedral arrangements of water molecules permeate through the ordered ice lattice, with typical bond angle closer to  $109.5^\circ$  [12], whereas those found in liquid

water are normally slightly smaller ( $<104.5$  degrees)[13][26]

Although water and ice are ubiquitous in nature, a substantial amount of research is still needed for us to fully understand the two phases and various transition states between them, particularly about the quasi liquid layer (QLL) a state approx. 0.10-0.15 nm [32] thick that lies between liquid and solid forms of ice.

Crystallization of ice is a two-step process composed of nucleation and ice crystal growth. Nucleation is described as the first step in the formation of a new phase, where small, stable clusters of ice-like water molecules are formed within a parent phase, like a body of liquid water. [5] The sizes of these clusters fluctuate whilst some grow or completely melt away. Once a cluster reaches a critical size, the embryo will continue to grow into crystal, which in turn facilitates the growth of a new phase. [24]

The classical nucleation theory is conventionally used to describe this whole process in terms of free energy [6][7][8]. Water molecules at the surface of the crystal have fewer neighbours and are not able to form the optimum number of hydrogen bonds so they experience a higher free energy than molecules within the crystal. [EX13] As the cluster grows, the surface-area-to-volume ratio is quite large and molecules that populate the surface are less well bound to their neighbours than those in the bulk. This causes the nucleus to be unstable, as a result, water molecules migrate back into the QLL and initial liquid phase. Beyond a critical size, a smaller surface-area-to-volume ratio is formed and the initial energy cost of maintaining a growing solid/liquid interface no longer outweighs the benefits of crystal growth. In fact, the old phase becomes increasingly metastable and a small fluctuation in density or another external property can spur the sudden growth of the new phase. [5]

Under laboratory conditions, pure water can withstand supercooling up to 236 K before ice is spontaneously formed. In this instance nucleation is described as homogenous and is unlikely to occur for two reasons. Firstly, the probability of nucleation is hampered by the free energy penalty for creating an interface. Secondly, water in most conditions contains small particles in them that act as seeds, templates or platforms for ice crystal nucleation, and so decrease the interfacial energy.

This type of nucleation is more likely to happen and is known as heterogenous nucleation. Depending on the particle's resemblance to the molecular structure and orientation of ice, they allow water molecules to become quickly oriented, similar to an ice nucleus. [15][18][TX14] Therefore these particles can actually be used to facilitate crystal growth and exert greater control over the rate of nucleation. [5][14] It is thought organisms also attempt to alter or manage ice crystal growth within

their systems using this same approach, which we discuss in greater detail in section (please insert your reference here)

Hexagonal ice (Fig. 1.1) has three different faces: the secondary prism plane, which is the fastest to grow and also the most stable [9], the prism plane, and finally the basal plane, which grows in the direction of the c-axis.

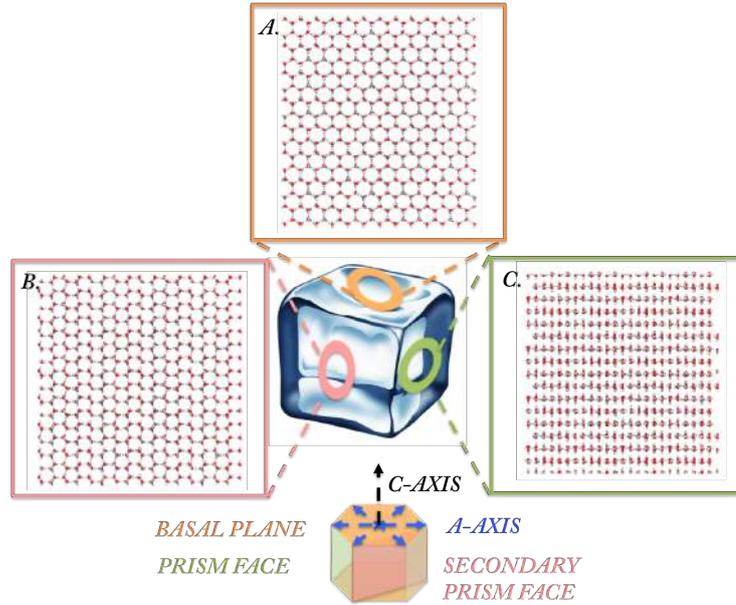


Figure 1.1: Three types of faces on the  $I_h$  crystal. Key: A. basal plane, B. secondary prism plane and C. prism face.

In this report we will refer to heating or supercooling ( $\Delta T$ ) as the difference between the melting temperature of ice ( $T_m$ ) and a temperature ( $T$ ) of interest. i.e.  $\Delta T = |T_m - T|$ , where positive and negative signs will be included to indicate higher and lower temperatures respectively. Ordinarily, at slight undercooling (up to  $\Delta T \approx -2 \text{ K}$ ) [3], ice crystals start to grow into spherical shapes in the absence of any ice shaping additives. However, at greater undercooling ( $\Delta T \approx -5 \text{ K}$ ) the anisotropic growth occurs due to increasing disparity in growth rates between the various faces of ice. For instance the secondary prism plane is known to grow much faster than the other planes and continued addition of water to the plane builds them out until it eventually grows itself out of existence. As a result the initial spherical crystals eventually turn into hexagonal prisms or columns with the smoothest, and slowest faces exposed (Fig. 1.2)[10] [3][11][16][17]. Its important to note that crystal growth rate is restricted by the degree of supercooling because the diffusion and re-orientation of water molecules become rate-limiting steps.

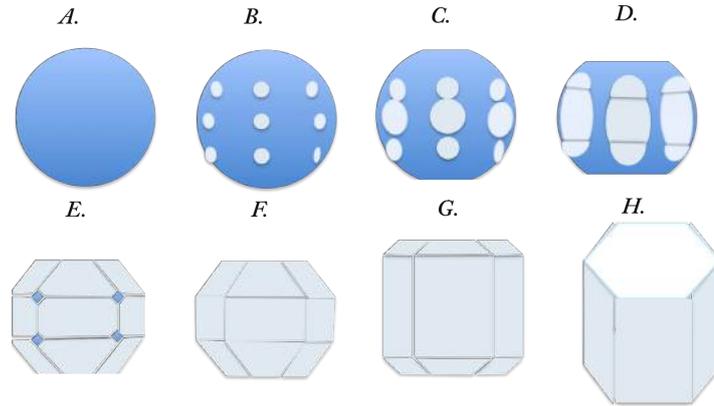


Figure 1.2: Schematic of growth process from spherical ice to a hexagonal disk. Adapted from Ref. [23] [23].

Using ellipsometry Furukawa *et al* also discovered that for each of the planes or faces of ice, their crystal habit and QLL thickness were also, unsurprisingly, dependent on the temperature.[19][20] Similarly Pruppachar reported that the growth rates of basal planes on small, free-standing crystals could be described by a square dependence on supercooling between 273.15 K –264.15 K . i.e.  $v(\text{cms}^{-1})=0.035 |\Delta T|^{2.2}$ [21][22]. More formulas for growth rates derived from various research are shown in Table [1]. However this piece of work is the most frequently cited in ice growth simulation papers. In this study the maximum growth rates were observed at  $\sim 255$  K with approximately  $10.5 \text{ cms}^{-1}$  velocity. Unlike earlier investigations, the study encompassed a large range between 272.65 K –253.15 K and showed that at temperatures lower than 261.15 K (higher supercooling), the square dependence was lost and became linear. At a narrow temperature range of 264.15 K –261.15 K , an unusual pattern appeared: the growth rates of individual repeats became noticeably more scattered. However its worth noting that exact standard deviations or errors were not shown for this study. [21] Slightly faster growth rates were also determined for the prism planes under sampled temperatures, but empirical relations were not reported in many of these studies [25].

Authors	Plane	Temperature	Rates	Ref
Hallet	Basal	273.15-253.15 K	$0.08 \Delta T ^{1.9}$	[25]
Pruppachar	Basal*	273.15-264.15 K	$0.035 \Delta T ^{2.2}$	[21][22]
Hillig & Turnbull	-	-	$0.16 \Delta T ^{1.7}$	[29][28]
Lindenmeyer & Chalmers	Basal*	271.15-266.65 K	$0.0228 \Delta T ^{2.39}$	[30][27]
Tirmizi & Gill	Basal & prism	272.95-272.15 K	$0.0187 \Delta T ^{2.09}$	[31]

Table 1.1: Note: The rates are plotted on a logarithmic scale

### 1.2.2 Three Macroscopic Antifreeze Properties

In the last section we discussed some of the crystal growth behaviors of ice in the absence of organic or polymeric additives. In the presence of additives, crystal ice growth behaves differently. In this section we cover three macroscopic properties of antifreeze molecules, which we will later use to characterize our polymers and compare to what has already been done in experiments. The three properties are ice recrystallization inhibition, thermal hysteresis and dynamic ice shaping.

#### Ice recrystallization Inhibition

During thawing or frozen storage, larger ice crystals normally grow at the expense of neighboring smaller ones in order to reduce the interfacial energy of the system. This process is known as ice recrystallization and is often likened to a form of Ostwald ripening. Antifreeze molecules that inhibit or slow this process give rise to the term ice recrystallization inhibition (IRI). [1][10] This particular antifreeze property is thought to be the most useful of the three, particularly for cryopreservation, because it can potentially reduce ice crystals to manageable sizes. Ice recrystallisation is affected by surface adsorbing polymers or proteins, which interfere with the movement of water molecules within the ice boundaries, e.g. ice surface modifications (e.g. during surface adsorption), disruption of water diffusion rates and manipulation of liquid channel thicknesses [1][2][10]. Because even non-ice-adsorbing macromolecules are sometimes able to manipulate water diffusion rates, the inclusion of small solutes like salt or sugars are often used to differentiate between adsorbing and non-adsorbing polymers whilst testing for IRI. The reason for this is that salts like NaCl are able to increase the size of the liquid channels, counteracting any non-specific, retarding effects on water mobility. [10] [1][2][11]

IRI is widely investigated using the “splat” assay first developed by Knight *et al* [8], although other techniques have also been developed.[7] The setup of a splat

assay is shown in Fig 1.3. In this test a syringe containing the solution of interest is suspended approximately 2 M above a 193.15 K chilled glass plate and a small drop is released. On immediate contact the drop is quickly frozen, and annealed at a warmer temperature close to 267.15 K , yielding a polycrystalline wafer after 30 minutes. A high-resolution camera is used to continuously monitor this droplet and its crystals from start to finish. At the end of the experiment, image analysis tools like ImageJ[12] are used to determine the mean ice grain size (MIGS) or the mean largest grain size (MLGS) from a number of randomly selected or largest crystals in the field of view respectively. The results are then compared to a negative control, typically a Phosphate Buffered Saline (PBS) solution, which serves as a solvent for the most part. The final results are normally reported as a percentage of the negative control. [3][4][5][6][8][9]

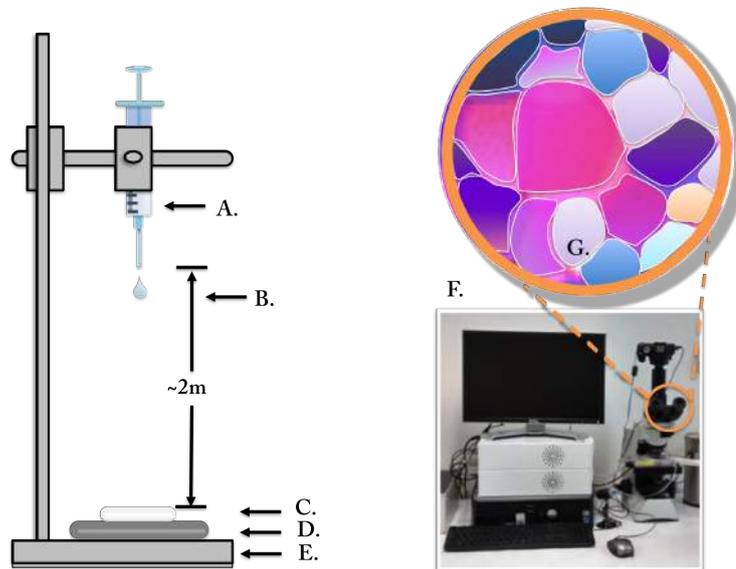


Figure 1.3: Splat assay set up. Key: A. syringe, B. distance, C. chilled glass plate, D. cooling aluminium plate, E. stand, F. imaging system G. schematic of micrograph.

### Thermal Hysteresis

Thermal hysteresis is possibly the most studied of the three properties. Ordinarily the melting and freezing temperature of ice crystals are one and the same. In the presence of antifreeze molecules this might change: thermal hysteresis describes when the freezing temperature falls beneath the melting temperature and the difference between the two is called the thermal hysteresis gap. TH is thought to be caused by direct and irreversible adsorption unto specific planes of a growing ice

crystal surface. [2] As a result of direct pinning unto the surface, the local area becomes microscopically curved and energetically unfavorable to grow any further (Fig.1.4). This is known as the Kelvin effect or sometimes Gibbs-Thomson effect [4]. The size of the TH gap varies from molecule to molecule, however at temperatures within the TH region, ice crystals don't grow or shrink. [5] In literature, TH is usually described as non-colligative, which means that TH antifreeze molecules don't lower the freezing point in proportion to their concentrations [3]. Rather, they typically work more effectively than expected at lower concentrations. In order to assess this antifreeze property, a nanolitre osmometer is also used to accurately monitor the temperature and ice crystal size at time intervals.[5] Alternatively a differential scanning calorimeter can be used to automate the process, without the need for microscopic observation [6].

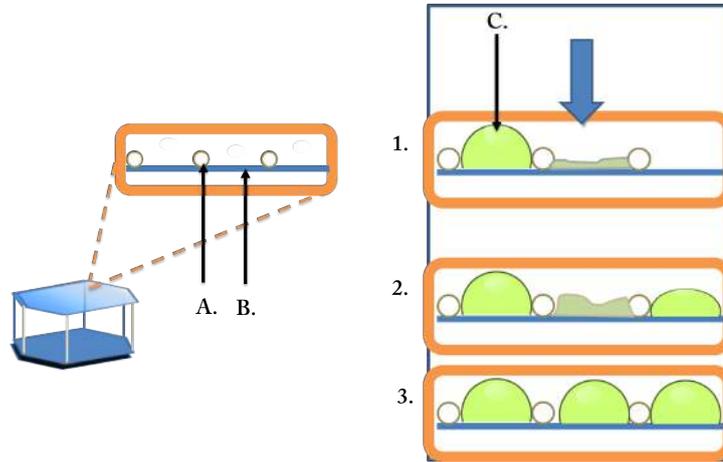


Figure 1.4: Splat assay set up. Key: A. antifreeze molecule, B. growing ice front, C. accumulating water molecules.

### Dynamics Ice Shaping

As mentioned earlier, ice crystals typically grow out into spherical shapes at temperatures close to the melting point ( $T_m$ ). However in the presence of antifreeze molecules the ice crystal shape becomes altered as it grows within the TH region (or at temperatures close to the  $T_m$ ), because the antifreeze molecules have the ability to interact closely with certain planes of ice.[1] The name for this process is dynamic ice shaping (DIS), and it can be monitored using a range of techniques, mainly hemispherical etching and a nanolitre osmometer. [1] The resulting ice crystals shapes are detailed in the following sections, which highlight various antifreezes

and the extent to which each of these three properties is exhibited.

### 1.2.3 Essential Polymer physics

The physical properties of large polymers can be described using statistics. These descriptions typically base themselves on one of various models of polymers' motion e.g. the Brownian motion, the Random walk, which is the simplest, or the Self-avoiding walk. The founder of this field is Paul Flory who also determined the radius of gyration ( $R_g$ ) for a polymer as a function of chain length (the Flory exponent ( $V$ )) under theta conditions (the conditions at which polymers will behave as an ideal chain [1]). The  $R_g$  describes how collapsed a polymer is in solution, and is calculated using Eq 1.1, where  $m_i$  is the mass of atom  $i$  and  $r_i$  is its position. It calculates the root mean square distance of a collection of atoms from the polymer' centre of mass ( $r_{mean}$ ). Equation 1.2 defines how  $R_g$  scales with chain length ( $N$ ) and molecular weight ( $M_w$ ). In a good solvent ( $V \approx 3/5$ ) [2] or theta solvent ( $V \approx 1/2$ ), a single polymer chain is fully extended or more swollen in solution, while in a bad solvent the chain is more compact or spherical in order to minimize polymer-solvent interactions ( $V \approx 1/3$ ) [3]. However for proteins, which typically have a well-defined secondary structure, it merely reflects how folded or unfolded it is in solution. Experimentally, the  $R_g$  is normally determined using (static) light scattering techniques like small angle neutron scattering or X-ray scattering [1][4]. It is important to note that the experimental  $R_g$  is reported as an ensemble average and that the polymers that make up the ensemble solution are in fact polydisperse, (although narrowly so depending on the type of polymerization used). Therefore direct comparison of experimental  $R_g$  and those obtained from computational simulations with a single polymer solution may be somewhat oversimplified.

$$R_g = \sqrt{\frac{\sum_{i=1}^N m_i (r_i - r_{mean})^2}{\sum_{i=1}^N m_i}} \quad (1.1)$$

$$R_g \propto M_w^V \quad (1.2)$$

### 1.2.4 AFGPs

In the north and south poles where the climate is continuously icy and the salty ( $\sim 0.45$  M) [14] sea waters rest at approximately 271.25 K [6], the evolution of biological antifreezes have enabled the survival of fish [2], insects [19], plants, bacteria and even fungi. [7][8][9] Two types of these biological antifreezes exist: antifreeze

proteins (AFP) and antifreeze glycoproteins (AFGP). From here on we shall refer to them both collectively as AF(G)Ps. To date AFPs have been investigated using a vast array of techniques; NMR spectroscopy [11][12], X-ray spectroscopy and chromatography to name a few, and these have been used to reveal their large range of structures. [10][14][15] As a result, AFPs have been characterized according to secondary structures, interactions and size, and thus far they are classified into 4 groups (Table 1)[3]. Briefly, type 1 and 4 are  $\alpha$ -helical and contain large amounts of alanine (>60%), whilst type 2 is cysteine rich and predominantly held together by disulphide bonds, and finally type 3 typically consists of a globular,  $\beta$ -sandwich structure. [1][10][14] AFGPs however are composed of regularly repeating structures, similar to synthetic polymers. Their monomers are glycotriptide units of [Ala-Ala-Pro/Thr]<sub>N=4-50</sub> [4][5], with a disaccharide ( $\beta$ -d-galactosyl-(1,3)- $\alpha$ -N-acetyl-d-galactosamine) attached onto the threonine residue via a glycosidic link (Table ??) [2][5]. To date only 8 classes have been identified and are often grouped as large (AFGP1-5) or small (AFGP6-8). Compared to AFPs, AFGPs are less studied because they are structurally complex and so are often problematic to synthesize or purify in the quantities needed for experimental studies to be made. [5][E3] As a result, there is actually less known about them despite the fact that they were the first to be isolated [6]. Both AFPs and AFGPs (AF(G)Ps) are able to exhibit thermal hysteresis, dynamics ice shaping and IRI activities – that is, antifreeze activities which are several orders of magnitude (300-500 times [10]) greater than their colligative properties. [13][7][6] In addition to this they have also been shown to alter the growth rates of ice. [22][23] AF(G)Ps do not necessarily display all three antifreeze properties at once; some may, for instance, only have IRI properties and negligible TH properties. [8] **[If it were not for their severe ice shaping properties, these would make them ideal and desirable macromolecules for applications involving cold storage of biological tissue. However, their ability to form harmful, “needle” — like crystal stand in the way. ]**

**[Some of the most promising cryoprotectant solutions incorporate AFGPs. Not only do they inhibit ice formation at any subzero temperature but they also interact with the cell membrane during the thermotropic phase transition state to avoid leakage, thereby addressing two issues at once. (The second effect will be discussed in Sect. 2.4.2.) Numerous factors that influence the success of post-thaw revival include the nature, the temperature, the concentration of the cryoprotectant, the rate of cooling and warming, the storage temperature, and the speed at which the cryoprotectant is added and removed.]**

[Insert key characteristics table here]

## 1.3 Key experimental studies

In this section we explore the key experimental studies for both biological antifreezes, for small and synthetic molecules.

### 1.3.1 Previous experiments on the mechanisms of AFGPs

#### Discovery of AF(G)Ps

The first BA was discovered in Notothonoid fish in the 1950s [A7] and elucidation of the mechanism first began in 1960s by Devries and Wolschang [A6]. Since then the structure function relationship of AF(G)Ps was explored in order to further understand these proteins and their mechanism(s). Thus far no one has been able to provide a theorem comprehensive enough to account for every experimental observation without some serious unanswered questions undermining its explanation. Rather, a few reasonable observations or theories have been put forward, all of which have some gaps that need closing. The predominantly accepted theory is the step pinning mechanism described earlier in the TH section, however some clarity is still needed to understand which interactions mediate the ice binding mechanism. [AX13] In this section we discuss what has been explored, and why. Please note that this section predominantly focuses on AFGP because they closely resemble polymers, and their structural differences to AFP suggest that their mechanism and interactions may also differ greatly.

#### Predominant mechanism

The adsorption inhibition mechanism was first mentioned by Raymond *et al* in 1977 and coined the name Step Pinning [A31]. Soon after, Knight *et al* extended the mechanism to give the mattress model, a competing theory with similar criticisms (to be mentioned later) [AX22][AX21]. The main difference between the two models is that the former assumes growth occurs in a stepwise fashion and suggests that AF(G)Ps block growth at steps, while the latter proposes that inhibition occurs perpendicular to the surface [(Fig. 1.5)] [A26]. In both models, the mechanism or interactions by which the molecules arrive at the ice interface is not declared in the theory. Rather they describe that the adsorption of AF(G)Ps unto the interface results in an energetically unfavorable, curved surface between adsorbed AF(G)Ps where waters molecules attempt to incorporate into the ice lattice (See TH section).

These water molecules cannot overgrow the AFPs whilst within the TH region [A24], so the ice crystal neither grows nor melts. However at temperatures lower than the TH the ice crystal will grow uncontrollably. This theory explains why it is possible for the freezing point depression to occur, and inherently explains why TH is always accompanied by DIS, while the reverse does not necessarily mean the presence of TH [AX1][AX2][EX3]

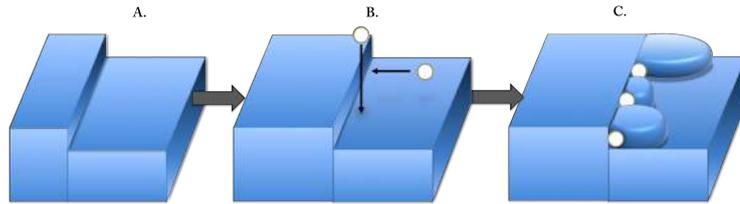


Figure 1.5: Splat assay set up. Key: A. and B. are at 273 K while C. occurs at  $T < 273$  K .

To investigate the mechanism of AF(G)P, ice hemisphere etching (IHE) studies are commonly employed. This technique is the process of freezing the dilute AF(G)P solutions just below their TH region so as to incorporate the molecules completely into the crystal. Normally the crystal is grown in such a way that all the interfacial orientations are equally exposed to the solution. On sublimation, the surfaces of ice crystals reveals faint patches (or etches) where AF(G)Ps are, or were once pinned unto specific planes of ice. [A1][A2][AX10?] The use of IHE demonstrated that the AF(G)Ps become irreversibly pinned. These experiments also reveal how single ice crystal growth habit changed around different AF(G)Ps because they had affinities for different planes of ice, normally the prism or secondary prism planes (a-axis), while growth along the basal plane (C-axis) remained uninhibited. If the temperature is dropped low enough, the basal plane will continue to advance gradually forming a tiered structure (Fig. 1.6). As a result the final crystal morphology are often “needle-like” crystals or hexagonal at low supercoolings rather than circular disks, as they would otherwise be. In exceptional cases, hyperactive AF(G)Ps are able to interact closely with the basal planes as well — however, these traits are often reserved for AF(G)Ps from insects which endure harsher climates than plants or notothonoid fish [AX19]. This multi-plane behavior is presumably responsible for their larger TH levels of up to 6°C and tendency to form hexagonal plates rather than needles [AX18][A32]. Both models propose that AF(G)Ps become trapped in the ice during freezing, and are therefore adsorbed irreversibly to the growing ice crystal. This is evidenced in experiments and so these models have become widely

accepted mechanisms. There is also recent evidence to suggest that some AFPs permit superheating ( $TH_m$ ) of ice crystals, which is in support of the adsorption inhibition mechanisms. [AX23][AX24][AX25]. Celik *et al* investigated superheating of ice crystals in the presence of AFPs and the correlation between TH and  $TH_m$  at a melting rate of  $0.02^\circ\text{C min}^{-1}$ . Under these conditions  $TH_m$  is considerably higher for hyperactive AFPs, just like with TH, and approximately a tenth of TH values.[AX23][AX24] Hall and Lips [A26] suggest that for the adsorption to be irreversible (and consequently accompanied with a large free energy), more AF(G)Ps must be adsorbed unto the ice at both small and large concentrations, rather than the concentration dependent profiles we actually see. The authors state that this process only appears to be irreversible they propose that the AF(G)Ps are instead slowly and temporarily blocking the addition of water molecules [A25][A26]. Their proposal explains why  $TH_m$  is a fraction of TH values and allows for a smaller free energy, however it does not explain why significant crystal growth is not observed prior to inhibition. To resolve debates like these, it is important to have a closer look at the types of interactions between ice and the antifreeze proteins or polymers.

**Importance of complementation** Both direct and indirect experimental methods have confirmed that AF(G)P binds to the surface of the growing ice crystal. [A15] [A16] [A17] [A18] [A19] [A20] The earliest proposals for the mechanism uses evidence that AFGPs compliment the ice lattice structures in certain planes, and so suggests that they are able to bind directly to the ice interface (or regions close to it).[A23][AX1][AX13][AX14] AFGPs are decorated with sugars, which carry restricted hydroxyl groups capable of hydrogen bonding with dangling atoms on the exposed ice interface. [A24] Upon testing this theory studies found that substitution of these hydroxyl groups could be tolerated but often resulted in reduced TH and DIS activity. [AX9][AX12][AX11]. However removal of 60% or more of the galactosyl residues resulted in complete loss of antifreeze activity [A15][A22]. Early CD studies revealed that AFGPs are left-handed helices, with hydrophilic groups bristling on one side of the helix while the hydrophobic residues line the opposite side. [A15][A21][AX9]. These studies also revealed that the steric N-acetyl group helped to maintain protein neutrality and to conserve the secondary structure necessary for both DIS and TH to function. [AX9][AX11][AX12]. Overall these provide a combined effort to keep the hydroxyl groups in place for efficient interactions with ice. NMR techniques were used to reveal that low molecular weight AFGPs are rod shaped at low temperatures and become increasingly flexible at higher temperatures and molecular weight, similar to polymers. [A15] This is an interesting contrast to

most AFPs that are often described as flat and /or rigid structures, and furthermore suggests that these two types of proteins (and even mimics) may have two completely different mechanisms. NMR spectroscopy was also used to probe the spacing between the hydroxyl groups and highlight that every tripeptide unit was separated by 0.931 nm [AX9][AX10], which corresponds to distances found between 2 units on the prism plane (0.904 nm ) [AX9][AX10] – coincidentally the same planes that growth would normally become inhibited. Similarly, AFPs have complimentary hydrophilic regions with each polar residue separated by identical distances (e.g. 0.45 nm for winter flounder) [A24]. These spaces vary from protein to protein, yielding different patterns that are attributed to inhibition of specific planes in different ways. As a result, they exhibit varying DIS behaviors and form spicular or hexagonal shaped crystals that are non-identical (Fig. 1.6) – even for proteins from the same AFP class. [A24] [A37] A combination of these attributes explained why, under laboratory condition, TH values as high as 1°C could be achieved by AF(G)Ps at concentrations as low as 0.1 mg ml<sup>-1</sup> [AX18]. Moreover, their ability to change the surface characteristics of a growing ice crystal is indicative of a high level of specificity [A14], especially as ~20.2 mg ml<sup>-1</sup> of NaCl salts would be required in order to achieve the same degree of undercooling [A36]. Similar observations were also found in studies of fish serum, which differentiated the contributions of salt and AFGPs to the overall antifreeze activity, and concluded that ~50-60% was a result of just 10–15 mg ml<sup>-1</sup> AFGPs. [A8][A9][A10][A11][A12] [A13][A24]

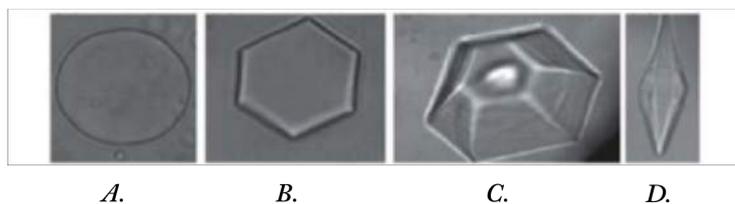


Figure 1.6: Different effects of dynamic ice shaping. Image taken from Ref. [A25]. [A25]. Key: A. no additives B. and C. show dilute solutions of AFP and D. shows a hexagonal bipyramid at high concentrations.

**Importance of Hydrophobicity** Due to a combination of the aforementioned attributes, AF(G)Ps were assumed to interact with ice directly via complimentary interactions such as hydrogen bonding. However there are disputes as to whether hydrogen bonding alone sufficiently explains the irreversibility seen by AFGP ice inhibition — particularly as the ice/water interface is not a sharp transition but a gradual one with the QLL at the center. Knight *et al* attempt to address this prob-

lem by suggesting the hydroxyl groups of the glycoproteins become incorporated into the growing ice lattice as described in [Fig x] and so the AFP/QLL or AFP/ice contact line effectively changes, resulting in more hydrogen bonding between the two structures. [A23] In spite of these suggestions, substantial considerations have recently been given to the involvement of hydrophobic interactions. [A25][A27][A28]. As mentioned earlier, most AF(G)Ps display some level of amphiphilicity; in fact the presence of hydrophilic and large hydrophobic regions (~60% alanine in some cases) suggest that the hydrophobicity also play a strong role in the antifreeze mechanism. It could not be coincident that most known antifreeze proteins have evolved independently with amphiphilic properties. [AY6]. To give an example, Devries and Lin found that the Type 1 winter flounder AFP has 0.45 nm separation between Threonine and Aspartate residues, with clusters of up to 7 hydrophobic residues separating them. [A24] Ultimately these indicate a use for hydrophobic regions as well, mainly pointing at a role that maintains fixed spacing between polar groups.

The recent ability to resolve the flat ice binding faces of Type 1 AFPs also revealed that they largely consist of hydrophobic groups, which creates an additional purpose to the seemingly resolved role of apolar groups. [Tis Ref] Substitution studies of hydrophilic amino acids with hydrophobic residues in Type 1 AFPs demonstrate that hydrogen bonding is not crucial for antifreeze activity. Replacement of Threonine with Valine (hydrophobic) retains some TH and DIS activity. [A33][A34] To give an example, the replacement of 4 out of 6 Thr residues\* (approx. 67% Thr) with Val at 10 mg ml<sup>-1</sup> retains ~50% TH activity. [A33]. Similar results were also found in another study where 2 of 4 Thr were substituted with Val and only resulted in a 10–20% loss in activity. [A34] However substitution with Serine (hydrophilic) results in little or no recognizable antifreeze activity, which is surprising. In the referenced studies [A33] [A34] no mention of IRI was made, but these results indicate that hydrophobic groups – specifically  $\gamma$ -methyl groups – are necessary or at least equally responsible for the adsorption of Type 1 AFP unto ice. It is possible that the van der Waals contributions from the non-polar groups and their hydrophobic interactions forces surrounding waters into a packing arrangement (presumably “ice-like”) so as to minimize interactions with the hydrophobic face and maximize the hydrogen bonding network between themselves [A23].

Alternatively, the mechanism may be driven by an entropic gain as a result of the exclusion of water molecules from the ice surface and also from the hydrophobic surface (desolvation). Additionally, upon binding to ice the hydrophobic solvent accessible area (PSASA) would become masked [A34]. In this instance the hydrophilic groups support this function by increasing or securing the comple-

mentarity of the AFP/Ice fit, ensuring buried surface area is exposed for maximum entropy gain. [A34] These hydrophobic contributions to the mechanism might also explain why it is possible some AFPs like TisAFP6 [Tis ref], with flat, but irregular beta structures can still have antifreeze activity (TH). Computer simulations have contributed greatly to these types of studies and, to investigate, the roles of both types of interactions are discussed further in the computational section.

**[In conclusion many questions arise, which face binds to ice? Is the mechanism irreversible? Does the polymer order water around it or disorder it or none of the above? What we can draw from all of these are that the mechanism of binding of AFGPs is still not well understood. ]**

### 1.3.2 Small molecules e.g. Sugars: C linked and O linked

#### Advantage of Synthetic mimics

Designing synthetic AF(G)P mimics with tailored antifreeze properties is an attractive prospect because it can be produced at batch levels required to meet the high demands for improved cold storage in medicine and commerce. [EX3] This route provides a better alternative to the exhaustive process of isolation and purification of AF(G)Ps. [EX3] A synthetic approach also means that a diverse range of structures can be built so they are ideal for systematic structure-function studies, in order to create an ideal antifreeze that might not be accessible in nature. Two other reasons prevent their use as cryoprotectants and cold storage of biological tissues:

1. TH values associated AF(G)Ps are hardly large enough for the low temperatures used ( $>93.15\text{ K}$ ). [EX13]
2. As a result the DIS abilities or explosive crystal growth at temperatures lower than TH present a mechanical hazard to the cells. [EX18][EX13].

#### C- and N- linked analogues

The production of AFGP in batch is limited by the orientation of the anomeric carbon-oxygen bond (O-link). [EX3][EX9][EX18] So Ben *et al* explored the inhibition of ice growth using synthetic C-linked and N-linked analogues [EX3]. These small mimics (N=4, 6 and 9) have little to no TH ( $\sim 0.06^\circ\text{C}$ ), and can form hexagonally shaped crystals, but still exhibit some IRI properties comparable to native AFGP-8. [EX13] These indicate that the mimics do not really interact with the ice lattice, but when they do it is on a preferential plane. [EX13] To some degree the latter is consistent with findings for native AF(G)Ps, and smaller chain lengths

(N=1 and 3) had no IRI activity at all. The authors systematically investigated the effects of backbone composition, lengths of side groups and the type of carbohydrate used. The results demonstrated that increasing distances between the side groups and the backbone beyond 2 or 3 carbons resulted in a complete loss of IRI activity. The same result was achieved by substituting galactose moieties with others sugars known to have cryoprotective properties (i.e. glucose, mannose and talose) [EX10][EX13]. This lays out the extent to which IRI relies on structure and orientation, particularly for the C2 hydroxyl group – where a change from an axial position to equatorial meant a 50% reduction in IRI activity and less seriously on the equatorial C4 hydroxyl group. [EX13] Based on these findings, they linked IRI performance to the hydration properties of sugars attached to these mimics. [EX4][EX3]. Their most potent analogues bore sugars that can produce larger hydration spheres, and so preoccupy surrounding water molecules from advancing the QLL or ice interface. Instead they help to distort the amorphous phase. [EX13][AX9] The lack of TH and DIS in C-linked and N-linked analogues translates to little or no cytotoxicity during cold storage of mammalian cells (i.e. human embryonic liver cells).[EX5][EX13] Studies of AF(G)P mimics like these have shown that IRI can be dissociated from the other 2 antifreeze properties. It has also been demonstrated that although DIS can sometimes be dissociated from TH, it is not possible to have TH without DIS [AX1][AX2][EX3]. This makes the two properties undesirable and shifts considerable attention to maximizing the third antifreeze property, IRI. In nature, similar approaches have already been employed by freeze tolerant organisms [EX19] which also use IRI active sugars whilst, freeze avoidant species rely more on AF(G)Ps. Seeing as DIS is most tolerant of structural modifications, aside from the need for a well-defined secondary structure and neutral moiety, one approach could be to subdue DIS by designing synthetic analogues with a different type of secondary structures. [EX9][EX11][EX12][EX13]

### **Non-protein “antifreezes”**

Biological tissues using non-protein cryoprotective agents can be cryopreserved using vitrification, the rapid cooling of a liquid to a glassy state without crystallization [EX18]. In order to protect mammalian cells, some conventional formations include molecules such as carbohydrates [EX13], salts, glycerol [EX15], or dimethyl sulfoxide (DMSO)[EX13][EX16]. The latter 2 are most commonly used, and are normally applied at concentrations of 5–10% (v/v) DMSO or glycerol, and 95–90% (v/v) Fetal Bovine Solution. [EX17] The disadvantage of using these cryoprotective agents are that they are only favorable for a few types of cells (erythrocytes and spermatocytes)

and non-ideal because of high cytotoxicity especially in the case of DMSO and glycerols weak cryoprotective abilities.[EX17]

### 1.3.3 Polymer antifreezes

In the search for an ideal antifreeze macromolecule, researchers assessed the antifreeze properties of a library of synthetic polymers [TX2][TX11], first started by Knight *et al* in 1995 [TX11]. Those which have received most attention are poly(vinyl alcohol) (PVA)[HIRI (high IRI) LTH (low TH) LDIS (low DIS)][TX11], poly(hydroxyproline) (PHYP)[TX11][ABC1], Poly-L-Histidine [TX11], zirconium acetate (ZA)[SX1][SX2][EX13], silane coupling agents (SCA) [EX15] and polyampholytes [EX17]. These have been largely studied rather than using synthetic glycopeptide conjugates like C-linked analogues because they are less challenging to make and often produce greater yields.

#### **Poly(vinyl alcohol)**

Inada *et al* first observed atactic PVA as an antifreeze polymer in 1995 [TX11]. They found that polydispersed; low grade PVA318-2273 (86-98%) exhibited relatively strong IRI in comparison to other polymers. These were qualitative observations and no comments were made for TH and DIS. [TX10] Later studies conducted by the group confirmed these findings that PVA exhibited small amounts of TH ( $\sim 0.037$  K at  $50 \text{ mg ml}^{-1}$ ) [TX4][TX5][TX2] and DIS by interacting with non-basal planes of ice. [TX6][TX2] They also showed that [TX3] PVA of 80-98 kDa could be as effective at IRI as AFGP 8 or AFP Type 1 (3.3 K Da) at equimolar concentrations [TX2][TX3]. However in a separate study Ben *et al*'s first generation C-linked analogues were almost 80 times less active in comparison. [EX13??][TX2] These highlight the effectiveness of PVA and consequentially PVA serves as a good starting point to study IRI in order to develop an ideal IRI polymer. This is further compounded by earlier studies on a range of structurally diverse polymers (carboxylic acids, Amides, amines, and alcohols), which revealed that only polymers with hydroxyl groups could achieve significant IRI activity. [TX2] Even PVAs structural isomer, PEG showed negligible IRI, TH or DIS activity [TX2][BX2][6], which makes sense because it has a functional group cannot meet the aforementioned requirements; the oxygen group neither protrude out of the polymer backbone, nor can it act as both hydrogen bond donor and acceptor [6] .

In 2013 Gibson *et al* carried out an extensive study on narrowly polydispersed PVA. [TX1] They argue that the purity and polydispersity of the polymer will affect

the integrity of these studies, which Inada previously demonstrated in the case of poly-l-histidine [TX11]. They investigated the effects of increased acetylation, PVP block copolymerization and the influence of 7 chain lengths (between 10-351 units) on the IRI activity. The concentrations tested ranged from 0.005-1 mg ml<sup>-1</sup>, and revealed that PVA<sub>N=351</sub> (34.5 kDa) could be IRI active from as low as 0.05 mg ml<sup>-1</sup> after just 30 minutes. At 1 mg ml<sup>-1</sup> PVA<sub>10</sub> produced crystals with 80% MLGS, while all other chain lengths (N= 19,30,56, 154, 246 and 351) produced 15-25% MLGS and had similar profiles for the concentration dependence of IRI activity. Somewhere between 10-19 repeat units; there exists a minimum chain length where IRI activity is “switched on”, but at longer chain lengths the IRI activity increased steadily until a critical point [TX1]. Larger polymers, (N= 246 and 351, 22.7-34.5 kDa) reached this point at much lower concentrations of 0.1 mg ml<sup>-1</sup>, in comparison to ~0.3-0.8 mg ml<sup>-1</sup> for much smaller polymers (N= 19–154, 1.7-14.6 kDa). [TX1] It therefore appears that, like AF(G)Ps, larger polymers may be more effective and so can be used at lower concentrations.

It is important to acknowledge three more things. Firstly, that these studies primarily deal with atactic PVA which means that their hydroxyl groups are not necessarily spaced evenly. It is easy to believe that combined with the positive size dependence [TX1][TX2] of IRI activity, this indicates that higher densities of hydroxyl groups may in fact influence a corresponding increase in IRI activity, rather than just a complementary fit to ice from OH-OH distances on PVA alone [EX3]. This notion is further corroborated by a decline in IRI activity by introducing acetyl spacer side groups into the PVA backbone [(Fig x)]. For instance a change from 10 to 30 mol% acetylation resulted in complete loss of IRI for a 2 mg ml<sup>-1</sup> solution. However this reasoning was questioned by findings that other glycopolymers with higher OH densities per unit mass than compared to PVA was found to have lower IRI [TX7][BX4]. It brings to light the influence and need for quantitative assessments of other factors such as the relative degrees of backbone flexibility, hydrophobicity, water hydration and water ordering or disordering. For instance, it is possible that as a result of greater flexibility, these PVA polymers can adopt plethora conformations and maximize their interactions with the ice. In the light of these considerations, it suddenly becomes unclear if the introduction of acetyl group adversely affects IRI because of a change in flexibility or the discontinuity within either the hydrophobic face or hydrophilic groups. [TX1]

Secondly, the splat assay shows both nucleation and the growth of ice crystals. It is difficult to distinguish which of the two PVA and other IRI inhibitors affect, however Wang *et al* demonstrated that both stages are affected. [TX12]

Further studies have also confirmed their finding and show that nucleators like bacterial ice nucleation proteins [TX13] or silver iodide [TX14] have had their activities suppressed by the presence of PVA.

And finally, another advantage of PVA is that it is non-toxic [TX16] and already in use for a range of pharmaceutical applications from eye drops to tablet coatings and even thickeners. Deller *et al* recently demonstrated that for both sheep and human erythrocytes, cell viability was enhanced by 0.1 wt% PVA (N=200, 9 kDa), which is almost 100 times less than typical fractions used for DMSO. This study tested the compatibility of PVA with human blood and the findings means that more blood cells will be rapidly available for transfusions after thawing, seeing as there is not an urgent need for PVA to be removed post surgery, unlike with DMSO. [TX15]

### **Poly(hydroxyproline)**

Poly-L-Hydroxyproline (PHYP) has also been shown to exhibit high degrees of IRI activity and nearly non-existent TH in pure water at moderate concentrations [TX11] or low degrees of polymerization. [ABC1]. In 2009, Gibson *et al* described the first quantitative studies of the IRI activities of HYP and they found that its IRI activity is concentration dependent. For N= 44 (5.7K Da) the IRI activities were comparable to PVA N= 80 at  $3.5 \text{ mmol L}^{-1}$ , producing a mean largest grain size of  $175 \mu\text{m}$ , which is approximately 34% smaller than the MLGS of the PBS negative control. [BBB1]. Interestingly, although the IRI activity has little apparent dependence on the secondary structure of the proteins or polymers, PHYP also has PPII helical secondary structure, similar to AFGPs [BBB2][BBB3]. And even more interestingly, compared to the structure of PVA polymer, albeit atactic in structure, PHYP is not considered facially amphiphilic and highlights the independence of IRI on this feature.

### **Zirconium (IV) acetate**

Normally ice in a colloid suspension grows into lamellar structures with dendritic surfaces however in the presence of 13-100g/L Zirconium acetate (ZrA), homogeneous sized hexagonal pores form along the direction of growth within the crystals [SX2][SX3]. At smaller cooling rates the pores are bigger and become faceted, which is believed to be a direct consequence of ZrA interacting with the ice crystal interface. For instance pores of  $4.5 \mu\text{m}$  are produced at a cooling rate of  $20^\circ\text{C min}^{-1}$  and increases to  $100 \mu\text{m}$  at  $0.5^\circ\text{C min}^{-1}$ . [SX2] ZrA or  $(\text{Zr}(\text{OH})_3\text{A}_2^-)$  and has re-

cently been revealed to have IRI activity [SX1], because it slows the growth of ice buffer solution (pH= 3.5–4.5) [SX1][SX2][EX13] while the true amount of thermal hysteresis activities still remain largely under debate [SX1][SX2][SX3]. Similar to PVA and AF(G)Ps its thought that the arrangement of the chemical groups like OH matches the periodicity of atoms on the ice surface. Due to its amphiphilic nature, it could also be stabilised by hydrogen bonding and so ZrA may also reside at the interface of the ice/water boundary or distort the diffusion rates of nearby migrating waters.

## 1.4 Key simulation studies

### 1.4.1 Simulations of small antifreeze molecules

Simulation of small antifreeze molecules in the presence of ice is rare. Rather they tend to be in systems of supercooled water, and focus on salts like NaCl or small organic molecules that are typically polyols e.g. carbohydrates or alcohols. Previous simulations of salt, sugars and AFP analogues are discussed here in this section, particularly findings which are relevant to later discussions.

#### Salts

MD simulation of salts in the presence of ice appear to be the most popular of the small “antifreeze” molecules, largely because of its application in other areas like the process of brine rejection. It is common knowledge that salts do not incorporate easily into the ice lattice, and so are rejected out of the growing ice region. In the presence of salts, the freezing point is depressed however the molecular mechanism has not yet been elucidated. Recently, Vrbka and Jungerwith [SALT1] studied the brine rejection mechanism in the presence of a growing cubic ice crystal. They found that during ion density fluctuations, a new layer of ice opportunistically forms in regions where the NaCl ions had temporarily vacated. Over a trajectory period of 60–630 ns , they studied freezing of dilute aqueous solutions (0.15 M and 0.3 M) at  $T_m+15$  and  $T_m-15$ . [SALT1] Consistent with expectations, a local concentration gradient was produced as the salts were rejected into an unfrozen region, forming a glassy area that would become even less likely to freeze. Complete freezing of a pure water system was obtained after 250 ns , while the time taken in the presence of salts typically decreased. In this study, a 0.15 M solution resulted in a 20% increase in the time taken to freeze (300 ns ), whereas doubling the concentration led to a remarkable 100% increase (500 ns ). [SALT1] These observations of the freezing

point depression and the glassy QLL are not uncommon, and have been reported elsewhere for different planes of ice, water models and a range of salt concentrations. [SALT6][SALT2]

For example, in 2006 Carignano *et al* [SALT6] studied the  $1_h$  ice growth of both the prism and the basal plane in contact with a saline solution with similar concentrations (0.2 M) and undercoolings. Their simulation was conducted at approximately  $T_m - 18$  K and used a 6 point water model for increased accuracy of the phase diagram of water. Peculiarly, the growth kinetics for the prism plane changed in the presence of salts and less so for the basal plane. The basal plane experienced approximately 30% reduction in growth rate and the prism plane was split into three stages. In the first stage, the initial growth rate was the same and faster than the basal plane however it came to a metastable state for about 2.5 ns . This period translates to approximately 30% of the simulation time needed for complete freezing of the system in the presence of salts. The author explains that these differences in kinetics are a direct result of the surface roughness on the two planes. Modelling work on the basal plane revealed a layer-by-layer growth mechanism, which produces a smoother surface, while the prismatic plane is much more random, disorganized and rough. [SALT8][SALT3] To place the observation into context, no isolated ions were found at the end of simulations and the authors concluded that complete freezing would not be possible unless the ions were fully paired or clustered together. The resulting decrease in entropy from pairing would encourage crystallization of the liquid phase. Ordinarily, the prism planes roughened surface offers an abundance of binding sites for faster ice growth however in this instance, it shelters ions from their pairs delaying the freezing process. It is interesting to note that the growth rate in the third stage is faster than the first, however the authors do not comment on this. It is plausible to suggest that in stage three, the collision with salt ions prior to pairing may roughen the surface even further and therefore speed up the growth rate at subfreezing temperatures. [SALT6] It also explains why concentrated solutions have been shown to increase the rate of melting at higher temperatures [SALT9] (particularly as the larger ions, which can do more damage to the lattice or roughening, has repeatedly been proven to approach much closer to the ice/water interface. [SALT9][SALT10][SALT11])

These types of simulations produce melting or freezing rates comparable to experiment [SALT6][SALT9], and so describe a reliable protocol for identifying the freezing point depression that can be applied to larger antifreeze molecules. They have also revealed important information for just plain simple ice/water simulations. For instance, there has been a long standing dispute as to whether super-cooled wa-

ter spontaneously crystallizes to cubic ice and quickly transitions into hexagonal ice. Recent studies have exposed that the introduction of stacking faults in both hexagonal ( $I_h$ ) and cubic ice ( $I_c$ ) polymorphs in contact with brine solution can produce layers of  $I_c$  or  $I_h$  respectively. [SALT6][SALT2][SALT3][SALT1]. In multiple studies Carignano *et al* [SALT6][SALT2] also noticed the presence of stacking faults as the ice crystal grew, but these were only observed for the basal plane [SALT2][SALT6] and not the prism plane [SALT6] The authors [SALT2] demonstrated that the differences between these two polymorphs could easily be identified in a snapshot by tracing a line through the center lattice honeycomb structure, layer by layer (ie from left to right).  $I_h$  is characterized by a zig-zag pattern and  $I_c$  is marked by a straight line. In the referenced simulation, they noted that this transition was a rare event, which occurred approximately every 50ns. Most importantly they identified that the cause of this transition was the constant freezing and re-melting of the ice layer because it produced the opportunity for the rearrangement of the lattice waters. These observations are important and cannot be readily dismissed as artifacts of the simulation system, mainly because so many different set-ups were used across these studies and still independently reach similar concentrations.

## Sugars

In 1989, Brady published [SUG1] the first molecular dynamics simulation of  $\alpha$ -d-glucose and shortly after forcefields/energy potentials was revised to improve calculations for more accurate conformations and descriptions of glycosidic links on proteins. [SUG5] These were mainly intended for protein-sugar or membrane-sugar complexes however some simulations of carbohydrates were soon used to study the effects on hydrating water because of the correlation with their cryoprotective capabilities. [SUG2] For instance, trehalose and sucrose are known cryoprotectants typically found in animals and plants respectively. [SUG10] So they are expected to have a lot in common, especially as both disaccharides are composed of fructose and glucose in a glycosidic link. [SUG10] Once such interesting study was produced ten years ago by Bordat *et al* . [SUG2] They studied three disaccharides – trehalose, sucrose and maltose – and the effects of varying concentrations (0%, 4%, 16%, 33%, 49%, 66%) or temperatures (273–373 K , in 20 K in intervals) on surrounding water. In a system consisting of 512 SPC/E water molecules, they investigate the size of the hydrogen bonded clusters and the tetrahedrality of water using an orientational order parameter. In accordance with experimental work [SUG2], and later simulations, [SUG6] they found that below a threshold concentration of 40% and at cold temperatures (273 K ), the disaccharide could not be distinguished in their disrupting

effect of the hydrogen bonding network (HBN).[SUG2] Above this threshold, which is close to the experimental value of 50%, trehalose was most disruptive, followed by sucrose and then maltose as expected [SUG2][SUG7][SUG6]. This expectation is largely based on NMR studies by Ablett *et al* [SUG3], and others [SUG7][SUG3] that have identified a correlation between a higher glass transition temperature with reducing ice recrystallization rates. Using the orientational order parameter Bordat *et al* [SUG2] demonstrated that trehalose has a greater hydration sphere than the other two disaccharides. As a result, it can disrupt more and so produces many more neighboring, small clusters of HBN than other disaccharides. [SUG2] This too, has been demonstrated before. [SUG8] Interestingly, the orientational order parameter did not reveal any shift among the three sugars profiles suggesting that in they adopt comparable conformations in solution. Instead, increasing the concentrations of trehalose produced a profile similar to that reported by Debenedetti as the temperature is increased. [SUG6][SUG15] To date, similar comparative studies of mono- and polysaccharides in unfrozen water continue to stress the protective capabilities of the solution directly correlates with concentration, and attribute their properties – including slowed water diffusion rates – to their water/sugar hydrogen bonding behaviors. [SUG4][SUG6] In terms of atomistic mechanism, the most widely accepted theory is the disruption of water interactions meaning that the sugar cannot easily be incorporated into the lattice. However prior to this Crowe *et al* first proposed an alternative called the water replacement theory. [SUG9] They suggest that sugars replaced hydrating waters around biomolecules thereby protecting them from freezing over and denaturing [SUG9][SUG2]. Sum *et al* [SUG11] considered this interaction possible with the lipid bilayer to protect it from the extracellular ice and suggested that trehalose could hydrogen bond to the phospholipid bilayer. They constructed a simulation that revealed miniscule changes when trehalose interacted with the phospholipid heads [SUG11] and ultimately expose Crowe’s assumption that biomolecules are hydrogen bonded to the sugars during the freezing process. [SUG2] Recently, similar attempts have been made to reconcile carbohydrate’s to protect proteins in solution from thermal denaturing, however focus on questionably high “dessicating” temperatures, presumably past the boiling points of the water model used and the  $T_g$  of the carbohydrates. [SUG12][SUG13].

#### 1.4.2 Simulations of antifreeze proteins or glycoproteins

Molecular dynamics simulations provide an atomistic view of the dynamics of a system and can therefore be used to look closely into the mechanism of antifreeze activity. This method is particularly useful as direct measurements of the Ice-AFP

interactions, which are not readily available to current solution, based experimental techniques. [CCC9] Molecular modelling of AFPs and free energy calculations lend some support to the range of mechanistic viewpoints which emphasize the importance of hydrophilic lattice matching groups [CCC6][CCC7], and more recently, hydrophobic groups [CCC4][CCC13] which have been discussed earlier. The third hypothesis, which was only touched on, is that the antifreeze molecules accumulates at the ice water interface, remodel the ice surface by influencing or assisting the development of the binding site. [CCC15][AX16][AX15][AX13] Each of these theories, births questions like: Is ice more hydrophobic than water? How does the protein differentiate between ice, QLL and water all of which are so chemically similar? [AX13] The most recent simulations aim to address these questions and select or rule out the aforementioned theories. Investigators typically draw their conclusions by studying the properties of the system such as hydrogen bonding, water ordering or disordering, hydration numbers or protein solvation, facial alignment toward ice or “recognition” of ice, distance from the QLL, hydroxyl group spacing, solvent accessible surface areas or contact areas, free energy calculations and so on. The work of McDonald, Jorgensen, Haymet, Madura, Wierzbicki and Granham, in particular, present landmark contributions in understanding the detailed mechanism using MD simulations.

### The first simulations

Two independent groups conducted and reported the first simulations studies of antifreeze proteins and glycoproteins in solution: McDonald *et al* [CCC7] in 1993 and Jorgensen *et al* [CCC6]. These largely focused on conformations and the spacing between the hydrophilic residues in room temperature (300 K ) [CCC7] and super-cooled water (273.15 K ) [CCC6] using 3-point water models and the CHARMM [CCC7] or AMBER [CCC6] forcefield. Like the experiments, early solution studies [CCC7][CCC6] pointed towards the key role of hydrophilic groups because they typically found that Thr residues had spacing, which matched those on specific (201) planes of ice. Take the study by Jorgensen *et al* [CCC6] for instance, following a short 0.12 ns NPT simulation, they found that that HPLC-6 winter flounder protein (the most studied AFP) had equally spaced Thr hydroxyls (close to 1.67 nm ) on one side of the helix, which matched the [0112] direction of ice. These findings are closely in line with etching results produced by Knight *et al* [CCC10] between 263 K –258 K , and other molecular modelling studies [CCC16] [CCC17][CCC18]. By taking a snapshot and calculating the root mean square deviation (RMSD) from the proteins initial structure, the authors were able to confirm that HPLC-6 was

linear for the most part, and maintained its helical structure at 273.15 K [CCC6]. Using the same, all-atom RMSD calculation at 300 K and similar number of water molecules- albeit a different forcefield and system set-up – McDonald *et al* [CCC7] found that the same HPLC-6 protein was stable but actually remained bent at the center, particularly at lower temperatures; 30° bent at 300 K and 60° at 237 K . Both referenced studies displayed average RMSDs between 0.225–0.325 nm after just 0.07–0.16 ns , however neither of the two studies confirmed these claims to secondary structure using Ramachandran plots, which are typically used in MD simulations of proteins. In order to capture a snapshot of the mechanism, [X] *et al* were the first to conduct simulations of AF(G)Ps in contact with ice in vacuo. [ref] In an effort to minimize the complexity of the simulation set up, these were traditionally simulated without the presence of liquid water and therefore the absence of the ideal QLL. [CCC12][ref] These systems were initially thought to be a reasonable approximation however a vacuum is actually an ill representation of water [CCC5], and further still increasing experimental evidence arose that the AFGPs might in fact work at the QLL interface. [A33] [Haymet ref/fishy proteins ref] Around the same time Karim and Haymet became the first to introduce the simulation set up which described the QLL – a 1–1.5 nm wide transition region between the two phases of water – using the TIP4P water model. [CCC1][CCC2] Several other ice/water simulations have been performed since then to confirm [ice ref] and these regions are typically identified using an order parameter such as self-diffusion, average density, translation order, all of which provides an inflection point whilst crossing between the two phases. [CCC1][CCC2][CCC12][ice ref] Seeing as accurate descriptions of the protein interactions would depend on the realistic description of the ice-water interface, Karim and Haymet's protocol was quickly adopted by McDonald *et al* [CCC3], and soon others, [CCC3][CCC4][CCC5] so that AFGP/Water/Ice interfaces became common practice for simulation of AFGPs; however at the time the ice lattice was typically fixed in position [CCC3][CCC4]. Later this was naturally followed up with the practice of unconstrained ice simulations in the presence of water, also first achieved by Dalal *et al* [cccdalal] / Karim and Haymet, in order to relieve simulations off of structuring artifacts, which propagate into the QLL, and could potentially result in altered freezing times [CCC2].

### **The introduction of the ideal QLL**

With the introduction of idealized interface, and better QLL description, investigators began to report that the lattice matching theory was not the predominant mechanism and moved away from this stand. In 1995 McDonald *et al* simulated the

proteins at the idealized interface for just 0.10 ns at 203 K using an NVT ensemble. [CCC3] They found a good fit for their AFP at the interface, however the protein was oriented in such a way that the Thr faced away from the interfacial region and so proposed that direct binding unto the ice via lattice matching may not be the only plausible mechanism to successfully bind unto the ice interface. They also noted that their starting structure formed a twist in in the protein that later “promoted” a bent helix for the remainder of the simulation. [CCC3][CCC12] In a similar study 5 years later, and using a longer trajectory of 0.40 ns, Madura *et al* argued that given sufficient time, the Thr residues would in fact reorient or remain oriented towards ice at 180 K. Its important to note that in these simulations the AFPs were already placed in hydrogen bonds with the ice in the starting position, whilst residing in the first QLL layer. [CCC12] They did not notice any bending of the helix however they found that one of the proteins “rolled” unto its side, moving the non-polar leucine residues from the QLL layer parallel to the ice surface, and allowing them to also interact with the ice. Similar to McDonald, they did find that the ends of the protein began to unwind at the N terminus, producing non-helical regions when close to the ice, while at the QLL, promoting helicity. [CCC12] In agreement with Houston *et al* [CCC15] they postulate that the third mechanism could be most appropriate; they imply that AFPs dont bind to any plane in particular but reorient in order help to generate the binding site.[CCC12]

### Recent breakthroughs

Also in agreement, Garnham *et al* [AX15] and Sharp [AX13] independently highlight that both types of interactions (hydrophobic and hydrophilic), most likely work in conjunction to mediate the ice binding mechanism. This stand negates the concept of a dominant interaction and may even place the 30-year debate at rest. In the past 2 years there has been a monumental discoveries and simulations about the antifreeze mechanism, which support this finding. Granham *et al* made one such discovery in reporting the first X-ray structure to capture ice-like water molecules hydrating a bacterial AFP known as MpAFP [AX15]. The reason why this is so rare and important is that crystal structures normally capture multiple AFPs, which are close enough to each other that they distort the few surrounding crystal waters that separate them. However on this occasion the authors found a region of unaffected and solvent exposed ice binding site. [AX13][AX15]. From studying the hydration of these proteins, the authors postulate that the proteins residing in the QLL order water around itself into an ice-like configuration using hydrophobic groups, (rather than binding to a plane and releasing surrounding waters for an entropic gain). The

ice-like structures are anchored by hydrogen bonds to polar groups on the protein backbone or side chain (at the edge) and so the protein can approach the ice lattice with its own “ice” ligand or binding site, which they termed an “anchored clathrate” [AX13][AX15] This theory was first proposed by Nutt and Smith in 2008 [AX16] and Graham *et al* go on to describe that the most effective anchoring would be achieved by polar groups spaced by apolar groups. This is not necessarily the same way that the polymers function however; this theory explains why most reported ice binding faces are flat and hydrophobic. [AX16] It explains the unusual concentration dependence. It explains why the type 3 AFPs functions as antifreezes, although it is globular and not bristled with many OH groups. It also explains occasions where it is possible to observe a decrease in TH during mutation studies by changing hydrophobic residues to hydrophilic, as well as the reverse across different AFP classes (e.g. Ala  $\rightarrow$  Thr in Type 3 or Thr  $\rightarrow$  Ala in Type 1). [AX13] A delicate balance of the two interactions might therefore be in order and if this were the correct route, a new question rises as to what this balance may be. Graham *et al* expand further that with this as a predominant mechanism one would expect to see high residence times of a number of water molecules at the ice binding face (IBS), which they do accross different classes of AFP, [AX15]. Once a certain number of these (a quorum) are achieved the ligand formation begins. Notably the mobility of liquid water is hampered at the QLL [CCC12][CCC1][CCC2] where the protein resides, which makes the concept of quorums forming at the IBS even more feasible at low temperatures. In solution, these quorums are likely to be affected by higher temperatures as well as by mutation studies, so Garnham *et al* also recommend that relationship between TH and mutation studies be re-examined on the basis of IBS hydration, rather than hydrophobicity alone. [AX15] [Noticeably this does not appear to be the case for hyperactive AFPs [AX17].]

Rightly so, Nada and Furukawa raised the question of inherent bias in setting up the AFPs initial environment in simulations how do we decide where to place the AFP at the start of the simulation? [CCC20] Ideally, the AFP would be unconstrained and allowed to migrate from the liquid phase itself as it would in experiments, rather than placing the AFP directly unto the QLL in order to save simulation time. [CCC20][CCC19] Particularly as experiments have no way of guiding investigators to identify how far into the QLL that the proteins may reside. Considering that in experiments, the AFP solutions are frozen and the ice-water interface begins to grow, it is crucial to observe the AFPs in these growth environments as well, rather than at static temperatures which maintained stable ice crystals alone. A recent study by Wierzbicki *et al* also already addressed this criticism. [CCC9]

They used potential mean force (PMF) profiles to show the free energy penalty in placing the Type 1 AFP at a range of distances from the midpoint of the QLL. If at all, a face specific adsorption was required they proposed that the approach with a hydrophobic face (Thr-Ala-Ala) was most realistic as it interacted more favorably with the ice than the hydrophilic face (Thr-Ala-Asx) and so could draw closer. The authors found that the lowest energy penalty was produced whilst approx. 0.75 nm away from the center of the ice/water interface, which coincides with the start of the QLL environment. Using the hydrophobic face, they calculated a  $-3.5 \text{ kJ mol}^{-1}$  free energy difference compared to bulk water as reference. This is approximately  $0.5 \text{ kJ mol}^{-1}$  less than for the hydrophilic face at its optimum distance of 0.9 nm, which happens to fall outside the QLL region and inside bulk water. [CCC9] They also found two interesting details. Firstly, that the hydrophobic face in contact with the QLL encouraged an additional number of 13 close contacts (or hydrogen bonds) than compared to the bulk, while no real difference was observed for the hydrophilic face. Similarly, they found that the contact surface area of the protein increased with the hydrophobic face from 85.7 nm<sup>2</sup> to 89.2 nm<sup>2</sup> and again, no real difference was observed for the Thr-Ala-Asx face.[CCC9] These findings, which were reported in also 2007 support the proposal and observations made later by Garnham *et al* [AX15] in 2011. In the past, the same faces had been studied [CCC9] and conflicting findings were reported, however this only goes to show how insufficient sampling from short trajectories or poor system set ups (eg in vacuo environments) can affect the representation of the real life mechanism. Newer studies by Nada and Furukawa [CCC21][CCC22][CCC23], Todde *et al* [CCC24][CCC25] and Kuiper[CCC26] *et al* are discussed in later sections. These authors have all made important contributions by making improvements to the system set up, the water models they select to accurately represent the phase diagram of water, and they thoroughly characterize the AFP in the presence these ideal environments under much longer time scales in the presence of stable crystals and during crystal growth.

### 1.4.3 Polymer simulations

#### Proline based polymers

Simulations of hydroxyproline or proline are typically produced to investigate the influence of substitution on the protein side chains or backbones to the secondary structure of a protein (or water structure around a protein). Take collagen for instance, where hydroxylation of prolines occurs and stabilizes the triple helix. Park *et al* (2005)[POS15] explored the importance of ring puckering conformations on

this process using MD simulation. They found that previous attempts to include hydroxyproline in calculations had not given the correct puckering as known to experiments.[POS][POS] So they proposed a new set of parameters, which reproduced the correct puckering preference at 300 K using the AMBER forcefield. [POS15]. A recent paper by Yu *et al* [POS16] demonstrated that the rigid, and hydrophilic proline molecules induced slow diffusion dynamics in vicinal water molecules. They also experienced increased distortion of hydrogen bonding network and a loss in tetrahedral orientation. For instance, a solution consisting of 96.2% water and 20% w/w proline, produced a 15% drop in tetrahedral orientation compared to bulk water. The reason for this behavior is that 5-membered proline ring, which is similar to poly(hydroxyproline), passes on its rigidity to vicinal waters. Clusters of surrounding water molecules exhibited O-O-O hydrogen bonding angles, which are restricted to  $79^\circ$  and  $99^\circ$  rather than the ideal  $109^\circ$  for tetrahedral orientations. In this study they used the OPLSAA forcefield, the TIP4P water model at room temperature to investigate the properties of proline concentration between 0-76% w/w.

Another recent study by Nizaga *et al* [POS17] (2013) examined the backbone solvation of three homopolymers poly(proline), poly(hydroxyproline) and poly(hydroxyproline- $\beta$ -galactose), each 9 units long. As is common in these types of simulations, they used AMBER and TIP3P water model for 100ns. Alpha, beta, PPI and PPII helices are three known helical secondary structures which proteins can adopt, and these are characterized by the respective backbone dihedral conformations  $(\phi, \psi)$ ;  $(60^\circ, 50^\circ)$ ,  $(120^\circ, 120^\circ)$ ,  $(75^\circ, 160^\circ)$  and  $(75^\circ, 145^\circ)$ . They probed the differences in solvent-polymer backbone interactions between PPI and PPII helical conformations within the three polymers. In agreement with experiments, they found that both PHYP and PHYP-Gal polymers preferred a PPII helical conformation in water because it obtained greater number of hydrogen bond interaction with the water solvent than with PPI.

Interestingly neither PHYP nor PPRO exhibited any intramolecular interactions in the PPII conformation, and as a result these polymers obtained similar, high levels of interactions with the solvent. In contrast, glycosylation reduced the number of hydrogen bond interactions with the solvent. The authors also found that aliphatic solvents could completely changed polymers preference to the PPI helix and so they concluded that a small or large change in hydrophilicity could be induced by adding groups like sugars and by careful solvent selection respectively. The group proposes that the galactosyl residue may alter the solvent accessible surface area, in order to interact with the protein backbone. Of the three homopolymers, it was the only one to exhibit intramolecular interactions (between the galactosyl O6 hydroxyl group

and the protein C=O carbonyl group) and these lasted for approximately 5% of the simulation trajectory. This study goes to show that polymer or protein secondary structure can be governed by solution-polymer interactions (or its side chains), at times more so than the intramolecular interactions. And reiterates that the changes in conformation might also be important in elucidating mechanisms such as the ice binding of protein.

### **Poly(ethylene glycol) polymers**

A number of simulations of polyethylene (aka polyethylene oxide or PEO for short) in solution have been reported in the literature. Due to their amphiphilic nature, these polymers are soluble in water as well a vast range of organic solvents like methanol, chloroform and benzene. MD simulations by Hezaveh *et al* [POS18] have been used to understand that interactions of PEO<sub>2-43</sub> with these solvents at 298 K . [POS18] They modeled the polymer using united atom models and the GROMSO/OPLS-UA forcefield. The PEO O-C-C-O dihedrals adopted gauche conformations in water and trans conformations in apolar solvents, which is consistent with available NMR data. For the radius of gyration ( $R_g$ ) the power law dependence with chain lengths for PEO also showed good agreement with experiments. They reported expansion of polymers in polar solvents (methanol) and contraction in apolar solvents (chloroform, carbon tetrachloride and n-heptane), relative to bulk water and as indicated by the radius of gyration calculations.

Lee *et al* also calculated the radius of gyration for PEO<sub>9-27</sub> oligomers using a revised C35r CHARMM forcefield. They obtained an exponent of  $0.515 \pm 0.023$  which relates the radius of gyration to molecular weight. [POS19] Although the value is also in agreement with experiments it is much smaller than the expected 0.588, and deviates further than the valued achieved whilst using the OPLS forcefield (0.59)[POS18]. Oeleimer *et al* [POS20] conducted a similar study using PEO<sub>6-81</sub> and the AMBER forcefield. They studied single polymer chains in solution for 10-30 ns and obtained. In agreement with these previous simulations, PEO was described as a random coil in solution with a preference for the gauche conformations. Unsuprisingly this forcefield also yielded a Flory exponent between 0.5-0.535, which is acceptable but still lower than the experimental value. [POS20]

## Polyvinyl alcohol

Simulations of PVA have been used to understand the miscibility of binary blends or polymer melts for a range of applications; coatings, adhesives and hydrogels. Hydrogels are of particular interest because they are used extensively in biomedical applications such as drug delivery, tissue engineering and pervaporation membranes (used to separate mixtures of liquids). [POS1]. They are formed from network of cross-linked polymers, which have the ability to retain water, so their structure is largely characterized by crosslinking density, and has the ability to influence the diffusion or release of molecules. Because the swelling of PVA membranes in aqueous solution results in a decrease in mechanical strength, early studies predominantly focused on the effect of percentage water content on the state and structure of water in PVA. These targeted crosslinking systems typically include multiple polymers or polymers with long chain lengths, which match the behavior of real polymeric systems. A few groups – Tamai *et al* [POS12][POS13][POS14], Muller Plathe *et al* [POS4] [POS6] [POS8] [POS11], Wu *et al* [POS5], Zhang *et al* [POS3] and Chiessi *et al* [POS9] – performed extensive MD simulations of PVA in water. Through their works, these authors have demonstrated that the solubility of PVA declined in the presence of supercooled water, and highlight that temperature may in fact have more of an influence on the structure of PVA than just the water content. [POS5] As a result, more recent studies concern themselves with the influence of water and temperature on the behavior of PVA, while only a handful focus on shorter chain lengths in dilute solutions. [POS7][POS5][POS8] Advantages of these new systems are largely due to the fact that there are fewer segmental motions and so the time required for relaxation/equilibration is drastically less demanding. **[It is also useful for efforts in characterizing and ameliorating the polymers as function of chain length obtaining scaling factors for properties used to characterize polymers]**. Below is a brief summary of the literature.

**PVA and swelling** Zhang *et al* [POS3] studied the swelling properties of PVA in binary solutions, and PVA hydrogels in MD simulations have previously been shown to swell with increasing water content in alcohol-water mixtures. [POS3][POS4] These findings were observed by calculating the free volume using a solvent sized probe to trace the contours of polymers and obtain the exposed O’Connell surface. As as a direct result of increased mobility and disorder of PVA chains, the densities of swollen PVA decreases with increasing swelling [POS3]. Their study found that a solution containing 15%, 30% and 45% weight water, caused syndiotactic PVA50 networks to swell by 22.13% 39.10% and 84.21% respectively. The reason

for this is that large holes formed, which water molecules inhabit and these increase in size for systems with greater water content. Diameters of up to 0.90-1.10 nm were found to occupy 0.36% of the system containing the most amount of water. Inside the membranes, they also found a selective absorption for 76.61–81.30%, water rather than the larger ethanol molecules. A continuation of the work revealed that the calculated diffusion coefficient of the carbon backbones were 2.95, 10.83 and  $22.7 \times 10^{-8} \text{ cm s}^{-1}$  at 323.2 K . However the diffusion rates of water were not consistent with experimental values or trends at low water contents, which casts a shadow over the reliability of hydration properties and highlights the importance of careful selection of water models used. [POS3]]

**PVA in the cold** In 1993, Tamai and Tanaka studied atactic PVA<sub>21–161</sub> hydrogels subjected to both supercooled and warm temperatures between 200-400 K using 28-100% w/w SPC/E water and the AMBER/OPLS forcefield. [POS12][POS14]

**PVA and water diffusion** Muller-Plathe reported a concentration (and temperature) dependence of water diffusion in atactic PVA400 for systems between 3-100% weight water. They found that at cooler temperatures and higher water content, PVA drastically reduced the rate of water diffusion. At 300 K , they found that the SPC water models self-diffusion constant was  $3.32 \times 10^{-5} \text{ cm s}^{-1}$  and in a 34% water solution fell by an appreciable 91%. At 375 K the constant dropped by 87% to  $1.3 \times 10^{-5} \text{ cm s}^{-1}$ . These studied employed really short timescales of 2-16ns [POS6]

**PVA and temperatures** Wu *et al* [POS5] investigated the influence of temperature on an atactic PVA<sub>300</sub> in 21% weight water (200 TIP4P water molecules) in 10 ns production runs. They studied a trajectory of 14 ns at each temperature between ranges of 200-420 K using 20 K intervals. They found that a cooperative behavior between the PVA chain and water were responsible for the  $T_g$ , by evaluating the total number of hydrogen bonds per hydroxyl group for water-water, water-polymer and polymer-polymer interactions. They arrived at this conclusion because a marked transition occurs in the hydrogen bond number of water-polymer interactions, at the  $T_g$  temperature, while negligible changes occur at this temperature for the other two categories. Although OPLSAA has some of its own partial charges that could be used for atoms in order to build PVA, Wu *et al* took partial charges from the COMPASS forcefield and applied them using the OPLSAA forcefield and the GROMACS simulation software. This approach is not recommended because it creates an untested hybrid of forcefield parameters which requires evalua-

tion. They found a  $T_g$  of 283 K , which is drastically different from the experimental value of 385 K , although it lies between the  $T_g$  values for pure water (134 K ) and PVA. It goes to show the importance of careful forcefield selection (Wu tries to attribute this to the large rate of cooling)[POS5]In another study, Tesei *et al* [POS9] investigated atactic PVA<sub>30</sub> in 97% (w/w) water and three different temperatures; 293 K , 303 K and 323 K over a 90 ns period. They explored the effects of two different GROMOS forcefield and the found the later version was more suitable when compared to aqueous experiments. The differences between the two were largely due to variations in intensity of interactions between the polymer and water. The chain size, polymer conformation and hydrogen bonding of PVA were studied. They used the  $R_g$  to describe polymer chain extension and found that bimodal or trimodal distribution fits were useful depending on the forcefield used. Their study revealed that temperature induced transitions between a globular form and extended coils in solution. The  $R_g$  distribution for the better forcefield was trimodal, with peaks at 0.80–0.84 nm , 1.03–1.06 nm and 1.34–1.60 nm , which begin to converge at elevated temperatures.

**PVA and tacticities** Although tacticities of a polymer greatly affects its properties, few simulations have addressed the effects on PVA, and even though it may be important for its IRI potency, remarkably this too has not yet been explored in experiments. In 2013 Norjahan and Choi [POS2] investigated the effects of tacticity as well as temperature (200–550 K ) on specific volumes, the thermal expansion coefficient, the glass transition temperatures and the solubility of the PVA<sub>400</sub> polymers. They used the OPLSAA forcefield and the TIP4P water model, which has a melting temperature of 232 K . This means that over 100 ns period, they explored a temperature range between  $T_m-32$  K and  $T_m+318$  K , using a 10 K interval. They found that aside from overestimated heat capacities, the systems reproduced the thermal properties found in experiments, and whilst within the amorphous phase all the properties of the stereoisomers produced similar values except for the solubility parameters. The isotactic structure was the most soluble, followed by the atactic and finally the syndiotactic structure. Surprisingly they found that the  $T_g$  was insensitive to tacticity in MD simulations, however an average  $T_g$  of 350 K was obtained, which falls within the experimental range of 325-375 K . [POS2]

### **Antifreeze protein analogues**

There are even fewer simulations investigating antifreeze protein analogues or polymers with the intent of investigating their properties for antifreeze purpose. To our

knowledge, and to this respect, Tam *et al* [ANA1] and Corzana *et al* [ANA3] presented the only simulations of antifreeze analogues in solution between 2006 and 2010. Tam *et al* [ANA1] also produced a complimentary review and study of C-linked antifreeze glycoprotein analogues, using both experiments and in silico methods. The AFGP analogue system was modelled in solution for 10 ns at 300 K using the AMBER ff. and the TIP3P water model, consisting of 2000-3000 water molecules. The authors found that the intermediate side chain lengths of three carbons was necessary for IRI activity, whilst values smaller or greater than this exhibited no IRI activity (i.e. one, two or four carbons). Their simulations studies reveal that the IRI active structures adopted a unique conformation in solution that permits folding into the peptide backbone, and buries the hydrophilic face of the carbohydrates away from the bulk solvents. This folding behavior characteristically brings the carbohydrate moiety closer to the peptide backbone, which can form a hydrophobic pocket that may be used to approach the QLL. Conversely, the IRI inactive peptides have their hydrophilic moieties completely exposed.

Interestingly they did not manage to identify any hydrogen bonds between the amide protons and other hydrogen bond acceptors. They also noticed that the more potent IRI active peptides (with longer side chains) hosted weaker hydrogen bonding between the HO4 and HO6 galactosyl hydroxyl groups, than than the less potent one. They suggest that, as with carbohydrate systems, the interaction of HO4/HO6 atoms with adjacent water molecules is crucial to IRI potency, rather than for interactions between AFGP-analogues. [ANA1] To test this theory further, the energy penalty was also calculated as a function of rotating the linker bond between the carbohydrate side groups in relation to the protein backbone. They found that the energy penalty was highest for the IRI active peptides, which is consistent with their hypothesis. They propose the stereochemistry of these hydroxyl groups, and the folding of the carbohydrate moiety is directly related to regulation of the hydration environment of the peptides (as previously shown with O4 hydroxyls on carbohydrates).[ANA1] Especially in the light of recent work by Davidovic *et al* [ANA2][ANA1], which reported, that cold induced denaturing of proteins is often accompanied with an unchanged hydration sphere at 273 K . If this is certainly the case then these IRI active C-AFGPs may be able to help with modulate the hydration state accordingly.

Although the referenced IRI experiments were undertaken at the same temperature as Davidovic *et al* [ANA2][ANA1], the same cannot be said for the simulation studies conducted at room temperature. In order to confirm their speculations new simulations would need to be conducted to reevaluate the behavior of the truncated

peptides at 273 K and possibly even in contact with ice. From their study it is clear that the side chain length is important for conformational changes in order to influence hydration of the AFGP analogues, which in turn affect their potency. [ANA1] Similar to Tam *et al* [ANA1], Corzana *et al* [ANA3] conducted modelling studies which also support these findings that the carbohydrate moiety's interaction with surrounding water influences the conformation of the peptide, its flexibility and the structuring of vicinal water molecules.

### Ice crystal growth simulations

[Move stuff from ice section here and add handwritten review here.] [In 1987, the ice crystal melt interface was studied for the first time [Karim Haymet 1987] for all three planes of ice  $xy$  and  $xz$ . In 1988 the authors expanded their study to further knowledge in the basal plane [karim Haymet 1988], using an ordered ice hexagonal ice crystal. In both studies they found that the ice interface length is  $x$  and found that the model was stable at supercooling of  $x$  K. It is important to acknowledge that these studies were conducted for TIP4P water model. Since then many more studies have been made for this water model [ref], TIP4P/Ice [ref] and other water models [] which agree with their findings.:

Things we want to cover:

1. Who first invented ice crystal growth simulations
2. What were the setbacks at the time:]
3. What were the new developments and who by and what led to adding these features or studying them

The first ice simulation was made in  $x$ . They found that the QLL was this and that they could get  $x$  informations. This was just the beginning with room for improvement such as the treatments of compressibility of ice, the random orientation of ice for bernal fowler rules.

The key issues with that this meant  $x$ . A step toward more similar simulations to experiments were made by considering the QLL interface. This was done by using a technique used in the simulation of metals called the direct coexistence method first developed by  $x$ .

Many recent interest have developed with ice crsyatl structure since then using an array of models and of the contributors Vega et al are key players. It is difficult. A lot of work has been done by them in caharacterising the phase diagram of different models of water and identifying melting points using different techniques like free energy, mc and md.

It is very difficult to simulate spontenaoues homogensous nucleation of ice crystals and its continued freezing prpocess from bulk water in the timescales of md simulations. So studies are often conducted using an ice/water interface or ice/gas or ice/vacuum interface which has been equilibrated. Even till this day supercooled water may even be studied instead (eg in the case of investigating the High density-low density phase transition) or and other studies focus on confined water within narrow spaces (x-xnm) of nanopores (hydrophobic and hydrophilic) and various other interfaces (MgO..etc). A revolutionary step in simulation of ice crystals was made in 2002 when matsumoto was able to simulate homogenous numcleation of ice crystals.

Average size of crystal cluster is 280 molecules minimum needed for nucleation]

### **Aims and Objectives (1 paragraph on general aims then summary of each chapter**

Previous studies have shown that PVA and HYP are antifreeze active while PEG is not. This work aims firstly to reconcile our findings with some of the experimental observations of our polymer in both solution and in ice/water environments. It then aims to disclose to what extent each of the following traits feature in systems with PVA, PEG and HYP in the presence and absence of an ice crystal:

1. Hydrophilicity and Hydrogen bonding
2. Hydrophobicity
3. Water ordering or disordering
4. Hydration numbers and long or short-range ordering [Our observations in solution are equally as valuable as those in an ice environment because antifreeze active polymers do not necessarily act in the same way in both environments. It will serve as a platform to observe pre-cursor requirements for an ideal an-

to freeze polymer using MD and will ultimately help us to design or propose a preliminary structure.] [Might move to pol/sol introduction]

[Some additional questions that we also consider are]:

1. Can we see a TH in MD comparable to experiments?
2. Can we foresee IRI from decreased freezing rates?
3. Is adsorption irreversible?
4. Can we observe superheating at  $x_C$ , comparable to experiments?
5. Can we see evidence of step pinning mechanisms or the mattress model or neither or both simultaneously? [A26]
6. Can we see differences between melting and freezing rates at 270K?
7. Can we see multi-step features different to pure systems?
8. Does the QLL change and if so, how?
9. How does the polymer fit into the QLL at time intervals and does this compare to Knights theory? [A23]
10. What are the suitable explanations for differences in TH and IRI activity?
11. Does the polymer actually differentiate between the QLL and Ice and if so, how? [ might move to conclusion and use past tense]]
12. Does polymer mobility affect the propensity for the ice growth rates []

## Chapter 2

# Computer Simulation Methodology

A central aim of scientific research is to enable a deeper knowledge of how things work. By studying the motion of their constituent molecules, researchers can begin to understand the mechanisms for chemical reactions, molecular recognition and many other processes. Molecular dynamics (MD) is a computational technique used to simulate the movement of molecules or atoms over time. It uses mathematical models to calculate the potential energy from interactions between large number of atoms in the system, and generates a trajectory by numerically integrating Newtons equations of motion (Eq. 2.1) at regular intervals, updating the atom positions.  $F_i$  is the force,  $m_i$  is the mass and  $a_i$  is the acceleration,  $v_i$  is the velocity and  $r_i$  is the position of the atom,  $i$ . MD simulations provide atomistic resolution and describes the dynamics and static properties of a system. Given the use of an accurate model system, sufficient sampling time and a suitable number of molecules to average over, these inherent features mean MD simulations provide strong predictive capabilities, and so it is also used to supplement experiments where sampling is untenable via traditional experiments or techniques.

$$F_i = m_i a_i = m_i \frac{\delta v_i}{\delta t} = m_i \frac{\delta^2 r_i}{\delta t^2} \quad (2.1)$$

### 2.1 Statistical mechanics

In experiments, we sample the bulk system and therefore gain a macroscopic view of a property of interest, also known as an ensemble average. MD simulations provide

atomistic details of the  $N$ -body particle systems and so in order to translate our findings into real life terms, we need to use ensembles time average of our properties. To achieve this this we employ statistical mechanics such as the ergodic hypothesis; given sufficient trajectory time, the system will sample every state, and so the time average will match the ensemble averages. In summary, statistical mechanics helps us to numerically represent the large number of configurations attainable by a systems state (which typically contains  $10^2$ – $10^4$  atoms) and so a full treatment of the many degrees of freedom can be more adequately described.

It is possible to use different types of ensembles. The most frequently used are the Microcanonical ( $NVE$ ), Canoninical ( $NVT$ ), Isobaric-Isothermal Ensembles ( $NPT$ ) and Grand Canonical Ensembles ( $\mu VT$ ). For each of these the letters in parenthesis,  $N$ ,  $V$ ,  $E$ ,  $T$ ,  $P$  and  $\mu$ , stands for a fixed number of atoms, volume, temperature, pressure and chemical potential respectively.

## 2.2 Forcefields

Interactions between atoms are important and a ubiquitous in everyday life, so they must be adequately described in MD simulations. They typically governed by a set of principles, which allow specific reactions take place under different conditions. A textbook definition of forcefields is mathematical formulas, which are used to describe the potential energy ( $U$ ) of a system of particles (Eq. 2.2). In a nutshell, they determine which atoms are allowed to interact with the other and how. These include descriptions of bonded and non-bonded terms, which are discussed in the next section, and they also detail constants such as the atomic masses and radii.

Among the vast range of forcefields, some of the most commonly used are CHARMM, AMBER, GROMOS and the OPLS forcefield, all of which were derived from a combination of experiments and quantum mechanical studies of small molecules. These in turn yield more reasonable models of proteins and organic systems. However, the onus lies on individuals to determine whether these parameters are transferrable to their own systems and molecules of interest.

$$U = U_{bonded} + U_{non-bonded} = (U_{bond} + U_{angle} + U_{dihedral}) + (U_{electrostatic} + U_{vdw}) \quad (2.2)$$

### 2.2.1 Bonded interactions

Bonded interactions are treated as stick and ball models of atoms, which oscillate harmonically. There are 3 types: bond stretching, angle bending and dihedral angles

(Fig. 2.1). For each unique pair of atoms, covalent bonds ( $U_r$ ) are described by the functional form in Eq. 2.3. The forcefield details an equilibrium distance ( $r_{ij-eq}$ ) between the two atoms and any deviations away from that distances incurs an energy penalty, which is added to the potential energy. The functional form for angles ( $U_\theta$ ) and dihedrals ( $U_\Phi/U_\Psi$ ) are also treated in a similar fashion (Eq. 2.4–2.6 respectively). In these equations  $k_r$ ,  $k_\theta$  are force constants,  $k_\Psi$  is the peak height,  $n$  is the frequency, and  $\Psi$  is the phase. All constants are highlighted in red.

$$U_r = \sum k_r (r_{ij} - r_{ij-eq})^2 \quad (2.3)$$

$$U_\theta = \sum k_\theta (\theta_{ijk} - \theta_{ijk-eq})^2 \quad (2.4)$$

$$U_\Phi = \sum k_\Phi (1 + \cos(n\Phi - \Phi)) \quad (2.5)$$

$$U_\Psi = \sum k_\Psi (\Psi_{ijkl} - \Psi_{ijkl-eq})^2 \quad (2.6)$$

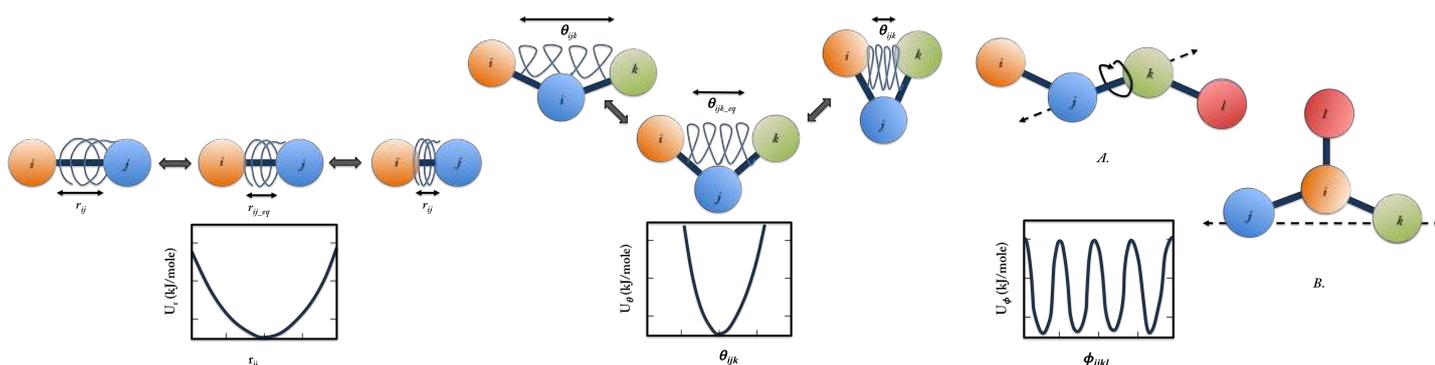


Figure 2.1: Schematic of the potential energy profiles for bonds (left) angles (center) and dihedral angles (right), where A. represents proper dihedrals and B. improper dihedrals

## 2.2.2 Non-bonded interactions

Non-bonded interactions are composed of two parts; van der Waals and Columbic term. Calculations for this part of the simulations are typically the most computationally expensive part of MD simulations, because interactions need to be evaluated between thousands of atoms at each time step. As users, we often have to make trade-offs between the accuracy and efficiency of these calculations. In order to run longer simulations we therefore aim to reduce the cost of computation and favor efficiency. In practice, one way of doing this is to calculate only the interactions between

pairs of atoms, rather than all possible combinations like triplets or quadruplets and so on (hence the frequently used textbook phrase, “pairwise additive functions”).

### Van der Waals

Van der Waals forces are described by Lennard Jones (LJ) 6-12 potential ( $U_{LJ}$ ); where attractive forces are long range ( $r_{ij}^{-6}$ ) while repulsive forces dominate at shorter ( $r_{ij}^{-12}$ ) distances (Eq. 2.7-2.9). To further reduce the cost of computation, cut-off distances are typically used with van der Waals interactions, and atoms that are separated by distances greater than the cut-offs are ignored or approximated. The reason why cut-offs can be used in this instance is because van der Waals interactions decay rapidly to zero as distances increase. As a result, by the time they reach the cut-off distances, their potential energies are so small that they are considered negligible. To date there are various methods for applying cut-offs: truncation, SHIFT and SWITCH are the most widely used. Truncation means that interaction energies are simply set to zero at the cut-off distance or discontinued. However, this abrupt change can lead to poor energy conservation and infinitely large artificial forces in the latter case. The SHIFT method essentially shifts the entire potential energy surface (i.e. the LJ potential curve) so that the interaction potential energy is zero at the cut-off distance and it typically results in higher energies and smaller equilibrium distances between the atoms. Finally, the SWITCH method serves as a combination of the two. Two cut-off distances are defined and within the small range between them the interaction potential gradually settles to zero, using a smooth scaling function. Today, computational power is much greater than that of our predecessors, and so cut-off distances are typically larger (from 0.75–1.5 nm ) depending on the system under study. A schematic of the cut-off schemes is shown in Fig. 2.2.

$$U_{LJ} = 4\epsilon_{ij}((\sigma_{ij}/r_{ij})^{12} - \sigma_{ij}/r_{ij})^6 \quad (2.7)$$

where a geometric average is used to calculate the two parameters in the OPLS-AA for:

$$\sigma_{ij} = (\sigma_{ii}\sigma_{jj})^{\frac{1}{2}} \quad (2.8)$$

$$\epsilon_{ij} = (\epsilon_{ii}\epsilon_{jj})^{\frac{1}{2}} \quad (2.9)$$

$$(2.10)$$

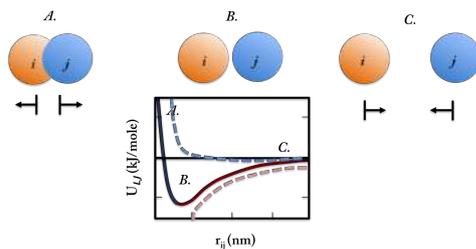


Figure 2.2: Schematic for the LJ, where (A) described the dominance of the repulsion term at short distances (blue), (B) is the optimum energy and equilibrium distance and (C) is the attractive term at long distances (red)

### Coulomb potentials

The coulomb potential is used as the functional form for modelling electrostatic interactions between a pair of charged atoms. The formula is shown in Eq. 2.11, where  $q_i$  is the charge for atom  $i$ ,  $r_{ij}$  is the distance between atoms  $i$  and  $j$  and  $\epsilon_0$  is the dielectric constant (also known as the permittivity). Compared to van der Waals, electrostatics is considered longer-range which means that it takes much longer to decline with distance. To be exact it declines as  $r^{-1}$ . It is therefore ideal to completely avoid truncation styled cut-offs, which can lead to huge errors and unphysical results. Even more so, as electrostatic interactions play a serious role for many properties seen in real systems such as transport properties and protein folding. Instead other methods are commonly used for these interactions, mainly switching functions (mentioned earlier) or Ewald summation techniques. [2]

$$U_{Coulombic} = \frac{q_i q_j}{4\pi\epsilon_0 r_{ij}} \quad (2.11)$$

### Ewald summations

Ewald summation is a technique for calculating electrostatic interactions, which also considers periodic images (see 2.4.5). It works by splitting summations into long- and short-range components. The short-range contributions fall within the cut-off distance and so are treated as normal (i.e. in real space). Next it divides the simulation box into a 3D grid and uses Fourier Transform (FT) to interpolate charges from the long-range contributions unto the nodes of the grid. Charges are calculated in Fourier space (i.e. reciprocal space aka momentum space) by assigning atoms their charges according to their weighted distance from the node. The resulting charges are then interpolated back into real space. Compared to traditional direct summa-

tion, Ewald summation is a much faster technique, so in recent years a multitude of similar methods have been developed to speed up calculations whilst incorporating the periodicity of the system. An example worthy of mention is the Particle Mesh Ewald method, which assigns charges to a grid using cardinal B-spline interpolation. The grid is then transformed using a Fast FT algorithm and the long-range energy term is calculated by summing over the grid in the reciprocal space. The reason why these grid-based methods are so widely accepted is that there are fewer nodes than there are atoms. As result a smaller number of calculations can be made for the same interactions and systems.

### 2.2.3 Water forcefields: 3-site and 4-site models

Water models and water forcefields are used interchangeably to describe how water molecules are represented in simulations. The most common water models are typically 3-site or 4-site rigid models (discussed in 3). These are ordinarily designed to mimic one or more properties of real water using a combination of experiments and computational techniques. Over the years various water models have been developed to account for different types of the seemingly never-ending properties of water such as dielectricity, density, heat of vaporization and the phase diagram of water. The different types of explicit water models are ordinarily classed according to the number of sites, flexibility (or rigidity) and polarisability. Both types of water models are depicted in the Fig.2.3. A 3-site water model is planar and has 3 atoms; an oxygen atom that carries a partial negative charge and two hydrogen atoms that are positive. In contrast, a 4-site water model has an added, massless dummy atom that resides at the bisector of the HOH molecule. In this model the negative charge is moved away from the oxygen and placed on the dummy atom in order to represent the lone pairs of the oxygen. Compared to the 3-site water models this change is known to improve the electrostatic distribution around the water molecules.

Today, the refinement of water models have advanced to support the different simulation parameters such as Ewald styled cut-offs, and not solely the properties of water. Unsurprisingly, the greater the models complexity, the longer the time taken for computation especially considering the enormous number of water molecules required for solvation. To avoid this, implicit solvation can also be used, where water molecules are represented by a distance dependent dielectric-screening function. This approach reproduces the average behavior of water molecules screening the interaction between two non-solvent molecules, however the disadvantage is that interesting interactions between water molecules and non-solvent molecules cannot

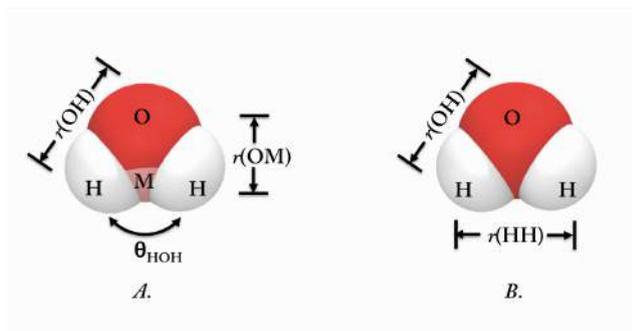


Figure 2.3: Schematic of (A) a 4-site water model and (B) a 3-site water model, where O-H bond lengths typically vary between 0.0957–0.100 nm and HOH angles between  $104.52^{\circ}$ – $109.5^{\circ}$

be reproduced. [1]

In 2002, Guillot *et al* [2] produced a comprehensive review of the older water models, in order to help us decide which one is best to use. The way that the polymers are able to interact with both ice and surrounding water is very much dependent on which model we pick. Seeing as water is ubiquitous in nature and plays a crucial role in almost every biological and synthetic process, we need to pick a reliable model to ensure our simulations are useful. In this report, we predominantly work with a 4-point water model called TIP4P/Ice [REF] water model, which has been designed to describe the phase diagram of water. Its use, advantages and forcefield parameters are discussed further in 3.

## 2.3 Potential energy minimization

Energy minimization methods (EM) are conducted to obtain the equilibrium configuration of molecules, which is often necessary to produce stable and experimentally relevant MD simulations. Initial structures are usually obtained from NMR or X-ray studies or built using a program, and then they are typically converted into lower-energy states of interest. EM can be achieved by using algorithms like the Steepest decent or the Conjugate gradient methods, which are used to search for local minima. These minimizations remove unreasonable or close contacts between given atoms in a system, which are often undesirable and unrealistic. Due to the complex nature of the potential energy surface, these algorithms typically bring the state to a local energy minimum, and not necessarily the global energy one. A schematic of this is shown in Fig. 2.5. To achieve the latter, simulated annealing and other computational techniques are often used. [3]

In this work, we use the steepest decent algorithm, which is a first derivative algorithm and so is much faster than second derivative alternatives. The steepest descent conceptually works by reevaluating the gradient of the potential energy as it is calculated. After each step, a displacement is added to the coordinates in the direction of the force and the size of this displacement or step is increased if the energy is lower than the previous calculated energy. Otherwise the step is reduced and the same loop is reiterated for a user-defined number of iterations. [4]

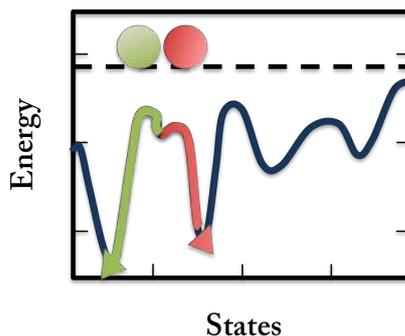


Figure 2.4: Schematic of a 1D cross section of the potential energy surface. The dashed line is the starting point for 2 different structures. The colours show their resulting end states at their different local minima, where green is the global minimum of the system.

## 2.4 Molecular dynamics

In molecular dynamics simulations, atoms are treated as balls with masses, and bonds are treated as springs with force constants as described earlier. Starting coordinates for each atom is taken from experiments and initial velocities are randomly assigned to every atom from the Maxwell-Boltzmann distribution (Eq. 2.12). This method for assigning initial velocities encourages equilibration distribution which would normally be expected at the temperature of interest. This approach is much better than initializing all the velocities to zero which is also fairly simple task, however would result in less stable structures because the atoms experience a thermal shock.

$$E_{kinetic} = \frac{1}{2}mv^2 = \frac{1}{2}k_B T \quad (2.12)$$

Consider an atom  $i$ , of mass  $m_i$ , which moves from position  $r_0$  to  $r_1$ , after a period of time  $\delta t$  because it has an initial velocity of  $v_0$  (Eq. 2.13). The move to

this new position,  $r_1$ , can be described by (Eq. 2.14).

$$v_0 = \frac{r_1 - r_0}{\delta t} \quad (2.13)$$

$$r_1 = (v_0 \delta t) + r_0 \quad (2.14)$$

The net force acting on atom  $i$  at this new position, is a direct consequence of all its interaction with every other particle in the system. The potential energy from bonded and non-bonded interactions are summed in order to produce a potential energy profile. Eq.2.15 describes the relationship between potential energy and force, which says that the force can be calculated from the negative gradient of the potential energy profile.

$$F_i = -\frac{\delta U}{\delta r_i} \quad (2.15)$$

Newtons second law of motion relates the calculated force with the acceleration of a particle of a given mass (Eq. 2.1). With this new force we can rearrange this equation to obtain a value for the atom's acceleration (Eq. 2.17), then calculate the new velocities at  $v_1$  using the old velocities (Eq. 2.18). Finally, from this new velocity one can calculate the new atom positions at time  $t_2$ , using the old positions (Eq. 2.19 and 2.20) and evolve the simulation in time for a predefined number of iterations.

$$\therefore a_0 = \frac{F_i}{m_i} = \frac{\delta U}{\delta r_i} m^{-1} \quad (2.16)$$

$$a_0 = \frac{\delta v}{\delta t} = \frac{v_1 - v_0}{\delta t} \quad (2.17)$$

$$\therefore v_1 = (a_0 \delta t) + v_0 \quad (2.18)$$

$$v_1 = \frac{\delta r}{\delta t} = \frac{r_1 - r_0}{\delta t} \quad (2.19)$$

$$\therefore r_2 = (v_1 \delta t) + r_1 \quad (2.20)$$

### 2.4.1 Time integration algorithms

In practice the trajectory of the atoms are not calculated directly from Newton's equation of motion due to the lack of an analytical solution, because the 3D potential energy surface is very complicated. Rather, in order to calculate the force (i.e. the gradient), an efficient time integration algorithms such as the verlet or leap-frog

algorithms is used. The leap-frog [1] algorithm is described by Eq. 2.21–2.22. It's given this name because the positions ( $r$ ) are obtained at time  $t$  while its velocities ( $v$ ) are determined at time  $t - \frac{1}{2}\Delta t$ . Then, it updates  $r$  and  $v$  using the forces determined by the positions at time  $t$ :

$$v_{t+\frac{1}{2}\Delta t} = v_{t-\frac{1}{2}\Delta t} + \frac{\Delta t}{m} F_t \quad (2.21)$$

$$r_{t+\Delta t} = r_t + \Delta t v_{t+\frac{1}{2}\Delta t} \quad (2.22)$$

Each of these schemes are approximations (using the Taylor series expansion) that may result in errors over time such as energy drifts. Its crucial for the algorithms to minimise these and to preserve energy, in order to accurately represent the system and it's forces. Due to the half-step calculations, the leap-frog algorithm is inherently more accurate than the verlet method because the kinetic energy and temperature of the system would be calculated more [frequently/closely] given the same same step size. This is a good example to highlight how truncations errors can be intrinsic to the algorithms used.

In order to minimize these errors, small time steps of 1–2 fs can be used, which also preserve the capture adequate sampling of the fastest vibrational fequency. It is important to understand that the size of time steps limits the length of the CPU time, because the processes that one may want to study could occur under microseconds, and so could require up to one million steps or iterations. In addition, these algorithms are also prone to rounding errors that are associated with the finite number of digits used in the computations. Although these arreors are unavoidable they can be kept at a minimum by using double precision which is widely used in workstations today.

### 2.4.2 Constant temperature MD

Thermostats are coupled to the system in order to modulate the average temperature of the systems, rather than by directly altering the temperatures because this would have unrealistic implications for the total kinetic energy of the system at the desired temperature. Instead the temperature is allowed to fluctuate about the average, and the thermostats function by by adding or removing energy from the system outside of the boundaries. Some of the most popular thermostats that are used in MD simulations are the Velocity rescaling [2], Andersen [3], Berendsen [4] and the Nosé-Hoover [1] thermostat, all of which help to reduce the energy drifts incurred by the calculations of the total energy over time.

### Berendsen thermostat

The Berendsen algorithm is described by Eq. 2.23 and it weakly couples the system to the external heat bath at temperature  $T_0$ . It works by scaling the velocities of atoms gradually in proportion to the deviation of the average system temperature to  $T_0$ . As a result, a slow correction is applied to the system temperature ( $T$ ), which makes this scheme suitable for equilibration. As a result the deviation decays exponentially with the coupling constant, i.e. the coupling strength ( $\tau$ ).

$$\frac{\delta T}{\delta t} = \frac{T_0 - T}{\tau} \quad (2.23)$$

### Nosé-Hoover thermostat

Although the Berendsen thermostat is an efficient method for relaxing a system, it is not considered suitable for representing a canonical ensemble because it also suppresses kinetic fluctuations. As a result the thermostat is not ideal for the purpose of calculating properties like the heat capacities. Once equilibrium has been reached it is important to probe the correct canonical ensemble and so the Nosé-Hoover [1] algorithm is more appropriate for this purpose because it is an extension of the Lagrangian approach to classical mechanics. Instead of constraining the kinetic energy of the system, the Nosé-Hoover thermostat acts to keep the energy of the system close to the desired value. In order to achieve this it introduces a virtual thermal reservoir at position  $r$  and a friction term in the equations of motion. The friction force is proportional to the product of each atom velocity and a frictional parameter ( $\xi$ ) — a dynamic property with its own virtual momentum ( $p_\xi$ ), virtual mass ( $m_\xi$ ) and equation of motion (Eq. 2.25), for which the time derivative is calculated from the difference between the  $T$  and the  $T_0$ . The Newtonian equations therefore have the additional terms in Eq. 2.24, where  $Q$  is the coupling constant.

$$\frac{\delta^2 r_i}{\delta t^2} = \frac{F_i}{m_i} - \frac{p_\xi}{Q} \frac{\delta r_i}{\delta t} \quad (2.24)$$

$$\frac{dp_\xi}{dt} = T - T_0 \quad (2.25)$$

Recall that  $T$  is proportional to the kinetic energy.

### 2.4.3 Constant pressure MD

Pressure coupling of an external barostat is added to the system in order to make it reminiscent of a physical experiment. The physical pressure is typically altered by adjusting the volume of the box analogous to the effects of a pressure piston. The most common thermostats used are the Berendsen [1] and Parinello-Rahaman [2] barostats. During the volume adjustments, the barostats also rescale the coordinates of the atoms which make up the systems so that they remain at the same relative distance from each other. Isotropic pressure coupling describes a uniform volume change, while semi-isotropic and anisotropic describe non-uniform scaling in a plane or separately in x, y and z directions respectively.

#### Berendsen barostat

Similar to its thermostat counterpart, the Berendsen barostat is also weakly coupled to the system and rescales both the box size and atom coordinates at every step towards a target pressure,  $P_0$  (Eq.2.26). It achieves this by using a matrix ( $\mu$ ), which incorporates the isothermal compressibilities of the bulk system.

$$\frac{\delta T}{\delta t} = \frac{T_0 - T}{\tau_P} \quad (2.26)$$

#### Parinello-Rahman barostat

The Parinello-Rahman barostat is different in that it allows for a pressure and temperature feedback mechanism. Similar to the Nosé-Hoover thermostat, it appropriates the langarian treatments and leads to more appropriate  $NPT$  ensemble. The box vectors are represented by the matrix  $b$  and its equations of motion are given by Eq.2.27. Where  $V$  is the volume of the box and  $W$  is the matrix parameter, which includes the isothermal compressibilities and determines the coupling strength ( $M$ ). Just as with the Nosé-Hoover thermostat the equations of motion for the particles are also changed as a result (Eq.2.28).

$$\frac{db^2}{dt^2} = VW^{-1}b'^{-1}(P - P_0) \quad (2.27)$$

$$\frac{\delta^2 r_i}{\delta t^2} = \frac{F_i}{m_i} - M \frac{dr_i}{dt} \quad (2.28)$$

#### 2.4.4 Bond constraints

Bond constraints are applied in order to maintain them within certain lengths. Different types of constraints algorithms exists for linear bond geometries of which SHAKE [1] and LINear Constraint Solver (LINCS) [3] algorithms are the most commonly used. These typically exclude certain intramolecular motions and so they offer much larger time steps which results in more efficient computation times. This is particularly useful when modelling hydrogen bonds, which are too important to neglect but produce high frequency vibrational motions. In order to apply constraints to non-linear geometries such as those found in 3-site water models, a more suitable protocol like the SETTLE [2] algorithm is typically applied.

#### 2.4.5 Periodic boundary conditions

In finite size systems it is important to remove undesirable surface artifacts, which would ordinarily arise from simulating close to the box edge or at the vacuum interface. Periodic boundary conditions (PBC) are rules used to treat a simulation cell containing a few number of particles, so that it reproduces properties of a bulk system. These conditions are achieved by surrounding the origin cell with copies of itself and its contents. As every copy perfectly mimics the original cell, if a particle were to exit through one side (and into a neighboring cell), a copy of the original particle would also re-enter the original cell simultaneously, but from the opposite side. PBC conditions means that each particle is able to experience forces from atoms within the origin cell and its neighboring cells. For simplicity, the minimum image convention is applied and a cut-off (smaller than half the smallest box vector) is applied so that a particle in the origin cell does not interact with its own image as though it were part of its solvation shell. It also states that each particle can only interact with 1 image of another particle at a given time, and it must always be the closest copy. A schematic of a PBC set up is shown in Fig. ??

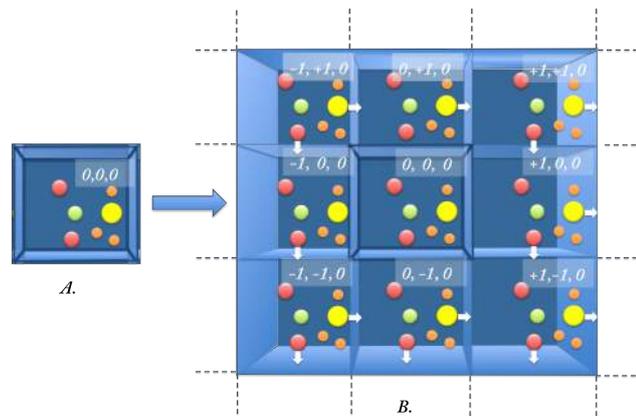


Figure 2.5: Schematic of a 2D cross section of periodic boundary conditions.

## Chapter 3

# Mixing and validation of OPLS-AA and TIP4P/Ice forcefields for simulations of ice growth in the presence of additives

### 3.1 Introduction

Modelling is now used to solve an incredibly wide range of problems from protein folding to adsorption and even crystal growth. The consequence is that vast amounts of forcefields for molecules have developed independently from one another. As some of these communities haven't been in collaboration in order to do so, this makes mixing forcefields an issue. It's therefore advised to check and quantify their compatibility before using these models and forcefields combined to address our simulation questions.

Many forcefields are developed to work with particular water models to obtain a proper balance between solvents and solute interactions. The OPLS-AA (Optimized Potential for Liquid Simulations - All Atoms) forcefield Jorgensen and Tirado-Rives [1988], was originally developed with 4-Point-Transferable-Intermolecular-Potential (TIP4P) Jorgensen et al. [1983] water and has been reported to work slightly better with this model than with TIP3P Jorgensen et al. [1983] and SPC Jorgensen and Tirado-Rives [1988] (simple point charge). It would therefore make it a preferential choice to a selected range of water models ?. Similarly, the CHARMM?? and AM-

BER? forcefields was developed with TIP3P. However practically many simulations still use other water models because of their various advantages and disadvantages which are thoroughly documented in literature Guillot [2002]; Vega et al. [2009]; Jorgensen et al. [1983].

For instance, the TIP4P/Ice water model is a modified version of the TIP4P water model, which was designed to adequately simulate freezing and captures the solid phase diagram of water Abascal et al. [2005]; ?; Bernal and Fowler [1933]; Guillot [2002]; Jorgensen et al. [1983]. Reports by Vega, Absacal, Weisser and others in The thickness of a liquid layer on the free surface of ice as obtained from computer simulations by MM Conde and Vega have repeatedly shown that unlike earlier water models (SPC, SPC/E?, TIP3P, TIP4P, TIP4P/2005), TIP4P/IceAbascal et al. [2005] does not significantly underestimate the melting temperature  $T_m$ , and has the closest value to experimental results at around 272.2 KAbascal et al. [2005]; García Fernández et al. [2006]; ? (compared to 190.5 KAbascal et al. [2005]; ?; ?, 215 KAbascal et al. [2005]; ?; ?; Conde et al. [2008], 145.6 K Vega et al. [2009]; ?; ?, 232 KAbascal et al. [2005]; ?; ?; García Fernández et al. [2006]; ?; ?; ?; Conde et al. [2008] and 252.1 K?García Fernández et al. [2006]; Conde et al. [2008] respectively) Conde and Vega [2010]; ?; ?. Although later models like TIP6P? have improved on this ability Abascal et al. [2005]; García Fernández et al. [2006]; ?; ?; ?, its additional sites makes it almost twice as computationally expensiveAbascal et al. [2005]. Understandably, these temperatures vary in literature as a result of using different computational methods and protocols. Additionally the 4-Point water models have been shown to predict the correct stable stable ice phase (Ice 1h), for which other models mistakenly predict other forms e.g. Ice II?García Fernández et al. [2006]; ?; ?; ?; Abascal et al. [2005]

In this study we aim to investigate the compatibility of the TIP4P/Ice water model with the OPLS-AA forcefield. These simulations are important for our investigations of ice growth in the presence and absence of additives like salts, antifreezes and synthetic macromolecules. Combining the OPLS-AA and TIP4P/Ice forcefields are particularly interesting to our studies as they were originally designed to improve the modelling and simulation of pure fluid systems. It is therefore important to quantify how well this and other water models perform with the OPLS-AA forcefield. In order to achieve an accurate description of our systems, we first employ the standard tests which check that the water models reproduce or resemble the well-documented dynamic and structural properties of water Levitt et al. [1997]; van der Spoel and Lindahl [2003]. Primarily the investigation explores the dependence of

the water properties on three temperatures and three cut-offs. Next we use the same tests to investigate how water behaves in the presence of salt ions (NaCl) and ethanol, a small organic molecule with known antifreeze activity.

The remainder of the paper is organised as follows. In section 1 we discuss the computational methodology. In section 2 we compare the properties of all the different water models at ambient conditions to one another and to experimental results. These properties include the radial distribution functions, diffusivity, densities, hydrogen bonding, orientational order parameter and the phase transition kinetics. Since each of these models are optimised to fit different physical properties of water, they are expected to behave differently. However comparisons are also made with reported theoretical values, which serve as valuable benchmarks. Next in section 3 we evaluate the effect of 3 temperatures (235 K, 298.15 K and 348 K) on the properties of the TIP3P, TIP4P and TIP4P/Ice water models. These simulation are relevant to systems which require the use of non-ambient temperatures such as the supercooling of water.

Fourthly we investigated the effects of 3 commonly employed cut-offs (0.85 nm, 1.00 nm and 1.5 nm) for the aforementioned models. In order to understand the importance of investigating the effect of cut-offs, let's take the example of simulating proteins, polymers and alcohols which are typically done using 1.0–1.5 nm cut-offs in a box of solvent. A recently published paper on the re-parametrized OPLS-AA forcefield recommended similar cut-offs with molecules like propan-1-ol and other hydrocarbons. Jorgensen et al. [1996]; Siu et al. [2012]. In contrast the TIP4P/Ice water was parametrised using 0.85 nm cut-offs and so simulating a system which require both models could prove difficult Abascal et al. [2005]. Which cut-off do we choose? Should we compromise between the two? Depending on the focus of the research, an informed compromise needs to be made which allows an accurate or reasonable description of both models. This is a difficult topic which only testing can give us the answer and our paper serves as a contribution to the wealth of work already done in this area.

The fifth and sixth sections of the paper focus on simulations of the water models in the presence of varying concentrations of salt (0.1–0.43 M) and ethanol (10%–100%). The TIP4P/Ice water model was not parametrized to interact with any additives, while the OPLS-AA forcefield has been re-parametrized for use with organic liquids like alkanes as well as aqueous solvents. It is important to demonstrate that the integrity of the reported properties of water are still maintained and that the experimental properties of the new mixtures can be reproduce within rea-

son. Finally the last section of this paper summarises the kinetics of the melting and freezing for TIP4P/Ice. It also reports the effects of cut-offs (and temperature) on the melting and freezing points as well as their rates.

## 3.2 Methods

### 3.2.1 Forcefield parameters

We use 5 different water models, SPC/E, TIP3P, TIP4P, TIP4P/Ice and TIP5P?, to describe the atomistic interactions and molecular dynamics of water. In each of these models, only the oxygen atom is considered to be a Lennard-Jones (LJ) interaction site and partial charges are distributed on atomic sites as well as virtual sites (Table 3.1). The 4-point and 5-point water models have massless dummy atoms, which we label  $M$  for simplicity. These sites are treated specially and their positions are calculated from the position of other sitesGarcía Fernández et al. [2006]. While, the OPLS-AA salt ions ( $\text{Na}^+$  and  $\text{Cl}^-$ ) and rigid ethanol parameters were taken from Jorgensen *et al* paper ? and GROMACS simulation package Hess et al. [2008]; ? respectively.

Model	SPC/E	F. SPC	TIP3P	TIP4P	TIP4P/Ice	TIP5P
$r(\text{OH})$ , Å	1.00	1.00	0.9572	0.9572	0.9572	0.9572
$\text{HOH}$ , $\theta$	109.47	109.47	104.52	104.52	104.52	104.52
$r(\text{OM})$ , Å	-	-	-	0.15	0.1577	0.70
$q(\text{O})$	-0.8476	-0.82	-0.834	-	-	-
$q(\text{H})$	+0.4238	+0.41	+0.417	+0.52	+0.5897	+0.241
$q(\text{M})$	-	-	-	-1.04	-1.1794	-0.241
Description	3-Point	3-Point	3-Point	4-Point	4-Pointt	5-Point

Table 3.1: Atomic charges ( $q$ ), bond angles ( $\theta$ ) and bond distances (Å) for various water models .

### 3.2.2 Simulation parameters

#### Pure liquid water simulations

The forcefields compatibility with various water models were validated by simulating a pure water system at ambient (298.15 K and 1 bar) and periodic boundary conditions. For these simulations 4,008 water molecules are placed in a cubic simulation box of approx.  $4.95 \text{ nm}^3$ . The simulations are performed in the NPT ensemble using

GROMACS and the OPLS-AA forcefield. In all simulations the Leapfrog algorithm was used to solve the equations of motion, and the SETTLES? algorithm was applied to constrain the bond length and angles of the water models in order to allow a larger integration time step of 2 fs. Lennard-Jones (LJ) potentials were cut-off at 0.85 nm and the particle-mesh Ewald (PME) algorithm Darden et al. [1993]; Essmann et al. [1995] was employed which meant that electrostatics could be treated accurately. The real part of the Coulombic potential was also truncated at 0.85 nm while the width of the mesh was 0.16 nm and a cubic order polynomial was used. The Nosé-Hoover thermostat Hoover [1985]; Nosé [1984] was used alongside the Parrinello-Rahman Barostat Parrinello and Rahman [1981] to maintain isothermic-isobaric conditions. The relaxation times for the thermostat and barostat were 0.2 ps and 2 ps respectively. Isotropic conditions were maintained in this simulation, using the compressibility of water at  $4.5 \times 10^{-5} \text{ bar}^{-1}$ . Unless stated otherwise, the temperature and cut-offs were maintained at 298.15 K and 0.85 nm respectively.

Each simulation lasted 3 ns and the first 1 ns of each run was discarded as equilibration. The remaining 2 ns was used for analysis. This range was selected by monitoring the conversion of the O-O RDFs of water over time (typically achieved by 0.9 ns, using short consecutive time blocks of 0.3 ns e.g.  $t_0=0-0.3$  ns,  $t_1=0.3-0.6$  ns ...  $t_{10}=2.7-3.0$  ns).

### Saline solution simulations

In order to validate the compatibility of the forcefields with small molecules like salts, and to compare the the simulations with the pure water systems, similar conditions as above were employed. The only exception was that water molecules were systematically and randomly replaced with Na and Cl ions, in order to achieve the specified concentration outlined in Table 3.2.

<b>Conc. (M)</b>	<b>0.10</b>	<b>0.14</b>	<b>0.21</b>	<b>0.28</b>	<b>0.35</b>	<b>0.43</b>
Water	3,994	3,988	3,978	3,968	3,956	3,946
Na	7	10	15	20	26	31
Cl	7	10	15	20	26	31

Table 3.2: The number of water molecules and ions in the saline solutions. Overall, the systems still remain neutral.

## Water/ethanol mixtures

A single ethanol molecule was energy minimised using the steepest decent algorithm for 20,000 steps. Then a box of water was made and solvated with multiple ethanol molecules as specified in Table 3.3. To avoid unrealistic initial forces in the simulation, the system was energy minimized again under the same prior conditions before the final production run. During this 3 ns simulation, all hydrogen bonds in the system were enforced using LINCS algorithm?. Although we use twice as many molecules in the system, the following protocol serves as comparison to the results from Wensink *et al* ?.

Mass % (M)	10	20	30	40	50	60	70	80	90	100
Water	1,800	1,600	1,400	1,200	1,000	800	600	400	200	0
Ethanol	78	156	234	312	392	470	548	616	704	782
$X$	0.042	0.088	0.143	0.206	0.281	0.370	0.477	0.610	0.779	1.000

Table 3.3: The mass (M) and mole ( $X$ ) fraction of ethanol in the solution and the number of water and ethanol molecules in the mixtures.

## Free Energy of Solvation of Ethanol

A single energy minimised ethanol molecule was arranged in a cubic box approximately  $2.4 \text{ nm}^3$  and was solvated by 310 water molecules. The system was energy minimized again and equilibrated for 3 ns at the same conditions employed for the ethanol-water mixtures. Finally the production run was conducted using the leap-frog stochastic dynamics integrator, using lambda values from 0 to 1 with 0.1 intervals. The first 1 ns of the simulation was discarded as equilibration and the last 2 ns were evaluated using the Bennett’s acceptance ratio.

## Freezing and melting simulations

These simulations are used to determine the melting and freezing kinetics of the TIP4P/Ice water model under the OPLS-AA forcefield. It is used to check that this particular model still maintains its good representation of the water phase diagram in conjunction with the OPLS-AA forcefield. For these simulations we followed a similar protocol to that of Weiss *et al*/Weiss *et al*. [2011] and Fernández *et al* García Fernández *et al*. [2006].

Essentially the simulation parameters are identical to the liquid simulation unless stated otherwise. A  $2.9 \times 3.1 \times 2.7 \text{ nm}$  box of hexagonal ice (768 water

molecules) is equilibrated in an NVT simulation at 10 K. The simulation lasts 100 ps, using the Nosé-Hoover thermostat to maintain the temperature. The reason for this is that any bad initial contacts are given the chance to resolve themselves, while the bulk of the ice crystal remains frozen at sub-freezing temperatures. In comparison to a single interface, this set up allows a more realistic environment where the ice seed is surrounded by significant amounts of water. Next the equilibrated ice is solvated so that it is sandwiched between two water layers of similar size (763 molecules each). It is oriented in such a way that the fastest growing ice plane (the 1120 face or the secondary prism plane?) is exposed to the liquid water, creating an interface at either side. This initial configuration is known as a two-phase coexistence or the direct co-existence method because both sides of the ice crystal are in direct contact with the liquid water García Fernández et al. [2006]???. This approach is important as the nucleation of ice is an activated process García Fernández et al. [2006]; ? and often requires a seed, so merely simulating pure liquid water below the freezing temperature usually produces supercooled water instead of ice. This has only once been successfully achieved following a 1 year NVT simulation by Matsumoto *et al* ?. Additionally the presence of this interface also removes the phenomena of bulk solid superheating, which is a erroneous computational occurrence???? and an experimental impossibility?.

The new system contains 2,294 molecules and a second NVT run with 10,000 kJ mol<sup>-1</sup> restraints is placed on the ice. This time the temperature is maintained at 283.15 K. It allows the water to equilibrate around the ice and the same conditions are conducted for a short consecutive NPT run using 10,000 kJ mol<sup>-1</sup> and 100 kJ mol<sup>-1</sup> respectively.

Once the NPT equilibration runs are complete, the restraints are removed and NPT production runs are conducted at different temperatures until melting or freezing is observed. In these simulations a 1 fs timestep was employed in order to accurately capture the shorter time scales reported for melting at high temperatures. During the NPT (equilibration and production) simulations, the box is allowed to fluctuate independently in all directions. This allows the system to freely freeze or melt, and the box shape can adjust accordingly. To accomplish this in GROMACS with Parinello-Rahman barostat, the compressibilities are specified and issued in such a way that it helps to scale the strength (or speed) of the pressure coupling and the box deformation. The compressibility of the X and Y planes are both set to the compressibility of real ice ( $1.0 \times 10^{-5}$  bar<sup>-1</sup>), while the Z direction is perpendicular to the ice/water interface so it remains set to the higher compressibility of real water. The angles are kept orthogonal and unlike the original Weiss simulations, the LJ

potential was truncated using 0.85 nm cut-offs. To compare the effects of cut-offs, this same protocol was also repeated for the 1 nm cut-off.

Prior to evaluating the results of the simulation, it was necessary to demonstrate that the system had reached equilibrium. This process is monitored by the convergence of the total energy. The melting process is identified by an increase in the total energy and freezing by a decrease in the total energy. Once this process is complete, a plateau is observed. Additionally, just before the plateau a characteristic spike to the baseline serves as an initial indicator that a few of the interface layers remain to be processed and so melting or freezing is especially fast. Similar to Weiss *et al* Weiss et al. [2011], we acknowledge that the direct monitoring of the potential energy would also have been sufficient to monitor the melting and freezing of the system, however in order to allow direct comparisons we decided to follow the approach reported by Fernández *et al* García Fernández et al. [2006].

### 3.3 Analysis methods

#### 3.3.1 Hydrogen Bonding

There are several methods for extracting the hydrogen bonding in computational simulations. The degree of hydrogen bonding in water is often estimated by calculating the coordination number of O-H RDFs, or by using an analysis script with a geometric criteria. We chose to use the latter for which GROMACS provides a suitable analysis tool. The criteria for a hydrogen bond is defined as when a donor-acceptor (e.g. oxygen and hydrogen) bond distance is less than 0.35 nm and the O-O bond angles are simultaneously less than 30 degrees. This cut-off was applied although subtle variations are present between models in their minima of the O-O RDFs. We do so in order to remain consistent in our approach for comparing the water models under the same specifications.

#### 3.3.2 Diffusion coefficients

The diffusion coefficients (D) of select molecules were calculated from the Einstein relation which is shown in Eq. 3.1. Calculations were performed on linear fittings to the average mean-squared displacement, MSD, ( $\langle |r_i(t) - r_i(0)|^2 \rangle$ ) at time t. D is readily taken from the gradient of the plot.

$$D = \lim_{t \rightarrow \infty} \frac{\langle |r_i(t) - r_i(0)|^2 \rangle}{6t} \quad (3.1)$$

### 3.3.3 Orientational Order Parameter

An important description of the structure of liquid water and ice is provided by the orientational order parameter,  $Q_4$  which is calculated by?:

$$Q_4 = 1 - \frac{3}{8} \sum_{(j=1)}^3 \sum_{(k=j+1)}^4 \left( \cos \theta_{jik} + \frac{1}{3} \right)^2 \quad (3.2)$$

For each water ( $i$ ), the nearest four oxygen atoms are selected and the angle ( $\theta_{jik}$ ) between three oxygen atoms:  $i$  and two of these neighbouring oxygen atoms,  $j$  and  $k$  are determined. With a criteria of four neighbouring oxygen atoms, there are six possible combinations which involve  $i$  at the center, and these are all summed. The  $3/8$  constant is used to normalise the distribution within the range of 0 to 1. If the average  $Q_4 = 1$ , the arrangement of water molecules is perfectly tetrahedral because the cosine of  $\theta_{jik}$  becomes equivalent to  $-1/3$ . For a random distribution (such as an ideal gas) the average  $Q_4 = 0$ , however limits for individual molecules have previously reported between  $-3$  and  $1$ ?. The same calculation can be performed for varying numbers of neighbours, and the normalisation constant must be adjusted accordingly:  $3/4$  for 3 neighbours ( $Q_3$ ),  $9/4$  for 2 neighbours ( $Q_2$ )

After equilibration the average number of hydrogen bonds, densities and diffusion coefficients for liquid simulations were calculated. In each instance, block averages were obtained by averaging over four blocks of 0.5 ns from the trajectory.

## 3.4 Results

### 3.4.1 All water models

We start by comparing the calculated properties of the different water models in order to confirm the validity of our OPLS-AA system for reproducing well tested properties of some older water models. We also use the same conditions to produce results for the less tested TIP4P/Ice water model. This benchmarks TIP4P/Ice in comparison to other water molecules when all other simulation parameters are kept identical.

The O-O, O-H and H-H radial distribution functions (RDFs) of all the simulated water models and experiment are shown in Fig.3.1. Work from Soper *et al* Soper [2000] was chosen as an experimental source because their spectral data are widely used and referenced in theoretical studies of water structure, particularly the oxygen-oxygen (O-O) RDF?. The reason for this is that it has often been used for

the development and refinement of structural properties of water models. [Cordeiro [1998]]

The most characteristic features of the O-O RDF are the position and heights of the first two peaks. With the exception of TIP3P all models predict the contours of these peaks well, and those of the other two RDFs. For the O-O RDFs, these models have a first peak at 0.275 nm but the height varies from 0.277 (TIP3P) to 0.325 (TIP4P/Ice). The closest model to this first peak is TIP3P while the remaining models overestimate it, particularly TIP4P/Ice. In contrast the second peak is absent in TIP3P and visibly broader at around 0.45 nm for the other water models. Its height varies from 0.110 (TIP4P) to 0.119 (TIP4P/Ice). While the other water models flatten out, both TIP4P/Ice and TIP5P are the most accurate at tailing last few shells more closely, exhibiting the long range ordering shown in the experiment. In this section, all models demonstrate a loss of structure with each successive hydration shell, comparable to theoretical literature [Abascal et al. [2005]; ?]; ?]; ? and experiment. As reported in the original TIP4P/Ice paper, although the model is not intended for the liquid state it is able to adequately reproduce the liquid water structure.

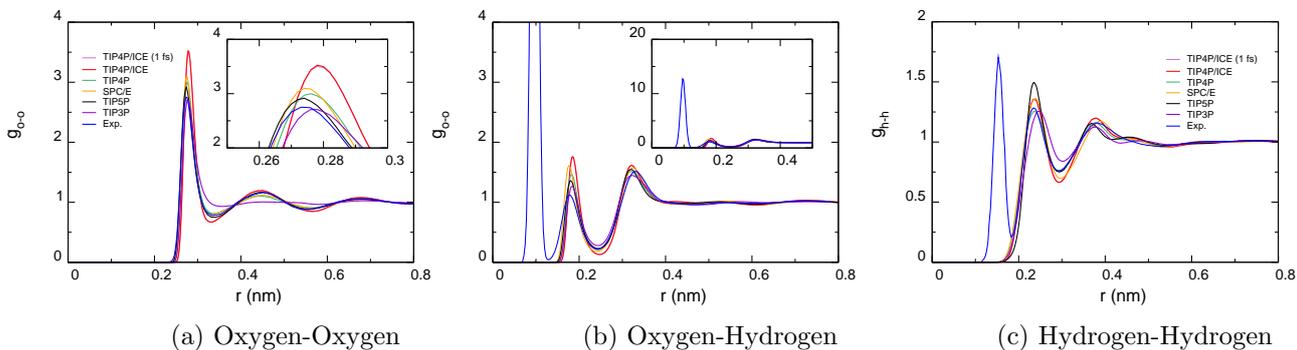


Figure 3.1: RDFs of all water models and experiment [Soper [2000]] at 298 K and 0.85 nm cut-offs. There are no noticeable structural differences between using 1 fs and 2 fs timesteps for the TIP4P/Ice water model. Inset (a) shows and (b) highlight the differences between water models and depict the full RDF profile respectively.

Diffusion coefficients,  $D$ , were calculated from the mean square displacements (MSD) of the whole water molecule using the Einstein relation. Their values are estimated from the gradient of the linear part of the MSD plot ( $D = \text{MSD}/6$ ) and are tabulated for various water models in Table 3.4. Experimentally at 298.15 K the diffusion constant of liquid water is  $2.3 \times 10^{-5} \text{ cm}^2 \text{ s}^{-1}$ . Of the water models

reported here, the models that give diffusion coefficients closest to experiments are SPC/E and TIP5P. The TIP4P/Ice model is approximately half this value, whilst TIP3P and TIP4P both severely overestimate it. The proposed diffusion coefficients of the models are as follow: SPC/E (2.49),???? TIP3P (5.06±0.09),Jorgensen et al. [1983] TIP4P (3.29±0.05)Jorgensen et al. [1983], TIP4P/Ice (1.12±0.05)Weiss et al. [2011] and TIP5P (2.62±0.04)? in units of  $10^{-5} \text{ cm}^2 \text{ s}^{-1}$ .

Property	Model					
	SPC/E	TIP3P	TIP4P	TIP4P/Ice	TIP5P	Expt.
$D, 10^{-5} \text{ cm}^2 \text{ s}^{-1}$	2.63 (0.035)	5.74 (0.100)	3.57 (0.032)	1.19 (0.023)	2.80 (0.089)	2.30 (-)
$\rho, \text{ kg m}^{-3}$	998.56 (0.19)	985.06 (0.07)	993.19 (0.22)	992.20 (0.12)	983.64 (0.37)	997.05 (-)
$N_H$	3.60	3.35	3.55	3.75	3.33	2-4

Table 3.4: Diffusion coefficient ( $D$ ), density ( $\rho$ ), the average number of hydrogen bonds per molecule ( $N_H$ ) and associated standard deviations.

The results for hydrogen bonding are also tabulated in Table 3.4. Studies of hydrogen bonding in different water models have extensively been investigated and it has previously been reported that SPC/E, TIP3P, TIP4P and TIP5P have 2.46–3.59???, 3.5Jorgensen et al. [1983]; ?, 2.36–4.0 ??Cordeiro [1998]; ?; ? and 2.08–3.7????? hydrogen bonds per water molecule respectively. Our results are very similar to theoretical findings and the estimated experimental number broadly ranges from 2–4??? at room temperatures which is suitably approximated by each model, including TIP4P/Ice. To our knowledge, hydrogen bonding for liquid TIP4P/Ice water has not previously been reported. It is apparent that TIP4P/Ice has the largest number of hydrogen bonds per molecule and this accounts for the overestimation of 28% seen in the first shell of its oxygen RDFs. The same rational can be used to explain its significantly slow diffusion coefficients.

The density of water at 298.15 K is  $997.05 \text{ kg m}^{-3}$  Lange and Dean [1973]; ? and this property is well represented by 3 water models: SPC/E, TIP4P and TIP4P/Ice. With the exception of SPC/E all models still provide underestimations, and surprisingly TIP5P provided one of the worst matches to the experimental value despite being parametrized to fit to the density properties of water?. For each models, the densities are widely recounted in literature and are typically quoted within the range reported here.Jorgensen et al. [1983]; ?; ?; ?

A comparison of the orientational order parameter was made for all the water models at 298.15 K. The results are presented in Fig. 3.2 and we can see that TIP4P/Ice and TIP5P, have the highest degree of tetrahedral arrangements with maxima at  $Q_4 = 0.81$ , while the rest have maxima at lower values (0.72 for SPC/E and TIP4P, and 0.48 for TIP3P). This compliments our previous findings (Fig.3.1) that TIP4P/Ice is predisposed to a more well defined tetrahedral arrangement, even as a liquid, while the reverse it true for TIP3P. As expected the  $Q_4$  distribution is bimodal for most models marked by a low  $Q_4$  peak at  $\sim 0.5$  (for liquid-like) and a high  $Q_4$  peak described earlier (for ice-like orientations).? However in our study the shoulder at 0.5 is faint for TIP4P/Ice and near unimodal for TIP3P. These profiles are more characteristic of lower and higher temperatures respectively ?? . Previously correlations between the bimodal features of  $Q_4$  distributions have been drawn to the highly debated polymorphism of water,??? which are typically seen at higher pressures and in supercooled systems, neither of which are the case here. Although these features are unusual, our results for TIP4P/Ice and TIP3P are the only two water models which rightly reflect this at ambient conditions. The high ordering present in TIP4P/Ice is most likely what makes the water model predisposed to a better representation of the phase diagram of water.

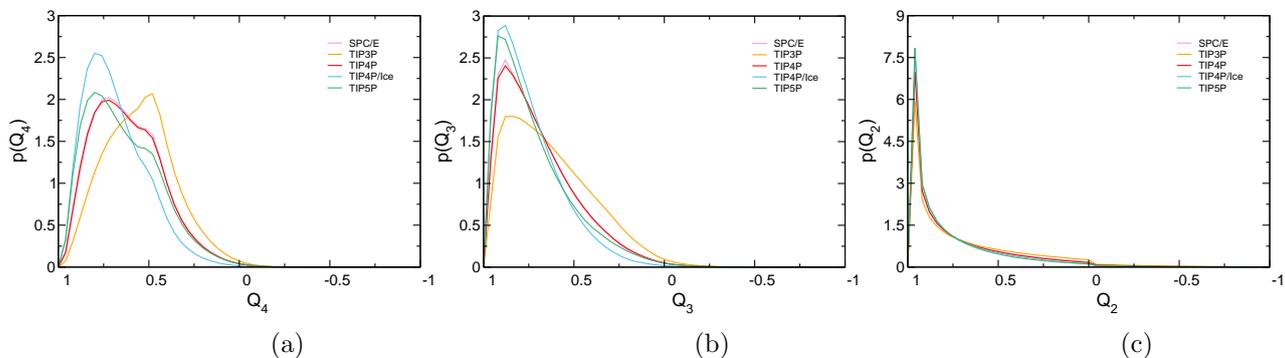


Figure 3.2: Orientational order parameter for all water models considering the (a) 4 nearest neighbours, (b) 3 nearest neighbours, (c) 2 nearest neighbours

In order to make this analysis suitable to other solutions where steric interference from solutes may prevent the approach of four neighbouring water molecules, we include analysis for 3 neighbours, 2 neighbours. Irrespective of the number of neighbours, TIP4P/Ice and TIP5P demonstrate much higher levels of tetrahedrality while the TIP4P and SPC/E water models display common features which almost overlap?. The effect of reducing the number of neighbours means that less waters

are simultaneously involved in attaining the ideal tetrahedral criteria and so we see a greater ability to adopt this orientation, as reflected by the shift from low  $Q$  peaks in  $Q_4$  to higher values in  $Q_2$ . It is also clear that the disparity between the models are drastically reduced with the transition from  $Q_4$  to  $Q_3$  and finally the  $Q_2$  parameter. Under similar conditions fewer neighbours (e.g. the  $Q_2$ ) may be used along side the  $Q_4$  parameter to assess the ordering effects of larger antifreeze molecules which cannot sterically accommodate four neighbouring water molecules. Snapshots of an isolated TIP4P/Ice water molecule is shown at 298 K with its neighbours in order to illustrate what the  $Q_4=1$  and  $Q_4=0.5$  shoulder correspond to in solution (Fig. 3.3).

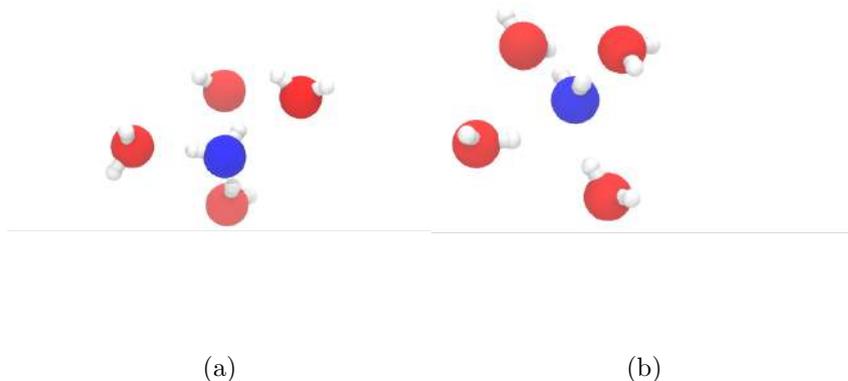


Figure 3.3: TIP4P/Ice water oxygens which exhibit  $Q_4$  values (a) at 0.5 and (b) at 0.81 are shown in blue, and their 4 closest neighbours at 298 K

### 3.4.2 Effects of temperatures on TIP4P and TIP4P/Ice

Here we investigate the effects of temperature on the structure and dynamics of the TIP4P/Ice water model under the OPLS-AA forcefield. The results will be used to assess the credibility of mixing forcefields under non-ambient conditions and the implications of varying cut-offs will be addressed in a later section. From here on, simulations in solution also focuses on TIP4P, a similar water model which the OPLS-AA forcefield was parametrized to work with.

All RDFs for TIP4P and TIP4P/Ice show that short and long range ordering is lost with increasing temperatures because more kinetic energy is available for water molecules to break out of their confined shells. As a result, the height and sharpness of each peak decreases with higher temperatures and the troughs become increasingly shallow. Similar observations have been reported for TIP4P and other

water models in independent studies??.

A 113 K increase in temperature from 235 K to 348 K corresponds to a 36 % drop in the first peak of the TIP4P/Ice O-O RDF. For TIP4P this value is much smaller at 31 %, and is to be expected since TIP4P/Ice exhibits the strongest peaks and troughs between the two (and in comparison to the other water models). At 235 K the faint but visible residual ordering can still be seen up 0.85 nm which is absent from the two higher temperatures. These stop at  $\sim 0.68$  nm and are shown in Fig. 3.5 and Fig. 3.4. The same qualitative trend is visible for the O-H and H-H RDFs and no significant shifts in any RDF peaks are apparent.

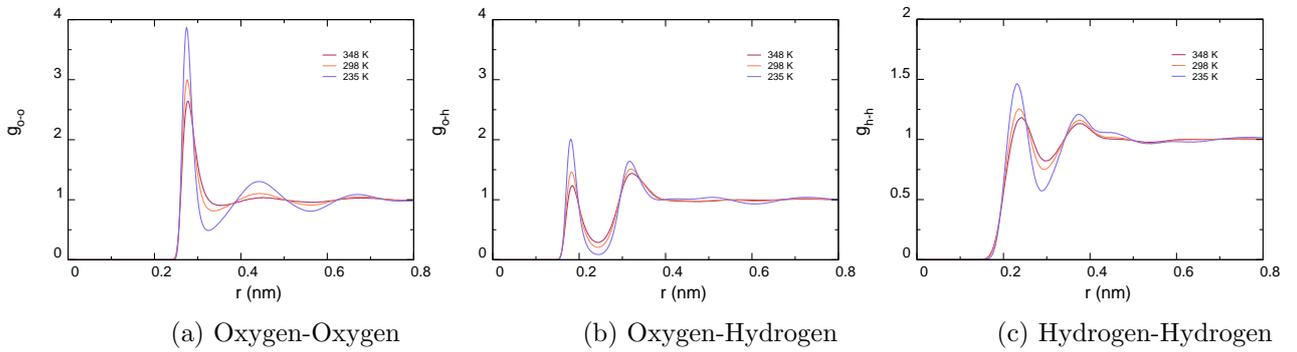


Figure 3.4: RDFs of the TIP4P water at 3 different temperatures and 0.85 nm cut-offs

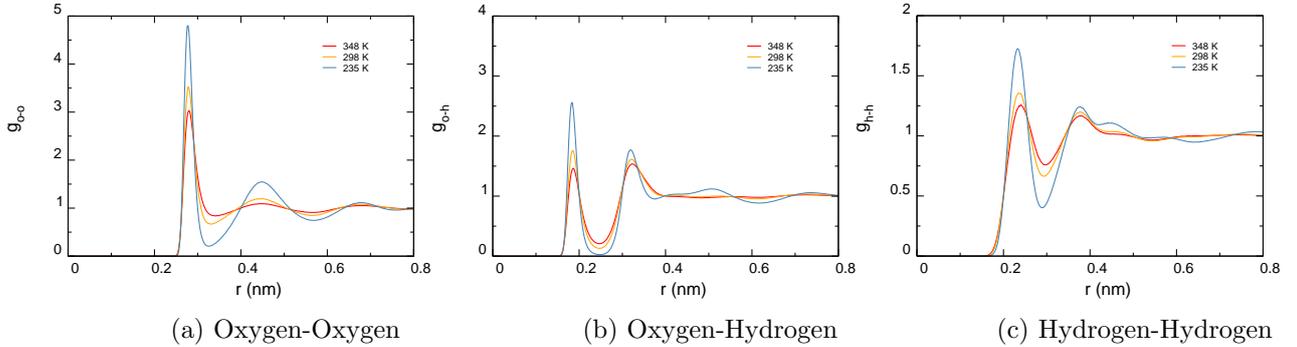


Figure 3.5: RDFs of the TIP4P/Ice water at 3 different temperatures and 0.85 nm cut-offs

The effects of temperature and cut-offs on the diffusion coefficients are shown in Table 3.5. At varying temperatures it is important that TIP4P/Ice diffusion coefficients can be reproduced close to experimental values seeing as these values have large impacts on the rate of ice growth at ice/water interface. Experimental diffusion

coefficients were taken from various reports?????? and an average was calculated for each temperature. With the exception of 235 K, we found that TIP4P/Ice consistently has half the diffusion coefficient values expected experimentally. Although the models cannot reproduce the exact diffusion constants, we are able to demonstrate that they relay the same qualitative trends seen experimentally. Moreover similar values have previously been calculated?. Overall TIP4P/Ice produces coefficients closest to the experimental values at lower temperatures (235 K and 298 K), while results for the TIP4P water model is the most suitable coefficient at 348 K.

<b>T (K)</b>	<b>Cut-off (nm)</b>	<b>TIP3P</b>	<b>TIP4P</b>	<b>TIP4P/Ice</b>	<b>Expt.</b>
235	0.85	1.58 (0.037)	0.46 (0.021)	0.015 (0.002)	0.187 (-)
235	1.00	1.63 (0.058)	0.45 (0.011)	0.015 (0.002)	
235	1.50	1.63 (0.0467)	0.45 (0.012)	0.016 (0.002)	
298	0.85	5.74 (0.100)	3.57 (0.032)	1.19 (0.023)	2.395 (-)
298	1.00	5.69 (0.109)	3.63 (0.052)	1.16 (0.014)	
298	1.50	5.66 (0.089)	3.63 (0.017)	1.20 (0.024)	
348	0.85	10.40 (0.097)	8.12 (0.197)	3.64 (0.070)	6.254 (-)
348	1.00	10.26 (0.186)	7.87 (0.083)	3.57 (0.054)	
348	1.50	10.35 (0.068)	8.07 (0.220)	3.55 (0.047)	

Table 3.5: Diffusion coefficient,  $D$  ( $10^{-5} \text{ cm}^2 \text{ s}^{-1}$ ) for various water models at 3 different temperatures and cut-offs.

In Table 3.6 we present the hydrogen bonding per molecule as a function of temperature for 3 water models. As expected we found that the hydrogen bonding network was promoted with lower temperatures, similar to observations of the RDFs. Such observations are not unfamiliar? because higher temperatures the water molecules are able to move more freely and so disrupt the local hydrogen bonding network of the system.

T (K)	Cut-off (nm)	TIP3P	TIP4P	TIP4P/Ice
235	0.85	3.67	3.84	3.96
235	1.00	3.67	3.84	3.96
235	1.50	3.67	3.84	3.96
298	0.85	3.35	3.55	3.75
298	1.00	3.35	3.55	3.76
298	1.50	3.35	3.55	3.76
348	0.85	3.08	3.29	3.56
348	1.00	3.08	3.29	3.56
348	1.50	3.08	3.29	3.56

Table 3.6: Average number hydrogen bonds per molecule ( $N_H$ ), for various water models at 3 different temperatures and cut-offs .

The densities of 3 water models are tabulated in Table 3.7. Each water model’s density increases with the temperature in line with experiment. At 298 K and 348 K the density of water is  $997.05 \text{ kg m}^{-3}$  Lange and Dean [1973]; ? and  $974.90 \text{ kg m}^{-3}$  ? respectively. At 235 K the simulated liquid water is sustained within a supercooled state which can only be achieved experimentally in pure water and in the absence of nucleation sites. Under experimental conditions this is difficult to achieve and so no comparisons are made to experimental densities at this temperature. The decline in density between 298 K and 348 K corresponds to a 2.2% drop. In the order of decreasing accuracy, the same calculation for TIP4P/Ice yields 1.21%, 3.80% for TIP4P and 5.07% for TIP3P. It is also known that the density of water increases non-monotonically with temperature, however this is not illustrated as we only investigate 3 temperatures ??

<b>T (K)</b>	<b>Cut-off (nm)</b>	<b>TIP3P</b>	<b>TIP4P</b>	<b>TIP4P/Ice</b>
235	0.85	1029.36 (0.36)	1001.99 (0.12)	946.65 (1.77)
235	1.00	1029.77 (0.54)	1002.72 (1.05)	945.91 (2.55)
235	1.50	1029.78 (0.17)	1002.29 (1.00)	945.71 (2.85)
298	0.85	985.06 (0.07)	993.19 (0.22)	992.20 (0.40)
298	1.00	986.13 (0.35)	994.34 (0.15)	992.73 (0.31)
298	1.50	986.22 (0.35)	994.18 (0.26)	992.64 (0.12)
348	0.85	935.16 (0.40)	955.45 (0.27)	980.14 (0.26)
348	1.00	936.79 (0.12)	956.56 (0.19)	981.18 (0.37)
348	1.50	937.26 (0.28)	956.41 (0.46)	981.13 (0.17)

Table 3.7: Densities for various water models at 3 different temperatures and cut-offs .

Figure 3.6 depict the distribution of  $Q_4$  order parameters as a function of temperature and different cut-offs. As the temperature changed, the distributions for the orientational order showed similar traits for all the water models. From low to high temperatures a shift in  $Q_4$  from high to low values indicates a clear decrease in tetrahedral order. The highest peaks can be found for the 235 K system at 0.72, 0.84 and 0.88 for TIP3P, TIP4P and TIP4P/Ice respectively. In the case of the 4-point water models, the increase in temperature is earmarked by a transition from a unimodal distribution to a bimodal peak which suggests 2 dominant orientations of waters. The presence of the strengthening second shoulder is caused by a transition between tetrahedral and non-tetrahedral environments. The  $Q_4$  peaks of 0.48 increased and an intermediary maxima is formed with an increase in the temperature, while the maxima at 0.84 or 0.88 vanishes. The shoulder peaks for TIP4P and TIP4P/Ice are found at 0.72 and 0.8 respectively while for TIP3P this transition is much sharper and non-existent before 0.48. It's absence and presence of a strong shoulder even at low temperatures indicates that the model is bias towards random arrangements while the reverse could be said for TIP4P/Ice which boasts of a higher degree of tetrahedral arrangements irrespective of temperature. The point at which these three temperatures cross indicates a cut-off value by which to identify ice-like water molecules. This point was most easily identified for TIP3P and TIP4P/Ice as 0.59 and 0.76 because the three distributions intersect at around the same point. For TIP4P we identified this cut-off at 0.68, however more temperature profiles may be required for confirmation.

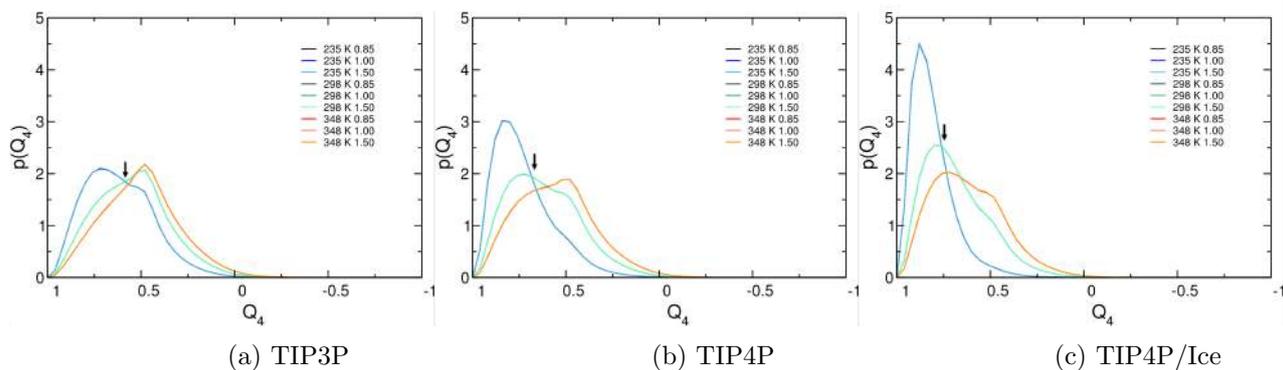


Figure 3.6: Orientational order parameter for TIP3P, TIP4P and TIP4P/ice at 3 different temperatures and cut-offs. Black arrows mark intersections that divide ice-like and fluid regions in each water model. Each cut-off finely overlaps.

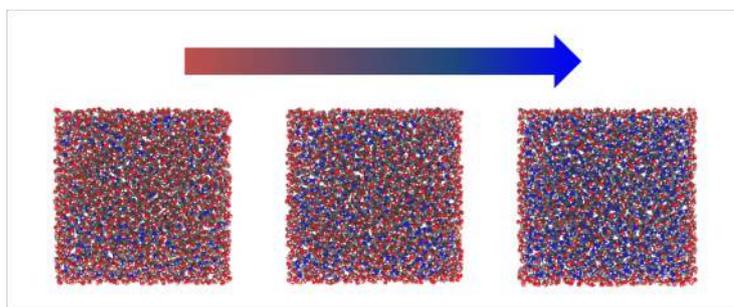


Figure 3.7: Snapshots of the TIP4P/Ice water at the end of 3 ns MD simulation at 3 different temperatures: 348 K 298 K and 235 K (from left to right). The water oxygens with  $Q_4 > 0.75$  are coloured blue and represents high tetrahedrality, while the others are coloured red. Hydrogen atoms are coloured white.

### 3.4.3 The effects of dilute salt concentrations on TIP4P/Ice and TIP4P

Often times simulations are conducted with with small solutes like salts in order to mimic biological systems or saline conditions. In order to investigate the effects of small charged particles on the mixing of the OPLS-AA forcefield with TIP4P/Ice we introduce sodium and chlorine into the system. The RDFs of the saline solutions are summarised in Fig. 3.8–3.11. In this section we primarily focus on systems with 0.85 nm cut-offs (although most observations are also applicable to 1.00 nm cut-offs). From here on we chose to limit the studies of TIP4P/Ice with additives to just these two cut-offs because future work will carry forward from these studies

as they are the less computationally expensive options. Consistent with the pure water systems, the RDFs for TIP4P/Ice reveal a more structured solution than for TIP4P. The models' salt-water RDFs, illustrates that increasing the salt concentration induces a subtle decreases in the intensities of the first and second shells. The position of the peaks in all water models are in good agreement with experimental data for these RDFs. For instance peaks are found at 0.266–0.244 nm for sodium-oxygen (Na-O) RDFs. The main peaks for the chlorine-oxygen (Cl-O) RDFs and the chlorine-hydrogen (Cl-H) RDFs are also visible at 0.29–0.327 nm and 0.225 nm respectively.

Noticeably with higher concentrations there isn't a significant shift in peaks, but rather a change in the heights which reflects that the introduction of such small salt concentrations doesn't have a marked affect in the water ordering. This is expected because there are sufficient water molecules available to solvate the hydration shell of each ion. As a result neighbouring salt molecules do not particularly compete for solvation and so the distribution of their respective shells (and that of the waters) are not unusually disturbed.

The results of Na-Cl RDFs are also included in the same figures and their noisy outlines confirms that the concentration of the solutions are very small. In comparison to the water-water RDFs, there are far fewer salt ions available for the calculation and the averaging of the distribution plot. These traits and the following are apparent for both water models. The Na-Cl RDFs have 2 clear peaks; the first is 0.28 nm which corresponds to the distance between the ionic pair, while the second peak at 0.45 nm corresponds to an ion pair separated by solvents. These results also agree with previous reported observation in similar simulations.

It is important to note that the Na-Cl RDF profiles are markedly taller for the TIP4P water model at 0.85 nm. This feature can be attributed to the larger negative charges used in the TIP4P/Ice water model so results in stronger interactions with the ions instead. The ion-water RDFs are good measures to describe how strongly bound water molecules are to the ion and can be used to calculate the coordination numbers. The calculated coordination numbers for our simulations are displayed in Table. 3.8. For sodium we deduced values from the Na-O RDFs and we obtain values between 5 and 6 from integrating the area under the first peak (from zero to the first minima). The position of the first minimum was taken at around 0.33 nm and 0.39 nm for Na-O and Cl-O respectively for both TIP4P and TIP4P/Ice. Experimental hydration numbers typically vary between 4–6 for  $\text{Na}^+$  but most simulations frequently predict a value of 6. While for  $\text{Cl}^-$  an estimated average of 6 to water molecules are found in the first hydration

shell (also consistent with our findings).

It is important to note that the range for accepted coordination numbers vary due to the use of different experimental techniques and at times varying concentrations. In spite of this leeway the TIP4P system consistently overestimates the coordination numbers for  $\text{Cl}^-$  anions, while predictions for TIP4P/Ice constantly fall within the accepted experimental range. We find, in agreement to the theoretical studies?, both models have coordination numbers which decline with increasing concentrations. This present work does not reveal stark contrasts between the structuring of TIP4P or TIP4P/Ice water in saline solutions, aside from the aforementioned and an artefact induced by changing the cut-offs which is discussed much later. We conclude that TIP4P/Ice water model performs just as good as TIP4P, if not better in saline concentrations at 0.85 nm.

Conc. (M)	Cut-off	TIP4P	TIP4P/Ice
<b>Na<sup>+</sup> ions</b>			
0.10	0.85	5.75 (0.112)	5.84 (0.084)
0.10	1.00	5.77 (0.094)	5.63 (0.144)
0.14	0.85	5.65 (0.064)	5.82 (0.063)
0.14	1.00	5.63 (0.195)	5.77 (0.086)
0.21	0.85	5.52 (0.178)	5.72 (0.079)
0.21	1.00	5.65 (0.039)	5.82 (0.055)
0.28	0.85	5.48 (0.105)	5.73 (0.061)
0.28	1.00	5.52 (0.046)	5.63 (0.062)
0.35	0.85	5.40 (0.065)	5.60 (0.076)
0.35	1.00	5.48 (0.043)	5.64 (0.075)
0.43	0.85	5.38 (0.107)	5.60 (0.073)
0.43	1.00	5.32 (0.036)	5.69 (0.070)
<b>Cl<sup>-</sup> ions</b>			
0.10	0.85	7.39 (0.029)	6.86 (0.024)
0.10	1.00	7.37 (0.038)	6.80 (0.046)
0.14	0.85	7.34 (0.025)	6.83 (0.023)
0.14	1.00	7.34 (0.066)	6.86 (0.035)
0.21	0.85	7.30 (0.068)	6.84 (0.036)
0.21	1.00	7.34 (0.025)	6.86 (0.016)
0.28	0.85	7.30 (0.040)	6.84 (0.025)
0.28	1.00	7.31 (0.019)	6.83 (0.025)
0.35	0.85	7.24 (0.026)	6.83 (0.033)
0.35	1.00	7.31 (0.017)	6.84 (0.025)
0.43	0.85	7.23 (0.089)	6.82 (0.039)
0.43	1.00	7.25 (0.020)	6.82 (0.039)

Table 3.8: The coordination number of Na<sup>+</sup> and Cl<sup>-</sup> ions in the saline solutions.

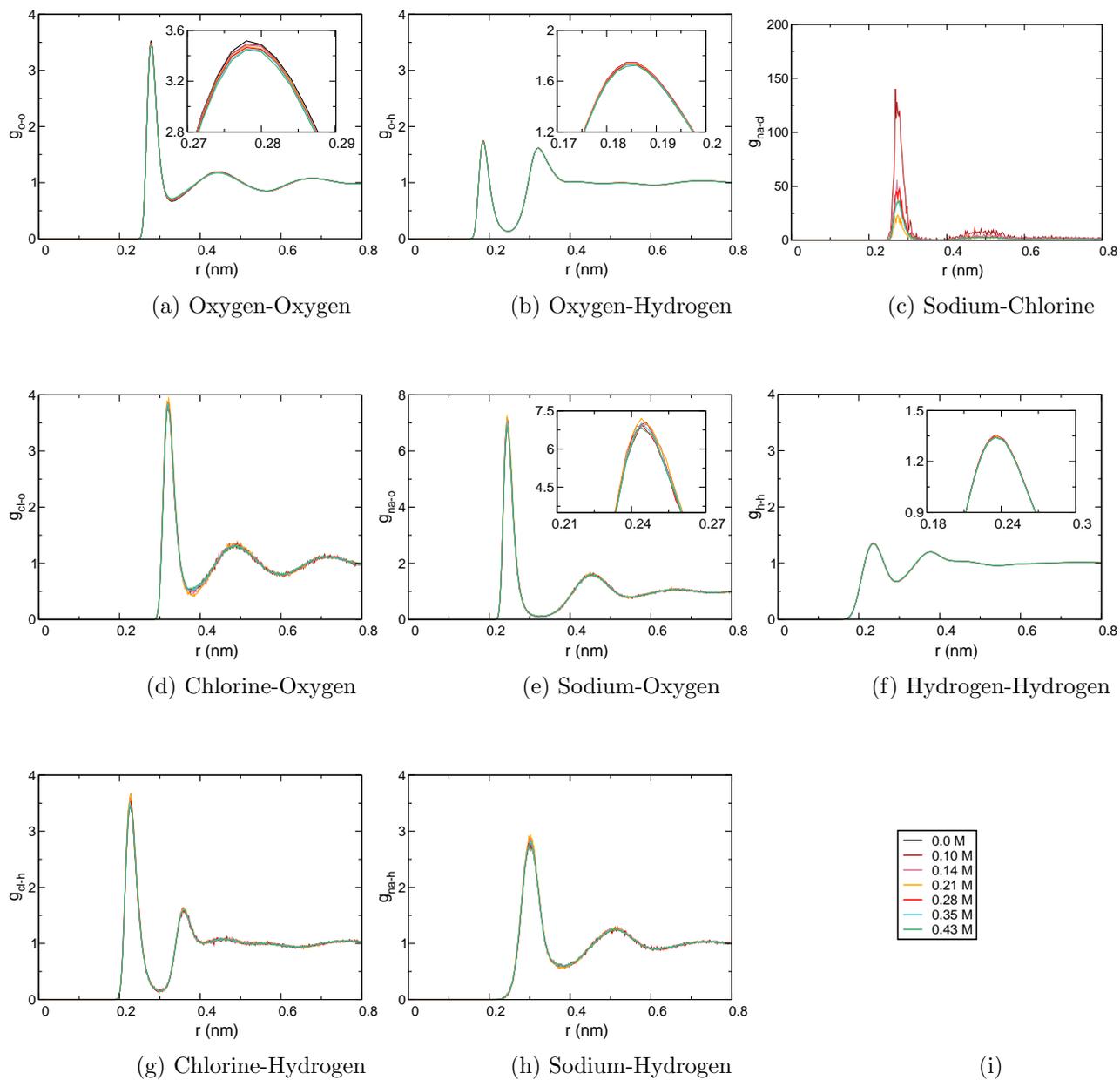


Figure 3.8: RDFs for different salt concentrations at 298 K with the TIP4P/Ice water model at 1 nm

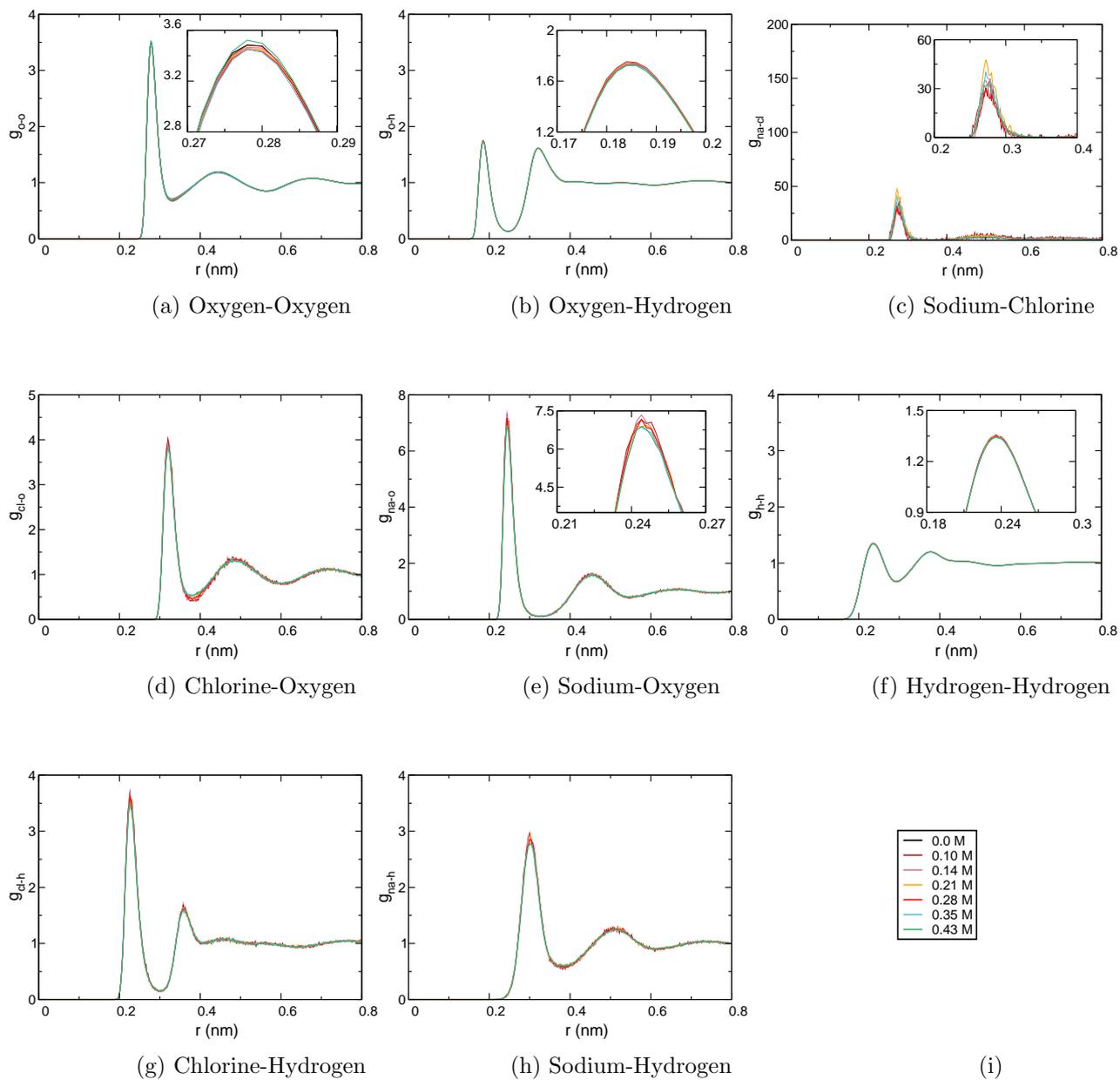


Figure 3.9: RDFs for different salt concentrations at 298 K with the TIP4P/Ice water model at 0.85 nm

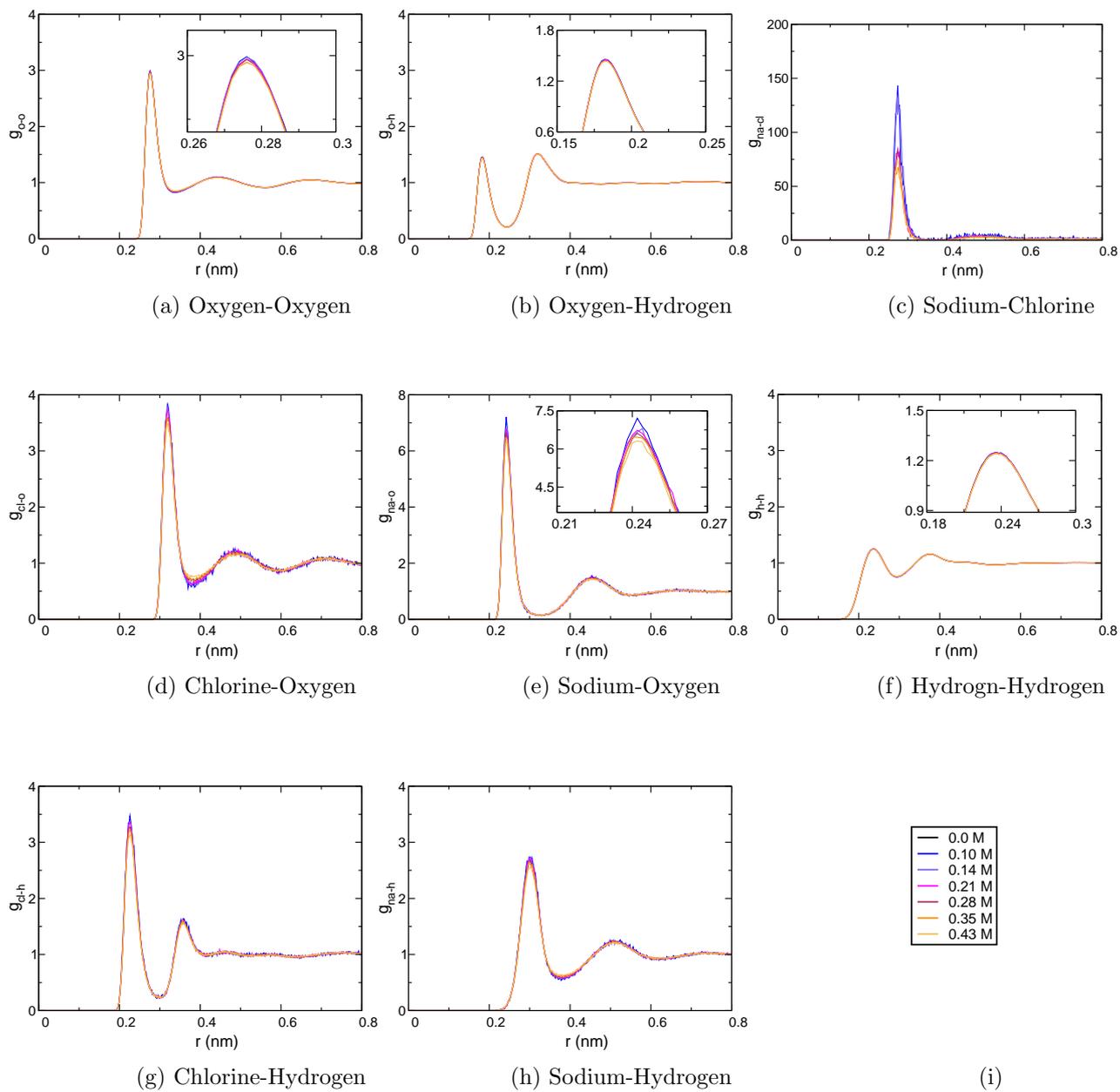


Figure 3.10: RDFs for different salt concentrations at 298 K with the TIP4P water model (1.00 nm cut-off)

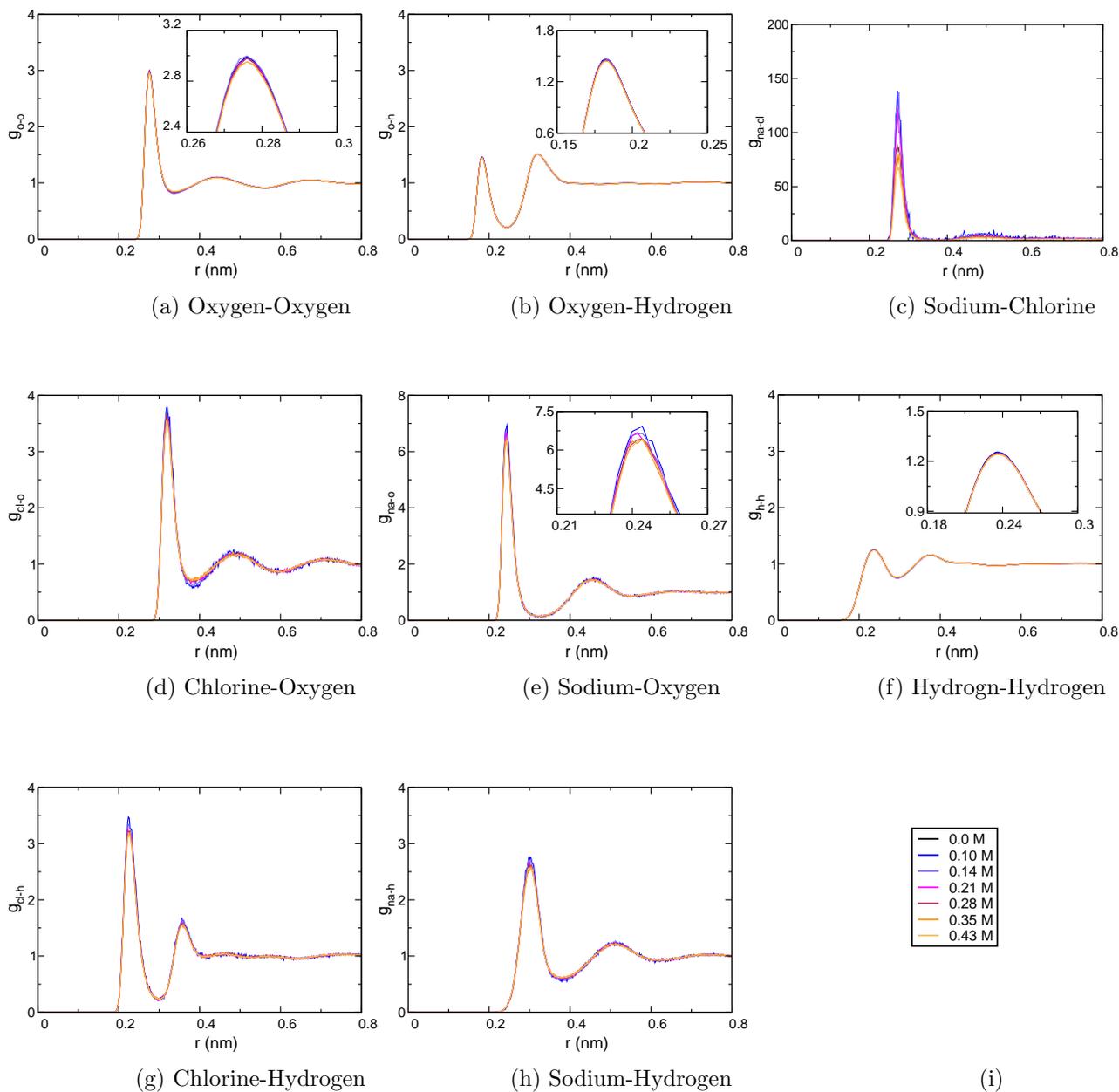


Figure 3.11: RDFs for different salt concentrations at 298 K with the TIP4P water model (0.85 nm cut-off)

The diffusion coefficients for the saline solutions are depicted in Fig. 3.12. In these dilute solutions as the concentrations increase we see fluctuations of the diffusion coefficients, however on average the values do not substantially change past 0.21 M for the TIP4P/Ice water models. The same cannot be said for the TIP4P water model, where the diffusion coefficients for both ions of the 0.85 nm simula-

tion appears to increase. At 298 K, NaCl is considered a “structure making” salt because it interacts strongly with water molecules and helps to maintain or build the hydrogen bonding network of water, reducing their mobility ?. It is therefore expected that the diffusion coefficients of constituents would continue to decrease with increase concentration however previous work have also shown that at dilute concentrations the diffusion constants do not follow this trend perfectly. ? As with the pure water systems, the diffusion coefficients of the TIP4P system is unsurprisingly higher than that of the TIP4P/Ice. Recent reports by Kim *et al* ? have found that the self-diffusion of TIP4P, TIP3P and SPC/E water in 1 M–4 M solutions can be suppressed around NaCl relative to that of neat water (i.e.  $D/D_{H_2O} < 1$ ). Extrapolation of their results indicate that at 0.5 M,  $D/D_{H_2O} \approx 0.975$  in experiments and  $\sim 0.95$  in TIP4P water. Here we show that our simulations successfully reproduce the qualitative trends observed experimentally using the TIP4P systems (with  $D/D_{H_2O}=0.95$  at both cut-offs). Although the same could not be said for TIP4P/Ice which remained fairly constant,  $D/D_{H_2O} = 0.93$  and  $0.98$  at  $0.85$  nm and  $1.00$  nm cut-offs respectively, which is still in line with relative reductions in the diffusion coefficient.

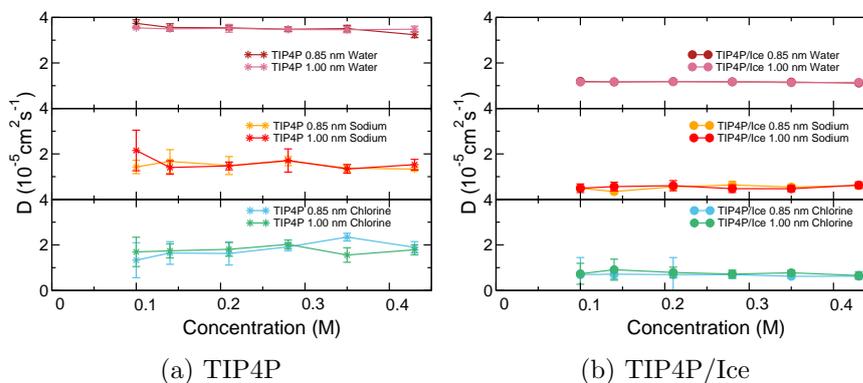


Figure 3.12: The self-diffusion coefficients for different salt concentrations at 298 K and a 1.00 nm cut-off.

The number of hydrogen bonds per water molecule is calculated for the saline solutions in Table 3.9. It shows a small decline in hydrogen bonds with increasing concentrations, which is a consequence of fewer water molecules and increasingly frequent occupations in salt-water interactions (see RDFs). For TIP4P and TIP4P/Ice water models a drop of 0.56 and 0.53% are seen in the hydrogen bonds from pure water to 0.1M respectively.

<b>Conc. (M)</b>	<b>Cut-off</b>	<b>TIP4P</b>	<b>TIP4P/Ice</b>
0.10	0.85	3.53	3.73
0.10	1.00	3.53	3.73
0.14	0.85	3.52	3.72
0.14	1.00	3.52	3.72
0.21	0.85	3.50	3.71
0.21	1.00	3.50	3.71
0.28	0.85	3.49	3.69
0.28	1.00	3.49	3.69
0.35	0.85	3.47	3.67
0.35	1.00	3.47	3.67
0.43	0.85	3.46	3.66
0.43	1.00	3.46	3.66

Table 3.9: The number of hydrogen bonds per water molecule ( $N_H$ ) in the saline solutions.

Fig. 3.13 shows the  $Q_4$  distributions for saline solutions. In comparison to the pure water systems and the effects of changing temperature, there are no substantial differences between the water ordering and the higher concentrations of salt for TIP4P. For the most part this is also the case for TIP4P/Ice water model. However at 0.43 M we begin to see a noticeable decline in the tetrahedral arrangement which indicates a higher model sensitivity with the addition of small molecules. Although this disruption can be attributed to the tetrahedral bias of TIP4P/Ice, the model presents a 8% drop in the ice-like region of water in 0.43 M compared to its pure systems, while TIP4P only declines by 6%.

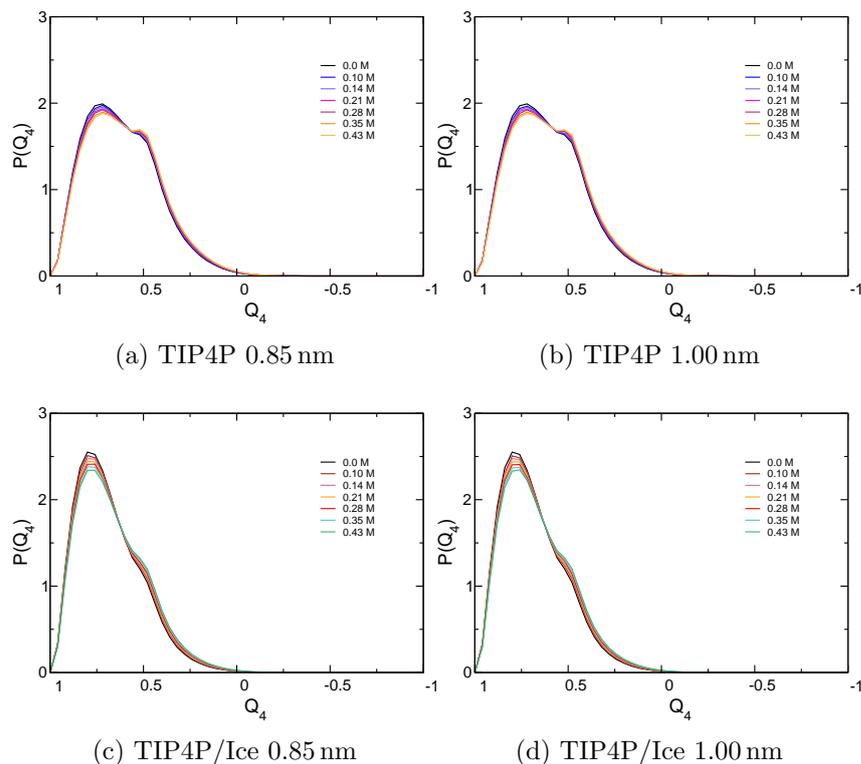


Figure 3.13: Orientational order parameter for TIP4P and TIP4P/ice at 298 K, from low salt concentrations (dark colours) to high concentrations (light colours).

### 3.4.4 The effect of ethanol concentrations on TIP4P/Ice and TIP4P

In this section we continue to assess the compatibility of the TIP4P/Ice water model with the OPLS-AA ff. using ethanol, which is an organic molecule that the OPLS-AA ff. has been parametrized to work with, unlike the smaller NaCl ions. The ethanol molecules are of special interest to us because its basic chemical structure closely resembles the PVA oligomers, which we wish to study in the next chapter. The RDFs for various ethanol concentrations are shown for TIP4P and TIP4P/Ice in Fig. 3.14–3.15 and the atomic labels for ethanol are subscripted with a capital  $E$ . For instance ethanol oxygen atoms are assigned  $O_E$ . In this report of TIP4P/Ice and TIP4P ethanol-water mixtures, the local structure appears similar to experimental and computational results Jorgensen et al. [1996]. For example, the expected peaks of the  $O_E$ - $H_E$  RDFs are confirmed at 0.18–0.19 nm and 0.34 nm. While  $O_E$ -H and  $H_E$ -O have characteristic peaks between 0.17–0.18 nm and 0.3–0.33 nm. As with previous simulations, the systems with TIP4P/Ice water consistently display much higher local short range ordering. Otherwise no appreciable difference is found

between the structures of the two water models. In both systems, the amplitudes of the first peaks grows visibly larger than that of the corresponding peaks in pure water as the ethanol concentration increases, however the position of the first peak doesn't move. The same could be said for all RDFs with second shoulders, except (O-O, O<sub>E</sub>-O). In these exceptions the heights of the second peak is largely conserved until high (greater than 90% ethanol) concentrations are attained, causing a shift in the neighbouring shells to the right. Overall, these observations suggest that the addition of ethanol extensively enhances hydrogen bonding network of the water molecules surrounding it, which is confirmed later.

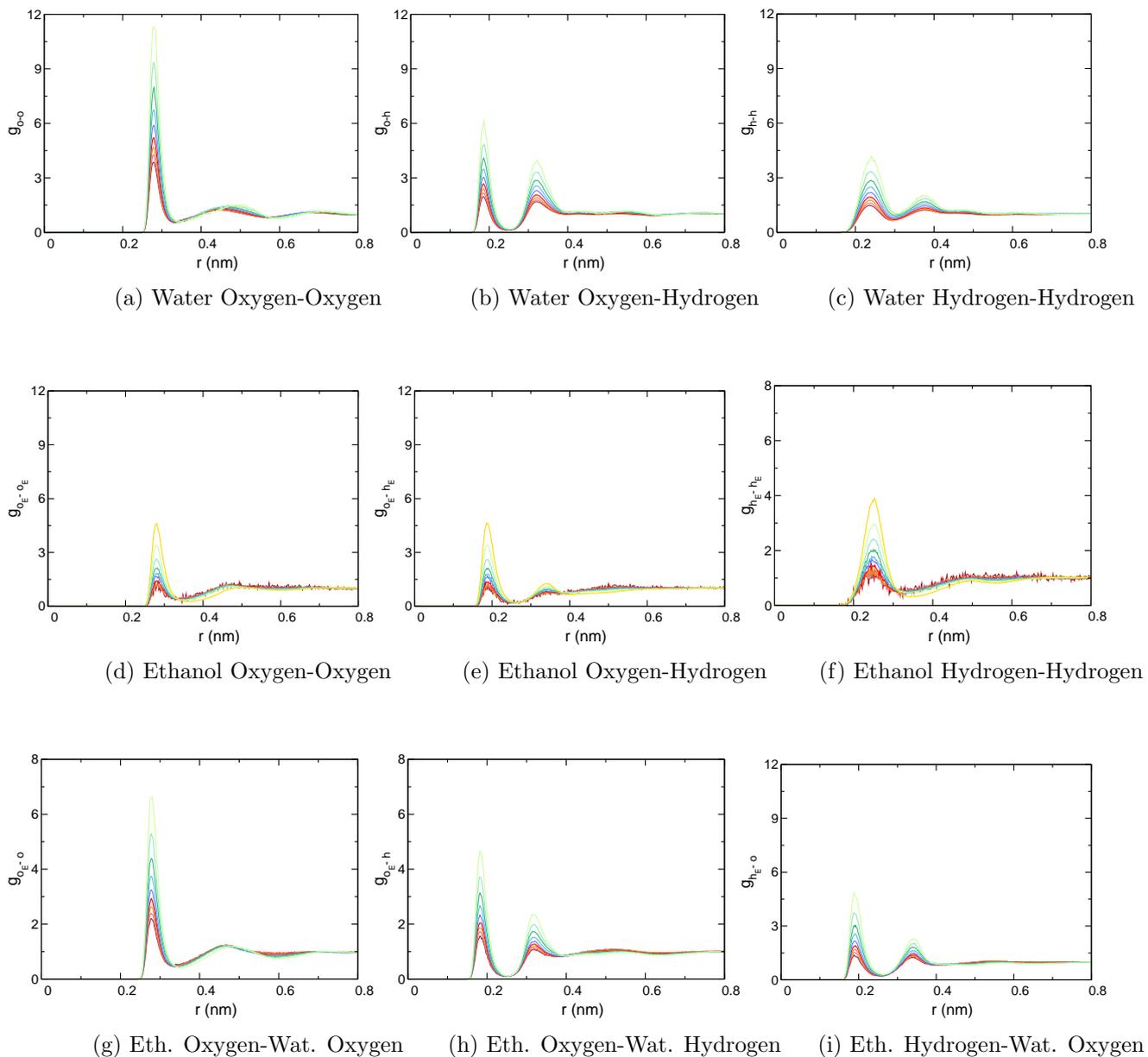


Figure 3.14: RDFs for different ethanol concentrations (10-100%) at 298 K with the TIP4P/Ice water model at 1.00 nm cut-off. Transition from dark red to blue, green and then yellow reflect low to high concentrations

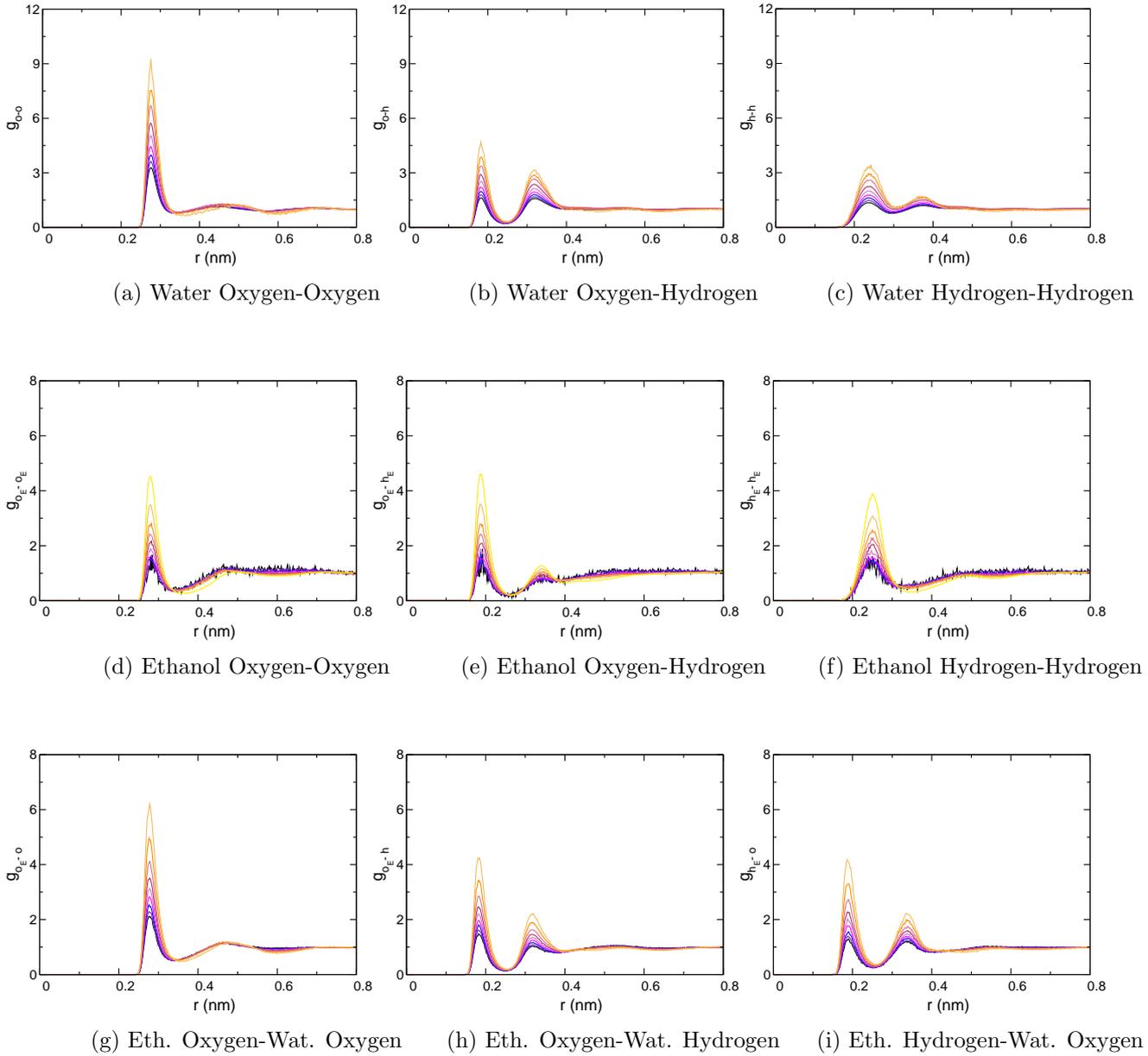


Figure 3.15: RDFs for different ethanol concentrations (10-100%) at 298 K with the TIP4P water model at 1.00 nm. Transition from purple, orange and then yellow reflect low to high concentrations

A wide range of diffusion coefficients were also obtained for an assortment of ethanol concentrations. Figure 3.16 shows the self diffusion coefficients for both water and for ethanol. We see that the model chosen to represent ethanol in these simulations do well to this respect and for both TIP4P and TIP4P/Ice systems the

coefficients are fairly close to experimental values. The ethanol diffusion coefficient for the TIP4P water systems is closest to the experimental values at almost every concentration however it fails to predict a similar contour and transition seen between concentrations. In contrast however, TIP4P/Ice system mimics the contours of the experimental diffusion rates for both water and ethanol, although it remains somewhat shifted. TIP4P does not provide an accurate prediction of the experimental profile of water’s self diffusion in the mixture. At best TIP4P/Ice is able to reproduce the experimental minimum close to 70% which is not achieved with TIP4P in our simulations nor in the cited literature???. Wensink *et al* began with several different parameters from our own systems. They used a Berendsen thermostat, a 1.1 nm cut-off and the SHAKE constraints algorithm but still wound up with the same results, which strongly suggests that these trends are not artefacts. We can therefore conclude that combining the TIP4P/Ice and OPLS-AA forcefields may present a better choice than a combination with the TIP4P water model.

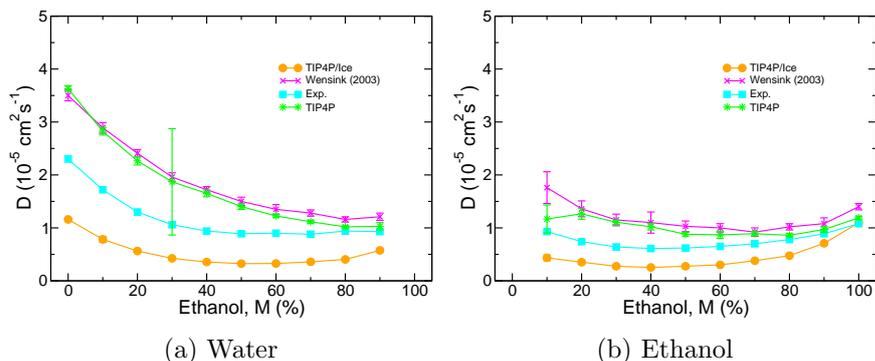


Figure 3.16: The self-diffusion coefficients for (a) water and (b) ethanol in ethanol/water mixtures at 298 K and a 1.00 nm cut-off. Results from Wensink *et al* ? (TIP4P model) and experimental values??? are also included.

Figure 3.17 reveals the results for  $N_H$  in ethanol-water binary mixtures. For both water and ethanol, these plots highlights the total  $N_H$  that the species participate in. The results are normalised for the number of species present, and so the units are effectively number of hydrogen bonds per species. Contributions from species-species and water-ethanol interactions are also shown.

As the ethanol concentrations rise, we observed an increase in the total  $N_H$  per water ( $N_{HW}$ ) and a decrease is seen in the total  $N_H$  per ethanol ( $N_{HE}$ ). In pure ethanol, the average  $N_{HE}$  has been reported as 1.85–1.92 and our solutions produce 1.83, which is in close agreement. ???? A simple calculation is made in order to determine an ideal average  $N_H$  per molecule ( $N_{H(W+E)}$ ) for the mixture at 70%.

Consider that pure liquid TIP4P/Ice water has a value of 3.76  $N_{HW}$ ,  $N_{H(W+E)}$  would therefore yield a value of 2.41 (i.e.  $(0.7 \times 1.83) + (0.3 \times 3.76)$ ). The same calculation for TIP4P yields a lower average of 2.35. At 70% the actual  $N_{HW}$  for TIP4P/Ice was 4.81 and  $N_{HE}$  was 3.49 yielding a higher average than expected at 3.88.

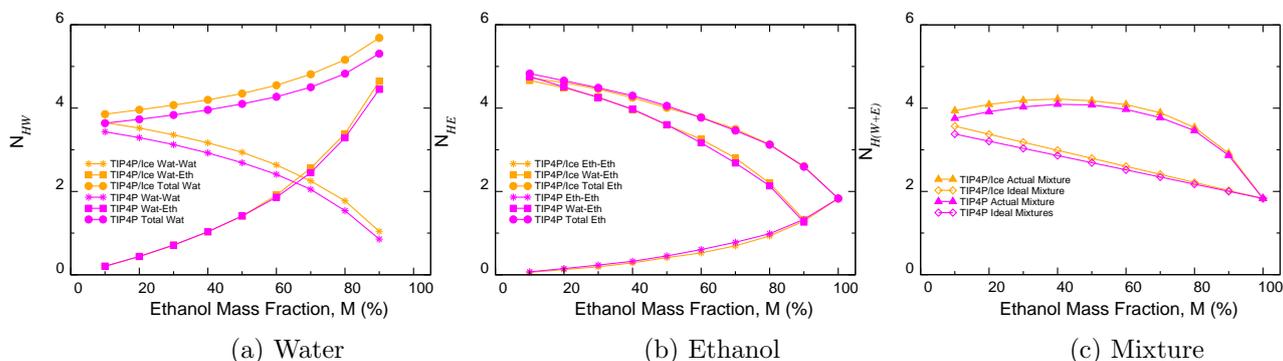


Figure 3.17: The number of hydrogen bonds, ( $N_H$ ) (a) per water, (b) per ethanol and (c) per molecule in the ethanol/water simulations.

As a whole, Fig. 3.17c demonstrates that in the mixtures, the average hydrogen bonds per molecule grow with concentration reaching a maximum of  $\sim 4$  at 60%, and then decreases to  $\sim 2$ . The maximum observed for TIP4P and TIP4P/Ice are 4.09 and 4.22 respectively. These maxima are achieved at a much lower concentration than the 70% experimental benchmark however the profile is consistent with observations in numerous other computational studies<sup>???</sup> and experimental investigations<sup>???</sup> on the hydrogen bonding in alcohol-water binary mixtures. In the region of 0–20% the hydrogen bonding structure of water is reinforced by the addition of small amounts of ethanol, induced by the formation of more water-ethanol hydrogen bonds. While in the range of 70% and onwards the water structures is disrupted by the increasing alcohol concentrations. It is important to note that significant variations in these profiles exist depending on the methods used to investigate the hydrogen bonding<sup>???</sup>.

Additionally, other binary ethanol-water studies report that at low alcohol concentrations, the binary mixture predominantly have 3 hydrogen bonds per molecule, whereas in the alcohol rich mixtures it is reduced to 2. Unlike the experiments our results proposes a higher average value at low concentrations which is further from the predictions of ideal mixing<sup>??</sup>. This is to be expected as both water models overestimate the hydrogen bonding of water, however this notion is validated by our hydrogen bonding and ignites an interest in the water ordering.

Per ethanol molecule, as the concentration increases we can see that the number of water-ethanol interactions decline due fewer water molecules present and as a result of the sterics associated with adding more ethanol molecules. Unsurprisingly, the resulting sterics shield the centring ethanol from surrounding water molecules and so the ethanol-ethanol interaction gradually increases as reflected by the growing hydrogen bonds. At most concentrations the water-ethanol interaction dominates ethanol's total hydrogen bonding per molecule, particularly at lower concentrations. Similar trends have been seen before for methanol, ethylene glycol and glycine in TIP3P water under a different forcefield.<sup>?</sup> These observations are also corroborated by RDF marginally closer distance of approach found in  $O_E-O_E$  compared to  $O_E-O$ . In contrast the number of hydrogen bonds per water molecule decreases with ethanol concentration as they are successively replaced with ethanol. Experimentally the maximum number of hydrogen bonds for water should be found at around 20%<sup>???</sup> however this limit is not present in the simulation studies due to the models' predisposition to overestimate hydrogen bonding and so the water-ethanol interactions rapidly incline.

The densities of the ethanol-water mixtures are depicted in Fig.3.18, and so are the number densities accross the box of the 50% ethanol system. Its is clear that the density decreases with the addition of ethanol to the system (Fig. 3.18a). The TIP4P water models closely overlap the experimental trail, while TIP4P/Ice overestimates the densities between the wide region of 30–80%. These features are largely due to close packing of molecules which may have consequences for our work with polymer simulations. It appears that mixing the OPLS-AA ff.and TIP4P/Ice may confine future simulations of poly(vinyl alcohol) to either low or high concentration solution studies in order to provide reliable estimates of the system density. Fig. 3.19 are snapshots of the system at the end of the simulation, which also confirm the disparities in density.

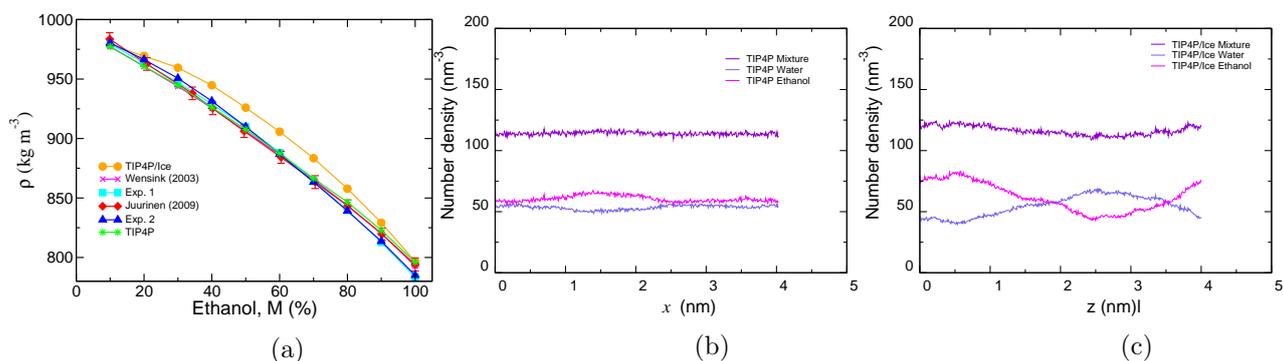


Figure 3.18: (a) Ethanol system density for TIP4P/Ice systems, TIP4P?? systems and Experiments?? as a function of concentration. Number densities for (b) TIP4P and (c) TIP4P/Ice at 50%

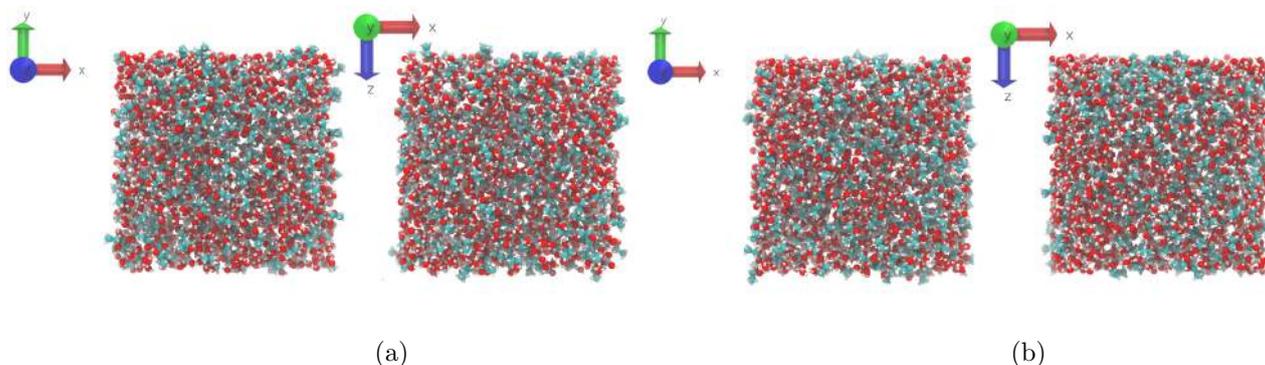


Figure 3.19: (a) Snapshot of packing density for (b) TIP4P and (c) TIP4P/Ice at the end of 3 ns trajectory.

Fig. 3.20 shows the distribution of the  $Q_4$  parameter as a function of water concentration (%) in the ethanol solutions. In both cases the degree of tetrahedrality declines with the introduction of ethanol molecules, similar to the effect of increasing the temperature. In line with previous observations, the ethanol molecules disrupt the hydrogen bonding network in the pure water system. The impact on the structuring of the waters is much greater than the effect of salt or temperature, particularly in TIP4P/Ice. The shoulders at 0.29–0.41 and -0.12 appear after 70% which indicates that another structuring may become prevalent with the introduction a certain level of hydrophobic molecules in the system. It is possible that the third shoulder indicates of a random and non-specific clustering of water molecules, which is in line with the hallmarks of the hydrophobic effect.

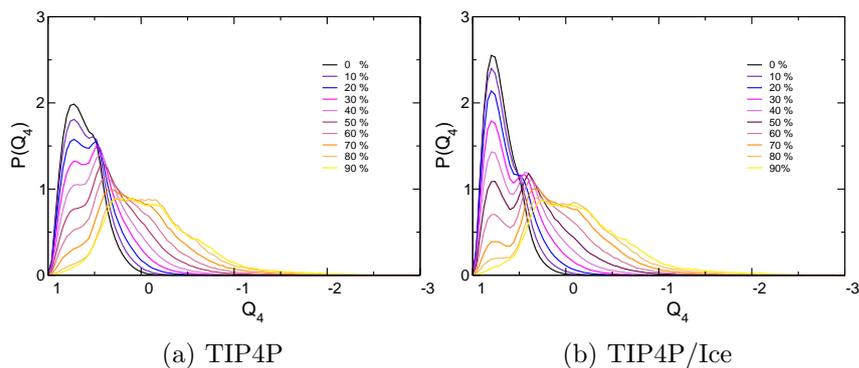


Figure 3.20: The tetrahedral order parameter for (a) TIP4P and (b) TIP4P/Ice.

Free energy of solvation of ethanol was also calculated for TIP3P, TIP4P and TIP4P/Ice water models and the results were found to be  $-18.86 \text{ kJ mol}^{-3} (\pm 0.16)$ ,  $-19.71 \text{ kJ mol}^{-3} (\pm 0.30)$  and  $-20.37 \text{ kJ mol}^{-3} (\pm 0.30)$  respectively. These are similar to experimental value  $(-20.9 \text{ kJ mol}^{-3})$  and is best reproduced by TIP4P/Ice. It is important that the ethanol molecules are

### 3.4.5 The effect of cut-offs on TIP4P and TIP4P/Ice

In this section we review the earlier data and report on the effects of using different sized cut-offs. For all TIP4P/Ice RDFs the three cut-offs (0.85–1.50 nm ) overlap finely and this is true at every sampled temperature (235 K, 298 K and 348 K). The oxygen-oxygen RDFs is depicted in Fig.3.21 and demonstrates that changing the cut-off within this range has negligible effects on the structure of the TIP4P/Ice water. This is also the case for the TIP4P water model which has also noted elsewhere Jorgensen et al. [1983].

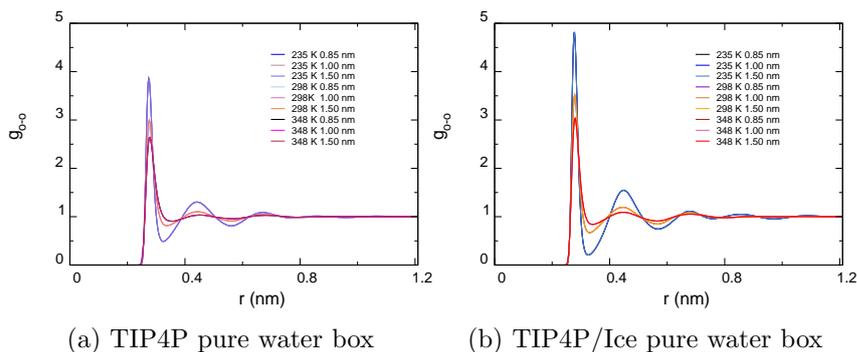


Figure 3.21: Oxygen-Oxygen correlation function for both water models at 3 temperatures and three cut-offs. The results of all three cut-offs overlap finely.

Similar approaches have previously been addressed, comparing popular water models for use in various forcefields van der Spoel and Lindahl [2003]; Hu and Jiang [2010] or the effect of changing a computational parameter Cordeiro [1998]. The same could be said for simple pure water box systems????, however to our knowledge none of these studies have focused on TIP4P/Ice. The results by Yonetani *et al* ?? summarised that with the exception of 1.5 nm, increasing the cut-offs from 0.9–1.8 nm for a pure TIP3P water system introduces undesirable artefacts in the orientational behaviour of water molecules over a 2 ns trajectory. These artefacts become apparent at 1.4 nm, and eventually manifests itself to an extent that an unphysical layer formation is observed. Although other water models were also investigated at smaller cut-off ranges, the study focused on TIP3P water model which is notorious for poorly representing the structural properties of liquid water??. They found that the artefact was dependent on the system size from 2,201 to 3,316 and 4,736 water molecules. With 2,201 waters the artefacts was present with 0.14, 0.16 and 0.18 nm cut-offs. While with the larger system this artefact was only seen at 0.18 nm. Moreover they found that the potential energy decreases and the density increases with the higher cut-offs, suggesting a cut-off dependence of the thermodynamic and structural properties of these water models. These artefacts were also visible in other water models, however they were found to be much lower. Similar conclusions have been deduced for simulation of 1.0 M aqueous NaCl solutions between 0.9–1.6 nm cut-offs?. This study focused on the SPC water model with and revealed that the cut-off approximation can severely affect the ion-ion radial distribution function at distances adopted by the cut-off length.

In also we found that for TIP4P/Ice there were some artificial attributes associated with the structuring of Na-Cl in solution, resulting in more pronounced peaks and

troughs as a result of a larger interaction sphere (see Fig. 3.8c). Aside from the effects of greater negative charges placed on the dummy atom, it is not clear why these same attributes were not detected with TIP4P, and are particularly true of the 0.10 M concentration.

For the observations of ionic simulations we see the following. Due to the noise visible in the Cl-Cl and the Na-Na plots, we assessed both water models for the artefact reported by Auffinger<sup>?</sup> using the Na-Cl RDFs. Our findings show that at both 0.85 nm and 1.00 nm, the tell-tale minima at the cut-off lengths were absent for all the concentrations. We did not conduct these simulation for larger cut-offs and are unaware if this is still the case for a system with a 1.5 nm cut-off. Rather what we do see is that reading from the graphs directly is much clearer to see the effect of concentrations with the 1.00 nm cut-off simulations than with the 0.85 nm cut-offs, particularly with respect to the Na-Cl and O-O RDFs. Except for the Na-Cl RDFs, there aren't substantial differences between the two cut-offs. For TIP4P/Ice water model, the Na-Cl RDFs of the longer cut-offs are considerably taller. For instance at 0.1 M solution it's almost 25% taller at 1.00 nm cut-off than it is at 0.85 nm. These changes cannot be attributed to larger sphere of interaction between the ions, since the same changes are not visible with the other water model.

Additionally we see that the water models do have small but different effects with the introduction of cut-offs. Generally in the TIP4P systems, the hydration number increases for both ionic species with an increase in the cut-offs. While TIP4P/Ice shows conflicting results with decreases for Na and increases for Cl. Nevertheless as the concentration is increased the standard deviation between the two cut-offs decreases for TIP4P/Ice (from 0.057 to 0 for Cl hydration) and so may be negligible at much higher concentrations. In contrast no clear pattern can be deduced for the TIP4P water model.

The cut-offs have negligible effects on the diffusion coefficients of pure water system. Reductions in the diffusion coefficients were observed between 0.85 nm and 1.5 nm cut-offs for the 3 water models at higher temperatures (Table 3.5). However these reductions are much smaller for TIP4P/Ice, particularly at 348 K, where the largest deviation is present. The onset of these deviations begins much earlier with TIP3P at 298 K and demonstrates a lower tolerance towards changes in the cut-offs.

It is clear from Table 3.9 that the hydrogen bonding is unaffected by changes in the cut-offs employed and so it is not surprising that this is reflected in the RDFs. This is true for pure water systems and those with low concentrations of salt (Table 3.6 and 3.9)

On average, at 298 K and 348 K the actual densities of water are best reproduced

by the 4-point water models. At 298 K the closest representative of density is consistently attained at the 1.00 nm cut-off for all models, while at 348 K it appears to be different for each model (1.5 for TIP3P, 1.00 for TIP4P and 0.85 for TIP4P/Ice). Between the cut-offs we calculate the standard deviations for the density of the models and discover that the differences between cut-offs are minor. We see that with TIP3P the variation between the three cut-offs increases with temperature from 0.24 to 0.65 and 1.10. This linear dependence on the temperature is absent for the 4-point water models, however by averaging these three values we can calculate the average standard deviation between cut-offs at different temperatures. We see that TIP4P/Ice has the smallest variation at 0.46 and TIP3P has the largest with 0.66, which means that the effect of changing the cut-offs on the density is smallest for TIP4P/Ice and greatest for TIP3P.

It has previously been remarked upon that the use of a cut-off radius leads to higher densities which displaces the the temperature of maximum density at 1 bar to lower temperatures Abascal et al. [2005]; [?]; [?]; [?]. In 2006 work by van der Spoel *et al* [?] previously reported the effect of 2 different cut-off treatments (0.9 and 1.2 nm) on the effect of the density, energy, dynamics and dielectric properties. This work was done on SPC, SPC/E, TIP3P and TIP4P water models however no report was made for TIP4P/Ice. These were done for 2 systems of 216 and 820 molecules, treating the long range electrostatic interaction with the reaction field approach. Their work serves as a good comparison to our work, however we have worked with larger systems of 4008 water molecules, longer simulation times (3 times), and did not alter the OPLS-AA forcefield.

In our simulations, increasing the cut-offs at 235 K decreases the density of TIP4P/Ice and at the higher temperatures (298 and 348 K) the densities increase (with an anomalous peak at 1.00 nm). This is not a linear trend and more points are needed to establish a more complete density profile. The aforementioned reports indicate that from 0.9 to 1.2 nm cut-offs, the density of the largest system increases for all water model at 298 K, regardless of the size of the reaction field used (1 and 78.5 kJ mol<sup>-1</sup>). Results for TIP3P are completely consistent with these observations, while it is much more difficult to draw a fine conclusion with the 4 point water models.

Again, no differences can be seen for the tetrahedral ordering as a result of changing the cut-offs rather the results for all finely overlap, as we see with the RDFs. This is evident in the absence (Fig. 3.6) and presence of small ions (Fig. 3.13) and such features are very important in studying the induced effects of the structuring by small polymeric, antifreeze molecules around the ice quasi-liquid layer. It is impor-

tant that we observed no artefacts and so far the results are promising under these conditions. As with the liquid solutions, the melting points of the water models as a function of the cut-offs have previously been discussed by many ?. The effects of cut off are primarily discussed in the section which follows (See Melting and freezing).

### 3.4.6 Melting and freezing

The melting points of the water models have previously been mentioned in this report and by many others ?García Fernández et al. [2006]; Abascal et al. [2005]; Conde et al. [2008] however we are primarily interested in the use of TIP4P/Ice for melting/freezing under reported conditions, Weiss et al. [2011]; García Fernández et al. [2006] as well as our own. Understandably these melting and freezing temperatures vary slightly in literature with respects to the methods used to simulate or calculate the melting and freezing points. Nevertheless the results are in reasonable agreement with the experimental melting point of ice (273 K) at 269 K. A range from 266–278 K have previously been found as the melting points for the TIP4P/Ice model and our simulations comfortably fall within this range for both cut-offs. Weiss et al. [2011]; García Fernández et al. [2006]; Abascal et al. [2005]; Conde et al. [2008]

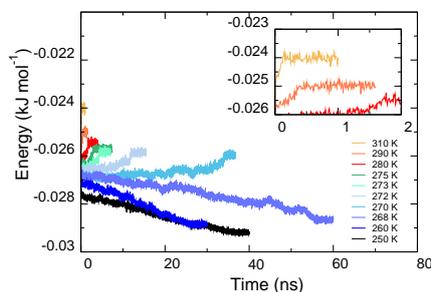


Figure 3.22: The total energy per water molecule from 250 K–310 K for 1 nm cut-off simulations.

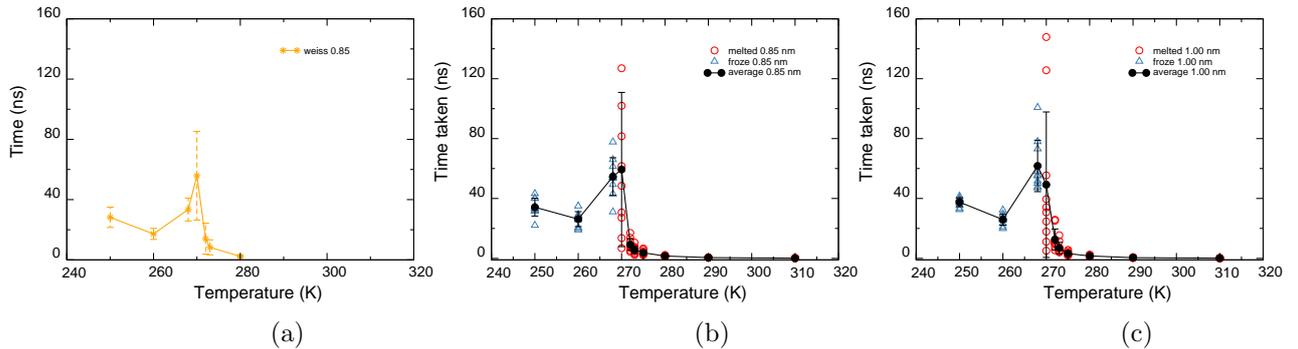


Figure 3.23: The time that has elapsed before completion of freezing or melting at different temperatures in studies (a) produced by Weiss *et al* Weiss et al. [2011] (at 0.85 nm ), (b) using at a 0.85 nm (c) and a 1.00 nm cut-off distance. Note that for easy comparison, we only show cited results for the temperatures that we have studied.

Figure 3.22 depicts the total energy as function of time at different temperatures (230–310 K) for the 1 nm the cut-off simulation, and it shows that at and above 270 K the TIP4P/Ice completely melts. Below 270 K the system completely freezes. The time taken for completion of freezing or melting of our own simulations is displayed in Fig. 3.23b and 3.23c for 0.85 nm and 1.00 nm cut-offs respectively. For comparison we include the the results produced by Weiss *et al* Weiss et al. [2011] (Fig. 3.23a). They determined that the melting point of TIP4P/Ice was approximately 270 K because at this temperature they observed complete melting as well as complete freezing of the simulation between  $\sim 32$ – $93$  ns and  $\sim 7$ – $14$  ns, respectively. We however did not find a temperature at which both complete melting and freezing could be observed. Rather we deduced 269 K to be the melting temperature for both cut-offs, since melting is observed above that and freezing below it. These temperatures are in good agreement with the expected  $T_m$  for the TIP4P/Ice water model and once again illustrate that the mixing of the two forcefields can still reproduce reliable thermodynamic data. Fernandez *et al* García Fernández et al. [2006] studied the effects of system size and different computational modelling techniques on the  $T_m$  [and freezing rates] of TIP4P/Ice using the direct coexistence protocol. They found that it is not uncommon to see small disparities in the values due to as a result of using different methods and found only a 4 K variation in using MD simulations rather than free energy calculations and a 2 K variation in doubling the system size. We observe only a 1 K difference between the work produced by Weiss *et al* Weiss et al. [2011] and our own, although they used a 0.85 nm cut-off distance, semi-isotropic pressure coupling, a system almost half the size of our own and 40%

ice composition (compared to our 67%).

The contours of these plots illustrate that at lower temperatures the freezing process typically takes much longer, as a result of depleting kinetic energy (needed to diffuse and reorient), and having approximately twice as many water layers to process than for melting. This is particularly true close to the melting point where we see a sudden spike in time required. Such observations are in line with theoretical findings?Weiss et al. [2011]. For the shorter cut-off the apex of the graph is marked by melting simulations while for the larger cut-off it is marked by the freezing simulations. Neither is incorrect and this feature can be explained by drawing parallels to the finite-size effects in freezing simulations. Liquid regions in smaller systems are more influenced by periodicity introduced by using 3D PBC and so exhibit some artificial ordering. This reflects in higher growth rates and larger  $T_m$  García Fernández et al. [2006]; Weiss et al. [2011]; ? In our studies, the first point of melting appears at 270 K using a shorter cut-off and this increases to a higher temperature (272 K) using the 1 nm cut-off. Using a greater cut-off radius provides more opportunities for the water molecules to interact with its neighbours, and the crystalline phase can induce further ordering into the liquid phase.

As mentioned earlier, the  $Q_4$  parameter is often used to determine the regions of hexagonal ice-like arrangements in liquid water???, and proves useful in the assessment of the growth of an interfaces in ice/water simulations?. This parameter is quite advantageous, seeing that hexagonal ice represents the limiting case of a well ordered tetrahedral network of hydrogen bonds in water. Figure 3.24 depicts the average orientational order parameter for different temperatures over time. As for the total energy plots, an inflection into a plateau signifies the completion of the melting or freezing process. The plot illustrate that the liquid water from the melting simulations can adopt a range of  $Q_4$  orientations (0.68–0.76 at 310–270 K) whereas the ice simulations, on completion of freezing are limited to a much more narrow range of  $Q_4$  values (0.94–0.95). Earlier, in section 3.4.2 we determined that liquid water that has been supercooled to 235 K can even obtain a values as high as  $Q_4= 0.88$ . As a result, it is better to use these  $Q_4$  values obtained from the ice simulations as the cut-off to distinguish between ice like and liquid like water, rather than identifying intersections in  $Q_4$  distributions from a large range of the solution studies of pure water conducted at a range of temperatures. Using the ice growth simulations at 268–260 K we propose a cut-off criteria of  $Q_4= 0.94$ . This method eliminates the need to repeat many simulations in solution, as is custom, and the uncertainty that accompanies the use of a smaller range of simulations.

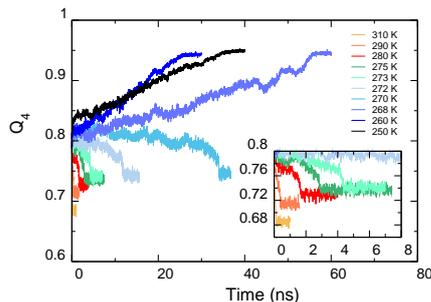


Figure 3.24: Orientational order parameter for different temperatures.

Figure 3.25 depicts the distribution of the  $x/y$  averaged density as a function of distance in the  $z$  direction. The entire sample was partitioned along the  $z$  direction into bins of  $0.2 \text{ \AA}$  wide. The density profile in Fig. 3.25f corresponds to liquid while Fig. 3.25e corresponds to a fully frozen system. The centre of the density profile in 3.25d is the middle of the ice region within the starting configuration (approx. 4 nm into the  $z$  direction). The density profiles show characteristic long spikes where water is present and almost zero values in-between which reflects a highly ordered and regularly spaced structures at the center of the box. In contrast the density of the liquid sections undergo random fluctuations and so peaks are consistently present at regions before  $\sim 2.5 \text{ nm}$  and after  $\sim 5.5 \text{ nm}$ . Additionally the troughs between the spikes do not approach zero at all. Notice that taking the average of the ice region and the liquid water region separately still shows that the density of liquid water is higher than the density of ice and unsurprisingly, upon melting the box shrinks slightly. Ordinarily the structure and dynamics of the interfacial water is reduced compared to those of other surrounding bulk environments, often an intermediate between the two. This structural change at the interface is evident in our density profiles,  $\sim 0.5 \text{ nm}$  from either bulk environments, which somewhat lower than the expected  $1.00\text{--}1.5 \text{ nm}$  ?, however it suffices to say that the QLL will gradually form over simulation time as a function of temperature.

### 3.5 Conclusion

In this paper, we have compared the different water models describing the liquid properties of water. Our results show that the density, structuring, hydrogen bonding and diffusiveness are fairly insensitive to the choice of the cut-offs between  $0.85 \text{ nm}$ ,  $1.00 \text{ nm}$  and  $1.5 \text{ nm}$ . Moreover TIP4P/Ice and the other water models have

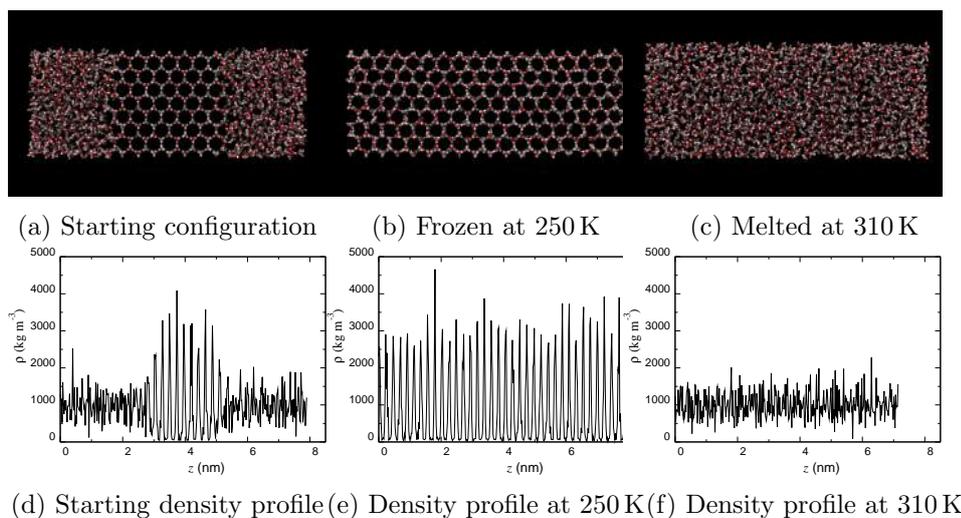


Figure 3.25: Snapshot of the system containing an ice slab sandwiched between liquid water, before and after simulations. Oxygen atoms are shown in red and hydrogen atoms in white.

been able to reproduce physical properties of water to similar previous reports of these models and in many cases similar to experimental results. To our knowledge this is the first time that in a molecular dynamics simulation that a variation of all water models have been compared and explored under the exact same conditions, over a systematic series of varying simulation conditions which includes later models like TIP5P and TIP4P/Ice.

The TIP4P/Ice appears suitable for use with the OPLS-AA forcefield and the use of larger cut-offs does not particularly impede the properties of water in the presence or absence of additives. However we foresee that for larger concentrations of NaCl, one should be cautious as the effects of cut-offs may induce problematic structuring features with small ionic additives like salts. The variations in properties like hydrogen bonding, RDFs, density and many features arise between different water models because of differences in molecular geometry and intermolecular potentials. From our assessment it is evident that the TIP4P/Ice water model was parametrized to be much more ordered than the other water models and is biased towards tetrahedral orientations even in liquid phase. The relatively strong and long range ordering of TIP4P/Ice water is crucial to what makes it particularly suitable for ice growth studies. Its indifference to cut-offs, over the range tested here, makes it reliable for investigations using the OPLS-AA forcefield and in studies of ice growth with additives like salts, ethanol and larger macromolecules such as

antifreeze polymers.

## Chapter 4

# MD simulations of antifreeze active and antifreeze inactive polymers in solution

### 4.1 Introduction

The solution properties of the antifreeze active and inactive polymers can provide information about the prerequisites for the antifreeze activities of PVA and why PEG remains inactive. MD simulation studies in solution can help to clarify which specific interactions the polymers partake in, and how these or other properties may change in the presence of the ice crystal. Studying these systems is also useful for addressing other aspects such as the importance of hydrophobic groups and the possibility of lattice matching or water ordering which may be required for significant interactions with ice. First we aim to validate the PVA and PEG models by comparing them to polymer theory and previous published modelling results. Next, we also investigated single polymer chains in solution in order to observe the effects of polymer type, oligomer size and temperature on the structure and dynamics of PVA and PEG. We aim to assess differences between the two polymers antifreeze properties and identify if any properties are switched on at a particular chain length or if they can be triggered by a new, colder environment. The respective chemical formulas of PVA and PEG are:  $X(-\text{CH}_2\text{CH}(\text{OH})-)_N\text{-Y}$  and  $X(-\text{CH}_2\text{CH}_2\text{-O-})_N\text{-Y}$  where  $N$  is the number of repeat units, while  $X$  and  $Y$  can be end groups of different types. In this study we have chosen  $X=Y=H$  because it omits experimental end groups which vary as a result of different synthetic approaches. The selected range of polymers and their degrees of polymerisations ( $N=5, 10, 19, 30$  and  $56$ ) studied here were chosen

for direct comparisons to recent ice recrystallisation experiments by Congdon *et al* .[TX1]

## 4.2 Methods

### 4.2.1 Forcefield parameters

In the previous chapter we demonstrated the compatibility of the OPLS-AA forcefield with the TIP4P/Ice water model which has not been done before. Following this demonstration we used the OPLS-AA forcefield to model the polymers and the TIP4P/Ice water model to represent water. We chose the OPLS-AA forcefield because it was developed for the simulation of organic liquids such alkanes, ethers and alcohols which are structurally similar to our polymers, and it has previously been used to successfully simulate PVA, PEG and related polymers.[POS2] [POS5] [POS12] [POS14] [POS16] [POS18] [POS19] The TIP4P/Ice water model was chosen in consideration of the next chapter where we will investigate the effects of the polymers on the phase change of an ice/water interface. These two forcefields were chosen because the hydration of polymers is central to both of these studies and so it is important to represent the polymers and water accurately within the limits of computational resources.

### 4.2.2 Simulation parameters

All simulations were performed using the Gromacs (v4.5.4) simulation package and can be described in the following stages. A simulation box was placed around a single equilibrated polymer of  $9.10 \times 10.34 \times 10.75$  nm and the system was solvated via random placement of 33207 TIP4P/Ice waters into free spaces. This equilibration procedure is crucial and outlined in the next section. It is important to ensure complete equilibration because polymer activities may rely greatly on their conformations as well as composition. The equilibration provides an opportunity for the models to find their ideal conformations in solution and the convergence of the system was therefore monitored using the radius of gyration over time. This approach is common, non-intensive calculation and is a fine way to measure polymer compactness over time. Table 4.2 summarises the parameters set for the 70 ns production simulation of the polymers in solution. During the NPT simulation the temperature was maintained at 298 K using the Nosé Hoover thermostat and the system was coupled to the thermostat every 0.2 ps . The pressure was held at 1 bar using the Parinello-Rahman barostat and a 2 ps coupling constant. Short range LJ

and coulombic interactions were cut-off at 1.00 nm , and long range electrostatics were controlled using using the PME protocol. These include the application of a 0.16 nm grid and a cubic interpolation. LINCS constraints were applied to all hydrogen bonds and periodic boundary conditions were used to remove surface effects and simulate a bulk environment. The lengths of interest were 5, 10, 19, 30 and 56 repeats units. Three simulations were conducted for each chain length and in sum 15 simulations were conducted for each polymer at 298 K . The last configurations from these runs were used to create initial configurations for simulations at 260 K . A subsequent 3 ns productions simulations were run at this temperature and aside for adjustments to the output frequency, the same simulation parameters were applied. The results of these simulations were used to validate the polymer models and to assess their physical properties before introduction to the ice environments.

<b>Parameter</b>	<b>Setting</b>
Forcefield	OPLS-AA
Lennard-Jones	1.00 nm
Coulombic (short)	1.00 nm
Coulombic (long)	PME
PME details	Cubic interpol. (0.16 nm grid)
PBC	Yes
Thermostat	Nosé Hoover
Barostat	Parinello Rahman
Constraints	LINCS (H-bonds)

Table 4.1: MD simulation parameters of production runs at 298 K and 260 K .

### 4.2.3 Initial configuration and equilibration

The range of selected oligomers were built linearly using a C-program [**written inhouse**]. The atom type assignments for our polymers are shown in Fig. 4.1 and Table. Starting configurations for each linear polymer was built isotactic in structure. In order to remove unreasonable or close contacts between given atoms, each polymer was subjected to a 20,000 step energy minimisation using the steepest descent algorithm. The new structures were placed in a cubic box accommodating their full length with an additional border of 3 nm . Next the system was solvated with TIP4P/Ice water and subjected to a second energy minimisation, with periodic boundary conditions. Each system was equilibrated for 30 ns with short consecutive

NPT simulations using the same conditions described in Table 4.2. This was performed in order to allow the polymers to collapse into their preferred conformations in solution and the process was monitored using the radius of gyration. A good convergence was reflected by negligible changes from a baseline. If convergence was incomplete, the polymer was reboxed, resolvated and the system was prepared for a second 30 ns simulation. This stage was repeated until convergence was observed and is preferred to the equilibration process commonly employed for collapsing polymers into a random coil in vacuum. Although the latter approach may be faster for larger polymers, our molecules are actually closer in size to oligomers and so the aforementioned approach is computationally affordable. Additionally it reduces risk associated with the assumption that the polymer will readily uncoil with the introduction of water in later stages. Since the solvent accessible regions are key properties to be investigated here, our approach is suitable for its purpose and the re-boxing permits us to work with smaller, inexpensive system sizes if the polymer is collapsed. Further reasons for this were discussed earlier in section [x afgp simulation section x]. It is important to stress that in the interest of time, this protocol restricts us from working with larger polymers which require a larger system size and typically equilibrate much more slowly.

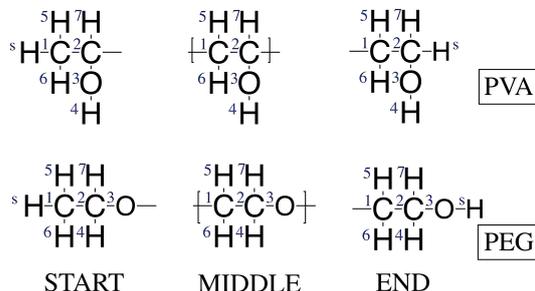


Figure 4.1: Numbering for atom type assignments of the terminal groups and repeat monomers used for PVA and PEG polymers

Label	PVA			PEG		
	Start	Middle	End	Start	Middle	End
1	OPLS135	OPLS136	OPLS136	OPLS135	OPLS182	OPLS182
2	OPLS158	OPLS158	OPLS157	OPLS182	OPLS182	OPLS157
3	OPLS154	OPLS154	OPLS154	OPLS180	OPLS180	OPLS154
4	OPLS155	OPLS155	OPLS155	OPLS185	OPLS185	OPLS140
5	OPLS140	OPLS140	OPLS140	OPLS140	OPLS185	OPLS185
6	OPLS140	OPLS140	OPLS140	OPLS140	OPLS185	OPLS185
7	OPLS140	OPLS140	OPLS140	OPLS185	OPLS185	OPLS140
S	OPLS140	-	OPLS140	OPLS140	-	OPLS155

Table 4.2: Atom type assignment of the terminal groups and repeat monomers used (4.3a) for PVA and (4.3b) PEG polymers in the OPLS-AA forcefield

## 4.3 Results

### 4.3.1 Macromolecules shape and size

The average radius of gyration ( $R_g$ ) was calculated for all oligomer and are shown in Fig. 4.2 and Fig.4.2. The results are also shown alongside a distribution of values explored by each oligomer over a full 70 ns trajectory (Fig. 4.4). We found that the  $R_g$  for our oligomers obtained a power law dependence with chain lengths, which is in good agreement with experiments ( $R_g \propto Mw^{0.588}$ ) and other modelling studies. [POS8] [POS18] [POS19] [POS20] Our PEG oligomers obtained a flory exponent of  $0.53 \pm 0.051$ , while PVA had one of  $0.63 \pm 0.027$ . Rossinsky *et al* previously investigated PVA oligomer melts within a similar range of chain lengths and determined an exponent of 0.65 [POS8]. Similarly, Lee *et al* calculated values of 0.52 for poly(ethylene oxide)[POS19] and Hezaveh *et al* obtained a value of 0.59 using the OPLS forcefield [POS19]. These calculations provide an insight into the relative shapes or sizes of the polymers in solution. The better the solvent, the more likely they are to adopt a random coil conformation, and the greater the flory exponent. In contrast, polymers in poorer solvents are typically more compact and globular in conformation, so they yield values that are much closer to 0.3. Therefore we conclude that both polymers are extended as random coils in solution.

It is important to note that unlike the polymers in experiments our oligomers are rather small and our system monodisperse, while a solution of polymers are typically several hundreds of units long and polydisperse. For instance Gibson *et al*

studied the ice recrystallisation inhibition (IRI) of PVA with polydispersity indices between 1.18–1.45.[TX1] In the light of this we acknowledge that these values may be somewhat overestimated, similar to findings from the aforementioned modelling studies.[POS8] [POS18] [POS19] [POS20]

The distributions of  $R_g$  reveal that larger oligomers and PEG explores a wider range of conformations while smaller oligomers and PVA sample a narrower range. In response to the cold temperature, these distributions continue to narrow and the PVA oligomers become less soluble in water, almost to the same extent as PEG at 298 K (Fig. 4.3), while the solubility of PEG remains approximately the same at both temperatures. At this present time we cannot tell whether or not this is part of the mechanism but with further study it may prove to be an important feature for antifreeze activity.

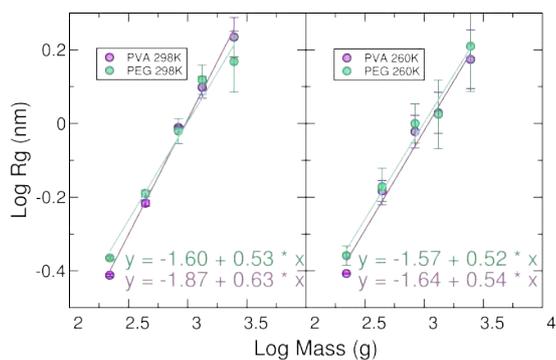


Figure 4.2: Linear fit to  $\log R_g$  vs.  $\log$  molecular weight at 298 K and 260 K . Error bars are standard deviations between repeats

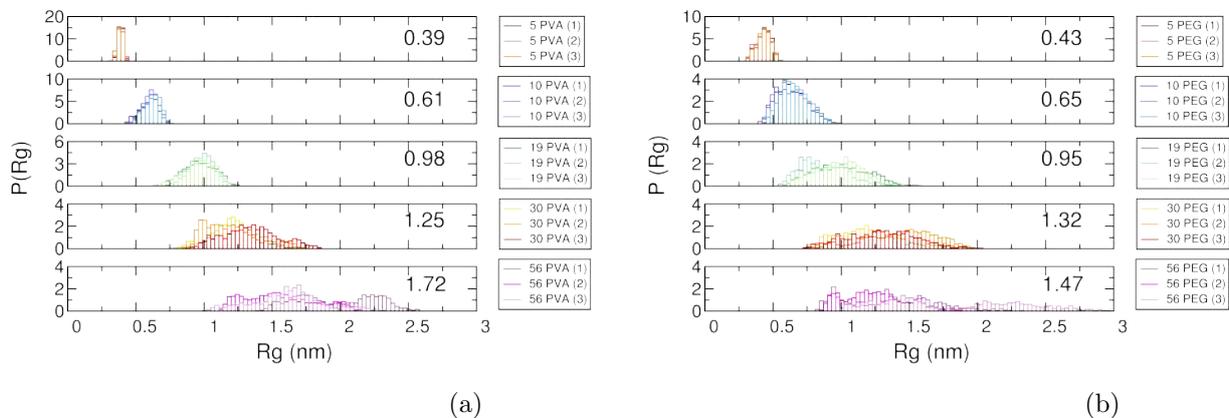


Figure 4.3: The probability distributions of the  $R_g$  for PVA and PEG oligomers at 298 K .

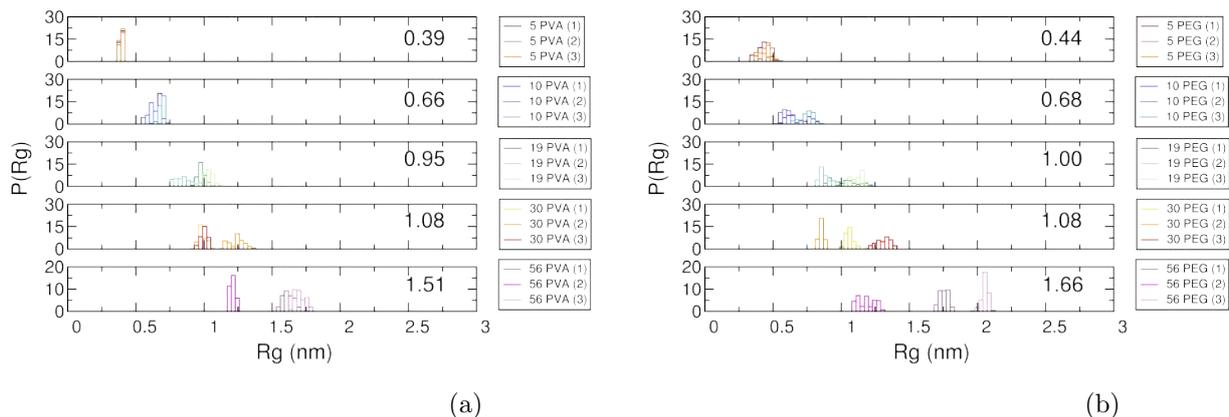


Figure 4.4: (4.3a) The probability distributions of the  $R_g$  for PVA (4.3b) and PEG oligomers at 260 K .

### 4.3.2 Water ordering or disordering properties

#### Through perturbation

One of the hypotheses mentioned earlier; is the possibility that AF(G)Ps and their mimics function by ordering or disordering surrounding water in order to inhibit growth of ice on the ice/water interface. [RES4] [SOLO3][SOLO4] This is one of the most important issues we aim to clarify during ice crystal growth as well as in solution. If these oligomers are capable of distorting the bulk or hydrating water

significantly, one would expect to observe some preceding activity in solution and at room temperature. In order to achieve this we characterised the water around the polymer by calculating the radial distribution functions (RDFs) of water oxygens (O) to the polymeric oxygens ( $O_X$ ), the hydration number and the orientational order of hydrating waters.

Fig. 4.6 shows the RDF results for each oligomer and highlight the probability of finding a water oxygen close to the polymeric oxygen. We find that these RDFs exhibit characteristic peaks for PVA at 0.28 nm (strong), 0.47 nm , 0.66 nm , 0.87 nm and 1.07 nm (weak), as well as for PEG at 0.28 nm , 0.48 nm (strong), 0.83 nm and 1.13 nm (weak). These positions agree with experiments and the probabilities are comparable to those reported by other similar studies. [POS3] [POS5] [POS7] [POS8] [POS9] the probability of finding a water oxygen close to the polymer oxygen is high at close distances, and declines on approach to the bulk water, which is further away from the polymer. It is apparent that PVA has multiple peaks, all of which are more pronounced and narrower than corresponding PEG polymers and indicates well defined hydration shells, with strongly confined water molecules. It also describes the longer-range ordering induced by interactions with PVA, which is unsurprising because PVA has two atoms which are capable of hydrogen bonding with surrounding water molecules, while PEG only has one.

The RDFs for PEG markedly has taller peaks for the second shell than the first one, which hints at another type of interaction being prioritized for short range interaction with the oligomers. As a result there is also a higher probability that water will be confined some distance away. One such example could be a preference for polymer-polymer interactions or some form of hydrophobic solvation. For both types of polymers the height of the peaks are reduced somewhat with increasing chain length. We attribute this to the ability of longer polymers to fold and temporarily mask or bury its own oxygen groups from waters around it.

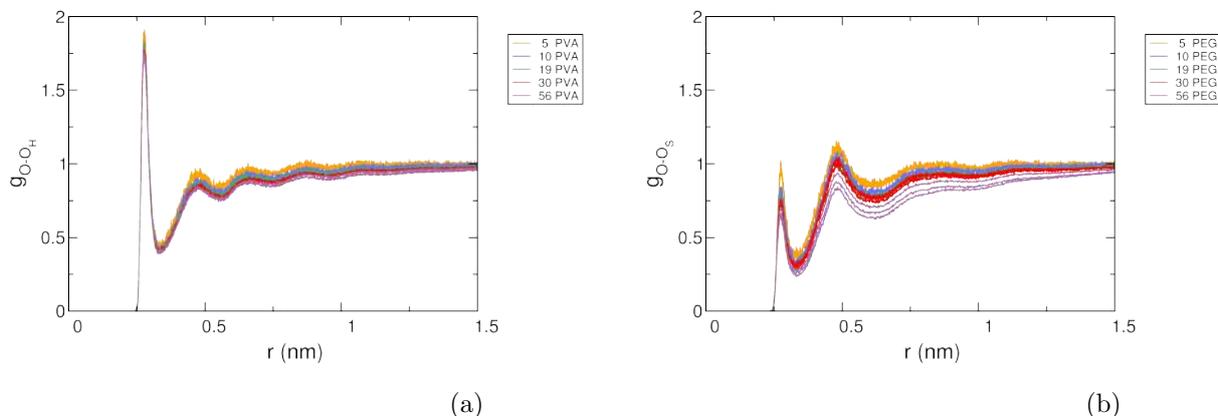


Figure 4.5: RDF between water oxygen and polymer oxygen atoms at 298 K

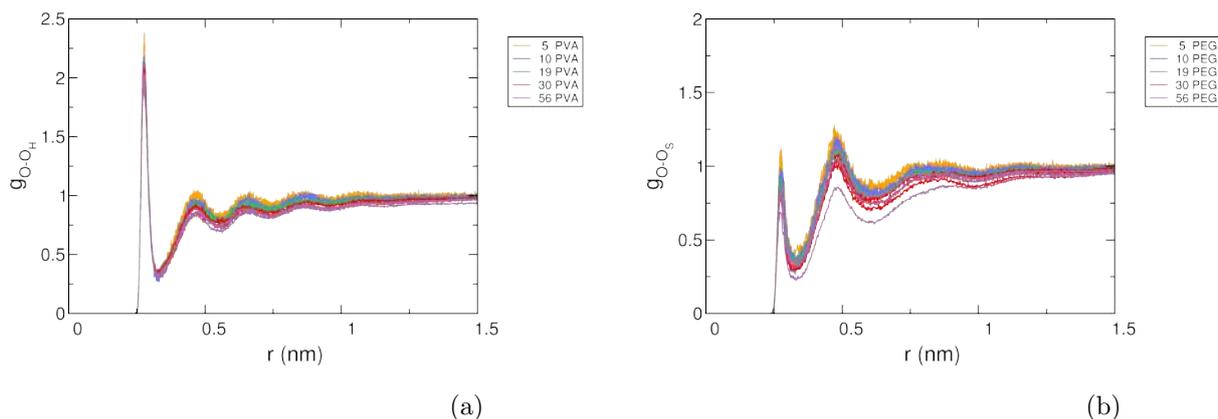


Figure 4.6: RDF between water oxygen and polymer oxygen atoms at 260 K

Because it is possible for the ordering or disordering of the polymers to occur via hydrogen bonding or through hydrophobic interactions or both, we plot the RDF for water oxygens with polymeric carbons in Fig. 4.7. The relative importance of these interactions can also be evaluated through several other types of analysis; the solvent accessible surface areas, the hydrogen bonding as well as the RDF of both hydrophilic and hydrophobic groups in water. The  $\text{O}-\text{C}_S$  RDFs for PEG reveal peaks at 0.38 nm which merges with a weak shoulders at 0.50 nm, and again for the next two weak peaks at 0.65 nm and 0.76 nm. In contrast, PVA oligomers have well separated peaks at 0.37 nm, 0.51 nm and 0.79 nm, all of which are lower in height than for PEG. A comparison of the two RDFs for PEG oligomers ( $\text{O}-\text{O}_S$

and O-C<sub>S</sub>) indicates that water oxygens can approach more closely to the oxygen atoms (0.28 nm ) than the carbon atoms (0.38 nm ), however the relative heights of the peak illustrate that they are more likely to be found interacting with the carbon atoms. This is further emphasised by the taller second shoulder in the O-O<sub>S</sub> RDFs, which confirm that the water oxygens are likely to be dispelled some distance away from the polymeric oxygen as a result of the hydrophobic repulsion. A similar comparison of the two RDFs of PVA reveals that waters largely reside at a distance of 0.28 nm and preferentially interact with its polymeric oxygen. We can therefore suggest that the PEG polymer largely interacts with solvent via hydrophobic interactions from which its backbone is predominantly composed of, while PVA does so via hydrophilic interactions provided by its hydroxyl side groups. These attributes remain true at cold temperatures, and the RDF peaks and troughs are simply more pronounced as a result of reduced thermal energy available for molecules to move outside of their respective shells.

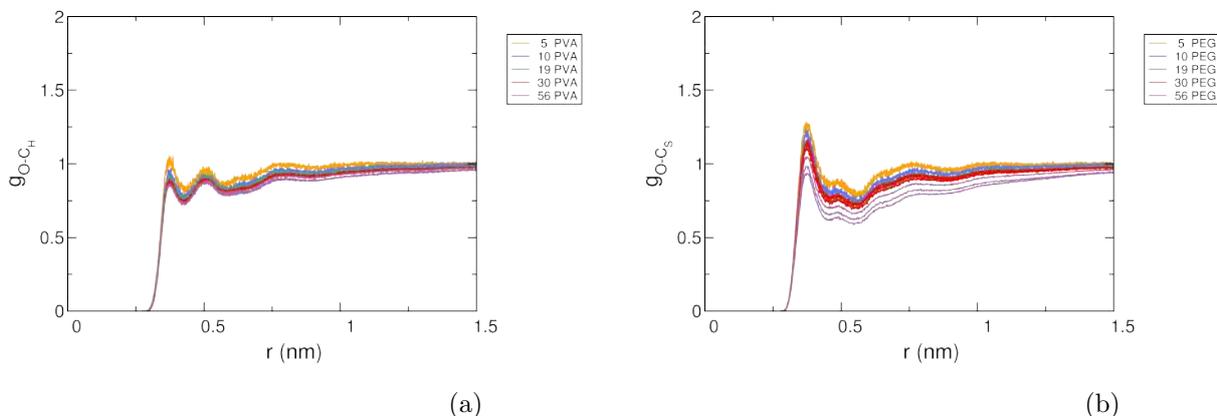


Figure 4.7: RDF between water oxygen and polymer carbon atoms at 298 K

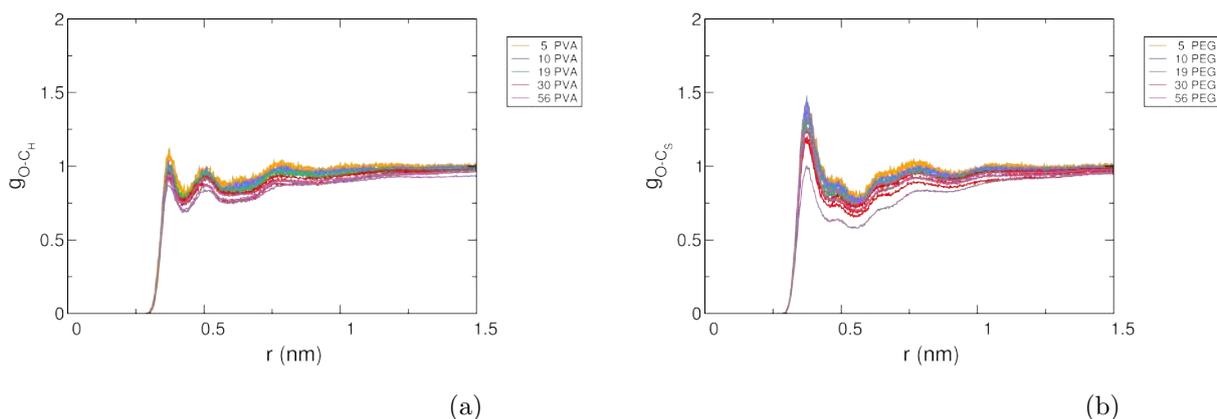


Figure 4.8: RDF between water oxygen and polymer carbon atoms at 260 K

### Through hydration indices

A histogram of hydration numbers is shown in Fig.4.9 for both types of oligomers. Carbohydrate have previously been used to demonstrate the importance of hydration in ice recrystallisation inhibition (IRI) activity. [EX13][EX7] Naturally, we investigate the hydration numbers of oligomers within a 0.35 nm cut-off distance and find that the profiles loosely mimic the distribution of the radius of gyration. The longer the polymer, the wider the distribution because more conformations are accessible and so the hydrating water will also be altered just as much. We find a much narrower distribution for PVA. Throughout the trajectory the 5-mers of

both polymer encounter approximately the same number of waters with the same frequency. With polymers larger than five units long, we find that PEG consistently has a larger hydration number than PVA. This too is reflected in the broader RDFs in Fig4.6 and is surprising as it indicates that more hydrophilic groups does not necessarily translate into larger hydration sphere, but could mean the reverse.

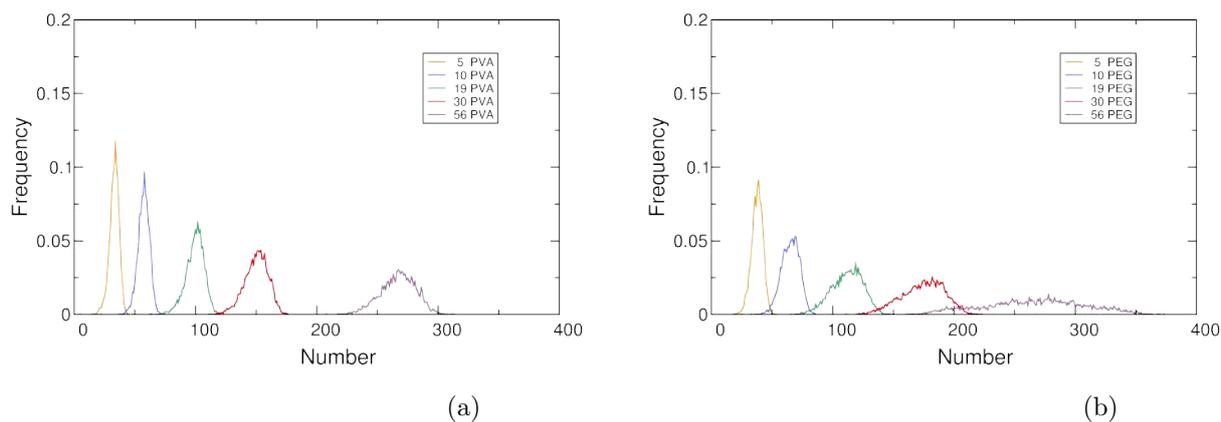


Figure 4.9: Probability density showing the average hydration number of the *whole* polymer at 298 K

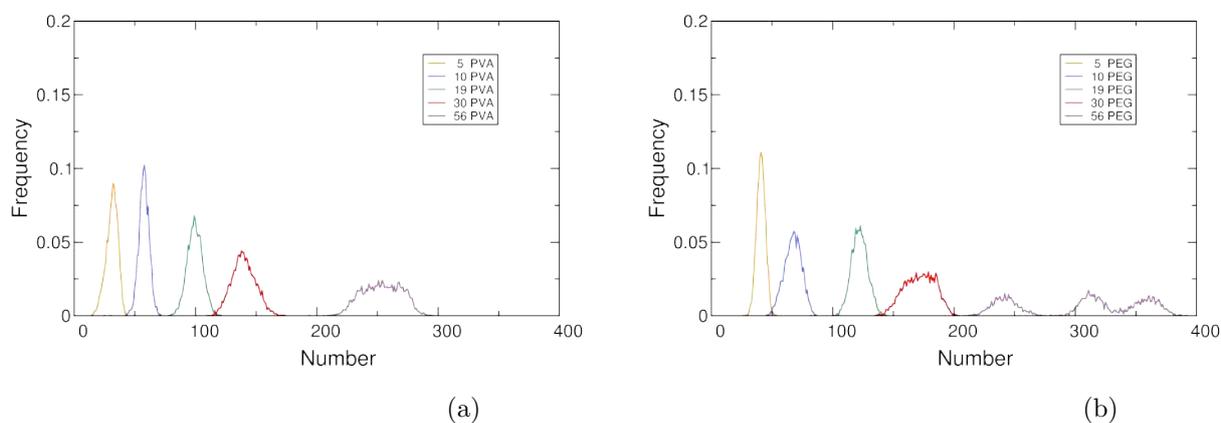


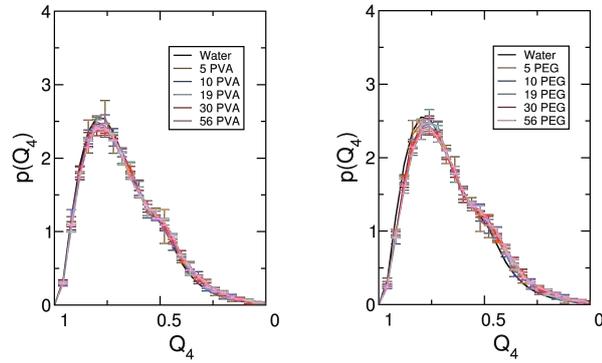
Figure 4.10: Probability density showing the average hydration number of the *whole* polymer at 260 K

### Through tetrahedral orientation

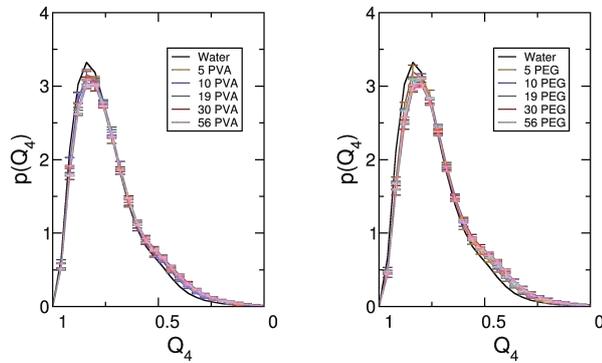
In Fig.4.11 we plot the tetrahedral orientation of water and polymeric oxygens. We find characteristic peak at 0.80 and a subtle shoulder at  $\sim 0.5$  for ordered and disordered regions respectively. In pure water, the transition from 298 K to 260 K caused an increase in tetrahedrality from 0.8 to 0.84, and in population from 2.552 to 3.322 (by approx. 30%). At both temperatures, we do not observe any remarkable ordering or disordering of the water oxygens within 0.35 nm of the polymeric oxygens, by either of the polymers nor as a function of chain length. A comparison of PVA and PEG polymers did reveal a weak 0.04 shift for the PEG simulations to disordered regions. No such shifts were not observed for any of the PVA systems. Rather PVA exhibits a miniscule decline in heights of the ordered peaks. This is reflected in the fact that the same transition from 0.80 (at 298 K ) to 0.84 (at 260 K ) in the presence of PVA<sub>5--56</sub> increased the population of the ordered region by just 25–28% and by 24–30% in the presence of PEG<sub>5--56</sub>. Generally, these disordering changes for both types of polymers are accompanied by a growing shoulder at  $\sim 0.5$ , but it is important to stress that these changes are very faint.

### 4.3.3 Lattice matching properties

Another interesting theory we explore is the concept of specificity and the ability of the PVA polymer to compliment the ice lattice spacing. So we calculated the RDFs between the polymeric oxygens and the results are shown in Fig. 4.12 and Fig.4.13. For PVA, we obtain large peaks at 0.30 nm , 0.46 nm and smaller ones at 0.62 nm , 0.73 nm and 0.81 nm . In total 3 distances could be found between PVA oxygens which compliment the lattice spacing on ice with deviations of just 0.01–0.02 nm, which is an excellent fit. The first two peaks are predominantly due to adjacent oxygen groups on the polymer backbone, one of which has the highest probabilities and perfectly matches the 0.45 nm oxygen spacing on the prism plane. The fourth and fifth populated peaks at 0.73 nm and 0.81 nm are formed by non-adjacent oxygens on the polymer backbone and these also coincide with oxygen separations of other planes of ice. The basal plane is characterised by spacings of 0.83 nm while oxygens on the other prism plane are distanced by 0.74 nm . In stark contrast, PEG only has two well defined peaks at 0.31 nm and 0.40 nm ; neither of which match the oxygen separation distances reported for any of the ice planes. Again, we report that with increasing chain length the peaks become increasingly broader and shorter. Peaks for oxygen separation also begin to form and merge at further distances because of an increased flexibility experienced by longer chain lengths.



(a)



(b)

Figure 4.11: Probability density distribution of the tetrahedral order parameter of interfacial water at 298 K and at 260 K. Error bars are standard deviations between repeats

In order to confirm our findings we also reported the RDFs for hydrogen groups from the PVA in Fig. 4.14. Hydrogen bond separations found on the oligomers also need to be considered because they are important for hydrogen bonding with the ice lattice. There are three main peaks at 0.26 nm, 0.52 nm and 0.71 nm. Although the hydrogen atoms do not necessarily point in the same direction as the oxygen it is attached to, we find that some of the hydrogen separation distances could be made to coincide with the ice lattice as well. Three consecutive separations of 0.26 nm (0.78 nm), or a combination of 0.26 nm plus 0.52 nm (0.79 nm) produces an appropriate match for the basal planes. Hydrogen atoms separated by 0.71 nm could also potentially interact with the oxygens prism plane as well. These findings are satisfactory and compliment our expectation that PVA is generally more spe-

cific than PEG and as a result it is capable adopting conformations which allow its hydroxyl groups to match and possibly incorporate into the ice lattice.

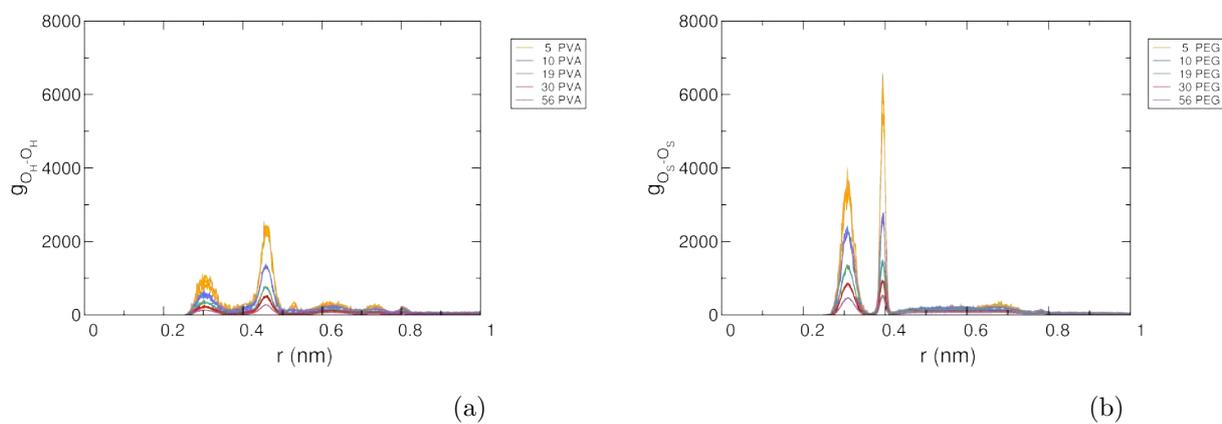


Figure 4.12: RDF between oxygen atoms in oligomers at 298 K

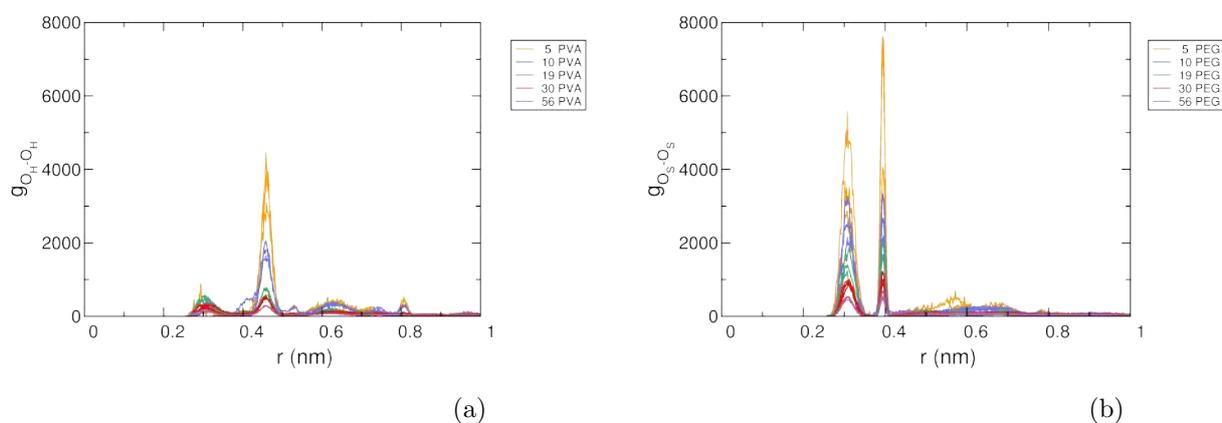


Figure 4.13: RDF between oxygen atoms in oligomers at 260 K

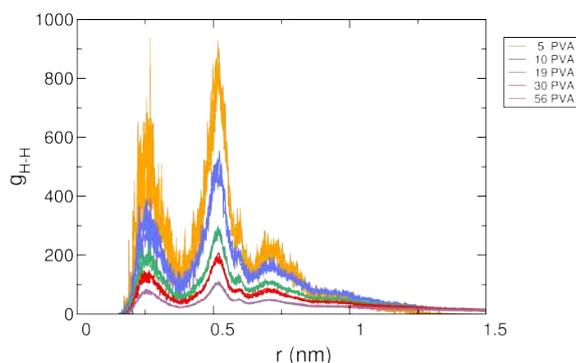


Figure 4.14: RDF between hydroxyl hydrogen atoms in PVA oligomers at 298 K

#### 4.3.4 Solvent accessibility

One other hypothesis that we tested in the solution studies is the importance of hydrophobicity and solvent accessible areas (SASA). Fig. 4.15 shows the total SASA of the polymers alongside their corresponding distributions. We find that PVA and PEG have approximately the same total SASA (TSASA) at each chain length and it increases linearly with size. Fig. 4.16a shows what proportion of total SASA that is hydrophobic and it shows that PVA oligomers are generally less hydrophobic than PEG. As the chain length increases we observe a decrease in percentage hydrophobic SASA, and a corresponding increase in the exposure of hydrophilic groups. As the hydrophobic SASA decreases for PVA it plateaus reaching a threshold at approximately 40% (and 60% for hydrophilic SASA) once the chain length is 19 units long or more. The reverse trend is observed for PEG, which also approaches a limit of 90% for hydrophobic residues and 10% for hydrophilic SASA at 19 units long. These opposing differences in surface exposure may be important to the IRI activities of PVA and the lack of it in PEG, particularly as the 19-mer has been shown to be the point at which the IRI activity is switched on. Further still the overlap for the values in hydrophobicity is consistent with qualitative trends seen in the percentage IRI activities produced by PVA which were reported by Gibson *et al*. They report that a  $1\text{mg ml}^{-1}$  solution of PVA<sub>10</sub> yields an IRI activity of  $\sim 80\%$  while the remaining polymers, PVA<sub>19–351</sub>, all overlap at  $\sim 20\%$  using the same concentration.[TX1]

At 260 K the average TSASA of PEG oligomers increases in solution compared to values obtained at room temperature, and the percentage of TSASA that is hydrophobic increased by only 0.21–0.30% (Fig. 4.15). This increase is not reflected

in the log-log  $R_g$  of PEG and so hints at a small change in shape rather than size (Fig. 4.2). In stark contrast the predominantly hydrophilic PVA oligomers contract in colder environments and the small proportion of hydrophobic SASA shrunk as well (by 0.3–0.75%). These changes are also corroborated by a general decrease in the log-log  $R_g$ . These observations support the ideas that the hydrophobic faces or regions of amphiphilic macromolecules become preferentially exposed to cold environments, which consist of moderate levels of tetrahedrally ordered water. Recent modelling studies on the sfAFP, which is also an amphiphilic protein with a hydrophobic ice binding face. The authors, Todde *et al* [RES10], report a small, but similar size contraction in the TSASA as the temperature is decreased or is placed in the presence of an ice crystal. A 21 K drop in temperature from 298 K to 277 K produced an 4.08% decrease in TSASA, and although the hydrophobic surface area declined (from 23.4 to 22.9 nm<sup>2</sup>) as a result, the proportion of hydrophobic SASA increased from 52.9% (at 298 K) to 53.8% (at 277 K) and 55.7% (in the presence of ice). All the while, at the expense of the hydrophilic SASA. These three examples and other similar studies [RES23] combined suggests that the variations in hydrophobicity at 298 K may facilitate the contraction or expansion of the TSASA in cold water.

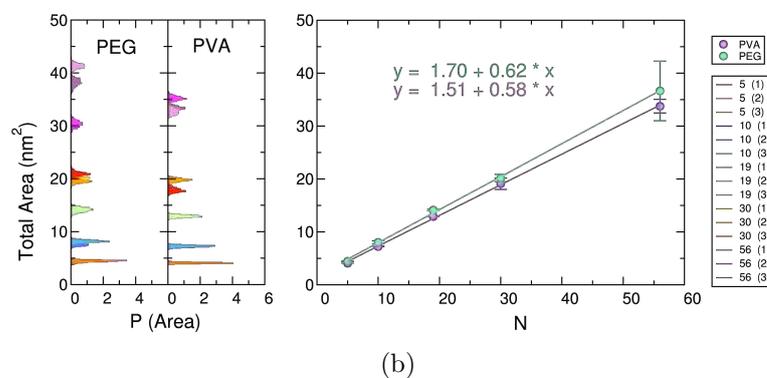
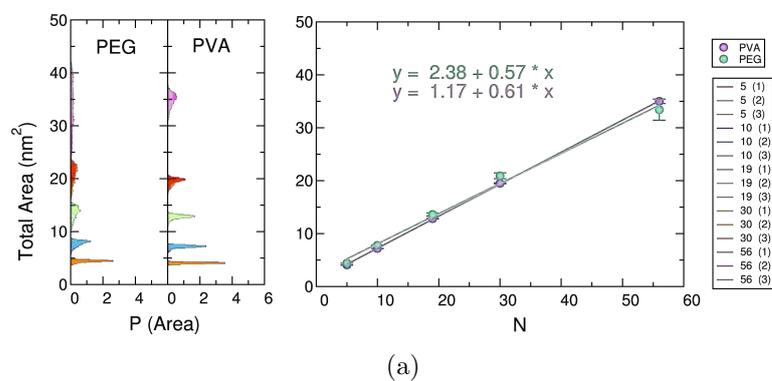


Figure 4.15: Distribution of the total solvent accessible surface area (TSASA) and plots of the average TSASA for each oligomer (4.15a) at 298 K and (4.15b) at 260 K . Error bars are standard deviations between repeats

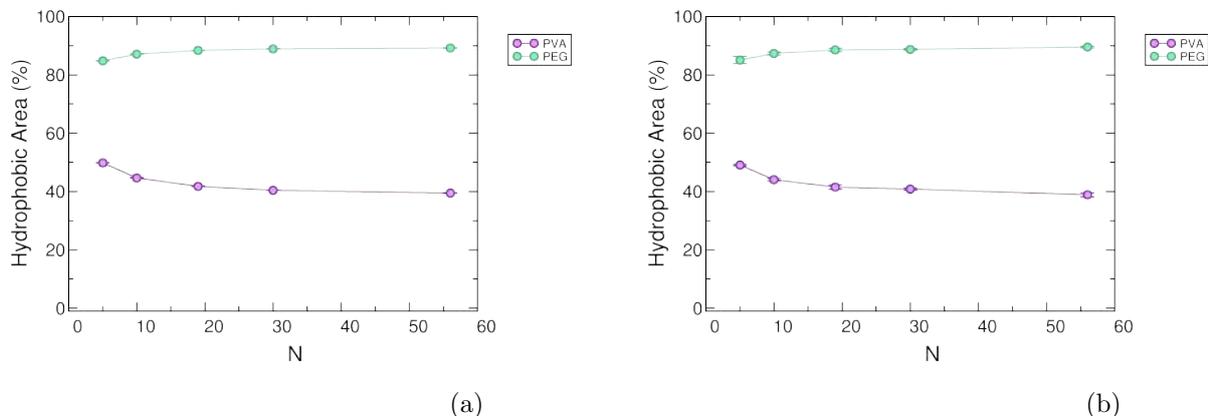


Figure 4.16: Proportion of the hydrophobic solvent accessible surface area in TSASA (in %) for each given oligomer at 298 K and 260 K. Error bars are standard deviations between repeats

Fig. 4.17 shows a snapshot of the polymers in solution at the end of the simulations in warm water. Aside from the smaller oligomers (PVA<sub>5-10</sub> and PEG<sub>5</sub>), none of the faces are particularly linear because they begin to kink and fold as the polymers increase in size. In both polymers, these bends systematically occur at the oxygen atoms for PEG or at carbon atoms carrying the hydroxyl groups for PVA. It is clear that PVA oligomers have hydroxyl groups that protrude out of the backbone to create an array of hydrophilic SASA. The PVA oligomers also appear to twist or kink in order to expose sections of 3-4 clustered hydrophilic groups and as a result produce some small regions that are flat. As the PVA chain length increases to N=56, these clusters increase and the hydrophobic groups begin to collapse at the centre with hydroxyl groups visibly decorating the outer region. PEG cannot behave in this fashion seeing as the oxygen atoms are inserted into the polymer backbone.

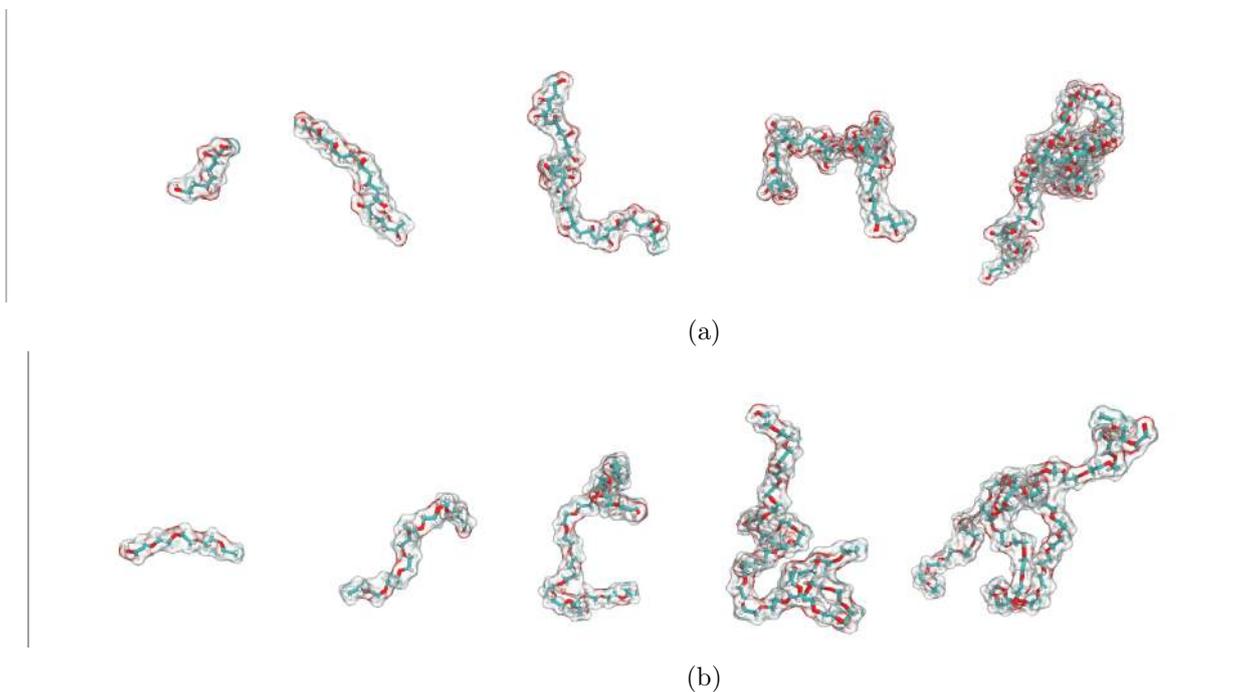


Figure 4.17: Snapshots of the polymers inside a box of 33207 water molecules, at the end of a 70 ns simulation and at 298 K . From left to right:  $N=5, 10, 19, 30, 56$ . The solvent accessible surface areas are shown by surf representation and displayed as a translucent material. Red highlights hydrophilic groups while the rest is hydrophobic. The chemical structures of the polymers are shown in detail at the center using dynamic bonds representations. Red represents oxygen, blue is carbon and white is hydrogen

#### 4.3.5 Hydrogen bonding

To further assess the influence of the polymer on the bulk water, we calculated the average number of hydrogen bonds of solvent water molecules. Fig. 4.18 shows the average number of hydrogen bonds per water molecule ( $N_H$ ) and there are no considerable changes to the ( $N_H$ ) of bulk water, irrespective of polymer type or size. PVA and PEG systems both produce values between 3.76–3.75, neither of which are substantial difference to 3.75 obtained in the pure water systems using the TIP4P/Ice water model. This means that on average, bulk water still hydrogen bonds with itself as it would normally. This holds true for both warm and cold temperatures; at 260 K , the average number of solvent hydrogen bonds increases to 3.88–3.89 for pure water systems and polymeric systems (3.89, 3.88 and 3.88 for pure water, PVA and for PEG systems).

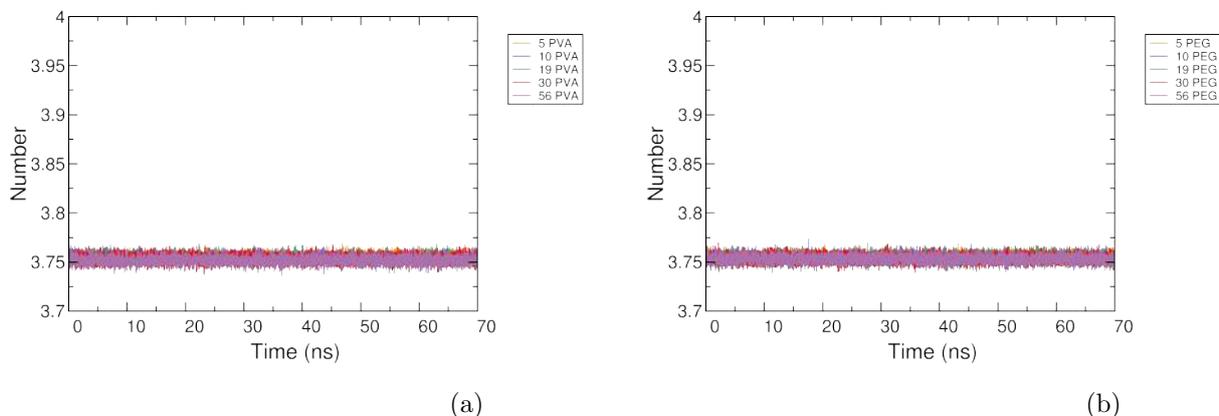


Figure 4.18: The number of solvent-solvent hydrogen bonds at 298 K

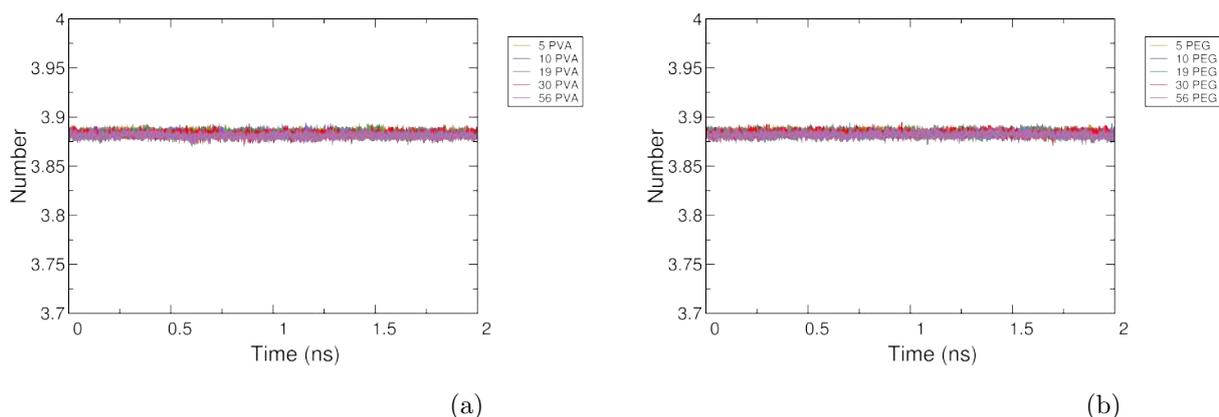


Figure 4.19: The number of solvent-solvent hydrogen bonds at 260 K

As mentioned earlier in the report, the inhibition of ice growth by antifreeze macromolecules was shown to involve some hydrogen bonds and it was first believed that antifreeze mechanism was dominated by hydrogen bonding interactions. From our assessment of the SASA of both polymers it appears as though this is the case. Thus, we also examine the importance of intramolecular hydrogen bonds formed by the polymer, and the number formed between the polymers and water. The results are shown in Fig. 4.20 and Fig. 4.21 respectively. Unsuprisingly, PVA is involved in several more hydrogen bonds than PEG at each chain length and in both graphs. The reason for this is that it carries two atoms capable of hydrogen bonding; one

of which is a donor (O) and the other an acceptor (H), while PEG only carries the former. In the PVA systems the number of polymer-solvent hydrogen bonds range from 21.60–224.71 while in PEG it ranges from 11.18–84.42, as a result. At 260 K these values only increase slightly to 22.92–229.55 and 9.28–86.17 respectively.

Despite having fewer solvent-polymer interactions, the number of intramolecular hydrogen bonds in the PEG oligomers are almost non-existent regardless of size, but it grows linearly with chain length for PVA. In both polymers an increase in size correlates positively with the level of hydrogen bonds because there are additional hydrophilic groups available to interact with itself or the surrounding waters. These findings suggest that the structure of PVA experiences greater internal stability due to a higher amount intramolecular interactions. Thus far these results are consistent with the narrower  $R_g$  distributions for PVA, and restrict the oligomer to fewer conformations. In contrast the reverse is true for PEG because it is weakly held together. These larger restrictions on conformation combined with the ability to hydrogen bond more with its solvent (more so than PEG), means that waters can be more easily drawn into small successive shells.

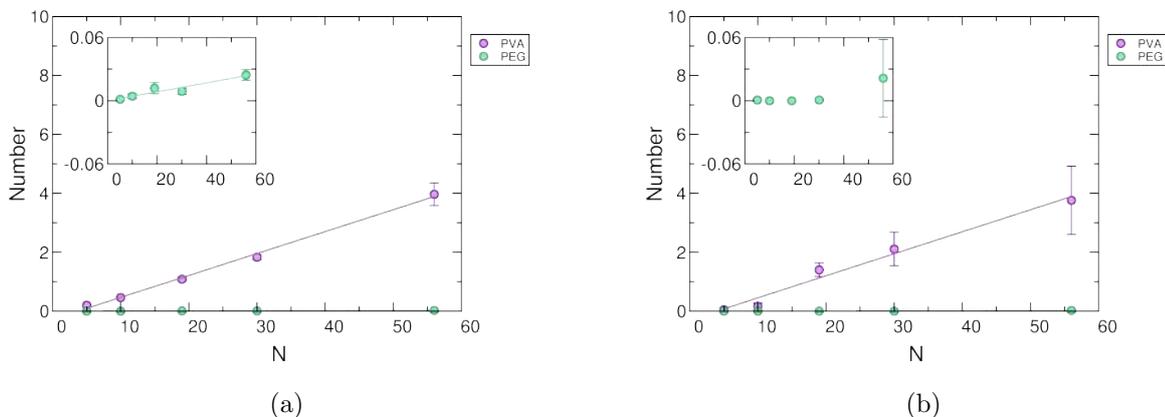


Figure 4.20: The number of intramolecular hydrogen bonds in PVA and PEG oligomers at (4.20a) 298 K and (4.20b) at 260 K. In colder simulations a non-linear relationship exists between the number of intramolecular hydrogen bonds and the chain length of PEG oligomers. Error bars are standard deviations between repeats

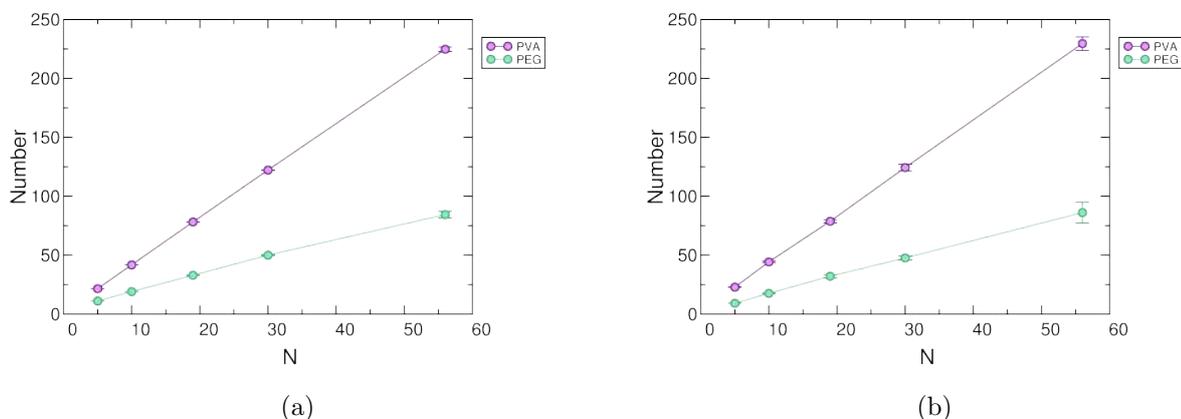


Figure 4.21: The number of intermolecular hydrogen bonds between PVA or PEG oligomers and TIP4P/Ice water (4.21a) at 298 K and (4.21b) at 260 K. Error bars are standard deviations between repeats

At 298 K the number of polymer-polymer bonds for PVA oligomers ranged from 0.20–3.98 and decreased to 0.05–3.77 at 260 K. At both temperatures, oligomers smaller than than  $N=19$  have values less than 1, which may be significant for antifreeze activity. Especially as PEG, which is antifreeze inactive, consistently has values below that threshold at both temperatures (0.00–0.02 at 298 K and 260 K). The trends reported here for hydrogen bonds are qualitatively similar to those reported in simulation studies of antifreeze molecules. [RES10] [] [] Notice that in this study we applied a simple geometric criteria for hydrogen bonding by using 0.35 nm and an angle cut-off of  $30^\circ$  deviation from an ideal Donor-Hydrogen-Acceptor angle of  $180^\circ$ . We acknowledge that hydrogen bonding criteria used may vary from study to study. As a result it means that direct comparisons to AFGP studies cannot always be made, particularly as our antifreeze macromolecules are not similar at all in size nor in the number of water molecules used, forcefields used, or the types of water models selected. [RES10][RES14][ANA1][ANA3] Our previous chapter has consistently demonstrated the level of hydrogen bonding in a system is very much dependent on the water potential used to model the interactions of the water solvent. To our knowledge TIP4P/Ice has not previously been used in AFP/water studies nor in AFP/water/Ice studies. They have however been used for the investigation of AFPs with the TIP4P/Ice/Methane clathrate studies produced by Bagherzadeh *et al* in 2015 [SOLO2].

### 4.3.6 Polymer flexibility and movement

To determine the structural deviations of atoms within the oligomers, the root mean square fluctuation (RMSF) of each atom was measured for all simulations (Fig. 4.22). The RMSFs of atoms indicates that the terminals of PVA and PEG oligomers are generally more flexible than the centre of the polymers. As the chain length decreases the flexibility of the oligomers declines rapidly and is also visually reflected by the standard deviations between repeats samples (displayed as error bars). The RMSF highlights which sections moved throughout the simulation, whilst the standard deviations reflect how much these regions can move. This flexibility is evidenced by multiple peaks and troughs, also highlighted by the error bars and in both polymers we find that oligomers with fewer than 19 units are not long enough for such features to become available.

The main difference between the two types polymers is that the PEG oligomers have greater RMSF values on average, however PVA oligomers have well defined regions of greater flexibility (taller troughs) and rigidity (deeper valleys). Both of these types of well defined regions span for longer across the PVA polymers. These differences in regions are most clearly seen in comparing PVA<sub>56</sub> and PEG<sub>56</sub>. These regions of pronounced flexibility and inflexibility reveal that the PVA<sub>56</sub> polymers maintain two hinges with approx. three dedicated areas that flap about. In PVA the rigid regions are approximately 10–15% units of the full chain length and in PVA<sub>56</sub> this translates to roughly 8–9 units long. If we treat the valleys as partitioning lines we can see that PEG<sub>56</sub> does not exhibit any long lasting regions of rigidity that could be described as hinges, however it does have several more segments. Frequent but small kinks or segments are promoted by oxygen atoms that lie in the PEG backbone. Alongside the effects of weak intramolecular hydrogen bonds, these segmental movements encourages PEG oligomers to create more conformations than PVA, discussed earlier. In both PVA and PEG the regions of lowest flexibility are typically half the amount of the maximum values. For example in PVA<sub>56</sub> the terminal ends exhibit RMSF values of 1.0 nm while the regions of lowest flexibility reach 0.5 nm. On closer inspection we found a fractal pattern, that revealed hydrogen and oxygen atoms moved the most, while carbon atoms moved the least. With regards to the ice lattice matching theory these movements would be ideal because it allows PVA oligomers to compensate for any small deviations from the ice-lattice spacing so that the atoms could orient accordingly. The atoms which moved the most in PEG oligomers however are hydrogen bonds and all other atoms moved less.

Fig. 4.22c–4.22d shows the same RMSF analysis for PVA and PEG oligomers in supercooled water at 260 K. Unsurprisingly we find that the flexibility for all

polymers declined dramatically in colder solutions, due to less available thermal energy. If present at all in PVA, small segmented regions of flexibility and inflexibility in extend for shorter periods and do not vary much from the baseline, unlike the studies at 298 K . Some additional features we find at reduced temperatures are that the average baselines for PVA<sub>5</sub> and PVA<sub>10</sub> is lower than 0.1 nm (0.062 nm and 0.087 nm), while larger PVA oligomers are at least 0.13 nm (0.13–0.16 nm for PVA<sub>19–56</sub>). These larger values are comparable to results reported for AFPs. For example, RMSF analysis of the 6.5 kDa snowflea AFP (sfAFP) was recently conducted by Todde et al [RES10] at 298 K and 277 K in a box of 1000 TIP4P waters, and trajectory of 100 ns. They found that aside from the C-terminals, which had the highest RMSF values, the sfAFP backbone had two distinct regions of high flexibility (approx. 0.25 nm and 0.5 nm) at room temperature and an average baseline value of 0.1 nm. [RES10] These values are consistent with reports in literatue [RES14] and they also report a similar trend in flexibility as a function of temperature. It is noteworthy that sfAFP is an insect AFP and therefore has a high TH, which is absent in PVA however what both molecules have in common is they are both IRI potent [RES15]. It is also important to note that PEG does not display either of these properties nor the temperature dependent periodic fluctuations or about the baseline found in PVA. Instead the average RMSF is reduced from 0.19–0.97 to 0.16–0.23 nm but doesn't decline lower than 0.16 nm. Although the mass of the sfAFP coincides with approximately 148 units of PVA, the flexibility of the protein is low compared to our synthetic polymers because its RMSF values are only comparable to ranges reported for the smaller oligomers (PVA<sub>5–30</sub>) in both warm and in cold temperature studies.

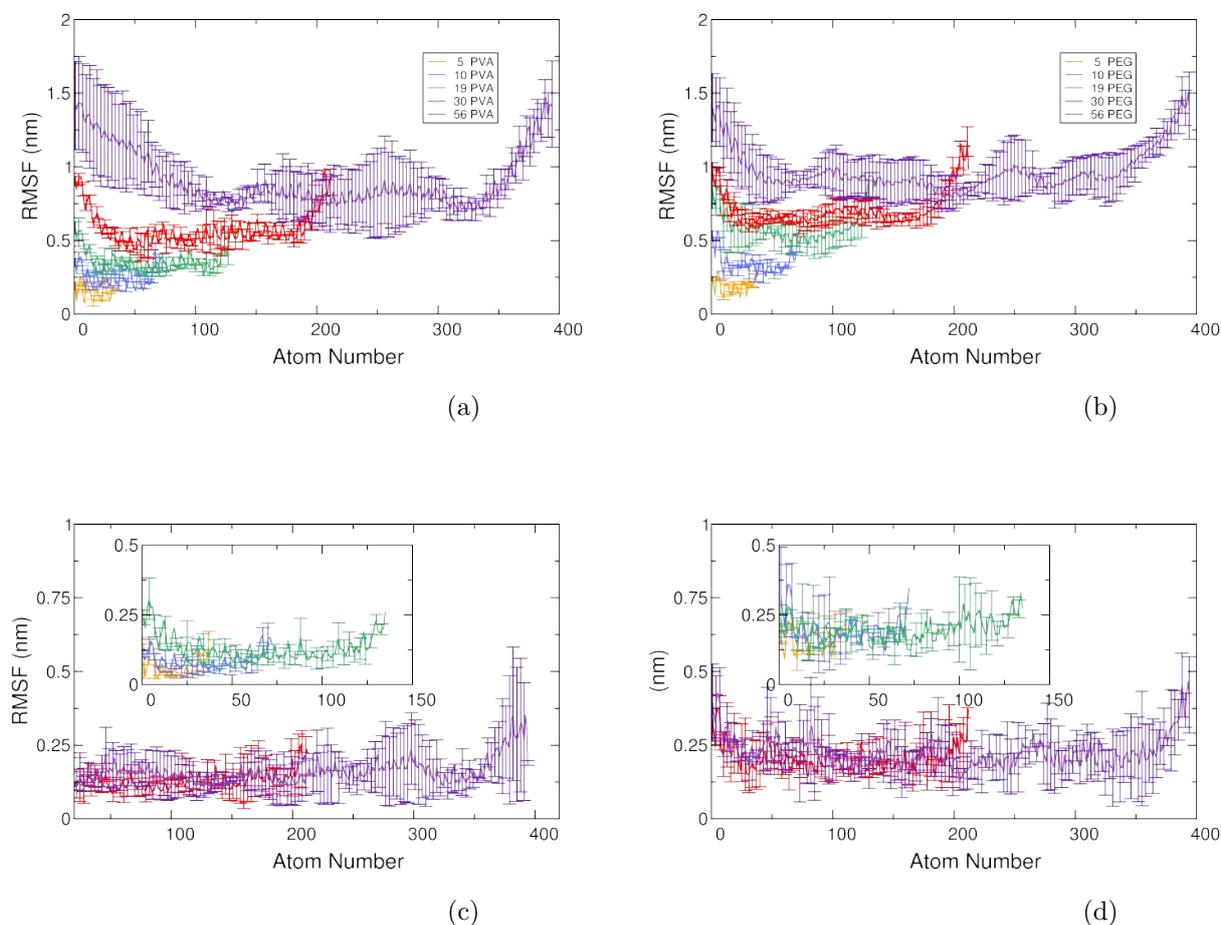


Figure 4.22: The relative fluctuation of each atom on the oligomer with respect to starting positions at 298 K (top panel) and 260 K (bottom panel). PVA oligomers preserve long stretches of important regions of low and high fluctuations at 298 K which is lost at 260 K. PEG is more flexible than PVA and also becomes rigid with a decline in temperature. Note that the polymers were not aligned prior to the RMSF calculation

## 4.4 Discussion

Thus far the results from our solution studies, especially the O-OH RDF separation distances support some form of complimentary ice matching mechanism in PVA. The oxygen-oxygen RDFs distances suggests that PVA can indeed lattice match and could potentially become incorporated into the ice lattice effortlessly because PVA already adopts the appropriate conformations in solution at room tempera-

ture. It will be interesting to see if this, or any other distribution, changes during the ice/polymer simulations.

After exploring the concept of specificity and ice lattice matching in solution we find that the  $Q_4$  results suggests that PEG<sub>56</sub> and PVA<sub>56</sub> polymers subtly disorder the surrounding water molecules at 298 K, however these changes are miniscule. In addition the hydration numbers for PEG are consistently higher than for PVA. Combined, there is no evidence for substantial ordering (or disordering) of surrounding water molecules into coherent tetrahedral arrangements or structures within a 0.35 nm radius from PVA oxygen atoms. Thus the hypothesis that synthetic PVA oligomer polymers poison the the ice/water interface via its own clathrate is unsupported by preexisting activity in solution at 298 K. This also remains true for studies conducted at 260 K, where water molecules are supercooled enough to crudely represent an unforzen ice-like region

It has become clear that the two polymer function through two very different ways in solution. PEG exhibits greater reliance on its hydrophobic interactions to interact with nearby waters while PVA uses hydrophilic interactions. For instance there are also many indications that PVA adopts and maintains a narrow range of, rigid, random coil conformations. Analysis which support these conclusions include the radius of gyration, RDFs ( $O-O_H$ ,  $O-H_H$ ,  $O-C_H$ ), the hydration numbers and RMSF calculations. Hydrogen bond analysis reveals that this is largely a result of hydrogen bonds to stabilise their internal structure of the polymer. In contrast PEG cannot act in this way and so is able to vary a lot in conformation. It is possible that a restricted or restrained structure may be important for lattice matching interaction with ice and other interactions. This is almost analogous to the importance of a protein with a well defined secondary structure in order to function properly. For instance, the folding or collapse of the PVA oligomers with increasing length allows clustering of neighbouring hydroxyl groups, which eventually bristle the larger oligomers and in the case of medium sized oligomers, they form small segments of arrays. A decrease in temperature from 298 K to 260 K encourages small contraction which, promotes this further.

The RMSF analysis is a testament to the versatility of the PVA oligomer, and its ability to tailor rigidity as a function of temperature. IRI potent PVA oligomer have characteristic regions of high and low flexibility with an apparent 0.13 nm RMSF as prerequisite in 260 K. The ability form regions of high and low flexibility

provides an allowed range of movement necessary for the IRI mechanism. For instance the polymers ability to bend or fold may be relevant features for alignment in lattice matching, regulation of the hydration sphere, or it may be crucial for selective exposure of specific faces in order to encourage select interactions as reported by Tam *et al* .[ANA1] Alternatively the mechanical motion of folding or flapping could disrupt the local ice crystal growth for an extended period of time, until adjacent regions are overgrown and the polymer is submerged in the ice lattice. Irrespective of the seemingly endless and possible reasons, the importance of these features is apparent across the literature for AFPs [RES10][RES14][CCC7][CCC3][CCC12][RES15] and for the simulations of glycopeptides [ANA1][ANA3]. In these referenced work, characteristic folding properties have also previously been reported for IRI potent molecules whilst the features are lacking in the IRI inactive molecules.

Although one can say that the PEG has greater number of waters around it, the PVA is better able to “control” its water than PEG due to its greater hydrogen bonding network and a delicate balance of exposed hydrophilic and hydrophobic regions, for which we have assign a 60:40 cut-off based on PVA<sub>19–56</sub>. As a result, PVA has a access to fewer conformations, which are restricted to similar shapes or sizes and this is strongly mirrored by its hydration shell, as evident by the narrow distributions of both hydration numbers and the well defined RDFs. The presence of an ice interface ordinarily induces a gradient of ordered water, [111] and so cold water alone is not wholly encompassing and a suitable representation of the *whole* QLL. We hypothesise that the ability of PVA to control surrounding water molecules may have greater implications for specific regions of the gradient, which have not been illustrated here. It is likely that large PVA oligomers could achieve this by disordering waters in the QLL that lie close to ice because these waters are weakly ordered by the ice/water interface. Waters in the QLL that lie further away from the ice will be even more weakly held in place and so these waters are better represented by the cold water studies. Due to their increased mobility, those waters may not be particularly influenced by the presence of the PVA oligomers. Similarly, waters at room temperature have enough kinetic energy to disperse away from the polymer’s sphere of influence, and this too is reflected by negligible changes in the Q<sub>4</sub> results.

Granting that the solution studied did not find substantial amounts of disordering or ordering of surrounding water molecules, we cannot rule out the possibility that PVA may significantly disorder waters in specific regions of the QLL. Particularly, as there is evidence to suggest that both types of polymers in warm and and cold temperatures hint at a decline in the Q<sub>4</sub> = 1 positions rather than an

increase, and lattice matching would prove useful for such a mechanism. Unlike the AF(G)Ps, our simulation studies of PVA in warm and cold water do not appear to have a single clear and definite hydrophilic or hydrophobic faces to interact with ice or water and this suggests that they function through a very different mechanism. As the temperature drops from 298 K to 260 K, the TSASA of PVA decreases and the polymer becomes more rigid. As a result PVA can more strongly govern an albeit smaller cluster of hydrating waters when the polymer resides in the QLL regions that are close to the ice interface.

## 4.5 Conclusion

We have established that the PVA simulations are suitable at 298 K and 260 K using the OPLS-AA forcefield, and the TIP4P/Ice water model to simulate model systems in solution. Thus far we find several distinguishing features between PVA and PEG which may be important for IRI activity in warm and cold simulation studies. Our hypothesis from these MD simulations are that the PVA polymer interact directly with the ice interface or with layers of the QLL very close to the ice lattice. They do so via their hydrogen bonds, as evidenced by their strong lattice matching features and hydrogen bond analysis. PVA appears to confine water close to itself. Narrow and well defined successive shells, and the  $Q_4$  parameter indicate that despite these confined regions, the waters are somewhat reduced in tetrahedral order but not more so than PEG. Overall our results indicate that hydrophilic interactions are most dominant with PVA while PEG largely interacts with hydrophobic interactions. When comparing the results of warm simulations to the cold ones we find that there is no real difference except for the rigidity, compactness and hydrophobicity of the polymers. These vary between the two polymers and there is evidence to suggest that the balance of properties like rigid/flexible regions, and of hydrophobic/hydrophilic groups is important for switching on IRI activity at 19 units long. We identified a cut-off at 40% hydrophobicity and 60% hydrophilicity which is earmarked by PVA<sub>19</sub>. Without a library of IRI active, synthetic polymers, it is difficult to say with certainty if the percentage hydrophobicity is important for IRI activity, or if the values themselves that are most relevant. Largely because PVA and PEG have similar TSASA values, while other polymers may vary enough in TSASA to confirm if the reported trend exists. In both temperatures, the bulk water-water hydrogen bonding network remains largely unaffected. Early simulation studies of AF(G)P in solution found that the MD study revealed hinges or twists in the helical structure of the

polymer using different water models.[CCC7][CCC3][CCC12]The same was found for later and recently improved modelling studies on analogues.[RES10][RES14][RES15] We too find that the IRI active PVA can produce hinge like regions that may be useful for aligning with different faces of ice. The heightened flexibility of hydrogen atoms and oxygen atoms, means that the PVA oligomers are also appropriately designed and spatially arranged for their functions. Using trajectory snapshots, we have revealed that regions of the PVA and PEG oligomers seem relatively flat, however only PVA has lattice matching groups that stick out more to interact with ice. Although the PVA oligomers twists to accomodate arrays of hydroxyl groups, and in spite of their isotactic structure these arrays are not restricted to largely one side of the polymer. Unlike AF(G)Ps, the known IRI active PVA oligomers do not appear to be divided into a largely hydrophilic and hydrophobic face, it is apparent that the two types of macromolecules may function in a different manner.

## Chapter 5

# The effect of single poly(vinyl alcohol) and poly(ethylene glycol) polymers on ice crystal growth rates

### 5.1 Introduction

Earlier in this thesis, we studied PVA and PEG polymers in solution and found some properties of interest that are chain length dependent, in a similar fashion to IRI potency. These properties are the RMSF and the hydrophobicity of the polymers. In this section, the aim is to communicate the effects of single oligomer chains on the secondary prism plane during the freezing and melting process using molecular dynamics simulations. To effectively model this system we previously studied the ice growth in the absence of any additives and investigated the effects of system size, 2 cut-offs and varying potential forcefields. Here we continue to use the direct coexistence methods which places the ice in contact with the water to produce two ice/water interfaces, with an immersed oligomer. As a result we can monitor and analyse the change in rates and dynamics for both melting and freezing processes. Ultimately this will help us to resolve some question surrounding the interactions between the oligomers and changing ice front.

In order to study crystal growth it is custom in literature to use direct coexistence method, and doing so will make our results more comparable to recent studies which probe the effects of AFPs on ice crystal growth[RES4] [RES5] [RES6] [RES7] [RES8] [RES10] [RES11] [RES12], especially as no such simulations of polymers have

been made yet. Based on the previous 3, we have established a computational protocol, which reproducibly simulates the freezing of an ice crystal in contact with liquid water. In this new section, we apply the previous work to molecular dynamics simulations of IRI active and inactive polymers in the presence of a larger size crystal. We chose to study ice growth instead of nucleation because the latter is spontaneous and will be more computationally demanding to observe. This is due to the energetic cost of forming a new interface, and so it may not be possible to observe nucleation under the simulation timescales accessible to us for atomistic studies. Another reason, mentioned earlier, is that PVA has been shown to inhibit the ice nucleation as well as growth [TX12] even in the presence of nucleators like bacterial proteins [TX13] and silver iodide [TX14]. Interestingly studies of the ice nucleating silver iodide revealed that inhibition of nucleation was also possible from other known antifreeze active macrostructures such as AFP (TH active) and poly(vinyl pyrrolidone) (PVP) (TH inactive). Surprisingly, inhibition was also achieved using antifreeze inactive polymers, poly(ethylene glycol) (PEG) [TX14]. MD simulations, are thus designed to reflect our interest in ice crystal growth and its inhibition, in order to investigate the mechanism for IRI activity, because PEG is known to not show any appreciable IRI activity. These are in line with observations by Gibson *et al* , that IRI activity is a consequence of slowed growth rather than nucleation [A25]. Congdon *et al* also showed that IRI activity of PVA switches on between  $N=10$  and  $19$  [TX1]. To keep our studies in line with experimental observations, isotactic PVA<sub>10</sub>, PVA<sub>19</sub>, PVA<sub>10×2</sub> and PEG<sub>10</sub> oligomers were chosen for our studies in order to better understand the difference between IRI activity and inactivity, as well as scales of IRI potency that are introduced by greater chain lengths and concentrations.

## 5.2 Methods

### 5.2.1 Forcefield parameters

The TIP4P/Ice water model was used along side the the OPLS-AA forcefield as with 3. The same forcefields were used for the polymers are reported in 4.

### 5.2.2 Simulation parameters

Simulations were preformed using the Gromacs (v4.5.4) package. Once the equilibration was complete, restraints on the ice were removed and the production runs were conducted at a range of 10 different temperatures (300 K , 290 K , 280 K , 275 K

, 273 K , 272 K , 270 K , 268 K , 260 K and 250 K ). A short timestep of 1 fs was used and the anisotropic pressure coupling was applied in order to permit independent fluctuations of the box in all directions during phase change. The compressibilities of the X and Y planes were set to those of real ice, while the Z direction was set to the higher compressibility of real water (See 3). The Parinello-Rahman barostat was used to maintain the pressure at 1 bar and the Nosé-Hoover thermostat was also used to control the temperature. All box angles were kept orthogonal and 1.00 nm cut-offs were applied to the short range Lennard Jones potentials and coulombic interactions. Long range electrostatics were controlled using the PME protocol using a 0.16 nm grid and a cubic interpolation. LINCS constraints were applied to all hydrogen bonds and periodic boundary conditions were used to mimic bulk conditions. Pure water simulations were repeated ten times for each temperature, while only nine repeats were made for the simulations conducted in the presence of all polymers [except the PVA10×2 which only has [six]]. The progress of each simulations was monitored by plotting the convergence of the total energy against the trajectory time. A negative drift in the total energy is indicative of freezing and a positive drift depicts melting of the system.

### 5.2.3 Initial configuration and Equilibration

A simulation box of hexagonal ice (768 water molecules) which was equilibrated in 3, was multiplied by two in each direction to create a box of 6,144 water molecules which would represent the ice phase. The dimensions of the ice crystal was  $5.8 \times 6.2 \times 5.5$  nm . The box dimensions were  $5.8 \times 6.2 \times 16.5$  nm and the ice lattice was placed in the centre of the box. If included, the starting configurations of the polymers used were obtained from the final 70 ns trajectory from the solution studies (3). The polymer were placed at approximately [0.2 nm ] away from the secondary prism plane (the 1120 face) [CITE 80] which is a plane of interest and is also reputed as the fastest growing plane. Next, the system was solvated in order to create an ice/water interface on either side, and interstitial waters were removed from the ice lattice. A total of 12,363 waters represented the fluid phase and a schematic of the simulation set up is shown in Fig. 5.1.

The same protocol as in 3 was used to simulate the melting and freezing of these larger systems. The new system contained 18,507 molecules and an NVT and an NPT run were conducted in sequence. The conditions were maintained at 300 K and 1 bar with  $10,000 \text{ KJ mol}^{-1}$  restraints placed on the ice. This allowed the water to equilibrate around the ice, as well as the polymer if included. Finally an NPT run was continued as the temperature is dropped to 283.15 K and the whole system is

allowed to equilibrate at supercooled conditions. Equilibration was monitored using two properties; the total energy and the Z-directional movement of the polymer's centre of mass. Once both of these features become stable then the production runs were started.

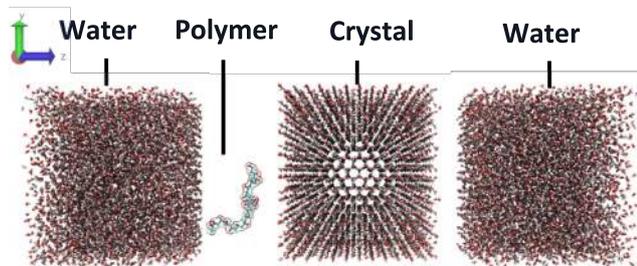


Figure 5.1: Schematic of the simulation system and the geometry of the ice and water positions.

## 5.3 Results

### 5.3.1 Total energy profiles

Figure 5.2-5.5 illustrates the evolution of the total energy over the duration of our simulation at 10 different temperatures, in the presence and absence of any polymers. In the absence and presence of any additives, we determined a melting point of 270 K, because the system generally remained stable and neither melting nor freezing was completely achieved for 300 ns. This is only  $1c_{irc}$  greater than for the smaller systems in 3. We can see that as the temperature declines, the time taken for the completion of freezing or melting process increases and so it become computationally demanding to study at lower temperatures. This occurs due to a decline in the kinetic energy required for the water molecules to diffuse about, in order to break away from the ice lattice or to reorient themselves accordingly as is necessary to build the growing ice lattice. To overcome this challenge recent studies which also incorporate an antifreeze macromolecule into the ice crystal during the movement of the ice/water interface, typically run simulations for shorter trajectory times. They choose to infer a phase change from the initial drift over just a few nanoseconds e.g. 0-10 ns [RES10][RES11]. In our pure water samples, we can see that there is no real harm in doing so. Using the initial drifts is accurate and the method is computationally efficient. However in the presence of antifreeze macromolecules it is not advisable because it is important to consider that the freezing simulations can take longer

than 300 ns and so it is possible miss important features that occur as a result of adsorption during the remaining 290 ns trajectory. Take for instance a two stage adsorption process which may be earmarked by a intermediate plateaus instead of just the single and final one. In the presented profiles select IRI active and inactive polymers both occasionally exhibit intermediate plateaus at 268 K, that are non-existent in the pure water studies. These plateaus last longer for the antifreeze active polymers with 52 ns, 125 ns, 114 ns and 182 ns for PEG<sub>10</sub>, PVA<sub>10</sub>, PVA<sub>19</sub> and PVA10×2 respectively. If these plateau features only occurred in one repeat, then they are highlighted in dark blue however if similar features have been identified elsewhere then they are highlighted in black as a representative sample would be. Notice that these plateaus not only last longer in IRI active polymers, but occur in other repeats of PVA<sub>19</sub> and PVA10×2. In some repeats they even occur multiple times for brief periods. These features indicate that larger and more PVA polymers somehow stabilise the growing ice lattice. It also appears that the inhibition of ice could occur in stages or in a fashion similar to the successive coordination associated with multidentate ligands, which is unsurprising.

### 5.3.2 Growth/melt kinetics

The time taken for the system to completely freeze or melt was recorded in the presence of each polymer and the results are summarised in Fig. 5.7. As with the pure water systems the time taken for the freezing, and for melting at temperatures closest to the  $T_m$  took the longest time. These temperatures also experienced the greatest delay (275-260 K), whereas at high temperatures (310-280 K) the delays were much smaller and there was generally no real difference between any of the types of polymers.

At four temperatures, 250 K, 260 K, 268 K and 270 K, a clear difference between the IRI active and inactive polymers appears. Firstly, no freezing or melting was typically observed at 270 K however PEG was the only polymer to have a single repeat that froze at 295.177 ns. At 260 K there is no significant difference between the time taken for the pure water samples to freeze on its own, and in the presence of PEG<sub>10</sub>. Both systems take ~52-53 ns whereas the addition of PVA<sub>10</sub>, PVA<sub>19</sub> and PVA10×2 delay the simulation by roughly same amount to ~69-70 ns. As the temperature is dropped even further to 250 K the the weakly IRI active and non-active oligomers — PVA<sub>10</sub> and PEG<sub>10</sub> respectively — freeze prematurely. These results are reminiscent of explosive growth rates that have previously been reported for AF(G)Ps outside of their TH range or range of antifreeze activity. Interestingly, the more potent IRI inhibitors (PVA<sub>19</sub> and PVA10×2) take the same amount of time

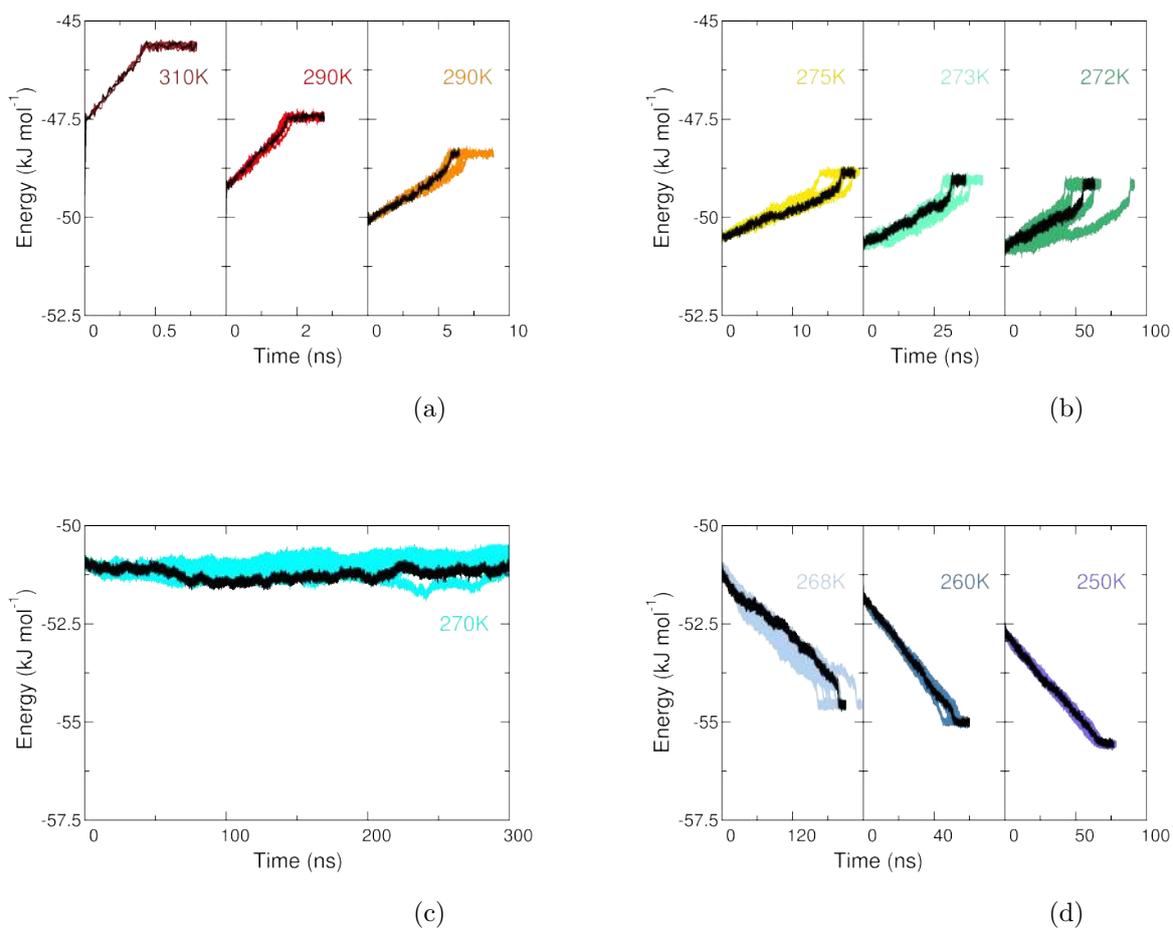


Figure 5.2: The evolution of the total energy per water molecule as a function of temperature. Complete melting is observed above 270 K (red, yellow, orange and green) and complete freezing occurs at temperatures below 270 K (blue). The  $T_m$  of TIP4P/Ice is identified as 270 K (cyan). Results for all ten repeats are included in each graph and a single representative is highlighted in black for clarity.

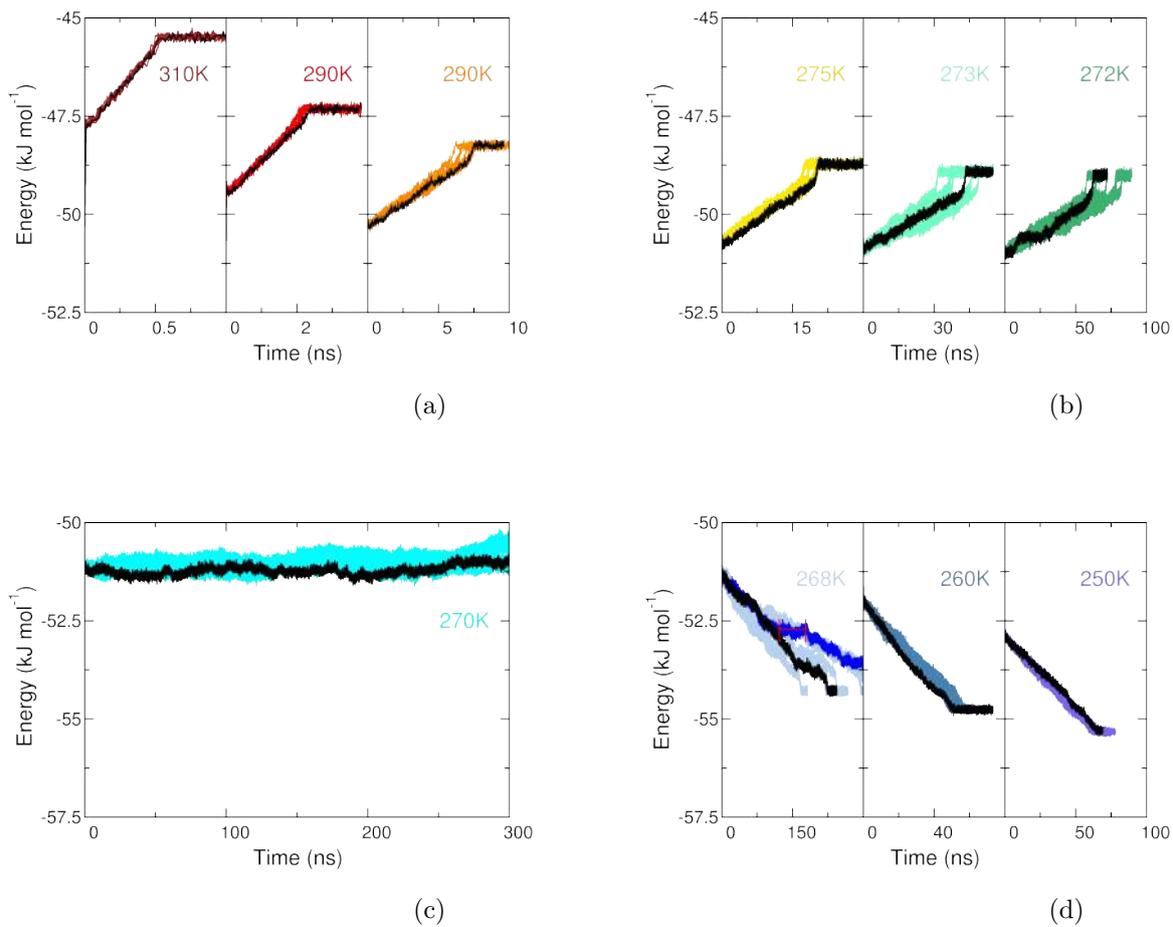


Figure 5.3: The evolution of the total energy per water molecule as a function of temperature for PEG<sub>10</sub>. Complete melting is observed above 270 K and complete freezing occurs for most simulations at temperatures below 270 K (blue). The  $T_m$  is still 270 K. Same colour key as in Fig. 5.2.

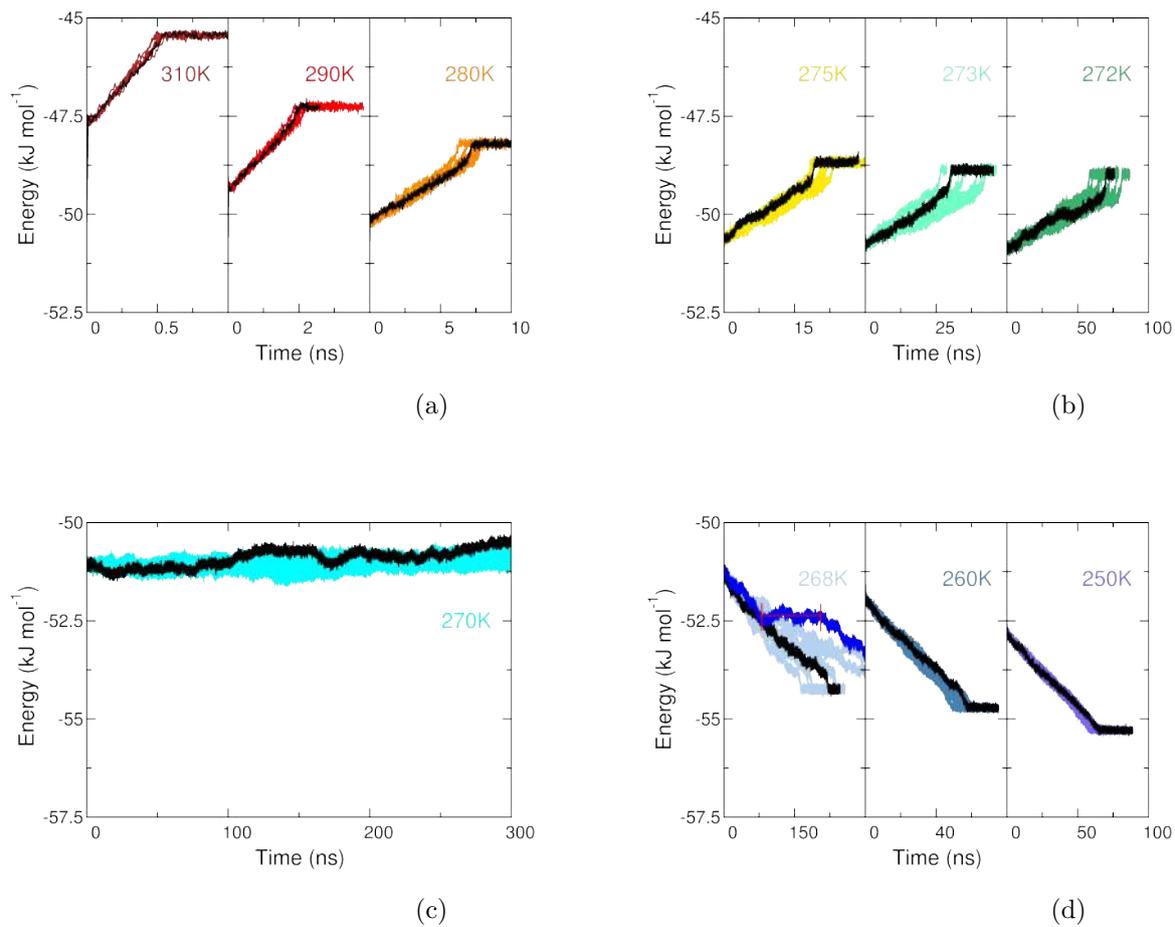


Figure 5.4: The evolution of the total energy per water molecule as a function of temperature for PVA<sub>10</sub>. Complete melting is observed above 270 K and complete freezing occurs for most simulations at temperatures below 270 K (blue). The  $T_m$  is still 270 K. Same colour key as in Fig. 5.2.

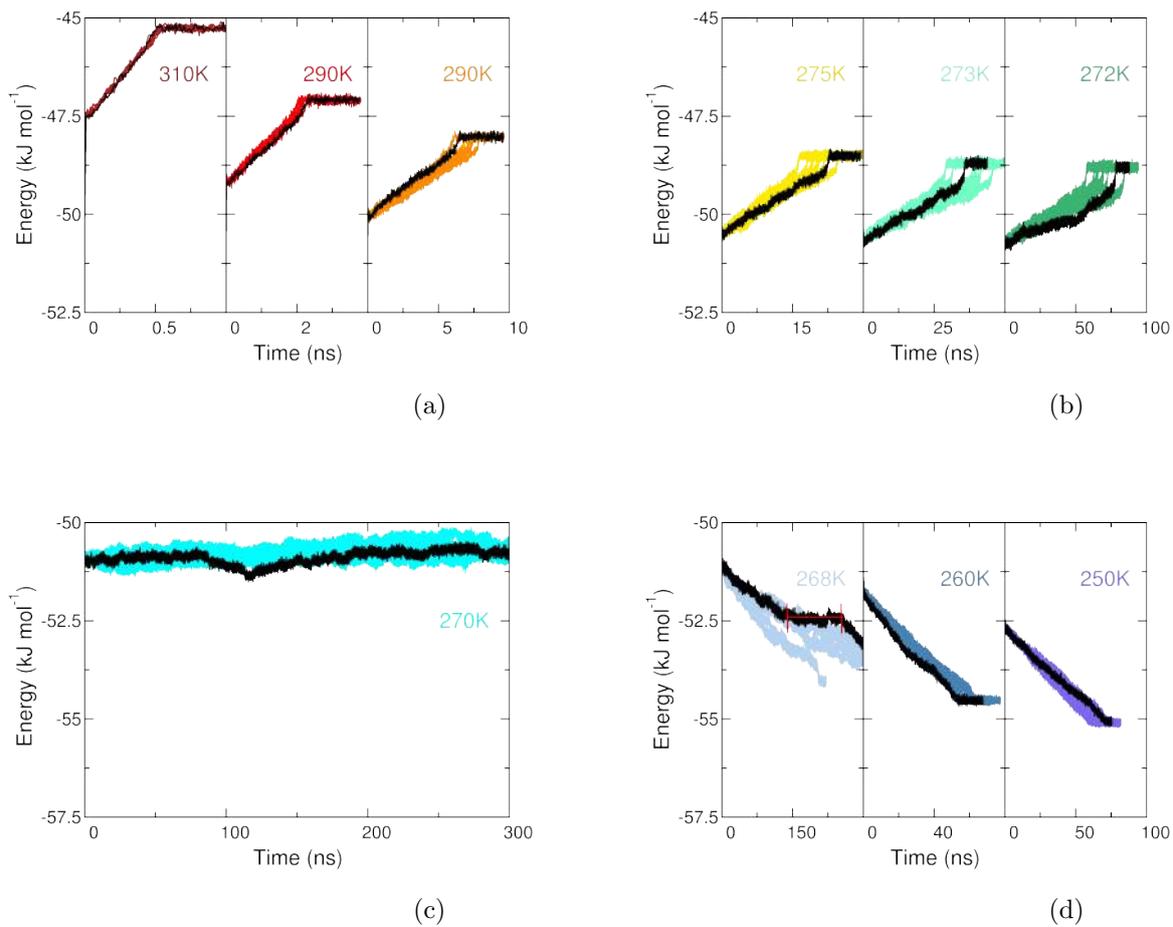


Figure 5.5: The evolution of the total energy per water molecule as a function of temperature for PVA<sub>19</sub>. Complete melting is observed above 270 K and complete freezing occurs for most simulations at temperatures below 270 K (blue). The  $T_m$  is still 270 K. Same colour key as in Fig. 5.2.

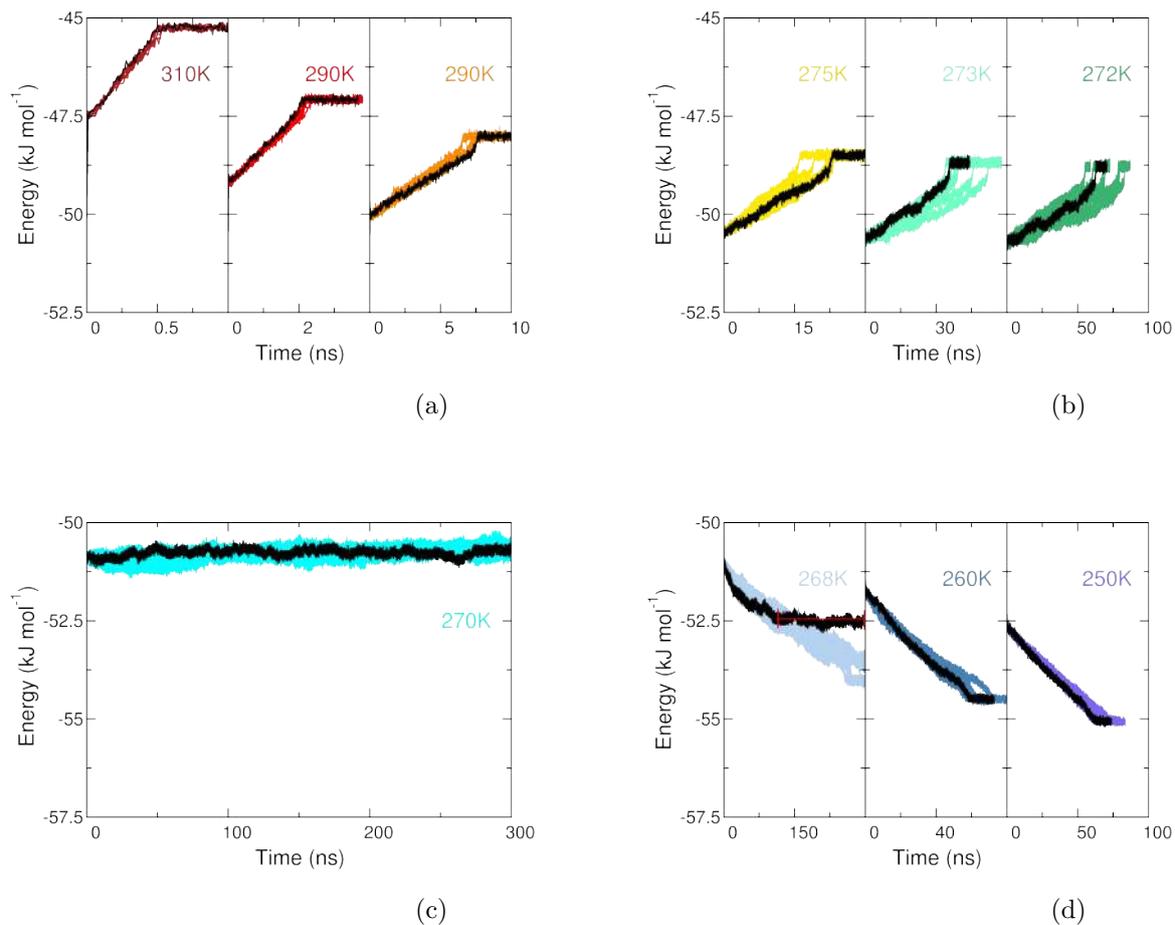


Figure 5.6: The evolution of the total energy per water molecule as a function of temperature for PVA10x2. Complete melting is observed above 270 K and complete freezing occurs for most simulations at temperatures below 270 K (blue). [The  $T_m$  is still 270 K]. Same colour key as in Fig. 5.2.

as for pure water systems to completely freeze, or slightly more. For comparison, Vrbka and Jungerwith found that at 250 K , a 0.15 M brine solution and yielded a 20% delay in freezing time compared to pure [SALT1] water which froze after 250 ns , compared to 300 ns . Similarly, Kuiper *et al* reported that a single [IRI and TH active (TH=6 K )][AX25]/inactive but TH active (TH=6 K )][AX25]] swbAFP at roughly the same temperature, (255 K ) delayed the freezing time by 67% from 150 ns in pure TIP4P water, to 250 ns .[RES12]. These suggest that the polymers function in a manner different to both salts and swbAFP.

At 268 K we found that the system typically achieved complete freezing for all 10 simulations under an average of 185.82 ns ( $\pm 27.8$  ns stdev). The presence of polymers did not change the melting point of 270 K , however fewer simulations at 268 K were completed within 300 ns in the presence of both PVA and PEG. These finding are tabulated in Table 5.1. Seven of nine simulations were completed with PEG<sub>10</sub>, and even fewer were completed in the presence of PVA (between two-six) in correlation with increasing IRI activity. This suggests that PVA oligomers are capable of considerably delaying the time taken for freezing simulations only, provided that the temperature is close to the melting temperature, however the freezing point is not noticeably reduced as a result because there is still a visible downward drift that is typically observed within the timescale of the simulations. These findings are somewhat reflected in the time taken for the repeats which completely froze at 268 K ; where remarkable slowed growth was observed in the presence of all polymers, and in particular the PVA<sub>10</sub> $\times$ 2 samples. Between the polymers these delayed times ranged from 19.14-88.63 ns at just 2 K below the melting temperature ( $T_m-2K$ ), whereas milder changes were observed at other temperatures. Take for instance at 272 K , ( $T_m+2K$ ) where no discernable difference in the time taken for melting was observed between PVA and PEG, nor between chain lengths or concentrations. Overall the number of completed simulations suggest that neither PVA nor PEG exhibit any significant thermal hysteresis, which is in agreement with experiments. [TX4] Another feature which is consistent with literature is that the fluctuations around the average energy drift is greatest at temperatures closest to the  $T_m$  (275-268 K ).[21] This is also reflected in the the standard deviations of times taken between repeats in Fig. 5.7 and these arise due to the heightened competing effects between the addition and removal of water molecules from the ice/water interface at these temperatures.

It is important to acknowledge that during the ice growth simulations, the ice/water interfaces grow in the Z-direction to fill out the regions that sandwich the ice crystal. The overall distance travelled is 11 nm in the this direction, whilst melting of the ice crystal only requires half the distance. The ice growth rates trans-

	Water	PEG <sub>10</sub>	PVA <sub>10</sub>	PVA <sub>19</sub>	PVA <sub>10×2</sub>
Completed <sup>a</sup>	10/10	7/9	6/9	2/9	4/9
Average, ns <sup>b</sup>	185.82	213.95	204.96	206.90	274.45
Stdev, ns <sup>c</sup>	(27.28)	(32.45)	(30.30)	(1.11)	(11.81)

Table 5.1: The number of simulations which completely freeze are recorded <sup>a</sup>. The average time taken <sup>b</sup> for complete freezing at 268, K from completed simulations. Standard deviations <sup>c</sup> between average times taken from completed simulations.

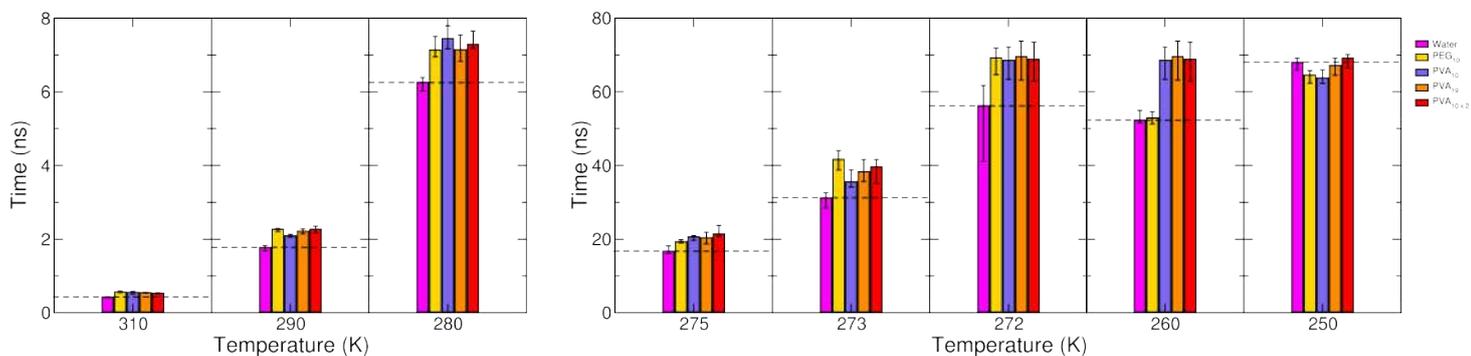


Figure 5.7: The timescales reached in MD simulations for freezing and melting processes in the absence and presence of PVA or PEG oligomers.

lates to distance travelled divided by the time taken. Thus pure water simulations at 268 K , 260 K and 250 K yield growth rates of  $0.06 \pm 0.011 \text{ nm ns}^{-1}$ ,  $0.21 \pm 0.011 \text{ nm ns}^{-1}$  and  $0.16 \pm 0.003 \text{ nm ns}^{-1}$  respectively. These are comparable to the maximum growth rates of  $0.11 \text{ nm ns}^{-1}$  reported by Pruppachar [21] [22] at  $\sim 255 \text{ K}$ . Similarly, Furukawa *et al* recently reported growth rates of the basal face at 268 K as  $0.27 \pm 0.12 \text{ nm ns}^{-1}$  [RES17] and prismatic face as  $0.40 \pm 0.18 \text{ nm ns}^{-1}$ , using TIP6P [RES18]. [RES16] While Kusalik and Rozmanov reported the growth rates at the secondary prismatic planes at 250 K and 260 K at  $0.04 \pm 0.018$  and  $0.54 \pm 0.034$  respectively. [RES16] These too are in line with our own results. Comparisons of our growth rates to empirical formulas listed in 1 were not extensively explored here because they describe tip velocities of free standing crystals, which are largely determined from studies of the basal plane. [25] [21] [22] [29] [28] [30] [27] [31] Whereas we investigate the growth rate of a planar face. In addition to that, caution must be taken in comparing experimental growth or melting rates to experiments because it is common knowledge that the growth rates of simulated systems are much faster than those reported by experiments. [RES16][RES19][RES17][RES18][RES7] This is largely due to the swift removal of latent heat by the thermostats used in our simulations and the use of the periodic boundary conditions. Ordinarily, latent heat would be removed by thermal diffusion from the ice/water interface and unto the bulk phases, and this process plays a determining role on whether the ice crystal will grow or melt. Thermostats work by rescaling the velocities of the molecules of the system in order to maintain an average pre-defined temperature, and so any deviations from this — even if it is induced by the latent heat — is quickly dealt with.

### 5.3.3 Evidence of superheating

At 273 K the PEG<sub>10</sub> oligomer unexpectedly slowed the melting process more so than the IRI active polymers (Fig. 5.7). Based on this and the above results, we now ask the question if superheating is present in the IRI active polymers. Our polymeric studies still obtain a  $T_m$  at 270 K and because the melting point did not increase, we did not identify any indications of superheating of an ice crystal in the presence of PVA (or PEG) oligomers. Superheating could also have been tested by conducting melting simulations of oligomer which have been completely overgrown by ice, rather than the melting of an ice crystal within the vicinity of a polymer.

### 5.3.4 Polymer residence regions

Fig. 5.8-5.10 shows the number density of water oxygens in homogenous systems and in the heterogenous systems at the midpoint of the simulations (30 ns or 150 ns), and towards the end (60 ns or 300 ns) of the simulation. These results are shown for just select repeats. The samples show that between 250 K and 270 K, the pure water systems typically have a QLL of 0.85-0.95 nm at the midpoint of the trajectory, which is slightly lower than the expected 1.0-1.5 nm. Our results show that at the midpoint of the freezing process, the oligomers typically reside within the icy regions at low temperatures (250 K), the QLL regions at 260 K and the liquid regions at temperatures approaching the  $T_m$  (268 K and 270 K). The oligomers normally become overgrown at the same positions they were found in at the midpoint, aside from two exceptions, at temperatures near  $T_m$ . At 268 K systems, PEG travels ahead of the growing ice front, whereas PVA travels from the liquid region in towards the icy regions. Both polymers appear equally mobile because they both travel the same distance ( $\sim 1$  nm) but in opposite directions. At 270 K where more kinetic energy should be available, the PEG oligomers moves  $\sim 5$  nm towards the ice (but still ahead of it), while PVA remains in the same distance from the ice centre position. Larger PVA oligomers or systems with more PVA molecules reside deeper in the ice like regions than their counterparts at 250 K and 260 K. An interesting feature we observe is that although they follow a similar trend with temperature — i.e. they move out from the icy regions and into the more fluid QLL regions — they are capable of extending across slightly into the two phases. As a result of this attachment, they only become freed into the liquid regions at the  $T_m$  however unlike their less IRI potent counterparts, they remain at approximately the same distance from the ice centre regardless of temperature, even up to the  $T_m$ .

### 5.3.5 Ice polymorphism

In Fig. 5.11–[5.13] snapshots of the growing ice crystal is shown in the presence and absence of the polymers at 268 K. Based on the findings from the ice growth simulations in 3 we determined that the value  $Q_4=0.94$  is a better cut-off criteria for identifying ice-like water molecules. We use this to print out only ice-like waters, as well as the polymers if they are present.

As mentioned in the introduction, the polymorphism of ice is currently under debate and of particular interest is the possibility that cubic ice ( $I_c$ ) can spontaneously form from supercooled water, and morph into hexagonal ice ( $I_h$ ). In support of

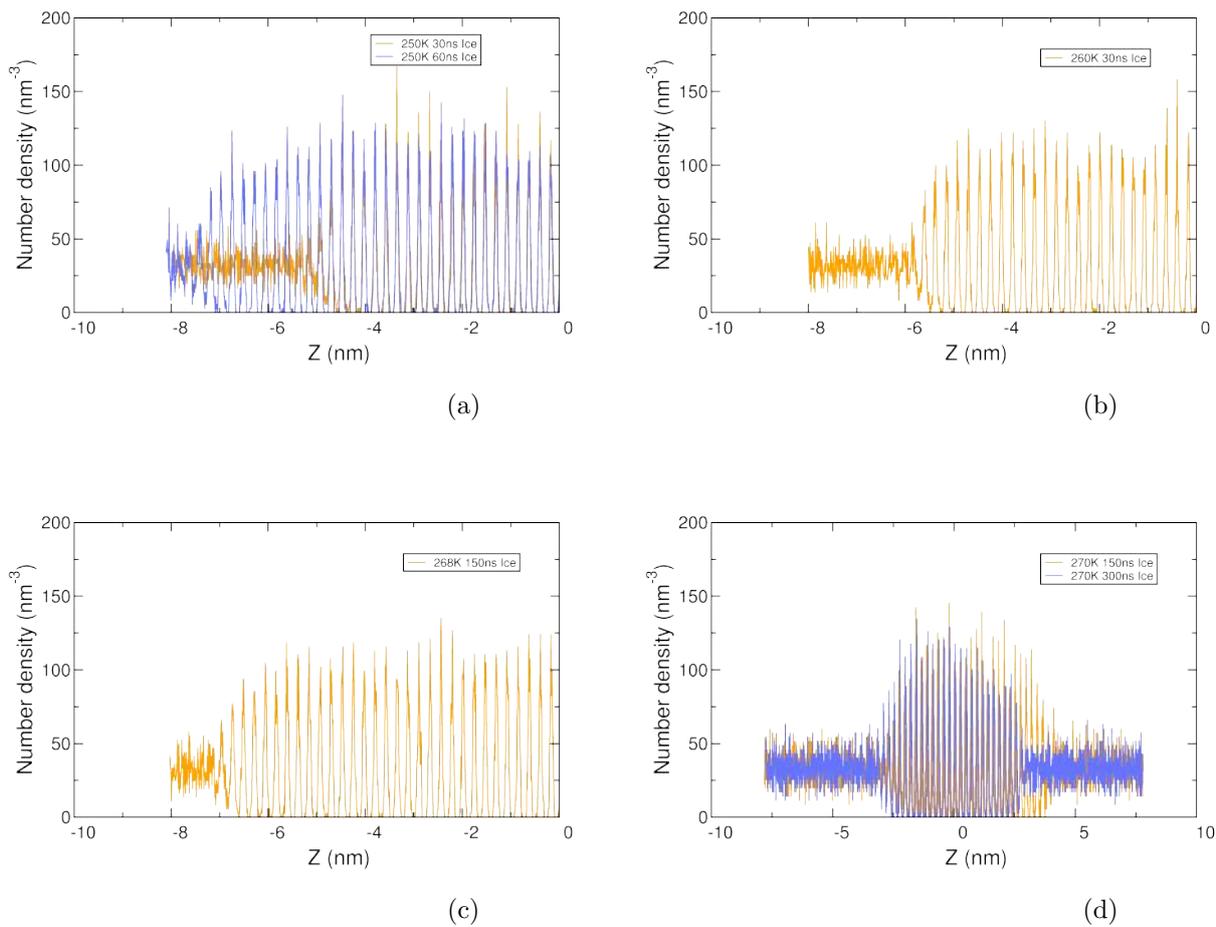


Figure 5.8: The density profile of the pure water systems at the middle (orange) of the trajectory and end of the simulations (blue), used for comparisons with the polymeric systems. At 270 K and 268 K the middle point is 150 ns and 300 ns . At 260 K and 250 K these points are at 30 ns and 60 ns respectively.

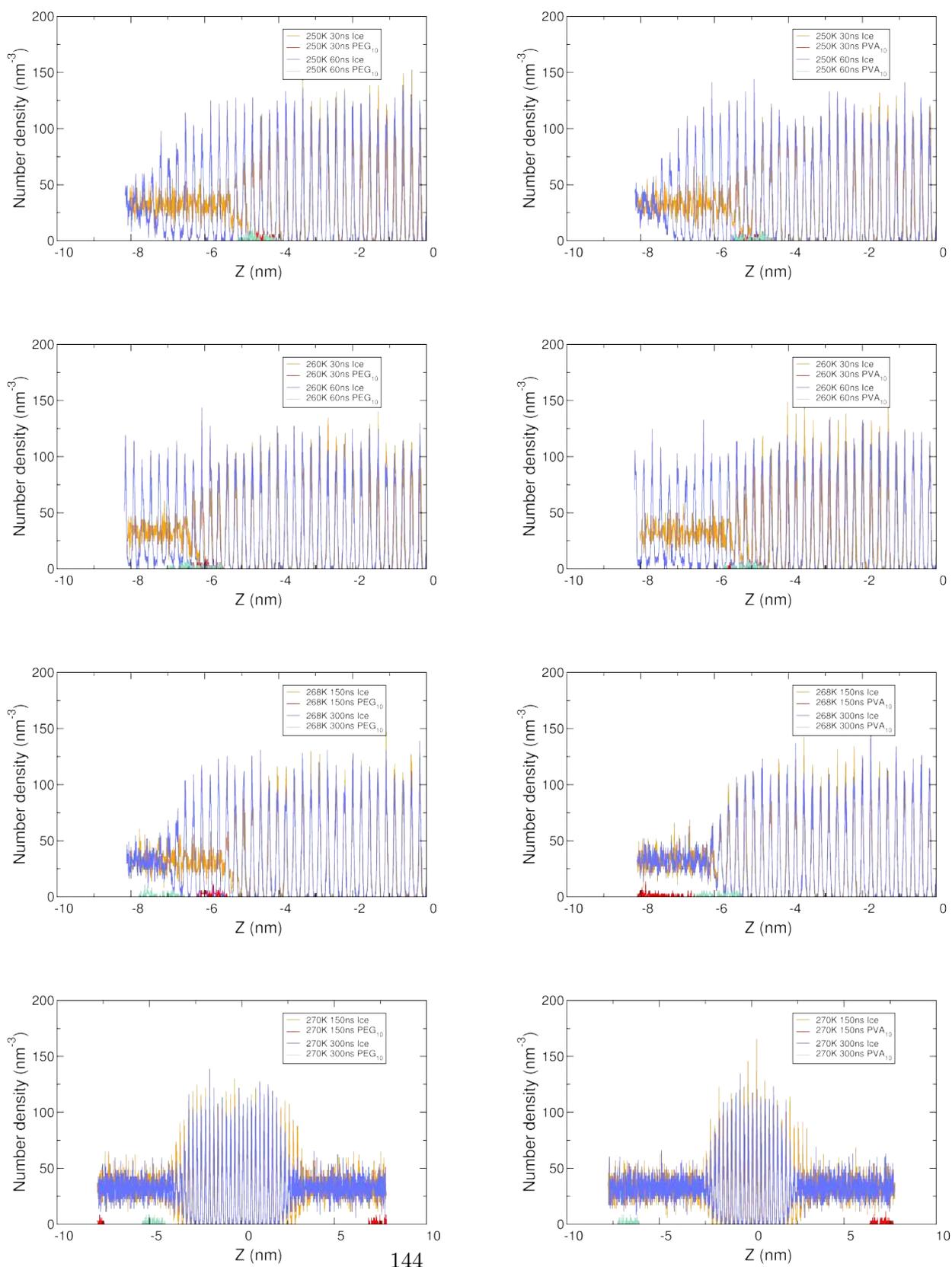


Figure 5.9: The density profile of the PEG<sub>10</sub> (left) and PVA<sub>10</sub> (right) systems at the middle (orange and red) of the trajectory and end of the simulations (blue and green)

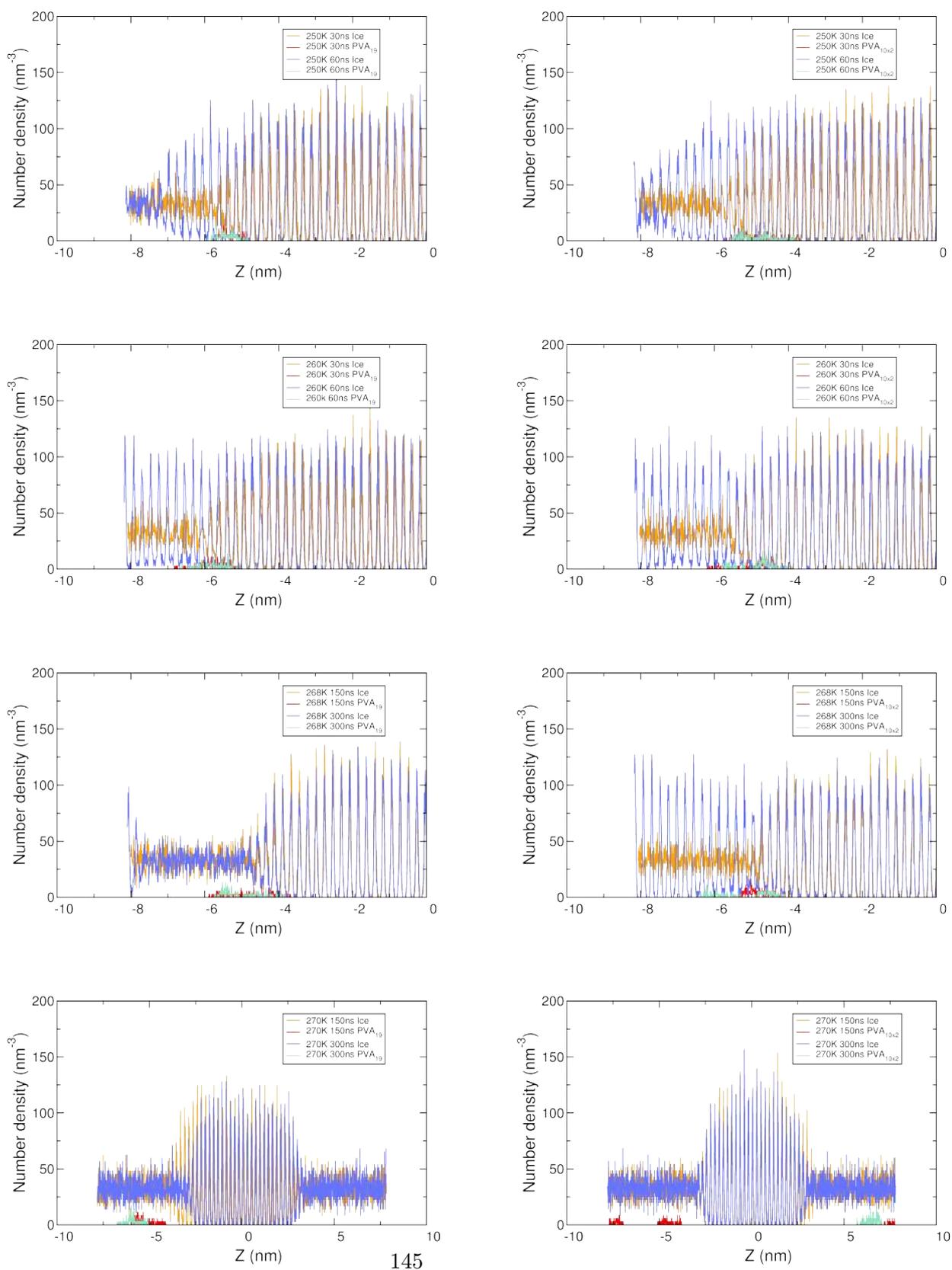


Figure 5.10: The density profile of the PVA<sub>19</sub> systems (left) and PVA<sub>10x2</sub> (right) at the middle (orange and red) of the trajectory and end of the simulations (blue and green)

this Carigano *et al* [SALT2][SALT6] and others [SALT1] used MD simulations and different water models, to show that it is possible to obtain stacking faults during the layer-by-layer growth of the basal plane of cubic ice in the presence of salts, which is not possible at the prism plane [SALT6]. These faults permit marked transitions between the two polymorphs. Similar observations have been made from  $I_h$  to  $I_c$ .? The snapshots show that in our freezing simulations at 268 K for pure water and polymeric systems we did not see any evidence of transitions from  $I_h$  to  $I_c$  in the new ice.

### 5.3.6 Ice growth around polymers

The TH properties of AF(G)Ps have long been attributed to the Gibbs-Thomson effect (or local curvature) but only recently Furukawa *et al* [RES7][RES5][RES8] and Kuiper *et al* [RES12] have both conducted novel simulations which demonstrate its involvement in the mechanism of AFPs. They modelled AFPs adsorbed unto a mobile ice/water interface at 10 K and 5 K below the melting point of two different water models (TIP6P and TIP4P) respectively and found marked convex curvature on either side of the AFPs. Fig. 5.11-5.16 support these findings and reveal that the Gibbs-Thomson effect is faint in all our modelled systems, especially for the PEG<sub>10</sub>. At 268 K and 150 ns, PVA<sub>10</sub> temporarily exhibits short-lived (<75 ns), curved ice growth around the polymer at 150 ns, which is the only snapshot to also coincide with the metastable state of ice for this system. [Similarly, the metastable states for PVA<sub>10×2</sub> at 268 K is explained by the existence of a greater curved ice front which slows the growth rate of ice]. This feature is absent in PVA<sub>19</sub> system with the metastable state larger than PVA<sub>10</sub>, for reasons which are discussed later. Notice that, we also provide snapshots of the pure water system for comparisons at 268 K because they serve as important benchmarks for understanding how much curvature is really introduced by the antifreeze macromolecule. Doing this is important, considering that recent modelling studies that have shown the prismatic planes are especially rough and uneven compared to other planes of ice [SALT2][SALT6], however they are often neglected in these types of studies.

These result is not suprising as non of the polymers are known to display significant TH activity, and is consistent with our interpretations of the results this far. This curvature is proportional to the degree of supercooling and can described by the Gibbs-Thompson equation (Eq. 5.1) [RES8][RES12].

$$R = \frac{A_g V \gamma T_m}{\Delta T \Delta H} \quad (5.1)$$

Assuming a spherical growth of the ice crystal between the periodic images of the oligomers on the ice/water interface,  $R$  represents the radius,  $T_m$  is the melting point (270 K),  $\Delta T$  the freezing depression ( $T - T_m$ ),  $V$  is the molar volume of the solid phase ( $\text{m}^3 \text{mol}^{-1}$ ),  $\Delta H$  is the molar latent heat of fusion ( $5.4 \times 10^3 \text{J mol}^{-1}$ )[RES20][RES21] and  $\gamma$  is the interfacial free energy ( $30.8 \times 10^{-3} \text{J mol}^{-2}$ )[RES20].  $V$  is calculated as the molecular weight of water ( $0.018 \text{kg mol}^{-1}$ ) divided by the density of ice at  $T_m$  ( $906 \text{K g m}^3$ )[RES20][RES21]. In the case of a spherical ice embryo,  $A_g=2$  or  $A_g=1$  for cylindrical geometries. If our polymers were highly TH active, at lower temperatures we would also expect a critical cylindrical radius of 3.06 nm and 1.53 nm at 260 K and 250 K respectively. At 260 K this can be approximated as 7 hexagons on the YZ plane and 14 lattice oxygens in the XZ plane however, we do not observe such large curvatures rising from the baselines in either of these temperatures for any of the systems.

For comparison, Kuiper *et al* found that the TH active sbwAFP achieved curvature with a radius of 4.6 nm, using a similar water model and undercooling.[RES12] In addition to this, the sbwAFP managed to sustain the curvature for much longer timescales (up to 100 ns). These differences confirm that the curvature is indeed mainly a consequence of TH activity, and not IRI.

Snapshots of the oligomers after 30 ns at 260 K also highlight that PVA oligomers are pinned much more deeply into the ice crystal than PEG, without being completely overgrown. The exact reason for this is unclear however there is a correlation with increasing IRI activity, where the curvature appears greater for systems with more hydrophilic groups. This is also reflected in the average slowed growth at 260 K by PVA polymers, whereas PEG<sub>10</sub> did not alter the timescales of freezing. This behaviour is not unexpected because PVA polymers have the added advantage of oxygen separation distances that coincide with the ice lattice arrangement. As results PVA oligomers are also less likely to strain or require significant rearrangement in order to adsorb/interact with ice via hydrogen bonding. At higher undercoolings, we find that the mobility for the oligomers decline. As a result the position and conformations of the oligomers remain largely the same even until 60 ns. This reduced mobility makes it easier for the ice to completely grow over the structures. It also explains why the freezing timescales in all of the the polymeric systems is not much higher or lower than for pure water simulations.

### **Lattice matching and polymer motion**

Metastable states at 268 K for the more IRI potent oligomers (PVA<sub>19</sub> and PVA<sub>10×2</sub>) were not earmarked by any curvature, so that means this stage can also be a

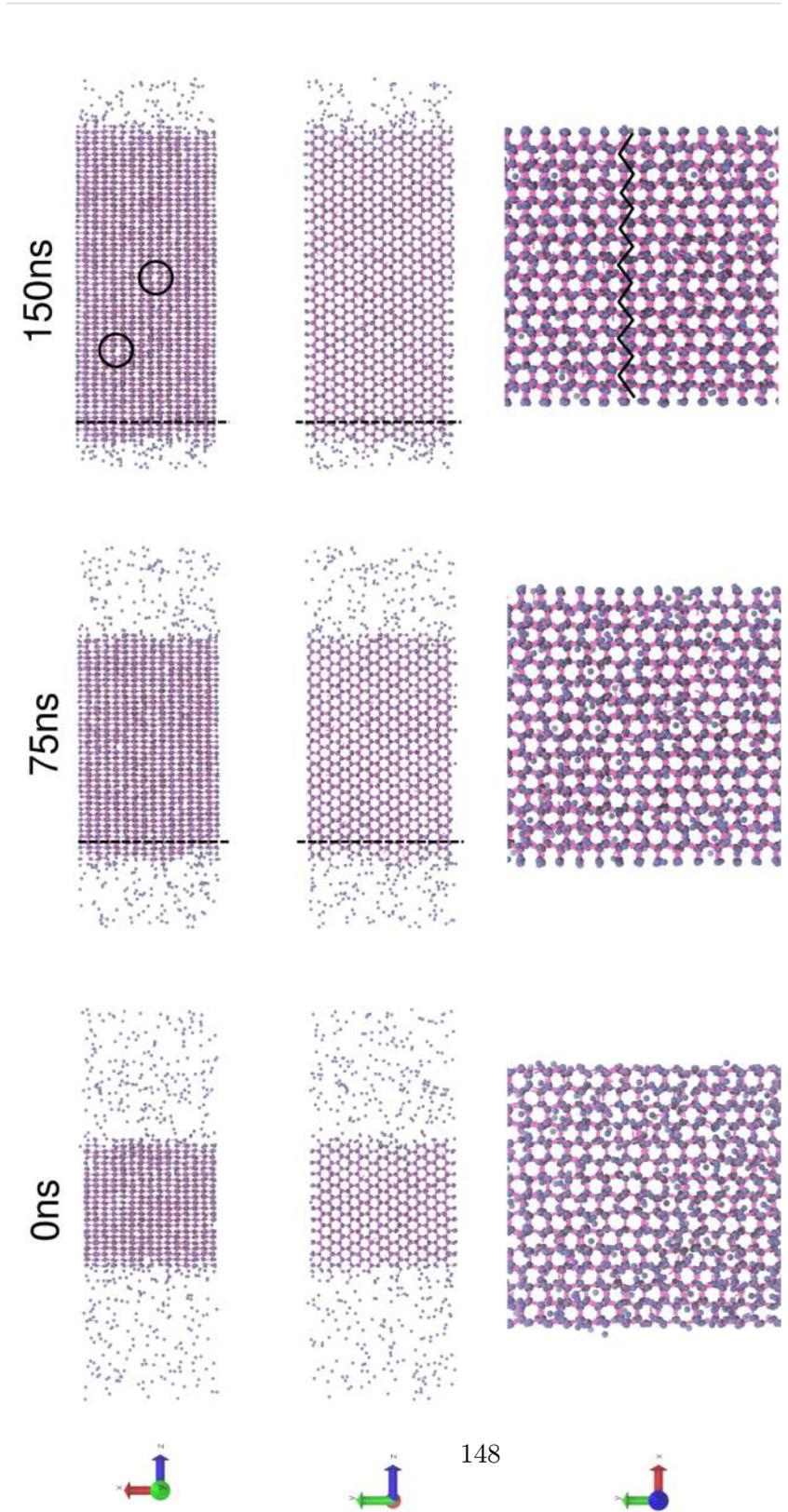


Figure 5.11: Snapshots of the evolution of freezing of pure water systems. XZ, YZ and YX are the top view (prism plane), front view (basal plane) and side view (1120 face) respectively. Oxygen atoms from ice-like water are shown in blue and are connected by hydrogen bonds. It is identified by a zig-zag pattern, whilst passing through the layers from left to right (as shown). Minor defects are circled and dotted lines mark lowest point on cusps using the YZ plane.

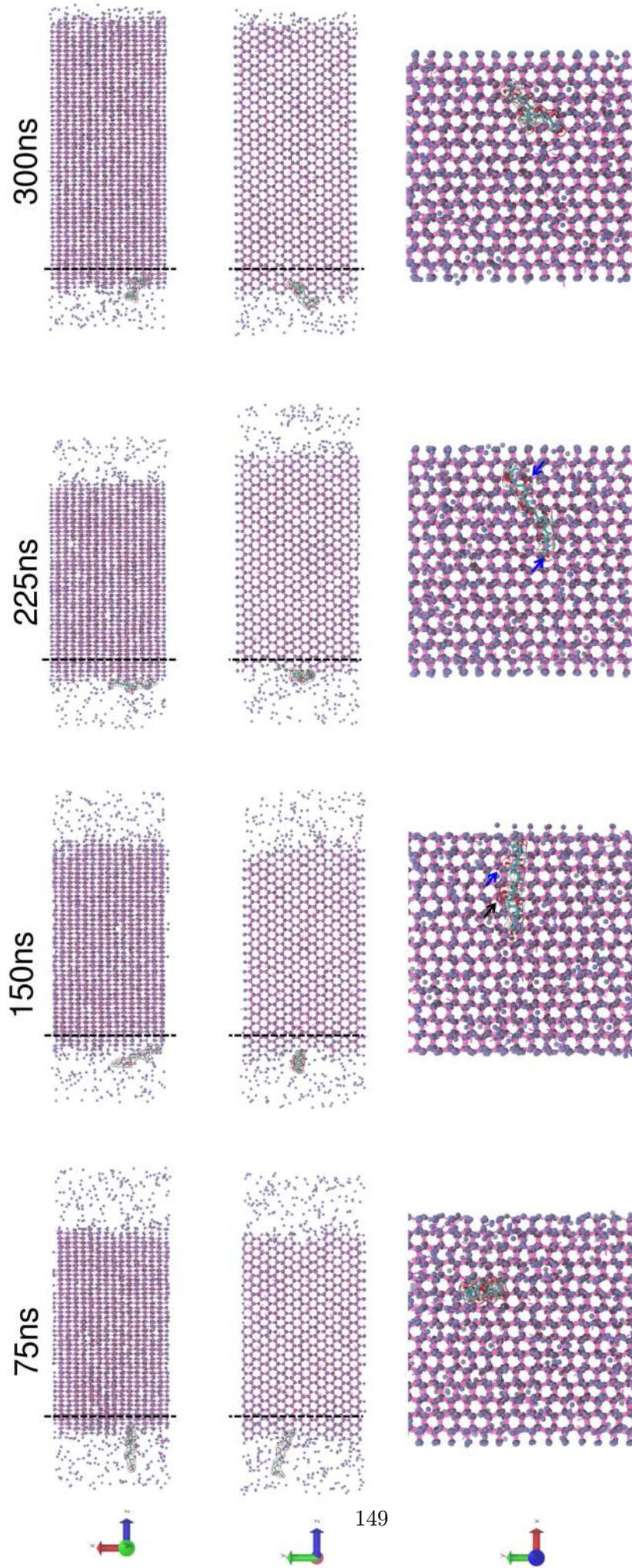


Figure 5.12: Snapshots of the evolution of freezing at 268 K for PVA<sub>10</sub>, including time frames with the metastable states. Blue arrows highlight examples where polymeric oxygen adopt positions normally occupied by oxygen in the ice lattice. Black arrows highlight regions where this adoption was not maximised.

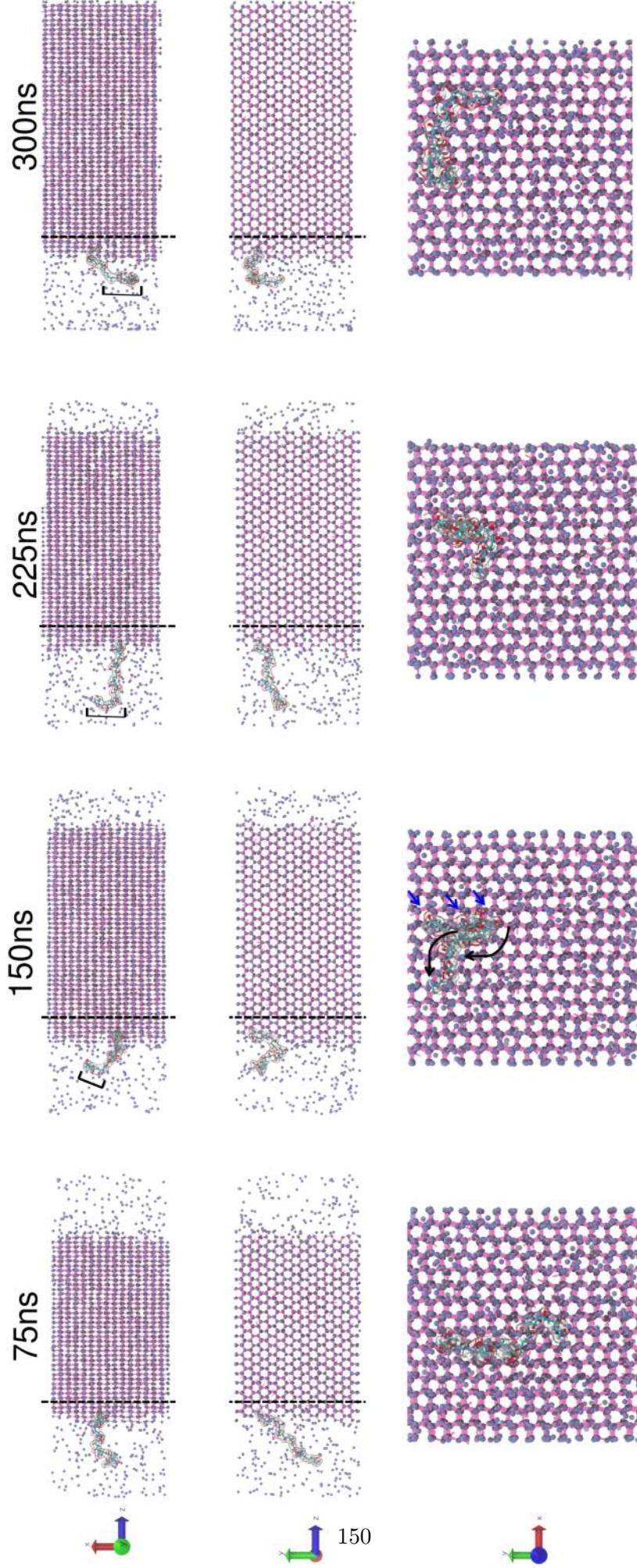


Figure 5.13: Snapshots of the evolution of freezing at 268 K for PVA<sub>10</sub>, including time frames with the metastable states. Same key as 5.12. Black brackets highlight 3-6 unit long regions that flap about.

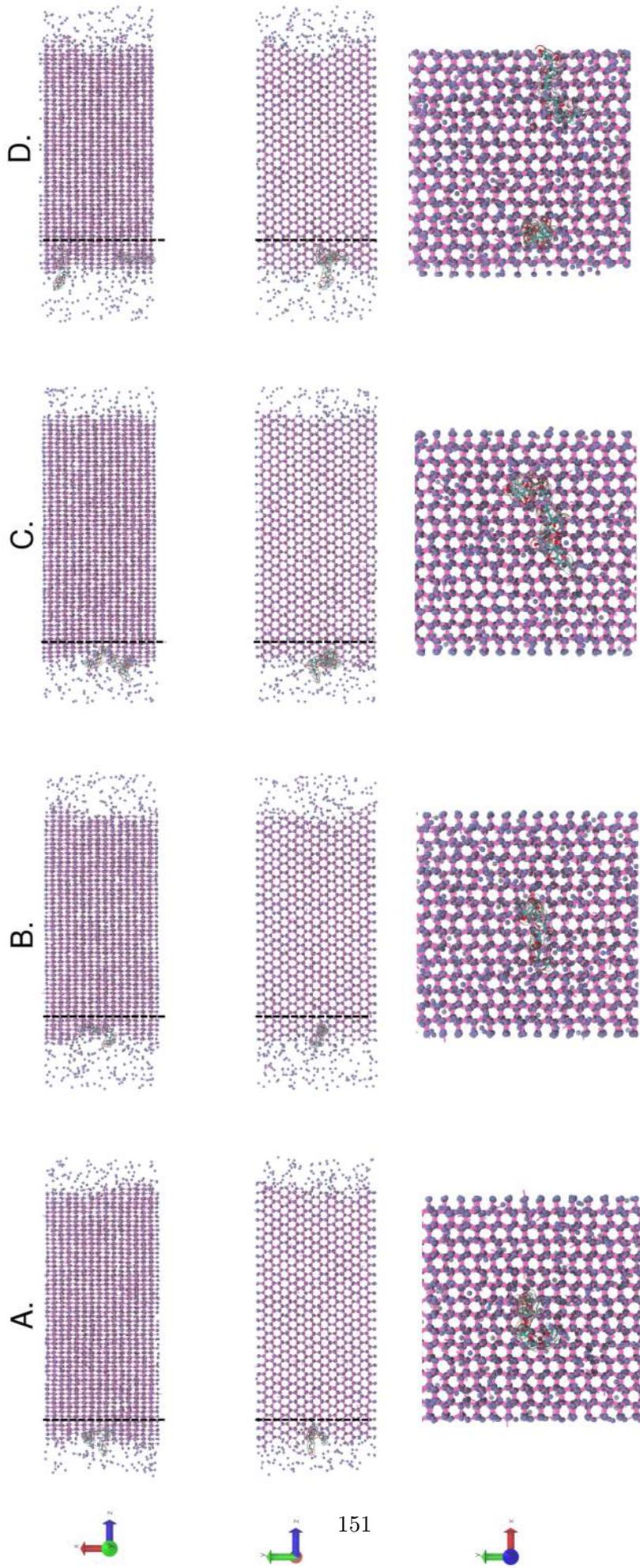


Figure 5.14: Snapshots of freezing at 260 K at 30 ns with A) PEG<sub>10</sub>, B) PVA<sub>10</sub>, C) PVA<sub>10</sub>×2 and D) PVA<sub>10</sub>×2. The black lines highlight the ice front at the position of adsorption (using YZ plane). Greatest curvature seen PVA<sub>10</sub>×2

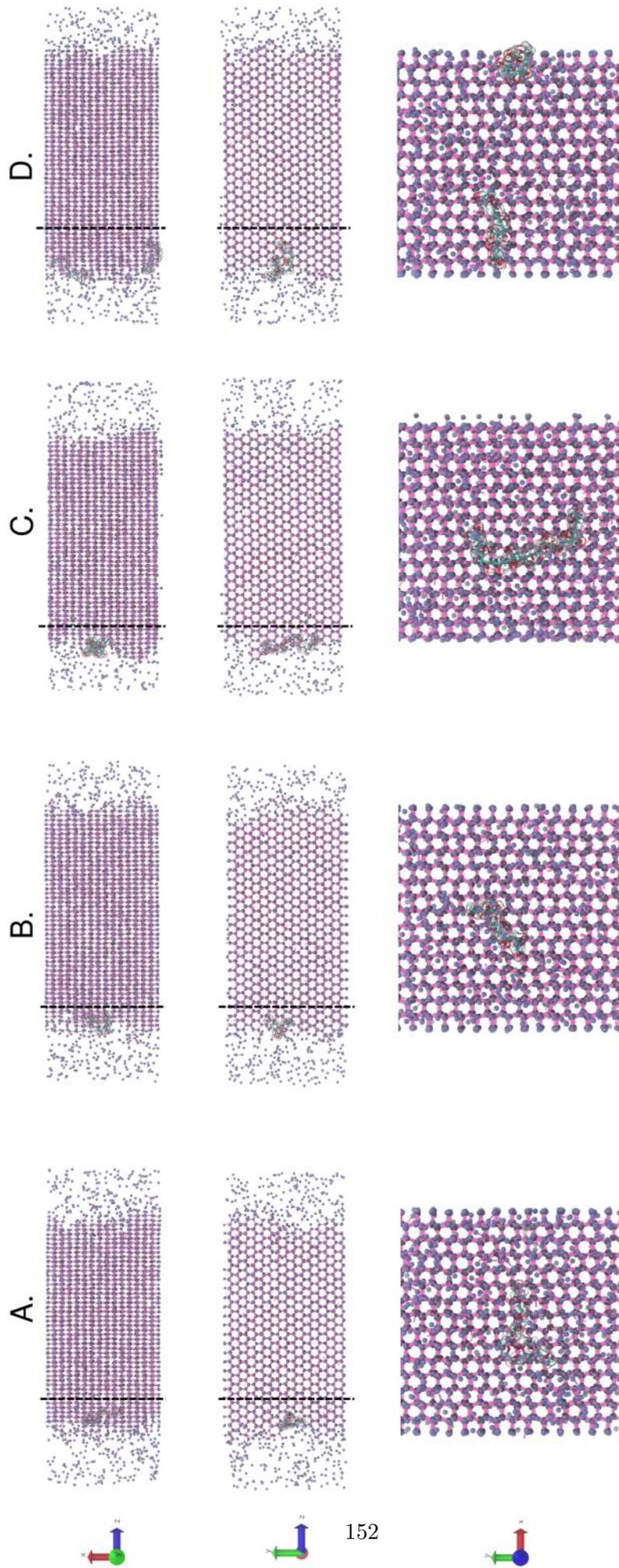


Figure 5.15: Snapshots of freezing at 250 K at 30 ns with A) PEG<sub>10</sub>, B) PVA<sub>10</sub>, C) PVA<sub>19</sub> and D) PVA<sub>10</sub>×2. The black lines highlight the ice front at the position of adsorption (using YZ plane). Greatest curvature seen PVA<sub>10</sub>×2

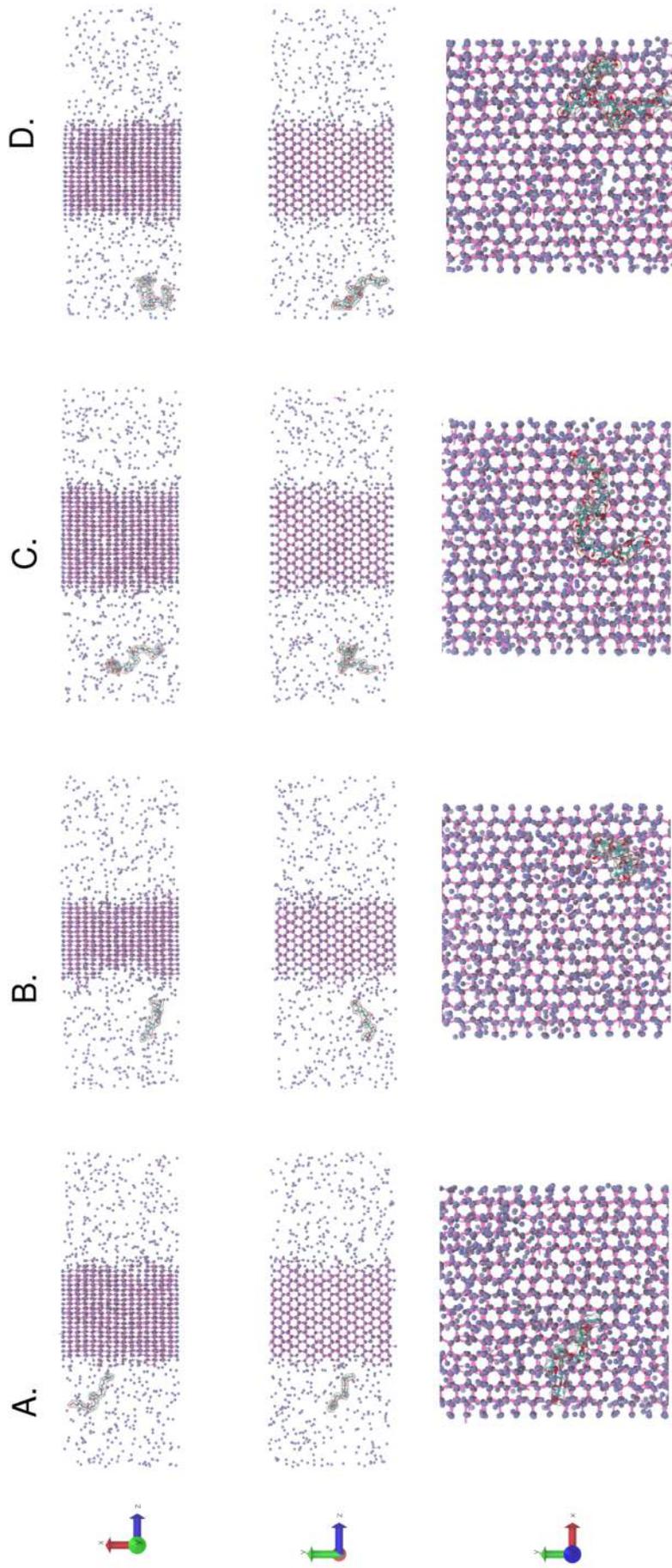


Figure 5.16: Snapshots of freezing at 272K at 30 ns with A) PEG<sub>10</sub>, B) PVA<sub>10</sub>, C) PVA<sub>19</sub> and D) PVA<sub>10</sub>×2. The black lines highlight the ice front at the position of adsorption (using YZ plane). Greatest curvature seen PVA<sub>10</sub>×2

consequence of something else (Fig. 5.13 and [x]). We believe this to be the flexible motion of the polymers. Snapshots of PVA<sub>19</sub> at 150 ns reveals that the oligomer becomes partly incorporated into the growing ice lattice. The incorporated section is linear, approximately 3 units long, and arranges itself to match its polymeric oxygen into free sites on the ice lattice that would normally be occupied by new ice oxygens. By folding upwards unto itself to form a hinge, a  $\sim$ 3-6 unit long tail is formed that is kept away from the growing ice lattice, and it sways from side to side. This mechanical motion, scatters surrounding water molecules and disrupts the build-up of the ice front. Similar kinks have been observed for AFPs [CCC7] [CCC3] [CCC12] and IRI active AFP analogues [ANA1] [RES10] [RES14] [RES15] in simulation studies, however this peculiar flapping mechanism has not been proposed elsewhere. The mechanism may have huge implications for movement of water at the grain boundary, by widening channels or keeping the water flow sustained. It may also explain the phenomena of switching on of IRI activity at PVA<sub>19</sub>. Earlier in the solution studies of polymers at 298 K and 260 K, we found that the hinges or flexibility of the polymers began at 19 units long and increased in correlation with IRI potency. Smaller sized oligomers were not large enough to form such features. As a result this confirms the use of these studies as predictive simulations. As the temperature dropped in those studies, we also found that the polymer contracted and so preferentially interacted with itself. At the same time a reduction in the number of polymer-polymer hydrogen bonds, and a corresponding contraction in the hydrophobic solvent accessible surface area suggests that hydrophobic interactions begin to play a more important role in the polymer's intramolecular interactions. Interestingly, snapshots of polymers in solution revealed that there are no distinct hydrophobic and hydrophilic faces, however there is clear preference for the polymers to approach the ice lattice with a hydroxyl group. As a result these regions can become hooked into the ice lattice during ice growth, if the polymer does not move much in solution. For instance, between 150 ns and 225 ns PVA<sub>19</sub> maintains just 3-units incorporated into the ice at 268 K. We propose that this gives rise to kinks that can be stabilised by intramolecular van der Waals interactions, and the hydrophilic groups on the tail may also simultaneously interact with the solvent to form the flapping motion. At around 225 ns growth of the interface resumes as the kink ceases to fold back unto the bound PVA regions, causing the tail to align with and venture towards the ice. Similar observations have been made on kink stabilisation via hydrophobic pockets of IRI active macromolecules and Tam *et al* also suggests that its exposure also allows the analogues access to conformations that modify its hydration shells, disrupting ice growth [ANA1].

## Ice recognition

The question is how does the polymer come to recognize the ice crystal from pure water? Is there a conformation change in the presence of ice? Our studies show that they already adopt ice like spacing in solution, in the absence of the ice, in both warm and cold solutions. [In the presence of ice the OH3-OH3 pair separations distance at 260 K still largely remain the same (Fig. [px]) compared to pure water studies at the same temperature.] So in sum, they do not recognise. Density profiles of select repeats indicate that the polymers remain in largely the same position over times, which leaves the onus on the ice front to advance towards the polymer in solution. As mentioned earlier, studies using different water models have shown that the prismatic planes are rougher than other planes due to differences in the way that they grow. Basal planes, grow layer-by-layer whilst prismatic planes grow more disorganised.[SALT8] [SALT19] Recent simulations studies of brine solution in contact with Ic ice at 15 K undercooling found that the rough regions on the prismatic planes are prone to trapping ions and as a result metastable state were also achieved until all ions were paired.[SALT6] This behavior is not at all unusual, and similar observations have also been made for simulations of CO<sub>2</sub> adsorption onto ice [RES22].

To characterise this further, we monitored the radius of gyration, the solvent accessible surface areas and the RMSF of the polymers over time at 260 K and 268 K. These temperatures were chosen because the differences observed between the IRI active and inactive polymers were most pronounced at these temperatures. The results are shown in Fig [x-x] and indication of recognition would be identified by a noticeable inflection point or features which coincide with the time taken for the oligomers to freeze over completely or for complete phase change to take place. [We find no such evidence of any recognition using any of these properties.] [Figure [x] plots the percentage hydrophobic SASA over time and we find that this too..]. [Compared to the solution studies/In agreement with the solution studies we find that  $R_g$  increases/decreases at the colder ice simulations. The SASA [increases/decreases] [as does the percentage hydrophobic SASA compared to solution.]]

[Mention the hydrophobic SASA region as a function of time for PVA<sub>19</sub> at 268 K. When the polymer folds, does the phobic SASA decrease or increase? Does that describe the flapping or [is it masked by the change in phase diagram/describes the change in phase diagram]. Can we assume that it reflects the flapping in other studies at 268K? And then what does that say about the flapping in other simulations at 268K?]

## Interactions during binding

### 5.4 Discussion

It is likely that the hinge mechanism is also capable of increasing the fluidity of water molecules by flapping about. Parallels of the hinge mechanism in larger IRI active polymers, can also be drawn to phospholipids found in cold-adaptive organisms. In these organisms the ratio of unsaturated to saturated fatty acids increases relative to their warmer counterparts and as a result of these added double bonds, kinks are formed in the tail. These kinks prevent tight packing of phospholipids in the membrane and as a result the optimal membrane fluidity can still be maintained, thereby permitting normal functions like transport across the cell membrane. We have also shown that the polymers do not need cooperative behaviour in order to function, and can function independently at the ice surface. It has also been demonstrated in some temperatures that the size of the polymer has the ability to affect the size of the convexity at the ice front. For smaller molecules the curvature affected the small delay that we saw, however for bigger polymers we did not see this curvature, rather the flapping of the polymer delayed the freezing time. In addition, we can see that multiple PVA molecules increase the coverage of the polymer and therefore the activity of AF, however they stayed close to each other, acting like a 19-mer, and so force a greater curvature at all instances. In a recent paper by Yu *et al* [SYu] they compared the differences in IRI activity of TH-hyperactive and TH-moderate AFPs. Using substitution studies, they found that both IRI and TH properties relied on the same lattice matching groups, and that the use of certain ions promoted both types of antifreeze activity. Despite these similarities and a 10-fold difference in TH, all the AFPs still had similar IRI activities. Rather correlations of IRI activity was better drawn based on AFP type with the greatest activity being Type 2 > Type 3 > Type 1. It is important to notice that these Types are not just classified in terms of interactions, but also the resulting shapes and size/length. Of the three types of AFPs, type 2 is the largest (11–24 kDa) and consists of a greater mix of linear and rigid regions, as well as flexible and globular regions ( $\alpha$  and  $\beta$  loop structures) which is in support of our hypothesis. However the other two types are either globular  $\beta$  sandwiches or  $\alpha$  helices, that are amenable to fewer structural variations.

**A. INSERT COMMENTS ON HAYWARD'S HINGE PAPER FINDINGS HERE AND PGK PROTEIN CROWDING EFFECTS ON FOLDING**

**B1. INSERT COMMENTS RECENT JULY 2015 PAPER ON AFGP**

## MD SIMULATION ON SCISSORSING FEATURES

**B2. INSERT COMMENTS ON SIMILAR OBSERVATIONS ELSEWHERE FOR DISTINGUISHING IRI ACTIVITY CF TH ACTIVITY [SH H SHAH PAPER (2012)] C. REPORT THE IRI OF PROTEINS FROM PAPER B1 & B2. FOR COMPARISON TO PVA [CONGDON PAPER]**

**D. INSERT COMMENTS ON RECENT 2015 PAPER ABOUT AFGP SIMULATIONS ARE BETTER SIMULATED AT LOW TEMP WHERE BIOLOGICAL RELEVANCE PAPER (IT HAS RMSF OF PROTEINS DISCUSSED IN PAPER B. SO WE CAN COMPARE HOW USEFUL RMSF IS IN TH AND IR IE HOW IT PLAYS A ROLE)**

**E. HOW DOES THE ABD PREDICT THE IRI OF THESE AFGPS AND HOW DOES THAT COMPARE TO OUR HYPOTHESIS QUALITATIVELY.**

**D. ALTERNATIVELY, SIMULATE A AFGP OURSELVES AT RELEVANT TEMP: 260K, 298K AND COMPARE TO PVA AND PEG. THEN DO ONE SET OF 3X260 K FREEZING SIMS**

**F. COMMENT AND CONTRAST THERMORESPONSIVENESS OF PVA AND PEG [CONGDON 2015 PAPER]**

Another paper commented on the reference [35]. They said that: “ The recent model proposed by Kristiansen and Zachariassen [35] can explain both irreversible adsorption and concentration dependence of thermal hysteresis activity. At 9 different temperatures, the interactions between AFPs and ice are different. At the melting point, there is reversible equilibrium exchange of antifreeze protein molecules between solution and the water-ice interface, resulting in concentration dependent of thermal hysteresis; while at temperatures lower than melting point, antifreeze proteins interact and bind to newly formed ice crystal surface; the adsorption becomes irreversible, as described in the adsorption-inhibition model [35].”

[35] E. Kristiansen, K. E. Zachariassen, The mechanism by which fish antifreeze proteins cause thermal hysteresis, *Cryobiology*. 51 (2005) 3: 262-280.

Source: one of the papers on Desktop.

Supporting source: Superheating of ice crystals in antifreeze protein solutions, Yeliz Celika, Laurie A. Grahamb, Yee-Foong Mokb, Maya Barc, Peter L. Daviesb, and Ido Braslavskya,1[AX25]

AT THIS RELEVANT TEMP OF STUDY DO WE SEE ANY OF THE POLYMERS POPPING IN AND OUT OF ADSORPTION NO, Just flapping about

Read: . Can, N. B. Holland. Modified Langmuir isotherm for a two-domain adsorbate: Derivation and application to antifreeze proteins. *Journal of Colloid and Interface Science*. 329 (2009) 1: 24-30.

[In this study, it is obvious that more temperatures that are also equally spaced, would have had to have been sampled closer to the freezing temperature or melting temperature- and completely achieved a process of melting or freezing which is computationally expensive to create a full closure on it. However a similar approach to x was taken where the simulations were partially conducted and inferred to calculate the plateau. Our calculations from these estimates are reliable and have been conducted on our completed processes to yield results which vary x% from our determined values.]

Note: We also acknowledge that in comparison to experimental results, bulky end groups may in fact affect the antifreeze activity of smaller oligomers like PVA<sub>10</sub>, which will explain any discrepancies seen in the TH or IRI being observed in simulations of experiments.

Studying the solution properties of the polymers are also useful to address other aspects such as the importance of hydrophobic groups, the possibility of cooperative interactions between multiple polymers which may be required for significant interactions with ice

The polymer residence location is in agreement/disagreement with the PMF profiles by wierzbicky for the hydrophobic/hydrophilic face, at x temperatures we see that they both reside approximately x nm. Earlier in the polymer simulation studies we suggested that the water may be good at controlling waters close to itself more so than peg which would make its flapping interactions much more disrupting than other polymers of equivalent length. At a considerable length of 19mer the hydrating waters in solution have been shown to still be large, except the polymer would have greater control over the larger quorum of waters.

## 5.5 Conclusion

[REEDIT: X et al made a study of the proteins in the presence of a growing ice crystal, using different faces to approach the ice close to the melting temperature. They found, that using a similar approach to the 268 K and 270 K simulation studies of ours, that the faces of ice were found to melt instead of freeze at would normally at  $T_m+x$  (227K) irrespective of which face was used to approach the ice except for in

the presence of the basal plane where partial meltin]

We have demonstrated that the interactions of polymers with themselves and with the water solvent under different temperatures can help to modulate the affinity for binding to an ice surface.

[It was shown that the Ice/polymer interactions and the thermal hysteresis was sensitive to the functional groups of the polymers, and to some extent as well as their spacing. A more definite study of the effect on sacing could be achieved by building polymers with spaces increased number of carbons separating them, but with the same number of atoms. As well as the effect of having the same number of hydroxyl groups and increased spacing.]

In this section, we have studied the effects of PVA and PEG oligomers in a range of 10 temperatures from 310 K -250 K in the presence of an ice crystal and we have found that in agreement with forerunning hypothesis, the ice growth rates are slowed as a result of their presence [], particularly for IRI active compounds. They do so by achieving larger metastable states as a result of greater curvature or hinge motion. The specific range at which IRI activity was most pronounced were at 260 K

The absence of superheating suggests that the polymers could not irreversibly adsorb to the ice water interface. Although we de not see any polymer popping back out, so this may also be observed under a greater timescale [density at 270K reveal that the polymer stays ahead of the ice front] binding may not be strong enough.

# Bibliography

- J. L. F. Abascal, E. Sanz, R. García Fernández, and C. Vega. A potential model for the study of ices and amorphous water: Tip4p/ice. *The Journal of Chemical Physics*, 122(23):234511, 2005. doi: <http://dx.doi.org/10.1063/1.1931662>. URL <http://scitation.aip.org/content/aip/journal/jcp/122/23/10.1063/1.1931662>.
- J. D. Bernal and R. H. Fowler. A theory of water and ionic solution, with particular reference to hydrogen and hydroxyl ions. *The Journal of Chemical Physics*, 1(8):515–548, 1933. doi: <http://dx.doi.org/10.1063/1.1749327>. URL <http://scitation.aip.org/content/aip/journal/jcp/1/8/10.1063/1.1749327>.
- M. M. Conde and C. Vega. Determining the three-phase coexistence line in methane hydrates using computer simulations. *The Journal of Chemical Physics*, 133(6):064507, 2010. doi: <http://dx.doi.org/10.1063/1.3466751>. URL <http://scitation.aip.org/content/aip/journal/jcp/133/6/10.1063/1.3466751>.
- M. M. Conde, C. Vega, and A. Patrykiewicz. The thickness of a liquid layer on the free surface of ice as obtained from computer simulation. *The Journal of Chemical Physics*, 129(1):014702, 2008. doi: <http://dx.doi.org/10.1063/1.2940195>. URL <http://scitation.aip.org/content/aip/journal/jcp/129/1/10.1063/1.2940195>.
- João Manuel Marques Cordeiro. Monte Carlo thermodynamic and structural properties of the TIP4P water model: dependence on the computational conditions. *Química Nova*, 21:698–701, 11 1998. ISSN 0100-4042.
- Tom Darden, Darrin York, and Lee Pedersen. Particle mesh ewald: An  $n\log(n)$  method for ewald sums in large systems. *The Journal of Chemical Physics*, 98(12):10089–10092, 1993. doi: <http://dx.doi.org/10.1063/1.464397>. URL <http://scitation.aip.org/content/aip/journal/jcp/98/12/10.1063/1.464397>.
- 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. doi: <http://dx.doi.org/10.1063/1.470117>. URL <http://scitation.aip.org/content/aip/journal/jcp/103/19/10.1063/1.470117>.

Ramón García Fernández, José L. F. Abascal, and Carlos Vega. The melting point of ice ih for common water models calculated from direct co-existence of the solid-liquid interface. *The Journal of Chemical Physics*, 124(14):144506, 2006. doi: <http://dx.doi.org/10.1063/1.2183308>. URL <http://scitation.aip.org/content/aip/journal/jcp/124/14/10.1063/1.2183308>.

Bertrand Guillot. A reappraisal of what we have learnt during three decades of computer simulations on water. *Journal of Molecular Liquids*, 101(1-3):219–260, November 2002. ISSN 01677322. doi: 10.1016/s0167-7322(02)00094-6. URL [http://dx.doi.org/10.1016/s0167-7322\(02\)00094-6](http://dx.doi.org/10.1016/s0167-7322(02)00094-6).

Berk Hess, Carsten Kutzner, David van der Spoel, and Erik Lindahl. Gromacs 4: Algorithms for highly efficient, load-balanced, and scalable molecular simulation. *Journal of Chemical Theory and Computation*, 4(3):435–447, 2008. doi: 10.1021/ct700301q. URL <http://dx.doi.org/10.1021/ct700301q>.

William G. Hoover. Canonical dynamics: Equilibrium phase-space distributions. *Phys. Rev. A*, 31:1695–1697, Mar 1985. doi: 10.1103/PhysRevA.31.1695. URL <http://0-link.aps.org.pugwash.lib.warwick.ac.uk/doi/10.1103/PhysRevA.31.1695>.

Zhongqiao Hu and Jianwen Jiang. Assessment of biomolecular force fields for molecular dynamics simulations in a protein crystal. *Journal of Computational Chemistry*, 31(2):371–380, 2010. ISSN 1096-987X. doi: 10.1002/jcc.21330. URL <http://dx.doi.org/10.1002/jcc.21330>.

Ref 36 37 39 40 41 42 43 44 45 46 in The thickness of a liquid layer on the free surface of ice as obtained from computer simulations by MM Conde and C Vega.

William L. Jorgensen and Julian. Tirado-Rives. The opls [optimized potentials for liquid simulations] potential functions for proteins, energy minimizations for crystals of cyclic peptides and crambin. *Journal of the American Chemical Society*, 110(6):1657–1666, 1988. doi: 10.1021/ja00214a001. URL <http://dx.doi.org/10.1021/ja00214a001>.

William L. Jorgensen, Jayaraman Chandrasekhar, Jeffrey D. Madura, Roger W. Impey, and Michael L. Klein. Comparison of simple potential functions for simulating liquid water. *The Journal of Chemical Physics*, 79(2):926–935, 1983. doi: <http://dx.doi.org/10.1063/1.445869>. URL <http://scitation.aip.org/content/aip/journal/jcp/79/2/10.1063/1.445869>.

- William L. Jorgensen, David S. Maxwell, and Julian Tirado-Rives. Development and testing of the opls all-atom force field on conformational energetics and properties of organic liquids. *Journal of the American Chemical Society*, 118(45):11225–11236, 1996. doi: 10.1021/ja9621760. URL <http://dx.doi.org/10.1021/ja9621760>.
- N.A. Lange and J.A. Dean. *Lange's Handbook of Chemistry*. McGraw-Hill, 1973. ISBN 9780070161900. URL <https://books.google.co.uk/books?id=bXYRAQAIAAJ>.
- Michael Levitt, Miriam Hirshberg, Ruth Sharon, Keith E. Laidig, and Valerie Daggett. Calibration and testing of a water model for simulation of the molecular dynamics of proteins and nucleic acids in solution. *The Journal of Physical Chemistry B*, 101(25):5051–5061, 1997. doi: 10.1021/jp964020s. URL <http://dx.doi.org/10.1021/jp964020s>.
- Shüichi Nosé. A molecular dynamics method for simulations in the canonical ensemble. *Molecular Physics*, 52(2):255–268, 1984. doi: 10.1080/00268978400101201. URL <http://dx.doi.org/10.1080/00268978400101201>.
- M. Parrinello and A. Rahman. Polymorphic transitions in single crystals: A new molecular dynamics method. *Journal of Applied Physics*, 52(12):7182–7190, 1981. doi: <http://dx.doi.org/10.1063/1.328693>. URL <http://scitation.aip.org/content/aip/journal/jap/52/12/10.1063/1.328693>.
- Shirley W. I. Siu, Kristyna Pluhackova, and Rainer A. Böckmann. Optimization of the opls-aa force field for long hydrocarbons. *Journal of Chemical Theory and Computation*, 8(4):1459–1470, 2012. doi: 10.1021/ct200908r. URL <http://dx.doi.org/10.1021/ct200908r>.
- A.K. Soper. The radial distribution functions of water and ice from 220 to 673 k and at pressures up to 400 mpa. *Chemical Physics*, 258(23):121 – 137, 2000. ISSN 0301-0104. doi: [http://dx.doi.org/10.1016/S0301-0104\(00\)00179-8](http://dx.doi.org/10.1016/S0301-0104(00)00179-8). URL <http://www.sciencedirect.com/science/article/pii/S0301010400001798>.
- David van der Spoel and Erik Lindahl. Brute-force molecular dynamics simulations of villin headpiece: Comparison with nmr parameters. *The Journal of Physical Chemistry B*, 107(40):11178–11187, 2003. doi: 10.1021/jp034108n. URL <http://dx.doi.org/10.1021/jp034108n>.
- C. Vega, J. L. F. Abascal, M. M. Conde, and J. L. Aragones. What ice can teach us about water interactions: a critical comparison of the performance of different

water models. *Faraday Discuss.*, 141:251–276, 2009. doi: 10.1039/B805531A.  
URL <http://dx.doi.org/10.1039/B805531A>.

Volker C. Weiss, Markus Rullich, Christof Kehler, and Thomas Frauenheim. Kinetic aspects of the thermostatted growth of ice from supercooled water in simulations. *The Journal of Chemical Physics*, 135(3):034701, 2011. doi: <http://dx.doi.org/10.1063/1.3609768>. URL <http://scitation.aip.org/content/aip/journal/jcp/135/3/10.1063/1.3609768>.