University of Strathclyde

Heterogeneous Nucleation at Hydrophobic Interfaces

Author:
David McKechnie

Supervisor:
Dr. Karen Johnston

Declaration

This thesis is the result of the authors original research. It has been composed by the author and has not been previously submitted for examination which has led to the award of a degree.

Signed:

November 3, 2021

Abstract

Crystal nucleation is a ubiquitous process that plays an important role in a range of environmental, biological and industrial processes. It is well accepted that nucleation most commonly occurs heterogeneously at interfaces, and a number of mechanisms have previously been explored that contribute to this effect. Nucleation experiments are often conducted in small scale experimental setups, where the interfaces present are significantly different to those present in normal, macroscale crystallisations. Despite the prevalence of heterogeneous nucleation, the effects of these very particular interfaces are often neglected when analysing the data from such experiments. This thesis will demonstrate the impacts that these interfaces can have on nucleation using a model system of aqueous glycine solution, and will demonstrate a novel concentration effect that facilitates heterogeneous nucleation and is distinct from previously investigated heterogeneous nucleation mechanisms.

Results of experiments will be reported for glycine solutions with and without contact with a tridecane oil interface. These results demonstrate that the presence of the oil significantly increases the nucleation rate of glycine. This is a surprising result as the non-polar hydrophobic tridecane interface would not be expected to enhance the nucleation of the highly polar, hydrophilic glycine. Classical molecular dynamics simulations reveal significantly enhanced vs depleted glycine concentrations at the oil–solution vs air–solution

interfaces, respectively. It is proposed that this interfacial concentration effect facilitates heterogeneous nucleation, and that it is due to dispersion interactions between the interface and the solution molecules.

To confirm this, model interfaces with tuneable interface–solution interactions were implemented to the molecular dynamics simulations. The solution composition at the interface was found to be strongly dependent on the strength of the dispersion interactions between the interface and the solution. In contrast, while the electrostatic interactions between the interface and the solution were also found to influence the interfacial solution composition, the observed effects are significantly weaker than those observed for the dispersion interactions.

These effects have been observed for glycine solutions at a tridecane interface, however it is expected that the same mechanism will be present in a wide range of solution–interface systems. Deeper understanding of these efffects will allow for control over the interfacial cooncentration in order to design effective nucleants for the enhancement of nucleation, or to suppress nucleation for anti-fouling purposes.

Acknowledgements

I would like to thank a number of people who have not only made this work possible but also helped to make it such an enjoyable experience. Firstly, I would like to thank my three supervisors, Karen Johnston, Jan Sefcik and Paul Mulheran. I am lucky to have had such an incredible team of supervisors, who not only were able to provide me with their support, expertise and guidance throughout my project but also have given me so many opportunities to develop as a researcher. I would particularly like to thank Karen for initially encouraging me to pursue a PhD which has allowed me to discover my love for research.

I would like to thank my partner Cecilia for all of her love and support throughout the years. Thank you for all of your encouragement and the countless hours you have endured listening to me talk (and occasionally complain) about crystals.

Completing my PhD at the University of Strathclyde has been made particularly special thanks to all of the fantastic people that I have met here. I have made so many friends throughout these years that is difficult to list them all, but I would like to thank Georgia, Martin, Scott, Paul, Rab, Carla, Russell, João and Sofia for all of the hours we've spent together between the office, the pub, dinners and trips away. All of these events have helped to make the time during my PhD so memorable.

I would like to thank my friends Mark, Marc, Shaun and Cowie for the times we spent virtually during lockdown, and for providing a greatly appreciated escape from writing my thesis.

Finally, I would like to thank my family for all of their support throughout my studies. I would like to especially thank my mum and Alex who managed to survive not one, but two, PhD students who were writing their theses in their home while locked down during a global pandemic.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Crystallisation is a ubiquitous process that plays a major role within environmental, biological and industrial processes. Crystallisation can be found in the formation of minerals [1], ice in the atmosphere [2], and cataracts and kidney stones within the body [3,4]. Crystallisation is utilised in a wide range of industries, including pharmaceuticals, fine chemicals and food, however, it plays a particularly important role within the pharmaceutical industry, where its ability to obtain material at a high purity and stability makes it particularly attractive. This is highlighted by the fact that more than 80% of marketed pharmaceutical products are formulated with the active pharmaceutical ingredient (API) in the solid crystalline form [5].

The crystallisation process can be separated into two key steps: nucleation, the formation of crystalline particles from the liquid phase, and growth. Nucleation influences a number of key quality attributes of the final crystalline product, but despite the importance of this step, our understanding of nucleation remains limited. While nucleation has long been described by classical nucleation theory, the nucleation pathway remains unclear

with growing evidence of pre-nucleation clusters, an intermediate state prior to nucleation, in a range of systems as part of a two-step process. Due to this lack of understanding, we cannot accurately control, or meaningfully predict, when nucleation will occur, where the nucleation will take place, or which crystal form will be obtained [6].

It is well understood that nucleation often occurs heterogeneously at interfaces, and that heterogeneous nucleation is significantly faster than homogeneous nucleation due to a number of contributing mechanisms. True homogeneous nucleation is rare, outside of carefully controlled conditions, and so a more complete understanding of heterogeneous nucleation will allow for greater control over crystallisation processes in the future.

This work aims to investigate how the presence of various interfaces influence the nucleation kinetics and polymorphic outcome of a model system of aqueous glycine solution. A combined experimental and computational approach will be used to provide atomic level insight at the interface to rationalise behaviours observed within the lab and understand how interfacial properties influence nucleation.

1.1 Thesis Layout

The next chapter will provide background information on crystallisation, the challenges related to nucleation and the model compound investigated in this work: glycine. The following chapter will give an overview of the theory related to the methods used throughout this work. This will be followed by a discussion of the results of nucleation experiments that will demonstrate the unexpected, but major increase in the nucleation rate of glycine solution in contact with a hydrophobic tridecane interface. This will be supplemented with similar results for isonicotinamde-ethanol solutions in contact with a fluorinated oil, highlighting that this effect is not unique to aqueous glycine solution. The results of molec-

ular dynamics simulations of oil–solution and air–solution interfaces will then be presented that show a novel concentration enhancement effect that contributes to the enhanced nucleation rates observed experimentally. It is proposed that this concentration enhancement effect is due to non-specific dispersion interactions between the interface and the solution molecules, and is distinct from previously explored mechanisms for heterogeneous nucleation. Finally, molecular dynamics simulations of glycine solutions in contact with model interfaces will be presented that investigate the effects of dispersion and electrostatic interactions on the solution composition at the interface. This will demonstrate that the solution composition is strongly influenced by dispersion interactions, with smaller effects observed for electrostatic interactions.

Chapter 2

Background

This chapter will begin with background information on crystallisation. This will be followed by a description of the challenges related to nucleation in the design of crystallisation processes. Finally, details of the model compound investigated in this work, glycine, will be provided.

2.1 Crystallisation

Crystallisation occurs when a solution is supersaturated i.e. when the amount of solute within the solution exceeds its solubility under the current conditions. Supersaturation is defined as:

$$S = \frac{C}{C^*} \tag{2.1}$$

where C is the concentration of the solution and C^* is the solubility at the current conditions. A supersaturated solution, if given enough time, will result in some of the solute

precipitating as a solid to allow the concentration to return to the solubility, and an equilibrium between the solution and solid phase is achieved. Crystallisation can be broken down into two key stages: nucleation, the initial formation of crystalline particles from the solution or melt, followed by growth of the crystalline particles.

2.1.1 Solubility Phase Diagrams

As supersaturation is the driving force for crystallisation, a good understanding of the phase diagram for a given solution system is vital to effectively design a crystallisation process. An example phase diagram is shown in Figure 2.1. Within the phase diagram there are two thermodynamically well-defined lines, the binodal (black solid line) and spinodal (black dotted line) lines.

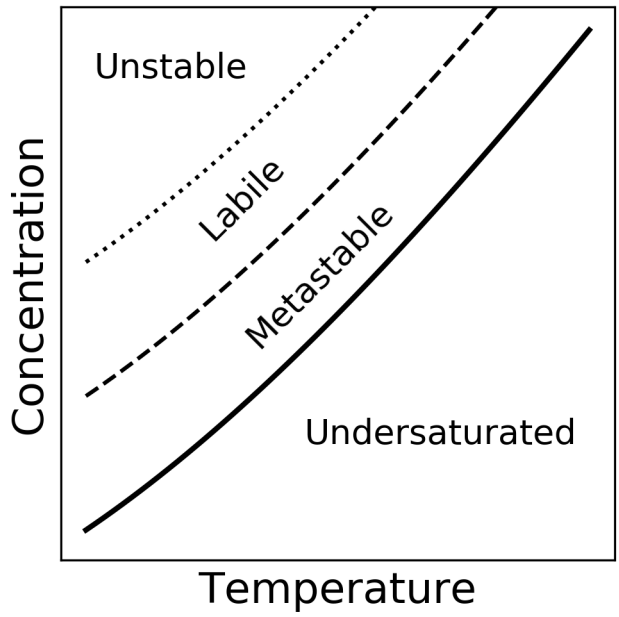

Figure 2.1: Example phase diagram demonstrating the undersaturated and supersaturated regions. The solid line represents the binodal line, the dashed line represents the metastable limit and the dotted line represents the spinodal line.

The binodal line, or solubility curve, defines the region where there is no driving force

for phase separation and the solution is stable. Below this line, a solution is undersaturated and no crystallisation will occur. Between the binodal and spinodal lines is a supersaturated, metastable region where, given enough time, nucleation and growth will occur to bring the system back to thermodynamic equilibrium. Finally, once the spinodal line is crossed, the system is unstable and phase separation occurs spontaneously through spinodal decomposition. During spinodal decomposition there is no barrier for phase separation and the process becomes diffusion limited, however it is unlikely that a solution of small molecules will reach the unstable region before nucleation and growth has occurred.

While supersaturation is the only thermodynamic requirement for nucleation to occur, there is also a kinetic barrier that must be overcome. This results in the metastable region of the phase diagram often being split into two further regions defined as 'labile' and 'metastable'. The metastable region is bound between the binodal line and the metastable limit (dashed black line). Within the metastable region nucleation is unlikely to occur as the supersaturation is not high enough to easily overcome the kinetic barrier. As crystals already present within the metastable zone will grow, it is possible to seed the system at this point to circumvent primary nucleation. It is important to note that the metastable limit is a kinetic parameter, rather than a thermodynamic one, and as such the metastable limit can be affected by conditions such as agitation, presence of impurities and the crystallisation vessel used. This means that while the metastable limit is a useful parameter in the design of crystallisation processes, it is not directly transferable between experimental setups and conditions.

The labile region is the final region of the phase diagram, and sits between the metastable limit and the spinodal line. Within this region spontaneous nucleation is likely to occur.

2.1.2 Generating Supersaturation

In order to reach the labile region to allow crystallisation to occur, it is necessary to generate supersaturation. The three most commonly used methods are cooling, evaporative and antisolvent crystallisation. Figure 2.2 provides schematics of typical routes through phase diagrams for each of these crystallisation types.

Cooling crystallisation can be used when the solubility of a given material is strongly influenced by temperature. Figure 2.2 (a) demonstrates a cooling crystallisation process. The process begins with an undersaturated solution at a high temperature (A). The solution is cooled moving the system to a supersaturated state (B). Given enough time, crystallisation will occur reducing the concentration and returning the system to equilibrium (C).

In some instances the solubility of a given material will vary little with temperature, such as with sodium chloride. In these instances cooling is not an effective method for generating supersaturation, and other methods are necessary such as evaporative crystallisation. Figure 2.2 (b) demonstrates an evaporative crystallisation process. The solution is initially saturated (A), and the solvent is allowed to evaporate. As the solvent evaporates the concentration increases generating supersaturation (B). Crystallisation then occurs to bring the solution back to equilibrium (C).

Antisolvent crystallisation involves the addition of another liquid into the system that is known as an antisolvent. An antisolvent is a liquid in which the solute has poor solubility and is also miscible with the initial solvent. While the addition of the antisolvent reduces the overall concentration of the solution, it also reduces the solubility resulting in an overall increase in the supersaturation. A typical antisolvent crystallisation process is shown in Figure 2.2 (c). It is important to note that this example is for a fixed temperature, and the x-axis of the phase diagram is now the antisolvent mass fraction of the solvent-antisolvent

mixture. The process begins as an undersaturated solution in pure solvent (A). As antisolvent is added to the solution the overall concentration decreases, but, as the solubility also decreases at a faster rate than the concentration, the solution becomes supersaturated (B). Crystallisation then occurs to bring the solution back to equilibrium (C). Similarly to evaporative crystallisation, antisolvent crystallisation can be used when the solubility does not vary strongly with temperature, however, antisolvent is frequently used in combination with cooling crystallisation to maximise the yield of a crystallisation process.

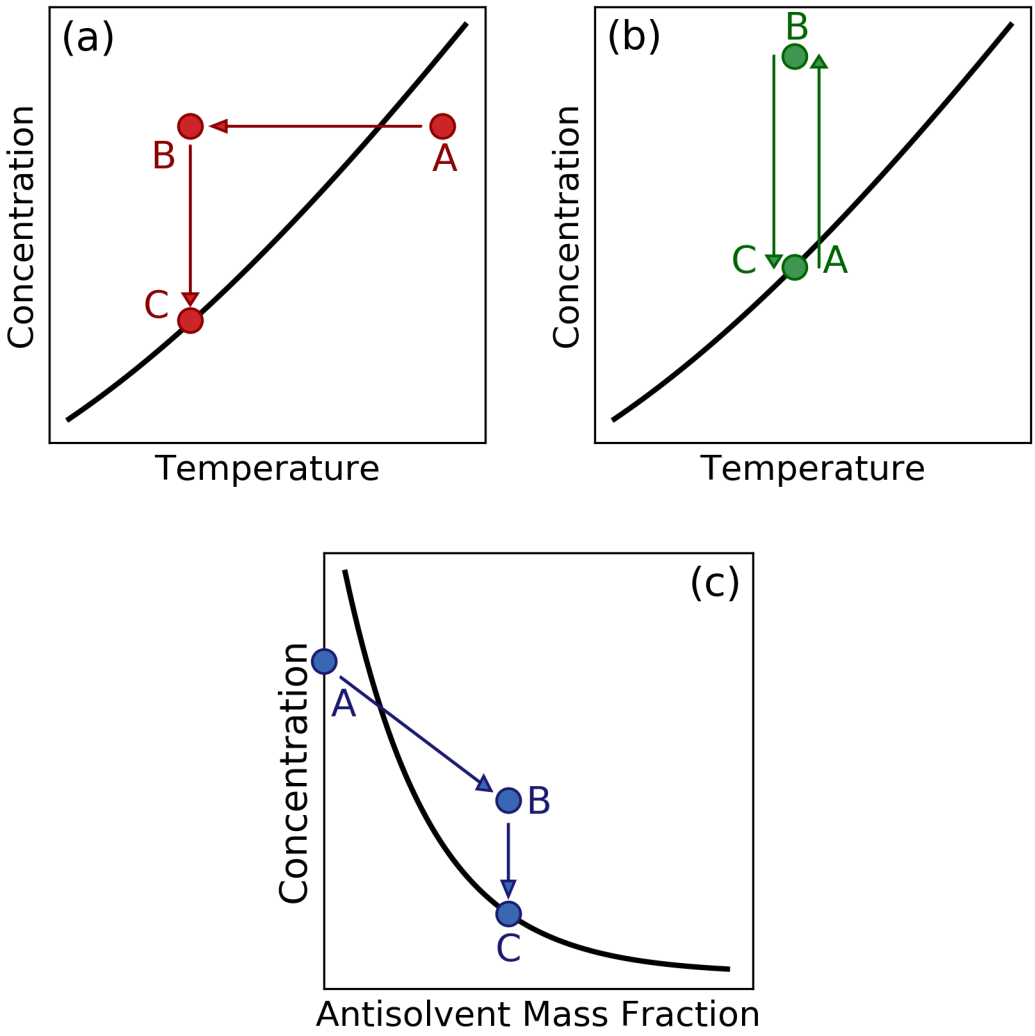

Figure 2.2: Schematics to demonstrate various methods of generating supersaturation: (a) cooling crystallisation, (b) evaporative crystallisation, and (c) antisolvent crystallisation. In each case the black line represents the binodal line.

2.1.3 Nucleation

Nucleation is the formation of crystalline particles from the supersaturated solution. This step influences a number of critical properties of the final crystalline product, such as the polymorph or particle size distribution, which influence both the efficacy and downstream processability of the resulting material. Despite the importance of this step, our understanding of nucleation is still limited.

Nucleation can be broadly broken down into two categories: primary and secondary nucleation. Primary nucleation is the first formation of a crystalline particle with no other crystalline particles of the material present. Secondary nucleation is the formation of new crystals in the presence of parent crystals of the material. Primary nucleation can be further split into homogeneous and heterogeneous nucleation. Homogeneous nucleation is the formation of a crystalline particle directly from the liquid phase, while heterogeneous nucleation occurs on a heterogeneous interface. Possible interfaces for heterogeneous nucleation include impurities within the system, such as dust; the walls of the crystallisation vessel; or probes inserted to monitor the crystallisation process. Heterogeneous nucleation occurs at a much greater rate than homogeneous nucleation, and it is accepted that the majority of nucleation events that are observed occur heterogeneously.

Classical Nucleation Theory

We know that a supersaturated solution is thermodynamically less stable than an equilibrium between a solid and saturated solution, however, it is possible to maintain a supersaturated solution for a significant amount of time without nucleation occurring. This means that while supersaturation is the only thermodynamic requirement for nucleation, there must also be a kinetic component to the nucleation process.

Classical nucleation theory (CNT) is the most widely used theory to describe nucleation from solution. It states the nucleation is caused in supersaturated solutions by stochastic density fluctuations. These fluctuations are a result of random collisions of the dissolved material. The formation of nuclei is dependent on the balance of the bulk and surface energy of the new phase, as the free energy change of the formation of a nucleus is the sum of the free energy change of the surface ($\Delta G_{interface}$) and the volume (ΔG_{bulk}):

$$\Delta G_{total} = \Delta G_{bulk} + \Delta G_{interface} \tag{2.2}$$

Molecules within the centre of the nucleus are fully surrounded by other molecules within the crystal lattice, in energetically favourable configurations, resulting in a favourable free energy change. If we assume a spherical nucleus this bulk term will scale with the volume of the sphere:

$$\Delta G_{bulk} \propto \frac{4}{3}\pi r^3 \tag{2.3}$$

While the free energy term associated with the bulk of the nucleus is favourable, there is also a cost associated with the creation of the interface between the two phases. This free energy cost will scale with the surface area of the spherical particle:

$$\Delta G_{interface} \propto 4\pi r^2 \tag{2.4}$$

As the favourable bulk term scales with the cube of the sphere's radius, r^3, and the unfavourable interfacial term scales with the square of the radius, r^2, the size of the nucleus will determine which term is dominant. The combined free energy change, ΔG_{total}, will pass through a peak at what is known as the critical radius, r_c, as shown in Figure 2.3. Below the critical radius the formed nucleus will be unstable and will redissolve, whilst nuclei with radii greater than the critical radius will stabilise and grow.

The presence of heterogeneous interfaces can reduce the nucleation barrier resulting in the increased rates observed for heterogeneous nucleation. If nucleation occurs on an

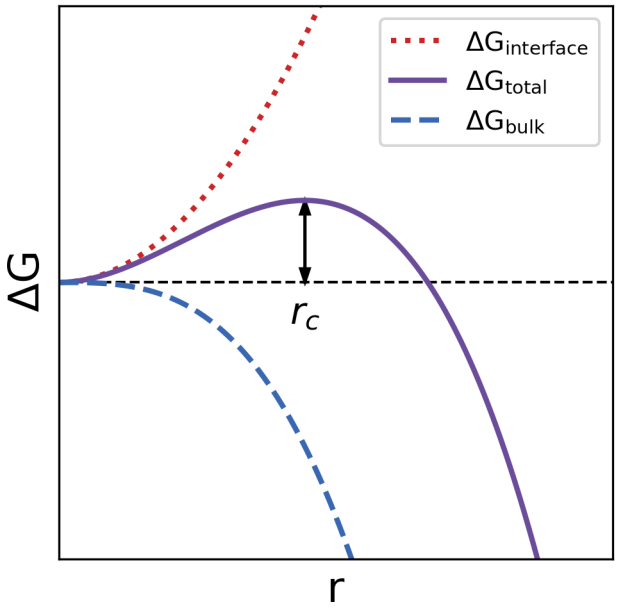

Figure 2.3: Free energy change related to the creation of a crystal nucleus from a supersaturated solution.

interface, the surface area between the nucleus and the surrounding solution is reduced, as is demonstrated in Figure 2.4. If the interaction between the heterogeneous interface and the crystal nucleus is more favourable than the interaction between the nucleus and the solution, this will result in an overall reduction of the free energy cost of the interfacial term in equation 2.2. This results in a reduced nucleation barrier and allows heterogeneous nucleation to occur more quickly than homogeneous nucleation.

Two Step Nucleation

In contrast to CNT, there is growing evidence for the existence of pre-nucleation clusters (PNCs) in solution [7]. These PNCs are dense, disordered, liquid-like regions, on the scale of hundreds of nanometres, that appear to act as an intermediate step in the nucleation process. Two-step nucleation theory incorporates these clusters in order to describe a new

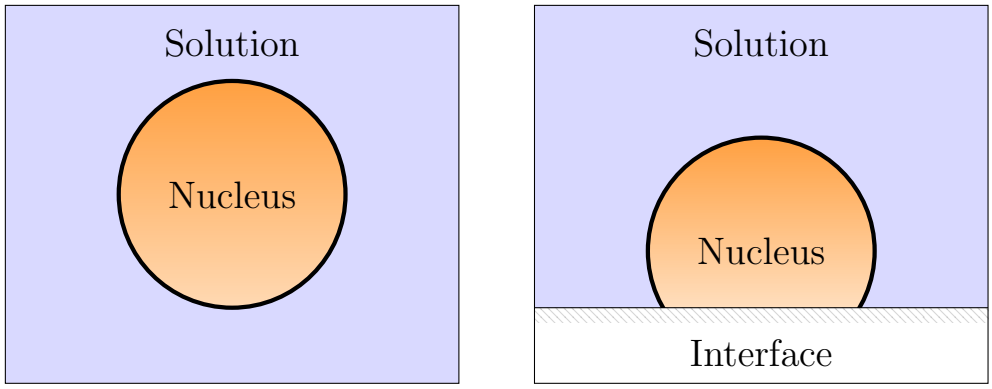

Figure 2.4: Schematic to demonstrate the reduced surface area between the nucleus and the surrounding solution for heterogeneous nucleation (right) in comparison to homogeneous nucleation (left).

pathway for nucleation: the initial aggregation of solute molecules to form PNCs, followed by a reordering of the molecules within these clusters to form the crystal nucleus [8]. Two-step nucleation describes a process with two energy barriers that must be overcome for nucleation: the initial formation of the PNC and the reordering of the PNC to form a crystal nucleus.

Evidence of pre-nucleation clusters has been obtained for proteins [9], biominerals [10] and small organic molecules [11, 12]. Fluorescence has been used to visualise the two-step nucleation of a dibenzoylmethane boron complex during evaporative crystallisation [13]. This growing evidence suggests that the nucleation pathway may be more complex than what is described by classical nucleation theory.

2.1.4 The Enhanced Rate of Heterogeneous Nucleation

As described previously, in section 2.1.3, the presence of interfaces within a crystallisation process can enhance the nucleation rate due to a reduction in the kinetic barrier for nucleation. However, there are a number of other mechanisms that have been explored that can also contribute to the enhanced rates observed for heterogeneous nucleation.

Physical templating, or epitaxy, can occur when the physical shape of the interface has an affinity for the crystal structure that is being formed, inducing spatial ordering of the solute molecules at the interface [14, 15]. In addition to physical templating, chemical templating can also occur. During chemical templating, the surface chemistry results in interactions between the functional groups of the interface and the functional groups of the solute molecules leading to binding and/or orientation of the solute at the interface [16]. The orientational and binding effects of these methods can result in an enhanced nucleation rate.

Finally, there have been numerous studies investigating the effect of confinement on nucleation in a wide range of systems and it was found that confinement can increase or decrease nucleation rates through a number of underlying mechanisms [2, 17, 18].

2.1.5 Challenges Related to Nucleation

Despite the importance of nucleation, our fundamental understanding of this process still remains limited. This prevents the meaningful prediction or control of where and when nucleation will occur. Uncontrolled nucleation can result in a number of problems for a crystallisation process. As drug molecules are becoming larger, more flexible and more complex there are a growing number of systems where nucleation is difficult. In these

instances materials can form an amorphous solid which lacks the stability that would be achieved by a crystalline form. In other cases systems will undergo a liquid-liquid phase separation (often referred to as 'oiling out') which can then subsequently nucleate, resulting in an overall slower crystallisation process that often produces lower quality particles [19].

In contrast to this, unwanted nucleation on the surfaces within the crystallisation vessel can occur which is known as fouling. Material that has fouled within the crystalliser may not necessarily be recoverable under tightly regulated conditions, such as within the pharmaceutical industry, and in the case of expensive API molecules can result in a significant loss of profit. In addition to this, fouling can impact the effectiveness of a crystallisation process. If fouling occurs on process analytical technology (PAT) probes this can influence the obtained measurements, resulting in a poorly controlled process. Significant build up within the crystalliser can also result in reduced performance as flow-rates and heat transfer coefficients can be impacted. In severe cases fouling can lead to complete failure of a crystallisation due to blockages within the system.

If the appropriate amount of nucleation is achieved it is also necessary to ensure that the correct crystal form is obtained. A crystal is a solid material that is made up of a perfectly repeating 3D pattern of atoms or molecules. Some materials have the ability to form multiple different crystal structures that have different 3D packing arrangements of the molecules, as demonstrated in Figure 2.5. This is known as polymorphism and has significant implications in the development of crystallisation processes. An analysis of 245 polymorph screens found that 90% of investigated compounds exist in multiple crystalline and non-crystalline forms, with 50% exhibiting polymorphism [20]. As the inter and intramolecular interactions will vary between different arrangements of molecules, different polymorphs of the same material will display differences in a number of key properties [21].

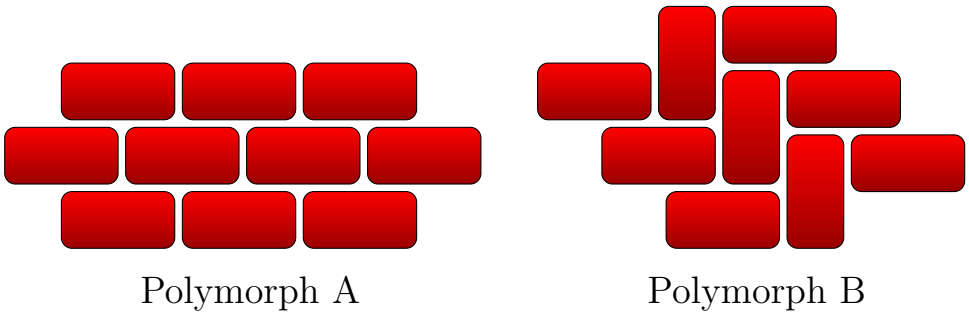

Figure 2.5: Schematic to demonstrate polymorphism. Each red rectangle represents one molecule within the crystal lattice. Recreated from [22].

The solubility will vary between polymorphs with the most stable form having the lowest solubility. The most stable form of drug molecules are often selected for development into pharmaceutical products as this ensures that no solid-solid transformations can occur after formulation. However, the water solubility of an API is directly related to the bioavailability of the drug and the most stable, but least soluble, polymorph will have a lower bioavailability than the other forms. As there is an increasing number of drug molecules currently in development that suffer from poor water solubility [23], this can result in the need for more creative crystal engineering and formulation solutions to counteract this problem [24].

A considerable effort is undertaken within the pharmaceutical industry to ensure that the possible polymorphism is well understood when developing a new drug. Extensive experimental polymorph screening [25] is performed alongside computational crystal structure predictions [26, 27] to ensure that as many crystal structures can be identified as possible. If a more stable crystal structure exists, and is not identified during screening, it can have disastrous results. The most famous example of such an incident is the case of Ritonavir, an anti-HIV drug, that suddenly started to nucleate in a previously undiscovered form with a significantly lower solubility rendering its initial formulation ineffective. This issue was compounded by the problem of 'disappearing polymorphs', the phenomena where

once a new, more stable, form of a compound has nucleated once, it becomes increasingly difficult to obtain the old, metastable form [28, 29]. Ultimately, a new formulation of the drug was required resulting in supply shortages risking patient well-being, and significant costs to the company [30].

The polymorph can also impact a number of physical properties of the final product such as the compressibility, flowability and packing density. A key example of this is paracetamol, where the stable form I has poor compressibility, due to its herringbone crystal structure, that results in low quality tablets. The metastable forms II and III have layered structures that result in much greater compressibility, however these forms are much more challenging to obtain, preventing their use for higher quality tablets [31].

From this range of properties that can vary for polymorphs, it is clear the the polymorph obtained can have a significant impact on both the efficacy of the final product and the downstream processing of the obtained crystalline particles. Therefore, it is vital to ensure that a crystallisation process is able to consistently obtain the desired polymorph.

2.1.6 Small Scale Experimental Setups

Nucleation is a stochastic process, and as such requires a significant volume of data to be collected before meaningful insight into the nucleation behaviour of a given system can be gained. This, alongside pressures within industry for crystallisation processes to be developed using less starting material and time, has resulted in an increase in the popularity of small scale experimental setups.

These small-scale setups have a number of attributes that make them attractive for the study of nucleation. A large number of identical experiments can be carried out

simultaneously to generate the necessary amount of data to gain a statistically significant insight into the nucleation behaviour of the system studied. In addition to this, the small scale of these experiments can allow for more careful control over key parameters, such as temperature or rate of evaporation, that can have a significant impact on the behaviour observed. As nucleation is a stochastic process, we cannot meaningfully predict where and when nucleation will occur. By performing nucleation experiments in these small-scale setups, the region that must be observed is reduced, allowing nucleation to be detected more quickly and efficiently. The small volumes of solution involved also reduces the amounts of solid material and solvent necessary to carry out the experiments.

While a useful tool, it is important to note that these small-scale setups differ from typical macroscale crystallisations in a number of ways. While the small volumes involved within these experiments allow for more careful control of temperature, macroscale crystallisations are more susceptible to thermal inhomogeneity which can result in temperature gradients within the crystalliser. While most macroscale crystallisations will be agitated, agitation is often not present or is poorly defined within small-scale experiments.

Interfacial effects become increasingly prominent with decreasing solution volume, as surface-to-volume ratio increases inversely with the container size. Small scale experimental setups often have very particular interfaces present at a very high surface area to volume ratio that would not be present in macroscale crystallisations. Examples include the oils present in microfluidic [32,33], millifluidic [34], and microwell [35] experiments; or polymers in tubings and stirrer coatings [36].

While these interfaces are often assumed to be inert, there is growing evidence that they can have a significant impact on the observed nucleation behaviour. Ildefonso *et al.* [37] demonstrated that discrepancies in the nucleation rates of lysozyme reported in the literature can be related to the interfacial energies of the oils used within the microfluidic setups,

while Selzer *et al.* [38] found that the choice of oil within their microfluidic experiments had a significant impact on the nucleation behaviour of potassium nitrate. It has been observed that the nucleation rate of isonicotinamide in ethanol is orders of magnitude larger when measured in a microfluidic setup with fluorinated oil in comparison to glass vials [39]. The presence of PTFE has been shown to increase the nucleation rate of glycine [36], and to reduce the metastable zone width of sodium chlorate [40].

These differences can have a significant impact on the nucleation behaviour that is observed. It is vital that these differences can be identified and accounted for to allow for data to be effectively transferred from a small-scale to a macroscale crystallisation during the scale up of crystallisation processes.

2.2 Glycine

This work has primarily focused on the nucleation of a model system of aqueous glycine solution. Glycine is the simplest amino acid and is frequently used in nucleation studies where it is a popular choice due to a number of its interesting properties. In neutral solutions glycine exists primarily as a zwitterions, the structure of which is shown in Figure 2.6.

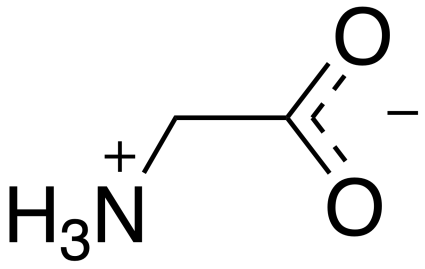

Figure 2.6: The molecular structure of zwitterionic glycine.

Glycine is polymorphic and exists as three crystal structures under ambient conditions:

α [41], β [42], and γ [43], with a further three phases obtained at high pressures: δ [44], ϵ [45] and ζ [46]. The three crystal structures obtainable under ambient conditions are shown in Figure 2.7. Of these three polymorphs, γ is the thermodynamically stable form, followed by α and finally the highly metastable β. Typical cooling crystallisations will result in the formation of α glycine, with agitation favouring the primary nucleation of this form [36, 47]. In order to obtain the other polymorphs of glycine, it is necessary to crystallise under specific conditions.

The most stable polymorph, γ, can be obtained by crystallisation from acidic or basic solutions [48]. It has also been obtained by crystallising under quiescent conditions, with higher concentrations favouring the formation of γ glycine [36]. α glycine will also undergo a solution-mediated transformation to γ glycine given enough time [49].

The highly metastable β polymorph is typically obtained using antisolvent crystallisation, as this allows for high supersaturations to be obtained. The β form has also been successfully obtained by crash cooling aqueous solutions [50], or by heterogeneous nucleation on substrates that stabilise the metastable form [51]. Due to its low stability, β will rapidly transform to α in the presence of water or heat [52], which can make the isolation of this form particularly challenging. This also leads to the possibility that β glycine is frequently nucleating within crystallisation experiments, but transforming into the more stable α form before detection and/or characterisation of the crystals.

Glycine has also been observed to form mesoscale clusters in solution, making it a compound of interest when considering CNT vs two-step nucleation [11, 12].

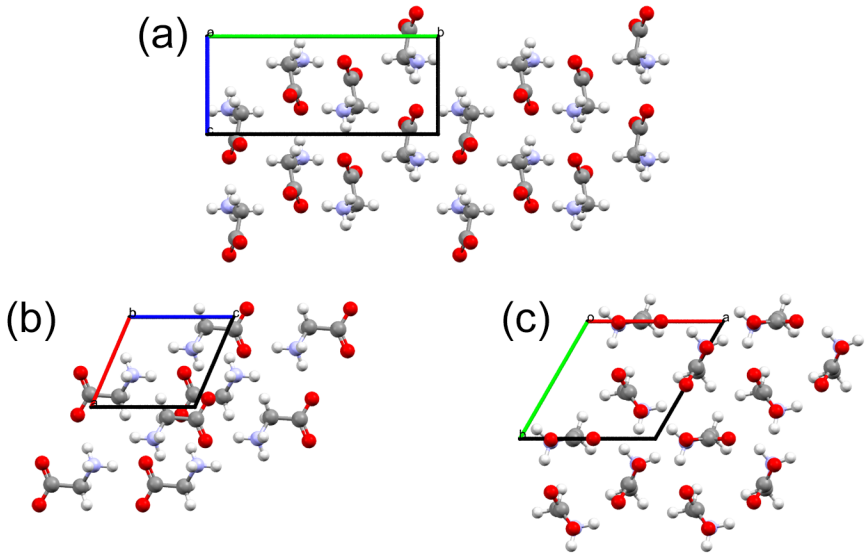

Figure 2.7: The crystal structures of glycine obtainable under ambient conditions: (a) α, (b) β, and (c) γ.

2.2.1 Solubility

As discussed in section 2.1.1, in order to effectively develop a crystallisation process it is vital to have a good understanding of the phase diagram. Despite the well-studied nature of glycine crystallisation, there are significant discrepancies in the literature regarding its solubility. An extensive review of the literature for solubility data revealed that the majority of studies have not determined which polymorph is being investigated, with only five of the 22 studies providing this information. As the solubility of each polymorph will vary (due to differences in their stability) this is an important detail that is missing from these works. The solubility data for no defined polymorph are shown in Figure 2.8 (a).

The remaining five studies that do state the polymorph also show significant variation, as shown in Figure 2.8 (c). This variation is particularly significant at higher temperatures, where solubility measurements are more challenging, however even at a temperature of 298 K the solubilities reported for α-glycine range from 182–248 g/kg (obtaining the values for Park [53] and Sun [54] at 298 K by interpolation using a second order polynomial) and

from 144–213 g/kg for γ-glycine (similarly interpolating for Igarashi [55]). The combined data sets are plotted in Figure 2.8 (b), with undefined polymorph data greyed out, to allow for comparison between the defined and undefined polymorph data. From this plot we see that the undefined polymorph data appears to be α-glycine with good agreement with the solubility data obtained by Park [53] and Sun [54]. This may be somewhat expected as cooling crystallisations from aqueous solution will typically result in the α polymorph.

These data highlight possible pitfalls when reporting experimental methodologies for crystallisation studies. Simply reporting the supersaturation is insufficient without also stating the solubility value used, while reporting concentrations directly will avoid ambiguity.

Table 2.1: Table of sources for solubility data and the temperature ranges and polymorphs investigated.

Author	**Polymorph**	**Temperature Range (K)**	**Reference**
Matsuo *et al.*	-	298	[56]
El-Dossoki et al.	-	298	[57]
Cao *et al.*	-	298	[58]
Dalton and Schmidt	-	273 - 373	[59]
Ferreira *et al.*	-	298 - 333	[60, 61]
Venkatesu *et al.*	-	298	[62]
Bonnin-Paris *et al.*	-	308 - 334	[63]
Cohn *et al.*	-	298	[64]
Ramasami	-	288 - 308	[65]
Dunn *et al.*	-	273 - 348	[66]
Kazimierczuk and Szydlowski	-	290 - 333	[67]
Yi *et al.*	-	296 - 319	[68]
Talukdar *et al.*	-	288 - 308	[69]
Nozaki and Telford	-	298	[70]
Romero and Oviedo	-	293 - 323	[71]
Mason	-	298	[72]
Sun *et al.*	α	293 - 343	[54]
Devi *et al.*	α & γ	274 - 353	[50]
Igarashi *et al.*	α & γ	278 - 343	[55]
Yang *et al.*	α & γ	278 - 334	[73]
Park *et al.*	α & γ	281 - 361	[53]

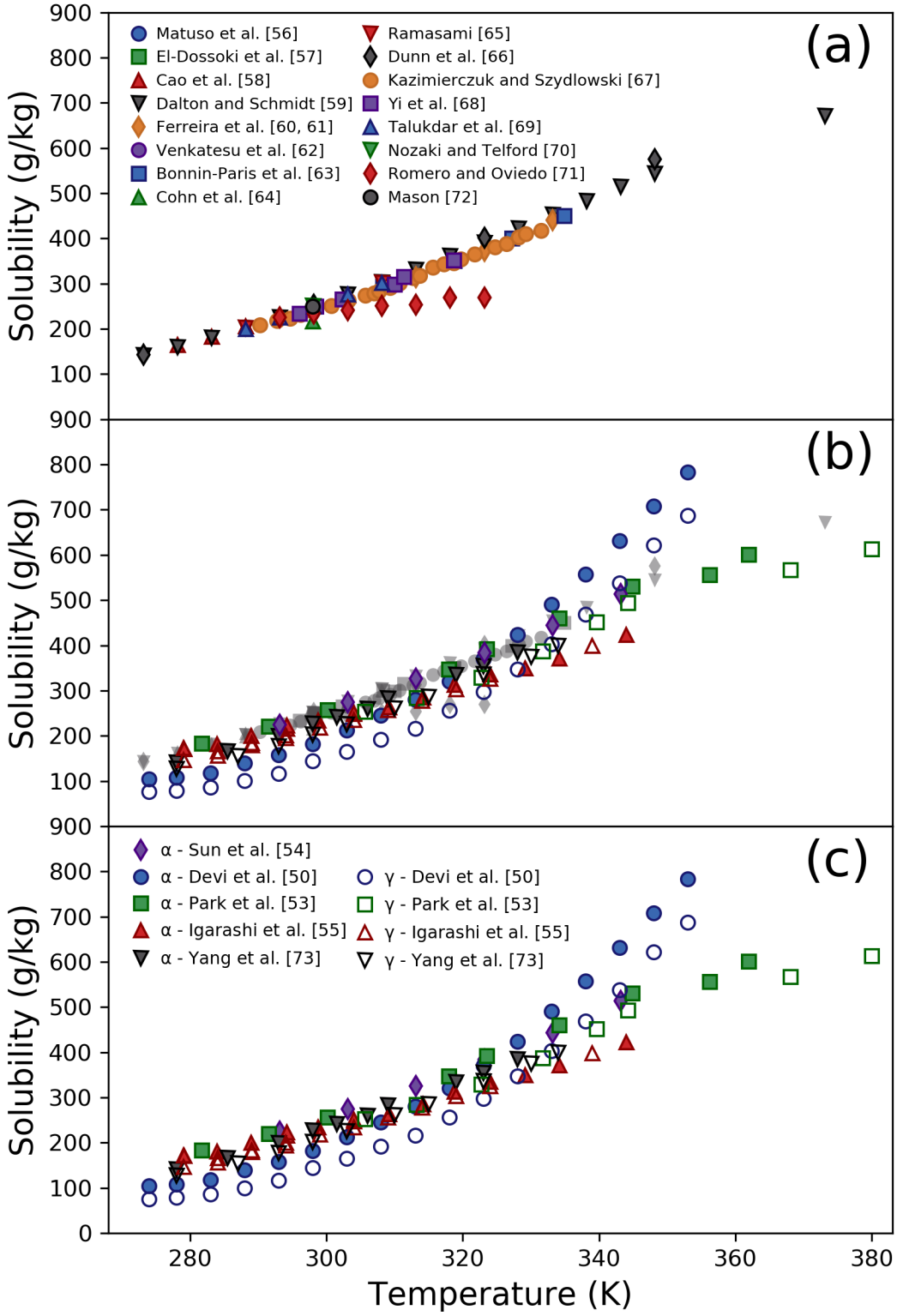

Figure 2.8: Solubility data for glycine obtained for the literature for (a) no defined polymorph, and (c) α and γ polymorphs. (b) shows the combined data sets, with undefined polymorph data greyed out, to compare defined and undefined polymorph data.

Chapter 3

Methodology

In this chapter the broad theory behind the methods that have been used throughout this work will be discussed, with detailed methodological information provided in the appropriate, subsequent chapters. First the theory relating to the experimental methods used to determine the nucleation rates and polymorphic outcomes of crystallisation experiments will be explored. Following this the theory of classical molecular dynamics simulations and the particular aspects used within the simulations performed here will be discussed. Finally broad simulation details that apply across all of the simulations performed, such as force field details and thermostat parameters, will be provided.

3.1 Experimental

3.1.1 Determining Nucleation Kinetics from Induction Times

It is possible to determine the nucleation rate of a system under a given set of experimental conditions by measuring the isothermal induction times of a series of identical experiments [74]. The isothermal induction time is the amount of time taken for a nucleation event to occur after the system has reached the temperature of interest, provided it is maintained isothermally at this temperature. By measuring the isothermal induction times of a series of experiments, a cumulative probability distribution of induction times can be constructed as shown in Figure 3.1.

Nucleation is a stochastic process and if there is a single, constant, nucleation rate then the probability of a nucleation event can be described by a Poisson distribution. The Poisson distribution states that the probability of m nuclei forming within a time interval, P_m, is:

$$P_m = \frac{N^m}{m!} \exp(-N) \tag{3.1}$$

where N is the average number of nuclei that form within the time interval. As such, the probability that no nucleation events have occurred, P_0, is given by:

$$P_0 = \exp(-N) \tag{3.2}$$

and the probability that at least one nucleus has formed within the time interval, $P_{\geq 1}$, is:

$$P_{\geq 1} = 1 - \exp(-N) \tag{3.3}$$

The average number of nuclei that form is related to the nucleation rate:

$$N = J \cdot X \cdot t \tag{3.4}$$

where J is the nucleation rate, t is the time and X is a scaling factor that will depend on the type of nucleation that is occurring. If the nucleation is homogeneous this will scale with the volume, V, of the sample; while heterogeneous nucleation will scale with the surface area, A, of the heterogeneous interface. Combining equations 3.3 and 3.4 we obtain the cumulative probability distribution function:

$$P(t) = 1 - \exp(-J \cdot X \cdot t) \tag{3.5}$$

It is important to note that the time at which a crystal is detected is not the time that was taken for nucleation to occur. Measured induction times include the time taken for nucleation to occur and the time for the nucleus to grow to a detectable size, t_d, which will vary depending on the detection method used and the growth rate of the crystals under the given experimental conditions. Including growth time in equation 3.5 gives:

$$P(t) = 1 - \exp\left[-J \cdot X \left(t - t_d\right)\right] \tag{3.6}$$

It is often convenient to define the characteristic time, τ, of nucleation where:

$$J = \frac{1}{X \cdot \tau} \tag{3.7}$$

giving another form of equation 3.6:

$$P(t) = 1 - \exp\left(\frac{-t_d - t}{\tau}\right) \tag{3.8}$$

These functions can then be fit to the cumulative probability distribution that has been constructed from the series of experiments, as demonstrated in Figure 3.1, to allow the nucleation rate to be determined.

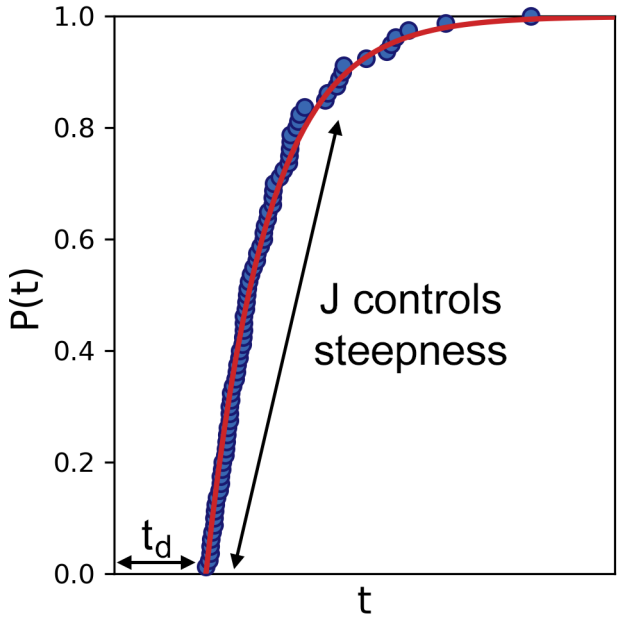

Figure 3.1: Example cumulative probability distribution function of induction times. The red line shows the fit to equation 3.8 that allows the nucleation rate to be obtained.

While this is a commonly used method to determine the nucleation rate of a given system it is important to be aware of the assumptions made and the associated limitations. Fitting the induction times of a series of experiments to a Poisson distribution assumes that there is a single, constant nucleation rate that applies to all samples. A Poisson distribution will not be able to accurately describe the nucleation behaviour of systems with multiple nucleation mechanisms occurring across the samples, or if there is some time dependence to the nucleation rate. It is also assumed that the time taken for a single nucleus to grow to a detectable size is constant across the samples. Variation in the growth rate of crystals between samples will also reduce the accuracy of the determined nucleation rates.

3.1.2 Polymorph Identification

As described in Chapter 2, it is possible for compounds to nucleate into a number of different polymorphs. As the polymorph obtained can impact a number of key properties of the final crystalline product, it is essential to be able to determine the polymorph that is obtained from a crystallisation. A number of analytical techniques are available that can differentiate between different polymorphs of the same material.

Spectroscopic techniques analyse the absorption and emission of light and other radiation to gain chemical information about the sample being analysed. Infrared (IR) and Raman spectroscopy are techniques that produce spectra that are dependent on the vibrational modes present within the sample. As these vibrational modes are sensitive to the molecular geometry, these spectra can then be used to identify polymorphs. Diffraction techniques, such as X-Ray diffraction (XRD), measure the elastic scattering of electromagnetic radiation from the crystalline material analysed, and produce diffraction patterns based on the constructive interference from the scattered radiation. The patterns obtained depend on the long range order of the crystalline material and can therefore be used to determine the polymorph of the sample.

In this work IR spectroscopy, Raman spectroscopy and X-Ray Powder Diffraction (XRPD) have all been used to analyse the glycine samples obtained from the nucleation experiments. Multiple techniques have been used to ensure consistency between the techniques and provide greater confidence in the obtained polymorphic distributions. The basic theory of these techniques will be described below.

IR and Raman Spectroscopy

IR spectroscopy measures the absorption of IR radiation by the analysed sample, and uses this to gain chemical information. There are multiple ranges of IR radiation that can be used (such as near-IR or mid-IR), however mid-IR (400-4000 cm^{-1}) is the most appropriate choice for polymorph identification as this region is related to the fundamental vibrations and is therefore influenced by the molecular geometry [75]. During IR spectroscopy, IR radiation of varying frequency is passed through the sample. When the frequency of the radiation matches a vibrational frequency within the molecule, if the vibration causes a change in the dipole moment of the molecule, energy is absorbed by the molecule changing its vibrational energy level [75]. The amount of radiation that is absorbed by the material is plotted for each frequency providing an IR spectra of the sample that is sensitive to the molecular geometry, and can therefore be used to distinguish between polymorphs. Mid-IR spectroscopy is often implemented as Fourier transform infrared spectroscopy (FTIR) as this allows spectral data to be collected over a wide range of frequencies simultaneously and is then deconvoluted to the spectra via a Fourier transformation [76].

Unlike IR spectroscopy, where radiation over a range of frequencies is used, in Raman spectroscopy monochromatic light is shone onto the material and is scattered by the electrons. The incident light distorts the electron cloud to form an unstable 'virtual state' that, due to this instability, quickly returns to a lower state and releases the light. The majority of these scattering events are elastic, returning to its original state and scattering light that is of the same frequency of the incident light. This elastic scattering is known as Rayleigh scattering. Rarely (approximately every 10^6 - 10^8 photons) nuclear motion is induced by the scattering process resulting in inelastic scattering and the electron returns to a different energy level than it was originally in, scattering light at a frequency that is higher or lower than the incident light [77]. If the electron returns to a higher energy level than its original position it is known as Stokes scattering, while returning to a lower

energy level is known as anti-Stokes. Figure 3.2 demonstrates the transitions occurring during Rayleigh, Stokes and anti-Stokes scattering. At room temperature the majority of molecules will exist in their lowest energy state, and so anti-stokes scattering will be significantly weaker than stokes scattering. As such, Raman setups will typically measure stokes scattering.

During Raman spectroscopy the light is focused onto the sample and the scattered light is collected in the spectrometer. Rayleigh scattered light is filtered and the remaining scattered photons are counted by their wavenumber shift providing the spectra for the sample.

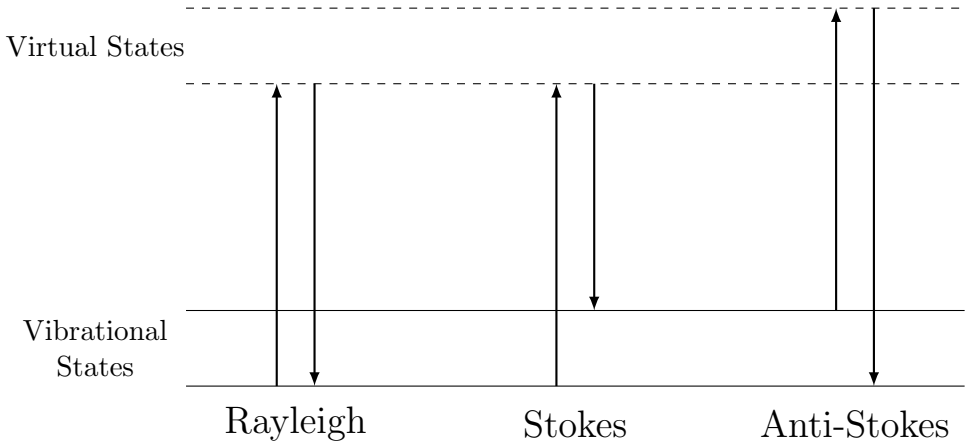

Figure 3.2: Energy diagram to demonstrate the Ralyleigh and Raman scattering processes.

For a material to be IR active it is necessary for a change in the dipole moment to occur as a result of the vibration that is caused by the absorption of the IR radiation. In contrast, a material is Raman active when the vibration results in a change in polarisability of the molecule. As such, not all compounds will be active for both IR and Raman spectroscopy, however they are often used as complimentary techniques.

X-Ray Powder Diffraction

X-Ray diffraction (XRD) is an important tool for the analysis of crystalline structures. When X-Rays are directed onto a sample they can be scattered by the atoms present. Within a crystal, the atoms are present in a perfectly repeating 3D structure and the scattered X-Rays can interfere with each other. While the majority of this interference is destructive, at certain incidence angles constructive interference occurs [78]. During X-Ray diffraction, X-ray radiation is directed onto the sample at a range of incidence angles and the diffracted X-rays are detected and counted. The angles that result in constructive interference are determined by the spacing of the planes within the crystal structure and result in spots in the obtained diffraction pattern. As the positions of these spots are related to the spacing of the planes of atoms within the crystal structure [78], these can be used to identify the crystal structure of a given sample.

Single crystal XRD requires a high quality, single crystal that is large enough to be mounted and aligned with the X-Ray source and detector to be obtained, which can be challenging. It is possible to grind a sample into a large number of fine crystals and to carry out X-Ray powder diffraction (XRPD). Provided that random orientation of the ground material is obtained (removing any preferred orientation effects) when an incidence angle that results in constructive interference is used the randomly oriented crystals will diffract the X-Rays in all directions resulting in concentric circles in place of the spots present in the patterns obtained from single crystal XRD. This can allow the polymorph to be determined without the need to obtain a single, high quality crystal.

As XRD techniques rely on the long range order of crystalline material, it is not necessary for vibrational modes within the sample to follow selections rules, as is the case for Raman and IR. However, both Raman and IR typically allow for faster data acquisition. XRPD relies on the random orientation of the sample and preferred orientation effects can

influence the intensities of the peaks in the obtained diffraction pattern [78]. XRPD is therefore more sensitive to sample preparation than IR or Raman.

3.2 Molecular Dynamics Simulations

Classical molecular dynamics (MD) simulations are a powerful tool that allow us to access information at the molecular level in a way that is unachievable by current experimental techniques. These simulations use statistical mechanics to determine macroscopic properties from the microscopic details of a many body system. MD simulations can be used to obtain thermodynamic and transport properties by examining the time evolution of a given system. These simulations work under the assumption that the motion of the atoms within the system obey the laws of classical mechanics.

Atoms are placed virtually within a simulation box and the potential energy, $U(r)$, for each atom, i, can be modelled based on its position, r_i. From the potential energies the force $F_i(r)$, acting on atom i can be determined from the derivative of the potential energy:

$$F_i(r) = -\frac{\partial U(r)}{\partial r_i} \tag{3.9}$$

This process is repeated in order to determine the forces acting on each atom within the simulation box.

Newtons equations of motion can then be integrated in order to determine the velocity of each atom. The system is moved forward by some small time step, resulting in a new set of positions, and the process is repeated to allow the system to progress through time. A number of algorithms can be used for integrating the equations of motion. The Verlet algorithm [79] uses an atoms current position, $r(t)$, current velocity, $v(t)$ and its current

acceleration, $a(t)$, in order to determine its position and velocity at a time $t + \delta t$:

$$r(t + \partial t) = r(t) + \partial t\, v(t) + \frac{1}{2} t^2 a(t) \tag{3.10}$$

$$v(t + \partial t) = v(t) + \frac{1}{2} \partial t \left[a(t) + a(t + \partial t) \right] \tag{3.11}$$

This requires the acceleration at times t and $t + \partial t$ to be stored, and as such will result in a loss of precision when calculating the velocities using equation 3.11. The velocity Verlet algorithm avoids this loss in precision through the use of a half step, $t + \frac{1}{2}\partial t$ [80]. The velocity is advanced by this half step:

$$v(t + \frac{1}{2} \partial t) = v(t) + \frac{1}{2} \partial t\, a(t) \tag{3.12}$$

and this allows the position of the atom at time $t + \delta t$ to be determined:

$$r(t + \partial t) = r(t) + \partial t\, v(t + \frac{1}{2} \partial t) \tag{3.13}$$

A force evaluation is then carried out in order to determine the acceleration at time $t + \delta t$ which then allows the velocity at time $t + \delta t$ to be obtained:

$$v(t + \partial t) = v(t + \frac{1}{2} \partial t) + \frac{1}{2} \partial t\, a(t + \partial t) \tag{3.14}$$

The velocity Verlet algorithm has been used to integrate the equations of motion for the simulations performed in this work, however other algorithms such as the leap-frog algorithm can also be used for this purpose [80].

3.2.1 Periodic Boundary Conditions

As mentioned above, MD simulations are performed on atoms that have been placed virtually within a simulation box. The system size of these simulations are limited by computational power, and due to these limited sizes surface effects can be significant. Surface effects can be overcome by employing periodic boundary conditions (PBC). Under PBC the simulation box is replicated into an infinite lattice of identical cells. There are no barriers between the simulation box and the periodic images and, as an atom leaves the box its periodic image well re-enter the box at the opposite side with the same velocity, allowing the system to maintain its number density. This allows bulk phases to be simulated with no surfaces present to induce surface effects. It is necessary to ensure that atoms will not interact with themselves across periodic boundaries and as such the length of the periodic box should be twice the size of the cut-offs used for intermolecular interactions. Interaction cut-offs will be discussed in further detail in section 3.2.2.

3.2.2 Force Fields

Once the simulation box has been populated by atoms, it is necessary to model all of the interactions that the atoms will experience within the system based on their positions. The model used to represent these interactions, and therefore determine their potential energy, is known as a force field. These models describe all of the interactions that will be experienced by atoms within the system and can largely be broken down into intramolecular and intermolecular interactions:

$$U = \underbrace{\sum_{bonds} E_{bond} + \sum_{angles} E_{angle} + \sum_{dihedrals} E_{dihed} + \sum_{impropers} E_{improp}}_{\text{Intramolecular}} + \underbrace{\sum_{\substack{non\text{-}bonded \\ pairs}} (E_{LJ} + E_{coul})}_{\text{Intermolecular}} \tag{3.15}$$

Here the potential energy, U, is the sum of all of the intramolecular and intermolecular interactions. Each of the components of equation 3.15 will be described in greater detail below.

Intramolecular Interactions

Intramolecular interactions describe the internal interactions of a molecule: bonds, angles, dihedrals and improper dihedrals. Bond stretching is typically modelled as a harmonic spring:

$$E_{bond} = k_b(r - r_0)^2 \tag{3.16}$$

where k_b is the bond strength, r is the bond length and r_0 is the equilibrium bond length, however other functional forms, such as the Morse potential [81], also exist.

Angle potentials are also typically in the form of a harmonic spring:

$$E_{angle} = k_a(\theta - \theta_0)^2 \tag{3.17}$$

where k_a is the angle strength, θ is the bond angle, and θ_0 is the equilibrium bond angle.

Dihedral potentials describe the torsion of a bond in the centre of four consecutively bonded atoms. If we consider four consecutively bonded atoms (atoms i, j, k and l) then the torsional angle, ϕ, is the angle between the planes that pass through atoms i, j and k and atoms j, k and l. There is much greater variation in the functional form of dihedral potentials between force fields than for bonds or angles, where harmonic springs are most dominant, and the functional form used within a given simulation will depend on the force field that has been selected.

Improper dihedrals are a type of dihedral between atoms that are not consecutively bonded, but instead involve a central atom bonded to three other atoms. These are used to maintain planar groups (e.g. aromatic rings) or to main the correct chirality of certain molecules. Impropers are often implemented using the same functional forms as proper dihedrals to allow them to be implemented without major changes required by an MD code. Figure 3.3 shows schematics to demonstrate the four types of intramolecular interactions.

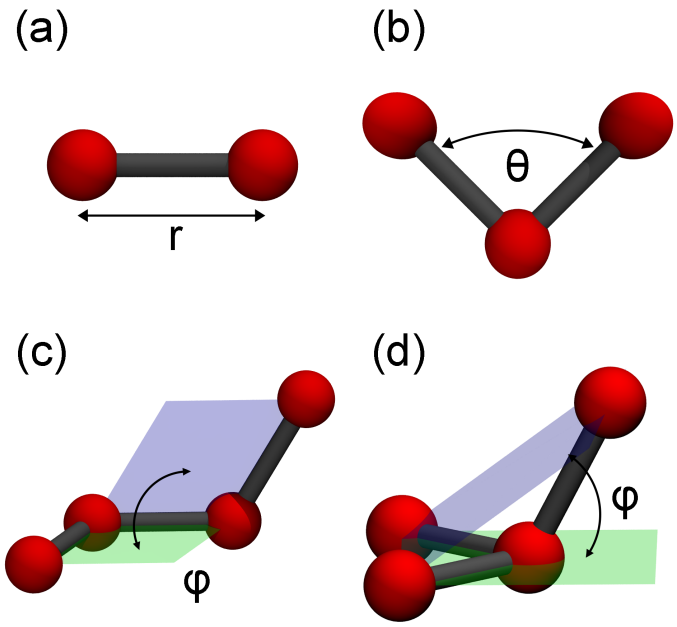

Figure 3.3: Schematic to demonstrate the four types of intramolecular interactions: (a) bonds (b) angles (c) dihedrals and (d) impropers. The blue and green areas in (c) and (d) indicate the two planes that determine the dihedral and improper angles.

Intermolecular Interactions

Intermolecular interactions include dispersion and electrostatic interactions. Dispersion interactions are often modelled in MD simulations using a Lennard Jones (LJ) 12-6 poten-

tial:

$$E_{LJ} = 4\epsilon_{ij}\left[\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6}\right] \tag{3.18}$$

where r_{ij} is the distance between the centres of the two interacting atoms, i and j; ϵ_{ij} is the well depth of the interaction and σ_{ij} is the distance where the interaction is equal to zero. Figure 3.4 demonstrates a typical LJ interaction with labels to indicate ϵ_{ij} and σ_{ij}.

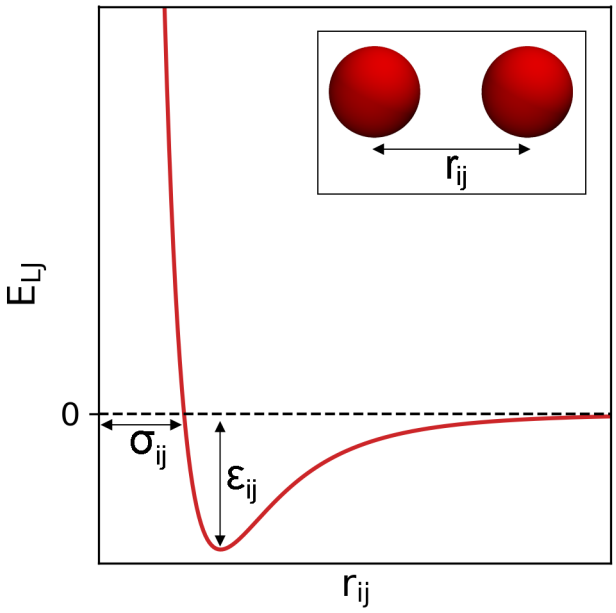

Figure 3.4: Plot of a Lennard Jones 12-6 interaction.

Force fields will typically only provide LJ parameters for interactions between atoms of the same type, and so it is necessary to use combining rules to create parameters for each combination of atom types within a given system. The Lorentz-Berthelot rules are the most frequently used combining rules for this purpose, however some force fields will specify that other combining rules should be used. Using the Lorentz-Berthelot rules, the arithmetic mean is used to combine σ parameters:

$$\sigma_{ij} = \frac{\sigma_{ii} + \sigma_{jj}}{2} \tag{3.19}$$

while the geometric mean is used to combine ϵ parameters:

$$\epsilon_{ij} = \sqrt{\epsilon_{ii} \cdot \epsilon_{jj}} \tag{3.20}$$

This method has been used to determine the LJ parameters for all atom type combinations within this work.

Typical force fields will represent the electrostatic interactions of atoms as fixed point partial charges that are positioned at the centre of the atom. The electrostatic interactions between atoms are then calculated using Coulomb's law:

$$E_{coul} = \frac{q_i \, q_j}{4\pi\epsilon_0 \, r_{ij}} \tag{3.21}$$

where q_i and q_j are the partial charges on the two interacting atoms i and j, r_{ij} is the distance between the atoms and ϵ_0 is the vacuum permittivity. As the partial charges are fixed throughout the simulation, these are non-polarisable models. In reality, the distribution of charge for a given molecule will adjust depending on its current environment. A number of polarisable force fields exist that will allow for this adjustment to the environment, however these models are considerably more expensive than the standard non-polarisable models [80]. In addition to this, standard non-polarisable models have undergone significantly more refinement due to their popularity and as such current polarisable models are not always more accurate despite accounting for polarisation effects [82].

Intermolecular interactions have a significant computational cost as, without any further restrictions applied, each atom will interact with every other atom within the system and not just those within the same molecule. Due to this high cost, and to prevent atoms from interacting with themselves across periodic boundaries, it is necessary to implement a spherical boundary for intermolecular interactions. Using these cut-off schemes any intermolecular interactions between atoms that are greater than the cut-off distance, r_c, from

each other are set to zero. Introducing a cut-off will result in a perturbation in the system and so it is necessary to choose a cut-off that is sufficiently large to ensure that this perturbation is small. In homogeneous systems tail corrections can be applied to LJ interactions to reduce this perturbation [83].

While LJ interactions decay significantly with distance, electrostatic interactions are still significant at long ranges and so it is not sufficient to apply tail corrections and it is necessary to account for these interactions at long ranges [80]. Instead, short range electrostatics (for $r < r_c$) are calculated explicitly using equation 3.21, whilst long range interactions are estimated using long range solvers. Within this work the particle-particle particle-mesh (PPPM) solver has been used for long range calculations, but other methods such as standard Ewald summations [83] or particle-mesh-Ewald (PME) [84] are also available. The PPPM method maps charges to a 3D mesh and uses fast Fourier transforms to solve Poisson's equation on the mesh and interpolates electric fields on the mesh points to the atoms within the system [85].

3.2.3 Temperature and Pressure Control

While the default ensemble for MD simulations is NVE (i.e. fixed number of particles, volume and energy) it is often more useful to perform simulations in ensembles that are more comparable to experimental conditions such as NVT (fixed number of particles, volume and temperature) or NPT (fixed number of particles, pressure and temperature).

To perform an MD simulation at a constant temperature (the system in thermal equilibrium with a heat bath at a fixed temperature) it is necessary to introduce a thermostat to the simulation. There are a wide range of thermostats available with common choices including the Berendsen [86], Andersen [87] and Nosé-Hoover [88, 89] thermostats. In this

work simulations have been performed using the Nosé-Hoover thermostat. This is a deterministic method that assigns a coordinate and momentum to the heat bath and carries out a simulation on this extended system. Energy is allowed to flow between the system and the reservoir, with the reservoir having a thermal inertia. Some consideration must be given to the thermal inertia that is selected, too low a value will result in wild temperature fluctuations, while too large a value with result in excessive equilibration times.

In order to simulate a system at a constant pressure (such as in the NPT ensemble) it is necessary to adjust the volume of the box throughout the simulation. This is achieved in a similar manner to maintaining a constant temperature, and a 'pressure bath' is coupled to the system and the volume of the box is allowed to adjust. Typically the volume of the box is allowed to adjust isotropically (i.e. all dimensions of the box adjust by the same amounts) however in some instances, such as simulating solids, it is necessary to allow the shape of the simulation box to change. In this work a Nosé-Hoover barostat has been used to control the pressure which, similarly to the thermostat, couples the system to a pressure bath with a set inertia [90].

3.3 Simulation Details

In this section general simulation details, such as force field choices, will be discussed. This section is intended to give an overview of settings and parameters that have been used throughout the simulations in this work, with finer details provided in the appropriate results chapter.

All simulations have been carried out using the LAMMPS MD code [91]. In all cases a timestep of 2.0 fs was used. Temperatures were maintained at 298 K using a Nosé-Hoover thermostat with a damping parameter of 1 ps. Pressures in simulations performed in the

NPT ensemble were maintained using a Nosé-Hoover barostat with a damping parameter of 2 ps. The damping parameters of the thermostat and barostat are related to the inertia of the related thermal and pressure baths, as discussed in section 3.2.3, and describe the length of time taken for the temperature and pressure of the system to relax. Detailed information about which simulations were carried out in each ensemble can be found in the appropriate chapters.

LJ interactions were truncated at a cut-off of 1.4 nm, while short-range electrostatics were calculated below 0.98 nm. Long-range electrostatics were calculated using PPPM with an accuracy of 1×10^{-6}. Lennard Jones 1–4 interactions were reduced to 0.5, while electrostatic 1–4 interactions were reduced to 0.833 as intended for AMBER style force fields.

3.3.1 Force Fields Used Within This Work

The accuracy of the results obtained via molecular simulation are highly dependent on the quality of the force field that is selected. In order to accurately simulate crystallisation, the force field must be able to accurately capture both the solution phase and solid phase behaviour of a given material. Cheong and Boon performed an extensive comparison of four forcefields: Charmm27 (Charmm), the general AMBER force field (GAFF), OPLS-AA (OPLS) and Gromos53a6 (Gromos) in combination with three water models: SPC, SPC/E and TIP3P and 6 different charge sets: default, CNDO, DNP, DZP, LCAO and 6-31G*, to determine their performance in reproducing the properties of crystalline α-glycine, glycine solutions and an α-glycine crystal in contact with a supersaturated solution [92].

Cheong and Boon determined the lattice energies and lattice parameters obtained for each force field and charge set combination for α-glycine. Following this they investigated

the diffusivity and densities of glycine solutions for the various force field, charge set and water model combinations. The solvation free energy for each force field in combination with the SPC/E water model (due to it's performance in the previous tests) was also determined. Based on these results a smaller selection of force field and charge combinations were selected (Charmm/CNDO, Charmm/6-31G*, GAFF/CNDO, GAFF/DNP, GAFF/DZP, OPLS/CNDO, and default OPLS) to simulate α-glycine in contact with a supersaturated glycine solution in an attempt to simulate crystal growth. From these simulations Cheong and Boon proposed that GAFF in combination with the CNDO charges is the most suitable choice for crystallisation studies as it is able to successfully reproduce the solution properties and heat of solution, however they do note that the lattice energy and parameters are significantly underpredicted for the bulk crystal.

The GAFF/CNDO glycine and SPC/E water force field combination has been further validated by Parks *et al.* who have determined the solubility curves of α and β-glycine for this model across a temperature range of 277.5–322.5 K [93]. When the calculated solubilities for α-glycine are compared to those obtained experimentally by Yang *et al.* [73], the values differ by 0.7%-23% across the whole temperature range, with particularly good agreement for values close to room temperature. While experimental solubility data is not available for β-glycine, due to its highly metastable nature, the GAFF/CNDO force field was found to correctly predict a higher solubility, and therefore lower stability, for β-glycine than α-glycine across the whole temperature range.

Interpolation by second order polynomial to the simulated solubility curve obtained for α-glycine by Parks *et al.* gives a solubility of 232 g/kg at 298 K. This value fits comfortably within the range of experimental values given in section 2.2.1 (182-248 g/kg), giving confidence that the levels of supersaturation within these simulations are representative of experimental values.

Glycine

Based on the recommendations of Cheong and Boon [92], we have modelled glycine molecules using GAFF [94] with the CNDO charge set [95]. Figure 3.5 demonstrates the atom types that have been selected for the glycine molecule labelled using the GAFF naming convention.

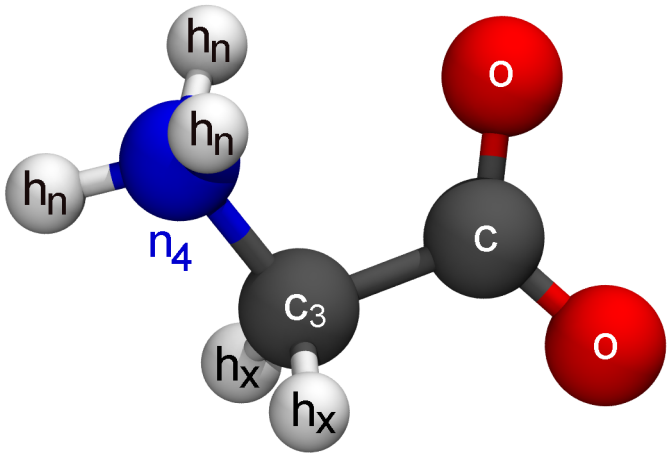

Figure 3.5: Diagram of a glycine molecule with each atom labelled with the GAFF atom type used within the simulations.

Within GAFF bond and angles are modelled using harmonic potentials (equations 3.16 and 3.17) while the dihedral potential is given by:

$$E_{dihed} = \sum_{i=1}^{m} k_i \left[1.0 + cos(n_i\phi + d_i)\right] \tag{3.22}$$

where k_i is the torsional strength, n_i is the multiplicity, ϕ is the torsional angle, and d_i

is the phase angle.

Non-bonded interactions are described using the Lennard-Jones potential and Coulomb interactions between point charges (equations 3.18 and 3.21). The two oxygen (o) atoms have charges of -0.526 and -0.483, the carboxylic acid group carbon (c) has charge +0.374, the aliphatic carbon (c3) has charge +0.374, aliphatic hydrogens (hx) have charges +0.033 and +0.030, the nitrogen (n4) has charge +0.022 and the amine hydrogens (hn) have charges +0.164, +0.208 and +0.199. All force field parameters for the glycine molecules are provided in Table 3.1

3.3.2 Water

Water molecules have been represented using the SPC/E model [96] as recommended by Cheong and Boon [92] due to its ability to reproduce the densities and diffusion coefficients of aqueous glycine solution. SPC/E is a three point model: a model with three interactions sites placed on the centre of each atom within the water molecule. The SPC/E model has a point charge on each of the three atoms and a LJ interaction site on the oxygen atom. The hydrogen atoms have no LJ interactions. This is a rigid model and as such the bond lengths and angle are fixed using the SHAKE algorithm [97]. The parameters for the SPC/E model are given in Table 3.2.

3.3.3 Tridecane

Tridecane molecules have been represented using the AMBER-ii [98] force field, which is an extension of the AMBER force field to provide a better description of long chain alkanes. This force field uses the same functional form as GAFF (as described above,

Table 3.1: GAFF force field parameters for glycine. Atom type X in the dihedral parameters denotes any atom type.

LJ Parameters		
Atom Type	ϵ (kcal/mol)	σ ($\AA$)
n4	0.1700	3.25
c3	0.1094	3.40
c	0.0860	3.40
o	0.2100	2.96
hn	0.0157	1.07
hx	0.0157	1.96

Bond Parameters			
Atom Type		$\mathbf{k_b}$	$\mathbf{r_0}$
1	2	(kcal/mol)	(Å)
hn	n4	369	1.066
n4	c3	293.6	1.499
c3	hx	338.7	1.091
c3	c	328.3	1.508
c	o	648.0	1.214

Angle Parameters				
Atom Type			$\mathbf{k_a}$	$\theta_\mathbf{0}$
1	2	3	(kcal/mol)	(°)
hn	n4	hn	40.5	108.11
hn	n4	c3	46.2	110.11
n4	c3	hx	49.0	107.91
n4	c3	c	65.8	111.54
c3	c	o	68.0	123.11
o	c	o	79.1	127.33
hx	c3	hx	39.0	110.74
c	c3	hx	47.4	108.64

Dihedral Parameters							
Atom Type				**i**	$\mathbf{k_i}$	$\mathbf{n_i}$	$\mathbf{d_i}$
1	2	3	4		(kcal/mol)		(o)
X	c3	n4	X	1	0.15	3	0
X	c	c3	X	1	0.00	2	180
X	o	c	X	1	1.10	2	180

Table 3.2: Force field parameters for the SPC/E water model.

Parameter	Value
r_{OH}	1.00 Å
$<_{HOH}$	109.47 °
ϵ_{OO}	0.1555 kcal/mol
σ_{OO}	3.166 Å
q_O	-0.8476
q_H	0.4238

equations 3.16 - 3.22) and the force field parameters are provided in Table 3.3. A charge of -0.012 is assigned to the CH_3 carbons and +0.012 is assigned to the connected CH_2 carbon. All other tridecane atoms are chargeless.

3.3.4 Other Materials

Simulations involving graphite and PTFE have also been performed in Chapter 6. Each of these materials have been represented using GAFF. In both of these cases these materials have been fixed in place and no dynamics for these molecules have been performed (see the methodology of Chapter 6 for specific details) and as such only intermolecular parameters were required. Lennard Jones parameters are provided in Table 3.4. The carbon atoms of graphite were chargeless, whilst the charges for the PTFE were obtained using the AM1-BCC method within Antechamber and are given in Table 3.5

Table 3.3: AMBER-ii force field parameters for tridecane. Note that the dihedral term for the carbon backbone of tridecane has three combined terms, whilst all other dihedrals have one.

LJ Parameters							
Atom Type					ϵ (kcal/mol)	σ (Å)	
C (CH2)					0.1560	3.12	
C (CH3)					0.1840	3.12	
H					0.0124	2.66	
Bond Parameters							
Atom Type					$\mathbf{k_b}$ (kcal/mol)	$\mathbf{r_0}$ (Å)	
C	C				240.0	1.52	
C (CH2)	H				340.0	1.096	
C (CH3)	H				340.0	1.092	
Angle Parameters							
Atom Type					$\mathbf{k_a}$ (kcal/mol)	θ_0 (°)	
H	C	H			33.0	107.0	
C	C	H			52.0	110.7	
C	C	C			30.0	109.5	
Dihedral Parameters							
Atom Type				**i**	$\mathbf{k_i}$ (kcal/mol)	$\mathbf{n_i}$	$\mathbf{d_i}$ (°)
H	C	C	C	1	0.16	3	0
H	C	C	H	1	0.16	3	0
C	C	C	C	1	0.16	3	0
C	C	C	C	2	0.09	2	180
C	C	C	C	3	0.06	1	180

Table 3.4: GAFF LJ parameters for graphite and PTFE.

Atom Type	ϵ (kcal/mol)	σ (Å)
Graphite		
ca	0.0860	3.4
PTFE		
c3	0.1094	3.40
f	0.061	3.12

Table 3.5: Charges assigned to the atoms within the PTFE molecules using the AM1-BCC method.

Backbone Position	$\mathbf{q_C}$	$\mathbf{q_F}$
1	0.6100	-0.2008
2	0.3477	-0.1762
3	0.3497	-0.1767
4	0.3517	-0.1762
5	0.3517	-0.1757
6	0.3517	-0.1757
7	0.3517	-0.1757
8	0.3517	-0.1757
9	0.3517	-0.1757
10	0.3517	-0.1757
11	0.3497	-0.1767
12	0.3477	-0.1762
13	0.6100	-0.2008

Chapter 4

Nucleation Experiments

4.1 Introduction

Primary nucleation is the formation of an initial crystalline particle from solution, and influences a number of the properties of the resulting crystalline product. Despite the importance of nucleation, there is still lack a fundamental understanding of this process. This limited understanding prevents the accurate control, or meaningful prediction, of the position at which nucleation will occur, which crystal structure will form or how long it will take for nucleation to take place [6]. A more complete understanding of the nucleation process would allow for greater control over crystallisation processes.

As described in Chapter 2, small-scale experimental set-ups are often used to investigate nucleation. These small-scale setups differ from macroscale crystallisations in a number of ways, and it is vital to understand the influence these differences can have on the observed nucleation behaviour in order to effectively scale-up crystallisation processes. One of these key differences is the presence of very particular interfaces, such as polymer

coatings/tubings or the oils in microfluidic/microwell type experiments, at high surface area to volume ratios. These interfaces are typically assumed to be inert, however their presence at such high surface area to volume ratios may allow heterogeneous nucleation to dominate the observed nucleation behaviour.

In this chapter, the impact of liquid-liquid interfaces on nucleation behaviour will be investigated using a model compound of aqueous glycine solution in contact with tridecane. A large number of small scale experiments of glycine solution with and without contact with the tridecane interface have been performed, and results investigating the impact of concentration, sample preparation method and oil-solution volume ratio will be presented. This will be followed by a quantitative assessment of the nucleation kinetics from the obtained induction times and a discussion of the issues related to such analysis.The polymorphic outcome of the nucleation experiments will then be presented and the results obtained using a range of techniques for polymorph identification will be compared. Finally, the results obtained for isonicotinamide nucleating from a ethanol solutions, a separate system which demonstrates similar behaviour when placed in contact with a fluorinated oil, will be shown.

4.2 Methodology

4.2.1 Glycine Solution Experiments

Glycine solutions were prepared using glycine powder (Sigma-Aldrich, for electrophoresis $\geq$ 99%) and deionized water from an in-house dispenser (Milli-Q, 18.2 MΩ cm). The oil used for the oil-solution interface experiments was tridecane (Sigma-Aldrich $\geq$99%).

Samples were prepared using two different methods in order to determine the impact of the preparation method on the nucleation behaviour. Method one involved glycine solution being prepared directly within individual vials in order to avoid issues related to transferring solutions at high temperatures. If solutions are transferred at high temperatures it is necessary to maintain all of the equipment at a temperature above the saturation temperature for the given concentration. If the pipette tips or vials cool significantly this can induce supersaturation within the solution and may trigger nucleation within the sample before it can be cooled to the desired isothermal conditions in a controlled manner. Similar issues were discussed by Little *et al.* [35] who reported that holding the glycine solution within the pipette tip for 30 *s* resulted in a significant increase in the nucleation of glycine at short time-spans and resulted in data with low reproducibility. Samples were prepared at a range of concentrations from 275 to 450 $g/kg_{solvent}$, denoted herein as g/kg. In each experiment new 1.5 mL glass vials (VWR 548-0018) were used. Vials were washed with deionised water and dried prior to the preparation of the solutions. The required amount of glyince powder was weighed directly into the glass vials and 1±0.01 mL of deionised water was pipetted into each vial. For the oil-interface experiments, 200±4 μL of tridecane was then pipetted on top of the water to create an oil-solution interface.

The vials were then transferred to a Polar Bear Plus Crystal. The Polar bear is a precision heating and cooling platform produced by Cambridge Reactor Design that uses interchangeable plate attachments to allow for accurate (±0.1°C) temperature control for a range of vessels from vials to round bottom flasks. The vials were held overnight to fully dissolve the glycine. Solutions prepared at 275-365 g/kg were held at 343 K, while those prepared at 400-450 g/kg were held at 363 K to ensure that they were safely below the solubility of γ-glycine which is poorly reported in the literature with few sources with contradictory results [36]. A total of 40 vials were prepared at each concentration of 275, 307, 333, 365, 400 and 450 g/kg without the oil interface, and 80 vials were prepared at each concentration of 275, 307, 333, 365 g/kg with the oil interface.

The vials were checked visually to ensure that the glycine had fully dissolved and were then cooled in the Polar Bear at a controlled rate of 1.5 K/min to 298 K. The vials were then transferred into vial racks placed within an incubator set to 298 K for temperature control. Webcams were used to capture images of the vials every 5 min to allow for the induction time of the vials to be measured. A schematic of the observation set-up is shown in 4.1. The full set of experimental conditions investigated using sample preparation method one are given in Table 4.1.

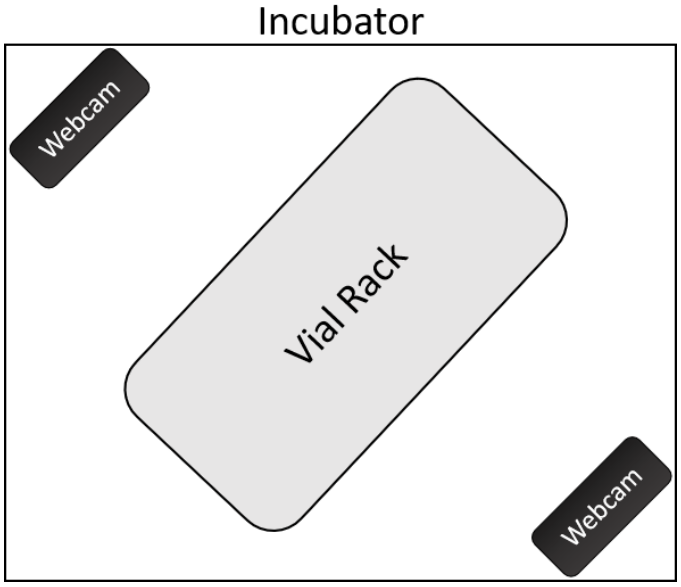

Figure 4.1: Schematic of the vial observation setup.

In method two samples were prepared using a stock solution. A 307 g/kg stock solution was prepared by weighing the required amounts of glycine and deionised water into a 100 mL glass bottle with a magnetic stirrer and sealed. A 200±4 μL aliquot of oil was pipetted into 40 pre-washed and dried glass vials and were left in an incubator at 333 K along with the solution, which was stirred overnight. Solution was then pipetted from the bottle into each vial on top of the tridecane, with the solution sinking below the layer of oil. A fresh pipette tip was used for each vial and pipetting was performed inside the incubator in order to prevent any crystallisation occurring during the solution transfer process. To ensure that temperatures remained high within the incubator, the solution

Table 4.1: List of experimental conditions investigated using preparation method one. Experiments A-K were prepared at 343 K, and experiments L and M were prepared at 363 K, to ensure dissolution of glycine.

Experiment	Concentration (g/kg)	Supersaturation	Interface type	Number of Vials
A	275	1.11	Oil	28
B	275	1.11	Oil	40
C	275	1.11	Air	40
D	307	1.24	Oil	40
E	307	1.24	Air	40
F	307	1.24	Oil	40
G	333	1.34	Oil	40
H	333	1.34	Air	40
I	333	1.34	Oil	39
J	365	1.47	Oil	40
K	365	1.47	Air	40
L	450	1.61	Air	40
M	400	1.81	Air	40
N	275	1.11	Oil	12
O	333	1.34	Oil	1
P	365	1.47	Oil	40

was transferred into vials in short bursts, closing the incubator between sets and allowing the temperature to return to 333 K before beginning the next set. Due to the temperature limits of the incubator, the solution preparation and transfer were carried out at 333 K. To ensure that any crystals that might have formed during the preparation of the samples were redissolved, the vials were transferred to the Polar Bear at 343 K where they were held for 2 h before being cooled at a controlled rate of 1.5 K/min and transferred to the incubator at 298 K, and observed via webcam as described above.

In order to investigate the effects of the volumes of solution and oil, a number of experiments were performed using method two at varying oil-solution volume ratios. In all experiments the total volume of both the oil and solution combined was 1.2 mL (to match the previous experiments). Oil-solution volume ratios of 1:5, 1:1, 1:2 and 2:1 were prepared at a glycine concentration of 307 g/kg. A total of 40 vials were prepared in each

Table 4.2: List of experimental conditions investigated using preparation method two. All samples prepared at a concentration of 307 g/kg. Note that Experiment R does not have a Oil:Solution volume ratio as this experiment has an air-solution interface to act as a control.

Experiment	Oil:Solution Volume Ratio	Number of Vials
Q	1:5	40
R	-	40
S	2:1	40
T	1:2	40
U	1:1	40

experiment. The full set of experimental conditions investigated using sample preparation method two are given in Table 4.2.

4.2.2 Polymorph Identification

Once a crystal formed within a vial, it was removed from solution and left at room temperature for 1 day to dry prior to identifying the polymorph. Polymorphs were primarily determined using Infrared (IR) spectroscopy and a selection of samples were also checked by Raman spectroscopy and X-ray powder diffraction (XRPD) to validate the results obtained by IR.

Raman microscopy allows for samples to be analysed directly without any pre-treatment or grinding of the crystals. Raman spectra were obtained using a Horiba XploRa+ microscope with a 532 nm laser at a grating of 2400 gr/mm. Due to the small spot size of the instrument (<1 μm) a number of measurements were taken for each crystal to be sure the measurements were representative. Two key spectral regions have been identified for determining the polymorph of glycine and are shown in Figure 4.2. In the first spectral region, from 1250-1600 cm^{-1}, α can be identified from the peak at 1455 cm^{-1}. In the

second spectral region, from 2900-3050 cm^{-1}, α glycine has two peaks at 2972 cm^{-1} and 3006 cm^{-1}, which shift down towards 2964 cm^{-1} and 2999 cm^{-1} for γ glycine.

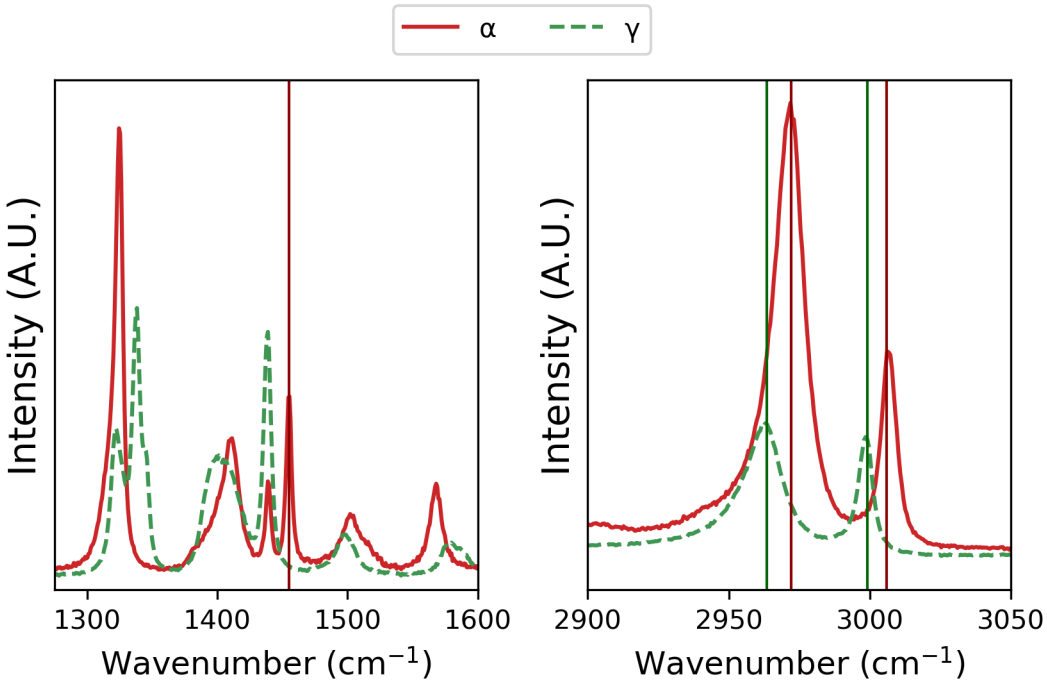

Figure 4.2: Example Raman spectras obtained for γ and α glycine in the spectral regions used to identify the polymorph obtained. The dark red and green vertical lines highlight the peak positions used to identify α and γ glycine respectively.

Prior to analysis by IR or XRPD it was necessary to grind the samples to a powder. IR spectra were obtained using an ABB MB3000 spectrometer at a resolution of 1 cm^{-1}. Absorbance spectra were averaged over eight scans in the wavenumber range 700-1000 cm^{-1}. All spectra were collected at ambient temperatures. The α polymorph can be identified by a characteristic peak at 907 cm^{-1}, while γ is identified by a peak at 927 cm^{-1}. Example spectra for each polymorph are shown in Figure 4.3

XRPD patterns were obtained by placing 10-50 mg of sample on a 28-well plate supported on a polyimide (Kapton 7.5 μm thickness) film. Data were collected on a Bruker AXS D8-Advance II transmission diffractometer equipped with θ/θ geometry, primary monochromated radiation (Cu Kα 1, $\lambda = 1.54056$ $\AA$), a Vantec 1D position sensitive de-

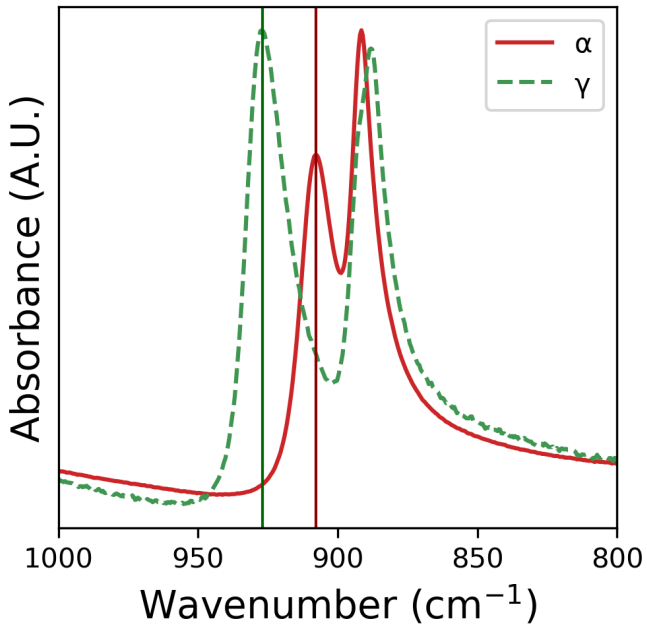

Figure 4.3: Example IR spectras obtained for γ and α glycine. The dark red and green vertical lines highlight the peak positions used to identify α and γ glycine respectively.

tector (PSD) and an automated multi-position x-y sample stage. Data were collected in the range 4-35° 2θ with a 0.015° 2θ stepsize and a 1 s step-1 count time. α glycine is identified by a characteristic peak at 30.04°, while γ is identified by peaks at 21.99° and 25.51°. Example diffraction patterns for each polymorph are shown in Figure 4.4.

4.2.3 Isonicotinamide Ethanol Solution Experiments

Isonicotinamide (INA) solutions were prepared using isonicotinamide powder (Sigma Aldrich, 99%) and ethanol absolute (99.8+%, Fisher Chemical). The oil used to create the oil-solution interface was Krytex GPL 106, a fluorinated synthetic oil.

Samples were prepared similarly to method one described in section 4.2.1 with some minor changes. New, pre-washed and dried 1.5 mL glass vials (VWR 548-0018) were used

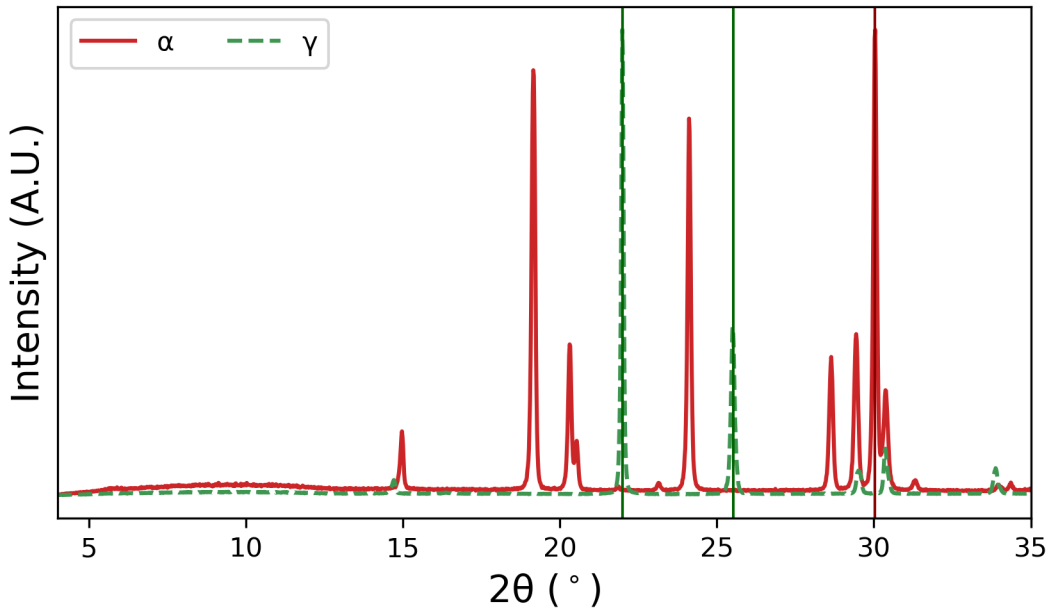

Figure 4.4: Example diffraction patterns obtained for γ and α glycine. The dark red and green vertical lines highlight the peak positions used to identify α and γ glycine respectively.

for each experiment and solutions were prepared directly within the vials. The required amount of INA powder was weighed directly into each vial and 1 mL of ethanol was pipetted into each vial. One experiment investigated the induction times of agitated vials and so a PTFE coated magnetic stirrer bar was added to each vial within this experiment. Samples were prepared at a concentration of 89 mg/ml for the agitated experiments, whilst the unagitated experiments were prepared at 137 mg/ml. The vials were sealed with a screw cap with a silicone, PTFE septum and parafilm to prevent evaporation. Due to the density of the fluorinated oil being greater than the density of ethanol it was not possible to add the oil to the vials prior to dissolution of the powder as was done previously. The vials were then transferred to the Polar Bear and held overnight at 333 K to allow the INA to dissolve.

The vials were then checked visually to ensure that the INA had fully dissolved. For the experiment investigating the effects of the oil-solution interface the fluorinated oil was injected through the septum of the vial lids to create a complete, thin layer at the bottom

of the vial. These vials were then further sealed with parafilm over the septum, to prevent evaporation, and held in the Polar Bear for a further 1 h to ensure that any crystals that may have formed during the oil addition redissolved. The vials were then cooled at a controlled rate of 1.5 K/min to 298 K and transferred to the incubator set at 298 K for temperature control. For the unagitated experiments the vials were placed into vial racks and observed via webcam as described previously. For agitated experiments the vials were placed on a stirring plate (2mag, Magnetic stirrer MIXdrive with control unit MIXcontrol, 15 stirring points), stirred at a controlled rate of 750 rpm and monitored via webcam. Unagitated vials were observed for 5 days, whilst agitated vials were observed for 10 h.

4.3 Results & Discussion

4.3.1 Nucleation Behaviour of Glycine at Air-Solution vs. Oil-Solution Interfaces

A large number of vials of glycine solution were prepared using sample preparation method one, as described in section 4.2, and monitored for 72 hrs. Vials were prepared without an oil interface at concentrations ranging 275-450 g/kg, and vials with the oil interface were prepared at concentrations 275-367 g/kg. Each concentration had a total of 80 vials topped with oil and 40 vials topped with air across a range of experiments. The full set experimental details and the number of vials that crystallised in each case can be found in Table 4.1

After observing the vials for 72 h it was found that probability of nucleation without the oil layer present was very low. At the majority of concentrations investigated no vials nucleated without oil, and at the higher concentrations very few crystals were formed

during the observation period. This agrees with previous work that found that below a concentration of 475 g/kg nucleation was a slow process, with few vials nucleating within a two week period [36]. In contrast to this, glycine was found to readily nucleate at moderate concentrations, 275, 307, 333 and 365 g/kg, where 60, 89, 94 and 93% of the vials nucleated respectively. This was a surprising result as the non-polar, hydrophobic tridecane interface would not be expected to favour the highly polar, hydrophilic glycine. The percentage of vials that nucleated at each concentration for each interface type is shown in Figure 4.5. Thermal history has been shown to impact nucleation kinetics [99], so to ensure that such effects are not impacting the observed differences, the 40 air topped vials were prepared alongside the first set of the oil topped vials for each concentration (experiments B and C, D and E, G and H, J and K). This results in identical thermal history between each pair or experiments, preventing thermal history effects from contributing to changes in the nucleation behaviour.

It was observed that the crystals were typically forming at the oil-solution interface, as demonstrated in Figure 4.6, however it is important to note that the crystals are only visible within the images once they have grown to sizes at the mm scale. It is possible that nucleation is occurring at the oil-glass-solution interface (point A in Figure 4.7) and the resulting crystals move down to the centre of the oil-solution interface (point B in Figure 4.7) prior growing to a size that is visible in the captured images. It is clear, from the stark difference in the percentage of vials nucleating between the oil and air-solution interface experiments, that if the nucleation is occurring at the oil-glass-solution interface the presence of the oil layer is necessary to generate this nucleation.

Cumulative probability distribution functions P(t) of induction times obtained using sample preparation method one are shown in Figure 4.8. In all cases there is similar behaviour, with a significant fraction of vials nucleating within the first few hours, followed by a smaller fraction of vials nucleating within the next few days, leaving a number of

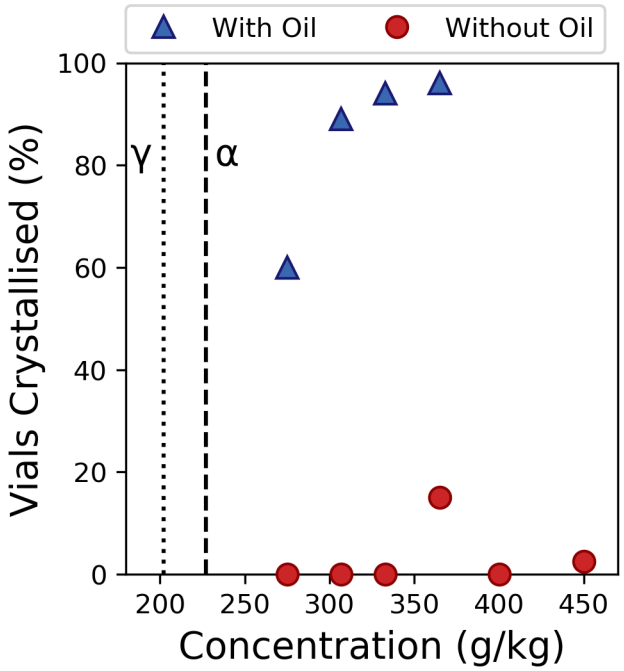

Figure 4.5: Percentage of vials where glycine crystallized within 72 h. Red symbols represent experiments with an air–solution interface and blue symbols represent experiments with an oil–solution interface. Note that experiments with the oil interface were only performed for concentrations below 400 g/kg. The dotted and dashed lines represent γ and α solubilities of 202 and 227 g/kg at 298 K respectively.

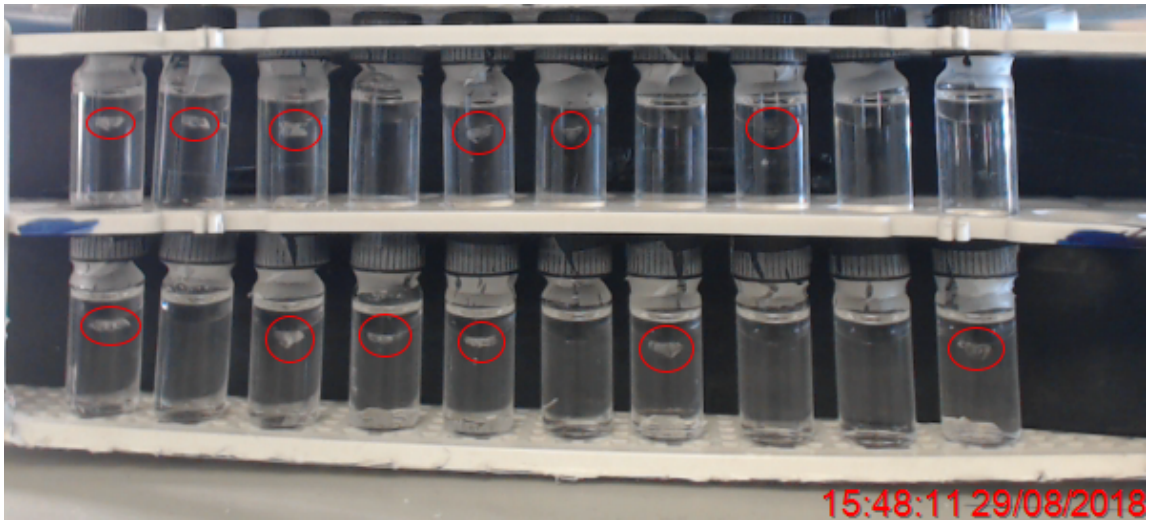

Figure 4.6: Image captured of vials during a nucleation experiment. The red circles highlight where a crystal has formed at the oil-solution interface.

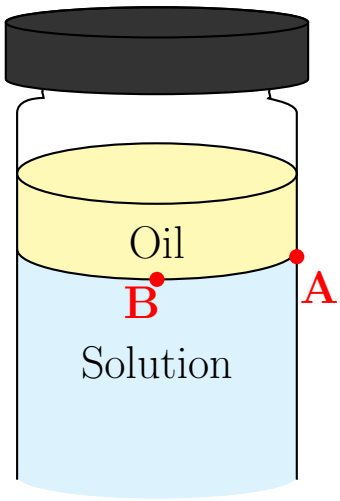

Figure 4.7: Schematic to demonstrate the possible nucleation sites at the oil-solution (A) and oil-glass-solution (B) interfaces.

vials without any visible crystals after 3 days. The nucleation probability increases with concentration as expected however, it can be clearly seen that the nucleation probability does not follow a Poisson distribution time dependence which would correspond to the expected stochastic outcome for a constant nucleation rate. This nucleation probability behaviour is consistent with those seen in previous studies of glycine nucleation under quiescent (non-agitated) conditions [35, 36, 100].

In Figure 4.9 the combined results from three experimental runs performed by Little *et al.* are shown. These results were obtained using a different experimental setup, but similarly using a layer of tridecane on top of glycine solution with the same relative supersaturation (Little: 333 g/kg at 294 K, present work: 365 g/kg at 298 K, relative supersaturation of 1.81 relative to γ glycine [73]). The experiments performed in this work had somewhat higher probabilities of nucleation than those observed by Little *et al*: this work sees 81% of the samples nucleating within the first hour and 95% nucleating within 2 days, compared to 52% and 79% nucleating, within the same time frames respectively, observed by Little *et al.* While each of these experiments were performed at the same supersaturation, there are a number of key differences between the experiments that may

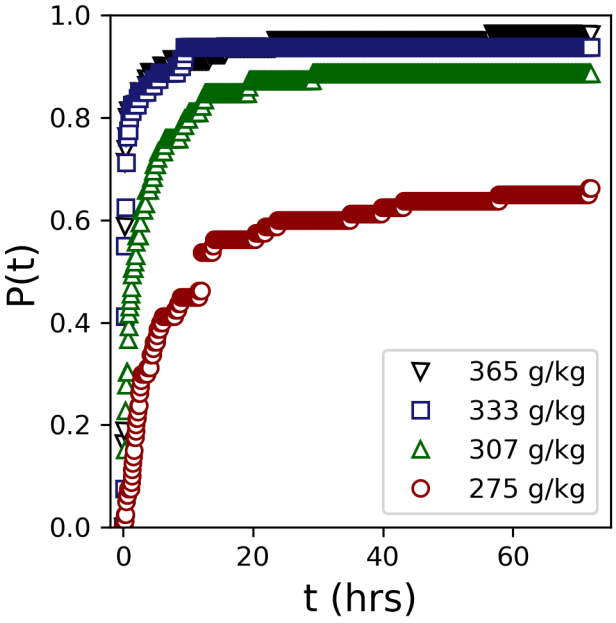

Figure 4.8: Cumulative probability distribution function of induction times of glycine crystallisation from aqueous solution with a tridecane interface at varying concentrations prepared using sample preparation method one.

contribute to the differences observed in the induction times. In addition to the different experimental temperatures, the samples in this work had a greater volume of solution (1 mL vs 100 μL), greater volume of oil (200 μL vs 100 μL) and greater oil-solution interfacial area (64 mm^2 vs 38 mm^2). The samples in this work were held in glass vials, while Little held their samples in treated polystyrene microplate wells. The increased volume of solution in the samples prepared here would result in reduced induction times if the nucleation was bulk-based, while the increased interfacial surface area would decrease the induction time for heterogeneous nucleation at the oil-solution interface. The increased volume of oil would only be expected to impact the induction time if impurities present within the oil influence nucleation.

Little *et al.* suggested that nucleation observed within 60 min of solution addition to the microplate wells may have been due to disturbances introduced by the addition of solution. In the experiments performed here the solutions are prepared directly within the

vials with the oil added before heating so that any crystals formed should be dissolved before the vials are cooled, and as such artefacts like those proposed by Little *et al.* would have been avoided. Despite this there are still large number of vials nucleating within the first 60 minutes in these experiments, similar to those in the experiments of Little *et al.*

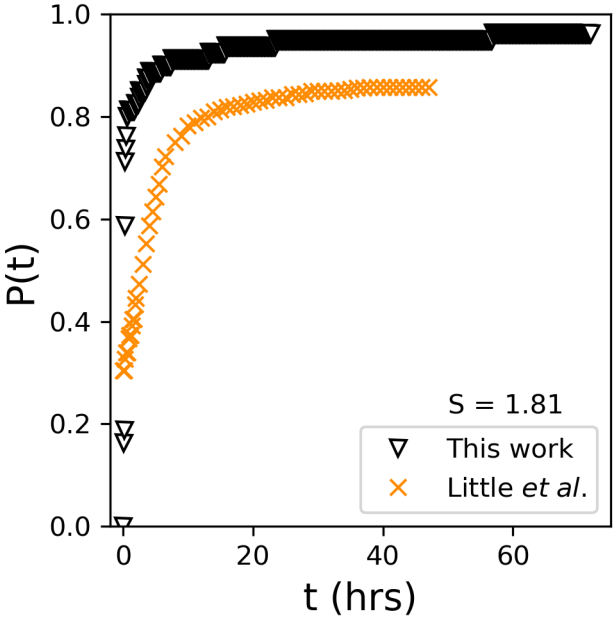

Figure 4.9: Cumulative probability distribution function of induction times of glycine crystallisation from aqueous solution with a tridecane interface at a relative supersaturation of 1.81 with respect to γ-glycine (365 g/kg at 298 K prepared using sample preparation method one and 333 g/kg at 294 K from the work of Little *et al.* [35]

The impact of the sample preparation method on the nucleation behaviour of glycine will now be examined. Experiments have been carried out using two separate sample preparation methods (method one and two) and the cumulative probability distribution functions of induction times obtained using each method at a concentration of 307 g/kg are presented in Figure 4.10. The samples prepared individually within the glass vials (method one) had lower nucleation probabilities than those prepared using a stock solution (method two) and all vials nucleated within a day when using the stock solution method. This highlights the importance of accounting for preparation method and thermal history of samples in nucleation studies in order to obtain accurate quantitative nucleation kinetic

data. A control experiment of 40 vials with an air-solution interface was also carried out at 307 g/kg using method two and, without the oil present, no vials nucleated within the three day observation period. This shows that while the preparation method clearly has an impact on the nucleation behaviour of glycine, the nucleation is strongly accelerated in the presence of the oil-solution interface, regardless of the preparation method.

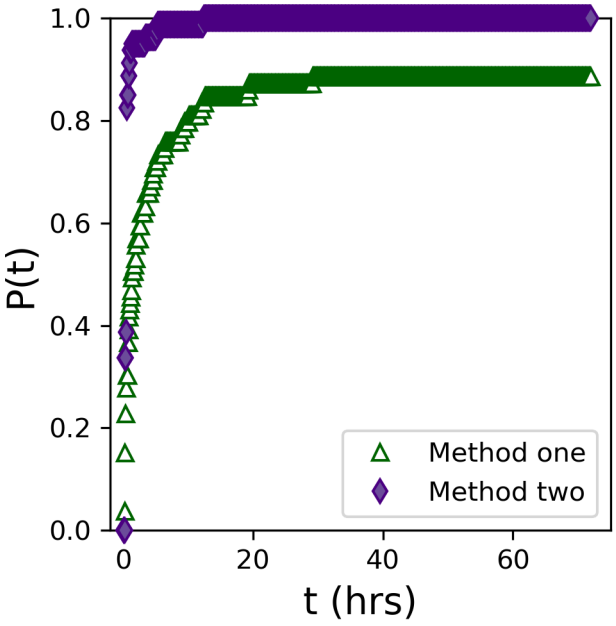

Figure 4.10: Cumulative probability distribution function of induction times of glycine crystallisation from aqueous solution with a tridecane interface at a concentration of 307 g/kg using sample preparation methods one and two

To further confirm that the oil-solution interface is the cause of accelerated nucleation, a range of oil to solution volume ratios have been investigated while keeping the surface area of the oil-solution interface constant. In these experiments the samples have the same total volume of solution and oil combined, and the same oil-solution interfacial area, but with varying oil to solution ratios. The induction times obtained for each ratio are shown in Figure 4.11. It can be seen that distributions of induction times are very similar for oil to solution volume ratios 1:5 (the original ratio), 1:2 and 1:1. In each experiment the volume of glycine solution and the glass-solution interfacial area were different for each ratio. If the overall nucleation rate was proportional to the solution volume, it would be expected

that the nucleation rate would scale with the solution volume in a given vial, and with oil-solution ratio within our experiments. In Figure 4.11 it is clear that the overall nucleation rate does not change significantly. This is what would be expected if the nucleation is controlled by the oil-solution interfacial area, which is constant, rather than the solution volume or glass-solution interfacial area [101]. We note that somewhat longer induction times were recorded at the ratio of oil to solution volume ratio 2:1, which may be due to an onset of concentration depletion as the solution volume becomes increasingly smaller, slowing down crystal growth so that crystal detection takes somewhat longer. Despite this it can still be seen that all of the vials nucleated in less than 10 h. These observations also rule out that nucleation would be due to impurities within the oil as increasing the volume of oil does not lead to faster nucleation.

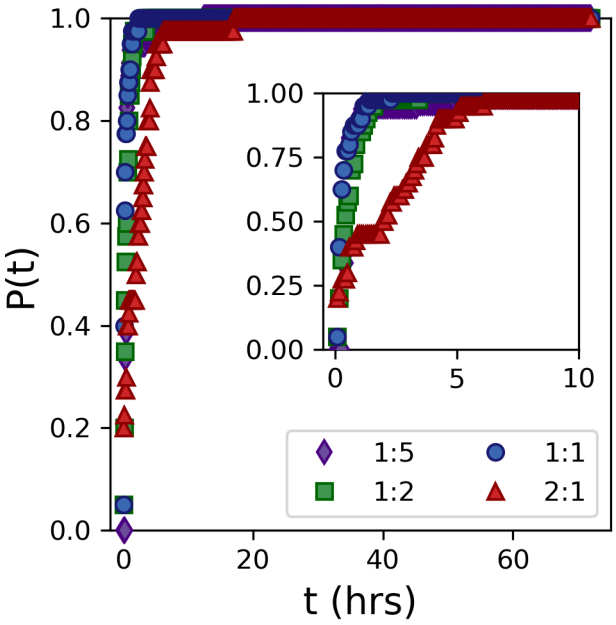

Figure 4.11: Cumulative probability distribution function of induction times of glycine crystallisation from aqueous solution with a tridecane interface at a concentration of 307 g/kg prepared using sample preparation method two at varying oil-solution volume ratios.

It has also been previously reported that PTFE coated magnetic stirrer bars placed in glycine aqueous solution, without agitation, strongly promotes glycine nucleation [36].

Both PTFE and tridecane are hydrophobic and both of these interfaces enhance glycine nucleation. Di Profio *et al.* [102] investigated the effects of polymer surfaces on heterogeneous nucleation from solution and concluded that chemical functionalities of the polymer surface dictate whether the surface promotes enhanced nucleation. Under similar conditions to the lower concentrations investigated here (Di Profio, 180.2 g/kg a t 278 K, present work , 275 g.kg at 298 K, with relative supersaturations of 1.42 and 1.36 respectively, relative to γ glycine), they found that hydrophilic polymers enhanced glycine nucleation from aqueous solutions in contrast to hydrophobic polymers (PP, co-PVDF). However, it is likely that wetting issues due to significant surface roughness may have been paramount in their work as reported by Di Profio.

4.3.2 Quantifying Nucleation Kinetics

As discussed in Chapter 3, it is possible to determine the nucleation rate from the induction times of a series of identical, small-scale experiments using equation 3.8. This method assumes that there is one nucleation rate that remains constant over time, applies to all samples and that all samples will nucleate. However, from the probability distributions obtained here using method one (see Figure 4.8), it is clear that the behaviour observed does not fit the standard Poisson distribution as a large number of samples remain unnucleated. Similar behaviour has been observed previously for the unagitated nucleation of glycine [36, 100], paracetamol [103] and potassium nitrate [38].

Equation 3.8 can be modified in a number of ways to better represent the data obtained in this work. If we assume that there are a number of vials that will fail to nucleate we can use a modified Poisson equation:

$$P(t) = A\left[1 - \exp\left(\frac{-t_d - t}{\tau}\right)\right] \tag{4.1}$$

Where A is the fraction of samples that will nucleate given enough time.

A biexponential distribution assumes that there are two nucleation rates occurring within the samples measured, and has been previously used to quantify nucleation rates of potassium nitrate in microfluidic experiments [38] and laser induced nucleation [100]. This equation may be suitable as these experiments show a large fraction of vials nucleating within the first few hours of the experiment, followed by a period of slower nucleation. The biexponential function is given as:

$$P(t) = A\left[1 - \exp\left(\frac{-t_d - t}{\tau_1}\right)\right] + (1 - A)\left[1 - \exp\left(\frac{-t_d - t}{\tau_2}\right)\right] \tag{4.2}$$

We can see that this equation takes a similar form to the Poisson distribution in equation 3.8, but now there are two nucleation rates (τ_2 is the characteristic time for the second nucleation rate) and a weighting factor, A, which determines which fraction of the samples nucleate under each regime.

Each of these alternative equations assumes there are two populations within the samples, in the case of the biexponential there are fast and slow nucleating samples, whilst the modified Poisson equation assumes that there are nucleating and non-nucleating samples. We can combine each of these to create a third function, a modified biexponential function:

$$P(t) = A\left[1 - \exp\left(\frac{-t_d - t}{\tau_1}\right)\right] + B\left[1 - \exp\left(\frac{-t_d - t}{\tau_2}\right)\right] \tag{4.3}$$

In this form, it is no longer required for the fraction of fast and slow nucleating samples, A and B, to sum to 1, allowing for a portion of non-nucleating samples to remain.

We have fit each of the above equations to the four cumulative distribution functions obtained using sample preparation method one as shown in Figure 4.12. As expected the standard Poisson distribution provides a poor fit to the data. However, each of the

alternative equations provides a good fit, with the modified biexponential providing the best fit to the data. This is expected as this function has the largest number of fitting parameters, and there is a risk that it is simply overfitting the data.

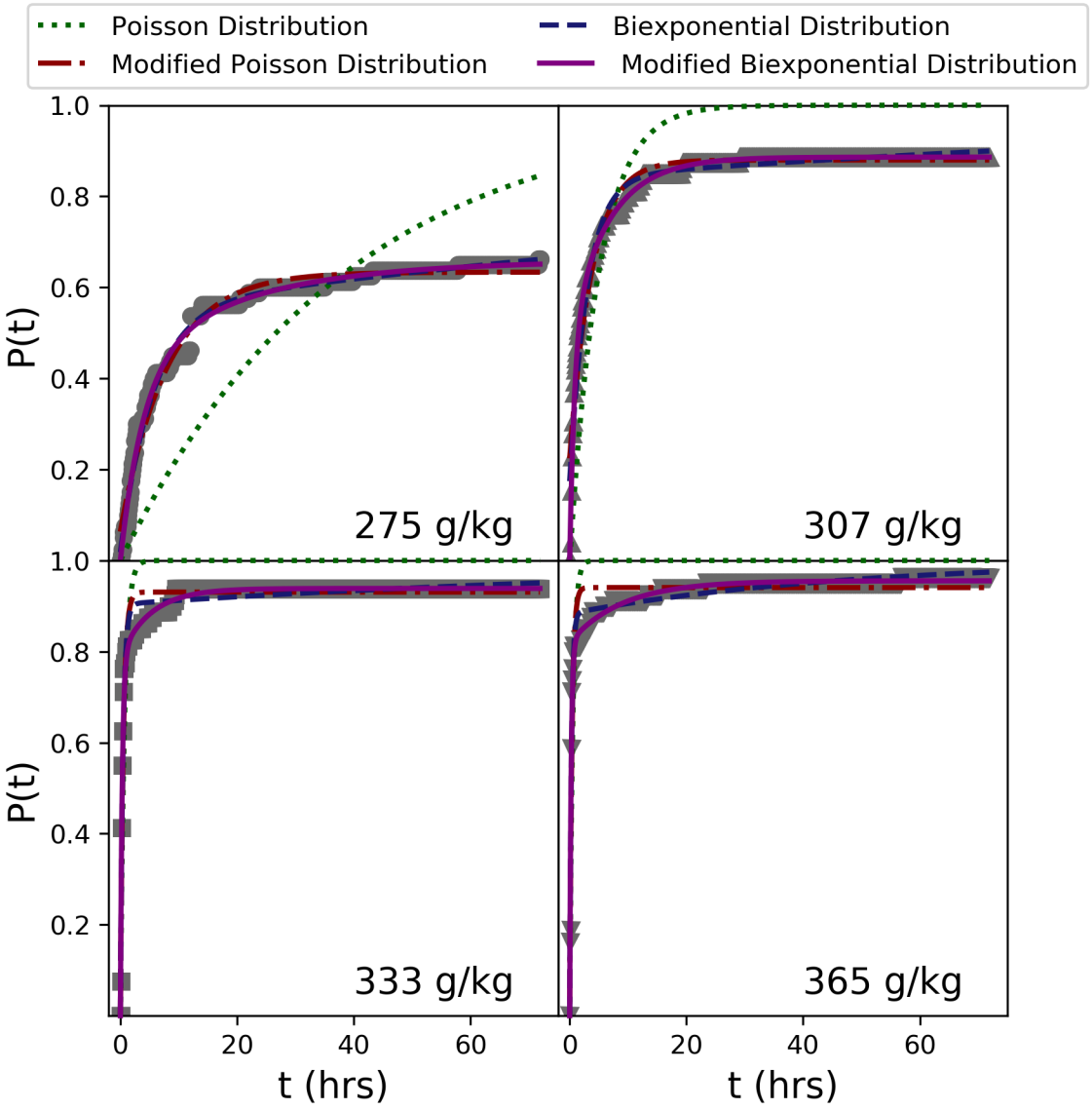

Figure 4.12: Fits obtained using each of the nucleation rate fitting equations to the cumulative probability distribution functions.

It is unclear what is causing the observed deviations from the expected Poisson behaviour within these experiments. Selzer *et al.* [38] made similar observations in their microfluidic experiments of potassium nitrate. They believe that the fast nucleating sam-

ples were due to impurities at, or close to, the interface between their solution and the immiscible oils used to maintain the droplets and that the slower nucleation is also heterogeneous, but of an unknown mechanism that is influenced by the liquid-liquid phase boundary. The oil-solution volume ratio experiments (see Figure 4.11) rule out impurities within the oil as enhancing the nucleation rates, as we would otherwise expect to see an increase in the nucleation rate with increasing oil volume. In contrast, if impurities present within the glycine solution were providing heterogeneous nucleation sites the opposite effect would be seen. We do see a slight decrease in the nucleation rate at the lowest solution volume experiment, however further investigation would be necessary to determine if impurities within the glycine solution are the source of the different nucleation rates observed between samples. Regardless of this, it is clear that if impurities within the glycine solution are contributing to the nucleation rate of glycine, these impurities alone are not enough to trigger nucleation and the presence of the oil-layer is necessary to unlock the nucleation at the concentrations investigated.

Interestingly, this deviation also appears to be sample preparation dependent. All experiments that used sample preparation method two displayed the expected Poisson behaviour (see Figure 4.10). One possible reason is that the contact time between oil and solution is greater in method one than in method two, which suggests that there may be a time dependent effect that acts to hinder nucleation occurring at longer times, for example an impurity present that, given enough time, poisons the interface, thus resulting in non Poisson behavior.

The nucleation rates of glycine have been determined using equations 4.1, 4.2 and 4.3, as shown in Figure 4.15, and errors have been determined via statistical bootstrapping. The cumulative probability distribution function represents data from 80 vials at each concentration. From these 80 vials a random sample of 80 induction times (or lack of induction time, for samples that did not nucleate) were selected with no restrictions in

place to prevent a given value being selected multiple times. This was used to construct a new cumulative probability distribution function that was fit to each of the equations. This process was repeated 500 times allowing for a distribution of the fitting parameters for the generated cumulative probability distribution functions to be obtained. As the nucleation rate appears to be driven by the oil-solution interface the interfacial area has been used as the scaling factor in equation 3.7.

The statistical bootstrapping identified significant statistical uncertainties for the obtained nucleation rates. Upon examination of the distributions of characteristic times, it was clear that these large errors were due to the presence of a small number of extreme outliers within the results obtained from the bootstrapping procedure. This is can be seen in Figure 4.13 (a) which shows the obtained distribution of characteristic times for the slow nucleation regime of the standard biexponential fit for the experiment carried out at 275 g/kg. In order to determine more representative nucleation rates and associated errors, these outliers were identified and removed from the datasets. Outliers were defined as values that were further than 1.5 interquartile ranges below the first quartile or above the third quartile of the obtained characteristic times. The distribution of characteristic times, with the outliers removed, for the slow nucleation at 275 g/kg are shown in Figure 4.13 (b).

Once the outliers are removed, a much more reasonable distribution is obtained, however it is clear that the resulting distribution is not Gaussian. This must be accounted for when determining errors, and symmetrical error bars will not be appropriate. To account for the non-normal distribution, the data has been fit to a gamma distribution and errors have been determined as 68% confidence intervals on the obtained distribution (equivalent to one standard deviation for a normal distribution). Figure 4.14 shows the gamma distribution fit to the filtered distribution of characteristic times for the biexponential slow regime at 275 g/kg experiment.

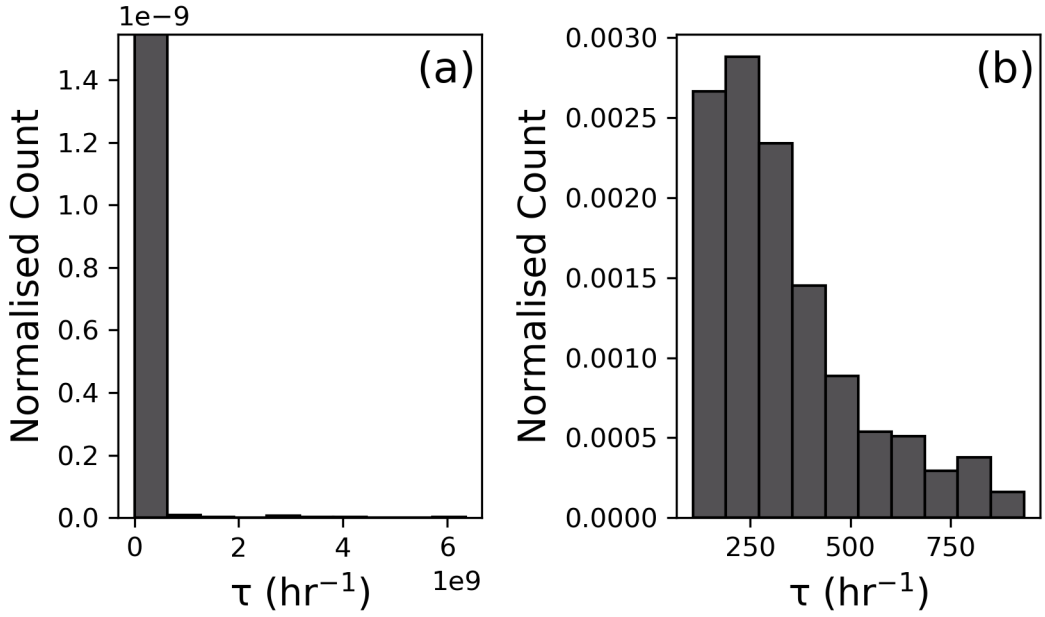

Figure 4.13: The distribution of characteristic times for the biexponential slow regime obtained from statistical bootstrapping for the 275 g/kg experiment. (a) shows the complete distribution of characteristic times, while (b) shows the distribution after outliers have been removed.

The resulting nucleation rates and associated errors are shown in Figures 4.15 and 4.16. For the fast nucleation rates (Figure 4.15 the rates generally increase with supersaturation as is expected. The modified Poisson nucleation rate at 365 g/kg is lower than the rate obtained at 333 g/kg, however the error for this value is significant. There is generally good agreement between the modified Poisson and biexponential nucleation rates, however the modified biexponential tends to give slightly higher nucleation rates at each concentration.

For the slow nucleation regime (Figure 4.16) the modified biexponential once again predicts higher nucleation rates. This is expected, as not all samples are required to nucleate using this function. As the standard biexponential function requires all samples to nucleate, the slower nucleation regime must be lower to account for the plateaus present within the cumulative probability distribution functions. It does appear that the biexponential functions provide more consistent nucleation rates than the modified Poisson distribution,

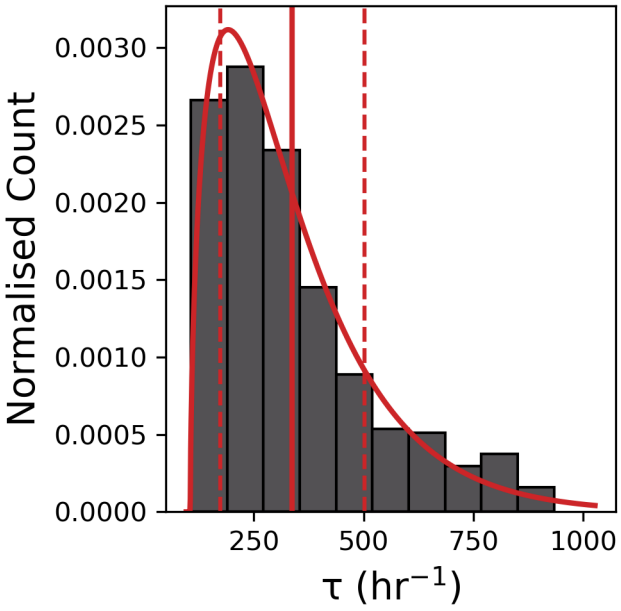

Figure 4.14: The gamma distribution fit to the filtered distribution of characteristic times for the biexponential slow nucleation regime for the 275 g/kg experiment. The solid, vertical red line shows the mean value of the fitted distribution, and the dashed, vertical red lines show the limits of the 68% confidence interval.

however it is unclear at this time why there are two apparent nucleation rates within the samples. It is clear however, that even with a reasonably large number of experiments performed (80 individual experiments at each concentration) there is still significant uncertainty in the obtained nucleation rates, regardless of the fitting used. This highlights the importance of ensuring that enough experiments have been performed in order to accurately determine quantitative nucleation kinetics. The full set of fitted parameters for each equation are provided in Tables 4.3–4.5.

4.3.3 Polymorphic Outcome

From the oil–solution interface systems that nucleated, there is a similar polymorphic distribution regardless of the concentration, as shown in Figure 4.17. 400 samples from the

Table 4.3: Table of parameters obtained from fitting the modified Poisson function to the cumulative probability distribution of induction times. τ values are reported to a larger number of decimal places as small changes in τ result in a significant difference in the final calculated nucleation rate.

Modified Poisson			
Concentration (g/kg)	**A**	$\mathbf{t_d}$ (hr)	$\boldsymbol{\tau}$ (hr^{-1})
275	0.64 ± 0.05	0.9 ± 0.8	7.871 [-1.976, + 1.972]
307	0.87 ± 0.04	0.1 ± 0.6	3.692 [-1.041, +1.039]
333	0.93 ± 0.03	0.1 ± 0.4	0.4978 [-0.0848, + 0.0847]
365	0.94 ± 0.02	1 ± 5	0.601 [-0.304, +0.307]

Table 4.4: Table of parameters obtained from fitting the biexponential function to the cumulative probability distribution of induction times. τ values are reported to a larger number of decimal places as small changes in τ result in a significant difference in the final calculated nucleation rate.

Biexponential				
Concentration (g/kg)	**A**	$\mathbf{t_d}$ (hr)	$\boldsymbol{\tau_1}$ (hr^{-1})	$\boldsymbol{\tau_2}$ (hr^{-1})
275	0.436 ± 0.003	0.21 ± 0.01	5.692 [-1.450, +1.447]	337.4 [-163.6, +163.9]
307	0.175 ± 0.002	0.54 ± 0.02	2.813 [0.866, +0.864]	194.0 [-95.3, +95.2]
333	0.32 ± 0.02	3 ± 1	0.44 [-0.062, +0.061]	113.3 [-69.1, +69.0]
365	0.44 ± 0.02	1.6 ± 0.6	0.28578 [-0.03824, +0.03817]	48.82 [-26.52, + 26.47]

Table 4.5: Table of parameters obtained from fitting the modified biexponential function to the cumulative probability distribution of induction times. τ values are reported to a larger number of decimal places as small changes in τ result in a significant difference in the final calculated nucleation rate.

Modified Biexponential					
Concentration (g/kg)	**A**	**B**	$\mathbf{t_d}$ hr	$\boldsymbol{\tau_1}$ (hr^{-1})	$\boldsymbol{\tau_2}$ (hr^{-1})
275	0.5 ± 0.3	0.5 ± 0.1	0.1 ± 0.2	4.321 [-1.976, +1.972]	167 [-157, +159]
307	0.4 ± 0.1	0.5 ± 0.1	0 ± 0.1	0.851 [-0.222, +0.221]	7.053 [-1.856, +1.856]
333	0.5 ± 0.3	0.5 ± 0.3	0 ± 0.1	0.32484 [-0.02942, +0.02940]	5.2179 [-0.7844, +0.7842]
365	0.4 ± 0.3	0.6 ± 0.3	0 ± 0.02	0.23901 [-0.02294, +0.02292]	8.86 [-3.72, +3.71]

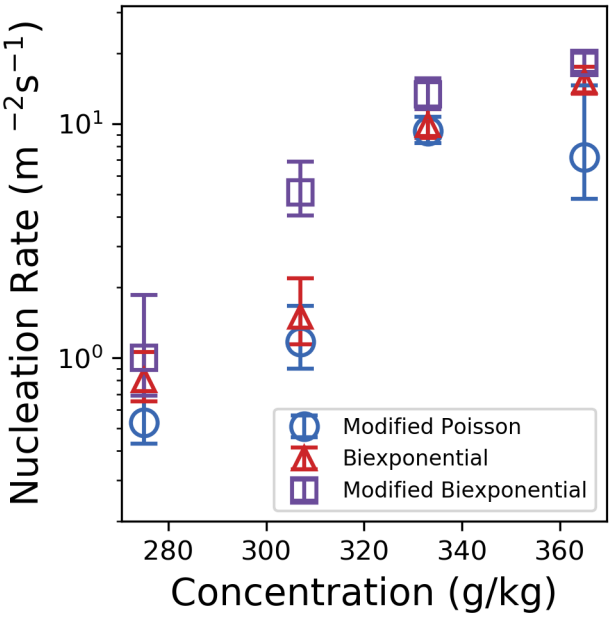

Figure 4.15: Nucleation rates obtained from fitting equations 4.1-4.3 to the cumulative probability distribution function of induction times of glycine crystallisation from aqueous solution with a tridecane interface at varying concentrations prepared using sample preparation method one. Errors are 68% confidence intervals obtained from statistical bootstrapping.

oil–solution interface experiments were analysed using IR spectroscopy and the resulting polymorphic distribution was 95% α, 1% γ and 4% a mixture of the two polymorphs. Of the seven samples that nucleated from the air–solution interface systems five were α, one was γ and one was a mixture of the two polymorphs. As only a small number of vials from the control experiments nucleated it is difficult to determine if the presence of the oil has a significant impact on the polymorphs obtained.

The polymorphs of the crystals obtained from the nucleation experiments have primarily been determined using IR spectroscopy, however to assess the validity of the obtained results, a number of samples have also been analysed using Raman microscopy and XRPD.

XRPD analysis was performed on 42 of the powdered samples that were analysed by

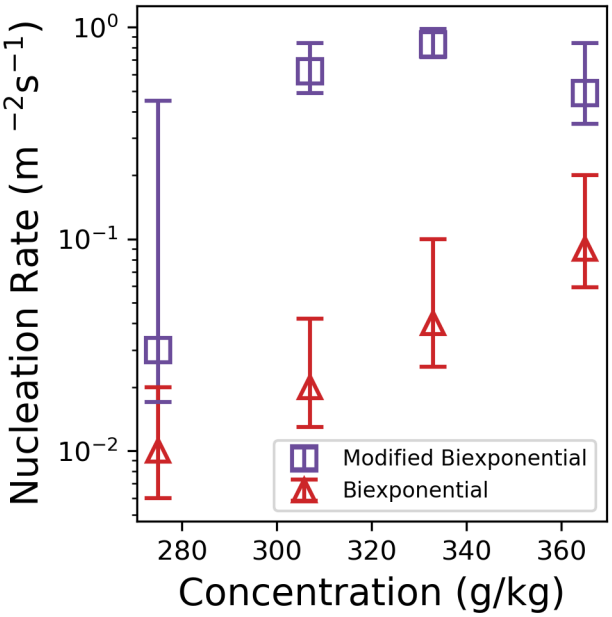

Figure 4.16: Slow nucleation rates obtained from fitting equations 4.2 and 4.3 to the cumulative probability distribution function of induction times of glycine crystallisation from aqueous solution with a tridecane interface at varying concentrations prepared using sample preaparation method one. Errors are 68% confidence intervals obtained from statisticla bootstrapping.

IR spectroscopy. There is good overall agreement between the two methods, however, XRPD did detect mixtures in 6 samples that IR identified as purely one polymorph, and IR detected a mixture in 1 sample that XRPD determined to be one polymorph. This is likely due to sampling issues, as each method would only be expected to detect polymorphs present in the range of 5-10% mass and above. Samples had been incorrectly identified as both pure α and pure γ so this effect is not limited to one polymorph. This suggests that while only a small number of mixtures of polymorphs (5%) were identified by IR spectroscopy in the larger dataset, these may be much more prevalent and missed due to sampling issues.

Raman microscopy can be used to determine the polymorph of the crystals directly from the crystallisation experiment without the need for any sample preparation or grinding.

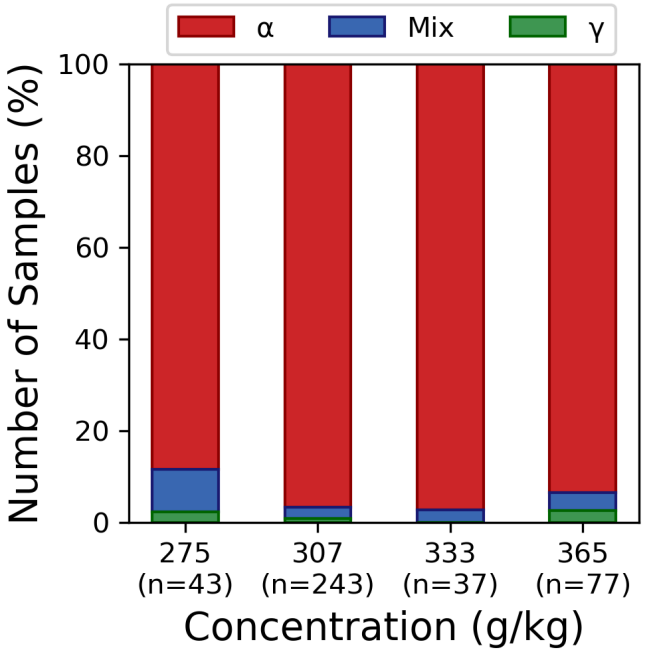

Figure 4.17: Polymorphic distribution of the 400 samples measured using IR spectroscopy for the oil-solution interface nucleation experiments.

However, due to the small spot-size of the Raman microscope (<1 μm) it would be necessary for the polymorph for each individual crystal obtained from a single vial to be measured separately to rule out mixtures of polymorphs forming in one vial confidently, and in cases where two crystals merge to form one particle, it would be necessary to measure each part of the resulting particle. This means that an excessive amount of time would be required to analyse every sample by Raman microscopy, but it does allow for the effects of grinding, which is a required step prior to analysis by IR spectroscopy or XRPD, on the polymorphism to be determined.

A total of 85 samples were analysed using both Raman microscopy prior to grinding and IR spectroscopy after grinding. A fraction of these samples were produced during method development, and as such are not included in the overall polymorph distribution discussed above due to issues with the experimental method. While these issues prevent their use in making meaningful observations of polymorphic outcome or nucleation kinetics

of samples nucleating at the oil-solution interface, they can still be used to provide insight into the reliability of the polymorph determination techniques. The polymorphic distribution obtained from these 85 samples using each of the methods is shown in Figure 4.18. From this data there is a clear shift in the distribution of polymorphs from α to γ. This was a surprising result, and at the time of writing, we have been unable to find any reports of the $\alpha \rightarrow \gamma$ transformation due to grinding in the literature. There are, however, reports of $\beta \rightarrow \alpha$ [43, 104] and $\gamma \rightarrow \alpha$ [105] transformations, and interestingly, one report that α was absolutely stable with respect to grinding [104]. The powdered samples that appeared to have undergone this transformation from $\alpha \rightarrow \gamma$ were re-analysed by Raman microscopy and were confirmed to now to be measured as α suggesting that this was not due to an issue with the Raman measurements.

This certainly raises some important questions with respect to the reliability of the results obtained by each technique. To better understand the extent of this effect, the number of samples that appeared to have transformed are shown in Figure 4.18. One possibility is that a previously unreported $\alpha \rightarrow \gamma$ transformation has occurred in some samples. From Figure 4.18, we can see that while the number of samples that were measured as a mixture of polymorphs by both Raman and IR are the same, a large portion of the samples identified as a mixture by Raman were later measured as pure γ by IR after grinding. These were then replaced by samples measured as pure α by Raman that were later measured as a mixture of the two polymorphs using IR.

These results show that, for this subset of samples, only 23% of the vials measured as pure γ by IR spectroscopy were also measured as pure γ prior to grinding. In addition to this, 23% of the vials that were measured as α prior to grinding were later measured as either γ or a mixture of the two polymorphs. If this is due to a polymorph transformation, it is unclear at this time which properties influence the probability of α transforming to γ during grinding, and the small amount of γ measured in the polymorphic distribution of

the full set of samples that have been analysed by IR spectroscopy suggests that not all samples are equally likely to transform.

Another possibility is that the discrepancies between each measurement technique are due to sampling issues. While multiple Raman spectra were captured from each crystal in an attempt to obtain a representative view of each sample, mixtures of polymorphs may have been missed by these measurements. This issue would be exacerbated if a core-shell type structure had formed, due to the low penetration depth of the Raman measurements. It is therefore possible that a number of the samples that appear to transform from α to a mixture were, in fact, a mixture prior to grinding, and that the fraction of the sample that was γ was missed by the small spot-size of the Raman measurement. The one vial that appears to have transformed from γ to a mixture is also most likely due to a sampling issue.

These results highlight the importance of validating analytical techniques, and demonstrates that commonly accepted methods may influence results.

With these issues in mind, the polymorphic distribution from the larger dataset can now be reflected on. It can stated that 95% of the samples that were measured were primarily α at the point of isolation, as no $\gamma \rightarrow \alpha$ transformations during grinding have been observed. While 1% of samples were determined to be γ, it is possible that a number of these transformed from either α or a mixture of the two polymorphs during the grinding necessary to prepare the sample for analysis. A mixture of polymorphs has been detected in 4% of the samples, however a number of these may have transformed from α into a mixture, or have transformed from a mixture to pure γ during the grinding. It is also possible that samples that have been identified as one polymorph are actually a mixture of the two polymorphs, with one of the polymorphs being missed due to sampling issues.

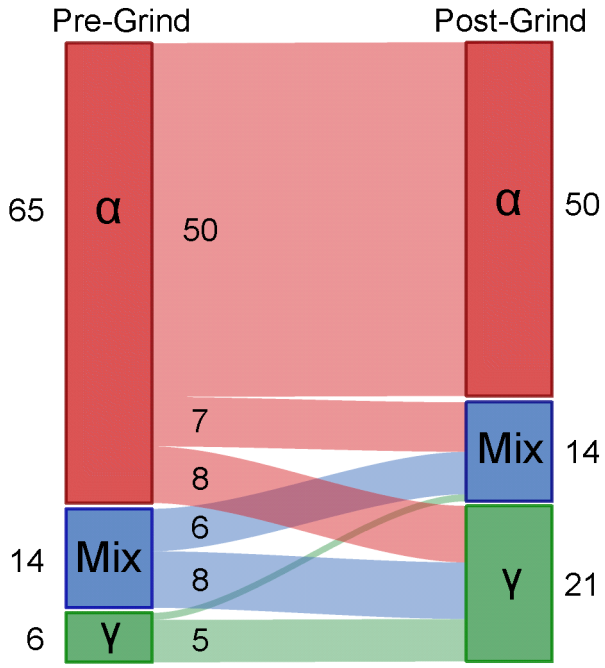

Figure 4.18: Sankey diagram of the polymorphic distribution of obtained glycine crystals before (measured by Raman microscopy) and after (measured by FTIR spectroscopy) grinding. The thin, unlabelled link from γ to mix represents one sample.

4.3.4 Isonicotinamide in Ethanol

While the enhanced nucleation observed for aqueous glycine solution in contact with tridecane is a surprising result, it is important to determine if similar effects will be observed for different systems. Nucleation enhancement at liquid-liquid interfaces has also been observed by Briuglia [39] for isonicotinamide (INA) ethanol solutions in contact with Krytox GPL 106, a fluorinated synthetic oil. In their thesis, the induction times of INA in ethanol was investigated using two different experimental setups: Crystal16 (Technobis), a multi-reactor setup, and in a microfluidic device. When transferred from Crystal16 to microfluidics, the measured nucleation rate increased by several orders of magnitude for a modest increase in supersaturation. The obtained nucleation rates are presented in Figure 4.19 alongside values reported by Maggioni *et al.* [106] and Kulkarni *et al.* [107] which were also obtained in a Crystal16 setup.

There is good agreement with the between the nucleation rates obtained using Crystal16 by Briuglia and Maggioni, with the obtained rates falling within the range of 100-1000 $m^{-3}s^{-1}$, however the rates obtained by Kulkarni do show some deviation at lower supersaturations. In contrast, nucleation is remarkably faster for the microfluidic experiments, with nucleation rates that are orders of magnitude higher than those obtained from agitated vials.

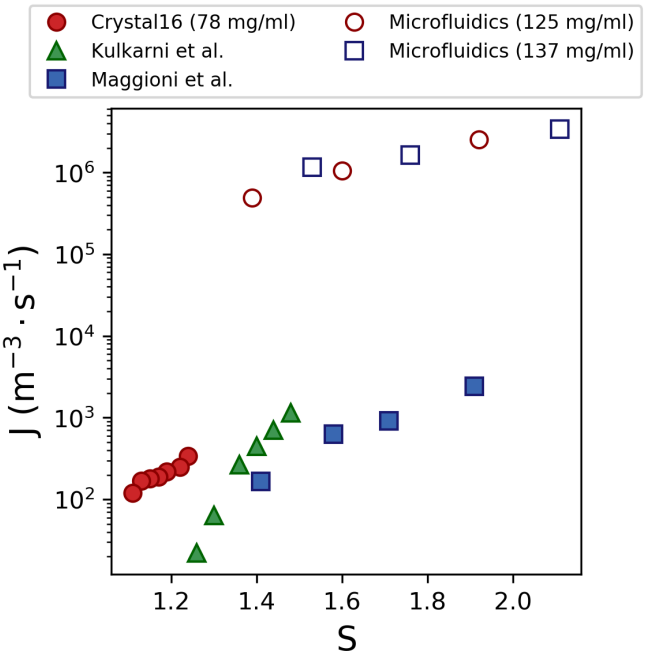

Figure 4.19: The volume based nucleation rates obtained for isonicotinamide in ethanol in Crystal16 setups (filled points) and microfluidics (open points) from various sources for a range of supersaturations.

To understand the clear disparity in the nucleation rates obtained using the two experimental procedures, the differences between the sets of experiments that may contribute to this must be considered. If we consider Briuglia's experiments, the volume of solution varies between the experimental setups with 1 mL of solution in the agitated vials compared to the 65 nL in the microfluidic experiments. The vial experiments have a tightly controlled cooling profile, whilst the microfluidic experiments simply are cooled by the environment. In the Crystal16 vials the solution is agitated, while there is no agitation within

the microdroplets of the microfluidics. In the microfluidic experiments the microdroplets of solution are in contact with the synthetic oil at a high surface area to volume ratio, whilst the agitated vials will have had air-solution and glass-solution interfaces.

Briuglia explored the effects of the cooling profile and demonstrated that a change in the nucleation rate of an order of magnitude can be obtained by changing the cooling rate from 5 K/min to 0.04 K/min. While this is an important result, it cannot fully explain the discrepancy between the Crystal16 and microfluidic results. In this section, the effects of agitation and the oil-solution interface on the nucleation kinetics of INA in ethanol will be explored.

Effect of Agitation

In order to determine the effect of agitation on the nucleation rate of INA, the nucleation of both agitated and unagitated vials has been observed using our experimental method and setup. To ensure that results would be transferable between the Crystal16 and incubator setups, an experiment was carried out with agitated vials within the incubator. Samples were prepared at a concentration of 89 mg/ml which results in a supersaturation that is comparable to those used by Briuglia (S = 1.14 with respect to the solubility reported by Briuglia). The cumulative probability distribution function is shown in Figure 4.20 and the data can be seen to follow a reasonably Poisson behaviour. The nucleation rate was obtained by fitting the standard Poisson model (equation 3.8) and a nucleation rate of 160 [+90, -60] $m^{-3}s^{-1}$ was obtained. This nucleation rate shows excellent agreement to those obtained in Crystal16, confirming the transferability between the Crystal16 and the incubator setup. The fit is presented in Figure 4.20.

The impact of agitation on the nucleation kinetics of INA was assessed by performing

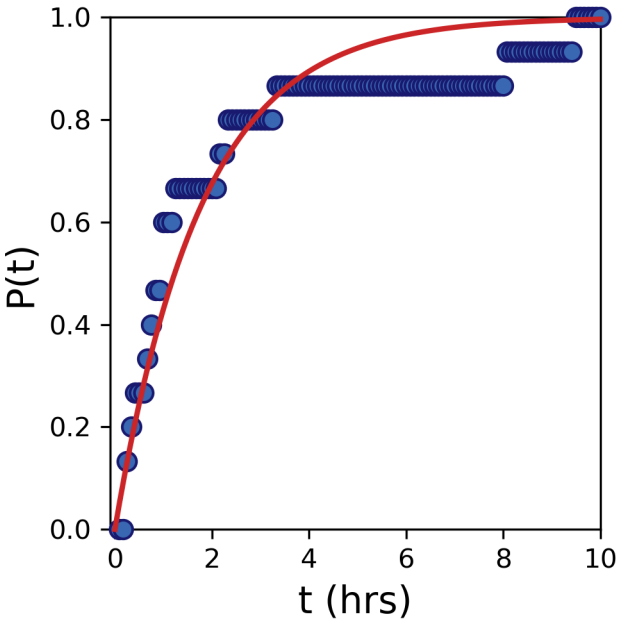

Figure 4.20: Cumulative probability distribution function of induction times of isonicotinamide crystallisation from agitated ethanol solution at a concentration of 89 mg/ml. The red line shows the fit to equation 3.8 used to determine the nucleation rate.

a similar experiment using unagitated vials that were observed for a period of 5 days. As agitation is expected to significantly increase the nucleation rate, and based on the experiences with the very slow nucleation of the unagitated glycine samples, the unagitated samples were prepared at 137 mg/ml in an attempt to ensure that nucleation would occur within a reasonable time frame. Within 5 days, only 30% of the vials had nucleated. In contrast, in the Crystal16 experiments with a similar cooling rate (1.5 K/min here, vs 1 K/min in Crystal16) 100% of the vials had nucleated within 4 hours, even at significantly lower concentrations (78 mg/ml). Due to the small number of samples that nucleated within the unagitated experiment, it is not feasible to quantify the nucleation rate, however it is clear qualitatively that agitation significantly increases the nucleation rate. This highlights the significance of the large increase in nucleation rate when moving from the agitated Crystal16 to the unagitated microfluidic experiment.

Effect of Oil Interface

A similar unagitated experiment was carried out with a layer of the fluorinated oil present at the bottom of each vial in order to determine the effect of the oil-solution interface. As previously observed for glycine, the oil interface was found to have a significant impact on the nucleation rate of glycine. The obtained cumulative probability distribution function, and fit to equation 3.8, are shown in Figure 4.21. While there is insufficient data to quantify the increase in the nucleation rate from the vial without oil to the vial with oil, there is clearly a significant increase in the nucleation rate with the oil interface present, with 100% of the vials in the oil-solution experiments nucleating within 3 days.

As the nucleation appears to be heterogeneous at the oil-solution interface, it is no long appropriate to scale the nucleation rate to the volume of solution. The obtained nucleation rate scaled with the interfacial surface area is 1.1 [+0.6, -0.1] $m^{-2}s^{-1}$

If the nucleation rates obtained by Briuglia in the microfluidics experiments are rescaled with the oil-solution interfacial area a fair comparison between the two experimental setups can be carried out. The microfluidics experiment resulted in a nucleation rate of 134 $m^{-2}s^{-1}$, which is two orders of magnitude larger than that obtained here. However, when compared to the volume scaled nucleation rates(1.6×10^{6} $m^{-3}s^{-1}$ obtained in the microfluidics, 67.7 $m^{-3}s^{-1}$ obtained here), scaling by surface area does appear to be more appropriate.

These results begin to bridge the gap in the observed nucleation rates between the experimental setups, however there is still a significant disparity in the obtained values. While differences in the cooling profiles between the microfluidic and glass vial experiments may contribute to the remaining difference, it is not expected to fully cover the two orders of magnitude difference between the final nucleation rates. These results demonstrate

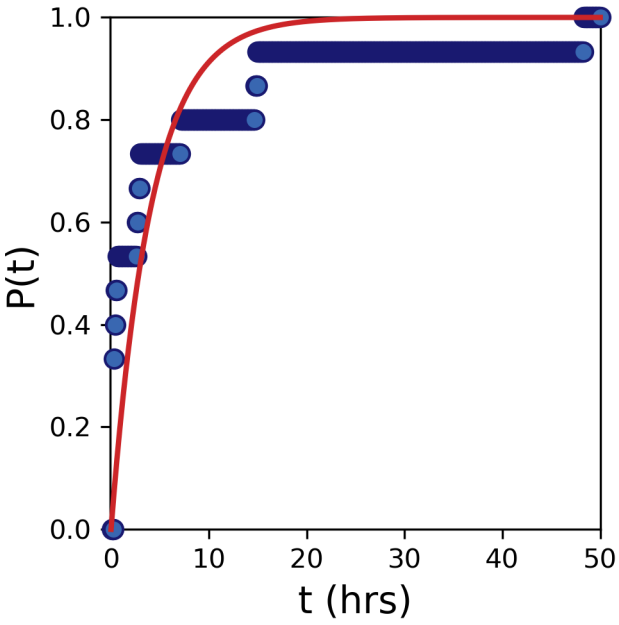

Figure 4.21: Cumulative probability distribution function of induction times of isonicotinamide crystallisation from ethanol solution in contact with the fluorinated oil interface at a concentration of 137 mg/ml. The red line shows the fit to equation 3.8 used to determine the nucleation rate.

that even with careful consideration of multiple experimental parameters, it is extremely challenging to obtain high quality, quantitative nucleation kinetic data that is transferable between experimental setups.

4.4 Conclusions

In this chapter the impact of an oil-solution interface on the nucleation kinetics of glycine has been investigated. It is observed that there is a vast increase in the nucleation rate in the presence of an oil-solution interface, when compared to vials with an air-solution interface. This is surprising as the hydrophobic oil would not be expected to enhance the heterogeneous nucleation of the highly polar, hydrophilic glycine. Current mechanisms widely used in the literature to describe the enhanced rate at which heterogeneous nucle-

ation occurs, such as physical and chemical templating, do not apply to this system due to the absence of specific functional groups and the liquid nature of the interface.

It has been demonstrated that sample preparation technique can have a significant impact on the observed nucleation kinetics however, regardless of these effects the nucleation of glycine is clearly accelerated in the presence of tridecane. The quantitative fitting of nucleation rates to induction times has been discussed, demonstrating deviations from the expected stochastic behaviour. The polymorphic distribution of the oil-solution experiments has been presented and a careful consideration of the analytical techniques used has highlighted potential sources of error.

Finally, the same enhancement of nucleation at oil-solution interfaces has been observed in a completely separate system of isonicotinamide crystallising from ethanol. It is therefore expected that similar effects will be present in a wide range of interface-solution interfaces. This necessitates a more careful consideration of interfaces that are widely present within nucleation experiments, such as oils and polymers, that have previously been considered inert.

Chapter 5

Molecular Simulations of Air–Solution and Oil–Solution Interfaces

This chapter aims to provide molecular-level understanding of the enhanced nucleation of aqueous glycine solution observed at oil–solution interfaces by investigating the solution structure and composition at air–solution and oil–solution interfaces using classical molecular dynamics (MD) simulations. First the solution composition at both oil–solution and air–solution interfaces will be investigated, demonstrating an enhanced concentration of glycine molecules at the oil–solution interface. Following this, the robustness of the concentration enhancement effect will be examined by testing the effects of system size, asymmetric film boundaries and simulating multiple independent starting configurations to increase sampling. The orientation of glycine molecules within the interfacial region will then be examined and an increased ordering of the glycine molecules at each interface, regardless of type, will be shown. Finally, the translational and rotational mobility of

glycine molecules will be compared between the interfacial region and the centre of the solution films, demonstrating that glycine molecules remain highly mobile at the oil–solution interface.

5.1 Introduction

In the previous chapter, a greatly enhanced nucleation rate of aqueous glycine solution was observed at the tridecane–solution interface. This is a surprising result, as the non-polar, hydrophobic tridecane interface would not be expected to favour the nucleation of the highly polar, hydrophilic glycine.

As discussed in Chapter 2, it is well understood that heterogeneous nucleation occurs at a much greater rate than homogeneous nucleation. A number of mechanisms have previously been explored that contribute to this effect, however, if the nature of the oil–solution interface is considered, it is unlikely that a number of these mechanisms will play a role in the nucleation enhancement observed here. Specific interactions between the functional groups of the interface and the solute are unlikely due to the interface being a simple alkane, while neither epitaxy or confinement apply due to the flat, liquid nature of the interface. This raises a question of the reason for the greatly enhanced nucleation rate observed here.

Nucleation is a rare event that happens on the nanometre scale. This small length scale, combined with our inability to predict where, or when, nucleation will occur, makes it particularly challenging to observe using current experimental techniques. Molecular dynamics (MD) simulations are an attractive tool for probing nucleation due to their ability to access information at the molecular level. While the length scales related to nucleation are well suited to MD simulations, there are significant issues related to time scales. As

can be seen from Chapter 4, in some instances days can pass before nucleation will occur from a supersaturated system. This presents a major challenge in the direct simulation of nucleation, as MD simulations are limited by computational power and can typically only access time scales up to microseconds. This issue is only exacerbated further in the case of nucleation from solution, where a number of solvent molecules will be required for each solute molecule within the simulation. Despite the time scale challenge, a number of successful 'brute force' MD simulations of nucleation have been performed. These are typically melt crystallisations, such as the simulations of the nucleation of ice [108, 109], however the successful simulation of NaCl [110] and KCl [111] nucleation from aqueous solution using brute force MD are reported within the literature.

An alternative approach to circumvent the time scale problem is the use of advanced simulation techniques that allow for better sampling of configurational space. These methods use an external bias on the system that guides it into unexplored areas of the free energy landscape. This allows the simulation to more readily overcome the energy barrier for nucleation to allow this process to be investigated at significantly reduced simulation times. One key example of these advanced sampling methods is metadynamics. During metadynamics simulations, one or more order parameters or collective variables are defined within the system. A history-dependent bias is then applied to the system, pushing it away from previously visited configurations [112]. Metadynamics simulations have been used successfully to investigate the nucleation of a range of systems including ice [113,114], biominerals [115] and small organic molecules [116, 117].

As the system investigated here involves small organic molecules in solution, achieving nucleation through a brute force approach is unfeasible. However, the use of advanced sampling techniques require the selection of collective variables to describe the system which is a non-trivial task. The chosen order parameters must be able to identify each of the relevant states within the system [118]. Ideally, multiple collective variables are

used to ensure that the pathways available are not overly constrained [112], however the computational cost of the simulation increases with each variable that is included. As such, this work will take a brute force MD approach, with the intention of investigating the solution behaviour at the air–solution and oil–solution interfaces prior to nucleation. The solution composition, molecular orientation and dynamics will be examined to gain insight into how differences in the solution behaviour at each interface leads to the drastically different nucleation behaviour observed experimentally.

5.2 Methodology

To investigate the differences in solution behaviour at oil–solution and air–solution interfaces, MD simulations have been performed of three different interfacial systems: one with tridecane–glycine solution interfaces; one with vacuum–glycine solution interfaces, which represents the air–solution interface in the control experiments; and an asymmetric film system with a tridecane–solution interface on one side and a vacuum–solution interface on the other. In all three cases, simulations were performed at 250 and 307 g/kg. Glycine solutions contained 240 glycine molecules and 4000 water molecules for 250 g/kg and 295 glycine and 4000 water molecules for 307 g/kg. A smaller solution film at 307 g/kg (147 glycine molecules, 2000 water molecules) was also simulated in an oil–solution interface setup in order to determine the effects of system size. Snapshots of each large system at 307 g/kg are shown in Figure 5.1.

The oil–solution interface system was constructed by placing two pre-equilibrated layers in contact with a box of glycine solution. The tridecane layers were prepared by simulating 128 tridecane molecules in the NVT ensemble for 1 ns, followed by 1.2 ns in the NPT ensemble. The simulation box was then modified to the desired cross section (3.45 nm$\times$3.45 nm in the xy directions), energy minimised with an energy tolerance (relative change in en-

ergy) of 1.0×10^{-6} and a force tolerance of 1.0^{-6} kcal/mol^{-1}Å^{-1}, and then simulated for a further 1 ns in NPT. A layer of tridecane was placed above and below the glycine solution box in the z direction and a further energy minimisation was performed. This combined system was then simulated for 0.2 ns in the NVT ensemble followed by 3.8 ns in NPT to equilibrate the system. A production run of 200 ns of NPT simulation was then performed.

The air–solution system used the same starting configurations of the glycine solution as the oil–solution system without the tridecane molecules added to either side. The total length of the box in the z direction was 30 nm. An energy minimisation was performed followed by 4 ns of NVT dynamics for equilibration. This was followed by a 200 ns NVT production run.

The mixed-interface system was prepared by placing two pre-equilibrated tridecane layers, to increase the oil thickness, below the glycine solutions and vacuum above the glycine solution and below the tridecane. An energy minimisation was performed followed by 4 ns of NVT dynamics for equilibration. This was followed by a 200 ns NVT production run.

Simulations of the oil–solution interface were performed in the NPT ensemble at a temperature of 298 K and pressure of 1 atm. The x and y dimensions were maintained at 3.45 nm, allowing the box to vary only in the z-direction to maintain the pressure. The air–solution and mixed interface simulations were performed in the NVT ensemble at a temperature of 298 K with the same x and y dimensions of the oil–solution simulation. Thermodynamic properties were sampled every 200 fs, while structural information was sampled every 20 ps.

Structural property profiles perpendicular to the oil–solution and air–solution interfaces have been calculated. The simulation box was separated into bins and for each snapshot

the position of each atom or molecule, as appropriate, was placed into the bin associated with its fractional position of the box for that timestep. To account for drifting of the interfaces in the z direction, the z component of the centre-of-mass (COM) of the solution phase was calculated at each timestep and distances were calculated from this point.

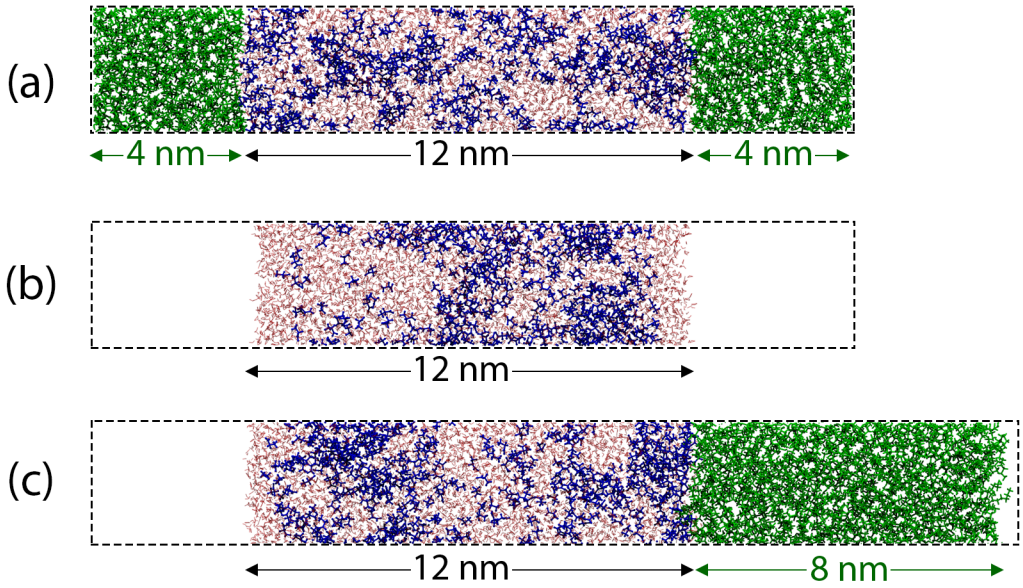

Figure 5.1: Snapshots of the simulated (a) oil-solution interface, (b) air-solution interface and (c) asymmetric film systems at a concentration of 307 g/kg. Glycine, water and tridecane molecules are coloured blue, red and green respectively. The dashed black box represents the boundaries of the simulation box.

5.3 Results & Discussion

The local solution composition has been investigated at both the air–solution and oil–solution interfaces. The density profile of each component within the simulated interface systems in the z direction perpendicular to the interface are shown in Figure 5.2.

It can be seen that there is a strong enhancement in the glycine density occurring in an interfacial region near the oil–solution interface, whilst at the air–solution interface there

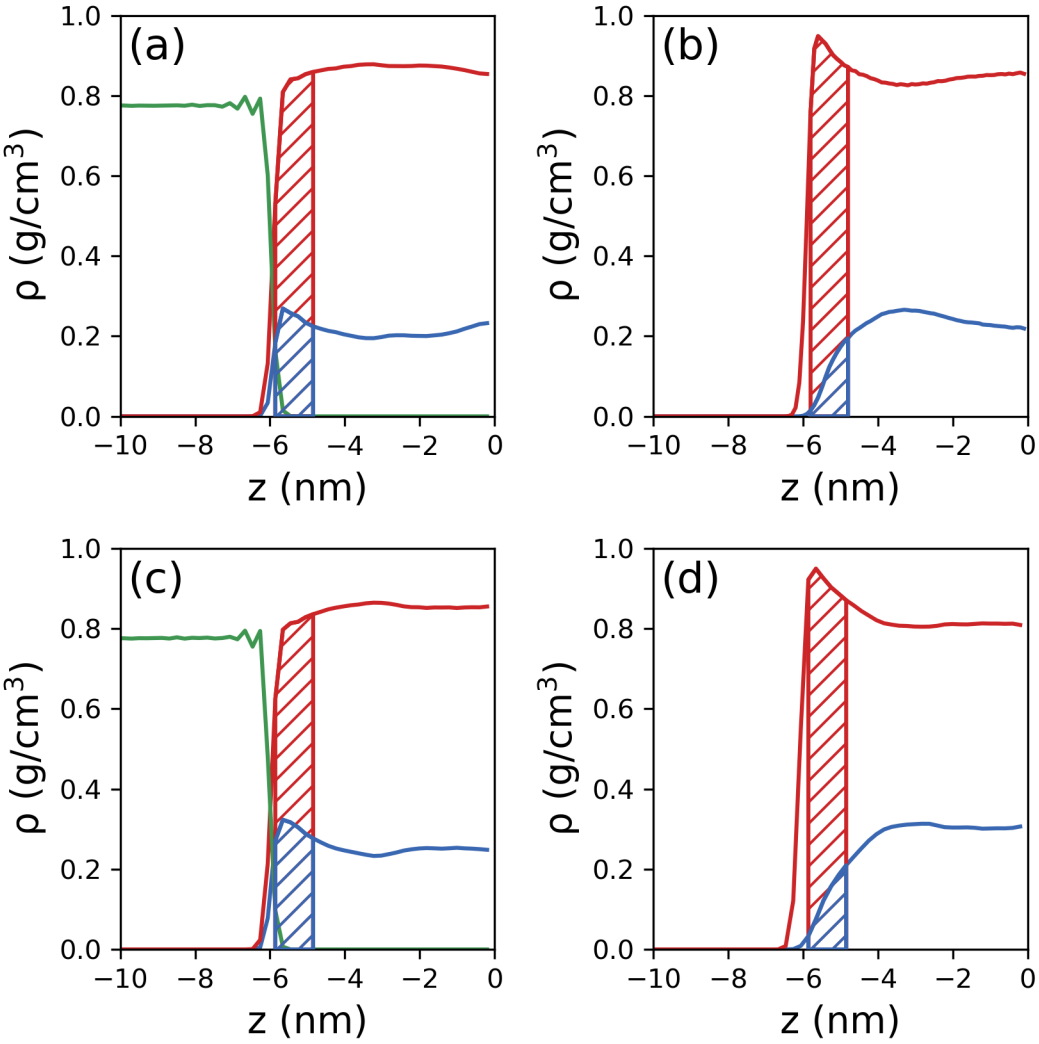

Figure 5.2: Density profiles of glycine (blue line), water (red line) and tridecane (green line) of the simulated oil–solution at (a) 250 g/kg and (c) 307 g/kg and air–solution interfaces at (b) 250 g/kg and (d) 307 g/kg in the z-direction (perpendicular to the interface). The patterned areas show the 1 nm interfacial regions. The centre-of-mass of the glycine solution is set to $z = 0$ and the data has been symmetrised over both interfaces.

is a strong depletion in the glycine density. Interfacial regions with thickness of 1 nm have been indicated to highlight that glycine density enhancement extends over length scales comparable to the expected magnitude of crystal nuclei dimensions. Interfacial regions for the oil–solution system were defined by taking a 1 nm region starting from the point of highest glycine concentration (the ratio of the density of glycine to water) at the interface

towards the centre of the glycine solution as demonstrated in Figure 5.3. The interfacial regions for the air–solution interface were defined as having the same position as the oil–solution interface system of the same concentration.

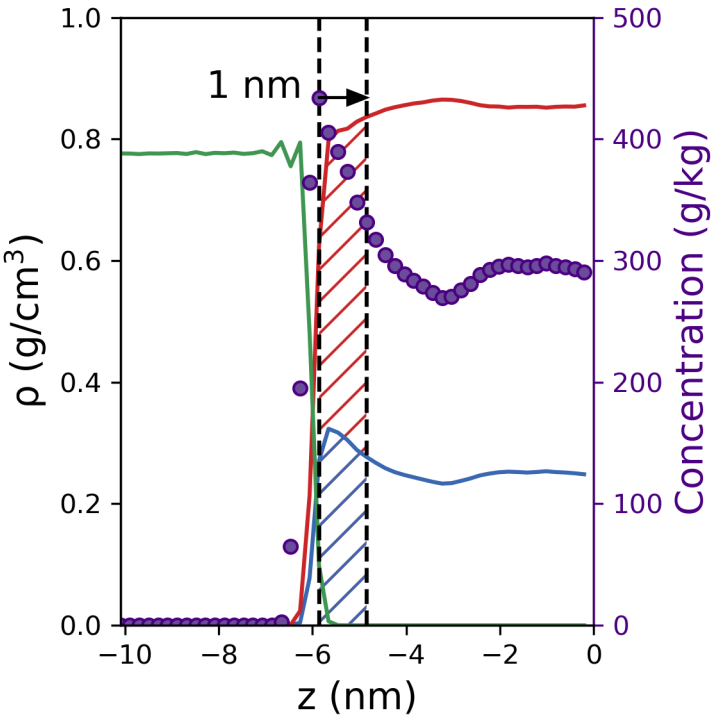

Figure 5.3: Density profiles of the oil-solution interface at 307 g/kg. The concentration is plotted as the purple circles on the secondary axis. The dashed black lines and arrow demonstrate how the 1 nm interfacial region was defined. The centre-of-mass of the glycine solution is set to $z = 0$.

The glycine concentrations of the interfacial regions are given in Table 5.1. The interfacial concentrations at the oil–solution interface are 1.21 and 1.26 times the overall system concentrations of 250 and 307 g/kg respectively. In contrast the interfacial concentrations of the air–solution interface are 0.54 and 0.51 times the overall system concentrations.

Within the nucleation experiments of Chapter 4, it was observed that nucleation was slow for overall concentrations below 400 g/kg in the absence of oil. The calculated glycine interfacial concentration of 386 g/kg at the oil–solution interface is therefore lower than

Table 5.1: Total and interfacial concentrations of glycine at the oil and air interfaces, with the concentration ratio showing enhancement and depletion at the oil and air interfaces, respectively.

Interface	Total Concentration (g/kg)	Interfacial Concentration (g/kg)	Concentration Ratio
Oil	250	303	1.26
Air	250	136	0.54
Oil	307	386	1.21
Air	307	156	0.51

the concentrations that would be expected to result in significant nucleation. Despite this, within the oil–solution interface experiments, a majority of samples nucleated within 3 days at an overall concentration of 307 g/kg. A number of factors may contribute to this disparity. It is difficult to compare the simulation results to those obtained experimentally, as these will depend on the accuracy of the force field used. While care has been taken to select an appropriate force field, certain properties of this model, such as the solubility, are unknown and may be significantly different from their experimental values. It is likely that in the samples without the oil layer present nucleation is occurring at the glass surface of the vial [47]. The interfacial concentration at the glass surface is unknown, and, as is seen from the results for the oil and air interfaces, it cannot be directly inferred from the bulk concentration.

System size may affect the density profiles obtained, as the finite size may act to reduce the enhanced interfacial concentration at the oil interface, due to depletion of glycine in the centre of the thin film. Conversely, the concentration at the air interface may be higher due to the finite size of the film. To determine how the system size influences the results obtained, a glycine solution film was prepared at a concentration of 307 g/kg at half the size of that was used in the full sized simulations. An oil–solution interface system was constructed using this reduced glycine solution film using the same method described above and a 200 ns NPT production run was carried out. As expected, the interfacial concentration enhancement was lower for the smaller solution film, with an interfacial

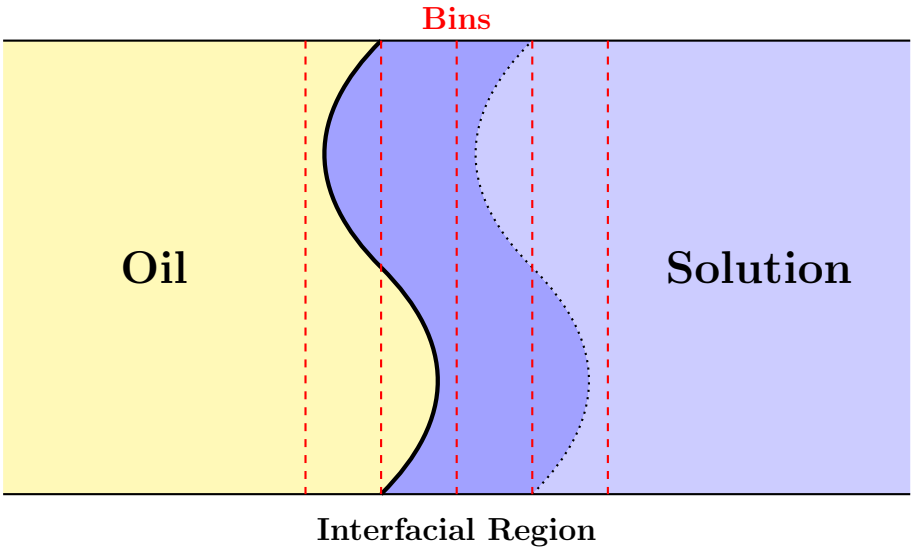

Figure 5.4: Schematic to demonstrate how capillary waves and interface fluctuations can influence interfacial densities. The dark blue region represents the high concentration region at the true undulating surface that can be split into multiple bins.

concentration of 344 g/kg (corresponding to a concentration ratio of 1.12). In vial-based experiments there will be an effectively infinite reservoir of glycine solution, when compared to the size of the interfacial region, and, therefore, it can be expected that the interfacial concentration effects would be even more significant under typical experimental conditions.

The obtained density profiles will also be affected by interfacial fluctuations. Density profiles are obtained by taking the average density of 2D slices along the z-axis. This does not account for capillary waves and interfacial fluctuations, and as such, the density of each component at the true undulating interface is smeared across multiple bins as shown in Figure 5.4. This may result in the interfacial concentration being higher than is obtained using the current 2D slices; however, to account for these effects, it would be necessary to employ more advanced interfacial analysis techniques, such as the generalised identification of truly interfacial molecules algorithm for non-planar interfaces [119].

In order to ensure robustness of the observed interfacial concentration effect, two additional tests have been performed: simulations using asymmetric films boundaries, and

increased sampling to test concentration fluctuations.

First the case of an asymmetric film will be considered. As air and oil interfaces have opposite effects on the interfacial glycine concentration, a third, asymmetric system was constructed containing an oil–solution interface on one side and an air–solution interface on the other. It was expected that the depletion of glycine concentration in the air–solution interfacial region would at least partially cancel the depletion of glycine in the central region caused by the enhancement at the oil–solution interface. The density profiles obtained are shown in Figure 5.5. The same effects can be observed at each of the interfaces as seen previously. The double-interface simulation results in an interfacial concentration of 1.33 times the overall concentration at the oil interface, and an interfacial concentration of 0.54 times the overall concentration for the air interface, for the overall concentration of 307 g/kg. Similar results were obtained for a glycine concentration of 250 g/kg, with a concentration ratio of 1.06 and 0.44 at the oil and air interfaces, respectively.

There are concentration fluctuations in the centre of the glycine solution films, as seen in Figure 5.2, that are still present after relatively long simulation times (200 ns). Due to these long lasting fluctuations, efforts were undertaken to improve sampling. The minimised, pre-equilibration starting configuration for the 250 g/kg solution had an independent set of velocities applied to the atoms. This was followed by the same equilibration procedure described in section 5.2, and a 10 ns production run. This process was repeated 10 times to provide a combined total of 100 ns simulation time from independent starting configurations. The density profiles obtained for each short run along with the average profile of all 10 runs are shown in Figure 5.6. Interfacial concentration enhancement is consistently present in all cases, giving an average interfacial concentration of 265 g/kg, corresponding to a concentration ratio of 1.06. This demonstrates that the enhanced concentration at the oil–solution interface is not due to fluctuations, which can be observed in the centre of the film.

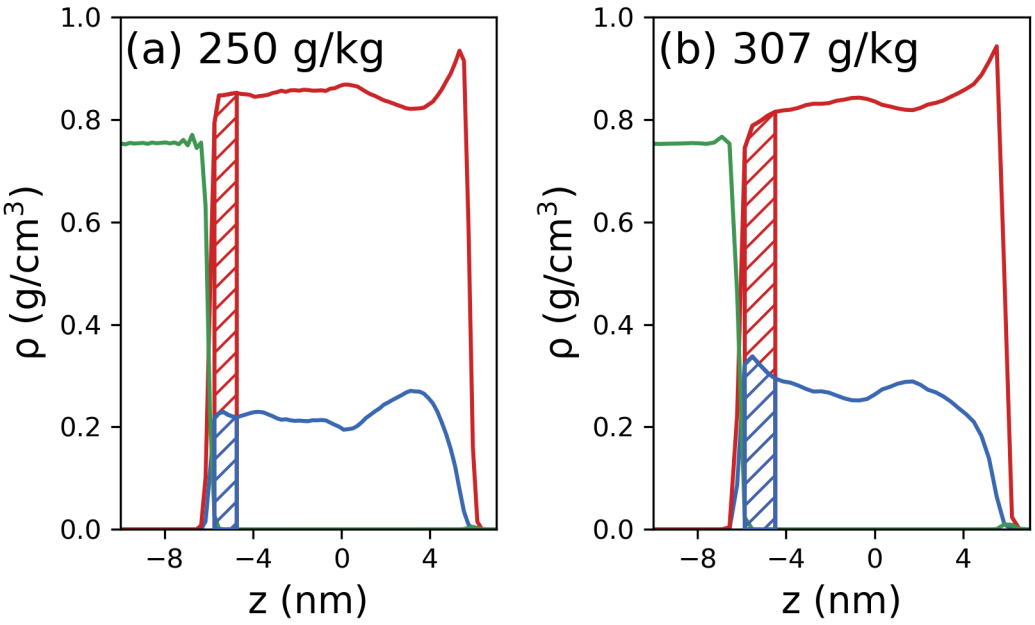

Figure 5.5: Density profiles of glycine (blue), water (red) and tridecane (green) of the mixed interface system at (a) 250 g/kg and (b) 307 g/kg in the z-direction (perpendicular to the interface). The patterned areas show the 1 nm interfacial region at the oil–solution interface. The centre-of-mass of the glycine solution is set to $z = 0$.

The contrasting effects of the oil and air interfaces on the interfacial concentration highlight that heterogeneous nucleation rates can be expected to vary significantly among different interfaces, even in the absence of specific interactions. It is important to note that this is in the absence of templating [14,15], physical confinement [2,17,18], or specific chemical interactions [16], which have been customarily implicated in heterogeneous nucleation mechanisms. Depletion of glycine near the air–solution interface is consistent with surface tension measurements of aqueous glycine solutions [120]. The increase in glycine concentration at the oil–solution interface is not a surfactant effect as glycine is zwitterionic and not amphiphilic. However, it is known that large, polarisable ions have an affinity for a water–oil interface due to cavitation and dispersion forces, whereas smaller ions remain hydrated [121]. By analogy with this effect, it is proposed that the enhanced interfacial concentration is likely due to non-specific van der Waals interactions between the interface material and the solute or solvent molecules. Specifically, for glycine aqueous solution, the

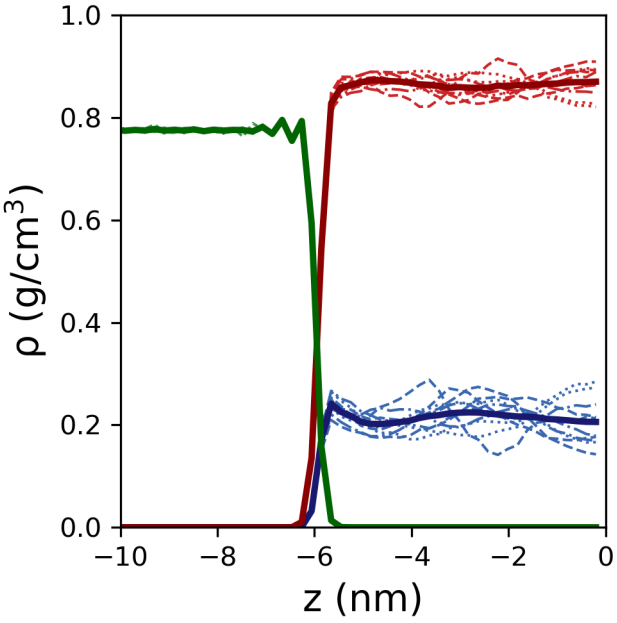

Figure 5.6: The density profiles for glycine (blue), water (red) and tridecane (green) obtained from 10 individual 10 ns runs (thin dashed and dotted lines). The average density profiles of the 10 runs are shown as thick solid lines. The centre-of-mass of the glycine solution is set to $z = 0$.

van der Waals interaction between glycine and tridecane is significantly stronger than the van der Waals interactions between water and tridecane, leading to an enhanced glycine interfacial concentration at the tridecane–solution interface. The same argument can also explain the enhanced nucleation of glycine that was previously seen at the liquid–solid interface of a PTFE stirrer bar [36].

5.3.1 Orientation

In addition to the concentration enhancement effect that is observed, orientation of the glycine molecules at the interface may contribute to the increased heterogeneous nucleation rate [17]. To investigate this effect, the orientation of the glycine molecules at the interface

has been examined using the bond orientation parameter P_2:

$$P_2 = \frac{3}{2}\langle cos^2\theta \rangle - \frac{1}{2} \tag{5.1}$$

where θ is the angle between the z-axis and the C-C bond vector. A P_2 value of 1.0 corresponds to the C-C bond being oriented perpendicular to the interface, while a value of -0.5 indicates the bond lies parallel to the interface. A P_2 value of 0 corresponds to random bond orientations. These orientations and their corresponding P_2 values are demonstrated in Figure 5.7.

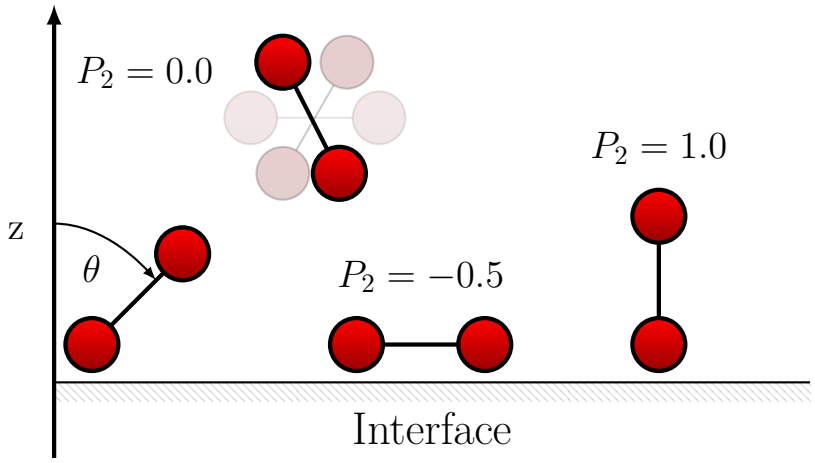

Figure 5.7: Schematic to demonstrate the bond orientation parameter given by equation 5.1.

The variation of the bond order with distance from the interface, taking the position of the molecule as its COM, has been investigated. The bond orientation profiles for the air and oil interfaces can be seen in Figure 5.8. For both the air and oil interfaces, it can be seen that the glycine molecule C-C bonds are mainly oriented parallel to the interface, with the orientation becoming random towards the centre of the solution. The orientation profiles at the air and oil interfaces are very similar, indicating that the orientation is mainly a steric or packing effect, rather than due to a specific interaction. While P_2 generally decreases within the interfacial region, there is a point in each graph where the value increases to above 0. These points occur in each graph where the average glycine density is less than

2×10^{-3} g/cm^3 and so they are not statistically relevant.

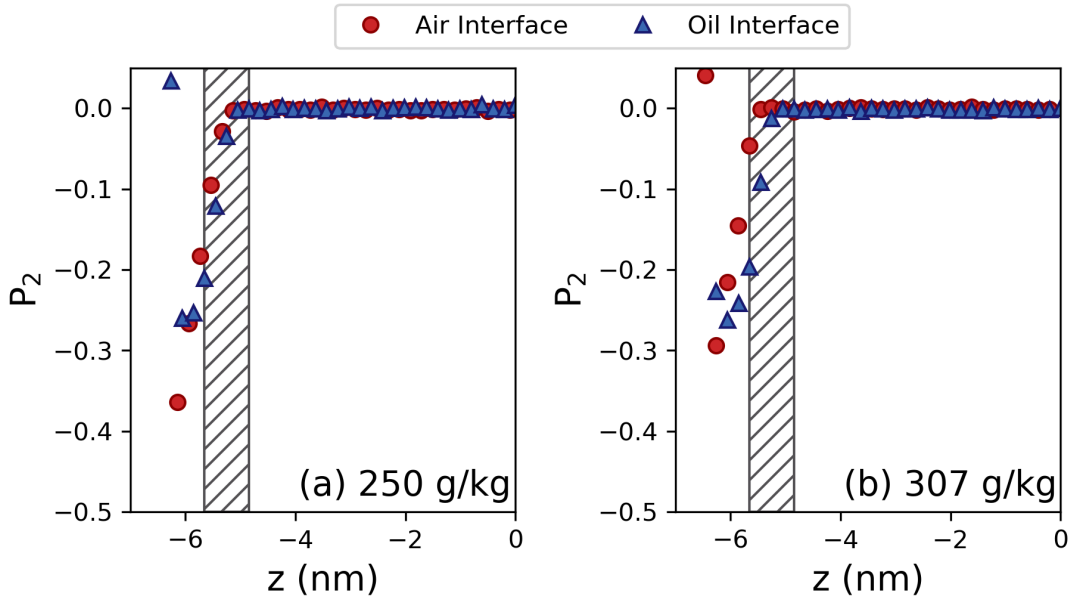

Figure 5.8: Bond orientation profile of the C-C bond vector of glycine at (a) 250 g/kg and (b) 307 g/kg for an air interface and oil interface. The patterned areas show the 1 nm thick interfacial regions.

5.3.2 Dynamics

The local translational and rotational mobility of the glycine molecules was also investigated in order to determine the mobility of the glycine molecules at the oil–solution interface. Translational mobility is often determined from MD simulations using the mean squared displacement (MSD), however as we are interested in a 1 nm thick region, the MSD in the z direction is not a useful metric for examining the mobility of the molecules perpendicular to the interface. Instead, the residence time of molecules remaining within the interfacial region has been compared to the residence time of molecules within an equally sized, 1 nm region within the centre of the film. The obtained residence time distributions are shown in Figure 5.9. From the results obtained it can be seen that for both concentrations investigated that there is a slight reduction in the mobility perpendicular to the

interface within the interfacial region, but that the molecules are still highly mobile.

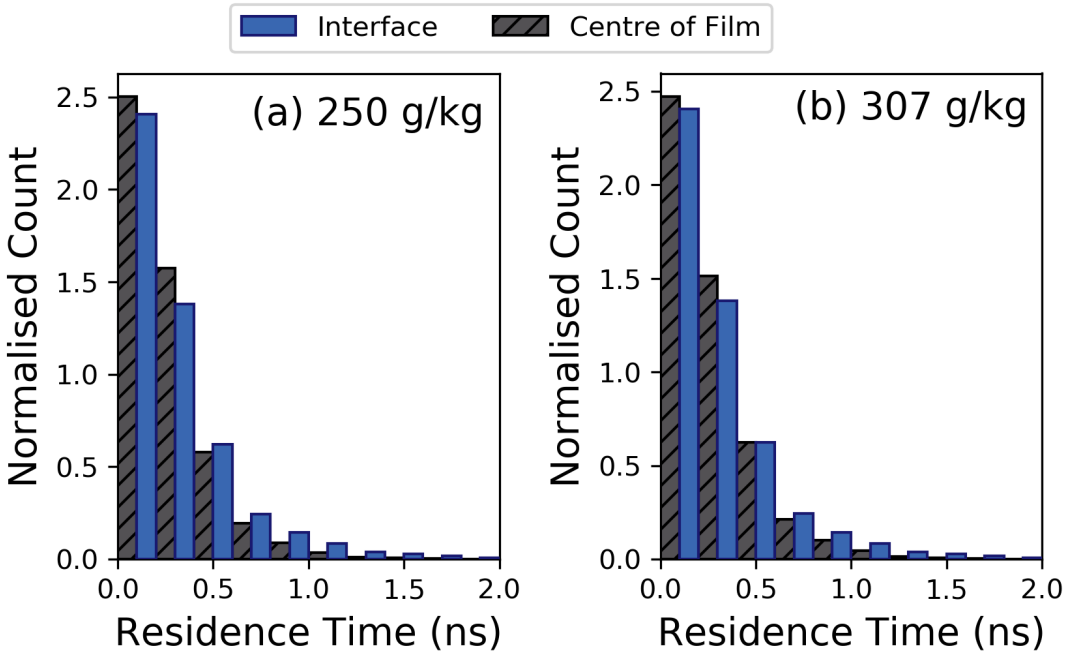

Figure 5.9: Residence times of the glycine molecules within the interfacial region and an equally sized slice at the centre of the film for the oil–solution interface system at concentrations of (a) 250 g/kg and (b) 307 g/kg.

While the MSD is not a suitable metric for the mobility perpendicular to the interface, it can used to compare the dynamics in the x-y direction (parallel to the interface) for the interfacial regions and the centre of the film as shown in Figure 5.10. Here, the full remainder of the solution film, in contrast to the 1 nm slice used for the residence time analysis, has been considered for the centre of film values to improve statistics. While typically the MSD can be used to determine diffusion coefficients, the short residence times of the molecules at the oil–solution interface results in insufficient statistics to obtain robust values. However, qualitatively it can be seen that there is a slight reduction in the parallel mobility of the glycine molecules, but that the molecules also remain mobile in this direction.

To determine the rotational mobility of the molecules, the autocorrelation function

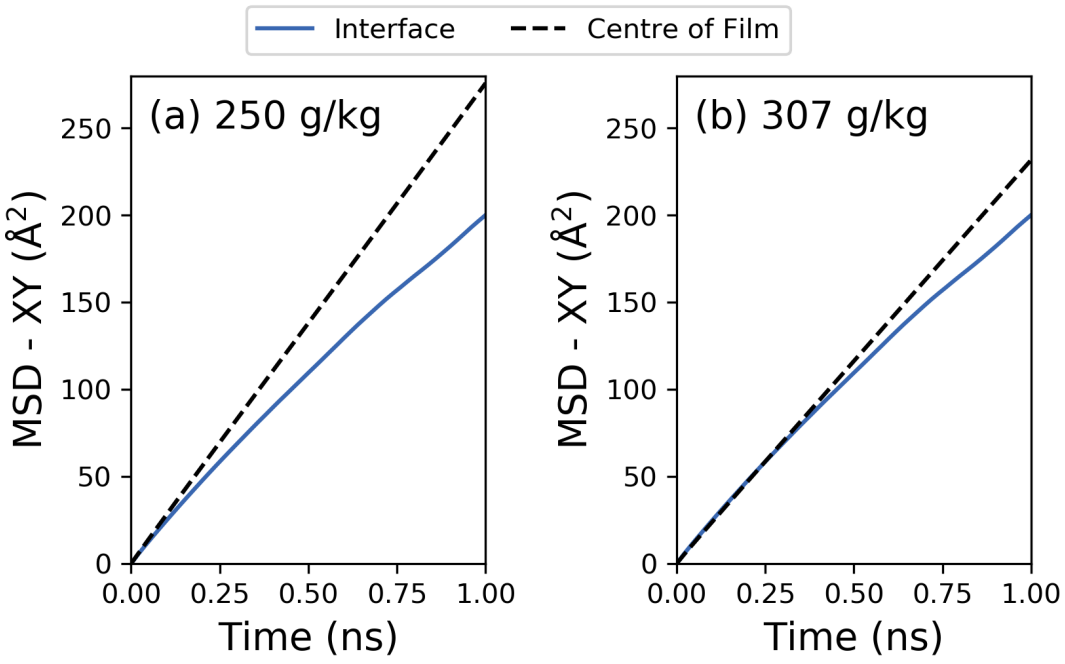

Figure 5.10: Mean Squared Displacement (MSD) of the glycine molecules within the interfacial region and centre of the film for the oil–solution interface system at concentrations of (a) 250 g/kg and (b) 307 g/kg. Note that the interfacial MSD is less linear due to worse statistics resulting from low residence times at the interface.

(ACF) of the C-C bond vector of the glycine molecules was calculated using:

$$ACF = \left\langle \frac{v(t_i) \cdot v(t_j)}{v(t_i) \cdot v(t_i)} \right\rangle \tag{5.2}$$

where $v(t_i)$ is the bond vector at time i and $v(t_j)$ is the bond vector at time j. All time windows have been included (i.e. t_i does not have to be 0) to improve statistics. The ACF of glycine molecules while they remain within the interfacial region is compared to the ACF of molecules within the centre of the solution film in Figure 5.11. Once more, there is a slight reduction in the mobility of the glycine molecules within the interfacial regions. By all three metrics there is a reduced mobility of glycine within the interfacial region however, this would be expected in a region of higher concentration [92, 122]. It is clear that the molecules remain highly mobile by all three measures and are not strongly immobilised at the interface. This is in contrast to previously observed effects such as

barite epitaxial growth that showed the water monolayer formation on a barite surface, where the adsorbed water creates a barrier for barium and sulfate ions approaching the surface [123].

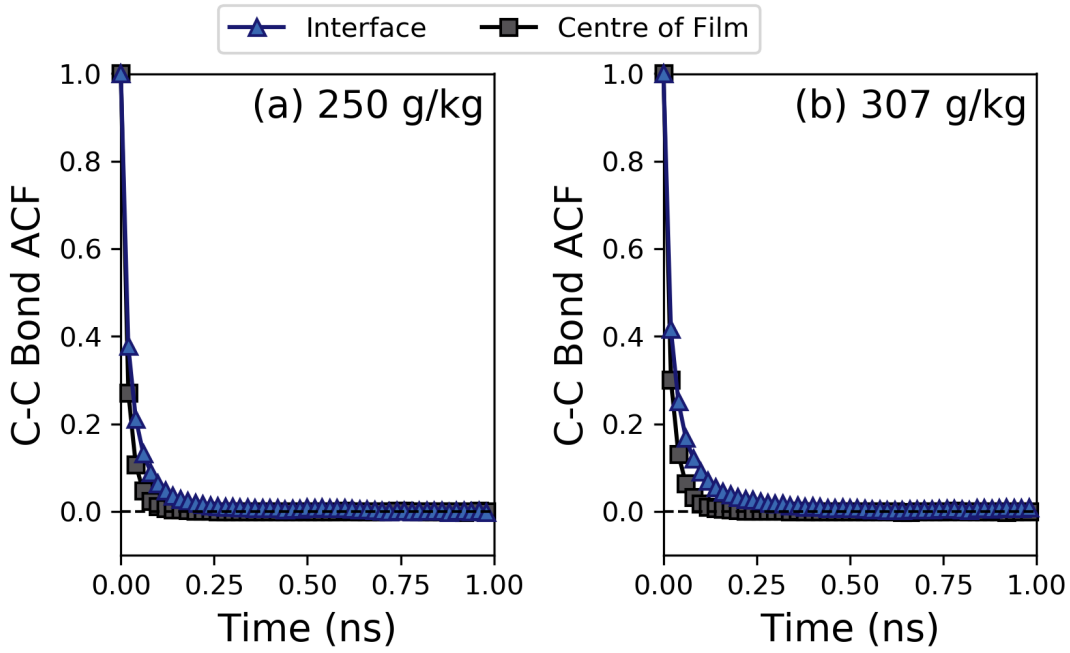

Figure 5.11: Autocorrelation function (ACF) of the C-C bond of glycine molecules within the interfacial region and centre of the film for the oil–solution interface system at concentrations of (a) 250 g/kg and (b) 307 g/kg.

5.4 Conclusions

In this chapter, MD simulations were used to gain insight into the solution structure and dynamics of glycine molecules at oil–solution and air–solution interfaces. This revealed an enhanced concentration and ordering of the glycine molecules at the oil–solution interface. At the air–solution interface the opposite effect is seen with a greatly reduced glycine concentration. These effects are likely due to non-specific net van der Waals interactions between the tridecane–water and tridecane–glycine molecules, which compete with solution interactions to determine the interfacial solution compositions.

While these effects were observed for glycine solutions at a tridecane interface, it is expected that the same mechanism will be present in a wide range of solution-interface systems. This has significant implications for the interpretation of nucleation experiments, as this suggests that the overall solution concentration is not necessarily representative of the environment in which heterogeneous nucleation occurs. Experiments performed at the same overall concentration may present different nucleation kinetics depending on the interfaces present. Further understanding of these effects may allow for the control interfacial concentrations in order to design effective nucleants for the enhancement of nucleation or to prevent heterogeneous nucleation in anti-fouling applications.

Chapter 6

Molecular Simulations with Model Interfaces

This chapter aims to determine how interface–solution interactions influence the solution composition within the interfacial region through the use of model interfaces. First, a description of a model interface with tuneable dispersion interactions, the Lennard Jones (LJ) 9-3 wall and details of the implementation will be provided. This will be followed by a description of the process that was developed to allow the LJ 9-3 wall to represent real materials. Following this, the results of the parameterisation process for three materials, heptane, tridecane and graphite, will be presented along with validation of the model interface representing tridecane. Simulations of glycine solution films in contact with LJ 9-3 walls of varying wall strength will be presented, demonstrating that the dispersion interactions have a strong influence on the solution composition at the interface. The effect of system size will be explored, to determine the influence of finite size effects on the obtained results. Finally, the impacts of electrostatic interactions will be examined and compared with the effects of dispersion interactions.

6.1 Introduction

In Chapter 5 it was demonstrated that the glycine concentration at interfaces can differ significantly from the average concentration in the solution, contributing to the enhanced rates of nucleation observed for heterogeneous nucleation. It is proposed that this concentration enhancement effect is due to dispersion interactions between the interface and the solution molecules. In order to explore the effects of dispersion interactions on the solution composition at the interface, it would be possible to simulate solutions in contact with interfaces with similar electrostatic properties, but differing dispersion interactions. However, it is difficult to decouple the effects of each of these types of interactions, as both properties are likely to change to some extent when changing between interfacial materials. In addition to this, the computational cost of MD simulations increases significantly with the number of atoms, and as such, simulations with fully atomistic interfaces can be computationally expensive. Model interfaces can allow for the direct tuning of interface–solution interactions, while also significantly reducing the computational cost of the simulations by representing the interface as a simple function of distance. This makes the use of model interfaces an attractive prospect, particularly for material screening purposes where it may be necessary to simulate a wide range of interfacial systems.

In this chapter, simulations have been performed with a model interface with tuneable dispersion interactions, a Lennard Jones (LJ) 9-3 wall. This allows the effect of interface–solution dispersion interactions on the interfacial concentration, molecular orientation and dynamics within the interfacial region to be determined at a significantly reduced computational cost. The model interface has been parameterised to represent specific materials (heptane, tridecane and graphite) enabling us to predict how these materials will influence the heterogeneous nucleation rate of glycine. This chapter will demonstrate that the concentration enhancement effect can be reproduced using a simple planar LJ wall, demonstrating the transferability of the phenomenon, and that these walls can be parameterised

from existing molecular models so that the simplified model is grounded in real materials. A methodology is presented that has predictive capabilities and can therefore be used in future work as a design tool for materials selection for the purposes of nucleation control. Finally, the effect of electrostatics will be examined and their significance compared to those observed for dispersion interactions.

6.2 Methodology

6.2.1 System setup & MD simulations

Within this chapter all MD simulations, with the exception of the bulk simulation performed in sections 6.3.5 and 6.3.6, have been carried out in the NVT ensemble, with periodic boundary conditions applied in the x and y directions, and non-periodic boundaries in the z-direction. Each simulation box was large enough in the z-direction to maintain a vacuum layer on one side of the glycine solution, to prevent compression of the glycine solution against the solid interface, as shown in the schematic in Figure 6.1. LJ interactions, including the LJ 9-3 wall potentials, were truncated at a cut-off of 1.4 nm. Short-range electrostatics were calculated below 0.98 nm and long-range electrostatics were calculated using a particle-particle-particle-mesh with a relative accuracy of 1×10^{-6}. A slab correction was applied to allow for long range electrostatics to be applied that treats the system as if it was periodic in z, but with an empty volume inserted between the slabs, effectively removing the electrostatic interactions between the slabs [124].

Glycine solution films were prepared at thicknesses ranging from 3.1 to 13.9 nm for concentrations of 296.7 and 500.7 g/kg, which are both supersaturated with respect to the experimental solubility of 202.1 g/kg for γ-glycine at 298 K [73]. The x-y cross-sectional

area for each film was 3.45×3.45 nm^2. Details of number of molecules in each film and the equilibration and production simulation times are given in Table 6.1.

Table 6.1: Table of system details and simulation times for each of the glycine solution films used in the MD simulations.

Concentration (g/kg)	$\mathbf{N}_{\text{glycine}}$	$\mathbf{N}_{\text{water}}$	**Film Thickness** (nm)	t_{equil} (ns)	t_{prod} (ns)
296.7	64	900	3.1	25	15
296.7	128	1800	5.9	35	25
296.7	256	3600	11.4	40	40
296.7	320	4500	13.9	60	70
500.7	108	900	3.2	25	15
500.7	216	1800	6.1	35	25
500.7	432	3600	12.2	40	40

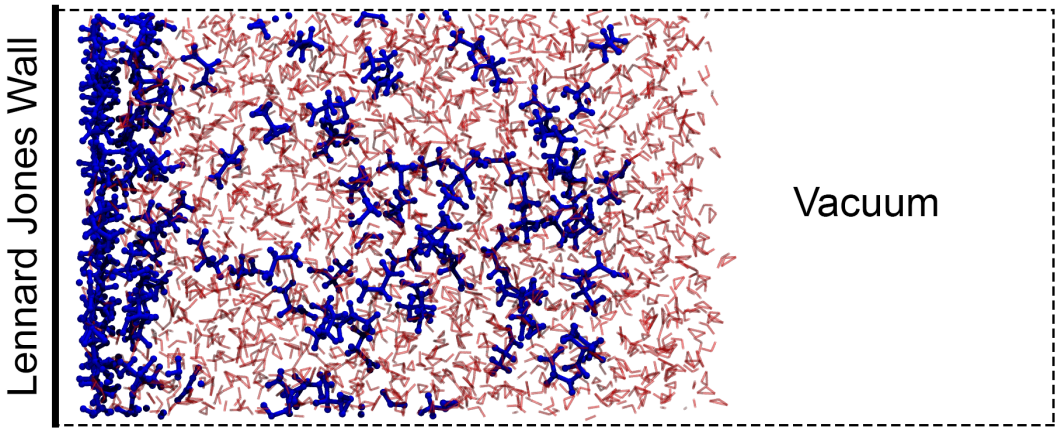

Figure 6.1: Schematic of the simulation setup. Glycine and water molecules are shown in blue and red respectively. The dashed black lines represent the boundaries of the simulation box, while the thick black line shows the position of the LJ interface.

To determine the impact of electrostatics on the solution composition at the interface simulations of the 296.7 g/kg 3.1 nm solution film were performed in contact with an atomistic, crystalline tridecane interface with varying charge sets and LJ parameters. LJ parameters for PTFE from GAFF [94] and tridecane from AMBER-ii [98] were used to vary the dispersion interactions. Three charge sets were used to vary the electrostatic interactions: charges for tridecane from AMBER-ii, charges for PTFE from GAFF and the PTFE charges from GAFF doubled.

6.2.2 Model interface and mixing rules

A LJ 9-3 wall potential was used as a model interface with tuneable dispersion forces. The functional form implemented within the LAMMPS code is given by:

$$E_{LJ\,9-3} = \epsilon_{iw}\left[\frac{2}{15}\left(\frac{\sigma_{iw}}{z_{iw}}\right)^9 - \left(\frac{\sigma_{iw}}{z_{iw}}\right)^3\right] \tag{6.1}$$

where ϵ_{iw} controls the depth of the well of the interaction between the wall w and a given atom i, σ_{iw} is related to the distance at which the potential crosses the x-axis and z is the distance perpendicular to the wall between the wall and the atom. The wall potential allows the interface to be represented without requiring the interactions for each atom within the interface to be calculated explicitly, resulting in a significantly reduced computational cost.

It is important to note that in equation 6.1 while ϵ and σ play a similar role to the standard LJ 12-6 potential, they correspond to different points on the potential which has implications on the use of combining rules to produce the ϵ_{iw} and σ_{iw} parameters. To demonstrate this, Figure 6.2 shows LJ 12-6 and 9-3 potentials of the same ϵ and σ parameters. The LJ 12-6 interaction crosses the x-axis at a distance equal to σ, while the LJ 9-3 interaction crosses the x-axis at $r = 0.715\ \sigma$. Additionally, while ϵ represents the depth of the potential well for the LJ 12-6 potential, the LJ 9-3 potential has a slightly deeper well of 1.05ϵ.

An LJ 9-3 potential can be obtained by integrating over a 3D half lattice of LJ 12-6 particles [125]. Other functional forms of wall potentials do exist, such as the LJ 10-4 potential, which can be obtained by integrating over the surface plane of the 3D lattice of LJ 12-6 particles. As the interactions are between the solute and an oil layer, the LJ 9-3 potential is most appropriate. This also provides the closest comparison to simulations of fully atomistic interfaces where all atoms within the LJ cut-off will interact with the

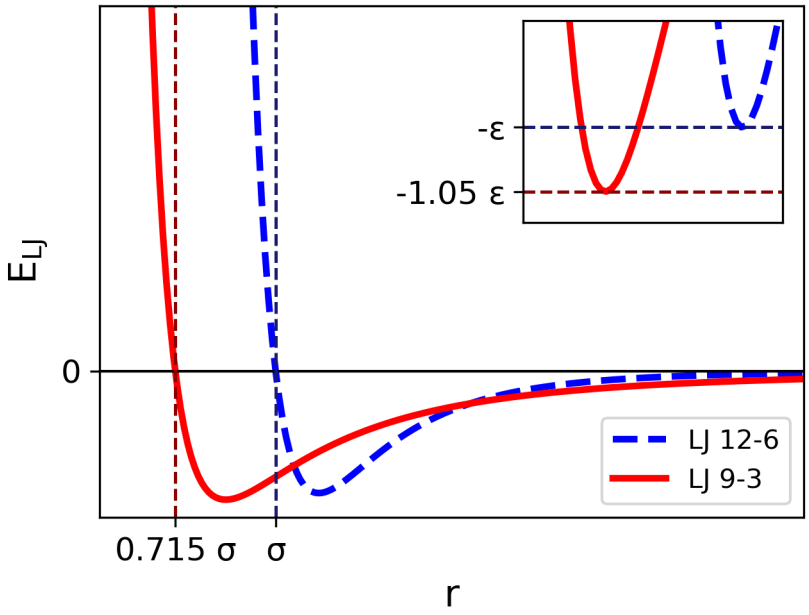

Figure 6.2: Lennard Jones 9-3 and 12-6 potentials with the same ϵ and σ parameters. Dashed lines on the main plot show the point at which the potential crosses the x-axis. The insert shows the region close to the minima of the two functions, and the dashed lines show the positions of these minima.

solution.

As discussed in Chapter 3, Lorentz-Berthelot rules are commonly used to produce ϵ and σ parameters for interactions between different particle types. The arithmetic mean is used to combine σ parameters, however, due to the different position of σ in equation 6.1 the adjusted equation 6.2 has been used.

$$0.715\,\sigma_{iw} = \frac{\sigma_{ii} + 0.715\,\sigma_{ww}}{2} \tag{6.2}$$

The geometric mean is used to combine ϵ parameters and as there is only a 5% difference between the LJ 9-3 and 12-6 potential well depths, as shown in Figure 6.2, the standard geometric mean combining, rule equation 3.19, has been used.

6.2.3 Model interface parameterisation

While is it possible in MD simulations to model a variety of interfacial materials atomistically, these simulations can be computationally expensive. Wall potentials, where the interaction depends only on the distance from the wall, can significantly reduce the number of interactions between particles and thus reduce the computational cost of the simulations. In order to relate the wall potential to specific materials, it is necessary to map the wall potential parameters to those of the atomistic representation. A procedure has been developed to parameterise wall potentials to specific materials, and has been applied to heptane, tridecane and graphite. In each case a slab of the material was created so that it was at least 3 nm in the x, y and z directions. To remove effects related to changes in the interface height (such as due to capillary waves) and surface variation, flat interfaces with periodic structure were prepared. For heptane and tridecane, pseudo-crystalline structures of periodic extended chains were prepared, so that the densities matched the experimental densities of the liquids at 298 K. For graphite the crystal structure was used directly. A snapshot of the crystalline tridecane slab is shown, with its dimensions, as an example in Figure 6.3.

In order to obtain suitable ϵ_{iw} and σ_{iw} parameters to represent the interactions between the interface and the solution molecules, the solution molecules were broken down into their constituent atoms. The interactions were mapped for each of the atom types within glycine, and the oxygen atom within the water molecules. The ϵ_{ii} and σ_{ii} parameters for each of these atom types are provided in Tables 3.1 and 3.2. The hydrogen atoms of the water molecules could not be included in the parameterisation process as they do not have LJ interactions with any atoms. An individual atom was placed at a set distance, z, from the interface and the total LJ 12-6 interaction between the atom and the interface was calculated. The distance z was then increased by a small amount and the process was repeated in order to map out the z-dependent LJ potential between the interface

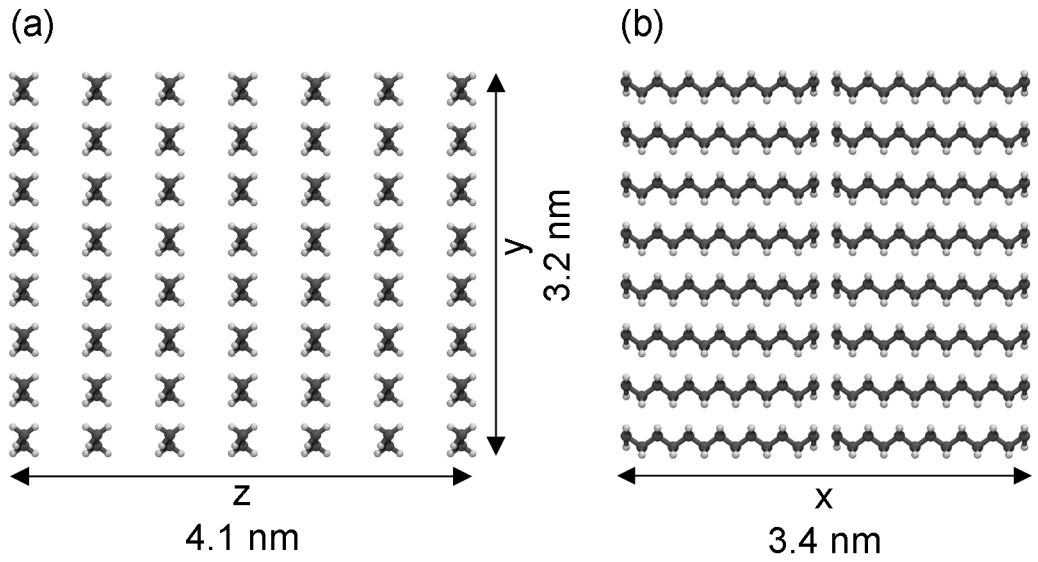

Figure 6.3: Snapshot of the crystalline tridecane slab used for parameterising the LJ 9-3 wall. (a) shows the z-y plane, while (b) shows the x-y plane. Carbon atoms are shown in grey, while hydrogen atoms are shown in white. The dimensions of the slab in are provided for each direction.

and the atom. Equation 6.1 was then fit to these values to obtain suitable ϵ_{iw} and σ_{iw} parameters. The LJ 9-3 potential is uniform along the interface plane, which can result in an underestimation of the potential strength of crystalline materials where the density is localised at lattice sites [126]. To account for this, the parameterisation process was repeated at 36 individual x,y positions across the interface and the average ϵ_{iw} and σ_{iw} was taken. A detailed example of the parameterisation of the LJ interaction between tridecane and the nitrogen atom of glycine is provided in section 6.3.1 for extra clarity.

This process was repeated for each atom type in order to build a full set of parameters for the interaction between the interface and the solution. The obtained ϵ_{iw} and σ_{iw} parameters can be plotted against the LJ parameters of the individual atoms, ϵ_{ii} and σ_{ii}. Equations 3.19 and 6.2 are then fit to this data in order to obtain suitable parameters that describe the interface, ϵ_{ww} and σ_{ww}. These wall parameters can now be combined with any given atom, with parameters ϵ_{ii} and σ_{ii}, using the above combining rules to obtain suitable parameters that describe the interface–atom interaction.

6.3 Results and Discussion

This section will demonstrate the effects of dispersion and electrostatic interactions on the structure and dynamics of glycine solutions within the interfacial region, allowing the identification of materials that would increase the nucleation rate.

Section 6.3.1 will provide a detailed example of the LJ 9-3 parameterisation process. This will be followed by the results of the model interface parameterisation for the heptane, tridecane and graphite. The model interface and parameterisation approach will then be validated by comparing the solution behaviour at the tridecane model and atomistic interfaces. In section 6.3.5, a range of model interface parameters will be explored that span the values of the three materials. In section 6.3.6 finite size effects and their influence on the results of the MD simulations will be investigated. Finally, the effect of electrostatic interactions and their relative impact in comparison dispersion interactions will be examined.

6.3.1 Interface–Atom Parameterisation Example: Tridecane and Nitrogen

In order to provide extra clarity on the LJ parameterisation process, this section will describe how the parameters ϵ_{iw} and σ_{iw} were obtained for the interaction between the LJ 9-3 wall representing tridecane and the nitrogen atom within the glycine molecules.

The nitrogen atom within the glycine molecule, n4, has LJ parameters of $\epsilon_{ii} = 0.17$ kcal/mol and $\sigma_{ii} = 0.325$ nm. 36 evenly spaced positions were selected in the x-y plane of the atomistic, crystalline tridecane slab. A nitrogen atom was placed at a distance, z, of 0.01 nm above the interface at the first x-y position, and the total LJ 12-6 interaction between the

tridecane slab and the nitrogen atom was calculated. The distance was then increased in 0.01 nm steps, up to a maximum of $z = 1.4$ nm, calculating the LJ 12-6 interaction at each step. Figure 6.4 (a) shows a snapshot of the nitrogen interacting with the tridecane slab.

This provides a z-dependent potential between the atom and the interface, such as the one shown in Figure 6.4 (b). Equation 6.1 is then fit to this potential to obtain ϵ_{iw} and σ_{iw} parameters that describe the interaction between the nitrogen atom and the tridecane interface at the specific x-y position. Measurements close to the interface resulted in extreme values of E_{LJ} due to repulsion between the nitrogen atom and the tridecane interface. Due to this, only values of $E_{LJ} < 3.0$ kcal/mol were included in the fitting of equation 6.1 to better capture the behaviour of the potential well. This does result in deviation from the z-dependent potential at z values close to the interface, however this will not influence the simulation results as atoms will be unable to reach this highly repulsive region under typical conditions. Fitting equation 6.1 to the data shown in Figure 6.4 (b) resulted in parameters $\epsilon_{iw} = 1.19$ kcal/mol and $\sigma_{iw} = 0.301$ nm.

This process was repeated for each of the 36 x-y positions across the face of the tridecane slab, and each of the z-dependent potentials are shown in Figure 6.4 (c) as the coloured data sets. The average ϵ_{iw} and σ_{iw} values are 1.22 kcal/mol and 0.316 nm respectively, and the resulting LJ 9-3 potential is shown as the black line in Figure 6.4 (c). The values ϵ_{iw} and $0.715\sigma_{iw}$ are used in the fitting of equations 3.19 and 6.2 in order to obtain the wall parameters, ϵ_{ww} and σ_{ww}, to represent tridecane. These values correspond to the red squares plotted at $\epsilon_{ii} = 0.17$ kcal/mol and $\sigma_{ii} = 0.325$ nm in Figure 6.5.

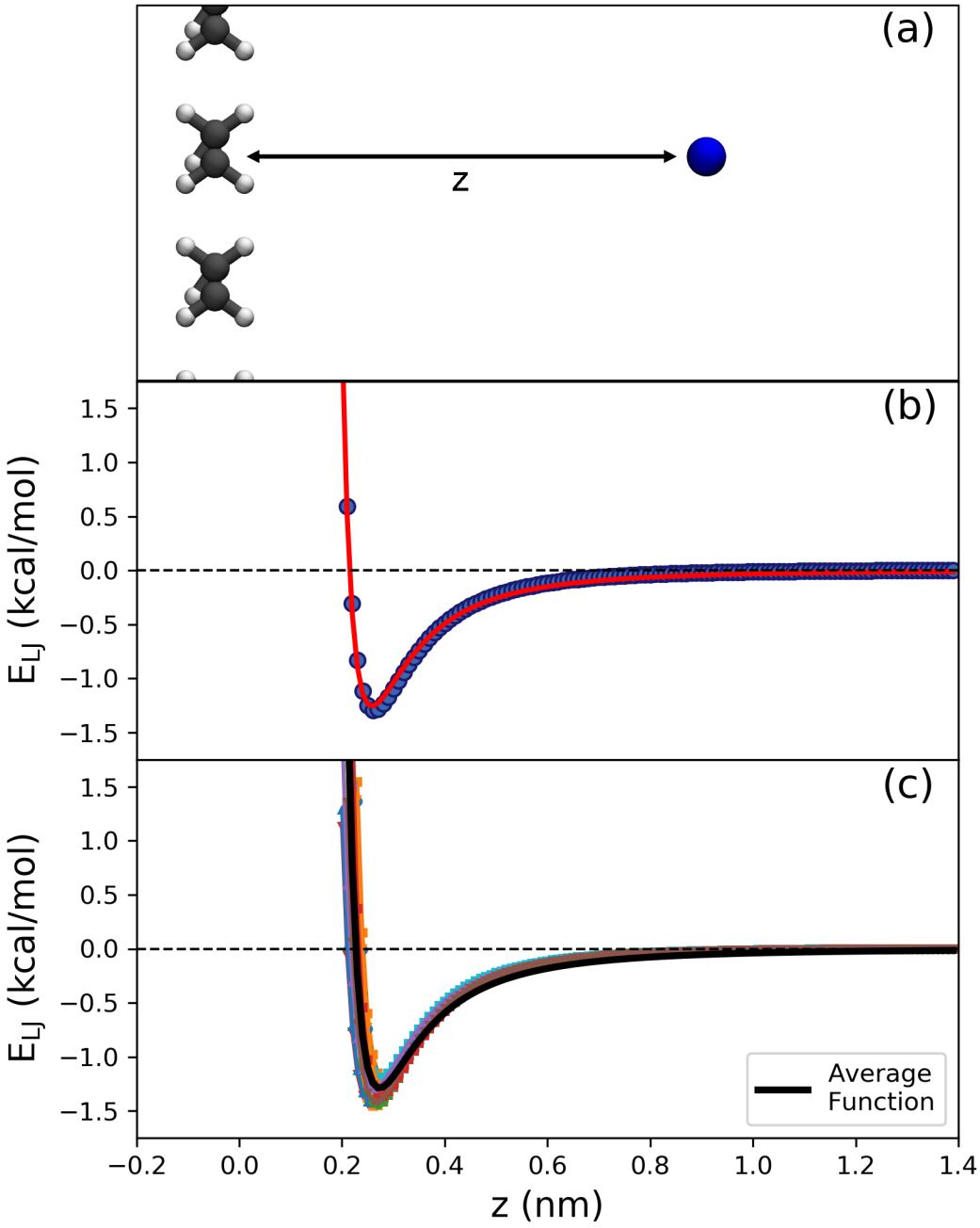

Figure 6.4: (a) Schematic to demonstrate the parameterisation process for the crystalline tridecane system. The LJ interactions between the tridecane slab and the blue nitrogen atom are calculated for the distance, z, from the interface. (b) The LJ interaction mapped out for a nitrogen atom interacting with the tridecane at varying distances at one x-y position. The red line shows the fit to equation 6.1 to obtain ϵ_{iw} and σ_{iw} (c) LJ interactions measured for a nitrogen atom interacting with the tridecane at each of the 36 x-y positions. Each coloured set of points represents one x-y position, while the black line shows the function using the average ϵ_{iw} and σ_{iw} values.

6.3.2 Model Interfaces for Tridecane, Heptane and Graphite

The parameterisation process described in section 6.3.1 was repeated for each of the atom types within glycine and the oxygen atom of water in combination with the tridecane, heptane and graphite interfaces. This allowed ϵ_{iw} vs ϵ_{ii} and σ_{iw} vs σ_{ii} plots to be constructed for each of the three interfacial materials as shown in Figure 6.5. Equation 3.19 and 6.2 were then fit to each of these data sets to obtain the ϵ_{ww} and σ_{ww} parameters that best represent each of the interfacial materials. The dashed lines in Figure 6.5 represent these fits and the resulting wall parameters are provided in Table 6.2. Good fits are obtained for each set of data, however it can be seen that at the lowest ϵ_{ii} and σ_{ii} parameters the fits deviate from the obtained values. As these values correspond to the hydrogen atoms of the glycine molecules, this deviation is not expected to significantly impact the behaviour of the system. In terms of wall strength heptane has the weakest interaction, followed by tridecane and finally graphite.

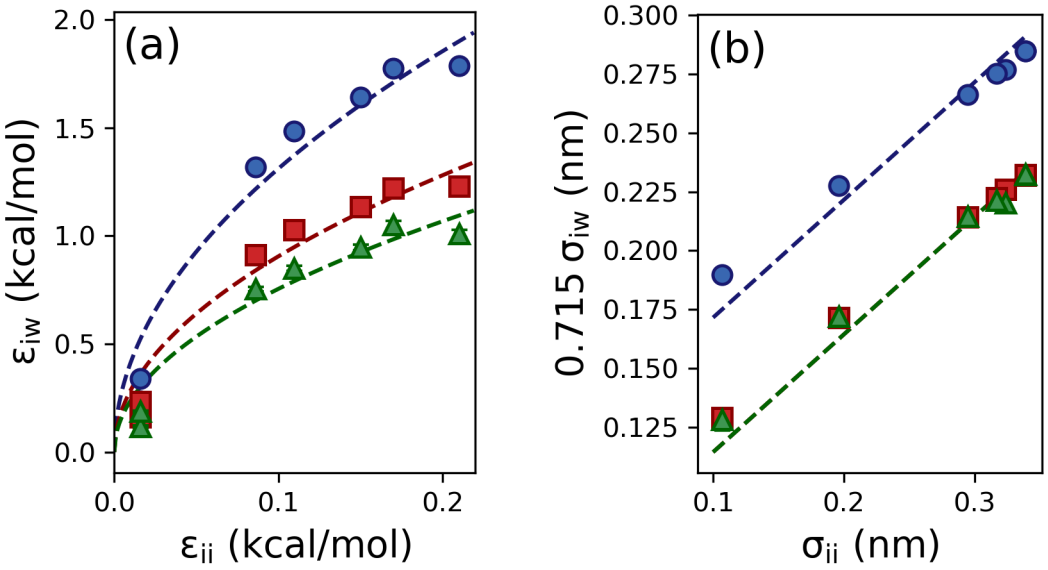

Figure 6.5: The (a) ϵ_{iw} and (b) $0.715\sigma_{iw}$ parameters for the interaction each atom type and graphite (blue circles), tridecane (red squares) and heptane (green triangles) plotted against their equivalent atom-atom parameters. Error bars represent the standard error. Dashed lines represent fits to equations 3.19 and 6.2.

Table 6.2: Table of fitted ϵ_{ww} and σ_{ww} parameters for selected materials.

Material	$\boldsymbol{\epsilon_{ww}}$ (kcal/mol)	$\boldsymbol{\sigma_{ww}}$ (nm)
Heptane	5.7	0.18
Tridecane	8.2	0.18
Graphite	17.2	0.34

6.3.3 Validation of the Model Interface

To validate the parameterisation of the model interfaces, the properties of the 296.7 g/kg, 3.1 nm film in contact with the atomistic, crystalline tridecane interface and the LJ 9-3 wall parameterised for tridecane have been compared. The density profiles obtained for each simulation are shown in Figure 6.6. There is excellent agreement between the density profiles, with peaks of similar magnitude in both cases, although the density peaks closest to the interface appears to favour glycine over water in the case of the LJ 9-3 wall.

Similarly to Chapter 5, we are particularly interested in the amount of glycine present within the interfacial region. Previously, due to the liquid nature of the interface, the interfacial region was defined based on the concentration profile of glycine within the solution (see Figure 5.3). In this instance, the position of the interface is fixed and well-defined. To ensure that the increasing size of the exclusion zone for increasing values of σ_{ww} did not influence the obtained results, the 1 nm interfacial region was defined from the point at which the LJ 9-3 potential crosses the x-axis ($z = 0.715\sigma_{ww}$).

Both density profiles in Figure 6.6 show significant layering that was not observed in the simulations of the liquid tridecane–glycine solution interface in the previous chapter. This layering is likely due to the fixed, flat nature of these interfaces and will be discussed further in section 6.3.4. In Chapter 5, the amount of glycine present within the interfacial region was quantified using the glycine concentration (the ratio of the density of glycine

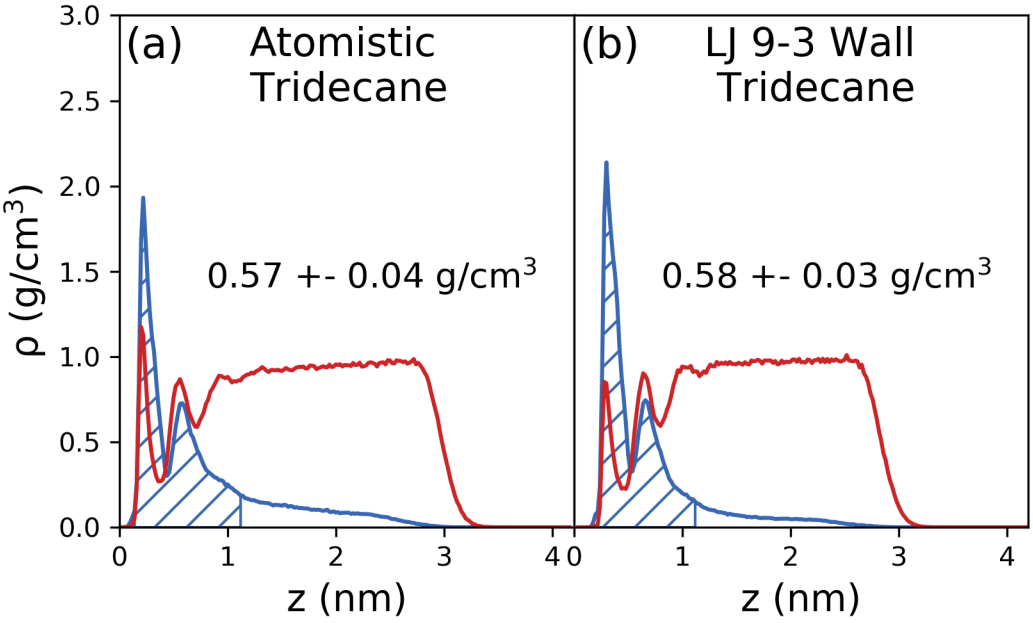

Figure 6.6: The density profiles for glycine (blue) and water (red) solution in contact with the (a) atomistic crystalline tridecane and (b) the corresponding LJ 9-3 wall (right). The hatched area indicates the 1 nm interfacial region. The average glycine density within the interfacial region are reported on each graph, with the standard deviation of the average density over time reported as the error.

to water), however, due to the presence of these highly dense glycine layers, particularly in cases where higher wall strengths are used, it no longer seems appropriate to describe the amount of glycine within the interfacial region by the concentration of the glycine solution. As such, for this chapter the average density of glycine within the interfacial region will be used to quantify the amount of glycine at the interface. The average glycine densities within the interfacial region of the atomistic crystalline tridecane and LJ 9-3 wall parameterised for tridecane are 0.57 and 0.58 g cm^{-3} respectively, showing that the LJ 9-3 wall is able to accurately reproduce the glycine interfacial density.

We can further compare the interfacial structural by analysing the orientation of the glycine molecules using the bond orientation parameter, P_2, as described in Chapter 5. The P_2 values for the atomistic tridecane and LJ 9-3 tridecane simulations are shown in

Figure 6.7. Both simulations demonstrate the expected random orientation of glycine in the bulk of the film, and ordering at both the solid-solution and solution-vacuum interfaces. At the solid-solution interface we see a strong alignment of glycine with the C-C bond parallel to the interface in the first layer and a weaker alignment within the second layer. The position and magnitude of the peaks in the P_2 values show excellent agreement between the two simulated systems, demonstrating that the LJ 9-3 wall successfully matches the structural details of the fully atomistic interface. Interestingly, we see stronger alignment with the solid interfaces than was observed in the previous liquid tridecane simulations, suggesting that the fluctuating non-planar interfaces also obscure the ordering of the molecules.

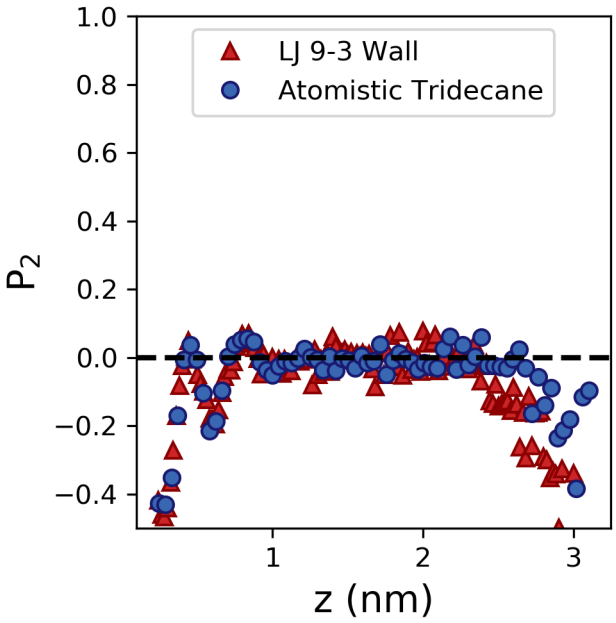

Figure 6.7: The obtained bond order profiles for glycine solution in contact with the atomistic crystalline tridecane (blue circle) and the LJ 9-3 wall fitted to represent the crystalline tridecane (red triangles).

We have investigated the translational and rotational mobility of the glycine molecules within the two layers that form at the interface. The translational mobility parallel to the interface has been determined using the MSD of the glycine molecules while they remain in a given layer. As the distance that the molecules can move in the z-direction is capped by

the size of the region specified for analysis it is inappropriate to use the MSD to determine their mobility perpendicular to the interface. Instead, the residence time of the molecules within each layer has been analysed to provide information of the molecules mobility in this direction. The rotational mobility of the molecules have been determined through the autocorrelation function (ACF) of the C-C bond vector of the glycine molecules while they remain in each of the interfacial layers.

The results of the dynamics analysis are presented in Figures 6.8 - 6.10. At both the LJ 9-3 wall and the atomistic interfaces we see a reduction in the mobility of the glycine molecules within the interfacial layers when compared to a simulation of bulk glycine solution at the same overall system concentration, with the rotational and movement perpendicular to the interface being more significantly reduced within the layer closest to the interface for both systems. The movement of the molecules parallel to the interface was found to be similar between the two layers, with the layer closest to the interface being marginally slower. The rotational and translational movement perpendicular to the interface show good agreement between the two systems, however the glycine molecules showed faster translation parallel to the interface within the atomistic interface system. This is a surprising result, as it would be expected for the LJ 9-3 system to result in faster movement along the interface due to reduced friction as has been seen previously for simulations of water in contact with a LJ 9-3 interface [127].

These results demonstrate that the LJ 9-3 wall is capable of successfully reproducing the solution structure and dynamics at the tridecane interface, however another important aspect is the computational cost. These simulations were performed on the ARCHIE-WeSt high performance computer (HPC), based at the University of Strathclyde, using 10 cores of an Intel Xeon Gold 6138 (Skylake) processor. Using this hardware, the atomistic tridecane interface simulation resulted in a performance of 5.71 ns/day, whilst the LJ 9-3 interface simulation achieved a performance of 10.19 ns/day, a performance increase of

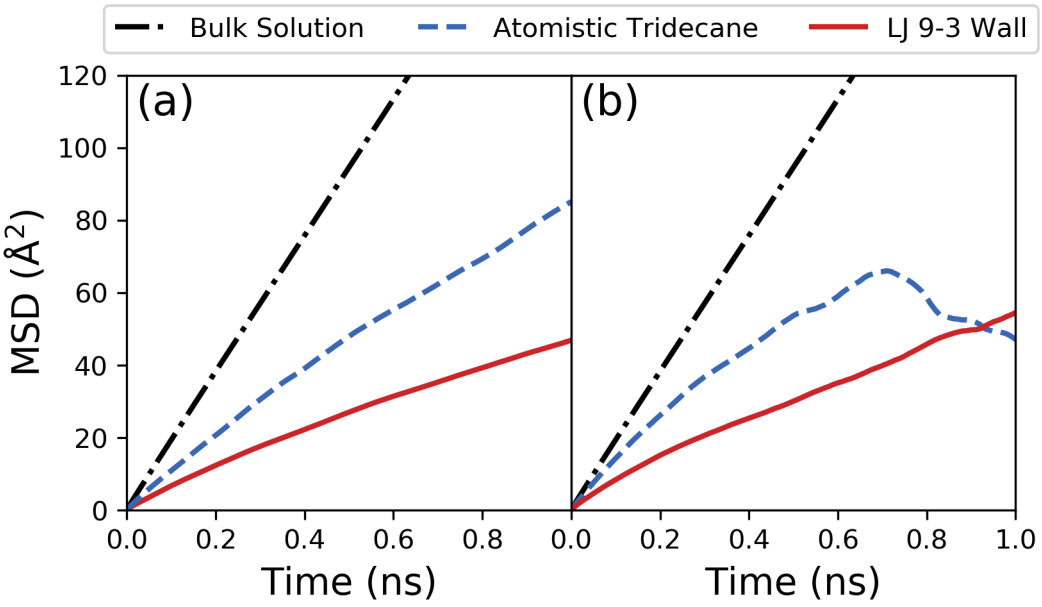

Figure 6.8: Mean squared displacement of the glycine molecules in the x-y direction (parallel to the interface) within (a) the first interfacial layer and (b) the second interfacial layer for the atomistic and LJ 9-3 interfacial systems. The black dash-dotted line represents the MSD in the x-y direction for glycine molecules within a homogeneous solution of the same concentration.

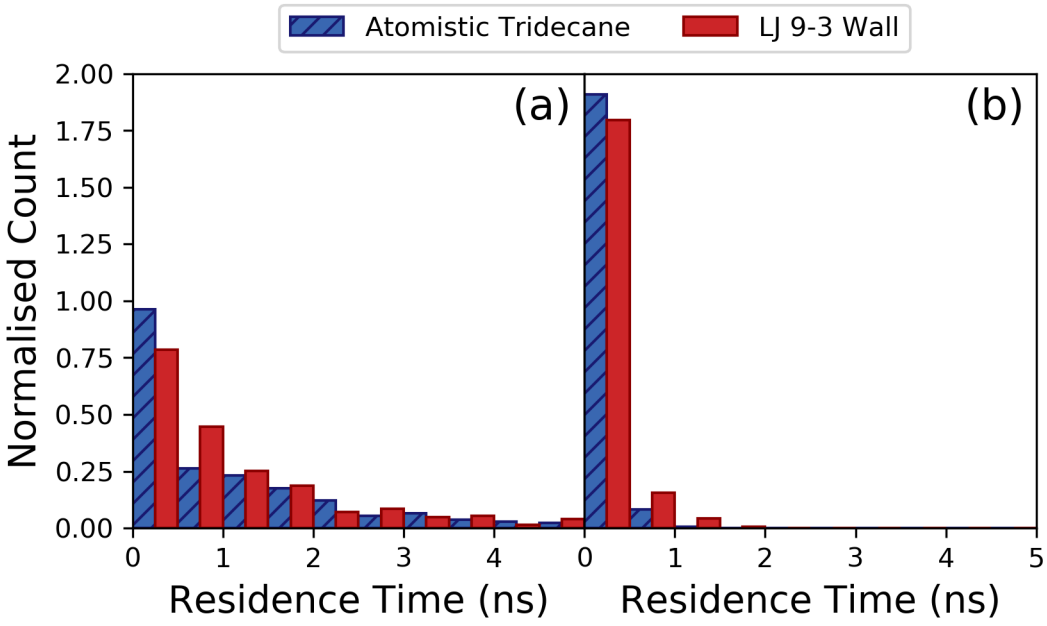

Figure 6.9: Distribution of residence times for glycine molecules within (a) the first interfacial layer and (b) second interfacial layer for the atomistic and LJ 9-3 interfacial systems.

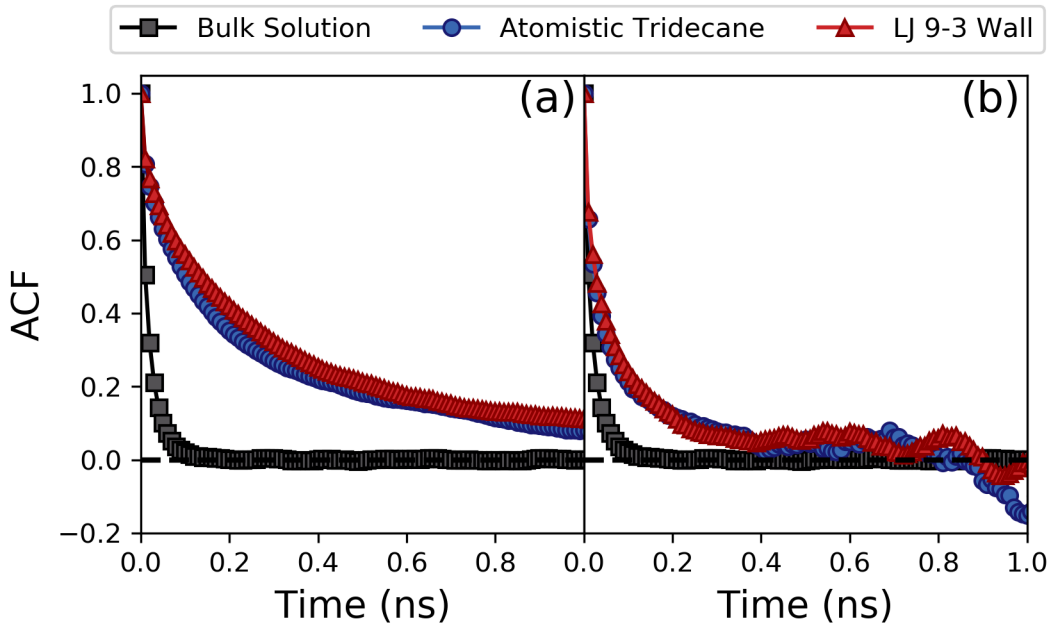

Figure 6.10: Autocorrelation function of glycine C-C bonds for molecules within (a) the first interfacial layer and (b) the second interfacial layer for the atomistic and LJ 9-3 interfacial systems. The black squares represent the autocorrelation function of glycine molecules within a homogeneous solution of the same concentration.

78%. While the performance increase will depend on a number of factors, such as the hardware used or the ratio of atoms between the interface and the solution, this clearly demonstrates that the LJ 9-3 wall provides a significant reduction in the computational cost.

6.3.4 Layering at the Interface

The density profiles obtained for the glycine solution in contact with both the atomistic crystalline tridecane and the LJ 9-3 wall representing tridecane, presented in Figure 6.6, show significant layering of the glycine solution in a manner that wasn't observed for the simulations of the liquid tridecane–glycine solution interface presented in Chapter 5. The observed layering is likely due to the fixed, flat nature of the interfaces used in these simu-

lations. As discussed in Chapter 5, the density profiles do not account for capillary waves and interfacial fluctuations. As such, the density of each component at the true, undulating interface is smeared across multiple bins. Due to this smearing, it is possible that these strong peaks are also present within the previous simulations, however it would be necessary to employ more advanced interfacial analysis techniques, such as the generalised identification algorithm for non-planar interfaces [119], to identify these.

To provide further insight into the effects of liquid vs solid interfaces on the obtained density profiles, a simulation of the smaller solution film at 307 g/kg, from Chapter 5, was carried out in contact with the liquid tridecane interface frozen in place to provide an amorphous, solid interface. This frozen, amorphous interface acts as an intermediate step between the fully liquid and crystalline tridecane interfaces. The density profile obtained for this simulation is shown in Figure 6.11 alongside the density profile obtained for the solution in contact with the fully liquid interface.

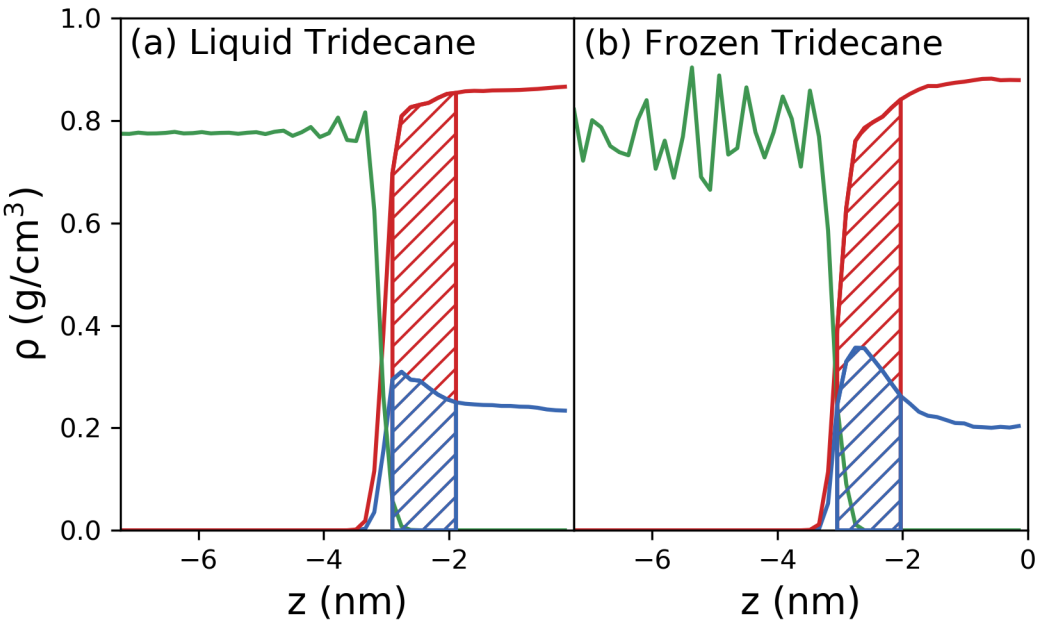

Figure 6.11: The density profiles obtained for glycine solution films of concentration 307 g/kg in contact with (a) fully liquid tridecane and (b) frozen, amorphous tridecane. The hatched area represents the defined 1 nm interfacial region.

The glycine density peak is both higher and sharper in the density profile obtained from the frozen tridecane simulation. It is important to note that the frozen interface will still vary in height, as it is simply the liquid interface frozen in place, and so there will still be smearing of the density profiles across bins. However, as the tridecane is now frozen in place, the surface of the interface is consistent throughout the simulation. Within these profiles the hatched area represents the 1 nm interfacial region that has been defined using the same method as in Chapter 5. The liquid tridecane interface results in an interfacial concentration of 344 g/kg, the frozen, amorphous tridecane results in an interfacial concentration of 442 g/kg.

6.3.5 Tuning the interfacial concentration

A range of model interface parameters will now be explored to show that they can be used to tune the interfacial concentration of glycine. We have carried out simulations of the 296.7 g/kg 3.1 nm film in contact with LJ 9-3 walls for each combination of ϵ_{ww} values 1, 2.4, 5, 10, 15 and 20 kcal mol^{-1} and σ_{ww} values of 0.17, 0.34 and 0.51 nm. This parameter range encapsulates the parameters of the three materials parameterised in section 6.3.2.

Solution Structure

The density profiles for the full grid of simulations are shown in Figure 6.12. As was previously seen in section 6.3.2, there is significant layering of the solution at the LJ 9-3 interface. As ϵ_{ww} increases the layering of glycine at the interface increases significantly. Increasing σ_{ww}, also results in an increase in the interfacial concentration of glycine, however this effect is much less pronounced than for ϵ_{ww}. For values of ϵ_{ww} greater than 2.4 kcal mol^{-1}, a highly dense layer of glycine forms at the interface, and at the highest

values of ϵ_{ww} the amount of water at the interface is reduced as the glycine becomes more concentration. There are only small increases in the interfacial density of glycine for values of ϵ_{ww} above 10 kcal/mol, however this is due a complete depletion of glycine within the remainder of the solution film. The results obtained here are therefore effected by finite size effects, and it is likely that the concentration enhancement effect would be greater for larger solution films. This will be explored further in section 6.3.6.

It is clear that the strength of the dispersion interactions between the interface and the solution significantly impacts the solution composition within the interfacial region. This suggests that during heterogeneous nucleation the overall solution concentration is not necessarily representative of the environment in which nucleation is occurring. This is a key insight when comparing nucleation rates in different experimental setups, as experiments with the same overall concentration may have significantly different interfacial concentrations. This means that the heterogeneous nucleation rates may vary significantly for experiments where the interfaces present differ significantly, such as oils in microfluidic experiments, glass vials, or polymer microwells, tubings and stirrers.

The orientation of molecules in the interfacial regions has been analysed by calculating the bond orientation parameter, P_2, as a function of distance from the interface. The full set of P_2 profiles are shown in Figure 6.13. The P_2 profiles show similar behaviour across the full range of parameters. Similarly to that observed for the previous simulations, there is a strong orientation of the glycine molecules at the LJ 9-3 interface with the C-C bond tending to parallel at the interface. Within the second interfacial layer there is a weaker orientation, also tending towards parallel, before the values return to their bulk-like value of 0 within the centre of the film. For values of $\epsilon_{ww} = 1$ or 2.4 kcal/mol there is less definition in the peaks of the P_2 profile between the interfacial layers. For ϵ_{ww} values 5 kcal/mol or more there is a small increase above $P_2 = 0$ between the interfacial layers. This enhanced orientation at the interface may contribute to increased nucleation, however

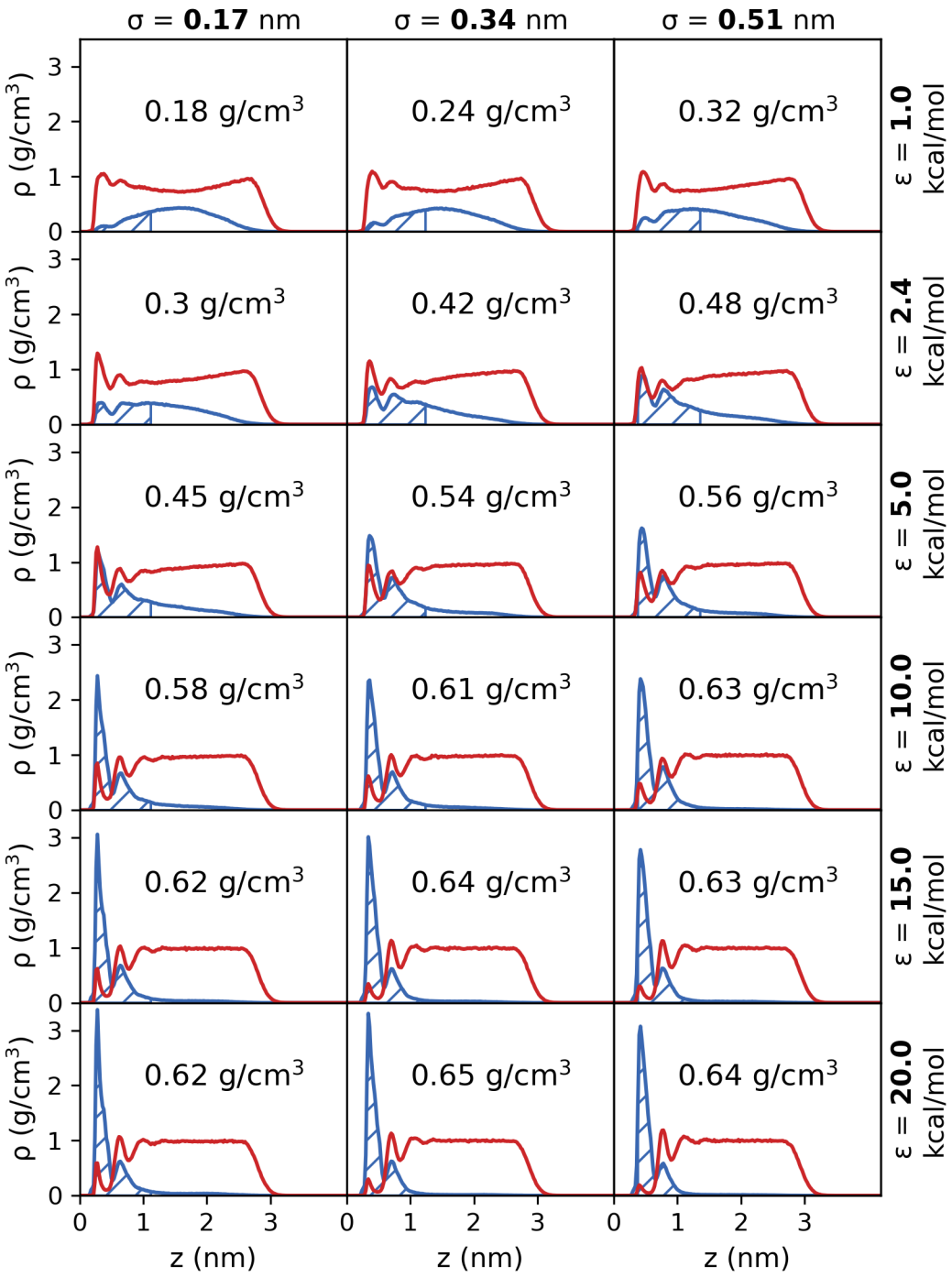

Figure 6.12: Density profiles for glycine (blue) and water (red) for solutions in contact with a LJ 9-3 wall with varying ϵ and σ parameters. The hatched area under the glycine profile indicate the 1 nm region defined as interfacial. The average glycine density within the interfacial region is given for each profile.

it appears to be present at all of the interfaces, regardless of the strength of the interaction. This is likely a steric or packing effect, rather than due to an inherent property of a given interface.

This behaviour of glycine near the interface is similar to the behaviour of physical adsorption of glycine on graphite through evaporation from ethanol that is reported within the literature. It was found that multiple layers of glycine formed initially on the surface, but over the course of two hours after the evaporation of ethanol, the overlayers of glycine evaporated under vacuum until two highly ordered layers of glycine remained [128]. This shows interesting parallels to the multilayer adsorption of glycine simulated here, however it is important to note that the ordering of the glycine molecules within the double layer formed experimentally was perpendicular to the interface in contrast to the parallel formation reported here.

Dynamics

The translational and rotational mobility of the glycine molecules within the first two layers that form at the interface has also been examined. The parallel movement has been investigated using the MSD in the x and y directions of the molecules while they remain within each of the interfacial layers. The parallel MSDs for the grid of simulations are shown in Figure 6.14. For lower wall strengths the molecules within the first layer have greater parallel mobility than those within the second layer, with a crossover in this order at wall parameters $\epsilon_{ww} = 5$ kcal/mol and $\sigma_{ww} = 0.51$ nm. As the wall strength increases, and the interfacial region becomes more concentrated, we see the mobility of the molecules within interfacial layers decreasing, as would be expected for a region of higher concentration [122]. The parallel mobility decreases up to ϵ_{ww} values of 10 kcal/mol, from which point onwards it remains fairly stable. This is due to the concentration of the two

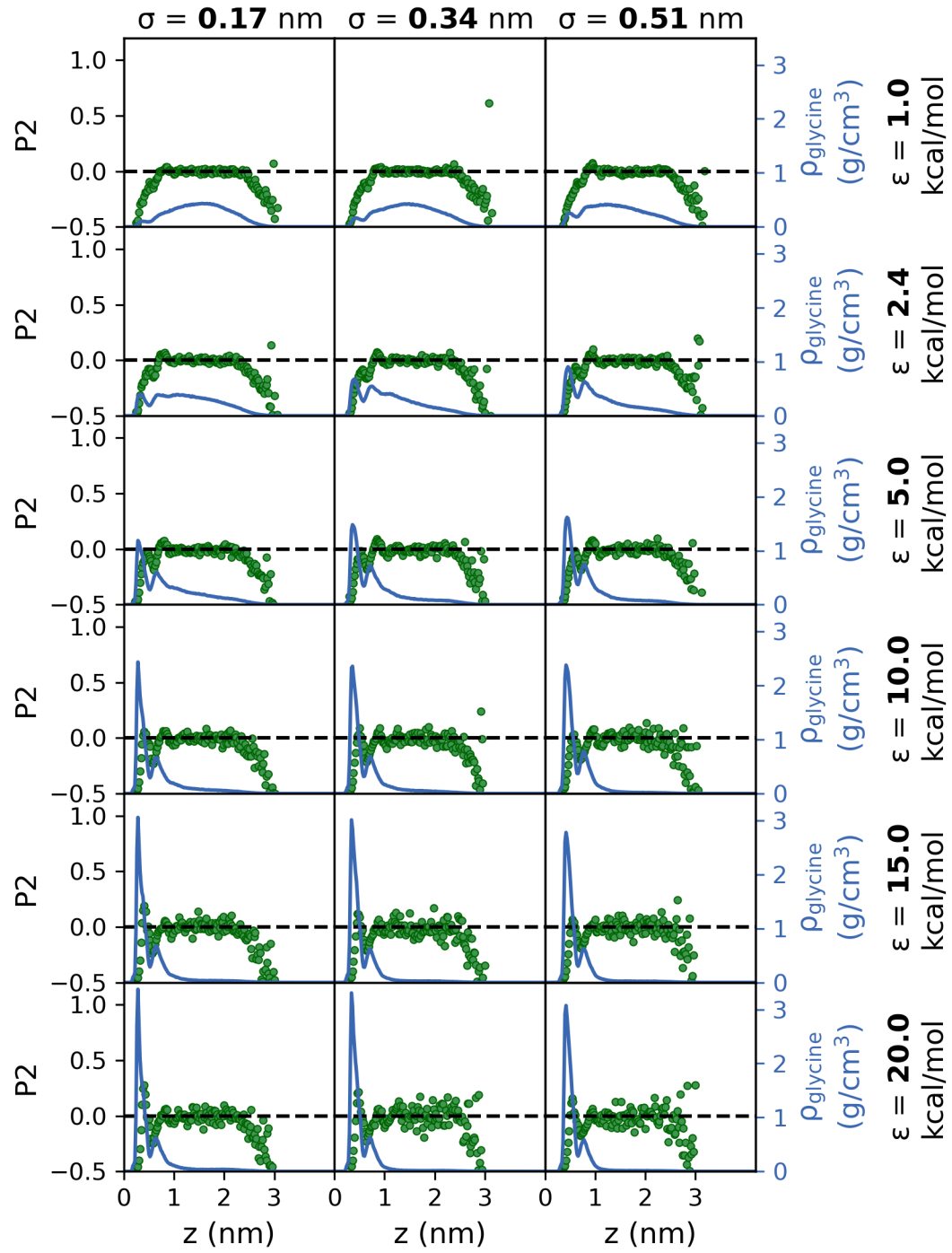

Figure 6.13: The bond order profiles for the glycine C-C bond within solutions in contact with a LJ 9-3 wall with varying ϵ and σ parameters. The glycine density profiles are shown on the secondary axis to demonstrate the alignment of the glycine molecules within each interfacial layer.

layers remaining largely stable across the higher wall strengths as the remainder of the solution is depleted of glycine. It is important to note that due to the short duration of time that the glycine molecules remain within the first interfacial layer at lower wall strengths, and the second layer across all wall parameters investigated, it is difficult to make any quantitative comments about the diffusion parallel to the interface, but the qualitative effects observed match our expectations.

Analysis of the residence time of the glycine molecules within the interfacial layers provides insight into the mobility of the molecules perpendicular to the interface. The distributions of residence times are shown in Figure 6.15. It can be seen that the molecules remain very mobile in the perpendicular direction within the second layer, with only a slight reduction in the mobility at the higher wall strengths ($\epsilon_{ww} \geq 10$ kcal/mol). Within the first layer we see a greater reduction in the perpendicular mobility, however we note that even at the higher wall strengths the molecules remain reasonably mobile, with a majority of molecules having a residence time of 0.5 ns or less.

Finally, we the rotational mobility of the glycine molecules has been examined using the ACF of the C-C bond within the glycine molecules, which are shown in Figure 6.16. In this case we see that across the full range of wall parameters there is a reduction in the rotational mobility of the glycine molecules when compared to bulk solution. For $\epsilon_{ww} =$ 1 kcal/mol the lack of data for molecules within the interfacial layers makes it challenging to compare the ACF with that of the bulk, similarly to the issues with the parallel MSD, but the limited data does suggest that even at the weakest wall strengths the rotational mobility is reduced. This is an expected result when the ordered nature of the interfacial layers that is revealed by the P_2 profiles is taken into consideration.

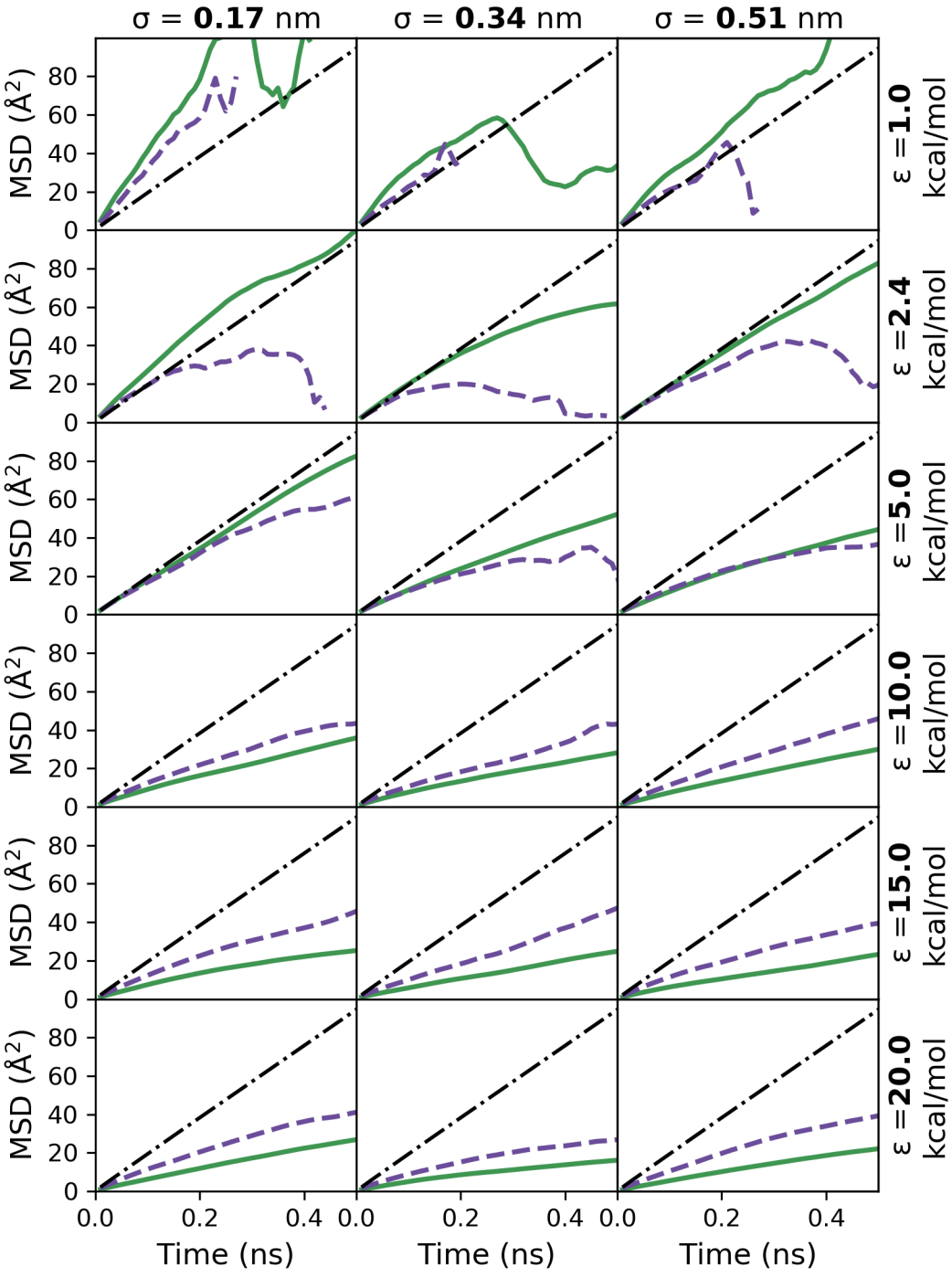

Figure 6.14: Mean squared displacement of the glycine molecules in the x-y directions (parallel to the interface) for solutions in contact with a LJ 9-3 wall with varying ϵ and σ parameters. The green solid, purple dashed and black dot-dashed lines represent molecules within the layer closest to the interface, molecules within the second layer and molecules within a homogeneous solution of the same concentration respectively.

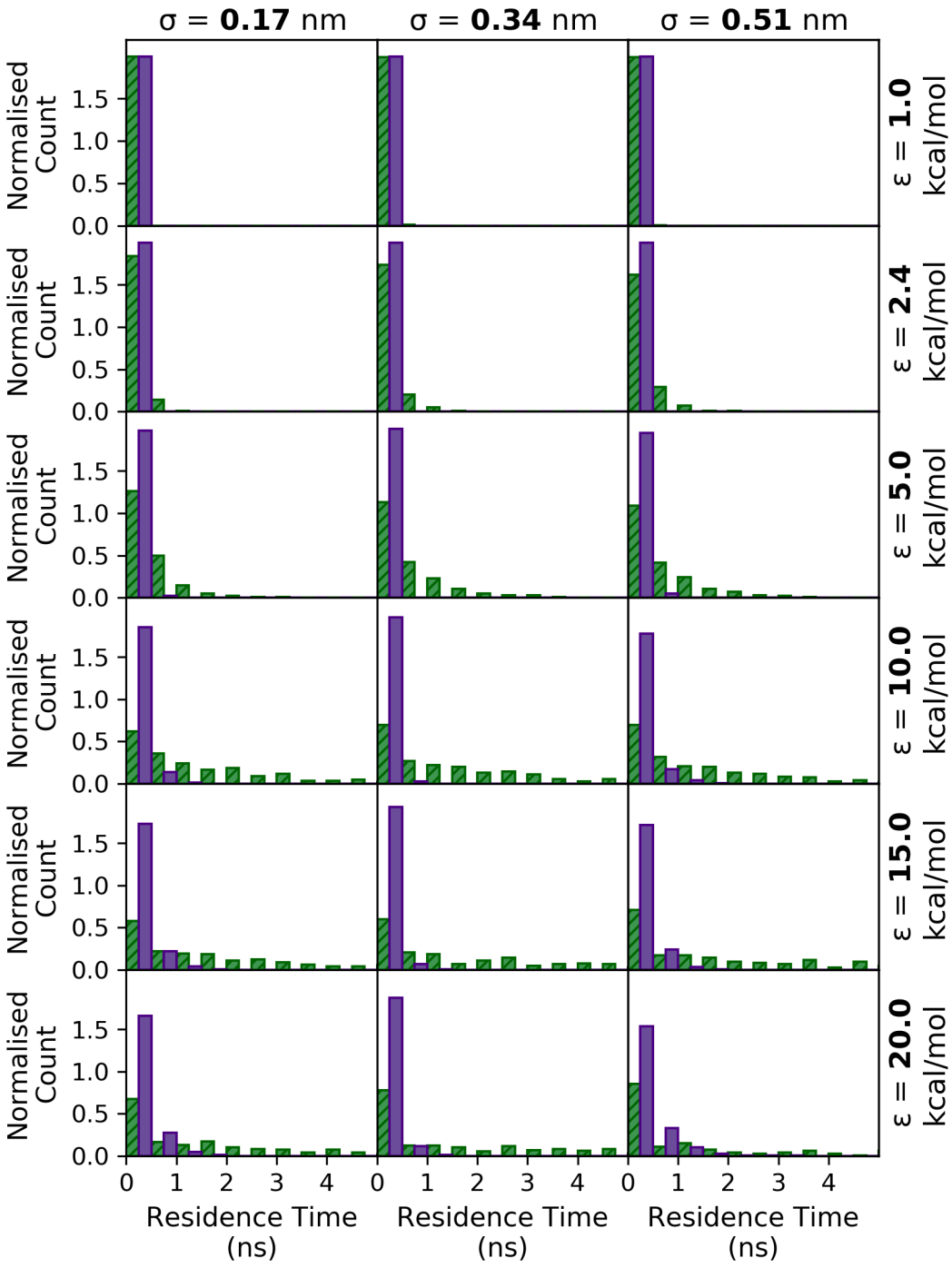

Figure 6.15: Distributions of residence times of glycine molecules within the layer closest to the interface (green hatched bars) and second layer (purple bars) for solutions in contact with a LJ 9-3 wall with varying ϵ and σ parameters.

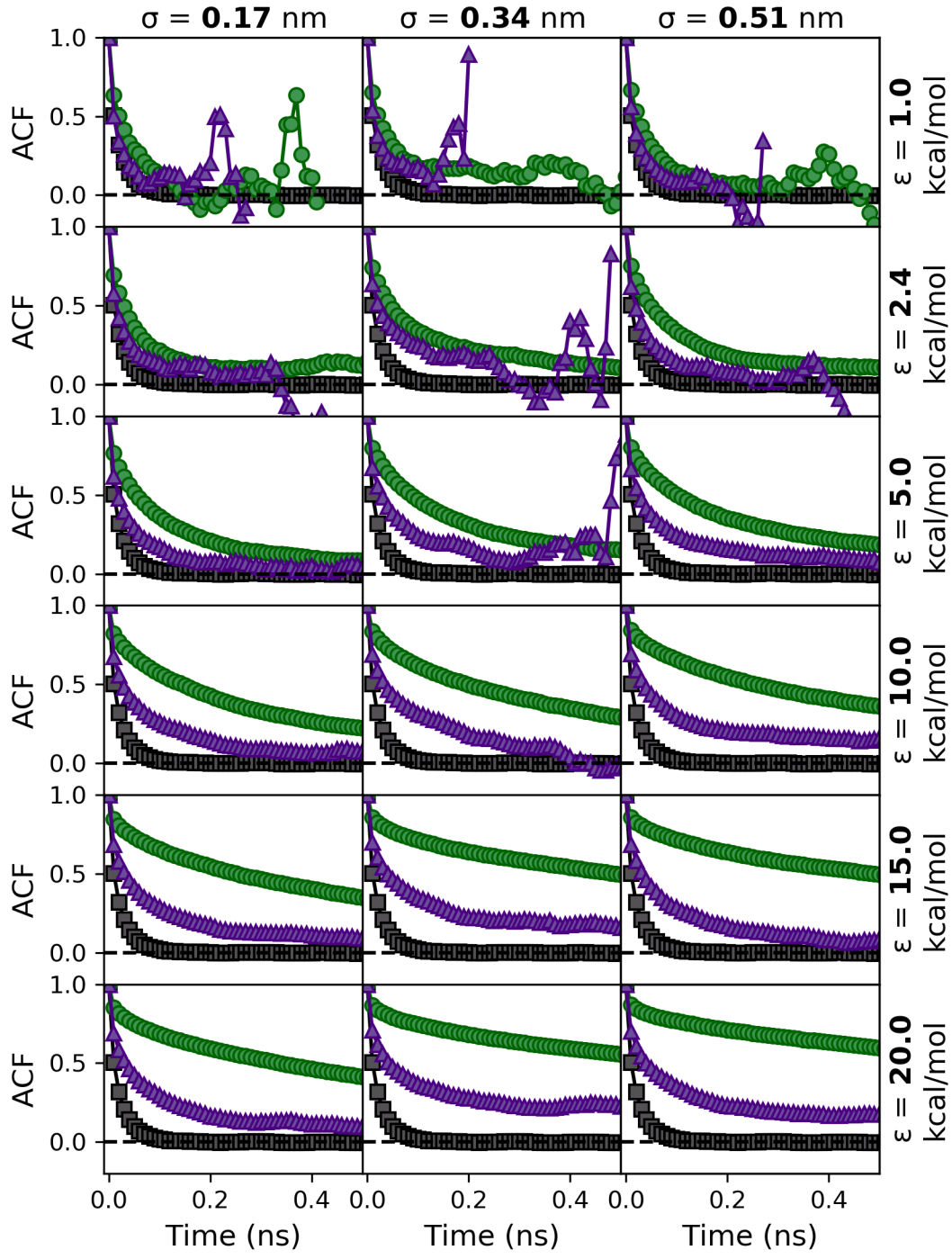

Figure 6.16: Autocorrelation functions of the glycine C-C bonds for solutions in contact with a LJ 9-3 wall with varying ϵ and σ parameters. The green circles, purple triangles and black squares represent molecules within the layer closest to the interface, molecules within the second layer and molecules within a homogeneous solution of the same concentration respectively.

Interfacial Densities and Comparison with Real Materials

Figure 6.17 shows the interfacial densities obtained for each combination of the wall parameters. Here the strong influence of ϵ_{ww} on the interfacial density is clear, with the values plateauing beyond an ϵ_{ww} of 10 kcal/mol as the bulk of the film becomes increasingly depleted. The interfacial densities obtained using the walls parameterised for heptane, tridecane and graphite are also included and can be seen to follow the trends set by the other points. This means that the interfacial densities induced by other materials can be estimated from the parameters obtained, allowing materials to be screened at a greatly reduced computational cost. We have fit the obtained values from the simulation grid to an asymptotic regression model:

$$\rho_I = a - (a - \rho_{vac}) \cdot \exp(-c \cdot \sigma_{ww} \cdot \epsilon_{ww}) \tag{6.3}$$

where ρ_I is the interfacial density, a is the maximum interfacial density for the current film size, ρ_{vac} is the interfacial density for the glycine solution film in contact with vacuum and c is proportional to the rate at which the interfacial density increases with ϵ. The resulting functions for σ_{ww} = 0.17, 0.34 and 0.51 nm are shown as the dotted lines in Figure 6.17. From the fit to the simulation grid data we obtain a and c parameters of 0.63 g/cm^3 and 0.13 mol/kcal $\cdot$ nm. We can see that we obtain a good fit for $\sigma_{ww} = 0.17$ nm and 0.34 nm, however the model deviates slightly at the point where the interfacial densities begin to plateau for $\sigma_{ww} = 0.51$ nm.

Table 6.3 shows the obtained interfacial densities for each of the LJ 9-3 walls that represent the real materials and the values predicted by the fit to equation 6.3. The model does underpredict the interfacial density for each of the materials, however it does provide a reasonable estimate. While the parameters obtained from this fitting will only apply

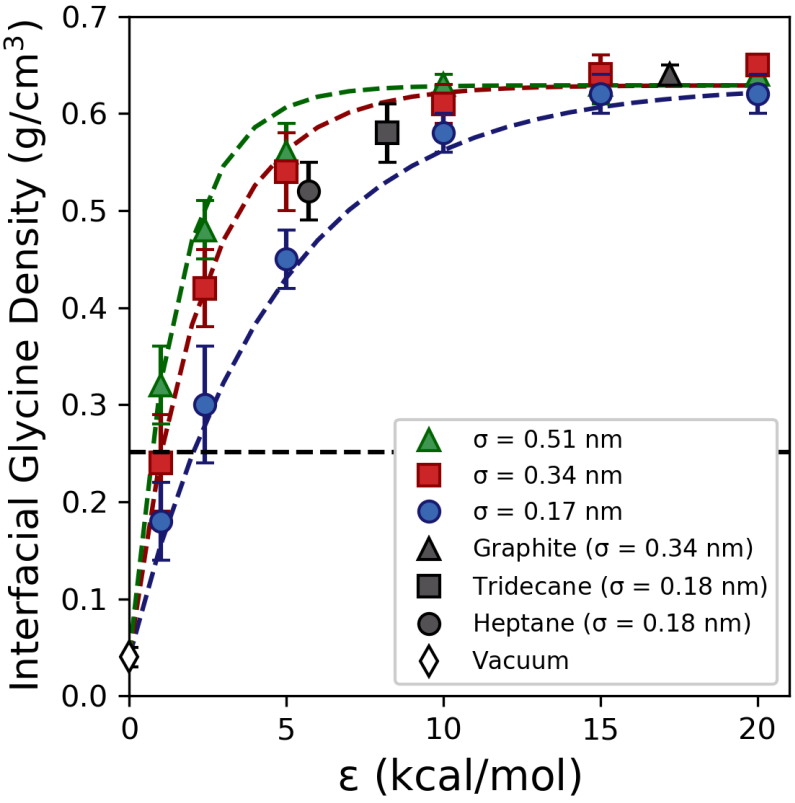

Figure 6.17: Interfacial glycine densities of the LJ 9-3 walls fitted to the real materials (black markers) plotted with the densities of the grid of LJ 9-3 parameter simulations (coloured markers). Coloured dashed lines are fits to equation 6.3 for σ_{ww} values of 0.17, 0.34 and 0.51 nm. The dashed black line shows the density of glycine for a bulk-like solution of the same concentration.

to glycine solutions films of this specific thickness and concentration, the accuracy of this simple model demonstrates that the strength of the dispersion interactions has a consistent and predictable effect on the solution composition at the interface. This will allow the effect of dispersion interactions on the interfacial composition of glycine solutions to be estimated using only the LJ parameterisation process at a significantly reduced computational cost compared to a full MD simulation of a solution film.

From these results we can predict that all three materials will enhance nucleation, compared to bulk solution, with graphite having the strongest effect. The effect of graphene

Table 6.3: Interfacial densities obtained for the real materials for various interface types and predicted by equation 6.3

Material	Atomistic Interface	LJ 9-3 Interface	Predicted Value
Heptane	-	0.52 +/- 0.03	0.47
Tridecane	0.56 +/- 0.04	0.58 +/- 0.03	0.54
Graphite	-	0.64 +/- 0.01	0.63

as both an additive and a substrate on the nucleation behaviour of glycine in small droplets has previously been investigated by Boyes *et al.* [129]. They found that while the presence of graphene as an additive within the droplet resulted in an decrease in the nucleation rate, graphene as a substrate resulted in an increase in the nucleation rate. It is important to note that in the graphene additive experiments a solvent mixture of water and ispropyl alcohol (IPA) is used, to stabilise the graphene dispersion, complicating the balance of solute-interface and solvent-interface interactions. Due to its lower polarity than glycine and water it is possible that IPA has more favourable interactions with graphene. Boyes hypothesises that the increase in the nucleation rate for the graphene substrate experiments, where the solvent is water only, is due to hydrogen bonding between oxidised graphene and glycine, however these results suggest that it may be due adsorption of glycine at the interface induced by dispersion interactions.

The interfacial densities in Figure 6.17 show a clear plateau for high values of ϵ_{ww}. We would expect a similar plateau at higher values of ϵ as the interface becomes saturated with glycine, however from the density profiles we can see depletion of glycine within the bulk of the film. This suggests that the film size within these simulations is limiting the interfacial density, rather than saturation of the interface. In real, bulk solutions, where there is an effectively infinite reservoir of glycine molecules, the interfacial concentration may be much greater than is observed here.

6.3.6 Effect of System Size

As discussed in the previous section, the results of these simulations will be influenced by finite size effects due to depletion of the bulk of the solution film. It is expected that the interfacial concentration enhancement is hindered by this depletion of the solution, and that if more glycine molecules were available within the system that there may be even greater enhancement of concentration at the interface. To confirm this, simulations have been performed with a larger number of glycine molecules. This can be achieved by increasing the overall concentration of the glycine solution, or increasing the thickness of the glycine solution films. The graphite LJ 9-3 wall was selected as a test system, as this is the interface representing a real material with the highest interfacial interactions, and as such, the most severe depletion of the solution film. Solution films of varying thicknesses at concentrations of 296.7 and 500.7 g/kg (all solution films listed in Table 6.1) were simulated in contact with the graphite LJ 9-3 wall to determine the effects of the system size on the interfacial solution composition. The density profiles for each of these simulations are presented in Figure 6.18.

The density profiles obtained for the thicker film sizes and higher concentrations show that, as expected, increasing the total amount of glycine within the solution results in a higher density of glycine at the interface. The first glycine density peak is relatively stable, with only as modest increase as the system size increases, which is likely due to saturation of the first layer. There is a significant increase in the second density peak, however it is likely that this layer would also eventually reach a saturated value and become stable, provided enough glycine molecules were included within the system. Interestingly, as the system size increases a third peak begins to form, with a small fourth peak forming at the thickest solution film. There is a similar effect with increasing concentration, and in the 500.7 g/kg, 12 nm film the dense glycine region extends far beyond the 1 nm interfacial region that has previously been defined. This extended dense region does not appear to

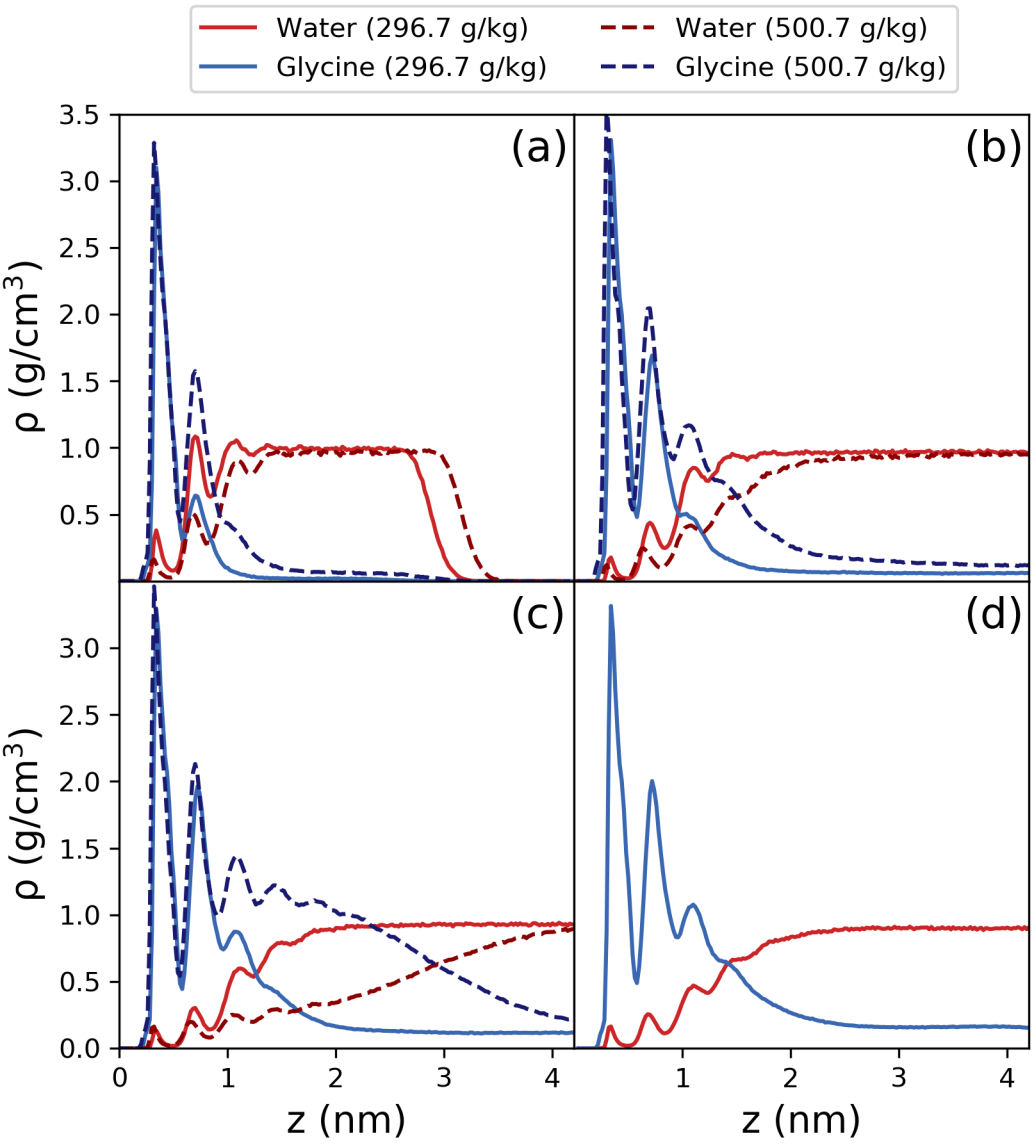

Figure 6.18: Density profiles for glycine and water for solutions of thickness (a) ≈ 3 nm, (b) ≈ 6 nm, (c) ≈ 12 nm and (d) ≈ 14 nm in contact with the LJ 9-3 walls parameterised to represent graphite.

be as structured as the well defined peaks that are present directly at the interface. The P_2 profile, shown in Figure 6.19, reveals that there is not significant ordering within this dense region, beyond the initial four peaks, with only a slight tendency for the molecules to arrange parallel with the interface. This is a dense and disordered region of concentrated glycine, similar to the pre-nucleation clusters that are a part of two-step nucleation theories.

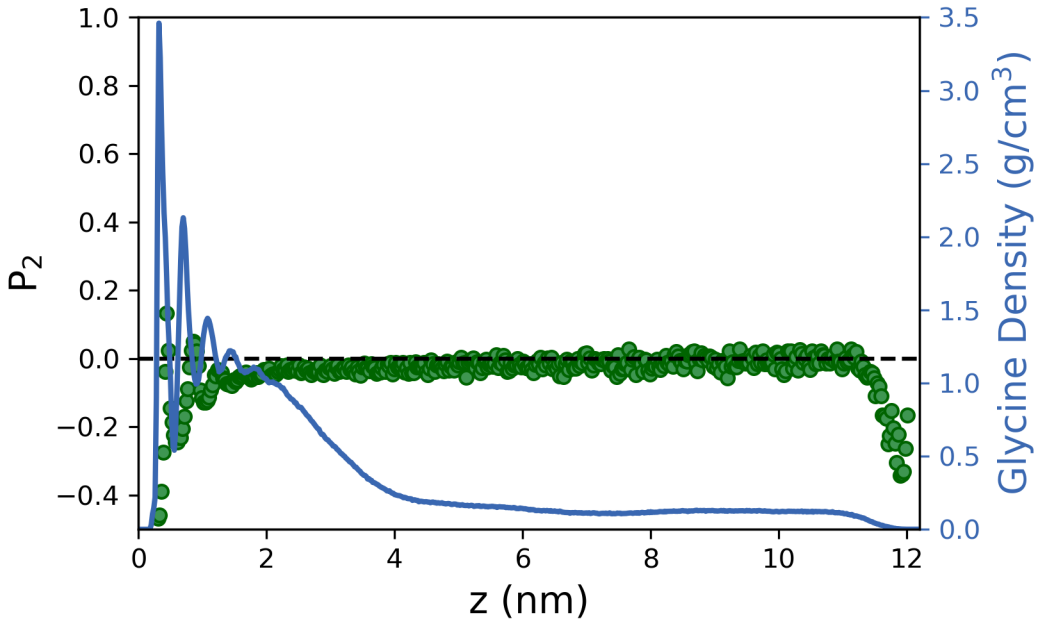

Figure 6.19: The bond order profile obtained for the 500.7 g/kg, 12 nm glycine solution film in contact with the LJ 9-3 wall representing graphite. The density profile of the glycine solution is plotted on the secondary axis.

From the density profiles in Figure 6.18, it is clear that the fixed 1 nm interfacial region is no longer suitable for quantifying the effects of the interface on the solution composition. It is therefore necessary to find a method to consistently define the interfacial region for films of different sizes. It would be expected that as you move away from the interfacial region into the bulk of the film the density will reach a constant value as it reaches a bulk region. The derivative of the density profile will reach a value of 0 when it reaches this constant value and can therefore be used to determine the interfacial region. As there is some noise within the density profiles, a Savitzky-Golay filter [130], that smooths the data by fitting a polynomial to a moving window throughout the data, was applied with third order polynomials and window size of five. An example of a smoothed density profile is shown in Figure 6.20 (a). The filter accurately captures the density profile, but it does slightly understate the peak closes to the interface. However, for determining the interfacial region it is important that the tail down to the bulk-like region is accurately captured, which the Savitzky-Golay filter achieves successfully.

To define the interfacial region, the derivative of the smoothed density profile is taken. The interfacial region is defined as the first point that the derivative of the density profile crosses the x-axis after the region containing the density peaks at the interface, as shown in Figure 6.20 (b).

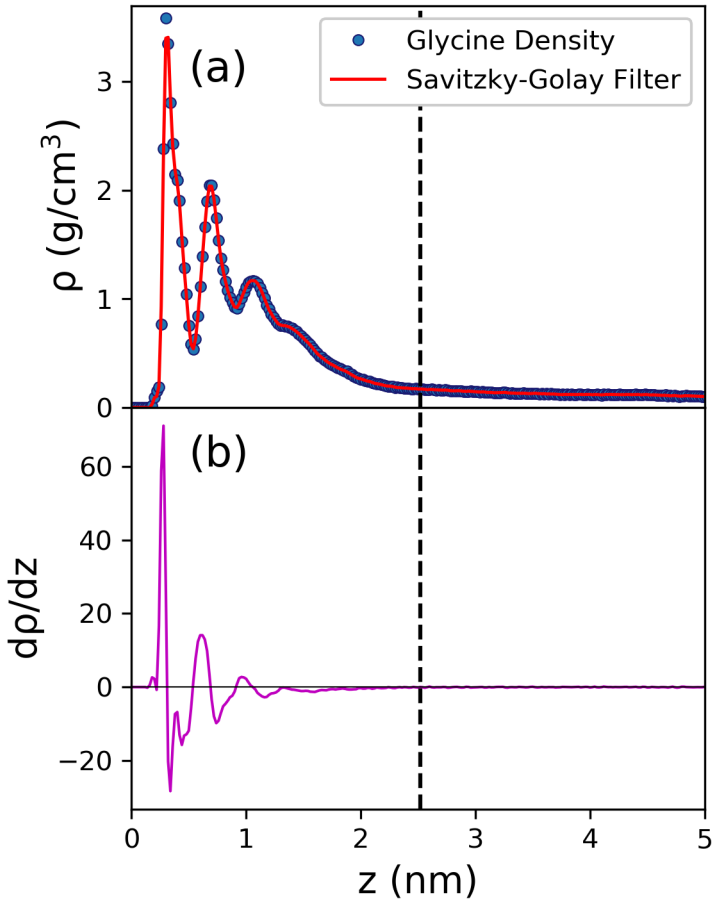

Figure 6.20: (a) Example of the Savitzky-Golay filter used to smooth the glycine density profile of the 500.7 g/kg, 6.1 nm film prior to taking the derivative. (b) The derivative of the smoothed density profile. The dashed black line represents the point where the derivative crosses the x-axis after the initial peaked area, defining the interfacial region.

This process was repeated for each of the films simulated in contact with the LJ 9-3 wall representing graphite. The resulting interfacial glycine densities for each film are shown in Figure 6.21. As expected, the interfacial density increases with both the solution thickness and concentration, demonstrating that depletion of the solution is limiting

the concentration enhancement effect observed within the simulations. In Figure 6.21 the interfacial densities are plotted against the inverse of solution film thickness. This allows for extrapolation back to zero, or a film of infinite thickness, to estimate the interfacial density for a bulk-like solution that would exist under experimental conditions. The data points appear to be fairly linear with the inverse of solution thickness, and a linear fit predicts interfacial densities of 0.81 and 0.97 g/cm^3 for films of infinite size at concentrations of 296.7 and 500.7 g/kg respectively. It is unclear if the data will continue linearly beyond the area where the current simulations have been performed, however, there is a major computational cost required to push further towards 1/Thickness = 0, and as such, testing of this assumption is challenging. Despite this, these results demonstrate that the concentration enhancement effect in larger systems is greater than is observed here.

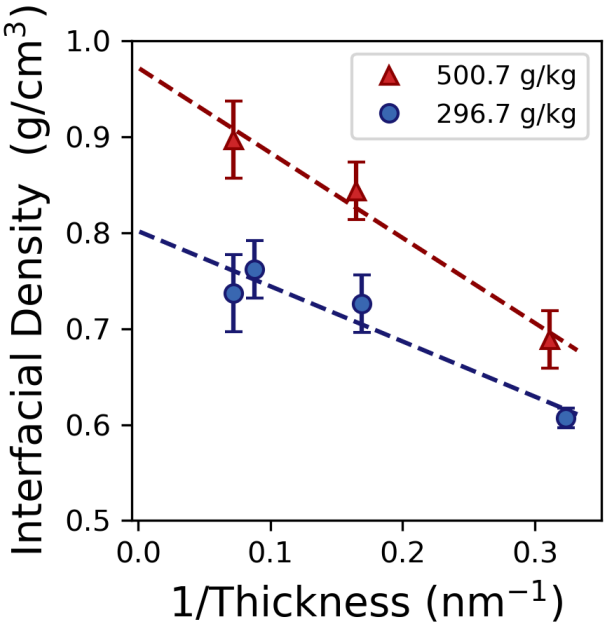

Figure 6.21: Interfacial glycine densities obtained for glycine solution films of varying thickness at 296.7 and 500.7 g/kg in contact with the LJ 9-3 wall representing graphite plotted against the inverse of the solution film thickness. Extrapolating back to zero will represent a solution film of infinite thickness which is representative of a true bulk solution.

6.3.7 Reversibility of Solute Adsorption

The formation of the highly dense, amorphous region within the simulation of the 500.7 g/kg, 12.2 nm film in contact with the LJ 9-3 wall representing graphite, shown in Figure 6.18 (c), was unexpected. Therefore, efforts were undertaken to ensure this was a reversible process, and that a permanent phase separation due to force field limitations had not occurred.

Of the three materials that have been studied, the graphite interface has the strongest interactions (ϵ_{ww} = 17.2 kcal/mol, σ_{ww} = 0.34 nm) and heptane has the weakest (ϵ_{ww} = 5.7 kcal/mol, σ_{ww} = 0.17 nm). In order to detrmine if similar solution behaviour would occur at an interface with weaker interactions, the 500.7 g/kg, 12.2 nm film was simulated in contact with the LJ 9-3 wall representing heptane. The density profiles obtained for the 500.7 g/kg, 12.2 nm film in contact with the graphite and heptane LJ 9-3 walls are shown in Figure 6.22 (a) and (c) respectively.

For the solution film in contact with the LJ 9-3 wall representing heptane, it can be seen that there is once again a dense region forming at the interface that extends beyond the initial layering, and tails off from a dense, amorphous region towards the bulk values of the film. In this instance, it can be seen that the heights of the two peaks closest to the interface are significantly lower for heptane (2.25 and 1.72 g/cm^3) than those obtained for graphite (3.46 and 2.13 g/cm^3). Despite the difference in peak heights between the two density profiles, the average interfacial densities obtained using the method described in section 6.3.6 are very similar for the two simulations. The interfacial density values are sensitive to the defined limits, and the values obtained for these simulations are influenced by the long tails from the interfacial region to the bulk of the film. This highlights that the interfacial density alone may not fully capture differences in solution behaviour at various interfaces, and that a more detailed analysis may be required to provide a more complete image.

The final configuration of the solution film in contact with the LJ 9-3 wall representing graphite was then placed in contact with the heptane LJ wall. If the process is reversible, replacing the walls in this manner should result in a reduction in the interfacial density towards the values obtained from the heptane LJ wall simulation. The solution film was simulated for a further 160 ns, and averaging over the final 80 ns of this simulation resulted in the density profile presented in Figure 6.22 (b).

Once again, similar values are obtained for the average glycine density within the interfacial region due to the issues discussed above. Despite this, there is excellent agreement between the peak heights obtained from the simulations using the heptane interface, which are significantly lower than those obtained at graphite. The heights of the peaks for the first, second and third interfacial layers for each simulation are provided in Table 6.4, demonstrating that the concentration enhancement at the interface can revert to lower values if a strongly interacting interface is replaced by a weaker one.

Table 6.4: Height of the density peaks for the first, second and third interfacial layers obtained for the 500.7 g/kg, 12.2 nm film in contact with the LJ 9-3 walls representing graphite and heptane.

System	**1st Peak Height** (g/cm^3)	**2nd Peak Height** (g/cm^3)	**3rd Peak Height** (g/cm^3)
(a) Graphite	3.46	2.13	1.44
(b) Graphite → Heptane	2.25	1.72	1.33
(c) Heptane	2.14	1.64	1.26

To further confirm that there is not a permanent phase separation occurring, a further simulation was performed where the wall was removed from the simulation. This simulation was performed on the final, well-equilibrated configuration of the 500.7 g/kg, 12.2 nm solution film in contact with the LJ 9-3 wall representing graphite. The LJ wall was removed from the simulation and the simulation box was made fully periodic in the x, y and z directions. As vacuum is known to repel glycine, the simulation was performed in the NPT ensemble, with the box allowed to adjust size in the z direction perpendicular

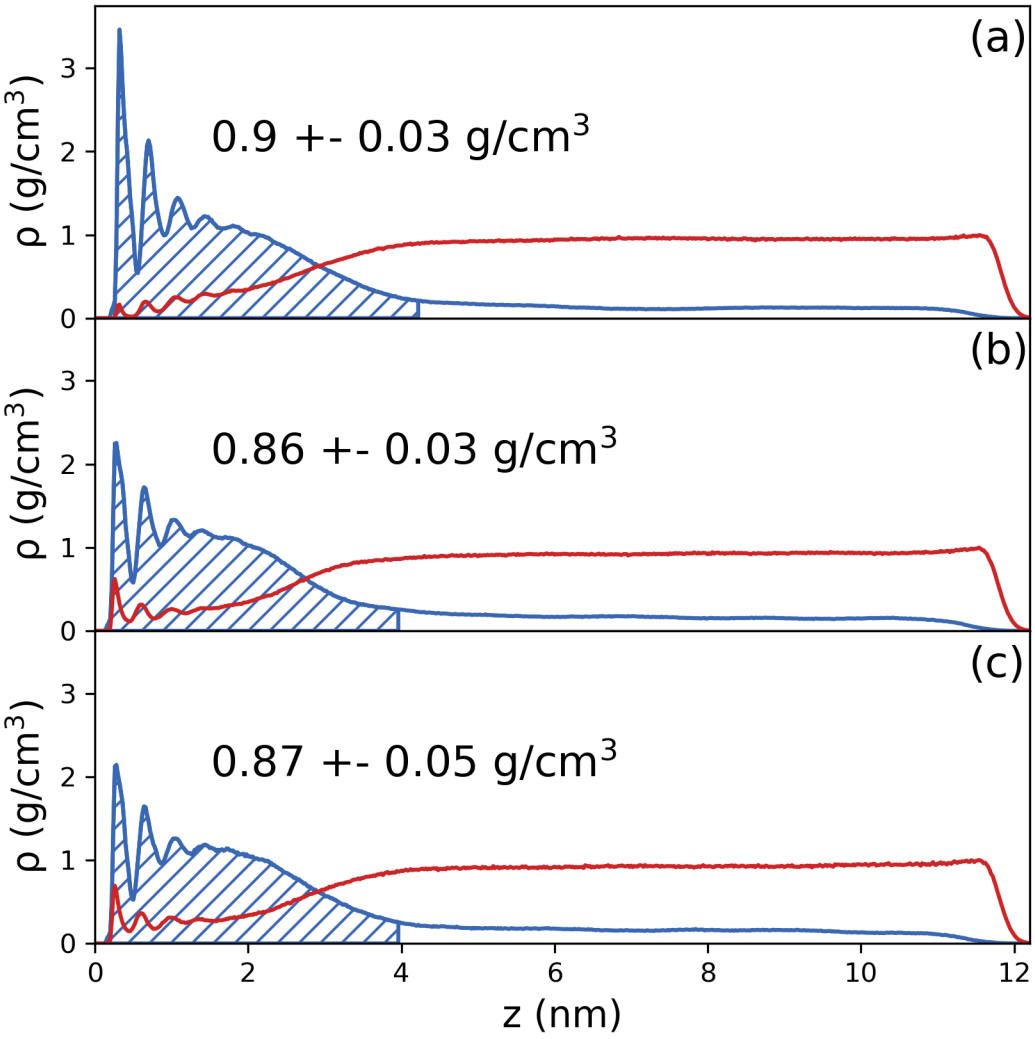

Figure 6.22: Density profiles obtained for glycine (blue) and water (red) from glycine solution films in contact with the (a) LJ 9-3 wall representing graphite and (c) the LJ 9-3 wall representing heptane. (b) shows the density profile obtained from simulating the final configuration from (a) in contact with the LJ 9-3 wall representing heptane. The hatched areas represent the regions defined as interfacial. The average glycine density within the interfacial area is shown on each plot, and the standard deviation of the average interfacial density with time is given as the error.

to the previous interface, to ensure that the repulsion from the vacuum interface was not artificially breaking up the dense glycine regions. The system was simulated for 130 ns in the NPT ensemble to determine if the concentration enhancement was reversible. If the enhancement was reversible, the dense, ordered region of glycine that had formed at the interface would be expected to break up, and redistribute throughout the remainder of the solution.

Figure 6.23 shows snapshots of the solution film throughout the simulation. Figure 6.23 (a) shows the initial configuration of the simulation, with the LJ 9-3 wall representing glycine on one side of the solution film, and vacuum on the other. Once the LJ wall is removed, and the system is simulated in NPT, the simulation box contracts in the z direction to form a bulk-like system within 0.003 ns. Figure 6.23 (b) shows the system at $t = 0.003$ ns to demonstrate that the dense, ordered region has been maintained for the time that was taken for the simulation box to reduce, removing the vacuum layer from the system. This dense region is crossing the periodic boundary in the z direction and so is split between the left and right hand sides of the snapshot. Figure 6.23 (c) shows a snapshot of the final configuration of the system after 130 ns of NPT simulation. While some dense regions remain within the glycine solution, the dense, ordered region is clearly breaking up and redistributing throughout the solution, demonstrating that the enhancement process is reversible and not a permanent phase separation. It is important to note that even after 130 ns of NPT simulation, a fully homogeneous solution has still not been obtained. The long simulation times that are necessary to completely redistribute the glycine molecules throughout the solution are consistent with the concentration fluctuations present after long simulation times (200 ns) within the oil–solution interface simulations reported in Chapter 5.

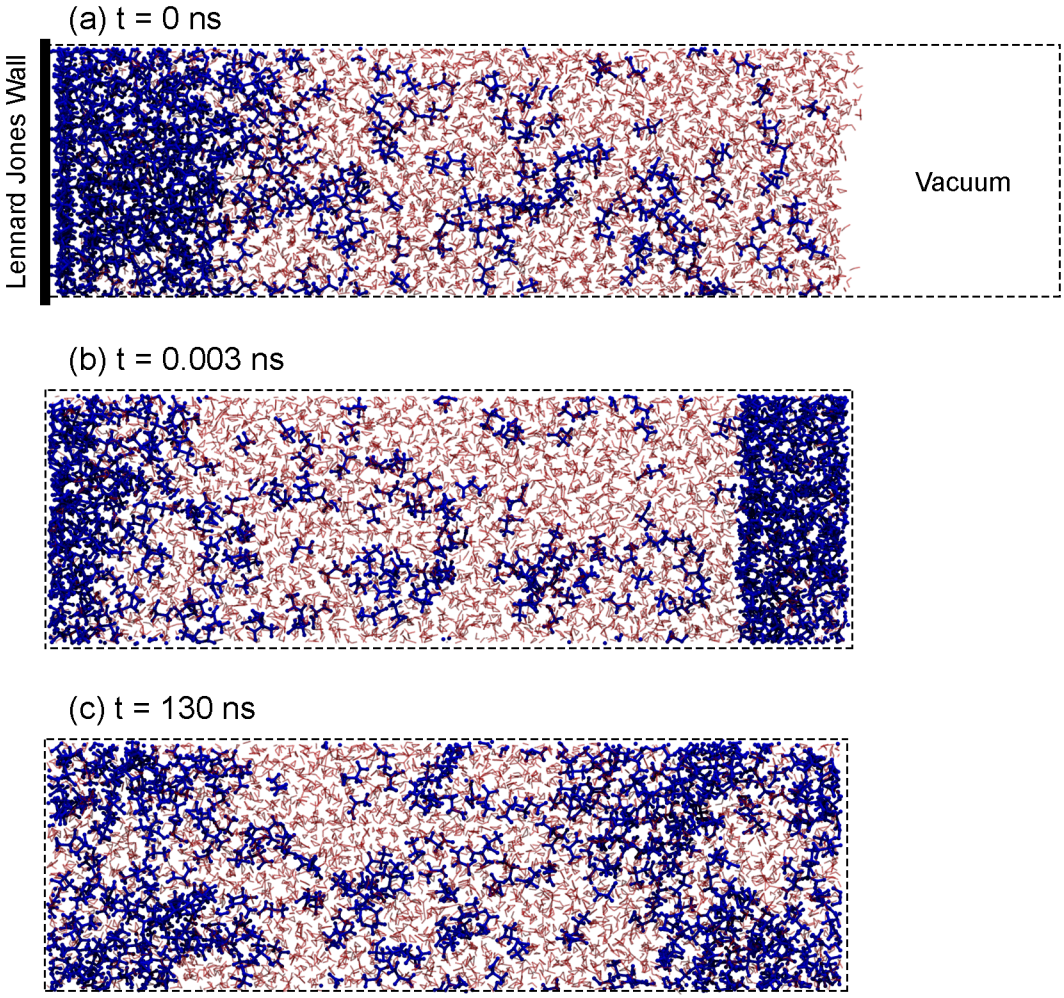

Figure 6.23: Snaphots from the simulation of the glycine solution after the LJ 9-3 wall has been removed. Glycine molecules are coloured blue and water molecules are coloured red. The dashed black lines show the boundaries of the simulation box. Note that in (b) and (c) the system is periodic in the z direction. The dense, layered glycine region in (b) is straddling the periodic boundaries, and so appears on the right hand side of the snapshot.

6.3.8 Effect of Electrostatics

Thus far, only the effects of dispersion interactions between the interface and the solution have been determined. In this section the impacts of electrostatics on the solution composition at the interface will be explored. This has been investigated through simulations of a glycine solution film in contact with atomistic interfaces with various combinations of LJ parameters and charge sets.

PTFE has a similar molecular structure to alkanes, however all hydrogen atoms are replaced with fluorine atoms. Due to the high electronegativity of fluorine, this results in much larger charges present within the molecule than those that are found within alkanes. These attributes made PTFE an attractive choice for this test. As the molecular structure is similar, it is possible to use the same atomic configuration for the two molecules, ensuring that the atomic density is equal between the two interfacial systems. The interface can be adjusted between tridecane and PTFE simply by adjusting the intermolecular potentials.

It is important to note that alongside the difference in electrostatic charges between the PTFE and tridecane molecules, there are also differences in the dispersion interactions for the two materials. For example, the fluorine atoms within PTFE have ϵ_{ii} and σ_{ii} parameters of 0.061 kcal/mol and 0.312 nm respectively, in contrast to the ϵ_{ii} and σ_{ii} parameters of 0.0124 kcal/mol and 0.266 nm for the hydrogens in tridecane (the full set of intermolecular parameters are provided in Tables 3.3–3.5). It is therefore necessary to decouple the effects of the electrostatic interactions and the dispersion interactions on the solution composition at the interface. This has been achieved by separating the LJ parameters and electrostatic charges into separate parameter sets that can be applied to the interface. Two LJ parameter sets have been considered, PTFE and LJ, and three charge sets: tridecane, PTFE and PTFE×2. As this section is interested in the effects of electrostatics on the concentration enhancement effect, the third charge set (PTFE×2)

was included to help determine if there was a trend between the interfacial glycine density with the charges present within the interface. While this third charge set is somewhat artificial, it does allow the effects of electrostatics to be explored further without the need to introduce a new material that would require another set of LJ parameters.

The 296.7 g/kg, 3.1 nm was simulated in contact with an interface with the atomic positions of the tridecane slab, shown in Figure 6.3, with each combination of the LJ parameter and charge sets. The obtained density profiles are presented in Figure 6.24. The left hand side column shows the simulations with the tridecane LJ parameters, while the right hand column shows the simulations with the PTFE LJ parameters. The first, second and third row show the simulations with the tridecane, PTFE and PTFE$\times$2 charges respectively. As such, moving through Figure 6.24 from left to right results in an increase in the LJ interactions, while moving from top to bottom results in an increase in the charges present within the interface.

For all three charge sets, the weaker tridecane LJ parameters have a lower interfacial glycine density when compared to the stronger PTFE LJ parameters. For the tridecane LJ parameters it can be seen that as the charges within the interface increase, the interfacial density of glycine decreases. The decrease in the interfacial density with charge is small, and the average values fall within each others errors, but it can be seen that the peak height of the first interfacial peak also decreases with the increasing charges. It is possible that the electrostatic interaction of the interface material is stronger with the water molecules, and if the dispersion forces with the glycine molecules weaken, the electrostatics increases the water interfacial density. In contrast, for the PTFE LJ parameter simulations there is a small increase in the interfacial glycine density when moving from the tridecane to the PTFE charge set, however all three values for the PTFE LJ parameter set fall within the errors of each other.

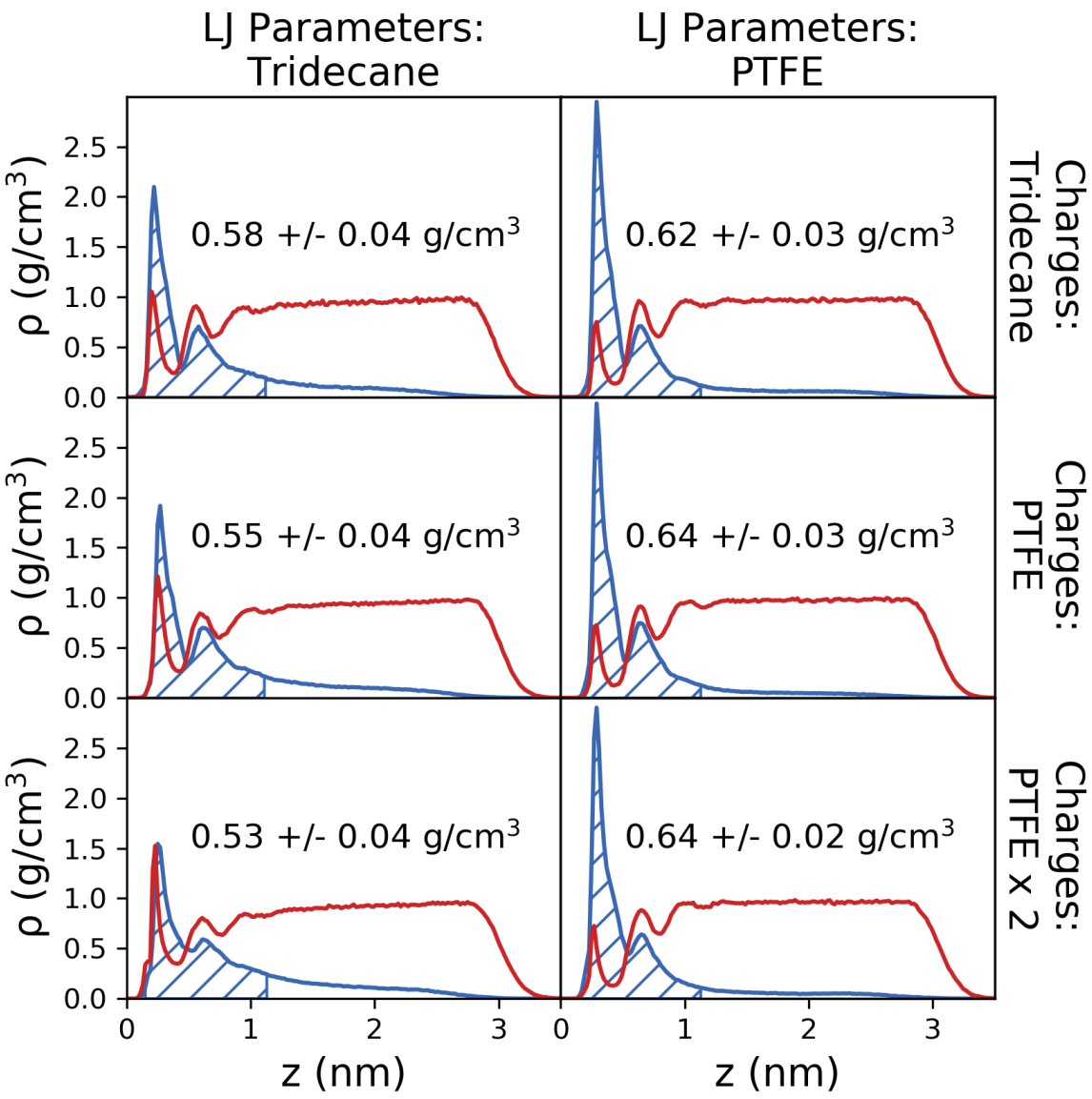

Figure 6.24: Density profiles for the tridecane/PTFE interface simulations. The left and right columns show the simulations with the LJ parameters for tridecane and PTFE respectively. The top, middle and bottom rows show the simulations with tridecane, PTFE, and doubled PTFE charges, respectively.

For these systems, it would appear that as the charges within the interface increase, the interfacial density of glycine decreases. This effect is much weaker than the effect that is observed for varying dispersion interactions, and have a much less significant impact on

the solution composition at the interface. It is possible however, that as the dispersion interactions weaken electrostatics may become more important.

6.4 Conclusions

In this chapter the effects of interface–solution interactions on the solution composition within the interfacial regions has been investigated through the use of model interface. Simulations using the LJ 9-3 walls reveal that the dispersion interactions between the interface and the solution have a major effect on the concentration of glycine at the interface, with the interfacial density of glycine increasing significantly with the strength of the interface–solution interactions.

A novel parameterisation process has been developed that allows for the LJ 9-3 potential to be mapped onto real materials. This parameterisation process has been validated using the tridecane interface, where it was demonstrated that the LJ 9-3 wall is capable of reproducing the solution composition, molecular orientation and dynamics at the interface. This procedure was also applied to heptane and graphite in order to generate suitable LJ 9-3 walls to represent each of these materials.

The interfacial glycine density was found to follow a simple trend with the ϵ_{ww} and σ_{ww} parameters of the LJ 9-3 walls. This allowed a simple model to be produced that can reasonably estimate the interfacial density of glycine for a solution film of the same size and concentration in contact with a LJ 9-3 wall of a given set of ϵ_{ww} and σ_{ww} parameters. This will allow for the effects of new materials on nucleation kinetics to be estimated using the computationally cheap LJ parameterisation process, without the need for expensive MD simulations. This can therefore be used as an effective tool for screening new materials for nucleation enhancement or suppression purposes.

Simulations have been performed at various system sizes and concentrations to demonstrate that the interfacial concentration enhancement effect is likely greater in real, bulk-like systems than is observed within the small scale simulations performed here. Tests have also been carried out to ensure that the concentration enhancement effect is reversible and that a permanent phase separation has not occurred due to force field limitations.

Finally, the influence of electrostatic interactions between the interface and the solution has been investigated and contrasted with the effects of dispersion interactions. While the effects on of the dispersion interactions on the solution composition are much greater than those observed for electrostatics, it is possible that for systems where dispersion interactions are weaker electrostatics may play an important role.

Chapter 7

Conclusions

Crystal nucleation from solution plays an important role in a wide range of environmental, biological and industrial processes. It is well understood that nucleation most often occurs heterogeneously at interfaces, and the aim of this thesis was to develop a better understanding of how various interfaces can influence nucleation behaviour. A better understanding of this process will provide opportunities for more accurate control and prediction of crystallisation processes in the future.

In Chapter 4, the effect of a tridecane interface, an oil interface like those found in microfluidic or microwell type experiments, on the nucleation behaviour of aqueous glycine solutions was investigated. It was shown that the the presence of the oil layer results in a major increase in the nucleation rate of glycine when compared to samples where the oil layer is not included. It was shown that the sample preparation method can also have a significant impact on the observed nucleation kinetics, however it is clear from the results that the nucleation rate of glycine is significantly accelerated in the presence of tridecane regardless of the preparation method used. This was an unexpected result,

as the non-polar, hydrophobic tridecane was not expected to favour the heterogeneous nucleation of the highly polar, hydrophilic glycine. Experimental results for a separate system, isonicotinamide in ethanol, show a similar enhancement in the nucleation rate when in contact with a fluorinated oil, demonstrating that this effect is not limited to aqueous glycine solutions.

The mechanisms that would normally be attributed to the enhanced nucleation rates observed for heterogeneous nucleation, such as physical or chemical templating, would not be expected to apply to the glycine solution–tridecane interface system. To better understand the enhanced nucleation rates observed, classical MD simulations were employed to investigate the solution behaviour at oil–solution and air–solution interfaces. The results of these simulations were reported in Chapter 5 and revealed a stark contrast in the solution composition at each interface, with enhanced and depleted glycine concentrations at the oil and air–solution interfaces respectively. This novel concentration enhancement effect is distinct from previously explored heterogeneous nucleation mechanisms, and it is proposed that this is due to dispersion interactions between the interface and the solution molecules. This effect is likely present in a range of solution–interface systems. Further analysis reveals ordering of the glycine molecules at each of the interfaces, and that the glycine molecules are still highly mobile at the oil–solution interface.

In order to confirm that dispersion interactions are the source of the concentration enhancement effect, model interfaces with tuneable solution–interface interactions were introduced to the MD simulations in Chapter 6. A new methodology was presented that allows Lennard Jones 9-3 walls to be mapped to realistic materials. This computationally inexpensive method was validated against an atomistic tridecane interface and successfully reproduced the solution composition, molecular orientation and dynamic properties at the interface. Simulations using the Lennard Jones 9-3 walls revealed that the strength of the dispersion interactions between the interface and the solution has a significant, but

consistent, effect on the solution composition at the interface. As the solution composition consistently follows a trend with respect to the wall parameters, the effects of new interfacial materials on the nucleation kinetics of glycine can be ranked against previously investigated materials using only the LJ 9-3 parameterisation process, without the need to carry out further computationally expensive MD simulations. Investigations into the effect of system size suggest that the interfacial concentration enhancement is likely greater in real bulk conditions than is observed within simulations. Finally, the effects of electrostatic interactions between the interface and the solution were also found to influence the solution composition at the interface, however these effects were much weaker than those seen for dispersion interactions.

This work has revealed a novel concentration enhancement effect that contributes to the enhanced nucleation rates that are observed for heterogeneous nucleation, and simulations suggest this effect is due to dispersion interactions between the interface and the solution, however, a number of questions remain. It is unclear why non-Poisson behaviour is observed within the nucleation experiments reported in Chapter 4. While the presence of the tridecane layer is certainly necessary for a significant amount of nucleation to occur, regardless of the sample preparation method, it is not obvious why there are what appears to be multiple nucleation regimes occurring within the samples. It is possible that there is some time dependence on the nucleation rate (e.g. due to impurities poisoning the interface after some time, or equilibration effects), but further investigation is necessary to better understand this behaviour.

While the simulations in Chapter 5 demonstrate an enhanced concentration at the oil–solution interface, the true interfacial concentration may be much higher than is currently captured by the analysis. Interfacial fluctuations can smear the high density regions of glycine across multiple bins, and more sophisticated interfacial analysis techniques may reveal that the concentration at the interface is much greater, similar to that seen for

the fixed, flat interfaces used in Chapter 6. Advanced sampling simulation techniques, such as metadynamics, could also be used to compare the pathways for nucleation at the oil–solution interface vs within bulk solution.

In Chapter 6, LJ 9-3 walls were used to predict that two other interfaces, heptane and graphite, would also result in nucleation enhancement of aqueous glycine solutions. Future experiments to quantify the nucleation kinetics at each of these interfaces would allow for this predicted nucleation behaviour to be tested. This would allow for the predictive capabilities of both the LJ 9-3 interfaces and the parameterisation process to be determined.

The simulations in Chapters 5 and 6 have focused on the effects of changing the interface, or the interactions between the interface and the solution. In reality, the concentration enhancement effect is likely a product of the balance between the solute–solvent–interface interactions. A more fundamental understanding of how this balance of interactions influences the solution composition will allow for rational solvent and interface selection for the purposes of nucleation control. It has been demonstrated that the influence of the dispersion interactions between the interface and the solution can be captured by a computationally inexpensive parameterisation process, and if this methodology can be expanded to account for solute–solvent interactions then it can be used as an effective design tool to quickly screen a wide range of interfaces and solvents for this purpose. This would allow for solvents and interfaces to be selected in order to either enhance nucleation in difficult to nucleate systems, or to suppress nucleation for anti-fouling purposes.

Chapter 8

Bibliography

[1] P Censi, C Inguaggiato, S Chiavetta, C Schembri, F Sposito, V Censi, and P Zuddas. The behaviour of zirconium, hafnium and rare earth elements during the crystallisation of halite and other salt minerals. *Chemical Geology*, 453:80–91, 2017.

[2] James M Campbell, Fiona C Meldrum, and Hugo K Christenson. Observing the formation of ice and organic crystals in active sites. *Proceedings of the National Academy of Sciences*, 114(5):810–815, 2017.

[3] Saeed R Khan, Margaret S Pearle, William G Robertson, Giovanni Gambaro, Benjamin K Canales, Steeve Doizi, Olivier Traxer, and Hans-Göran Tiselius. Kidney stones. *Nature Reviews Disease Primers*, 2(1):1–23, 2016.

[4] Ajay Pande, Jayanti Pande, Neer Asherie, Aleksey Lomakin, Olutayo Ogun, Jonathan King, and George B. Benedek. Crystal cataracts: Human genetic cataract caused by protein crystallization. *Proceedings of the National Academy of Sciences*, 98(11):6116–6120, 2001.

[5] Brice Martin Couillaud, Philippe Espeau, Nathalie Mignet, and Yohann Corvis. State of the art of pharmaceutical solid forms: from crystal property issues to nanocrystals

formulation. *ChemMedChem*, 14(1):8–23, 2019.

[6] Roger J Davey, Sven LM Schroeder, and Joop H ter Horst. Nucleation of organic crystals—a molecular perspective. *Angewandte Chemie International Edition*, 52(8):2166–2179, 2013.

[7] Denis Gebauer, Matthias Kellermeier, Julian D Gale, Lennart Bergström, and Helmut Cölfen. Pre-nucleation clusters as solute precursors in crystallisation. *Chemical Society Reviews*, 43(7):2348–2371, 2014.

[8] Peter G Vekilov. The two-step mechanism of nucleation of crystals in solution. *Nanoscale*, 2(11):2346–2357, 2010.

[9] Maria A Vorontsova, Dominique Maes, and Peter G Vekilov. Recent advances in the understanding of two-step nucleation of protein crystals. *Faraday discussions*, 179:27–40, 2015.

[10] Denis Gebauer, Antje Völkel, and Helmut Cölfen. Stable prenucleation calcium carbonate clusters. *Science*, 322(5909):1819–1822, 2008.

[11] Anna Jawor-Baczynska, Jan Sefcik, and Barry D Moore. 250 nm glycine-rich nanodroplets are formed on dissolution of glycine crystals but are too small to provide productive nucleation sites. *Crystal growth & design*, 13(2):470–478, 2013.

[12] Georgina Zimbitas, Anna Jawor-Baczynska, Maria Jazmin Vesga, Nadeem Javid, Barry D Moore, John Parkinson, and Jan Sefcik. Investigation of molecular and mesoscale clusters in undersaturated glycine aqueous solutions. *Colloids and Surfaces A: Physicochemical and Engineering Aspects*, 579:123633, 2019.

[13] Fuyuki Ito, Yukino Suzuki, Jun-ichi Fujimori, Takehiro Sagawa, Mitsuo Hara, Takahiro Seki, Ryohei Yasukuni, and Marc Lamy De La Chapelle. Direct visualization of the two-step nucleation model by fluorescence color changes during evaporative crystallization from solution. *Scientific reports*, 6(1):1–6, 2016.

[14] Christine A Mitchell, Lian Yu, and Michael D Ward. Selective nucleation and discovery of organic polymorphs through epitaxy with single crystal substrates. *Journal of the American Chemical Society*, 123(44):10830–10839, 2001.

[15] Jose V Parambil, Sendhil K Poornachary, Jerry YY Heng, and Reginald BH Tan. Template-induced nucleation for controlling crystal polymorphism: from molecular mechanisms to applications in pharmaceuticals processing. *CrystEngComm*, 2019.

[16] Ying Diao, Allan S Myerson, T Alan Hatton, and Bernhardt L Trout. Surface design for controlled crystallization: The role of surface chemistry and nanoscale pores in heterogeneous nucleation. *Langmuir*, 27(9):5324–5334, 2011.

[17] Ying Diao, Kristen E Whaley, Matthew E Helgeson, Mahlet A Woldeyes, Patrick S Doyle, Allan S Myerson, T Alan Hatton, and Bernhardt L Trout. Gel-induced selective crystallization of polymorphs. *Journal of the American Chemical Society*, 134(1):673–684, 2012.

[18] Fiona C Meldrum and Cedrick O'Shaughnessy. Crystallization in confinement. *Advanced Materials*, 32(31):2001068, 2020.

[19] Jie Lu, Yi-Ping Li, Jing Wang, Guo-Bin Ren, Sohrab Rohani, and Chi-Bun Ching. Crystallization of an active pharmaceutical ingredient that oils out. *Separation and purification technology*, 96:1–6, 2012.

[20] G Patrick Stahly. Diversity in single-and multiple-component crystals. the search for and prevalence of polymorphs and cocrystals. *Crystal growth & design*, 7(6):1007–1026, 2007.

[21] Rolf Hilfiker and Markus von Raumer. *Polymorphism in the pharmaceutical industry*, volume 308. Wiley Online Library, 2006.

[22] Dario Braga and Fabrizia Grepioni. Making crystals from crystals: a green route to crystal engineering and polymorphism. *Chemical Communications*, (29):3635–3645, 2005.

[23] Thorsteinn Loftsson and Marcus E Brewster. Pharmaceutical applications of cyclodextrins: basic science and product development. *Journal of pharmacy and pharmacology*, 62(11):1607–1621, 2010.

[24] Nicholas Blagden, Marcel de Matas, Pauline T Gavan, and Peter York. Crystal engineering of active pharmaceutical ingredients to improve solubility and dissolution rates. *Advanced drug delivery reviews*, 59(7):617–630, 2007.

[25] P Láng, V Kiss, R Ambrus, G Farkas, P Szabó-Révész, Z Aigner, and E Várkonyi. Polymorph screening of an active material. *Journal of pharmaceutical and biomedical analysis*, 84:177–183, 2013.

[26] Yuriy A Abramov. Current computational approaches to support pharmaceutical solid form selection. *Organic Process Research & Development*, 17(3):472–485, 2013.

[27] Marcus A Neumann and Jacco van de Streek. How many ritonavir cases are there still out there? *Faraday discussions*, 211:441–458, 2018.

[28] Jack D Dunitz and Joel Bernstein. Disappearing polymorphs. *Accounts of chemical research*, 28(4):193–200, 1995.

[29] Dejan-Krešimir Bučar, Robert W Lancaster, and Joel Bernstein. Disappearing polymorphs revisited. *Angewandte Chemie International Edition*, 54(24):6972–6993, 2015.

[30] John Bauer, Stephen Spanton, Rodger Henry, John Quick, Walter Dziki, William Porter, and John Morris. Ritonavir: an extraordinary example of conformational polymorphism. *Pharmaceutical research*, 18(6):859–866, 2001.

[31] Richard Telford, Colin C Seaton, Alexander Clout, Asma Buanz, Simon Gaisford, Gareth R Williams, Timothy J Prior, Chidera H Okoye, Tasnim Munshi, and Ian J Scowen. Stabilisation of metastable polymorphs: the case of paracetamol form iii. *Chemical communications*, 52(81):12028–12031, 2016.

[32] Philippe Laval, Jean-Baptiste Salmon, and Mathieu Joanicot. A microfluidic device for investigating crystal nucleation kinetics. *Journal of Crystal Growth*, 303(2):622–628, 2007.

[33] Jung-uk Shim, Galder Cristobal, Darren R Link, Todd Thorsen, and Seth Fraden. Using microfluidics to decouple nucleation and growth of protein crystals. *Crystal Growth and Design*, 7(11):2192–2194, 2007.

[34] Anuradha R. Pallipurath, Pierre-Baptiste Flandrin, Lois E. Wayment, Chick C. Wilson, and Karen Robertson. In situ non-invasive Raman spectroscopic characterisation of succinic acid polymorphism during segmented flow crystallisation. *Mol. Syst. Des. Eng.*, 5:294–303, 2020.

[35] Laurie J Little, Richard P Sear, and Joseph L Keddie. Does the γ polymorph of glycine nucleate faster? a quantitative study of nucleation from aqueous solution. *Crystal Growth & Design*, 15(11):5345–5354, 2015.

[36] Maria J. Vesga, David McKechnie, Paul A. Mulheran, Karen Johnston, and Jan Sefcik. Conundrum of γ glycine nucleation revisited: to stir or not to stir? *CrystEngComm*, 21:2234, 2019.

[37] Manuel Ildefonso, Nadine Candoni, and Stéphane Veesler. Heterogeneous nucleation in droplet-based nucleation measurements. *Crystal Growth & Design*, 13(5):2107–2110, 2013.

[38] Daniel Selzer, Corinna Frank, and Matthias Kind. On the effect of the continuous phase on primary crystal nucleation of aqueous kno3 solution droplets. *Journal of Crystal Growth*, 517:39–47, 2019.

[39] Maria L. Briuglia. Primary and secondary crystal nucleation of pharmaceutics. PhD thesis, 2017.

[40] Ru-Ying Qian and Gregory D Botsaris. Nuclei breeding from a chiral crystal seed of naclo3. *Chemical engineering science*, 53(9):1745–1756, 1998.

[41] Richard E Marsh. A refinement of the crystal structure of glycine. *Acta Crystallographica*, 11(9):654–663, 1958.

[42] Yoichi Iitaka. The crystal structure of β-glycine. *Acta Crystallographica*, 13(1):35–45, 1960.

[43] Yoichi Iitaka. The crystal structure of γ-glycine. *Acta Crystallographica*, 14(1):1–10, 1961.

[44] Alice Dawson, David R Allan, Scott A Belmonte, Stewart J Clark, William IF David, Pamela A McGregor, Simon Parsons, Colin R Pulham, and Lindsay Sawyer. Effect of high pressure on the crystal structures of polymorphs of glycine. *Crystal growth & design*, 5(4):1415–1427, 2005.

[45] Stephen A Moggach, William G Marshall, David M Rogers, and Simon Parsons. How focussing on hydrogen bonding interactions in amino acids can miss the bigger picture: a high-pressure neutron powder diffraction study of ε-glycine. *CrystEngComm*, 17(28):5315–5328, 2015.

[46] Craig L Bull, Giles Flowitt-Hill, Stefano De Gironcoli, Emine Küçükbenli, Simon Parsons, C Huy Pham, Helen Y Playford, and Matthew G Tucker. ζ-glycine: insight into the mechanism of a polymorphic phase transition. *IUCrJ*, 4(5):569–574, 2017.

[47] Carol Forsyth, Iain S Burns, Paul A Mulheran, and Jan Sefcik. Scaling of glycine nucleation kinetics with shear rate and glass–liquid interfacial area. *Crystal Growth & Design*, 16(1):136–144, 2015.

[48] In Sung Lee, Ki Tae Kim, Alfred Y Lee, and Allan S Myerson. Concomitant crystallization of glycine on patterned substrates: The effect of ph on the polymorphic outcome. *Crystal Growth and Design*, 8(1):108–113, 2008.

[49] Xia Yang, Jie Lu, Xiu-Juan Wang, and Chi-Bun Ching. Effect of sodium chloride on the nucleation and polymorphic transformation of glycine. *Journal of Crystal Growth*, 310(3):604–611, 2008.

[50] K Renuka Devi, V Gnanakamatchi, and K Srinivasan. Attainment of unstable β nucleation of glycine through novel swift cooling crystallization process. *Journal of crystal growth*, 400:34–42, 2014.

[51] Ensieh Seyedhosseini, Maxim Ivanov, Vladimir Bystrov, Igor Bdikin, Pavel Zelenovskiy, Vladimir Ya Shur, Andrei Kudryavtsev, Elena D Mishina, Alexander S Sigov, and Andrei L Kholkin. Growth and nonlinear optical properties of β-glycine crystals grown on pt substrates. *Crystal growth & design*, 14(6):2831–2837, 2014.

[52] Elena S Ferrari, Roger J Davey, Wendy I Cross, Amy L Gillon, and Christopher S Towler. Crystallization in polymorphic systems: the solution-mediated transformation of β to α glycine. *Crystal growth & design*, 3(1):53–60, 2003.

[53] Kyungho Park, James MB Evans, and Allan S Myerson. Determination of solubility of polymorphs using differential scanning calorimetry. *Crystal growth & design*, 3(6):991–995, 2003.

[54] Hua Sun, Lu Wang, and Baoshu Liu. Solubility of α-glycine in water with additives at a temperature range of (293.15–343.15) k: Experimental data and results of thermodynamic modeling. *Fluid Phase Equilibria*, 434:167–175, 2017.

[55] Koichi Igarashi, Yuichi Sasaki, Masayuki Azuma, Hideo Noda, and Hiroshi Ooshima. Control of polymorphs on the crystallization of glycine using a wwdj batch crystallizer. *Engineering in life sciences*, 3(3):159–163, 2003.

[56] Hiroshi Matsuo, Yoshihisa Suzuki, and Seiji Sawamura. Solubility of α-amino acids in water under high pressure: glycine, l-alanine, l-valine, l-leucine, and l-isoleucine. *Fluid phase equilibria*, 200(2):227–237, 2002.

[57] Farid I El-Dossoki. Effect of the charge and the nature of both cations and anions on the solubility of zwitterionic amino acids, measurements and modeling. *Journal of solution chemistry*, 39(9):1311–1326, 2010.

[58] Zheng Cao, Yonghong Hu, Jiaojiao Li, Yumei Kai, and Wenge Yang. Solubility of glycine in binary system of ethanol+ water solvent mixtures: Experimental data and thermodynamic modeling. *Fluid Phase Equilibria*, 360:156–160, 2013.

[59] John B Dalton and Carl LA Schmidt. The solubilities of certain amino acids in water, the densities of their solutions at twenty-five degrees, and the calculated heats of solution and partial molal volumes. *Journal of Biological Chemistry*, 103(2):549–578, 1933.

[60] Luisa A Ferreira, Eugénia A Macedo, and Simão P Pinho. Solubility of amino acids and diglycine in aqueous–alkanol solutions. *Chemical engineering science*, 59(15):3117–3124, 2004.

[61] Luisa A Ferreira, Eugénia A Macedo, and Simão P Pinho. Kcl effect on the solubility of five different amino acids in water. *Fluid phase equilibria*, 255(2):131–137, 2007.

[62] Pannur Venkatesu, Ho-mu Lin, and Ming-Jer Lee. Counteracting effects of trimethylamine n-oxide and betaine on the interactions of urea with zwitterionic glycine peptides. *Thermochimica acta*, 491(1-2):20–28, 2009.

[63] Johanne Bonnin-Paris, Stephane Bostyn, Jean-Louis Havet, and Henri Fauduet. Determination of the metastable zone width of glycine aqueous solutions for batch crystallizations. *Chemical Engineering Communications*, 198(8):1004–1017, 2011.

[64] Edwin J Cohn, Thomas L McMeekin, Jesse P Greenstein, and John H Weare. Studies in the physical chemistry of amino acids, peptides and related substances. viii. the relation between the activity coefficients of peptides and their dipole moments. *Journal of the American Chemical Society*, 58(12):2365–2370, 1936.

[65] P Ramasami. Solubilities of amino acids in water and aqueous sodium sulfate and related apparent transfer properties. *Journal of Chemical & Engineering Data*, 47(5):1164–1166, 2002.

[66] Max S Dunn, Frank J Ross, and Lee S Read. The solubility of the amino acids in water. *Journal of Biological Chemistry*, 103(2):579–595, 1933.

[67] M Jelińska-Kazimierczuk and J Szydłowski. Isotope effect on the solubility of amino acids in water. *Journal of solution chemistry*, 25(12):1175–1184, 1996.

[68] Yongjin Yi, Dimitri Hatziavramidis, Allan S Myerson, Michael Waldo, Vladimir G Beylin, and Jason Mustakis. Development of a small-scale automated solubility measurement apparatus. *Industrial & engineering chemistry research*, 44(15):5427–5433, 2005.

[69] Himansu Talukdar, Sibaprasad Rudra, and Kiron K Kundu. Thermodynamics of transfer of glycine, diglycine, and triglycine from water to aqueous solutions of urea, glycerol, and sodium nitrate. *Canadian journal of chemistry*, 66(3):461–468, 1988.

[70] Yasuhiko Nozaki and Charles Tanford. The solubility of amino acids and two glycine peptides in aqueous ethanol and dioxane solutions: establishment of a hydrophobicity scale. *Journal of Biological Chemistry*, 246(7):2211–2217, 1971.

[71] Carmen M Romero and Carlos D Oviedo. Effect of temperature on the solubility of α-amino acids and α, ω-amino acids in water. *Journal of Solution Chemistry*, 42(6):1355–1362, 2013.

[72] LS Mason. The solubilities of four amino butyric acids and the densities of aqueous solutions of the acids at 25. *Journal of the American Chemical Society*, 69(12):3000–3002, 1947.

[73] Xia Yang, Xiujuan Wang, and Chi Bun Ching. Solubility of form α and form γ of glycine in aqueous solutions. *Journal of Chemical & Engineering Data*, 53(5):1133–1137, 2008.

[74] Shanfeng Jiang and Joop H ter Horst. Crystal nucleation rates from probability distributions of induction times. *Crystal Growth & Design*, 11(1):256–261, 2010.

[75] Peter Larkin. *Infrared and Raman spectroscopy: principles and spectral interpretation*. Elsevier, 2017.

[76] Peter R Griffiths and James A De Haseth. *Fourier transform infrared spectrometry*, volume 171. John Wiley & Sons, 2007.

[77] Ewen Smith and Geoffrey Dent. Modern raman spectroscopy: a practical approach. 2005.

[78] Challapalli Suryanarayana and M Grant Norton. *X-ray diffraction: a practical approach*. Springer Science & Business Media, 2013.

[79] William C Swope, Hans C Andersen, Peter H Berens, and Kent R Wilson. A computer simulation method for the calculation of equilibrium constants for the formation of physical clusters of molecules: Application to small water clusters. *The Journal of chemical physics*, 76(1):637–649, 1982.

[80] Michael P Allen and Dominic J Tildesley. *Computer simulation of liquids*. Oxford university press, 2017.

[81] Philip M Morse. Diatomic molecules according to the wave mechanics. ii. vibrational levels. *Physical review*, 34(1):57, 1929.

[82] Christopher M Baker. Polarizable force fields for molecular dynamics simulations of biomolecules. *Wiley Interdisciplinary Reviews: Computational Molecular Science*, 5(2):241–254, 2015.

[83] Daan Frenkel and Berend Smit. *Understanding molecular simulation: from algorithms to applications*, volume 1. Elsevier, 2001.

[84] 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.

[85] R Hockney and J Eastwood. Computer simulations using particles mcgraw-hill. *New York*, 1981.

[86] Herman JC Berendsen, JPM van Postma, Wilfred F van Gunsteren, ARHJ DiNola, and Jan R Haak. Molecular dynamics with coupling to an external bath. *The Journal of chemical physics*, 81(8):3684–3690, 1984.

[87] Hans C Andersen. Molecular dynamics simulations at constant pressure and/or temperature. *The Journal of chemical physics*, 72(4):2384–2393, 1980.

[88] Shūichi Nosé. A molecular dynamics method for simulations in the canonical ensemble. *Molecular physics*, 52(2):255–268, 1984.

[89] William G Hoover. Canonical dynamics: Equilibrium phase-space distributions. *Physical review A*, 31(3):1695, 1985.

[90] Glenn J Martyna, Douglas J Tobias, and Michael L Klein. Constant pressure molecular dynamics algorithms. *The Journal of chemical physics*, 101(5):4177–4189, 1994.

[91] Steve Plimpton. Fast parallel algorithms for short-range molecular dynamics. *Journal of computational physics*, 117(1):1–19, 1995.

[92] Daniel W Cheong and Yi Di Boon. Comparative study of force fields for molecular dynamics simulations of α-glycine crystal growth from solution. *Crystal Growth & Design*, 10(12):5146–5158, 2010.

[93] Conor Parks, Andy Koswara, Frank DeVilbiss, Hsien-Hsin Tung, Nandkishor K Nere, Shailendra Bordawekar, Zoltan K Nagy, and Doraiswami Ramkrishna. Solubility curves and nucleation rates from molecular dynamics for polymorph prediction–moving beyond lattice energy minimization. *Physical Chemistry Chemical Physics*, 19(7):5285–5295, 2017.

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

[95] JL Derissen, PH Smit, and J Voogd. Calculation of the electrostatic lattice energies of. alpha.-,. beta.-, and. gamma.-glycine. *The Journal of Physical Chemistry*, 81(15):1474–1476, 1977.

[96] HJC Berendsen, JR Grigera, and TP Straatsma. The missing term in effective pair potentials. *Journal of Physical Chemistry*, 91(24):6269–6271, 1987.

[97] Jean-Paul Ryckaert, Giovanni Ciccotti, and Herman JC Berendsen. Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *Journal of computational physics*, 23(3):327–341, 1977.

[98] Alexei M Nikitin, Yury V Milchevskiy, and Alexander P Lyubartsev. A new amber-compatible force field parameter set for alkanes. *Journal of molecular modeling*, 20(3):2143, 2014.

[99] Manuel Kuhs, Jacek Zeglinski, and Åke C Rasmuson. Influence of history of solution in crystal nucleation of fenoxycarb: kinetics and mechanisms. *Crystal Growth & Design*, 14(3):905–915, 2014.

[100] Nadeem Javid, Thomas Kendall, Iain S Burns, and Jan Sefcik. Filtration suppresses laser-induced nucleation of glycine in aqueous solutions. *Crystal Growth & Design*, 16(8):4196–4202, 2016.

[101] Richard P Sear. Quantitative studies of crystal nucleation at constant supersaturation: experimental data and models. *CrystEngComm*, 16(29):6506–6522, 2014.

[102] Gianluca Di Profio, Enrica Fontananova, Efrem Curcio, and Enrico Drioli. From Tailored Supports to Controlled Nucleation: Exploring Material Chemistry, Surface Nanostructure, and Wetting Regime Effects in Heterogeneous Nucleation of Organic Molecules. *Crystal Growth and Design*, 12:3749–3757, 2012.

[103] Isaac Jerome C Dela Cruz, Jem Valerie Perez, Bryan Gencianeo Alamani, Gerard Capellades, and Allan S Myerson. Influence of volume on the nucleation of model

organic molecular crystals through an induction time approach. *Crystal Growth & Design*, 2021.

[104] EV Boldyreva, VA Drebushchak, TN Drebushchak, IE Paukov, Yu A Kovalevskaya, and ES Shutova. Polymorphism of glycine, part ii. *Journal of thermal analysis and calorimetry*, 73(2):419, 2003.

[105] GL Perlovich, L Kr Hansen, and A Bauer-Brandl. The polymorphism of glycine. thermochemical and structural aspects. *Journal of thermal analysis and calorimetry*, 66(3):699–715, 2001.

[106] Giovanni Maria Maggioni, Luca Bosetti, Elena dos Santos, and Marco Mazzotti. Statistical analysis of series of detection time measurements for the estimation of nucleation rates. *Crystal Growth & Design*, 17(10):5488–5498, 2017.

[107] Samir A Kulkarni, Somnath S Kadam, Hugo Meekes, Andrzej I Stankiewicz, and Joop H ter Horst. Crystal nucleation kinetics from induction times and metastable zone widths. *Crystal growth & design*, 13(6):2435–2440, 2013.

[108] Masakazu Matsumoto, Shinji Saito, and Iwao Ohmine. Molecular dynamics simulation of the ice nucleation and growth process leading to water freezing. *Nature*, 416(6879):409–413, 2002.

[109] JY Yan and GN Patey. Heterogeneous ice nucleation induced by electric fields. *The Journal of Physical Chemistry Letters*, 2(20):2555–2559, 2011.

[110] Debashree Chakraborty and GN Patey. How crystals nucleate and grow in aqueous nacl solution. *The journal of physical chemistry letters*, 4(4):573–578, 2013.

[111] Soroush Ahmadi, Yuanyi Wu, and Sohrab Rohani. Molecular dynamics simulation of homogeneous nucleation of supersaturated potassium chloride (kcl) in aqueous solutions. *CrystEngComm*, 21(48):7507–7518, 2019.

[112] David Quigley and PM Rodger. A metadynamics-based approach to sampling crystallisation events. *Molecular Simulation*, 35(7):613–623, 2009.

[113] D Quigley and PM Rodger. Metadynamics simulations of ice nucleation and growth. *The Journal of chemical physics*, 128(15):154518, 2008.

[114] Hiroki Nada. Pathways for the formation of ice polymorphs from water predicted by a metadynamics method. *Scientific reports*, 10(1):1–8, 2020.

[115] D Quigley, PM Rodger, CL Freeman, JH Harding, and DM Duffy. Metadynamics simulations of calcite crystallization on self-assembled monolayers. *The Journal of chemical physics*, 131(9):094703, 2009.

[116] Marco Lauricella, Simone Meloni, Niall J English, Baron Peters, and Giovanni Ciccotti. Methane clathrate hydrate nucleation mechanism by advanced molecular simulations. *The Journal of Physical Chemistry C*, 118(40):22847–22857, 2014.

[117] Federico Giberti, Matteo Salvalaglio, Marco Mazzotti, and Michele Parrinello. Insight into the nucleation of urea crystals from the melt. *Chemical Engineering Science*, 121:51–59, 2015.

[118] Federico Giberti, Matteo Salvalaglio, and Michele Parrinello. Metadynamics studies of crystal nucleation. *IUCrJ*, 2(2):256–266, 2015.

[119] Marcello Sega, Sofia S. Kantorovich, Pál Jedlovszky, and Miguel Jorge. The generalized identification of truly interfacial molecules (ITIM) algorithm for nonplanar interfaces. *J. Chem. Phys.*, 138(044110), 2013.

[120] Diana M Rodríguez and Carmen M Romero. Surface tension of glycine, alanine, aminobutyric acid, norvaline, and norleucine in water and in aqueous solutions of strong electrolytes at temperatures from (293.15 to 313.15) k. *Journal of Chemical & Engineering Data*, 62(11):3687–3696, 2017.

[121] Alexandre P dos Santos and Yan Levin. Ions at the water–oil interface: interfacial tension of electrolyte solutions. *Langmuir*, 28(2):1304–1308, 2012.

[122] Jun Huang, Thomas C Stringfellow, and Lian Yu. Glycine exists mainly as monomers, not dimers, in supersaturated aqueous solutions: Implications for understanding its crystallization and polymorphism. *Journal of the American Chemical Society*, 130(42):13973–13980, 2008.

[123] Stefano Piana, Franca Jones, and Julian D Gale. Assisted desolvation as a key kinetic step for crystal growth. *Journal of the American Chemical Society*, 128(41):13568–13574, 2006.

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

[125] William Arthur Steele. *The interaction of gases with solid surfaces*, volume 3. Pergamon, 1974.

[126] Daniel W Siderius and Lev D Gelb. Extension of the steele 10-4-3 potential for adsorption calculations in cylindrical, spherical, and other pore geometries. *The Journal of chemical physics*, 135(8):084703, 2011.

[127] R_ Sonnenschein and K Heinzinger. A molecular dynamics study of water between lennard-jones walls. *Chemical physics letters*, 102(6):550–554, 1983.

[128] William R Salaneck, Ingemar Lundström, and Bo Liedberg. Photoelectron spectroscopy of amino acids adsorbed upon surfaces: glycine on graphite. In *Surfactants, Adsorption, Surface Spectroscopy and Disperse Systems*, pages 83–88. Springer, 1985.

[129] Matthew Boyes, Adriana Alieva, Jincheng Tong, Vaiva Nagyte, Manuel Melle-Franco, Thomas Vetter, and Cinzia Casiraghi. Exploiting the surface properties of graphene for polymorph selectivity. *ACS nano*, 14(8):10394–10401, 2020.

[130] Abraham Savitzky and Marcel JE Golay. Smoothing and differentiation of data by simplified least squares procedures. *Analytical chemistry*, 36(8):1627–1639, 1964.