Molecular Dynamics Studies of Lipase-Surface Interactions

Nathalie Willems

Department of Biochemistry

University of Oxford

A thesis submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy in Biochemistry at the University of Oxford

Kellogg

Michaelmas 2016

Look here brother, who you jiving with that cosmik debris? - Frank Zappa

You run and you run to catch up with the sun, but it's sinking, racing around, to come up behind you again. - Pink Floyd

Acknowledgements

I would like to sincerely thank my supervisor Prof. Mark Sansom, who not only gave me the opportunity to work in his lab, but also provided exceptional support and guidance throughout my project. I would also like to thank Dr. Mickäel Lelimousin and Dr. Heidi Koldsø, for their endless patience and guidance; your advice and direction kept me on track and ultimately made me into a better scientist. I also thank Prof. Gail Preston for the opportunity to join the DTP, and for being very supportive over the last four years. I also acknowledge the BBSRC for funding for the project.

Further thanks go to the rest of the lab, particularly Matteo, Ben, Maria, Jemma, Erin, Dave and Jo. Our weekly pub trips and coffee breaks provided a source of much needed amusement, which made me feel at home in the lab. Also, a huge thank you to my house mates, Mike, Jonny, and Gemma, for all the scientific and political discussions, and for being funniest, most loving bunch of people I have ever met.

I would also like to thank my parents for raising me to be curious, and providing me with the endless advice and encouragement that got me to where I am today; this thesis would not exist without you. Finally, I would like to thank my partner Evie, for her unconditional support, kindness, and encouragement throughout my DPhil. I am eternally grateful.

Publications

The following publications are mentioned in this thesis:

- Willems, N., Lelimousin, M., Koldsø, H., Sansom, M.S.P. "Interfacial enzymes and their interactions with surfaces: molecular simulation studies." *Understanding Enzymes; Function, Design, Engineering and Analysis.* Pan Stanford Publishing: Jan 2016;
- Willems, N., Lelimousin, M., Koldsø, H., Sansom, M.S.P. "On the interfacial activation of the M37 lipase: a multi-scale simulation study." (*accepted, BBA Biomembranes*);
- Willems, N., Urtizberea, A., Verro, A.F., Iliut, M., Lelimousin, M., Hirtz, M., Vijayaraghavan, A., Sansom, M.S.P. "Biomimetic Phospholipid Membrane Organization on Graphene and Graphene Oxide Surfaces: a Molecular Dynamics Simulation Study." (*submitted to ACS Nano*).

Abstract

Lipases are enzymes that play fundamental roles in fat digestion and metabolism, and function at the interface formed between hydrophobic molecules and the surrounding aqueous environment. These interfacial interactions are thought to induce conformational changes in a "lid" region of the lipase, leading to a dramatic increase in activity. This thesis aims to provide insight into the interactions that govern lipase association with interfaces of different structural characteristics, and the possible conformational changes that arise as a function of these interactions. A multi-scale molecular simulation approach (combining atomistic and coarse-grained methods) was applied to study two different lipases with a range of interfaces, including "soft" biological surfaces and "hard" non-biological surfaces.

Three major insights were gained from these studies. First, interactions of a small bacterial lipase (M37) with lipid interfaces resulted in substantial structural changes in a lid region, uncovering of the underlying active site. A mechanism of interfacial activation is proposed for this lipase. Second, the interaction of M37 with non-biological interfaces differ from lipid interfaces, leading to altered interfacial orientations with possible functional consequences. Third, the amino acid composition of the lid region of a yeast lipase (TLL) is shown to play crucial roles in lipase activation and structural stability.

Table of contents

List of figures

List of tables

Chapter 1

Introduction

Lipases are enzymes of biotechnological importance that function at the interface formed between hydrophobic and aqueous environments. Their physiological role is to break down triglycerides during fat digestion, performing key roles in metabolism and homeostasis [1]. In addition to triglyceride hydrolysis, lipases can catalyse a wide range of other reactions on a variety of natural and unnatural substrates, which has resulted in the application of lipases in a range of biotechnological industries [2–4]. The interaction of lipases with hydrophobic interfaces can induce structural transitions that result in a dramatic increase in enzyme activity, a topic of relevance concerning their industrial potential. Although it is known that both the nature of the interface and the inherent structural dynamics of the enzyme are important in the interfacial activation of lipases, a detailed mechanism of this process is currently not well understood for many lipases. This thesis aims to further the general understanding of factors affecting lipase activation, including the structural nature of the interface and the conformational properties of the lipase itself. A comprehensive study is presented in which the interfacial interactions of different lipases are investigated with a range of surfaces of varying structural characteristics, aiming to provide molecular insight into mechanisms of interfacial activation.

1.1 Proteins at Interfaces

About half the proteins in the cell are thought to be membrane-associated and therefore function at the interface formed between the hydrophobic molecules and the surrounding aqueous environment [5]. Proteins at the interface play numerous key biological roles within the cell, including structural, catalytic, signalling, immunological, transportation, and developmental roles [1, 6–8]. A very important class of proteins includes membrane proteins, which either insert directly into the hydrophobic core of the cellular membrane (integral proteins), or anchor to the membrane surface (peripheral proteins) [9–11]. Other classes include interfacial proteins, such as phospholipase enzymes, superficially interact with membrane surfaces but are not physically anchored to them [12, 13]. While membrane proteins function at the interface between the interior and exterior of the cell, for the purposes of this thesis, the definition of *interfacial proteins* will be limited to proteins that transiently associate with the interface, such as phospholipases and lipases. The mechanisms of action of these interfacial proteins, such as the hydrolysis of phospholipids by phospholipases, can differ greatly, and depend heavily on the nature of the interface. Consequently, the extent to which interfacial proteins associate with the interface, their level of penetration within the interface, and their functional mechansims once bound to the interface remain active areas of study. The key processes performed by interfacial proteins highlight their importance in regulating normal cell function. Unravelling how these proteins interact with biological interfaces such as lipid membranes is therefore of fundamental biological interest.

1.1.1 Studies of Proteins at Interfaces

The reconstitution of biological interfaces in model systems is a common method to the study properties such as protein association and unravel specific interactions

with the interface. For example, supported lipid membranes were used to show how different peripheral protein constructs interact with a planar bilayer surface, revealing how specific protein-lipid interactions affect bilayer penetration and protein diffusion on the surface [14, 15].

From an applied perspective, the study of protein interactions with non-biological interfaces is important in the fields of nanotechnology, biotechnology, and biomaterials. In nanotechnology, understanding protein-surface interactions are key in generating functionalised devices, such as biosensors and activators [16]. For example, protein assemblies of bacteriorhodopsin molecules interfaced with a silicon dioxide surface were shown to maintain their optical properties, enabling the fabrication of a switchable optical device that could be tuned by using light [17]. In biomaterials, the first step in generating a human implantable devices is to investigate how proteins such as blood proteins (e.g fibrinogen) interact with the surface in question [18]. One study utilised surface-functionalised polystyrene nanoparticles to investigate how charge affected the adsorption of human blood serum proteins [19].

The range of surfaces used to investigate interfacial protein function can therefore differ greatly in physicochemical characteristics [20, 21], representing the considerable interest in understanding protein-surface interactions from a biological and applied perspective. This thesis will focus on how interfacial enzymes interact with biological and non-biological surfaces.

1.2 Interfacial Enzymes

Enzymes are unique molecules that catalyse chemical reactions by lowering the energy required to initiate the reaction [22]. A hugely diverse range of enzymes have been identified to date, which are currently classed according to the framework presented by the Enzyme Commission (E.C)[23]. The first level of this classification includes

six broad classes of enzymes, each corresponding to the particular type of chemistry involved in the reaction (Table 1.1).

Enzyme classes (E.C)	Class Description
A. Oxidoreductases	Substrate is oxidised (regarded as electron donor)
B. Transferases	Transfer of a group from one substrate to another
C. Hydrolases	Hydrolytic cleavage of a bond
D. Lyases	Cleavage of a bond by elimination
E. Isomerases	Isomerisation reaction
F. Ligases	Enzyme catalysing the joining of molecules in parallel with ATP hydrolysis

Table 1.1: Description of the six major enzyme classes according to the Enzyme Commission (E.C). Each class is described according to the chemical reaction catalysed.

Enzymes can either function as soluble entities in the cytoplasm, or associate with an interface. Interfacial enzymes interact with biological surfaces including the cellular membrane environment, into which non-polar enzymatic substrates can partition, or aggregates of non-polar substrates that form micelle-like structures, liposomal dispersions, or emulsions [5, 6]. Interfacial enzymes are crucial in cellular processes as diverse as cell signalling, membrane remodelling, fat digestion, endocytosis, and inflammation [1, 12]. For example, phospholipases catalyse the hydrolysis of phospholipid molecules, maintaining cellular metabolism and homeostasis [6]. Phosphatases such as Phosphatase and Tensin Homologue (PTEN) modify specific membrane lipids, regulating the cell cycle and cell differentiation [24].

The catalytic behaviour of interfacial enzymes is influenced by the overall dynamic and ever-fluctuating state of biological membranes [25]. Unravelling the dynamic interplay between the enzyme and the interface is important in understanding the non-

equilibrium processes that govern overall membrane topology and cellular function. Additionally, interfacial enzymes have considerable potential for industrial applications, such as lipases in food, detergent, and related industries [2]. Characterising the dynamic mechanisms of interfacial enzymes is therefore key in furthering their industrial potential, as well as our fundamental understanding of cellular processes.

1.2.1 Interfacial enzyme kinetics

The fluctuating, heterogeneous nature of biological interfaces previously mentioned imposes physical constraints on interfacial enzymes and catalysis, including effects on substrate accessibility, enzyme orientation, substrate and product partitioning within the interface, and the association/dissociation equilibrium of enzyme with the interface [5]. The kinetics that govern interfacial enzyme catalysis therefore deviate from classic Michaelis-Menten kinetics, typically used to describe soluble enzymes and substrates that are present within the same phase (Fig. 1.1) [26].

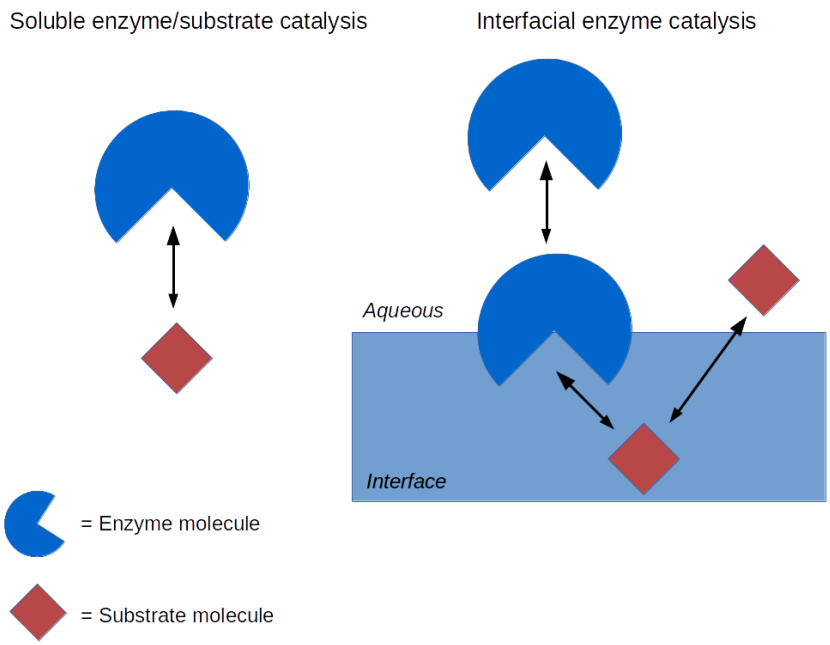

Fig. 1.1: *Left*: Typical substrate catalysis by non-interfacial soluble enzymes. *Right*: Typical substrate catalysis by interfacial enzymes. Substrate partitioning with the interface affects the catalytic cycle and kinetics of interfacial enzyme catalysis, differing from substrate recognition by soluble enzymes in the aqueous phase. The figure represent the general case of substrate partitioning within an interface. The position of the substrate within the interface, and thus its proximity to the associated enzyme, is difficult to determine experimentally and will differ depending on the nature of the interface.

The first kinetic model of interfacial enzyme catalysis was presented by Verger & de Haas, who adapted the Michaelis-Menten model to reflect interfacial lipolysis by studying soluble phospholipases (sPLA) [27]. This model included additional equilibrium terms that describe enzyme association/dissociation with the interface, and formation of the enzyme-substrate complex [27]. Further extension of the interfacial kinetics model also accounted for substrate partitioning at the lipid interface, dissociation of the enzyme-product complex, and equilibrium dissociation of the enzyme from the interface [26]. Other kinetic models are more system-specific and include the potential mode of action of the enzyme on a particular interface e.g mixed mi-

celles, monolayers, and vesicular systems, as well as product accumulation within the interface (Fig. 1.2) [28]. The association process of interfacial enzymes, depth of penetration at the interface, and location of the substrate within the interface, are difficult parameters to determine experimentally. Consequently, Fig. 1.1 and 1.2 present the general case of processes involved in interfacial enzyme kinetics. It is thought that the substrate orientation within the interface must be such that it is accessible by the interfacial enzyme, and thus most likely occupies locations near the interfacial region where the hydrophobic molecules meet the surrounding aqueous environment [1]. However, substrate localisation at the interface will also depend on other parameters, such as substrate concentration and the nature of the interactions between the substrate and other interfacial molecules. Therefore, the location of the substrate within the interface is not specifically commented in this thesis, but is assumed to be accessible to the associated enzyme.

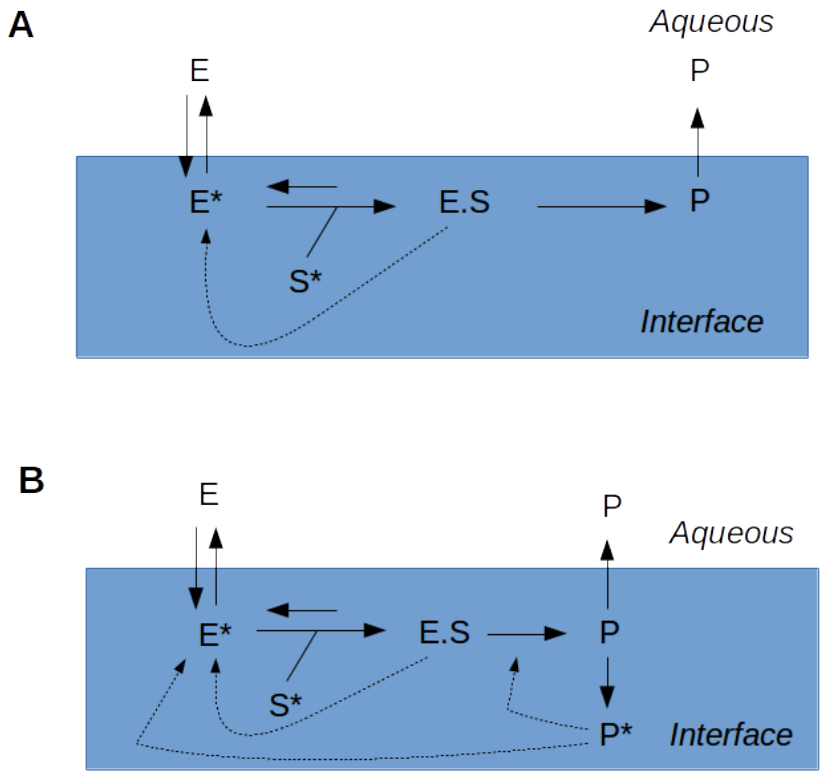

Fig. 1.2: **(A)** Verger-de Haas kinetic model describing the kinetics of interfacial lipolysis catalysed by soluble phospholipases at micelles interfaces [27]. **(B)** Extension of the Verger-de Haas kinetic model to describe interfacial lipid hydrolysis specifically by Phospholipase D at a phospholipid monolayer interface [28]. E = enzyme, S* = interfacial substrate, P = product, E* = interfacially bound enzyme, P* = product that remains at/within the interface. The model depicts the possible mechanisms (dotted arrows) by which product accumulation at the interface could modify individual kinetic steps within the reaction mechanism. The solid arrows indicate the direction of the reaction.

The system-specific kinetic models also highlight the importance of the nature of the interface itself during interfacial enzyme catalysis. For example, when the lipid substrate consists of micelles, liposomal dispersions, or emulsions, the possible exchange of substrate, enzyme, and product molecules between lipid particles must be taken into account [29, 30]. Early studies using liposomal dispersions addressed this by considering enzyme "processivity" at the interface. Processivity describes whether

or not the enzyme dissociates from the interface after one or a few catalytic cycles, and then re-associates for successive reactions, or if the enzyme remains bound to the interface, hydrolysing multiple substrate molecules successively for the duration of the catalytic phase. For phospholipases, two mechanisms were considered: the enzymes either adopts a "scooting" mode of action, where the interacting molecules (substrate, enzyme, and product) do not exchange between interfaces/particles and remain associated for multiple turnover cycles; or, the enzyme adopts a "quasi-scooting" mode, whereby replenishment of substrate from other aggregates occurs on the same time scale as an individual catalytic turnover event. In 1986, studies using pig pancreatic sPLA2 provided concrete evidence for interfacial catalysis via scooting mode, where the enzyme remained associated with DMPMe (ester) vesicles, hydrolysing all available substrate without measurable dissociation from the interfacial surface [31]. Subsequent studies supported this mechanism, laying the groundwork for quantitative analysis of turnover events via the defined kinetic path taken by these interfacial enzymes [5, 30, 32, 33].

Other factors such as interfacial tension are also known to affect interfacial kinetics [32]. Pancreatic lipase, for example, is known to deactivate at hexane/water interfaces during triglyceride hydrolysis, attributed to denaturation of the enzyme as the interfacial tension increases with continuing hydrolysis [34]. Importantly, this behaviour can be reversed by adding more amphiphiles to increase the surface pressure of the interface. Other environmental contributions such as substrate orientation, electrostatic interactions, and hydration forces are also known to affect interfacial properties, and therefore the associated interfacial enzyme [35, 36].

The pioneering work on interfacial enzyme catalysis of lipid substrates laid the groundwork for studying more complex systems: lipase-interface systems.

1.3 Lipases

Lipases are amongst the best-studied examples of interfacial enzymes. This is due to their biologically relevant roles in fat metabolism and human disorders, such as Wolman's disease, as well as their industrial potential [2, 37]. The physiological function of lipases is to hydrolyse the ester bonds in triglycerides (Fig. 1.3). They are found in most organisms ranging from microbial, plant, and animal kingdoms [38, 39]. They belong to the serine hydrolase family of enzymes, possessing the α/β protein fold and a Ser/His/Asp catalytic triad [40, 41].

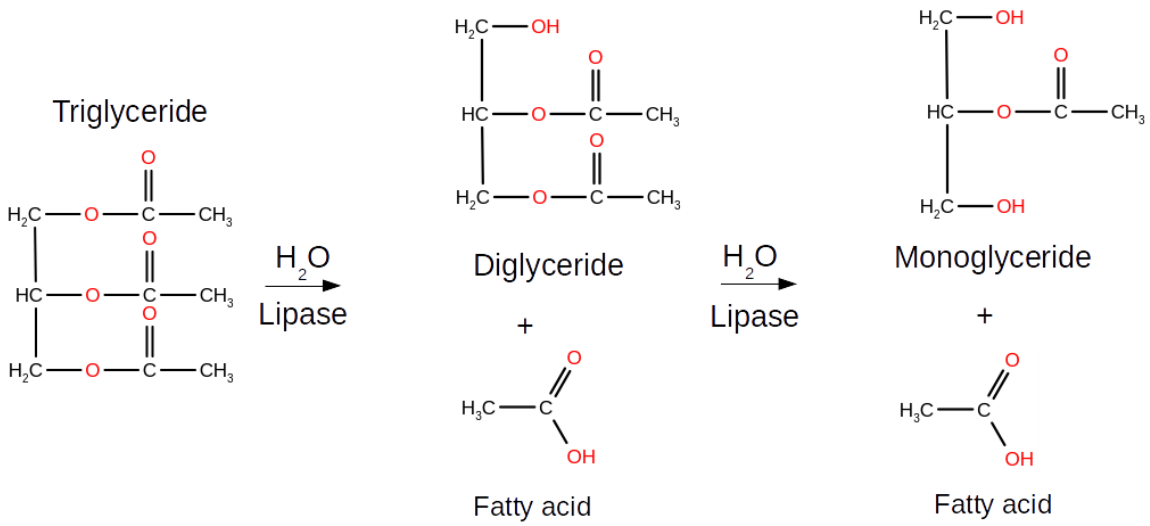

Fig. 1.3: Reaction scheme of lipase-catalysed triglyceride hydrolysis. The arrows indicate the progression of the reaction, in which complete hydrolysis of one triglyceride molecule will generate one glycerol and three fatty acid molecules.

Similar to phospholipases, lipases exhibit a dramatic increase in activity in the presence of an interface, a phenomenon called "interfacial activation" [36, 42]. Several theories have been proposed to explain this increase in activity, including a localised increase in substrate concentration at the interface, better orientation of the ester bond due to the geometrical packing of triglyceride molecules at the interface, better

access to the active site through reduction of the water shell around the substrate (the *substrate* theory) [43, 44], as well as a conformational transition in the lipase upon interaction with the interface (the *enzyme* theory) [42].

The enzyme theory was first proposed in the early 1990s, after several X-ray crystallography studies reported evident conformational changes in the crystal structures of several different lipase molecules solved in different crystallisation conditions [45–50]. Among these, the crystal structures of the *Rhizomucor miehei* lipase (RML) and human pancreatic lipase (HPL), solved in the presence of substrate molecules at 3 Å resolution, indicated a shift in the position of a structural region directly overlying the active site, termed the *lid* region [45, 49]. This mobile region was identified through comparison with the previously published structures of the same or related lipase molecules, solved in aqueous solution [51, 52]. This comparison revealed that the conformational transition in the lid region resulted in exposure of the active site to the substrate molecules (Fig. 1.4). For example, analysis of the "open" RML crystal structure, solved in the presence of a covalent inhibitor (N-hexyl chlorophosphonate ethyl ester), revealed substantial electron density corresponding to the inhibitor within the active site, suggesting that the lid repositioned itself to allow substrate access to the active site [45]. The conformational change resulted in exposure of hydrophobic residues, both within the lid region and the underlying active site, to the surrounding aqueous environment. Simultaneously, polar residues within the lid region became buried within a hydrophilic pocket next the active site, previously occupied by well-ordered water molecules, as identified from the closed crystal structure of RML [53]. Similarly, incubation of HPL with mixed micelles of phosphotidyl choline and bile salt resulted in co-crystallisation of the HPL with the micelles, revealing a transition within the lid region overlying the active site as compared with the closed crystal structure of a related lipase, previously solved in aqueous solution [49, 51]. Electron density was

identified within the active site of HPL, attributed to a phosphotidyl choline molecule, indicating that displacement of the lid region allowed access of the substrate to the active site. Structural analysis revealed that specific hydrophobic residues within the main helix of the lid region became exposed, as well as residues within the active site region [49]. This change in hydrophobic solvent accessible surface area is thought to be the driving force behind interfacial activation of the lipase, and initiated upon interaction with a hydrophobic interface.

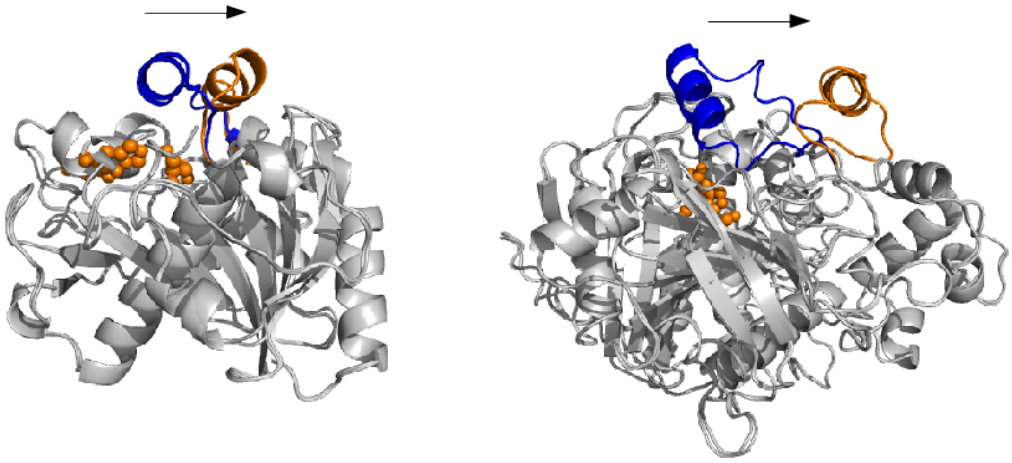

Fig. 1.4: Crystal structures of *Rhizomucor miehei* lipase (RML) [52, 53] and *Candida rugosa* lipase (CRL) solved in both the open and closed enzyme form [54]. The structures were aligned (grey cartoons); the open position of the lid region is shown in blue, whilst the closed position is shown in orange. The catalytic residues are shown as orange van der Waals spheres. The arrows indicate the direction of lid motion upon interfacial activation. Open RML = PDB: 4TGL [53], closed RML = PDB: 3TGL [52]; open CRL = PDB: 1CRL, closed CRL = PDB: 1TRH [54].

These early studies therefore supported the enzyme theory, suggesting that lipases occupy an closed conformation in solution, where hydrophobic residues are shielded from the aqueous environment, and transition to an open conformation upon interaction with an interface, exposing underlying hydrophobic residues in the active site, as well as in the lid region itself. On the other hand, crystal structures of *Thermomyces lanuginosus* lipase (TLL) and *Rhizopus delemar* lipase (RDL) revealed conformational

lability of the lid region in aqueous solution, suggesting the region may occupy multiple conformational states in aqueous solution [55]. In particular, the lid region of TLL was found to have poorly resolved electron density, suggesting high flexibility of this region in solution, whilst the lid region of RDL occupied a "semi-open" conformation in the crystal. The semi-open conformation of the RDL lid region was thought to be stabilised by crystal contacts between two symmetrical, neighbouring RDL molecules within the crystal, influencing the observed position of the lid region [55]. Furthermore, the study highlighted that the previously solved closed structure of RML, crystallised in aqueous solution, involved substantial crystal contacts between neighbouring molecules, perhaps artificially stabilising the closed position of the lid region identified in this molecule in aqueous conditions [52].

Interestingly, a systematic crystallographic study of TLL in media containing detergent or substrate analogues identified discrete lipase structures that exhibited different positions of the lid region [56]. These conformations were distinguished according to the degree of active site accessibility as a function of the lid position, allowing identification of a low activity, activated, and fully active lipase forms. Furthermore, biophysical studies of TLL used tryptophan fluorescence quenching to demonstrate lid motion in solvents with varying dielectric constants [57]. At low solvent polarity, the lid region is more open based on the quenching of a fluorescent probe attached near the lid region, whereas the lid adopts a semi-closed position in aqueous solution [57].

Considering this body of research, it is likely that the interfacial activation of lipases involves both substrate and enzyme components, where conformational changes within the enzyme are necessary to expose the underlying active site, but perhaps not sufficient to account for the dramatic increase in lipase activity in the presence of an interface. As such, lipases may exhibit a range of conformations in solution, of which the open and closed structures are the conformational extremes.

Once the lipase is activated however, the mechanism of action for hydrolysis is similar to that exhibited by all serine hydrolases [40, 58]. The serine residue within the catalytic triad initiates the reaction via nucleophilic attack on the carbon of the ester bond within the triglyceride substrate. A tetrahedral intermediate is formed, stabilised by the the oxyanion hole present in most lipases, usually comprising two or three amino acids neighbouring the active site that transiently interact with the negatively charged intermediate [1, 3, 59]. Loss of an alcohol group from the triglyceride intermediate then results in formation of the acyl-enzyme complex, which is followed by nucleophilic attack by a water molecule to give a second tetrahedral intermediate that finally loses the acid product molecule to restore the native enzyme form. The proposed reaction mechanism of CalA is shown in Fig. 1.5, in which residues Asp95 and Gly185 form the oxyanion hole [60].

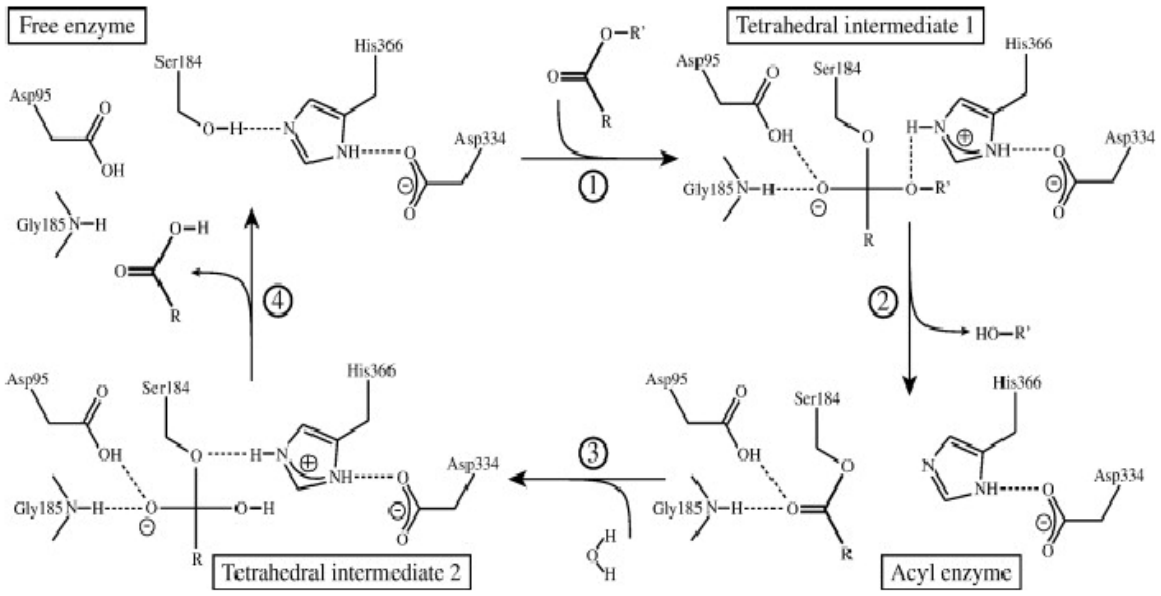

Fig. 1.5: Proposed reaction mechanism for *Candida antartica* lipase A during triglyceride hydrolysis. Figure reproduced from [60].

In addition to triglyceride hydrolysis, lipases are also able to catalyse additional reactions including trans-estrification, aminolysis, alcoholysis, aldol-additions, and epox-

idation reactions [3]. Furthermore, these enzymes are quite stable in non-aqueous organic solvents, resulting in the ability to fine-tune reaction conditions to promote hydrolysis reactions or ester synthesis, depending on the desired product [3]. This versatility has resulted in the exploitation of lipases in food, detergent, pharmaceutical, and other industries [4, 61]. For example, TLL is widely exploited in the detergent industry due to its ability to hydrolyse many different triglyceride substrates, as well as high stability at a range of temperature and pH values, resulting in commercially available forms of the enzyme such as Lipozyme TL IM and Lipolase [2]. Lipolase is added to several major detergent formulations to remove fat-containing stains including salad oils, fried fat, butter, and lipsticks, exemplifying the versatility of lipase promiscuity [2].

Conversely, applications of lipases in the pharmaceutical industry rely heavily on the ability of the enzyme to specifically discriminate between different substrates, particularly in the resolution of racemic mixtures. For example, the PS-30 lipase from *Pseudomonas cepacia* is used in the production of the anticancer drug Taxol (paclitaxel), where it selectively hydrolyses racemic acetate to form the desired (3R)-acetate [62]. Lipases are also used in the resolution of (R,S)-ibuprofen racemic mixtures, in order to produce the much more potent S-enantiomer of ibuprofen (160x) [62]. The lipase selectively synthesises the (S)-ester from the racemic mixture, allowing isolation of (R)-ibuprofen, which is then chemically transformed to (S)-ibuprofen. In the food industry, lipases are used to selectively synthesise "structured" lipids by modifying fats or oils in order enhance desirable properties, such as a lower resistance to stress, making fats easier to spread and rapidly melt [63]. These modifications involve changes in the position of the fatty acid on the glycerol backbone, degree of saturation, or changes to the length of the fatty acid tail. For example, lipases are used to produce Econa (diacylglycerol) oil from natural oils, resulting in 80% more diacylglycerol content,

which possesses the same energetic value as triglyceride oils, but are not transformed into body neutral fat [63].

Three major groups distinguish the ability of lipases to hydrolyse triglycerides (or other substrates), related to which of the three carbon tails in triglycerides is preferentially hydrolysed: 1) *Non-specific lipases* catalyse reactions on all carboxyl groups of the triglyceride, resulting in fatty acid release from any position on the glyceride [64]; 2) *1,3-specific lipases* only attack primary hydroxyl groups thus preferentially releasing fatty acids from position 1 and 3 [65]; 3) *Fatty-acid specific lipases* prefer hydrolysis of esters formed from long-chain fatty acids with double bonds between C9 and C10 e.g lipases from *Geotrichum* [66].

Clearly, substrate, regio- and stereospecificity is central to the application of lipases as industrial biocatalysts. For example, lipases that exhibit 1,3-regiospecificity are used to catalyse interestrification of natural triglycerides to produce different lipid structures, such as cocoa butter equivalents [67]; lipases that discriminate against omega-3 polyunsaturated fatty acids can be used in fatty acid production, such as *Candida rugosa* lipase-catalysed hydrolysis of salmon oil [68]. The industrial relevance of lipases has generated great interest in characterising lipase-catalysed reactions and particularly their interfacial interactions with different interfaces.

1.4 Lipase interactions with surfaces

As mentioned previously, the nature of the interface is important in the mode of action of interfacial enzymes, including lipases. For example, factors such as differences in surface curvature of liposomes of identical lipid composition can greatly affect the catalytic activity of lipases, as has been shown for TLL and pancreatic lipases [69–71]. Several experimental and computational studies have therefore focussed on investigating lipase interactions with different interfaces, indicating that lipases are able to

functionally interact with surfaces varying in hydrophobicity and topology [72–74]. For example, *Rhizomucor miehei* lipase (RML) interactions with silica supports resulted in a two-fold increase in specific activity relative to the free enzyme [75]. The addition of silanes to the silica surface (increases hydrophobicity) also resulted in improved activity of CalB and *Bulkholderia cepacia* lipase (BCL) [76]. Lipase interactions and activities have also been studied with inorganic surfaces such as carbon-based supports, including graphene, graphene oxide, and carbon nanotubes, as well as organic polymers, sol-gel materials, lipid aggregates, and organic solvent-aqueous solution interfaces [3, 72–74].

In contrast to the hydrophobic nature of the supports often used for lipase adsorption, hydrophilic materials such as alginate have also shown beneficial results regarding lipase activity, as indicated by studies of *Psuedomonas cepacia* lipase [77]. Furthermore, RML prepared with an anion-exchange resin in the presence of a surfactant resulted in high enzyme activity, thought to be due to activation of the enzyme prompted by the surfactant, which was subsequently stabilised by "conformational fixing" through ionic interactions with the support [78]. These studies indicate that the nature of interactions between the enzyme and the interface very important, as enzyme inhibition can occur via electrostatic interactions that cause conformational changes and/or substrate aggregation [79]. The considerable body of research characterising lipase-surface interactions reveals the versatility of lipases, although it is not always clear how the structural dynamics of the enzyme are affected, including enzyme orientation at the interface and substrate accessibility. Computational investigation has been useful in elucidating these details, providing insights into the structural effects of interfacial interactions and the molecular motions that result in lipase interfacial activation.

1.4.1 Computational Studies of Lipases

Computational methods such as molecular dynamics (MD) simulations are used to study numerous biological systems, including complex lipid membranes, protein-lipid interactions, protein-protein interactions, and virus particles [80–83], as well as systems involving non-biological components such as carbon nanomaterials [84–88]. Other computational methods such as quantum mechanical calculations are also useful in studying enzyme properties [89, 90]. These investigations have provided mechanistic insight into reaction mechanisms, substrate specificity, association of lipases with interfaces, and conformational transitions.

Importantly, MD simulation data can aid in the interpretation of experimental information. For example, simulations of both the open and the closed conformations of TLL at an air-water interface provided insight into experimental X-ray reflectivity data of the same system [91]. The orientation and conformation of TLL at the interface was deduced as a function of this data, showing that the lipase adopted a semi-open lid conformation pointing towards the interface. Furthermore, an atomistic MD study of TLL at a tributyrin interface revealed that the presence of a polar channel within the enzyme structure could play a role in the access of water to the active site, affecting the type of reaction catalysed (hydrolysis *vs* transestrification). Interestingly, the study reported the loss of α-helicity within the lid region in certain repeat simulations, indicating conformational lability of this small α-helical region, corresponding to crystallographic data [55]. Investigation of lipase interactions with non-biological surfaces have also revealed interesting conformational affects. For example, simulations of a small yeast lipase called Lip2 from *Yarrowia lipolytica* with carbon nanotubes provided insight into the experimentally observed enhanced activity exhibited by immobilised Lip2 [92, 93]. Interestingly, the simulation data showed that the interaction site of the lipase with the carbon nanotube was distal to the lid region,

and yet still resulted in an opening motion of the lid region. It was suggested that these interfacial interactions were propagated to the lid region of Lip2, initiating a chain of interactions leading to structural rearrangement of the lid region to an open conformation. The suggestion that lipases may be able to bind interfaces in a number of different orientations has implications for lid movement and activation, as seen in other experimental studies [71, 94–96].

Lipases are also active at the interface between organic media and the aqueous environment, and advantageously exhibit higher activity and stability relative to most other enzymes [3]. MD simulations of lipases in organic solvent have revealed pathways of conformational transitions during interfacial activation for three different lipases (RML, TLL, and *Candida rugosa* lipase (CRL)), indicating important differences in the activation mechanisms of homologous lipases (Fig. 1.6.1) [97, 98]. Similarly, simulations of lipase interactions with alkane mixtures showed a dependence on the type of alkane interface in conformational transitions to an open state of the lipase molecule (Fig. 1.6.2) [99]. Furthermore, MD simulations have aided in elucidating the structural elements that are relevant in lipase activation, even identifying an additional second lid region for a particular lipase not identified from the original crystal structure (Fig. 1.6.3) [100].

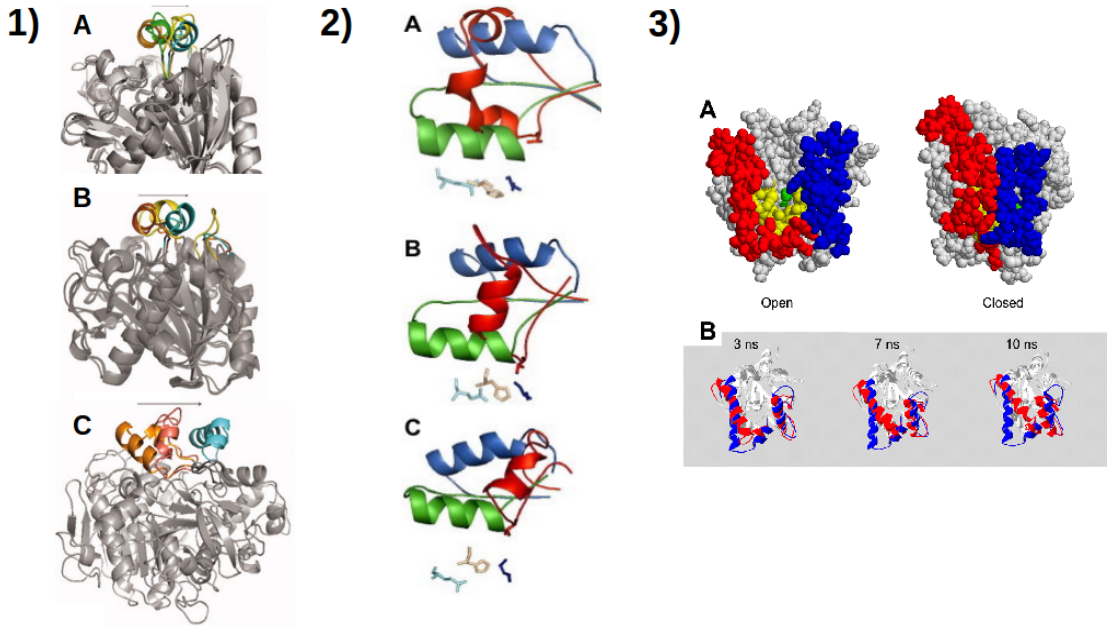

Fig. 1.6: MD simulation studies of lipases. **1)** Conformational transition pathways from closed to open structures observed during simulations of RML (**A**), TLL (**B**), and CRL (**C**) [97]. Coloured regions correspond to the lid region of each lipase structure, showing its position at 0 ns, 25 ns simulation time, and in the open position for each lipase (orange, yellow and blue for RML; orange, yellow, and cyan for TLL; orange, pink and cyan for CRL). **2)** Position of the lid region of CRL in the closed (green) and open (blue) position; red represents the position of the lid region after 20 ns of atomistic simulation of the lipase positioned at hexane- (**A**), octane- (**B**), and decane-aqueous (**C**) interfaces [99] . The catalytic triad is shown as sticks. **3) (A)**Van der Waals representations of the open crystal structure of PAL (*left*) and the structure after 20 ns of atomistic simulation in water (*right*) [100] . Two lid regions were identified after simulation, coloured in red and blue. Binding pocket residues = yellow and green. **(B)** Images representing the simulated structures at 3 ns, 7 ns, and 10 ns. Blue coloured regions = the open crystal structure, while red = the position of the lid regions during simulation. Figures were reproduced from [97] (**1**), [99] (**2**), [100] (**3**).

In this thesis, MD is used to study how different lipases, namely M37 and TLL, interact with different surfaces, including lipid membranes, triglyceride interfaces, and graphene/graphene oxide surfaces. A combination of atomistic (AT) and coarse-grained (CG) methodologies are applied, resulting in a comprehensive multi-scale study investigating how varying surface properties affect lipase dynamics.

1.5 Lipid interactions with support surfaces

Interfacial protein interactions with surfaces can be studied using supported lipid membranes (SLMs) [101]. SLMs tend to form fluid, two-dimensional structures allowing free rotational and translational diffusion of lipid molecules, making them ideal test systems to study protein adsorption, protein function, and lipid phase effects on protein localisation. For example, it was recently shown that alkaline phosphatase enzymes, tethered to the membrane via a glycosylphosphatidylinositol (GPI) anchor, tend to localise at the boundary of gel-phase domains in mixtures of DOPC and sphingolipid membranes [102]. Furthermore, SLMs generated using nanolithography techniques were used to show that proteins can specifically interact with functionalised lipid molecules within the SLM [103]. Specifically, fluorescence studies indicated that the proteins (streptavidin) interacted with biotinylated lipids, resulting in altered spreading behaviour and dynamics of the lipid membranes [103]. These studies indicate that protein-lipid interactions result in considerable effects on membrane fluidity and lipid diffusion on the surface, in addition to revealing the preferential localisation of specific proteins [102]. Given the established experimental methods to probe protein-lipid interactions using SLM systems, these could constitute interesting test systems to study lipase adsorption, focussing on how fat digestion could alter the dynamic properties of the membrane, such as surface tension, which is thought to affect lipase association at the interface [7, 104].

Interestingly, computational studies of supported lipid membranes have revealed effects of the support surface on lipid diffusion, ordering, and lateral pressure profiles, indicating that the nature of the support is important regarding the molecular properties of the membrane [105–108]. In order to investigate the potential of SLM systems in studying lipase interactions, initial molecular dynamics studies investigating lipid behaviour at different support surfaces were performed in this thesis. Specifically,

lipid behaviour at graphene and graphene oxide surfaces was investigated in different environments, and compared to similar experimental systems. The systems were investigated with a view towards establishing a model for the study of protein interactions, e.g lipases, with supported lipid membranes. The simulations are thought to constitute a first step in this direction.

1.6 Thesis Overview

This thesis presents four results chapters focussing on computational studies of lipase interactions with different surfaces, as well as lipid interactions with graphene/graphene oxide surfaces.

Chapter 3 presents a detailed investigation of how a bacterial lipase (M37) interacts with model membranes and triglyceride surfaces, focussing on the structural and functional consequences. The study provides insight into a possible activation mechanism of this lipase as a function of interfacial interactions, using enhanced sampling methods and docking studies to verify this mechanism. The results are discussed with comparison to related lipases, ending with a perspective on future studies.

Chapter 4 utilises CG and AT simulations to investigate the functional consequences of mutations in the lid region of the well-characterised *Thermomyces lanuginosus* lipase (TLL). The interfacial interactions of TLL with a triglyceride surface are studied and related back to affect of mutation, commenting on binding orientations at the interface and conformational dynamics. Finally, this chapter presents systems that mimic lipase hydrolysis of triglycerides at the interface, focussing on potential mechanisms of lipase inhibition. This data is presented in the context of experimental studies performed by collaborators at Novoyzmes, Copenhagen.

Chapter 5 builds upon the findings presented in Chapter 3, using the structural models of the M37 lipase to investigate how the enzyme interacts with graphene and

graphene oxide surfaces. M37-graphene interactions are compared to the interaction mechanisms of M37 with the phospholipid bilayers and triglyceride interfaces, focussing on the nature of the interactions and mobility on the surface.

Chapter 6 presents CG simulation data of lipid interactions with graphene and graphene oxide surfaces. Both small and large CG systems are modelled, investigating the effects of interactions with the supports on overall membrane structure. Emphasis on the molecular properties of the supported lipid membranes is given, with insight into lipid ordering and diffusion. The simulations match experimental data provided by collaborators at the University of Manchester and the Karlsruhe Institute of Technology.

Finally, Chapter 7 highlights the main findings of each of these projects, and discusses future directions.

Chapter 2

Methods

Molecular dynamics simulations were used to generate all the data presented in this thesis. Both atomistic and coarse-grained simulation methods were used, as well as enhanced sampling algorithms, to investigate the molecular interactions of lipases and lipids with different surfaces. The following presents an overview of these methods, whilst more detailed descriptions can be found in [109] and [110].

2.1 Classical Molecular Dynamics Simulations

Molecular dynamics (MD) simulations involve the propagation of a classical many-particle system based on Newton's equations of motion. The term *classical* refers to the treatment of atoms within the system, which are considered to obey the laws of classical mechanics. In this interpretation, it is assumed that the motions of electrons adapt instantaneously to the position of the nuclei they are associated with, given that the masses of atomic nuclei are much greater than the masses of electrons (Born-Oppenheimer approximation). This approximation allows atoms to be considered as single points in space, ignoring the positions of electrons, thereby also dramatically reducing the number of particles to be evaluated within a simulation. Quantum me-

chanical simulations on the other hand, do consider electrons, and are therefore limited to system sizes that rarely expand to more than multiple amino acid residues, directly constrained by the computational power necessary to propagate these systems at this level [109].

The classical treatment of atoms as single points in space allows description of atomic position by a set of Cartesian coordinates. In this description, an MD simulation can be performed by applying Newtonian equations of motion to evolve the atomic positions within a system over time, relating the position of each particle to the force acting on it:

$$F_i(t) = m_i \ddot{r}_i(t) \tag{2.1}$$

Where F_i represents the force acting on particle i, of mass m_i, with acceleration r_i at time t. This equation can be rewritten as:

$$\frac{F_{x_i}}{m_i} = \frac{d^2 r_i}{dt^2} \tag{2.2}$$

to find the motion of a particle of mass m_i along coordinate x_i in function of force F_{x_i} acting in the x direction. The MD simulation therefore predicts the positions of the particles relative to the change in their velocities, in response to the calculated forces acting on the particles. Solving the equation of motion for each particle at various time points during the simulation thus allows the generation of a series of coordinates that describe the evolution of the system over time.

2.1.1 Forcefields

In order to calculate the forces acting on each particle within a system, a description of the interactions between all the particles is necessary [109]. MD forcefields provide

such a description [109]. Each forcefield possesses a variation of the potential energy function (V), which evaluates the interaction energies between particles as a function of their atomic coordinates, r, and a set of parameters s used to describe the interactions within a system. These parameters are then used to estimate the force on atom i (F_i), calculated as the negative derivative of the potential energy function with respect to the atom coordinates, r_i:

$$F_i = -\frac{\partial}{\partial r_i} V(r_2, r_2, ..., r_n; s) \tag{2.3}$$

The potential energy of a system made up of N particles is calculated as the sum of all bonded and non-bonded interactions (s), between the atoms as a function of their coordinate positions, r:

$$V(r^N; s) = \sum V_{bonded}(r; s) + \sum V_{non-bonded}(r; s) \tag{2.4}$$

2.1.2 Bonded Interactions

Bonded interactions refer to all the interactions that occur between atoms that are linked, or bonded, to each other within a particular molecule (e.g an ethane molecule). In classical MD, the bonds between these atoms are approximated by a harmonic potential (Hooke's law), and are thus treated as particles (atoms) connected by springs (bonds) [109]. Different terms are considered when describing the bonding interactions within a molecule, including pairwise interactions between two neighbouring atoms, as well as 3-body and 4-body interactions. These terms correspond to:

- Bond stretching ($\mathrm{V_{bonds}}$);
- Bond angle bending ($\mathrm{V_{angles}}$);
- Bond torsions, or proper dihedral angles ($\mathrm{V_{torsions}}$);

- Improper dihedral angles ($V_{impropers}$).

The rotation of dihedral angles are therefore described by the torsion terms between bonded atoms, whilst improper dihedral angles are used to describe, and preserve, the stereochemistry the molecule. The total bonded contribution to the potential energy of the system can therefore be represented as:

$$\sum V_{bonded} = \sum V_{bonds} + \sum V_{angles} + \sum V_{torsions} + \sum V_{impropers} \tag{2.5}$$

Bond Stretching

As mentioned previously, the bond stretching between two atoms can be approximated by a harmonic potential. The energy of this bond varies proportionally relative to the square displacement from a reference value that describes the distance between the atoms connected by the bond [109]. This reference value is defined by the parameters in the particular forcefield used for simulation. Bonds can oscillate around their reference value, as defined by:

$$V_{bonded}(r_{ij}) = \frac{1}{2} k_{ij}(r_{ij} - b_{ij}^{eq})^2 \tag{2.6}$$

where r_{ij} represents the bond length between atoms i and j at a particular time point, b_{ij}^{eq} represents the reference bond length, and k_{ij} is the spring force constant used to restrain the bond at the reference bond length (Fig. 2.1).

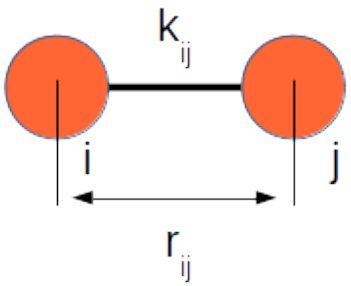

Fig. 2.1: Bond stretching between two atoms (i and j). Atoms are represented as red spheres and bonds shown as black rods.

Bond Angle Bending

The angle between three bonded atoms, *i, j*, and *k* (θ_{ijk}), is defined by a harmonic oscillator, where deviation from the reference angle value, θ_{ijk}^{eq}, results in an increase in the potential energy describing the angle:

$$V_{angles}(\theta_{ijk}) = \frac{1}{2} k_{ijk}(\theta_{ijk} - \theta_{ijk}^{eq})^2 \tag{2.7}$$

where k_{ijk} is the force constant used to maintain the angle at the reference value (Fig. 2.2).

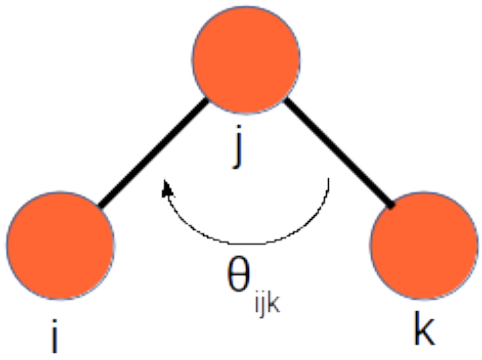

Fig. 2.2: Angle bending between three atoms (i, j, and k). Atoms are represented as red spheres and bonds shown as black rods.

Proper Dihedrals

A proper dihedral angle, or torsion, is defined between four atoms, *i, j, k*, and *l* (Fig. 2.3) and is calculated as the angle between the plane going through atoms *i, j*, and *k*, and the plane going through atoms *j, k*, and *l* (ϕ_{ijkl}). The torsion potential is derived from a cosine function where multiple reference values can be assigned:

$$V_{torsions}(\phi_{ijkl}) = k_\phi(1 + cos(n\phi_{ijkl} - \phi^{eq}_{ijkl})) \tag{2.8}$$

where ϕ^{eq}_{ijkl} is the reference value for the angle between the atoms, k_ϕ is the force constant, and n defines the multiplicity of the function.

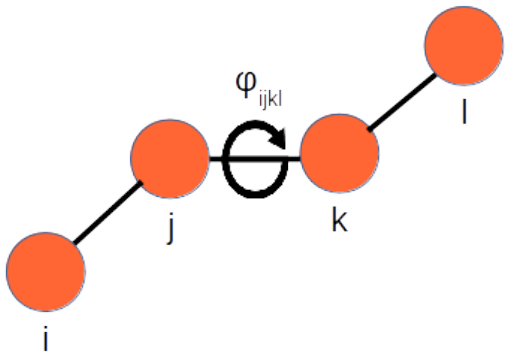

Fig. 2.3: Proper dihedral angle (torsion) between four atoms (i, j, k and l). Atoms are represented as red spheres and bonds shown as black rods.

Improper Dihedrals

The improper dihedral function is used to maintain chirality on a tetrahedral extended heavy atom, or to maintain planarity of certain bonded atoms (Fig. 2.4). This function is defined by three sequential atoms, *i*, *j*, and *k*, which are centred on a fourth atom, *l*. The function also employs a harmonic potential used to restrain the angles along the four atoms, as defined by the planes of *ijk* and *jkl*. The improper dihedral potential takes on the form:

$$V_{impropers}(\xi_{ijkl}) = \frac{1}{2}k_{\xi}(\xi_{ijkl} - \xi_{ijkl}^{eq})^2 \tag{2.9}$$

where ξ_{ijkl}^{eq} is the reference value of the angle, ξ_{ijkl}, and k_{ξ} is the force constant.

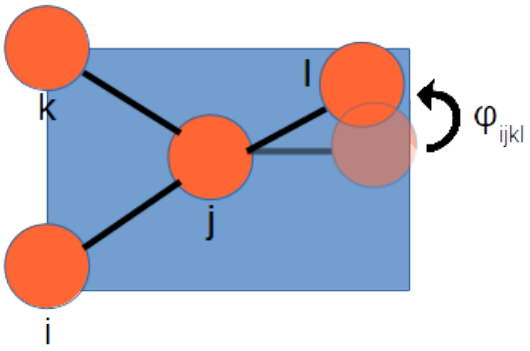

Fig. 2.4: Improper dihedral angle (torsion) between four atoms (i, j, k and l). Atoms are represented as red spheres and bonds shown as black rods.

2.1.3 Non-Bonded Interactions

Non-bonded interactions occur between all atoms within a system, independent of the specific bonding relationships between atoms [109]. Non-bonded terms are usually defined by two components: van der Waals interactions and electrostatic interactions. These terms are added together to calculate their contribution to the potential energy of the system:

$$\sum V_{non-bonded} = \sum V_{VDW} + \sum V_{electro} \tag{2.10}$$

Van der Waals Interactions (Lennard-Jones Potential)

The Lennard-Jones (LJ) potential (12-6 function) is the best known approximation used to calculate van der Waals interactions in MD simulations [109]. This potential can be rapidly evaluated, which is important given that calculation of the non-bonding interactions is the most computationally demanding task within an MD simulation. Two components are considered within this potential: an attractive component (r^6) which acts over long-range distances, and originates from instantaneous dipole in-

teractions (London dispersive forces), and a repulsive component (r^{12}), which acts over short distances, and is related to overlapping electron orbitals (Pauli exclusion principle) [109]. These terms are described by the following function:

$$V_{VDW}(r_{ij}) = 4\epsilon_{ij}\left(\left[\frac{\sigma_{ij}}{r_{ij}}\right]^{12} - \left[\frac{\sigma_{ij}}{r_{ij}}\right]^{6}\right) \tag{2.11}$$

The interaction energy between two atoms, i and j, is most favourable at a particular inter-atom distance ($\frac{\sigma_{ij}}{r_{ij}}^{6}$). However, the interaction energy rapidly increases when these atoms come closer together, resulting in the repulsive part of the function ($\frac{\sigma_{ij}}{r_{ij}}^{12}$) . The depth of the well describing the interaction energy between two particular atoms is defined by ϵ_{ij}, representing the attractive forces between atoms i and j. The distance between the atoms is defined by r_{ij} (Fig 2.5) and σ_{ij} defines the distance at which the energy between the two atoms equals zero (collision distance).

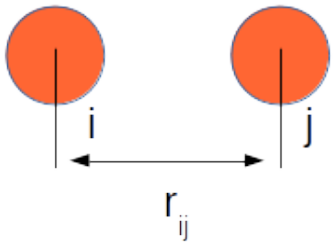

Fig. 2.5: Van der Waals interactions between a pair of atoms (i and j). Atoms are represented as red spheres and bonds shown as black rods.

One disadvantage of the LJ potential, in terms of computational cost, is that it has an infinite range, where the forces become infinitely weak at long distances. This is highly demanding to compute as the forces must be evaluated for every pair of particles within a simulation. A cut-off is therefore introduced to reduce the number of calculations performed for atom pairs at large distances. In this thesis, a cut-off is placed at 1.2 nm and a smoothing of the potential from 0.9 nm to avoid artefacts

arising from an abrupt cut-off.

Coulombic Interactions

Electrostatic interactions within a system arise through differences in electronegativity of atoms, resulting in an unequal charge distribution across the molecule (polarisation). Charge distribution is often represented as fractional point charges throughout a molecule, which are designed to reproduce the electrostatic properties of the molecule. In classical MD simulations, the point charges are often confined to the nucleus of the atoms, referred to as partial charges (Fig. 2.6) [109]. The potential energy arising from the electrostatic contributions between two atoms, i and j, at distance r_{ij} within a system is usually evaluated using Coloumb's law:

$$V_{electro}(r_{ij}) = \frac{q_i q_j}{4\pi\epsilon_0\epsilon_1 r_{ij}} \tag{2.12}$$

Where the partial charges on each atom are represented by q_i and q_j, whilst ϵ_0 is a term related to the permittivity of free space, and ϵ_1 is related to the relative permittivity of the medium.

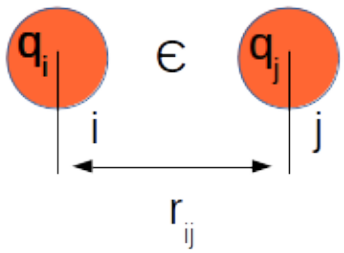

Fig. 2.6: Electrostatic interactions between a pair of atoms (i and j), where ϵ represents the dielectric constant of the surrounding medium. Atoms are represented as red spheres and bonds shown as black rods.

Electrostatic interactions generally extend over a much larger distances than van der Waals interactions, resulting in a huge computational cost if they were to be

calculated for every atom pair within a system. The treatment of electrostatics can be addressed by employing different methods, which generally involve the use of a cut-off. These are discussed later on.

Potential energy function

Considering all bonding and non-bonding interaction terms, the general form of the potential energy function applied in most forcefields corresponds to [109]:

$$V(\mathbf{r}^N) = \sum_{bonds} \frac{k_{ij}}{2}(r_{ij} - b_{ij}^{eq})^2 + \sum_{angles} \frac{k_{ijk}}{2}(\theta_{ijk} - \theta_{ijk}^{eq})^2 + \sum_{torsions} k_\phi(1 + cos(n\phi_{ijkl} - \phi_{ijkl}^{eq})) + \sum_{impropers} \frac{k_\xi}{2}(\xi_{ijkl} - \xi_{ijkl}^{eq})^2 + \sum_{i=1}^{N}\sum_{j=i+1}^{N}\left(4\epsilon_{ij}\left[\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6}\right] + \frac{q_i q_j}{4\pi\epsilon_0\epsilon_1 r_{ij}}\right) \tag{2.13}$$

2.1.4 Integration Methods

The potential energy function provides a description of the forces between the atoms in a particular system. Once this description is obtained, MD simulations employ finite difference methods to integrate the equations of motion to compute the positions of atoms accordingly. Integration is broken down into many discrete, small stages, each separated by a fixed time δt [109]. The total force on each particle is then calculated as the vector sum of its interactions with other particles. The accelerations of each of the particles is then calculated as a function of these forces, which are combined with the positions and velocities of the atoms at time t to calculate the positions and velocities at time $t + \delta t$.

A variation of the *Verlet* algorithm is used to integrate the equations of motion in this thesis, called the leap-frog algorithm [111]. The leap-frog algorithm calculates the atomic positions ($\mathbf{r}$) at time $t + \delta t$ by evaluating atomic velocities ($\mathbf{v}$) at half time

step intervals ($\frac{1}{2}\delta t$), taking on the following forms:

$$\mathbf{r}(t+\delta t) = \mathbf{r}(t) + \mathbf{v}(t+\frac{1}{2}\delta t)\delta t \tag{2.14}$$

$$\mathbf{v}(t+\delta t) = \mathbf{v}(t-\frac{1}{2}t\delta t) + \delta t\mathbf{a}(t) \tag{2.15}$$

where atomic accelerations ($\mathbf{a}$) are calculated by applying equation 2.2. The algorithm is implemented by first calculating the velocities at t - $\frac{1}{2}\delta t$, and the accelerations at time t to predict the velocities at time t + δt (2.15). These velocities are then used to derive the positions of the atoms at $\mathbf{r}(t+\delta t)$ using the positions at $\mathbf{r}(t)$ (2.14). Therefore, the velocities are calculated between t - δt and $t + \delta t$, thus "leap-frogging" over the atomic positions.

When initiating a MD simulation from an initially static state, random velocities are assigned to each particle to provide "starting" velocities, obtained from a Maxwell-Boltzmann distribution at a user-defined temperature.

2.1.5 Forcefields

The ability of an MD simulation to provide useful results relies heavily on the accuracy of the interaction parameters employed by the forcefield used to perform the simulation. These differ depending on the forcefield used for simulation. The development and parameterisation of forcefields, as well as their appropriate use, is therefore of great importance. Generally, forcefields are parameterised on as much experimental data as possible. The references for bonds and angle values are generally obtained from X-ray crystallography data of small and large molecules, and their corresponding force constants from spectroscopy data, such as infrared spectroscopy [109, 112]. Furthermore, thermodynamic data are used in the parameterisation of small molecules

in the OPLS and GROMOS forcefields, whilst NMR data are used in the AMBER forcefields [112–114]. Quantum mechanical (QM) calculations are often used to supplement the available experimental data, for example in generating torsion parameters and atomic partial charges [109].

Given that individual forcefields can employ different parameterisation methodologies, often optimised to be used under specific conditions, the mixing or transfer of parameters between forcefields should be avoided [115]. Furthermore, forcefields can inherently differ in their description of particular molecules, specifically in their treatment of hydrogen atoms. For example, all-atom forcefields such as CHARMM and AMBER treat hydrogen atoms explicitly [116, 117], whereas "united-atom" forcefields such as GROMOS treat hydrogen atoms implicitly by combining their parameters with their associated parent carbon atom parameters [112]. This approach reduces the computational cost of performing the simulation, whilst maintaining accuracy in simulating systems for which the forcefield was intended. It is therefore important to consider the general applicability of a particular forcefield to a specific system before conducting the simulation.

2.2 Simulation Methods

Many techniques employed in MD simulations focus on optimising accuracy and/or reducing the computational cost of running the simulation. The methods employed in this thesis are listed below.

2.2.1 Time Step

The time step determines how frequently the equations of motion are integrated during the simulation, and can differ depending on the simulation method, forcefield, and

system. There is a balance in choosing a time step that is both accurate and efficient; too large a time step will result in simulation errors, whilst very small time steps will considerably limit the amount of phase space sampled within a particular time frame. A generally applied rule recommends using a time step that is ~10 times smaller than the fastest motion occurring in the system. In flexible molecules, the highest-frequency vibrations are generally due to bond stretches, particularly those of bonds to hydrogen atoms. For example, a C-H bond vibrates with a repeat period of ~10 fs [109].

As will be discussed in a future section, constraint algorithms can be applied to increase the time step used to integrate Newton's equations of motion, reducing the computational load of simulating the system by reducing the number of times the equations must be integrated. In this thesis, a time step of 2 fs is applied for all atomistic simulations, whilst a time step of 10-20 fs is applied for all coarse-grained simulations. The use of a larger time step for coarse-grained simulations will be discussed in a future section.

2.2.2 Periodic Boundary Conditions

A simulation system can often be restricted in size regarding computational feasibility. Periodic boundary conditions (PBC) are a method commonly employed to reduce the number of particles simulated in a system, whilst still mimicking bulk behaviour. To achieve this, PBC replicates a central simulation box in all dimensions around the central box and allows particles to diffuse in and out of the box, reappearing on the opposite side (Fig. 2.7). By applying this method, there are no edge restrictions on the central simulation box that would otherwise produce artefacts in the system. For example, simulation of a box of water particles of finite size without PBC would result in the water coming into contact with the surrounding vacuum at the box edge, producing unnatural forces on the atoms and spurious results [109].

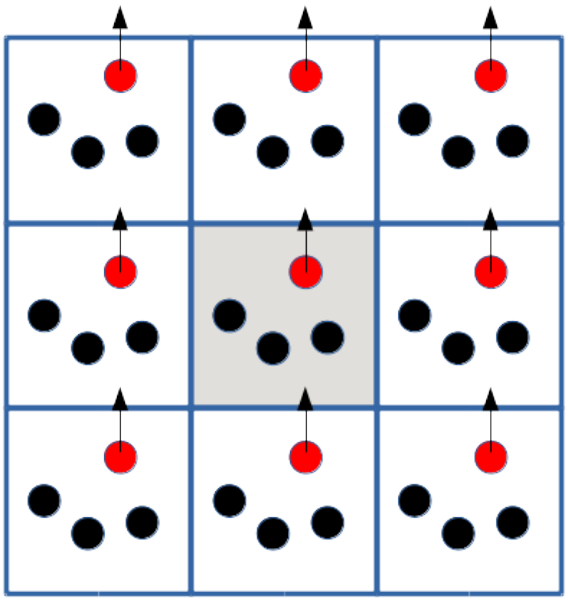

Fig. 2.7: Image demonstrating the concept of periodic boundary conditions. The central simulation box (shaded grey) is replicated in the x, y, and z dimensions. The particles (black and red dots) are able to diffuse from one simulation box into the next, through the periodic boundary conditions. Their periodic image is then reproduced in the neighbouring box, as demonstrated by the red dot.

Given that particles are able to diffuse through the periodic boundaries, the interactions occurring between atoms within the system also function across these boundaries. Care must be taken to ensure that the initial size of the simulation box is large enough such that one particular atom does not see its own image through the periodic boundary, resulting in unphysiological interactions. Here, the minimum image convention can be applied in combination with cut-off schemes that limit the number of particle interactions considered within a particular time step, avoiding artefacts.

2.2.3 Temperature and Pressure Coupling

In order to replicate general experimental conditions, MD simulations are usually performed in the NPT ensemble, in which the number of particles (N), pressure (P),

and temperature (T) of a simulation are kept constant [109]. Other commonly sampled ensembles include the NVT ensemble, in which the volume (V) of the simulated system is kept constant, and pressure is allowed to fluctuate. All simulations performed in this thesis were performed using either the NPT or NVT ensemble.

Temperature and pressure algorithms are used to maintain an overall constant energy of the system, and can be set to reproduce various conditions to match experiments. In this thesis, atomistic simulations were performed using the Velocity-rescale thermostat and the Parrinello-Rahman barostat [118, 119]. Coarse-grained simulations were performed using the Berendsen thermostat and Berendsen barostat [120].

2.2.3.1 Thermostat Algorithms

The effects of temperature on the evolution of a system in MD simulations are implemented by relating temperature to the time average of system's kinetic energy [109]. The kinetic energy (K) of a system is a function of the momenta ($\mathbf{p}$) of the individual particles and their masses (m), as shown by:

$$K(\mathbf{p}; m) = \sum_{i=1}^{N} \frac{\mathbf{p}_i^2}{2m_i} = \sum_{i=1}^{N} \frac{1}{2} m_i \mathbf{v}_i^2 \tag{2.16}$$

where v represents the velocity of particle i. Temperature can therefore be approximated in the MD simulation by scaling the velocities of the particles in function of their kinetic energy at a defined temperature. Thermostat algorithms achieve this by coupling the system to an external heat bath that is fixed at the desired temperature (T_{bath}), acting as a source of thermal energy, supplying or removing heat from the system as appropriate. The velocities are scaled at each time step, employing a coupling constant (τ) that determines how tightly the bath and the system are coupled

in maintaining the system at a particular temperature (T) at time t:

$$\frac{dT(t)}{dt} = \frac{1}{\tau}(T_{bath} - T(t)) \tag{2.17}$$

The change in temperature (ΔT) between successive time steps (δt) is therefore given by:

$$\Delta T = \frac{\delta t}{\tau}(T_{bath} - T(t)) \tag{2.18}$$

and the scaling factor for the velocities is defined as:

$$\lambda = 1 + \frac{\delta t}{2\tau}\left(\frac{T_{bath}}{T(t)} - 1\right) \tag{2.19}$$

If τ is large, the coupling between the external bath and the system is weak, allowing temperature to fluctuate, whereas if τ is small, the coupling is stronger and fluctuations are minimised. This relatively simple scheme is employed by the Berendsen thermostat [120], which is stable and easy to implement. However, this thermostat does not possess a conserved quantity and is not associated to a well-defined ensemble. Despite these limitations, it has been shown that the Berendsen thermostat yields approximately correct results for many calculated properties for large systems (on the order of 100s-1000s of particles) [121]. Furthermore, this thermostat was used in the parameterisation of the coarse-grained forcefield employed in this thesis, as will be discussed below. Consequently, for reasons of consistency and efficiency, the Berendsen thermostat was employed for all coarse-grained simulations performed in this thesis, which consisted of larger systems composed of several thousands of particles. Importantly, separate coupling baths are used for different parts of the system (e.g water molecules are coupled separately from protein molecules) in order to avoid the problem of "hot solvent, cold solute", in which the "temperature" of the solvent is higher than that of the solute, even though the overall temperature of the system is at the target

value.

Variations of the Berendsen thermostat have been developed to address the aforementioned limitations, particularly for smaller systems. One such thermostat is the Velocity-rescale thermostat, which introduces an additional stochastic term that enforces sampling of the correct kinetic energy distribution [118]. Specifically, a defined quantity is implemented in the algorithm that can be used to verify that the simulation generates configurations belonging to the desired ensemble [118]. This is achieved by applying the following procedure:

1. Evolve the system for a single time step using the equations of motion;
2. Calculate the kinetic energy;
3. Evolve the kinetic energy using continuous stochastic dynamics;
4. Rescale the velocities in order to enforce this new value of the kinetic energy.

Rescaling of the velocities is implemented by calculating a rescaling factor (λ) that enforces a canonical distribution of the kinetic energy (K). The target value of the kinetic energy (K_t) is selected using the stochastic procedure, and the rescaling factor is defined as:

$$\lambda = \sqrt{\frac{K_t}{K}} \tag{2.20}$$

This method then rescales the velocities according to λ, simulating the effect of temperature within the system. The Velocity-rescale thermostat is employed in all atomistic simulations performed in this thesis.

2.2.3.2 Barostat Algorithms

Just as a macroscopic experimental system changes volume in response to pressure, the volume of a simulation box can be scaled to maintain a constant pressure in the NPT ensemble. Similar to the thermostat algorithms mentioned above, constant pressure

can be implemented by coupling the system to an external pressure bath:

$$\frac{dP(t)}{dt} = \frac{1}{\tau_p}(P_{bath} - P(t)) \tag{2.21}$$

in which *P(t)* is the pressure component at time t, P_{bath} is the pressure of the pressure bath, and τ_p the coupling constant. The volume of the simulation box is then scaled by a factor λ and is related to the isothermal compressibility of the system (κ):

$$\lambda = 1 - \kappa\frac{\delta t}{\tau_p}(P(t) - P_{bath}) \tag{2.22}$$

This method is employed in the Berendsen barostat [120], which is applied in coarse-grained molecular dynamics simulations reported in thesis. This was chosen, similar to the Berendsen thermostat, for reasons of consistency and efficiency, matching the algorithms used in parameterisation of the coarse-grained forcefield used for simulation. Furthermore, in order to maintain the correct structural properties of phospholipid bilayers (e.g surface tension, area/lipid), semi-isotropic pressure coupling is used for all phospholipid bilayer simulations (x-y plane of the bilayer is coupled to a separate pressure bath than the z plane), in accordance with the simulation parameters for similar systems using the coarse-grained forcefield [122].

Conversely, the Parrinello-Rahman barostat was employed in all atomistic simulations in this thesis. This barostat extended the Berendsen method by making each unit vector of the simulation box independent, allowing dynamic shape change of the box in an anisotropic manner [123]. The algorithm was initially formulated to study systems under conditions of stress, such as external pressure [123]. In order to allow dynamic shape change of the simulation box in response to that stress, the shape and volume of the box was described by three vectors (**a, b, c**), represented by a 3x3 matrix (**h**) with columns representing the components of **a, b** and **c**. The box volume

(V) was therefore given by:

$$V = \det \mathbf{h} = \mathbf{a} \cdot (\mathbf{b} \times \mathbf{c}) \tag{2.23}$$

where det = the determinant of the **h** matrix. Consequently, the position of particle i ($\mathbf{r}_i$) could be written in terms of **h** and a column vector $\mathbf{s}_i$, with components ξ_i, η_i, and ζ_i, as:

$$\mathbf{r}_i = \mathbf{h}\mathbf{s}_i = \xi_i\mathbf{a} + \eta_i\mathbf{b} + \zeta_i\mathbf{c} \tag{2.24}$$

where $0 \leq \xi_i,\ \eta_i,\ \zeta_i \leq 1$. The square of the distance between particles i and j is given by:

$$r_{ij}^2 = \mathbf{s}_{ij}^T \mathbf{G} \mathbf{s}_{ij} \tag{2.25}$$

where **G** represents the metric tensor at temperature (T):

$$\mathbf{G} = \mathbf{h}^T\mathbf{h} \tag{2.26}$$

Using these representations of box volume and shape, the equations of motions to implement the Parinello-Rahman barostat were derived from the Lagrangian (L):

$$L = \frac{1}{2}\sum_{i=1}^{N} m_i \dot{\mathbf{s}}_i^T \mathbf{G} \dot{\mathbf{s}}_i - \sum_{i=1}^{N}\sum_{j>i}^{N} U(r_{ij}) + \frac{1}{2} W Tr \dot{\mathbf{h}}^T \dot{\mathbf{h}} - pV \tag{2.27}$$

where p is the pressure imposed on the system, W is a constant with the dimensions of mass, m_i is the mass of particle i, U represents the potential energy of the system, and Tr is the trace operation on matrix **h**.

Use of the Parinello-Rahman barostat matches the algorithms employed in parameterisation of the GROMOS forcefields used for atomistic simulations within this thesis. All simulations were performed at a reference pressure of 1 bar.

2.2.3.3 Particle Mesh Ewald

One of the most computationally demanding tasks in an MD simulation is calculation of the electrostatic interactions between the atoms within the system. These interactions are long-range and scale exponentially with the number of particles within the system. The small rate of decay of the electrostatic potential additionally prevents the use of hard cut-offs when considering methods to reduce the number of calculations required to accurately model the system. The Particle Mesh Ewald (PME) method is an electrostatic summation method that addresses this issue [124]. Here, the interaction is split into two parts, a short range potential, and a long range potential. The short range potential is calculated in real space, which converges quickly, whereas the long range potential is calculated in Fourier space, allowing this component to also converge quickly. Given that both these components converge relatively rapidly, the number of electrostatic interactions considered for a particular atom can still be limited by imposing a cut-off, thus only losing slightly in accuracy, but gaining considerably in efficiency.

All atomistic simulations were performed using the PME method for calculation of electrostatic interactions.

2.2.3.4 Constraints

Atomistic systems generally contain flexible molecules, which involve translation, rotation, torsion, and vibrational motions. However, these high-frequency motions (e.g bond stretches) within the system are usually of less interest than lower frequency modes (e.g conformational changes). Constraint algorithms may be applied to constrain, or "fix", certain degrees of freedom within a system, allowing the use of a larger time step and reducing the frequency at which Newton's equations of motion are integrated during the simulation. For example, the distance between two hydrogen atoms

within a water molecule can be constrained to maintain a certain value, which is simply reset at each time step such that this value remains constant. Constraints are particularly useful in simulations of biological systems that involve large numbers of water molecules, in which evaluation of each particular H-O bond at each time step would impose a large computational demand.

However, it is important that the constrained degrees of freedom are only weakly coupled to the remaining degrees of freedom, such that the constraint algorithm does not compromise the overall flexibility and motion of the molecule. In this thesis, constraints are only applied on bonds involving hydrogen atoms (high-frequency) using the LINCS algorithm [125], enabling the time step to be increased from 1 fs to 2 fs for atomistic simulations.

2.2.3.5 Position Restraints

Similar to constraints, position restraints can be applied to restrain particular atoms in their x, y, or z coordinates within a system. Position restraints work by applying harmonic potentials to minimise the deviation of a particular angle or bond from a specific value by imposing a large energetic penalty that scales with a user-defined force constant. Position restraints are typically applied during the equilibration phase of preparing a system for simulation, and are usually applied to the backbone atoms of, for example, a protein molecule, allowing the rest of the system to relax around this molecule. Restraints can also be used during the production simulation.

2.3 Energy Minimisation

Energy minimisation is an essential method used to generate an acceptable system configuration prior to simulation using MD [109]. Energy minimisation allows the particles within a system to adjust their positions and spatial relationships, such that

the forcefield parameters defining these interactions are more accurately matched, without performing a full MD simulation. For example, the side chains within a protein molecule are able to adapt their locations to find the closes energy minimum during energy minimisation, reducing steric clashes with other system particles such as water molecules. This occurs in an iterative manner, prompting the system to find respective energy minima until a user-defined number of iterations is reached, or until the algorithm converges to a value that is smaller than a user-defined maximum force (e.g 10 kJ mol^{-1}).

The steepest descent algorithm is applied in this thesis. The method uses a line-search approach to find energy minima, which is independent from the previous step, and most effective for energy minimisation of systems far from their minima.

2.4 Coarse-grained Forcefields

Much of this chapter has described approaches used to reduce the computational time required to perform an MD simulation. Despite these efforts, atomistic simulations, in which each atom within a system is explicitly represented, are typically restricted to several hundred nanoseconds on readily available computational resources. Whilst this allows certain molecular properties to be investigated in detail, larger-scale systems often require longer simulations times to achieve equilibrium and/or to reveal interesting properties. Many biological behaviours, such as protein dimerisation or association with interfaces, occur on longer timescales, often up to tens of microseconds, and thus remain inaccessible using atomistic simulations on standard computational resources. Furthermore, depending on the property of interest, many systems require simulation over multiple repeats in order to obtain convergence and provide adequate statistical sampling, thus necessitating the generation of large simulation ensembles [126]. Atomistic simulations are therefore not only limited by the total number of particles

within a system, but also by the number of repeat simulations necessary to obtain convergence for certain properties, such as the packing mode of two transmembrane helices [126].

Coarse-grained (CG) forcefields have been developed in order to address some of the limitations imposed by the computational demands of atomistic simulations. These methods tend to employ more approximate representations of the particles within a system to significantly reduce the number of particles to be simulated, and commonly implement short-range potentials for evaluation of non-bonding interactions, thereby dramatically reducing the computational time required to perform an MD simulation. Consequently, large-scale systems and events, such as bacterial signalling [127], lipid packing within whole virus particles [128, 129], and larger sampling of protein-surface binding interactions [130, 131], become accessible on relatively standard computational setups. These simulations allow evaluation of system properties, such as clustering of lipids and diffusion within complex membranes [128, 130], that would otherwise be very difficult to access using atomistic simulations. Naturally, the results from CG simulations need to be validated by comparison to experimental measurements, or indeed atomistic simulations of (smaller) systems. Encouragingly, it has been shown that certain coarse-grained systems perform well, reproducing experimental properties such as protein-lipid clustering and membrane dynamics [130, 132,], demonstrating the utility of CG simulations. In this thesis, the Martini coarse-grained forcefield (version 2.2) is used to simulate all CG systems.

2.4.1 Martini Forcefield

The Martini forcefield is an off-lattice model that, on average, maps 4 heavy atoms into a single interaction site or "bead" [133]. Importantly, hydrogen atoms are not considered within the Martini forcefield, given their small size and mass. For example,

a butane molecule would be represented as one CG particle, such that the hydrogen atoms associated with the carbon atoms are not taken into account within the 4:1 mapping scheme.

The first version of the Martini model was parameterised based on structural, dynamical, and thermodynamic data for simulations of lipid systems. Four different bead types were introduced in order to represent the underlying atomistic structure, based on chemical nature:

- polar (P) beads representing neutral atom groups easily dissolved in water;
- nonpolar (N) beads representing hydrophobic moieties;
- apolar (C) representing mixed groups of both polar and apolar nature;
- charged (Q) representing ionised groups.

In addition to these four bead types, a further four subtypes can be defined for the polar and charged particles (0, d, a, da) representing hydrogen bonding capabilities, where d = donor, a = acceptor, 0 = neither, da = both. For example, a non-polar group able to accept and donate hydrogen bonds, such as 1-butanol, is represented by an Nda bead. The masses of the individual beads are also set to 72 atomic mass units (amu) to enhance computational efficiency [133].

Given the 4:1 mapping scheme, most lipid molecules are represented by 12-16 bead particles, depending on the length of hydrocarbon tails (Fig 2.8).

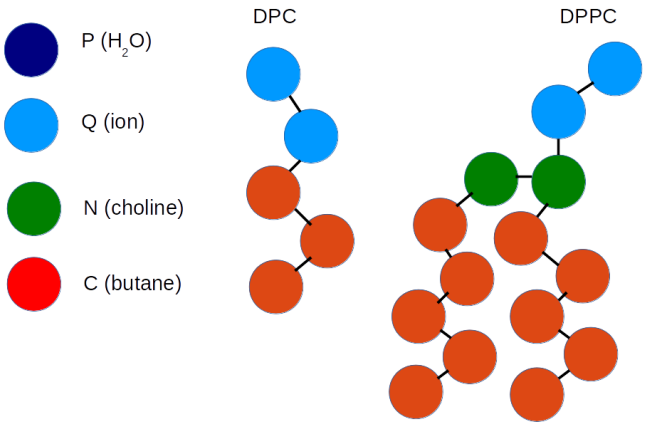

Fig. 2.8: Representations of the bead types within the Martini forcefield, showing water, ions, DPC (dodecylphosphocholine), and DPPC (dipalmitoylphosphatidylcholine). Figure based on [133].

Bonded interactions between particles within the same molecule are treated using the same principles as atomistic forcefields, and are modelled by harmonic potentials and restraints. Non-bonded interactions are described using the LJ potential, where the interaction strength between individual bead types is defined by altering the σ (effective minimum distance) and ϵ energy well depth parameters of this potential. An interaction matrix is used to define how each bead interacts with all the other bead types. Electrostatic interactions are treated using the Coulomb potential where only groups bearing a full, or close to full, charge are evaluated by the Coulomb potential; other small partial charges are implicitly represented within the LJ potential interactions. LJ and electrostatic interactions are shifted to a value of zero between 0.9 nm to 1.2 nm for LJ interactions, and between 0 and 1.2 nm for Coulomb interactions. Therefore, long range electrostatic interactions are not taken into account and a dielectric constant is imposed to mimic the surrounding environment outside of this cut-off [133].

Parameterisation of the non-bonding interactions terms was mainly based on thermodynamic data such as free energy of hydration, vaporisation, and partitioning free

energy between polar and apolar phases for each of the bead types. These data were compared to experimental values of similar molecules, and optimised to reproduce the experimental values as closely as possible [133]. For example, the partitioning free energy of water in hexadecane of 24 ± 2 kJ mol^{-1} calculated from CG simulation compared well to the experimental value of 25 kJ mol^{-1}. Simulations of self-assembled lipid systems allowed the calculation of other properties, such as area/lipid and lipid order parameters, which were compared to atomistic simulations and experimental values of the same systems. Both thermodynamic and structural properties were semi-quantitatively reproduced within the model, allowing accurate reproduction of, for example, the phase transitions of lipid systems [133].

However, although the variables of the CG system (e.g densities, energies, temperature, length scales) keep their physical meaning compared to atomistic and experimental systems, this is not necessarily true for the time scale. Due to the reduced number of interactions sites within CG representations, the dynamics of the CG system are faster compared to atomistic systems. Specifically, the reduced number of degrees of freedom and larger particle sizes result in smoothening of the free energy landscape underlying system, such that the friction arising from the atomic degrees of freedom are simply missing [134]. Importantly, this factor affects *all* of the dynamics in the system. For example, the diffusion of CG water particles (representing four individual atomistic water molecules) is effectively sped up by a factor of 4 compared to real water. This means that the diffusion of CG water only compares well to that of real water after scaling by a factor of 4. Calculation of other diffusion constants in CG systems reveal that the effective time sampled by the Martini CG model can be up to 2- to 10-fold larger than atomistic systems [134]. Interestingly, global events such as the aggregation of lipids into vesicles, also occur on a comparable time scale to atomistic simulations after a 4-fold scaling [135], however sampling of the configurational

space of liquid hydrocarbons is observed to be 5- to 10-fold faster in CG simulations *vs* atomistic simulations [136]. Consequently, the interpretation of time scale within CG simulations is not straightforward. As a first approximation, a standard conversion factor of 4 is generally applied to compare the time scales of sampling in CG simulations with atomistic simulations, however it is important to be aware that time scale has a different meaning in CG simulations compared to experimental/atomistic data [133, 134]. Thus, all of the simulation times reported in this thesis correspond to the unscaled simulation times, and were not subject to any conversion factors.

Furthermore, the simulations described in this initial version of the Martini forcefield were performed using a time step of 50 fs, which is much larger compared to 2-5 fs time step generally applied in atomistic simulations. The reduced number of particles, use of short-range potentials, and increased dynamics due to the nature of the CG model allows a speed up of 3-4 orders of magnitude compared to atomistic simulations. However, as will be detailed below, 10-20 fs time steps are more commonly applied in order to maintain numerical stability of the simulations, particularly as system complexity is increased.

Martini Forcefield: Extension to Proteins

Further parameterisation of the Martini forcefield over the years has resulted in elaboration of the systems that can be simulated using this forcefield. With the publication of the Martini 2.1 forcefield, additional bead types were introduced that were able to model ring particles, such as the side chain of histidine residues, allowing the simulation of proteins [122]. The mapping scheme was adjusted for these structures, employing either a 3:1 or 2:1 scheme. Consequently, the van der Waals radii (σ) of beads representing ring structures were reduced from 0.47 nm to 0.43 nm, and the well-depth (ϵ) of the LJ potential was scaled to 75% of the original value. These beads

were termed "special" beads, designated with an "S" prefix in their classification (e.g SC1).

Furthermore, an additional description of 5 subtypes was implemented denoting the degree of bead polarity, ranging from 1 (low) to 5 (high) for each of the standard Martini bead types (P, N, C, and Q). A mapping scheme was devised for all 20 amino acids, mainly based on water/oil partitioning data of amino acid analogues compared to experimental measurements of similar molecules, providing a basis for parameterisation of the non-bonding interactions [122]. Potential of mean force calculations were employed in cases where different particles produced similar results, allowing refinement of the parameters of different amino acids. Consequently, the total range of potential energies describing the interactions were increased from 5 levels (I = attractive - V = repulsive) to 10 separate levels (O - IX) to more accurately differentiate between different interaction energies. The resultant bead representations for all the amino acids modelled within the forcefield, classified according to chemical nature, are shown in Fig. 2.9.

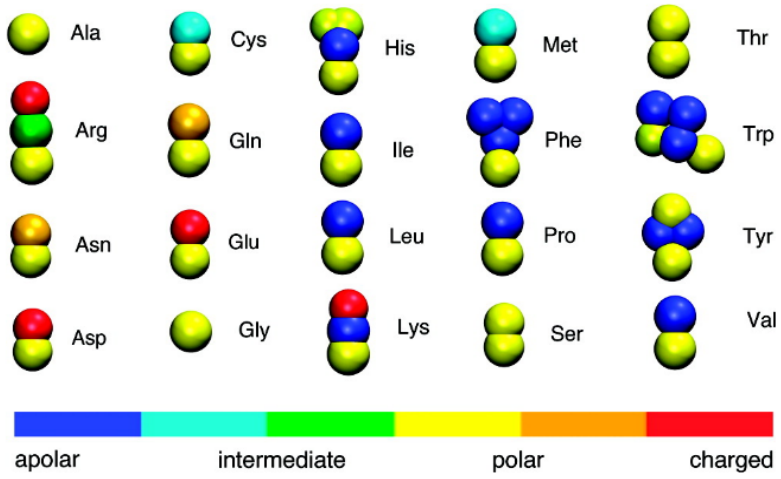

Fig. 2.9: Representations of the amino acid bead types within the Martini 2.1 forcefield, highlighting the difference in chemical nature between each representation. Figure reproduced from [122].

Parameterisation of the bonding interactions between amino acids was based on a set of 2000 protein structures extracted from the PDB. The secondary structure of every residue was determined using the DSSP program [137], providing distributions for bond lengths, bond angles, and dihedral angles for all combinations of amino acids and secondary structures [122]. The bonded parameters for the CG models were then assigned to match the bond length, angle, and dihedral distributions for each secondary structure type calculated from the reference PDB protein structure basis set, and further analysed by extracting another set of protein structures from the PDB and comparing the CG model distributions to this new parameter set. The results were found to be nearly indistinguishable from the original basis set. Subsequently, multiple simulations were performed on short test peptides, in which the CG bead was placed at the centre of mass for the mapped group (on average, 4 atoms), in order to provide parameter values for the force constants employed in the bonded potentials.

Some approximations were made in modelling a particular secondary structure type in order to minimise the number of parameters, such that all beta structures are approximated as extended, while 3_{10}-helix and π-helix are modelled as α-helices. Furthermore, the backbone-backbone bond lengths were all set to 0.35 nm regardless of secondary structure, matching the average calculated from PDB crystal structures [122]. Other bonded parameters, such as angles, were set to be dependent on secondary structure for backbone-backbone bonds, however dihedral angles were only imposed when four interacting beads corresponded to the same secondary structures, as assigned by the DSSP program [137]. Conversely, backbone-side chain and side chain-side chain bond lengths and force constants were independent of secondary structure, and amino acid dependent. The dihedral angle interaction potentials thus attempt to restrain protein secondary structure in order to match the starting, input structure of the protein. Currently, a freely available program (martinize.py) can be used to

"coarse-grain" an input protein structure, and was used to generate all CG protein structures in this thesis.

Due to the substantial extension of the parameter set within the Martini 2.1 forcefield, the recommended simulation time step is reduced to 10-20 fs depending on system complexity and stability. The CG simulations performed in this thesis, which involved proteins, were performed with this in mind.

Further extension of the Martini 2.1 forcefield to the Martini 2.2 forcefield addressed errors and inaccuracies in the earlier version [138]. Most importantly, the forcefield reparameterised numerous protein side chain beads to more accurately match atomistic simulations, particularly for side chains with rings such as phenylalanine residues. Additionally, the optimised parameter set addressed binding behaviour of proteins and peptides, including interfacial orientation, at membrane surfaces [138]. All CG simulations in this thesis were performed with the Martini 2.2 forcefield.

Elastic Networks

Elastic network models were originally introduced to describe the motions and flexibility within biomolecules, including protein conformational changes and dynamics, as an alternative to normal mode analysis. Traditional elastic network models describe the protein structure as a network of point masses connected by springs, representing the bonded interactions. These models were strongly influenced by previous work describing the dynamics of polymers [139], and were shown to reproduce the frequency spectrum of proteins remarkably well [140]. The Gaussian Network Model was introduced as a variant of the elastic network model at the residue level, aiming to predict amino acid fluctuations based on experimental β-factors (thermal fluctuations), thus assuming that amino acid fluctuations are isotropic and gaussian [141]. These interactions are applied to residues located within a particular range, or cut-off distance. Consequently, the dynamics of the protein is entirely defined by the network

topology [141].

Simulations employing the Martini forcefields (v2.1 and v2.2) have revealed that the dihedral angle potential is inadequate to model, and maintain, protein secondary structure, functioning poorly when applied to large proteins, or soluble proteins that are not stabilised by a membrane environment. This problem was addressed by the application of elastic network models. Currently, two main elastic network models can be applied for CG systems [142, 143]:

- Traditional Elastic Network Models: these apply additional harmonic bonds to any two backbone particles that are within a certain distance cut-off. It is possible to tune the elastic bonds by altering the force constant used for the harmonic bond potential, as well as the upper and lower distance cut-off that defines where the harmonic bonds are placed. For most systems, the spring force constant and the distance cut-off remains the same for all particles over the network of points.
- ElNeDyn Elastic Network Model: this newer model incorporates a change in the location of the backbone beads within the mapping scheme, placing them where the Cα carbons usually reside [143]. Previously, the beads were positioned at the centre of mass of the protein backbone, within the context of the 4 atoms to which the bead is mapped [122]. This model emphasises the use of a global elastic network, and aims to remain close to the overall input protein structure than traditional elastic networks. Harmonic bonds are only applied to backbone beads, and two backbone beads are only linked by the spring if the distance between them in the crystal structure is less than the user-defined cut-off, and if they are separated by at least two positions in the protein structure. Additionally, the force constant applied within the harmonic bond potentials can be user-defined (e.g 50 kJ mol^{-1} Å^{-2}). Both the distance cut-off and the spring force constant are therefore tuneable within the ElNeDyn model, and are often parameterised

based on atomistic simulations or comparisons with experimental data, such as NMR data.

In this thesis, the ElNeDyn network model is used to model protein secondary structure [143]. The parameters for cut-off and force constants are individually specified depending on the system, and optimised based on atomistic simulations of the same system.

Nanoparticle Simulations

Over the last 10 years, the Martini forcefields have been applied to simulate nanoparticle systems, including fullerenes, carbon nanotubes, and graphite [144–147]. In 2013, Gobbo *et al* published a CG model of graphite that could be used to study the adsorption and dynamics of long-chain organic molecules from solution onto the hydrophobic surface [84]. The model aimed to provide insight into self-assembly of biological molecules on support surfaces for nanopatterning and design of composite materials. The model was parameterised based on adsorption enthalpies of small molecules from the gas phase and on wetting enthalpies of pure compounds, including butane, octane, hexadecane, benzene, butanol, etc. Following this, the model was able to reproduce the order-disordered transition of hexadecane and hexadecanol (Fig. 2.10), as well as preferential adsorption of long-chain organic compounds from organic solvents, forming lamellar arrangements on the surface [84].

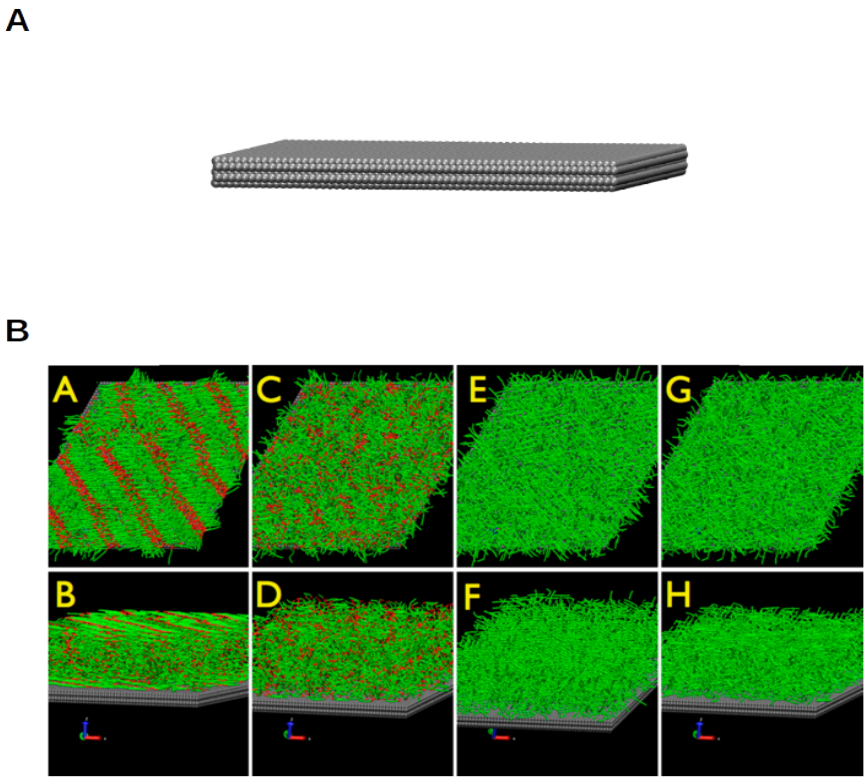

Fig. 2.10: **(A)** Van der Waals representation (grey) of the Martini graphite model (5 layers). **(B)** Order-disorder phase transition of pure hexadeconal at 358 K (A, B) and 368 K (C,D) and hexadecane at 328 K (E, F) and 338 K (G, H). Both organic compounds are shown as licorice representations, with alkyl C-type beads coloured green, and alcohol P-type beads coloured red. Order-disorder transitions in pure hexadecanol and hexadecane on a graphite surface. Snapshots of hexadecanol at 358 K (A, B) and 368 K (C, D). Snapshots of hexadecane at 328 K (E, F) and 338 K (G, H). Image reproduced from [84].

Three new bead types were added to the Martini 2.2 forcefield during this parameterisation effort [84]:

- C1S beads, modelling alkyl groups in the middle of long-chain organic molecules e.g docosane. The beads possess stronger lateral interactions with each other to favour the adsorption of long-chain molecules over short-chain molecules (e.g heptane) on the surface. Bond lengths were also increased from 0.47 nm to 0.51 nm, and a higher force constant for the angle bending potential (30 kJ mol^{-1} *vs*

25 kJ mol^{-1}) was implemented to reproduce the straightened chain behaviour on the surface.

- C1E beads, representing the end group of long-chain organic molecules, which also interact with each other more strongly than other bead types.
- SG4 beads, modelling the graphite surface. The bead falls within the S-type class due to the 2:1 mapping scheme, resulting in a reduced size and LJ potential energy well depth (scaled to 75% of the original value). The scheme allows the surface to be stabilised by non-bonded interactions only, enabling the simulation of very large graphite surface areas [84].

The graphite surface was composed of hexagonally packed SG4 beads with a bead-bead distance of 0.27 nm. A very strong LJ ϵ well-depth parameter of 102 kJ mol^{-1} was used to maintain this distance, and a modified σ parameter of 0.27 nm. The graphite was also treated as a static surface, modelled as "frozen" using a temperature of 0 K for part of the system during simulation.

A variation of the graphite model was employed in this thesis, studying adsorption of lipids and protein molecules to the support.

Polarisable Water

In order to capture screening of charged interactions by surrounding solvent in a local environment, a CG polarisable water model was developed for simulations with the Martini forcefield [148]. In contrast to the standard Martini water model, three beads were used to represent four water molecules in the polarisable model, rather than one (Fig. 2.11). The central bead was modelled as neutral and interacted with other system particles through LJ interactions, identical to the standard Martini water model. Additional beads, WP and WM, are bound to this central particle and carry a positive ($+q$) and negative ($-q$) respectively (Fig. 2.11). These beads interact with other beads only via a Coulomb function, lacking any LJ interactions. The W-WP and W-WM bonds were restrained at distance l, whilst self-interactions within the same water bead are excluded, such that the WM and WP particles are "invisible" to each other. Consequently, the charged beads were able to rotate around the central W bead, modelled by a harmonic angle potential with an equilibrium angle (θ) and force constant (K_θ) to control this rotation and adjust the distribution of dipole momentum [148].

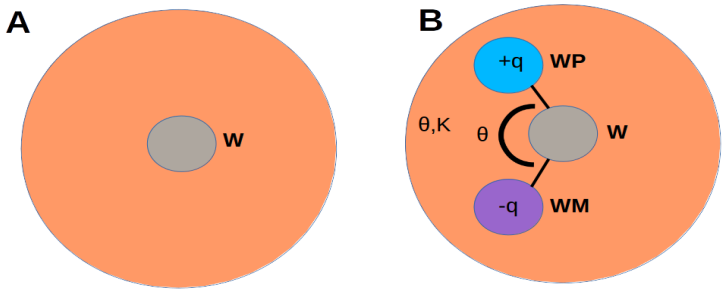

Fig. 2.11: **A)**: The standard Martini water model. **B)**: The polarisable water model. The orange spheres represented the van der Waals radii of the center W particles. Figure inspired by [148].

The polarisable water model more accurately represents the dielectric screening properties of bulk water, and also models phase-transition temperatures more precisely.

This water model was used for simulations with the aforementioned graphite model in this thesis.

2.4.2 Multi-scale Simulations

Given that secondary structures of proteins were restrained by the elastic network model employed in CG simulations, CG simulations were unable to capture large scale structural motions or transitions in proteins. Methods have been developed that allow the conversion of CG systems to atomistic resolution (AT), which are able to capture these transitions [149, 150]. The conversion methods commonly involve mapping the CG system configuration back onto atomistic descriptions of the same molecules. This is the approach taken by the "CG2AT" method, which utilises a fragment-based fitting approach in order to convert CG protein-membrane systems to atomistic detail, supporting simulations with either GROMOS, OPLS, and CHARMM36 forcefields [149].

The CG2AT method was employed to convert enzyme-lipid membrane systems from CG to AT resolution in this thesis.

2.5 Non-equilibrium Simulation Methods

Non-equilibrium methods are often employed to sample states of a system that are inaccessible using classical MD methods. This can include the characterisation of an ensemble of conformational states of a peptide in solution, or sampling a α-helix structural transition [151, 152]. These system properties can be very difficult to sample on the time scales of atomistic MD simulations, particularly if the system is trapped in an energetic minimum (e.g an equilibrium state) [151, 152]. Non-equilibrium simulation methods have been developed to address such limitations, providing insight into "rare"

events that usually occur on time scales of milliseconds to seconds.

2.5.1 Steered Molecular Dynamics

Steered molecular dynamics simulations are useful to accelerate biological events and explore slow and low frequency motions, such as the rotation of the γ-subunit of ATPase, or the unfolding of fibronectin [153–156]. This acceleration is achieved by applying an external force to a defined part of the system described by a collective variable, which is used to drive a particular event (e.g a conformational transition) within reasonable simulation times [157]. The collective variable (CV) describing the event of interest can be defined as, for example, a dihedral angle, a set of distances, or a root mean square deviation of a protein structure. Once the CV has been defined, the external force is applied by setting a predefined moving harmonic restraint along the CV vector. This harmonic restraint potential fixes the CV to a particular point in space, and is then shifted along the direction of the CV over time, pulling the system away from its initial configuration [157]. This force is implemented by modifying the Hamiltonian (H) describing the system into H_λ, adding the harmonic potential term (centered on the CV) which moves linearly with time:

$$\begin{aligned} H_\lambda(R,t) &= H(R) + U_\lambda(R,t) \\ &= H(R) + \frac{k(t)}{2}(s(R) - \lambda(t))^2 \\ &= H(R) + \frac{k(t)}{2}(s(R) - s_0 - vt)^2 \end{aligned} \tag{2.28}$$

where k is the spring constant of the harmonic potential, *s(R)* defines the coordinates of the pulling group (S) which is the CV, v is the velocity at which the harmonic potential moves along the CV vector, s_0 is the initial position of the CV, and t is the simulated time.

The force constant of the harmonic spring can be tuned such that the CV closely

follows the centre of the moving potential during the SMD simulation, allowing the system to sample a specific motion. The total work performed to pull the CV during the SMD simulation can then be calculated as the integral of the modified H_λ Hamiltonian:

$$W = \int_0^{t_s} dt \frac{\partial H_\lambda(t)}{\partial t} \tag{2.29}$$

This quantity can be useful in estimating the structural forces underlying the simulated event, providing information about the system. Although the technique introduces a biasing force in the system, equilibrium thermodynamic properties such as a free energy profile, or a potential of mean force, should in principle be possible to calculate from the simulation [158].

In this thesis, constant velocity SMD simulations were performed using the Plumed 2.1 plugin for GROMACS [159–161].

2.6 Software and Analysis

All the simulations within this thesis were performed with the GROMACS biomolecular simulation package, using GROMACS 4.6.x versions (www.gromacs.org) [161]. All CG simulations were performed using the Martini 2.2 forcefield [122], and all atomistic simulations were performed using the GROMOS 53A6 and 54A7 forcefields [112, 162].

The Visual Molecular Dynamics (VMD) program was used for visualisation of trajectories and image rendering [163]. All analysis was performed using either GROMACS tools or locally written Python code, occasionally employing the MDAnalysis module for analysis of trajectory data [164].

Chapter 3

M37 Lipase Interactions with Lipid Interfaces

This chapter is based on the following publication: Willems, N., Lelimousin, M., Kolds∅, H., Sansom, M.S.P. "On the interfacial activation of the M37 lipase: a multi-scale simulation study" (*accepted, BBA Biomembranes*). It will discuss the use of multi-scale molecular dynamics simulations to study the interfacial interactions between M37 lipase and different lipid surfaces. The simulations reveal a possible activation mechanism for M37, which is supported by enhanced sampling simulations and docking studies.

3.1 Introduction

Among the fifty or so bacterial lipase proteins that have recently been identified, the M37 lipase from *Photobacterium lipolyticum* represents an interesting example to explore interfacial activation. M37 exhibits a low activation energy towards triglyceride substrates, stability in non-aqueous solvents, and catalytic activity at low temperatures, due to its apparent ability to function at low temperatures (psychrophilicity)

[165]. Some unique structural features of M37 are thought to underlie these properties, including the existence of the active site cavity and a wide oxyanion hole [166]. These structural factors are regarded as destabilising factors, possibly contributing to a lower activation barrier for structural rearrangement, allowing catalysis at lower temperatures. As is commonly the case in lipase structures, the protein contains an amphipathic lid region that covers the catalytic residues of the active site, as well as a so-called α-helical flap region (Fig. 3.1).

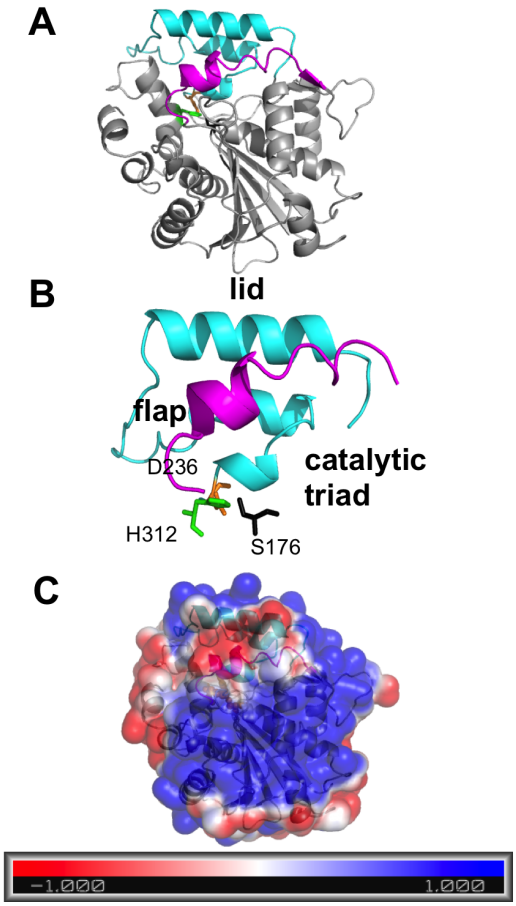

Fig. 3.1: **(A)** Crystal structure of M37 lipase (PDB: 2ORY, 2.2 Å resolution) highlighting the lid (cyan; residues 235-283) and active site flap (magenta; residues 94-110) regions. **(B)** Close up of the lid and active site flap regions of M37 with the catalytic triad (Ser174, Asp236 and His312) residues. **(C)** Electrostatic surface of M37 (calculated using APBS [167]) shown in the same orientation as in A (blue: positive (+1), red: negative (-1)).

The active site flap and lid regions of M37 were identified on the basis of structural comparisons with related lipase enzymes. A DALI search for structurally similar proteins revealed that the *Rhizomucor miehei* lipase (RML) and *Thermomyces lanuginosus* lipase (TLL) produced the highest Z-scores (measurement of structural similarity; RML = 20.7, TLL = 20.3) [166]. Alignment of these structures revealed the positions of the catalytic triad in M37, consisting of Ser172, Asp236, and His312. The structural regions overlying this cavity were thus identified as the lid and active

site flap region. In particular, the active site flap region aligned well to the previously identified lid regions of both the RML and TLL structures. Consequently, it is thought that the lid and flap regions determine the interfacial activation of M37 [166]. In general, interactions of lipases with hydrophobic interfaces are thought to result in a conformational transition in the lid region, exposing the active site and allowing substrate molecules to bind [42, 168]. However, the exact nature and order of these steps is thought to differ for different lipases and thus a general mechanism of interfacial activation remains elusive. In this context, molecular dynamics simulations provide a valuable approach to probe the underlying dynamic processes that determine the activity of enzymes and their behaviour in different environments [169, 170]. Simulations have been used to identify solvent-induced effects on interfacial activation for lipases closely related to M37 (e.g TLL) [97]. Computational approaches have also provided important insight on lipase interfacial interactions, which contributed to unveiling functionally important motions in these enzymes [97, 100, 171–173]. Therefore, the analysis of M37 dynamics over different simulation time scales might be expected to reveal key aspects of its activation mechanism.

In this chapter, a multi-scale simulation framework was used to investigate the dynamics of M37 in different environments, including lipid interfaces and aqueous solution. First, the interfacial behaviour of the lipase was analysed with two different phospholipid bilayers of varying surface charge. These simulations suggested that M37 preferentially interacts with negatively charged membranes, correlating with experimental studies of a related fungal lipase [70, 94, 174]. Second, interfacial interactions with a natural substrate (tributyrin) revealed functionally relevant motions of the enzyme. A large-scale motion of the lid region was identified to open the initially closed conformation of M37, uncovering the entry pathway of the catalytic site to the natural substrate. This mechanism of activation provides new insights to rationally engineer

the psychrophilic M37 lipase.

3.2 Methods

3.2.1 Coarse-grained MD Simulations

The crystal structure of M37 (PDB: 2ORY) was downloaded from the Protein Data Bank and converted to coarse-grained (CG) representation using the Martinize script (available at http://md.chem.rug.nl/index.php/tools2/proteins-and-bilayers). The Martini 2.2 forcefield was used for the CG simulations employing the ElNeDyn (EDM) elastic network model to maintain the secondary and tertiary structure of the protein [134, 138, 143]. A molecular dynamics (MD) simulation of M37 in aqueous solution was initially performed at atomistic (AT) level. Different EDM parameters were tested in CG simulations, using the AT simulation as a reference for protein dynamics. The root mean square fluctuations (RMSF) and root mean square deviations (RMSD) of Cα atoms of M37 were calculated at both AT and CG levels and then compared. This comparison suggested that a cut-off radius of 0.95 nm and a force constant of 550 kJ mol^{-1} resulted in the optimal reproduction of protein dynamics observed from the AT simulations (Fig. 3.2).

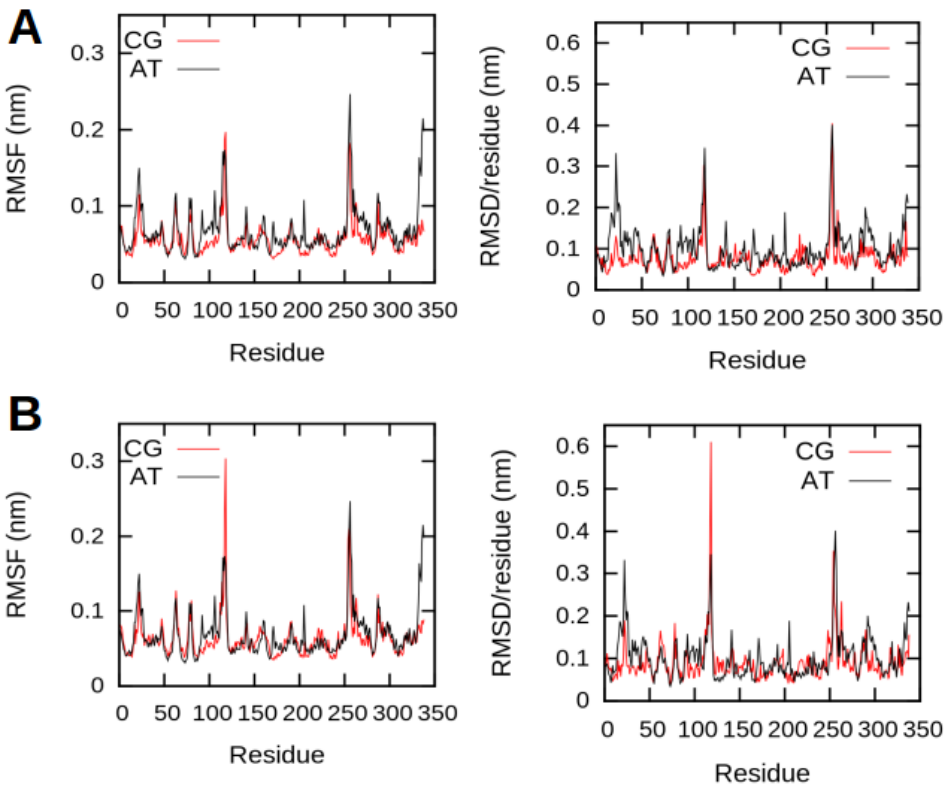

Fig. 3.2: RMSF and RMSD values calculated for Cα particles of M37 after AT simulation in water, and for backbone particles after a CG simulation in water simulation. The following ENM parameters for the M37 CG structure model were tested: **(A)** 550 kJ mol^{-1} force constant and a 0.95 nm cut-off radius, or **(B)** a 500 kJ mol^{-1} force constant and a 0.95 nm cut-off radius.

Self-assembly simulations were first performed to construct two solvated bilayers containing 512 lipids each. A purely zwitterionic bilayer (100% DPPC) and an anionic bilayer (20% DPPG: 80% DPPC) were built to investigate the effect of surface charge on M37 binding. Both bilayers were symmetric in terms of lipid composition of the upper and lower leaflets, with an distribution of 257 lipids (191 PC + 66 PG) in the upper leaflet and 255 lipids (193 PC + 62 PG) in the lower leaflet (measured after 1 µs simulation time). The enzyme was initially placed 3 nm above the pre-formed bilayers and then solvated. Counter ions were added to neutralise the charge of the systems. Two different starting orientations of M37 were simulated to minimise any

bias occurring from the initial position (Fig 3.3). Energy minimisation using the steepest descent algorithm was performed for 200 steps followed by equilibration for 10,000 steps. Ten replicas of each starting orientation were simulated in the NPT ensemble at 310 K and 1 bar, resulting in a set of 20 CG-MD simulations of 5 µs each. A leapfrog algorithm was used to integrate Newton's equations of motion, with a time step of 20 fs. Van der Waals interactions were evaluated using a buffered Verlet scheme, applying a cutoff of 1.1 nm, and the electrostatic interactions evaluated using the reaction-field method, also applying a cutoff of 1.1 nm [175, 176]. The Berendsen thermostat was used for temperature coupling (310 K) with a weak coupling constant of 1.0 ps [120]. Semi-isotropic pressure coupling was applied using the Berendsen barostat with a 1.0 ps coupling constant and a compressibility of $3x10^{-4}$ bar [120].

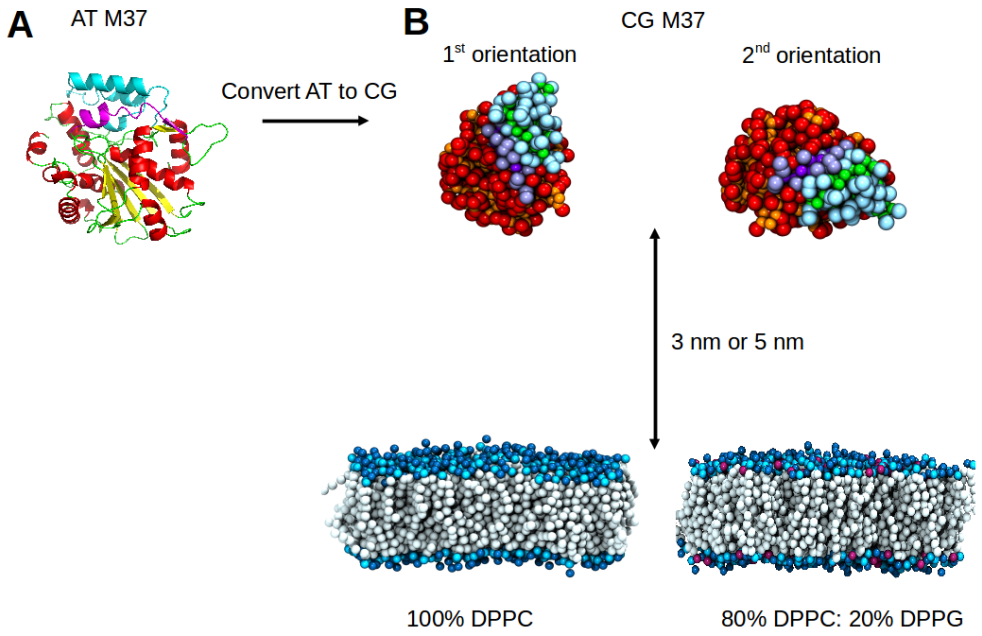

Fig. 3.3: **(A)** Crystal structure of M37 (cartoon representation), highlighting the lid region (cyan) and active site flap region (magenta). The rest of the protein is coloured in (α-helices), yellow (β sheets) and green (loops). **(B)** CG representations of M37 and lipid bilayers. Two different starting orientations of M37 positioned above the bilayers were tested (orientation 1 and 2), at varying distances (3 nm or 5 nm) as measured from the bottom of the protein to the lipid headgroups (choline). Individual simulations were performed with either a pure PC bilayer (*left*) or 80% PC:20% PG bilayer (*right*). All CG particles are shown as van der Waals spheres. Lipid choline groups are coloured dark blue, phosphate groups coloured cyan, and glycerol groups coloured purple. Lipid carbon tails are coloured white. Protein particles are coloured red (backbone) and orange (side chains), whilst the lid is shown as green (backbone) and cyan (side chains) groups, and the active site flap as light purple (backbone) and dark purple (side chains).

3.2.2 Atomistic MD Simulations

Lipid Bilayer Simulations

The final frames of selected CG M37-anionic bilayer simulations were converted to atomistic (AT) representations using a fragment-based approach [149]. The GROMOS 53A6 forcefield and SPC water model were used for all AT simulations, conducted in

the NPT ensemble (Table 1) [112, 177, 178]. Non-bonding interactions were considered through a buffered Verlet scheme [175]. Long-range electrostatic interactions were treated using the particle mesh Ewald method with a real-space cutoff of 1 nm, a Fourier spacing of 0.12 nm and a fourth-order spline interpolation [124]. Short-range van der Waals interactions were treated within a cut-off range of 1 nm. The V-rescale thermostat was used for temperature coupling and the Parrinello-Rahman barostat for semi-isotropic (lipid simulations) or isotropic (tributyrin simulations) pressure coupling [118, 123]. The LINCS algorithm was used to constrain bond lengths [125]. A leapfrog algorithm was used to integrate Newton's equations of motion, with a time step of 2 fs. All simulations were equilibrated for 1 ns during which all the atoms in the protein were position restrained using a force constant of 1000 kJ mol^{-1}, prior to performing 200 ns of unrestrained simulations.

Triglyceride (tributyrin) Simulations

The tributyrin GROMOS 54A7 AT forcefield topology and coordinate files were downloaded from the Automated Topology Builder website (http://compbio.biosci.uq.edu.au/atb/) [162, 179–181]. A tributryin layer was formed by randomly inserting 731 tributyrin molecules into a 10x10x5 nm cubic box. After steepest descent energy minimisation, 1 ns of NVT equilibration was performed at 298 K, after which the z-dimension of the box was extended to 10 nm and the system solvated with SPC water. An additional 1 ns of NPT simulation ensured that the density of tributyrin equilibrated to 1,026 g/L (experimental value at 298 K = 1,027 g/L) [182]. M37 was then positioned 1 nm above the tributyrin layer in the centre of the box (distance measured from the bottom of the protein to the top of the layer), before performing 200 ns of MD simulation.

Steered MD Simulations

Steered MD (SMD) simulations were performed using the Plumed 2.1 plugin for GROMACS [159, 183]. The equilibrated structure of the M37 simulations in water was used as the starting point, applying identical simulation parameters as described for the AT-MD simulations mentioned above. All SMD simulations were performed using the GROMOS 54A7 forcefield. Different collective variables (CVs) were tested to probe the lid and active site flap motions occurring though opening of the closed form of M37 (crystal structure). The chosen CV was a distance CV defined between the COM of the lid helix closest to the active site flap (residues 264-279) and the helix of the active site flap itself (residues 94-110; Fig. 3.1). Different distances were tested, ranging from 1 nm to 3 nm over individual 10 ns simulations. A harmonic bias potential was applied to the CVs, moving at a constant velocity of 1 nm ns^{-1} with a spring stiffness ranging from 500-1000 kJ mol^{-1} (Table 1). Additionally, a pseudo-dihedral CV was tested, defined by specified Cα atom positions of residues within the hinge regions of the active site flap. The limit values of both the distance and the pseudo-dihedral CVs were estimated by producing an opened M37 model based on structural alignment of the active site flap region with an open form of TLL using the sculpting tool in PyMOL (www.pymol.org) [56].

Docking Calculations

All docking studies were performed using the GOLD 5.0 suite [184, 185]. Protein models consisted of either the last frame of the M37-tributyrin simulation or an equilibrated structure of the closed form. The protein models were prepared by adding

hydrogens with the default settings of the GOLD program. Docking of one tributyrin molecule in the active site cavity, defined by a 1 nm radius from the position of residue Ile235, was performed using the CHEMPLP scoring function [186].

3.3 Results

3.3.1 Interfacial Interactions of M37 with Lipid Bilayers Explored via CG Simulations

Initial coarse-grained (CG) simulations of M37 positioned above self-assembled bilayers were performed to investigate interfacial interactions with lipid surfaces. Both zwitterionic (100% PC) and anionic (20% PG:80% PC) were tested to explore the effect of surface change on the enzyme binding and interfacial interactions (see Table 3.1)

Simulation	Forcefield	Replicates
CG-MD, M37+PC bilayer	Martini 2.2	6 x 1 µs
CG-MD, M37+PC/PG bilayer	Martini 2.2	20 x 5 µs
AT-MD, M37+PC/PG bilayer	GROMOS 53A6	9 x 200 ns
AT-MD, M37 in water	GROMOS 53A6	3 x 200 ns
Steered AT-MD, M37 in water: CV=Distance[1]	GROMOS 547	5 x 10 ns
Steered AT-MD, M37+tributyrin: CV=Distance[2]	GROMOS 54A7	3 x 15 ns

Table 3.1: Summary of the main simulations performed in this chapter. PG/PC bilayer = 20% PG:80% PC. Different CVs were tested by steered MD simulations. The distance CV was defined as the distance between the lid helix and the active site flap of M37 in water. Distance[1] = A spring force constant of 500 kJ mol^{-1} was used and no position restraints were applied to the protein atoms. Additional CVs were tested (detailed below). Distance[2] = The distance between the C4 atom of a tributyrin molecule and the Oγ atom of the catalytic Ser174 was used to study substrate binding in M37, performing SMD simulations with a spring force constant of 2000 kJ mol^{-1}.

Only transient interactions occurred with the zwitterionic bilayer and thus no binding events were observed (Fig. 3.4). A further three simulations of M37 positioned in a different orientation above the zwitterionic bilayer were performed, which also did not generate any bound configurations of the M37 enzyme. In contrast, preferential and long-lasting interactions of the protein with the anionic bilayer (PC/PG) were observed. Therefore, the following analysis refers to simulations with anionic bilayers only. Binding events occurred over the course of all the 20 replicate CG-MD simulations, within a time ranging from 100 ns to 4 µs (Fig. 3.5). Once bound, the protein did not dissociate from the anionic lipid bilayer.

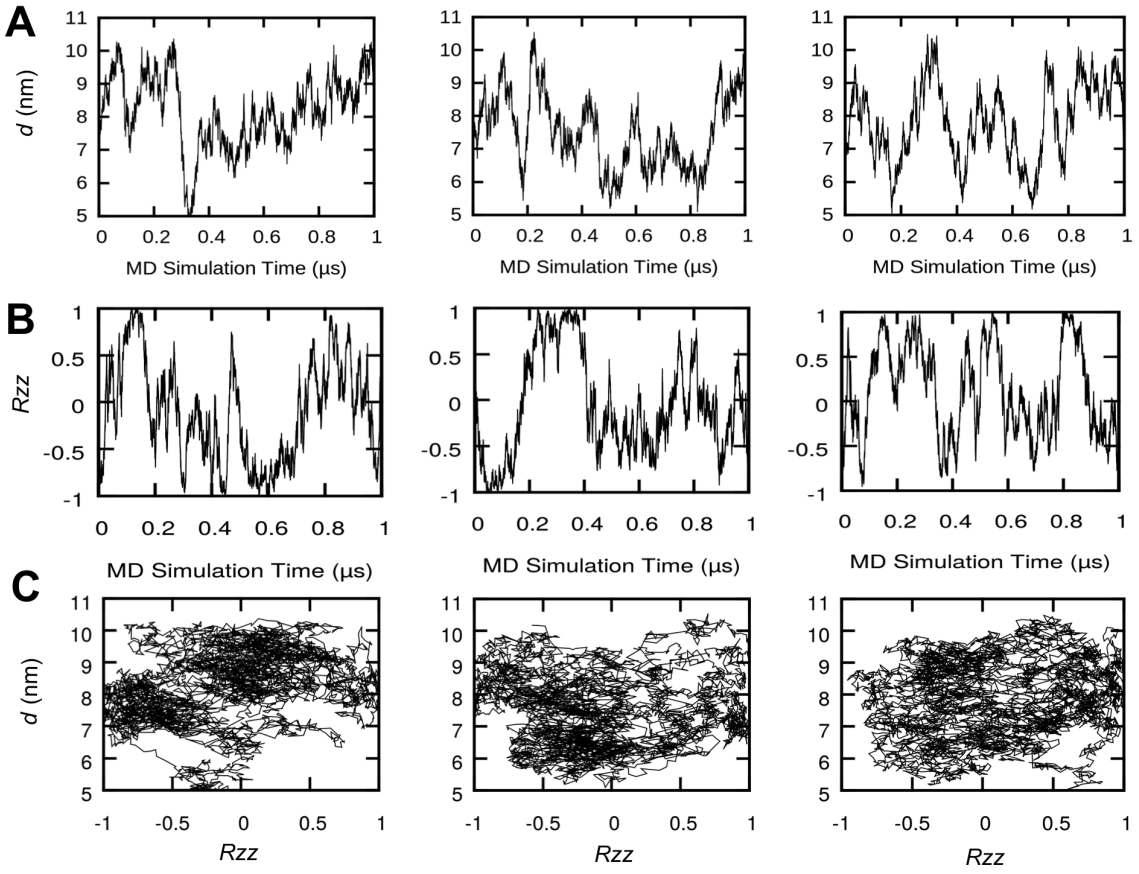

Fig. 3.4: Interaction analysis of CG-MD simulations of M37 with pure DPPC membranes for 3 individual replicate simulations. **(A)** Time evolution of the centre of mass (COM) distance between the M37 and the lipid bilayer (d). **(B)** Time evolution of the Rzz value from a rotation matrix calculation. The rotation matrix was calculated relative a reference binding orientation of M37 with anionic bilayers (please see Fig. 3.6 for more information). **(C)** The distance metric (d) plotted against Rzz, showing the evolution of the Rzz value as a function of the distance between M37 and the lipid bilayer for each of the simulations. The left, middle, and right panels correspond to repeat simulation 1, 2, and 3 respectively. These plots show the random diffusion of M37 throughout the aqueous solvent, as is reflected in the Rzz value. The lipid headgroups of the upper leaflet of the bilayer corresponds to $d \sim 2$ nm, however binding is expected to occur when $d \sim 4$ nm, accounting for the COM of the protein. The protein therefore never interacted with the zwitterionic bilayer.

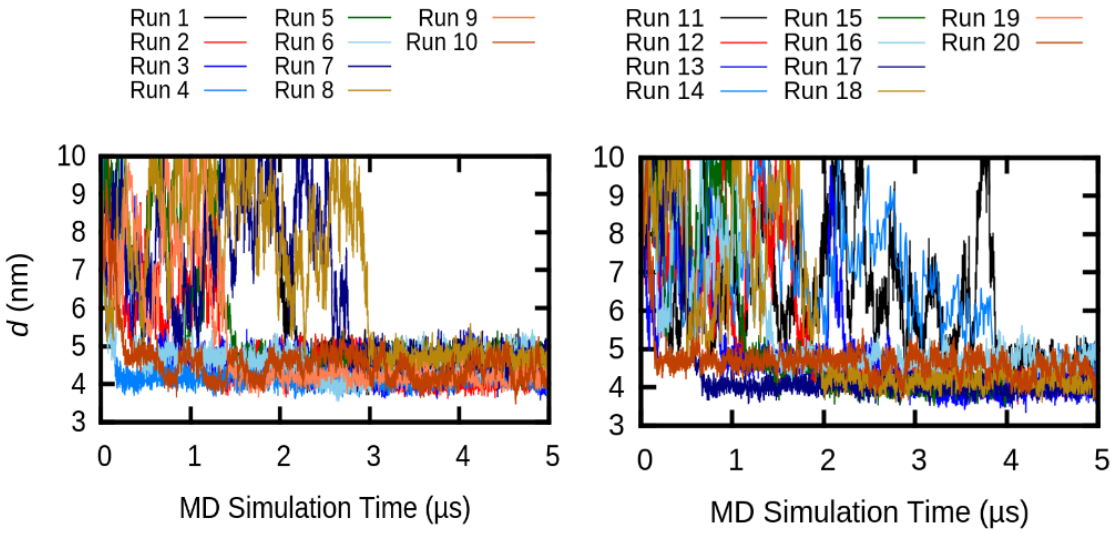

Fig. 3.5: *Left*: Time evolution of the z-component of the COM distance between M37 and the bilayer, calculated for the first 10 replicate simulations of M37 interactions with an anionic bilayer. *Right*: Time evolution of the z-component of the COM distance between M37 and the bilayer for the last 10 replicate simulations. The lipid headgroups of the upper leaflet of the bilayer correspond to $d \sim 2$ nm, however binding occurs when $d \sim 4$ nm, accounting for the COM of the protein.

In order to analyse the binding mechanisms and the interfacial orientations of M37 on the anionic membrane surface in more detail, two collective variables (CV) were defined: a translational CV (d), defined as the z-component of the distance between the centre of mass (COM) of the protein and of the bilayer, and a rotational CV (Rzz), obtained from the rotation matrix calculated with respect to a reference orientation of the protein (Fig. 3.6). Before binding to the lipid bilayer, the protein underwent translational and rotational motions in solution as anticipated, assessed by variations in d and Rzz. In contrast, upon binding to the anionic membrane, the lipase showed a small number of canonical binding orientations, although conversions between these configurations were observed (Fig. 3.6).

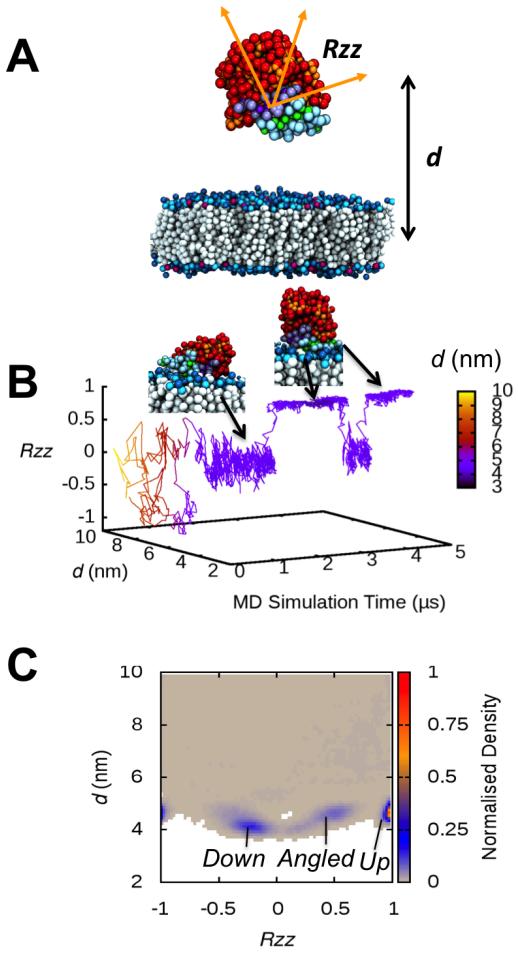

Fig. 3.6: **(A)** Initial system configuration for the CG-MD simulations of M37 binding to lipid bilayers. The colour scheme is the same as in Fig. 3.3. The figure schematically represents the CVs (d and Rzz) used to describe the motion and orientation of M37. The metric d was calculated as the z-component of the distance between the COM of the protein relative to the COM of the bilayer. A rotation matrix was calculated as a function of the Rzz angle that defines the transition from a given orientation of the enzyme (orange arrows) to a reference orientation. When Rzz = 1, the enzyme adopts the orientation of the reference structure. **(B)** Time evolution of d and Rzz during one of the CG-MD simulations. The enzyme switched from the *Up* to the *Angled* orientation during the simulation, as is reflected in the value of Rzz. The inset images show the orientation of M37 at the corresponding Rzz value. **(C)** The normalised density of each binding orientation calculated as a function of d and Rzz derived from the ensemble of the 20 CG-MD simulations. The lipid headgroups of the upper leaflet of the lipid bilayer correspond to the bottom of the map ($d \sim 2nm$).

The ensemble of 20 CG-MD simulations allowed calculation a 2-dimensional land-

scape representing the normalised density of the binding orientations, as a function of *d* and *Rzz* (Fig. 3.6C). Together with visual inspection of the simulations, this analysis revealed three major binding configurations of the lipase on the bilayer surface, which I have termed *Up, Angled* and *Down* (Fig. 3.7). The density map indicates that the *Up* orientation is the most frequently sampled orientation within the simulation ensemble. All trajectories (20) were pre-processed with the *trjconv* tool from GROMACS, in order to remove the rotation of the protein in the x-y plane. Consequently, any variation in the x-y plane of the protein relative to the phospholipid bilayer plane is reflected in the *Rzz* value, which represents the orientation of the z-axis of the reference structure after fitting to the z-axis of the protein in the trajectory frame. This can be seen from the representative binding orientations of M37 producing the corresponding *Rzz* values in Fig. 3.6B, in which the protein does not produce the same *Rzz* value for a different binding orientation. For example, if M37 adopts an orientation similar to the *Angled* orientation (Fig. 3.7A) but the active site flap region was oriented towards the aqueous solvent rather than towards the bilayer during the CG simulation, this would produce a different *Rzz* value than if the true *Angled* orientation was sampled during the simulation (active site flap oriented toward bilayer). Therefore, only very small variations in protein binding orientation will produce the same Rzz value, validating the use of this metric to assess the interfacial binding orientations of M37 with anionic phospholipid bilayers. This is validated further in future chapters.

The three identified orientations were further characterised by cluster analyses, *i.e* extracting and concatenating all trajectory frames corresponding to a specific orientation from the whole simulation ensemble. This resulted in three compiled trajectories, each representing all the frames that correspond to a particular binding mode within the ensemble. First, the average number of contacts between the protein and the lipid headgroups were calculated for each orientation (Fig. 3.7B and Fig. 3.8). Second,

average radial distribution functions (RDF) were computed considering the two lipid types (PC and PG) individually (Fig. 3.7C).

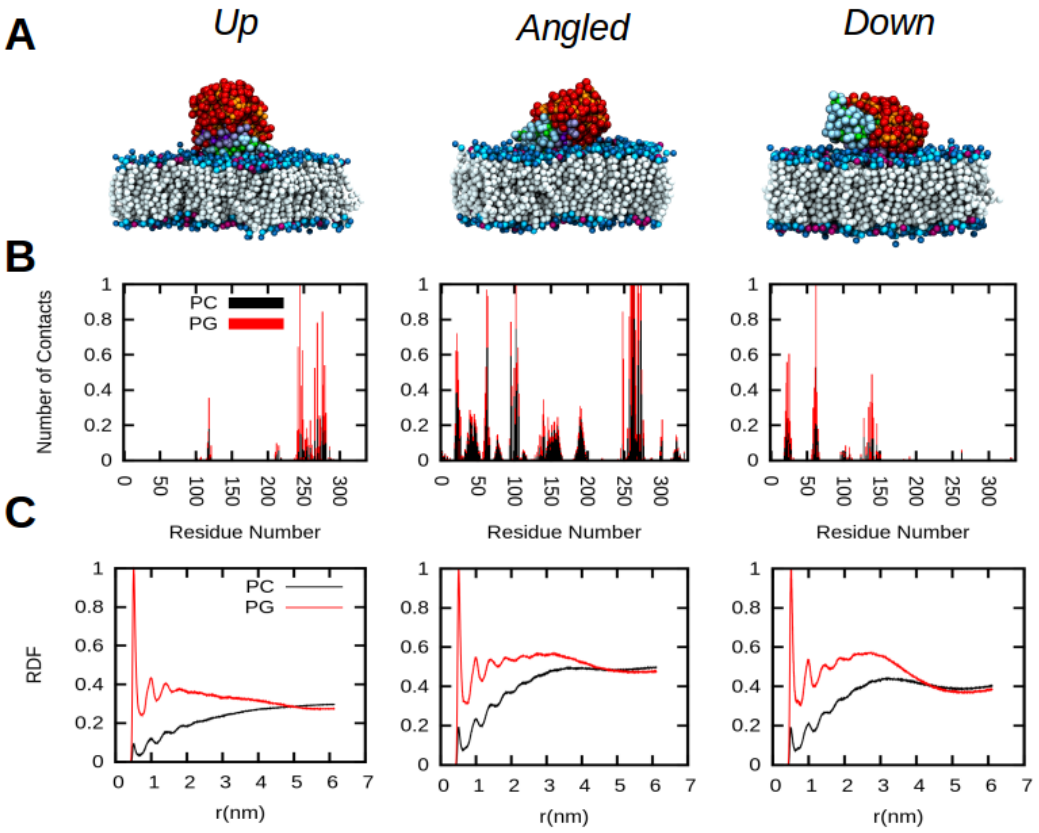

Fig. 3.7: **(A)** Representative structures of the three different orientations of M37, observed in the CG-MD simulations, when bound to the anionic lipid bilayer (20% PG: 80% PC). The same colour scheme as in Fig. 3.3 is used. **(B)** Normalised number of contacts made between the residues of M37 and either PC or PG lipids, as a function of the protein residue number, calculated within a 0.8 nm cut-off. **(C)** Normalised radial distribution functions of lipid density as a function of the distance of the protein to either the PC or the PG headgroups (choline and glycerol, respectively). All frames corresponding to one binding orientation from the 20 replicate ensemble were combined to calculate the statistics for each individual binding orientation. The analyses thus represent the entire ensemble rather than an individual simulation.

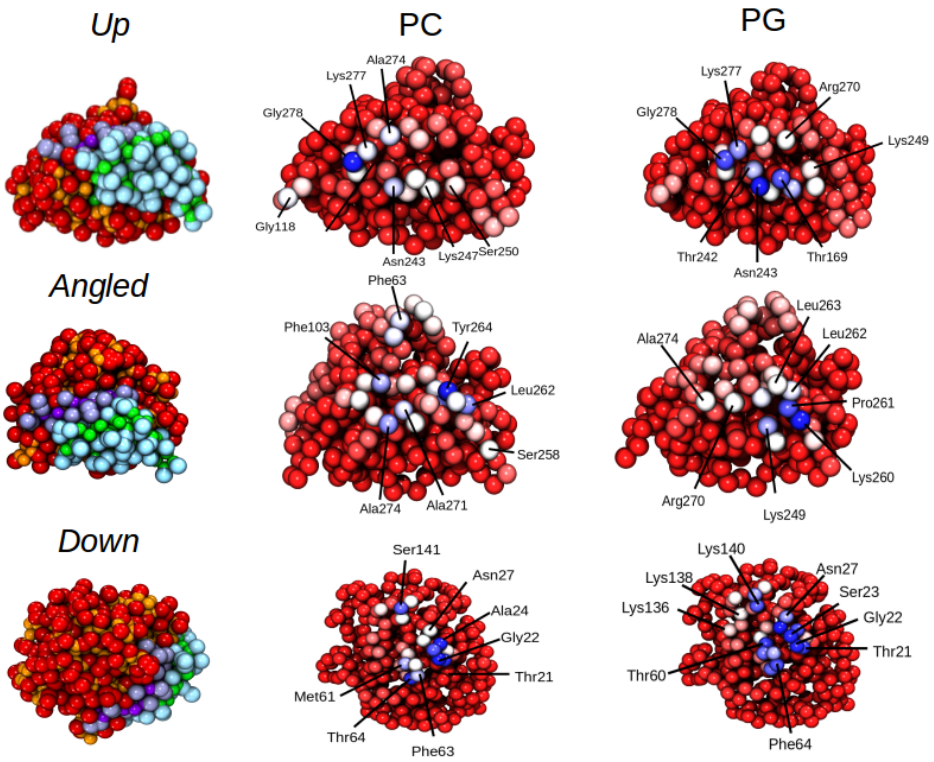

Fig. 3.8: Contacts between M37 residues and the phospholipids (PC *vs* PG lipids) calculated for the 20 replicate ensemble of CG simulations. All frames corresponding to a particular binding orientation were combined to calculate the normalised contacts. The residues that mediate contacts with the bilayer are shown on van der Waals representations of M37 using a blue/white/red colour scale (red = no contact, blue = greatest number of contacts). A cut-off radius of 0.8 nm was used for. A reference structure showing the position of the lid and active site flap regions of M37 with respect to the contact residue maps is shown on the left.

Contact analysis indicated that hydrophobic and basic residues mediated interactions with PG lipids, while mainly hydrophobic amino acids formed contacts with PC molecules (Fig. 3.8). The *Up* and *Down* binding orientations were mediated by a larger number of basic residues interactions with anionic PG lipids relative to the *Angled* orientation. Additionally, the RDF calculations indicated a higher density of PG lipids in the vicinity of the protein, relative to PC lipids, for all three of the binding orientations (Fig. 3.7C). Overall, these data suggests that electrostatic interactions promoted the interfacial interactions between M37 and the lipid bilayers, as has been

suggested experimentally for other related lipases [70, 95, 187].

3.3.2 Atomistic (AT) Simulations of the Conformational Dynamics of Bound M37

Atomistic (AT) simulations were initiated from representative CG simulations frames for each binding orientation, with the aim of studying the interfacial interactions and conformational dynamics of M37 in more detail [149]. The AT simulations showed a general agreement with the CG simulations regarding the protein-membrane interactions for each of the three binding orientations (Fig 3.9).

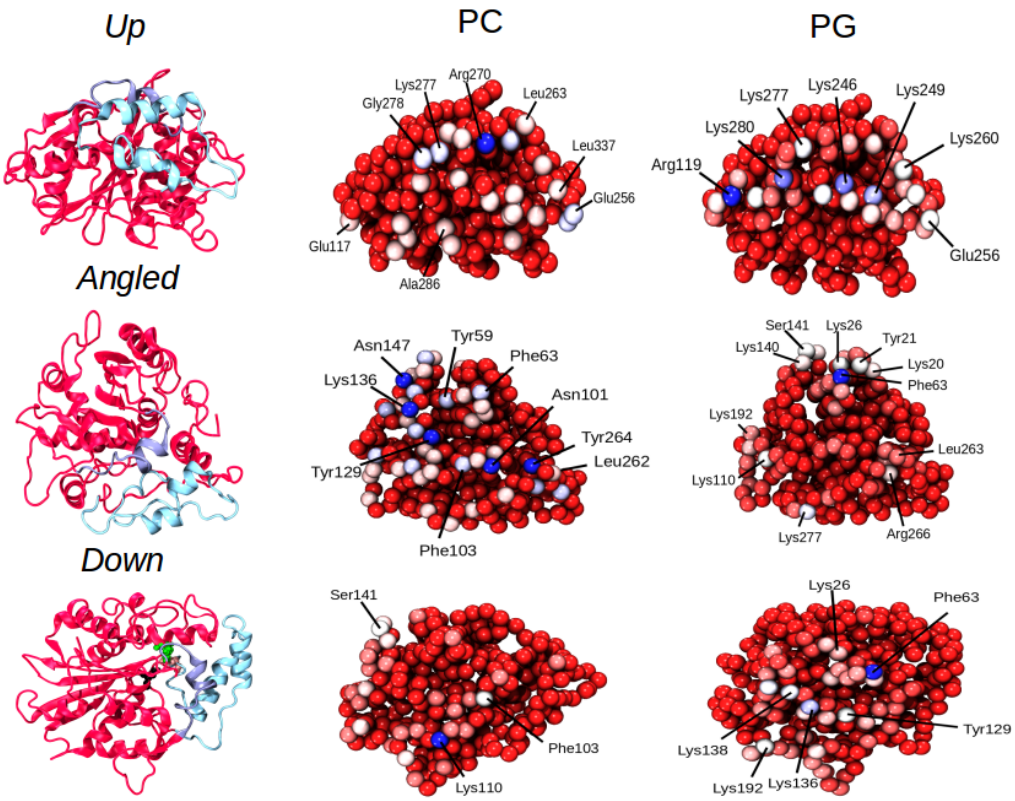

Fig. 3.9: Contact residue maps calculated for of AT-MD simulations of M37 with anionic bilayer simulations. 3 repeat simulations were performed for each binding orientation. The contacts between protein residues and the lipids within the bilayer were calculated within a 0.4 nm cut-off. The interacting residues are represented on van der Waals representations of M37 using a blue/white/red colour scale (red = no contacts, blue=greatest number of contacts). Reference structures showing the position of lid and flap regions with respect to the contact residue maps are shown on the left.

The AT simulations (200 ns) allowed analysis of the conformational dynamics of M37 in the three binding orientations, and comparison to simulations of the protein in water. Interestingly, all binding orientations, except for the *Up* orientation, exhibited larger structural fluctuations in the lid region compared to the active site flap region, as reflected by RMSD calculations of these regions (Fig. 3.10B and C). In particular, the *Down* orientation resulted in the largest RMSD value for the lid region, most likely related to its considerable solvent exposure, inducing fluctuations

similar to those observed for the protein in solution. The increased interactions with the membrane lipids for the protein in the *Up*, and to a lesser extent the *Angled* orientation, resulted in comparatively smaller fluctuations, most likely suppressed through additional interfacial interactions by the lid region.

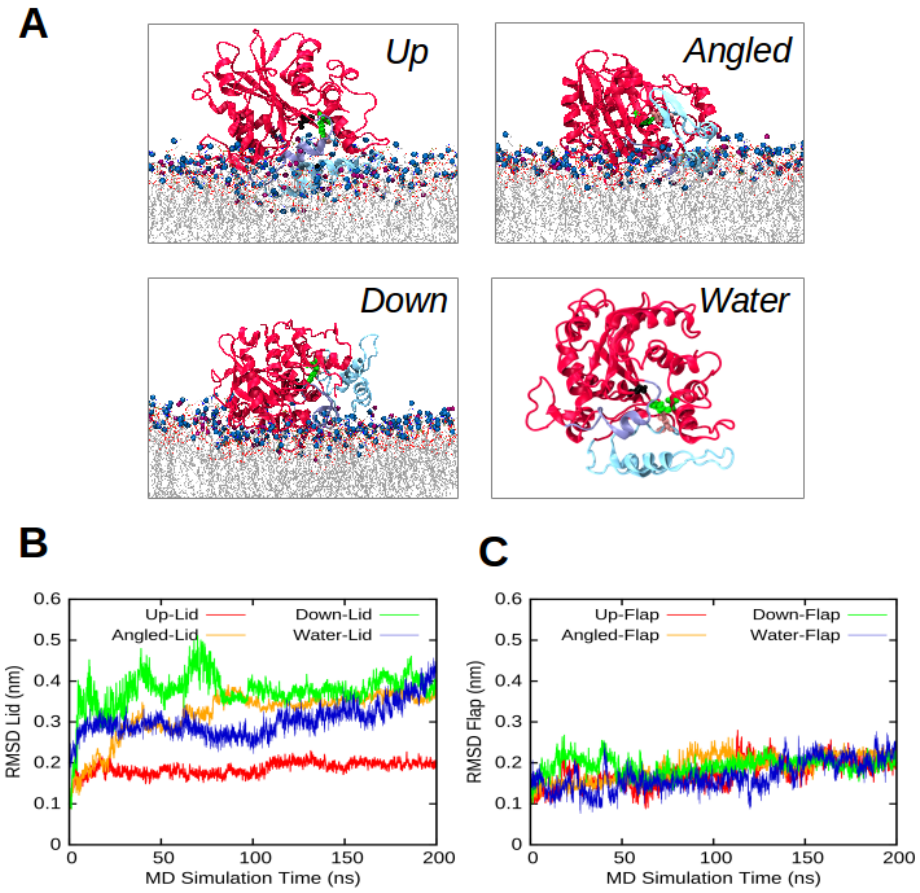

Fig. 3.10: **A)** Final binding orientations of the AT-MD M37-bilayer simulations (after 200 ns) representing the *Up, Angled*, and *Down* binding orientations, as well as the final structure of a simulation of M37 in water. The lid region of M37 is coloured in cyan and the active site flap region in purple; the rest of the protein is coloured in red (cartoon representations). The lipid tails are shown as sticks; PC headgroups are shown as points coloured in blue, and PG headgroups are coloured in purple. Below, time evolution of the RMSD of Cα atoms within the lid **(B)** and **(C)** active site flap regions for each of the simulation shown in **(A)**.

Overall, these results suggest that the interfacial interactions of M37 can affect the

dynamics of the lid but not of the active site flap region. By inspecting the activation mechanisms of related enzymes (e.g TLL or *Rhizomucor miehei* lipase (RML)), one might rather expect functionally important motions to occur in the active site flap region of M37 [166]. However, alignment of the final structures from the end of the AT bilayer simulations with the crystal structure of M37 indicates that neither the active site flap nor lid regions were significantly displaced as a function of these interfacial interactions (Fig. 3.11). Thus, in order to identify functional motions of M37, I investigated longer timescale motions of the lid and the active site flap regions.

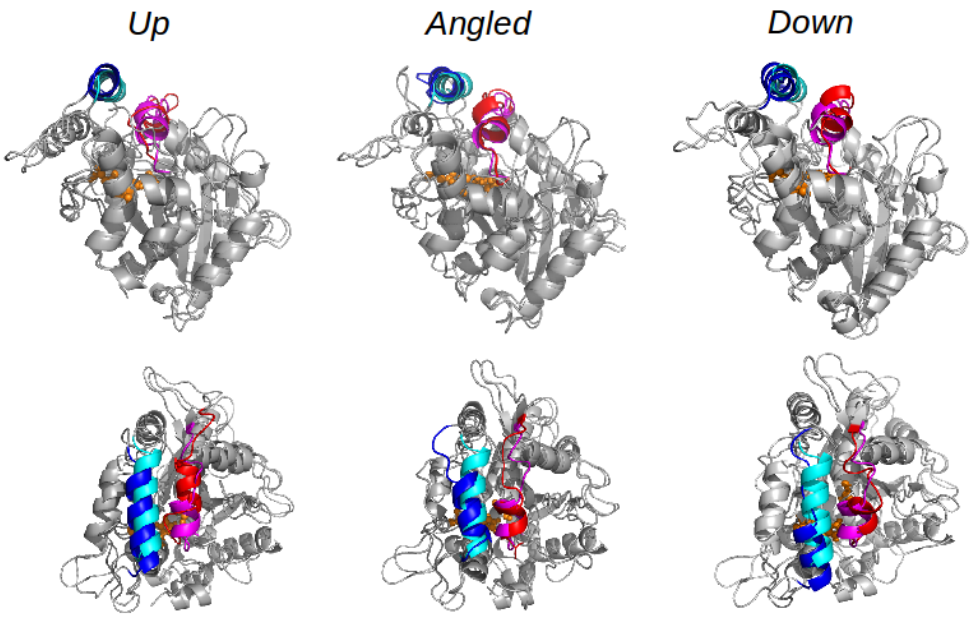

Fig. 3.11: Structural alignment of the final structures from AT-MD simulations of M37 with anionic bilayers at 200 ns, aligned with the M37 crystal structure (PDB: 2ORY [166]). The alignment is shown for a side view (top panel) and top down view (bottom panel). The images represent the final frame of simulations of M37 adopting one of the three different binding orientations (*Up, Angled*, and *Down*). The proteins are shown as cartoon representations (grey), the lid is shown in blue and the active site flap in red for the simulated structures (200 ns). The same regions are shown in cyan and magenta, respectively, for the crystal structure. The catalytic triad is shown as orange van der Waals spheres.

3.3.3 Functionally Relevant Motions of M37 in Water

Steered molecular dynamics (SMD) simulations may be employed to investigate slow and low frequency motions of biomolecules. A steered MD simulation accelerates biomolecular motions by applying an external force along a collective variable (CV) of interest, driving biologically relevant motions within shorter simulation times [188].

The aim was to define CVs able to model the possible long time scale motions of the lid (residues 235-283) and the active site flap (residues 94-110) regions, which could lead to opening of the closed form (crystal structure) of M37 in water. Since there are no crystal structures of the open form of M37, the closed form was compared to the open conformation of the structurally similar TLL (PDB: 1DT5) and *Rhizomucor miehei* lipase (RML; PDB: 4TGL) [53]. Alignment and comparison of the two structures (Fig. 3.12A) revealed that the lid region observed in M37 is not present in TLL or RML. However, the active site flap region within M37 could be aligned to the lid region of both TLL and RML, showing an evident potential shift in its position. This information was used as a starting point to define suitable CVs for the SMD simulations.

The first CV tested involved defining a pseudo-dihedral angle describing four consecutive Cα atoms within the active site flap front hinge region (CV_{front}), spanning residues 90-93, residues 91-94, or residues 92-95 (3 different individual simulations). These simulations were unable to drive any significant motion of the active site flap, or lid region, of M37 (Fig. 3.12). I therefore extended the CV_{front} by including an additional term describing the pseudo-dihedral angle between consecutive Cα atoms within the back of hinge region of the active site flap (CV_{back}). The CV_{back} spanned either residues 106-109, or residues 107-110, resulting in two independent SMD simulations. These simulations were unstable and thus did not complete successfully.

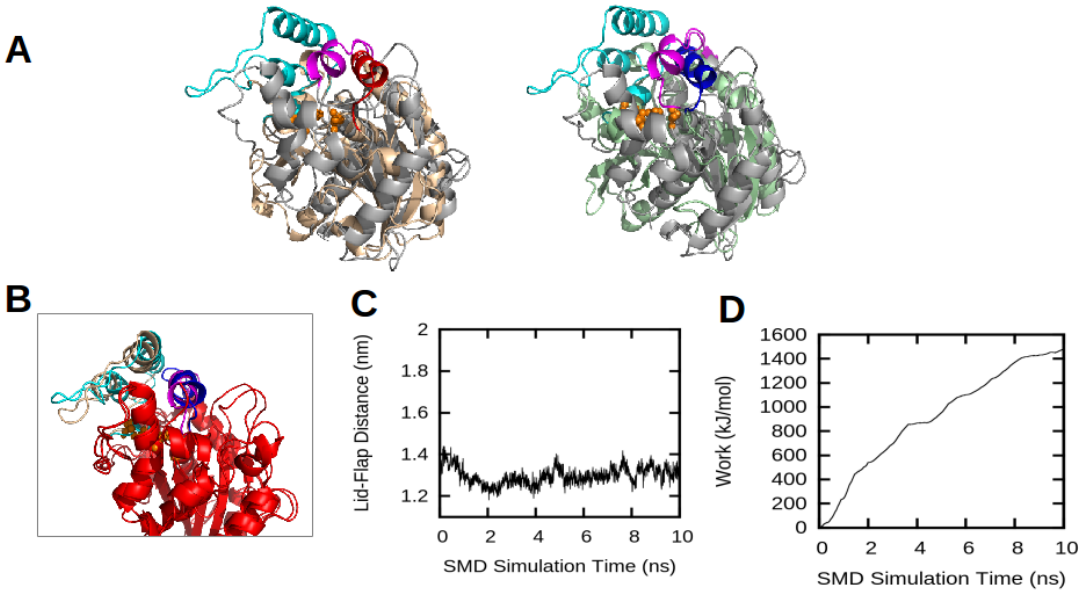

Fig. 3.12: **(A)** Structural alignment of the closed M37 crystal structure (PDB: 2ORY; grey [166]) with the open TLL crystal structure on the left (PDB: 1DT5; light orange [56]), and the open RML crystal structure on the right (PDB: 4TGL; light green [53]). The M37 lid and active site flap regions are coloured in cyan and magenta, respectively. The lid regions of TLL and RML are shown in red and blue respectively. **(B)** Superimposed cartoon structures of M37 (red) showing the position of the lid helices and active site flap extracted from the first (cyan and magenta respectively) and final (orange and blue respectively) frames of the CV_{front} SMD simulation (10 ns). **(C)** COM Distance between the lid helix (residues 264-279) and the active site flap helix, as a function of the SMD simulation time. **(D)** Time-evolution of the cumulated work.

Next, a distance-based CV was investigated, defined as the distance between the COM of the helical regions of the lid (residues 264-279) and of the active site flap (94-110). This SMD protocol enabled a gradual extension of the distance between lid and flap helices (Fig. 3.13). Multiple SMD simulations were performed using this distance CV, all of which showed very similar structural motions (5 replicates). In agreement with previous results, the active site flap did not significantly move from its original position. Instead, the lid region showed a large displacement, maintaining its α-helical fold during the simulation.

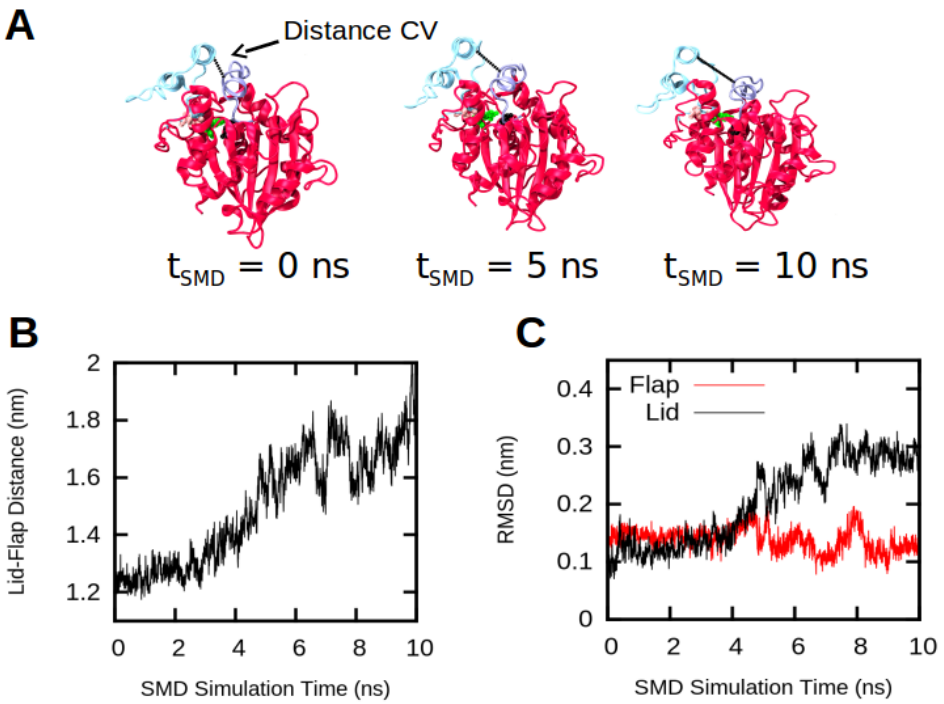

Fig. 3.13: **(A)** Structures of M37 in water (omitted for clarity) extracted from different time frames from a selected unrestrained SMD simulation using the lid-flap distance as a CV (schematically represented by a dotted line). **(B)** Distance CV calculated between the COM of the lid and the active site flap helices, as a function of the simulation time. **(C)** Time evolution of the RMSD of the lid and flap regions.

The apparent rigidity of the active site flap region was further investigated by performing additional SMD simulations. These applied the same protocol, however also included position restraints applied to all the atoms within the protein except those in the active site flap helix (Fc 1000 kJ mol^{-1}). Interestingly, the position restrained SMD simulations also resulted in negligible movement of the active flap region, even when using spring force constants up to 2000 kJ mol^{-1} (Fig. 3.14).

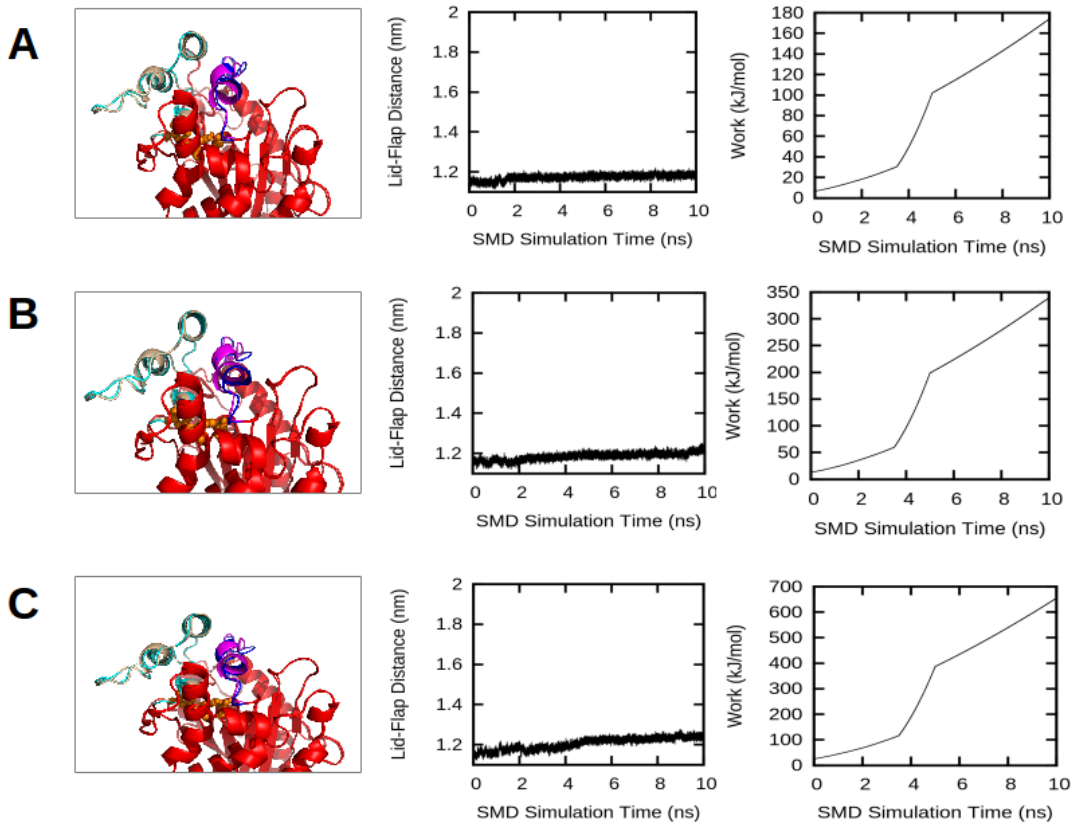

Fig. 3.14: SMD simulations of M37 in water using a distance CV defined between the lid helix and the active site flap helix. Position restraints were applied to protein atoms, except for those in the active site flap helix. Three different spring force constant for the pulling force were tested: **(A)** 500 kJ mol^{-1}, **(B)** 1000 kJ mol^{-1} and **(C)** 2000 kJ mol^{-1}. *Left*: Superimposed cartoon structures of M37 showing the position of the lid and active site flap regions at the start (cyan and magenta respectively) and end (orange and blue respectively) of one of the SMD simulation. *Middle*: Distance between the lid helix and the active site flap helix, as a function of the SMD simulation time. *Right*: Time evolution of the cumulated work performed during the SMD simulation.

The SMD simulations suggested that the lid region is more flexible than the active site flap region. Initial contacts between the lid and flap regions mainly involved polar and hydrophobic residues and these contacts break as the helices moved away from each other (during unrestrained SMD simulations) (Fig. 3.15). The final exposure of the hydrophobic residues to the surrounding solvent suggests that interactions with

hydrophobic interfaces are necessary to bring about lipase activation. Overall, this indicates that the displacement of the lid region is a likely structural determinant of M37 activation.

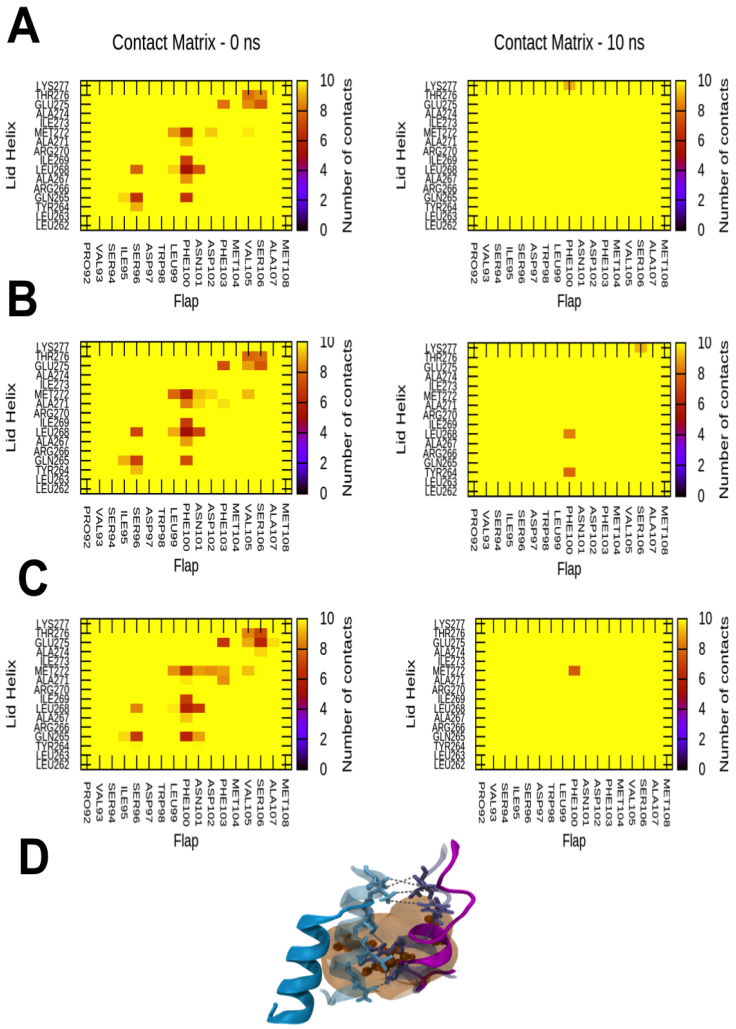

Fig. 3.15: Contact matrices calculated for three of five unrestrained SMD simulations of M37 in water using the lid-flap distance CV, probing the activation mechanism of M37 (**A-C**). The matrices show contacts for the first and final frames of the SMD simulations. A 0.4 nm cut-off was used to calculate the contacts between residues of the lid and active site flap helices. **(D)** Position of the lid helix (cyan) and active site flap (purple) region at the start (transparent) and end (opaque) of an SMD simulation. The active site pocket is shown as an orange surface, and the underlying catalytic residues as orange van der Waals spheres. The hydrophobic residues within the functional helices are shown as sticks. Initial contacts are indicated by the dotted lines.

3.3.4 Interfacial activation of M37 with a natural substrate

Beyond the intrinsic dynamics of M37 observed both in solution and at the interface with lipid bilayers, it is important to characterise the motions of the protein in contact with an interface formed by its natural substrate. It has been reported that the related TLL enzyme could maintain an open active state at a triglyceride interface [189]. Therefore, I performed atomistic simulations with an interface composed of tributyrin molecules, a triglyceride substrate of M37 [165].

The lipase was initially positioned above the layer of tributyrin. Within the first 20 ns of the atomistic simulation, M37 associated with the tributyrin interface, which induced a large-scale conformational change of the lipase (Fig. 3.16). The evolution of the distance between the COM of the lid region and of the active site flap observed was similar to that observed in the SMD simulations of M37 in water (Fig. 3.13B). Opening of the lipase structure occurred within the first 100 ns of simulation time, and was maintained for the duration of the simulation. This indicates that interfacial interactions with the tributyrin layer resulted in stabilisation of M37 in an open state. Consistent with the previously identified structural motions of the lipase, the conformational change corresponded to a displacement of the lid region, while the active site flap region remained mainly static (Fig. 3.16C). These (unbiased) simulations with a natural substrate confirm the potential mobility of the lid region and its involvement in the likely activation mechanism of the M37 lipase.

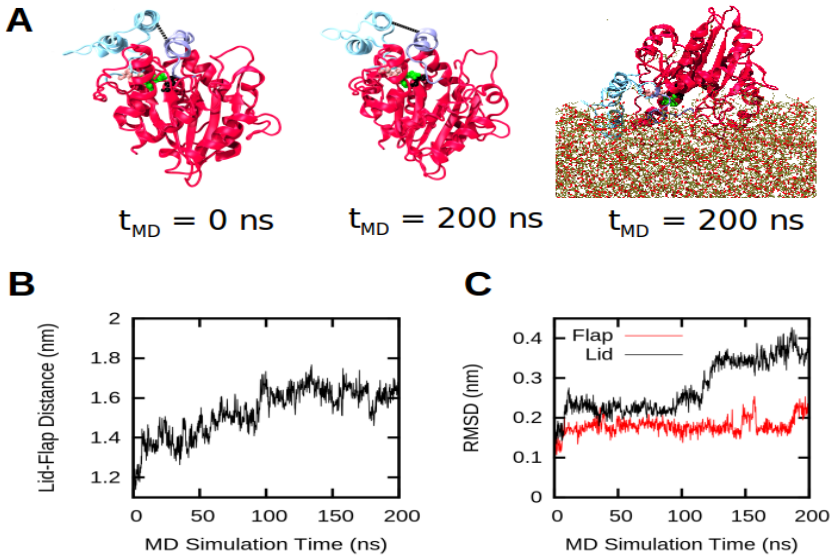

Fig. 3.16: **(A)** Structures of M37 extracted from the first and final frames of AT M37-tributyrin simulation. Position of the lid (cyan) and active site flap (purple) regions are shown at 0 ns (*left*) and 200 ns (*middle*); the water and tributyrin molecules are omitted for clarity. *Right*: Image of the final frame of an AT simulation of M37 interactions with a tributyrin interface, showing the tributyrin layer (brown sticks) and the binding orientation of M37 at 200 ns. The lipase is coloured in red (cartoon representation), the active site flap in purple, and the lid region in cyan. **(B)** Time evolution of the distance CV calculated between the COM of the lid and the active site flap helices. **(C)** Time evolution of the RMSD of Cα atoms within the lid and active site flap regions.

The binding orientation of M37 on the tributyrin surface was similar to the *Angled* orientation identified from simulations with an anionic lipid bilayer (Fig. 3.7). However, the large-scale displacement of the lid region captured on the tributyrin layer was not observed for the related orientation on the anionic bilayer (Fig. 3.11). These results show that the nature of the interface is essential for triggering functional motions of the bound lipase.

Docking calculations were also performed to predict binding poses of tributyrin substrate with both the closed form of M37 (crystal structure) and the open form observed at the end of the unbiased simulations with the tributyrin layer. Whereas

several binding modes of the substrate molecule were identified in the active site of the open form (Fig. 3.17), the calculations failed to find any viable docking poses for the closed form of the lipase. This indicates that the open form of M37, in which the lid is displaced, likely corresponds to an active state of the enzyme.

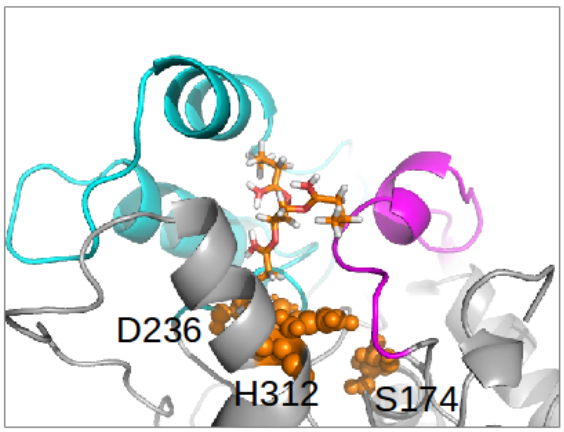

Fig. 3.17: The highest ranking docking pose calculated for a single tributyrin molecule within the opened structure of M37. Tributyrin is shown as orange (carbon), red (oxygen) and grey (hydrogen) sticks. The protein is shown in cartoon representation; the lid region is coloured in cyan and the active site flap region in magenta. The catalytic triad is shown as orange van der Waals spheres. The rest of the protein structure is coloured in grey.

3.3.5 SMD Simulations of Tributyrin Entry Into the Active Site Region

I have identified functionally relevant motions of the lid region that enable the opening of the closed conformation of M37. However, in order to determine whether the M37 structures bound to the tributyrin layer represent a partially or fully open state, additional SMD simulations were performed to explore possible entry pathways of the tributyrin substrate molecule into the enzyme active site. A similar protocol has previously been applied to study activation pathways and substrate binding mechanisms of other lipases [190, 191]. Here, M37 was defined as strictly *opened* if the lid region

did not impede substrate entry into the active site.

A tributyrin molecule from the layer to which the enzyme was bound was used to simulate substrate binding. This tributyrin molecule was initially wedged between the active site flap and the lid region at the end of an unbiased AT simulation, and pulled into the binding pocket during the SMD simulation, specifically towards the catalytic Ser174 residue (Fig. 3.18). The SMD simulations showed complete entry of the substrate into the binding pocket (Fig. 3.18B). During this process both the lid and the active site flap regions moved only minimally (Fig. 3.18C). Consequently, the distance between the lid and flap regions showed only small variations (Fig. 3.18D). Therefore the M37 conformation observed at the end of the unbiased tributyrin simulations corresponds to a fully open form, which could enable catalytic activity.

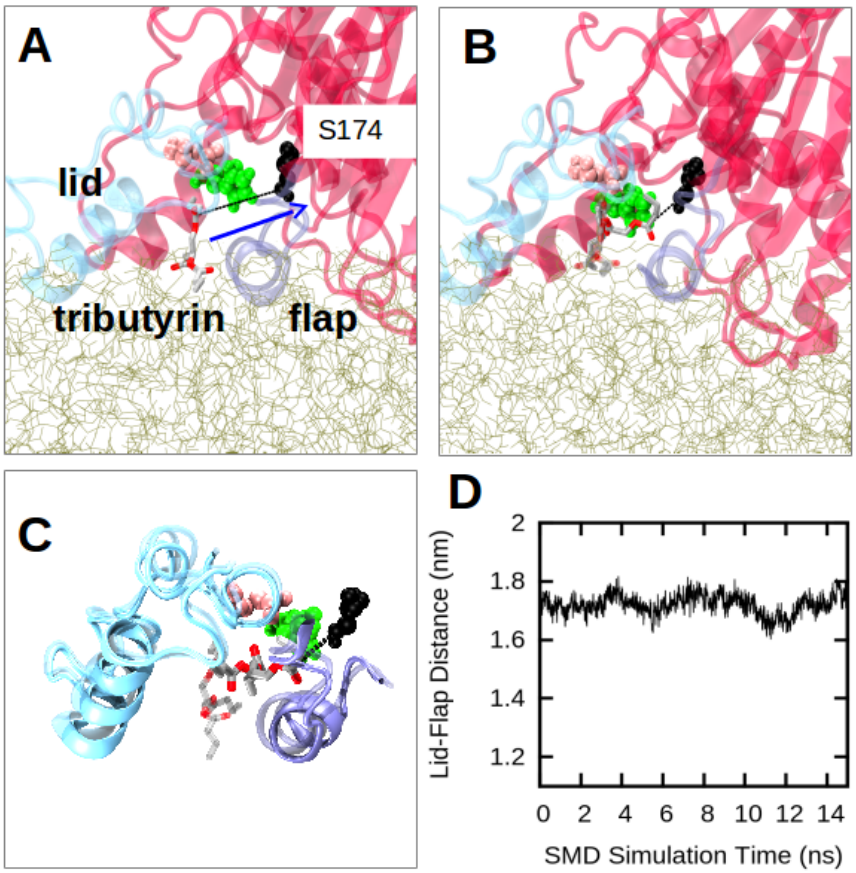

Fig. 3.18: **(A)** Setup of the SMD simulations in which a tributyrin molecule (grey and red sticks) was pulled from the tributyrin layer (light brown sticks) toward the active site Ser174 residue of M37. The catalytic residues Ser174 (black), Asp236 (pink), and His312 (green) are shown as van der Waals spheres. The arrow represents the direction of the pulling force and a dashed line indicates the distance between the tributyrin and Ser174 molecules. **(B)** Starting (transparent sticks) and end positions (opaque sticks) of the tributyrin molecule from a selected SMD simulation. **(C)** Cartoon representations of the starting (transparent) and end positions (opaque) of the lid (cyan), the active site flap (purple), and the tributyrin molecule (sticks) from the selected SMD simulation. **(D)** Time evolution of the distance CV calculated between the COM of the lid and the active site flap.

3.4 Discussion

Interactions of lipases with a range of different interfaces including lipid bilayers have been investigated using many experimental and computational techniques [71, 96, 187, 192, 193]. However, the mechanisms of their association with lipid bilayer and related interfaces remain incompletely understood, especially in the context of lipase activation, although electrostatic interactions are generally considered important [187, 192]. This is exemplified by studies reporting that the TL lipase interacts with both POPG and POPC lipid vesicles, but is only active in the presence of POPG vesicles [94, 95]. This correlates nicely with the presented CG simulations, which showed that the related M37 enzyme could extensively interact with anionic PG-containing bilayers, but not with zwitterionic PC bilayers.

Studies of TLL using fluorescence microscopy suggested that this lipase could adopt different orientations on the lipid vesicles, comparable to the different binding orientations exhibited in the simulations of M37 with anionic lipid bilayers (*i.e Up, Angled, Down*). Other computational studies have also revealed that lipases may exhibit varying orientations at different interfaces, which have different functional consequences regarding lipase activation [91, 194]. The amino acids mediating interfacial interactions were found to be different when comparing the interactions with PC *vs* PG molecules within the anionic bilayer. Hydrophobic residues mainly showed interactions with PC molecules, whereas basic residues mediated interactions with the PG molecules, particularly for the *Angled* and *Up* states (Fig. 3.10). Therefore, the balance of hydrophobic and predominantly electrostatic interactions at the interface determined, for the most part, the orientation adopted by the enzyme (Fig. 3.7). The agreement with atomistic simulations suggests that the Martini model is able to correctly distinguish the various types of lipid-enzyme interactions at bilayer interfaces [195, 196].

Whilst the CG simulations were able to identify interfacial binding of M37 with

anionic bilayers, simulations of these systems at atomistic resolution revealed that M37 interfacial interactions did not result in displacement of either the lid or active site flap region of the lipase (Fig. 3.11). The varying binding orientations exhibited by M37 are thus non-functional regarding lipase activation, as is similar to TLL interactions with large unilamellar vesicles (POPG) relative to small unilamellar vesicles [94]. It would therefore be of interest to investigate the effects of membrane curvature, allowing a more detailed characterisation of interfacial interactions of M37 with membrane surfaces.

The atomistic simulations allowed identification of a possible activation pathway of M37 induced by lid displacement. This could be promoted by interfacial interactions with a tributyrin layer, a natural substrate for M37 (Fig. 3.16 and 3.17). The lipase adopted a similar *Angled* orientation that was identified as a binding orientation in simulations with anionic lipid bilayers. This suggests a selection process in the interfacial binding orientations for the enzyme to become activated when a natural substrate interface is present. During activation, the lid motion is required to uncover the underlying catalytic site for subsequent entry of substrates (Fig. 3.18). This conformational change is somewhat unexpected as activation of structurally similar enzymes (e.g. TLL and RML) instead involves a motion of lid regions that align to the active site flap of M37 [53, 56]. However, it is reasonable to suggest that M37 may display a novel activation pathway, given that the lid region of M37 (residues 235-283) is not found in these related enzymes. Furthermore, the active site flap also exhibited less flexibility relative to the lid region in the lipid bilayer simulations, highlighting the possible functionality of the lid region, and the rigidity of the active site flap region. This was further confirmed by position restrained SMD simulations in which the active site flap region was barely displaced relative to the lid region (Fig. 3.14).

In addition, the hydrophobic residues that mediated contacts between the lid and

the active site flap regions in the closed form of M37 became exposed to the surrounding solvent during the opening motion, induced by the lid region displacement (Fig. 3.15). Consequently, hydrophobic interfaces can play a decisive role in stabilising the active form of the lipase, thus enhancing its activity. The nature of the interface is clearly important, given that the lid opening motion was only seen at the substrate interface (tributyrin) and not at the lipid bilayer interface. Similarly, soluble phospholipase A2 enzymes are allosterically modulated to adopt an "active" state by specific interfacial interactions with their natural substrates [191].

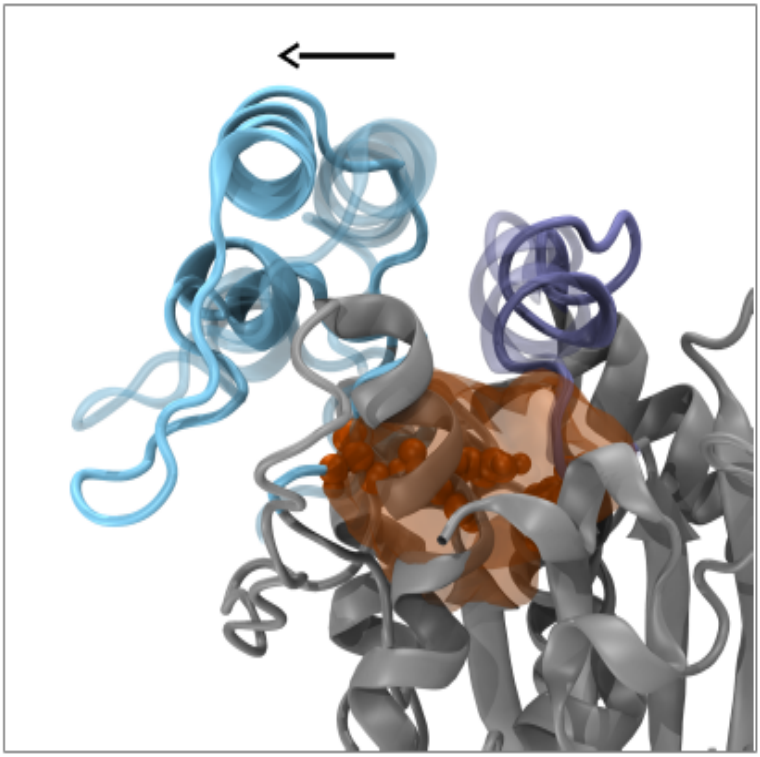

Fig. 3.19: Proposed activation mechanism of M37 induced by lid displacement (cyan). The catalytic residues are shown in orange van der Waals spheres, and the binding pocket as an orange surface. The active site flap is coloured in purple. The position of the lid region and active site flap region in the closed crystal structure of M37 are transparent, while those in the "open" model are opaque.

3.5 Conclusions

In this chapter, I have presented a multi-scale simulation approach that enabled the characterisation of the interfacial interactions between M37 and hydrophobic surfaces. The exposure of hydrophobic residues within the functional regions of M37 rationalises the observation of lid displacement during interactions with a substrate interface. An open model of M37 is proposed that reflects lid motion as a possible structural determinant of substrate entry in the catalytic site.

The characterisation of a possible activation mechanism for this relatively understudied lipase serves as a basis for subsequent experimental and computational studies. These could focus on stabilisation of the open form, perhaps through mutagenesis, to optimise lipase activation, most likely resulting in higher catalytic efficiency in differing environments. This approach has been used for the *Burkholderia cepacia* lipase (BCL), in which in silico mutations of hydrophobic to polar amino acids were observed to stabilise the open form of BCL in solution [197]. The role of other environmental factors, such as temperature and pH, could also be investigated in the context of lipase activation. As mentioned previously, M37 is a psychrophilic lipase, which exhibits maximal activity at around 25°C for short-chain triglyceride substrates [165]. Characterisation of temperature effects on M37 activation has been attempted, as will be detailed in Chapter 7. These studies could be extended, and combined with mutational experiments to investigate which structural features underlie this psychrophilicity.

Naturally, further investigation of the open model of M37 presented here relies on the assumption that this is a true representation of the active state of the lipase. Currently, the suggested model is reliant on the ability of the MD forcefield (GROMOS 54A7) and methodology to accurately reproduce lipase dynamics as would be observed in experiment. The model should therefore be validated by additional structural studies of the M37 lipase, including characterisation of the open form of the

lipase. Furthermore, it would be of general interest to characterise the free energy surface underlying the activation of M37 in order to produce a more rigorous mechanistic model for the proposed open structure. This would also probe the stability of the open model of the lipase at a hydrophobic interface, although current AT-MD simulations of M37-tributyrin interactions showed that the lipase was able to maintain its opened structure for the duration of multiple 200 ns simulations (lid displacement occurred within the first 20 ns). Additionally, analysis of the energetics of lid displacement at the interface and in aqueous solution would provide enhanced insight into how these different environments are able to stabilise/destabilise the proposed open form of M37. Importantly, the steered MD simulations exploring possible activation mechanisms of M37 were performed in water, which would not be expected to naturally produce an interfacially activated form of the lipase [42]. This is thought to be related to unfavourable solvent exposure of the hydrophobic amino acids present in the lid and active site flap regions, as has been observed for other lipases [55–57, 97, 197]. Indeed, unbiased AT simulations of the open M37 model in water resulted in closing of the opened lid region, moving back to its original position as identified in the crystal structure. Attempts have been made to investigate the free energy surface of M37 activation, revealing a complex underlying energy landscape, as detailed in Appendix A. The proposed open structure of the M37 lipase thus provides a basis for further investigation into M37 lipase dynamics and energetics at the interface, and possibly other environments.

Chapter 4

Thermomyces lanuginosus Lipase Interactions With Triglyceride Surfaces

This chapter discusses the affect of mutating the lid residues in *Thermomyces lanuginosus* lipase (TLL) on the binding and activation mechanisms upon enzyme interaction with triglyceride surfaces. Coarse-grained and atomistic simulations were used to study these properties, whilst enhanced sampling methods were employed to estimate the energetics of lid opening for different mutants. Finally, the affect of triglyceride hydrolysis on lipase binding and enzyme dynamics was investigated. This work was done in collaboration with colleagues at Novozymes, Copenhagen (Dr Jakob Skjold-Jorgensen and Allan Svendsen), who provided details of experimental data to aid in the design of the simulations.

4.1 Introduction

Altering the amino acids present in the lid regions of lipases has been shown to affect lipase activity, enantioselectivity, and thermostability [198]. Site-directed mutagenesis of the lid region can generate lipase variants with advantageous properties, such as enhanced catalytic efficiency, altered substrate specificity, and superior stability in non-native environments, such as organic solvents [198–200]. This chapter aims to provide insight into how mutating the lid residues of TLL affects the conformational dynamics of the lipase, particularly of the lid region, and whether or not this affects association of the lipase with triglyceride interfaces. Specifically, how does mutating the lid region of a lipase to the lid region present in, for example, an esterase affect the structure and conformational flexibility of the lipase, and what does this mean for the interfacial activation of the lipase?

TLL is a well-characterised lipase with biotechnological applications in the detergent, cosmetics, and pharmaceutical industries, making it an interesting candidate to study the affect of mutations within the lid region [2]. Furthermore, structural studies of the enzyme have aided in elucidating the role of the lid region in lipase activation, with crystal structures available for both the closed (PDB: 1DT5) and open (PDB: 1DTE) forms of the enzyme (Fig. 4.1) [55, 56]. The lid region is comprised of residues 82-98 and the catalytic triad consists of S146, H258, and D201 residues [56]. Furthermore, the lipase contains a glycosylation site at position N33, distal to the lid region, consisting of the N-high mannose sugar type, thought to play a role in lipase binding properties to micelles [201, 202].

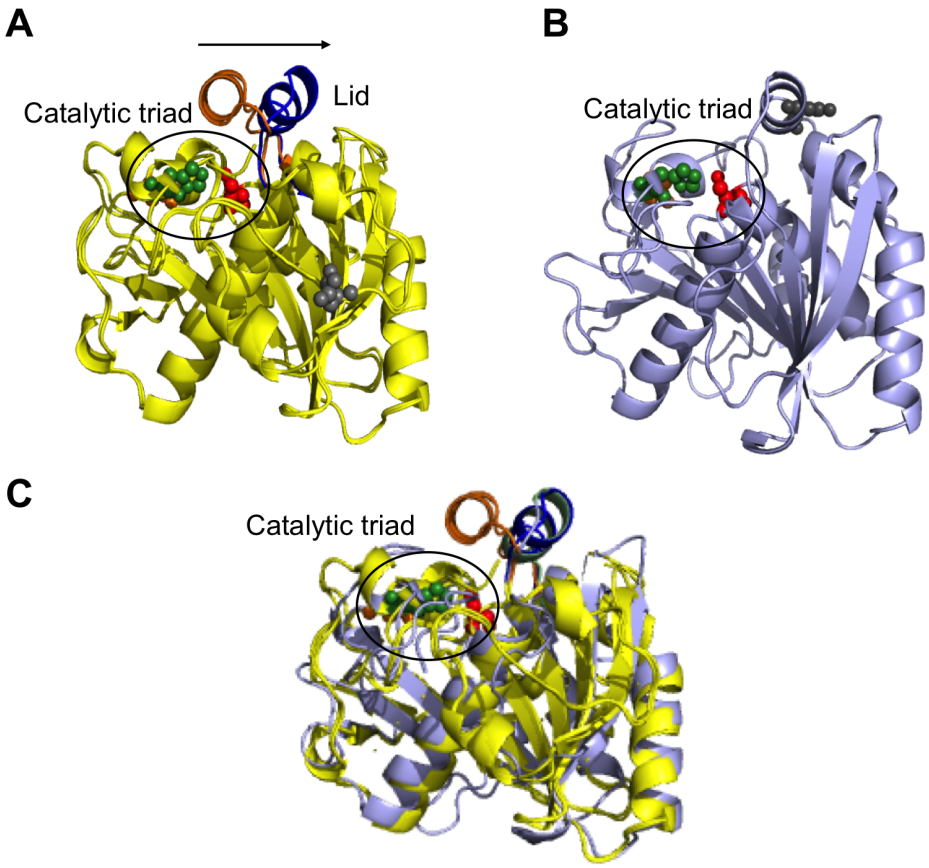

Fig. 4.1: **(A)** Structural alignment of the crystal structures of the open (orange lid; PDB: 1DTE) and closed (blue lid; PDB: 1DT5) forms of TLL shown as cartoon representations [55, 56]. The catalytic triad comprises S146, H258, and D201 residues shown as red, green, and orange van der Waals spheres, respectively. The N33 glycosylation site is shown as grey van der Waals spheres. The arrow indicates the motion of the lid region upon interfacial activation of TLL. **(B)** The crystal structure of ferulic acid esterase (FAEA) shown as a light blue cartoon representation [203]. The catalytic triad (S133, D194, H247) is shown in the same respective colours and representations as in **(A)**. The N79 glycosylation site is shown as grey van der Waals spheres. **(C)** Structural alignment of the TLL and FAEA structures shown in **(A)** and **(B)**, using the same colour scheme. Only the catalytic triad to TLL is shown. The N-glycosylation sites are omitted for clarity. The helix region of FAEA (residues 70-71) is coloured in green, and aligns well to the open position of the lid region of TLL (coloured in blue).

The interfacial activation of TLL has been characterised by previous experimental studies [204, 205]. Recently, it was shown that mutating the residues within the lid region of this lipase resulted in altered catalytic properties in different environments

[57, 206]. The purpose of these investigations was to generate variants of TLL with enhanced biocatalytic properties, such as the ability to act on water-soluble substrates. To this end, five different lipase variants were constructed based on the lid region present in an esterase called ferulic acid esterase (FAEA) (Fig. 4.2) [203, 206]. FAEA possesses a remarkably similar structure to TLL (RMSD: 0.13 nm over 205 aligned Cα atoms) but does not exhibit interfacial activation, and instead hydrolyses water-soluble substrates [207]. Furthermore, the esterase possesses a glycosylation site at N79 consisting of the N-high mannose type, which resides within a helix region (residues 71-77) that aligns to the lid region of TLL (Fig. 4.1B). The esterase thus provides an interesting test case for investigating the role of the lid region in TL lipase activation and catalysis.

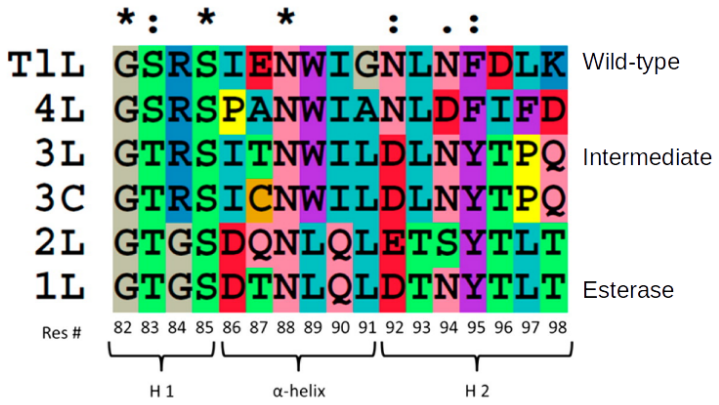

Fig. 4.2: Multiple amino acid sequence alignment of the lid region (residue 82-98) of the wild-type, 1L, 2L, 3L, and 4L TLL mutants produced in [206]. The lid domain is divided into three motifs: residues 82-85 correspond to the anterior hinge domain (H1); residues 86-91 correspond to the α-helical domain; residues 92-98 correspond to the posterior hinge domain (H2). Colour code: grey for Gly, green for Ser and Thr, yellow for Pro, orange for Cys, pink for Asn and Gln, purple for Phe, Tyr, and Trp, red for Asp and Glu, blue for Arg and Lys, and cyan for Iso, Ala, and Leu. Asterisks denote a single fully conserved residue; colons denote regions of conserved residues that are highly similar, and the period denotes a region of conserved residues with a low degree of similarity. The predominant enzymatic activity is listed for selected mutants, as observed in [206]. Image reproduced from [206].

This chapter will focus on three of the five lipase variants investigated in this experimental study: the wild-type, 1L, and 3L lipase variants. The 1L protein was generated by replacing the entire lid region of TLL by that of the FAEA enzyme (residues 71-79; Fig 4.1B), whereas the 3L lipase was generated by only mutating selected residues in the lid region (sequences shown in Fig. 4.2). The experimental study showed a completely altered activity profile for the 1L mutant relative to the wild-type, with more pronounced activity on water-soluble substrates, and very little activity at the interface [206]. This activity profile is thought to reflect the esterase activity conferred by the mutations within the lid region of 1L, in which the native residues were entirely replaced by those in the helix of FAEA (Fig. 4.2). Interestingly, the 3L variant exhibited higher interfacial activity than the wild-type lipase, as well as a relatively high activity on water-soluble substrate, therefore exhibiting both lipase and esterase characteristics [206].

In this chapter, both coarse-grained (CG) and atomistic (AT) MD simulations were performed in order to provide additional molecular insight into the structural and functional consequences of the lid mutations in the 1L and 3L lipases compared to the wild-type TLL. The interfacial binding properties and structural dynamics of the lipase variants were investigated in the presence of a natural triglyceride substrate interface, and compared to simulations in water. It was found that lid flexibility was affected by the mutations within this structural region, particularly for the 1L variant, influencing lid conformation both in water and at the interface. These observations were confirmed by performing steered MD simulations to investigate the lid opening process for all variants, indicating a higher activation barrier for the 1L lipase. Finally, attempts are made to characterise how lipase-catalysed hydrolysis of triglyceride molecules at the interface affects lipase association and structural dynamics. Accumulation of hydrolysis product at the interface is thought to play a role in lipase inhibition

[170, 208].

4.2 Methods

4.2.1 CG-MD Simulations

Crystal structures for the open (PDB: 1DTE) and closed (PDB: 1DT5) form of wild-type TLL were downloaded from the PDB [56]. These structures were used to generate atomic coordinates for the TLL variants by altering the lid composition of wild-type TLL using the mutagenesis tool in PyMOL (www.pymol.org). The side chain rotamer with the highest occupancy and least number of clashes with the surrounding protein environment was chosen for each mutation. All structures were then energy minimised using steepest descent for 1000 steps, and converted to CG representations using the the same protocol as presented in Chapter 3. The EDM elastic network model was used to maintain secondary and tertiary structure, and was parameterised for each variant as described in Chapter 3 [134, 138, 143]. All CG simulations were performed using the Martini 2.2 forcefield and the Martini polarisable water model [122, 148].

Triglyceride System Setup

A supported triglyceride system was constructed to mimic the systems used in the experimental study (microtitre plate assays) [206]. The triglyceride interface was generated by performing self-assembly simulations of 501 trioleate triglyceride molecules that were randomly inserted in a simulation box (final dimensions: 10x10x15 nm). The systems were solvated and simulated for 500 ns, allowing self-assembly to occur. The parameters for the CG trioleate model were kindly supplied by Vattulainen *et al* [209]. The self-assembled triglyceride layer was then placed directly above a hydrophobic support and the system was solvated [84]. Energy minimisation was performed

using the steepest descent algorithm for 1000 steps, followed by an equilibration simulation of 10,000 steps at constant temperature and pressure (NPT ensemble). The equilibrated system was then simulated for 500 ns in the NPT ensemble using a time step of 20 fs, indicating that the layer was stable on the surface. The Berendsen thermostat was used to maintain temperature at 298 K with a coupling constant of 0.3 ps; the Berendsen barostat was applied to maintain pressure at 1 atm, with a coupling constant of 3.0 ps, using anisotropic pressure coupling, and a compressibility of 0.5×10^{-5} bar^{-1} in x and y planes, and 3.0×10^{-5} bar^{-1} in the z plane [120]. Non-bonding interactions were modelled using shift functions; LJ interactions were evaluated within a 1.2 nm cut-off and shifted within a 0.9 nm cut-off distance, Coulombic interactions were evaluated within a 1.2 nm cut-off, and shifted within a 0.0 nm cut-off distance. These interaction parameters reflect those applied in previous simulations of molecules with the support surface [84].

The CG lipase structures were then individually placed 3 nm above the supported triglyceride layer and counter ions were added to neutralise the charge of the systems. Three different starting orientations were simulated to minimise any bias occurring from the initial orientation of the enzyme (Fig. 4.3). The systems were subsequently energy minimised, and equilibrated for 10 ps in the NPT ensemble. Five replicate simulations of 2 μs each were performed for each lipase variant, including the open and closed form of the wild-type TL lipase, resulting in a fifteen replicate ensemble for each of the lipase molecules.

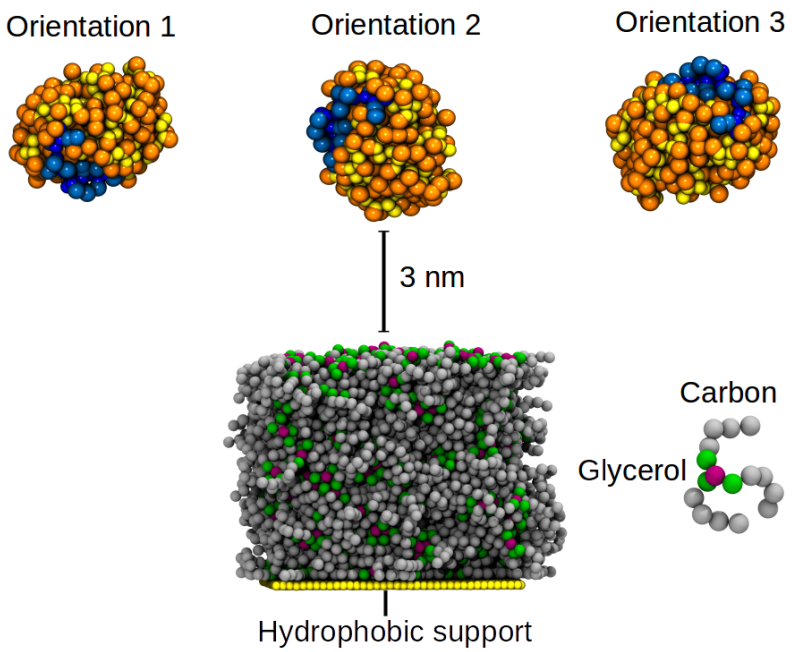

Fig. 4.3: CG setup of TL lipase-triglyceride simulations. Three different starting orientations of TLL positioned above the supported triglyceride layer were simulated (2 μs). The lipase molecules are shown as yellow and orange van der Waals representations; the lid regions are coloured in blue. The triglycerides are shown as green (esters), grey (carbons), and purple (glycerol) van der Waals spheres. The support surface is shown as yellow van der Waals spheres.

4.2.2 AT-MD Simulations

All AT simulations were performed using the GROMOS 54A7 forcefield and SPC water model [112, 177, 178]. A triglyceride layer composed of tributyrin molecules was constructed using the same method described in Chapter 3, applying identical simulation parameters described therein. Atomistic models of the individual lipase variants were generated by using the *pdb2gmx* tool in GROMACS. The AT protein models were then individually placed 1 nm above the pre-formed tributyrin layer, and water molecules were added. The systems were then energy minimised and equilibrated for 100 ps in the NPT ensemble. Production simulations consisting of three repeat simulations for each variant were performed for 100 ns each, totalling nine simulations.

Three repeat simulations of 100 ns were also performed for the lipase variants in water, allowing comparison of the interfacial behaviour.

4.2.3 Steered MD simulations

Steered MD (SMD) simulations were performed using the Plumed 2.1 plugin for GROMACS [159, 183]. The equilibrated, initially closed, structures of each lipase variant in water were used as a starting point, applying the same AT simulation parameters as above. A distance collective variable (CV) was chosen to investigate the opening motion of the lid region of the TL lipase variants. This CV was defined as the distance between the centre of mass (COM) of the Cα atoms of I255, positioned distal to the lid region, and the 87^{th} residue, present in lid region (further details in Section 4.3.5). A harmonic bias potential was applied to the CV, moving at a constant velocity of 0.06 nm ns^{-1} with a spring stiffness of 1000 kJ mol^{-1}. The limit values for the distance CV were measured from the closed and open crystal structures of wild-type TLL.

Additional SMD simulations were also performed for the 1L lipase associated with the tributyrin layer. These were initiated from the final frame of an unbiased AT-MD simulation of 1L with the tributyrin interface (100 ns), applying identical SMD simulation parameters as above. The SMD simulations were performed by either applying position restraints to the protein atoms (Fc = 1000 kJ mol^{-1}) except those within the lid region (residues 82-98), or no position restraints at all.

4.3 Results

4.3.1 Interfacial Interactions of TL Lipase Variants with Triglycerides Explored via CG Simulations

The affect of lid mutations on the catalytic properties of the TL lipase has been characterised experimentally [57, 206]. Here, CG simulations were used to investigate the interfacial binding properties of selected TLL mutants in more detail. An interface composed of a natural triglyceride substrate of TLL (trioleate) was used to study this. Three lipase variants were considered: the wild-type TLL, exhibiting lipase activity, a 1L lid variant, exhibiting esterase activity, and a 3L variant, exhibiting superior esterase *and* interfacial lipase activity (Fig. 4.2) [206]. Additionally, both the open and closed forms of the wild-type TL lipase were considered in order to capture the binding behaviour an interfacially activated protein.

All the lipase variants exhibited association with the triglyceride layer, and did not dissociate from the interface once bound. However, the number of simulations in which the lipase associated with the interface differed for each variant (Table 4.1). Interestingly, the 1L variant displayed the least number of binding events within the 15 replicate ensemble, where one binding event represents association of the lipase with the interface in one repeat simulation. The 3L variant displayed the most binding events, associating with the triglyceride surface in 14/15 simulations (Table 4.1).

Simulated system	Binding events
Wild-type: Open	10/15
Wild-type: Closed	13/15
1L	9/15
3L	14/15

Table 4.1: The number of binding events shown for CG-MD simulations of TLL variants (1L, 3L, wild-type) with a triglyceride layer. One binding event represents the association of the lipase variant with the interface within one simulation. Once bound, the lipase did not dissociate form the interface in any of the simulations (2 μs). 15 replicates were simulated per variant.

Representative structures for the bound lipase molecules are shown in Fig. 4.4A. These indicate that the orientation of the lid region at the interface differed for the bound lipase variants. The lid region of the wild-type closed lipase and the 1L lipase were much more solvent exposed, typically pointing away from the interface, whilst the lid region of 3L and wild-type open variants embedded within the interface, buried from the surrounding solvent (Fig. 4.4A). The interfacial orientation can also be evaluated from partial density profiles, calculated by averaging over the last 25% of simulation time, where there is wider density peak for the lid region of 1L, as well as the closed form of the wild-type lipase, within the density of the interface (Fig. 4.4B and C).

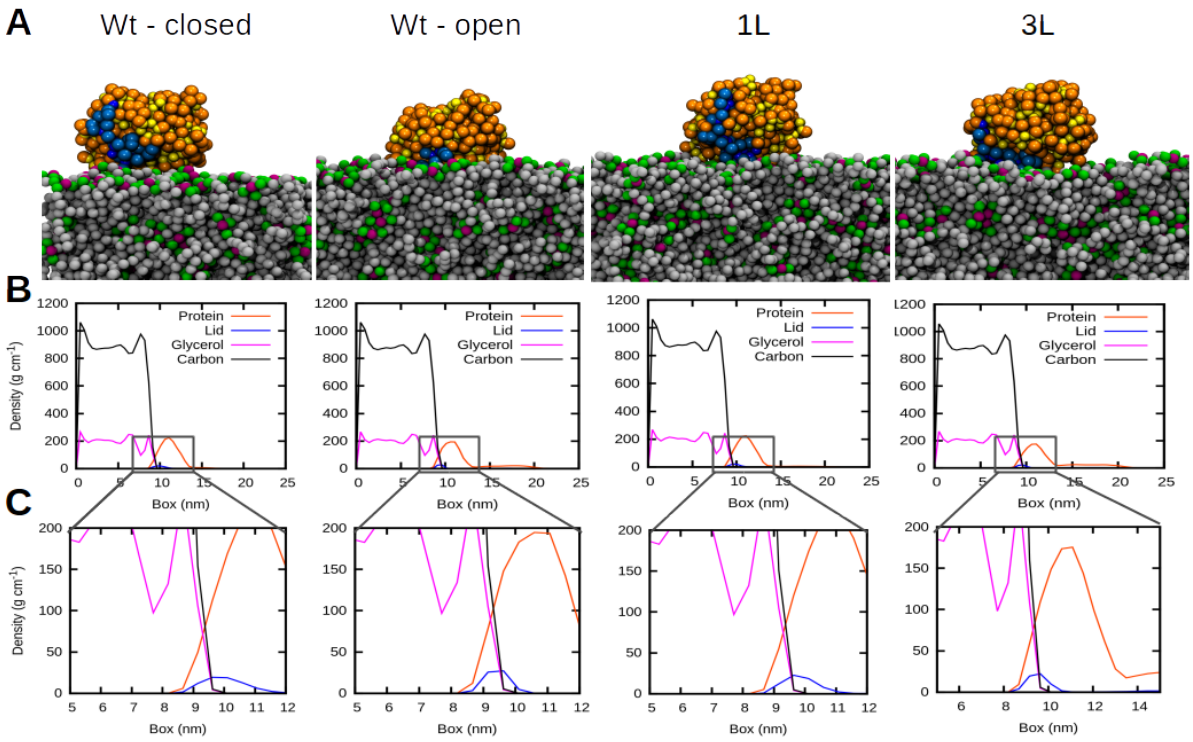

Fig. 4.4: **(A)** Images from CG-MD simulations of lipase-triglyceride systems showing representative binding orientations of the lipase variants at the triglyceride interface. All components are shown as van der Waals spheres; carbon tails of the triglyceride molecules are coloured in grey, ester groups in green, and glycerol in purple. The lipase molecules are coloured in yellow (backbone) and orange (side chains) and the lid region is shown in blue. Water molecules are omitted for clarity. **(B)** Partial density profiles showing the densities of triglyceride components (carbon tails and glycerol group) and the lipase protein within the simulation box, calculated for the last 25% of simulation time, for each individual simulation shown in **(A)**. **(C)** Close up of the same density plots shown in **(B)**, showing the difference in the densities of the lid region within the triglyceride interface.

The interfacial interactions exhibited by the lipase variants were further characterised by calculating the average number of contacts between the lipase and the triglyceride molecules (Fig. 4.5). All simulations in which the lipase associated with the interface were considered for each variant. The interaction data was normalised by averaging the total number of contacts across the repeat simulations within each ensemble.

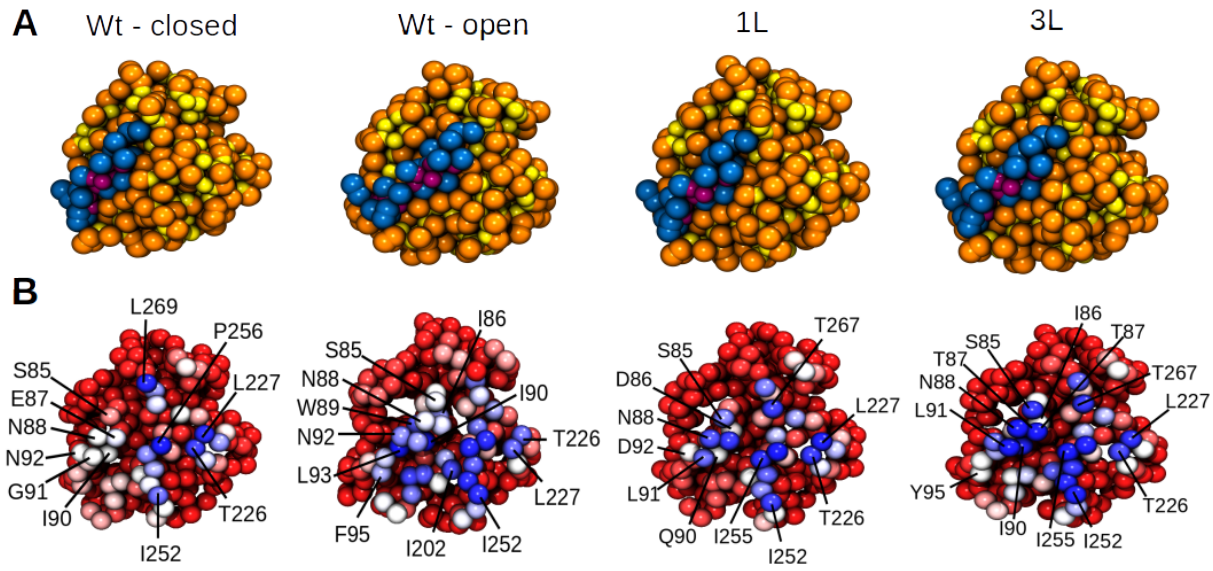

Fig. 4.5: Contacts calculated between the TLL variants and the triglyceride molecules during CG simulations of the variants with a triglyceride interface. **(A)** Reference orientations of the lipase variants for the contact residue maps shown in **(B)**. All structures are shown as van der Waals representations; the lid is coloured in blue, and rest of the protein in orange (amino acid side chains) and yellow (amino acid backbone). **(B)** Contact residue maps showing the interacting residues for each lipase variant, shown in the exact same orientation as in **(A)**. The residues are coloured according to a blue/white/red scale (blue = greatest number of contacts, red = none or very little contacts). Contacts were calculated within a 0.8 nm cut-off radius.

Both polar and hydrophobic amino acids mediated contacts with the interface for all of the variants, representing interactions with both polar glycerol groups and non-polar carbon tails of the triglyceride molecules. Importantly, the interfacial contacts residues consistently mapped to the lid region and surrounding protein residues for each of the lipase variants, emphasising the role the lid region plays in mediating interfacial interactions. However, the overall number of contacts made by residues within the lid region differed for each of the variants, which is consistent with the different interfacial orientations observed for the variants (Fig 4.4A). For example, the lid region of the 3L lipase, of both esterase and lipase character, displayed the largest number of interactions with the interface relative to other residues in the 3L protein,

suggesting that the lid mainly remained in contact with the interface once bound (Fig. 4.6). Interestingly, the most commonly interacting residues included the mutated lid residues, such as G91L and E87T, but also the wild-type residues such as I86 and N88. Conversely, the mutated residues within the lid region of the 1L variant, such as I86D, I90Q, and N92Q, mediated relatively less contacts with the interface than other residues within the lid region, suggesting that these negatively charged substitutions could lead to diminished interfacial contact via the lid region of 1L.

The different interfacial orientations of the lipase variants on the triglyceride surface were further analysed by evaluating the rotational and translation motions exhibited by the variants during association with the interface (Fig. 4.6). This was done by applying the same protocol as presented in Fig 3.6 in Chapter 3. Briefly, the rotational motions of the lipase were described by a defined angle between the lipase and the interface (Rzz). This was obtained by calculation of a rotation matrix with respect to a reference orientation of the protein at the interface (Fig. 4.6A). The translational motions were described by measuring the centre of mass (COM) distance (d) between the protein and triglyceride interface over simulation time.

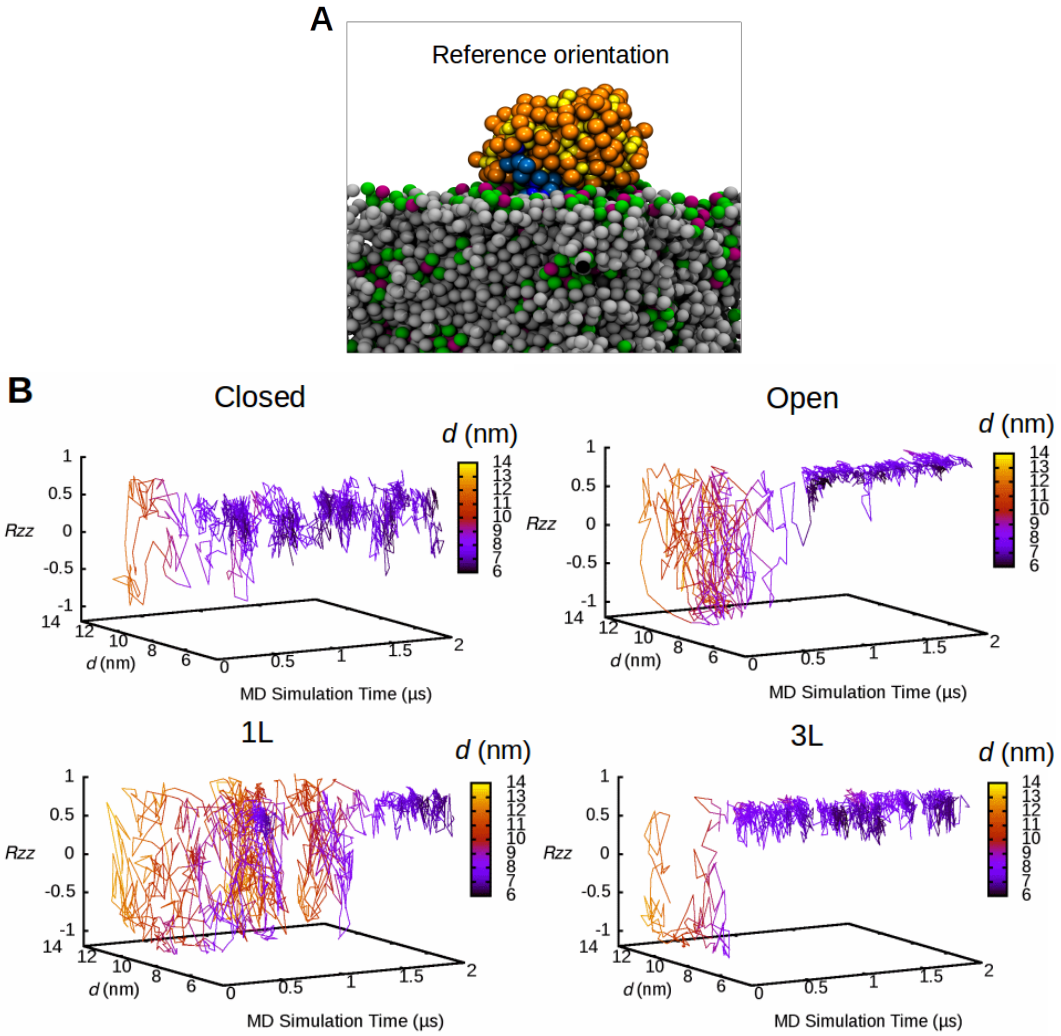

Fig. 4.6: **(A)** Reference orientation of the TLL wild-type variant bound to the triglyceride interface. This reference orientation was used for calculation of rotation matrices evaluating the rotational motions of TLL variants during CG simulations with a triglyceride interface. The same colour scheme is used as in Fig. 4.4. **(B)** Time evolution of d and Rzz for representative CG-MD simulations of lipase variants with a triglyceride surface. Rzz defines the transition of the lipase variant relative to the reference orientation; when $Rzz = 1$, the lipase variant adopts the orientation of the reference structure.The d metric is calculated as the z-component of the distance between the lipase and the triglyceride interface.

Consistent with previous observations, the association process and interfacial orientations differed for the lipase variants. Specifically, the 1L protein exhibited a considerably slower association time relative to the 3L protein, reflected by substantial

variations in Rzz and d between 0-1.5 μs (Fig. 4.6B). Furthermore, once bound, the interfacial orientation of 1L deviated from the reference structure in which the lid region is buried within the interface ($Rzz \sim 0.5$). This is reflective of a more solvent exposed orientation of the 1L lid region, in accordance with previous analyses. Similarly, the closed wild-type structure exhibited a greater distribution of interfacial binding orientations relative to its open counter part, resulting in similar Rzz values as the 1L variant.

Extension of the rotation and translational analysis to the remaining replicates in the individual ensembles resulted in very similar observations regarding interfacial orientation. The collective data allowed calculation of a 2-dimensional density landscape representing the normalised density of the binding orientations exhibited by the variants, as a function of d and Rzz (Fig. 4.7). Accordingly, the density landscapes indicate a substantial difference in the average interfacial enzyme orientation adopted by the 1L variant relative to the 3L variant. As noted previously, this can be seen from variation in Rzz value for the 1L lipase and closed wild-type lipase, showing a distribution between Rzz = 0-0.5, compared to Rzz = 0.5-1 for the wild-type open and 3L variants (Fig. 4.7).

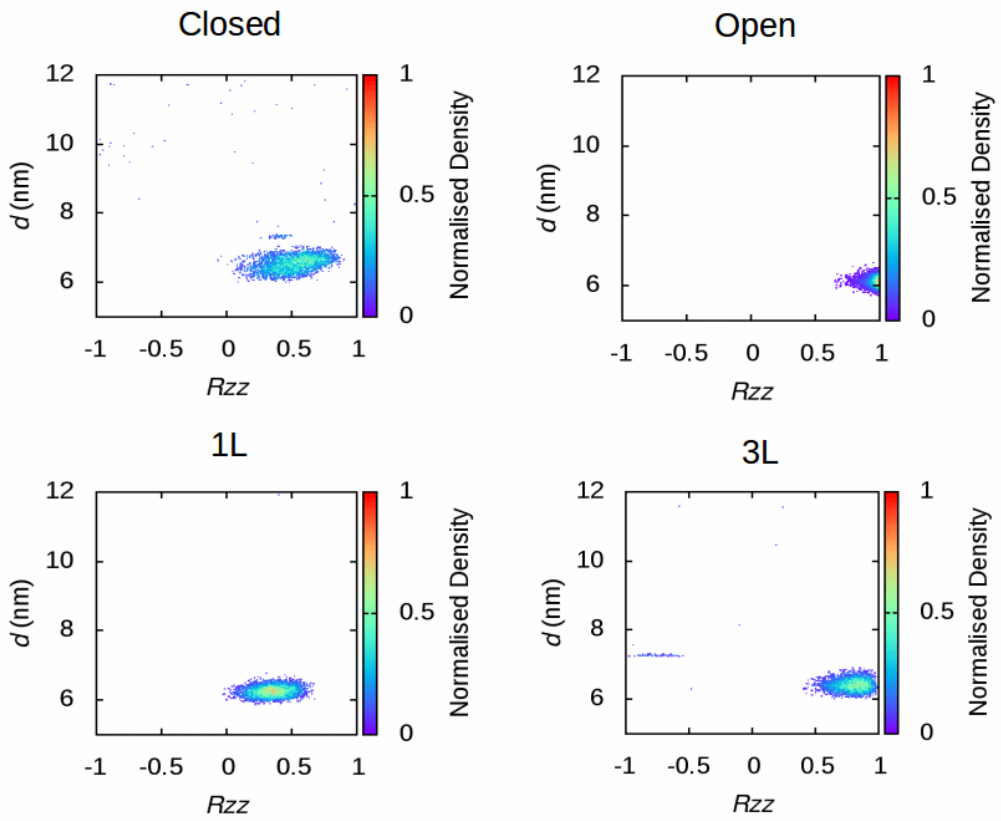

Fig. 4.7: 2D density landscapes calculated for CG-MD simulation ensembles (15 repeats) for wild-type, 1L, and 3L TLL variants simulated with a triglyceride interface, as a function of d and Rzz. Representative data is shown in Fig. 4.6. The top of the triglyceride layer corresponds to the bottom of the map (textitd $\sim$ 5).

The overall larger distribution of binding orientations for 1L and the closed wild-type variants suggest that both the nature of the residues within the lid region, as well as their orientation (e.g exposed in the open wild-type variant), are important determinants of overall interfacial orientation. These data suggest that an altered binding orientation could be a factor in the diminished interfacial activity of the 1L variant observed experimentally, relative to the wild-type and 3L variants. Furthermore, the frequently sampled buried orientation of 3L lid region within the interface also suggests that lipase activation is favoured in this variant as a function of its lid residue

composition. Together, these results correspond well with the experimental activity data presented in [206].

4.3.2 Conformational Dynamics of Interfacially Bound TLL Variants

Atomistic (AT) simulations of the lipase interactions with triglycerides surfaces were performed in order to study the structural motions of the lipase variants in more detail. A triglyceride interface composed of tributyrin was used to investigate this, which a natural substrate for TLL [71]. The AT simulations (3 repeats) showed a general agreement with the CG simulations regarding lipase orientation at the interface (Fig. 4.8).

Interestingly, both the initially closed forms of the wild-type and 3L variants (Run 1 & 3) displayed partial interfacial activation during the simulations, whereas the lid region of 1L remained firmly closed for all replicate simulations. This can be seen from alignment of the final lipase structures with the closed and open crystal structures of wild-type TLL (Fig. 4.8), and is reflected in calculation of the COM distance travelled by the lid region during the simulations (Fig. 4.9A & C). The similar structural motions exhibited by the wild-type and 3L variants suggests that the lid mutation within 3L did not impact lid mobility or flexibility to a great extent. Furthermore, the overall protein structure remained relatively stable for each of the variants, although a degree of α-helicity in the lid region was lost in simulations of the wild-type and 1L variants, as is suggested by time evolution of the RMSD of the lid region for these variants (Fig. 4.9A & B, right panels).

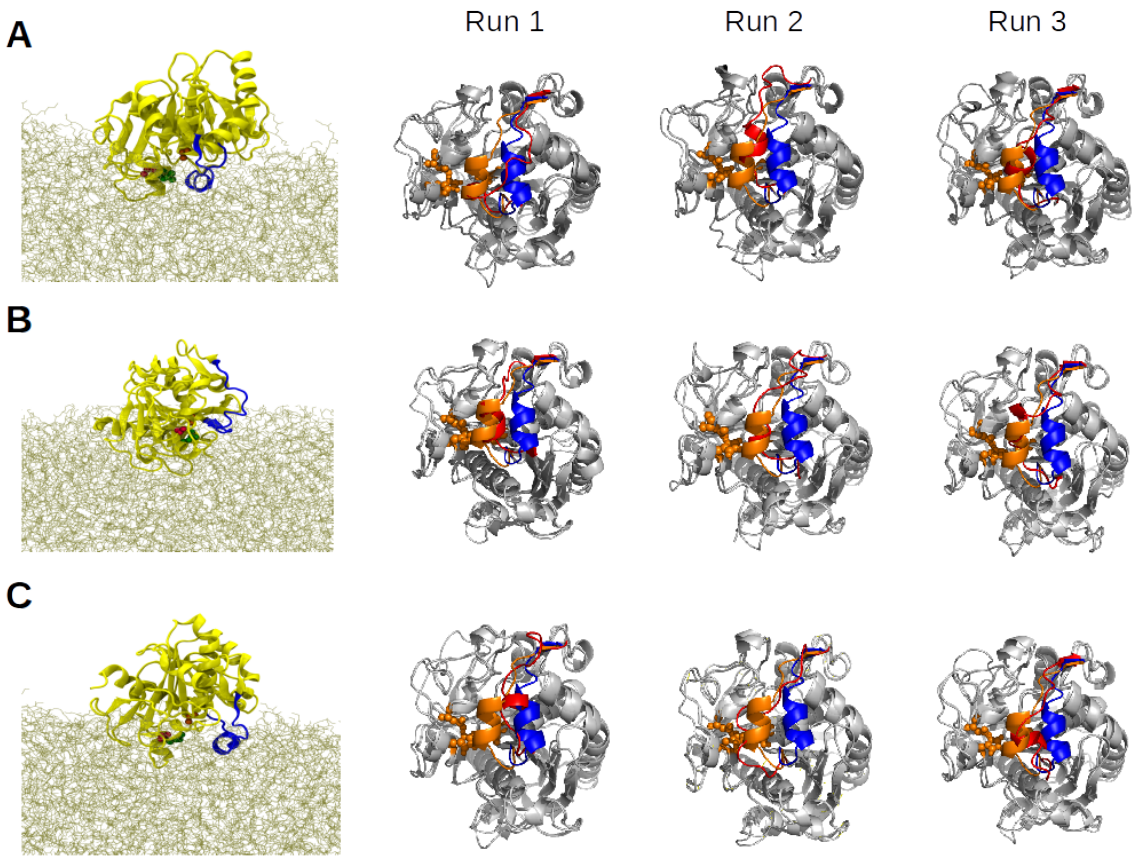

Fig. 4.8: *Left*: Bound orientations of initially closed wild-type **(A)**, 1L **(B)** and 3L lipase **(C)** variants extracted from the final frame of AT simulations with a tributyrin interface. The lipase is shown as a yellow cartoon; the lid is coloured blue. The catalytic residues (S172, H258, D201) are shown as van der Waals spheres coloured in red, green, and orange respectively. *Right*: Top down view of the simulated lipase structures (100 ns; lid coloured red) aligned with the closed (lid coloured orange) and open (lid coloured blue) crystal structures of wild-type TLL.

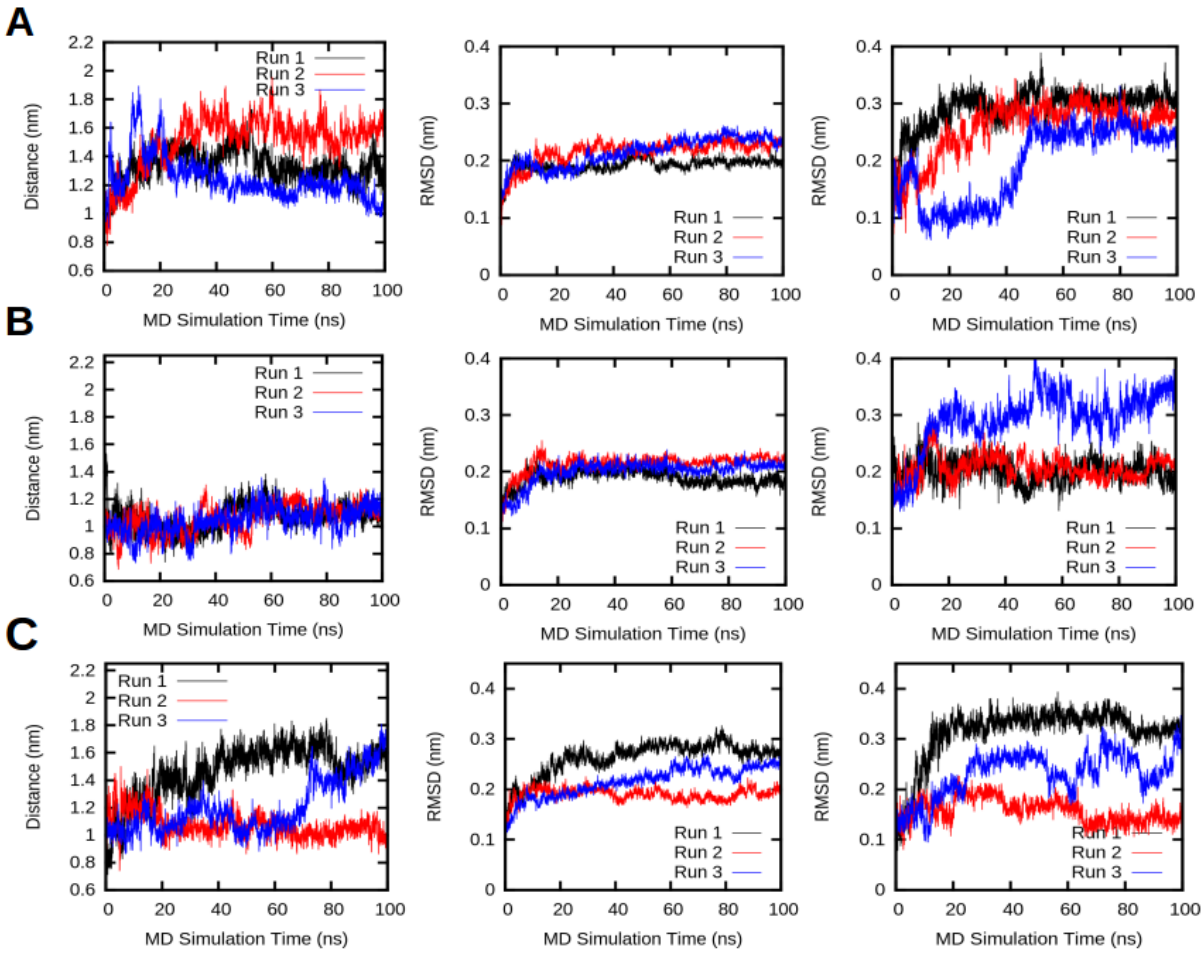

Fig. 4.9: Structural analysis of AT-MD simulations of initially closed the wild-type **(A)**, 1L **(B)**, and 3L variants **(C)** with a tributyrin layer. *Left*: Time evolution of the distance between the Cα atoms of residue 87 (within lid region) and residue 255 (neighbouring region) measuring the closed-open lid transition. *Middle*: Time evolution of the RMSD of protein Cα atoms. *Right*: Time evolution of the RMSD of Cα atoms within the lid region of the protein.

The observation that the 1L variant did not exhibit any interfacial activation upon association with the triglyceride interface correlates with the reduced activity observed for this mutant experimentally [206]. To investigate lid flexibility as a possible factor underlying reduced interfacial activity, in addition to altered binding orientations, three replicate AT simulations were also performed for initially open models of the lipase variants positioned above the tributyrin layer. These models were generated

based on the open crystal structure of the wild-type lipase [56]. The final interfacial binding orientations of the lipase variants are shown in Fig. 4.10.

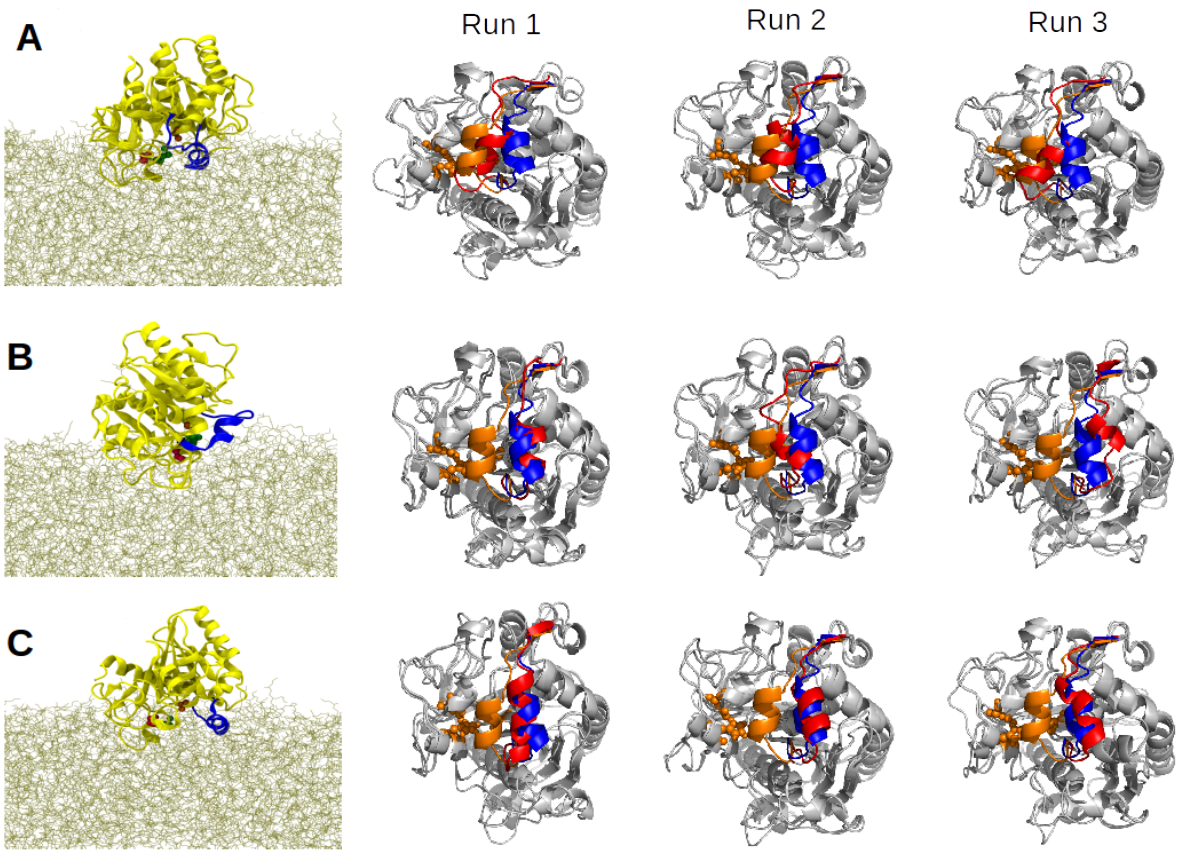

Fig. 4.10: *Left*: Bound orientations of initially open wild-type **(A)**, 1L **(B)** and 3L lipase **(C)** variants extracted from the final frame of simulations with a tributyrin interface. The same colour scheme as in Fig. 4.8 is used. *Right*: Top down view of same simulated lipase structures (100 ns; lid coloured red) aligned with the closed (lid coloured orange) and open (lid coloured blue) crystal structures of wild-type TLL.

Interestingly, the lid region of 1L did not return to the closed position in any of the simulations, correlating with previous observations of reduced lid flexibility, although some structural deviation could be observed, particularly for the third repeat simulation (Fig 4.10). In this simulation, the α-helical structure of the lid region is not completely conserved, as seen in time evolution of the RMSD, resulting in

reduced alignment with the open position of the lid region in the wild-type TLL crystal structure (Fig. 4.11).

Conversely, the lid region of the wild-type lipase exhibited significant closure in each repeat simulation, most likely due to initial solvent exposure of the hydrophobic residues within this region prior to lipase binding, which is thought to drive lid closure (Fig. 4.11A) [97]. Interestingly, this closing motion was not exhibited by simulations of the 3L variant, where the initially open lid conformation was apparently stable even prior to lipase association with the interface (Fig. 4.10C). This stability is also reflected in a constant value of the distance metric used to measure the lid transition, indicating that the 3L lid region maintained its open position at ∼2.1 nm for the duration of the simulations (distance measured as I255-R87 Cα distance) (Fig 4.11C). The altered structural dynamics of the 3L lid region compared to the wild-type variant are therefore a likely determinant of the enhanced activity observed against water-soluble substrate [206].

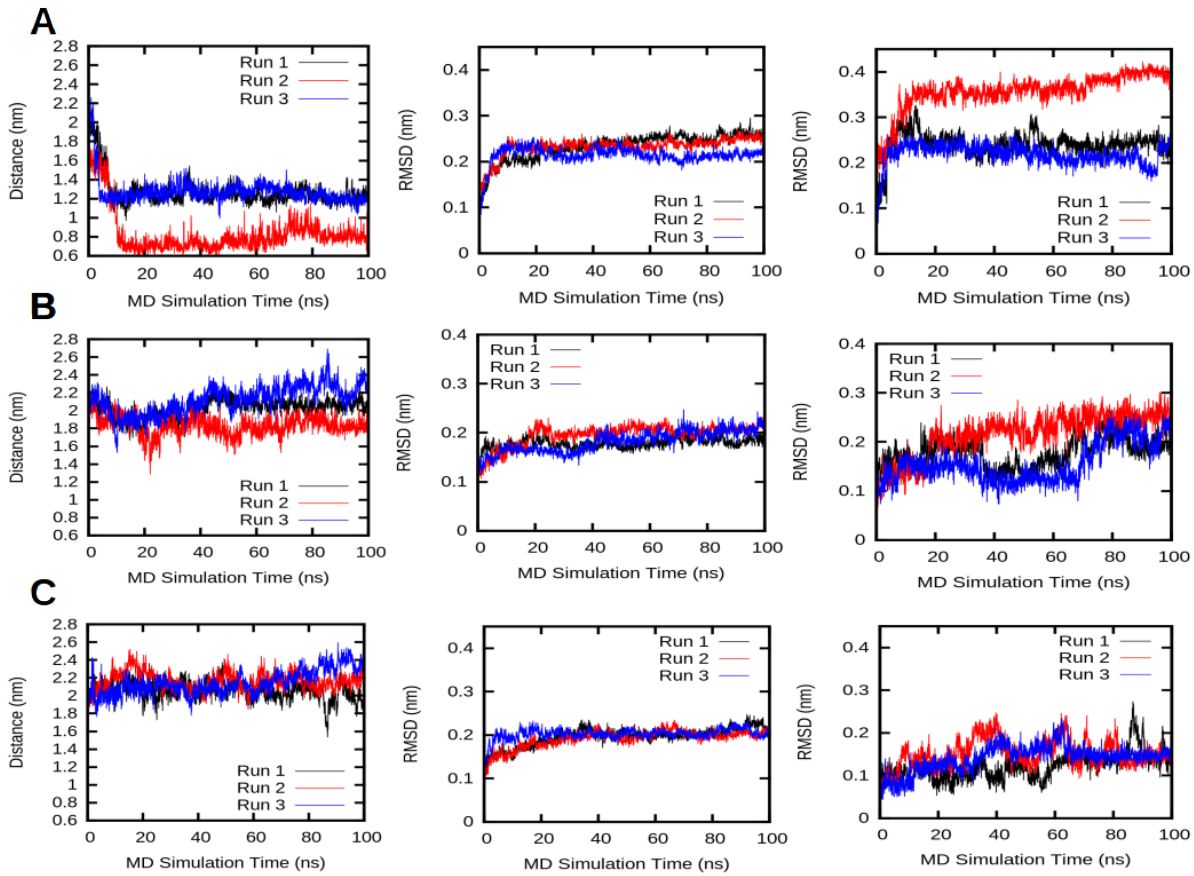

Fig. 4.11: Structural analysis of AT-MD simulations of initially open wild-type **(A)**, 1L **(B)**, and 3L variants **(C)** with a tributyrin layer. *Left*: Time evolution of the distance between the Cα atoms of residue 87 (within lid region) and residue 255 (neighbouring region) measuring the closed-open transition. *Middle*: Time evolution of the RMSD of the Cα atoms of the lipase variants. *Right*: Time evolution of RMSD of Cα atoms within the lid region of the protein.

4.3.3 Conformational Dynamics of Lipase Variants in Water

In order to compare the structural motions observed for the lipase variants at the tributyrin interface, additional AT simulations were performed in water. Similarly, three repeat simulations were performed for initially open and closed models of the variants. The simulations showed good agreement with earlier observations of the different structural dynamics exhibited by the lid regions of the lipase variants (Fig.

4.12). Specifically, simulation of the initially closed form of the 1L variant did not result in significant lid displacement in any of the repeat simulations, although slightly increased lid flexibilty was observed in the aqueous environment (Fig. 4.12B). This coincides with measurement of low activity on water-soluble substrate [206]. Similarly, the wild-type lipase also maintained a closed conformation in this high dielectric environment, consistent with observations that the lipase is inactive in the absence of an interface (Fig. 4.12A) [206]. In contrast, the lid region of 3L was more mobile, even exhibiting partial activation during one of the replicate simulations (Fig. 4.12C).

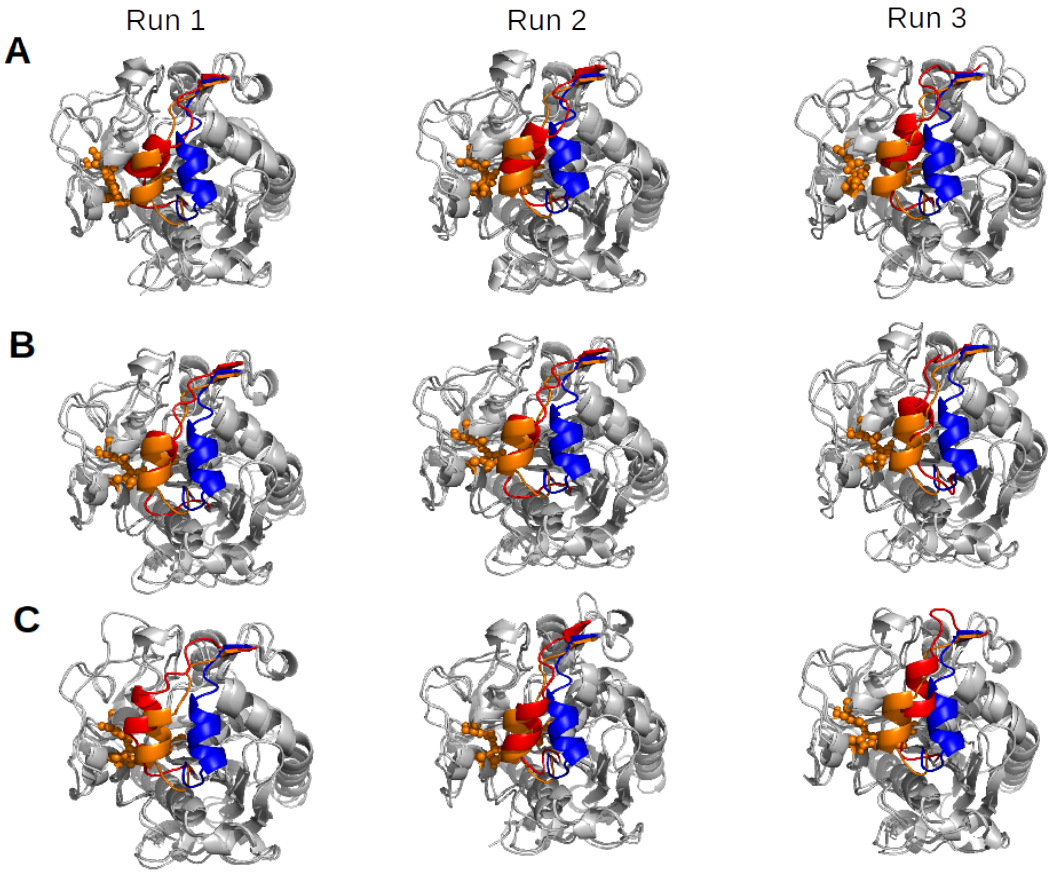

Fig. 4.12: Top down view of initially closed lipase structures extracted from the final frame of AT-MD simulations of the TL lipase variants in water. The simulated structures of the wild-type **(A)**, 1L **(B)** and 3L **(C)** lipase variants (100 ns; lid coloured red) are aligned with the closed (lid coloured orange) and open (lid coloured blue) crystal structures of wild-type TLL.

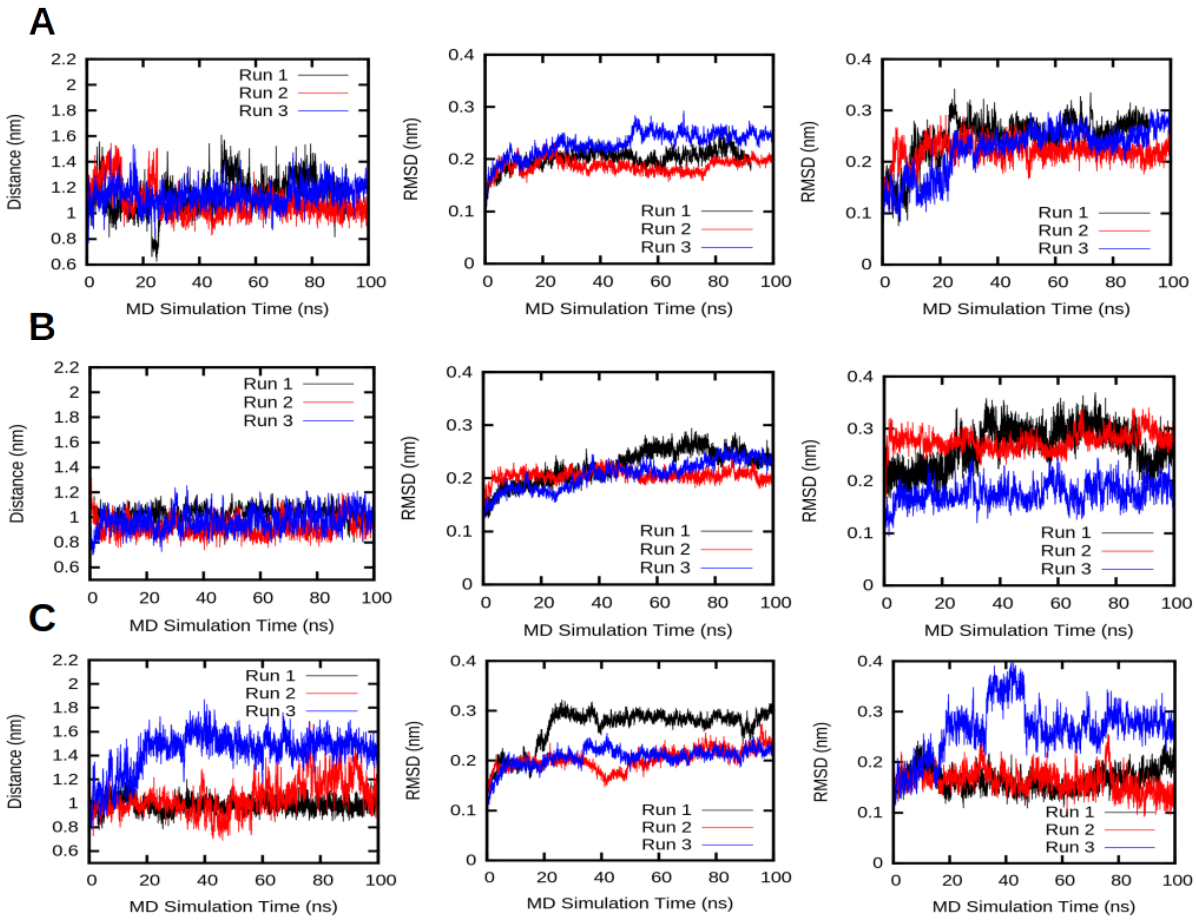

Fig. 4.13: Structural analysis for AT-MD simulations of initially closed wild-type **(A)**, 1L **(B)**, and 3L variants **(C)** in water. *Left*: Time evolution of the distance between the Cα atoms of residue 89 (within lid region) and residue 255 (neighbouring region) measuring the closed-open transition. *Middle*: Time evolution of the RMSD of protein Cα atoms. *Right*: Time evolution of the Cα atoms within the lid region of the protein.

However, more variation in lid displacement was observed when starting from an initially open conformation, particularly for the wild-type and 1L variants (Fig. 4.14). Here, the 1L protein remains fully open for 2/3 repeat simulations, whilst the wild-type lipase exhibited either partial or full closure of the lid region in all repeat simulations (Fig. 4.14A and B). It is noted that the secondary structure of the lid helix was disrupted during simulation of these variants, as seen from the increasing RMSD of lid region relative to its initial α-helical structure (Fig. 4.15A & B, right panel)

for both the wild-type and 1L variants. This could affect lid flexibility and result in greater alignment with the closed position of the lid for these variants. Conversely, the lid region of the 3L variant tended to maintain its α-helical structure, and remained relatively open for 2/3 replicate simulations, coinciding with previous observations of superior stability of the open conformation of this lipase in solution (Fig. 4.14C) [206].

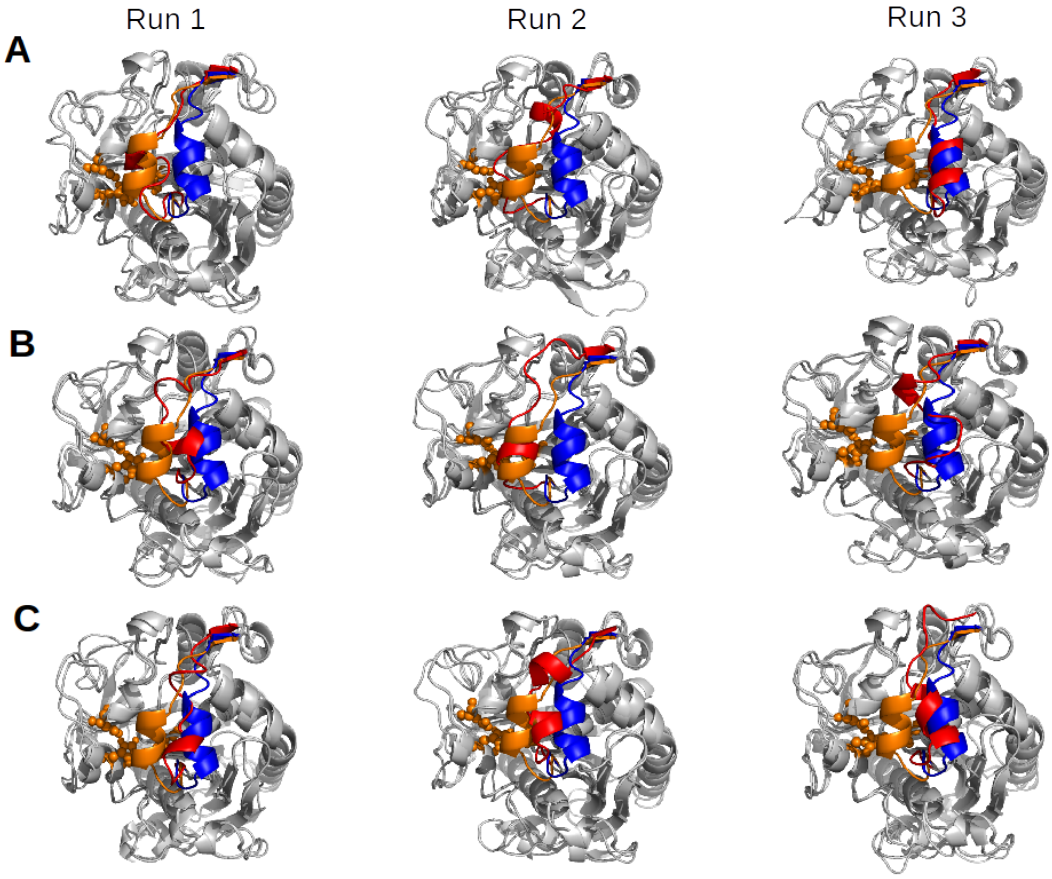

Fig. 4.14: Top down view of initially open structures of lipase variants after AT simulation in water. The simulated structures of wild-type **(A)**, 1L **(B)** and 3L **(C)** lipase variants (100 ns; lid coloured red) are aligned with the closed (lid coloured orange) and open (lid coloured blue) crystal structures of wild-type TLL.

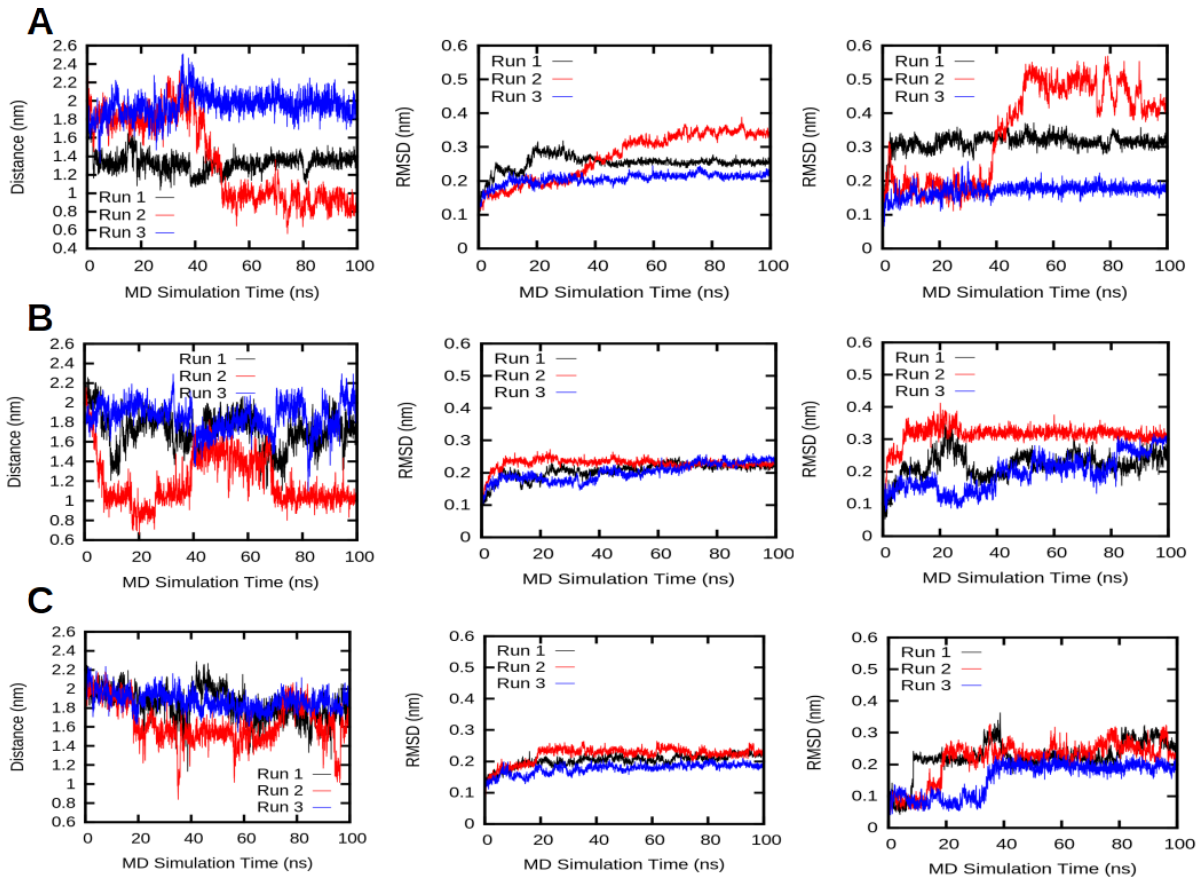

Fig. 4.15: Structural analysis of AT-MD simulations of initially open wild-type **(A)**, 1L **(B)**, and 3L variants **(C)** in water. *Left*: Time evolution of the distance between the Cα atoms of residue 87 (within lid region) and residue 255 (behind lid region) measuring the closed-open transition. *Middle*: Time evolution of the RMSD of Cα atoms within the lipase variants. *Right*: Time evolution of the Cα atoms within the lid region of the lipase variants.

The observed differences in lid dynamics of the lipase variants both at the interface and in solution highlight the importance of residue composition on overall structural stability and lid mobility. Of particular interest is the observation that the 3L variant is able to maintain an open conformation, even when the lid region is solvent exposed. Contrastingly, the wild-type lid region exhibited a closing motion in solution. The distinct distribution of hydrophobic residues within the α-helix of the 3L lid domain is similar to that of wild-type TLL, suggesting that the hinge domains of the lid region

play a key role in lid mobility [206]. Experimentally, the differences in activity between the variants have been attributed to the underlying energetic barrier to lid opening, which appears to be sufficiently low in 3L to allow catalysis of water-soluble substrate [206]. This was explored further using steered MD simulations.

4.3.4 Steered MD Simulations of Lipase Variants in Water

Steered MD simulations (SMD) were employed to provide initial estimates of the energetic barrier associated with the lid opening process for each of the variants. To capture the opening motion of the lid, a distance CV was defined between the Cα atom of residue 87 in the lid region (E87 in wild-type) and the Cα atom of a distally located residue (I255) (Fig. 4.16). These residues were previously identified as a good metric for lid activation in fluorescence studies of TLL in different solvents [57]. The upper and lower limits for the distance travelled by the lid region during opening were estimated from the open and closed crystal structures of wild-type TLL.

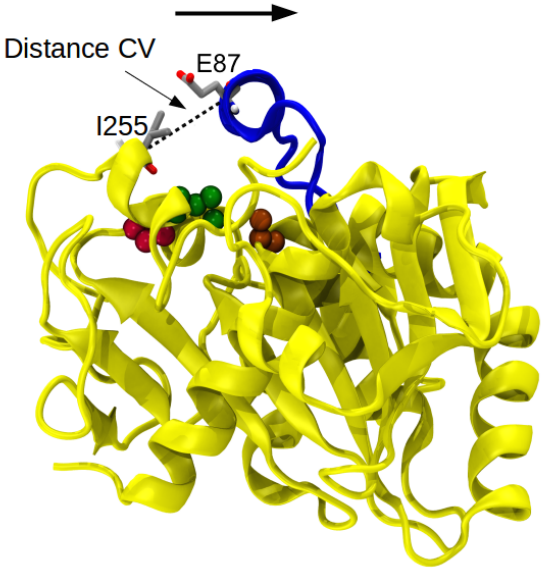

Fig. 4.16: Definition of a distance CV for SMD simulations of lipase variants in water, investigating the lid opening motions of TLL. The CV measures the COM distance between the Cα atoms of I255 and E87 (wild-type sequence). The lipase is shown as a yellow cartoon; the lid helix is shown in blue. The catalytic triad is shown as van der Waals spheres: S146 = red, D201 = orange, H258 = green.

15 ns SMD simulations were performed for each of the lipase variants in water (Fig. 4.17). The resulting work profiles indicated that the largest amount of work was performed for simulation of the 1L variant, whilst the smallest work was performed for the 3L variant. This is in line with previous observations of reduced lid flexibility in the 1L lipase, where structural analysis shows the 1L lid region was barely displaced during the SMD simulation (Fig. 4.17). Evolution of the distance CV for 1L is therefore due to the biasing force actually displacing a neighbouring loop region close to the 87th residue within the lid, rather than displacement of the lid region itself. This suggests that the intermolecular forces mediating lid position are stronger than the biasing force applied during the SMD simulation. In contrast, the lid region of the 3L variant was most easily displaced in comparison with the wild-type and 1L variants, coinciding with previous observation of increased lid mobility for 3L. Furthermore, the

stabilisation of the open form of the 3L variant also coincides with the experimental observations of superior esterase activity exhibited by 3L compared to the wild-type variant [206].

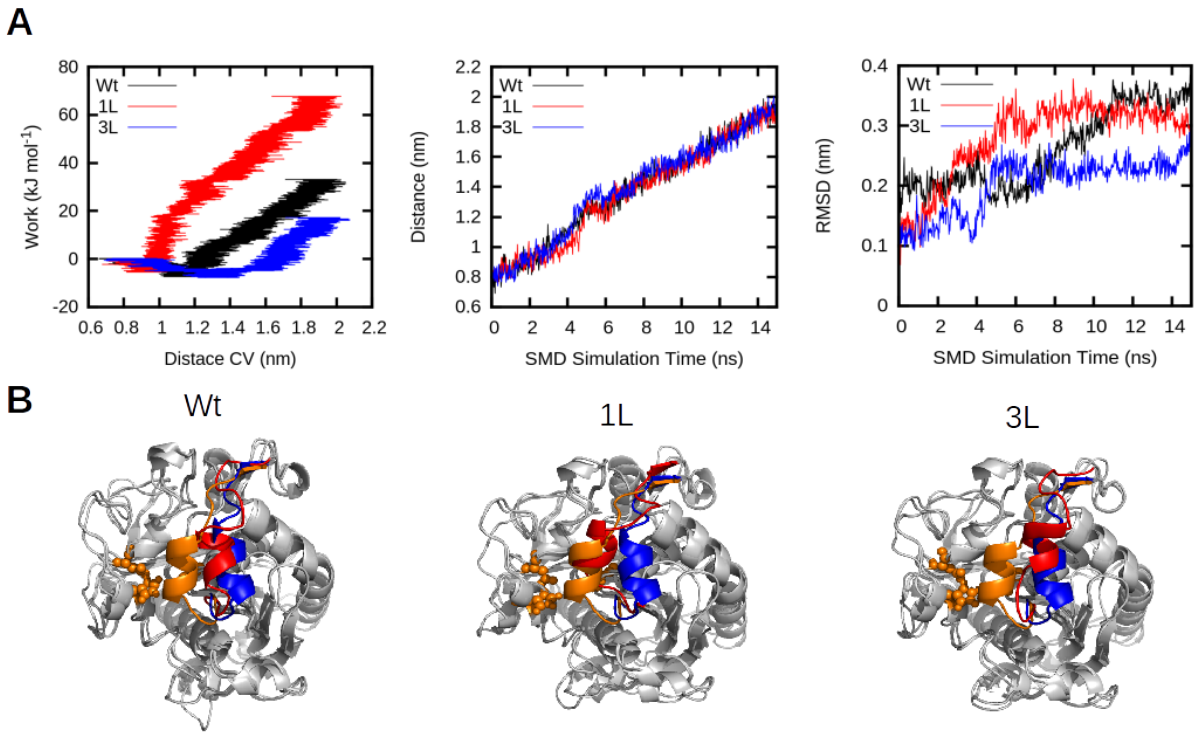

Fig. 4.17: SMD simulations were performed for wild-type (Wt), 1L, and 3L lipase variants in water using a distance CV to investigate the lid opening motion (see Fig. 4.16). **(A)** The work profiles (*left*), time evolution of the CV (*middle*), and time evolution of the RMSD of the Cα atoms within the lid region (*right*) were calculated for the SMD simulations. **(B)** The final simulated structures of the proteins at 15 ns (lid coloured red) aligned with the closed (lid coloured orange) and open (lid coloured blue) crystal structures of the wild-type lipase.

Analysis of the contacts between the lid region and the neighbouring protein residues indicate that hydrophobic interactions between residues such as I90-I255 and L93-I202 were broken as the lid was pushed away during the SMD simulations, exposing these hydrophobic residues to the solvent for the wild-type and 3L variants (Fig. 4.18).

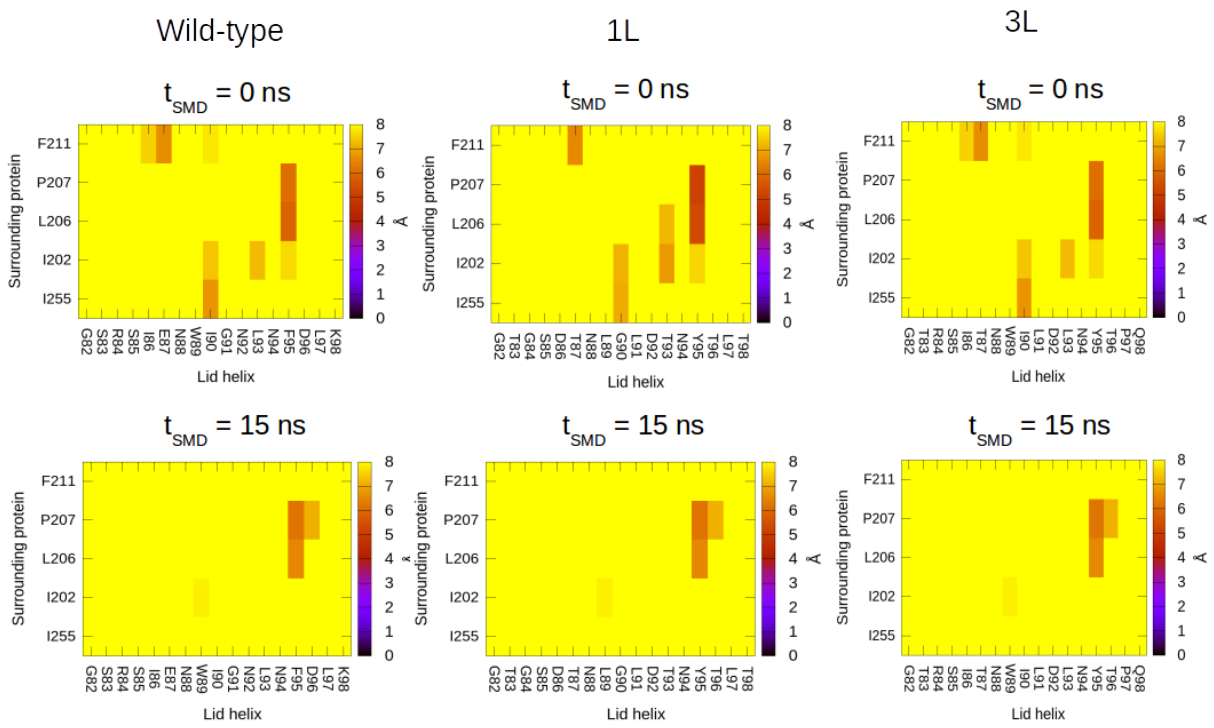

Fig. 4.18: Contacts calculated between the lid region of TLL variants and surrounding protein residues for SMD simulations of the variants in water. The matrices show contacts for the first (*top panel*) and final (*bottom panel*) frames of the SMD simulations. A 0.4 nm cut-off was used.

The altered positions of the hydrophobic residues within the lid region of the simulated structures are shown in Fig 4.19, indicating they become more solvent exposed. Bulky residues such as W89 however do not shift very much relative to their original position in the closed crystal structure for the wild-type and 3L lipases. This is most likely related to the aqueous environment surrounding the lipase molecules, where unfavourable interactions of hydrophobic lid residues with the solvent could affect their final positions relative to their positions in the open crystal structure of TLL, which was solved in the presence of a substrate interface [56].

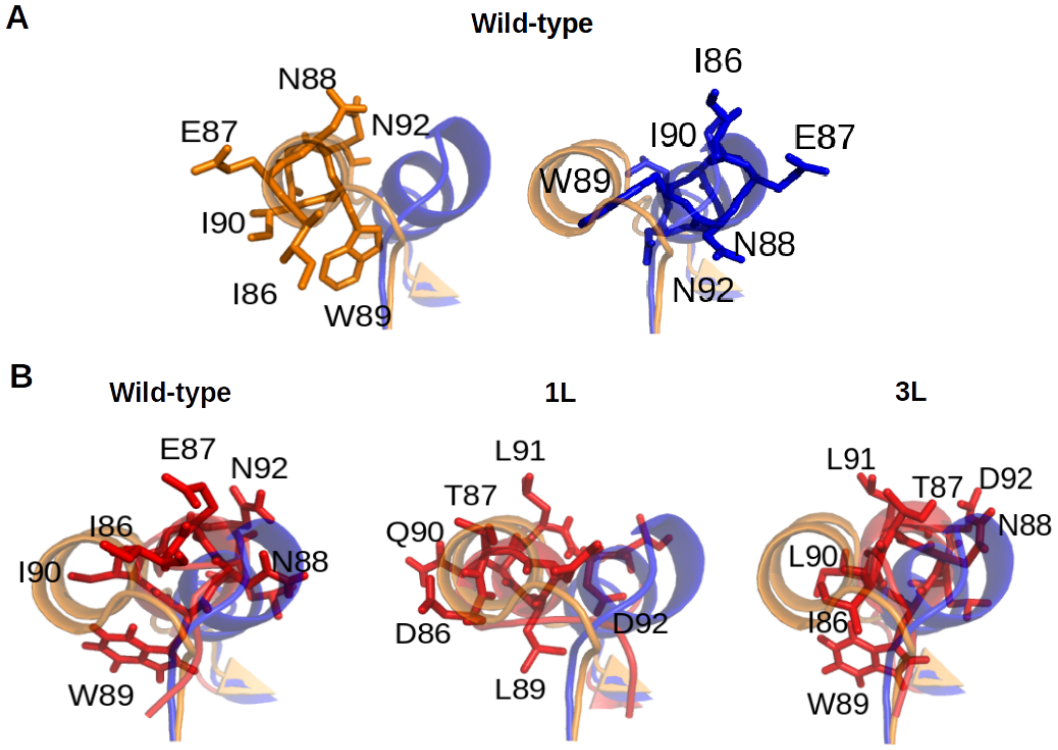

Fig. 4.19: **(A)** Position of amino acid side chains within the lid region of the closed (orange) and open (blue) crystal structures of wild-type TLL [56]. Only the lid region (cartoon representation) is shown for the aligned structures. The residues are shown as sticks. **(B)** Position of amino acid side chains within the lid regions of lipase variants after SMD simulations (15 ns) probing lid displacement in water. The lid region and residues are coloured in red. These are aligned with the closed (orange) and open (blue) positions of the lid region in the crystal structures of TLL.

The resultant work profiles therefore support previous observations from unbiased AT-MD simulations, suggesting constraints on lid dynamics for the 1L variant, compared to high lid mobility for the 3L variant.

4.3.5 Steered MD Simulations of Lipase Variants at a Triglyceride Interface

In addition to simulating the lid opening process of the lipase variants in water, it is important to characterise the motions of the lid region when the lipase is in contact with the triglyceride interface. Additional SMD simulations were performed to investigate these motions for the 1L variant. These were initiated from the final structure of the bound lipase molecule from unbiased AT simulations with the natural substrate interface (Fig 4.20). The same SMD protocol was performed as detailed in the previous section. SMD simulations of the bound forms of the wild-type and 3L variants were not performed given that these variants displayed degrees of interfacial activation during unbiased simulations with the triglyceride interface. The SMD simulations would therefore not sample the full lid opening process, preventing comparisons with SMD simulations performed in an aqueous environment, in which the lid region occupied the initially closed state for all the variants.

SMD simulations of the 1L variant at the interface resulted in an increase in the amount of work performed to sample lid displacement compared to SMD simulations in water. However, the lid region was still not significantly displaced at the interface. Similar to the SMD simulations in water, the biasing force resulted in displacement of the neighbouring loops that contain the I255 residue, resulting in evolution of the I255-E87T distance CV (Fig. 4.20).

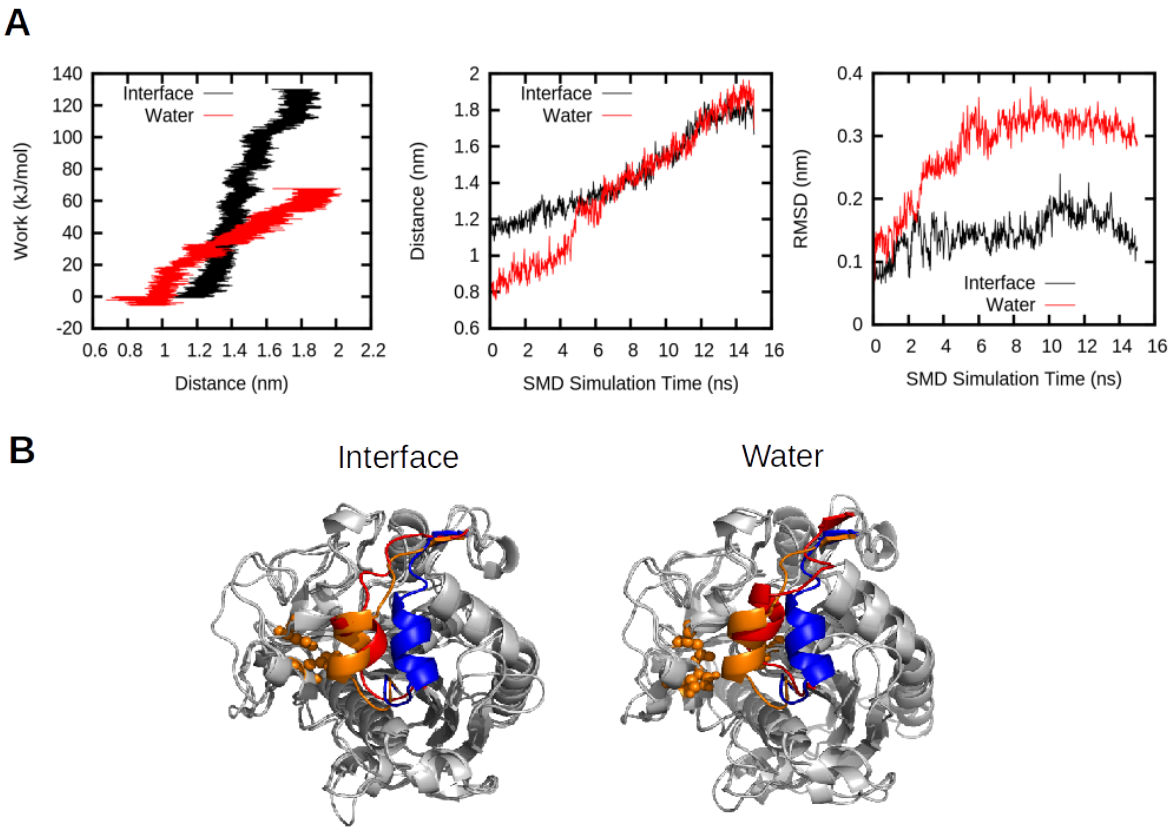

Fig. 4.20: SMD simulations were performed for the 1L lipase variant in water and at a tributyrin interface in order to investigate lid opening motions during interfacial activation. **(A)** The calculated work profiles (*left*), time evolution of the lid distance CV (*middle*), and time evolution of the RMSD of the Cα atoms within the lid region (*right*). **(B)** The 1L lipase structures at the end of the SMD simulations are shown for both the water and interface simulations (lid coloured red), aligned with the closed (lid coloured orange) and open (lid coloured blue) crystal structures of the wild-type lipase.

In order to overcome the issue of sampling the lid opening motion within the 1L variant, additional SMD simulations were performed in which position restraints were applied to all the atoms in the lipase except for atoms within the lid region. This protocol enabled greater extension of the distance CV and subsequent displacement of lid region in the 1L variant in water and at the interface. However, position restrained simulation of 1L at the triglyceride interface resulted in a similar work profile compared to the *unrestrained* SMD simulation of 1L at the interface, whilst much more work

was performed for position restrained SMD simulations of 1L in water ($\sim$ 250 kJ mol^{-1} for one replicate).

These results indicate that the SMD simulations did not converge to provide a complete description of the lid opening process for the different lipase variants. Consequently, comparisons of lid activation between the 1L variant and the wild-type and 3L lipases were hindered regarding the energetic barrier for lid opening as a function of the lid residue composition. More rigorous enhanced sampling methods are necessary to provide a good estimate for the activation energy required to displace the lid regions of these mutant lipases. Additionally, a comparison between the energy profiles of the wild-type and 3L variants in water *and* at the interface would be necessary to determine how lid mutations affect lid activation when the lipase is interfacially bound.

4.3.6 Lipase Hydrolysis of Triglyceride Systems

The mechanisms of lipase inhibition have been reported in various experimental contexts, and include effects from interfacial tension of the surface to which the lipase is bound [199], interactions that affect substrate access and the overall lipase conformation [205, 210], as well as interactions of charged product molecules with the lipase binding pocket [170, 208]. Additionally, the accumulation of water molecules, as well as fatty acid products, at the interface alters the surrounding dielectric constant detected by the associated lipase. As shown by previous spectroscopic studies, the dielectric constant of the environment is an important determinant of lipase activation and lid mobility [57].

To investigate how product accumulation at the interface might affect lipase association, substrate binding, and lipase conformational dynamics, I set up systems that mimic tributyrin hydrolysis. Three different systems were constructed to model three

different hydrolysis percentage: 10%, 25%, and 50%. Here, I considered one lipase-catalysed hydrolysis cycle as the conversion of one tributyrin molecule into three fatty acid components. Configuration of each system was initiated from final snapshots of the lipase variants bound to the tributyrin interface from the previous unbiased AT-MD studies. The bound positions of the lipase variants were used a reference point to identify all tributyrin molecules within 1 nm radius of the protein. Out of these identified molecules, either 10%, 25%, or 50% of the tributyrin molecules were each converted to three fatty acid molecules (Fig. 4.21). The fatty acid molecules possessed a negative charge group on the ester moiety (O^-). This was to investigate the effect of surface charge on lipase binding and structure. 50 ns AT-MD simulations of these hydrolysed systems were performed for each of the lipase variants, initiated from their bound configurations from unbiased AT-MD studies.

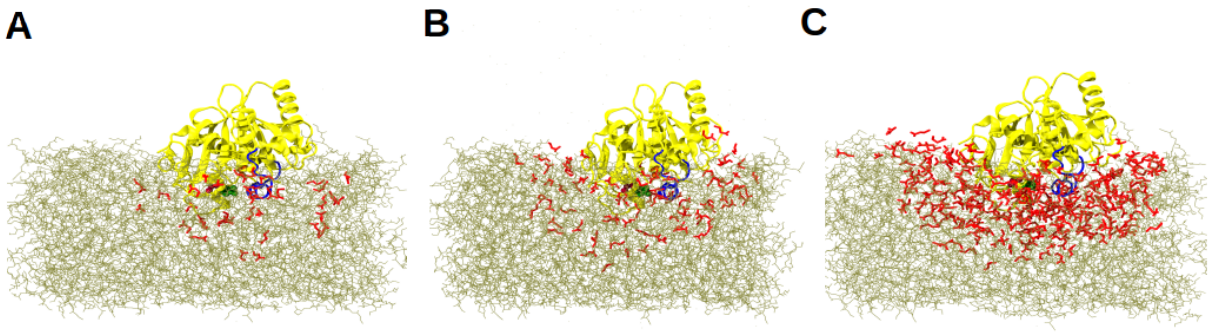

Fig. 4.21: Configurations of hydrolysed systems reflecting 10% **(A)**, 25% **(B)**, and 50% **(C)** hydrolysis. Hydrolysis was modelled by identifying all the tributyrin molecules (brown sticks) within 1 nm of the bound lipase (yellow and blue cartoon representation; blue = lid region), and converting a percentage of these into 3 fatty acid molecules each (red sticks). The catalytic triad (S146, H258, and D201) is shown as red, green, and orange van der Waals spheres, respectively. The images represent the last frame of an AT simulation of the wild-type TLL variant at 100 ns.

None of the simulations resulted in dissociation of the lipase variants from the interface within the simulated time. The final binding orientations of the wild-type

lipase for each system (10%, 25%, or 50%) are shown in Fig. 4.22 (top panel), and indicated that the fatty acid moieties initially occupying the surface of the layer quickly dissociated over time. This is due to the high solubility of the fatty acid functional groups in water, which includes a negatively charged oxygen atom, overcoming weak hydrophobic interactions between the short carbon tails (4 carbon atoms) of the surrounding tributyrin molecules. Consequently, there were only small structural effects of the hydrolysis products on the conformation of the bound lipase variants for the layers containing 10% or 25% fatty acids. However, the 50% hydrolysed system seemed to have a greater affect on overall protein structure, as can be seen from increasing Cα RMSD values over time, as well as the larger radius of gyration (Fig. 4.23).

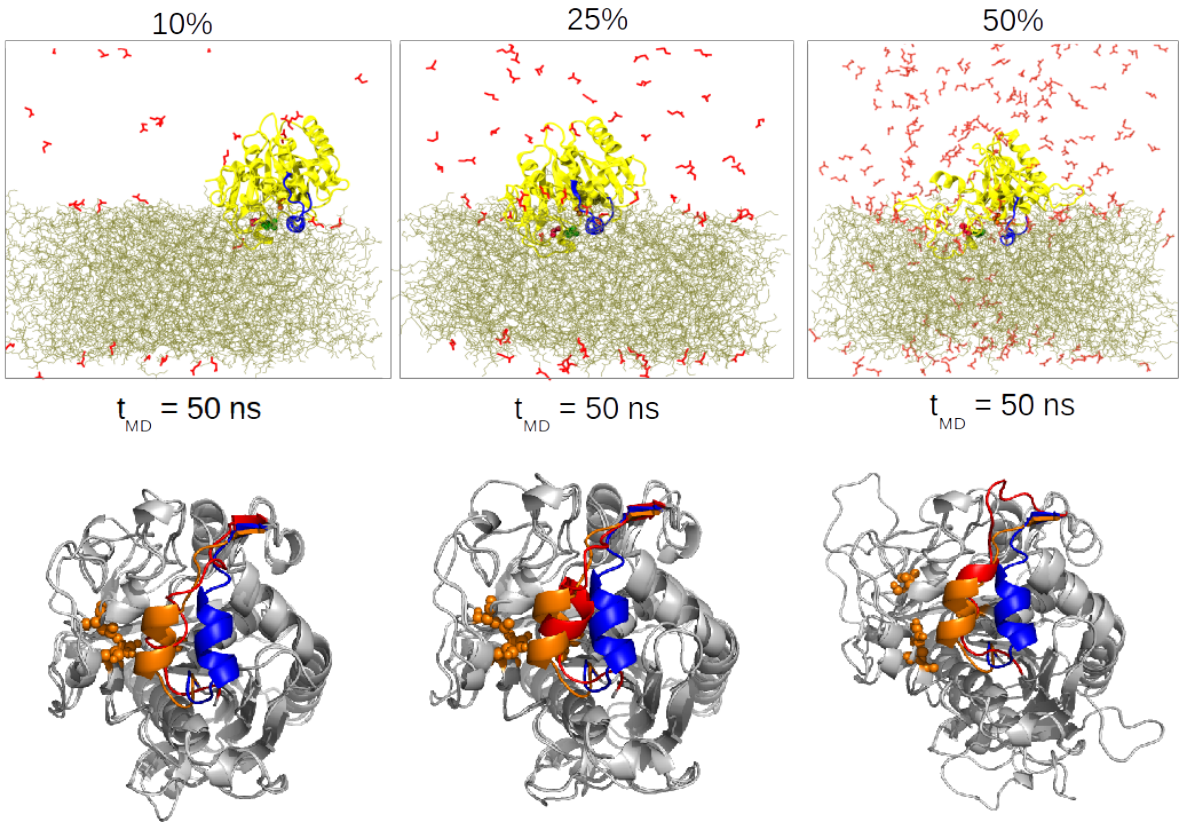

Fig. 4.22: *Top panel*: Final binding orientations at the end of AT-MD simulations of the wild-type lipase bound to a tributyrin layer containing either 10%, 25%, or 50% fatty acid molecules. *Bottom panel*: The final lipase structure at the end of the simulations in **(A)** (50 ns; lid coloured red) aligned with the closed (lid coloured orange) and open (lid coloured blue) crystal structures of wild-type TLL.

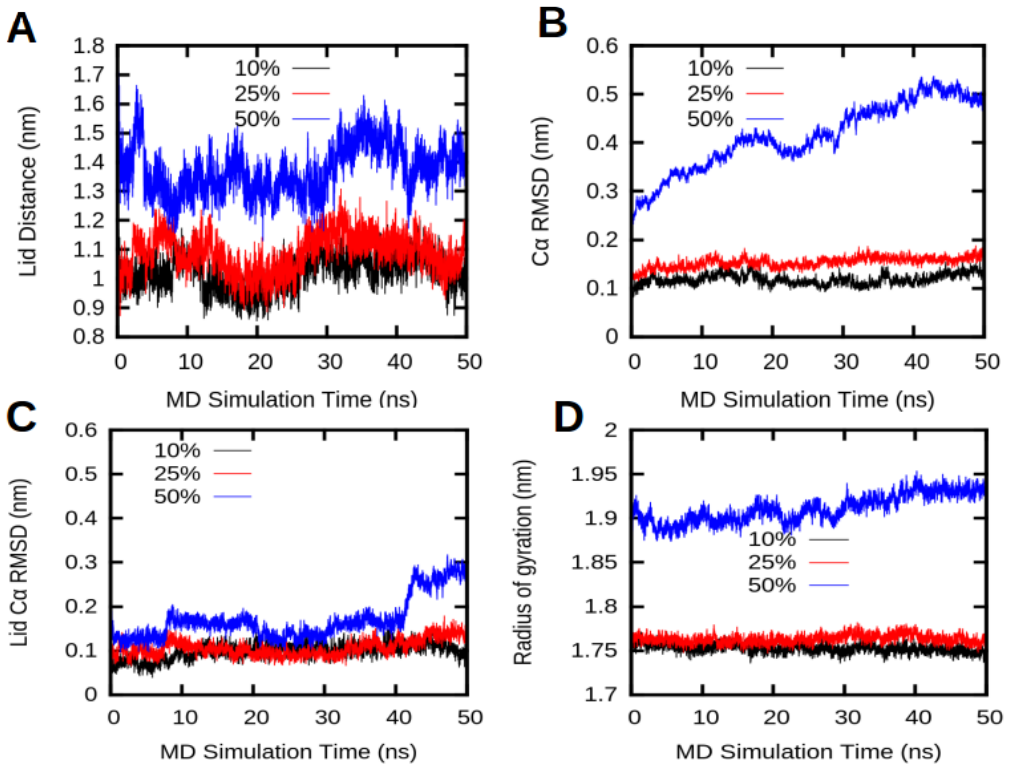

Fig. 4.23: Analysis of AT-MD simulations of hydrolysed tributyrin systems with wild-type TLL. **(A)** Time evolution of the E87-I255 distance measuring lid displacement, calculated from simulations of the lipase with tributyrin layers containing either 10%, 25%, or 50% fatty acid molecules. **(B)** Time evolution of the RMSD calculated for all Cα in the lipase structure. **(C)** Time evolution of the Cα atoms within the lid region. **(D)** Time evolution of the radius of gyration of the protein.

Similar structural effects were also observed for the 1L and 3L variants for the 50% hydrolysed systems (Fig. 4.24 and 4.26). This is most likely due to electrostatic interactions between the charged oxygen groups of the fatty acid molecules disrupting stabilising bonds within the protein structure. Additionally, the 50% hydrolysed systems also seemed to cause some displacement of the lid region for both the wild-type and 1L variants (Fig. 4.22, Fig. 4.24).

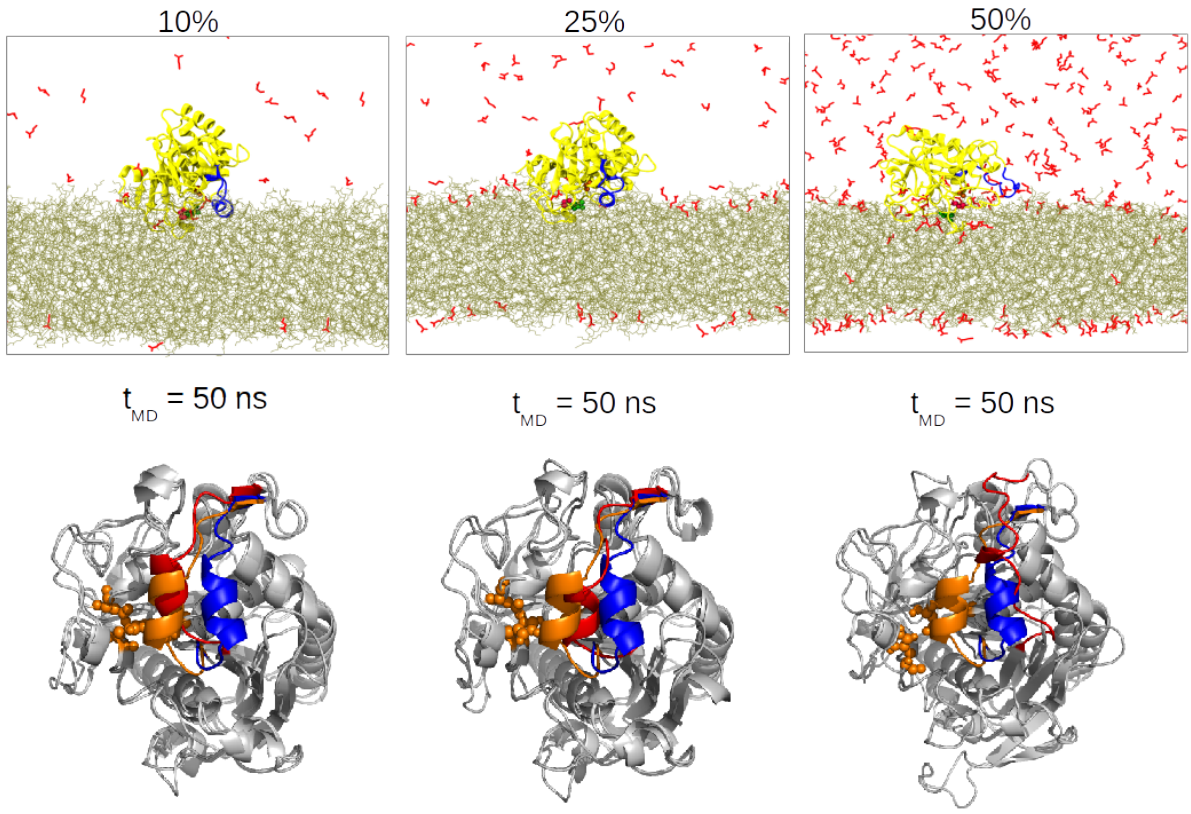

Fig. 4.24: *Top panel*: Final binding orientations at the end of AT-MD simulations of the 1L variant bound to a tributyrin layer containing either 10%, 25%, or 50% fatty acid molecules. *Bottom panel*: The final lipase structure (50 ns; lid coloured red) aligned with the closed (lid coloured orange) and open (lid coloured blue) crystal structures of wild-type TLL.

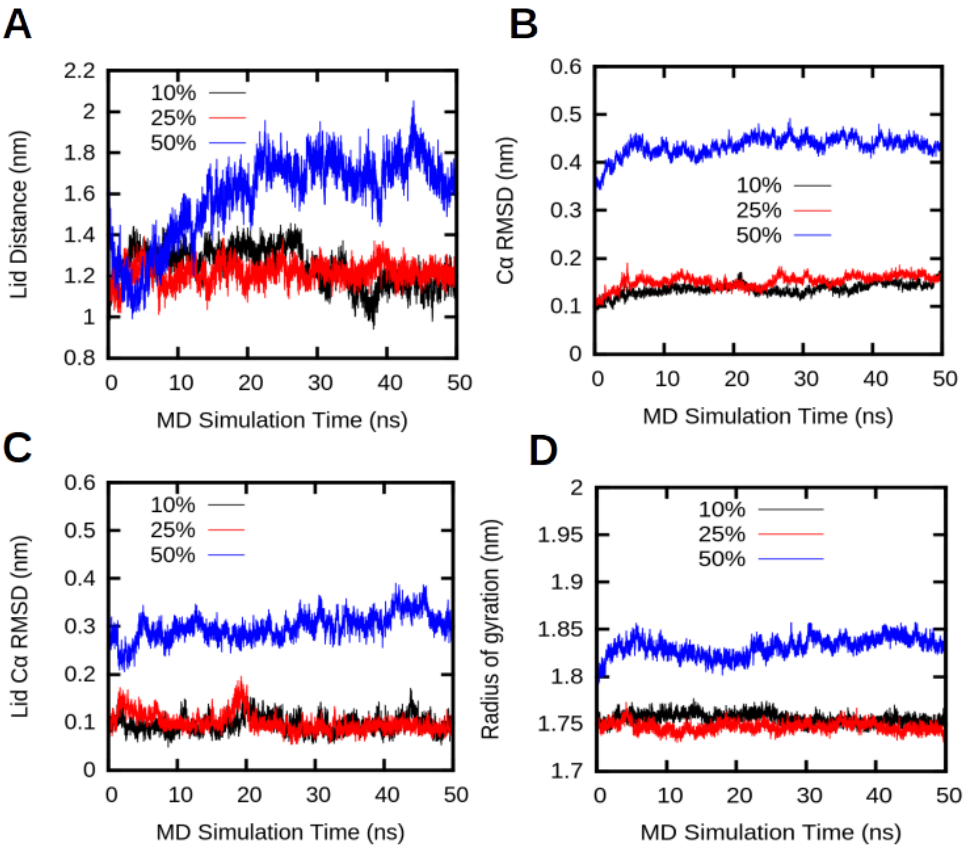

Fig. 4.25: Analysis of AT-MD simulations of hydrolysed tributyrin systems with the 1L variant. **(A)**: Time evolution of the T87-I255 distance, measuring lid displacement, calculated from simulations of the lipase with tributyrin layers containing either 10%, 25%, or 50% fatty acid molecules. **(B)** Time evolution of the RMSD of all Cα atoms in the lipase structure. **(C)** Time evolution of the Cα atoms within the lid region. **(D)** Time evolution of the radius of gyration of the bound lipase molecule.

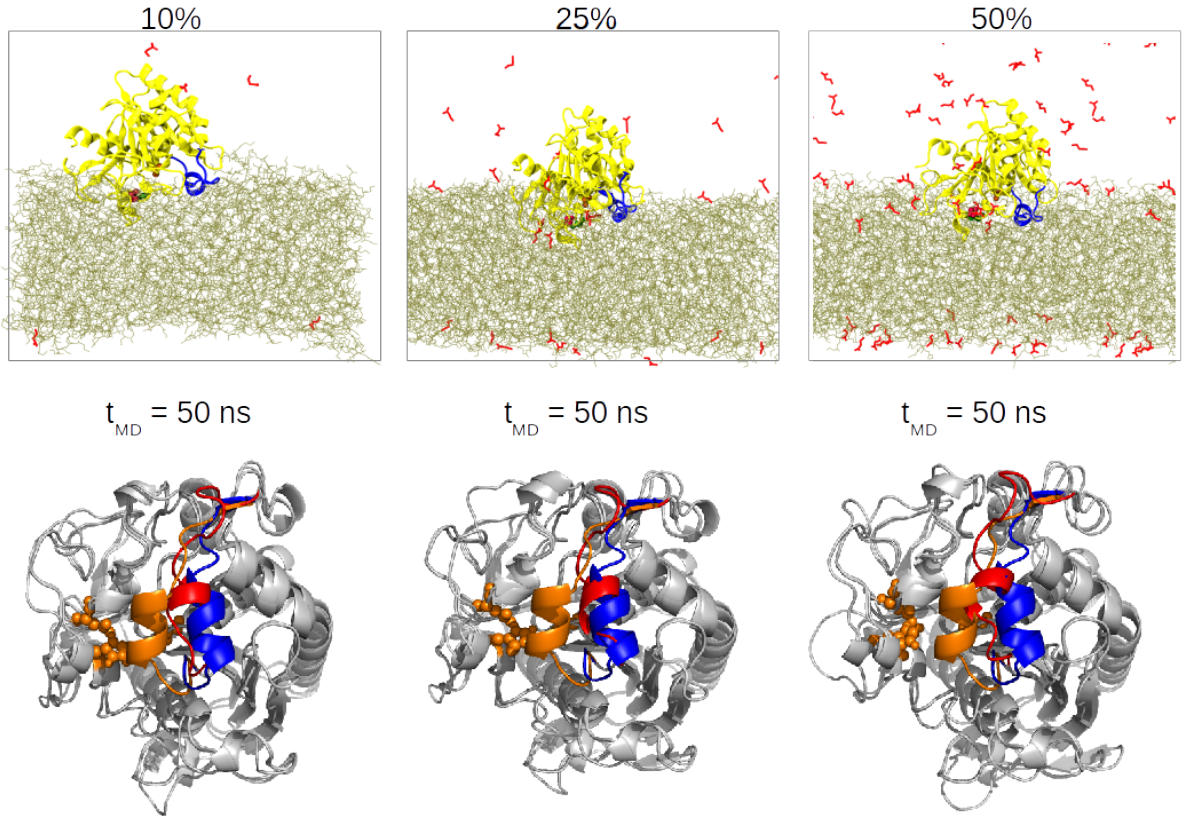

Fig. 4.26: *Top panel*: Final binding orientations at the end of AT-MD simulations of the 3L variant bound to a tributyrin layer containing either 10%, 25%, or 50% fatty acid molecules. *Bottom panel*: The final lipase structure (50 ns; lid coloured red) aligned with the closed (lid coloured orange) and open (lid coloured blue) crystal structures of wild-type TLL.

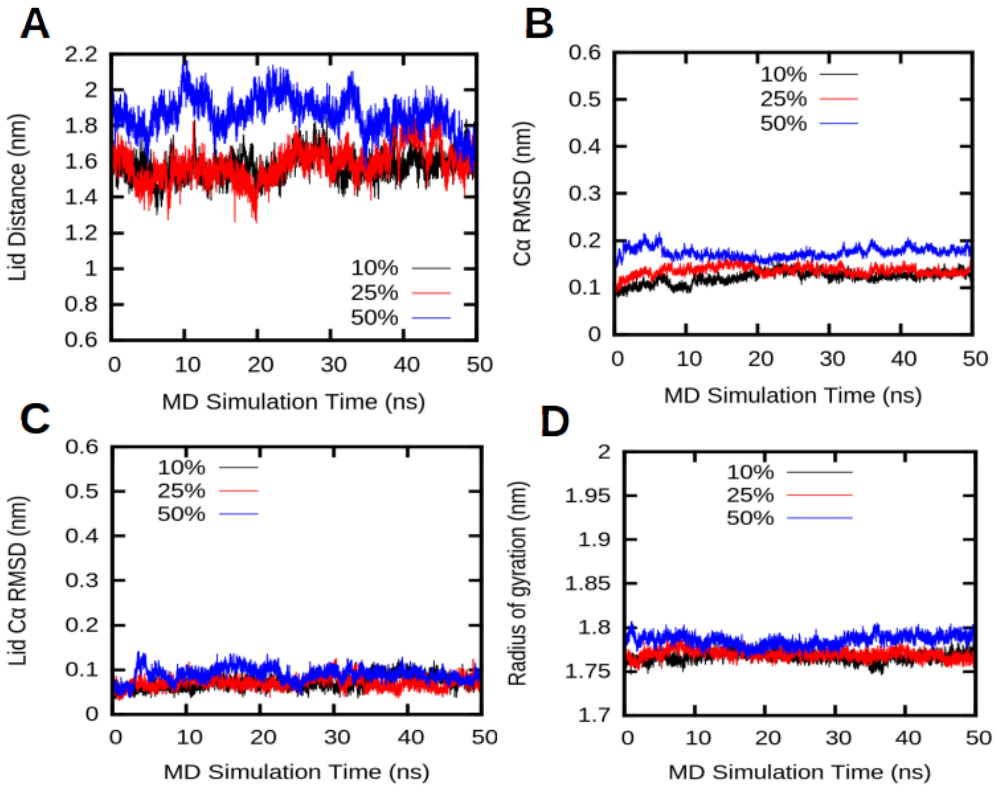

Fig. 4.27: Analysis of AT-MD simulations of hydrolysed tributyrin systems with the 3L variant. **(A)**: Time evolution of the T87-I255 distance measuring lid displacement, calculated from simulations of the lipase with tributyrin layers containing either 10%, 25%, or 50% fatty acid molecules. **(B)** Time evolution of the RMSD of all Cα atoms in the lipase structure. **(C)** Time evolution of the Cα atoms within the lid region. **(D)** Time evolution of the radius of gyration of the bound lipase.

Calculation of the normalised contacts between the fatty acid molecules and bound lipase variants indicate that these vary throughout the simulations. There is a steady decrease in the number of contacts between all lipase variants and the fatty acid molecules in the 50% hydrolysed systems, most likely due to rapid dissociation of the molecules from the interface and the subsequent random diffusion throughout the aqueous solution. The number of contacts seem to fluctuate more randomly for the 10% and 25% hydrolysed systems, presumably because the smaller number of free fatty acids tended to diffuse through the solution prior to interacting with the

associated lipase, and may encounter the surface of the lipase less frequently. Overall, the interactions between the free fatty acid molecules and the lipase seem to be random, given the continuous fluctuation in the number of contacts over time, indicating these interactions may not have a functional impact on the bound lipase molecules.

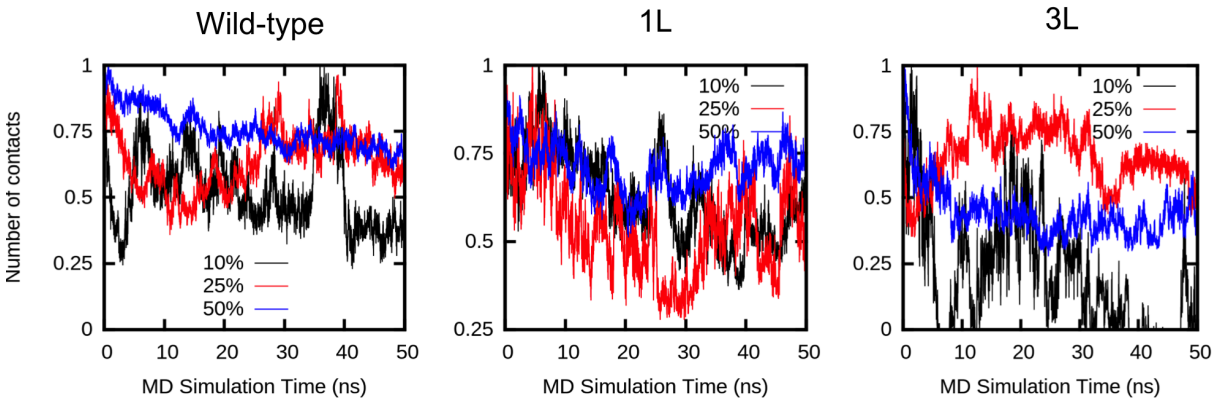

Fig. 4.28: Time evolution of the number of contacts between fatty acid molecules and the wild-type, 1L, and 3L lipase variants (0.4 nm cut-off) for AT-MD simulations of hydrolysed tributyrin systems (10%, 20%, and 50%.

4.4 Discussion

Existing experimental data suggest that amino acid substitutions within the lid region of the TL lipase affects interfacial activation and catalytic activity [206]. Other studies of different lipases have demonstrated similar effects, highlighting the critical role of the lid region in enzyme activity and substrate specificity [198, 200, 211]. Here, CG and AT simulations were used to investigate how mutation of the lid region of TLL affects interfacial interactions with a natural substrate, as well as conformational dynamics at the interface and in solution.

Both the CG and AT simulations of the lipase mutants with the triglyceride interface suggested a distribution of interfacial binding orientations. Specifically, the 1L protein, which exhibited esterase activity, consistently displayed an orientation in

which the lid region pointed away from the triglyceride interface, interacting with the surrounding solvent [206]. In contrast, the 3L variant, of lipase and esterase character, displayed an average orientation in which the lid region was buried within the interface, consistent with the open form of the wild-type lipase (Fig 4.4A) [206]. This buried interfacial orientation is in line with experimental studies of wild-type TLL at phospholipid interfaces, as well as other hydrophobic surfaces [95, 96]. Furthermore, contact analysis indicates that the nature of the amino acid sequence within the lid region most likely underlies the different orientations exhibited by the variants at the interface (Fig 4.5).

The different amino acid sequences within the mutated regions also appeared to affect the structural dynamics of the lid regions, particularly for the 1L variant. AT simulations of both an initially closed and open form of the variants revealed that lid displacement and flexibility was significantly diminished in the 1L variant. Similarly, experimental studies of other lipases with mutated lid regions were seen to affect the structural dynamics and thermostability of the mutated enzymes, implying that amino acid substitutions can affect lid region dynamics [198, 199]. In contrast, the lid region of the 3L variant appeared to be more mobile relative to the wild-type and 1L proteins, and exhibited an open conformation in both interfacial and aqueous environments. These trends correlated with steered MD simulations investigating the lid opening motion of the variants in water. Accordingly, the largest amount of work was evolved in simulations of the 1L lipase, whereas the smallest work was performed in simulations of the 3L variant.

The observed superior stability of the open conformation of the 3L variant relative to the wild-type variant most likely underlies its unique activity profile [206]. This has been related to a reduced energetic barrier for lid displacement in the 3L protein compared to the wild-type, and is thought to be due to the specific residues within the

hinge-regions of the lid [206]. Additionally, the difference in lid dynamics displayed by the 3L and wild-type variants can also be related the substitution of the wild-type E87 residue to a threonine residue in the 3L variant. Previous computational studies of wild-type TLL suggested that displacement of the lid region to an open position in a high dielectric medium was unfavourable due to electrostatic repulsion between the E87 residue and negatively charged residues on the opposite side of the protein [187]. Mutation of the E87 residue to a histidine in this study decreased the overall energy required to displace the lid region, whilst mutation to a positively charged lysine residue resulted in an overall increase in energy gain upon lid displacement [187]. The electrostatic interactions between residues within the lid region and the surrounding protein are thus important in governing the lid opening process. The E87T mutation within the lid region of 3L could therefore result in reduced electrostatic repulsion experienced by the lid region upon lid activation relative to the wild-type variant. This could explain the differing behaviour exhibited by the variants in unbiased AT simulations, particularly in simulations of initially open structures in water.

4.5 Conclusions

In this chapter, a combined simulation and experimental approach has provided insight into the conformational dynamics and interfacial interactions of TL lipase variants at triglyceride interfaces, and in solution. Factors such as altered interfacial binding orientations and lid dynamics could underlie the experimentally determined activity profiles for the wild-type, 1L and 3L variants, and are tightly related to the particular residue composition of the different lid regions. Furthermore, attempts to characterise the affect of substrate hydrolysis on lipase binding and structural dynamics were made. The results from these simulations provide a basis for further investigation with different triglyceride interfaces, such that factors including surface charge and product

accumulation on lipase activation can also be considered. Additionally, lid glycosylation is known to play a part in lipase binding to the lipid interfaces [201, 202]. The lid region of the 1L lipase contains a glycosylation site at the N89 residue of the N-high mannose type, as well as at the N33 residue of the same type [203, 206]. The latter is also present within the wild-type and 3L variants [206]. Simulations of the glycosylated variants will thus be important in providing a comprehensive description of lipase binding dynamics, and for comparison with the experiments [206].

Moreover, it would be of interest to characterise the energetics of the lid activation process for each of the variants in more detail. Although attempts were made to investigate this using SMD methods, these suffer from limitations. First, the SMD simulations were unable to displace the lid region of the 1L variant relative to the 3L and wild-type variants in water. Instead, the neighbouring loops were displaced such that the lid opening process was not adequately sampled. When position restraints were applied to all the atoms within the protein except for the lid region, the CV was able to capture lid motion for the 1L lipase in water, resulting in a much larger work profile. In contrast, similar work profiles were generated during both unrestrained and position restrained SMD simulations of 1L bound to the interface. These results indicate the both the distance CV and the nature of the SMD methods were insufficient in describing the lid activation process accurately. More rigorous enhanced sampling methods are required to firmly establish how lid mutations affect the energetic landscape of lid displacement for the mutated lipases, both at the interface and in solution. Furthermore, subsequent simulations of variants with partially regenerated wild-type amino acid sequences would provide insight into which particular residue substitution, or indeed a combination of these, is responsible for altering the energy landscape of lid activation. These data could then be generalised for other lipases, providing a predictive tool that could be used in the rational design of lipase variants with attractive

properties.

Chapter 5

Lipase Interactions With Graphene and Graphene Oxide

The previous two chapters have focussed on lipase interactions with "soft" biological interfaces, such lipid bilayers and triglyceride surfaces. This and the following chapter will shift the focus to how molecules interacts with "hard" non-biological interfaces. Characterisation of protein interactions with nanomaterials, such as silica nanoparticles and carbon-based supports, has provided fundamental insight into the mechanisms that govern interfacial protein function. The result has been the generation of systems with applications in biosensing, drug delivery, biocatalysis, and decontamination [103, 212–215]. Furthermore, investigations of protein interactions within multicomponent systems, such as supported phospholipid membranes, reveal that specific protein-lipid interactions influence supported membrane spreading behaviour and topology [103].

This chapter and chapter 6 will present results on how lipases and lipids interact with a carbon-based support called graphene. The studies aim characterise how both lipases and lipids behave on this interface, commenting on binding, topology, and molecular properties. Future studies could focus on the merging of these two sys-

tems, investigating how enzymes interact with supported lipid membranes on graphene surfaces.

5.1 Introduction

Graphene is a single layer of graphite, composed of sp^2-hybridised carbon atoms, arranged in a hexagonal pattern [216, 217]. Despite its geometric simplicity, graphene has generated huge scientific interest, particularly since its isolation is 2004 [216]. The surface possesses a vast, modifiable surface area and unique physico-chemical properties that have resulted in applications in fields including biocatalysis, biosensing, drug delivery, and energy storage [213–215]. Furthermore, graphene can be functionalised by exposing it to oxidising agents, resulting decoration of the surface with hydroxyl, epoxide, and carboxylic groups [218]. Oxidised derivatives of graphene are called graphene oxide (Fig. 5.1).

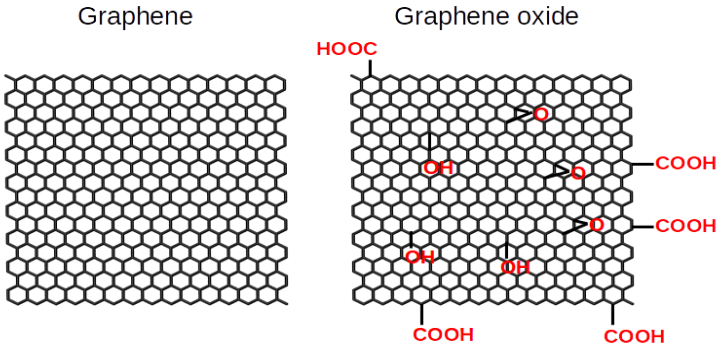

Fig. 5.1: Licorice representations (grey) of graphene and graphene oxide.

Biomolecules, such as enzymes, can interact with graphene nanomaterials via electrostatic, van der Waals, π-π stacking, and hydrophobic interactions [219, 220]. The study of these interactions has led to the generation of biocatalytic systems with advantageous properties, such as enhanced enzyme thermostability, greater control over

reaction conditions, ease of substrate and product isolation, and in some cases, increased enzyme activity [74, 212, 221].

Recently, the interfacial interactions of different lipase and esterase enzymes were investigated with functionalised graphene supports [212, 221]. Esterases perform the same function of hydrolysing triglycerides as lipases, but do not exhibit interfacial activation, rather acting on water-soluble substrate [2, 207]. These studies revealed that interactions of the lipases with graphene oxide resulted in enhanced catalytic efficiency, and in some cases a 2-fold increase in hydrolytic activity relative to the soluble enzyme [212]. Conversely, the interaction of the esterases with the supports severely reduced hydrolytic activity, as well as catalytic efficiency [212]. Furthermore, the addition of hydrophobic hexamethylenediamine groups to the graphene surfaces resulted in a further increase of lipase catalytic efficiency [212]. The differences in catalytic activity was attributed to the conformational dynamics and structural stability of the enzymes, where the esterase were observed to denature in the presence of the graphene surfaces [212]. This behaviour has also been observed in other enzyme-carbon support systems, highlighting the important interplay between the physico-chemical properties of the surface and the conformational dynamics the enzymes [222–224].

In this chapter, I used CG-MD to investigate the adsorption behaviour of the M37 lipase from *Photobacterium lipolyticum* (presented in Chapter 3) on graphene and graphene oxide surfaces. Given the overall planar, rigid nature of the graphene support surface, as well as previous experimental studies with related lipases, this interface provides an interesting example for comparison of lipase interactions with "hard" non-biological *vs* "soft" biological surfaces [73, 212, 219]. To this end, the interfacial interactions of M37 with the graphene supports were compared to simulations of M37 with phospholipid bilayer surfaces (presented in Chapter 3). Furthermore, the influence of surface charge on the lipase binding with graphene oxide surfaces was

studied, revealing differences in the adsorption process and interfacial orientation. Finally, I conclude with future directions, including parameterisation of an atomistic graphene oxide model for simulation with the GROMOS 54A7 forcefield.

5.2 Methods

5.2.1 Graphene and Graphene Oxide Models

Graphene Models

Two different CG models of graphene were initially tested to determine which would be most suitable for simulation. The first CG model applied a 1:4 mapping scheme of CG beads to atomistic (AT) particles, employing Martini forcefield parameters developed for a CG fullerene model [144, 225]. This model was called the 1:4 graphene model. The atomic coordinates for a graphene layer (armchair configuration) were generated using the Carbon Nanostructure Builder plug-in in VMD (version 1.5) [163]. The atomistic structure was then used to build the 1:4 CG model, placing a CG bead at the centre of every four atoms of the AT graphene model (Fig. 5.2). The parameters for the 1:4 model were downloaded from Luca Monticelli's website (http://perso.ibcp.fr/luca.monticelli/MARTINI/index.html) [145, 225].

The second graphene model applied a 1:2 mapping scheme, employing the Martini parameters developed for a CG model of graphite [84]. The coordinates and parameters for this model were downloaded from the Martini forcefield website (http://www.cgmartini.nl/). This model was called the 1:2 graphene model. Both CG graphene models are represented by van der Waals spheres in Fig. 5.2. The radius of the van der Waals spheres was scaled such that the underlying atomistic structure of graphene was visible, and so that the mapping scheme between the two models could be distinguished. Conse-

quently, the CG model may appear permeable, however this is purely due to size of the van der Waals representations. These are later adjusted to their "true" size within the CG model (based on σ parameter used for the LJ interaction), such that the CG graphene sheets are not permeable within the CG simulations.

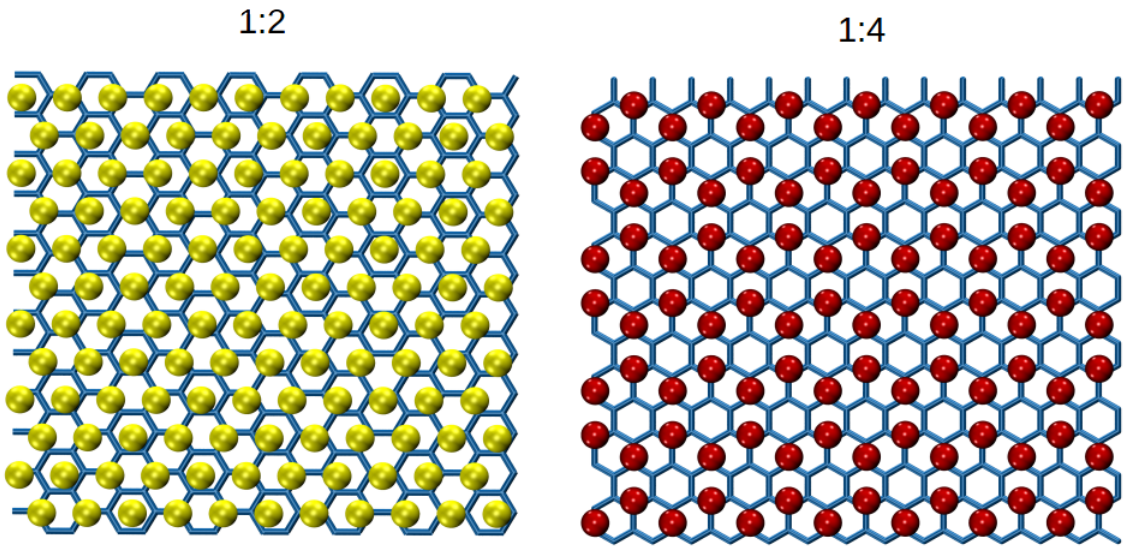

Fig. 5.2: The mapping scheme applied for the 1:2 and 1:4 graphene models. The underlying atomistic graphene sheet is shown in licorice and coloured in dark blue. Van der Waals representations of the CG particles show the position of each CG bead for the 1:2 (yellow beads) and 1:4 models (red beads). The radius of the van der Waals spheres for the CG models was scaled such that the underlying atomistic structure of graphene remained visible, allow the mapping scheme between the two models to be distinguished. Consequently, the CG models may appear permeable, however this is purely due to size of the van der Waals representations. This adjustment is not present during the production simulations, where the CG graphene models are simulated using the original LJ parameters, such that the CG graphene sheets are not permeable within the simulations [84, 225].

Each CG graphene model was individually placed in the centre of a 10x10x10 nm simulation box and solvated with standard Martini CG water [134]. In order to maintain planarity of the graphene sheets, the models were position restrained using a force constant of 10,000 kJ mol^{-1} for the 1:4 model, and 5000 kJ mol^{-1} for the 1:2 model. Additionally, the 1:2 graphene sheet beads were kept frozen during simulation

(temperature of 0 K) [84]. The systems were then energy minimised, and simulated at a constant temperature of 303 K and a constant pressure of 1 bar for 250 ns each (NPT ensemble). Non-bonding interactions were modelled using shift functions; both LJ and Coulombic interactions were evaluated within a 1.2 nm cut-off, and shifted within a 0.9 nm cut-off distance for the LJ interactions, and within a 0.0 nm cut-off distance for the Coulombic interactions [84]. Temperature was regulated using the Berendsen thermostat and a coupling constant of 0.3 ps [120]. Anisotropic pressure coupling was applied using the Berendsen barostat with a coupling constant of 3 ps and a compressibility of 0.5×10^{-5} bar^{-1} in the x/y direction, and 3.0×10^{-5} bar^{-1} in the z direction [120]. These parameters reflect those employed in initial parameterisation studies of the 1:2 graphene model [84].

CG-MD simulations of either the 1:2 or 1:4 graphene models resulted in rapid freezing of the standard Martini water particles, forming ordered layers that spanned the entire simulation box (Fig. 5.3A). Although experimental and computational studies have suggested that water can form ordered layers directly above the graphene surface, these are thought to only extend to two-three layers hydration layers above the surface [226–228].

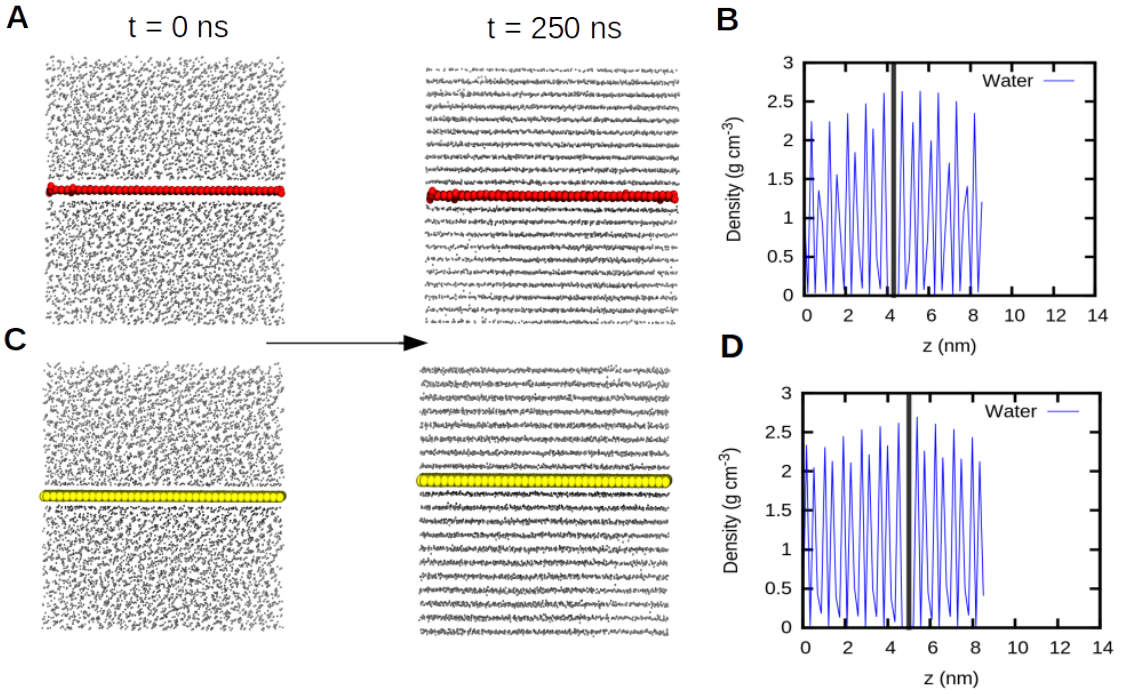

Fig. 5.3: Images of the first and final frames of CG-MD simulations of the 1:2 **(A)** and 1:4 **(C)** graphene models in water using the standard Martini water model. The 1:4 graphene model is shown as red van der Waals representations, the 1:2 model is shown as yellow van der Waals representations. Water is shown as grey CPK representations. **(B)** and **(D)** show partial density profiles calculated for the last 25% of simulation time for the simulations in **(A)** and **(C)**. The position of the graphene surface within the simulation box is shown by the solid black line.

The freezing effect has been observed in other CG Martini simulations with graphene surfaces [105, 107]. The standard Martini water model is known to exhibit a freezing temperature that is too high compared to real water; depending on the system, the water can freeze between 280 and 300 K, particularly if a nucleation point such as a solid surface is present [134]. "Anti-freeze" particles were introduced to address this issue. These particles possess a larger σ Lennard-Jones (LJ) parameter (0.57 nm instead of 0.47 nm), and exhibit stronger interactions with the standard water particles than water-water self-interactions [134]. Anti-freeze particles would therefore disturb lattice packing of the uniformly sized standard water particles, the effect of which should be seen when randomly replacing 10% of the standard water particles with the

anti-freeze particles [134]. This protocol was adopted for both 1:2 and 1:4 models, however this did not alleviate the freezing effects observed. Varying other simulation parameters such as increasing temperature or adding salt (Na^+ or Cl^- ions) also did not result in liquid water behaviour.

Eventually systems were setup using the polarisable Martini water model, which is said to address the higher freezing temperature exhibited by the standard water model compared to real water [148]. Water freezing effects were minimised in CG simulations with the 1:2 graphene model, but persisted with the 1:4 graphene model (Fig. 5.4). All subsequent simulations were therefore performed with the 1:2 graphene model, given that further parameterisation of the 1:4 model would be required to more accurately model a liquid phase aqueous system.

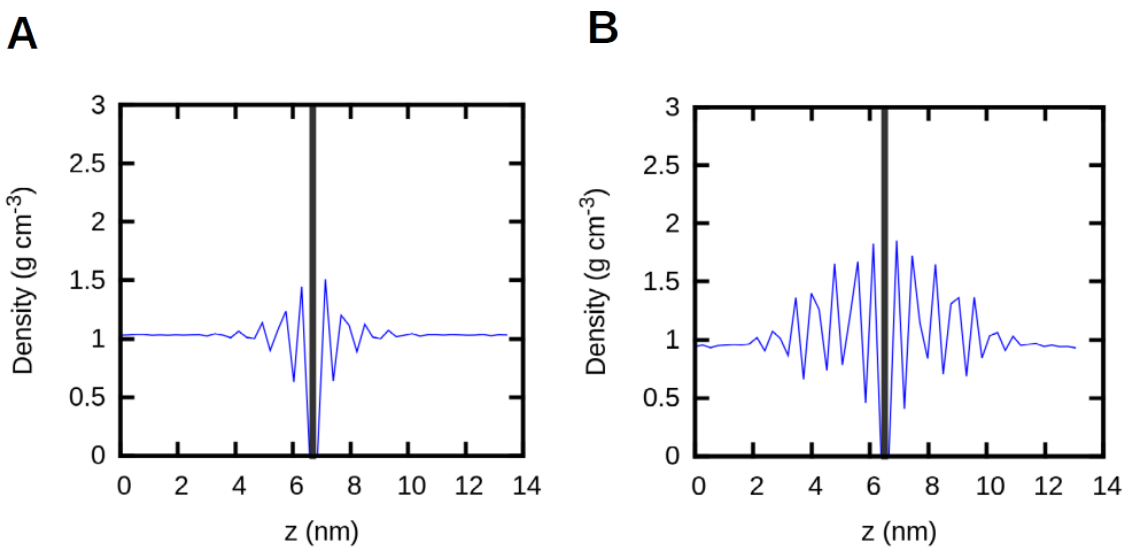

Fig. 5.4: Partial density profiles calculated for the last 25% of simulation time for CG simulations of the 1:2 **(A)** and 1:4 **(B)** graphene models with polarisable water (250 ns). The position of the graphene surface within the simulation box is shown by the solid black line. The profiles indicate that the 1:4 CG graphene model induces formation of several ordered water layers (∼5), whereas the 1:2 model induces formation of ∼2. Therefore, simulation with the 1:2 model resulted in the least freezing of water, as was regarded as most suitable for further simulations.

As can be seen from Fig. 5.4, a small number of ordered water layers (∼2) are

formed above and below the 1:2 CG graphene surface. Experimentally, the formation of ordered hydration layers is thought to be due to hydrogen-bonding, electrostatic, and dipolar interactions between the hydrogen-carbon atoms of the graphene sheet and the hydrogen-oxygen atoms of the water molecules, forming a network that results in an ordered water structure [228, 229]. Naturally, the nature of the CG water model, which maps 4 atomistic water particles within one CG water bead, implies that the ordered layers seen with the CG models extend beyond the number that is observed experimentally (2-3 hydration layers) [226–228]. A direction conversion of one CG water bead to 4 water molecules implies that 2 CG water layers could translate to 8 atomistic water layers depending on the relative orientation of the water molecules, which again is not captured within the CG water model. Furthermore, the CG graphene and graphene oxide models assume that the graphene surface is entirely flat and planar, which is known to not be the case experimentally [230]. Such a planar surface would naturally introduce order into the system, and may result in artificial stabilisation/formation of the ordered water layers observed in the CG model, which may not be present if the graphene surface was more kinked or contained planar defects. Consequently, it remains difficult to compare the CG model to experimental or atomistic systems, however the observation that ordered water layers are known to form above graphene and silicon dioxide surfaces remains valid in understanding the natural behaviour of such systems.

Graphene Oxide Models

CG graphene oxide models were built by replacing the carbon beads in the 1:2 graphene model with beads representing oxygen-containing functional groups. Hydroxyl and epoxy groups were modelled using the SP1 bead and negatively charged

carboxylic acid groups (COO^-) were modelled by the SQa bead [134]. Following on from experimental and computational observations that hydroxyl groups tend to be found on carbons atoms that neighbour already oxidised carbons atoms (e.g attached to epoxy groups) [85, 231], the CG graphene oxide models were generated using the following protocol: the first carbon bead to be "oxidised" was chosen, and replaced by an SP1 (hydroxyl/epoxy) bead, the next bead to be replaced was then chosen in the proximity of the now converted SP1 bead, and subsequently replaced. This procedure was not random as the script used to convert the graphene model to the graphene oxide model performed the conversion for every 5 lines within the coordinate file. A more random approach would have consisted of randomly choosing a line within the coordinate file, and converting the SG4 bead to a SP1 bead. This approach was implemented later, generating a more randomly "oxidised" graphene oxide model. Test simulations with this more "random" model indicated that the behaviour of the systems did not alter much compared to model generated by the less random method. This could be a consequence of the CG nature of the model, such that the distribution of identical polar beads (SP1) may not affect the behaviour of interacting particles due to the short-range non-bonding interaction potentials applied. Therefore, the simulations detailed within this chapter concern the graphene oxide model presented in Fig. 5.5, generated by the non-random procedure.

The stoichiometry of oxygen functional groups relative to carbon atoms can differ depending on the synthesis method [232–234], and several groups have proposed different models for the structure of graphene oxide [218, 234, 235]. Based on models proposed in the experimental literature, two differing percentages of oxygen content were modelled: 20% and 23% (Fig. 5.5) [236]. The 20% oxidised model consisted only of carbon and oxygen (SP1) beads. To investigate the affect of surface charge, the 23% model was built by randomly replacing a number of carbon beads at the edges

of the 20% oxidised surface with SQa beads representing COO^- groups, resulting in a final oxidation percentage of 23%.

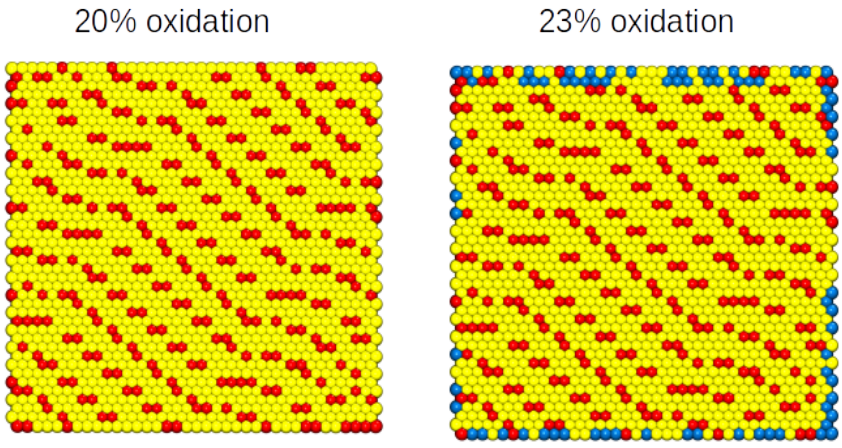

Fig. 5.5: CG graphene oxide models for simulation with lipase enzymes. Two different oxidation percentages were modelled: 20% and 23%. Surfaces are shown as van der Waals spheres. Carbon beads are coloured yellow, oxygen beads are coloured red, and carboxyl (COO^-) beads are coloured blue.

5.2.2 Protein Models

The same CG model of M37 and parameters as presented in the Chapter 3 were used for the simulations with graphene and graphene oxide.

Positive control simulations for protein interactions with graphene were performed with a small fungal protein called hydrophobin (HFBI), which is known to interact with pristine graphene surfaces [237, 238]. The crystal structure of HFBI from *Trichoderma reesei* (PDB: 2FZ6) was downloaded from the PDB [239] and converted to CG Martini representations employing the Martini 2.2 forcefield [134, 138, 143]. The ElNeDyn (EDM) parameters that maintain secondary and tertiary protein structure were calibrated for HFBI, following the same protocol as presented in Chapter 3.

5.2.3 System Setup and Parameters

M37 or HFBI were individually placed 3 nm above the graphene or graphene oxide surface in a simulation box (10x10x15 nm) and solvated with polarisable water (Fig. 5.5). These systems were energy minimised, and then equilibrated using the following protocol: 1) position restraints were applied to both the protein and the surface, 2) an initial time step of 1 fs was applied for 10,000 steps, and incrementally increased to the final 20 fs time step by performing consecutive simulations using time steps of 2 fs, 5 fs, 10 fs, and 20 fs for 10,000 steps each, totalling 5 independent equilibration simulations for each system. All systems were then simulated at 303 K, releasing protein position restraints [212].

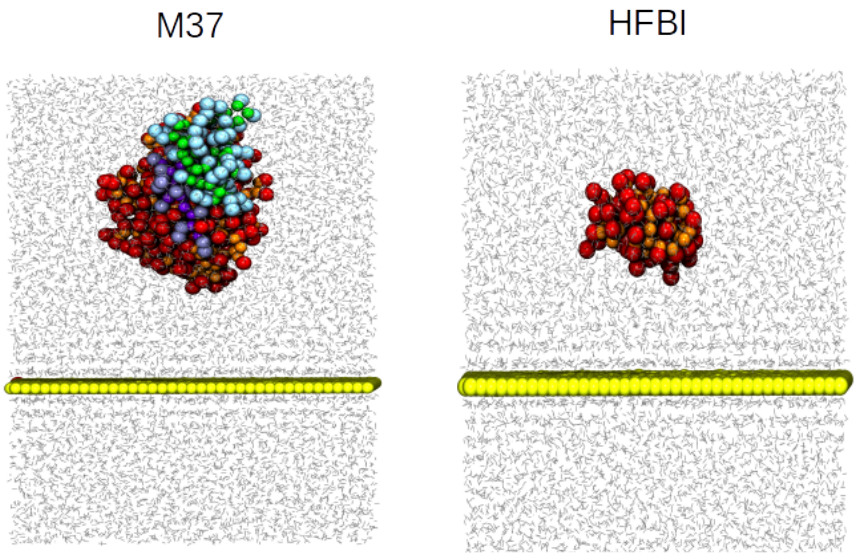

Fig. 5.6: The starting configurations of M37 (*left*) and HFBI (*right*) simulations with the 1:2 graphene model. The proteins and graphene (yellow) are represented as van der Waals spheres. The lid region of M37 (residues 235-283) is shown in cyan (side chain particles) and green (backbone particles), and active site flap region (residues 94-110) is shown in dark purple (backbone) and light purple (side chains); the rest of the protein is coloured in orange (backbone) and red (side chain). HFBI is shown in orange (backbone) and red (side chain). Water is shown as grey lines.

5.2.4 Diffusion Analysis

The diffusion of M37 on different surfaces was analysed using documented open-source code (http://dx.doi.org/10.5281/zenodo.11827). The code employs an algorithm which calculates the mean square displacement values of individual lipid centroids over a range of time sampling windows, including: 1, 2, 5, 10, 20, 40, 100, and 140 ns. The diffusion coefficients are then calculated by fitting the MSD *vs* time data using a linear diffusion equation [240], using a least squares first degree fit of the data. Standard deviations of the diffusion coefficients were estimated as the difference of the slopes from the first and second halves of the MSD *vs* time data. This is the approach used by the GROMACS tools function *g_msd*. In addition to linear fitting, the code also allows fitting of the MSD data with an anomalous diffusion equation, distinguishing it from the *g_msd* tool, which only performs linear fitting of the MSD data. For this reason, the open source code was used rather the *g_msd* tool. The anomalous approximation uses a non-linear least squares fitting to a two parameter equation of the form:

$$MSD = 4D_{\alpha}t^{\alpha}, 0 < \alpha < 2 \tag{5.1}$$

where D_{α} is the fractional diffusion coefficient, measured in units of length^2/time$^{\alpha}$. The scaling exponent (α) was estimated as the slope of the log MSD *vs* log time data. The standard deviations of both parameters (D_{α} and α) were calculated from the square root of the diagonal of the covariance matrix from the anomalous fit. This code has been used by previous simulation studies of lipid diffusion in virus particles [128, 129].

5.3 Results

The study of enzyme interactions with carbon materials has resulted in the generation of biocatalytic systems with advantageous properties, such as increased enzyme thermostability and, in the case of lipases, enhanced enzyme activity [212, 241]. Here, CG-MD simulations were used to investigate how the M37 lipase interacts with pristine graphene and graphene oxide surfaces. The simulations were compared to previous simulations of M37 interactions with biological surfaces, such as phospholipid membranes (presented in Chapter 3).

5.3.1 Lipase Interactions with Graphene Surfaces

Three replicate CG-MD simulations of M37 positioned above a pristine (unfunctionalised) graphene surface were performed for 2 μs each. None of these simulations resulted in interfacial interactions between the lipase and the graphene surface. The lipase was observed to randomly diffuse through the surrounding aqueous solvent, occasionally approaching the surface within two-three hydration layers, however direct interfacial contacts were not observed in any of the simulations within the simulated time.

Additional simulations of a small protein molecule called hydrophobin (HFBI; see Methods), known to interact with pristine graphene, were performed as a positive control [237, 238]. These simulations did result in interactions of HFBI with the graphene surface (Fig. 5.7), mediated by mainly hydrophobic contacts with the surface (Fig. 5.8). HFBI adsorbed quickly to the surface and remained relatively static once bound, most likely due to strong hydrophobic interfacial interactions with the hydrophobic support. This analysis suggests that the M37 lipase does not interact with pristine graphene within the time scales studied here, perhaps requiring polar or electrostatic interactions to generate a bound configuration, as was seen with zwitterionic *vs* an-

ionic bilayer surfaces in Chapter 3. The following results thus only refer to simulations of M37 with graphene oxide.

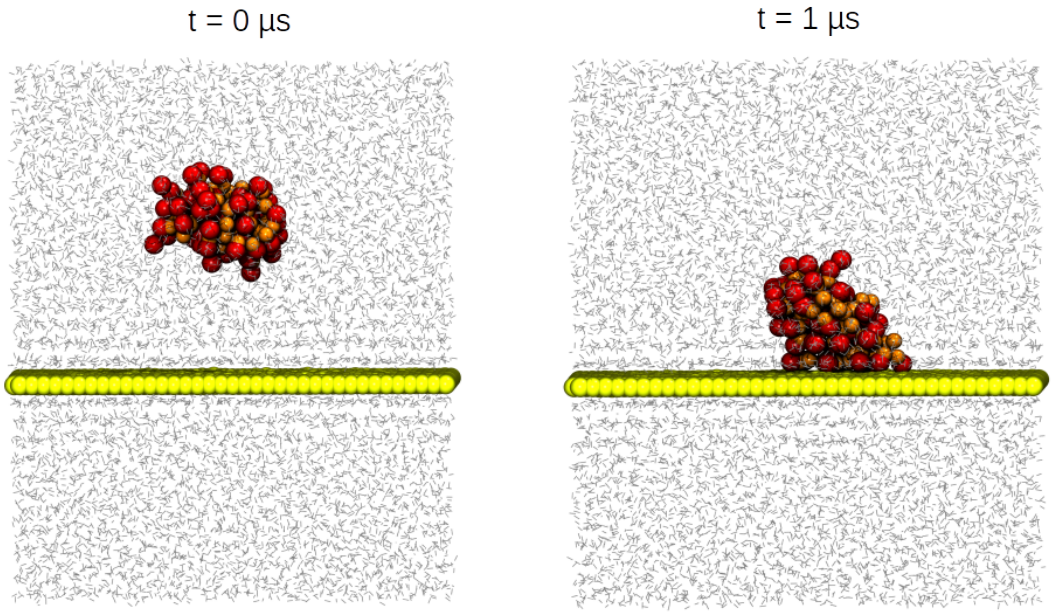

Fig. 5.7: Starting and end configuration of a CG-MD simulation of HFBI with a pristine graphene surface. The protein (orange and red) and graphene surface (yellow) are represented as van der Waals spheres, water is shown as grey lines.

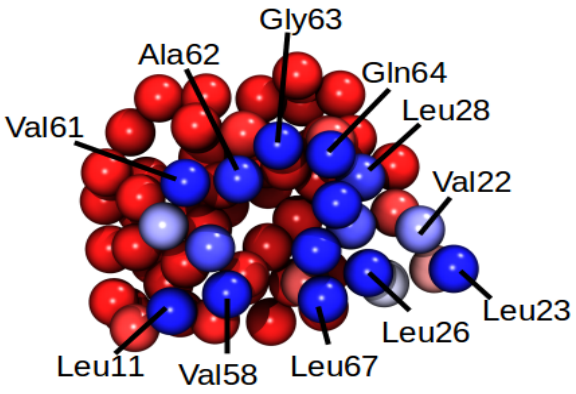

Fig. 5.8: Contact map showing the residues that mediated interactions between HFBI and the pristine graphene surface during a CG-MD simulation. Contacts are coloured in blue for largest number, white for middle, and red for lowest or no contacts. Interactions were calculated within an 0.8 nm cut-off.

5.3.2 Lipase Interactions with Graphene Oxide Surfaces:

Previous experimental studies suggest that lipase interactions with graphene *oxide* results in enhanced hydrolytic activity and catalytic efficiency relative to their soluble counterparts (*i.e* in the absence of any nanomaterials) [73, 212, 221]. CG-MD simulations of M37 individually placed above graphene oxide surfaces of differing oxidation percentages (see Methods) were performed to investigate lipase interactions with this functionalised surface.

20% graphene oxide

Contrasting simulations with the pristine graphene support, initial simulations of the lipase-graphene oxide systems showed that M37 was able to adsorb to this surface, remaining bound to the support for the duration of the simulation (4 μs). Further replicate simulations of M37 with graphene oxide were performed, resulting in a 10 replicate ensemble, in which the lipase exhibited binding interactions with the interface in 8/10 simulations. Before binding to the surface, the lipase explored the surface hydration layers known to form above graphene and graphene oxide interfaces [226–228]. M37 interactions with the ordered water layers resulted in reduced random diffusion through the surrounding bulk water environment. Once the lipase bound to the surface, the protein did not dissociate within the simulated time frame in any of the simulations except for Run 9. In this replicate, M37 made a small number of initial contacts with the surface, followed by dissociation, and subsequent reassociation with the surface, in which many more interfacial contacts were formed. Similar to CG-MD simulations of M37 interactions with anionic phospholipid bilayers (512 lipids; 20% PG:80% PC), three different binding orientations of the enzyme within this replicate ensemble could be identified (Fig. 5.9). Table 5.1 reports which binding orientation

was adopted by M37 for each simulation within the ensemble.

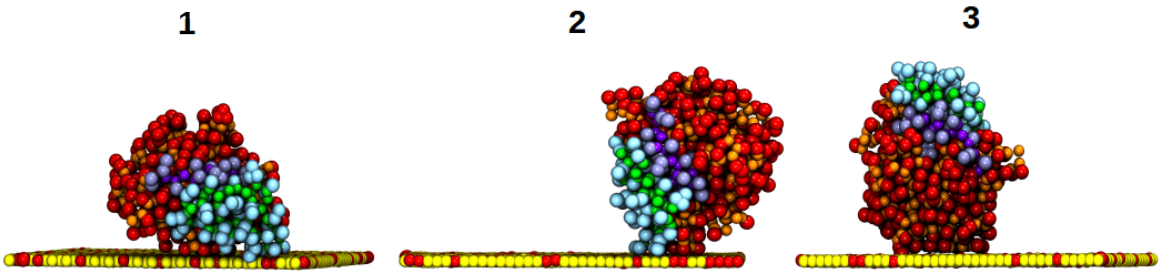

Fig. 5.9: Three different binding orientations were identified from CG-MD simulations of M37 with graphene oxide (20% oxygen content), denoted by the numbers. The lid region of M37 (residues 235-283) is shown in cyan (side chain particles) and green (backbone particles), and active site flap region (residues 94-110) is shown in dark purple (backbone) and light purple (side chains); the rest of the protein is coloured in orange (backbone) and red (side chain). The graphene oxide surface is shown as yellow (carbon beads) and red (oxygen beads) van der Waals representations. Water is omitted for clarity.

Binding orientation	Replicate simulation
1	1, 2, 3, 5, 6, 8
2	9
3	7

Table 5.1: Number of individual simulations (10 replicates) corresponding to a particular binding orientation of M37 during interactions with graphene oxide. The binding modes are presented in Fig 5.8. Two simulations did not result in binding of the M37 lipase with the graphene oxide surface. The binding orientations were not observed to convert during the span of a single simulation. Once bound to the surface, the protein did not dissociate in any replicate simulation, except for simulation 9.

These three orientations were characterised by analysis of the rotational and translational motions exhibited by the enzyme as it adsorbed on the graphene oxide surface, revealing differences in the *Rzz* angle value (Fig. 5.10A-B). This metric is calculated as the angle between a given orientation of the enzyme in simulation relative to a reference orientation (same protocol as presented in Chapter 3 and 4). Given that orientation

1 was the most frequently sampled orientation within the 10 replicate ensemble, this served as the reference orientation for calculation of the *Rzz* value (rotation matrix analysis). The resultant analysis reflects the dissociation/association event exhibited in Run 9 (orientation 2), as is seen from the distance metric (d) measuring the centre of mass (COM) distance between the protein and the surface. Furthermore, the *Rzz* value clearly changes as the final binding orientation is adopted throughout the CG simulations for each of the different binding orientations. This is particularly true for the CG simulation which M37 adopted binding orientation 3 (Fig. 5.10C), in which a clear shift in *Rzz* value is observed from ∼0.7 μs to ∼1 μs, as the lipase further adsorbs onto the graphene oxide surface (see inset images). The *Rzz* is clearly sensitive to such small changes, indicating only small variations in lipase binding orientations will produce the same or a similar Rzz value. Thus, the rotational matrix analysis is able to distinguish between different binding orientations in the analysis of M37 binding behaviour.

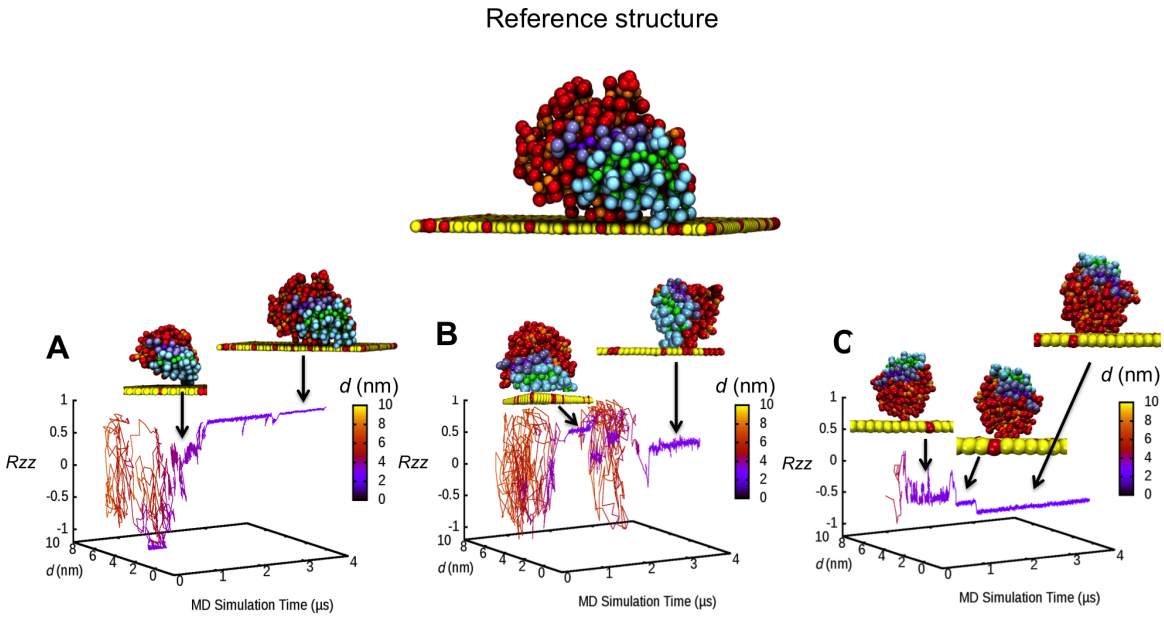

Fig. 5.10: Rotational and translation analysis of the M37 adsorption process on 20% graphene oxide, calculated for 3 representative CG-MD simulations. The metric d was calculated as the z-component of the distance between the COM of the protein relative to the COM of the graphene oxide surface. A rotation matrix was calculated (Rzz) that defines the transition of an enzyme orientation during the simulation to a reference orientation (*top panel*). This analysis is shown for individual simulations in which M37 adopts either binding orientation 1 (Run 1) (**A**), 2 (**B**), and 3 (**C**). The graphical analysis is accompanied by images showing the orientation of M37 on the graphene oxide surface during the simulation, correspond to changes in Rzz values. The colour scheme used for the reference orientation image is the same as in Fig. 5.9; water is omitted for clarity.

It should be noted that binding orientations 2 and 3 are only exhibited by 1 respective simulation each (Table 5.1). Furthermore, binding orientation 1 and 2 are very similar in that the lid region (residues 235-283) of M37 mediates the initial interactions of the lipase with the interface. It is possible that extension of the simulation exhibiting orientation 2 would result in further adsorption of the lipase onto the surface, eventually adapting to orientation 1. The distinction between the three identified orientations is also reliant on the number of repeat simulations performed within the ensemble. It is possible that extension of the simulation ensemble could result in iden-

tification of more binding orientations, or indeed increase the sampling of the already defined orientations. The analysis of M37 interactions with graphene oxide is therefore limited in sampling and the three identified binding orientations of M37 should be considered as three of possibly several other binding orientations that M37 can adopt through interactions with the graphene oxide surface. Indeed, other methods, such as parallel-tempering Monte Carlo simulations have been useful in identifying the optimal binding orientations of lipases with graphene surfaces, prior to CG simulation [173]. This could be alternative method to identify the most likely binding orientation without limitations imposed by performing multiple CG simulations.

Overall however, the lipase appeared structurally stable on the graphene oxide surface for all observed binding orientations, as shown by constant RMSD values calculated for protein backbone particles, peaking at 0.1 nm, throughout the simulations (Fig. 5.11). Naturally, the elastic network employed in the CG simulations restricts further conformational changes that might occur as a function of protein-surface interactions.

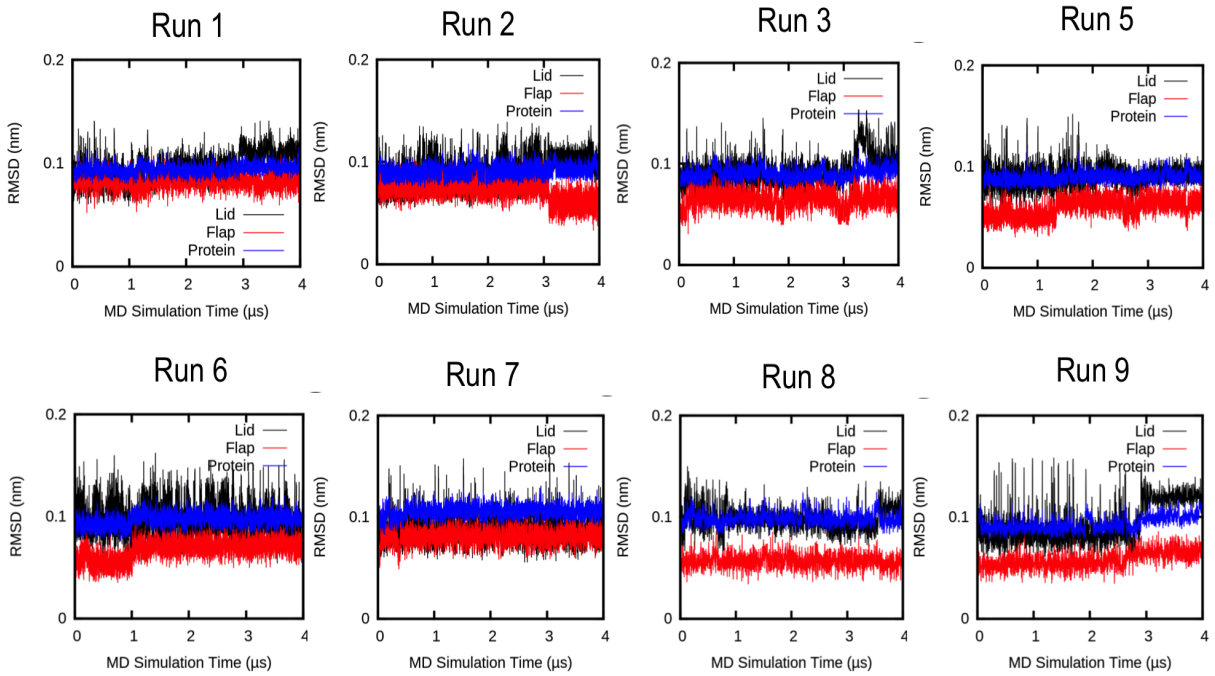

Fig. 5.11: Time evolution of the RMSD of the protein, lid and and active site flap backbone particles, calculated for CG simulations of M37 with a graphene oxide (20%) surface.

Interestingly, binding orientation 1 is similar to the *Angled* orientation adopted by M37 through interfacial interactions with anionic bilayers and triglyceride interfaces (Fig. 5.12). However, the active site flap region (residues 94-110) does not engage the oxide surface, resulting in an interfacial orientation that is nearly the mirror image of the *Angled* orientation identified from simulations with anionic bilayers (Fig. 5.12). Consequently, the contacts mediating the main interactions between the protein and the graphene oxide surface differ from those identified from the M37-bilayer simulations (Fig. 5.12).

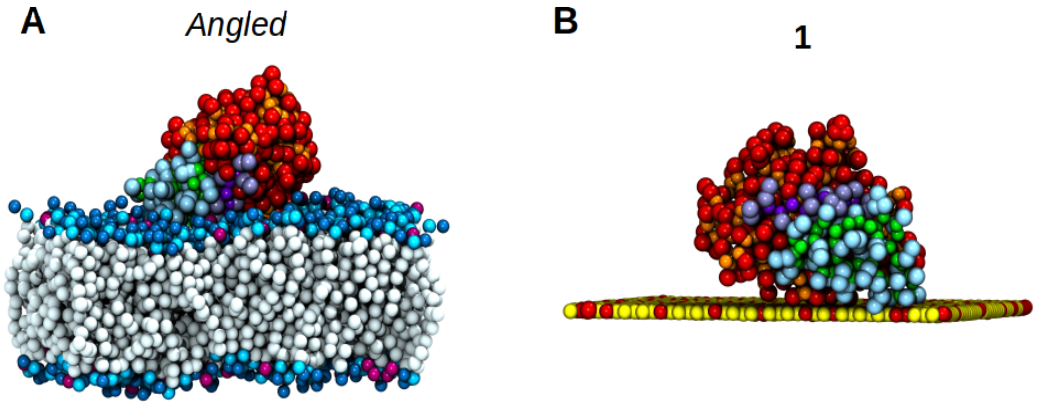

Fig. 5.12: **(A)**: The *Angled* binding orientation identified in simulations of M37 with an anionic bilayer (20% PG:80% PC; PG = purple headgroups, PC = dark blue headgroups). **(B)**: Binding orientation 1 of M37 identified in simulations with a graphene oxide surface. All of the components are shown as van der Waals spheres. The same colour scheme as in Fig. 5.9 is used for the M37. Water is omitted for clarity.

The interactions between the protein and graphene oxide surface were mainly mediated by polar and charged amino acids, such as Ser254, Gln256, Asp216, and Glu278 (orientation 1 & 2) (Fig. 5.13B and C), whilst M37-bilayer contacts were mainly mediated by basic and hydrophobic amino acids (Fig. 5.13A). Binding orientation 3 showed a larger distribution of residue interactions, and also included contacts mediated by charged residues such as Lys157, Lys164, Glu80, and Glu156 (Fig. 5.13). These interactions represent the larger binding surface exhibited by orientation 3, resulting in a larger number of contacts relative to orientations 1 and 2. Importantly, a significant proportion of polar and charged residues were found to mediate interfacial contacts in all binding orientations, suggesting an important role for polar and electrostatic interactions in generating bound configurations of M37. In contrast, mainly hydrophobic interactions were identified in simulations of the HFBI protein with the *pristine* graphene surface (Fig. 5.8), indicating that the absence of polar groups on the surface could underlie the lack of binding events observed for M37 with this hydrophobic surface. This is in line with experimental studies of similar systems [223, 224].

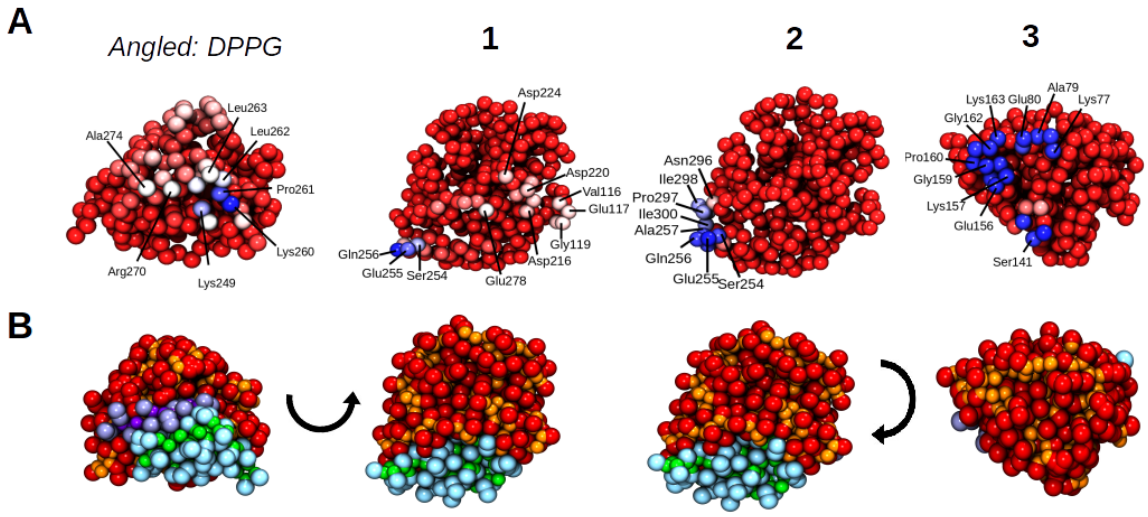

Fig. 5.13: **(A)** Contact residue maps calculated for CG-MD simulations of M37 with a bilayer surface (*Angled*), and a graphene oxide surface, adopting either orientation 1, 2, or 3 (denoted by numbers) (see Fig. 5.9). **(B)** Reference orientations of the lipase for each of the contact residue maps. Contacts were calculated within a 0.8 nm cut-off. Contacts are coloured in blue for largest number, white for middle, and red for lowest or no contacts. The same colour scheme as in Fig. 5.9 is used for M37.

Unlike the M37-bilayer simulations, the lipase did not exhibit conversion between different binding orientations once adsorbed to the surface in any of the replicate simulations (within the simulated time). This is likely due to the nature and strength of the interactions between the lipase and surface, in which initial protein-surface contacts were fully maintained throughout the simulations, and additional interactions resulted in further adsorption of the lipase onto the surface. This can be seen from distance analysis between the protein and surface within the replicate ensemble, resulting in incremental decreases in distance over simulation time throughout the adsorption process (Fig. 5.14). This is also observed in related studies of protein interactions with a graphene surface [242]. The dynamic nature of the lipid bilayers on the other hand, allowed the lipase to embed within the interface, and alter its binding orientation as a function of protein-bilayer interactions.

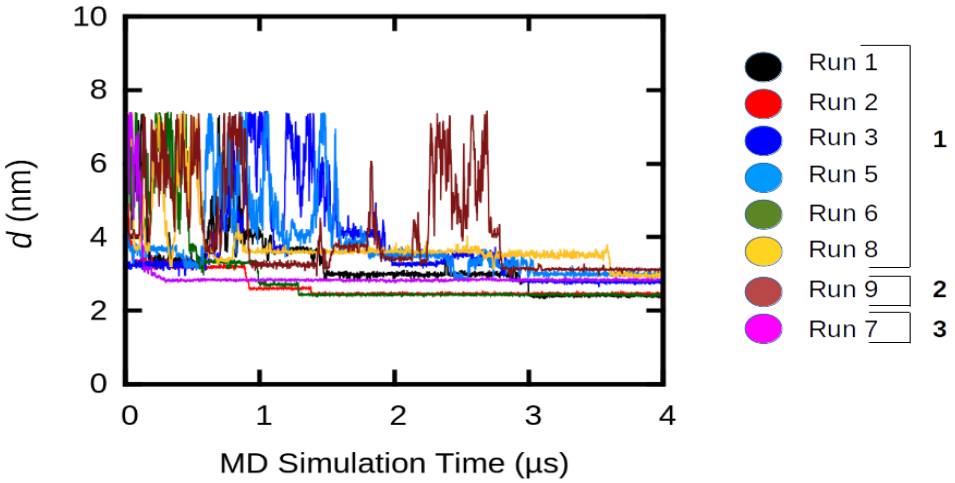

Fig. 5.14: Time evolution of the z-component of the COM distance between M37 and the 20% graphene oxide surface, shown for 8/10 CG-MD simulations in which the lipase associated with the graphene oxide surface. The binding orientation adopted by M37 in the simulations is shown by the numbers in the key (1, 2, or 3). The graphene oxide surface is positioned at $d \sim 1$.

Furthermore, the lateral mobility of M37 once adsorbed to the surface was minimal. This is observed by tracing the COM of the bound protein on the graphene oxide surface as a function of simulation time (Fig. 5.15). While the lipase is able to diffuse across the planar surface in binding orientation 1, and to a lesser extent binding orientation 2, diffusion is extremely limited for orientation 3. The difference in protein mobility can be related to the larger number of protein-surface interactions mediating the binding orientation 3 in comparison to binding orientations 1 and 2, highlighting the role of interfacial contacts in affecting surface diffusion (Fig. 5.15B). For example, initial interactions between M37 and graphene oxide in orientation 1 (t_{MD} = 1.5-3 μs) resulted in ~50 total contacts, which corresponds to the white part of the bound trajectory time trace (Fig. 5.15A). Further lipase-surface interactions resulted in the enzyme fully adopting binding orientation 1, such that the number of contacts increased and protein-surface distance decreased, which corresponds to the blue part of

the trajectory. When comparing this behaviour to M37 in binding orientation 3, the number of protein-surface contacts is almost double, resulting in very little protein diffusion on the graphene oxide surface (Fig. 5.15). Furthermore, extension of the Run 1 replicate simulation (binding orientation 1) to 10 μs did not result in additional surface diffusion of the adsorbed enzyme, indicating that once a number of contacts have been made, M37 is relatively static on the graphene oxide support, at least within the simulated time scales.

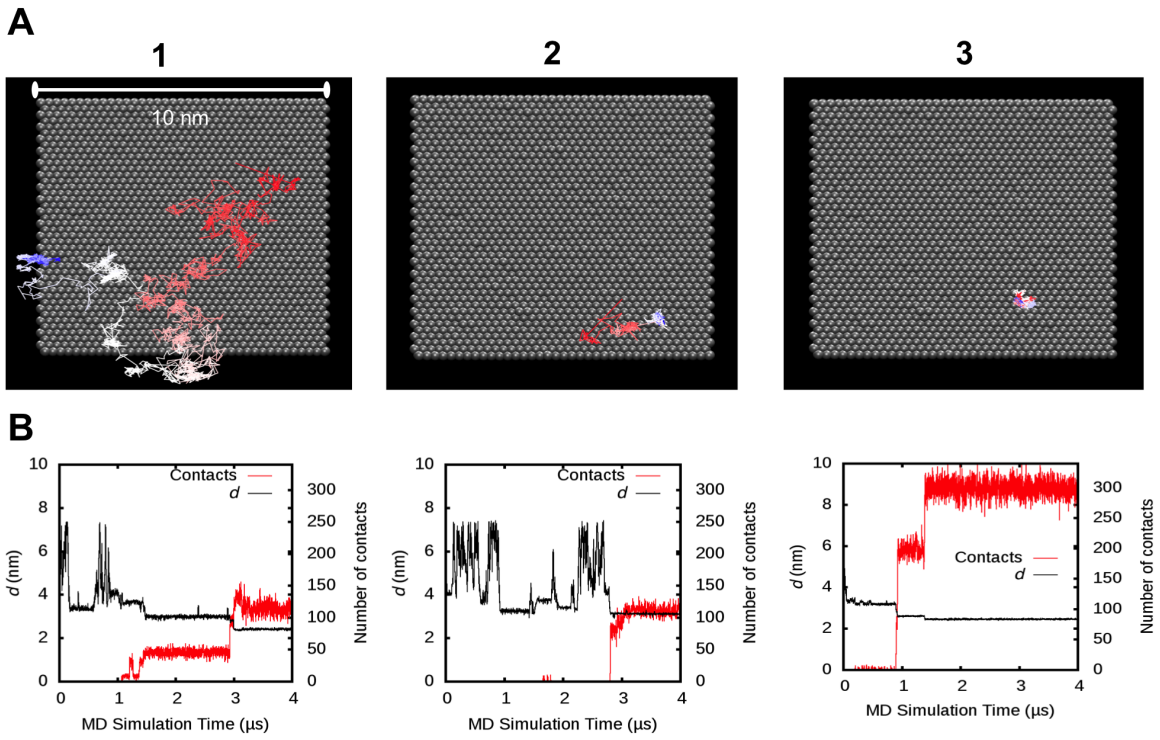

Fig. 5.15: **(A)** Time traces showing the position of M37 during CG simulations of the lipase with a graphene oxide surface. The lines trace the COM of M37 once adsorbed onto the graphene oxide surface, visualised for representative CG simulations in which M37 adopts orientation 1 (Run 1), 2, or 3 (indicated by numbers). The time traces are coloured in a red-blue scale, where red corresponds to the beginning of the trajectory, white to the middle, and blue to the end of the trajectory. Image perspective is from the top of the simulation box looking down onto the surface. The graphene oxide surface (10x10 nm) is shown as van der Waals spheres coloured in grey. **(B)** Time evolution of d (left axis), measuring the COM distance between the protein and graphene oxide, and number of contacts between the protein and the surface (right axis) calculated for each of the simulations shown in **(A)**.

The limited lateral diffusion of M37 on the graphene oxide surface is further highlighted through comparisons with the lateral diffusion of M37 bound to anionic bilayers (Fig. 5.16). Diffusion was calculated by measuring the mean square displacement (MSD) of the protein on the surface for different time sampling points, and fitting the resulting data with a linear diffusion equation [240] (see Methods).

In particular, the MSD data indicates almost no mobility of the protein when

adsorbed on the graphene oxide surface within the sampled time frames, whereas the MSD data for M37 bound to bilayer surfaces samples much higher values (Fig. 5.16). Correspondingly, a diffusion coefficient of 0.2 ± 0.1 x10^{-7} cm^2/s was estimated for M37 bound to the graphene oxide support, compared to 2.1 ± 0.2 x10^{-7} cm^2/s for M37-bilayer simulations (unscaled diffusion coefficients).

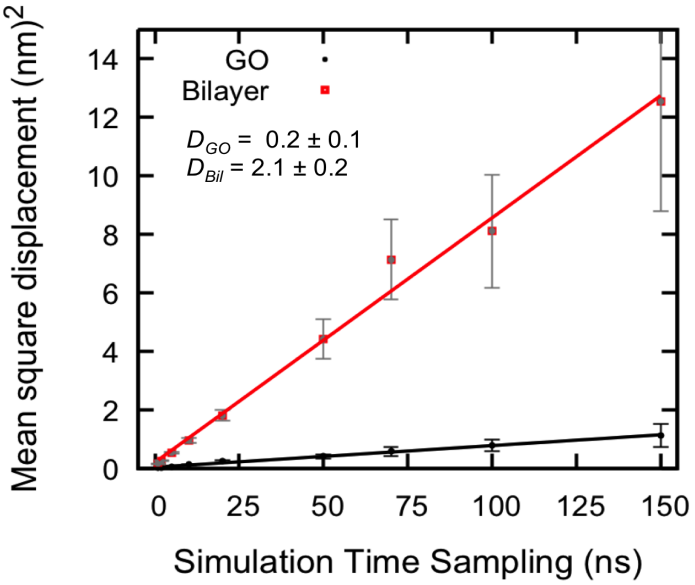

Fig. 5.16: MSD data plotted against different time sampling windows (ns), calculated for CG-MD simulations of M37 with graphene oxide (GO) and with an anionic lipid bilayer (Bil) (Chapter 3). The MSD data was fitted with a linear diffusion equation; the resultant diffusion coefficients are shown (D_{GO}, D_{Bil}) in units: x10^{-7} cm^2/s. Errors bars show standard error of the mean.

Overall, the CG results suggest that M37 interacts differently with graphene oxide surfaces relative to anionic lipid bilayers, in which polar and electrostatic interactions are important in driving adsorption of the lipase on to the graphene oxide surface. Conversely, hydrophobic and basic residues were important in mediating interfacial interactions between M37 the anionic lipid bilayers (Fig. 5.13). Furthermore, the limited diffusion of the protein on the oxide surface appears to be related to the number of protein-surface contacts, contrasting the higher mobility of the M37 lipase

when bound to anionic bilayers. This could be related to the strength of interactions between the protein and the "hard" graphene oxide support compared to the "soft" bilayer systems.

23% Graphene Oxide

Previous experimental studies of graphene oxide surfaces reveal that the edges of the sheet are decorated with carboxyl groups [218, 243, 244]. In order to investigate the affect of surface charge on the adsorption behaviour of the M37 lipase, CG-MD simulations were performed with a graphene oxide model that included additional negatively charged carboxyl groups randomly placed at the edges of the surface (oxygen content = 23%). Three replicate CG-MD simulations of M37 positioned 3 nm above the oxidised surface were performed (4 μs), all of which resulted in lipase interactions with the graphene oxide model (Fig. 5.17). Once the lipase adsorbed to the surface, it did not dissociate in any of the simulations. M37 mainly interacted at the edge of the graphene oxide surface, exhibiting similar binding orientations as identified from simulations with the 20%-oxidised graphene model (Fig. 5.17A). These orientations were termed 1*, 2*, and 3*.

The lipase adsorption process tended to occur more rapidly with the 23% oxidised surface, indicating that the presence of the carboxyl groups and charged interactions influenced M37 association with the interface, particularly for binding orientations 1* and 2* (Fig. 5.17B). This can be seen from enzyme binding pathways in which M37 exhibited initial interactions with the surface within the first 500 ns of CG-MD simulation time for all replicate simulations (Fig. 5.17B). However, additional simulations are necessary to confirm the potential difference observed in the enzyme-surface binding process for this graphene model relative to the 20% graphene oxide

model, for which 10 replicates were performed.

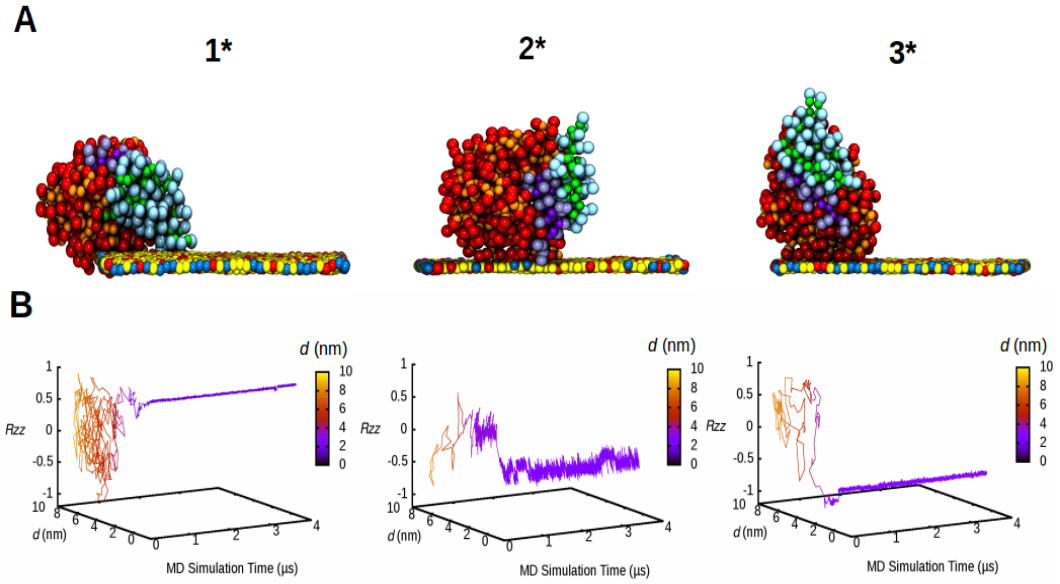

Fig. 5.17: **(A)** Three different binding orientations identified from CG-MD simulations of M37 interactions with the 23% graphene oxide model. Binding orientations are identified by 1*, 2*, and 3*. **(B)** Translation and rotational analysis performed for each of the 3 simulations. The rotation matrices (*Rzz*) are calculated relative to binding orientation 1 identified from simulations of M37 with the 20% graphene oxide model (Fig. 5.9) for comparative purposes. The same colour scheme is used as in Fig. 5.9. The graphene oxide surface is shown as yellow (carbon), red (epoxy/hydroxyl) and blue (carboxyl) van der Waals spheres.

The structural stability of M37 did not seem to be affected by the presence of charged groups at the edge of the oxidised sheets, displaying very similar RMSD profiles as calculated for M37 interactions with the uncharged graphene oxide surface (Fig. 5.18). Only small variations in the RMSD values are seen for the lid (residues 235-283) and active site flap (94-110) regions of M37, maintaining a constant RMSD for each of the three simulations performed.

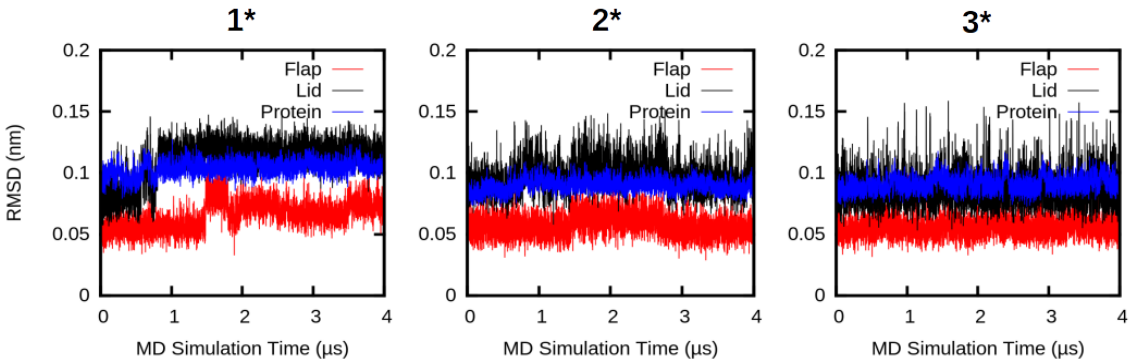

Fig. 5.18: Time evolution of the RMSD calculated for the backbone particles of the lid, active site flap, and rest of the protein structure during CG-MD simulations of M37 with a 23% graphene oxide model; M37 exhibited different three binding orientations through interactions with the sheet denoted by 1*, 2* and 3*, as shown in Fig. 5.17.

The different binding orientations exhibited by the M37 lipase however did differ in the nature of interactions mediating protein-surface contacts (Fig. 5.19). The enzyme predominantly displayed contacts with the basal plane of the graphene oxide sheet when adopting binding orientation 3*, mediating very little interactions with the carboxyl groups displayed on the edges (Fig 5.18C). Interestingly, in this simulation, M37 initially made contact with the edge of the sheet, and then diffused a small distance across the graphene oxide plane, resulting in an increased interfacial contacts mediating protein-surface interactions with the graphene oxide plane (Fig. 5.18).

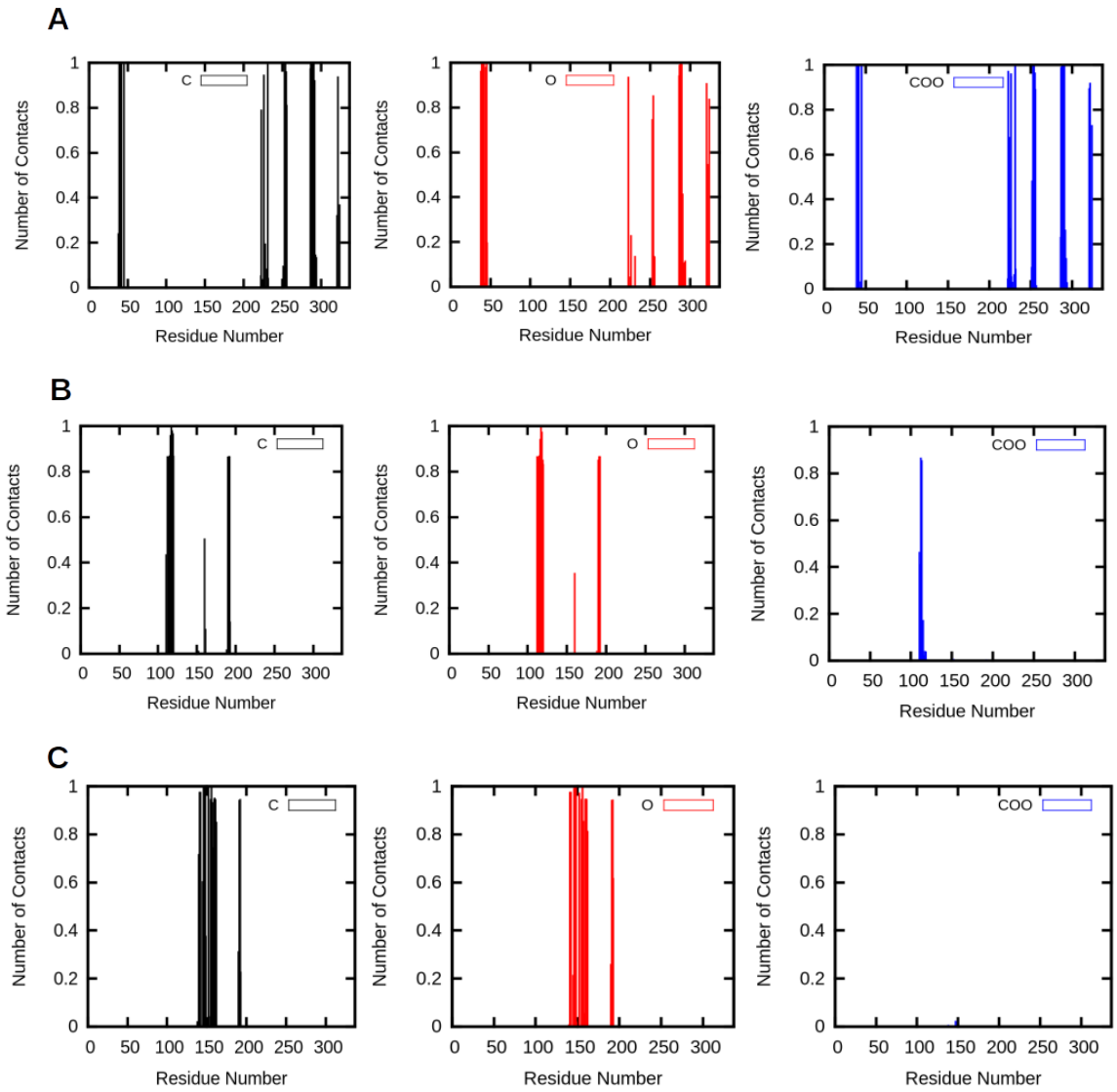

Fig. 5.19: Normalised contacts calculated for CG-MD simulations of M37 with a graphene oxide model displaying carboxyl groups at its edges (23%). Interactions were calculated for bindings orientations 1* **(A)**, 2* **(B)**, and 3* **(C)**, displayed in Fig. 5.17.

Conversely, M37 exhibited the most extensive interactions with the surface edge in through binding orientation 1*, as measured from protein residue contacts made with the carboxyl groups, in addition to interactions with the oxygen-containing beads and carbon beads on the surface (Fig 5.19A). The enzyme maintained its initial adsorbed

orientation, and remained at the edge of the surface for the duration of the simulation, and thus did not exhibit surface diffusion Fig 5.20A). Conversely, protein lateral diffusion on the graphene oxide surface was most prevalent in binding orientation 2*, where initial contacts made with the basal plane of the interface allowed the protein to diffuse laterally (t_{MD} = 0-0.5 μs), until additional protein-surface contacts were made at the surface edge, corresponding to a reduction in enzyme lateral mobility (Fig 5.20B). This behaviour is in line with observations of protein diffusion on the 20% graphene oxide surface, indicating a correlation between increasing protein-surface contacts and reduced protein diffusion on the interface.

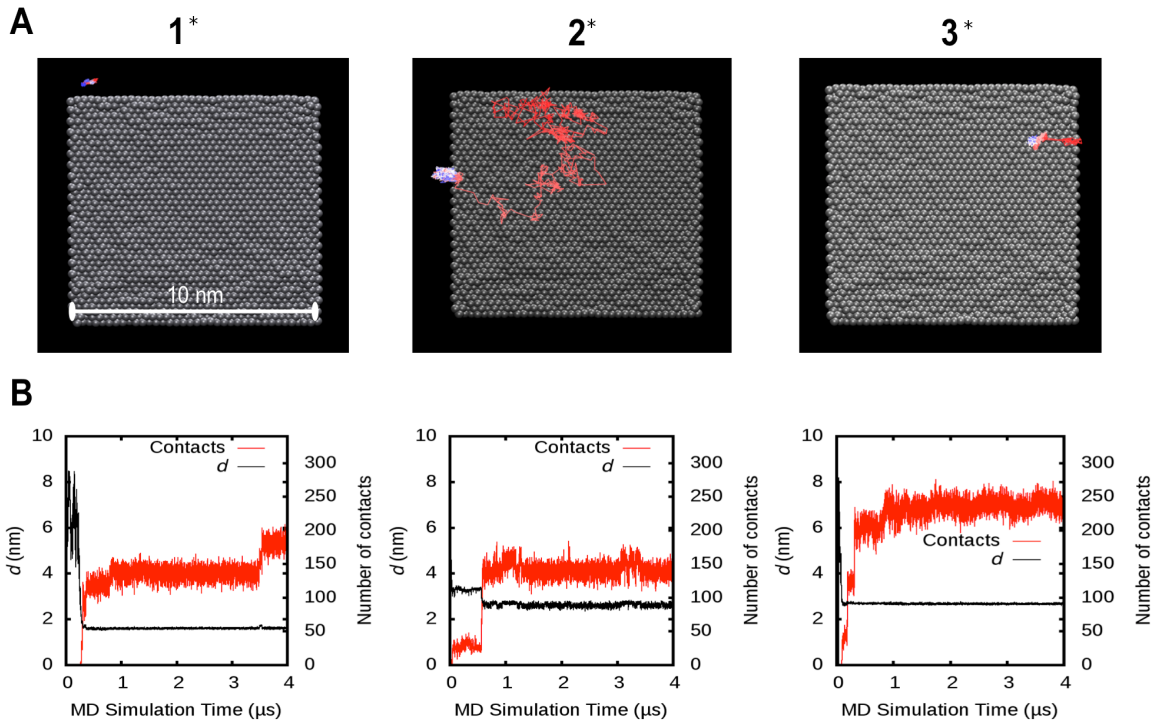

Fig. 5.20: **(A)** Time traces showing the position of M37 during CG simulations with a charged graphene oxide surface. The lines trace the COM of M37 once adsorbed onto the graphene oxide surface, visualised for CG simulations in which M37 adopts orientation 1*, 2*, or 3* (indicated by numbers). The time traces are coloured in a red-blue scale; red = the beginning of the trajectory, white the middle, and blue the end. Image perspective is from the top of the simulation box looking down onto the surface. **(B)** Time evolution of d (left axis), measuring the COM distance between the protein and graphene oxide, and number of contacts between the protein and the surface (right axis).

5.4 Discussion

The application of useful biocatalytic systems involving biological enzymes depends heavily on the nature of the support surface, such as its surface chemistry, as well as the structural nature of the enzyme, particularly its conformational dynamics and stability. Previous experimental studies of lipase interactions with graphene and graphene oxide-based supports indicate that lipases interact favourably with carbon nanomate-

rials, characterised by increased enzyme activity and catalytic efficiency (relative to soluble enzyme) [73, 212, 219, 221]. The molecular mechanisms that govern lipase adsorption and activation on these surfaces, and how this is affected by variation in the surface chemistry of graphene supports, however have only in part been characterised [92, 173]. Computational studies of the adsorption process of other enzymes (e.g lysozyme) on related surfaces suggest factors such as protein hydration and initial binding orientation are key in the adsorption process of enzymes on surfaces [245–247].

In this chapter, I used CG-MD methods to investigate the adsorption process of the M37 lipase on graphene and graphene oxide. Simulations were performed to study how this lipase interacts with graphene-based surfaces, characterising the interfacial interactions that govern lipase association. Additionally, the affect of surface charge on binding and interfacial orientation were investigated through simulation of a CG graphene oxide model displaying negatively charged carboxyl groups on the surface edges. The collective simulations were compared to simulations of M37 with anionic phospholipid bilayers, a "soft" biological surface (Chapter 3).

Preliminary results indicate that M37 preferentially interacted with graphene oxide rather than pristine (unfunctionalised) graphene. This was substantiated by performing positive control simulations with HFBI, a small protein known to bind graphene surfaces [237, 238]. Contrastingly, simulations of M37 with a graphene oxide surface indicated extensive interfacial interactions, resulting in identification of three different binding orientations (Fig. 5.9). Binding orientation 1 was similar to the *Angled* orientation exhibited by M37 through interactions with bilayer and triglyceride surfaces. The interactions governing M37 association with the oxidised surface were mainly mediated by polar and electrostatic interactions, contrasting the hydrophobic interactions observed HFBI simulations with pristine, hydrophobic graphene. Indeed, van der Waals interactions are known to be critical in protein adsorption to pristine

graphene surfaces, as has been characterised extensively both computationally and experimentally for other peptide/protein-graphene systems [220, 223, 242, 248, 249]. Similarly, previous investigations have also marked the importance of electrostatic and polar interactions of protein interactions with graphene *oxide* surfaces, matching the observations made for M37 [224, 248, 250].

Interestingly, two of the three binding orientations of the lipase on graphene oxide involved substantial interactions between the lid region of M37 and the surface, suggesting potential structural affects in terms of lid displacement and activation. The most commonly adopted binding orientation (orientation 1) was similar to the *Angled* orientation exhibited by M37 through interactions with phospholipid bilayer and triglyceride interfaces. One important difference however, was that M37 did not exhibit interactions through the active site flap region with graphene oxide, resulting in different protein-surface contacts compared to bilayer interactions (Fig. 5.13). Furthermore, a similar binding orientation to binding orientation 1 was identified in simulations with the charged graphene oxide model (1*), suggesting a common binding motif in M37-surface interactions involving the lid region.

Interestingly, lateral diffusion of M37 once adsorbed to these surfaces was only minimal and seemed to correlate with an increased number of protein-surface contacts, as exhibited by binding orientation 3 in simulations with 20% graphene oxide (Fig. 5.15). This suggests that protein-surface interactions were strong, and only minimal variation in binding orientation was observed once the lipase adsorbed onto the surface. Similar observations were made by a computational study of *Candida rugosa* lipase on a pristine graphite surface [173]. Furthermore, comparisons of the diffusive behaviour of M37 on graphene oxide surfaces with bilayer surfaces also suggested the protein was much more static on the surface, which again could be attributed to the nature and strength of the interfacial interactions between M37 and graphene oxide.

The functional consequences of M37 interactions with graphene oxide supports are difficult to predict from CG simulations, in which structural changes are not captured. Given that the potential activation mechanism of M37 involves displacement of the lid region, rather than the active site flap (Chapter 3), interfacial interactions involving the lid region with the graphene oxide supports may indeed affect the lipase activation process. If lid displacement were to occur through interfacial interactions with graphene oxide in, for example, binding orientation 1 (20% model), the underlying catalytic site would indeed become exposed to the surrounding solvent, potentially allowing substrate access. Conversely, previous simulation studies of a related lipase (Lip2) with carbon nanobtubes suggested that interactions of the protein through sites located away from the lid region were still able to result in displacement of the the lid region [92]. It was suggested that these interfacial interactions of the protein could be propagated to the remote lid subdomain, inducing a structural change [92]. It is therefore possible that any of the binding orientations observed in this study could lead to conformational changes in the functional regions of M37. Furthermore, the nature of substrate binding and substrate access in immobilised enzyme systems is currently not well understood or characterised, which is related to the difficulty in experimentally characterising the orientation of the bound enzyme molecules. Substrate access and entry into the cataltyic site would therefore be of interest to investigate for the M37-graphene oxide systems as well.

5.5 Conclusions

The nature and strength of the interactions between M37 and graphene oxide surfaces are clearly important in both interfacial binding orientation, and protein mobility on the surface. Unlike the M37-bilayer simulations, M37 was not observed to alter binding configuration once adsorbed to the surface, which may have implications for interfacial

activation of this enzyme. In order to study the conformational changes that may occur as a function of M37-graphene oxide interactions, a preliminary atomistic model of graphene oxide was parameterised (Appendix B). However, further development and validation studies are necessary before the model can be usefully simulated.

Furthermore, other, non-functional, structural affects might occur through extensive interactions of M37 with the planar graphene oxide surface, such as loss of secondary structure (e.g α-helicity). This has been observed for other enzyme-graphene systems [219, 222]. It would therefore be interesting to compare M37 interactions with carbon nanomaterials of different curvatures, such as carbon nanotubes (CNTs), which effectively correspond to a rolled up sheet of graphene. Indeed, previous experimental data suggests that lipase activity is affected by the topological differences between graphene (oxide) and CNTs, in which up to 20-fold increase in lipase activity is observed for functionalised CNTs relative to graphene oxide [219]. The difference in enzyme activity is thought to be related to structural denaturation of the enzyme through increased interactions with a flat graphene oxide surface, compared to the curved CNT surface, which has also been observed for other enzymes [219, 222, 223].

Other factors such as surface hydration and the formation of ordered water layers above planar graphene surfaces are thought to play an important role in the binding and orientation of adsorbed molecules [229, 251]. The ability of a CG water model to capture the dynamic processes that occur during protein adsorption that influence the surrounding water structure are limited [250]. Furthermore, the electronic structure of graphene is known to affect molecular adsorption, orientation, and interactions energy, where π-π stacking interactions (aromatic amino acids) and charge accumulation are very important [220, 251–253]. These effects are clearly not captured by the CG models presented here, and thus might affect the interfacial orientations observed for M37. The ability of the computational models of graphene/graphene oxide and MD

forcefields to predict this behaviour accurately will be important in generating relevant data, particularly in the characterisation of possible conformation changes of bound enzyme molecules. This will also be important in order to make useful comparison to the available experimental data.

Future studies employing atomistic MD simulations would be useful in elucidating both the structural and functional affects of lipase-surface interactions, both in terms of conformational stability of M37 and lipase activation. These would form a comparative basis for the possible activation mechanism suggested for M37 in Chapter 3, and more generally provide insight into enzyme binding and interfacial orientations at the interface in the context of experimental literature.

Chapter 6

Lipid Interactions With Graphene Surfaces

The interfacial interactions of the M37 lipase with graphene support surfaces were discussed in the previous chapter. In this chapter, these systems are extended to investigate how lipid molecules interact with graphene surfaces. Coarse-grained MD simulations were used to model phospholipid bilayers interactions with pristine graphene and graphene oxide surfaces. These results were compared to experimental measurements of the same systems, provided by colleagues (Dr. Michael Hirtz and Dr. Aravind Vijayaraghavan) at the Karlsruhe Institute of Nanotechnology and the University of Manchester. The simulation results agree well with the experimental data, suggesting that bilayer topology is significantly affected depending on both the chemical nature of the support surface, as well as system solvation. By establishing a model system exploring the behaviour of supported lipid membranes on these surfaces, the groundwork is laid for future investigations with lipases. These systems are amiable to experimental characterisation and thus constitute a basis for investigating how the properties of the supported interface are affected by lipase interactions, with possible further characterisation of lipase dynamics and activation mechanisms. Hence, this

integrated study represents a basis for future simulations of supported bilayer systems with lipases.

6.1 Introduction

Supported lipid membranes (SLMs) are generated by the deposition of lipid molecules on a support surface to investigate molecular properties such as lipid organisation, protein interactions, and overall membrane dynamics [101]. A recent approach to generating SLMs is lipid dip-pen nanolithography (L-DPN). This technique involves covering the tip of an atomic force microscope (AFM) with a phospholipid species or mixture (the "ink"), and directly depositing the molecules onto a support surface (the "paper") (Fig. 6.1) [254]. L-DPN has the benefit of direct and precise spatial control during lipid deposition, with the ability to tailor the lipid ink to the desired composition [255, 256]. The nature of the support surface, amongst other factors, plays an important role in the observed molecular properties of SLMs [103, 257, 258].

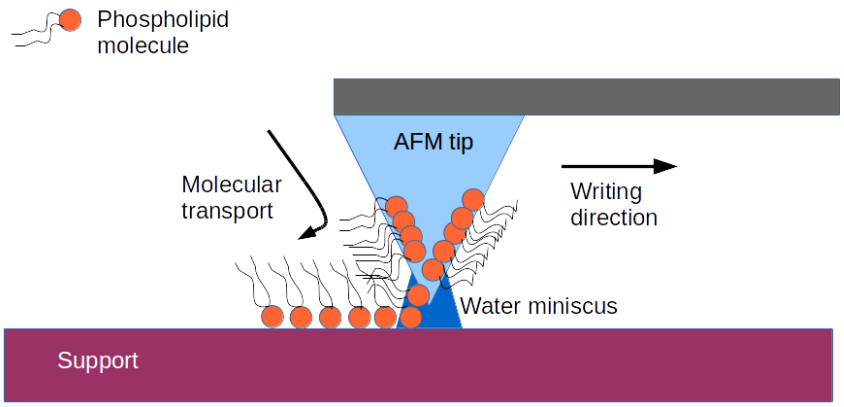

Fig. 6.1: Schematic of the basic molecular transport in lipid dip-pen nanolithography (L-DPN) to generate supported lipid membranes (SLMs). The tip of an atomic force microscope is covered with lipid "ink", represented by orange circles (headgroups) and black lines (tails), which comes into contact with a support surface via a water miniscus that forms between the tip and the surface, allowing molecular transport of lipid molecules onto the surface. The AFM tip is moved across the surface in a particular writing direction, generating lines on the surface.

Recently, L-DPN was used to investigate the molecular organisation of different lipid inks on pristine graphene and silicon dioxide surfaces [259]. It was found that the molecular topology of SLMs on graphene differed substantially to lipid formations on silicon dioxide in air. Furthermore, the addition of aqueous solution to the graphene systems resulted in significant re-organisation of the lipid structures, suggesting important roles for hydration in stabilisation of lipid membranes on the support. The authors further demonstrate the binding of small proteins to functionalised lipids, confirming membrane stability, as well as biosensing capabilities of the SLM systems [259]. These results describe a modular, discretely patterned system in which molecular interactions between different biological compounds, for example lipids and proteins, can be studied.

Here, I utilised CG-MD simulations to investigate how phospholipid bilayers interact with pristine graphene (unfunctionalised) and graphene oxide surfaces in vacuum

and in water. The CG-MD results were compared to experimental measurements, including atomic force microscopy (AFM) studies, of the same systems. These correlated well, observing similar effects on membrane topology in function of surface polarity and system hydration [259]. Further characterisation of the molecular properties of the SLMs indicated substantial effects on lipid diffusion within the supported bilayers, as well as lipid ordering, in comparison to free-standing bilayers. These simulations constitute initial characterisation of lipid-graphene systems, with a view towards further investigation of protein interactions with SLMs, focussing on lipases.

6.2 Methods

6.2.1 CG Surface Models

The CG graphene model presented in Chapter 4 was utilised in this chapter for simulations of phospholipids with pristine graphene. Briefly, the model involved a 2:1 mapping scheme of carbon atoms to CG graphene beads (SG4 bead type) (Fig. 6.2) [84]. Original parameterisation of the graphene model was based on reproducing the adsorption and topological behaviour of long-chain alkane molecules on graphite, suggesting the model is suitable to study the behaviour of lipids on graphene surfaces [84].

The CG graphene model was used to build the graphene *oxide* model. To do this, the carbon beads (SG4) comprising the pristine graphene surface were randomly substituted by oxygen beads (SP1), which represent either a C-O-C epoxy group or a C-O-H group [122], eventually reproducing the oxygen content of the graphene oxide substrates used in the experiments (2.82% C-O-C: 49.3% C-O-H based on X-ray photoelectron spectroscopy data).

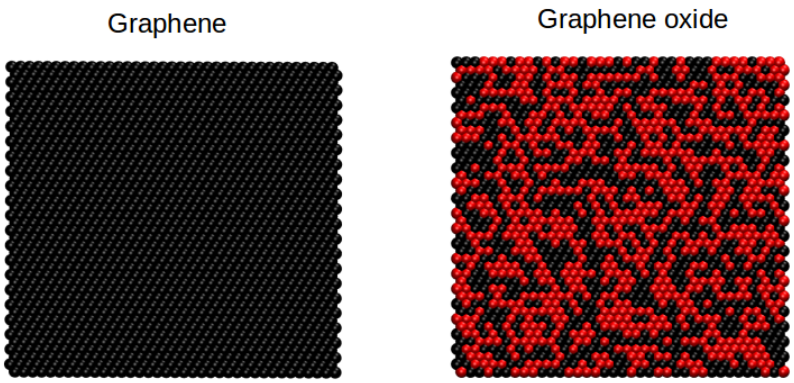

Fig. 6.2: CG graphene and graphene oxide models used in this Chapter. The graphene oxide model reflects the same oxidation percentage as the graphene oxide flakes used by experimental collaborators. The surfaces are shown as van der Waals spheres; carbon beads (SG4) graphene beads are coloured black, oxygen beads (SP1) are coloured red.

6.2.2 Lipid Bilayers

Lipid bilayers consisting of 1,2-dioleoyl-sn-glycero-3-phosphocholine (DOPC) molecules were constructed by self-assembly simulations in water using the same method presented in Chapter 3. The DOPC lipid composition reflects the main lipids employed in the experiments and previous, related studies [259, 260]. The self-assembled bilayers were used to build inverted bilayers that mimic structures observed in previous AFM studies of lipid-graphene interactions [259]. The inverted bilayers were obtained by isolating the bottom leaflet of a regular bilayer and placing this on top the the upper leaflet, such that the lipid headgroups were oriented towards each other (Fig. 6.3). The *editconf* tool from the GROMACS simulation package was used to do this.

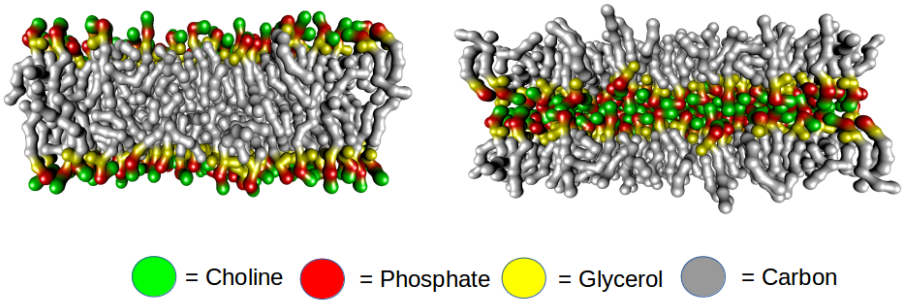

Fig. 6.3: Regular (*left*) and inverted (*right*) bilayers used for CG-MD simulations with graphene and graphene oxide surfaces. Each bilayer is composed of 512 DOPC molecules. The lipid choline, phosphate, glycerol and carbon tail groups are shown as licorice representations coloured in green, red, yellow, and grey respectively. The Hyperballs program was used to generate the images [261].

6.2.3 Simulation Details

The Martini 2.2 forcefield was used for all CG-MD simulations [122]. All systems containing water were performed with the polarisable Martini water model [148]. Simulations of the graphene (oxide)-lipid systems in vacuum were conducted at constant temperature and volume (NVT ensemble), applying the Berendsen thermostat with coupling constant of 0.3 ps at 298 K [120]. Fully solvated simulations were performed at constant temperature and pressure (NPT ensemble) utilising the Berendsen barostat with a coupling constant of 3.0 ps, applying anisotropic pressure coupling, and a compressibility of 0.5×10^{-5} bar^{-1} in x and y, and 3.0×10^{-5} bar^{-1} in the z-direction [120]. Non-bonding interactions were modelled using shift functions; both LJ and Coulombic interactions were evaluated within a 1.2 nm cut-off. The LJ interactions were shifted within a 0.9 nm cut-off distance, the Coulombic interactions were shifted from a 0.0 cut-off distance. These parameters reflect those applied in parameterisation studies of the CG graphene model [84]. The lipids, water, and graphene were coupled to separate external baths for all of the simulations.

6.2.4 Lipid Order Parameter Analysis

The 2nd rank lipid order parameters were calculated for select simulated systems using the following expression:

$$S_n^{CC} = \frac{1}{2}\langle 3cos^2(\theta) - 1\rangle = \langle P_2(cos(\theta))\rangle \tag{6.1}$$

where θ is the angle between the bond connecting carbon beads B_{n-1} and B_n and the z-axis (normal to the bilayer). The angle brackets indicate the ensemble average. The 2nd rank term refers to the second order Legendre polynomial used to describe the order parameter [262]. This approximation has been used in experimental and computational studies of bilayer systems [263–266], and can be compared to S_{CD} order parameters calculated from NMR deuterium experiments using the recursion relation [267].

6.2.5 Lipid Diffusion Analysis

The diffusion of lipids within the simulated systems was analysed using documented open-source code (http://dx.doi.org/10.5281/zenodo.11827). The code employs an algorithm which calculates the mean square displacement values of individual lipid centroids over a range of time sampling windows, including: 1, 2, 5, 10, 20, 50, 70, 100, and 200 ns. The diffusion coefficients are then calculated by fitting the MSD *vs* time data using a linear diffusion equation [240], or an anomalous diffusion equation [268]. As mentioned in the Methods section of Chapter 5, fitting of the MSD data using an anomalous diffusion equation distinguishes this code from the GROMACS tools function *g_msd*, which only performs linear fitting of the MSD data. For this reason, the open source code was used rather the *g_msd* tool. The linear approximation uses a least squares first degree fit of the data, whereas the anomalous approximation uses

a non-linear least squares fitting to a two parameter equation of the form:

$$MSD = 4D_{\alpha}t^{\alpha}, 0 < \alpha < 2 \tag{6.2}$$

where D_{α} is the fractional diffusion coefficient, measured in units of length^2/time$^{\alpha}$.

For the linear diffusion fit, standard deviations of the diffusion coefficients were estimated as the difference between the slopes from the first and second halves of the MSD *vs* time data. For the anomalous diffusion fit, the scaling exponent (α) was estimated as the slope of the log MSD *vs* log time data. The standard deviations of both parameters (D_{α} and α) were calculated from the square root of the diagonal of the covariance matrix from the anomalous fit. This code has been used by previous simulation studies of lipid diffusion in virus particles [128, 129].

6.3 Results

Previous studies have shown the formation of stable lipid structures, such as multi-layered lipid membranes, on graphene, silicon dioxide, and other substrates can be generated by L-DPN, and are influenced by surface morphology and experimental conditions [103, 255, 259]. Here, CG-MD simulations were used to investigate the molecular details of supported lipid membrane interactions with pristine graphene and graphene oxide surfaces. The simulated systems were constructed to reproduce the experimental conditions used during generation and analysis of the supported membranes by L-DPN and AFM [259]. These experiments were either performed for systems in air or systems in solution. Experiments in air were typically performed at 20-30% relative humidity (R.H). The simulated systems were however of a much smaller size, resulting in a negligible number of water particles (< 1) occupying the volume of the box at 20-30% R.H, even for the largest systems (30x20x20 nm). Exper-

iments in air were therefore approximated by simulations in vacuum. Furthermore, in all cases, the x and y-axes of the simulation box were larger than the total width and length of the graphene/graphene oxide surface area. This was to ensure there were no periodic boundary effects regarding lipid interactions with the surrounding period images.

6.3.1 Lipid Interactions with Graphene

L-DPN generated lipid membranes on pristine graphene surfaces are thought to form "inverted" bilayer structures in air [259, 269]. I investigated the stability of these inverted bilayer structures on a graphene surface in vacuum and in water. Two different graphene surface areas were modelled: a small (10x10 nm) and large (30x30 nm) surface. Pre-formed inverted bilayers consisting of 512 DOPC lipids (256 lipids per leaflet) were placed above the individual small and large graphene surfaces and simulated for 0.5-1 μs.

Simulations of the systems in vacuum resulted in rapid adsorption of the lipid layers onto the surface, mediated by strong hydrophobic interactions between the graphene beads and the lipid tails (Fig. 6.4A). The lipid structures maintained their inverted configurations as shown by monitoring the centre of mass (COM) distance between the lipid phosphate headgroup and the graphene surface (Fig. 6.4E). Polar interactions between the headgroups resulted in clustering of the headgroups, stabilising the inverted bilayer topology. Encouragingly, CG-MD simulations of lipid monolayers with the small graphene surface resulted in spontaneous lipid re-organisation to form an inverted bilayer structure in vacuum (Fig. 6.3C). Furthermore, similar partial densities of the lipid groups can be observed for the inverted bilayer and monolayer simulations, suggesting the different starting structures converged to similar end structures (Fig. 6.3B and 3D). The CG results corroborate AFM measurements of L-DPN deposited

lipid layers on pristine graphene surfaces, revealing that phospholipids form a flat inverted bilayer of uniform thickness on the hydrophobic support in vacuum (simulation) or in air (experiment) (Fig. 6.5A) [259]. The AFM measurements were performed by Dr. Michael Hirtz and colleagues at the Karlsruhe Institute of Technology.

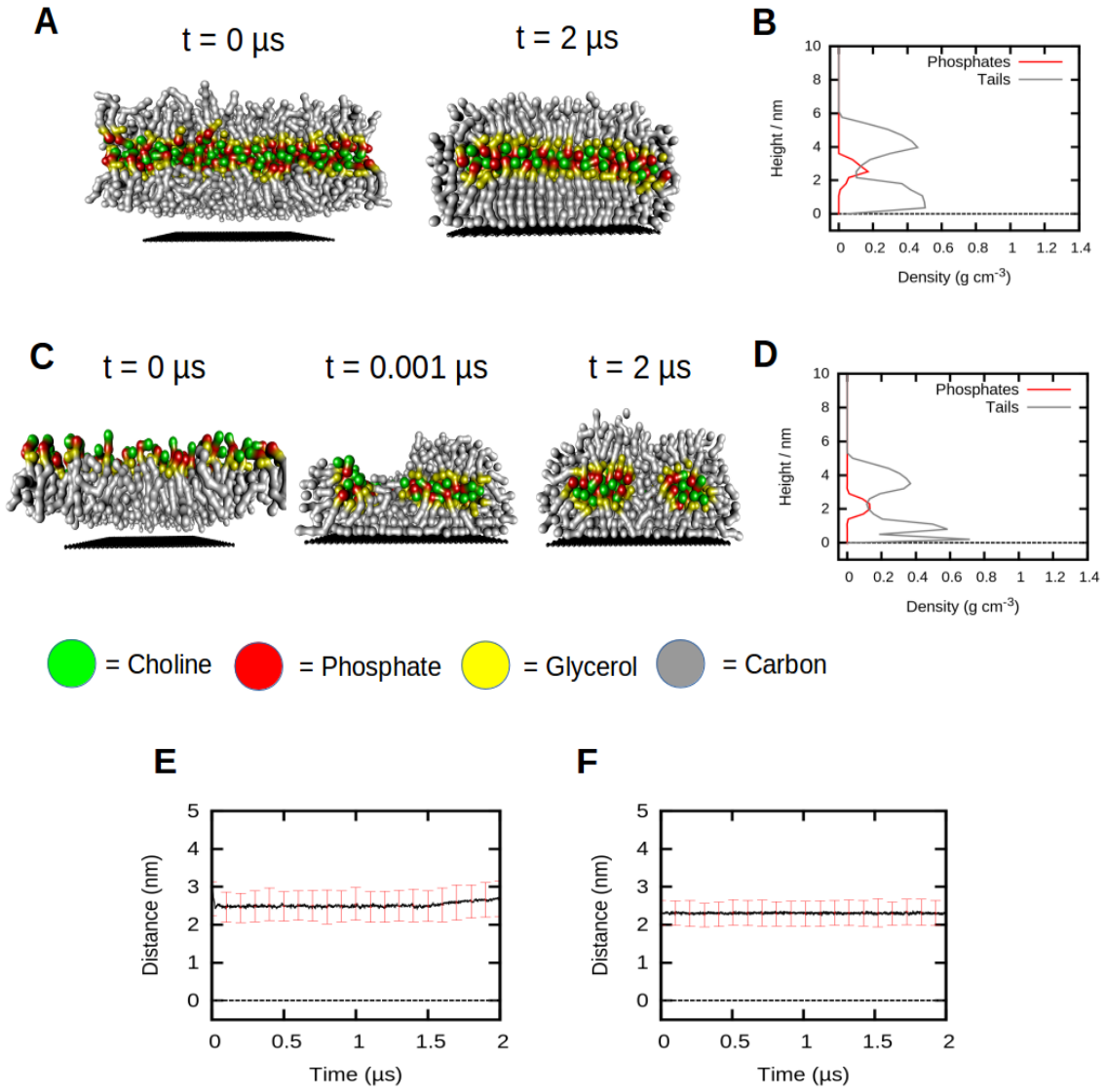

Fig. 6.4: Images of CG-MD simulations of a pre-formed inverted bilayer (**A**) and lipid monolayer (**C**) with a small (10x10 nm) graphene surface in vacuum. The lipid choline, phosphate, glycerol and carbon groups are shown as licorice representations in green, red, yellow, and grey respectively. Graphene is represented as black licorice. The partial densities of phosphate headgroups and carbon tails were calculated for the last 25% of simulation time for the inverted bilayer (**B**) and monolayer simulations (**D**). The dashed line represents the position of graphene in the simulation box. Lipids are only shown from the graphene surface upwards. However, some lipids occasionally transferred to the bottom of the graphene surface. Given that these were only a few lipids, they are not included in the images. Time evolution of the average centre of mass (COM) distance between lipid phosphate headgroups and graphene surface (dashed line) for the same bilayer (**E**) and monolayer simulations (**F**).

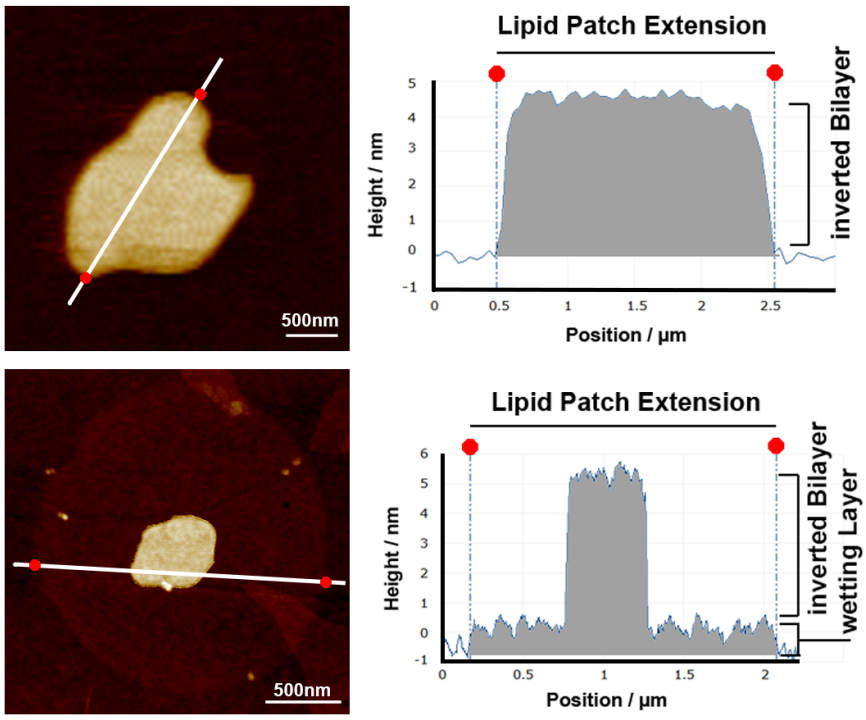

Fig. 6.5: *Left*: AFM images of L-DPN generated lipid membranes (DOPC) on pristine graphene (**A**) and graphene oxide (**B**) surfaces in air. *Right*: AFM height measurements of the same patches, measured between the red dots shown on the left images. The top section profile shows a smooth single step membrane compatible with the thickness of a single inverted bilayer on graphene (in air) [259]. The bottom profile section reveals a bilayer membrane on top of a monolayer (wetting layer). The wetting layer is similar to lipid layers observed on silicon dioxide, though thinner, probably due to reduce layer density [270]. This data was collected and provided by colleagues (Dr. Michael Hirtz) at the Karlsruhe Institute of Nanotechnology.

CG-MD simulations of inverted bilayers with the large graphene surface (30x30 nm) however, were less stable regarding bilayer configuration. Here, the strong hydrophobic interactions between the lipid tails and graphene surface dominated lipid organisation, and resulted in disassembly of the layer and spreading of lipids across the available surface area, aligning with the support in a lateral fashion (Fig. 6.6A). The same behaviour was also observed in simulations of pre-formed regular bilayers (Fig. 6.6B). Interestingly, simulation of larger inverted bilayers (2110 lipids) exhibited

higher stability and maintained the inverted configuration to a larger extent compared to the smaller bilayer systems (512 lipids) (Fig. 6.6E). However, lipid tail entanglement and lateral interactions with the surface, as well as headgroup-headgroup clustering, resulted in deviation from the more "ideal" inverted bilayer configuration observed in simulations of the smaller bilayers on the small graphene surface in vacuum (10x10 nm) (Fig. 6.4). Importantly, the increased stability of larger lipid layers indicates that lipid density on the graphene surface is another determinant of supported membrane topology.

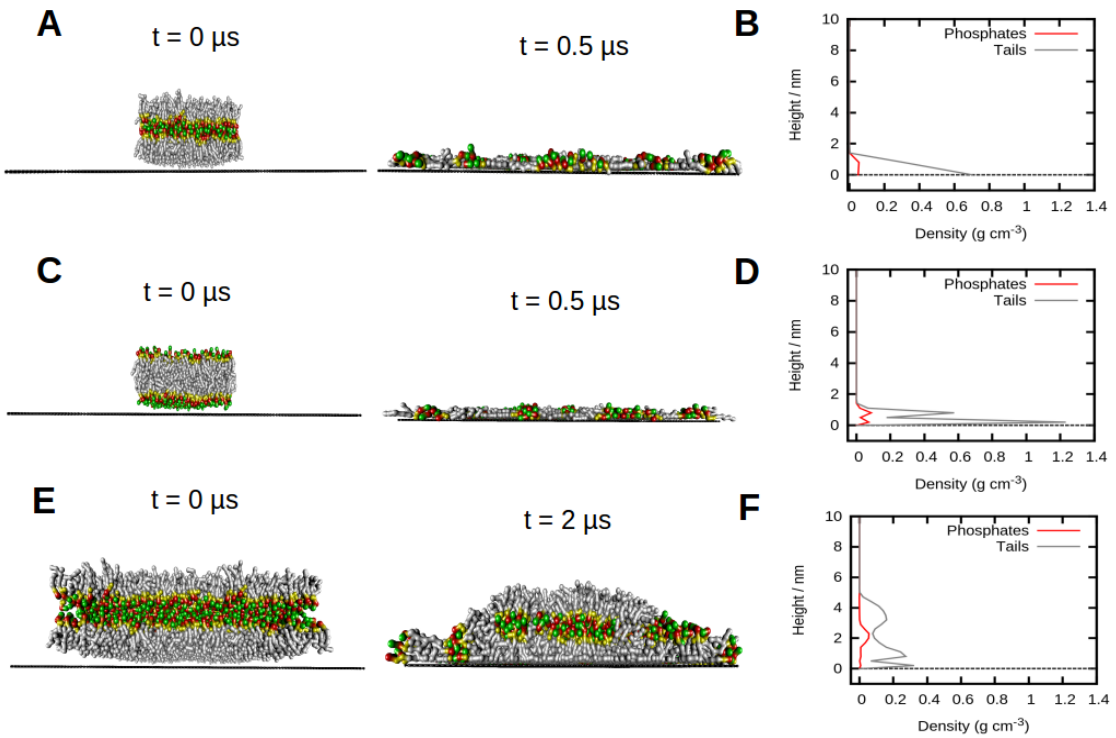

Fig. 6.6: *Left*: Images of the first and last frames of CG-MD simulations of small pre-formed inverted (**A**) and regular (**C**) bilayers (512 lipids), and a larger inverted bilayer (**E**; 2110 lipids) on large graphene in vacuum. The same colour code as images in Fig 6.4 is used. *Right*: Partial densities of the phosphate and lipid tail groups calculated for the small inverted (**B**), regular (**D**) and large inverted bilayer (**F**) simulations. The x- and y-axes of the simulation box are larger than the total width and length of the graphene surface, in order to minimise periodic effects through lipid interactions with the period image. Lipids are also only shown from the graphene surface upwards. However, some lipids occasionally transferred to the bottom of the graphene surface. Given that these were only a few lipids, they are not included in the images.

6.3.2 Lipid Interactions with Graphene Oxide

Surface polarity is known to affect the molecular properties of supported lipid bilayers, such as overall topology, diffusion, and lipid ordering [107, 269, 271–273]. In order to compare lipid behaviour on graphene surfaces to other surfaces, L-DPN was also performed on graphene oxide substrates in air. AFM measurements of the deposited lipid layers on the hydrophilic substrate indicated that the lipids organised into a "1.5"

bilayer configuration (Fig. 6.5B), as has been suggested for lipid structures on silicon dioxide surfaces in air [259, 270, 274]. Two distinct lipid layers could be distinguished: a "wetting" layer composed of phospholipids with their headgroups oriented towards the hydrophilic support surface, and a second inverted bilayer that formed on top of this wetting layer (Fig. 6.5B).

The same lipid organisation is seen in CG-MD simulations of pre-formed regular bilayers with graphene oxide in vacuum. The lipids rapidly interacted with the surface, undergoing substantial re-organisation to form a 1.5 bilayer on top of the oxidised surface (Fig. 6.7A). This rearrangement is thought to be driven by polar headgroup interactions with the hydrophilic graphene oxide support, and is further stabilised by hydrophobic interactions between the lipid tails. Encouragingly, simulation of a pre-formed *inverted* bilayer configuration positioned above the oxide surface also converged to a similar 1.5 bilayer structure (Fig. 6.7B). The initially inverted structure underwent significant rearrangement to facilitate formation of the 1.5 bilayer structure, resulting in very similar partial density profiles calculated for the initially inverted and regular bilayer simulations (Fig. 6.7C). The measurements confirm that the differing starting structures were able to converge to very similar end structures. These results corroborate the interpretation of a 1.5 bilayer configuration by AFM height measurements for supported lipid membranes on graphene oxide, as well as silicon dioxide supports in air, suggesting that this arrangement is the preferred molecular state of lipid layers on oxide surfaces in air [259, 270, 274].

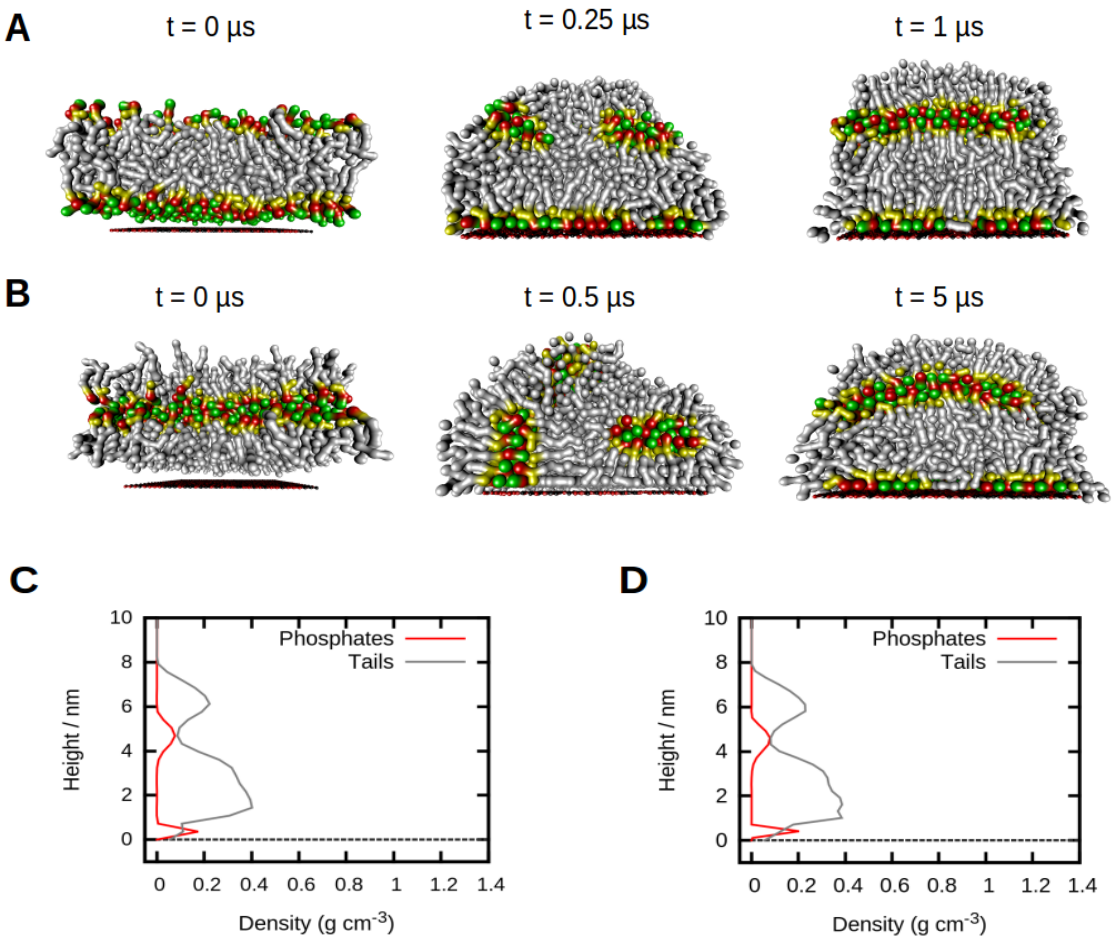

Fig. 6.7: Images of a CG-MD simulation of a pre-formed regular lipid bilayer (**A**) or inverted bilayer (**B**) on a small (10x10 nm) graphene oxide surface in vacuum. Partial densities of the lipid tail and phosphate headgroups were measured for the last 25% of the simulation time, shown for the regular (**C**) and inverted (**D**) simulations. The dashed line represents the position of the graphene oxide surface in the simulation box. The lipid choline, phosphate, glycerol and carbon groups are shown as licorice representations in green, red, yellow, and grey respectively. Graphene oxide is represented as black (carbon) and red (oxygen) licorice. Only the lipids interacting with the top of the graphene surface are shown. Very few lipids transferred to the bottom of the graphene oxide surface during the simulations and are thus not shown.

Interestingly, the 1.5 lipid bilayer configuration was not observed in CG-MD simulations of lipid-graphene oxide systems in water. Instead, pre-formed lipid structures reorganised to form bicelle-like bilayer configurations on both small and large graphene

oxide surfaces in water (Fig. 6.8B). Similar to previous observations, the rearrangement of the pre-formed inverted bilayers was driven by lipid headgroup interactions with the underlying support, either directly or through bridging water particles, as well as by interactions with the surrounding solvent, resulting in very stable bicelle structures with little overall surface diffusion.

Conversely, the addition of water to lipid-graphene systems in the CG-MD simulations resulted in destabilisation of the previously formed inverted bilayer, and triggered disassembly of the structure (Fig. 6.8A). Consequently, the lipids spread over the available surface area with headgroups oriented toward the surrounding solvent molecules, suggesting that the disassembly is most likely triggered by favourable polar interactions between lipid headgroups and the solvent, in addition to hydrocarbon tail interactions with the surface. Importantly, the same behaviour is observed for lipid-covered graphene transferred to aqueous solution by AFM measurements [259].

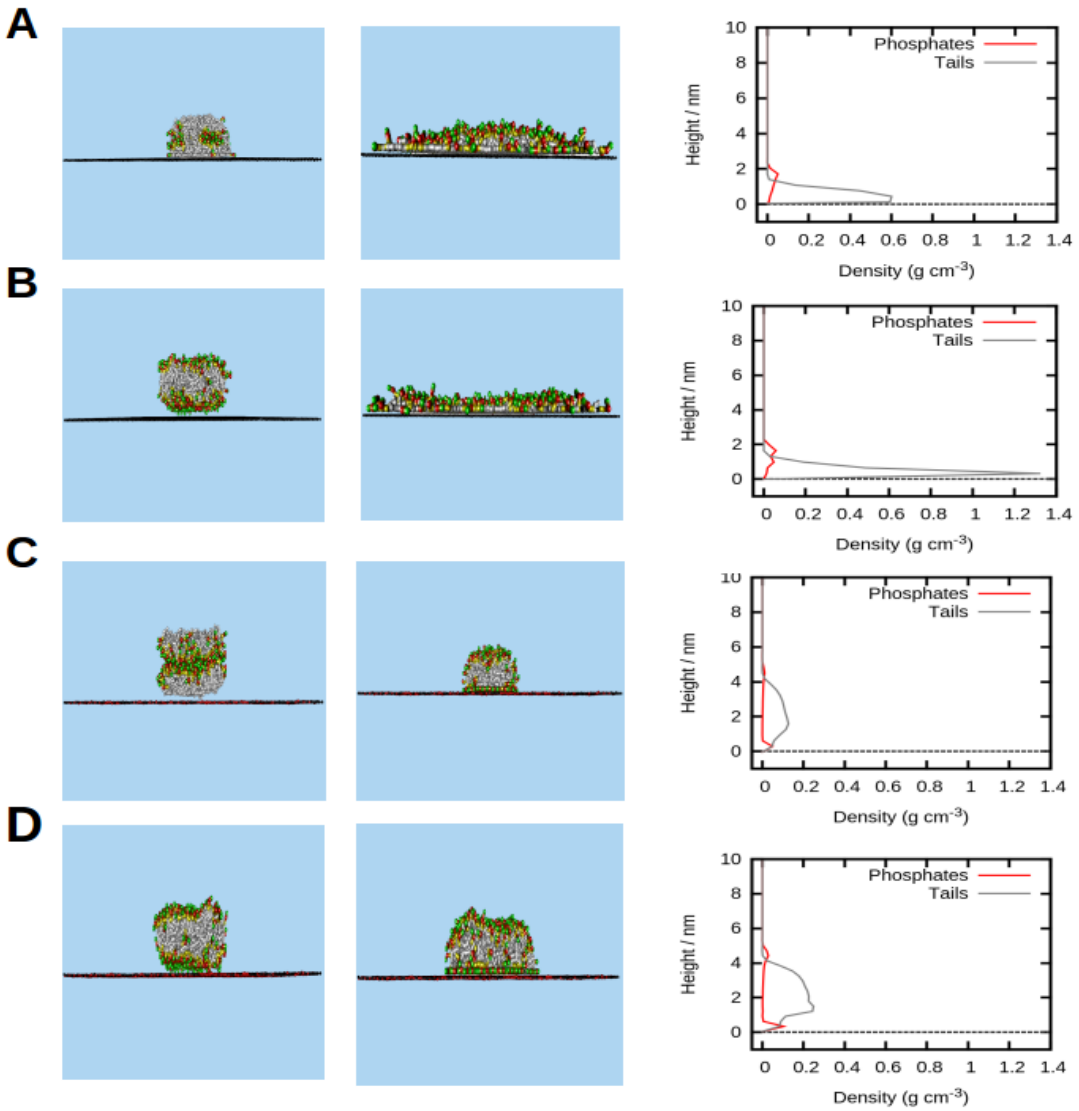

Fig. 6.8: Images of the first and final frames of a CG-MD simulations of inverted and regular bilayers on large graphene (**A** and **B**) and graphene oxide (**C** and **D**) surfaces in water. Water is represented as a light blue background. The colour scheme for the system components is the same as in Fig. 6.7. The right panel shows partial density profiles calculated for the last 25% of simulation time for the same simulations shown in the left images. Only lipids remaining on top of the graphene/graphene oxide surfaces are shown.

Two observations can be made from the collective simulation and experimental data. First, the addition of aqueous solvent can result in altered lipid configurations on the same substrate surface e.g bicelle-like formations on graphene oxide in

an aqueous environment relative to a 1.5 bilayer configurations in air (or vacuum) [259, 270, 274]. The different configurations in solution were stabilised by interactions of the lipid headgroups with the surface and the surrounding solvent, suggesting a thermodynamic component to the generation of different membrane topologies [274]. Second, the polarity difference resulting from the oxygen-containing functional groups in graphene oxide was able to stabilise lipids layers in different configurations relative to the hydrophobic graphene surface (e.g inverted bilayers *vs* 1.5 layers). This implies that supported bilayer configurations can be tuned as a function of system solvation as well as underlying surface polarity, resulting in relatively different end structures.

Direct observation of lipid re-organisation on graphene and silicon dioxide in liquid by AFM was reported by a previous study [259]. These observations necessitated the use of bovine serum albumin (BSA), which binds to the support surface, to block excessive lipid spreading and stabilise the membrane patches in aqueous solution. Here, direct AFM measurements of the lipid layer structures on graphene oxide in liquid has been hindered by the blocking layers of BSA needed to stabilise the small membrane patches during scanning. Given that specific binding interactions in L-DPN generated SLM systems usually occur through functionalisation of the lipid headgroups, it was inferred that lipid membrane re-organisation took place on graphene oxide, exposing the otherwise buried functional lipid headgroups to the aqueous phase [275].

6.3.3 Supported Membrane Properties

Lipid Order

So far, I have given an overview of the topological differences of supported lipid structures as a function of interactions with surfaces of varying polarity as well as system solvation, but how does this affect their molecular properties? Lipid tail order

parameters may provide a metric to track lipid rearrangements on the support surface (see Methods). Order parameters are a measure of lipid ordering within a layer, calculated by estimating the angle between the bonds connecting the lipid particles and the z-axis of the simulation box (taken to be the bilayer normal). Full alignment of the lipid particles with the z-axis gives an order parameter value of 1, whereas anti-alignment gives a value of -0.5, and 0 represents randomly orientated lipids. Order parameters were calculated for the simulated lipid molecules (B_n-B_{n+1} bonds, where B represents a bead) for selected systems and compared to simulations of a free-standing bilayer in water.

Systems in which substantial lipid re-organisation was observed resulted in changes in the lipid order parameters over time, particularly for inverted bilayer simulations with the graphene oxide surface in vacuum (Fig. 6.9). In this system, the average order parameter values change significantly during the first 2 μs of simulation time, particularly for the phosphate-glycerol bond, and consecutive bonds for both sn-1 and sn-2 lipid tails up to the third bead. The changes represent altered configurations of the lipids as the molecules re-organised to form the 1.5 bilayer structure. Specifically, the average order parameters for particular lipid bonds decrease from around 0.3 to 0-0.1 during the first μs of the simulation time, indicating a completely random arrangement for many lipids as these re-organised to form an inverted bilayer on top of the monolayer (wetting layer). Towards the end of the rearrangement ($\sim$2 μs), the average order parameters for e.g the phosphate-glycerol bond returns to 0.3, suggesting better alignment with the z-axis after the transition (Fig. 6.9). Other order parameter values, such as the C2-C3 bond however, do not completely recover their initial value, attributed to the level of bilayer distortion brought about by micelle-like lipids within the bilayer formed on top of the wetting layer (Fig. 6.7A). This value ($\sim$0.1) is reflected by subsequent bonds (e.g C3-C4) in both sn-1 and sn-2 chains, reflecting the random

arrangement of lipids tails within the 1.5 bilayer structure on graphene oxide (Fig. 6.11C).

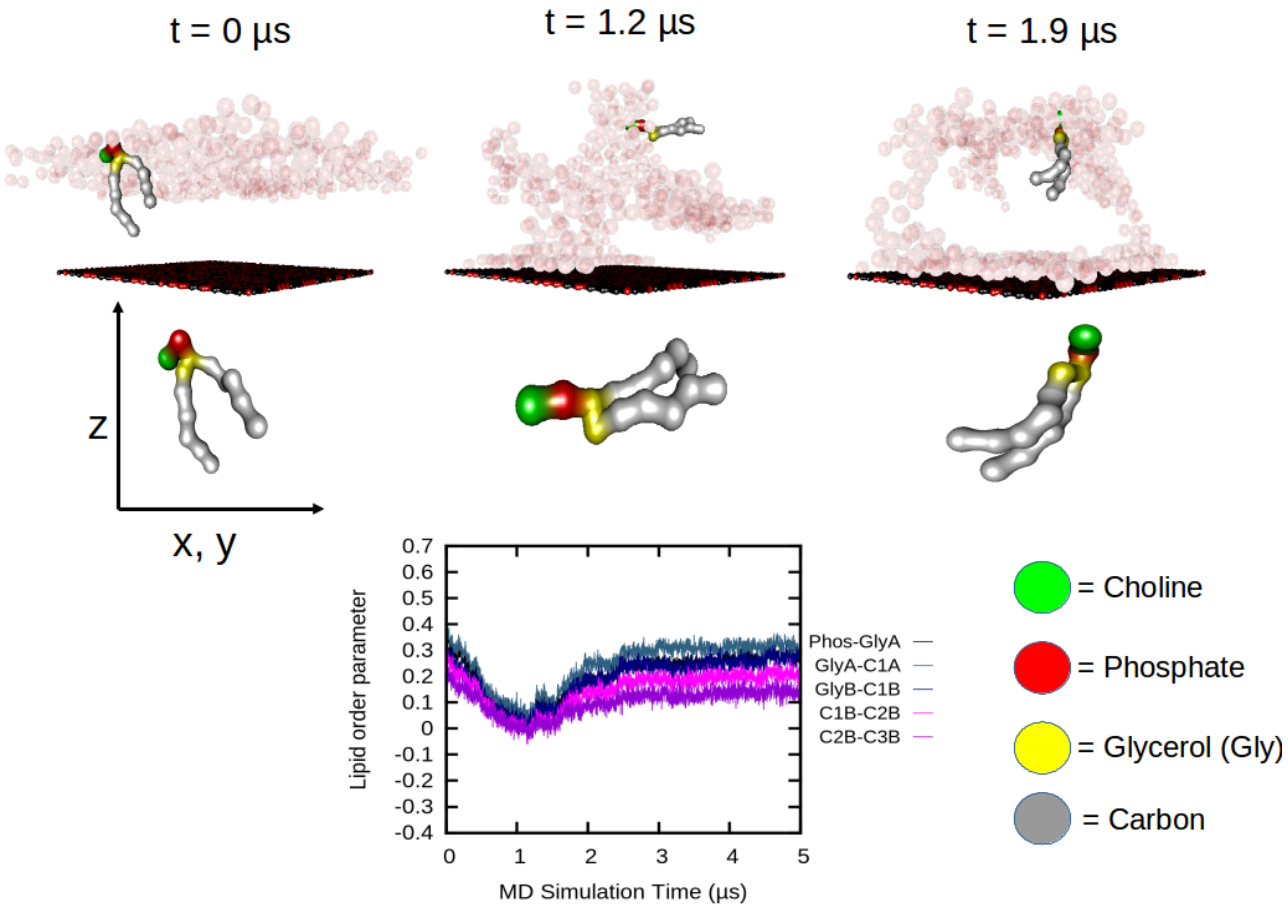

Fig. 6.9: Lipid order parameters calculated for B_n-B_{n+1} bonds, where B represents a bead, from a CG-MD simulation of an inverted bilayer with small graphene oxide in vacuum. Time evolution of order parameters for selected bonds are shown in the bottom panel. The top panel shows the position of all the lipid headgroups in the system (grey van der Waals spheres), as well as the position a single lipid in the layer, at different time frames during the simulation. The second panel shows the alignment of this single lipid with the z-axis of the simulation box, during lipid reorganisation. The same colour scheme is used as in Fig 6.7.

In general, the average lipid tail order parameters for all systems suggest that the tails were less ordered on both graphene and graphene oxide compared to a free-standing bilayer in water (Fig. 6.10). For example, an average order parameter of 0.18 is calculated for lipids in an inverted bilayer-graphene simulation in vacuum (small

surface area), compared to 0.36 calculated for lipids within the free-standing bilayer.

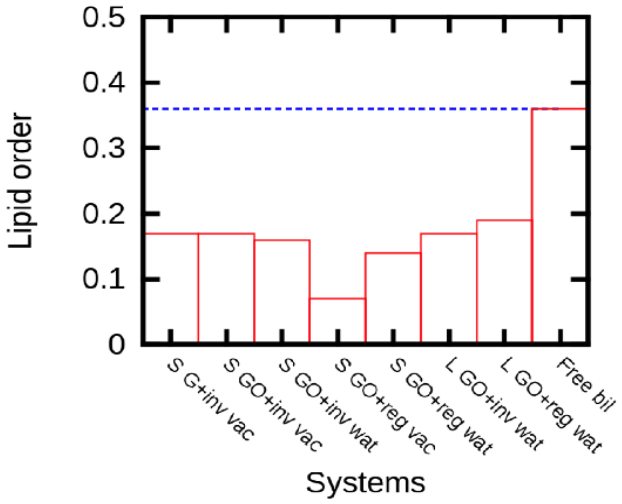

Fig. 6.10: Average lipid order parameters calculated for the simulated systems of lipid-surface interactions, compared with a free-standing bilayer in water. S = small, L = large, GO = graphene oxide, G = graphene, inv = inverted bilayer, reg = regular bilayer, vac = vacuum, wat = water. The dashed blue line indicates the average order parameter calculated for lipids within the simulated free-standing bilayer.

However, inspection of the time evolution of the order parameters from the inverted bilayer-graphene simulation reveal a slight gain in order within the last 0.5 μs of simulation time, particularly for the sn-2 chain of the lipids (e.g C1B-C5B) (Fig. 6.11A). This transition indicates the increase in alignment between the lipid tails, as the inverted bilayer configuration is stabilised on the pristine graphene surface, as a function of the lipid density on the surface. Furthermore, a good overlap between the order parameters for simulations of different starting structures with graphene oxide can be observed, indicating convergence to the similar 1.5 bilayer end structure on this surface in vacuum (Fig 6.11C and D).

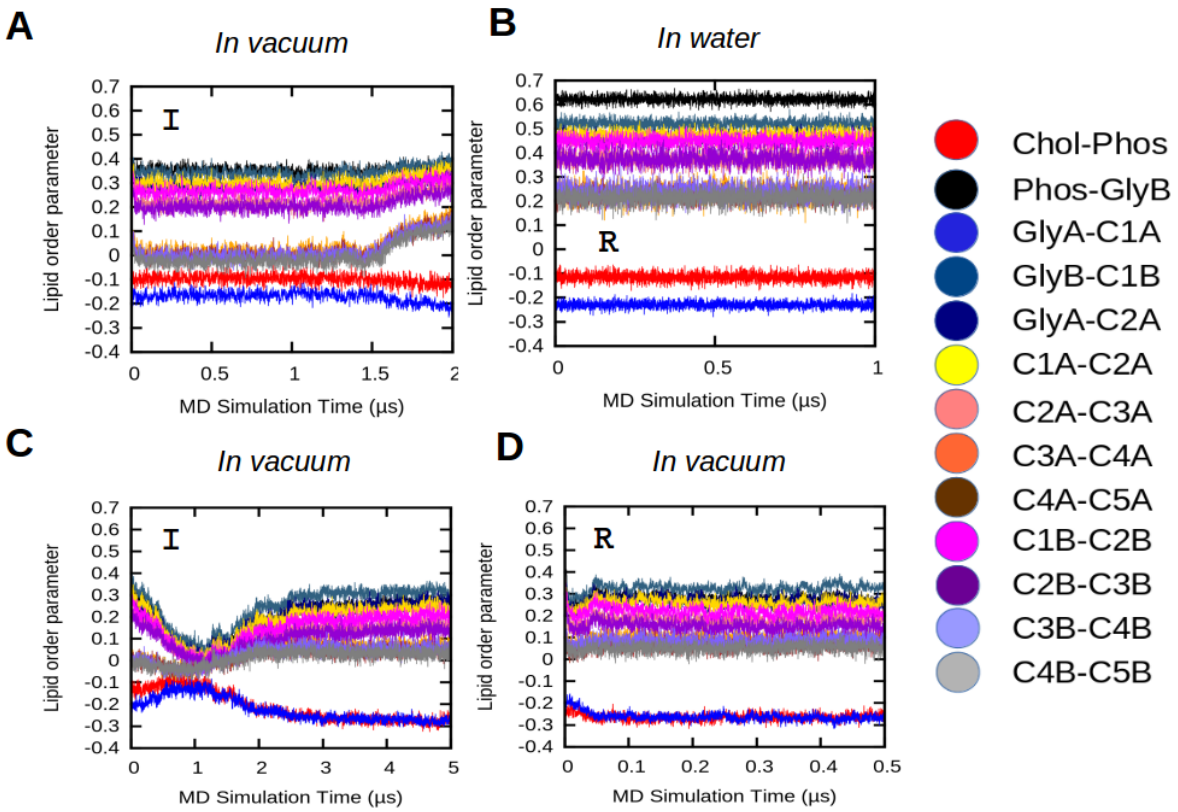

Fig. 6.11: Time evolution of average lipid order parameters calculated for CG-MD simulations of lipid interactions with small graphene (**A**) and graphene oxide (**C** and **D**), and a free-standing bilayer (**B**). The initial simulated lipid structure is indicated by the letter within the plots; I = inverted bilayer, R = regular bilayer.

The overall reduced lipid order on the graphene (oxide) supports could be a result of lateral interactions between the lipid tails and the surface (partial alignment of the lipid tails with the surface), as is observed in the pristine graphene systems (due to strong hydrophobic forces), or headgroup interactions with the graphene oxide surface, disrupting lipid order in the layer. Additionally, only small deviations in the average tail order parameter are observed from the hydrated systems compared to simulations in vacuum, suggesting that solvent has little affect on lipid order (Fig. 6.12).

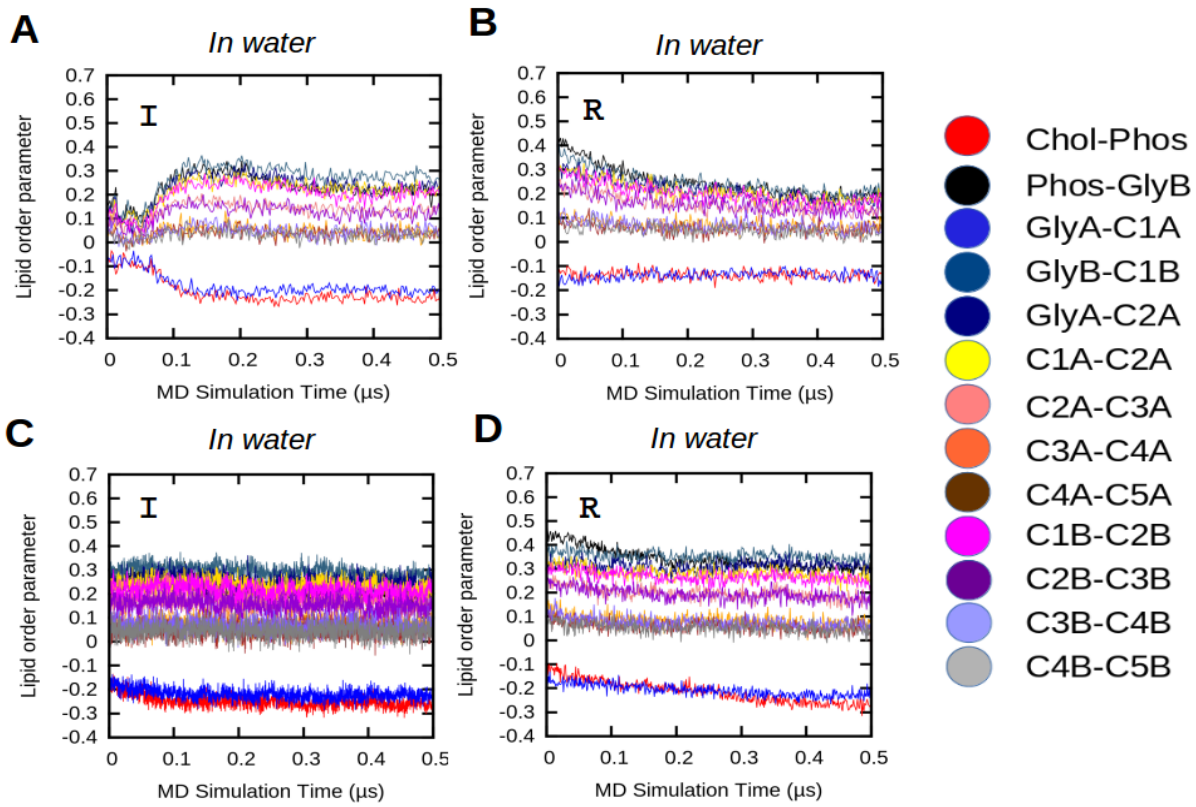

Fig. 6.12: Time evolution of average lipid order parameters calculated for CG-MD simulations of lipid interactions with small (**A**, **B**) and large graphene oxide (**C** and **D**) surfaces. The initial simulated lipid structure is indicated by the letter within the plots; I = inverted bilayer, R = regular bilayer.

Lipid Diffusion

Characterisation of the diffusion of lipids within supported membrane systems is fundamental to understanding the dynamic properties of cellular membranes [276–279]. Experimental studies report both linear and anomalous diffusion regimes of lipids depending on the systems, and increasingly highlight the importance of the time and length scale in which the diffusion data is collected for distinction between these regimes [272, 277, 278, 280]. In order to provide an estimate of lateral diffusion of lipids within the simulated supported bilayer systems presented here, both linear

and anomalous diffusion regimes were considered. Mean square displacement (MSD) data was collected for individual systems by tracking the displacement of lipids over different simulation time sampling windows (see Methods). The resultant MSD data was then fitted with either a linear diffusion equation [240] or an anomalous diffusion equation [268].

Comparison of the linear and anomalous diffusion fits revealed some interesting differences in the lateral displacement behaviour of lipids on surfaces, particularly relative to lipid diffusion in free-standing bilayers. For example, the measured lipid diffusion coefficients in supported regular bilayers on graphene oxide suggests that lipid mobility is considerably slower in the presence of the support compared to free-standing bilayers (Fig. 6.13). The addition of aqueous solution appeared to result in increased lipid diffusion on the surface (Fig. 6.13B), suggesting the water layers could act as a lubricant for lipid diffusion within the layer. This has been suggested for experimental systems of supported lipid membranes on hydrophilic surfaces [277].

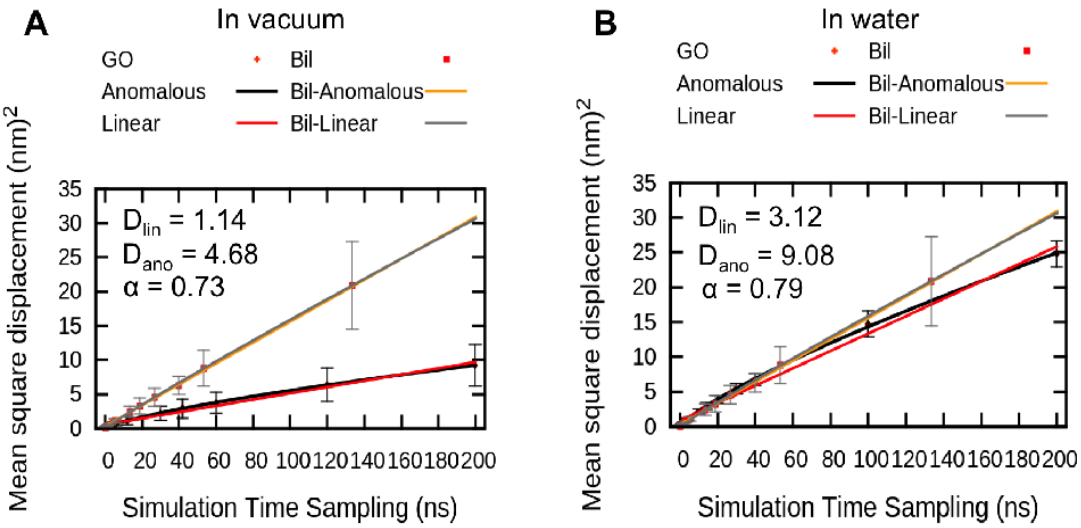

Fig. 6.13: MSD data plotted against different time sampling windows (ns), calculated for lipids in a CG-MD simulation of a regular bilayer (512 lipids) with a small graphene oxide surface in vacuum (**A**) and in water (**B**). MSD data was also calculated for a simulation of a free-standing bilayer in water (Bil). MSD data was averaged for all lipids in the simulation and sampled at window frames ranging from 1-200 ns in simulation length. The data was fitted with either a linear or anomalous diffusion equation; the resultant diffusion coefficient (D_{lin}, D_{ano}) are shown (units: x10^{-7} cm^2/s). Errors bars show standard deviation.

Furthermore, fitting MSD data using the anomalous diffusion equation produced significant α values for all simulated systems, suggesting that the lipids within the supported membranes may exhibit different diffusion regimes in function of surface interactions (Fig. 6.14B). In particular, sub-diffusion was measured for the majority of supported membrane simulations, in which the diffusion of mobile lipid particles decreased with time (α = 0.7-0.8) (Fig. 6.14B). Conversely, α values > 1 were calculated for select simulations of lipid bilayers with the large graphene and graphene oxide surfaces, in which the initial lipids structures disassembled and spread across the available surface area. This is indicative of fast spreading of the lipids, and as such is not reflective of lipid diffusion within stable membrane structures.

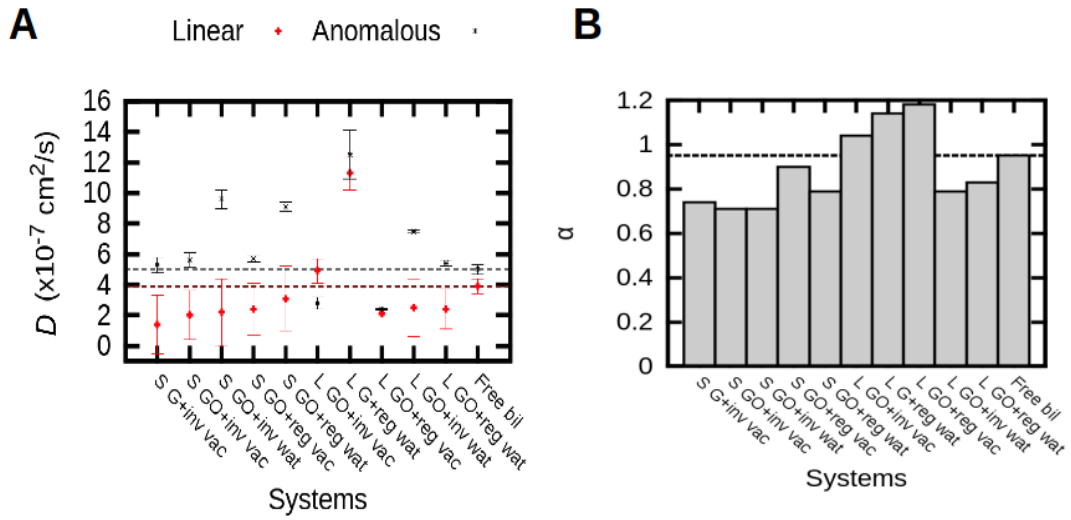

Fig. 6.14: **(A)** Diffusion coefficients (D) calculated from either linear or anomalous fits of MSD data for each of the simulated systems in this chapter. **(B)** α values (scaling exponent) calculated from anomalous diffusion fits for the simulated systems shown in **A**. S = small, L = large, GO = graphene oxide, G = graphene, inv = inverted bilayer, reg = regular bilayer, vac = vacuum, wat = water, Free bil = free-standing bilayer. Dashed lines represent the diffusion properties calculated from the CG simulation of the free-standing bilayer.

Importantly, the displacement of lipids within the free-standing bilayer were consistent with the linear diffusion model, producing a value of 3.9 x 10^{-7} cm^2/s ($\pm$ 0.5) (α = 0.95). This value can be scaled appropriately (divide by four) to account for the reduced degrees of freedom within CG simulations, resulting in D_{lin} = 0.98 x 10^{-7} cm^2/s, which compares well to atomistic simulations of DOPC bilayers in water (1.5 x 10^{-7} cm^2/s) [133, 264]. However, it remains difficult to compare these diffusion data with experimental values, or even other simulation data, given that the dependence on both the time scale at which diffusion is measured, and the length scale of the system [272, 276, 278, 280]. MSD data and linear and anomalous diffusion fits are presented for all systems in Fig. 6.15-6.17.

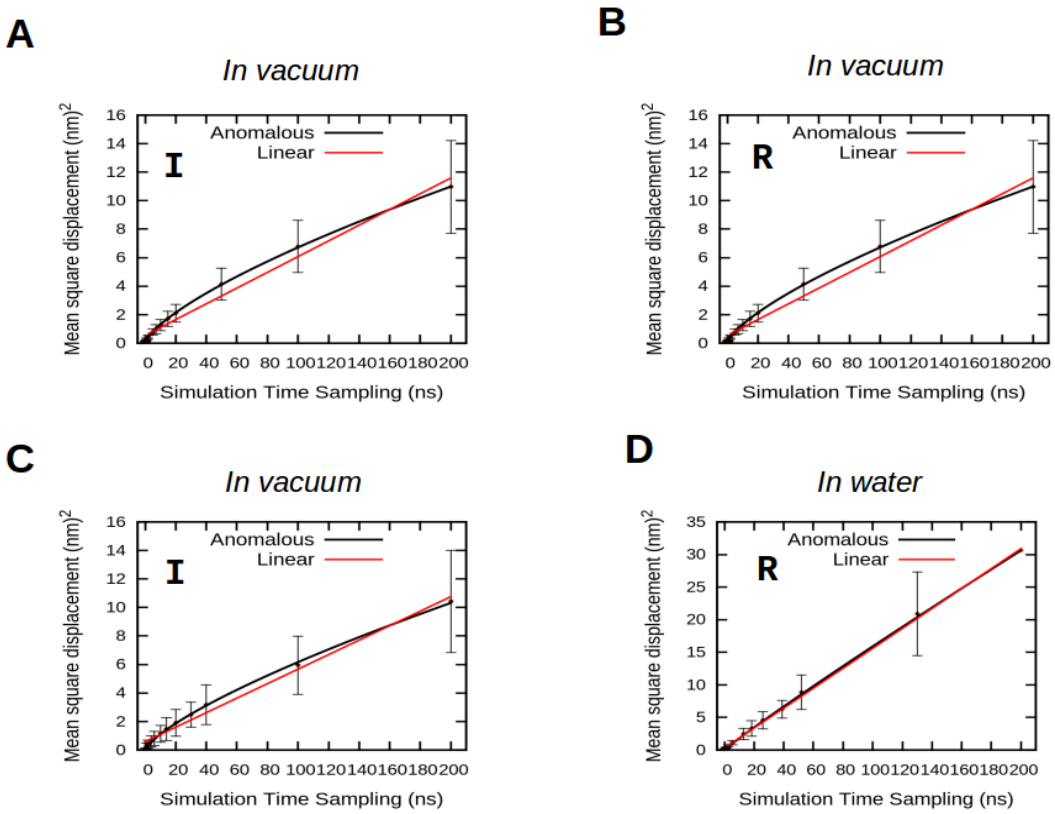

Fig. 6.15: Average lipid MSD data points calculated for different simulation sampling times for lipid interactions with small graphene oxide (**A** and **B**), and small graphene (**C**) and a free-standing bilayer (**D**). The initial simulated lipid structure is indicated by the letter within the plots; I = inverted bilayer, R = regular bilayer. The MSD data points were fitted with both linear and anomalous diffusion equations. Error bars show standard deviation.

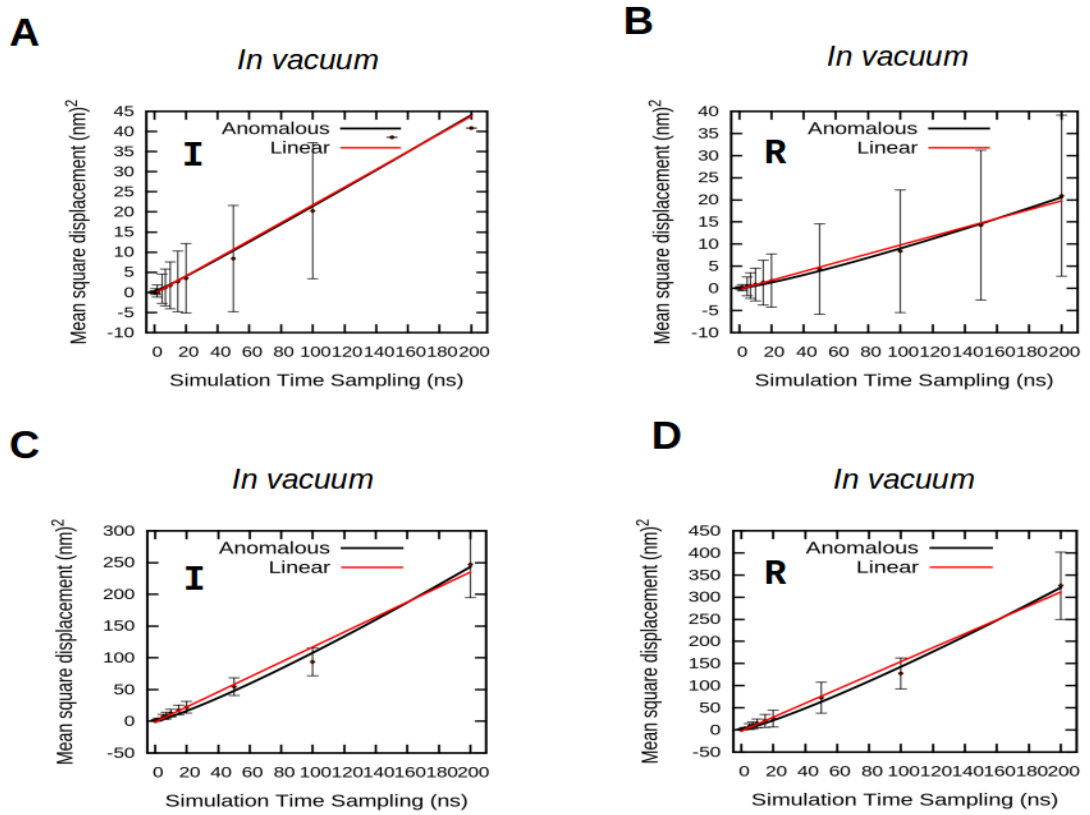

Fig. 6.16: Average lipid MSD data points calculated for different simulation sampling times for lipid interactions with large graphene oxide (**A** and **B**) and large graphene surfaces (**C** and **D**). The initial simulated lipid structure is indicated by the letter within the plots; I = inverted bilayer, R = regular bilayer. The MSD data points were fitted with both linear and anomalous diffusion equations. Errors bars show standard deviation.

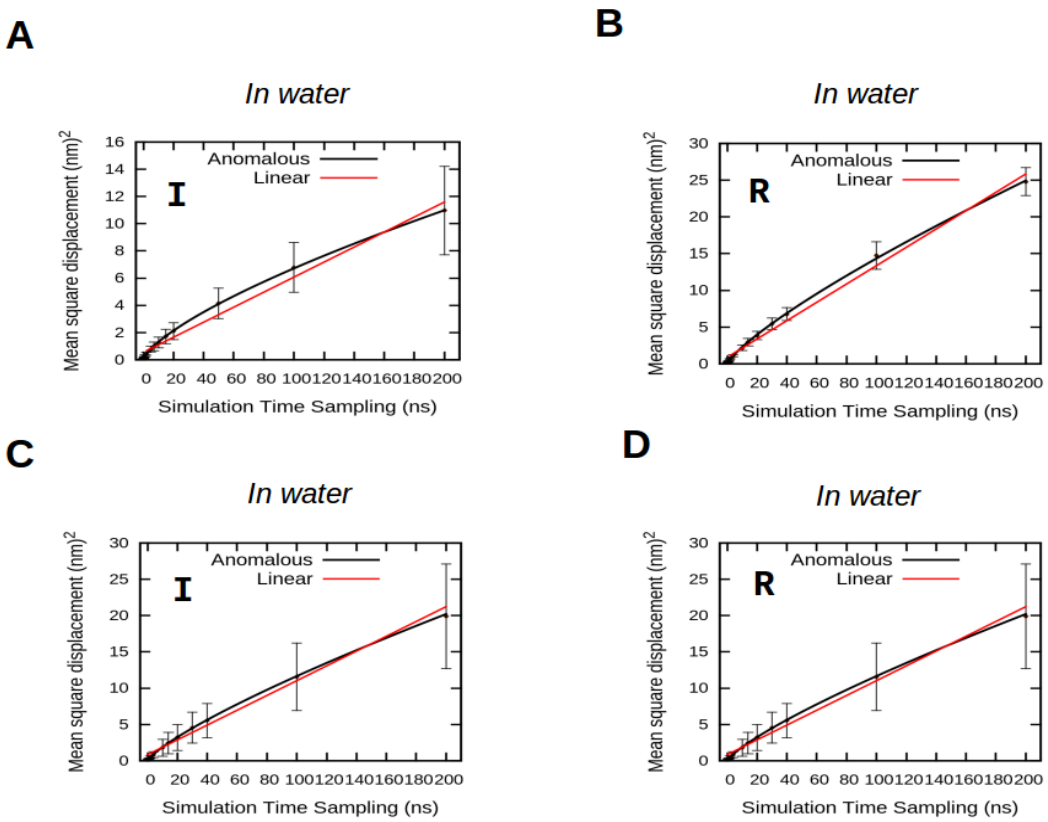

Fig. 6.17: Average lipid MSD data points calculated for different simulation sampling times for lipid interactions with small (**A** and **B**) and large (**C** and **D**) graphene oxide surfaces. The initial simulated lipid structure is indicated by the letter within the plots; I = inverted bilayer, R = regular bilayer. The MSD data points were fitted with both linear and anomalous diffusion equations. Errors bars show standard deviation.

6.4 Discussion

In this chapter, lipid organisation on graphene and graphene oxide supports was investigated using both experimental and computational techniques. Similar to other studies of supported lipid structures, the physical properties of the membrane structures were affected by surface polarity, as well as system hydration [105, 107, 281–283]. Specifically, different bilayer topologies were observed to form on pristine graphene compared to graphene oxide surfaces. AFM measurements suggested the formation

of inverted bilayer topology on graphene, whilst a 1.5 lipid structure was observed on graphene oxide in air (Fig 6.5B) [259]. The same membrane organisation was observed in CG simulations, in which spontaneous rearrangement of different initial lipid configurations could be observed (e.g monolayers *vs* regular bilayers), converging to very similar end structures (Fig. 6.4 and 6.7). Furthermore, the CG simulations suggested that stability of the observed lipid structure was related to lipid density on the surface. Whilst small inverted bilayers were observed to disassemble on large graphene surfaces due to strong hydrophobic interactions with the surface, larger bilayers were more stable (Fig. 6.6). Importantly, initial AFM characterisation of phospholipid patches on pristine graphene in air indicated that lipids are more mobile on this surface than on hydrophilic surfaces such as silicon dioxide, also attributed to the strong hydrophobic interactions between lipid hydrocarbon tails and graphene [259]. Spreading of a very small membrane patch on a large graphene surface is thus perhaps not unexpected. Furthermore, the number of lipids deposited on the graphene surface in the L-DPN experiments is likely to be much higher than the number of lipids simulated in these much smaller systems, making direct comparisons between system behaviour more difficult, particularly when larger graphene surface areas were simulated (relative to overall lipid surface area). Additionally, the experiments were performed in humid conditions, which is likely to change the conditions and interactions between the lipid molecules on the support surface, such as water cluster formation around the lipid headgroups, which could potentially stabilise lipid structure formation. Indeed, L-DPN functions by transfer through a water meniscus formed between the AFM tip and the surface, and is directly influenced by the humidity of the surrounding air [259]. The approximation of these conditions to a completely dry vacuum in the simulations are thus likely to affect the lipid behaviour on the graphene and graphene oxide surfaces. While the comparisons made with the AFM height measurements of

the different lipid topologies (e.g inverted bilayers) seem to correspond with the observations made from the simulations, these differences in conditions must be taken into account, particularly when attempting to quantify molecular properties.

Interestingly, the addition of water to inverted lipid structures stabilised on pristine graphene also resulted in disassembly and spreading on the surface, both experimentally and computationally (Fig. 6.8A) [259]. Conversely, very stable bicelle-like structures were observed in simulations with graphene oxide in water, correlating with experimental AFM measurements of supported bilayers on graphene oxide and silicon dioxide in solution [271, 284]. The CG simulations also suggested that lipids were dynamic on either support surface, as is evidenced from monitoring the time evolution of lipid tail order parameters. This is particularly true for simulations of regular bilayers with the small graphene oxide surface, which resulted in complete restructuring of the bilayer to form the more favourable 1.5 bilayer configuration (Fig. 6.9). Exchange between lipids from upper and bottom leaflets with respect to the surface was also observed, emphasising the evolving nature of these lipid structures on the supporting surface.

Furthermore, characterisation of lipid mobility within the supported membrane structures suggested that lipid diffusion might reside within the sub-diffusion regime for most systems, where it is expected that interactions between the lipid and the support affect longer-term MSD values that might result in anomalous diffusive behaviour. Sub-diffusion of lipids within supported DOPC membranes on hydrophilic titanium oxide was also suggested by Tero *et al*, related to the atomic topology of the support surface [272]. Furthermore, experimental measurements of lipid diffusion in supported membranes have reported both linear and anomalous diffusion regimes, and relate this to lipid interactions with the support surface, surface topology of the support, membrane composition and topology (e.g planar bilayer *vs* vesicles), as well as system

hydration [272, 277, 278, 280]. The observed linear diffusion coefficients range from $\sim 0.4 \pm 0.05$ x 10^{-7} cm^2/s for DOPC planar membranes on mica supports to 0.25 ± 0.02 - 0.64 ± 0.06 x 10^{-7} cm^2/s for supported membranes on polymer surfaces [277, 279]. Importantly, many of the experimental measurements of lipid diffusion coefficients are reported for supported membranes in aqueous solution. Given that system hydration can be directly related to measured lipid mobility, as is suggested by the simulations presented here (Fig. 6.13), this is clearly an important consideration when comparing simulation data to experimental measurements of lipid mobility [277]. Furthermore, the distinction of varying diffusion regimes is also directly related to both the temporal and spatial scale at which diffusion is measured [276]. It would therefore be of interest to extend the initial simulations presented here to bigger length scales and longer time scales, ideally with the availability of experimental data for similar systems. Bigger systems would also provide insight into effects of lipid-surface interactions on lipid phasing, and elucidate how lipid density affects bilayer stability.

To conclude, the collective simulations and experimental results highlight the effects of both surface polarity and the solvent environment on supported phospholipid membrane structures. These observations imply that the supported bilayer configuration can be tuned as a function of system solvation and surface polarity, resulting in different final structures. This has implications for discrete patterning of support surfaces with biological assemblies, allowing superior spatial control of membrane formations on the support. Future studies could therefore also focus on including different, functionalised lipid molecules within the supported membranes, exploring affects on membrane topology and molecular properties. These systems could be combined with other molecules, such as proteins, including lipases. Characterising these interactions would provide important insight into the applications of graphene/graphene oxide within biotechnology, including sensor devices, drug delivery systems, and catalytic

systems [275, 285, 286].

Chapter 7

Conclusions and Future Directions

This thesis represents a study of how different lipases interact with surfaces of varying structural properties. Specific questions regarding the structural mechanisms underlying lipase activation are asked, and initial results suggest possible answers. In addition to this body of work, an integrated modelling and experimental study is presented that investigated how lipids interact with graphene and graphene oxide surfaces. The aim of the thesis was to characterise lipase structural dynamics and to establish modular systems for exploration of this aspect. A particular outcome is the suggestion of a possible activation mechanism for the M37 lipase, previously unidentified for this industrially relevant lipase.

7.1 Major Conclusions

The first results chapter in this thesis (Chapter 3) begins with a study of the recently identified M37 psychrophilic lipase. Initial investigations aimed to characterise how this lipase interacts with biological interfaces, including lipid bilayers and triglyceride surfaces. In particular, the chapter demonstrated the importance of electrostatic interactions in driving lipase association with lipid bilayer interfaces. Further investi-

gations identified functionally relevant structural motions of the enzyme, which likely underlie the M37 lipase activation mechanism. Specifically, the lid region was observed to undergo large-scale displacement using both unbiased and biased simulation techniques. First, steered MD simulations indicated structural flexibility within the lid region, showing that the lid could be displaced from its original closed position, contrasting the apparently rigid active site flap region. Second, unbiased simulations of M37 interactions with a natural substrate interface resulted in opening of the closed lipase structure, uncovering the entry pathway of the catalytic site. The interfacial interactions induced a conformational change within the lid region, which was very similar to the motions identified from the steered MD simulations. Finally, substrate binding with the opened lipase structure was simulated, suggesting that the opened lipase conformation could allow entry of a substrate molecule into the catalytic site without hindrance.

Continuing the investigation of lipase interactions with biological surfaces, Chapter 4 focussed on how mutations in the lid of the well-characterised *Thermomyces lanuginosus* lipase affected interfacial interactions with triglyceride surfaces. First, CG simulations were used to study the binding mechanisms of the wild-type TLL and two mutant TLL proteins, the 1L and 3L proteins. Multiple CG simulations identified a general difference in binding orientation of the lipase with the triglyceride interface depending on the lipase variant. Atomistic simulations confirmed this trend, indicating a tendency of the 1L variant to bind such that the lid region pointed away from the substrate interface. Extension of the atomistic simulations of the proteins in water displayed a general inflexibility of the 1L lid region, contrasting the 3L protein, which displayed partial activation at both the substrate interface and in water. Enhanced sampling methods confirmed the rigidity of the lid region of the 1L protein, relative to the wild-type and 3L proteins. Finally, it is suggested that the particular residue sub-

stitutions within the 1L lid region underlie its low interfacial activity profile, whereas the mutations within the lid region of the 3L variant most likely underlie its enhanced interfacial and soluble activation profile.

Chapter 5 marks a shift in the focus on lipase interactions with "soft" biological interfaces, to "hard" non-biological interfaces, such as graphene and graphene oxide. This chapter presents a study of how the M37 lipase interacts with pristine graphene (unfunctionalised) and graphene oxide models. Initial CG simulations indicated the protein only exhibited long-lasting interactions with the graphene oxide models and not the pristine graphene model. The protein adsorbed to the graphene oxide surface in a distribution of binding orientations. A correlation between the number of interfacial contacts and lipase mobility/diffusion on the surface was suggested; the larger the number of contacts, the less mobile the protein was on the surface. Further investigation with a charged graphene oxide surface model indicated differences in the association behaviour of M37 relative to the uncharged GO model. The same contact/mobility correlation was found, with reduced mobility on the surface in function of a higher number of protein-surface contacts. The initial simulations confirmed that graphene oxide might be an interesting system to investigate M37 lipase dynamics with, particularly given that the results from Chapter 3 form a comparative basis.

An extension of the graphene and graphene oxide surface systems is presented in Chapter 6. This chapter however focussed on how lipid bilayers interact with these planar surfaces. The CG simulations are presented in the context of experimental information on the same systems. The simulations were shown to accurately reproduce the experimental data, suggesting altered lipid bilayer configurations depending on the support surface. Whilst inverted bilayers were observed to be stable on pristine graphene, a "1.5" bilayer topology were formed on graphene oxide surfaces in vacuum. Furthermore, addition of water to the systems resulted in substantial reorganisation

of the pre-formed lipid systems on both surfaces. The aqueous environment also significantly altered the calculated diffusion properties of lipids within the supported bilayers, enhancing their mobility relative to "dry" systems. Furthermore, lipids were less ordered on the surfaces compared to free-standing bilayers. The integrated study represents a basis for future simulations of modular systems with lipase enzymes.

7.2 Future Directions

A major area of interest in the study of lipase enzymes is the characterisation of the energetics underlying their activation mechanisms. This is a particular direction in which the results presented in Chapter 3 (M37) and Chapter 4 (TLL) could be extended. Regarding Chapter 3, an initial body of simulation work has already been performed in attempts to measure the energetics underlying the suggested mechanism of M37 lipase activation (Appendix A). Classical and well-tempered metadynamics simulations indicated that this free energy landscape is quite complex. In particular, the initial simulations using either method did not result in convergence of lid sampling, despite simulating many different parameter combinations. Extension of these methods using umbrella sampling simulations confirm this complexity.

With reference to Chapter 4, the initial steered MD work suggested that the lid region of the mutated 1L TLL variant was less flexible compared to the 3L lid region. The increased interfacial activity profile for the 3L variant in the related experimental study was related to a lowering of the activation energy barrier for opening of the lipase structure. It is thus of interest to characterise the free energy landscapes for the activation of the 1L, 3L, and wild-type variant, both in aqueous solution and at interface. Another area of particular interest would be to perform rescue mutations for the 1L and 3L variants, eventually restoring the wild-type amino acid sequence of the lid region. In particular, restoration of the T87 mutation to the original glutamate

residue in the 3L variant would provide information about the driving forces for its increased activity in aqueous solution, and at the interface. Furthermore, both the wild-type and 1L TLL variants contained glycosylated residues in the experiments. Incorporation of the sugar groups in the lipase structures could be an avenue of future research as well.

Going on to Chapter 5, the initial CG simulations of M37 with graphene and graphene oxide should be extended to atomistic simulations. Attempts to parameterise an atomistic model of graphene oxide for simulation with the GROMOS 54A7 forcefield has been performed (Appendix B). This model should be refined and validated regarding the its parameters. Using the atomistic graphene oxide model, the dynamics of the M37 protein could be studied and compared to its interactions with biological surfaces from Chapter 3. It would be of particular interest to observe if the M37 lipase displays any similar structural transitions as were observed in Chapter 3, possibly adding more data for the proposed activation mechanism.

Regarding the lipid bilayer simulations with graphene and graphene oxide in Chapter 6, these simulations could be extended in both time and size scales. This would allow for better comparison of molecular with experimental data for similar systems, particularly regarding lipid diffusion. The initial and simple DOPC bilayers that were simulated could be extended to more complex lipid mixtures, investigating how the lipid itself affects supported membrane properties. Finally, these more complex systems could be simulated at different temperatures to monitor possible effects on phase transitions when interacting with a support surface, compared to free-standing bilayers. This is also of interest to characterise experimentally on graphene and graphene oxide surfaces.

References

[1] P. Reis, K. Holmberg, H. Watzke, M. E. Leser, and R. Miller, "Lipases at interfaces: a review.," *Advances in colloid and interface science*, vol. 147-148, pp. 237–50, 2009.

[2] A. Houde, A. Kademi, and D. Leblanc, "Lipases and their industrial applications," *Applied biochemistry and biotechnology*, vol. 118, pp. 155–170, 2004.

[3] P. Adlercreutz, "Immobilisation and application of lipases in organic media.," *Chemical Society reviews*, vol. 42, pp. 6406–36, jul 2013.

[4] F. Hasan, A. A. Shah, and A. Hameed, "Industrial applications of microbial lipases," *Enzyme and Microbial Technology*, vol. 39, pp. 235–251, jun 2006.

[5] O. Berg and M. Jain, *Interfacial Enzyme Kinetics*. John Wiley & Sons, 2002.

[6] M. H. Gelb, M. K. Jain, A. M. Hanel, and O. G. Berg, "Interfacial enzymology of glycerolipid hydrolases: lessons from secreted phospholipases A2.," *Annual review of biochemistry*, vol. 64, pp. 653–88, 1995.

[7] P. Reis, K. Holmberg, R. Miller, *et al.*, "Competition between lipases and monoglycerides at interfaces," *Langmuir*, vol. 24, no. 14, pp. 7400–7407, 2008.

[8] H. Lodish, A. Berk, S. L. Zipursky, *et al.*, *Molecular Cell Biology*. New York: w.H Freeman, 4th ed., 2000.

[9] W. Cho and R. V. Stahelin, "Membrane-protein interactions in cell signaling and membrane trafficking.," *Annu. Rev. Biophys. Biomol. Struct.*, vol. 34, pp. 119–151, 2005.

[10] B. Soullam and H. J. Worman, "Signals and structural features involved in integral membrane protein targeting to the inner nuclear membrane.," *The Journal of Cell Biology*, vol. 130, pp. 15–27, jul 1995.

[11] P. Walter, "Signal Sequence Recognition and Protein Targeting To the Membrane," *Annu. Rev. Cell Biol*, vol. 10, pp. 87–199, 1994.

[12] O. G. Berg, M. H. Gelb, M. D. Tsai, and M. K. Jain, "Interfacial enzymology: The secreted phospholipase A2-paradigm," *Chemical Reviews*, vol. 101, no. 9, pp. 2613–2653, 2001.

[13] J. M. Winget, Y. H. Pan, and B. J. Bahnson, "The interfacial binding surface of phospholipase A2s.," *Biochimica et biophysica acta*, vol. 1761, pp. 1260–9, nov 2006.

[14] B. P. Ziemba and J. J. Falke, "Lateral Diffusion of Peripheral Membrane Proteins on Supported Lipid Bilayers Is Controlled by the Additive Frictional Drags of 1) Bound Lipids and 2) Protein Domains Penetrating into the Bilayer Hydrocarbon Core," *Chemistry and physics of lipids*, vol. 0, pp. 67–77, may 2013.

[15] M. J. Richards, C.-Y. Hsia, R. R. Singh, *et al.*, "Membrane Protein Mobility and Orientation Preserved in Supported Bilayers Created Directly from Cell Plasma Membrane Blebs," *Langmuir*, vol. 32, pp. 2963–2974, mar 2016.

[16] M. D. Marazuela and M. C. Moreno-Bondi, "Fiber-optic biosensors - An overview," *Analytical and Bioanalytical Chemistry*, vol. 372, no. 5-6, pp. 664–682, 2002.

[17] P. Ormos, L. Fabian, L. Oroszi, *et al.*, "Protein-based integrated optical switching and modulation," *Applied Physics Letters*, vol. 80, no. 21, pp. 4060–4062, 2002.

[18] D. G. Castner and B. D. Ratner, "Biomedical surface science: Foundations to frontiers," *Surface Science*, vol. 500, pp. 28–60, mar 2002.

[19] A. Gessner, A. Lieske, B.-R. Paulke, and R. H. Müller, "Functional groups on polystyrene model nanoparticles: Influence on protein adsorption," *Journal of Biomedical Materials Research Part A*, vol. 65A, pp. 319–326, jun 2003.

[20] P. Walstra and A. L. Deroos, "Proteins at Air-Water and Oil-Water Interfaces – Static and Dynamic Aspects," *Food. Rev. Int.*, vol. 9, no. 4, pp. 503–525, 1993.

[21] J. L. Brash and T. A. Horbett, "Proteins at Interfaces," in *Proteins at Interfaces II*, vol. 602 of *ACS Symposium Series*, p. 1, American Chemical Society, may 1995.

[22] S. Freilich, R. V. Spriggs, R. A. George, *et al.*, "The Complement of Enzymatic Sets in Different Species," *Journal of Molecular Biology*, vol. 349, pp. 745–763, jun 2005.

[23] K. Tipton and S. Boyce, "History of the enzyme nomenclature system," *Bioinformatics*, vol. 16, pp. 34–40, jan 2000.

[24] T. Maehama and J. E. Dixon, "The Tumor Suppressor, PTEN/ MMAC1, Dephosphorylates the Lipid Second Messenger, Phosphatidylinositol 3,4,5-Trisphosphate*ressor," *Journal of Biological Chemistry*, vol. 273, no. 22, pp. 13375–13378, 1998.

[25] O. G. Mouritsen, T. L. Andresen, A. Halperin, *et al.*, "Activation of interfacial enzymes at membrane surfaces.," *Journal of physics: Condensed matter*, vol. 18, no. 28, pp. S1293–S1304, 2006.

[26] A. Aloulou, J. a. Rodriguez, S. Fernandez, *et al.*, "Exploring the specific features of interfacial enzymology based on lipase studies.," *Biochimica et biophysica acta*, vol. 1761, pp. 995–1013, sep 2006.

[27] R. Verger and G. H. de Haas, "Enzyme reactions in a membrane model 1: A new technique to study enzyme reactions in monolayers," *Chemistry and Physics of Lipids*, vol. 10, pp. 127–136, feb 1973.

[28] S. Majd, E. C. Yusko, J. Yang, D. Sept, and M. Mayer, "A model for the interfacial kinetics of phospholipase D activity on long-chain lipids," *Biophysical Journal*, vol. 105, no. 1, pp. 146–153, 2013.

[29] R. R. Annand, M. Kontoyianni, J. E. Penzotti, *et al.*, "Active Site of Bee Venom Phospholipase A2: The Role of Histidine-34, Aspartate-64 and Tyrosine-87," *Biochemistry*, vol. 35, pp. 4591–4601, jan 1996.

[30] O. G. Berg, B. Z. Yu, J. Rogers, and M. K. Jain, "Interfacial catalysis by phospholipase A2: determination of the interfacial kinetic rate constants.," *Biochemistry*, vol. 30, no. 29, pp. 7283–97, 1991.

[31] M. K. Jain, J. Rogers, D. Jahagirdar, J. F. Marecek, and F. Ramirez, "Kinetics of interfacial catalysis by phospholipase A2 in intravesicle scooting mode, and heterofusion of anionic and zwitterionic vesicles," *Biochimica et Biophysica Acta (BBA) - Biomembranes*, vol. 860, pp. 435–447, sep 1986.

[32] G. C. Upreti, S. Rainier, and M. K. Jain, "Intrinsic differences in the perturbing ability of alkanols in bilayer: Action of phospholipase A2 on the alkanol-modified phospholipid bilayer," *The Journal of Membrane Biology*, vol. 55, no. 2, pp. 97–112, 1980.

[33] M. K. Jain, C. D. Krause, J. T. Buckley, T. Bayburt, and M. H. Gelb, "Characterization of interfacial catalysis by Aeromonas hydrophila lipase/acyltransferase in the highly processive scooting mode.," *Biochemistry*, vol. 33, pp. 5011–5020, may 1994.

[34] H. Brockerhoff, "On the function of bile salts and proteins as cofactors of lipase.," *Journal of Biological Chemistry*, vol. 246, no. 18, pp. 5828–5831, 1971.

[35] C. J. O'Connor and P. Walde, "Interactions of human milk lipase with sodium taurocholate and other surfactants," *Langmuir*, vol. 2, pp. 139–146, mar 1986.

[36] M. Wickham, M. Garrood, J. Leney, P. D. G. Wilson, and A. Fillery-Travis, "Modification of a phospholipid stabilized emulsion interface by bile salt: effect on pancreatic lipase activity," *Journal of Lipid Research*, vol. 39, pp. 623–632, mar 1998.

[37] A. D. Patrick and B. D. Lake, "Deficiency of an Acid Lipase in Wolman's Disease," *Nature*, vol. 222, pp. 1067–1068, jun 1969.

[38] E. Gilbert, "Pseudomonas lipases: Biochemical properties and molecular cloning," *Enzyme and Microbial Technology*, vol. 15, pp. 634–645, aug 1993.

[39] K. E. Jaeger, S. Ransac, B. W. Dijkstra, *et al.*, "Bacterial lipases.," *FEMS microbiology reviews*, vol. 15, pp. 29–63, sep 1994.

[40] L. Brady, A. M. Brzozowski, Z. S. Derewenda, *et al.*, "A serine protease triad forms the catalytic centre of a triacylglycerol lipase," *Nature*, vol. 343, pp. 767–770, feb 1990.

[41] N. M. Packter, "Lipases — their structure, biochemistry and application," *Biochemical Education*, vol. 22, p. 216, oct 1994.

[42] L. Sarda and P. Desnuelle, "Actions of pancreatic lipase on esters in emulsions," *Biochemica et Biophysica Acta*, vol. 30, no. 3, pp. 512–521, 1958.

[43] J. H. Law and J. K. I. Zdy, "Catalysis by Adsorbed Enzymes," *Journal of Biological Chemistry*, vol. 248, no. 14, pp. 4965–4970, 1973.

[44] H. Brockerhoff, "Substrate specificity of pancreatic lipase," *Biochimica et Biophysica Acta (BBA) - Enzymology*, vol. 159, no. 2, pp. 296–303, 1968.

[45] A. M. Brzozowski, U. Derewenda, Z. S. Derewenda, *et al.*, "A model for interfacial activation in lipases from the structure of a fungal lipase-inhibitor complex.pdf," *Nature*, vol. 351, pp. 491–494, 1991.

[46] B. Vasel, H.-J. Hecht, R. D. Schmid, and D. Schomburg, "3D-Structures of the lipase from Rhizomucor miehei at different temperatures and computer modelling of a complex of the lipase with trilaurylglycerol," *Journal of Biotechnology*, vol. 28, pp. 99–115, mar 1993.

[47] Z. S. Derewenda and A. M. Sharp, "News from the interface: the molecular structures of triacylglyceride lipases.," *Trends in biochemical sciences*, vol. 18, pp. 20–25, jan 1993.

[48] H. van Tilbeurgh, L. Sarda, R. Verger, and C. Cambillau, "Structure of the pancreatic lipase-procolipase complex," *Nature*, vol. 359, pp. 159–162, sep 1992.

[49] H. Van Tilbeurgh, M. P. Egloff, C. Martinez, *et al.*, "Interfacial activation of the lipase-procolipase complex by mixed micelles revealed by X-ray crystallography," *Nature*, vol. 362, no. 6423, pp. 814–820, 1993.

[50] D. M. Lawson, a. M. Brzozowski, S. Rety, C. Verma, and G. G. Dodson, "Probing the nature of substrate binding in Humicola lanuginosa lipase through X-ray crystallography and intuitive modelling.," *Protein engineering*, vol. 7, pp. 543–50, apr 1994.

[51] J. D. Schrag, Y. Li, S. Wu, and M. Cygler, "Ser-His-Glu triad forms the catalytic site of the lipase from Geotrichum candidum," *Nature*, vol. 351, pp. 761–764, jun 1991.

[52] Z. S. Derewenda, U. Derewenda, and G. G. Dodson, "The crystal and molecular structure of the Rhizomucor miehei triacylglyceride lipase at 1.9 Å resolution," *Journal of Molecular Biology*, vol. 227, no. 3, pp. 818–839, 1992.

[53] U. Derewenda, A. M. Brzozowski, D. M. Lawson, and Z. S. Derewenda, "Catalysis at the interface: the anatomy of a conformational change in a triglyceride lipase.," *Biochemistry*, vol. 31, no. 5, pp. 1532–1541, 1992.

[54] P. Grochulski, Y. Li, J. D. Schrag, and M. Cygler, "Two conformational states of Candida rugosa lipase.," *Protein science : a publication of the Protein Society*, vol. 3, pp. 82–91, jan 1994.

[55] U. Derewenda, L. Swenson, Y. Wei, *et al.*, "Conformational lability of lipases observed in the absence of an oil-water interface: crystallographic studies of enzymes from the fungi Humicola lanuginosa and Rhizopus delemar," *Journal of Lipid Research*, vol. 35, no. 3, pp. 524–534, 1994.

[56] a. M. Brzozowski, H. Savage, C. S. Verma, *et al.*, "Structural origins of the interfacial activation in Thermomyces (Humicola) lanuginosa lipase.," *Biochemistry*, vol. 39, pp. 15071–82, dec 2000.

[57] J. Skjold-Jørgensen, V. K. Bhatia, J. Vind, *et al.*, "Enzymatic activity of lipases correlates with polarity-induced conformational changes: a Trp-induced quenching (TrIQ) fluorescence study," *Biochemistry*, p. 150618153605000, 2015.

[58] C. Chapus and M. Semeriva, "Mechanism of pancreatic lipase action. 2. Catalytic properties of modified lipases.," *Biochemistry*, vol. 15, pp. 4988–4991, nov 1976.

[59] F. K. Winkler, A. D'Arcy, and W. Hunziker, "Structure of human pancreatic lipase," *Nature*, vol. 343, pp. 771–774, feb 1990.

[60] D. J. Ericsson, A. Kasrayan, P. Johansson, *et al.*, "X-ray structure of Candida antarctica lipase A shows a novel lid structure and a likely mode of interfacial activation.," *Journal of molecular biology*, vol. 376, pp. 109–19, feb 2008.

[61] S. Naik, A. Basu, R. Saikia, *et al.*, "Lipases for use in industrial biocatalysis: Specificity of selected structural groups of lipases," *Journal of Molecular Catalysis B: Enzymatic*, vol. 65, pp. 18–23, aug 2010.

[62] R. Sharma, Y. Chisti, and U. C. Banerjee, "Production, purification, characterization, and applications of lipases," 2001.

[63] M. K. Walsh, *Novel Enzyme Technology for Food Applications*. Cambridge: Woodhead Publishing, 2007.

[64] B. D. Ribeiro, A. M. de Castro, M. A. Z. Coelho, and D. M. G. Freire, "Production and Use of Lipases in Bioenergy: A Review from the Feedstocks to Biodiesel Production," *Enzyme Research*, vol. 2011, pp. 1–16, 2011.

[65] M. Matori, T. Asahara, and Y. Ota, "Reaction conditions influencing positional specificity index (PSI) of microbial lipases," *Journal of Fermentation and Bioengineering*, vol. 72, pp. 413–415, jan 1991.

[66] S. H. Krishna and N. G. Karanth, "Lipases and Lipase-Catalyzed Esterification Reactions in Nonaqueous Media," *Catalysis Reviews*, vol. 4940, 2016.

[67] A. R. Macrae, "Lipase-catalyzed interesterification of oils and fats," *Journal of the American Oil Chemists' Society*, vol. 60, no. 2, pp. 291–294, 1983.

[68] D. Kahveci and X. Xu, "Repeated hydrolysis process is effective for enrichment of omega 3 polyunsaturated fatty acids in salmon oil by Candida rugosa lipase," *Food Chemistry*, vol. 129, no. 4, pp. 1552–1558, 2011.

[69] O. G. Berg, Y. Cajal, G. L. Butterfoss, *et al.*, "Interfacial Activation of Triglyceride Lipase from Thermomyces (Humicola) lanuginosa : Kinetic Parameters and a Basis for Control of the Lid," *Biochemistry*, vol. 37, no. 19, pp. 6615–27, 1998.

[70] Y. Cajal, a. Svendsen, V. Girona, S. a. Patkar, and M. a. Alsina, "Interfacial control of lid opening in Thermomyces lanuginosa lipase.," *Biochemistry*, vol. 39, pp. 413–23, jan 2000.

[71] Y. Cajal, A. Svendsen, J. D. Bolós, S. Patkar, and M. Alsina, "Effect of the lipid interface on the catalytic activity and spectroscopic properties of a fungal lipase," *Biochimie*, vol. 82, pp. 1053–61, 2000.

[72] I. V. Pavlidis, T. Tsoufis, a. Enotiadis, D. Gournis, and H. Stamatis, "Functionalized Multi-Wall Carbon Nanotubes for Lipase Immobilization," *Advanced Engineering Materials*, vol. 12, pp. B179–B183, feb 2010.

[73] Q. Li, F. Fan, Y. Wang, W. Feng, and P. Ji, "Enzyme Immobilization on Carboxyl-Functionalized Graphene Oxide for Catalysis in Organic Solvent," *Industrial & Engineering Chemistry Research*, vol. 52, pp. 6343–6348, may 2013.

[74] D. Brady and J. Jordaan, "Advances in enzyme immobilisation.," *Biotechnology letters*, vol. 31, pp. 1639–50, nov 2009.

[75] P. R. Burney and J. Pfaendtner, "Structural and dynamic features of Candida rugosa lipase 1 in water, octane, toluene, and ionic liquids BMIM-PF6 and BMIM-NO3.," *The journal of physical chemistry. B*, vol. 117, pp. 2662–70, mar 2013.

[76] Q. Jin, G. Jia, Y. Zhang, Q. Yang, and C. Li, "Hydrophobic Surface Induced Activation of Pseudomonas cepacia Lipase Immobilized into Mesoporous Silica," *Langmuir*, vol. 27, pp. 12016–12024, oct 2011.

[77] K. Mondal, P. Mehta, B. R. Mehta, D. Varandani, and M. N. Gupta, "A bioconjugate of Pseudomonas cepacia lipase with alginate with enhanced catalytic efficiency," *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics*, vol. 1764, pp. 1080–1086, jun 2006.

[78] M. Filice, M. Marciello, L. Betancor, *et al.*, "Hydrolysis of fish oil by hyperactivated Rhizomucor miehei lipase immobilized by multipoint anion exchange.," *Biotechnology progress*, vol. 27, pp. 961–8, jul 2011.

[79] M. T. Neves Petersen, P. Fojan, and S. B. Petersen, *How do lipases and esterases work: the electrostatic contribution.*, vol. 85. feb 2001.

[80] W. F. van Gunsteren and H. J. C. Berendsen, "Computer Simulation of Molecular Dynamics: Methodology, Applications, and Perspectives in Chemistry," *Angewandte Chemie International Edition in English*, vol. 29, no. 9, pp. 992–1023, 1990.

[81] S. C. L. Kamerlin, S. Vicatos, A. Dryga, and A. Warshel, "Coarse-Grained (Multiscale) Simulations in Studies of Biophysical and Chemical Systems," *Annual Review of Physical Chemistry*, vol. 62, pp. 41–64, mar 2011.

[82] M. Karplus and J. A. McCammon, "Molecular dynamics simulations of biomolecules.," *Nature structural biology*, vol. 9, pp. 646–52, sep 2002.

[83] S. a. Adcock and J. A. McCammon, "Molecular dynamics: survey of methods for simulating the activity of proteins.," *Chemical reviews*, vol. 106, pp. 1589–615, may 2006.

[84] C. Gobbo, I. Beurroies, D. de Ridder, *et al.*, "MARTINI Model for Physisorption of Organic Molecules on Graphite," *The Journal of Physical Chemistry C*, vol. 117, pp. 15623–15631, aug 2013.

[85] Y. Tu, M. Lv, P. Xiu, *et al.*, "Destructive extraction of phospholipids from Escherichia coli membranes by graphene nanosheets.," *Nature nanotechnology*, vol. 8, pp. 594–601, aug 2013.

[86] F. Zhu and K. Schulten, "Water and proton conduction through carbon nanotubes as models for biological channels.," *Biophysical journal*, vol. 85, pp. 236–44, jul 2003.

[87] M. Lelimousin and M. S. P. Sansom, "Membrane Perturbation by Carbon Nanotube Insertion: Pathways to Internalization.," *Small*, vol. 9, pp. 3639–3646, feb 2013.

[88] E. J. Wallace and M. S. P. Sansom, "Carbon nanotube/detergent interactions via coarse-grained molecular dynamics.," *Nano letters*, vol. 7, pp. 1923–8, jul 2007.

[89] M. P. Frushicheva and A. Warshel, "Towards quantitative computer-aided studies of enzymatic enantioselectivity: the case of Candida antarctica lipase A.," *Chembiochem : a European journal of chemical biology*, vol. 13, pp. 215–23, jan 2012.

[90] P. Schopf and A. Warshel, "Validating computer simulations of enantioselective catalysis; reproducing the large steric and entropic contributions in Candida Antarctica lipase B.," *Proteins*, pp. 1–14, jan 2014.

[91] M. Ø. Jensen, T. R. Jensen, K. Kjaer, *et al.*, "Orientation and conformation of a lipase at an interface studied by molecular dynamics simulations.," *Biophysical journal*, vol. 83, pp. 98–111, jul 2002.

[92] W. Feng, X. Sun, and P. Ji, "Activation mechanism of Yarrowia lipolytica lipase immobilized on carbon nanotubes," *Soft Matter*, vol. 8, no. 27, p. 7143, 2012.

[93] P. Ji, "Lipase Covalently Attached to Multiwalled Carbon Nanotubes as an Efficient Catalyst in Organic Solvent," *American Institute of Chemical Engineers Journal*, vol. 56, no. 11, pp. 3005–3011, 2010.

[94] E. M. K. Hedin, P. Høyrup, S. a. Patkar, *et al.*, "Implications of surface charge and curvature for the binding orientation of Thermomyces lanuginosus lipase on negatively charged or zwitterionic phospholipid vesicles as studied by ESR spectroscopy.," *Biochemistry*, vol. 44, pp. 16658–71, dec 2005.

[95] E. M. K. Hedin, P. Høyrup, S. a. Patkar, *et al.*, "Interfacial orientation of Thermomyces lanuginosa lipase on phospholipid vesicles investigated by electron spin resonance relaxation spectroscopy.," *Biochemistry*, vol. 41, pp. 14185–96, dec 2002.

[96] A. W. Sonesson, T. H. Callisen, H. Brismar, and U. M. Elofsson, "Adsorption and activity of Thermomyces lanuginosus lipase on hydrophobic and hydrophilic surfaces measured with dual polarization interferometry (DPI) and confocal microscopy," *Colloids and Surfaces B: Biointerfaces*, vol. 61, no. 2, pp. 208–215, 2008.

[97] S. Rehm, P. Trodler, and J. Pleiss, "Solvent-induced lid opening in lipases: a molecular dynamics study.," *Protein science : a publication of the Protein Society*, vol. 19, pp. 2122–30, nov 2010.

[98] P. Trodler, R. D. Schmid, and J. Pleiss, "Modeling of solvent-dependent conformational transitions in Burkholderia cepacia lipase.," *BMC structural biology*, vol. 9, p. 38, jan 2009.

[99] J. J. James, B. S. Lakshmi, A. S. N. Seshasayee, and P. Gautam, "Activation of Candida rugosa lipase at alkane-aqueous interfaces: a molecular dynamics study.," *FEBS letters*, vol. 581, pp. 4377–83, sep 2007.

[100] S. L. Cherukuvada, A. S. N. Seshasayee, K. Raghunathan, S. Anishetty, and G. Pennathur, "Evidence of a double-lid movement in Pseudomonas aeruginosa lipase: insights from molecular dynamics simulations.," *PLoS Computational Biology*, vol. 1, p. e28, aug 2005.

[101] H. McConnell, T. Watts, R. Weis, and A. Brian, "Supported planar membranes in studies of cell-cell recognition in the immune system," *Biochimica et Biophysica Acta (BBA) - Reviews on Biomembranes*, vol. 864, pp. 95–106, jun 1986.

[102] P. E. Milhiet, M. C. Giocondi, O. Baghdadi, *et al.*, "Spontaneous insertion and partitioning of alkaline phosphatase into model lipid rafts," *EMBO Reports*, vol. 3, no. 5, pp. 485–490, 2002.

[103] S. Lenhert, F. Brinkmann, T. Laue, *et al.*, "Lipid multilayer gratings.," *Nature nanotechnology*, vol. 5, no. 4, pp. 275–279, 2010.

[104] P. Reis, K. Holmberg, R. Miller, *et al.*, "Lipase reaction at interfaces as self-limiting processes," *Comptes Rendus Chimie*, vol. 12, pp. 163–170, jan 2009.

[105] C. Xing and R. Faller, "Interactions of lipid bilayers with supports: a coarse-grained molecular simulation study.," *The journal of physical chemistry. B*, vol. 112, no. 23, pp. 7086–7094, 2008.

[106] R. Guo, J. Mao, and L.-T. Yan, "Computer simulation of cell entry of graphene nanosheet.," *Biomaterials*, vol. 34, pp. 4296–301, jun 2013.

[107] A. R. Mhashal and S. Roy, "Self-assembly of phospholipids on flat supports," *Phys. Chem. Chem. Phys.*, vol. 17, pp. 31152–31160, 2015.

[108] J. Mao, R. Guo, and L. T. Yan, "Simulation and analysis of cellular internalization pathways and membrane perturbation for graphene nanosheets," *Biomaterials*, vol. 35, no. 23, pp. 6069–6077, 2014.

[109] A. R. Leach, *Molecular Modelling: Principles and Applications*. Dorset: Pearson Education, 2nd ed., 2001.

[110] D. Frenkel and B. Smit, "Understanding Molecular Simulation: From Algorithms to Applications," 2001.

[111] R. W. Hockney, S. P. Goel, and J. W. Eastwood, "Quiet high-resolution computer models of a plasma," *J. Comput. Phys.*, vol. 14, pp. 148–158, 1974.

[112] C. Oostenbrink, A. Villa, A. E. Mark, and W. F. van Gunsteren, "A biomolecular force field based on the free enthalpy of hydration and solvation: the GROMOS force-field parameter sets 53A5 and 53A6.," *Journal of computational chemistry*, vol. 25, pp. 1656–76, oct 2004.

[113] K. Lindorff-Larsen, S. Piana, K. Palmo, *et al.*, "Improved side-chain torsion potentials for the Amber ff99SB protein force field," *Proteins: Structure, Function and Bioinformatics*, vol. 78, no. 8, pp. 1950–1958, 2010.

[114] D. Kony, W. Damm, S. Stoll, and W. F. Van Gunsteren, "An improved OPLS-AA force field for carbohydrates," *Journal of Computational Chemistry*, vol. 23, no. 15, pp. 1416–1429, 2002.

[115] A. D. Mackerell, "Empirical force fields for biological macromolecules: overview and issues.," *Journal of computational chemistry*, vol. 25, pp. 1584–604, oct 2004.

[116] W. D. Cornell, P. Cieplak, C. I. Bayly, *et al.*, "A Second Generation Force Field for the Simulation of Proteins, Nucleic Acids, and Organic Molecules J. Am. Chem. Soc. 1995, 117, 51795197," *Journal of the American Chemical Society*, vol. 118, p. 2309, jan 1996.

[117] A. MacKerell and D. Bashford, "All-atom empirical potential for molecular modeling and dynamics studies of proteins," *The Journal of Physical Chemistry B*, vol. 5647, no. 97, pp. 3586–3616, 1998.

[118] G. Bussi, D. Donadio, and M. Parrinello, "Canonical sampling through velocity rescaling," *The Journal of chemical physics*, vol. 126, no. 014101, pp. 1–7, 2007.

[119] M. Parrinello and A. Rahman, "Polymorphic transitions in single crystals: A new molecular dynamics method," *Journal of Applied Physics*, vol. 52, no. 12, p. 7182, 1981.

[120] H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, A. DiNola, and J. R. Haak, "Molecular dynamics with coupling to an external bath," *The Journal of Chemical Physics*, vol. 81, no. 8, pp. 3684–90, 1984.

[121] T. Morishita, "Fluctuation formulas in molecular-dynamics simulations with the weak coupling heat bath," *The Journal of Chemical Physics*, vol. 113, no. 8, p. 2976, 2000.

[122] L. Monticelli, S. K. Kandasamy, X. Periole, *et al.*, "The MARTINI goarse-grained force field: extension to proteins," *Journal of Chemical Theory and Computation*, vol. 4, pp. 819–834, may 2008.

[123] M. Parrinello and A. Rahman, "Crystal Structure and Pair Potentials: A Molecular-Dynamics Study," *Physical Review Letters*, vol. 45, pp. 1196–1199, oct 1980.

[124] U. Essmann, L. Perera, M. L. Berkowitz, *et al.*, "A smooth particle mesh Ewald method," *The Journal of Chemical Physics*, vol. 103, no. 19, p. 8577, 1995.

[125] B. Hess, H. Bekker, H. J. C. Berendsen, and J. G. E. M. Fraaije, "LINCS: A linear constraint solver for molecular simulations," *Journal of Computational Chemistry*, vol. 18, no. 12, pp. 1463–1472, 1997.

[126] B. A. Hall, K. B. A. Halim, A. Buyan, B. Emmanouil, and M. S. P. Sansom, "Sidekick for membrane simulations: Automated ensemble molecular dynamics simulations of transmembrane helices," *Journal of Chemical Theory and Computation*, vol. 10, no. 5, pp. 2165–2175, 2014.

[127] B. A. Hall, J. P. Armitage, and M. S. P. Sansom, "Mechanism of bacterial signal transduction revealed by molecular dynamics of Tsr dimers and trimers of dimers in lipid vesicles.," *PLoS Computational Biology*, vol. 8, no. 9, p. e1002685, 2012.

[128] T. Reddy and M. S. P. Sansom, "The Role of the Membrane in the Structure and Biophysical Robustness of the Dengue Virion Envelope," *Structure*, vol. 24, no. 3, pp. 375–382, 2016.

[129] T. Reddy, D. Shorthouse, D. L. Parton, *et al.*, "Nothing to Sneeze At: A Dynamic and Integrative Computational Model of an Influenza A Virion," *Structure*, vol. 23, no. 3, pp. 584–597, 2015.

[130] A. C. Kalli, G. Morgan, and M. S. P. Sansom, "Interactions of the auxilin-1 PTEN-like domain with model membranes result in nanoclustering of phosphatidyl inositol phosphates.," *Biophysical journal*, vol. 105, pp. 137–45, jul 2013.

[131] A. Buyan, A. C. Kalli, and M. S. P. Sansom, "Multiscale Simulations Suggest a Mechanism for the Association of the Dok7 PH Domain with PIP-Containing Membranes," *PLoS Computational Biology*, vol. 12, no. 7, pp. 1–14, 2016.

[132] P. J. Stansfeld, E. E. Jefferys, and M. S. P. Sansom, "Multiscale simulations reveal conserved patterns of lipid interactions with aquaporins.," *Structure (London, England : 1993)*, vol. 21, pp. 810–9, may 2013.

[133] S. J. Marrink, A. H. de Vries, and A. E. Mark, "Coarse grained model for semiquantitative lipid simulations," *The Journal of Physical Chemistry B*, vol. 108, pp. 750–760, jan 2004.

[134] S. J. Marrink, H. J. Risselada, S. Yefimov, D. P. Tieleman, and A. H. de Vries, "The MARTINI force field: coarse grained model for biomolecular simulations.," *The journal of physical chemistry. B*, vol. 111, pp. 7812–24, jul 2007.

[135] D. Van Der Spoel, E. Lindahl, B. Hess, *et al.*, "GROMACS: Fast, flexible, and free," *Journal of Computational Chemistry*, vol. 26, no. 16, pp. 1701–1718, 2005.

[136] R. Baron, A. H. de Vries, P. H. Hünenberger, and W. F. van Gunsteren, "Comparison of Atomic-Level and Coarse-Grained Models for Liquid Hydrocarbons from Molecular Dynamics Configurational Entropy Estimates," *The Journal of Physical Chemistry B*, vol. 110, pp. 8464–8473, apr 2006.

[137] W. Kabsch and C. Sander, "Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features.," *Biopolymers*, vol. 22, pp. 2577–637, dec 1983.

[138] D. de Jong and G. Singh, "Improved parameters for the martini coarse-grained protein force field," *Journal of Chemical ...*, vol. 9, pp. 697–697, 2012.

[139] P. J. Flory, M. Gordon, and N. G. McCrum, "Statistical Thermodynamics of Random Networks [and Discussion]," *Proceedings of the Royal Society of London. A. Mathematical and Physical Sciences*, vol. 351, pp. 351 LP – 380, nov 1976.

[140] M. M. Tirion, "Large Amplitude Elastic Motions in Proteins from a Single-Parameter, Atomic Analysis," *Physical Review Letters*, vol. 77, pp. 1905–1908, aug 1996.

[141] A. J. Rader, C. Chennubhotla, L.-w. Yang, and I. Bahar, "The Gaussian Network Model: theory and applications," *Normal Mode Analysis - theory and applications to biological and chemical systems*, vol. 10, no. 20, pp. 41–64, 2006.

[142] A. R. Atilgan, S. R. Durell, R. L. Jernigan, *et al.*, "Anisotropy of fluctuation dynamics of proteins with an elastic network model.," *Biophys. J.*, vol. 80, pp. 505–515, 2001.

[143] X. Periole, M. Cavalli, S.-J. Marrink, and M. a. Ceruso, "Combining an Elastic Network With a Coarse-Grained Molecular Force Field: Structure, Dynamics, and Intermolecular Recognition," *Journal of Chemical Theory and Computation*, vol. 5, pp. 2531–2543, sep 2009.

[144] J. Wong-Ekkabut, S. Baoukina, W. Triampo, *et al.*, "Computer simulation study of fullerene translocation through lipid membranes.," *Nature nanotechnology*, vol. 3, pp. 363–8, jun 2008.

[145] S. Baoukina, L. Monticelli, and D. P. Tieleman, "Interaction of pristine and functionalized carbon nanotubes with lipid membranes," *Journal of Physical Chemistry B*, vol. 117, no. 40, pp. 12113–12123, 2013.

[146] E. Wallace and M. Sansom, "Blocking of carbon nanotube based nanoinjectors by lipids: A simulation study," *Nano letters*, vol. 8, pp. 2751–6, sep 2008.

[147] D. Sergi, G. Scocchi, and A. Ortona, "Coarse-graining MARTINI model for molecular-dynamics simulations of the wetting properties of graphitic surfaces with non-ionic, long-chain, and T-shaped surfactants.," *The Journal of chemical physics*, vol. 137, pp. 094904–10, sep 2012.

[148] S. O. Yesylevskyy, L. V. Schäfer, D. Sengupta, and S. J. Marrink, "Polarizable water model for the coarse-grained MARTINI force field.," *PLoS computational biology*, vol. 6, p. e1000810, jun 2010.

[149] P. J. Stansfeld and M. S. Sansom, "From Coarse Grained to Atomistic: A Serial Multiscale Approach to Membrane Protein Simulations," *Journal of Chemical Theory and Computation*, vol. 7, pp. 1157–1166, apr 2011.

[150] T. Wassenaar, "Going backward: A flexible geometric approach to reverse transformation from coarse grained to atomistic models," *Journal of Chemical Theory and Computation*, vol. 10, pp. 676–690, 2013.

[151] A. Barducci, M. Bonomi, and M. Parrinello, "Linking well-tempered metadynamics simulations with experiments," *Biophysical Journal*, vol. 98, no. 9, pp. L44–L46, 2010.

[152] F. Palazzesi, A. Barducci, M. Tollinger, and M. Parrinello, "The allosteric communication pathways in KIX domain of CBP.," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 110, pp. 14237–42, aug 2013.

[153] J. Ma, T. C. Flynn, Q. Cui, *et al.*, "A dynamic analysis of the rotation mechanism for conformational change in F1-ATPase," *Structure*, vol. 10, no. 7, pp. 921–931, 2002.

[154] R. a. Böckmann and H. Grubmüller, "Nanoseconds molecular dynamics simulation of primary mechanical energy transfer steps in F1-ATP synthase.," *Nature structural biology*, vol. 9, no. 3, pp. 198–202, 2002.

[155] E. Paci and M. Karplus, "Unfolding proteins by external forces and temperature: the importance of topology and energetics," *Proceedings of the National Academy of Sciences*, vol. 97, pp. 6521–6526, 2000.

[156] B. K. Ho and D. a. Agard, "An improved strategy for generating forces in steered molecular dynamics: The mechanical unfolding of titin, e2lip3 and ubiquitin," *PLoS ONE*, vol. 5, no. 9, 2010.

[157] H. Grubmüller, B. Heymann, and P. Tavan, "Ligand binding: molecular mechanics calculation of the streptavidin-biotin rupture force.," *Science (New York, N.Y.)*, vol. 271, no. 5251, pp. 997–999, 1996.

[158] J. a. Lemkul and D. R. Bevan, "Assessing the stability of Alzheimer's amyloid protofibrils using molecular dynamics," *Journal of Physical Chemistry B*, vol. 114, pp. 1652–1660, 2010.

[159] M. Bonomi, D. Branduardi, G. Bussi, *et al.*, "PLUMED: A portable plugin for free-energy calculations with molecular dynamics," *Computer Physics Communications*, vol. 180, pp. 1961–1972, oct 2009.

[160] G. a. Tribello, M. Bonomi, D. Branduardi, C. Camilloni, and G. Bussi, "PLUMED 2: New feathers for an old bird," *Computer Physics Communications*, vol. 185, no. 2, pp. 604–613, 2014.

[161] B. Hess, C. Kutzer, D. van der Spoel, and E. Lindahl, "GROMACS 4: Algorithms for Highly Efficient, Load-Balanced, and Scalable Molecular Simulation.," *Journal of Chemical Theory and Computation*, vol. 4, no. 3, pp. 435–447, 2008.

[162] N. Schmid, A. P. Eichenberger, A. Choutko, *et al.*, "Definition and testing of the GROMOS force-field versions 54A7 and 54B7.," *European biophysics journal : EBJ*, vol. 40, pp. 843–56, jul 2011.

[163] W. Humphrey, A. Dalke, and K. Schulten, "VMD - Visual Molecular Dynamics," *J Molec. Graphics*, vol. 14, pp. 33–38, 1996.

[164] N. Michaud-Agrawal, E. J. Denning, T. B. Woolf, and O. Beckstein, "MDAnalysis: A toolkit for the analysis of molecular dynamics simulations," *Journal of computational chemistry*, vol. 32, pp. 2319–2327, 2011.

[165] H. S. Ryu, H. K. Kim, W. C. Choi, *et al.*, "New cold-adapted lipase from Photobacterium lipolyticum sp. nov. that is closely related to filamentous fungal lipases.," *Applied microbiology and biotechnology*, vol. 70, pp. 321–6, apr 2006.

[166] S.-K. Jung, D. G. Jeong, M. S. Lee, *et al.*, "Structural basis for the cold adaptation of psychrophilic M37 lipase from Photobacterium lipolyticum.," *Proteins*, vol. 71, pp. 476–84, apr 2008.

[167] N. A. Baker, D. Sept, S. Joseph, M. J. Holst, and J. A. McCammon, "Electrostatics of nanosystems: Application to microtubules and the ribosome," *Proceedings of the National Academy of Sciences*, vol. 98, pp. 10037–10041, aug 2001.

[168] P. Reis, R. Miller, J. Kragel, *et al.*, "Lipases at interfaces: Unique interfacial properties as globular proteins," *Langmuir*, vol. 24, no. 13, pp. 6812–6819, 2008.

[169] K. A. Henzler-Wildman, M. Lei, V. Thai, *et al.*, "A hierarchy of timescales in protein dynamics is linked to enzyme catalysis," *Nature*, vol. 450, pp. 913–916, dec 2007.

[170] S. K. Ramakrishnan, V. Krishna, K. V. Kumar, *et al.*, "Molecular dynamics simulations of lipases," *International journal of integrative biology*, vol. 2, no. 3, pp. 204–213, 2008.

[171] J. Lee, S. W. Suh, and S. Shin, "Computational studies of essential dynamics of Pseudomonas cepacia lipase.," *Journal of biomolecular structure & dynamics*, vol. 18, pp. 297–309, oct 2000.

[172] M. R. Ganjalikhany, B. Ranjbar, A. H. Taghavi, and T. Tohidi Moghadam, "Functional motions of Candida antarctica lipase B: a survey through open-close conformations.," *PloS one*, vol. 7, p. e40327, jan 2012.

[173] D. Zhao, C. Peng, and J. Zhou, "Lipase adsorption on different nanomaterials: a multi-scale simulation study," *Phys. Chem. Chem. Phys.*, vol. 17, no. 2, pp. 840–850, 2015.

[174] E. M. K. Hedin, P. Høyrup, S. a. Patkar, *et al.*, "Interfacial orientation of Thermomyces lanuginosa lipase on phospholipid vesicles investigated by electron spin resonance relaxation spectroscopy.," *Biochemistry*, vol. 41, pp. 14185–96, dec 2002.

[175] S. Páll and B. Hess, "A flexible algorithm for calculating pair interactions on SIMD architectures," *Computer Physics Communications*, vol. 184, pp. 2641–2650, dec 2013.

[176] J. A. Barker and R. O. Watts, "Monte Carlo studies of the dielectric properties of water-like models," *Molecular Physics*, vol. 26, pp. 789–792, sep 1973.

[177] C. Oostenbrink, T. a. Soares, N. F. a. van der Vegt, and W. F. van Gunsteren, "Validation of the 53A6 GROMOS force field.," *European biophysics journal : EBJ*, vol. 34, pp. 273–84, jun 2005.

[178] H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, and J. Hermans, "Interaction Models for Water in Relation to Protein Hydration," in *Intermolecular Forces* (B. Pullman, ed.), vol. 14 of *The Jerusalem Symposia on Quantum Chemistry and Biochemistry*, pp. 331–342, Springer Netherlands, 1981.

[179] W. Huang, Z. Lin, and W. van Gunsteren, "Validation of the GROMOS 54A7 force field with respect to β-peptide folding," *Journal of Chemical Theory and Computation*, vol. 7, pp. 1237–1243, 2011.

[180] A. K. Malde, L. Zuo, M. Breeze, *et al.*, "An Automated force field Topology Builder (ATB) and repository: Version 1.0," *Journal of Chemical Theory and Computation*, vol. 7, pp. 4026–4037, 2011.

[181] S. Canzar, M. El-Kebir, R. Pool, *et al.*, "Charge Group Partitioning in Biomolecular Simulation," *Journal of Computational Biology*, vol. 20, no. 3, pp. 188–198, 2013.

[182] M. Eiteman and J. Goodrum, "Density and viscosity of low-molecular weight triglycerides and their mixtures," *Journal of the American Oil Chemists' Society*, vol. 71, no. 11, pp. 1261–1265, 1994.

[183] G. A. Tribello, M. Bonomi, D. Branduardi, C. Camilloni, and G. Bussi, "PLUMED 2: New feathers for an old bird," *Computer Physics Communications*, vol. 185, pp. 604–613, feb 2014.

[184] M. L. Verdonk, J. C. Cole, M. J. Hartshorn, C. W. Murray, and R. D. Taylor, "Improved protein–ligand docking using GOLD," *Proteins: Structure, Function, and Bioinformatics*, vol. 52, no. 4, pp. 609–623, 2003.

[185] J. W. M. Nissink, C. Murray, M. Hartshorn, *et al.*, "A new test set for validating predictions of protein–ligand interaction," *Proteins: Structure, Function, and Bioinformatics*, vol. 49, no. 4, pp. 457–471, 2002.

[186] O. Korb, T. Stützle, and T. E. Exner, "Empirical Scoring Functions for Advanced ProteinLigand Docking with PLANTS," *Journal of Chemical Information and Modeling*, vol. 49, pp. 84–96, jan 2009.

[187] G. H. Peters, S. Toxvaerd, O. H. Olsen, and A. Svendsen, "Computational studies of the activation of lipases and the effect of a hydrophobic environment.," *Protein engineering*, vol. 10, pp. 137–47, feb 1997.

[188] B. Isralewitz, J. Baudry, J. Gullingsrud, D. Kosztin, and K. Schulten, "Steered molecular dynamics investigations of protein function," *Journal of Molecular Graphics and Modelling*, vol. 19, pp. 13–25, feb 2001.

[189] S. Santini, J. M. Crowet, a. Thomas, *et al.*, "Study of Thermomyces lanuginosa lipase in the presence of tributyrylglycerol and water.," *Biophysical journal*, vol. 96, pp. 4814–25, jun 2009.

[190] F. Bordes, S. Barbe, P. Escalier, *et al.*, "Exploring the conformational states and rearrangements of Yarrowia lipolytica Lipase.," *Biophysical journal*, vol. 99, pp. 2225–34, oct 2010.

[191] V. D. Mouchlis, D. Bucher, J. A. McCammon, and E. a. Dennis, "Membranes serve as allosteric activators of phospholipase A2, enabling it to extract, bind, and hydrolyze phospholipid substrates.," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 2, jan 2015.

[192] G. H. Peters, O. H. Olsen, a. Svendsen, and R. C. Wade, "Theoretical investigation of the dynamics of the active site lid in Rhizomucor miehei lipase.," *Biophysical journal*, vol. 71, pp. 119–29, jul 1996.

[193] A. Aloulou, D. Puccinelli, A. De Caro, Y. Leblond, and F. Carrière, "A comparative study on two fungal lipases from Thermomyces lanuginosus and Yarrowia lipolytica shows the combined effects of detergents and pH on lipase adsorption and activity.," *Biochimica et biophysica acta*, vol. 1771, pp. 1446–56, dec 2007.

[194] C. C. Gruber and J. Pleiss, "Lipase B from Candida antarctica binds to hydrophobic substrate–water interfaces via hydrophobic anchors surrounding the active site entrance," *Journal of Molecular Catalysis B: Enzymatic*, vol. 84, pp. 48–54, dec 2012.

[195] C. N. Lumb and M. S. P. Sansom, "Defining the membrane-associated state of the PTEN tumor suppressor protein.," *Biophysical journal*, vol. 104, pp. 613–21, feb 2013.

[196] C. L. Wee, K. Balali-Mood, D. Gavaghan, and M. S. P. Sansom, "The interaction of phospholipase A2 with a phospholipid bilayer: coarse-grained molecular dynamics simulations.," *Biophysical journal*, vol. 95, pp. 1649–57, aug 2008.

[197] S. Barbe, V. Lafaquière, D. Guieysse, *et al.*, "Insights into lid movements of Burkholderia cepacia lipase inferred from molecular dynamics simulations.," *Proteins*, vol. 77, pp. 509–23, nov 2009.

[198] F. Secundo, G. Carrea, C. Tarabiono, *et al.*, "The lid is a structural and functional determinant of lipase activity and selectivity," *Journal of Molecular Catalysis B: Enzymatic*, vol. 39, pp. 166–170, may 2006.

[199] S. Bezzine, F. Ferrato, M. G. Ivanova, *et al.*, "Human pancreatic lipase: Colipase dependence and interfacial binding of lid domain mutants," *Biochemistry*, vol. 38, no. 17, pp. 5499–5510, 1999.

[200] X.-W. Yu, S.-S. Zhu, R. Xiao, and Y. Xu, "Conversion of a Rhizopus chinensis lipase into an esterase by lid swapping," *The Journal of Lipid Research*, vol. 55, no. 6, pp. 1044–1051, 2014.

[201] G. H. Peters, A. Svendsen, H. Langberg, *et al.*, "Glycosylation of Thermomyces lanuginosa lipase enhances surface binding towards phospholipids, but does not significantly influence the catalytic activity," *Colloids and Surfaces B: Biointerfaces*, vol. 26, pp. 125–134, sep 2002.

[202] C. Pinholt, M. Fanø, C. Wiberg, *et al.*, "Influence of glycosylation on the adsorption of Thermomyces lanuginosus lipase to hydrophobic and hydrophilic surfaces," *European Journal of Pharmaceutical Sciences*, vol. 40, no. 4, pp. 273–281, 2010.

[203] K. E. McAuley, A. Svendsen, S. A. Patkar, and K. S. Wilson, "Structure of a feruloyl esterase from Aspergillus niger," *Acta Crystallographica Section D: Biological Crystallography*, vol. 60, no. 5, pp. 878–887, 2004.

[204] M. Martinelle, M. Holmquist, and K. Hult, "On the interfacial activation of Candida antarctica lipase A and B as compared with Humicola lanuginosa lipase," *Biochemica et Biophysica Acta*, vol. 1258, pp. 272–276, 1995.

[205] J. E. Mogensen, P. Sehgal, and D. E. Otzen, "Activation, inhibition, and destabilization of Thermomyces lanuginosus lipase by detergents," *Biochemistry*, vol. 44, no. 5, pp. 1719–1730, 2005.

[206] J. Skjold-Jørgensen, J. Vind, A. Svendsen, and M. J. Bjerrum, "Altering the Activation Mechanism in Thermomyces lanuginosus Lipase," *Biochemistry*, vol. 53, pp. 4152–4160, 2014.

[207] F. O. Aliwan, P. A. Kroon, C. B. Faulds, R. Pickersgill, and G. Williamson, "Ferulic acid esterase-III from Aspergillus niger does not exhibit lipase activity," *Journal of the Science of Food and Agriculture*, vol. 79, pp. 457–459, mar 1999.

[208] R. Valivety, P. Halling, and A. Macrae, "Water as a competitive inhibitor of lipase-catalysed esterification in organic media," *Biotechnology letters*, vol. 15, no. 11, pp. 1133–1138, 1993.

[209] T. Vuorela, A. Catte, P. S. Niemelä, *et al.*, "Role of lipids in spheroidal high density lipoproteins.," *PLoS computational biology*, vol. 6, p. e1000964, jan 2010.

[210] R. A. Wahab, "Engineering catalytic efficiency of thermophilic lipase from Geobacillus zalihae by hydrophobic residue mutation near the catalytic pocket," *Advances in Bioscience and Biotechnology*, vol. 03, no. 02, pp. 158–167, 2012.

[211] G. Wang, Z. Liu, L. Xu, H. Zhang, and Y. Yan, "Probing role of key residues in the divergent evolution of Yarrowia lipolytica lipase 2 and Aspergillus niger eruloyl esterase A," *Microbiological Research*, vol. 178, pp. 27–34, 2015.

[212] I. V. Pavlidis, T. Vorhaben, D. Gournis, *et al.*, "Regulation of catalytic behaviour of hydrolases through interactions with functionalized carbon-based nanomaterials," *Journal of Nanoparticle Research*, vol. 14, pp. 1–10, apr 2012.

[213] D. Du, Y. Yang, and Y. Lin, "Graphene-based materials for biosensing and bioimaging," *MRS Bulletin*, vol. 37, no. 12, pp. 1290–1296, 2012.

[214] S. Goenka, V. Sant, and S. Sant, "Graphene-based nanomaterials for drug delivery and tissue engineering," *Journal of Controlled Release*, vol. 173, pp. 75–88, jan 2014.

[215] D. Bitounis, H. Ali-Boucetta, B. Hee Hong, M. Dal-Hee, and K. Kostarelos, "Prospects and challenged of graphene in biomedical applications," *Advanced Materials*, vol. 25, no. 16, pp. 2258–2268, 2013.

[216] K. Novoselov, A. Geim, S. Morozov, *et al.*, "Electric Field Effect in Atomically Thin Carbon Films," *Science*, vol. 306, pp. 666–668, 2004.

[217] K. S. Novoselov, D. Jiang, F. Schedin, *et al.*, "Two-dimensional atomic crystals," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 102, no. 30, pp. 10451–10453, 2005.

[218] A. Lerf, H. He, M. Forster, and J. Klinowski, "Structure of Graphite Oxide Revisited," *Journal of Physical Chemistry B*, vol. 102, no. 23, pp. 4477–4482, 1998.

[219] I. V. Pavlidis, T. Vorhaben, T. Tsoufis, *et al.*, "Development of effective nanobiocatalytic systems through the immobilization of hydrolases on functionalized carbon-based nanomaterials.," *Bioresource technology*, vol. 115, pp. 164–71, jul 2012.

[220] F. De Leo, A. Magistrato, and D. Bonifazi, "Interfacing proteins with graphitic nanomaterials: from spontaneous attraction to tailored assemblies," *Chem. Soc. Rev.*, 2015.

[221] R. a. Silva, M. L. Souza, G. D. Bloisi, P. Corio, and D. F. S. Petri, "Bioconjugation of lipase and cholesterol oxidase with graphene or graphene oxide," *Journal of Nanoparticle Research*, vol. 17, no. 4, 2015.

[222] A. S. Campbell, C. Dong, F. Meng, *et al.*, "Enzyme catalytic efficiency: A function of bio-nano interface reactions," *ACS Applied Materials and Interfaces*, vol. 6, pp. 5393–5403, 2014.

[223] Y. Zhang, J. Zhang, X. Huang, *et al.*, "Assembly of graphene oxide-enzyme conjugates through hydrophobic interaction," *Small*, vol. 8, no. 1, pp. 154–159, 2012.

[224] J. Zhang, F. Zhang, H. Yang, *et al.*, "Graphene oxide as a matrix for enzyme immobilization.," *Langmuir : the ACS journal of surfaces and colloids*, vol. 26, pp. 6083–5, may 2010.

[225] L. Monticelli, "On Atomistic and Coarse-Grained Models for C60 Fullerene," *Journal of Chemical Theory and Computation*, vol. 8, no. 1, pp. 1370–1378, 2012.

[226] J. G. Vilhena, C. Pimentel, P. Pedraz, *et al.*, "Atomic-Scale Sliding Friction on Graphene in Water," *ACS Nano*, p. acsnano.5b07825, 2016.

[227] K. Suzuki, N. Oyabu, K. Kobayashi, K. Matsushige, and H. Yamada, "Atomic-resolution imaging of graphite-water interface by frequency modulation atomic force microscopy," *Applied Physics Express*, vol. 4, no. 12, 2011.

[228] M. C. Gordillo and J. Martí, "Water on graphene surfaces.," *Journal of physics. Condensed matter : an Institute of Physics journal*, vol. 22, p. 284111, jul 2010.

[229] K. C. Jena and D. K. Hore, "Water structure at solid surfaces and its implications for biomolecule adsorption.," *Physical chemistry chemical physics : PCCP*, vol. 12, no. 43, pp. 14383–14404, 2010.

[230] a. K. Geim and K. S. Novoselov, "The rise of graphene.," *Nature materials*, vol. 6, pp. 183–91, mar 2007.

[231] W. Cai, R. D. Piner, F. J. Stadermann, *et al.*, "Synthesis and solid-state NMR structural characterization of 13C-labeled graphite oxide.," *Science (New York, N.Y.)*, vol. 321, pp. 1815–7, sep 2008.

[232] B. C. Brodie, "On the Atomic Weight of Graphite," *Philosophical Transactions of the Royal Society of London*, vol. 149, no. 9, pp. 249–259, 1859.

[233] W. S. Hummers and R. E. Offeman, "Preparation of Graphitic Oxide," *Journal of the American Chemical Society*, vol. 80, no. 6, pp. 1339–1339, 1958.

[234] T. Szabo, O. Berkesi, P. Forgo, *et al.*, "Evolution of Surface Functional Groups in a Series of Progressively Oxidized Graphite Oxides Evolution of Surface Functional Groups in a Series of Progressively Oxidized Graphite Oxides," *Chem. Mater.*, vol. 18, no. 11, pp. 2740–2749, 2006.

[235] T. Nakajima and Y. Matsuo, "Formation process and structure of graphite oxide," *Carbon*, vol. 32, no. 3, pp. 469–475, 1994.

[236] T. Cassagneau, F. Gurin, and J. H. Fendler, "Preparation and Characterization of Ultrathin Films Layer-by-Layer Self-Assembled from Graphite Oxide Nanoplatelets and Polymers Preparation and Characterization of Ultrathin Films Layer-by-Layer Self-Assembled from Graphite Oxide Nanoplatelets and Polyme," *Langmuir*, vol. 16, no. 18, pp. 7318–7324, 2000.

[237] W. Yang, Q. Ren, Y.-N. Wu, *et al.*, "Surface functionalization of carbon nanomaterials by self-assembling hydrophobin proteins.," *Biopolymers*, vol. 99, pp. 84–94, jan 2013.

[238] P. Laaksonen, M. Kainlauri, T. Laaksonen, *et al.*, "Interfacial engineering by proteins: exfoliation and functionalization of graphene by hydrophobins.," *Angewandte Chemie (International ed. in English)*, vol. 49, pp. 4946–9, jul 2010.

[239] Z. A. R. Szilvay, H. Kaljunen, and J. Hakanpa, "Two crystal structures of Trichoderma reesei hydrophobin HFBI — The structure of a protein amphiphile with and without detergent interaction," *Protein science*, vol. 15, no. Ebbole 1997, pp. 2129–2140, 2006.

[240] A. Einstein, "Über die von der molekularkinetischen Theorie der Wärme geforderte Bewegung von in ruhenden Flüssigkeiten suspendierten Teilchen," *Annalen der Physik*, vol. 322, pp. 549–560, jan 1905.

[241] I. V. Pavlidis, M. Patila, U. T. Bornscheuer, D. Gournis, and H. Stamatis, "Graphene-based nanobiocatalytic systems: Recent advances and future prospects," *Trends in Biotechnology*, vol. 32, no. 6, pp. 312–320, 2014.

[242] J. Guo, X. Yao, L. Ning, *et al.*, "The adsorption mechanism and induced conformational changes of three typical proteins with different secondary structural features on graphene," *RSC Advances*, vol. 4, no. 20, pp. 9953–9962, 2014.

[243] A. Rodríguez and P. Jiménez, "Some new aspects of graphite oxidation at 0°c in a liquid medium. A mechanism proposal for oxidation to graphite oxide," *Carbon*, vol. 24, pp. 163–167, jan 1986.

[244] D. Hadzi and A. Novak, "Infra-red spectra of graphitic oxide," *Transactions of the Faraday Society*, vol. 51, no. 0, pp. 1614–1620, 1955.

[245] C. M. Nakano, H. Ma, and T. Wei, "Study of lysozyme mobility and binding free energy during adsorption on a graphene surface," *Applied Physics Letters*, vol. 106, no. 15, p. 153701, 2015.

[246] T. Wei, M. A. Carignano, and I. Szleifer, "Lysozyme adsorption on polyethylene surfaces: Why are long simulations needed?," *Langmuir*, vol. 27, no. 19, pp. 12074–12081, 2011.

[247] T. Wei, M. A. Carignano, and I. Szleifer, "Molecular dynamics simulation of lysozyme adsorption/desorption on hydrophobic surfaces," *Journal of Physical Chemistry B*, vol. 116, no. 34, pp. 10189–10194, 2012.

[248] J. Chen, X. Wang, C. Dai, S. Chen, and Y. Tu, "Adsorption of GA module onto graphene and graphene oxide: A molecular dynamics simulation study," *Physica E: Low-dimensional Systems and Nanostructures*, vol. 62, pp. 59–63, 2014.

[249] G. Raffaini and F. Ganazzoli, "Protein adsorption on a hydrophobic surface: a molecular dynamics study of lysozyme on graphite.," *Langmuir : the ACS journal of surfaces and colloids*, vol. 26, pp. 5679–89, apr 2010.

[250] L. Baweja, K. Balamurugan, V. Subramanian, and A. Dhawan, "Hydration Patterns of Graphene-Based Nanomaterials (GBNMs) Play a Major Role in the Stability of a Helical Protein: A Molecular Dynamics Simulation Study.," *Langmuir : the ACS journal of surfaces and colloids*, nov 2013.

[251] N. Dragneva, W. B. Floriano, D. Stauffer, *et al.*, "Favorable adsorption of capped amino acids on graphene substrate driven by desolvation effect," *Journal of Chemical Physics*, vol. 139, no. 17, 2013.

[252] Y.-n. Guo, X. Lu, J. Weng, and Y. Leng, "Density Functional Theory Study of the Interaction of Arginine-Glycine-Aspartic Acid with Graphene, Defective Graphene, and Graphene Oxide," *The Journal of Physical Chemistry C*, vol. 117, no. 11, pp. 5708–5717, 2013.

[253] W. Qin, X. Li, W.-W. Bian, X.-J. Fan, and J.-Y. Qi, "Density functional theory calculations and molecular dynamics simulations of the adsorption of biomolecules on graphene surfaces.," *Biomaterials*, vol. 31, no. 5, pp. 1007–1016, 2010.

[254] R. D. Piner, J. Zhu, F. Xu, S. Hong, and C. A. Mirkin, ""Dip-Pen" Nanolithography," *Science*, vol. 283, pp. 661–663, jan 1999.

[255] S. Lenhert, P. Sun, Y. Wang, H. Fuchs, and C. A. Mirkin, "Massively parallel dip-pen nanolithography of heterogeneous supported phospholipid multilayer patterns.," *Small*, vol. 3, pp. 71–75, jan 2007.

[256] C. C. Wu, D. N. Reinhoudt, C. Otto, V. Subramaniam, and A. H. Velders, "Strategies for patterning biomolecules with dip-pen nanolithography," *Small*, vol. 7, no. 8, pp. 989–1002, 2011.

[257] H. Lei, X. Zhou, H. Wu, *et al.*, "Morphology change and detachment of lipid bilayers from the mica substrate driven by graphene oxide sheets.," *Langmuir : the ACS journal of surfaces and colloids*, vol. 30, pp. 4678–83, apr 2014.

[258] R. Frost, G. E. Jönsson, D. Chakarov, S. Svedhem, and B. Kasemo, "Graphene oxide and lipid membranes: interactions and nanocomposite structures.," *Nano letters*, vol. 12, pp. 3356–62, jul 2012.

[259] M. Hirtz, A. Oikonomou, T. Georgiou, H. Fuchs, and A. Vijayaraghavan, "Multiplexed biomimetic lipid membranes on graphene by dip-pen nanolithography.," *Nature communications*, vol. 4, pp. 1–8, jan 2013.

[260] M. Hirtz, A. Oikonomou, N. Clark, *et al.*, "Self-limiting Multiplexed Assembly of Lipid Membranes on Large-area Graphene Sensor Arrays," *Nanoscale*, vol. 00, no. 1, pp. 1–3, 2016.

[261] M. Chavent, A. Vanel, A. Tek, *et al.*, "GPU-accelerated atom and dynamic bond visualization using hyperballs: A unified algorithm for balls, sticks, and hyperboloids," *Journal of Computational Chemistry*, vol. 32, no. 13, pp. 2924–2935, 2011.

[262] M. Abramovitz and I. Stegun, *Handbook of mathematical functions*. Washington: Dover, 10 ed., 1972.

[263] L. S. Vermeer, B. L. de Groot, V. Réat, A. Milon, and J. Czaplicki, "Acyl chain order parameter profiles in phospholipid bilayers: computation from molecular dynamics simulations and comparison with 2H NMR experiments," *European Biophysics Journal*, vol. 36, no. 8, pp. 919–931, 2007.

[264] A. S. Reddy, D. T. Warshaviak, and M. Chachisvilis, "Effect of membrane tension on the physical properties of DOPC lipid bilayer membrane," *Biochimica et Biophysica Acta - Biomembranes*, vol. 1818, no. 9, pp. 2271–2281, 2012.

[265] J.-P. P. Douliez, A. Léonard, E. J. Dufourc, and A. Leonard, "Restatement of order parameters in biomembranes: calculation of C-C bond order parameters from C-D quadrupolar splittings.," *Biophys J*, vol. 68, no. 5, pp. 1727–39, 1995.

[266] J. P. Douliez, a. Léonard, E. J. Dufourc, *et al.*, "Conformational order of DMPC sn-1 versus sn-2 chains and membrane thickness: An approach to molecular protrusion by solid state 2H-NMR and neutron diffraction," *J. Phys. Chem.*, vol. 100, no. 96, pp. 18450 – 18457, 1996.

[267] J. P. Douliez, A. Ferrarini, and E. J. Dufourc, "On the relationship between C-C and C-D order parameters and its use for studying the conformation of lipid acyl chains in biomembranes," *Journal of Chemical Physics*, vol. 109, no. 6, pp. 2513–2518, 1998.

[268] G. R. Kneller, K. Baczynski, and M. Pasenkiewicz-Gierula, "Communication: Consistent picture of lateral subdiffusion in lipid bilayers: Molecular dynamics simulation and exact results," *Journal of Chemical Physics*, vol. 135, no. 14, pp. 2–5, 2011.

[269] K. Tsuzuki, Y. Okamoto, S. Iwasa, *et al.*, "Reduced Graphene Oxide as the Support for Lipid Bilayer Membrane," *Journal of Physics: Conference Series*, vol. 352, p. 012016, 2012.

[270] M. Hirtz, R. Corso, S. Sekula-Neuner, and H. Fuchs, "Comparative height measurements of dip-pen nanolithography-produced lipid membrane stacks with atomic force, fluorescence, and surface-enhanced ellipsometric contrast microscopy.," *Langmuir*, vol. 27, no. 18, pp. 11605–11608, 2011.

[271] Y. Okamoto, K. Tsuzuki, S. Iwasa, *et al.*, "Fabrication of Supported Lipid Bilayer on Graphene Oxide," *Journal of Physics: Conference Series*, vol. 352, p. 012017, 2012.

[272] R. Tero, G. Sazaki, T. Ujihara, and T. Urisu, "Anomalous diffusion in supported lipid bilayers induced by oxide surface nanostructures," *Langmuir*, vol. 27, no. 16, pp. 9662–9665, 2011.

[273] R. Tero, "Substrate Effects on the Formation Process, Structure and Physicochemical Properties of Supported Lipid Bilayers," *Materials*, vol. 5, pp. 2658–2680, dec 2012.

[274] A. Urtizberea and M. Hirtz, "A diffusive ink transport model for lipid dip-pen nanolithography," *Nanoscale*, vol. 7, pp. 15618–15634, 2015.

[275] S. Sekula, J. Fuchs, S. Weg-Remers, *et al.*, "Multiplexed lipid dip-pen nanolithography on subcellular scales for the templating of functional proteins and cell culture," *Small*, vol. 4, no. 10, pp. 1785–1793, 2008.

[276] R. Machan and M. Hof, "Lipid diffusion in planar membranes investigated by fluorescence correlation spectroscopy," *Biochimica et Biophysica Acta - Biomembranes*, vol. 1798, no. 7, pp. 1377–1391, 2010.

[277] L. Renner, T. Osaki, S. Chiantia, *et al.*, "Supported lipid bilayers on spacious and pH-responsive polymer cushions with varied hydrophilicity," *Journal of Physical Chemistry B*, vol. 112, no. 20, pp. 6373–6378, 2008.

[278] T. V. Ratto and M. L. Longo, "Anomalous subdiffusion in heterogeneous lipid bilayers," *Langmuir*, vol. 19, no. 5, pp. 1788–1793, 2003.

[279] A. Benda, M. Benes, V. Marec, A. Lhotsky, and M. Hof, "How To Determine Diffusion Coefficients in Planar Phospholipid Systems by Confocal Fluorescence Correlation Spectroscopy," vol. 9, no. 14, pp. 4120–4126, 2003.

[280] G. J. Schütz, H. Schindler, and T. Schmidt, "Single-molecule microscopy on model membranes reveals anomalous diffusion.," *Biophysical journal*, vol. 73, no. 2, pp. 1073–1080, 1997.

[281] C. Liu and R. Faller, "Conformational, dynamical. and tensional study of tethered bilayer lipid membranes in coarse-grained molecular simulations," *Langmuir*, vol. 28, no. 45, pp. 15907–15915, 2012.

[282] C. Xing and R. Faller, "Coarse-grained simulations of supported and unsupported lipid monolayers," *Soft Matter*, vol. 5, no. 22, p. 4526, 2009.

[283] A. Lamberg and T. Taniguchi, "Coarse-Grained Computational Studies of Supported Bilayers: Current Problems and Their Root Causes," *J Phys Chem B*, vol. 118, pp. 10643–10652, 2014.

[284] S. R. Tabaei, W. B. Ng, S.-J. Cho, and N.-J. Cho, "Controlling the Formation of Phospholipid Monolayer, Bilayer, and Intact Vesicle Layer on Graphene," *ACS Applied Materials & Interfaces*, p. acsami.6b02837, 2016.

[285] F. Wang, B. Liu, A. C.-F. Ip, and J. Liu, "Orthogonal adsorption onto nanographene oxide using different intermolecular forces for multiplexed delivery.," *Advanced Materials*, vol. 25, pp. 4087–4092, 2013.

[286] E. Sackmann, "Supported membranes: Scientific and practical applications," *Science*, vol. 271, no. 5245, pp. 43–48, 2007.

[287] A. Laio and F. L. Gervasio, "Metadynamics: a method to simulate rare events and reconstruct the free energy in biophysics, chemistry and material science," *Reports on Progress in Physics*, vol. 71, no. 12, p. 126601, 2008.

[288] A. Barducci, M. Bonomi, and M. Parrinello, "Metadynamics," *Wiley Interdisciplinary Reviews: Computational Molecular Science*, vol. 1, pp. 826–843, sep 2011.

[289] A. Barducci, G. Bussi, and M. Parrinello, "Well-tempered metadynamics: a smoothly-converging and tunable free-energy method," *Phys Rev Lett*, vol. 100, no. 2, pp. 1–4, 2008.

[290] B. Roux, "The calculation of the potential of mean force using computer simulations," *Computer Physics Communications*, vol. 91, no. 1-3, pp. 275–282, 1995.

[291] J. Chen, L. Chen, Y. Wang, and S. Chen, "Molecular dynamics simulations of the adsorption of DNA segments onto graphene oxide," *Journal of Physics D: Applied Physics*, vol. 47, no. 50, p. 505401, 2014.

[292] A. Bagri, C. Mattevi, M. Acik, *et al.*, "Structural evolution during the reduction of chemically derived graphene oxide.," *Nature chemistry*, vol. 2, pp. 581–7, jul 2010.

[293] D. Stauffer, N. Dragneva, W. B. Floriano, *et al.*, "An atomic charge model for graphene oxide for exploring its bioadhesive properties in explicit water," *The Journal of Chemical Physics*, vol. 141, no. 4, p. 044705, 2014.

[294] N. Patra, B. Wang, and P. Král, "Nanodroplet activated and guided folding of graphene nanostructures.," *Nano letters*, vol. 9, pp. 3766–71, nov 2009.

Appendix A

This appendix presents supplemental data relating to Chapter 3 of this thesis. In Chapter 3, the interfacial activation mechanism of the M37 lipase was explored. The following details the use of enhanced sampling methodologies to assess the overall energetics associated with proposed lid movement and activation of M37.

Methods

Classical and well-tempered metadynamics simulations were performed for M37 positioned in the centre of a simulation and solvated with water (10x10x10 nm). The same AT forcefield, water model, and simulation parameters were applied as presented in Chapter 3.

Classical metadynamics involves adding a history-dependent bias potential in the shape of a Gaussian function to the Hamiltonian of the system, in the space of a few selected degrees of freedom, or the collective variable (CV) [287, 288]. The adapted potential can thus be written as a sum of Gaussian functions deposited along the trajectory of the system, in the space of the CV, discouraging the system from revisiting previously sampled configurations. The metadynamics bias potential can be written

as:

$$V_G(S,t) = \int_0^t dt' \omega \exp\Big(- \sum_{i=1}^{d} \frac{(S_i(R) - S_i(R(t')))^2}{2\sigma_i^2} \Big) \tag{A.1}$$

where S is a function of the coordinates of the system R, ω is an "energy rate" and σ_i is the width of the Gaussian for the ith CV. In classical metadynamics, the energy rate (ω) is constant and expressed in terms of a Gaussian height (W) and the deposition stride τ_{G}:

$$\omega = \frac{W}{\tau_G}$$

Well-tempered metadynamics however involves an adaptation of the classical metadynamics approach in which the bias deposition rate decreases over simulation time. This is achieved in practice by rescaling the Gaussian height W. This is termed the biasfactor and scales the Gaussian deposition rate [289]. The parameters used are presented in the following section.

Enhanced sampling methods for M37 activation

Metadynamics is an enhanced sampling technique that biases a system to visit previously unvisited areas of its free energy landscape [287, 288]. A history-dependent bias potential is added to a CV that describes a particular event, e.g a molecular motion, thus accelerating the sampling of that particular CV and forcing the system to escape from local energy minima. The free energy landscape is then reconstructed in the space of the CV, providing an estimate of the free energy associated with the occurrence of the particular event. Here, metadynamics were applied to probe the free energy landscape of opening of the lid region of M37 in water, switching it from a closed to open state.

The metadynamics simulation system was set up as an extension of previous SMD

simulations of lid opening in M37 in water. Consequently, an identical distance CV describing the COM distance between the lid and active site flap region was applied. It is important to note that the accuracy and convergence of metadynamics simulations relies on selecting a CV that incorporates the slowest degrees of freedom in the system [288].

Initial classical metadynamics (continuous direct metadynamics [288]) simulations were performed for 100 ns to test the properties of the CV and the behaviour of the system. Table A.1 and A.2 depicts all the metadynamics simulations, and variants thereof, performed in this study.

Simulation result	**Method**	**Parameters**
Over extension of lid region and unfolding of lid and active site flap helices – did not complete successfully.	Classical metadynamics	Hill height = 6.27 Sigma = 0.196 Pace = 1500
Unfolding and collapse of lid and active site flap helices – did not complete successfully.	Classical metadynamics	Hill height = 1.25 Sigma = 0.196 Pace = 1500
Overextension and collapse of lid helix – 10 ns.	Classical metadynamics	Hill height = 0.125 Sigma = 0.196 Pace = 1500
Slight increase in CV_d and overall helix stability – 10 ns.	Classical metadynamics with a lower wall at CV_d = 0.7	Hill height = 0.125 Sigma = 0.3 Pace = 5000 Fc lower wall = 1000 kJ mol^{-1}

Table A.1: All the classical metadynamics simulations performed for M37 in water, probing the energy landscape for lid opening. The collect variable (CV_d) used is the COM distance between the lid helix (residues 264-278) and the active site flap (94-110). Fc = force constant.

Simulation result	Method	Parameters
Very slight increase in COM lid and active site flap helix distance (CV) and overall helix stability – 10 ns.	Well-tempered metadynamics	Hill height = 0.1 Sigma = 0.3 Pace = 5000 Biasfactor = 10
Rotation and overextension of lid helix – 20 ns.	Well-tempered metadynamics	Hill height = 0.1 Sigma = 0.3 Pace = 5000 Biasfactor = 20
Increased COM lid and active site flap helix distance, however loss of lid alpha helicity – 40 ns.	Well-tempered metadynamics with upper wall at CV_d = 2.0	Hill height = 0.1 Sigma = 0.3 Pace = 5000 Biasfactor = 20 Fc lower wall = 1000 kJ mol^{-1}
One complete closed-open transition by the lid helix, very little sampling – 100 ns.	Well-tempered metadynamics with upper wall at CV_d = 2.0	Hill height = 0.1 Sigma = 0.3 Pace = 5000 Biasfactor = 20 Fc lower wall = 2000 kJ mol^{-1}
No overall increase in COM distance between lid and active site flap helix – 20 ns.	Well-tempered metadynamics with upper wall at CV_d = 2.0	Hill height = 0.1 Sigma = 0.3 Pace = 3000 Biasfactor = 30 Fc lower wall = 2000 kJ mol^{-1}
Successful opening of lid region, but no sampling of open-closed transition – 20 ns.	Well-tempered metadynamics with upper wall at CV_d = 2.0	Hill height = 0.1 Sigma = 0.3 Pace = 1000 Biasfactor = 20 Fc lower wall = 2000 kJ mol^{-1}
Gradual displacement of lid region to open position, and very gradual return to closed position but unfolding of actie site flap helix and lid helix instability – 250 ns.	Well-tempered metadynamics with upper wall at CV_d = 2.0	Hill height = 0.1 Sigma = 0.3 Pace = 3000 Biasfactor = 20 Fc lower wall = 2000 kJ mol^{-1}

Table A.2: All the well-tempered metadynamics simulations performed for M37 in water, probing the energy landscape for lid opening. The collect variable (CV_d) used is the COM distance between the lid helix (residues 264-278) and the active site flap (94-110). Fc = Force constant.

The initial results suggested that the parameters used in these test runs overflooded the minima associated with M37 lid opening. Consequently, the lid helix and active site flap region lost their α-helical fold and collapsed onto each other within the first nanoseconds of simulation time (Fig. A.1). Variation of the input parameters for subsequent classical metadynamics simulations resulted in similar behaviour of the protein.

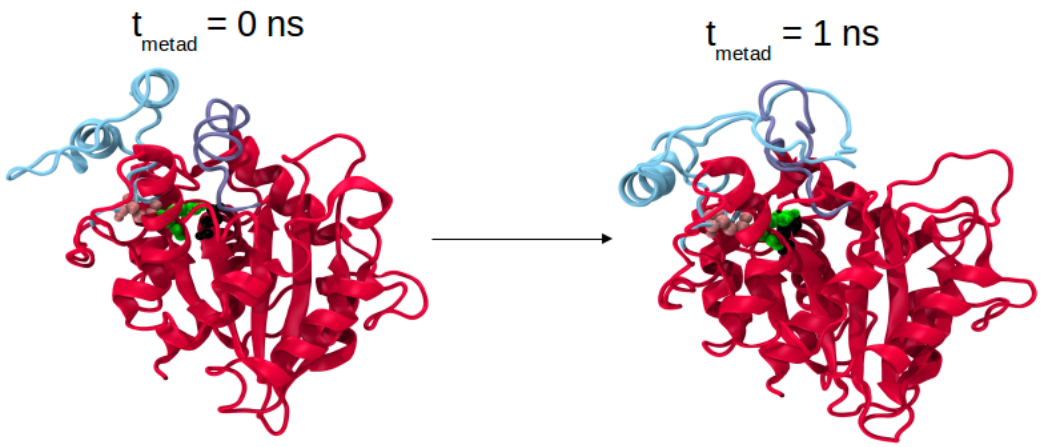

Fig. A.1: Structures of M37 during a classical metadynamics simulation of the protein in water. The CV defined the COM distance between the lid and active site flap helix.

Next, the well-tempered metadynamics ensemble was employed to promote system stability [289]. This method involves automatic scaling of the bias potential deposition rate according to a user-defined scale factor, to avoid possible overflooding of the energetic minima describing lid motion. Initial simulations showed that M37 was much more stable in this ensemble, particularly with respect to lid helix stability. Extension of these simulations up to 500 ns however, resulted in very little overall sampling of the motions involved in lipase activation (Fig. A.2), producing only one successful closed-open transition.

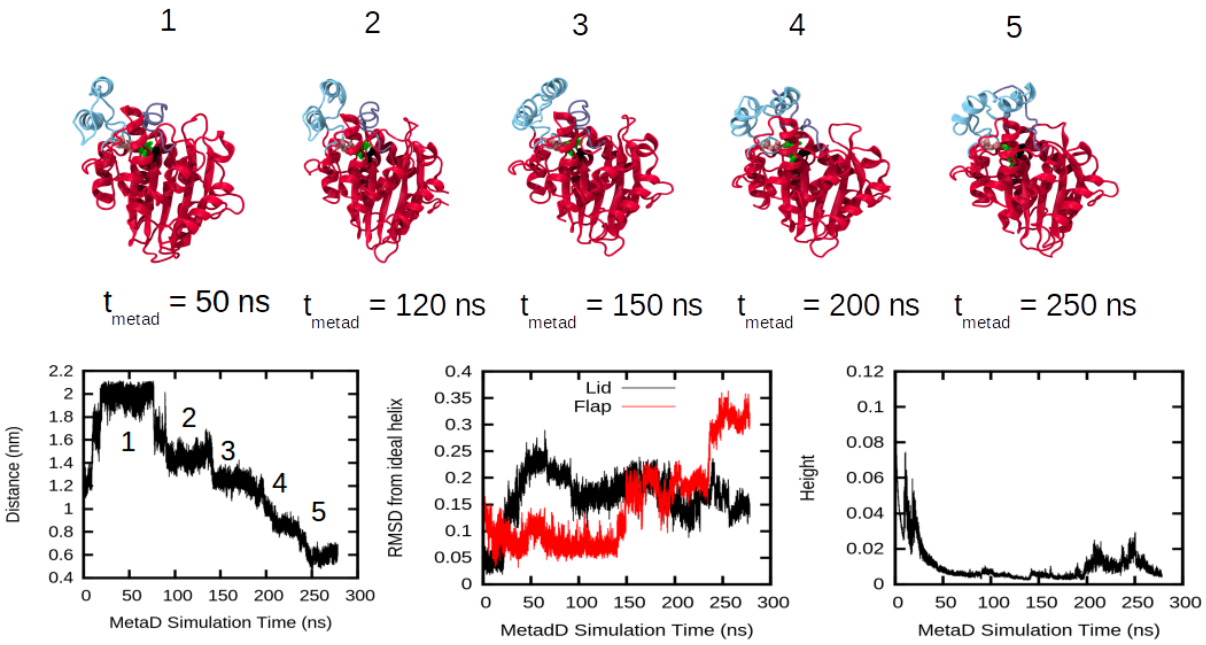

Fig. A.2: Top panel: structures of M37 at different time points during a well-tempered metadynamics simulation of M37 in water. Bottom left panel: Time evolution of the COM distance between the lid and active site flap helices. The numbers correspond to the structures shown in the top panel. Bottom middle panel: The RMS deviation from an ideal α helix calculated for both the lid and active site flap helices over simulation time. Bottom right panel: Time evolution of the gaussian height with a biasfactor of 20.

I next attempted to estimate M37 activation energetics using the common umbrella sampling technique to calculate the potential of mean force (PMF) [290]. Using a successfully completed SMD simulation of lid opening in M37, umbrella sampling was performed on the same trajectory, using windows spaced 0.1 nm apart, simulated for 10 ns each (Fig. A.3). These simulations were also unsuccessful, noticeable from only small overlap in the chosen window set, particularly at larger distances of the lid region relative to the active site flap helix. By monitoring the restrained position of the lid helix within the windows at these larger distances, I deduced that perhaps the force constant used for the restraints was not strong enough to allow for sufficient sampling of the lid helix at these larger distances (Fig. A.4). Applying large force constants

however, resulted in the same small overlap of the windows. It is also very likely that the relatively short simulation time of the windows (10 ns) did not allow for sufficient sampling of lid helix motion at these position, thus affecting the overlap between the windows.

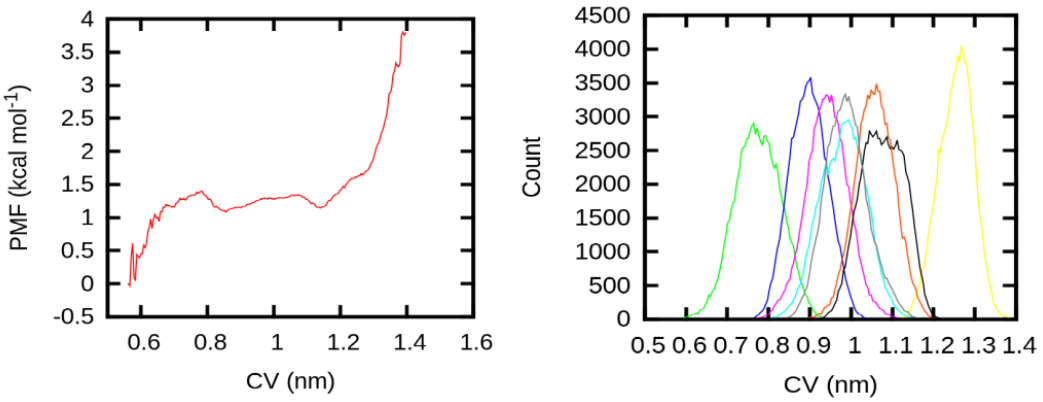

Fig. A.3: Representative structures of M37 at the end of AT-MD simulations (three replicates) with a tributyrin interface, aligned with the closed (crystal structure) and open structures of the lipase. The lid and active site flap are coloured in blue and red respectively for the simulated structure; the same regions are coloured in cyan and magenta for the closed form, and in light blue for the open form of the enzyme.

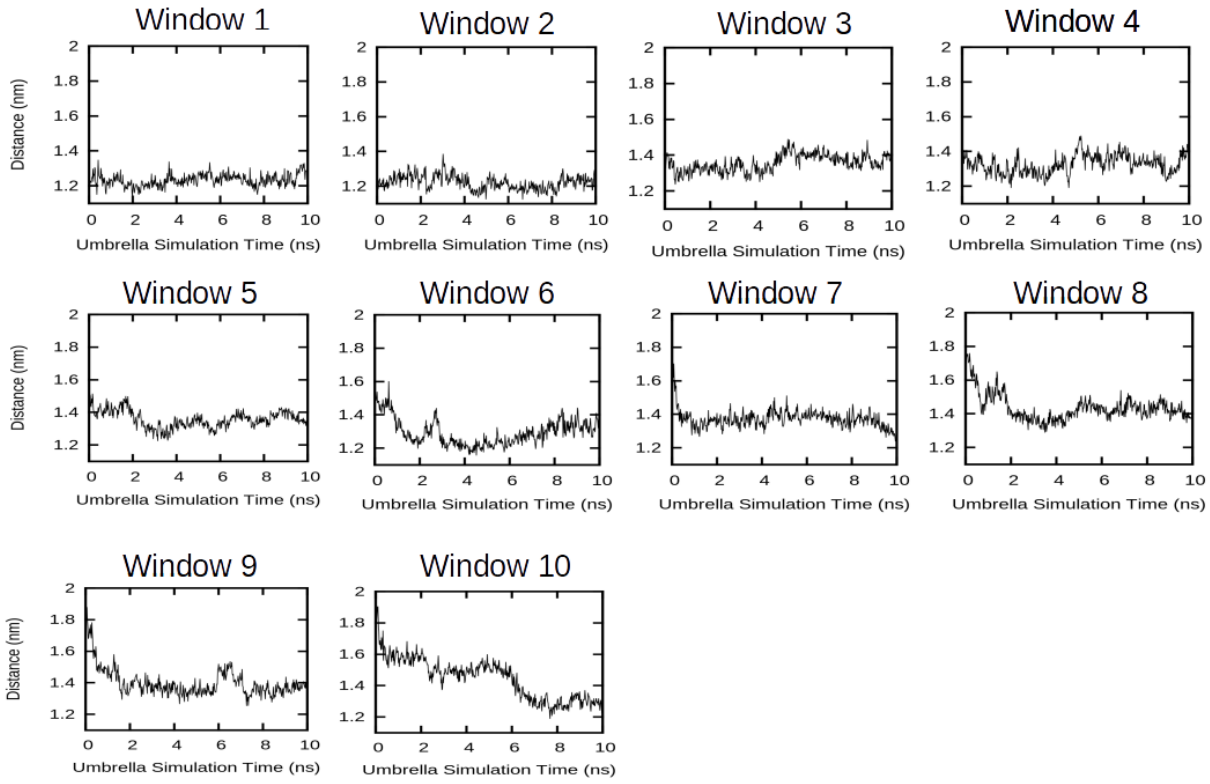

Fig. A.4: Time evolution of the COM distance between the lid and active site flap helices for each of the windows simulated during the umbrella sampling simulation of M37 in water.

The cumulative results from both the metadynamics and umbrella sampling simulations suggest that the underlying free energy landscape describing M37 activation is quite complex. It is possible that the relatively simple distance CV chosen here is not adequate to probe this free energy landscape. Alternatively, the aqueous environment will most likely affect the energetics of the opening process, and it not a true reflection of the interfacial environment that triggers spontaneous interfacial activation of the enzyme. Due to time constraints, I did not continue in this characterisation. It would be of interest however to pursue alternative CV's and enhanced sampling methods to provide insight on the factors governing the energetics of M37 activation. Furthermore, extending these investigations to an interfacially bound M37 system would

provide interesting details about the quantitative differences in the energy required for lid displacement in water relative to an interface.

Appendix B

This appendix presents supplemental data relating to Chapter 5 of this thesis. An atomistic graphene oxide model is parameterised for simulation with the GROMOS 54A7 forcefield. The parameters used for this model are listed in Tables 1-4.

The oxygen composition of the surface was based on a ratio of $C_{10}O_1(OH)_1(COOH)_{0.5}$, used in many computational studies of protein interactions with the oxidised interface [248, 291, 292]. The coordinates for the model were kindly provided by the authors of [293] (Fig. B.1).

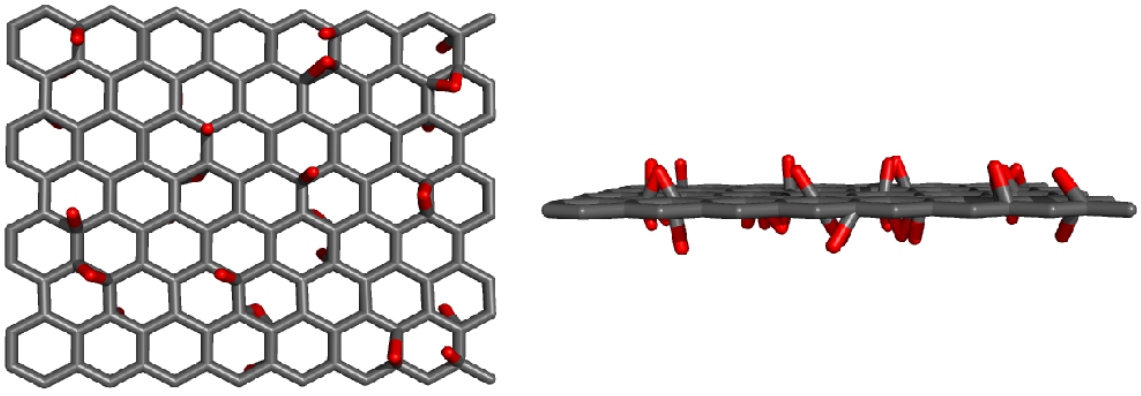

Fig. B.1: Atomistic graphene oxide model displaying a C:O:OH:COOH ratio of 10:1:1:0.5, the coordinates for which were provided by [293]. The surface is displayed as sticks; carbons are coloured in dark grey and oxygen groups in red. The unit extends 1.84 x 1.55 nm in the x and y dimensions. The left image perspective is looking down on the surface; the right image is a side-on view.

An initial topology describing the bonded and non-bonded interactions for this model was generated by parameterising the small functional unit of GO presented in Fig. B.1, using the Automated Topology Builder website (ATB) (https://atb.uq.edu.au). A script was used to alter the parameters of this initial topology based on existing parameters within the GROMOS 54A7 forcefield. A similar approach has been used in previous computational studied of graphene oxide [248, 291, 294]. The partial charges for the oxygen functional groups were take from [293].

AT simulation of the newly parameterised graphene oxide sheet in SPC water resulted in a slight loss in graphene planarity, which is also observed in [293] (Fig. B.2). This affect is attributed to the electrostatic interactions between the oxygen groups on the surface and the coordination change for the functionalised carbon atoms from sp^2 to sp^3 hybridisation [293] .

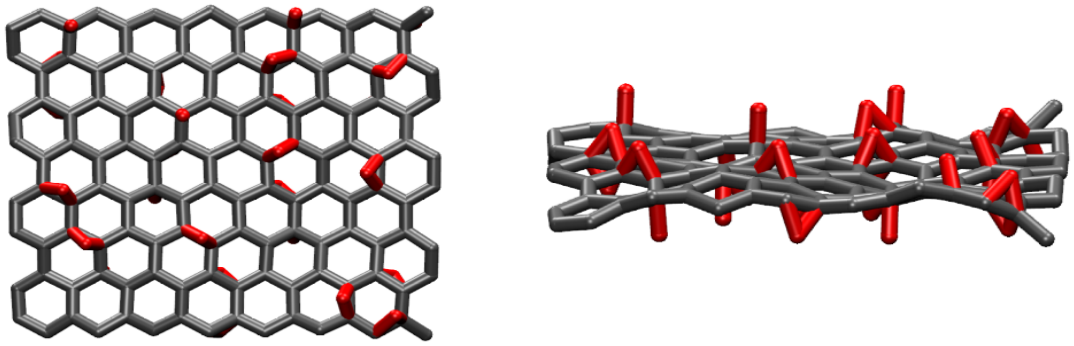

Fig. B.2: 5 ns AT-MD simulation of a newly parameterised graphene oxide model in water (omitted for clarity). The left image perspective is looking down on the surface; the right image is a side-on view.

Atom	Atom	Bond length (nm), Fc (10^6kJ mol^{-1}nm^{-4})	GROMOS building block
C	CH_3	0.153, 7.15	C,CHn-C,CHn
C	CH_2	0.139, 10.8	C,CR1-CH2,C,CR1 (6-ring)
C	CH_1	0.139, 10.8	C,CR1-CH2,C,CR1 (6-ring)
C	O(H)	0.143, 8.18	ATB parameters
C	O(C)	0.143, 6.16	ATB parameters
C	H	0.109, 12.3	HC-C
O	H	0.100, 15.7	H-OA

Table B.1: Bond interaction parameters for a graphene oxide unit (Fig. 5.16) based on [112] and ATB assigned parameters

.

Atom	Atom	Atom	Angle (°), Fc (kJ mol^{-1})	GROMOS building block
C	C	C	120, 560	N,C,CR1 (6-ring, no H)
H	C	C	120, 505	HC-6-ring
C	O(H)	C	110, 530	ATB parameters
C	O(H)	H	109.5, 450	X-OA
C	O(C)	C	Differ	ATB parameters
H	C	O(C)	Differ	ATB parameters

Table B.2: Angle interaction parameters for a graphene oxide unit (Fig. 5.16) based on [112] and ATB assigned parameters

.

Atom	Atom	Atom	Atom	Angle (°), Fc (kJ mol^{-1}), M	GROMOS building block
C	C	C	C	0, 3.77, 6	-CHn-C(ring)-
C	C	C	O(C)	180, 1.00, 3	ATB parameters
C	C	O(H)	H	0, 1.26, 4	-CHn-OA-

Table B.3: Torsional dihedral interaction parameters for a graphene oxide unit (Fig. 5.16) based on [112] and ATB assigned parameters. M = multiplicity.

.

Atom	Atom	Atom	Atom	Angle (°), Fc (kJ mol^{-1} degree^{-2})	GROMOS building block
C	C	C	C	0.0, 0.051	planar groups

Table B.4: Improper dihedral interaction parameters for a graphene oxide unit (Fig. 5.16) based on [112] and ATB assigned parameters

.