On the Chemical Reactivity of Molecules with Membrane Lipids: An Experimental and Theoretical Approach

YESHURUN AMARASINGHAM TANNA

On the Chemical Reactivity of Molecules with Membrane Lipids: An Experimental and Theoretical Approach

Yeshurun Amarasingham Tanna

Department of Chemistry

University of Durham, United Kingdom

A thesis presented for the degree of

Doctor of Philosophy

June 2025

Abstract

It is now well-established that the reactivity of membrane associated molecules, such as CADs (cationic amphiphilic drugs), with membrane lipids is influenced by the disposition of each molecule in the membrane rather than its lipophilicity. The reactions involving CADs and phospholipids such as POPC include lipidation and hydrolysis. Some CADs promote lipid hydrolysis, whereas others either decree the rate or have no effect on lipid chemical stability. Lipidation involves the transfer of an acyl group from the glycerol part of the lipid to a reactive group on the CAD. Lipidation and hydrolysis reactions both lead to the formation of lysolipids. This project aims to investigate the factors of drug reactivity towards lipid membranes.

This project involves both experimental and theoretical work to understand drug reactivity factors with POPC lipids. A range of isotopically enriched reactive compounds have been synthesised. These compounds vary in the rates of lysolipid formation that they induce in POPC membranes. Furthermore, a novel synthetic methodology has been developed to incorporate an ^{15}N isotope into molecules with aniline functionality, and its applicability to clinically relevant molecules. 1D and 2D solid-state NMR data is presented which enables the interactions of each labelled molecule with POPC lipids to be probed via close contact distance measurements. Atomistic molecular dynamics (MD) simulations provide corroborating information on the preferred depth and orientation of the molecules in the lipid bilayer. QM/MM simulations are used to locate the reactive intermediates and the transition state in each bond forming and bond breaking process along the reaction pathway. These calculations were successful in being able to predict reactive and non-reactive conformations of drugs in the membrane interface. The MD and DFT results correlate well with ssNMR data in showing how orientation and depth of partitioning influence drug-lipid reactivity.

Declaration

I declare that the content of this thesis has not previously been submitted for a degree at Durham University or any other university. The research described in this thesis has been conducted solely by the author.

Acknowledgements

First and foremost I would like to thank my supervisors, John Sanderson and Mark Wilson. Their guidance, support and encouragement have been of the greatest value when conducting both research and in the writing of this thesis. I would like to thank members of the Sanderson group, including our former post-doc Genevieve Duche and PhD student Bethany Henderson. Gen provided insightful discussions into both synthetic work and biophysical methods. Beth also contributed to helpful discussions on synthetic work. Thanks also is given to Sam Page, head of solid-state NMR, who was remarkably helpful in running ssNMR experiments on my samples. His insight and detailed conversations provided enormous help with ssNMR analysis. I would like to thank members of the Wilson Group, including former post-doc Fred Pontes, and fellow PhD students Jack Taylor and Ajeeth Kanagarajan. Fred was paramount in aiding my computational methods and subsequent analysis of results, Jack was helpful in useful DFT conversations when discussing functionals and theory, and Ajeeth was instrumental to good practices on python scripting.

Outside of academia, I thank firstly my Mother, Father, Brother and Grandma, who were and continue to be great sources of love, support and comfort to me whilst writing this thesis. I also thank my undergraduate university friends, particularly Robert Simpson, Pete Baker, Kristian Herath and Rushil Chugani for always being supportive and providing encouragement during difficult seasons of research. Finally, I thank my postgraduate university friends, particularly Roo Birch, Alex Keith, Alan Dodgson, Ian Dolby and other ChristChurch Durham friends for being calm and supportive. Above all else, I thank Almighty God who has given me strength, resilience and determination to complete this PhD research and thesis.

"Great are the works of the Lord, studied by all who delight in them." Psalm 111:2

List of Abbreviations

AA All-Atom

AMBER Assisted Model Building with Energy Refinement

B3LYP Becke, 3-parameter, Lee–Yang–Parr

BD Bürgi-Dunitz

BOC *tert*-butyloxycarbonyl

CAD Cationic Amphiphilic Drug

CHARMM Chemistry at Harvard Macromolecular Mechanics

COM Centre of Mass

CPU Central Processing Unit

CSA Chemical Shift Anisotropy

DFT Density Functional Theory

DIPL Drug Induced Phospholipidosis

DMAP Dimethylaminopyridine

DMPC Dimyristoylphosphatidylcholine

DOPC 1,2-dioleoyl-*sn*-glycero-3-phosphocholine

DPPC 1,2-dipalmitoylphosphatidylcholine

EA Ethyl Acetate

EDAC 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide

Et_3N Triethylamine

GAFF General AMBER Force Field

G-D3 Grimme D3 Dispersion

GPU Graphics Processing Unit

HBF_4 Tetrafluoroboric acid

HMBC Heteronuclear Multiple Bond Correlation

HOBt 1-hydroxybenzotriazole

HSQC Heteronuclear Single Quantum Coherence

LC-MS Liquid chromatography-Mass Spectrometry

m/z mass to charge ratio

MAS Magic-angle Spinning

MD Molecular Dynamics

MM Molecular Mechanics

MsCl Methanesulfonyl chloride

NMR Nuclear Magnetic Resonance

nOe nuclear-Ovenhauser effect

NOESY Nuclear Overhauser Effect Spectroscopy

OL Oleic Acid

ONIOM Our own N-layered Integrated molecular Orbital and Molecular mechanics

PA Palmitic Acid

pbc Periodic Boundary Conditions

PBE0 Perdew–Burke–Ernzerhof 0

PC Phosphatidylcholine

PE Petroleum Ether

PM6 Parameterisation Method 6

PMF Potential of Mean Force

POPC 1-palmitoyl-2-oleoyl-*sn*-glycero-3-phosphocholine

PUFA Polyunsaturated Fatty Acids

QM Quantum Mechanics

QM/MM Quantum Mechanics/Molecular Mechanics

RDF Radial Distribution Function

RF Radio-frequency

ssNMR Solid State NMR

TBN *tert*-butyl nitrite

TLC Thin-layer Chromatography

TS Transition State

TsCl 4-Toluenesulfonyl chloride

UV-Vis Ultraviolet–Visible

WHAM Weighted Histogram Analysis Method

Contents

Chapter 1

Introduction

1.1 Motivation

Lipids are a diverse group of biomolecules derived from fatty acids that have significant roles in cellular function. They include sterols, phospholipids, fats and waxes. Lipids are insoluble in water but soluble in organic solvents. Phospholipids and sterols form the structure of the semi-permeable barrier, known as the cell membrane, which provides stability, function and compartmentalisation for the cell[53]. Furthermore, lipids play an essential role in both cell physiology and pathology. An example would be the role of lipid rafts, composed of sphingolipids, to promote bacterial phagocytosis during infection[69]. Membranes provide the cellular platform for a plethora of other biomolecules to attach themselves onto. Lipids contribute to cell signalling by glycoproteins, stability by glycolipids forming lipid rafts with cholesterol, and control of the passage of substances across the bilayer[113]. Whilst historically, the cell membrane was thought to be a chemically stable and inert environment with very few interactions between the phospholipids and proteins embedded within the bilayer, this assertion in more recent years has been challenged. New evidence has pointed to intrinsic reactivity of peptides, proteins and low molecular weight organic compounds, such as drugs, with membrane lipids[54]. The reactivity of these molecules with membrane lipids is structure dependent. The motivation of this project is to understand the contributing factors of drug reactivity towards the cell membrane.

1.2 Introduction to Membranes

1.2.1 Structures and Properties of Lipids

Lipids are broadly categorised as either hydrophobic or amphiphilic molecules. All lipids possess a hydrophobic core as part of their structure, formed from fatty acids in the case of phospholipids[62] (see table 1.1, figure 1.1). Lipids can be classified as either simple or complex. Simple lipids are those which, upon hydrolysis reactions, form a maximum of two products, whilst complex lipids are those which hydrolyse to form a maximum of three products. In the former class, the hydrolysis products are fatty acids and glycerol, whilst in the latter class, the products are glycerol, fatty acids and phospholipids[193].

Lipid Class	Example of Lipid	Chemical Name
Fatty Acyls	**Oleic Acid**	(9Z)-Octadec-9-enoic acid
Glycerolipids	**Glyceryl Tristearate**	Propane-1,2,3-triyl tri(octadecanoate)
Glycerophospholipids	**DOPC**	1,2-Dipalmitoylphosphatidylcholine
Sterol Lipids	**Cholesterol**	Cholest-5-en-3β-ol
Prenol Lipids	**Farnesol**	(2E,6E)-3,7,11-Trimethyldodeca-2,6,10-trien-1-ol
Sphingolipids	**Sphingosine**	(2S,3R,4E)-2-Aminooctadec-4-ene-1,3-diol

Table 1.1: Classes of lipids

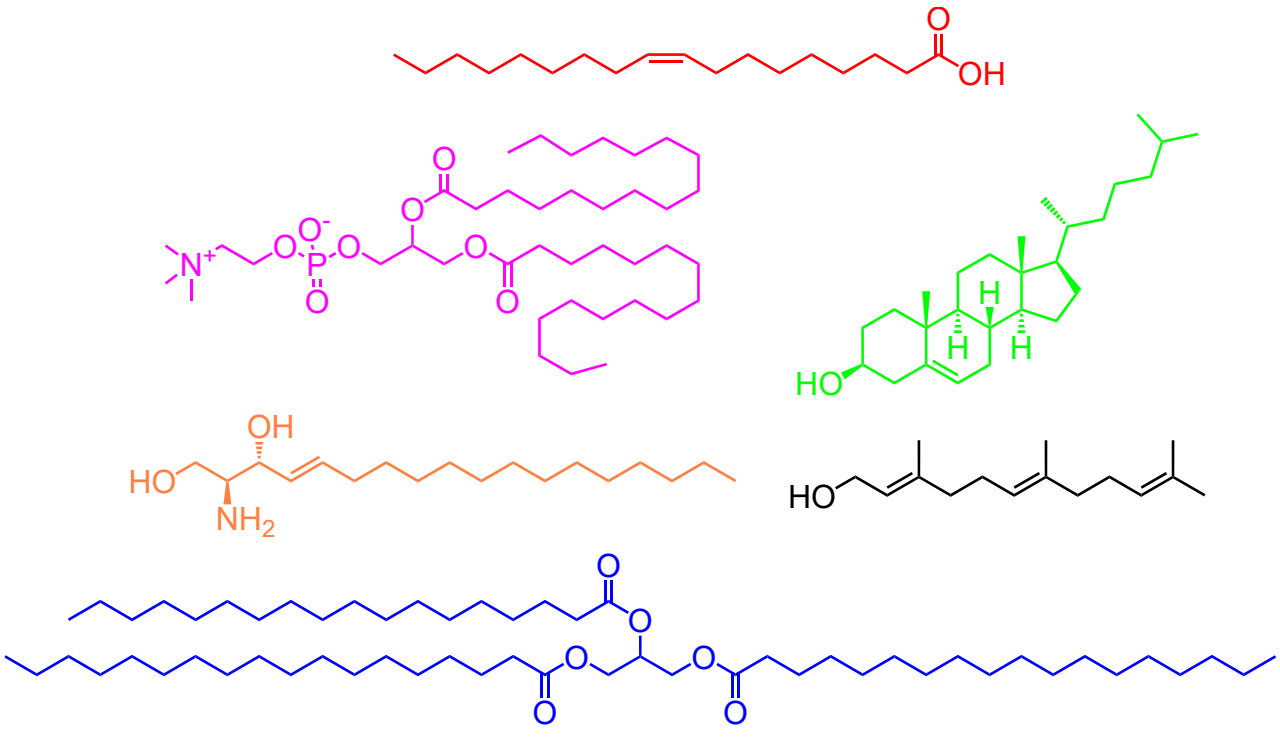


Figure 1.1: Chemical structures of lipids

Fatty acids are chemically composed of a single methyl group and a single carboxyl group at either end of a long hydrocarbon chain. They have a systematic naming convention which aids in identifying their properties. The α carbon is one attached directly to the carboxyl group, with the β directly attached to the α carbon. The letter ω signifies the number of carbon atoms away from the double bond, itself being represented by the n (figure 1.2). Finally the letter Δ relates to the double bond position along the fatty acyl chain. Saturated fatty acids, having no double bonds and often long, are straight hydrocarbon chains with carbon atoms ranging from n = 12-22 in length[177]. Whilst saturated fatty acids are straight chained, unsaturated ones are often branched and adopt geometrical isomerism, being either cis or trans in configuration. Often these unsaturated fatty acids contain a single double bond (mono-unsaturated) or multiple double bonds (poly-unsaturated (PUFA)). These geometrical configurations result in the trans form being similar to a saturated chain (straight), whilst a slight kink in the cis form causes it to be more branched. The most occurrent polyunsaturated fatty acids are ω-6 and ω-3, thus the double bonds being 6 and 3 carbons away, respectively, from the methyl end. As humans are unable to synthesise these types of PUFAs, we mammals must acquire them from alternative food sources. Among these PUFAs, linoleic and α-linolenic acids are the major unsaturated fatty acids (see figure 1.2). Metabolism, by desaturasing enzymes, then causes oxidation of the double bond to ensure nutrients are more easily up-taken by the body[173].

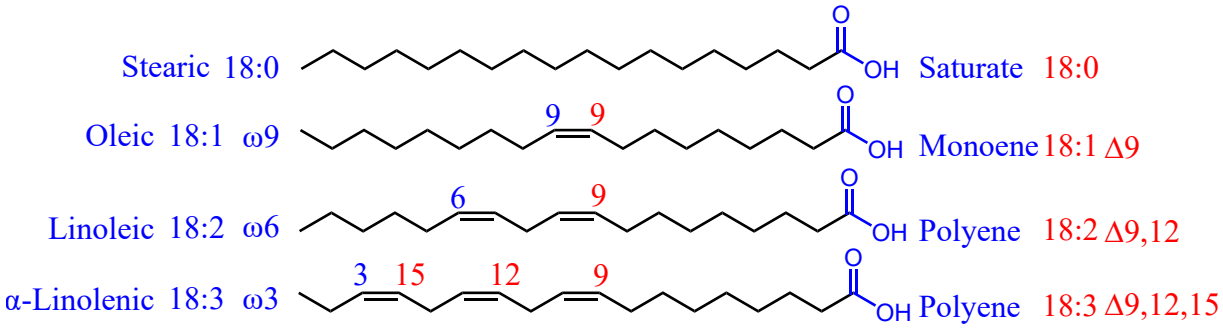


Figure 1.2: Structures of saturated and unsaturated fatty acids

The three major lipids found within all mammalian cells which regulate membrane structure and function are phospholipids, glycolipids and sterols. Diacylglycerol (DAG) molecules act as proto-lipids which, during cellular processes, form phospholipids. These molecules possess a glycerol-like backbone with two of the carbon

atoms being ester-bonded to fatty acids at positions 1 and 2 (*sn*-1 and *sn*-2)[96] (see figure 1.3). This proto-lipid undergoes metabolic transformations which contributes to cell regulation and natural source of fatty acids in fatty acid metabolism. Upon cellular stress, diacylglycerols are hydrolysed into phosphatidic acid (PA). PA molecules regulate the formation of phosphatidylinositol (PI) lipids; themselves being vital in lipid and cellular signalling. In a cyclical nature, generated PA then undergoes de-phosphorylation to reform DAG which reforms phospholipids. Thus there is a fine balance between phospholipid formation and cellular signalling with this lipid[169].

Figure 1.3: Diacylglycerol structure with ester-bond at *sn*-1 (R_1) and *sn*-2 (R_2) positions

Phospholipids are chemically formed by esterification of the glycerol backbone found in DAG lipids with different phosphate headgroups. These lipids exhibit a wide range of different headgroups, including phosphatidic acid (PA), phosphatidyl choline (PC), phosphatidylserine (PS), phosphatidylglycerol (PG), phosphatidylinositol (PI) and cardioplipin (CI)[138] (see figure 1.4). The simplest of these is phosphatidic acid (PA) which acts as a template for the other lipids. The choline headgroup makes PC, the most chemically abundant headgroup lipid in mammalian cell membranes, serine contributes to PS, ethanolamine to PE, glycerol to PG which can then be further modified to CI and inositol to PI. There exist two final minor lipids which are important in some metabolic functions. CDP-diacylglycerol (CDP-DAG) and bis(monoacylglycerol) phosphate (BMP). BMP is stereochemically distinct to all other phospholipids as its phosphate group is bonded at the *sn*-1 position rather than the conventional *sn*-3 position. It is found to aggregate in lysosomes and this chemical feature protects it from lysosomal degradation. CDP-DAG is a vital precursor to PC and CI lipids, aggregating in the mitochondria, and an intermediate to PI lipids, accumulating in the endoplasmic reticulum[39].

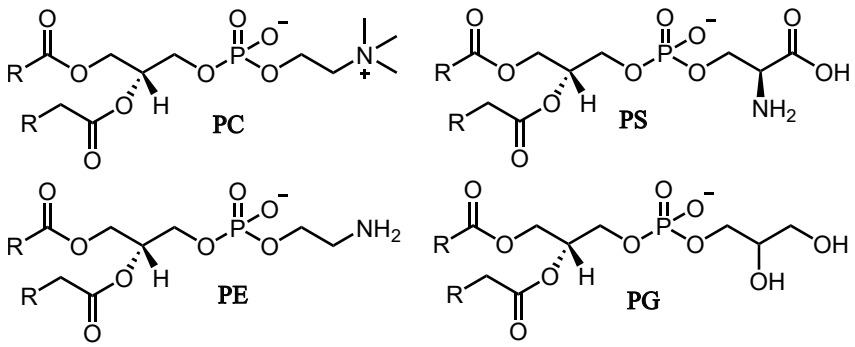


Figure 1.4: Chemical structures of phospholipid headgroups

Sterols are paramount for structural support in mammalian cells through rigidity of the cell membrane (cholesterol) and plant cell membranes (stigmasterol). Its chemical composition is a saturated fused four-ring steroid structure with a terminal hydroxyl group with a confined hydrocarbon tail[225]. In mammals, cholesterol is the most abundant sterol. Its primary role is to rigidify the membrane and withstand mechanical stress; both contributing to continuous diffusion of small molecules across the membrane bilayer. Interestingly, thermodynamics affects this rigidifying effect: at very low temperatures this effect is negated whilst at physiological temperatures it is maintained. Another important role of cholesterol is in the formation of lipid-like 'rafts' which support enzymes in their extensive role as signalling molecules in cellular processes[169].

1.2.2 Cell Membrane

Lipid membranes are a vital part of the cell. They form the semi-permeable barrier which control the passage of substances through the cell, thus ensuring cellular integrity. They also play an important role in compartmentalisation of cells to enable specific chemistry to take place leading to cell function and stability[182].

The most abundant membrane lipids are phospholipids, with phosphatidylcholine being the most common type of phospholipid found ubiquitously in mammalian cells. The chemical nature of phospholipids, namely the hydrophobic, long fatty acyl chains and the hydrophilic, small charged headgroups spontaneously generate lipid bilayers in the presence of water[23]. The active process of phospholipid movement, catalysed by flippases/floppases, results in an asymmetric distribution of lipids

across two leaflets within the bilayer (see figure 1.5). The aforementioned choline-like lipids distribute to the external leaflet whilst the amine-like lipids distribute to the cytoplasmic leaflet[44]. Another comparison of lipid composition in each of the leaflets is the relative abundance of high melting vs low melting temperature lipids. Gangliosides, cholesterol and sphingomyelin (high melting temperature) lipids are present in the external leaflet and embedded in a liquid-like matrix with PC (low melting temperature) lipids[87]. Homeostatic control of the lipid asymmetry across the bilayer is maintained by enzymes. Firstly, floppases catalyse the migration of phospholipids from the cytoplasmic to external leaflet and vice-versa. Secondly, flippases catalyse the reversible process across the leaflets and scramblases catalyse protein transport across the bilayer[182]. This is known as interleaflet diffusion, trans-bilayer diffusion, or more colloquially 'flip-flop'. This process is key to maintaining lipid composition across both bilayers and thus structural integrity. The timescale for flip-flop various for different lipids. Bulky, polar lipids such as PC only traverse the bilayer in a longer timescale (hours) whilst small, non-polar lipids such as cholesterol migrate across leaflets in a sub-second timescale[47].

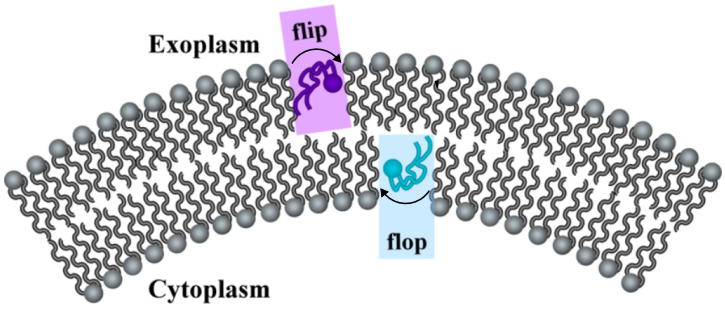


Figure 1.5: Flip-flop mechanism arrangement in lipid membrane[47]

Owing to the diversity of lipid types, each cell membrane may adopt different lipid compositions with varying headgroups, extent of saturation, and carbon chain lengths. The average diameter of the 1-palmitoyl-2-oleoyl-*sn*-glycero-3-phosphocholine (POPC) phospholipid bilayer has been measured as 4.2 nm, with its lipids spanning the diameter of the membrane, further integrated amidst transmembrane proteins and peptides[190]. The biological importance of these proteins extends to not only cell structural integrity, but also to cell signalling. Moreover, the lipid composition of such membranes is dynamic and susceptible to alteration by external factors, such as temperature or pH. Enzymatic activity is responsible for maintaining structural

support of the lipid bilayer, both internally and externally (cytosol). An equilibrium exists between the formation and degradation of lipids which also contributes to this cellular stability. Ubiquitously, within all eukaryotic cells, at least half of the lipid composition of the bilayer is composed almost exclusively of the glycerophospholipid PC[212].

These lipids share a diacylglycerol group, being found at the centre of the membrane, within the hydrophobic region. Each fatty acyl chain within the diacylglycerol group contains cis/unsaturated chains of differing lengths. The PC lipid can undergo supramolecular self-assembly, contributing to the formation of the flat bilayer structure (see figure 1.6). This chemical assembly enables the PC to adopt a cylindrical geometry with hydrophobicity being preserved within the core of the cell membrane, and hydrophilicity existing within the aqueous extracellular environment. Due to the presence of a single cis/unsaturated acyl chain within certain lipids, such as PC and PE, they exist as lyotropic liquid crystals at room temperature. These membrane lipids each have distinct physical properties[101]. PE lipids possess smaller hydrophilic heads, and thus have preference for a conical geometry during molecular self-assembly. Research has found that PE lipids also play a critical role in the organisation of membrane proteins which traverse the bilayer. Sterols, discussed earlier, also constitute part of the lipid bilayer. Within mammalian cells more than 90% of these lipids are in the form of cholesterol, whereas within fungal cells, more than 90% of these lipids are in the form of ergosterol[94].

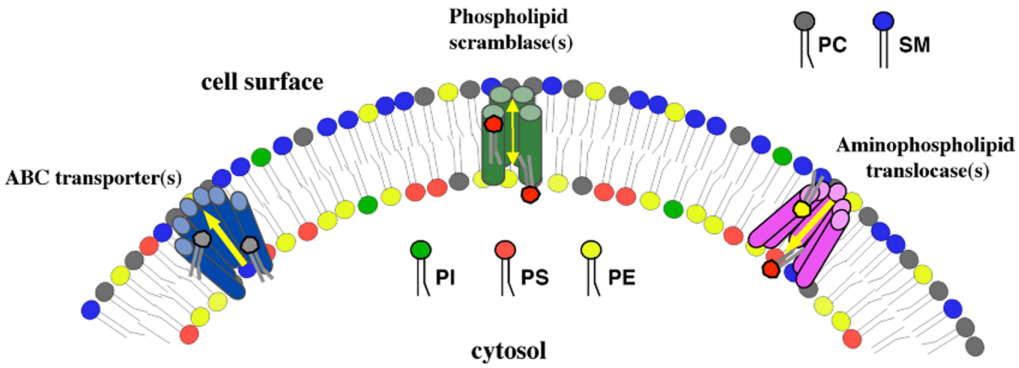


Figure 1.6: Asymmetric lipid distribution and composition of cellular membranes, PC, SM, PS, PE and PI being the most abundant membrane lipid headgroups[101]

1.2.3 POPC Membrane

POPC was chosen as a lipid bilayer within this project to model the interaction and reactivity of drugs with cell membranes (see figure 1.7). POPC heavily contributes to the plasma membrane of most vertebrate cells, and thus can be useful as a comparison with other lipid bilayers, as to the effectiveness and reactivity with varying drugs. POPC is chemically distinguishable from other bilayer lipids, as one of its hydrophobic tails (C_{18} chain) is shorter than its 2^{nd} tail (C_{16} tail). Cholesterol is a sterol which can have profound effects on the bilayer properties of cells and is thus of interest in terms of its impact to drug reactivity with POPC. From a physical perspective, its influence on POPC is to increase the bilayer mechanical strength, reduce the permeability of small organic/drug molecules across the bilayer and finally control bilayer fluidity[175].

Research has shown that whilst in a physiologically controlled state, cholesterol favours lipid ordering, but once the phase changes to a gel-like state, such ordering is reduced. A suppressed effect of cholesterol on POPC may occur by any slight alteration to the structural components of cholesterol, namely its steroid ring, its hydroxyl residue, and its hydrocarbon chain, attached at the opposite end of the ring structure to the hydroxyl group. Moreover, cholesterol is believed to act in the production of lipid rafts, in microdomains composed of high melting temperature lipids[133].

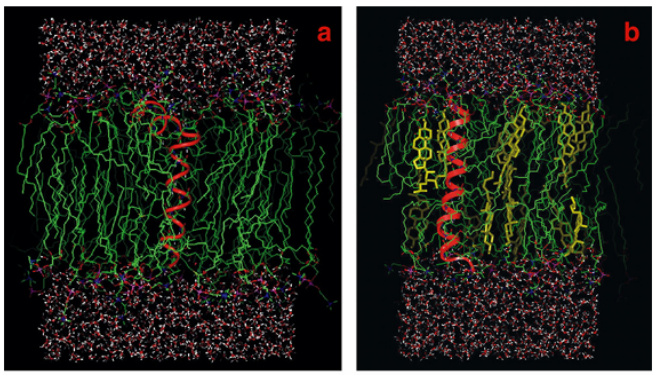


Figure 1.7: Molecular dynamics snapshot of a) POPC, b) POPC with cholesterol, Colours: carbon-green (lipid chains), oxygen-red and hydrogen-white (water), nitrogen-blue and phosphorus-purple (lipid headgroups), cholesterol-yellow (ribbons)[175]

The presence of cholesterol, embedded within the POPC bilayer with no additional lipid components, results in a phase formation similar to an ordered lipid domain (L_o) at 310 K (physiological temperature). Whilst historic evidence has shown that only lipid mixtures present in the outer leaflet of the lipid bilayer have been able to form segregated, ordered lipid domains, recent observations have shown that lipid rafts, with an abundance of ordered lipid domains, possess a transmembrane-like effect, whereby ordering in one leaflet induces ordering an another[222]. Thus, the inner leaflet of the cell membrane may experience alterations/fluctuations in the frequency of ordered lipid domains. However, there remains much speculation as to the favourable thermodynamics which may influence the overall organisation of lipids (L_o) within the inner leaflet of the cell bilayer.

POPC though, has been proven to undergo phase changes to an ordered (L_o) phase, when influenced by a high level of cholesterol molecules (50–60%). Furthermore, cholesterol can distinguish between membrane phospholipids, by their difference in acyl chain structure and length. This in turn would promote segregation of L_o and L_d domains in the leaflets of the cell membrane. Because L_o and L_d have different bilayer widths, their coexistence is more likely to lead to defects, at the interface, which may be able to promote membrane reactivity[106] (see figure 1.8). Thus the effect of cholesterol would be an interesting subject to study.

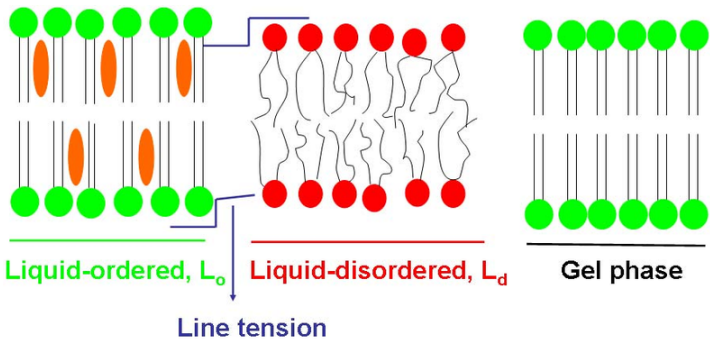


Figure 1.8: Schematic of ordered (L_o), with embedded proteins (orange ovals), and disordered (L_d) lipid domains, L_o acting as an intermediate between the gel and L_d phases[106]

1.2.4 Membrane Models

One of the earliest proposed structures of the Fluid-Mosaic (FM) model was in 1972 by Singer and Nicolson. This original paper suggested that the most abundant

lipids found within the bilayer are glycerophospholipids, alongside minor lipids such as sterols or sphingolipids. Phospholipids, which form the ordered bilayer structure, ensure balance between hydrophobic and hydrophilic interactions[212]. Certain phospholipids with a net neutral charge, such as isoelectric (zwitterionic) phosphatidylcholine, can form covalent dipole-dipole interactions between ion pairs, which leads to stabilisation of the bilayer structure. This can be seen in figure 1.9 at the terminal ends of the embedded proteins. Although other non-covalent interactions such as London dispersion forces, electrostatic interactions, or hydrogen bonding, impact the lipid bilayer formation, they remain minor in comparison to the hydrophobic interactions which dominate structure formation[196]. This hydrophobic effect is favoured when the fatty acids orient on the inside, whilst the polar groups remain on outside. This results in polar groups that are well solvated and hydrocarbon parts which avoid water.

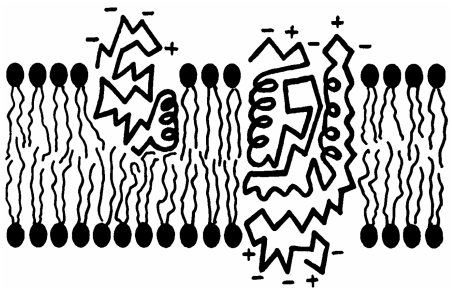


Figure 1.9: Interactions of proteins and lipids in the Fluid-Mosaic model, stabilisation of lipid components by non-covalent interactions[196]

Furthermore, the phospholipids, along with certain transmembrane proteins (TM), are posited to be amphipathic (containing both hydrophobic and hydrophilic elements). Thus, the constitutive lipids are physically asymmetric, each composed of one highly polar group and another non-charged group. This classical amphipathic model (see figure 1.10), which is thermodynamically favoured in its structure formation, is determined largely by the amino acid structures, in turn influenced by sequence, within the phospholipid bilayer. Cumulatively, this results in a low Gibbs free energy, and thus an optimal energetic state. Structural evidence, such as X-ray diffraction, points to the cell membrane having a mosaic structure, whereby the integral proteins are embedded within a 'matrix' of phospholipids[146]. This model is an example of the three-dimensional, long-range structure of the bilayer, with continual repeating units along the z axis in three-dimensional space.

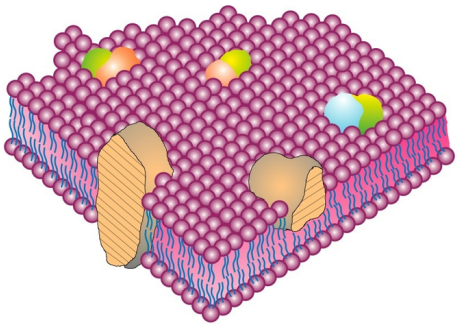


Figure 1.10: Amphipathic model of cell membrane, lipid head and tails (purple) with integral and transmembrane proteins (orange)[146]

Whilst the fluid-mosaic model has provided success in describing the structure and composition of biological membranes, there are certain biological functions which it fails to capture. The lipid whisker model has been proposed as an extension to the fluid-mosaic model which takes into account the role of oxygen in lipid peroxidation reactions within the membrane bilayer. The discovered presence of oxidised phospholipids within the bilayer led to the development of the lipid whisker model (see figure 1.11). They found that oxidised lipids have polar groups inserted at the depth of the alkenes. Some of these also fragment to form shorter chains capped by aldehydes. New research has shown that the classic organisation of lipids in the fluid-mosaic model differs from the changed conformation present in oxidised phospholipids. This chemical oxidation causes the lipids to adopt a 'whisker-like' structure and significantly change conformational shape. Experimental evidence has shown that the interaction of monoatomic oxygen with the hydrophobic chain forces the lipid tails to adopt this whisker structure rather than the bilayer structure[75].

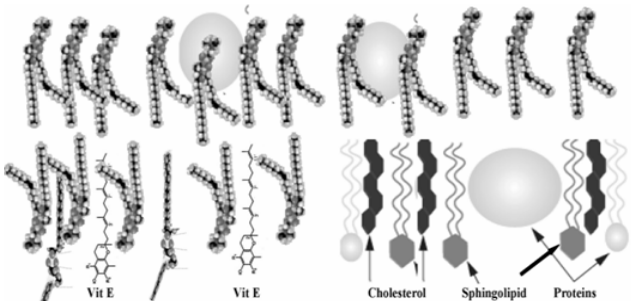


Figure 1.11: Lipid whisker model, with additional vitamin E, cholesterol and protein between the 'whiskers'[75]

Another historically important model of the cell membrane is the Paucimolecular model (see figure 1.12), first proposed in the 1930s. The radical change to the clas-

sic FM model was the determination of the relative positions of proteins within the membrane. In 1935 Danielli and Davson stated that the lipids in fact are constrained between two monolayers of proteins; above and below. Whilst holding to this model, they also suggested the presence of TM proteins embedded throughout the membrane would account for experimental permeability measurements across the bilayer. This picture of the cell membrane permeated scientific thinking for nearly a half-decade. Initially measurements on surface-tension led to the development of this model. The inclusion of amphipathic molecules, such as sterols or glycerophospholipids, to non-miscible fluids caused a significant reduction in surface tension between these hydrophobic molecules and very polar substances. Thus the necessity of these hydrophobic proteins to establish a lower surface tension gave credence to the Paucimolecular model[127].

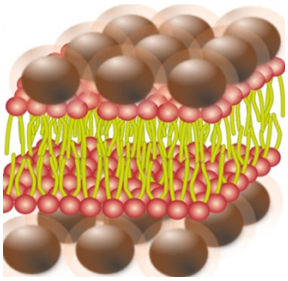


Figure 1.12: Paucimolecular model, proteins (brown) above and below the lipid bilayer (red/yellow)[127]

A major shortcoming of the fluid-model was the lack of structural detail on the effect of transmembrane and integral membrane proteins on the lipid arrangement. As a result, Engelman in 2005 revisited this concept of protein crowding (see figure 1.13). Although Singer and Nicolson believed the impact of transmembrane proteins on lipids was minimal due to the large molecular excess of lipids over proteins, new biophysical evidence has shown that there exists a much higher abundance of proteins which can significantly affect the lipid structure. Three new features have come to light which fundamentally change our perspective on the lipids within the membrane, namely the higher amount of protein contacts through non-covalent interactions, irregular bilayer thickness and the abundance of external membrane proteins. This novel understanding has led scientists to favour the mosaic over the fluid nature of the membrane[73].

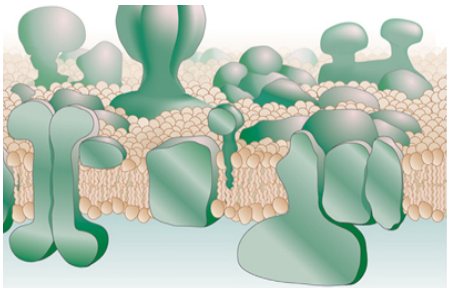


Figure 1.13: Engelman model of membrane protein crowding, transmembrane and integral proteins (green) traversing through the membrane bilayer[73]

1.3 Chemical Stability of the Membrane

1.3.1 Reactivity at the Membrane Interface

Reactivity of small molecules with membrane lipids remains an ongoing and multi-disciplinary project. The factors which contribute to solubility, bioavailability and absorption of molecules across the bilayer continue to be developed in understanding. Many commercially available drugs exhibit poor solubility which can often lead to poor availability and uptake of these molecules, thus understanding of factors which lead to reduced solubility is important in drug development research[35]. There are various mechanisms by which molecules interact and cross the membrane barrier. Firstly, the simplest and most common method by which molecules migrate across the bilayer is by passive transport, often being referred to as simple diffusion. Whilst this process is true for hydrophobic molecules, hydrophilic molecules require more time than hydrophobic ones to traverse the bilayer. Electrostatic interactions between water and these hydrophilic molecules makes them more difficult to enter the membrane interior. Thus, ion channels are required to move these substances across the membrane (facilitated transport). This process of simple diffusion can be quantified as its rate being proportional to both the maintained membrane concentration gradient $(X_1 - X_2)$, as the molecule diffuses across the membrane, and the partition coefficient (P) of the molecule across the lipid-water interface, alongside the drug diffusion coefficient (D) from the bilayer centre to the aqueous region, all divided through by the membrane bilayer thickness (d)[190]:

$$\text{Rate} = \frac{D \cdot P}{d}(X_1 - X_2) \tag{1.1}$$

Secondly, facilitated transport is required to carry molecules, such as amino acids and sugars, across the membrane. This process relies entirely upon TM proteins which traverse the span of the bilayer. One type, namely carrier proteins undergo conformational change in order to bind appropriately to the molecule and bring it into the cell, whilst the other type, namely channel proteins, control the passage of substance by opening and closing of their 'pores'. The opening/closing mechanism is activated by external stimuli such as ligand binding and thus selectively control the passage of only specific molecules across the bilayer. This process is unidirectional along the concentration gradient and requires ATP for its cellular function[158].

Membrane partitioning of a drug molecule is also a vital process which affects lipid stability. When considering partition coefficients historically, the oil-water interface was thought to be a good model which could be applied to drug-membrane partitioning. As the oily (hydrophobic) layer was in direct contact with the water (hydrophilic) layer this represented both an aqueous and organic phase and the partition coefficient of a molecule was determined as the ratio of concentration of this molecule in both these two immiscible liquids[170]. A unifying principle known as the Overton rule was established in 1899 for quantifying the membrane permeability. The main caveat of the rule is that it excludes drug migration by facilitated diffusion, requiring carrier or channel proteins, thus is limited to simple diffusion of molecules. From a mathematical modification of Fick's law of simple diffusion, the so called transmembrane flux density (J) can be calculated if the partition coefficient (K_p) of the molecule diffusing from the lipid to the water phase is known[139]:

$$P_M = \frac{K_p D_M}{d} \tag{1.2}$$

$$J = P_M(c_{2w} - c_{1w}) \tag{1.3}$$

$$J = -D_M \frac{dc_m}{dx} = -D_M \frac{c_{1m} - c_{2m}}{d} \tag{1.4}$$

Where P_M defines the membrane permeability derived experimentally, D_M denotes the diffusivity, c_{1m} and c_{2m} respectively represent the drug permeability concentration between the two phases, c_{1w} and c_{2w} respectively are the aqueous concentrations

for both phases, x represents the distance from the first interface and once more d represents the membrane bilayer thickness[84].

1.3.2 Methods for Quantifying Drug Association with Membranes

Owing to the diverse complexity of membrane bilayers, various biomimetic models have been proposed which can be more easily probed through biophysical techniques. A well established approach uses liposomes. These particles are spherical and act as model lipid vesicles with an internal aqueous compartment (see figure 1.14). For biophysical studies, there are three different sizes of large unilamellar vesicles (LUV). Firstly, small unilamellar vesicles (SUV) ranging from 20 - 50 nm in diameter. Secondly, large unilamellar vesicles (LUV) ranging from 50 - 100 nm in diameter. Lastly, giant unilamellar vesicles (GUV) ranging from 10 - 100 μM in diameter[158].

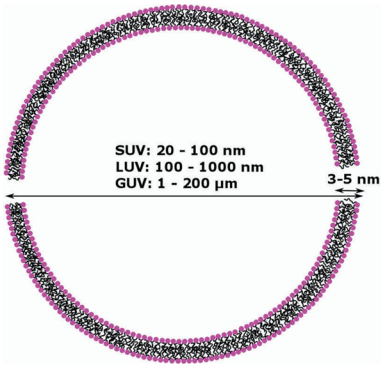


Figure 1.14: Comparison of the size of GUV, LUV and SUV liposomes, alongisde bilayer thickness[158]

The next series of biophysical techniques quantifies the affinity of drug molecules with membranes. This is important to establish the most useful technique for my own project. A biophysical technique to study model membranes is through derivative spectrophotometry. This spectroscopic technique can be used to calculate the partition coefficient for the liposome in direct contact with the aqueous region (see figure 1.15). The main experimental requirement for this calculation is the segregation of the liposome vesicles and drug (solution) into their respective lipid and aqueous regions by either dialysis or centrifugation. This is due to the excessive signals which are produced by liposomes when they scatter light in the spectrophotome-

ter. This technique produces a line shift corresponding to any potential interactions between the drug and liposome, which then can calculate the partition coefficient. The apparatus is sufficiently accurate to be able to cancel the excess noise from light scattering by the vesicle and accurately quantify the spectral change[103].

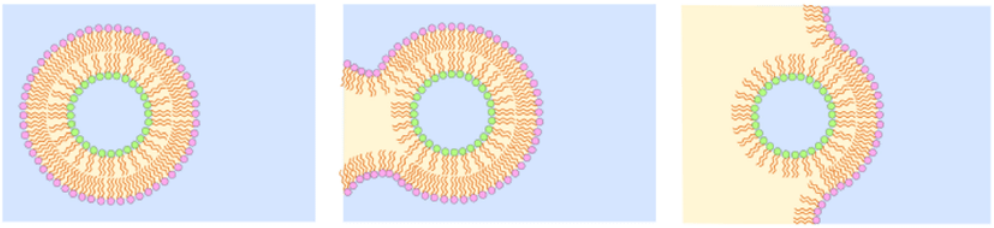


Figure 1.15: Liposome rearrangement in water[103]

Another biophysical technique is the use of fluorescence spectroscopy to calculate the partition coefficient of the relevant molecule across the aqueous phase. This calculation is achieved through the use of a hydrophilic fluorescent indicator which can give a direct measurement of the log P value. As with all techniques there are limitations to the practical use of their method. In this case, fluorescence measurements are only possible if the quantum yield (fluorescence parameter) is sufficiently different for the molecule in the aqueous phase relative to being partitioned in the phospholipid bilayer. Also, there must be a measurement possible from the fluorescence indicator upon partitioning of the molecule in the membrane interface. Mathematically, the partition coefficient is derived from either the steady-state anisotropy (r) or fluorescence emission intensity (I), both requiring simple algebra to obtain K_p (partition coefficient):

$$I = \frac{I_w + K_p \gamma_L [L] I_L}{1 + K_p \gamma_L [L]} \tag{1.5}$$

$$r = \frac{r_w((\gamma_L[L])^{-1} - 1) + \dfrac{r_L K_p \epsilon_L \phi_L}{\epsilon_w \phi_w}}{(\gamma_L[L])^{-1} - 1)\dfrac{K_p \epsilon_L \phi_L}{\epsilon_w \phi_w}} \tag{1.6}$$

where ϕ is the fluorescent quantum yield and L is the lipid concentration. I_L and I_w relate to the fluorescence emission intensity of the lipid and aqueous regions respectively, ϵ is the molar absorption coefficient and γ refers to the molar volume[185].

Another biophysics technique is called Isothermal Titration Calorimetry (ITC). Whilst previous methods have analysed lipophilicity, ITC can directly predict thermodynamic properties of drug-membrane binding at the level of individual bond formation and breaking between lipids and small organic molecules. Therefore this technique which analyses energetics can account for drug-lipid binding at equilibrium conditions. The key parameters for this calorimetry are the equilibrium constant (K_{eq}) of binding and the molar stoichiometry in the reaction equation:

$$K_{eq} = \left(\frac{\text{complex}}{\text{receptor}} \cdot \text{ligand}\right)_{eq} \tag{1.7}$$

$$\Delta G = -RT\ln K_{eq} \tag{1.8}$$

$$\Delta G = \Delta G + RT\ln\left(\frac{\text{complex}}{\text{receptor}} \cdot \text{ligand}\right)_{\text{actual}} \tag{1.9}$$

$$\Delta G = \Delta H - T\Delta S \tag{1.10}$$

where ΔG is the standard Gibbs Free Energy, R the ideal gas constant, T the temperature, ΔH the enthalpy and ΔS the entropy. This non-destructive method is simple, quick and utilises a series of titrations. The enthalpic contribution reflects the electrostatics, hydrogen bonding and van der Waals interactions between the drug and lipid molecules. The entropic contribution is a measure of the degree of disorder in the system, thus providing a picture of the average distribution of the molecular system. Graphically, the saturation point of the curve reflects constant rate of binding whilst steeper gradients on the curve point to higher affinity (see figure 1.16). This method utilises multiple titrations to build up a quantitative description of the binding[184].

Nuclear Magnetic Resonance (NMR) Spectroscopy is a biophysical technique which relies upon the intrinsic spin and magnetism of the atomic nucleus. The individual nuclei possess a magnetic moment which causes a change in atomic energy levels and frequencies when in close proximity to an externally applied magnetic field. Owing to a diverse array of atomic environments within a lipid, such as ^{1}H, ^{14}N, ^{31}P, NMR spectroscopy can be utilised to measure the binding of drug molecules to these lipids. These measurements result in a chemical shift change for the drug in

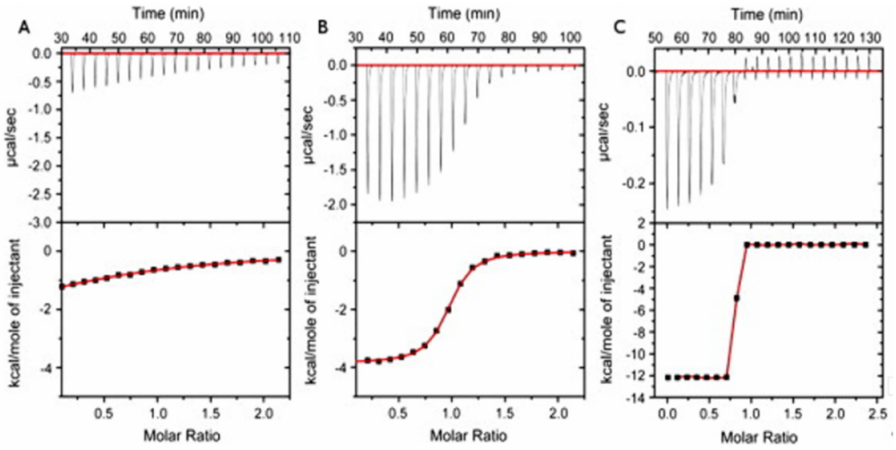


Figure 1.16: Diagram of ITC curves showing binding of drug to membrane over time. Increase in enthalpy of binding from (A)-(C)[184]

solution compared to the drug bound to liposomes (see figure 1.17). The ability of NMR as a non-destructive, non-invasive and quantitative method makes it a prime biophysical technique when considering the structure and dynamics of drug-lipid intermolecular interactions[121], and thus of great importance in this project.

The crucial quantities considered in biomembrane NMR dynamics include the linewidth ($\Delta v1/2$), T_1 relaxation time (longitudinal), T_2 relaxation time (transverse), and the self-diffusion coefficient (D). $\Delta v1/2$ is dependent on the molecular motions and refers to the extent of narrowness/broadness of different chemical species, whilst also being dependent upon the correlation time. For lipid membranes, these signals are much broader than for small organic molecules due to much higher levels of heterogeneity. T_1 and T_2 provide information on segmental motions and the combination of $\Delta v1/2$, T_1 and T_2 can be used to calculate correlation times which directly reflect the dynamics. The final parameter D represents the translation motion of atomic nuclei in the system[149].

1.3.3 Simulation Techniques to study Membranes

Theoretical chemistry can be considered as looking down a computational microscope, where atomistic level detail and high resolution can be obtained on molecular systems. The main advantage of simulations is to shed light on chemical reactivity inaccessible through experiments. The choice of length and timescales is important when considering the size of the molecular system. If one wishes to consider the

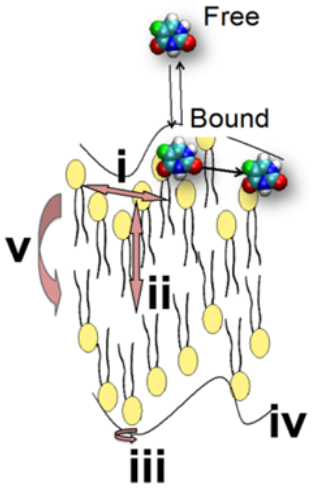


Figure 1.17: Lipid motion in membrane bilayers, (i)-lateral diffusion in parallel direction to membrane surface, (ii)-protrusion along the bilayer normal, (iii)-rotational diffusion, (iv)-undulation, (v)-flipflop, drug transport between free and bound states reflect dynamic motion of lipid molecules[121]

electronic structure of a molecule then quantum mechanics (QM) is the only viable method, whereas if one wishes to measure partitioning coefficients of molecules at the lipid-water interface then a coarse-grained approach is far more suitable[215].

Performing a molecular simulation to study drug-lipid interactions requires a systematic methodology. Whilst quantum mechanics has historically proved successful in determining the electronic properties of molecules by solving the Schrödinger equation, in practice this is numerically impossible for any many-body system. The so called 'Wave Function' is the quantum mechanical equivalent to the force calculated in Newton's Second Law of Motion and entirely describes the dynamics of the system as a function of time.

To mitigate this difficulty, an approximation known as the Born-Oppenheimer (BO) approximation was developed. This assumed the nuclei were fixed and calculated the electronic wave-function, which was then fed into the overall calculation of the nuclear wave-function. As the time to simulate a system increases dramatically with size, the BO approximation provides a computationally feasible method for studying large biomolecular systems.[4].

A far simpler and much more computationally feasible approach has been developed

based upon classical mechanics. This methodology, known as molecular dynamics, solves for the potential energy as a function of atomic positions in the drug-lipid system. The force of each atom exerted on another atom is calculated from the 1st order derivative of the potential energy as a function of distance between two atoms. This potential energy captures all covalent (bonding) and non-covalent (non-bonding) interactions in a so called 'Force Field'[48]. From a software perspective, biomolecular force fields contain all required parameters for nucleic acids, carbohydrates, lipids and proteins. Thus, they have a huge capability of accurately simulating biomolecular interactions of small molecules. The force fields for the purpose of biological interactions are AMBER, CHARMM, GROMOS and OPLS. For my project, I have exclusively used AMBER, including with my QM/MM calculations. The AMBER force field contains a derivative GAFF (general AMBER force field) for the parameterisation of small organics molecules, LIPID17 for lipids and TIP3P for water molecules[128].

There is a wide application of molecular dynamics to study drug-lipid interactions. Firstly, to measure drug localisation within a membrane. Simulations show that amphipathic (both hydrophilic and hydrophobic ends) molecules prefer to localise at the interface between the lipid headgroups and the aqueous region whilst fully hydrophobic drugs prefer to orient themselves in the membrane centre surrounded by lipid tails. Furthermore their partitioning takes place in the regions they migrate towards (see figure 1.18). Another trend seen by simulations is the preferred migration of charged molecules to this interface vs their non-ionised counterparts[110]. Also, the use of biased simulations can be used to calculate the Gibbs Free Energy of the migration of a molecule across the lipid bilayer. This gives quantitative predictions into the stability and preferred partitioning of the molecule in the membrane by measuring the difference in free energy between the aqueous region, interfacial region and bilayer centre.

Secondly, simulation techniques can be used to measure the drug binding kinetics in the membrane. The binding of a drug to a receptor target in the membrane is quantified as the association (binding) or dissociation (unbinding) constants. Whilst

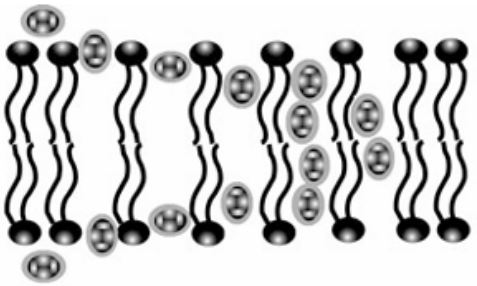


Figure 1.18: Diagram of the effect of membrane lateral pressure on the localisation of drug molecules (oval grey beads) in the cell membrane[110]

thermodynamics is sufficient to quantify the absolute binding enthalpy, the binding is also dependent upon intermediate states which rely upon more subtle kinetic effects in the membrane. It has recently been shown that molecular dynamics can accurately calculate these so called association and dissociation constants, represented as k_{on} and k_{off}. The k_{on} rate constant is an important pharmacological tool in the drug discovery process, as drug molecules which have a higher k_{on}, thus a higher rate of binding, are more likely to be suitable drug candidates[153]. These simulation tools safeguard against potentially poor pharmacological properties exhibited by supposedly high affinity receptor binding targets, but with poor equilibrium conditions in the membrane. For a typical drug-receptor binding event, the rate equation could be shown as:

$$\frac{d[DR]}{dt} = \kappa_{\text{on}}[D][R] - \kappa_{\text{off}}[DR] \tag{1.11}$$

Kinetically, the forward rate of reaction is of 2^{nd} order, whereas the reverse rate of reaction is of 1^{st} order. This binding can then be algebraically rearranged to provide a final differential equation which describes the rate of drug-receptor binding at equilibrium in the membrane. $[D]$ refers to the drug concentration and $[R]$ relates to the receptor concentration. This 1^{st} order differential equation computes the rate of drug-receptor ($[DR]$) binding.

From a computational perspective, the kinetics of binding/unbinding could be quantified as a 1D potential of mean force (PMF) calculation with a double well present

in the free energy landscape profile (see figure 1.19). The free energy barrier height, between the two minima, is calculated as high enough to conclude that the transition from one basin to the other basin of the double well is a statistically rare or unlikely event[48]. Transition State Theory computes that the rate constant (κ) is proportional to the exponential of the activation free energy barrier ($\Delta G^{\ddagger}$), divided by the product of the Boltzmann constant (k_B) and temperature (T), with a pre-exponential coefficient known as the proportionality constant (κ_0):

$$\kappa = \kappa_0 e^{\frac{-\Delta G^{\ddagger}}{k_B T}} \tag{1.12}$$

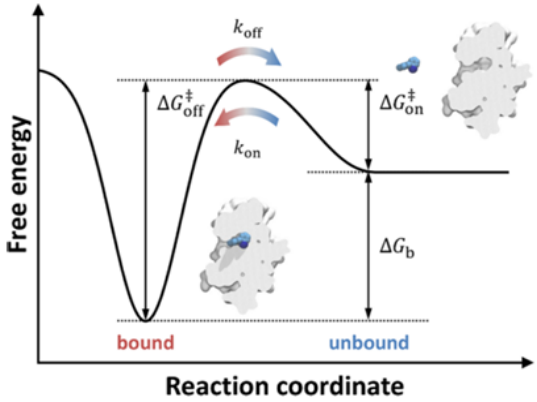


Figure 1.19: Free energy landscape of drug-receptor binding event, between bound and unbound states[48]

1.4 Introduction to Drug-Lipid Interactions

1.4.1 Cationic Amphiphilic Drugs

Cationic Amphiphilic Drugs (CAD) are a class of drug molecules which ubiquitously share a hydrophobic aromatic ring system and a hydrophilic side chain which includes amine functionality either in its termination or side chain. These amine groups have a high propensity to be ionised to a protonated form within certain highly acidic cellular compartments, such as lysosomes, which can lead to their sequestration in these organelles[214]. CAD molecules are either neutral or protonated. The neutral form has greater permeability than the protonated form, so they become trapped in the lower pH environment[147]. As this drug partitioning mechanism across the membrane is pH dependent, the cellular uptake of these CAD molecules

is widely varied in timescale and accumulation rate. Partitioning of the molecules as physiological pH, across the membrane, can lead to physiological events such as formation of lysolipids and potential drug-induced phospholipidosis[178]. The scope of CAD molecules include a huge array of clinically relevant drug molecules such as beta-blockers, antidepressants, obesity-lowering drugs, antibiotics and cholesterol-lowering drug molecules. The specific CAD molecules of interest (figure 1.20) in this project are (A) propranolol, (C) fluoxetine and (D) phentermine. A small organic molecule, (B) 2-aminomethylbenzimidazole, is also of interest to probe its aminolysis chemistry with POPC membranes.

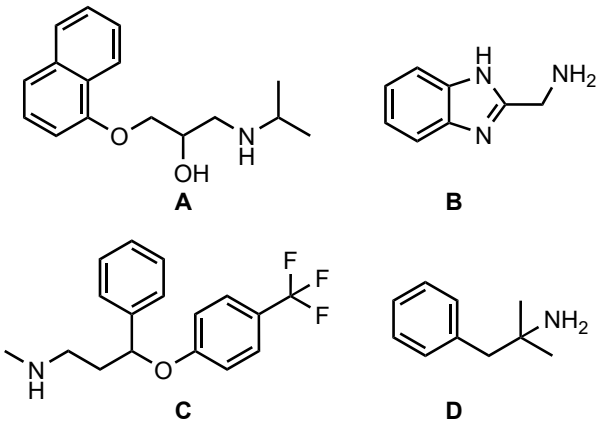


Figure 1.20: Chemical structures of CAD molecules

CADs share identical transport methods across the cell membrane, which is useful in understanding the interaction of these small organic molecules with the POPC bilayer. The most common mode of action of drug transport across the cell membrane is passive diffusion. This mechanism, which occurs spontaneously, is more favourable for hydrophobic organic molecules, as opposed to charged ones, and fortuitously requires no energy barrier for the molecule to cross. Diffusion has been found to be the main pathway for drug transport of neutral small organic molecules, across the lipid bilayer[11].

The ratio of charged to neutral drug abundance relies upon both the overall pK_{a} of the drug, as well as the pH of the aqueous environment on either side of the POPC bilayer (see figure 1.21). The cationic part of CAD is typically a primary, secondary, or tertiary amine with a high pK_{a} (9). Computational results have shed light on the change in pK_{a}, for amine-like groups (pK_{a} 11), as they traverse the

POPC membrane[190]. MD simulations have shown as large as a 4.5 change in the pK_a of strong bases. Although, this process may be energetically unfavourable and thus have a much lower probability of crossing the membrane. Further, for CADs that have a lower pK_a than most secondary/tertiary amines, diffusion alone would be sufficient for their transport across the cell membrane, as these CADs would be neutral at physiological pH. Interestingly, interaction of a neutral form of a drug with the intracellular environment would lead to a higher equilibrium concentration[35].

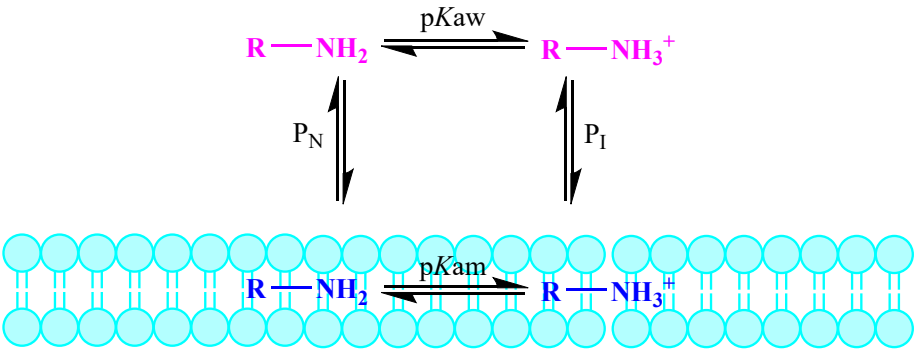


Figure 1.21: Amine/ammonium ion equilibria, P_N: amine water-membrane partition coefficient, P_I: ammonium water-membrane partition coefficient, pK_{aw}: -$\log_{10}$ of the acid-base dissociation constant in water, pK_{am}: -$\log_{10}$ of the acid-base dissociation constant in membrane interface

Drug-induced phospholipidosis (DIPL) was first observed in 1948, after the administration of the CAD molecule chloroquine (antimalarial) to a rodent. DIPL is a lysosomal storage disorder which results in the accumulation of phospholipids by the adverse effect of certain CAD molecules interacting with cellular membranes. These phospholipids aggregate and are stored in intracellular compartments. Since 1948, more than 50 CAD molecules have been found to induce phospholipidosis across many different organs, including the kidney and liver[7]. Physiologically, DIPL results in the formation of lamellar bodies. These lamellar bodies consist of concentric myelin-like structures and are often referred to as cytosolic inclusions, or intracellular substances which accumulate in various different cells[22]. The timescale after induction of these CAD molecules across the membrane that can cause these lamellar bodies to form can be as short as a few hours. *In vivo* studies, within a range of different animals, have shown that many factors such as dosage level, exposure duration and drug functionality all contribute to the effect and severity of drug-induced phospholipidosis[148]. The effect of certain CAD molecules, such as propranolol, pro-

caine and tetracaine, have all shown the capability of inducing phospholipidosis (see table 1.2). Thus study into their chemical reactivity with membrane lipids is important to understand the underlying factors which may affect this process.

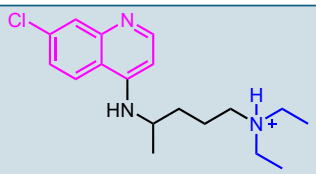
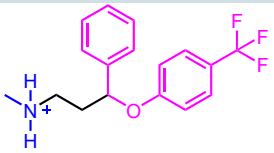
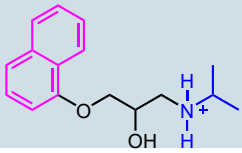
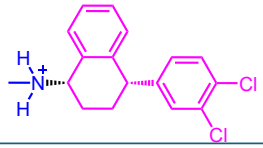
Drug	Biological Action	pK_a	Chemical Structure	clogP	DIPL in vivo	DIPL in vitro
Chloroquine	Antimalarial	10.3		4.3	**+ (Humans, Rat)**	**+**
Fluoxetine	Antidepressant	9.8		4.2	**+ (Humans, Rat)**	**+**
Propranolol	Antiarrythmic	9.7		2.6	**(+)**	**+**
Sertraline	Antidepressant	9.9		5.2	**(+)**	**+**

Table 1.2: Table of DIPL effects in vivo and in vitro for different drug candidates

1.4.2 Neutral and Protonated CAD molecules at the Membrane Interface

Bilayers composed of POPC have been chosen as a model in this study for the interactions of drugs with lipid membranes. These molecular interactions can help to shed light on pharmacokinetic properties such as absorption, distribution, metabolism, and elimination (ADME). These properties are also dependent upon molecular interactions, such as the binding selectivity of drugs to cell membranes. Therefore, both the binding affinity and selectivity of drug molecules to cell receptors can be influenced by the drug partitioning characteristics in membranes[206]. Thus, from longer term applications, molecular simulations help provide a more detailed understanding of favourable drug binding sites to membranes (pre-clinical research).

Alongside the clinical benefits of propranolol as a beta-blocker, this drug molecule also displays non-specific membrane effects (NME), which could influence its pharmacological properties, specifically the favoured binding interaction of a specific

membrane receptor with the drug, linked not to lipophilicity but to partitioning behaviour[111]. Both neutral (amine) and protonated (ammonium) forms of propranolol can be constructed, atomistically, to compare membrane permeability of these forms. Historically, the pH-partition hypothesis has proposed that solely the neutral form can permeate across the lipid membrane[34] (see figure 1.22). Whilst this classical model fits for many neutral molecules, recent research has shown that it also fits for many protonated molecules. A prime example is the reported permeability coefficients for charged quaternary ammonium compounds, diffusing across cell membranes[186].

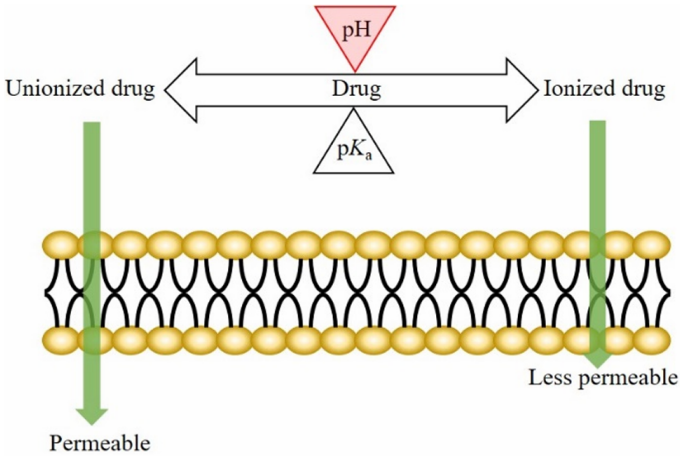


Figure 1.22: pH partition hypothesis, neutral drug candidates are more permeable across the bilayer than charged molecules[34]

Mounting experimental evidence has shown that the pH partition hypothesis has been challenged in recent years, as certain 'CAD' molecules have been shown to migrate across the aqueous phase into the membrane, in their protonated state. MD simulations have shown that these drug molecules seem to localise in the interface where the hydrophilic lipid heads are in high abundance. Once a charged drug molecule, propranolol, migrates to the interface, its ionisable groups are stabilised by their molecular interaction with isoelectric lipid heads, consequently forming an ion pair. Thus, the neutral form of propranolol has been tested to predict whether similar or different molecular behaviour is exhibited. The equilibrium position is also observed with the neutral form of propranolol, but with the drug molecule sitting in a deeper position within the interface[233]. Whilst experimentally, it is impossible to ascertain whether ionised CAD molecules can neutralise before diffusing across the cell membrane, and ultimately uncovering the precise location where this event

occurs, computationally this detail can be provided. Part of this project will help to uncover this molecular detail, extending to both neutral and protonated forms of other CAD molecules, namely fluoxetine and phentermine.

1.4.3 Lipidation and Hydrolysis at the Membrane Interface

The intrinsic reactivity of molecules towards membrane lipids is the central, dominant theme in this research project. The factors which control the chemical interaction between these CAD molecules and membrane lipids are poorly understood, and thus uncovering these reactivity factors is paramount in this research. The conventional Fluid-Mosaic model of a chemically inert and stable membrane environment has been challenged by experimental evidence pointing toward chemical instability in the membrane induced by small organic molecules at the interface. Some molecules undergo lipidation, whilst others favour hydrolysis and some exhibit no reactivity with the membrane. These molecules encompass CADs, small organics and even proteins/peptides. New evidence has come to light that membrane lipids with an ester bonded fatty-acyl chain, such as POPC or other diacylglycerophospholipids, can undergo nucleophilic attack of their carbonyl carbon atom by nucleophilic atoms present in some CAD molecules[183].

Lipidation is a type of lytic reaction, or direct acyl transfer reaction, where the CAD molecule directly interacts with the lipid carbonyl. This reaction causes the fatty acyl chain to be directly transferred from the lipid to the drug either through transesterification (figure 1.23) or aminolysis (figure 1.24) mechanism pathways. Aminolysis products are formed when a nitrogen nucleophile attacks the carbonyl and the transesterification products are formed when an oxygen nucleophile attacks the carbonyl group of the lipid within the membrane. In the case of CAD molecules such as propranolol, this process occurs in the membrane interface, alongside lysolipid formation. Propranolol has the potential to undergo both transesterification and aminolysis reactions based upon its amine and alcohol functionalities[24].

The kinetically competing reaction to lipidation is hydrolysis. Certain other CAD molecules, such as fluoxetine and sertraline, promote lipid hydrolysis without un-

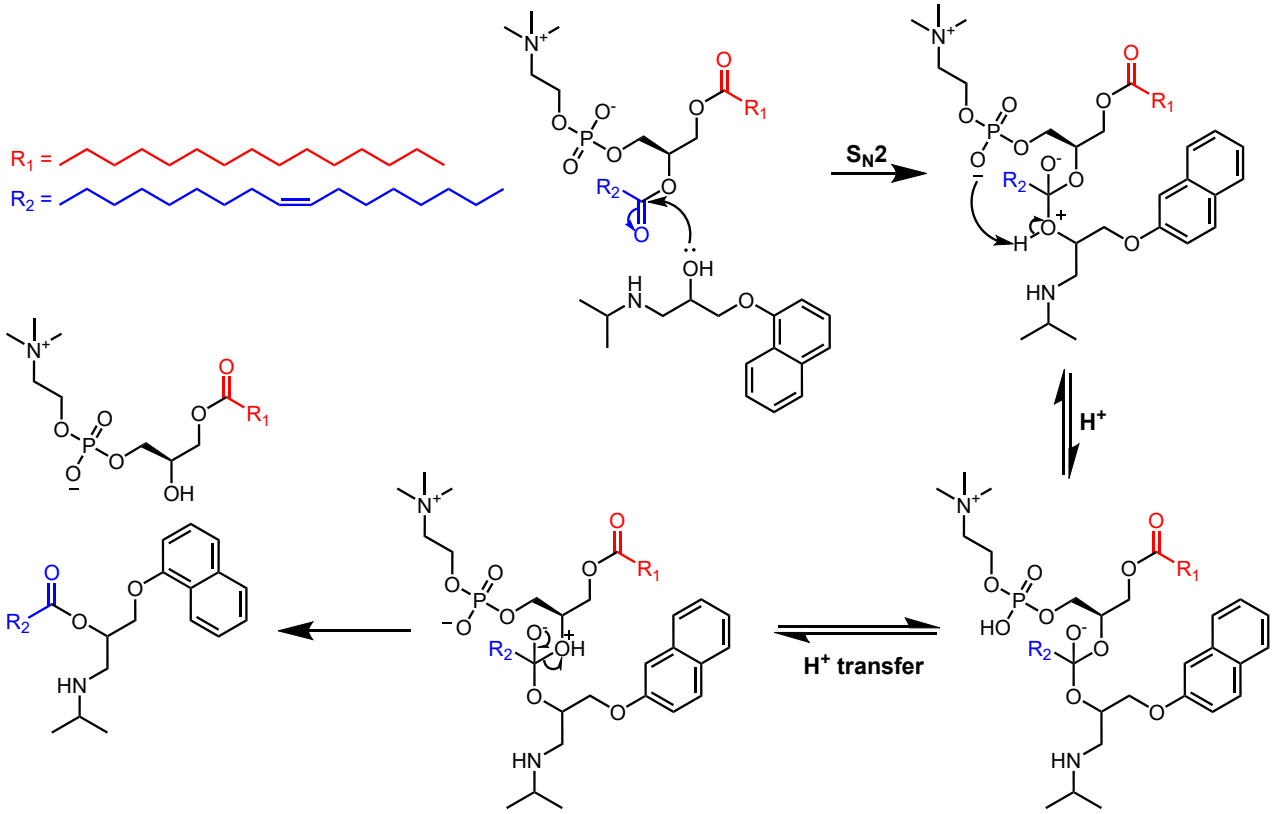


Figure 1.23: Transesterification mechanism of propranolol, leading to ester formation

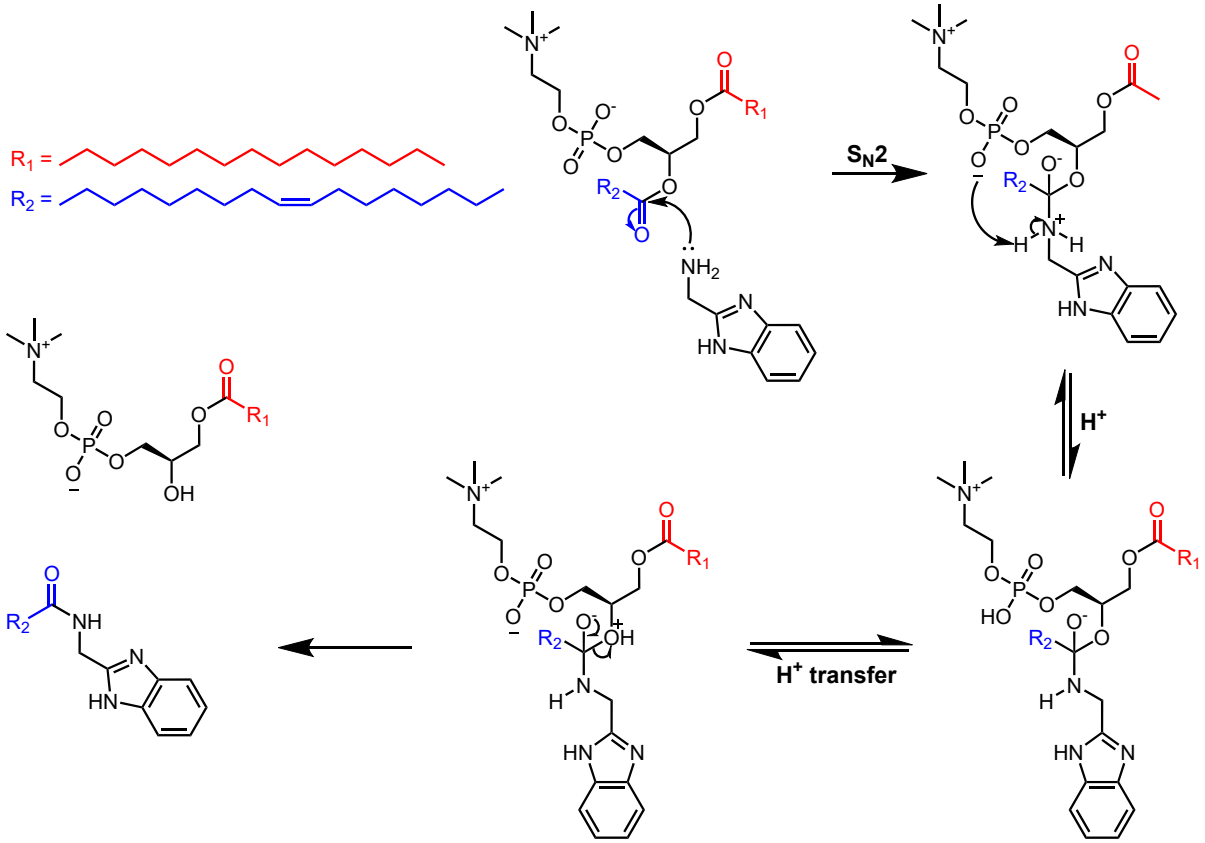


Figure 1.24: Aminolysis mechanism of 2-aminomethylbenzimidazole, leading to amide formation

dergoing lipidation reactions. Alongside lipidation, this process also results in the formation of lysolipids. In the case of fluoxetine, this CAD has been proposed to act as a phase transfer catalyst to promote hydrolysis in the membrane[54]. In part, this is due to a higher LogD value and thus the drug is more lipophilic. This leads to increased solubility in high lipid environments which result in increased hydrolytic activity. There also exists the potential for lipidated products from direct lipidation reactions to then promote further lysis of lipids. It is therefore desirable to quantify the kinetics of these two processes and further probe the mechanism by which one molecule undergoes lipidation over hydrolysis[25] (see figure 1.25).

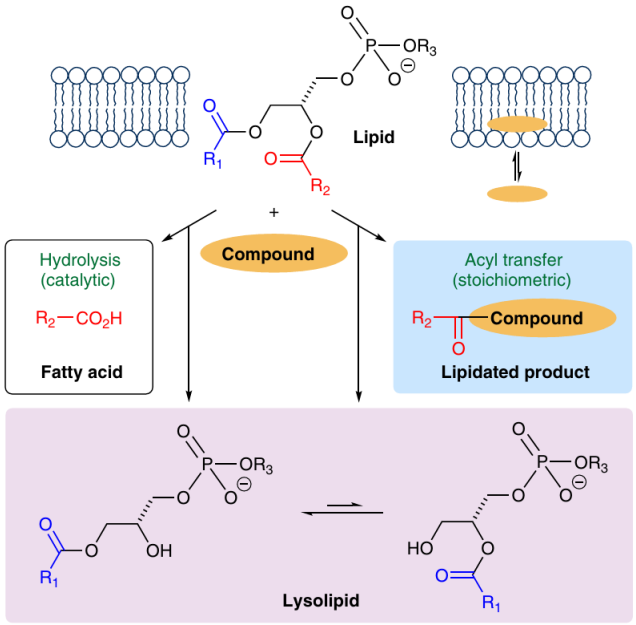


Figure 1.25: Comparison of lipidation, by means of acyl transfer, with hydrolysis, both favouring lysolipid formation, the former leads to lipidated product whilst the latter favours fatty acid production[25]

Fundamentally, the effect of drug lipidation or hydrolysis on the membrane seems to have no dependence on membrane affinity but purely on the depth/orientation of partitioning of these small reactive CAD molecules. Interestingly for propranolol, experimental evidence shows that lipidation occurs almost entirely on the oxygen nucleophile, and only when the drug has its amine group protonated at physiological pH. A reason is due to a minor component of this existing as the neutral amine form, and preferred depth/orientation enables the oxygen to be in closer proximity than the nitrogen and thus more likely to react. Whilst probing further small CAD molecules, it has been shown the lipidation reactivity is disfavoured if the reaction site is close to bulky side chains and thus sterics play an important role in the competing reactions. Furthermore, a change in the reactive site from an amine to an alcohol almost exclusively negates reactivity, as measured by a negligible lysolipid formation. Thus, there is sufficient credence to state that the localisation and positioning of the molecule in the membrane is of far more significance for reactivity than its membrane affinity or partitioning behaviour[54].

1.5 Aims and Objectives

The overarching theme of this project is to uncover the underlying factors which affect the reactivity of molecules with membrane lipids. A hypothesis would be that the orientation and depth of partitioning of these molecules affects reactivity. As has been discussed in the literature, there are a whole host of biophysical and analytical techniques which can be utilised to assess drug-lipid interactions. I have broadly divided this thesis into experimental and computational chapters and have sought to show the relevance and bridging of these two separate techniques to add confidence to our understanding of reactivity.

Initially I describe the synthesis of the relevant ^{15}N labelled CAD molecules, to aid in our understanding of which molecules undergo lipidation versus hydrolysis behaviour. Having researched the economy of purchasing ^{15}N labelled precursors for the synthesis of labelled CAD molecules, I have undertaken an additional synthetic methodology project to incorporate ^{15}N labelled molecules into any aniline-derivatised CAD molecule. I show this can be applied to a relatively wide substrate scope, with suitable yields to be used in drug-liposome solid-state NMR experiments. The main experimental analysis of these interactions is covered in a chapter dedicated to solid-state NMR analysis. Here, the appropriate NOESY, HMBC and HSQC data are presented for a few CAD molecules. This data will show the key interactions between the drug molecule and POPC lipid and highlight the significance of depth and orientation of the molecule in the membrane.

The next chapter will use atomistic simulations as a comparison and complementary technique to solid-state NMR, to show the preferred orientation, stability and depth of the molecule in the membrane. The final results chapter will use QM/MM simulations to model the reactive intermediates in the overall reaction pathway, to help identify the rate determining step in the reaction. The combination of experimental and theoretical work makes this project insightful and gives clarification on the ultimate factors underpinning molecular reactivity in membrane lipids.

Chapter 2

Chemical Synthesis of ^{15}N Molecules and Development of ^{15}N Labelling Methodology

2.1 Proposed Synthetic Targets

The scope of my synthesis includes both CAD and small organic molecules. The motivation for making these targets is to study their behaviour once embedded within membrane lipids. As alluded to in the previous chapter, some of these molecules exhibit direct lipidation, whilst others demonstrate hydrolysis behaviour and still others show no reactivity[24]. Propranolol has been found to undergo direct lipidation with membrane lipids in a transesterification mechanism, resulting in lysolipid formation. Both 2-Aminomethylbenzimidazole and 4-amino-*N*-phenylbutanamide also undergo lipidation with membrane lipids, but via an aminolysis mechanism. These are appropriate synthetic targets, as propranolol **14** shows lysolipid increase, 4-amino-*N*-phenylbutanamide **23** shows no lysolipid change and 2-aminomethylbenzimidazole **19** shows lysolipid decrease[54]. The distinct difference in lysolipid formation can be quantified with solid-state NMR measurements. Thus, the objective is to synthesise these targets in good yield and high purity for subsequent ssNMR experiments. The presence of an ^{15}N enriched molecule greatly improves signal to noise ratio in ssNMR experiments, and thus each ^{15}N analogue of

these molecules were synthesised (figure 2.1). Additionally, ^{15}N enrichment enables improved HMBC and NOE distance measurements in ssNMR for drug-lipid reactivity profiles. The distinct functionality of the naphthalen-1-yloxy and benzimidazole rings in propranolol and 2-aminomethylbenzimidazole respectively furnishes good reactivity for these molecules with wide literature precedence for their synthesis[14].

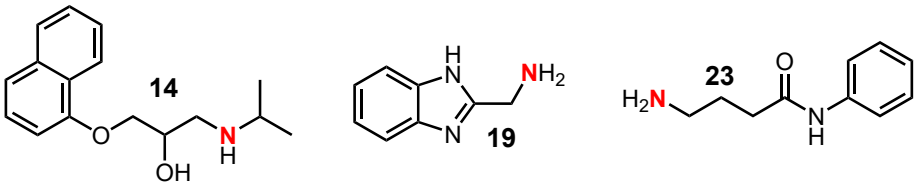


Figure 2.1: ^{15}N synthetic targets

2.1.1 Review of CAD Synthetic Literature

CADs are organic molecules which are defined by having a spatial separation between a cationic group and a hydrophobic segment. In the drug discovery process, structure-activity relationships (SARs) involve making systematic structural changes and relating these numerically to their activity. They are formed from precise and efficient syntheses of smaller building blocks into larger functional molecules[188]. Historically, there have been two key organic strategies used in drug synthesis. Firstly, the total synthesis of natural products. This area of synthesis is inspired from natural products. The major clinical use of these natural products is as novel medicines. Since the 1940s, it has been calculated that nearly half of all anti-cancer drugs were inspired from either natural products or semi-synthetic structures, in the design of novel chemotherapeutics[21]. An example of such is the piperazine fragment which selectively binds to a protein (figure 2.2), which is then incorporated into a lead compound:

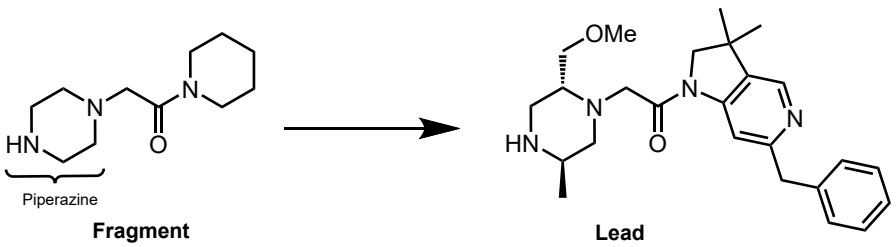


Figure 2.2: Anti-cancer agent, piperazine fragment present in the lead compound

The second strategy, namely the one employed in this project, takes its inspira-

tion from the chemical diversity and structural complexity of different classes of molecules, rather than specific natural products. This approach is known as modular/small molecule synthesis. Here, the chemist is interested in potential stereogenic sites, chemical functionality and biological action. Heterocyclic molecules have been found to be appropriate building blocks for the synthesis of more complex CAD molecules. In 2014, an FDA (Food and Drug Administration) report found that heterocyclic building blocks represent a huge structural array of different drug molecules. Piperidine, pyridine, piperazine, cepham, pyrrolidine and thiazole were reported to account for over 300 of around 1,000 FDA approved drugs on the market. This provided confidence that a larger portion of pharmaceutics could be synthesised from a common set of precursor building blocks[117].

Both five and six membered heterocyclic molecules have been found to have remarkably high pharmaceutical benefits and thus are routinely marketed as clinical medicines. A classic example of a piperidine-type drug is nifedipine (figure 2.3). In the early 1970s this was one of the first synthesised drugs marketed as an antihypertensive, resulting in a reduction in high blood pressure. The synthesis of this drug utilises the Hantzsch reaction. This is a condensation reaction involving 2-nitrobenzaldehyde and methylacetoacetate with an ammonia source. The far more stable magnesium nitride is used *in situ* with the reagents, and upon hydrolysis, releases the gaseous ammonia reagent[14]. The reaction is in high yield (84%) with limited reagents, showing the success of a small molecules synthesis:

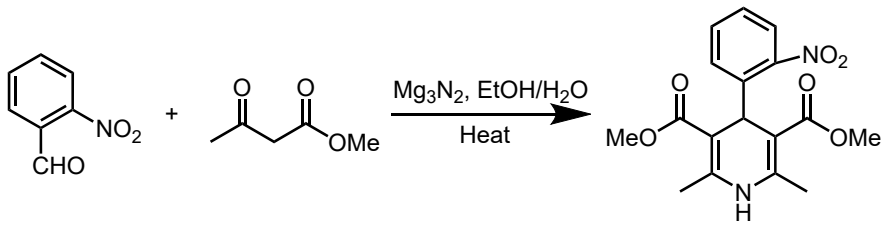


Figure 2.3: Synthesis of nifedipine

A prime example of a five membered carbazole-type drug is carvedilol. Carvedilol, like propranolol, is a beta-blocker with non-selective binding properties. The presence of a carbazole ring is crucial to its biological function in the treatment of congestive heart failure, from acute to chronic effects. Clinical evidence points to carvedilol having a higher effect than propranolol, and also having potential an-

tioxidant properties. Its mode of action, by the carbazole ring, is in capturing harmful oxygen radicals, and thus reducing heart damage. This pharmaceutical is made across a four step synthesis (figure 2.4). The first step is the condensation of phenylhydrazine with 1,3-cylohexanedione to form cyclohexane-1,3-dione monophenyl hydrazone (64%). The second step, being a typical Fischer indole synthesis, forms tetrahydro-4-oxocarbazole from an acid-catalysed reaction. A high yield of 83% is then achieved by the dehydrogenation of this compound to 4-hydroxy-9*H*-carbazole[14]. This reaction utilises catalytic Raney Ni in the presence of aqueous potassium hydroxide. This key intermediate then undergoes substitution in the final step to form the drug compound:

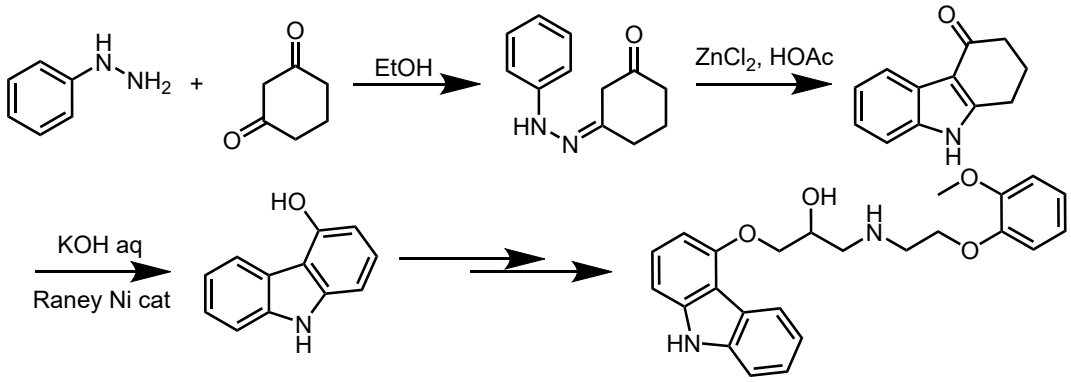


Figure 2.4: Synthesis of carvedilol

A final significant example of a five membered heterocyclic benzimidazole ring is the drug candesartan. This molecule has a core benzimidazole ring, similar to 2-aminomethylbenzimidazole studied in this project. This drug molecule acts as an angiotensin II antagonist. Unlike the previous examples, this synthesis is quite complex both in reagents and number of synthetic steps (figure 2.5). The initial step involves mono-selective acid-catalysed esterification of the carboxylic acid compound to the methyl ester compound (95%). The second step involves a curtius rearrangement, with the diphenyl phosphorazidate reagent. The third step then involves selective alkylation with the biphenylmethyl bromide compound. This reaction proceeds with acid and base conditions. This furnishes a good 75% yield with a bulky biphenyl substituent attached to the main ring. The fourth step makes use of tin chloride as a reagent to promote nitro reduction to the amine. Closing of the benzimidazole ring and attachment of the ethoxy group on the two position of ring then takes place to form the precursor[164]. The final drug is then the triazole

substituted molecule candersartan:

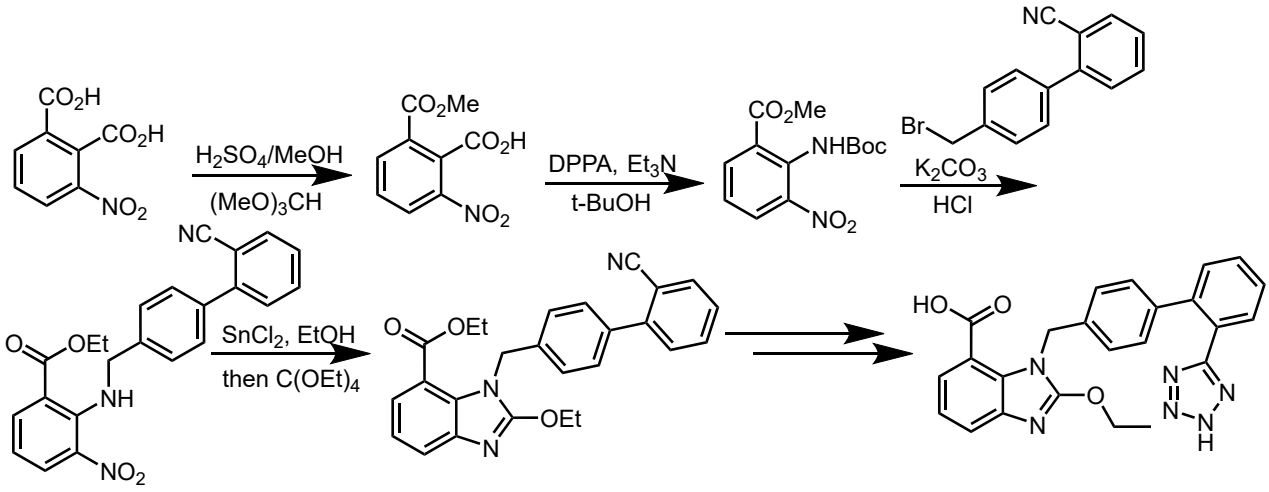


Figure 2.5: Synthesis of candesartan

2.1.2 Synthesis of ^{15}N analogues of propranolol, 2-aminomethyl benzimidazole and 4-amino-*N*-phenylbutanamide

^{15}N propranolol

Previous syntheses of propranolol have involved a multi-step synthesis to form epichlorohydrin as the intermediate in the overall propranolol synthesis (figure 2.6) [13]. These reactions require purification at each step and often with a poor yield at the end of the reaction. It was first desirable to improve the synthetic route to epichlorohydrin, and then this synthesis was optimised by a reduction of the number of synthetic steps to achieve the final ^{15}N product without purification and in high yield.

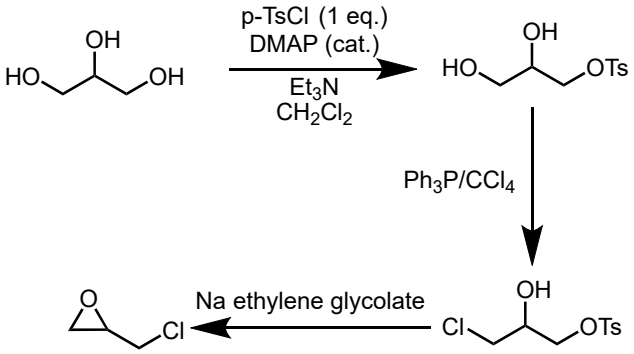


Figure 2.6: Literature synthesis of epichlorohydrin intermediate

Various synthetic pathways were tested to find the optimal route for the synthesis of ^{15}N propranolol. Initially, epichlorohydrin was considered an important inter-

mediate in the synthesis of propranolol (see figure 2.7). The initial synthesis was a di-tosylation reaction of the glycerol to form glycerol-1,3-ditosylate. This was achieved by using two equivalents of 4-toluenesulfonyl chloride. Triethylamine was used to deprotonate the two hydroxyl groups, which could then undergo reaction with TsCl to form the di-tosylated compound. The presence of a sharp doublet with an integration of six protons indicated successful formation of this intermediate **2** compound.

The second step, to form the epoxide, involved a selective deprotonation of the hydroxyl (pK_a = 15-16) group with K_2CO_3 (pK_a = 10.33) base. The hydroxyl anion could then act as a nucleophile to form an epoxide ring through ring closure. The presence of a singlet methyl group, with an integration area of three, alongside the CH_2 groups of the epoxide, confirmed formation of the glycidyl tosylate **3**. Whilst purification of this compound through column chromatography was required, the epoxide unfortunately stuck to the column. Thus an alternative synthetic route was proposed.

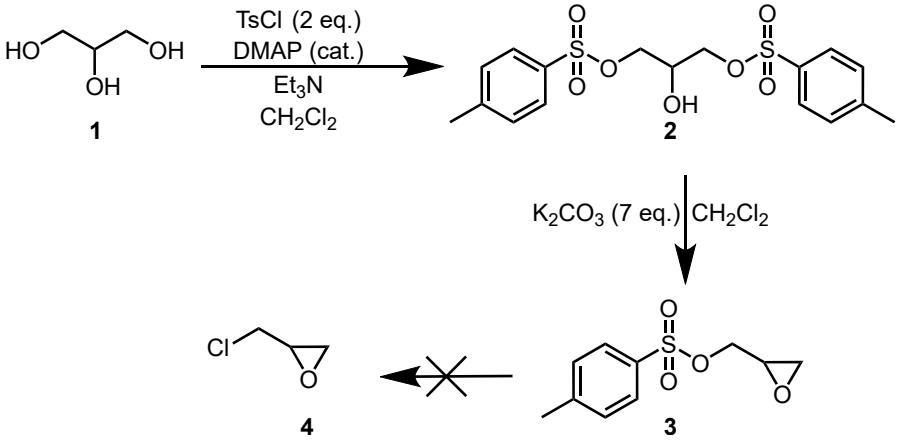


Figure 2.7: Synthesis of epichlorohydrin

This alternative method leads to the synthesis of a precursor mesylate compound, which can react with ^{15}N isopropylamine to form ^{15}N Propranolol. The first step involves selective monotosylation of glycerol with one equivalent of 4-toluenesulfonyl chloride to form the substituted glycerol. This compound was then naphthylated with 1.2 eq. of 1-naphthol, whereby the tosylated group was subsituted for the naphthyl group. The last step involved mono-mesylation at the least sterically hindered hydroxyl group (figure 2.8). Having analysed the NMR spectra, a sharp single

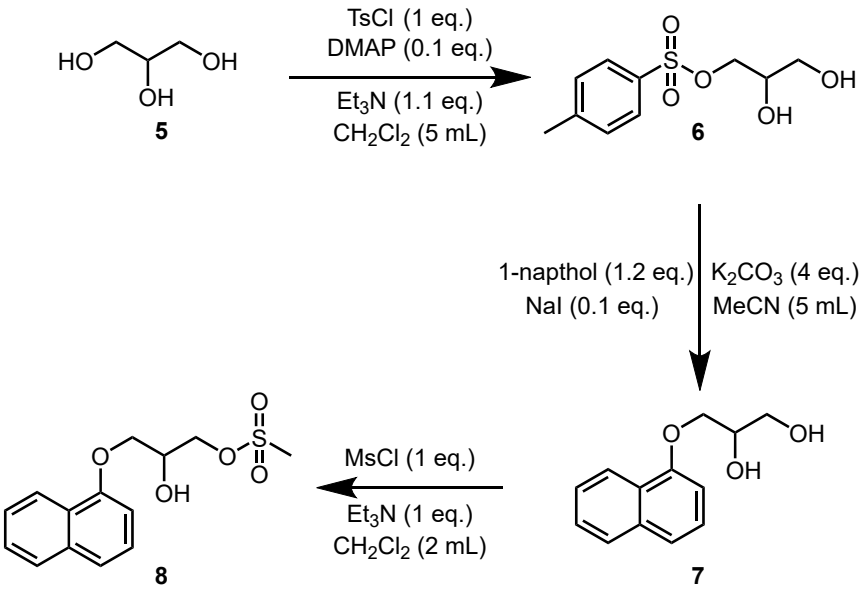


Figure 2.8: Synthesis of mesylated intermediate

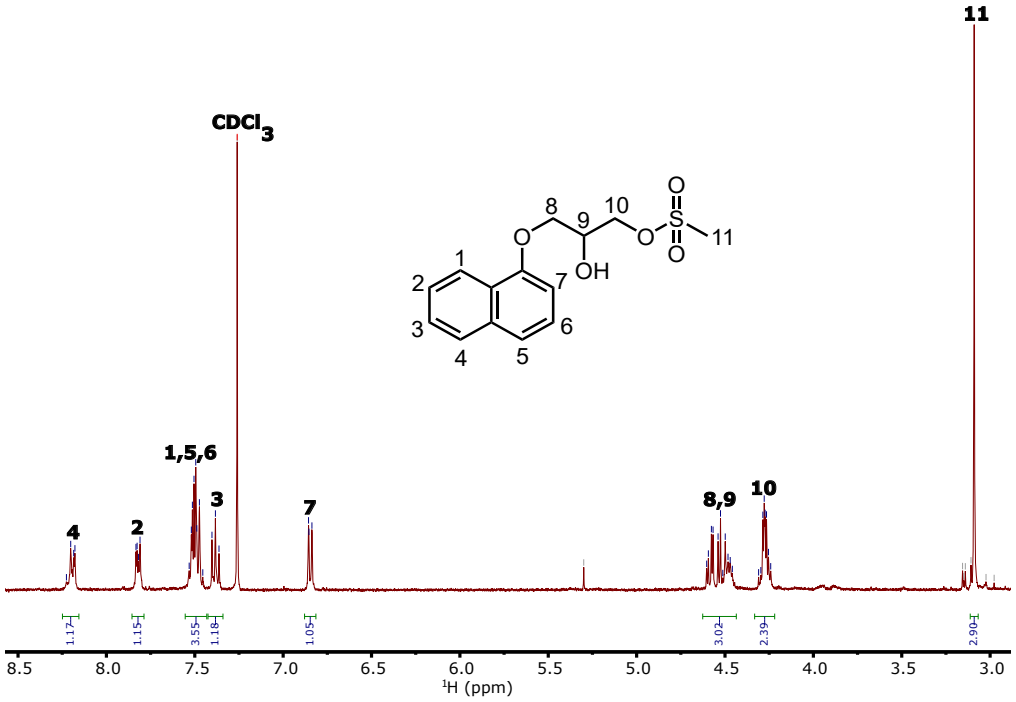


Figure 2.9: ^{1}H NMR spectrum ($CDCl_3$) of 2-hydroxy-3-(2-naphthalenyloxy)propyl methanesulfonate

peak for both the tosylated **6** and mesylated **8** (figure 2.9) compounds indicates successful synthesis of these molecules. The major disadvantage of this synthetic route is a large number of steps, thus a final efficient two step method was used in the synthesis of ^{15}N propranolol.

The final method, taken from the literature[24], involved initially the synthesis of ^{15}N isopropylamine. 2-Bromopropane was refluxed with ^{15}N potassium phthalimide to form ^{15}N isopropylphthalimide in the Gabriel synthesis. With the use of hydrazine hydrate in MeOH, this phthalimide group was deprotected to form the ^{15}N isopropylamine intermediate **11**. For the propranolol precursor **13**, the two step synthesis involved *in situ* formation of the napthylated intermediate with a chloride attached to the least sterically hindered part of the molecule (figure 2.10). This synthetic step was efficient in forming both the napthylated part of the molecule, alongside the commercially available epoxide **12**, which could then be ring opened with the pyridine base. The second step was simply the nucleophilic substitution of the chloride with the ^{15}N isopropylamine hydrochloride. Both the chloride intermediate and ^{15}N propranolol compounds were pure, with no need for purification, and resulted in a final yield for ^{15}N propranolol (figure 2.11) of 37%.

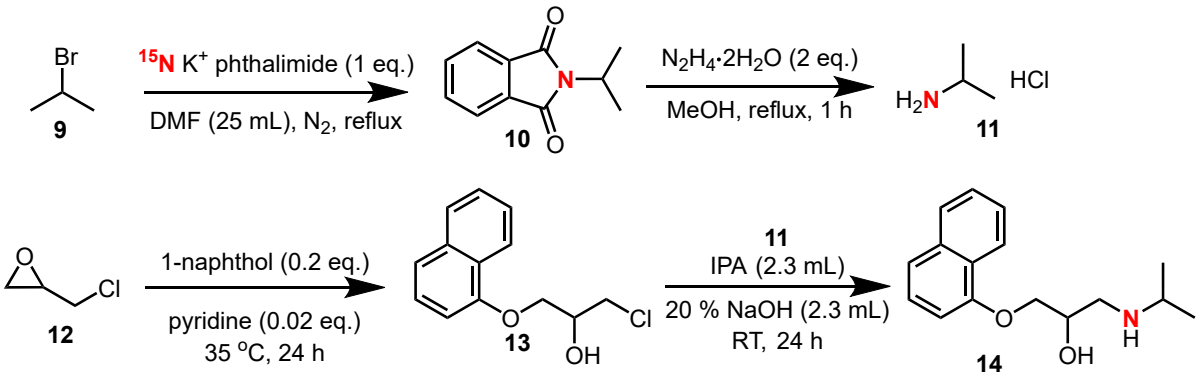


Figure 2.10: ^{15}N propranolol synthesis

^{15}N 2-aminomethylbenzimidazole

Previous syntheses of 2-aminomethylbenzimidazole have employed a condensation reaction of glycine with 1,2-diaminobenzene (figure 2.12)[36]. This reaction gives modest yields and a challenging purification, with the principal by-products forming by the addition of multiple glycine molecules. Nevertheless, as ^{15}N labelled glycine is readily available, it was desirable to use this as a stating material and to attempt

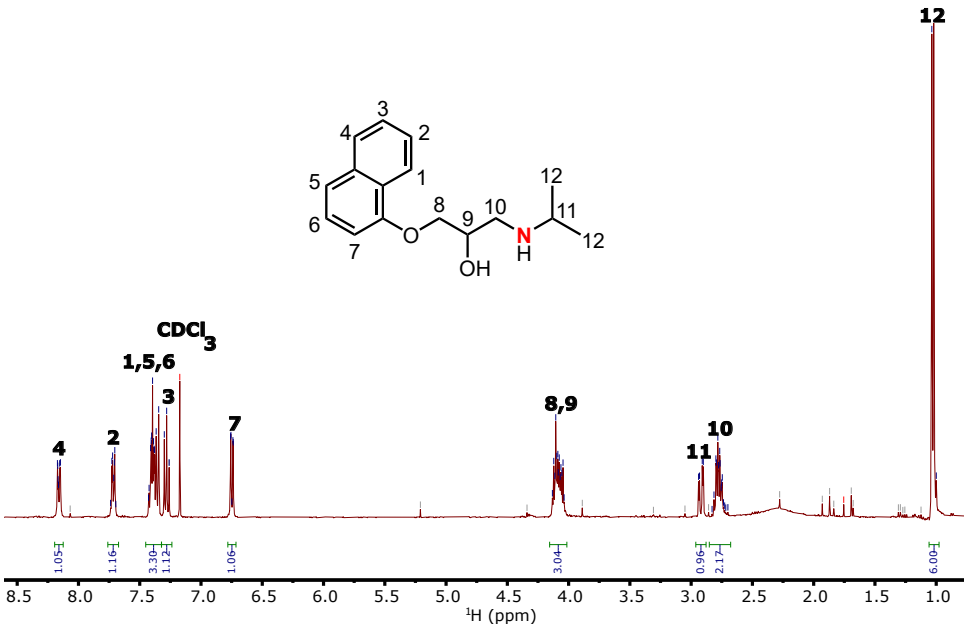


Figure 2.11: ^{1}H NMR spectrum ($CDCl_3$) of ^{15}N propranolol

to improve this reaction.

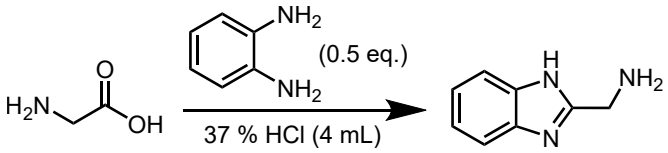


Figure 2.12: Literature synthesis of 2-aminomethylbenzimidazole[36]

Attempts were made to perform this reaction with a suitably amino-protected glycine (figure 2.13). A variety of different protecting groups were used in the synthetic pathway of this molecule to optimise yield and purity of the final labelled product. In the initial synthesis of ^{15}N 2-aminomethylbenzimidazole, the amine group of glycine had to be protected. This was achieved by protection with *tert*-butoxycarbonyl (BOC), benzyloxycarbonyl (Cbz) or fluorenylmethoxycarbonyl (Fmoc) chemistry (figure 2.14). Each protecting group was added to ^{15}N labelled glycine with different quantities to ensure single protection. Whilst protection, addition (figure 2.16) and cyclisation proved to be effective for the Cbz and Fmoc protected molecules (figure 2.13), their final deprotection was found to be far more challenging than anticipated. For Cbz deprotection, hydrogenation with Pd/C (10%) in EtOH was chosen as suitable conditions. This deprotection resulted in many impurities in the final molecule, too many to accurately purify. For Fmoc deprotection, likewise 20% piperidine in DMF was found as an appropriate deprotecting agent.

However, many impurities were retained in the final product.

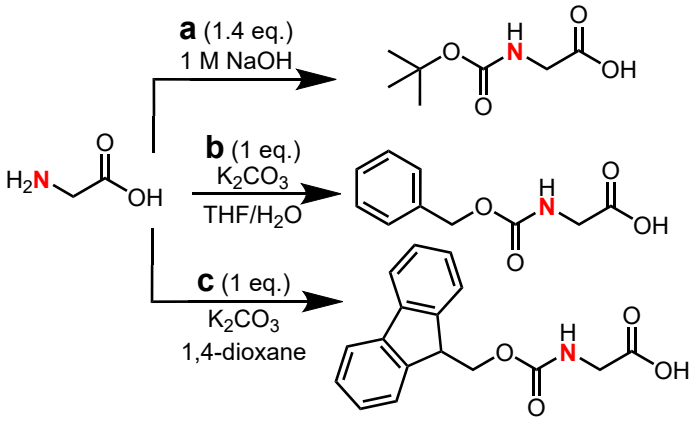


Figure 2.13: Protecting group reactions

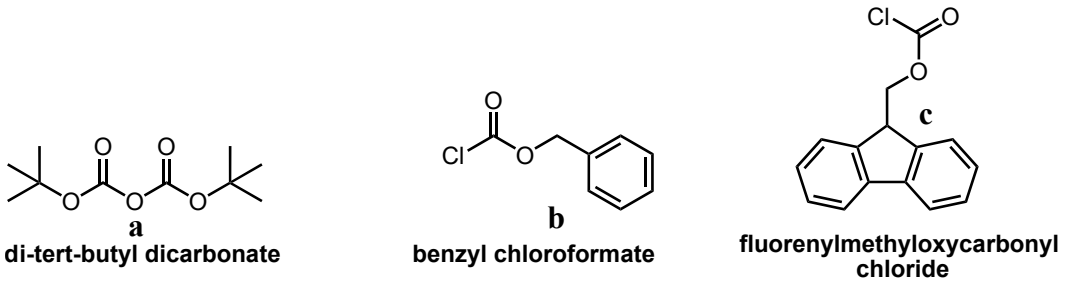


Figure 2.14: Reagents for installing protecting groups

The key to successful deprotection of the BOC group was the use of 2 M HCl in MeOH (figure 2.15). Upon trituration of the final product in ether, the impurities were soluble in the organic phase, whilst the product was able to precipitate out. After prep-HPLC, there was clean ^{15}N 2-aminomethylbenzimidazole (figure 2.17). A direct comparison of the LC-MS trace between non-labelled 2-aminomethylbenzimidazole and ^{15}N labelled 2-aminomethylbenzimidazole **19** shows a difference of 148.14 m/z (non-labelled) compared with 149.21 m/z (labelled), confirming isotopic enrichment (figure 2.18). As the expected m/z for the labelled molecule is 149.18, the actual value is within a good error margin to the expected.

^{15}N 4-Amino-*N*-phenylbutanamide

The final molecule in this set of CAD syntheses is the molecule ^{15}N 4-amino-*N*-phenylbutanamide. Fortunately, the synthesis of both this non-labelled and ^{15}N molecule utilises similar protecting group chemistry as ^{15}N 2-aminomethylbenzimidazole. The initial step utilises only 1.2 equivalents of BOC anhydride to ensure full BOC protection of the amine group. CDI is a coupling agent, often used in peptide

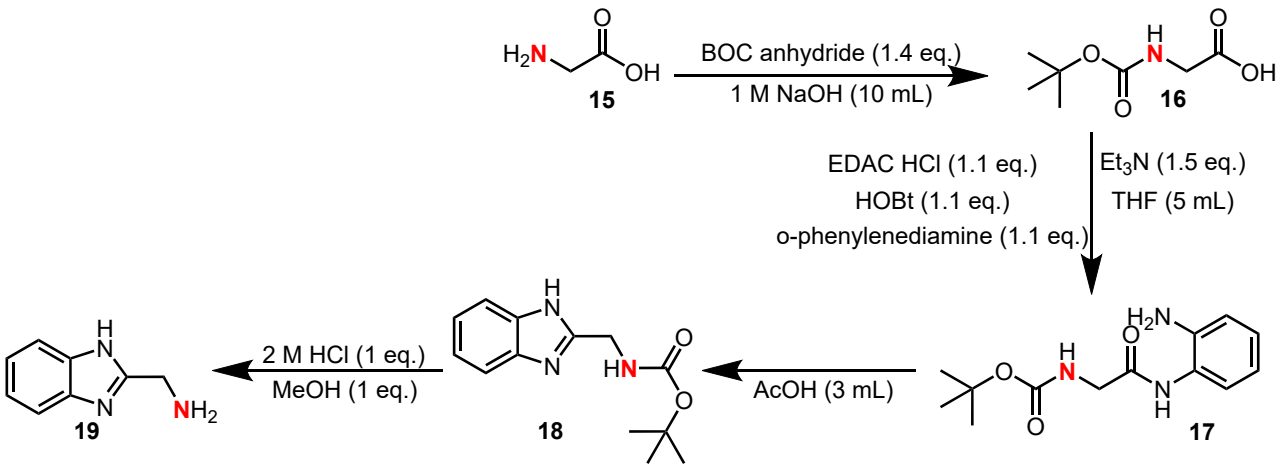


Figure 2.15: Optimised ^{15}N 2-aminomethylbenzimidazole synthetic pathway

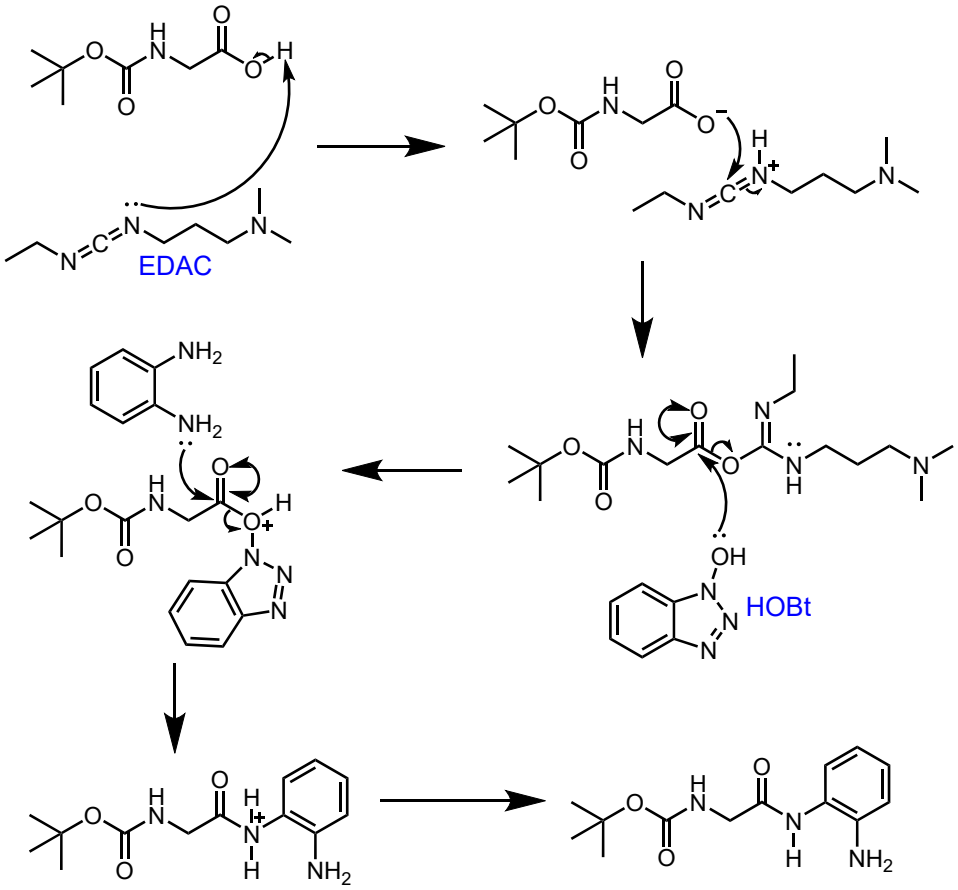


Figure 2.16: Role of EDAC and HOBt in formation of precursor molecule to 2-aminomethylbenzimidazole

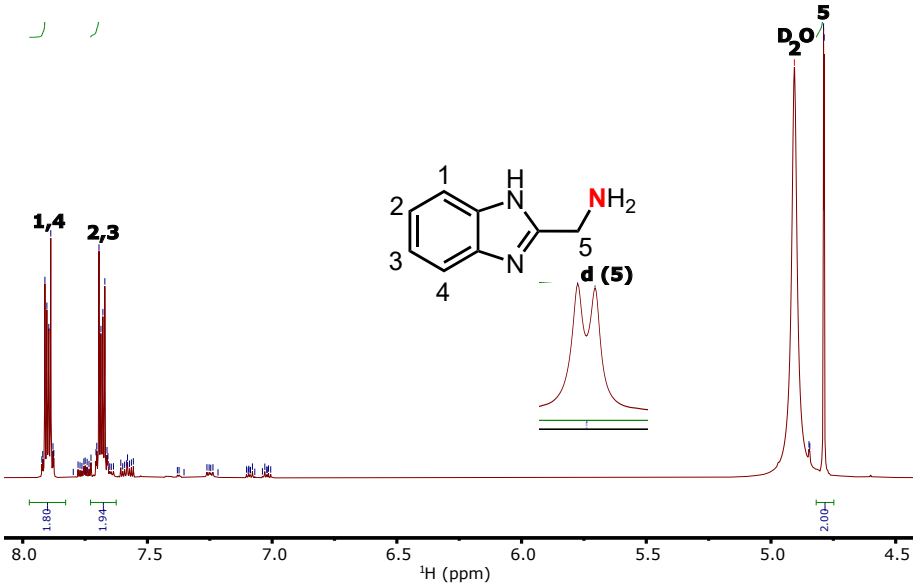


Figure 2.17: ^{1}H NMR (MeOD) of ^{15}N labelled 2-aminomethylbenzimidazole

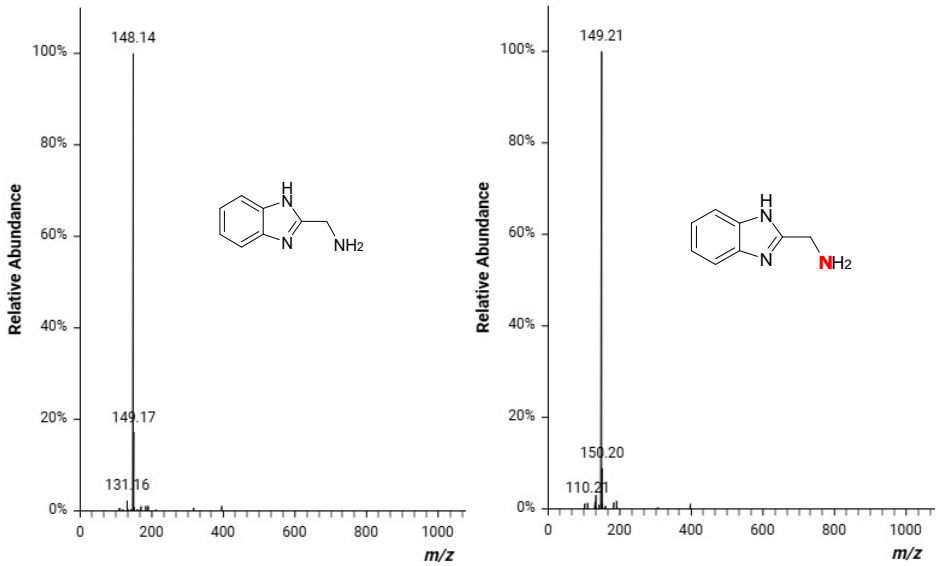


Figure 2.18: Mass spectrometry comparison between non-labelled and ^{15}N labelled 2-aminomethylbenzimidazole

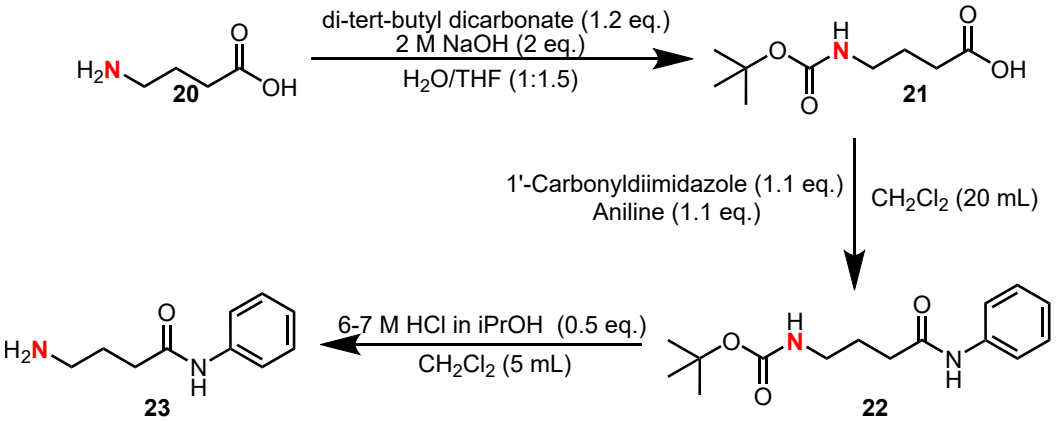


Figure 2.19: ^{15}N 4-amino-*N*-phenylbutanamide synthetic route

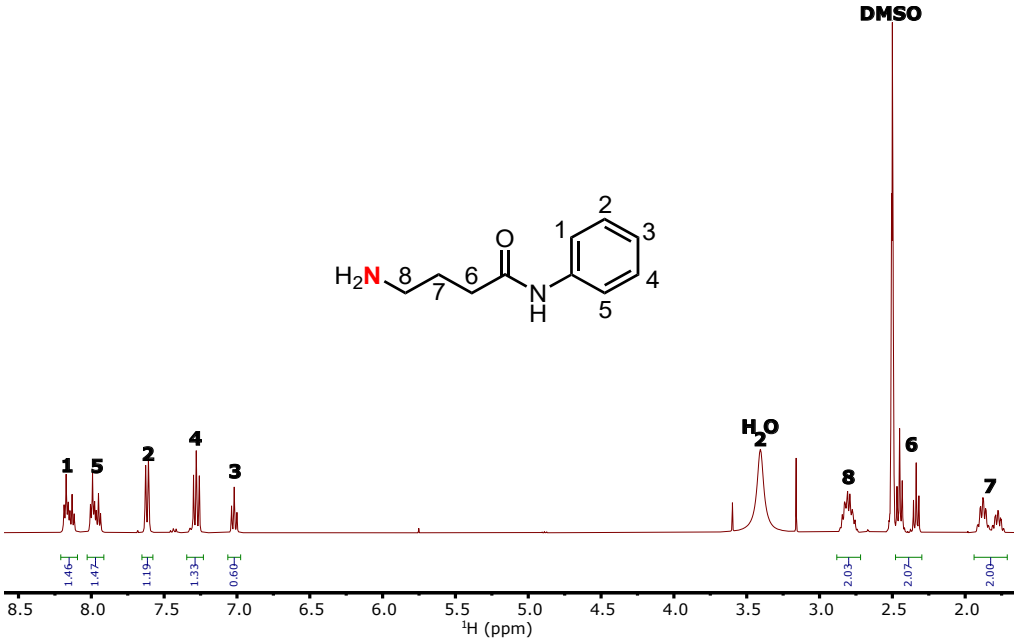


Figure 2.20: ^{1}H NMR (MeOD) of ^{15}N 4-amino-*N*-phenylbutanamide

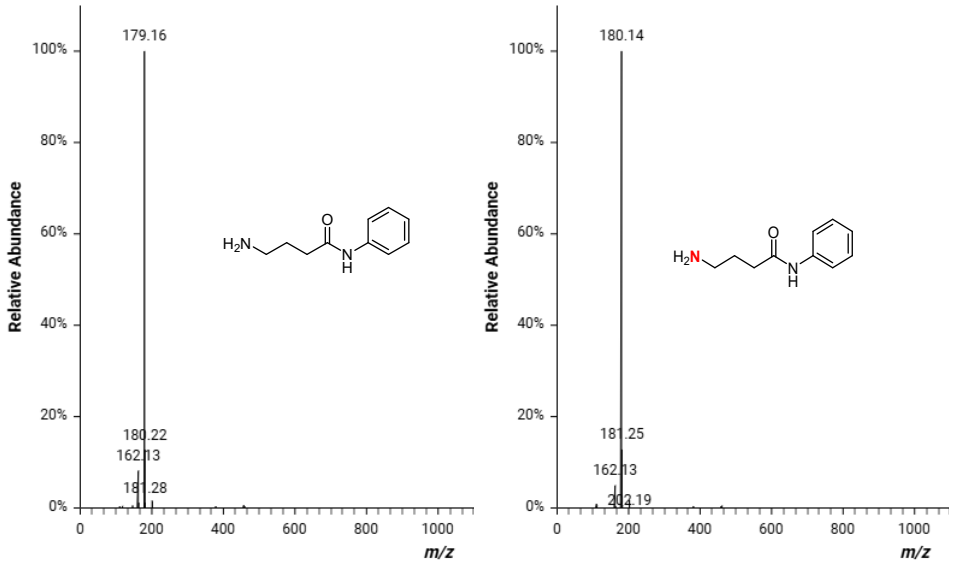


Figure 2.21: Mass spectrometry comparison between non-labelled and ^{15}N labelled 4-amino-*N*-phenylbutanamide

synthesis, which activates the carboxylic acid to favour reactivity with the aniline, in forming the amide. The final step utilises similar conditions to ^{15}N 2-aminomethylbenzimidazole, whereby the basic BOC group is hydrolysed in the presence of 6 - 7 M HCl in isopropyl alcohol (figure 2.19). Similarly, this molecule is triturated with diethyl ether. The impurities are soluble in the organic phase and the product precipitates out leaving a pure grey solid. LC-MS analysis confirms that non-labelled 4-amino-*N*-phenylbutanamide has one mass unit lower (179.16 m/z) than its ^{15}N enriched analogue **23** (180.14 m/z) (figure 2.21). Analysis of the ^{1}H NMR spectrum reveals successful synthesis of ^{15}N 4-amino-*N*-phenylbutanamide (figure 2.20).

2.2 Proposed ^{15}N Labelling Methodology

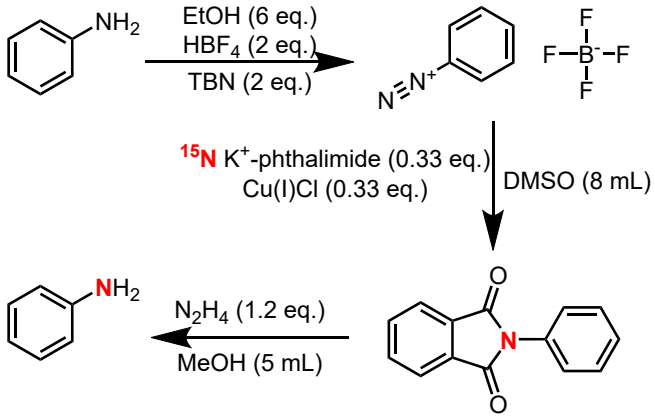


Figure 2.22: Proposed synthetic route for ^{15}N labelling method

In addition to the chemical synthesis of ^{15}N labelled molecules, I have also developed a novel synthetic methodology for the incorporation of an ^{15}N isotope into a set of aniline derivatised molecules. Specifically, a copper catalyst is required for C-N bond formation which couples an ^{15}N enriched phthalimide group to a range of substituted diazonium tetrafluoroborates. This labelled coupled molecule can then be deprotected to form the final labelled product (figure 2.22). This side project was motivated by the need to obtain ^{15}N molecules for ssNMR experiments of drug-lipid reactivity. This reaction pathway (figure 2.23) provides a cost effective, with fairly mild reaction conditions, methodology to obtain labelled molecules. This can also be applied to CAD and other biologically active molecules which have an aniline functionality.

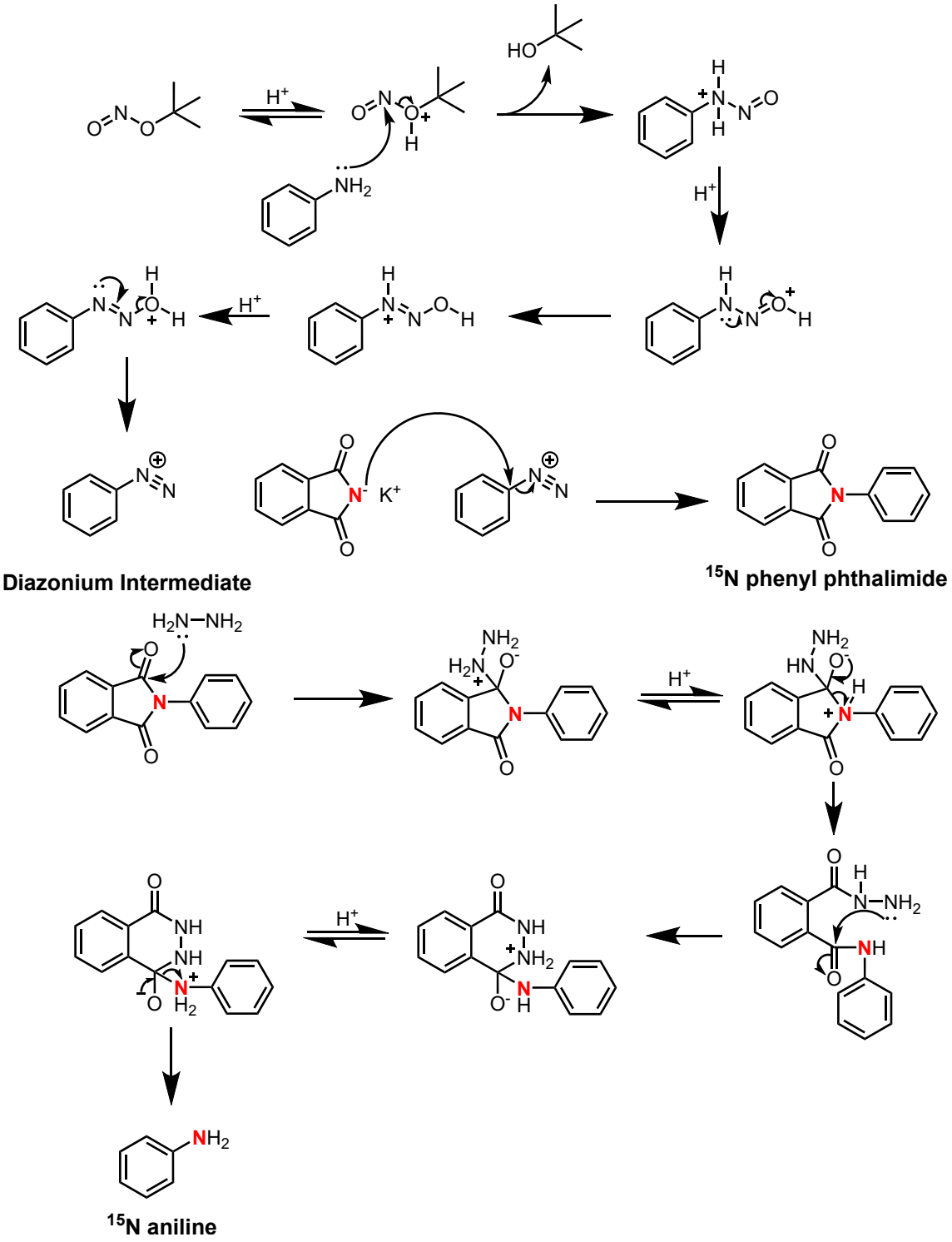


Figure 2.23: Mechanism for ^{15}N labelling method

2.2.1 Review of C-N Bond Formation Literature

C-N bond formation is found throughout all drug synthetic procedures, with more than 80% of all clinically available drug molecules having at least a single C-N bond. The development of new functional C-N bonds is therefore of interest, and its formation by the use of transition metal catalysts have transformed the field of organometallic chemistry. Historically, the scientists Ullman and Goldberg contributed to copper catalysed C-N coupling reactions, the so called Ullman synthesis makes use of this chemistry. Also, the scientists Buchwald and Hartwig developed palladium catalysed C-N coupling reactions, aptly named the Buchwald-Hartwig amination synthesis. Whilst these methodologies have proven successful in the construction of novel C-N bonds, these reactions often require prohibitively high refluxing temperatures, expensive ligands and complex starting materials. It is thus prudent to develop new C-N bond forming methods with far more mild conditions[236]. Although the use of HBF_4 is a potential environmental concern, my methodology has sought to capture greener and more sustainable chemistry.

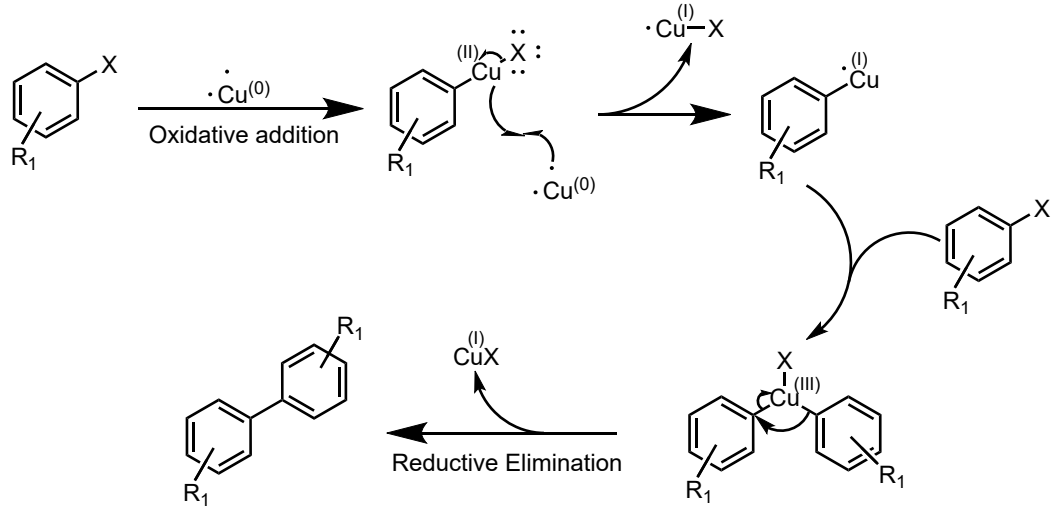


Figure 2.24: Mechanism involving aryl copper intermediate to form the biaryl compound

This field began in 1901, where Ullman discovered that copper could be used to catalyse the formation of C-C bonds between two aryl halides to form a biaryl group[3]. Mechanistically, this reaction proceeds by formation of an organocuprate from a single aryl halide (figure 2.24). This intermediate then reacts with a second aryl halide to form the biaryl. The first reaction is thus an oxidative addition and

the second a reductive elimination. This understanding was then refined to the formation of C-N bonds from *N*-aryl amines and ethers in the succeeding few years (1903 - 1905). Goldberg then successfully reported the copper catalysed C-N bond formation of an aryl amide (figure 2.25), clearly an important biological functional group[180]. Whilst these discoveries were important, harsh conditions, such as very high temperatures, were required for the formation of these compounds:

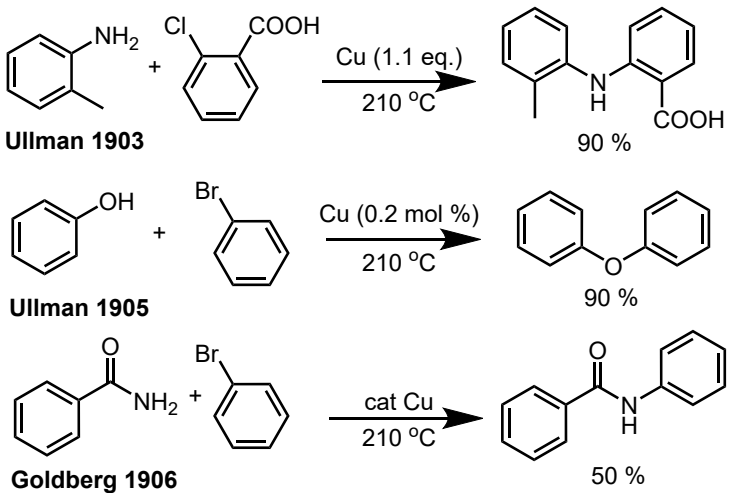


Figure 2.25: Development of methods for C-N bond formation

For the purpose of my project, copper mediated arylation of aromatic amines is of importance. In recent years, greener conditions for the Ullman coupling have been developed. These include ligand free, minimal catalyst and shorter reaction times to achieve appropriate conversion and good yields of arylated compounds. In 2007, the scientists Correa and Bolm created a ligand-free style Ullman *N*-arylation coupling. This chemistry was applied to the coupling of alkyl halides with *N*-heterocycles. The chosen catalyst was Cu_2O, providing both a catalyst and suitable oxidative conditions. Optimal reaction conditions were found to utilise 20 mol% of CuI catalyst at a reaction temperature of 35-40°C[124]. This work was further developed in 2010 to include a 3-bromopyridine coupling with a substituted amine[83]. Interestingly the yield was found to be relatively reasonable for this compound, but substitution for the 2-bromopyridine rendered a very low yield. This system utilised a Cu (solid) 10 mol% as the catalyst, DMSO solvent at high reflux temperature (189°C), with CsOAc as the base, again ligand free. The most successful synthesis in this area was discovered by the Wei group in 2011. They found that an aryl halide could react with an alkyl amine without the need for either a ligand or solvent

(figure 2.26). The catalyst requirement was only 5% copper powder, aerated, in the presence of water. The only disadvantage is the high temperature of 100°C. This produced fully coupled anilines with relatively modest yields[17].

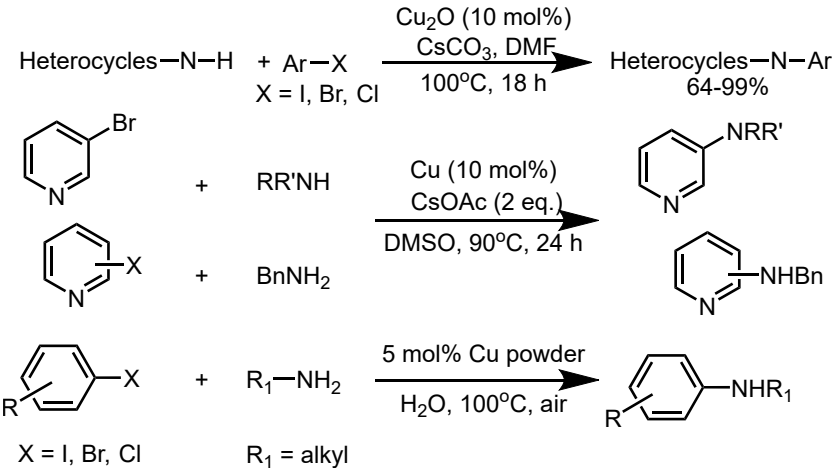


Figure 2.26: Additional methods for C-N bond formation

Whilst trying to develop a suitable ^{15}N synthetic methodology, phenylboronic acid was used as a potential reagent for new C-N bond formation from a phthalimide reagent. In 1998, Chan et al. reported the application of the Ullman coupling on the cross coupling of arylboronic acids with *N*-nucleophiles, again through the use of a suitable copper catalyst[171]. The advantage of this method provides much milder conditions and a broader range of copper concentrations. The novelty was that this coupling could be achieved with anilines, carbamates, amides, imides and a host of different *N*-substituted nucleophiles (figure 2.27), in the presence of air at room temperature. Historically, they found that stochiometric amounts of copper acetate could be used in the coupling of aromatic *N*-heterocycles with arylboronic acids resulting in good chemical yields. Although this required large amounts of copper catalyst, a couple of years later Collman et al. demonstrated that the same coupling could be achieved with only catalytic amounts of copper. They discovered that with only 10% copper catalyst, arylboronic acids could be efficiently coupled with imidazoles, with a higher yield achieved in DCM solvent rather than water. The scientists also reported that the presence of molecular sieves in the reaction could improve chemical yield by a reduction in unwanted side-product formation[204].

The choice of an oxidising agent, to ensure oxidative addition of Cu(I) to Cu(II), can

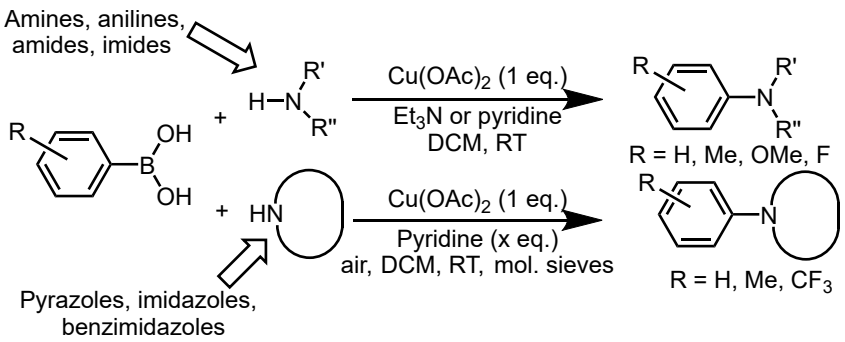


Figure 2.27: Application of C-N bond formation to different *N*-substituted heterocycles

contribute to a greater yield of the coupled product. Compounds such as 2,2,6,6-tetramethyl-1-piperidinyloxy (TEMPO) or pyridine *N*-oxide (PNO) were proposed as suitable oxidising agents, thus devising reaction conditions for the coupling of substituted phenylboronic acids with a range of different *N*-nucleophiles, including phthalimides[161], under aerated and 10% copper catalysed reaction conditions with DCM solvent. This work culminated in the use of catalytic copper chloride/acetate. This synthetic method was both ligand and base free (figure 2.28). Efficient coupling of arylboronic acids to substituted amine compounds was achieved in good yields. Furthermore, the group of Sreedhar et al. discovered that this reaction could be performed at room temperature in the presence of catalytic Cu_2O. The benefit of this second method is that the catalyst can be recycled simply without affecting reactivity[5].

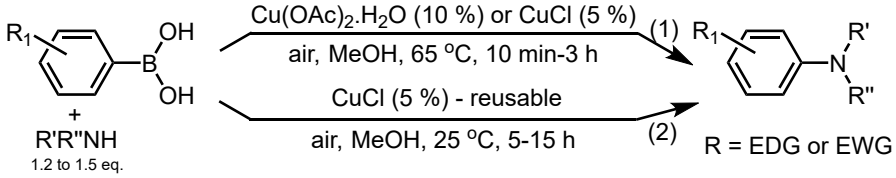


Figure 2.28: Application of C-N bond formation to coupling of substituted phenylboronic acids with substituted amines

Having established that this C-N bond formation can be utilised in the coupling of aryl boronic acids to phthalimides (*N*-heterocyclic systems), it has been shown that the same coupling can take place between phthalimides and arenediazoniums, hence the need to develop an efficient ^{15}N coupling methodology. This chemistry is known as the Meerwein arylation (figure 2.29). The catalysis by Cu(I)Cl is cyclical and was first reported with the coupling of an arenediazonium tetrafluoroborate to

maleimide. The oxidation of Cu(I) to Cu(II) results in the diazonium being reduced to a radical species and the generation of N_2 gas. This radical species is then coupled with the maleimide to form an unstable intermediate. Reduction of Cu(II) to Cu(I) then causes the hydride radical to form the stable coupled compound[232]. This cyclical mechanism is the basis of my own methodology, where the *N*-heterocyclic compound is coupled with ^{15}N potassium phthalimide.

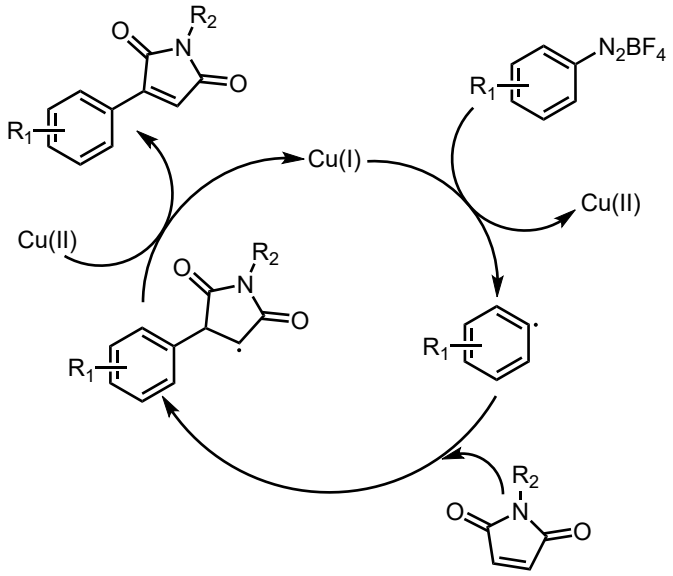


Figure 2.29: Mechanism for coupling of diazonium with maleimide via a radical species

The inspiration for this ^{15}N synthetic method came from the 2017 paper by Ouyang et al. where they reported the copper catalysed cross coupling of arenediazonium tetrafluoroborates with a range of substituted succinimides and phthalimides[151] (figure 2.30). This paper provided an adequate comparison of different catalysts, solvents and reaction times, ultimately to report whether chemical yields were efficient. Initially, DMSO solvent, 25°C and aerated conditions were used in the reaction. No product formation was reported when the reaction was both base and catalyst free. Having undergone screening of this method, the use of catalytic CuI yielded only a 33% chemical yield of the coupled product, compared with 63% (CuCl). Inclusion of a base (Et_3N) significantly boosted this yield up to 76%. The optimal conditions for this reaction, which subsequently gave 93% chemical yield, were using K_2CO_3 (base), DMSO (solvent) and CuCl (catalyst). Solubility of all reagents is far higher in DMSO solvent and the choice of CuCl implies greater yield than other copper halide catalysts[151]. These conditions were utilised in my own synthetic methodology,

with modifications to improve yield and accuracy of experimentation.

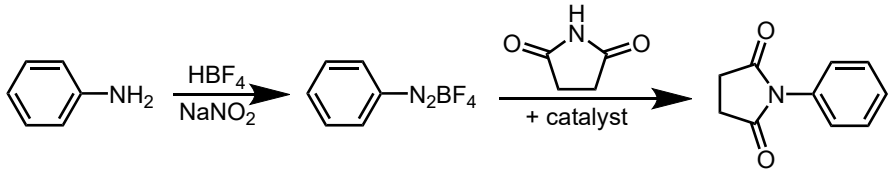


Figure 2.30: Literature synthesis route for coupling of succinimide with diazonium

2.2.2 ^{15}N Synthetic Methodology

Method 1 for ^{15}N Coupling Reaction

Having been reported in the literature[95,134], a viable diazotisation reaction was proposed to ensure formation of diazonium chloride. These papers reported *in situ* formation of the diazonium in a continuous flow reactor (figure 2.31). In this method, stock solutions of trimethylsilyl chloride (TMS chloride), isobutyl nitrite and *tert*-butyl-4-aminophenylcarbamate were prepared and stored safely, alongside a 2:1 mixture of DCE and MeCN solvent. A flow rate of 0.25 mL/min was maintained and pre-mixed, reacting with the aniline derivative molecule at 20°C. An initial 0.3 M conc.of *tert*-butyl-4-aminophenylcarbamate was added in the flow reactor which yielded a greater than 98% conversion to the diazonium chloride[95].

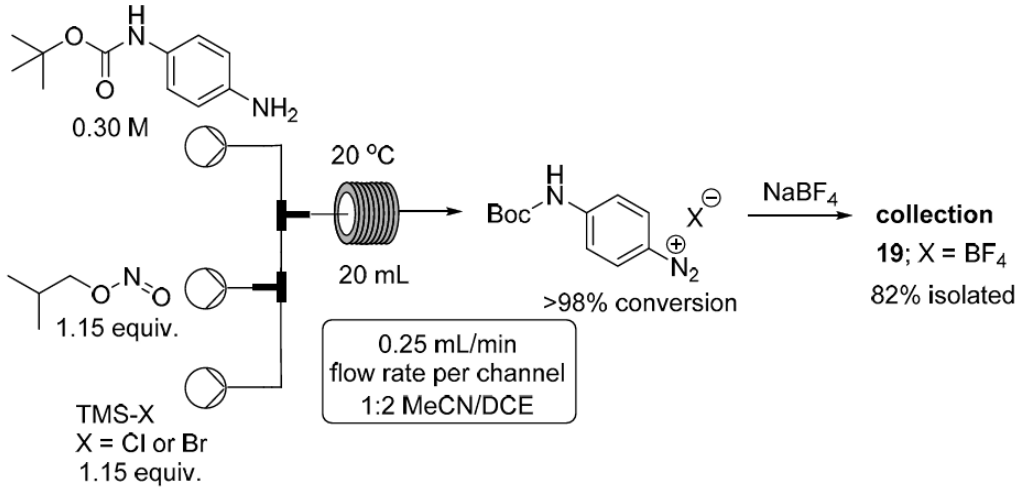


Figure 2.31: Literature flow synthetic route for diazotisation reaction[95]

It was therefore reasonable to adapt this method in the formation of the diazonium chloride. *Tert*-butyl nitrite (TBN) was chosen as a more readily available reagent than isobutyl nitrite in the diazotisation of aniline. Having accounted for the flow rate in the literature method, stock aniline was reacted with anhydrous MeCN for

precisely 6.4 min with dropwise addition of TMS chloride. This ensured cation formation on the amine. Again to ensure consistent flow rate, stoichiometric TBN was added dropwise to form the nitrosonium ion (diazotisation) with the counter chloride ion (figure 2.32). As this reaction was not flow-rate controlled, and to ensure safety, the vessel was cooled down to 0°C and maintained during the reaction (9.6 min) to form the diazonium chloride.

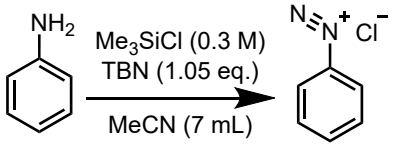


Figure 2.32: Proposed synthesis of diazonium chloride

Once the diazonium had formed, additional literature was sought to confirm reaction of this intermediate to form the ^{15}N labelled aniline. Although much literature was available for the so called Sandmeyer reaction[92], forming an aryl halide from an aryl diazonium salt, very little was present for the formation of an aryl amine. Indeed it showed a return to the starting material.

As no convincing literature could be found to support the reaction of the diazonium chloride with a suitable nucleophile, various schemes were devised. To incorporate an ^{15}N isotope into the molecule, ^{15}N ammonium chloride was proposed as a good nucleophile to add into the ring and form the labelled aniline product. As this labelled reagent is relatively cheap compared with other ^{15}N reagents, it was chosen as a suitable nucleophilic agent. The reaction would require a non-nucleophilic base to deprotonate the ^{15}N ammonium chloride and form ^{15}N ammonia, which could then undergo direction addition into the ring (figure 2.33).

The first attempt, with water (solvent), failed to facilitate successful nucleophilic behaviour from the ammonium chloride. A competing reaction of the water to form the phenol may have been preferred. The second attempt, with diisopropylamine (base), also failed to form the ^{15}N aniline product. Although this base was sterically hindered and ^{15}N ammonium chloride had good solubility, upon NMR analysis, evidence showed that the high nucleophilicity of this base caused it to be substituted *in lieu* of the amine. The third attempt sought to use K_2CO_3 (base) to deprotonate

the ^{15}N ammonium chloride. As this reaction was only allowed to stir for 3-4 h, a potential cause of the reaction not working may have been heterogeneity in the reaction. The fourth attempt used *tert*-butoxide as the base. This was chosen as a more sterically hindered base would reduce its nucleophilicity. However, similar issues arose from this reaction. The final attempt was with 2,2,6,6-tetramethylpiperidine (TMP). Whilst this base was the most sterically hindered, and effective to deprotonate ^{15}N ammonium chloride, the reaction failed. Only catalytic quantities (0.2 eq.) of TMP were used in this reaction, thus limiting its nucleophilic potential. Stoichiometric quantities may have facilitated more successful nucleophilicity. After all these reactions had been attempted (see table 2.1), a different method was investigated.

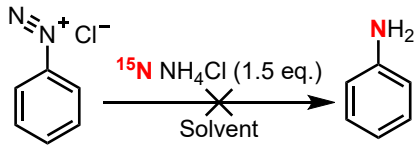


Figure 2.33: Synthesis of ^{15}N aniline using ammonium chloride as the nitrogen source

Attempt No.	Nucleophilic Agent	Base	Solvent	Yield (%)
1	$^{15}NH_4Cl$ (1.5 eq.)	H_2O (2 mL)	H_2O	0
2	$^{15}NH_4Cl$ (1.5 eq.)	DIPA (1.1 eq.)	THF	0
3	$^{15}NH_4Cl$ (1.5 eq.)	K_2CO_3 (2 eq.)	THF	0
4	$^{15}NH_4Cl$ (1.5 eq.)	KOtBu (1 eq.)	THF	0
5	$^{15}NH_4Cl$ (1.5 eq.)	TMP (0.2 eq.)	THF	0

Table 2.1: Conditions examined for reactions using ammonium chloride

Method 2 for ^{15}N Coupling Reaction

Sodium azide was proposed as a potential nucleophilic reagent in the substitution of the diazonium chloride. This phenyl azide intermediate could then be reduced to the amine. The advantage of this literature method is that the diazonium chloride is once more formed *in situ* with water (solvent) and the sodium azide is highly soluble when added to this solution. In this literature method[208], 3,5-dimethoxyaniline underwent a diazotisation reaction with sodium nitrite in water. This reaction proceeded at 0°C with 37% HCl solution and formed the diazonium chloride within 10 min. Whilst this reaction was stirring, sodium azide was added in small portions and left to stir

for 1 h. An aqueous work-up and removal of solvent yielded a pure phenyl azide (figure 2.34).

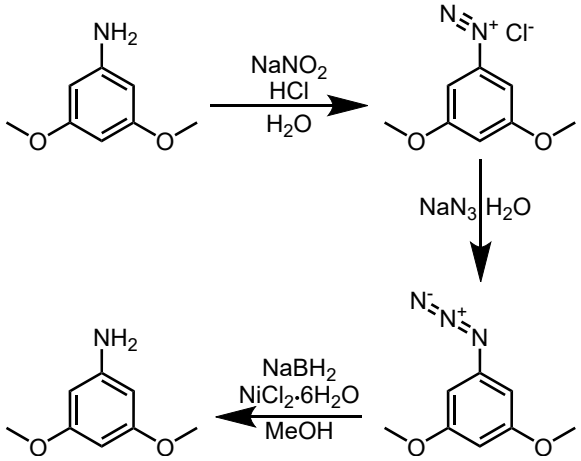


Figure 2.34: Literature route, via azide intermediate, for synthesis of aniline derivative

The second step in this literature method[165], namely the reduction of the azide to the amine, utilised sodium borohydride with catalytic quantities of hydrated nickel(II) chloride. This reaction was simple and successfully reduced the azide within 1 h.

This method was applied to aniline to ensure that the same conditions could result in formation of the labelled aniline. The diazonium chloride was formed as an intermediate in the synthesis, with 2 M HCl and 1.5 eq. of sodium nitrite, both having high solubility in water. This resulted in the diazonium intermediate being formed *in situ*, and whilst it was formed, stoichiometric sodium azide, in water, was added and left to stir for 1 h. The second step successfully formed the labelled amine with mild reducing conditions of sodium borohydride (2.5 eq.) and a nickel complex in catalytic quantities (0.5 mol%). Anhydrous MeOH solvent was used to ensure minimal impurities and side reactions. The drawback was that this labelled product was formed in low yield (3%). Whilst this synthesis is effective, and can thus be applied for incorporation of the ^{15}N isotope (see figure 2.35), ^{15}N sodium azide is very expensive to purchase, and thus alternative ^{15}N reagents were sought. Furthermore, purchasing ^{15}N sodium azide often comes as the isotope attached at either the terminal (3) position or the required (1) position, and thus not always possible to ensure the isotopic enrichment is controlled on the correct nitrogen post reduction of the azide to the amine.

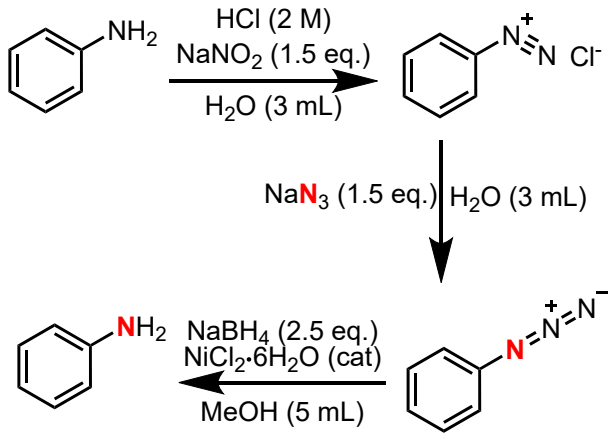


Figure 2.35: Synthetic route, via ^{15}N azido intermediate, to form ^{15}N aniline

Method 3 for ^{15}N Coupling Reaction

As the difficulty of coupling reactions grew, a different approach was needed. ^{15}N phthalimide was then proposed as an alterative reagent for ^{15}N incorporation. ^{15}N phthalimide was found to be the most economically feasible reagent for substitution of the diazo intermediate. To ensure the copper-catalysed cross coupling of the diazonium with the ^{15}N phthalimide, this method was first tested with phenylboronic acid. As has been mentioned previously, aryl boronic acids can couple with phthalimides in a copper catalysed Ullman style reaction to furnish a substituted phthalimide intermediate[166]. A literature method was found which coupled phthalimide with phenylboronic acid catalysed with copper(I) chloride and heated at reflux for 2-6 h under aerated conditions (figure 2.36). After purification with 2:1 PE:EA, the product was reported with a 65% yield and high purity[125].

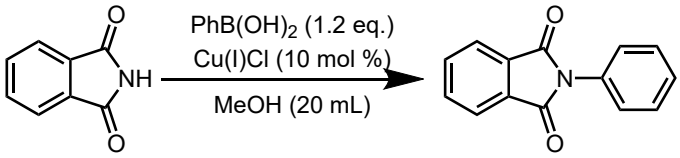


Figure 2.36: Literature synthesis for phthalimide coupling with boronic acid[125]

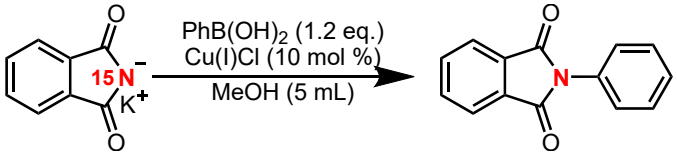


Figure 2.37: ^{15}N phthalimide coupling reaction with phenylboronic acid

Once ^{1}H NMR analysis had confirmed successful coupling of non-labelled phthalimide with boronic acid to form *N*-phenyl phthalimide (figure 2.38), these reaction

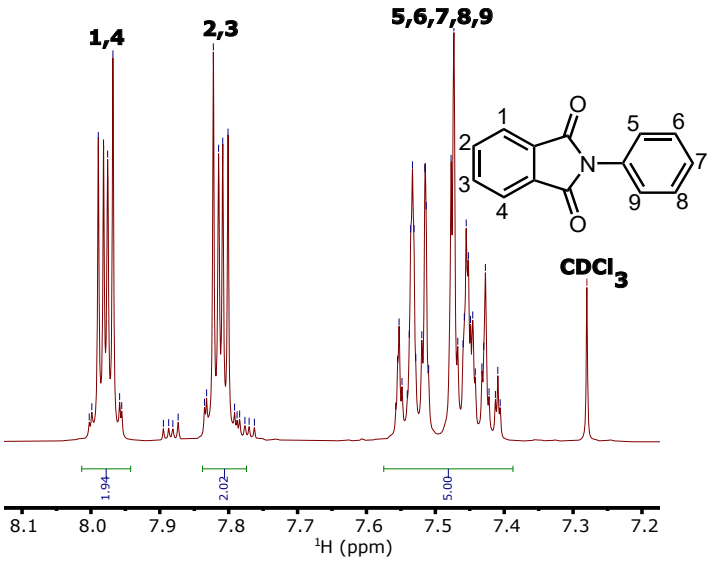


Figure 2.38: ^{1}H NMR ($CDCl_3$) of phenyl phthalimide

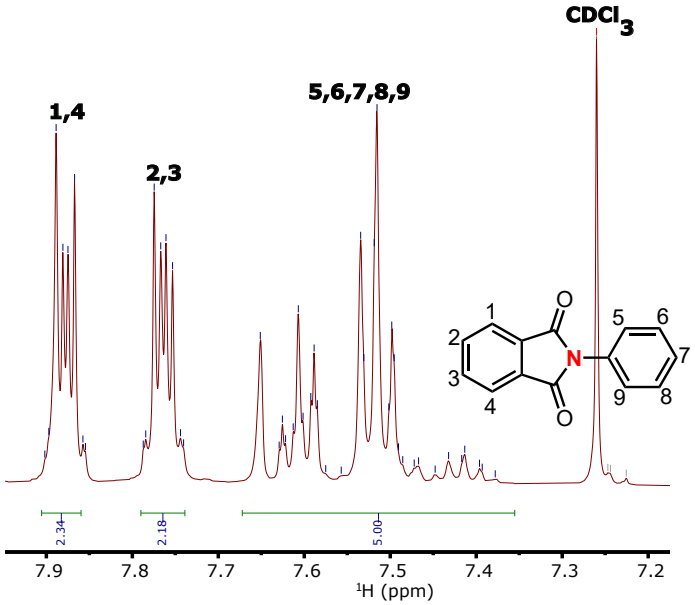


Figure 2.39: ^{1}H NMR ($CDCl_3$) of ^{15}N phenyl phthalimide

conditions were then applied to the labelled analogue, being the coupling of ^{15}N potassium phthalimide with boronic acid (figure 2.37). Catalytic quantities of copper(I) chloride were used in the reaction and heated at reflux. Unlike the literature method, this reaction was left to stir overnight with compressed air bubbling through the reaction vessel. After having filtered off the catalyst, and having done an aqueous work-up, 2:1 PE:EA were found as suitable TLC conditions. However, 4:1 PE:EA column conditions were found to provide better separation during purification and were thus employed. ^{1}H NMR analysis showed nearly identical splitting patterns compared with the non-labelled spectrum, confirming successful formation of ^{15}N phenyl phthalimide (figure 2.39).

Method 4 for ^{15}N Coupling Reaction

Once it had been established that this coupling was successful for both ^{15}N labelled and non labelled phthalimide with phenylboronic acid, this methodology was then applied directly to the diazonium intermediate. There was an abundance of literature precedence to show the facile synthesis of a diazonium tetrafluoroborate from an aniline molecule, and thus the tetrafluoroborate salt was chosen as the diazo intermediate[234]. Also, diazonium tetrafluoroborate is more stable at RT than the diazonium chloride, so was more favourable to synthesise from a safety perspective[65]. A recent literature method was found for the coupling of a diazonium tetrafluoroborate with phthalimide to form the coupled product[151] (figure 2.4). This method reported the reaction of benzene diazonium tetrafluoroborate with phthalimide in the prescence of catalytic Cu(I)Cl. Having screened various reaction conditions, they reported the use of K_2CO_3 (base) and DMSO (solvent) with aerated conditions stirring for 12 h. K_2CO_3 acted as a base to deprotonate the phthalimide, so this reagent could as a nucleophile in the reaction. All reagents were found to have good solubility in DMSO solvent. Once complete, the mixture was filtered off, washed with water and extracted with ethyl acetate in an aqueous work-up. This molecule was purified with 2:1 PE:EA column conditions and formed a white solid with a reported yield of 93%[151].

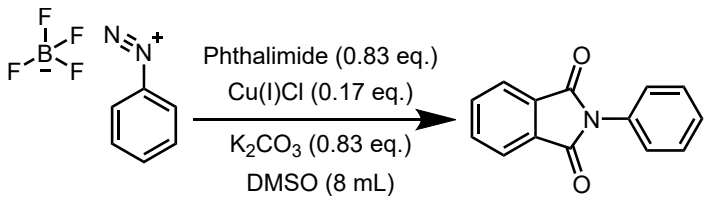


Figure 2.40: Literature synthesis for the diazonium coupling reaction [151]

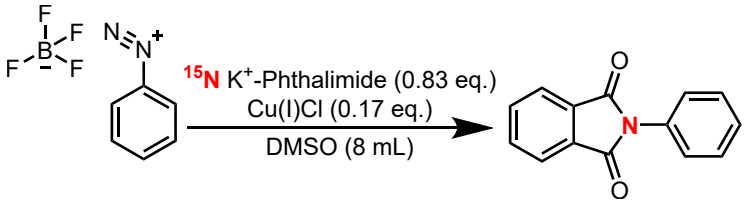


Figure 2.41: ^{15}N diazonium coupling reaction

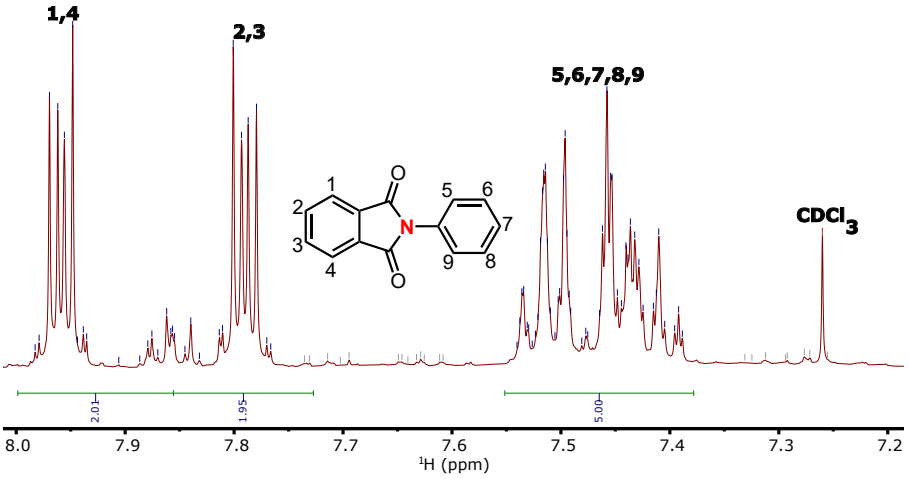


Figure 2.42: ^{1}H NMR ($CDCl_3$) of ^{15}N phenyl phthalimide

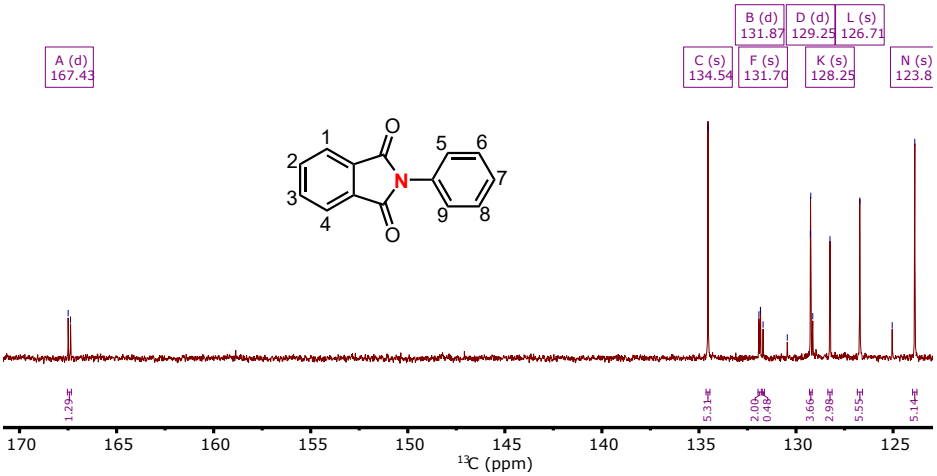


Figure 2.43: ^{13}C NMR ($CDCl_3$) of ^{15}N phenyl phthalimide

After having confirmed success, this reaction was repeated with ^{15}N potassium phthalimide. To attempt to improve yield and purity, modifications were made on this method (figure 2.41). Firstly, ^{15}N potassium phthalimide could act directly as a nucleophile in the reaction without the need for deprotonation by K_2CO_3 (base). Secondly, owing to method vagueness from the paper, the reaction was bubbled with compressed air for 12 h to ensure maximal reactivity. Once the reaction was finished, the catalyst was filtered off to remove unwanted copper(I) chloride. Thirdly, at least five aqueous extractions were applied to completely remove any residual DMSO in the organic phase. Lastly, although 2:1 PE:EA were good TLC conditions, 4:1 PE:EA were even better column conditions to ensure good separation of unreacted phthalimide from the coupled product. ^{1}H NMR reveals an almost identical splitting pattern to the non-labelled phenyl phthalimide (figure 2.42). Furthermore in the ^{13}C NMR, additional splitting of the singlet (carbonyl) into a doublet (figure 2.43) shows the ^{15}N isotope within the molecule, and thus successful coupling conditions.

Electron donating or withdrawing groups on a aniline ring can affect the yield of the coupling reaction. As the *para* methyl group on 4-methylbenzenediazonium tetrafluoroborate is electron donating, it seemed prudent to test these conditions to ensure yield was maintained and reaction was successful. An important alteration to this reaction was the use of excess diazonium tetrafluoroborate (figure 2.44). A 3 M excess of diazonium was found to favour a higher yield than stoichiometric amounts. Thus, the molar quantity of phthalimide, base and copper catalyst was 0.33 eq. This reaction was again conducted at RT, overnight, with compressed air bubbling through the reaction vessel. 4:1 PE:EA column conditions were maintained throughout. ^{1}H NMR analysis showed a singlet at 2.4 ppm (figure 2.45). This confirmed successful formation of the methyl substituted *N*-phenyl phthalimide coupled product.

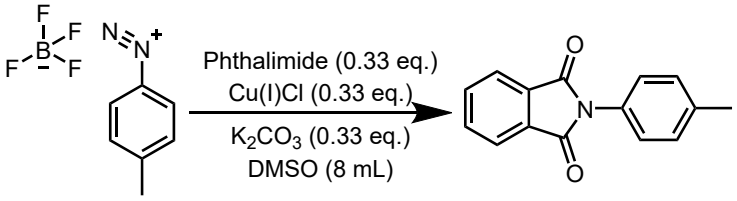


Figure 2.44: Synthesis of 4-methyl-*N*-phenyl phthalimide

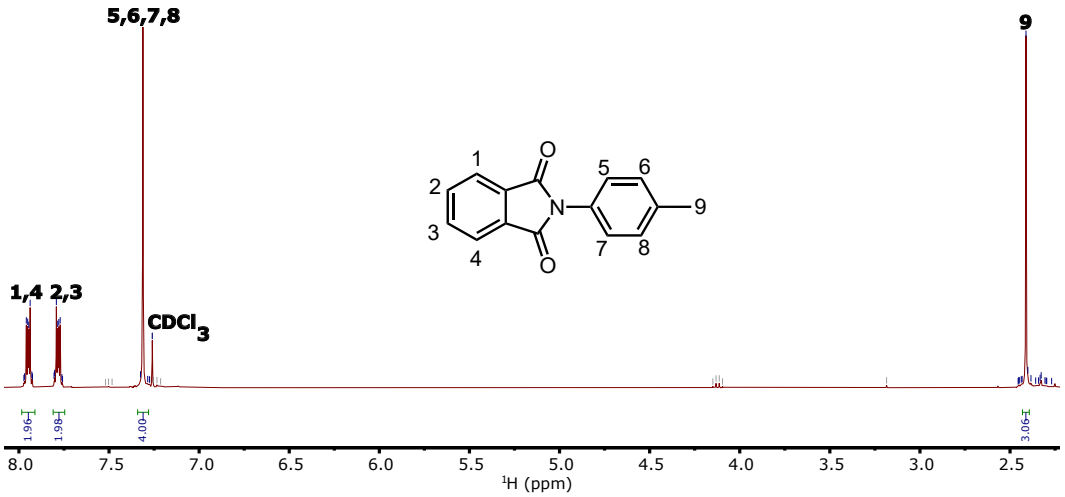


Figure 2.45: ^{1}H NMR ($CDCl_3$) of 4-methyl-*N*-phenyl phthalimide

Optimised Methodology

The optimised method was then applied to a wide substrate scope of substituted aniline molecules (figure 2.46). The optimal conditions for the first step (diazotisation) utilised 2 equivalents of both HBF_4 and TBN. Unlike previously, the reaction vessel was kept under ice (0°C) and TBN was added dropwise to ensure maximum yield. The mixture was then triturated with diethyl ether to ensure that the diazonium precipitates out of solution. The average yield of this first step was in the range of 75-95%, thus deemed suitable to continue to the second step.

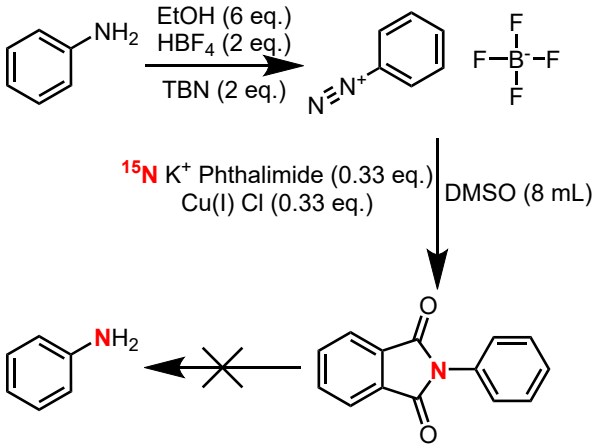


Figure 2.46: Synthetic route of ^{15}N aniline

This challenging second step was further optimised to ensure higher yield. It had now been established that a 3 M excess of diazonium tetrafluoroborate to ^{15}N potassium phthalimide and Cu(I)Cl would render a higher yield (figure 2.47). As seen in the literature, a slower rate of addition of diazonium to the reaction vessel produced a higher yield of coupled product[151]. Thus, instead of adding all reagents together in a flask, each diazonium tetrafluoroborate was dissolved in DMSO and added dropwise. Furthermore, the reaction was cooled down to 0°C. This was to improve the rate of substitution of the N_2^+ leaving group, and maximise the addition of the ^{15}N potassium phthalimide into the heterocyclic ring. The major difficulty however was the abundance of impurities in the product. An aqueous work-up was applied to ensure impurities were removed, and the aqueous phase was washed at least 5 times with H_2O to ensure residual DMSO had been fully removed from the organic phase.

Optimisation of column conditions proved to be extremely difficult. Purification with 4:1 PE:EA column conditions removed unreacted phthalimide alongside other impurities. Whilst good separation was evident on the TLC plate, more in-depth ^{1}H NMR analysis revealed retained impurities. Thus, a second purification with 8:1 PE:EA column conditions were applied to remove these remaining impurities. This ensured all ^{15}N coupled products were pure, albeit at the cost of reduced yield. As the average yield of all ^{15}N coupled products was between 8-19% (table 2.2), this was an insufficient quantity to use as a reagent in the final step. As the last step was a phthalimide deprotection to form the labelled starting material, the yield would have suffered to the point of being almost negligible and unable to characterise.

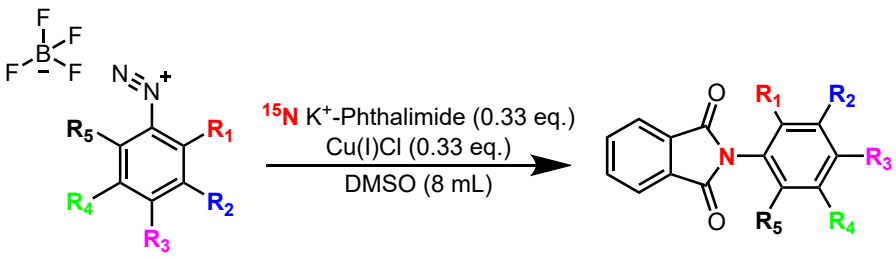


Figure 2.47: Synthetic route for ^{15}N coupling reaction

Compound No.	R group	Yield (%)
62	R_2 = Br	**9.9**
63	R_3 = Br	**9.9**
64	R_3 = Cl	**19**
65	R_3 = F	**10**
66	R_3 = I	**8.6**
67	R_2 = F	**15**
68	R_1 = R_3 = R_5 = Me	**9.8**
69	R_2 = R_4 = F	**10**
60	R_2 = F, R_3 = Me	**14**
61	R_2 = Cl, R_3 = F	**10**

Table 2.2: Yield entries for ^{15}N coupled molecules

2.3 Summary

In conclusion, this chapter firstly presented findings on efficient synthetic routes for production of ^{15}N labelled CAD molecules, and secondly presented results on the development of a novel ^{15}N labelling methodology. In the former case, propranolol, 2-aminomethylbenzimidazole and 4-amino-*N*-phenylbutanamide were all synthesised successfully, both labelled and non-labelled forms, and subsequently utilised in the succeeding chapter. In the latter case, formation of the diazonium, as supported by a high literature precedence, was a simple step and resulted in a high yield. The second synthetic step, namely the incorporation of the ^{15}N isotope, proved far more challenging. Subsequently, the yield in the second step was too low for phthalimide deprotection and thus not possible to continue.

Future work will need to be conducted into optimising reaction and purification conditions for formation of the ^{15}N phthalimide derivatives to ensure a high enough yield for the final synthetic step. Re-columning of each molecule is unviable in synthetic methodology development, and thus screening should be taken in varying quantities of reaction time, reaction temperature, catalyst and solvent. Once optimised, this methodology can then be applied directly to clinically important drug molecules.

Chapter 3

Solid State NMR of Drug-Lipid Interactions

3.1 Background

3.1.1 Application of ssNMR for Drug-Lipid Reactivity

Solid state NMR (ssNMR) is a technique which can determine the preferred orientation and alignment of membrane associated constituents, such as CAD molecules. Chemical information can be obtained for different atomic species of molecules embedded within lipid membranes[140]. This part of the thesis will utilise the synthesised ^{15}N molecules from Chapter 2 for drug-lipid reactivity measurements. Lipid membranes composed of POPC are used as model systems in which to probe the positioning of these molecules by ssNMR. Methods used include ^{1}H NOE, for probing close contacts, and ^{1}H-^{13}C and ^{1}H-^{15}N HMBC, for establishing connectivity and aiding in assignment of peaks. A comparison of the POPC liposomes, molecule and mixed liposomes with molecules have been compared to gauge the interaction and specifically target ^{1}H-^{1}H contacts for which kinetic data can be extracted. A combination of this data with the computational models will point to preferred orientation and distribution of these molecules within the POPC membrane.

3.1.2 NMR Theory

Solid State NMR spectroscopy is a useful analytical tool in the study of drug-lipid interactions. This spectroscopic technique can reveal atomic level detail on the interactions of drug molecules with model liposomes. The basic principle of nuclear magnetic resonance is the interaction of atomic nuclei, placed in a magnetic field, with electromagnetic radiation. A requirement of this oscillatory response is that the nuclei spin are non-zero and thus are subjected to the effects of the external magnetic field[172]. This process results in energy excitation of the nuclear spins, with the change in energy between lower and upper states given by:

$$\Delta E = \gamma\hbar(1-\sigma)B_0 \tag{3.1}$$

ΔE is the energy change between the lower and upper states, σ here corresponds to chemical shielding, the $\hbar$ to the reduced Planck constant, B_0 to the strength of the magnetic field, and γ to the gyromagnetic ratio, essentially the ratio of the nuclear magnetic moment to the nuclear angular momentum[120].

The transitions between different nuclear states occurs within the radio-frequency range of the EM spectrum (3 kHz - 300 GHz), with these frequencies corresponding to the ^{1}H Larmor frequency range (213-1200 MHz) for magnetic field strengths ranging from 1.5-3 T. The electrons surrounding the nucleus shield the effect of an applied magnetic field on the nuclei, which these transition frequencies take into account. The chemical shielding constant, σ, can vary substantially between different chemical nuclei, and thus the resonant frequency between nuclei varies. This difference in NMR transition frequencies provide information on the number of equivalent types of nuclei present within a chemical structure. Within drug-liposome interactions, these NMR frequencies are translated into chemical shift values, these being a marginal difference in the frequency of a pre-set standard compound, such as trimethylsilane, and the chosen nucleus of interest[61].

The oscillating magnetic field induces an electric field, or flow of electrons, which

then causes further small magnetic fields to be generated. The difference in magnetic field strength between the external and applied fields results in an effective magnetic field change felt by the nucleus, and thus also a difference in nuclear frequency. This frequency change, which also results in a change in orientation, gives rise to anisotropy[115].

In an NMR experiment at thermal equilibrium, a chemical species such as protons are perturbed by a radio-frequency pulse pre-set to the Larmor frequency. Excitation and relaxation of the protons to thermal equilibrium takes place if the pulse is applied or removed. Upon relaxation, these protons emit an NMR signal which can be detected by the spectrometer[16]. As this signal is emitted across 3D coordinates, the energy of the NMR signal is given by:

$$E_{x,y}(t) = E_0 \sum_i f_2 e^{-\frac{t}{T_{2i}}} \tag{3.2}$$

$$E_z(t) = E_0(1 - \sum_i f_1 e^{-\frac{t}{T_{1i}}}) \tag{3.3}$$

In equations 3.2 and 3.3, the transverse and longitudinal parts of the NMR signal are given by $E_{x,y}(t)$ and $E_z(t)$ respectively. E_0 represents the initial signal magnitude and scales with increasing number of protons. T_1 and T_2 importantly reflect the relaxation times in the longitudinal and transverse directions of the signal. f_1 and f_2 is the proportion of the magnetic field relaxing in the longitudinal and transverse directions respectively.

The simplest form of an NMR experiment is when a single pulse is applied, oscillating at the Larmor frequency, for a short time interval. When considered on a large scale, this applied pulse results in the magnetisation vector 'tipping' away from the longitudinal part of the NMR signal. This angle is known as the tip angle. This angle varies based on pulse intensity and time duration of the pulse[26]. If the tip angle is set at 90°, this experiment is called a free induction decay (FID).

As can be seen in figure 3.1, the grey (dead time) reflects the non-zero time between

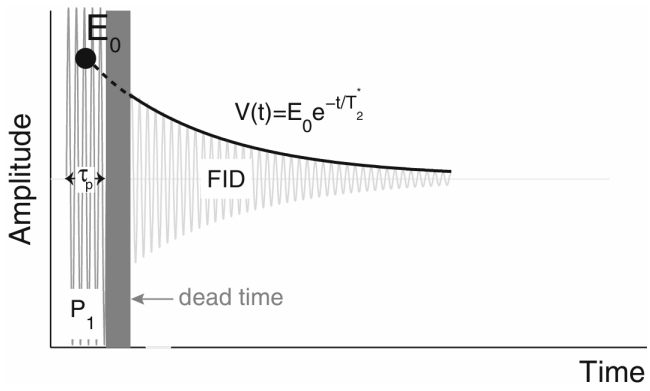


Figure 3.1: Pulse sequence of an FID[26]

the termination of the pulse and measurement. As the FID is incredibly sensitive to magnetic field inhomogeneities, the presence of these inhomogeneities greatly enhances the relaxation time. This altered relaxation time in the transverse direction is denoted by T_2^*. If these inhomogeneities are removed, T_2^* is effectively equal to T_2[91]. However due to the dynamic nature of the Earth's magnetic field, rarely does the static magnetic field remain homogeneous. T_2^* is defined by three relaxation processes:

$$\frac{1}{T_2^*} = \frac{1}{T_{2B}} + \frac{1}{T_{2S}} + \frac{1}{T_{2IH}} \tag{3.4}$$

In equation 3.4, T_{2B}, T_{2S} and T_{2IH} reflect the bulk fluid, surface and inhomogeneous magnetic field transverse relaxation times respectively. As T_1 can only be calculated in the transverse direction, it relies upon two pulse sequences, namely the inversion and saturation recovery pulse sequence. The former pulse causes rotation of the magnetisation vector whilst the latter aligns this vector in the x, y rather than the z plane. This process begins with recording the FID pulse, applying a second 90° pulse (saturation pulse sequence), the exponential recovery, or T_1 can be measured[218]. The benefit of this measurement is that T_1 is unaffected by magnetic field inhomogeneities:

$$\frac{1}{T_1} = \frac{1}{T_{1B}} + \frac{1}{T_{1S}} \tag{3.5}$$

where T_{1B} and T_{1S} are the bulk and surface longitudinal relaxation times respectively. Once the FID has been acquired, Fourier transforms are utilised to convert the signal into a spectrum which reflects the chemical shift of different atomic nuclei[80]. NMR spectra can help to determine chemical structure.

The interactions of nuclear spins with other nuclei are broadly known as internal and external interactions. Internal interactions directly quantify the local chemical environment of atomic nuclei, and thus greatly affect the resulting NMR spectra acquired from FID. Of these internal interactions, chemical shielding represents the interaction between nuclei and the electrical currents, which are generated by the external magnetic field (figure 3.2). This in turn decreases the effective magnetic field experienced by the nucleus. A change in this magnetic field induces a change in the molecular orientation, which gives rise to chemical shift anisotropy (CSA)[145]. Anisotropy increases with nuclear complexity. The extent of CSA is determined by the electron density distribution around bonding and non-bonding orbitals. Often, this electron distribution along bonds is elongated to form a geometrical ellipsoid, and the dependence of this on the resonant frequency is entirely reliant on the orientation of this electron distribution relative to the direction of the external magnetic field[8].

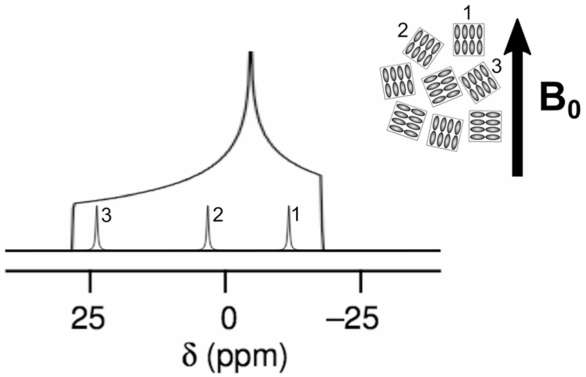


Figure 3.2: 1D spectrum of a powder sample due to chemical shielding interactions, each molecule, 1,2 or 3, is present in the spectrum with a sharp, narrow peak, which resonates at a frequency dependent upon orientation with respect to the direction of the magnetic field, B_o[145]

This orientation change can be seen in figure 3.3. Here, the most de-shielding effect occurs when the narrowest part of the carbonyl group aligns itself in the direction of the magnetic field, thus resulting in a very large chemical shift. In contrast, the most

shielding effect occurs when the widest part of the carbonyl group aligns itself in the direction of the magnetic field, thus resulting in a very small chemical shift[115]. These two configurations, alongside another perpendicular one, are referred collectively as chemical shift tensors, being represented by the principal components: δ_{11}, δ_{22}, δ_{33}. In conventional solid-state NMR, the $\delta_{11} >= \delta_{22} >= \delta_{33}$. Thus, the shielding factor (δ_{iso}) is calculated from these respective chemical shift tensors.

$$\delta_{\text{iso}} = \frac{1}{3}(\delta_{11} + \delta_{22} + \delta_{33}) \tag{3.6}$$

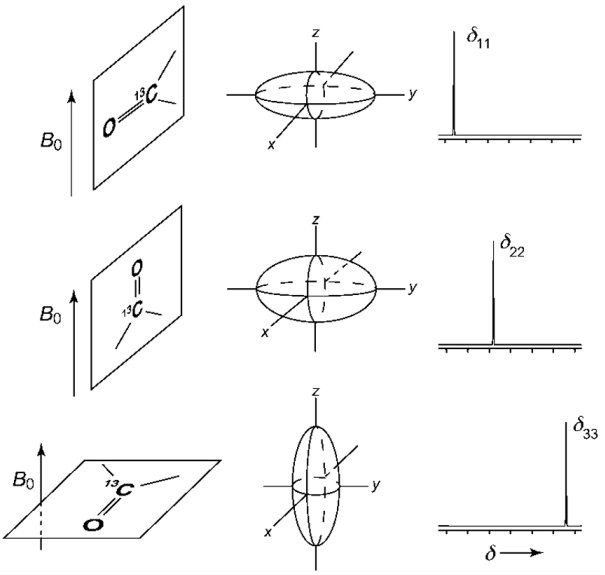


Figure 3.3: Diagram of chemical shift tensors[115]

Molecular tumbling is defined as the random, rotational motion of molecules within the NMR spectrometer. This tumbling is much faster in solid-state than in solution-state NMR, resulting in much sharper and narrower spectral bands that have a Gaussian profile. This tumbling motion enables the CSA to be averaged out by the so called magic angle spinning (MAS). Whilst an advantage of this method over solution-state NMR is that it greatly improves the signal-to-noise ratio, as the signal intensities are integrated over a far more narrow spectral range[194], the main distinction is that without MAS, the signals are too broad to be of any practical use. Thus within my project, this is the critical biophysical technique used to capture interactions between the drug molecule and liposomes.

As liposomes are more solid than liquid in state, they are often used in solid-state NMR. Whilst anisotropic components of the internal interactions are averaged out by the much faster tumbling of molecules in solution state NMR, the same is not true of solid-state NMR. The spectral lines are far broader in solid-state NMR, as the anisotropic components are observed by much slower rates of molecular tumbling[142]. To circumvent this line broadening, the sample is rotated at an angle of 54.7°, relative to the magnetic field, termed the magic angle. By spinning the sample at this precise angle, the dipolar interactions between different atomic nuclei are effectively reduced to zero, and thus only anisotropic interactions need be considered. This value is derived from half of the tetragonal angle (109.5°) taken from the angle calculated between a body diagonal in a cube and each of its faces[89].

Alongside chemical shielding and magic angle spinning, there are other important internal interactions worth noting. Dipolar interactions, in ssNMR spectroscopy, or simply dipole-dipole interactions, is the phenomenon of direct coupling of two atomic nuclei as a function of their intrinsic magnetic moment[230]. The interaction can take place independent of the spin integers of the individual nuclei, and is only reliant on the so called dipolar coupling tensor, itself being a function of the internuclear distance between the two nuclei (figure 3.4). This coupling takes into account all pairwise interactions (both homo and heteronuclear). This interaction is dominated by the nuclear magnetic moment; itself being derived from the angle θ (between rotational axis and magnetic field) and the gyromagnetic ratio. Thus, this dipolar coupling is important for nuclei with large magnetic moments and half-integer spin. The nuclei isotopes of ^{1}H and ^{15}N, utilised in this project, have predominantly dipolar coupling interactions. Whilst there is a high abundance of ^{1}H-^{13}C dipolar interactions, due to the relatively high abundance of ^{1}H (99.98%) nuclei, there is a very low amount of ^{13}C-^{13}C dipolar coupling because of the small abundance (1.1%) of ^{13}C nuclei[29].

For heteronuclear dipolar coupling, such as ^{1}H-^{15}N or ^{1}H-^{13}C, the total interaction energy is given by:

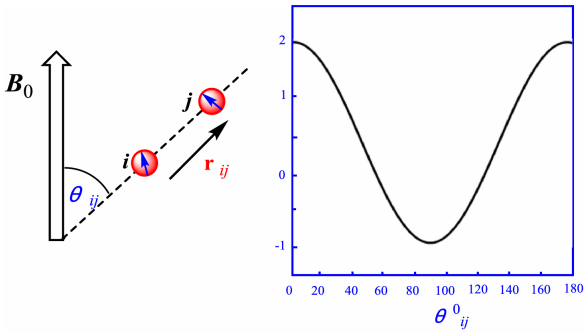


Figure 3.4: Diagram of two nuclear spins, i, j, joined by internuclear vector, r_{ij} in the presence of a magnetic field[230]

$$E_{\text{dipolar}} = -d(3\cos^2\theta - 1)I_z S_z \tag{3.7}$$

$$d = \left(\frac{\mu_0}{4\pi}\right)\frac{\hbar\gamma_I\gamma_S}{r_{IS}^3} \tag{3.8}$$

In these equations, I represents the nuclear spin of abundant nuclei (^{1}H) and S reflects the nuclear spin of the rare nuclei (^{13}C or ^{15}N). Thus, r_{IS} is the internuclear distance between the two nuclei, γ_I and γ_S the gyromagnetic ratios of the respective nuclei, and μ_0 the permeability of free space. Within the overall Hamiltonian expression, both I_z and S_z reflect the magnetic field induced by the spins I and S respectively, and thus the net effect of spin S experiencing the magnetic field from spin I. The dipolar coupling energy is thus entirely dependent upon the dipolar coupling constant, d.

3.1.3 Review of ssNMR Literature for the study of drug-lipid interactions

Previous literature has shown the extensive use of solid-state NMR in studying the interactions of various pharmaceuticals with phospholipid bilayers. Since the 1990s, a variety of ^{1}H NMR techniques have been developed to measure intermolecular distances between drug molecules and membrane lipids, thus confirming distribution and localisation within the membrane[68]. MAS is typically used with the caveat that the MAS frequency must be greater than the dipolar coupling interaction. Evidence has shown that much faster reorientation of lipid molecules in the membrane allows for far lower MAS frequency to achieve the same good resolution in the ^{1}H NMR

spectrum[226]. The overarching advantage of this ^{1}H NMR MAS technique is that the high resolution enables both 1D and 2D spectra to be acquired for understanding of the intermolecular interactions.

The Gawrisch group were the first pioneers into the use of ^{1}H MAS NOESY NMR to understand the lipid organisation in both pure and mixed DMPC membrane lipids, both with and without cholesterol. The high degree of disorder and lipid mobility in this membrane resulted in multiple cross peak signals between different lipid environments in the DMPC bilayer. Having analysed the contour lines of the 2D ^{1}H-^{1}H NOESY spectrum, they found larger cross-relaxation rates for lipid segments in close proximity (such as the multiple CH_2 groups on the lipid tails), and the reverse for lipid segments far away (such as the polar headgroup from the fatty acyl chain). They concluded that these relaxation rates must then be of entirely intermolecular origin within the membrane lipids. Having included small molecules in the DMPC membrane, they found the additional difficulty of distinguishing ^{1}H-^{1}H NMR peaks between the drug and lipid from the lipid-lipid contacts in a typical 2D ^{1}H NOESY spectrum. Although they did not report the concentration ratio of drug to lipid, for practical purposes in my project, a 1:10 ratio of drug to lipid is required. The advantage of aromatic drug molecules with membrane lipids, is that the chemical shift difference between the aromatic range and the phospholipid range is large enough to analyse quantitatively the proton-proton distances between the drug and lipid, whilst negating the closer ^{1}H-^{1}H lipid-lipid contacts. A difference of more than 6 ppm between the aromatic and phospholipid resonance range is large enough to characterise the respective interactions with the lipid[187].

3.2 NMR Methodology

3.2.1 1D and 2D NOE Measurements

The principal technique to measure drug-lipid interactions used in this work was to analyse proton-proton interactions with the Nuclear Overhauser Effect (nOe). This effect is a form of direct, through space, magnetic interactions between different atomic nuclei. Thus it is a form of dipolar coupling alluded to in the background

earlier. The mathematical definition of the nOe is the intensity change from a single resonance when the nuclear spin transitions of a different resonance are perturbed from their original state. This so called 'resonance saturation' or perturbation arises from inversion of the population differences across these nuclear spin transitions. This nOe intensity change, $\eta_I(S)$, is thus calculated as a function of the perturbed spin state S, such that:

$$\eta_I S = \frac{I - I_0}{I_0} \times 100\% \tag{3.9}$$

The spin S always refers to the perturbed spin whilst the spin I refers to the enhanced spin and I_0 to the equilibrium intensity. The magnitude of the nOe intensity changes are dependent upon the relative motional properties of the molecules alongside their respective nuclear gyromagnetic ratios. Steady-state nOes are the predominant type of nOes measured in this project. These nOes arise from the aforementioned perturbation by saturating the spin S transitions (figure 3.5). Practically, this is achieved by short, discrete bursts of RF irradiation to the S resonance[38].

Figure 3.5: Population intensity plot for two spins, S and I (dipolar coupling), at thermal equilibrium (left) and after instantaneous saturation of the S spins (right)[38]

To obtain high-resolution steady-state nOe spectra, there has to be an appropriate time for presaturation of a specified resonance region, achieved by short, discrete bursts of RF irradiation to the targeted S resonance. A basic requirement for these steady state measurements is that they are derived from simple 1D spectra. As such, nuclear spins within the nOe resonate with perturbed intensities relative to the standard 1D spectrum. The term 'nOe difference' takes its origin from subtracting the peaks in the nOe spectrum from the control spectrum. In principle this displays only the nOe enhancement peaks. This spectrum is acquired by application of the presaturation to a target resonance, such as the aromatic protons of a drug molecule, for a specified time period (figure 3.6). Once this time has elapsed, the presaturation

RF pulse is switched off and the 1D plotted with a 90° pulse. The control spectrum is acquired similarly without presaturation[18].

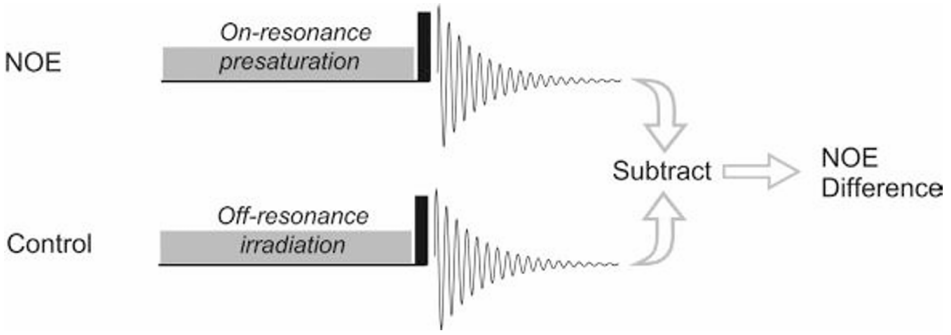


Figure 3.6: nOe difference measurement from subtraction of the nOe signal from the control signal[18]

Specifically for this project, transient nOes are required to measure internuclear distances between interacting atoms of the drug relative to the lipid. Here, homonuclear two-dimensional nOe spectroscopy (2D NOESY) and its analogous one-dimensional nOe spectroscopy (1D NOESY) are useful experiments. The greatest advantage this method has over nOe difference is its ability to be non-selective and thus measure all potential nOe signals in a molecule. Further, nOe difference measurements often take a longer time than more routinely available NOESY experiments. The key difference here is that transient, rather than steady-state, enhancements are captured. Crosspeaks in 2D NOESY experiments show all possible nOe interactions between correlated spins, thus are useful for capturing distance measurements between protons on the drug relative to the lipid. Direct distance measurements through space are captured by peaks in 2D NOESY which arise from incoherent magnetisation transfer between proton-proton nuclear spins as a function of mixing time within the nOe[63].

The significant parameter for nOe development is the mixing time, τ_m. Post excitation as a function of time t_1, the magnetisation vector reorients along the transverse axis. Another 90° pulse reorients this vector onto the inverse z axis, leading to population inversion as mentioned earlier. This in turn generates the transient nOe as a function of mixing time. This sequence (figure 3.7) is then iterated until an appropriate signal-to-noise ratio is captured with high resolution peaks[131].

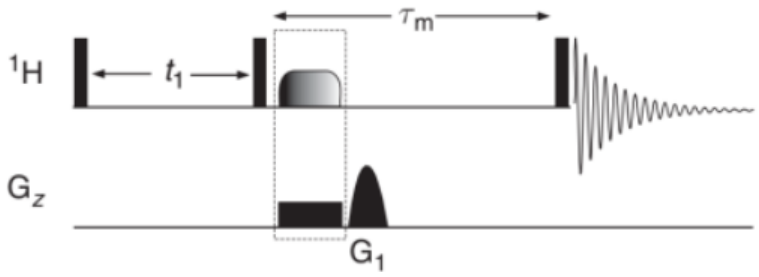


Figure 3.7: Diagram of the NOESY sequence, with the frequency gradient G_1, with transient nOe measured during mixing time τ_m [131]

Quantitatively, the distance measurements can be extracted directly from 2D NOESY spectra. Having known a reference distance from a set of two interacting nuclei and assuming isotropic molecular tumbling in the system, the ratio of the nOe intensity, A, relative to another intensity, X, can be calculated as:

$$\frac{\eta_A B}{\eta_X Y} = \frac{r_{AB}^{-6}}{r_{XY}^{-6}} \tag{3.10}$$

In this calculation, if the reference inter-nuclear distance between X and Y is known, then a direct ratio between two nOe intensities provides the unknown inter-nuclear distance (A and B). nOe intensity changes (η_{AorX}) are proportional to $1/r^6$.

There are two methods for this calculation. Firstly, the nOe comparison between the reference and unknown distance can be calculated from the volume intensity change of 2D NOESY cross-peaks as these spectra evolve over τ_m. This build up curve can then be plotted and mapped onto the relevant internuclear distances, showing nOe growth over a wide range of time intervals (figure 3.8). Secondly, the cross-peak intensities, having assumed they fall within the linear regime, are compared at a specified mixing time rather than over a range of times [228].

Lastly in this section, we turn to 1D NOESY experiments (figure 3.9). These measurements are based upon nOe difference to measure nOe enhancement between different nuclei and thus its disadvantage in experimentation is the possibility of difference artifacts. A derivative of this method, namely the 1D gradient-selected experiments, have found far more utility than 1D NOESY. The advantage of this

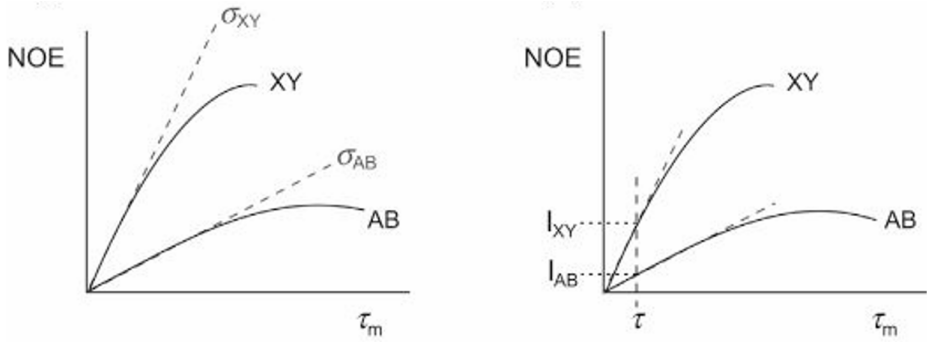


Figure 3.8: Build-up plots for nOe intensity change as a function of mixing time, XY and AB plateaus off which shows saturation to a constant inter-nuclear distance[228]

method is that it is independent of difference measurements and considers only nOe enhancement signals by automatically deleting non-nOe signals in the spectra. This ensures far more accurate spectra in a much shorter acquisition time. Similarly to 2D, a build-up curve can be obtained through these measurements to show nOe intensity changes as a function of mixing time[52].

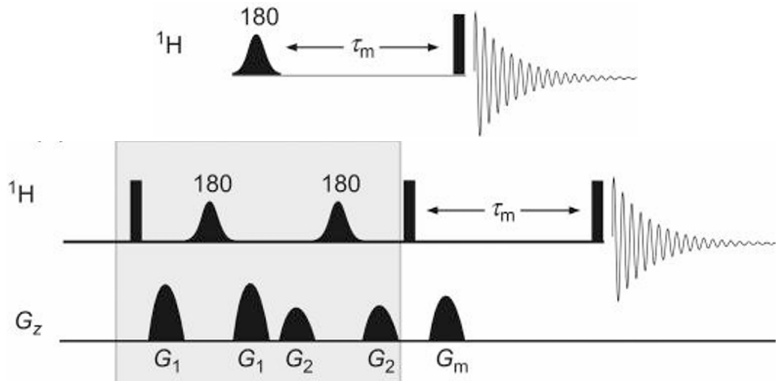


Figure 3.9: Diagram of the 1D gradient nOe sequence, a series of gradients ($G_{1/2/m}$) applied during the 180° pulses, pre mixing-time τ_m[52]

3.2.2 ^{1}H-^{13}C, ^{1}H-^{15}N HMBC Measurements

Heteronuclear shift correlations are also important measurements, alongside NOESY experiments, to determine the potential drug-lipid interactions. Proton-carbon and proton-nitrogen correlation measurements can be ascertained as indirect couplings. These high sensitivity measurements can provide chemical shifts of a whole host of different nuclei without any direct observation. Multiple bond or long range correlations are the primary technique used in these heteronuclear experiments to demonstrate connectivity between the drug molecule and membrane lipid[119]. Heteronuclear multiple-bond correlation (HMBC) spectroscopy is therefore of importance.

Predominantly, and in this project, proton-carbon and proton-nitrogen connectivity can be measured over at least two to three bonds, providing structural information to support reactivity between the drug and lipid molecules.

HMBC detects correlations by small coupling of atomic nuclei such as proton-carbon or proton-nitrogen. This sensitivity is improved particularly by proton detection. To achieve good resolution with HMBC, the long-range coupling constant (Δ_{LR}) has to be run sufficiently long to capture proton-carbon couplings; often being small and long-range through multiple bonds. This results in vector displacement which gives rise to hetero nuclear multiple-quantum coherence (HMQC). To circumvent the far smaller proton-carbon coupling in multiple bond correlations, when compared with single bond correlations, the Δ_{LR} has to be set to a minimum of 100 ms. Additional homonuclear proton-proton couplings of a magnitude comparable to that of multiple bond heteronuclear proton-carbon couplings arise as cross-peaks in the spectrum[130]. Thus, often even smaller Δ_{LR} times are chosen to negate these peaks from the data. To avoid phase errors from these additional couplings, often HMBC experiments are phase modulated with respect to the t_1 (figure 3.10).

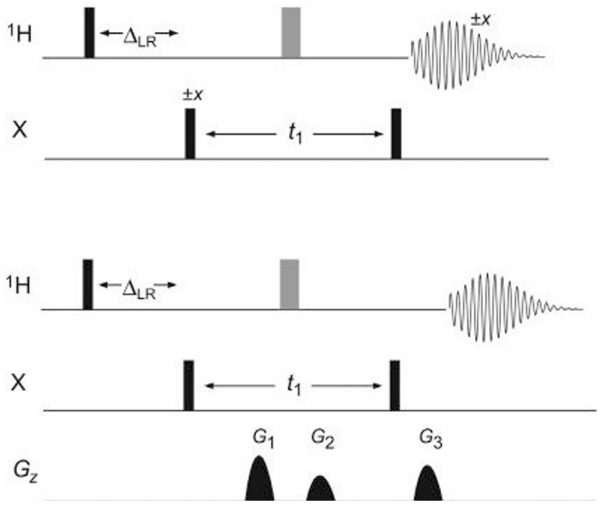


Figure 3.10: HMBC sequence, top sequence is without, whereas bottom sequence is with a pulsed gradient G_z, Δ_{LR} is optimised for long-range coupling constants in each sequence[119]

When interpreting HMBC spectra, the crosspeak intensities arise from both the long-range coupling effect and the chosen Δ_{LR} values. When the molecular system is saturated, proton-carbon couplings are often in the range of 10-25 Hz, whilst

when unsaturated these couplings rarely exceed 5 Hz. Thus, the Δ_{LR} value of 100 ms is appropriate given these coupling values. To account for relaxation losses, the Δ_{LR} is reduced to around 60 ms which is suitable for an average coupling constant value of about 8 Hz. Generally, smaller molecules adopt a Δ_{LR} value of up to 200 ms to circumvent the slower relaxation times, whereas the inverse is applicable for larger molecules, often around 40 ms (Δ_{LR}). Thus, there is a fine balance required when considering molecular system relative to relaxation time and thus choice of the Δ_{LR}. The only disadvantage is the difficulty to distinguish between two and three bond correlations in molecular systems. However, certain chemical groups such as a methyl display intense correlations as all three protons are detected synchronously [15]. The typical J coupling values for common chemical species are given in table 3.1.

Coupling Pathway	$^{2}J_{CH}$	Coupling Pathway	$^{3}J_{CH}$	Coupling Pathway	$^{4}J_{CH}$
H–C–**C**	(±) ≤ 5	**H**–C–C–**C**	≤5	**H**–C=C–C=**C**	(±) ≤ 1
H–C=**C**	≤10	**H**–C=C–**C**	≤15[a]	**H**–C–C–C–**C**[b]	≤1
H–C≡**C**	40–60	**H**–C≡C–**C**	≤5		
H–C(=O)–**C**	20–25				

Table 3.1: J coupling values for different chemical species in HMBC measurements [104]

3.3 Results and Discussion

3.3.1 ssNMR Systems: General Protocol

Once the ^{15}N molecules had been synthesised and fully characterised to ensure a good yield and high purity, these molecules were then combined with POPC liposomes in ssNMR molecule-lipid reactivity studies. The aims here were to measure proton-proton distances between the molecule and lipid in 1D and 2D nOe experiments, and establish connectivity through HMBC experiments.

A protocol was developed to ensure a meaningful physiological pH of the molecule, when combined with liposomes. A pH value of 7.2-7.4 was used to mimic physiological conditions and POPC was used as the lipid of choice due to its high abundance in many biological membranes. Thus, a buffer solution of 0.1 M citric acid, 0.2 M $NaHPO_4$ at pH 7.4 was chosen as a suitable pH controlled solution. This stock drug-buffer solution could then be used for all ssNMR experiments. Once the drug

concentration had been verified by UV-Vis absorption measurements, the POPC liposomes were then prepared. Firstly, a stock POPC-$CHCl_3$ solution was made up and stored at 0°C. After letting the solution mix for a few hours, the appropriate volume was extracted with a Hamilton syringe into a dry flask. The solvent was evaporated and the remaining POPC thin lipid film left to further dry in the desiccator for 1-3 days.

The protocol for ssNMR sample preparation was then conducted. The appropriate amount of drug-buffer solution was extracted with a Hamilton syringe and added to the POPC liposomes in the flask. Multilamellar dispersions of POPC liposomes were prepared by hydrating a thin film of lipids using buffer and repeated freeze thaw cycles. [compound]:[lipid] ratios were determined from the mass of lipid used and the volume of stock drug. Essentially for all drug-lipid systems, the first set of experiments were carried out on a low concentration of lipid to drug, whereas the second set of experiments were conducted on a higher lipid to drug concentration ratio. The liposomes were then subjected to a series of freeze-thaw cycles. The mixed sample was cooled in liquid N_2 for 30 seconds, left to equilibrate to room temperature for a few seconds and then submerged in a water bath preset at 37°C (physiological conditions). This sequence was repeated 5 times to ensure maximal lipid dispersion. Once complete, this sample was then extracted and placed in the rotor for ssNMR experiments.

3.3.2 ssNMR System I: ^{15}N 2-Aminomethylbenzimidazole with POPC Liposomes

The first set of experiments were 1D and 2D NOESY to probe close contacts between the molecule and lipid. Initially a high-field proton NMR spectrum was obtained, showing the 2 aromatic regions, at δ 7.25 ppm and 7.6 ppm for the 4 aromatic protons, and a single aliphatic region, at δ 3.95 ppm for the 2 aliphatic protons of the methyl group. Ubiquitously across all these ssNMR experiments, the 2 doublets from δ 2.4-2.6 ppm correspond to the proton chemical shifts of the citrate buffer.

All experiments were run in D_2O solvent as this ensured good solubility of the drug

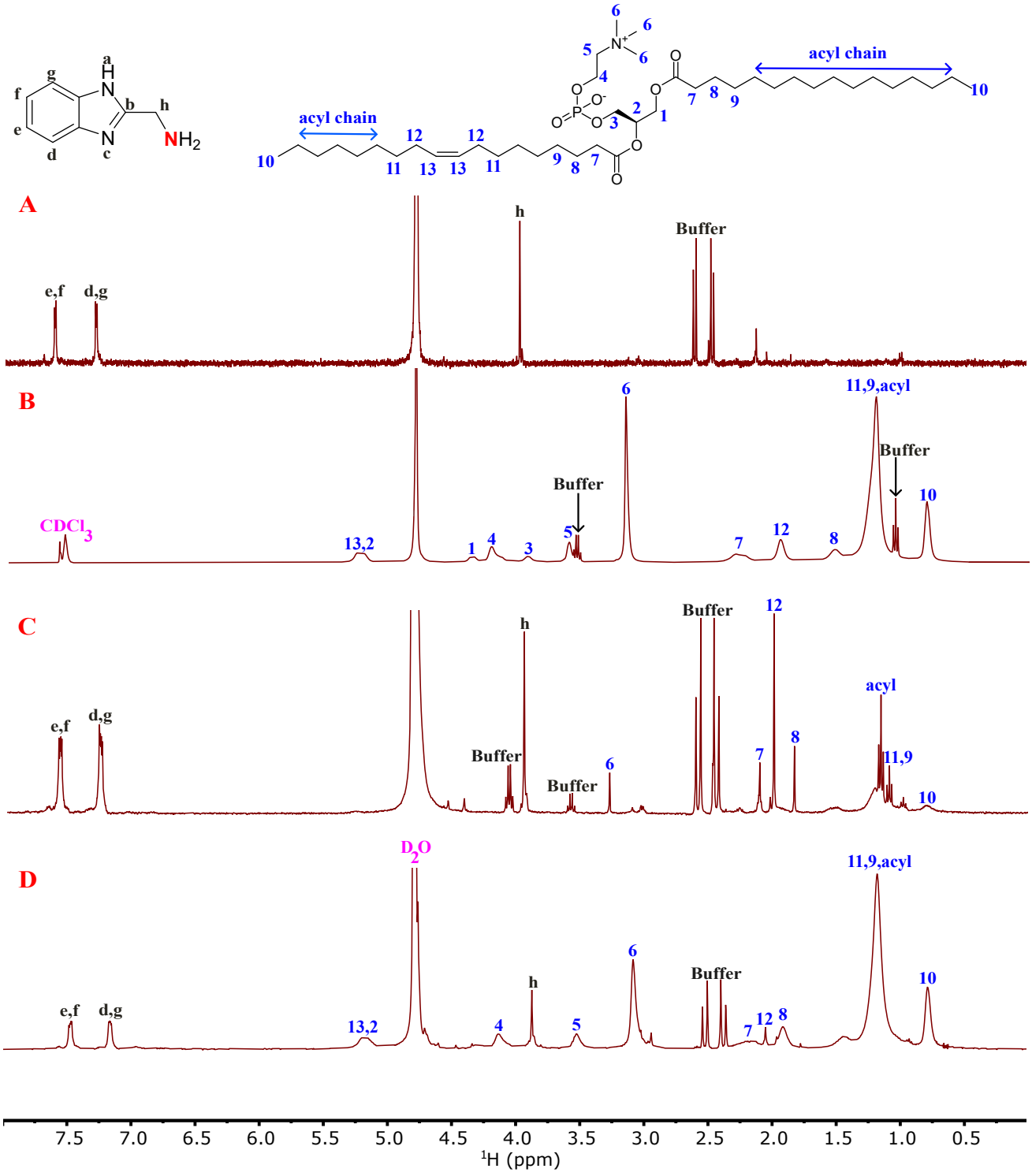


Figure 3.11: ^{1}H NMR spectra (run in buffer): (A) 2-aminomethylbenzimidazole, (B) POPC liposomes, (C) liposomes+compound at low concentration ratio [0.07:1], (D) liposomes+compound at higher concentration ratio [2.45:1]

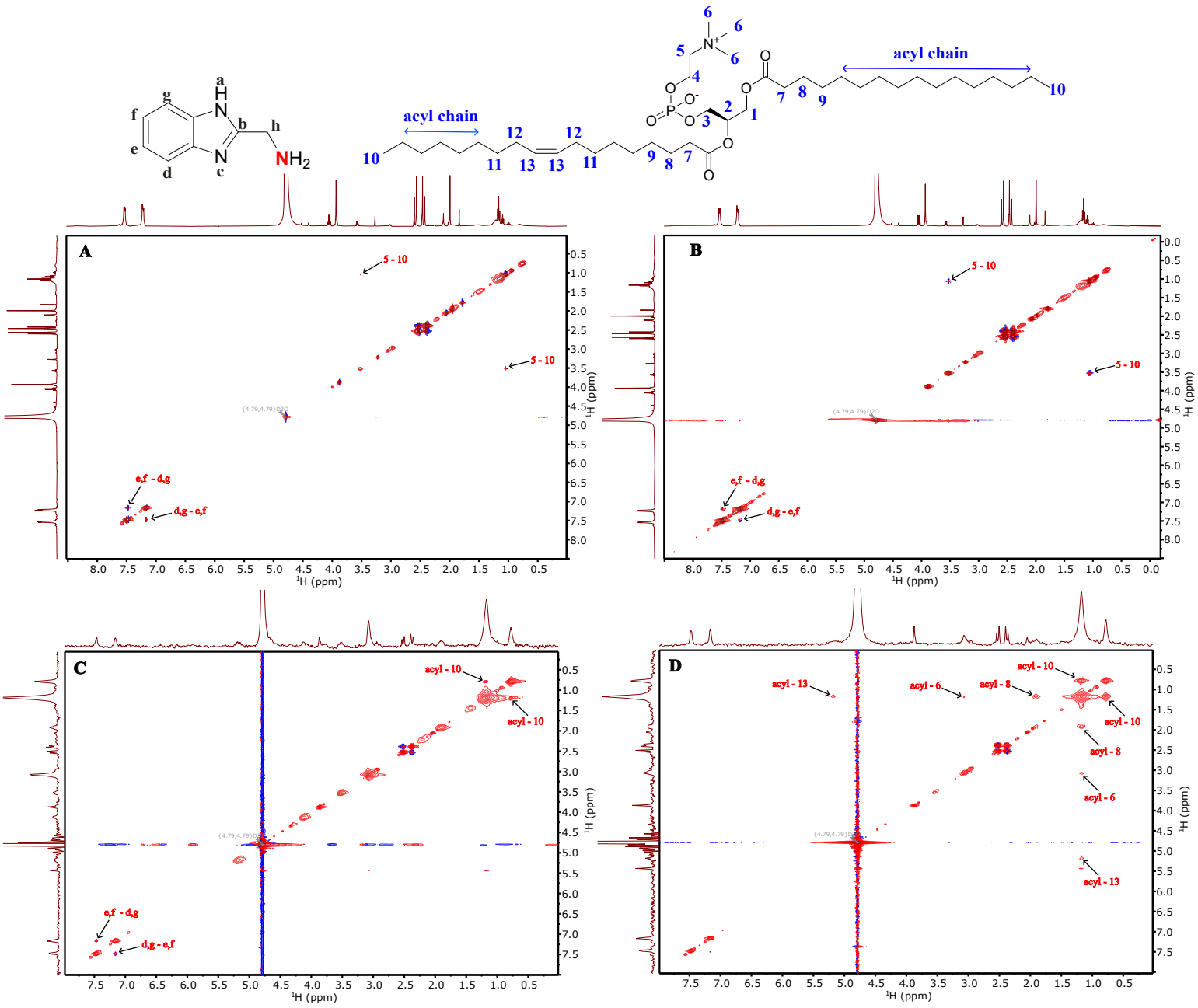


Figure 3.12: 2D NOESY spectra (run in buffer): (A) compound+liposomes [1:0.07, 3 ms], (B) compound+liposomes [1:0.07, 300 ms], (C) compound+liposomes [1:2.45, 3 ms], (D)compound+liposomes [1:2.45, 300 ms]

in buffer solution. Once the proton chemical shifts of the drug had been established, the proton chemical shifts of the POPC liposome were then ran. As was expected, the presence of 2 long fatty acyl chains caused line broadening in the NMR spectrum with an integration value of 20 fold relative to other proton chemical shift peaks in the lipid. The presence of the choline headgroup accounted for an integration value of 9 for the 3 methyl groups present in the lipid at a value of δ 3.15 ppm. Figure 3.11C shows the proton spectrum of a 0.07:1 concentration ratio of [lipid]:[molecule]. Evidently the lipid concentration was very low, exhibited by an integration value of 0.07 for the phosphocholine headgroup protons (6) relative to 1 for the drug aromatic protons (e,f). Importantly, there is a slight chemical shift change from δ 3.95 to 3.9 ppm for the -CH_2 (h). This difference is even greater in the final plot, showing a lipid-molecule concentration ratio of 2.45:1 (figure 3.11D). The chemical shift of the -CH_2 methyl group in the drug (h) is now δ 3.8 ppm. Having synthesised a control, namely the oleoylated 2-aminomethylbenzimidazole **74** (see Chapter 6), the -CH_2 in this molecule was found to have a lower chemical shift (3.55 ppm) than its non-acylated+liposome counterpart (3.95 ppm at low lipid-molecule conc.). Thus, this reduction in chemical shift for the -CH_2 group indicates formation of the amide, and therefore likely reactivity via an aminolysis mechanism.

To further confirm this aminolysis mechanism, the second set of experiments were 2D NOESY experiments across a series of time intervals from a mixing time of 3-300 ms. As alluded to in the methodology, this increase in mixing time can then be plotted as a volume intensity curve to show possible close contact interactions between protons on the molecule relative to the lipid. All these plots indicate intramolecular interactions, either between the molecule or the lipid itself. At low concentration in figure 3.12, the intensity of the cross peak between the -CH_2 (5) and -CH_3 methyl (10) increases from 3-300 ms. At higher concentration, in figures 3.12C and 3.12D, both the intensity of the crosspeaks between the acyl chain and lipid protons (10) increases and additional cross-peaks emerge between the acyl chain of the lipid and aliphatic protons on the lipid chain (8,13) from mixing time 3 to 300 ms. Furthermore, the intensity of cross-peaks between the aromatic protons on the molecule decrease over time. The lack of nOe interactions between the

molecule and lipid might testify to a reduction in lysolipid formation, as previously examined in the literature. However, as nOe measures interactions through space, more cross-peaks implies closer proximity and thus potential reactivity between the nitrogen (molecule) and carbon (lipid) to form the amidated product.

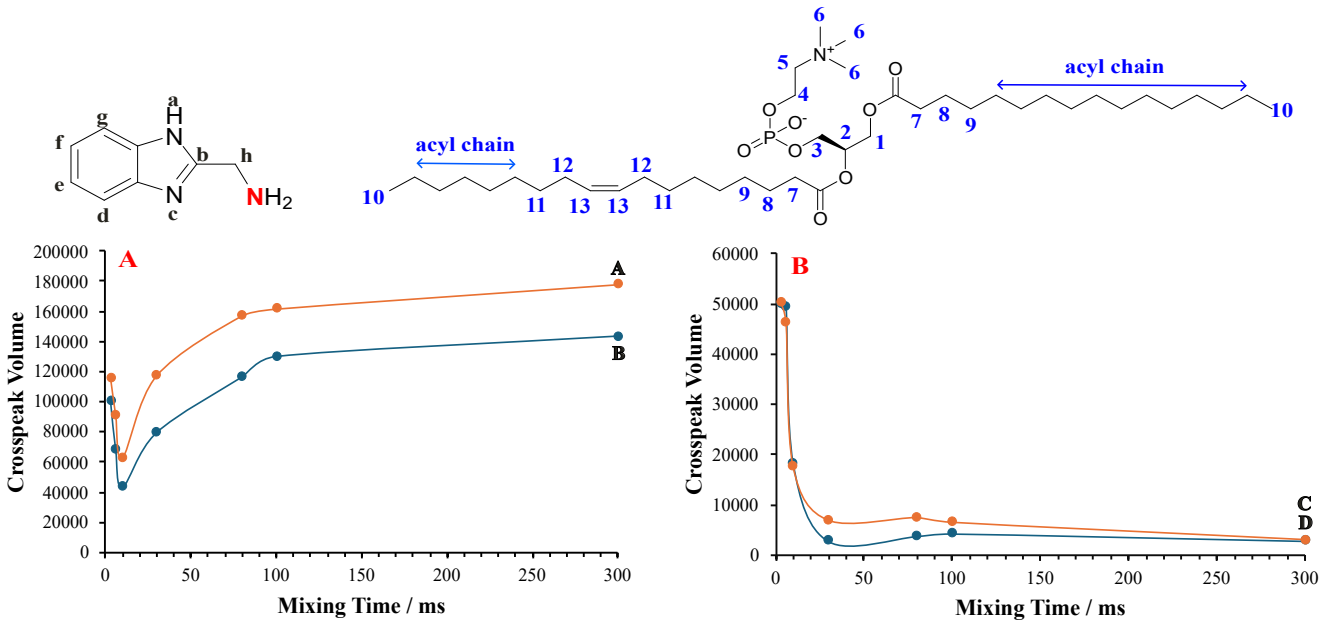


Figure 3.13: nOe build-up plots: A - two cross peaks between lipid acyl chain ('acyl') and lipid terminal methyl ('10'), B - two cross peaks between drug aromatic protons ('e,f') and drug aromatic protons ('d,g'), both plots showing lipid intramolecular interactions

Lastly for this molecule, these nOe build-up curves measure the change in crosspeak volume as a function of mixing time. In figure 3.13A, the plot shows an initial steep descent from 0-10 ms, then a gradual increase from 10-100 ms, until the line plateaus off. Importantly between 30-80 ms there is a linear regime indicating cross-relaxation between the acyl chain protons (6) and lipid terminal methyl protons (3). This implies intramolecular interactions, specifically within this region, between the two lipid species. Figure 3.13B shows an exponential reduction in crosspeak volume between the aromatic protons of 2-aminomethylbenzimidazole between 0-30 ms, but then plateauing from 30-300 ms. This non-linear behaviour suggests only intramolecular interactions being the dominant ones, so whilst reactivity may occur, subsequent lysolipid formation may be reduced. This data would need to be repeated several times to confirm no additional cross peaks between the molecule and lipid are present.

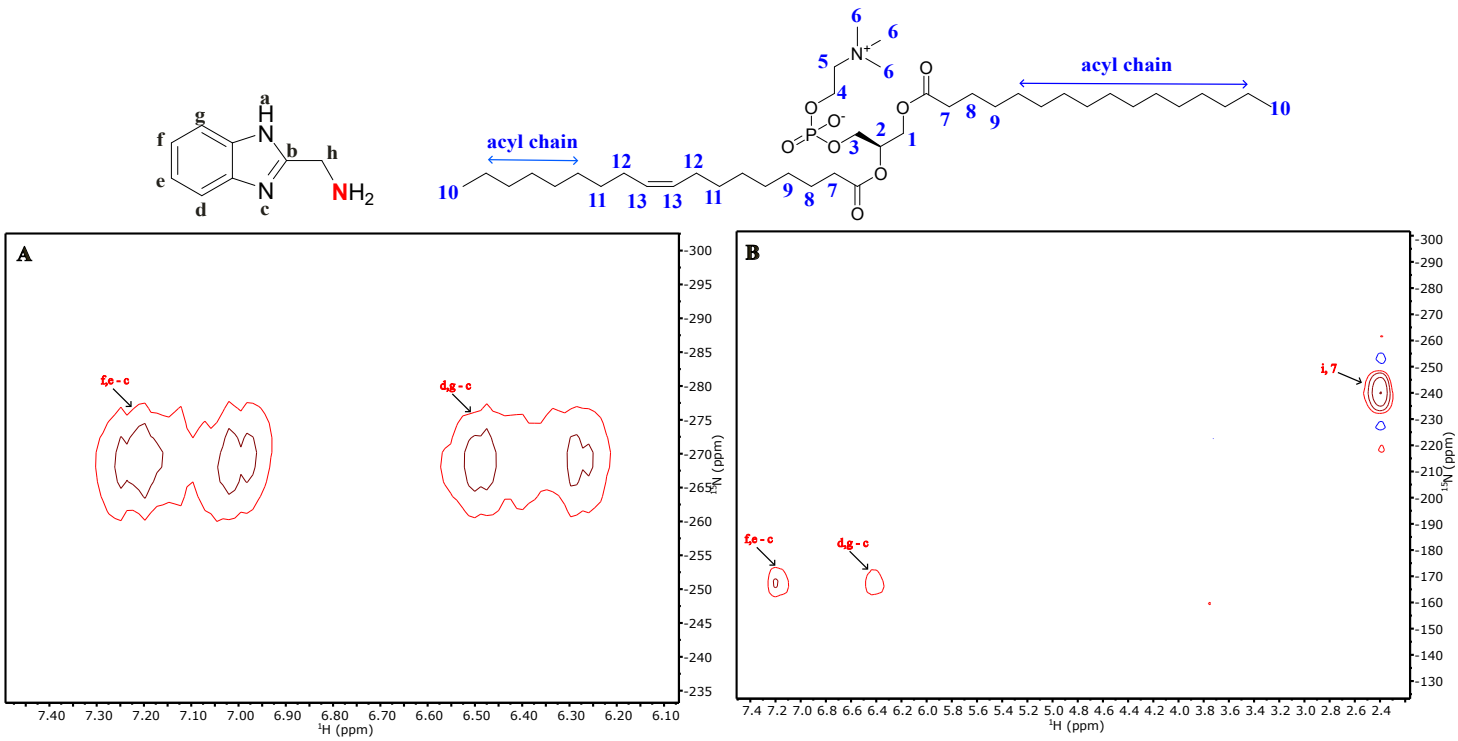


Figure 3.14: ^{1}H-^{15}N HMBC spectra: (A) compound+liposomes (3 ms), (B) compound+liposomes (300 ms)

The ^{1}H-^{15}N HMBC spectra in figure 3.14 confirms correlation between the nucleophilic nitrogen of the compound and the lipid carbon. The spectrum at 3 ms mixing time shows two cross peaks. Both show correlation, separated by three bonds, between the aromatic protons of 2-aminomethylbenzimidazole and the benzimidazole lone nitrogen atom (c). As the system evolves in time up to 300 ms, an additional crosspeak appears. This shows a correlation between the ^{15}N enriched nitrogen of 2-aminomethylbenzimidazole and the -CH_2 (7) of the lipid. This crosspeak is found at δ 2.3 ppm and -240 ppm. This nitrogen chemical shift is in the correct range for an amide group and shows a three-bond connection between the molecule and lipid. A combination of all this data points to aminolysis, with reduced rate of lysolipid formation, as being the likely mechanism for reactivity.

Connection with Simulations

The 1D spectra (figure 3.11) are significant in pointing towards interactions between the molecule and lipid. The chemical shift change between low and high (figure 3.11C-D) [lipid]:[molecule] of the -CH_2 (h) of the molecule from δ 3.9 to 3.8 ppm shows binding of the molecule with the lipid. This is further confirmed by the

chemical shift change in the aromatic region of the molecule (e,f/d,g) from 7.55-7.45 and 7.25-7.20 ppm respectively. Lastly, the chemical shift change of the phosphocholine protons (6) from low to high [lipid]:[molecule] shows a directional interaction between the lipid headgroup and aromatic protons of the molecule.

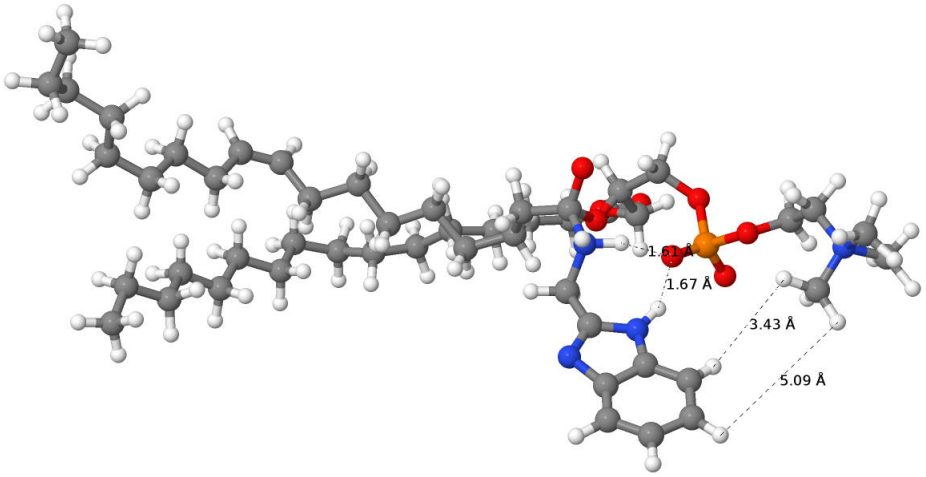


Figure 3.15: MD snapshot of close contact interactions between the aromatic protons of the molecule and lipid headgroup protons

The simulation snapshots presented here are taken from chapter 5, where QM/MM simulations have been conducted to locate the energy minimum, and thus stable conformation in the energy landscape. The MD simulation (figure 3.15) here shows a close contact interaction between the -CH_2 (h) of the molecule and the phosphate oxyanion of the lipid. Furthermore, this shows a strong interaction between the aromatic protons of the molecule (f,g) and the phosphocholine head group protons (6). This theoretical model correlates well with experimental findings. The chemical shift change of the phosphocholine headgroup implies close contact with the molecule and likewise the change of the -CH_2 (h) of the molecule indicates close contact with the lipid head group. Thus, these 1D spectra are a useful comparison with simulations to confirm molecule reactivity with the lipid. The different intermediates in the reaction pathway of this molecule with the lipid are quantified in Chapter 5.

3.3.3 ssNMR System II: ^{15}N Propranolol with POPC liposomes

The first set of experiments were 1D and 2D NOESY to probe close contacts between the drug and the lipid. In figure 3.16A, the proton NMR spectrum shows the 2

methyl peaks at an integration of 6 and the aromatic protons in the correct chemical shift region between δ 6.8-8.25 ppm. The proton NMR spectrum for liposomes (figure 3.16B) is identical to the other molecule-lipid systems. In both the low and high [lipid]:[drug] proton NMR spectrums, the integration of the acyl chain peak occurs at a similar resonance to the 2 methyl groups of the drug (l,m), hence a total integration value of 21. The proton spectrum in figure 3.16C shows broadening of peaks, due to the liposome, but also far more overlap of the aliphatic chemical shift of the drug with the lipid. An increase in [lipid]:[drug] concentration ratio from [1.21:1] to [8.03:1] causes considerable chemical shifts. Noticeably, the drug aromatic protons are broader and their chemical shifts are more de-shielded. Also, the aliphatic -CH_2 protons on the backbone of propranolol (h,j,k) are more de-shielded at higher concentration (figure 3.16D). These chemical shifts indicate interactions between the polar part of propranolol and the lipid acyl chains. Ubiquitously, this behaviour may be primary evidence of transesterification behaviour with increased lysolipid formation. Further 2D NOESY experiments were then conducted to investigate close contacts between the drug and lipid.

In figure 3.17A, there are no significant cross-peaks and thus limited information at a short 3 ms mixing time. Upon time evolution to 300 ms (figure 3.17B), this spectrum shows cross peaks between the drug and lipid. Initially there are cross-peaks between aromatic protons of the drug, but more noticeable are the interactions between the drug aliphatic protons (h,k) and the lipid headgroup protons attached to the nitrogen cation (6). There are also interactions between the isopropyl protons (l,m) of the drug and both the lipid headgroup protons (6) and -CH_2 protons (8) close to the electrophilic carbonyl group, indicative of interactions through space. Furthermore, an increase in the lipid concentration yields far more cross-peaks. At 3 ms (figure 3.17C) there are already the emergence of cross-peaks between the aliphatic -CH_2 (h) of propranolol and -CH_2 (5) attached to the choline headgroup. This implies close proximity of the drug to the lipid headgroup for reactivity. Further, the 2 isopropyl protons have a cross-peak with the -CH_2 protons (8) close to the electrophilic carbonyl group, as seen before. Lastly in figure 3.17D at 300 ms, a plethora more of cross-peaks emerge, all between either drug aliphatic -CH_2 protons

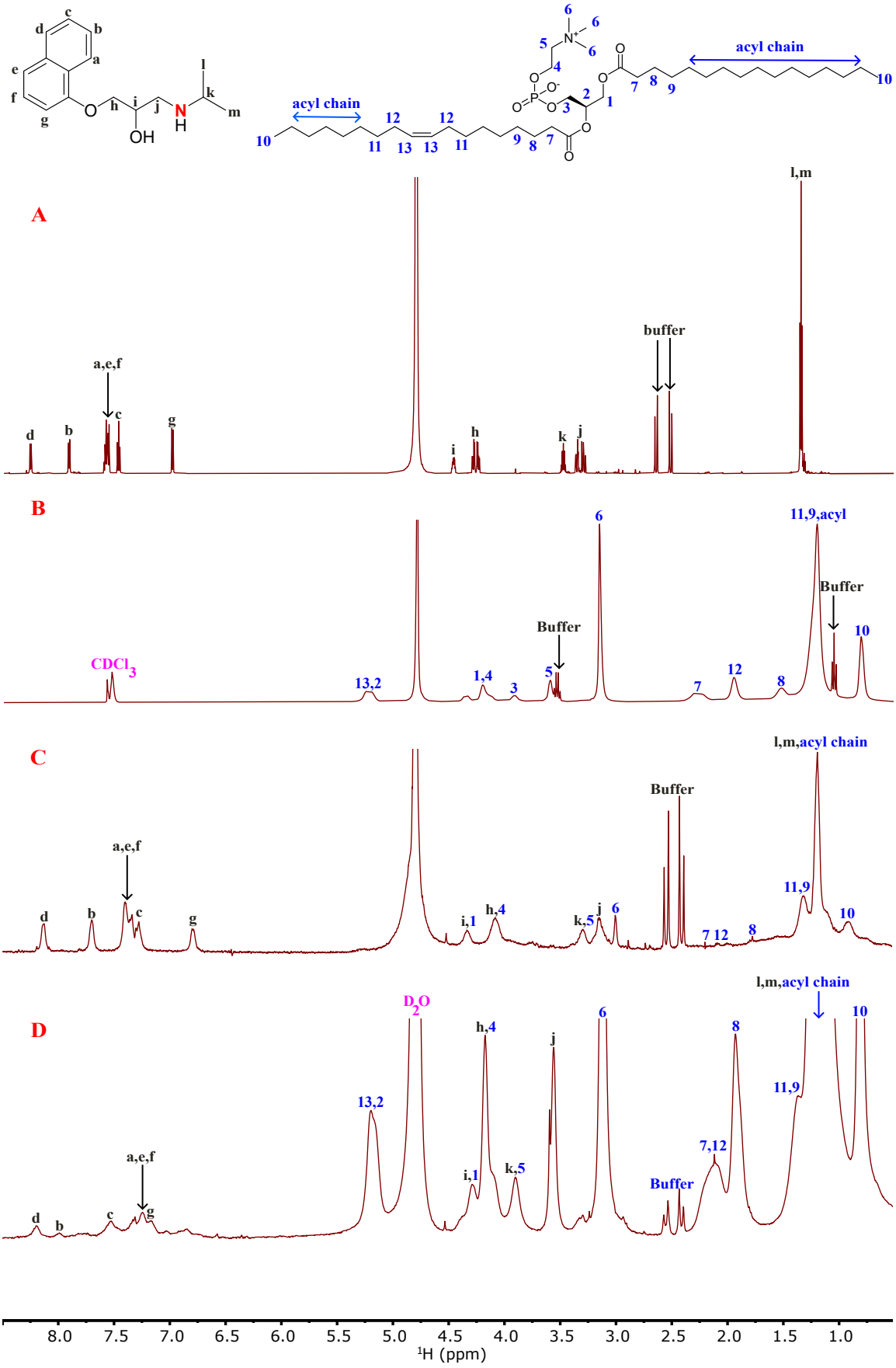


Figure 3.16: ^{1}H NMR spectra (run in buffer): (A) propranolol, (B) POPC liposomes, (C) liposomes+compound at low concentration ratio [1.21:1], (D) liposomes+compound at higher concentration ratio [8.03:1]

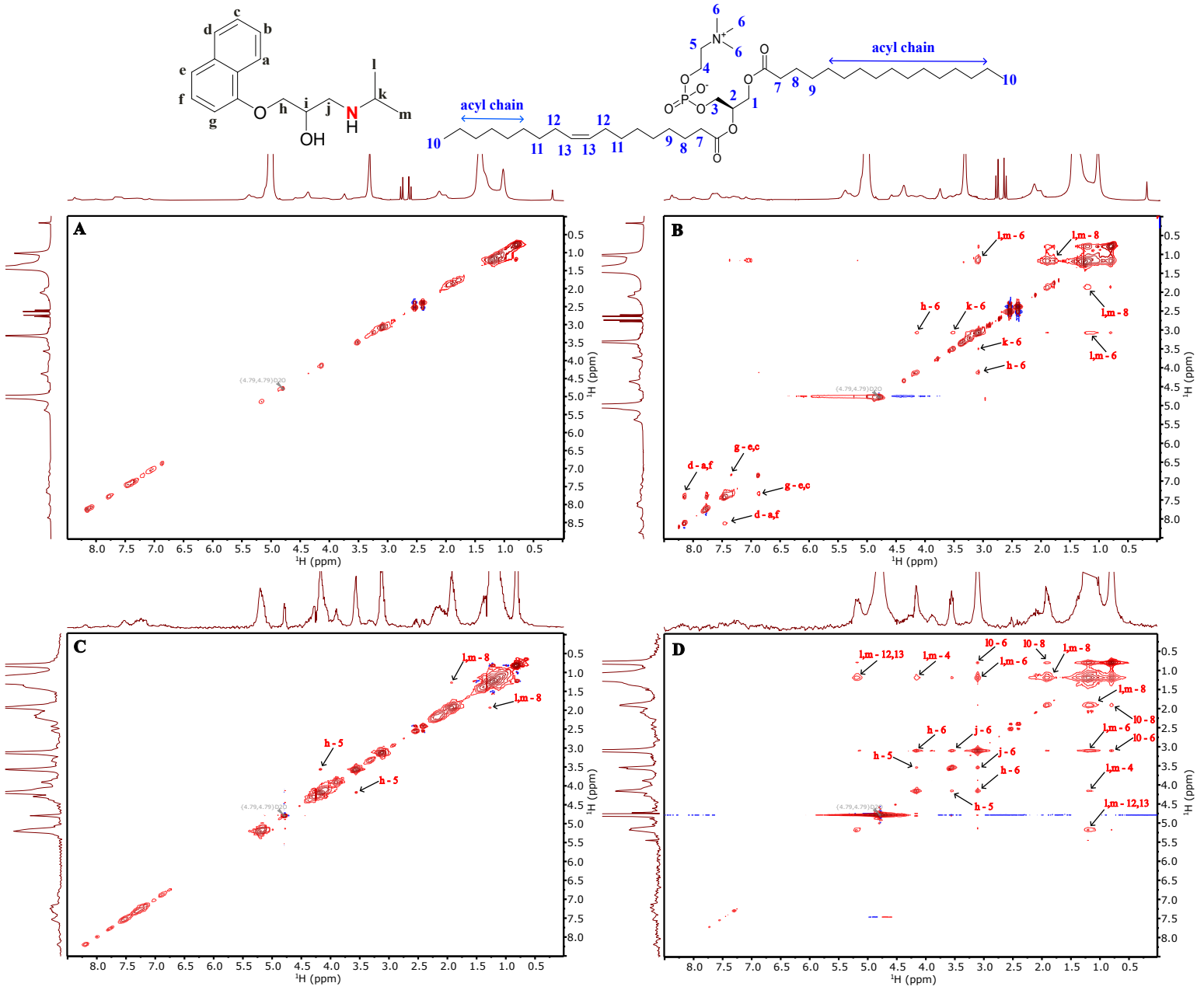


Figure 3.17: 2D NOESY spectra (run in buffer): (A) compound+liposomes (1:1.21, 3 ms), (B) compound+liposomes (1:1.21, 300 ms), (C) compound+liposomes (1:8.03, 3 ms), (D) compound+liposomes (1:8.03, 300 ms)

(h,j) and lipid headgroup protons (5,6), or between drug isopropyl protons (l,m) and lipid fatty acyl protons (8,12,13). Intermolecular interactions between protons on the drug and protons on the lipid acyl chain imply that the drug localises parallel to the lipid. The sheer increase in number of interactions points to transesterification taking place at the oxygen nucleophilic centre with the electrophilic carbon, alongside increase in lysolipid formation, as confirmed with nOe distance measurements.

Lastly for this molecule, these nOe build-up curves measure the change in cross peak volume as a function of mixing time. In figure 3.18A, the plot shows an initial steep descent from 0-10 ms, then a gradual increase from 10-100 ms, until the line plateaus off. Importantly between 30-80 ms there is a linear regime indicating cross-relaxation between the aliphatic -CH_2 (h) of the drug and the lipid headgroup methyls (6). This demonstrates close contact between the polar part of the drug and lipid headgroup. Figure 3.18B shows similar behaviour between the aliphatic -CH_2 (8) of the lipid and the isopropyl methyl groups of the drug (l,m), further showing close distance between the lipid protons, close to the reactive site, and drug methyl groups. Both linear regimes for these two cross-peaks show intermolecular interactions between the drug and lipid and confirms transesterification as the most viable reactivity pathway, alongside an increase in the rate of lysolipid formation.

The ^{1}H-^{15}N HMBC spectrum in figure 3.19 shows two cross-peaks at a mixing time of 3 ms and an increase in the intensity of the cross-peak from 3-300 ms. Having analysed the nOe plots, both of these cross peaks refer to an interaction between the ^{15}N nitrogen of propranolol and the protons attached to the acyl chain of the lipid, implying a close through-bond contact between the drug and the lipid. As the literature demonstrates that propranolol exists as a major protonated and minor neutral form, these two peaks could correspond to different depths of membrane partitioning of the two forms of propranolol. Protonated propranolol would only form the esterified product, whereas neutral propranolol would form both the esterified and aminated product, with the former being at a higher concentration. Thus, there would be stark difference in ^{1}H-^{15}N correlation for an amide versus an ester, accounted for by these two peaks.

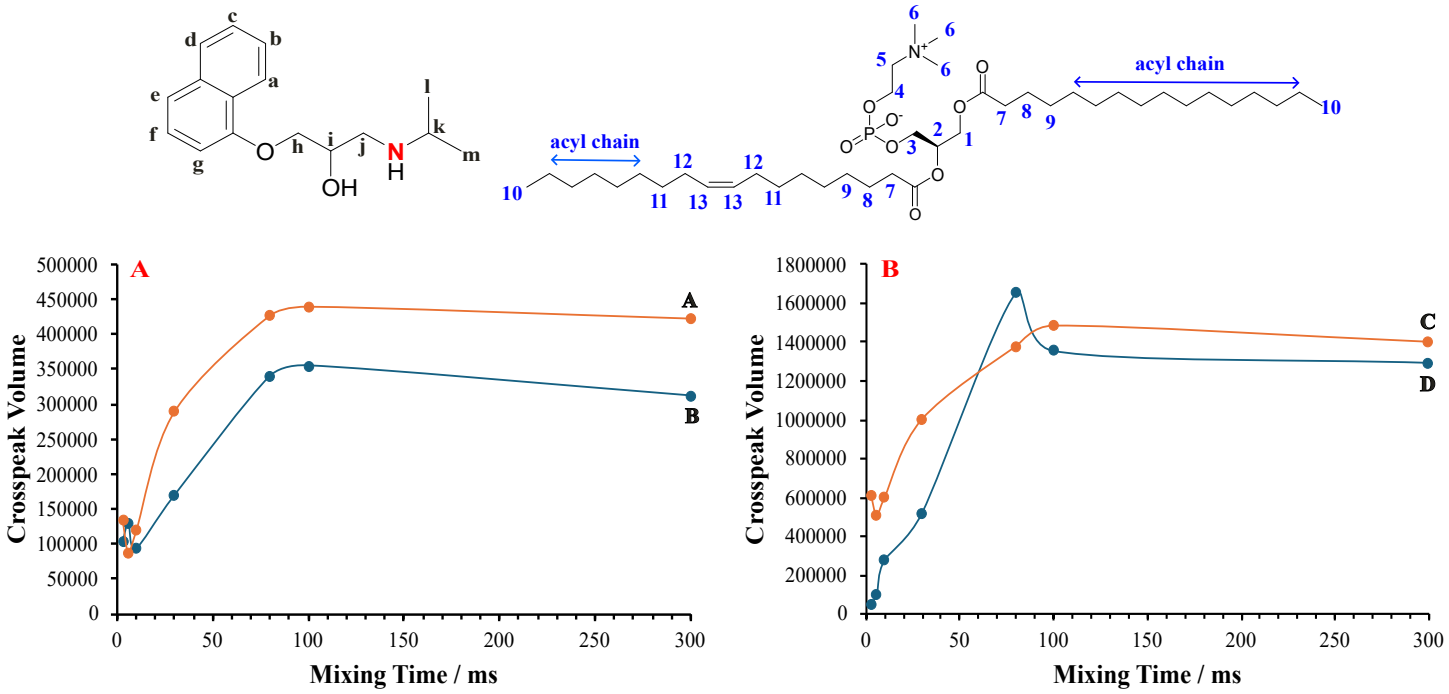


Figure 3.18: nOe build-up plots: A - two cross peaks between drug -CH_2 ('h') and lipid headgroup methyl ('6'), B - two cross peaks between drug methyl ('l,m') and lipid aliphatic -CH_2 ('8'), both showing drug-lipid intermolecular interactions

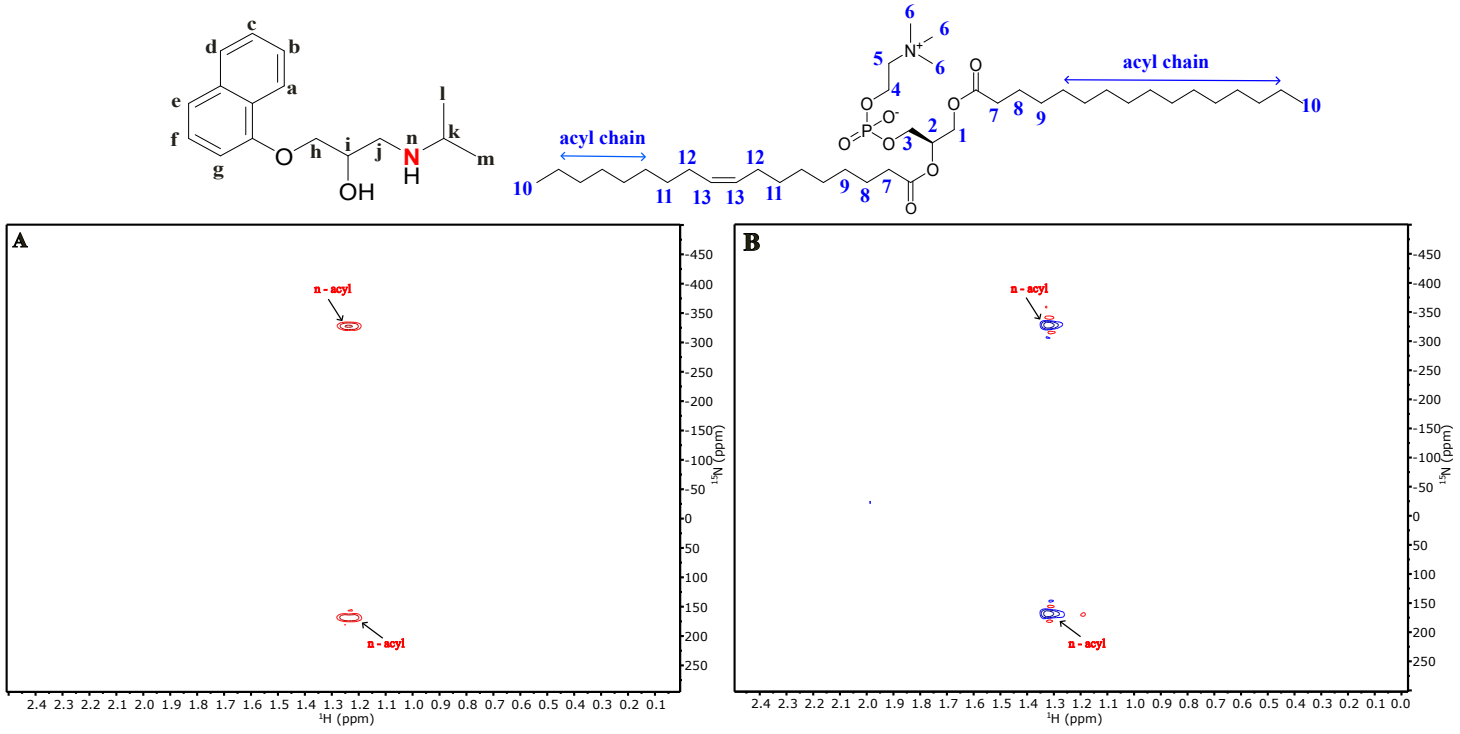


Figure 3.19: ^{1}H-^{15}N HMBC Plots: (A) compound + liposomes (3 ms), (B) compound + liposomes (300 ms)

Connection with Simulations

The 1D spectra (figure 3.16) and 2D NOESY plots (figure 3.17) show good qualitative information on the orientation of propranolol in the membrane. Cross-peaks have shown that the polar parts of propranolol (h,j) interact strongly with the lipid head group region (6). Additional cross-peaks have shown the interaction of the methyl groups of propranolol (l,m) with the lipid acyl chains (12,13). This experimental data points to reactivity and binding of the drug with the lipid across a mixing time of 3-300 ms. Figure 3.20 shows an MD simulation of the drug propranolol reacting with a POPC lipid. This theoretical model correlates well with experimental findings in showing that the drug orients itself near the membrane-water interface. Importantly from the model, $\pi - \pi$ stacking interactions between the aromatic rings of propranolol and lipid head group contribute to stability within the membrane. The interaction distance between the aromatic drug protons and lipid phosphocholine head group protons decrease over time and shows that the aromatic ring is in close proximity to the lipid head groups, as shown experimentally, and the remainder is closer to the membrane-water interface. Furthermore, interactions between the polar part of propranolol (h) and lipid head group (5) are correctly predicted by the model.

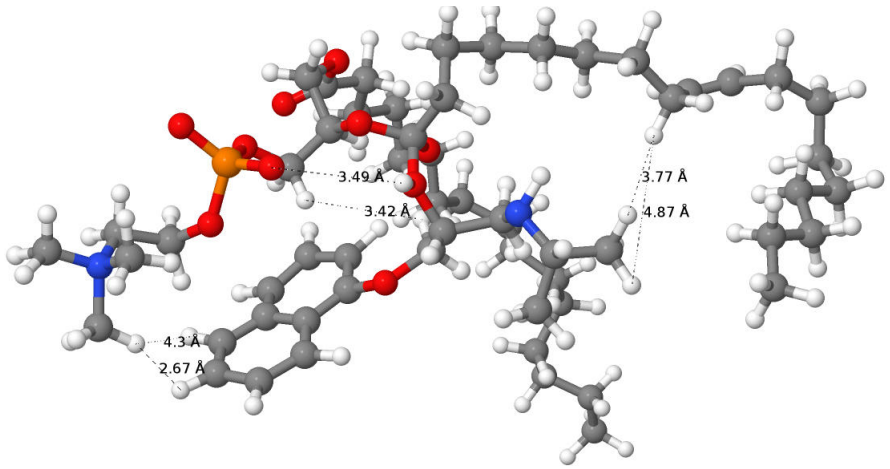


Figure 3.20: MD snapshot of close contact interactions between the aromatic protons of the molecule and lipid headgroup protons, polar part of drug and lipid headgroup protons, methyl groups of propranolol and lipid acyl chain

Previous literature has shown that for neutral CADs, such as propranolol or tetracaine, these molecules are bound to the membrane provided that the charged part of the molecule localises closer to the membrane surface, whilst the hydrophobic part

is localised closer to the acyl chain of the lipid[9]. Furthermore, additional literature has shown that drug-lipid reactivity is stabilised if the reactive CAD has H-bond donors as part of its structure[90]. Additional pi-pi stacking interactions have been shown to contribute to stability of the drug at the membrane interface[49].

In accordance with this literature, my simulations have shown the crucial role of the phosphate oxyanion in stabilising the drug-lipid system, especially as the hydroxyl proton is in close proximity with the lipid headgroup. Additional hydrogen bonding of the aromatic drug protons, seen experimentally and computationally, with the lipid head group protons confirm reactivity and localisation of the polar part of the drug closer to the lipid head groups. Finally, ssNMR data predict favoured orientation and binding of propranolol with the lipid. This is confirmed through close distance between the drug methyl groups and lipid acyl chain which also support the literature that the hydrophobic part of the drug is buried closer to the lipid acyl chain. Further quantitative measurements of drug intermediates have been calculated in Chapter 5.

3.3.4 ssNMR System III: ^{15}N 4-Amino-*N*-phenylbutanamide with POPC Liposomes

The first set of experiments were 1D and 2D NOESY to probe close contacts between the drug and the lipid. In figure 3.21A, the spectrum shows the split methyl protons in the aliphatic chain of the drug, with the citrate protons close to the chemical shift region of these aliphatic protons. The aromatic protons are all accounted for in the δ 7.25-7.4ppm region of the spectrum. In figure 3.21C at low concentration, the chemical shift for -CH_2 (8) protons changes from δ 2.9-3.1 to 2.85 ppm and from 2.1 to 1.8 ppm for -CH_2 (7) protons. As acylated 4-amino-*N*-phenylbutanamide **75** (see Chapter 6) was synthesised as a control, and its chemical shifts show a reduction in ppm for -CH_2 (7,8) proton environments, this behaviour correlates with the proton spectrum obtained for the liposome+drug at low concentration ratio ([0.60:1]). The liposome+drug at higher concentration ratio ([2.95:1]) (figure 3.21D) reveals chemical shifts of the drug and lipid in essentially the same position as with lower concentration, but with an increased integration for the protons both on

the lipid acyl chain (acyl) and the phosphocholine headgroup (6). Importantly, the chemical shift change for the aliphatic protons (7,8) of the drug point to formation of the amide product, similarly to 2-aminomethylbenzimidazole. This would suggest an aminolysis mechanism, with no change in lysolipid formation, as confirmed in the literature.

In figure 3.22, these spectra are 2D NOESY plots for the interaction of ^{15}N 4-amino-*N*-phenylbutanamide with POPC liposomes. figure 3.22A at a low mixing time (3 ms) already shows cross-peaks between the aliphatic -CH_2 protons of the drug (h,g,f) and -CH_2 protons of the lipid acyl chain (8,12). An important cross-peak is the one between the drug (h) and lipid (8), this suggest close contact between the ^{15}N amine and carbonyl carbon of the lipid. An increase in mixing time to 300 ms reveals additional cross-peaks between the lipid, showing intramolecular interactions between the protons on the acyl chain (10) and terminal protons (11). Also, there is a significant increase in intensity of drug-lipid peaks when compared with the 3 ms spectrum. In figure 3.22C at low mixing time (3 ms), identical cross-peaks between the drug and lipid are present, suggesting that the drug localises close to the lipid and multiple interactions between the aliphatic protons of the drug and lipid acyl chain further confirms aminolysis as the most likely reaction mechanism. Interestingly, an increase in mixing time to 300 ms results in new cross-peaks. There are now interactions between the aliphatic drug protons (h) and protons closer to the phosphocholine headgroup (3,5), suggesting close contact interaction at the headgroup site. Further, interactions between the drug and acyl chain are now present, suggesting that the drug would lie parallel to the acyl chain in reactivity.

Lastly for this molecule, the nOe build-up curves were obtained from the higher [lipid]:[drug] concentration ratio [2.95:1] 2D NOESY spectra. Figure 3.23A shows a small decrease in cross-peak volume from 0-10 ms, then an initial smooth increase in volume from 10-100 ms, then the curve plateaus off from 100-300 ms. These peaks refer to the interaction between the drug -CH_2 (g) protons and lipid acyl chain protons (acyl). As the linear regime is maintained from 30-80 ms, this spectrum

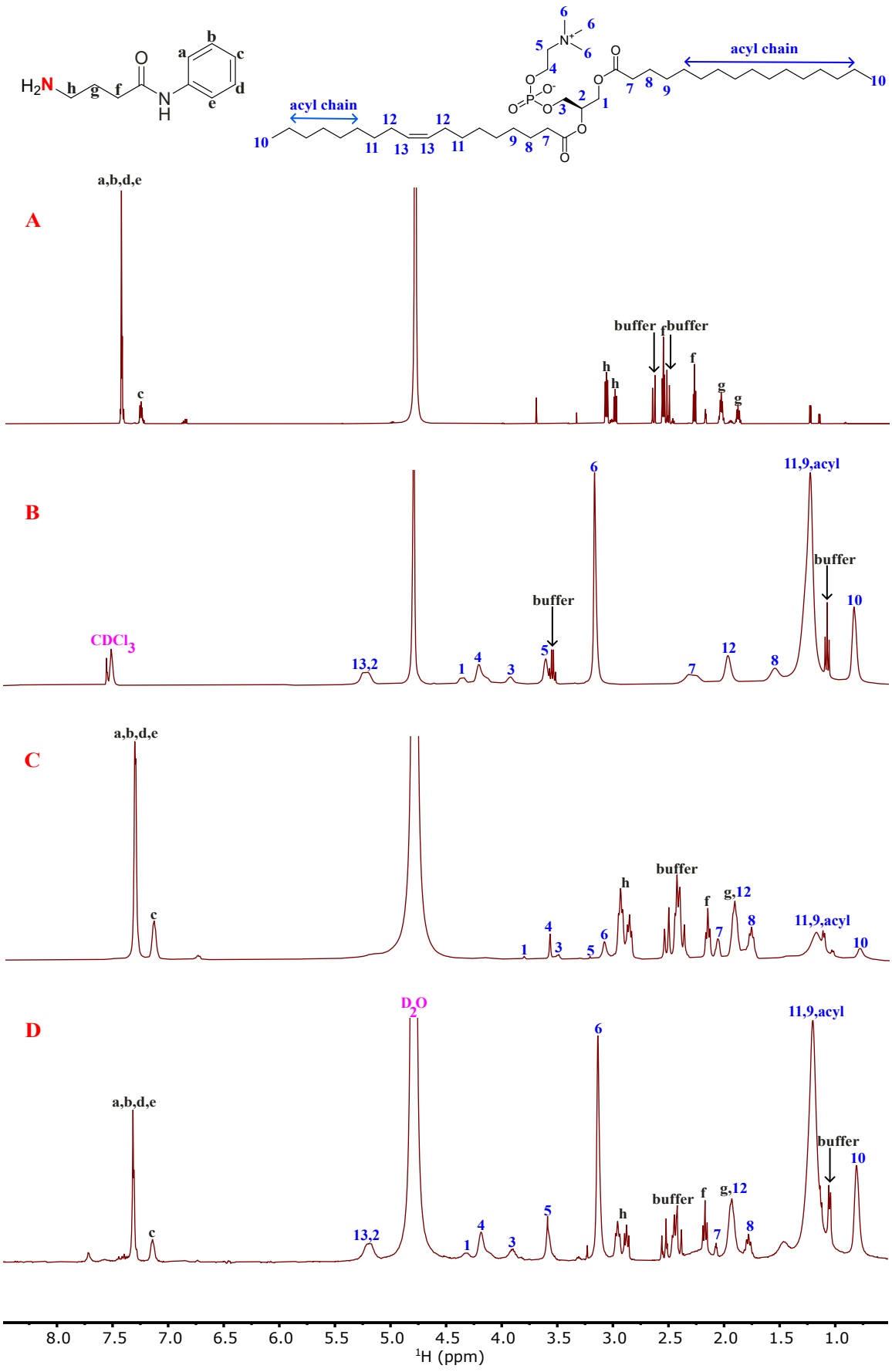


Figure 3.21: ^{1}H NMR spectra (run in buffer): (A) 4-amino-*N*-phenylbutanamide, (B) POPC liposomes, (C) liposomes+compound at low concentration ratio [0.60:1], (D) liposomes+compound at higher concentration ratio [2.95:1]

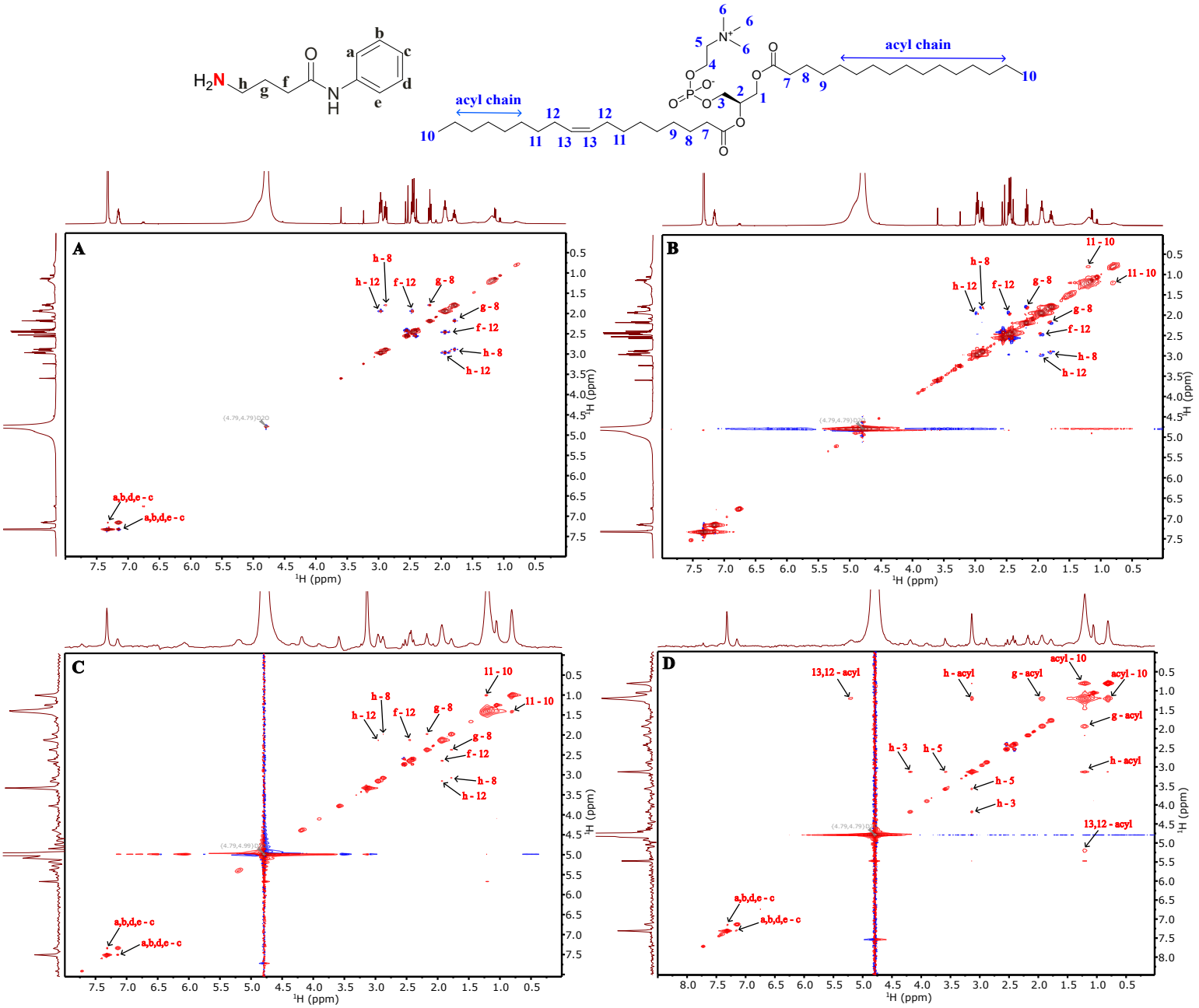


Figure 3.22: 2D NOESY spectra (run in buffer): (A) compound+liposomes (1:0.60, 3 ms), (B) compound+liposomes (1:0.60, 300 ms), (C) compound+liposomes (1:2.95, 3 ms), (D) compound+liposomes (1:2.95, 300 ms)

suggests close contact between the drug aliphatic protons and lipid chain. Figure 3.23B shows an initial steep decrease in cross-peak volume as a function of mixing time from 0-10 ms, a gradual increase from 10-100 ms then once again the graph plateaus off from 100-300 ms. This linear regime is present from 30-80 ms, showing a close contact between the drug -CH$_2$ aliphatic protons (g) and lipid -CH$_2$ protons close to the lipid headgroup (5). This behaviour suggests that reactivity happens close to the headgroup, namely the aminolysis behaviour between the nucleophilic drug and membrane lipids. All these cross-peaks suggest intermolecular interactions between the drug and lipid resulting in the formation of a stable acylated amide product.

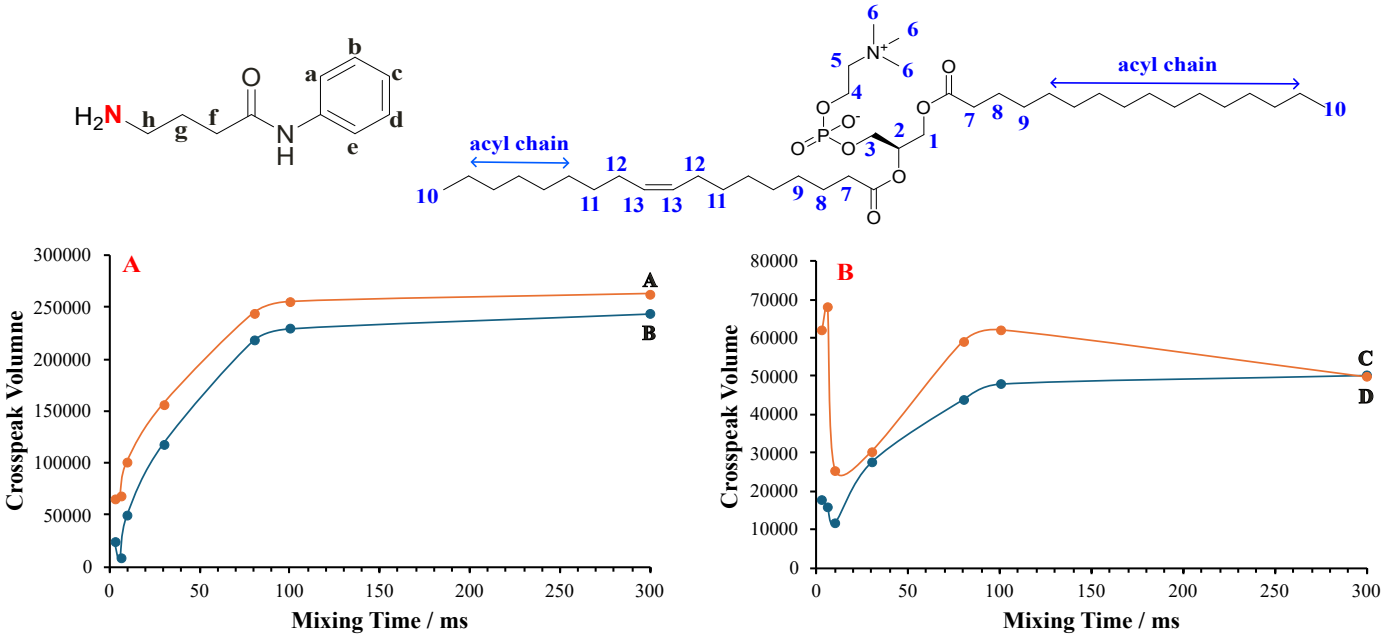


Figure 3.23: nOe build-up plots: A - two cross peaks between drug -CH$_2$ (‘g’) and lipid acyl chain (‘acyl’), B - two cross peaks between drug -CH$_2$ (‘h’) and lipid -CH$_2$ (‘5’), both showing drug-lipid intermolecular interactions

Finally, ^{1}H-^{15}N HMBC spectra were obtained to establish connectivity between the drug and lipid. As can be seen in figure 3.24A, multiple cross-peaks appeared at 3 ms mixing time, between the aromatic protons of the drug (a-e) and amine nitrogen bonded to the benzyl ring (i). However an increase in mixing time to 300 ms resulted in a new cross-peak at δ 3.0 ppm in the proton and δ -340 ppm in the nitrogen spectra. This chemical shift is in alignment with formation of an amide, and this cross-peak correlates with a through-bond connection between the ^{15}N amine and phosphocholine headgroup protons (6). This data, coupled with nOe measurements,

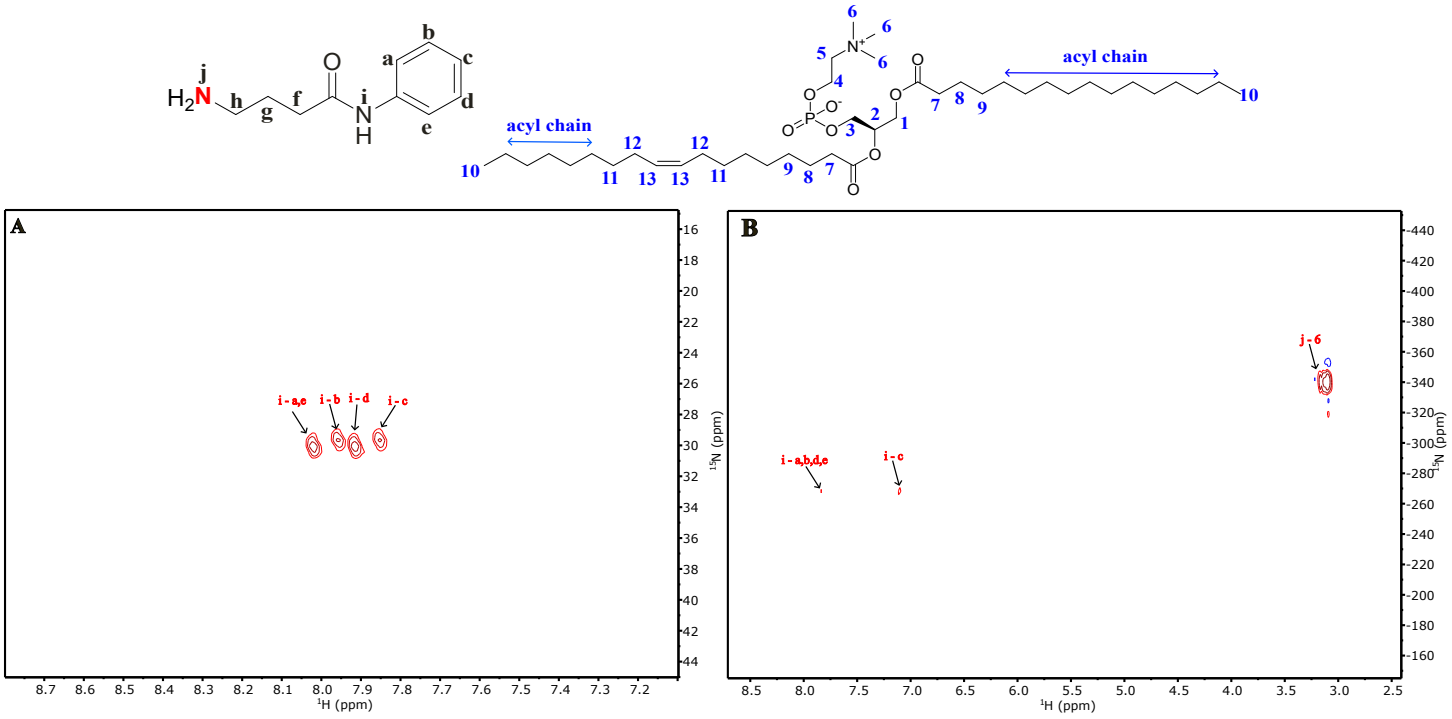


Figure 3.24: ^{1}H-^{15}N HMBC plots: (A) compound + liposomes (3 ms), (B) compound + liposomes (300 ms)

demonstrates aminolysis behaviour to form the stable acylated amide product, in a similar fashion to 2-aminomethylbenzimidazole.

Connection with Simulations

In a similar manner to 2-aminomethylbenzimidazole, both the 1D spectra (figure 3.21) and 2D NOESY spectra (figure 3.22) point towards reactivity of the drug with the lipid. A chemical shift change from 2.0 to 1.8 ppm for the drug aliphatic -CH_2 protons (g) and 3.0 to 2.8 ppm for the drug -CH_2 protons (h) when mixed with POPC liposomes show binding of the drug with the lipid. Furthermore, the emergence of multiplet signals for the aliphatic drug protons -CH_2 (h,g) (cf. figure 3.21A to figure 3.21C,D) confirm that the isomerism has been resolved and formed the stable acylated drug. The chemical shift change of the aromatic drug protons (a-e) also show closer proximity of the drug with the lipid, and thus more likely reactivity between them. Furthermore, the chemical shift change for both the lipid acyl chain protons and lipid head group protons (3,5) indicate a favourable interaction with the drug molecule.

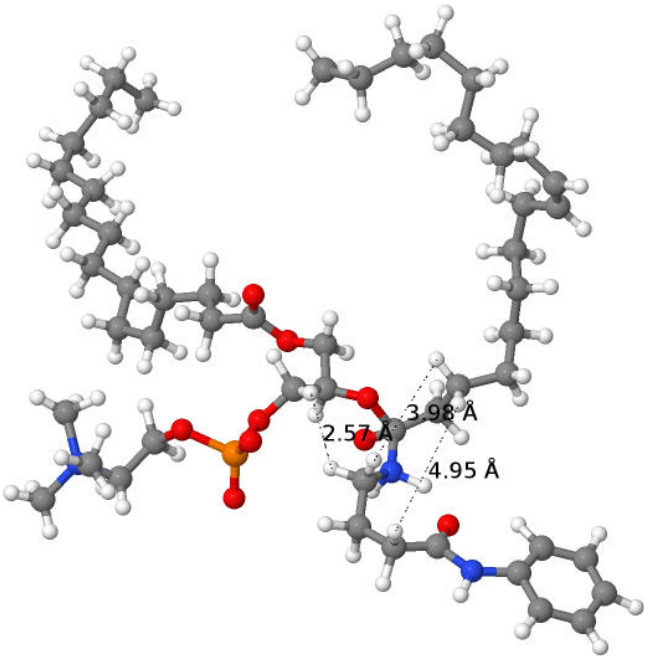


Figure 3.25: MD snapshot of close contact interactions between the aliphatic drug protons of the molecule and lipid headgroup protons, and aliphatic protons close to the reactive lipid carbonyl

Figure 3.25 shows an MD snapshot of the interaction of this drug with the POPC lipid. This theoretical model correlates well with experimental ssNMR data in showing the close proximity, and subsequent reactivity, of this drug with the membrane lipids. It correctly predicts a close contact interaction between the drug protons (h) and lipid head group protons (3), as well as additional interactions between the drug protons (h,f) and lipid protons close to the carbonyl reactive centre (8). These simulation results nicely complement the ssNMR data in showing that aminolysis is the most favourable reaction pathway of this drug molecule with the lipid. The presence of intermolecular interactions between the drug and lipid is captured by this MD simulation, and is verified with the nOe build-up curves which show a favoured interaction between the drug aliphatic protons and both the acyl chain and lipid head group protons (3,5), the former showing that the drug localises parallel to the fatty acyl chain. Subsequent intermediates in this reaction pathway have been quantified in Chapter 5.

3.4 Summary

In conclusion, a combination of 1D, 2D NOESY and 2D HMBC measurements have been utilised to probe the potential close contacts between ^{15}N 2-aminomethyl benzimidazole, ^{15}N propranolol and ^{15}N 4-amino-*N*-phenylbutanamide, each with POPC liposomes in buffer solution.

The first system demonstrated a high number of intramolecular interactions between the 2-aminomethylbenzimidazole molecule and POPC liposomes, further confirmed by the nOe build-up curves reaching an equilibrium between different parts of the lipid. A potential conclusion could be that a reduction in lysolipid formation, as previously demonstrated in the literature, may result in higher intramolecular, rather than intermolecular interactions. The aminolysis mechanism was confirmed by a close contact in the ^{1}H-^{15}N HMBC spectrum and thus demonstrates reactivity between the drug and lipid. The simulation results (Chapter 4) correlate well with these experiments in showing that the lipid head group protons are in close proximity with the drug aromatic protons, and thus contribute to favourable reactivity.

The second system showed a high number of intermolecular interactions between the propranolol and POPC liposomes, confirmed alongside linear regime nOe build-up curves between polar parts of the drug and the lipid acyl chain. A conclusion for this behaviour is that propranolol orients itself close to the electrophilic carbonyl atom and additional hydrogen bonding/pi-pi stacking contributes to stability of this drug at the membrane interface. This behaviour seen experimentally is complemented nicely with simulation data which shows intermolecular interactions between the drug aromatic protons and lipid head group protons. Furthermore, the oxyanion phosphate contributes to stabilisation of the drug-lipid system by hydrogen bonding with protons attached the polar part of propranolol. ^{1}H-^{15}N HMBC spectra confirmed different depths of drug partitioning of both the protonated and neutral forms.

Lastly, there are an abundance of both inter and intramolecular interactions between 4-amino-*N*-phenylbutanamide and POPC liposomes, confirmed once again

by the linear regime nOe build-up curves between the aliphatic protons on the drug molecule and lipid acyl chain. A conclusion can be drawn that although the stable amide product is formed through aminolysis, there results in no lysolipid formation and thus an equal balance of inter and intra molecular interactions. MD simulations once more correlate well with this aminolysis behaviour in displaying close contact between the drug aliphatic protons and both the lipid head group protons and aliphatic protons close to the electrophilic carbonyl reactive centre of the lipid. ^{1}H-^{15}N HMBC spectra confirm formation of the amide after a mixing time of 300 ms, and thus follow the same aminolysis mechanism as with 2-aminomethylbenzimidazole.

To expand this work further, a range of different solvents, pH, temperature and molecules should be tested to ascertain whether similar transesterification and aminolysis mechanisms are present. Kinetic mechanisms can be further explored by continuation of this work.

Chapter 4

Atomistic Simulations of Drug-Lipid Interactions

4.1 Background

Simulations have the ability to shed insight into fundamental reactivity of organic molecules with lipids. For historical context, the earliest computer simulations of lipids can be traced back to 1980. Whilst still in its infancy, the soft-matter systems studied computationally were small solvated micelles and monolayers. From the early 1980s to the 1990s, Scott et al.[189] utilised the capability of molecular dynamics as the most routinely applied computational technique to understand the dynamics of lipid membranes. As this field developed, so too did the accuracy of these models[58]. The first realistic, detailed membrane model was provided by Pastor and Venable in 1993 specifically focused on PC lipid headgroups. The timescale of these simulations was limited to picoseconds but still very computationally demanding for the time, and did reveal molecular motions and membrane viscosities at these timescales[118]. By 1997, Berger had fully developed and characterised the so called 'Berger Lipid Model'. This was the golden standard in lipid simulations until the early 2010s. To ensure the accuracy and validation of this universally applied force field, Berger et al. sought to re-parametrise the non-bonded interactions

present within the lipid force field. They altered the Lennard-Jones parameters across each CH_2 and CH_3 group in the lipid chain and then tried to fit each data point with the lipid enthalpy of vaporisation found from experiment[58]. The system size of a lipid membrane in the last century was limited to perhaps 72-128 lipids (4,000-15,000 atoms) and a system timescale of maybe a couple of 100 picoseconds. Thanks to the advancement of computational power, current computing enables access to timescales ranging from nanoseconds to microseconds with sizes ranging from 10,000 to 150,000 atoms (figure 4.1). The accuracy of computational power enables efficient GPU machines to utilise atomistic simulations to simulate multi-scale, complex, realistic models of lipids[135]. Thus the primary computational tool to analyse drug-lipid interactions via computational means is atomistic simulations.

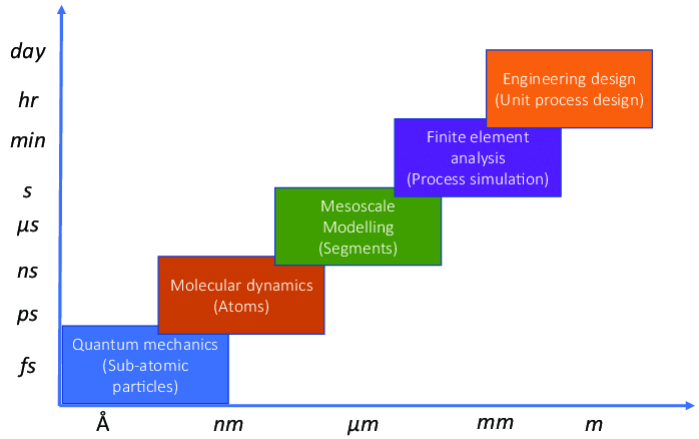


Figure 4.1: Comparison of length and timescales across different simulation techniques[135]

In this project, detailed comparison with experiment requires a solvated membrane bilayer with a single reactive drug molecule. Hence, fully atomistic molecular dynamics is the most appropriate simulation technique for characterising the lipid dynamics and quantifying reactivity of the lipid headgroups with the drug molecule. Atomistic simulations compute the properties of individual atoms within the system by considering each interaction site. Usually two categories of atomistic simulations are considered, namely the all-atom (AA) model and the united-atom (UA) model. The former assigns each individual atom with an interaction site whilst the latter excludes hydrogen atoms with the system. Often these H atoms are treated collectively with heavier atoms as a separate interaction site[144].

To ensure well represented, high quality atomistic models, there must be continual development and parameterisation of accurate force fields accounting for both bonded and non-bonded interactions. These parameters can be obtained from quantum chemical calculations, which are then later fitted to molecular mechanical parameters to ensure good experimental validation. The most common AA force fields used in biomolecular simulations are CHARMM, AMBER, GROMOS and OLPS. Among the most versatile of these software packages is AMBER. AMBER is not only associated with a set of protein and lipid force fields but also with specific molecular dynamics software. Whilst historically AMBER could only provide force fields for cholesterol, PC and PE lipids, the most recent LIPID21 software package includes all major acyl chains and lipid headgroups, due to multiple updates in the force fields[50]. CHARMM is also a set of reliable force fields for the molecular building of complex membrane systems, with CHARMM36 being the most recent software update. A major advantage of this force field over others is the high compatibility of this lipid force field with CHARMM additive force fields for smaller organic molecules. In the *NPT* ensemble, the CHARMM36 force field was found to produce accurate models for hydrophilic headgroups in the POPC, DMPC and DPPC membrane lipids. Akin to AMBER, this set of force fields also possess parameters for sterols and a host of different hydrocarbons[155].

Within this chapter, molecular dynamics (biased and static) is the primary simulation technique used to calculate the dynamical, structural and thermodynamical properties of the interaction between a reactive drug molecule and the corresponding lipid headgroup in the membrane. To ensure physiological conditions are maintained, each simulation would be under constant temperature (310 K) and constant pressure, whilst the bilayer is fully solvated. In all atom molecular dynamics (AA-MD) each system is placed in an appropriate simulation box with periodic boundary conditions and subjected to Newtonian mechanics to quantify the interacting forces between the particles[179]. For my system, the timestep between interacting particles is on the order of 1-2 fs with the system size being in the range of 35-40,000 atoms including water. Forces between individual atoms account for both bonded and non-bonded interactions. The bonded interactions are modelled as classical

harmonic springs with associated force constants for the stretching of bonds, the bending of molecular angles and the dihedral potentials. Whilst the non-bonded interactions account for the electrostatics and Van der Waals terms; these being the most computationally demanding part of the simulation[72]. With these mathematically constructed force fields, properties such as partial density, radial distribution function and molecular orientation can be extracted from these simulation trajectories. These analyses can help to uncover the preferred depth of orientation and positioning of a drug molecule within the membrane.

A complementary but distinct form of MD is biased-MD. This specific simulation techniques is utilised to calculate the free energy barrier of the drug whilst it traverses the membrane barrier. A bias or pulling force is required to translate the drug molecule through the membrane, while conformational space can be sampled through MD. This can inform on the stability of the drug molecule at the interface relative to the hydrophobic centre and aqueous regions of the membrane[30]. The first and foremost method to calculate the free energy utilises a statistical tool called umbrella sampling. This technique involves a biasing potential which improves the sampling of the thermodynamic quantity, in this case the free energy. This added potential limits the fluctuation of the sample as a function of coordinate distance, to a prescribed value, which improves the average configurational sampling of the drug at the membrane interface[176]. The second technique is known as thermodynamic integration to compute the free energy, typically between two different Hamiltonians. A constant known as the coupling parameter is assigned to each Hamiltonian respectively and the derivative (change) of the free energy as a function of coupling parameter is equal to the canonical average of the derivative of the Hamiltonian as a function of coupling parameter[98]. The calculation of free energy greatly demonstrates the preferred localisation and stability of the drug molecule in the membrane by comparing the free energy in the centre of the membrane compared with the interface[100].

4.2 Computational Methodology

4.2.1 Force Fields

In classical molecular mechanics, the force field is a mathematical summation of all the individual potential energy functions in the molecular system giving rise to a total potential energy. Ubiquitously, the force field accounts for all bonded and non-bonded interactions within the molecular system[162]. Thus, interactions involving atoms that are covalently and non-covalently bonded are added to give the total energy:

$$E_{\text{total}} = E_{\text{covalent}} + E_{\text{non-covalent}} \tag{4.1}$$

The covalent interactions are categorised by the energy contribution from atomic bond stretching, molecular angle bending and a torsion angle around a central bond. The non-covalent interactions account for the electrostatics, represented by Coulomb's Law and the van der Waals (vdw) interactions, represented by a Lennard-Jones Potential; thus taking the functional form:

$$E_{\text{covalent}} = E_{\text{bond}} + E_{\text{angle}} + E_{\text{dihedral}} \tag{4.2}$$

$$E_{\text{non-covalent}} = E_{\text{electrostatics}} + E_{\text{vdw}} \tag{4.3}$$

Within the context of the AMBER Force Field, the GAFF (general AMBER force field) is the bedrock upon which this set of force fields take its functional form. This force field has been fully parameterised and has been found to model simple pharmaceutical and bio-organic systems reasonably well[220]. Mathematically, the bond and angle terms are computed by harmonic potentials, whilst the torsion angles follow a Fourier series. The electrostatics are modelled by Coulomb's Law and the van der Waals intermolecular interactions are computed by a 6-12 Lennard-Jones potential[231]:

$$E_{\text{total}} = \sum_{\text{bonds}} K_b(b - b_{eq})^2 + \sum_{\text{angles}} K_\theta(\theta - \theta_{eq})^2 + \sum_{\text{dihedrals}} \frac{V_n}{2}[1 + \cos(n\phi - \gamma)] + \sum_{i<j} \left[\frac{A_{ij}}{(R_{ij})^{12}} - \frac{B_{ij}}{(R_{ij})^6} + \frac{q_i q_j}{\epsilon R_{ij}} \right] \tag{4.4}$$

Bonded Interactions

The bond stretching energy contribution is modelled as a classical harmonic oscillator with an associated spring constant. The important terms in this calculation are the difference in bond length (b) from the equilibrium bond distance, b_{eq} and the bond force constant, K_b. To obtain accurate values of b_{eq}, data is collated from both experimental sources (X-ray crystal structures) and theoretical calculations (MP2 *ab initio* calculations and the AMBER protein FF). Further, to model accurate K_b values, an empirical rule is applied from reference tabulated force constants, to calculate the modified force constant in the GAFF force field, thus obtaining both the K_b and the bond stretching energy[213]:

$$K_b = K_{ij} \left(\frac{r_i}{r_j} \right)^6 \tag{4.5}$$

$$E_{\text{bond}} = \frac{1}{2} K_b (b - b_{eq})^2 \tag{4.6}$$

The angle bending energy contribution is likewise computed as a classical harmonic oscillator with an associated angle force constant. The crucial terms in the computation are the difference in bond angle (θ) from equilibrium bond angle, θ_{eq} and the angle force constant, K_θ. Whilst for bond stretching there exists a 2-point interaction, in this case there is a 3-point interaction. A similar empirical approach has been utilised to obtain accurate values for both K_θ and the angle bending energy[86]:

$$K_\theta = 143.9 Z_a C_b Z_c (r_{ab} + r_{bc})^{-1} (\theta_{abc})^{-2} e^{-2D} \tag{4.7}$$

$$E_{\text{angle}} = \frac{1}{2} K_\theta (\theta - \theta_{eq})^2 \tag{4.8}$$

The dihedral angle energy contribution is modelled as a sinusoidal curve which follows a modified cosine wave, with an associated dihedral force constant. The force constant, V_n, is actually the amplitude of the cosine wave. n is the dihedral periodicity and γ is the phase of the dihedral angle, ϕ. The continual iteration of this series of cosine waves leads to a Fourier series, describing energy contributions for a four point atomic system with proper torsions contributing to the motion around the bond and improper torsions accounting for planarity of the molecule[93]. The general form of this dihedral energy contribution accounts for both proper and improper dihedrals and may be modified based upon the number of terms per single dihedral calculation in the molecule, but the calculation of the dihedral energy remains empirically derived:

$$E_{\text{dihedral}} = V_n[1 + \cos(n\phi - \gamma)] \tag{4.9}$$

Non-bonded Interactions

These forces account for intermolecular interactions, often where atoms are separated by at least a distance of three atoms. The most computationally demanding part of any molecular simulation is the calculation of electrostatics. Coulomb's law is used to compute all electrostatic interactions between pairs of non-bonded atoms (i and j) (equation 4.10). Each non-bonded atom is assigned with a partial charge denoted by q, and the overall electron density in the system is the net sum of all these partial charges separated by an interaction distance r. The electric constant ϵ_0 and π are an additional term which correct for unit dimensions[112]. Coulomb's law describes long-range interactions which decay as $\frac{1}{r^2}$, with the potential energy being $\frac{1}{r}$. This decaying strength of charged interactions with distance ($d \propto \frac{1}{r}$) is more than counterbalanced by the increasing number of interactions ($\frac{1}{r^2}$). Therefore this calculation requires a special treatment called the Particle-Mesh-Ewald (PME). This method mitigates the problem by the inclusion of opposing Gaussian charges to the system, thus the effect of these (artificial) charges is now corrected. This method divides the net electrostatic potential into a direct and reciprocal part of

the Ewald summation. The direct summation is one of electrostatics in Cartesian space whilst the reciprocal sum is one of Poisson's electrostatics (equation 4.11) in Fourier space[59]. The Laplacian operator (∇^2) relates the electrostatic potential (ϕ) to the charge distribution ($-\frac{\rho_f}{\epsilon}$). Here, ρ signifies the electric charge density and ϵ is the permittivity of free space. Thus, the combination of these two equations combined accounts for both short and long range electrostatics:

$$E_{\text{electrostatics}} = \frac{1}{4\pi\epsilon_0}\frac{q_i q_j}{r_{ij}} \tag{4.10}$$

$$\nabla^2\phi = -\frac{\rho_f}{\epsilon} \tag{4.11}$$

The final term in the total potential energy of the system is the intermolecular van der Waals energy contribution. This intermolecular potential is usually represented as a 12-6 Lennard-Jones potential. This potential accounts for repulsion, attraction and dispersion of the intermolecular energies in the system. The typical 12-6 potential has a repulsion term, $1/r^{12}$, over short distances up to a minimum whilst the attraction term, $1/r^6$, over much longer distances and eventually becomes asymptotic along the x axis in Cartesian space[123]. For the GAFF force field, the 12-6 LJ potential seems to represent the repulsion and attraction well, however other modified potentials have been developed. The 9-6 LJ potential is far more compatible with the force fields COMPASS and PCFF, and this class of Mie potentials have been developed with varying exponent terms in the attractive and repulsive terms[88]. Within the 12-6 LJ potential, there is a minimum well depth, ϵ_0, a minimum distance, r_0, and an attractive term, $-1/r^6$ alongside a repulsive term, $1/r^{12}$, to give the functional form of the vdW term:

$$E_{\text{vdW}} = 4\epsilon_0\left[\left(\frac{r_0}{r}\right)^{12} - 2\left(\frac{r_0}{r}\right)^{6}\right] \tag{4.12}$$

4.2.2 Molecular Dynamics

Equations of Motion

The calculation of total potential energy from the force field can be directly used to compute the interacting forces between particles in a molecular dynamics simulation. The force $\boldsymbol{f}_i$ is calculated from the negative first order derivative of the potential energy ($E(\boldsymbol{r}_i)$) as a function of inter-atomic distance $\boldsymbol{r}_i$ accordingly:

$$\boldsymbol{f}_i = -\frac{\partial E(\boldsymbol{r}_i)}{\partial \boldsymbol{r}_i} \tag{4.13}$$

Once the force is calculated, Molecular Dynamics simulations involve solving Newton's equations of motion[85] by the precise integration of these equations to compute positions $\boldsymbol{r}_i$ and velocities $\boldsymbol{v}_i$ of 'N' interacting particles in a molecular trajectory:

$$\boldsymbol{F}_i = m_i \boldsymbol{a}_i = \frac{\partial \boldsymbol{p}_i}{\partial t} = m_i \frac{\partial \boldsymbol{v}_i}{\partial t} = m_i \frac{\partial^2 \boldsymbol{r}_i}{\partial t^2}, \tag{4.14}$$

for $i = 1,...,N$

Varied integration algorithms have been developed to compute these forces and calculate the velocities and positions of particles between discrete time intervals. These collectively use the finite difference method and the simplest of these is known as the Verlet algorithm. This simple algorithm updates the positions and accelerations of particles at time $t + \Delta t$ by calculating the positions $\boldsymbol{r}$ and accelerations $\frac{d^2\boldsymbol{r}(t)}{dt^2}$ of particles at time t and the former positions and accelerations of particles at time $t - \Delta t$ to give the overall equation:

$$\boldsymbol{r}(t + \Delta t) = 2\boldsymbol{r}(t) - \boldsymbol{r}(t - \Delta t) + \frac{d^2\boldsymbol{r}(t)}{dt^2}\Delta t^2 \tag{4.15}$$

A more advanced form of this algorithm is called the Leap-frog algorithm[2]. In this instance the positions $\boldsymbol{r}$ at time $t + \Delta t$ and the velocities $\frac{d\boldsymbol{r}(t)}{dt}$ at time $t - \frac{\Delta t}{2}$ are

required to calculate future positions and velocities in a similar manner to the Verlet algorithm. Overall these equations are complex but preferred in MD simulations:

$$\boldsymbol{r}(t+\Delta t) = \boldsymbol{r}(t) + \frac{d\boldsymbol{r}(t)}{dt}\left(t + \frac{\Delta t}{2}\right)\Delta t \tag{4.16}$$

$$\frac{d\boldsymbol{r}(t)}{dt}\left(t + \frac{\Delta t}{2}\right) = \frac{d\boldsymbol{r}(t)}{dt}\left(t - \frac{\Delta t}{2}\right) + \frac{d^2\boldsymbol{r}(t)}{dt^2}\Delta t \tag{4.17}$$

Thermodynamic Ensembles

Within the context of an MD simulation, in a well-defined periodic simulation box, an ensemble is a set of thermodynamic parameters which remain fixed whilst other parameters can vary. From a statistical mechanics perspective, an ensemble is a collection of configurations λ of the macroscopic state. The probability P_λ of finding a molecule with energy E_λ is equal to the fraction of the molecules n_λ divided by a collection of N molecules. Although the microscopic interactions between the atoms are identical in energy, the energy E of the macroscopic system is calculated by an ensemble average[77]. This average energy $\langle E \rangle$ is the product of the energy and probability, all summed over λ:

$$P_\lambda = \frac{n_\lambda}{N} \tag{4.18}$$

$$E = \sum_{\lambda=1} E_\lambda P_\lambda = \langle E \rangle \tag{4.19}$$

Each part of the MD simulation requires a specific ensemble based upon the equilibration and dynamics of the system. The key thermodynamic parameters considered in the simulation are the number of particles (N), the temperature (T), the volume (V) and the pressure (P). The minimisation step does not require an ensemble, as the kinetic energy of the system is equal to zero. The heating step requires a constant NVT (canonical) ensemble, where the box size is kept fixed as with the NVE ensemble, but the temperature is now constant with the addition of a thermostat. The equilibration and production steps require the use of a constant NPT

(isothermal-isobaric) ensemble where the box size is allowed to change and thus the average pressure remains constant. The thermostat is now coupled with a barostat to ensure NPT conditions are maintained[6].

The motivation for change in the ensemble from NVE to NVT is that longer simulations may cause a significant alteration in energy and thus the NVE ensemble becomes inaccurate at constant energy. The drawback with the NVT ensemble is that whilst temperature is constant, perturbations in the velocities of the particles during a longer simulation mean that temperature fluctuations have to be taken into account. Rescaling of the velocities can be achieved through the use of the Berendsen thermostat. This velocity rescaling utilises a scaling factor as a function of discrete timesteps Δt with time constant t_B in the simulation[143]. A reference temperature T_0 is divided by the actual temperature T in part of this equation. The scaling factor λ is then calculated within the system in the heat bath:

$$\lambda = \left[1 + \frac{\Delta t}{t_B}\left(\frac{T_0}{T} - 1\right)\right] \tag{4.20}$$

The main disadvantage with this thermostat occurs when simulations are extended to nanoseconds or longer. The identical scaling factor of particle velocities would not allow sufficient kinetic energy fluctuations as a function of time evolution. A more appropriate thermostat for longer simulations is the Nosé-Hoover thermostat. This thermostat enables sampling across all particle energies in the system. An additional friction term was included within their algorithm which provided a more accurate Boltzmann distribution of the particle energies. The fluctuations in the state variables don't affect non-equilibrium or equilibrium systems and thus makes this a more deterministic and simpler algorithm to apply[60]. The added friction constant, ζ, then transforms the Nosé-Hoover equations of motion, with scaled momentum ($\boldsymbol{p_i}$) and scaled time (t) to recalculate the force $\boldsymbol{F}_i$:

$$\boldsymbol{F}_i - \zeta \boldsymbol{p}_i = \frac{\Delta \boldsymbol{p}_i}{\Delta t} \tag{4.21}$$

Similarly, constant pressure can be maintained by the use of a barostat. Historically, the Berendsen barostat operates a linear scaling of atomic coordinates as a function of discrete timestep in the simulation. This results in a 1st order derivative of the pressure calculated from the difference in reference pressure (P_0) from the current pressure (P) divided by the pressure time constant τ_P:

$$\frac{dP}{dt} = \frac{P_0 - P}{\tau_P} \tag{4.22}$$

Whilst this barostat is fine for NPT ensembles during a shorter period of equilibration, the Berendsen barostat fails to produce the correct statistical distribution of volumes, and thus before gathering statistics, a different barostat must be employed for longer equilibrations. The so called Parrinello-Rahman barostat ensures minimal pressure deviations by the inclusion of an additional term to keep pressure constant. The independence of each unit vector within each box size of the simulation ensures the shape and volume of the box size can alter dramatically. This results in complete internal equilibration of the system by an applied external stress, accordingly with the formulae[105]:

$$\frac{d^2\boldsymbol{r}_i}{dt^2} = \frac{\boldsymbol{F}_i}{m_i} - M\frac{d\boldsymbol{r}_i}{dt} \tag{4.23}$$

$$\frac{db^2}{dt^2} = VW^{-1}b'^{-1}(P - P_{ref}) \tag{4.24}$$

In equation 4.24, b represents a matrix for the calculation of the box vector, W is also a matrix for the calculation of the system mass, which ultimately computes the coupling strength. The pressure term results from subtraction of a reference pressure P_{ref} from the current pressure P.

Periodic Boundary Conditions

In MD simulations, the need of a periodic boundary condition (pbc) is critical when considering large molecular structures. pbc are important as they represent the

physical repeating unit cell of a simulation box which has a well defined 3D structure in Cartesian space and thus a finite molecular size across the x, y and z dimensions. Visually, a single simulation box is surrounded by identical repeating simulation boxes within the 3D Cartesian space of the simulation. The advantage of this method, coupled with the minimum image convention (MIC), ensures that if an atom leaves one face of the box, it will re-enter the identical box on the parallel or opposite side, thus maintaining the atomic arrangement. The box needs to be large enough for $g(r)$ values to converge. With this method, homogeneity is assumed across the molecular system[27].

4.2.3 Analysis Techniques

Mass Density Profiles

Partial density is a physical measure of the distribution of molecular mass across the membrane bilayer as a function of axis, laterally often the z axis. In all classical systems, the partial density profile follows an almost identical trend to the electron density profile where the highest electron density is found in the most dense areas of the plot. This calculation typically involves two parts. First, the atomic coordinates of the centre of mass (COM) of the membrane bilayer are calculated by an average of the COM of atomic coordinates on each bilayer leaflet. The atomic coordinates of all other molecules are then calculated by the same method and referenced against the membrane bilayer COM. The smooth nature of the density profile implies a fairly homogeneous medium, which of course is an approximation, but predicts clearly the segregation of the inner hydrophobic core of the membrane from the outer hydrophilic layer dominated by polar lipid headgroups. This analysis is pertinent in determining preferred localisation and position of the drug molecule in the membrane bilayer and thus informing on potential reactivity[141].

Radial Distribution Function Profiles

The radial distribution function ($g(r)$) is a statistical quantity which measures the average local density change as a function of distance from a set reference particle.

Often in a radial pair distribution, the density change is calculated between two set particles rather than a collection of particles. This quantity is calculated by considering the difference between a number of particles at distance r relative to a distance $r + \Delta r$ and thus computing the pair potential. This pair distance is calculated from the COM of a set atom relative to another set atom to inform on relative distance between the two atoms and thus potential for reactivity. Also, RDFs have the ability to identify potential hydrogen bonds formed between HB donors and HB acceptors. This is useful in drug-lipid simulations for showing the potential interactions between the hydrophilic headgroups and the drug molecule in stabilising reactive intermediates. Graphically, if the peak appears at a shorter distance than the cut-off for HB formation, there is a high probability that a hydrogen bond could be formed. Furthermore, the height of the peak could be an indication of a strongly localised interaction[46]. Mathematically, $g(r)$ is defined as:

$$\rho(r) = \rho g(r) \tag{4.25}$$

This is computed as the probability of finding a particle $\rho(r)$ relative to the average density ρ of the surrounding particles. Graphically this is represented as a histogram with bin width of dr.

Potential of Mean Force

The potential of mean force is a specific free energy calculation which calculates the change in free energy of a drug molecule as it traverses a membrane bilayer across a defined axis; laterally along the z axis. This quantity helps define the stability and localisation of the drug molecule in the membrane by examining where the thermodynamic minimum can be found in the free energy profile. This quantity also provides thermodynamic evidence of the partitioning of the molecule in the membrane and thus can be used to locate the most likely site of chemical reactivity between the drug molecule and lipid.

The PMF is also versatile in being able to be verified through the experimentally

derived partition coefficient and the position dependent diffusion coefficient ($D(z)$) gaining molecular insights into both the partitioning behaviour and permeability of the molecule across the membrane[137]. The main setback with this method in atomistic molecular dynamics is the often poor sampling that arises as a result of this calculation. This can be somewhat mitigated by increasing the simulation time per window of the PMF calculation, but usually is significantly improved by a technique called Umbrella Sampling. This techniques improves sampling of the PMF calculation by the inclusion of biasing potentials. Each sampling region, or window, is enforced with a biasing potential which improves the individual sampling. Through this method a probability distribution is obtained which accounts for all sampling windows. These individual distributions are then superimposed to ensure full overlap between each window which creates good sampling[81].

The PMF is computed from the force (F) acting on the drug molecule's COM at a defined partitioning depth along a chosen coordinate axis as a function of timesteps in the simulation[19]. The free energy (ΔG) as a function of coordinate (z) is then computed as:

$$\Delta G(z) = -\int_{\text{outside}}^{z} \langle F(z') \rangle_t dz' \tag{4.26}$$

$$\Delta G(z) = -RT \ln K(z) \tag{4.27}$$

In equation 4.26, the average force $\langle F(z') \rangle_t$ applied to the molecule as a function of its depth in the membrane (z') is integrated along the z axis from the aqueous region (outside) to a defined depth (z). equation 4.27 shows an alternative calculation of the free energy from the partition coefficient ($K(z)$)[152]. Here, the partition coefficient is computed from the average mass density of the chosen molecule after normalisation of the density, with R being the ideal gas constant and T being the temperature. Either of these equations are used in PMF calculations to accurately compute the free energy.

4.2.4 Simulation Details

Atomistic simulations were carried out using the AMBERtools23 molecular dynamics software package for all drug-membrane systems. Small organic molecules were energy minimised with AM1 and PM3 semi-empirical force fields within SCIGRESS software. CHARMM-GUI software was utilised to construct 1-palmitoyl-2-oleoyl-*sn*-glycero-3-phosphocholine (POPC) lipids, 64 lipids on the upperleaflet and 64 on the lowerleaflet. Cholesterol was added to the system and appropriately solvated. Membrane models were obtained with and without cholesterol. AM1-BCC partial charges of all individual atoms were obtained from the AmberTools23 package. This program compared atomic equivalences based upon bond and atomic types. The tLEaP command was used to assign the appropriate AMBER force fields to all chemical types: GAFF2 for the organics and cholesterol, LIPID17 for the lipids, TIP3P for water. Ions were added with the addions2 command in AmberTools to ensure a net neutral charge for the system. Once fully equilibrated bilayers had been obtained, molecular dynamics simulations could be employed.

The system was first energy minimised (2,000 time-steps) to remove any potential close contacts, heated (0 to 310 K over 20 ps) within a constant NVT ensemble, pre-equilibrated (310 K) also within a constant NVT ensemble to ensure simulation stability; after which a series of longer runs were employed. The first involved a critical equilibration step to enable the drug molecule to be brought to room temperature within the membrane lipids. This was conducted for 20 ps, the second for 5 ns and the final production run for 200 ns. Each production step was carried out within the constant NPT ensemble framework. Constant NPT conditions are more favourable than NVT for comparison with experimental data (usually the latter are conducted at constant pressure). To maintain a constant temperature and pressure in the longer production runs, a Nosé-Hoover thermostat and Berendsen barostat were utilised as viable thermodynamic control parameters, with periodic boundary conditions being conserved throughout all simulations.

PMF calculations were undertaken, through umbrella sampling, to predict the preferred location of each drug molecule in the bilayer. A biased potential was used to

pull each drug molecule from the bilayer centre out to the aqueous region under a pre-defined force constant. The pulling force was defined as a function of distance between the centre of mass of the drug molecule and the centre of mass of the N31 lipid headgroups. The total system was simulated for 50 ns from 0 to 50 Å at a pulling force constant of 1100 kJ mol^{-1} nm^2; at a constant pulling rate of initially 2.5 and then reduced to 1.1 Å ns^{-1}, across the z axis. To improve convergence and enable good sampling, 50 individual windows were extracted from the output file, with 1 Å spacing between each frame. To optimise convergence, the simulation length per frame was increased from 5 to 50 ns with 1 Å spacing between each frame. Furthermore, the z position was reduced from being written every 1000^{th} step to every step, thus capturing more of the free energy fluctuations per frame. Upon completion, the so called 'WHAM' method[79] was utilised to construct the free energy profile and its associated histogram (as a measure of sampling quality).

4.3 Results and Discussion

4.3.1 Mass Density Profiles

All unbiased molecular dynamics simulations were run for 200 ns in the NPT ensemble at a constant temperature of 310 K (physiological conditions). A python script was written which reoriented each drug molecule to be positioned in the centre of the periodic box. This calculation was essentially a difference between the centre of mass of the atomic coordinates of the membrane and the centre of mass of the atomic coordinates of the drug molecule. The simulation box dimensions were constant throughout at 78.756×75.599×84.077 for all drug-membrane simulations. Periodic boundary conditions were maintained throughout all simulations and the longest production run of 200 ns was sufficient in length to capture the essential physics and dynamics of each drug-lipid system.

As can be seen in figure 4.2, the peak height of each drug molecule was a measure of the molecular weight of each drug. Fluoxetine, having the highest molecular weight, had the highest partial density at 650 kg m^{-3}, with its protonated analogue having almost the exact same density. Whereas for phentermine, which has the lowest

molecular weight, its peak height is only at 350 kg m^{-3}, with propranolol having a peak height of about 570 kg m^{-3}. Cholesterol was inclusive in all these membranes to add lipid rigidity and provide structural support for the membrane. Water, at pH 7, was present in all membrane systems to ensure a realistic physiological model of the cell membrane and the 3 major lipid sections of POPC were present: palmitic acid (PA), oleic acid (OL) and phosphatidylcholine (PC). The lipid tails are extremely hydrophobic and thus remain close to the centre of the simulation box, thus the highest density of OL and PA lipids are found near the centre, whilst the lipid headgroups are very hydrophilic and thus point outwards and are at the interface between the aqueous region and the membrane. Hence, the PC lipids were at their highest density at the membrane-water interface.

As is evident from the simulation plots, all neutral drug molecules preferred to position themselves at the interface, close to the PC lipid headgroups, whereas all protonated drug molecules preferred to position themselves closer to the water region (figure 4.2). The centre of the simulation box is at 3.8 nm, with the drug fluoxetine being centred at 4.9 nm; which is in the interfacial region between the lipid headgroups and tails, with the PC headgroups at 5.6 nm. As fluoxetine exhibits lipid hydrolysis by acting as a phase transfer catalyst, the protonated form is actually more crucial in terms of its positioning in the membrane. The protonated form must be deprotonated by the water molecules before undertaking hydrolytic activity. It is centred at 5.2 nm sitting slightly further into the headgroup region. The protonated form of propranolol is known to be more likely to undergo direct lipidation than the neutral form. It also sits closer in distance to the centre of the lipid headgroups than the neutral form (table 4.1). Phentermine exhibits virtually no reactivity with the membrane lipids, and sits closer to the lipid tails than either fluoxetine or propranolol. 2-aminomethylbenzimidazole undergoes direct lipidation, whilst its protonated form sits closer to the centre of the headgroup distribution.

Ultimately, these mass density profiles suggest that positioning of the drug molecule is central to their reactivity. Propranolol and 2-aminomethylbenzimidazole, which localise close to the membrane-water interface where there are a high abundance of

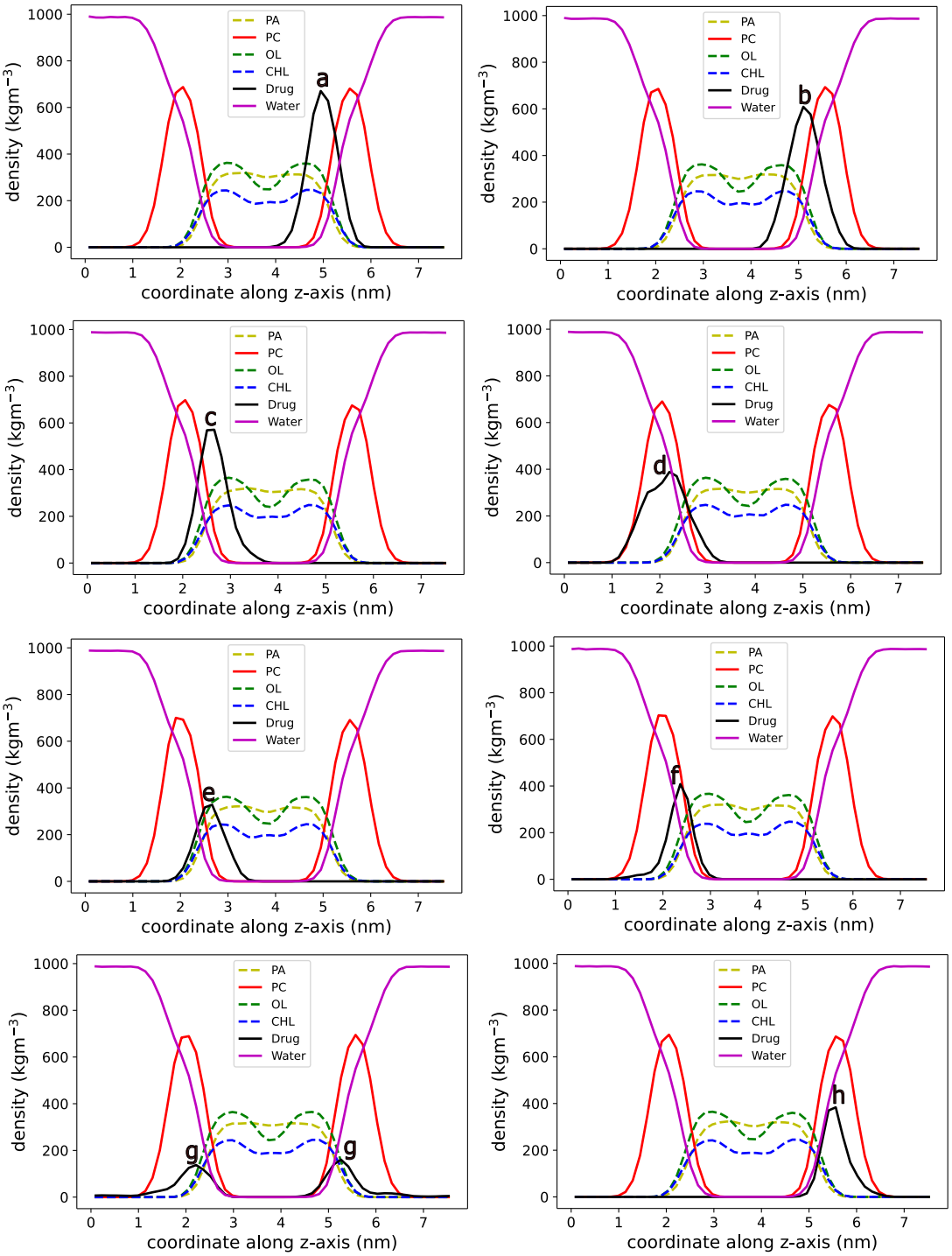


Figure 4.2: Partial density plots for (a) fluoxetine, (b) protonated fluoxetine, (c) propranolol, (d) protonated propranolol, (e) phentermine, (f) protonated phentermine, (g) 2-aminomethylbenzimidazole, (h) protonated 2-aminomethylbenzimidazole in a solvated POPC membrane bilayer

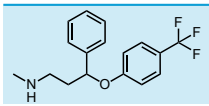
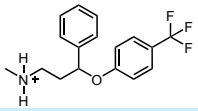
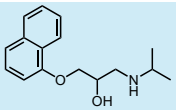
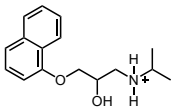
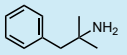
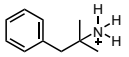
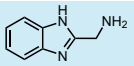
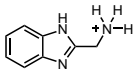
Chemical Structure	Drug Name	PC Lipid position / nm	Drug position relative to the PC Lipid / nm
	Fluoxetine	**5.56**	**-0.61**
	Protonated fluoxetine	**5.56**	**-0.46**
	Propranolol	**2.05**	**+0.53**
	Protonated propranolol	**2.05**	**+0.16**
	Phentermine	**1.98**	**+0.61**
	Protonated phentermine	**1.98**	**+0.39**
	2-aminomethylbenzimidazole	**5.57**	**-0.3**
	Protonated 2-aminomethylbenzimidazole	**5.57**	**-0.09**

Table 4.1: Comparison of PC lipid position and drug position relative to PC lipid for neutral and protonated CAD molecules, from the membrane centre (3.81 nm)

PC lipids, are setup well for reactivity, whilst the other CAD molecules (fluoxetine and phentermine) exhibit lipid hydrolysis and no reactivity respectively. These latter two molecules are located further from the centre of the lipid headgroups.

A quantitative measure of the position and localisation of the drug molecule in the membrane was then applied to a series of additional molecules. Firstly a series of indoles, indazoles, benzimidazoles and triazoles were positioned in the centre of the membrane and molecular dynamics simulations were run to understand how they localised in the membrane. As can be seen in the partial density plots for all compounds, the neutral molecules preferred to orient themselves at the interface, close to the PC lipid headgroups, whereas the protonated analogues preferred to position themselves closer to the aqueous region where deprotonation would be more likely.

Although these molecules have the potential for direct lipidation with membrane lipids, their orientation along the z axis in the membrane is not directionally pre-

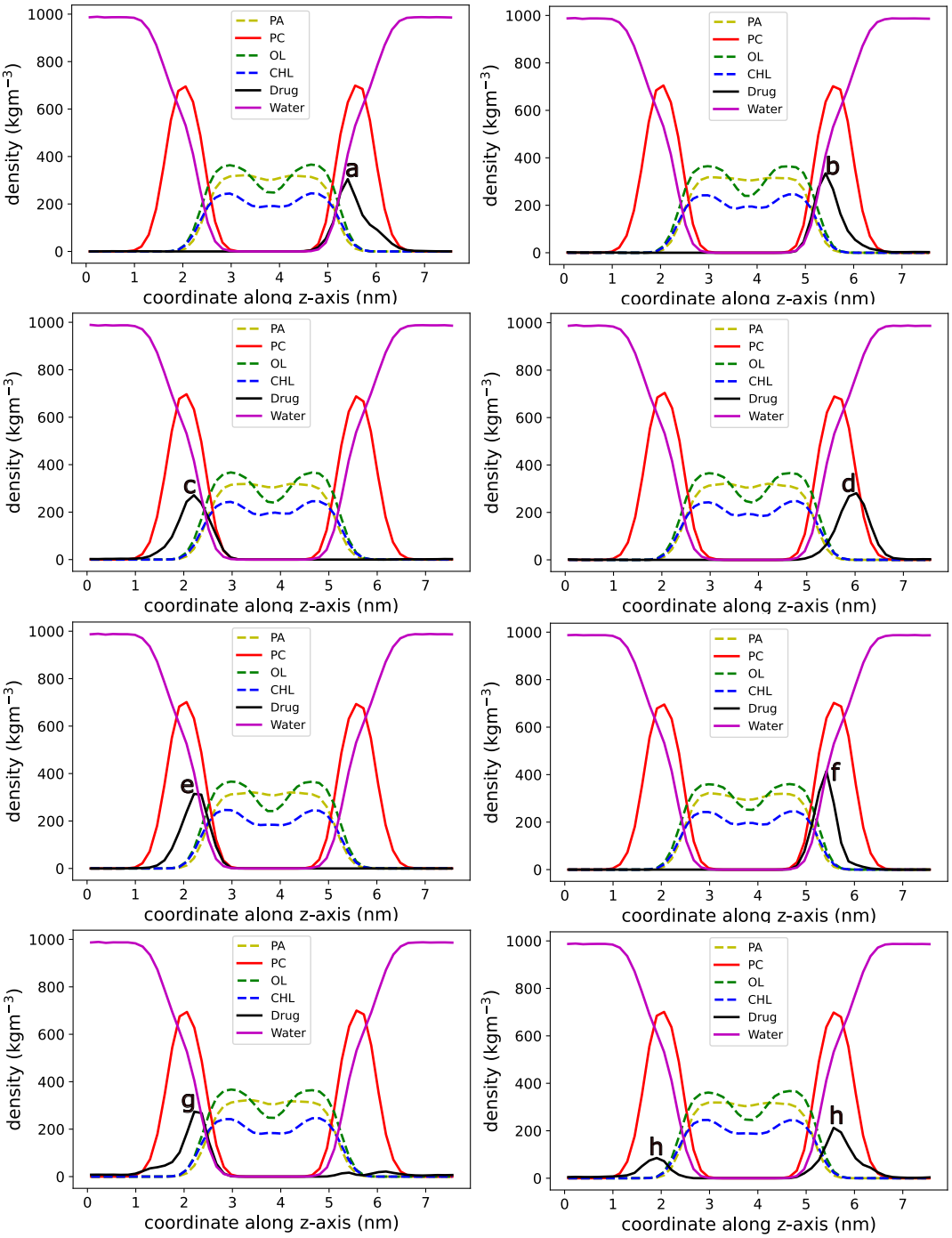


Figure 4.3: Partial density plots for (a) 5-aminoindole, (b) protonated 5-aminoindole, (c) 3-aminoindazole, (d) protonated 3-aminoindazole, (e) 5-aminoindazole, (f) protonated 5-aminoindazole, (g) 5-aminobenzimidazole, (h) protonated 5-aminobenzimidazole in a solvated POPC membrane bilayer

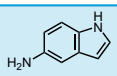
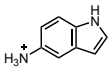
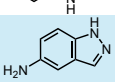
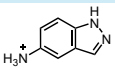
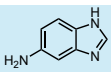
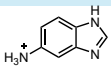
Chemical Structure	Drug Name	PC Lipid position / nm	Drug position relative to the PC Lipid / nm
	5-aminoindole	**5.57**	**-0.15**
	Protonated 5-aminoindole	**5.57**	**-0.14**
	3-aminoindazole	**2.06**	**+0.15**
	Protonated 3-aminoindazole	**6.03**	**-0.46**
	5-aminoindazole	**2.06**	**+0.24**
	Protonated 5-aminoindazole	**5.58**	**-0.16**
	5-aminobenzimidazole	**2.06**	**+0.23**
	Protonated 5-aminobenzimidazole	**2.06**	**-0.15**

Table 4.2: Comparison of PC lipid position and drug position relative to PC lipid for neutral and protonated CAD molecules, from the membrane centre (3.81 nm)

ferred. Previous experimental data with DOPC membranes[25] implicates no observed lipidation for 5-aminoindole, 5-aminoindazole, 5-aminobenzimidazole and 4-aminobenzotriazole. However, there is evidence for some lipidation present for 3-aminoindazole with a DOPC membrane. Experimental conditions were maintained at 37°C with pH 7.4, close to the conditions for the computational model at 310 K and pH 7 (physiological conditions). Chemically, lipidation in DOPC is comparable with POPC. A small intensity/trace in the LC-MS spectrum indicates some lipidation present in the DOPC membrane[25]. This is further correlated with the peak (figure 4.3) for the 3-aminoindazole distribution being closer to the PC lipid headgroups for this compound than the other indoles, indazoles and benzimidazoles (table 4.2). POPC provides a good comparison with DOPC as they share a common lipid headgroup.

For the next set of compounds localised in the membrane, a different trend is exhibited. These methyl-substituted indazoles are chemically identical, apart from the amino group being positioned at different points in the aromatic ring (see table 4.3). For all 5 compounds, their protonated analogues orient themselves closer to the PC lipid headgroups than their neutral counterparts (figure 4.4). Computationally, this

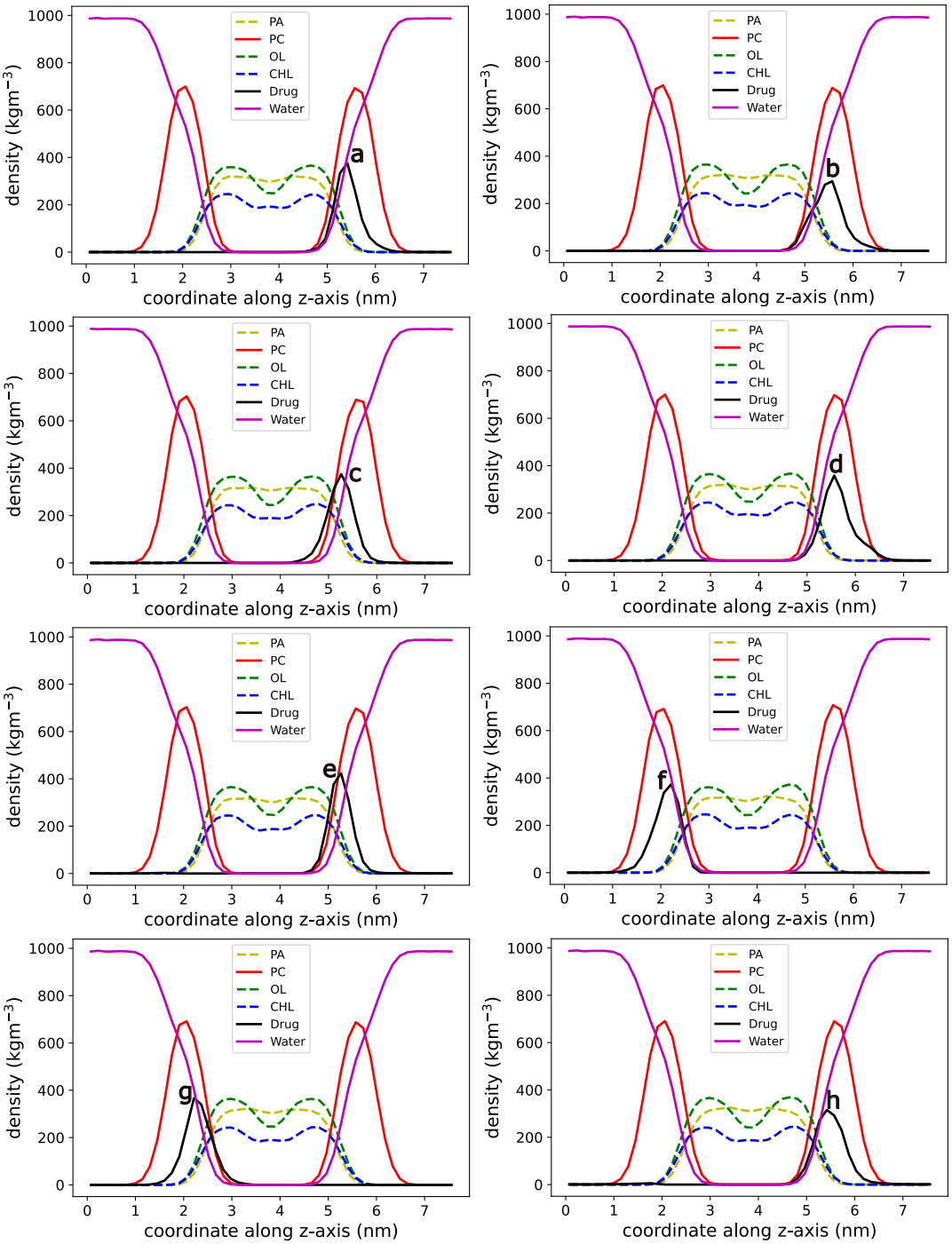


Figure 4.4: Partial density plots for (a) 4-aminobenzotriazole, (b) protonated 4-aminobenzotriazole, (c) 3-(aminomethyl)-1*H*-indazole, (d) protonated 3-(aminomethyl)-1*H*-indazole, (e) 4-(aminomethyl)-1*H*-indazole, (f) protonated 4-(aminomethyl)-1*H*-indazole, (g) 5-(aminomethyl)-1*H*-indazole, (h) protonated 5-(aminomethyl)-1*H*-indazole in a solvated POPC membrane bilayer

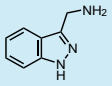
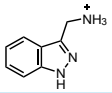
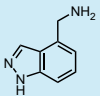
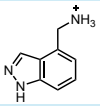
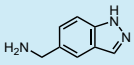
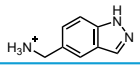
Chemical Structure	Drug Name	PC Lipid position / nm	Drug position relative to the PC Lipid / nm
	4-aminobenzotriazole	**5.57**	**-0.15**
	Protonated 4-aminobenzotriazole	**5.57**	**0**
	3-(aminomethyl)-1H-indazole	**5.57**	**-0.3**
	Protonated 3-(aminomethyl)-1H-indazole	**5.57**	**0**
	4-(aminomethyl)-1H-indazole	**5.57**	**-0.3**
	Protonated 4-(aminomethyl)-1H-indazole	**2.06**	**+0.16**
	5-(aminomethyl)-1H-indazole	**2.06**	**+0.16**
	Protonated 5-(aminomethyl)-1H-indazole	**5.57**	**-0.15**

Table 4.3: Comparison of PC lipid position and drug position relative to PC lipid for neutral and protonated CAD molecules, from the membrane centre (3.81 nm)

would imply that the protonated molecules have a better geometry (nucleophilic distance and angle from the drug to the lipid carbonyl carbon) for reactivity than their neutral counterparts. In fact experimental data, again with DOPC membranes, implies a high level of lipidation for all these molecules. Thus, whilst the protonated forms are closer to the interface, they also position themselves closer to the aqueous region. They can therefore undergo deprotonation by the water molecules and then act as molecules in direct lipidation with the POPC membrane lipids.

The next set of computational models tested were for a series of increasingly substituted indoles in the POPC membrane (table 4.4). Experimental results for these molecules in the DOPC membrane shows no lipidation present for 3-aminoindole, 3-(2-aminoethyl)-indole and 3-(*N*-ethyl-amino)-methylindole. The highly substituted amine in 3-(*N*-ethyl-amino)-methylindole is chemically dominated by steric effects from the bulky ethyl group attached to the amine. This would reduce the potential for direct lipidation reactivity from the NH to the lipid carbon. Interestingly, the 3-(aminomethyl)indole compound, in agreement with the literature, shows ev-

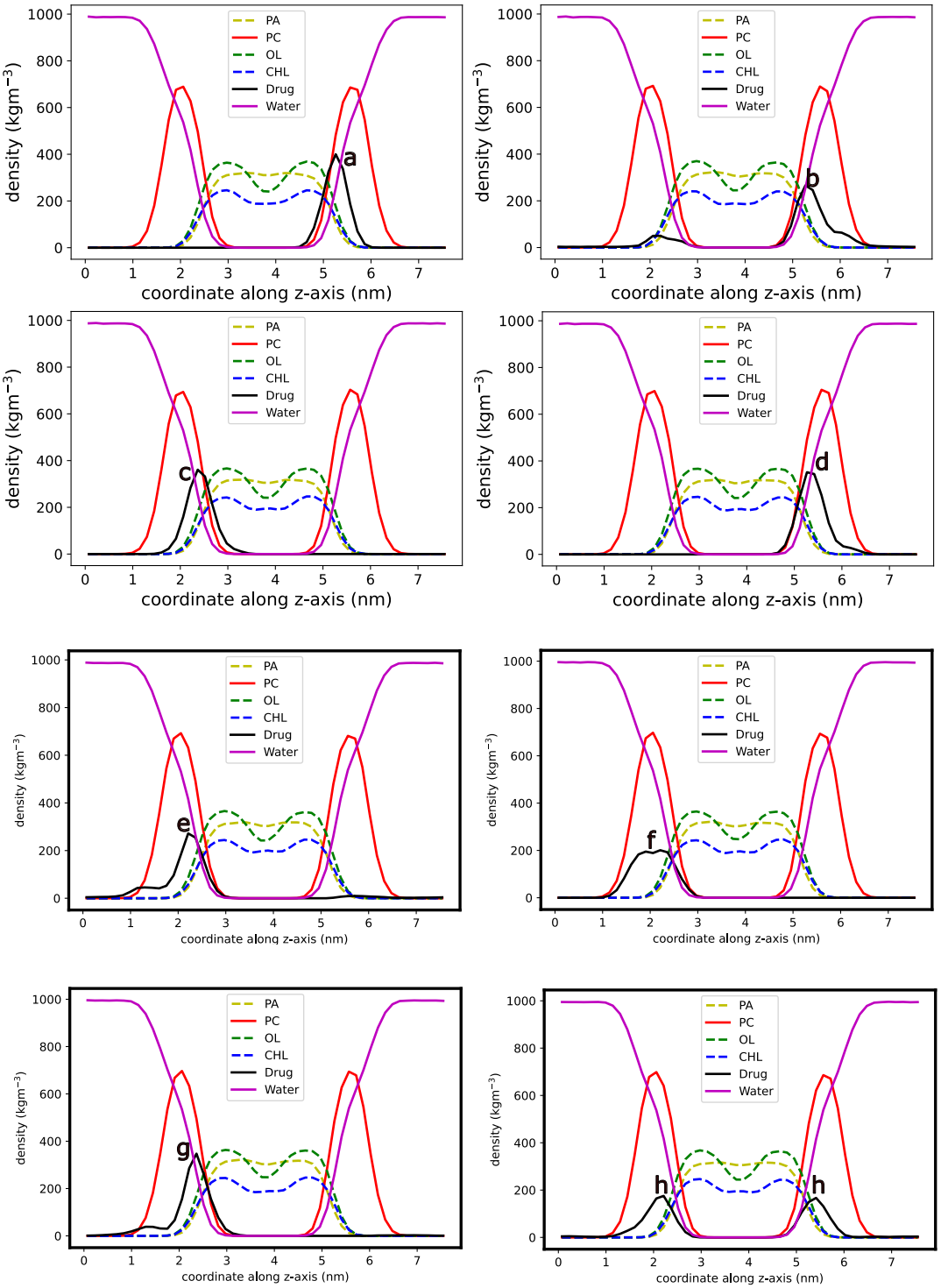


Figure 4.5: Partial density plots for (a) 6-(aminomethyl)-1*H*-indazole, (b) protonated 6-(aminomethyl)-1*H*-indazole, (c) 7-(aminomethyl)-1*H*-indazole, (d) protonated 7-(aminomethyl)-1*H*-indazole, (e) 3-aminoindole, (f) protonated 3-aminoindole, (g) 3-(aminomethyl)-indole, (h) protonated 3-(aminomethyl)-indole in a solvated POPC membrane bilayer

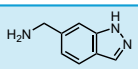
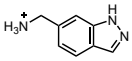
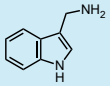
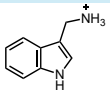
Chemical Structure	Drug Name	PC Lipid position / nm	Drug position relative to the PC Lipid / nm
	6-(aminomethyl)-1H-indazole	**5.56**	**-0.3**
	Protonated 6-(aminomethyl)-1H-indazole	**5.56**	**-0.29**
	7-(aminomethyl)-1H-indazole	**2.06**	**+0.32**
	Protonated 7-(aminomethyl)-1H-indazole	**5.57**	**-0.3**
	3-aminoindole	**2.06**	**+0.16**
	Protonated 3-aminoindole	**2.06**	**0**
	3-(aminomethyl)-indole	**2.06**	**+0.3**
	Protonated 3-(aminomethyl)-indole	**2.06**	**+0.16**

Table 4.4: Comparison of PC lipid position and drug position relative to PC lipid for neutral and protonated CAD molecules, from the membrane centre (3.81 nm)

idence for membrane positioning being correlated with lipidation in POPC (figure 4.5). This is corroborated with the former results for all substituted indazoles, where the aminomethyl functionality on the ring results in direct lipidation with the membrane. Thus, chemical functionality between indoles or indazoles seems to be independent of potential reactivity with the membrane, more so the extent of substitution on the ring system. Further evidence of this is that the drug density of 3-(aminomethyl)indole is closest to the PC lipid headgroups when compared to the other 3 systems (table 4.4).

The next set of compounds display a similar trend in reactivity. As for 2-aminomethyl benzimidazole, where this benzimidazole substituted compound shows high lipidation, the 2-aminomethylindole (table 4.5) analogue still exhibits lipidation, but at a smaller rate. Its neutral form is localised closer to the PC lipid headgroup of POPC, and thus can undergo direct lipidation independent of its protonated form. Experimentally, in DOPC membranes, there is a smaller LC-MS intensity trace for this compound, relative to 2-aminomethylbenzimidazole, thus correlating well

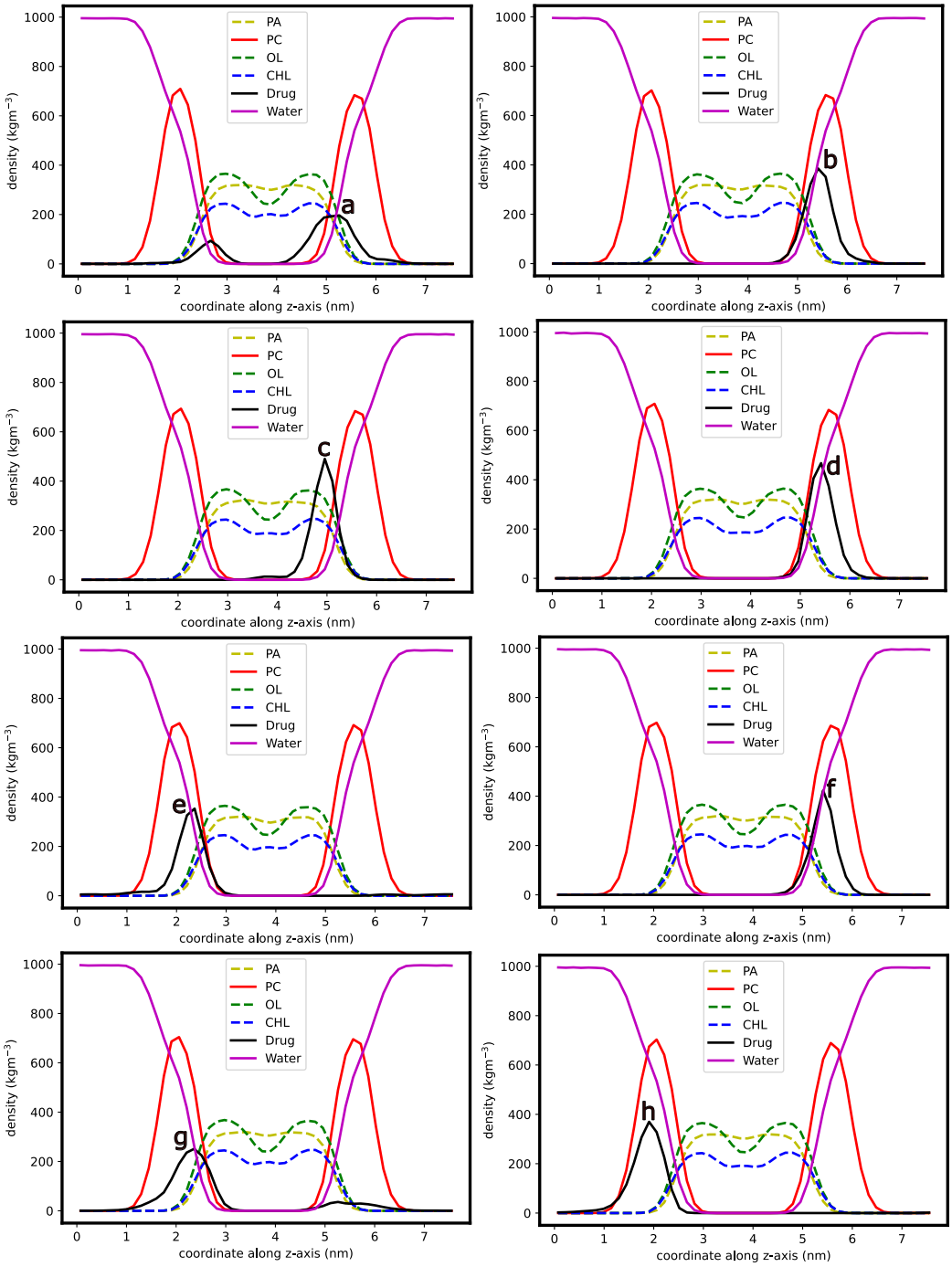


Figure 4.6: Partial density plots for (a) 3-(1*H*-indol-3-yl)propanamine, (b) protonated 3-(1*H*-indol-3-yl)propanamine, (c) *N*-indol-3-ylmethylethylamine, (d) protonated *N*-indol-3-ylmethylethylamine, (e) 2-aminomethylindole, (f) protonated 2-aminomethylindole, (g) 5-methoxy-1*H*-indol-4-amine, (h) protonated 5-methoxy-1*H*-indol-4-amine in a solvated POPC membrane bilayer

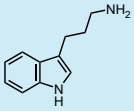
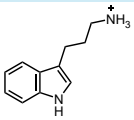
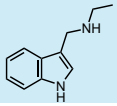
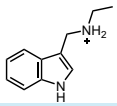
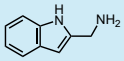
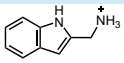
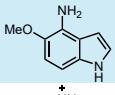
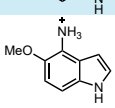
Chemical Structure	Drug Name	PC Lipid position / nm	Drug position relative to the PC Lipid / nm
	3-(1H-indol-3-yl)propan-1-amine	**5.57**	**-0.44**
	Protonated 3-(1H-indol-3-yl)propan-1-amine	**5.57**	**-0.15**
	N-indol-3-ylmethylethylamine	**5.56**	**-0.6**
	Protonated N-indol-3-ylmethylethylamine	**5.57**	**-0.15**
	2-aminomethylindole	**2.06**	**+0.31**
	Protonated 2-aminomethylindole	**5.57**	**-0.15**
	5-methoxy-1H-indol-4-amine	**2.06**	**+0.3**
	Protonated 5-methoxy-1H-indol-4-amine	**2.06**	**-0.15**

Table 4.5: Comparison of PC lipid position and drug position relative to PC lipid for neutral and protonated CAD molecules, from the membrane centre (3.81 nm)

with theoretical predictions of this molecule in the membrane. Experimentally, 5-methoxy-1*H*-indol-4-amine (figure 4.6) exhibits no reactivity with either DOPC or POPC membrane lipids. A cause of this may be due to the bulky methoxy group *ortho* to the nucleophilic amine group. This essentially could negate the effect of lipidation from the free amine and thus result in no reactivity.

4-Amino-*N*-phenylbutanamide displays a remarkably high level of direct lipidation, as confirmed experimentally with DOPC lipids. Unlike with propranolol and 2-aminomethylbenzimidazole, this compound results in no lysolipid formation once reacted with membrane lipids. Its maximum in the density plot is closer to the centre of the membrane than its protonated analogue (figure 4.7), implying a good rate of lipidation and localised relatively close to the PC lipid headgroup. Serotonin is a monoamine neurotransmitter chemical and exhibits some lipidation with POPC membrane lipids, as exemplified with experimental data in DOPC lipids. Its neutral form positions itself closer to the lipid-water interface than its protonated analogue and thus has a higher potential to act as a nucleophile in direct lipidation (table 4.6).

The last set of molecules also exhibit lipidation behaviour. The substituted sulfonated compound experimentally with DOPC demonstrates some lipidation, and this is verified computationally with the drug density being localised close to the PC lipid headgroups in POPC (figure 4.7). The free amine is not affected by sterics and thus has the potential to undergo some direct lipidation. Procaine is a local anaesthetic which results in a high lysolipid formation experimentally in DOPC membranes, and thus a good candidate for direct lipidation in membranes. Computationally, this is exemplified with good positioning relative to the PC lipid headgroups in the POPC membrane and thus can undergo direct lipidation with the appropriate nucleophilic distance and angle of attack to confirm reactivity (table 4.6).

Tetracaine, on the other hand, shows no reactivity with the POPC membrane. Although both its neutral and protonated forms position themselves close to the membrane interface where there is an abundance of PC lipids (figure 4.8), the absence of

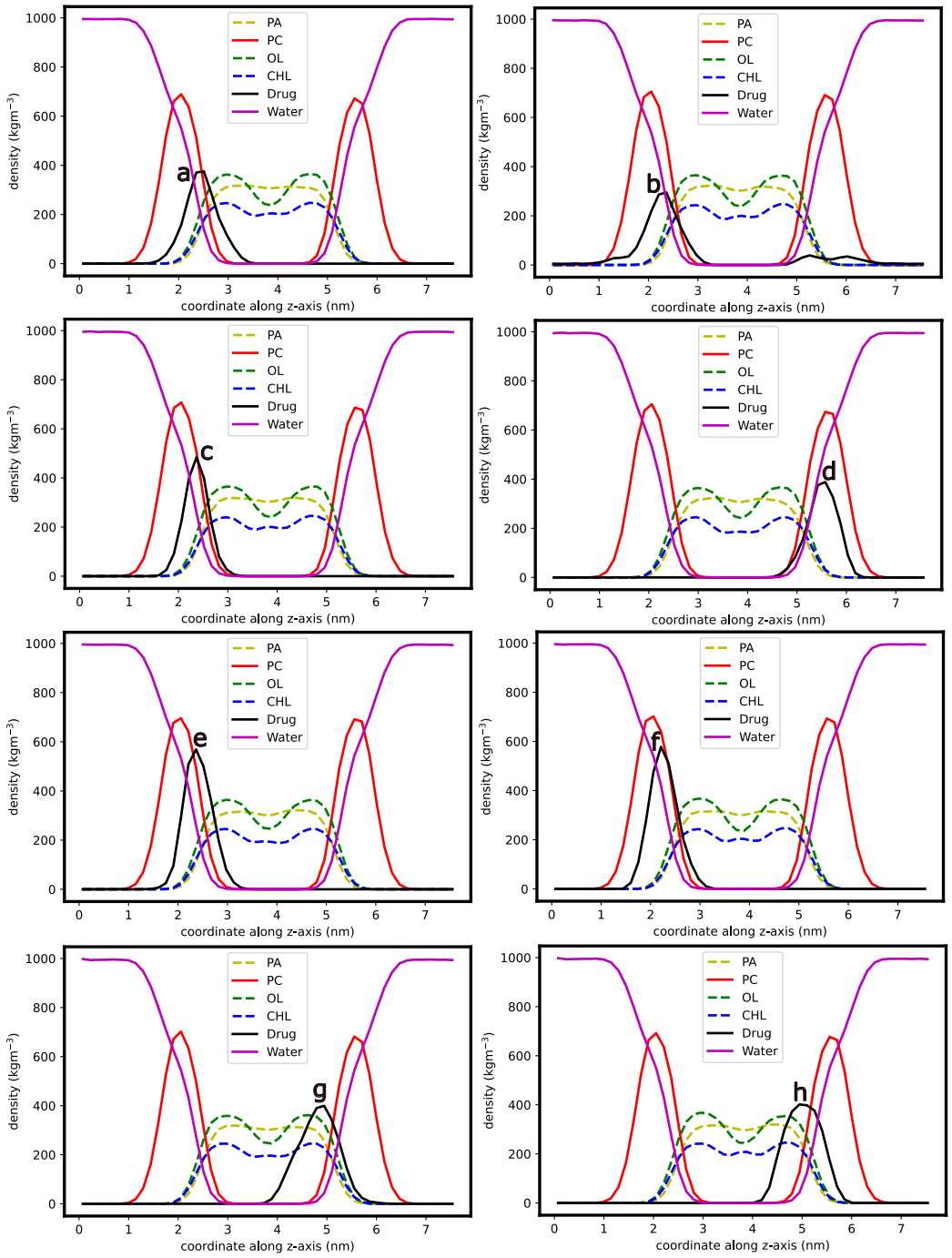


Figure 4.7: Partial density plots for (a) 4-amino-*N*-phenylbutanamide, (b) protonated 4-amino-*N*-phenylbutanamide, (c) serotonin, (d) protonated serotonin, (e) *N*-(4-aminophenyl)-1,1,1-trifluoromethanesulfonamide, (f) protonated *N*-(4-aminophenyl)-1,1,1-trifluoromethanesulfonamide, (g) procaine, (h) protonated procaine in a solvated POPC membrane bilayer

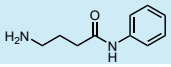
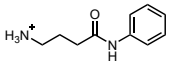
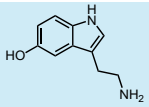
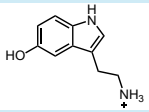
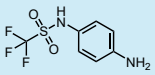
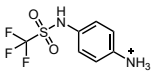
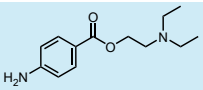
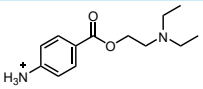
Chemical Structure	Drug Name	PC Lipid position / nm	Drug position relative to the PC Lipid / nm
	4-amino-N-phenylbutanamide	**2.06**	**+0.38**
	Protonated 4-amino-N-phenylbutanamide	**2.06**	**+0.23**
	Serotonin	**2.06**	**+0.31**
	Protonated Serotonin	**5.57**	**-0.07**
	N-(4-aminophenyl)-1,1,1-trifluoromethanesulfonamide	**2.06**	**+0.3**
	Protonated N-(4-aminophenyl)-1,1,1-trifluoromethanesulfonamide	**2.06**	**+0.16**
	Procaine	**5.56**	**-0.61**
	Protonated Procaine	**5.55**	**-0.52**

Table 4.6: Comparison of PC lipid position and drug position relative to PC lipid for neutral and protonated CAD molecules, from the membrane centre (3.81 nm)

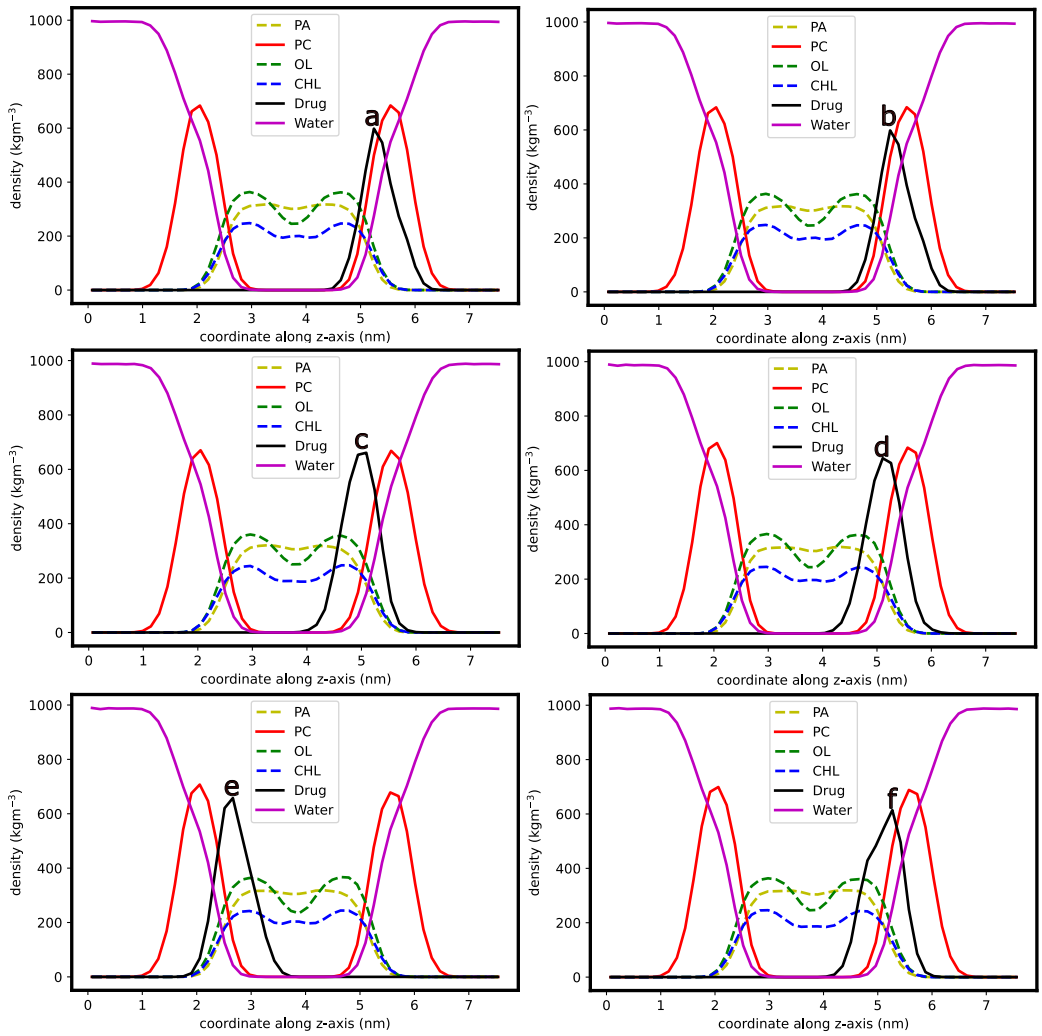


Figure 4.8: Partial density plots for (a) tetracaine, (b) protonated tetracaine, (c) chloroquine, (d) protonated chloroquine, (e) sertraline, (f) protonated sertraline in a solvated POPC membrane bilayer

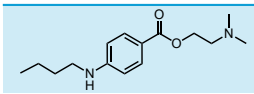
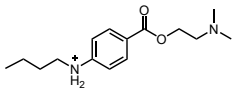
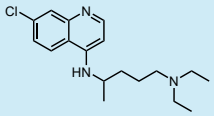
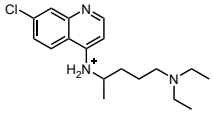
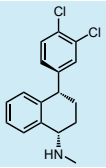
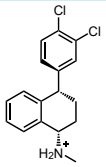
Chemical Structure	Drug Name	PC Lipid position / nm	Drug position relative to the PC Lipid / nm
	Tetracaine	**5.55**	**-0.34**
	Protonated Tetracaine	**5.55**	**-0.3**
	Chloroquine	**5.54**	**-0.52**
	Protonated Chloroquine	**5.57**	**-0.46**
	Sertraline	**2.06**	**+0.61**
	Protonated Sertraline	**5.58**	**-0.3**

Table 4.7: Comparison of PC lipid position and drug position relative to PC lipid for neutral and protonated CAD molecules, from the membrane centre (3.81 nm)

an unsubstituted amine results in no potential reactivity with the membrane lipids. Furthermore, a high level of substitution causes low reactivity at a chemical level. Experimentally, the drug chloroquine results in a similar lysolipid formation as with propranolol in POPC membranes. Computationally, the density plot shows that the distribution of this drug in the membrane is not close to the PC lipid headgroups (figure 4.8), and thus does not undergo direct lipidation. The absence of any potential nucleophilic amine groups precludes this molecule from reactivity at the membrane interface. Sertraline, experimentally with POPC, results in a higher lysolipid formation than chloroquine, but again exhibits no reactivity at the membrane interface. Being chemically and structurally similar to fluoxetine, sertraline does exhibit lipid hydrolysis behaviour, and its protonated form prefers to localise closer to the aqueous region where the molecule can be more likely deprotonated (table 4.7).

4.3.2 Radial Distribution Function Profiles

RDF calculations provide a prediction of potential reactivity by looking at the distribution of distances between two reactive atoms. They are a measure of the average distance change from a reference set of atoms, and thus can be directly extracted from simulation trajectories to confirm distance, angle and geometry of the molecule for nucleophilic attack and subsequent chemical reactivity.

For propranolol, intuitively the nitrogen atom is a more nucleophilic site for reactivity. However, from the RDF plot in figure 4.9, the oxygen to carbon peak in $g(r)$ is at a shorter distance than the nitrogen to carbon peak, and is higher. This implies that the oxygen is the more likely centre for reactivity, as seen experimentally[54]. As previously alluded to, the partitioning of propranolol at the membrane interface biases the oxygen to be at a closer distance to the lipid carbonyl headgroups than the nitrogen, and thus becomes geometrically favoured for nucleophilic attack. Furthermore, in the membrane, propranolol exists in the protonated rather than neutral state. Thus the peak in $g(r)$ of the oxygen (protonated) to carbon distance is shorter than the oxygen (neutral) to carbon, and this places it close to the site of nucleophilic attack.

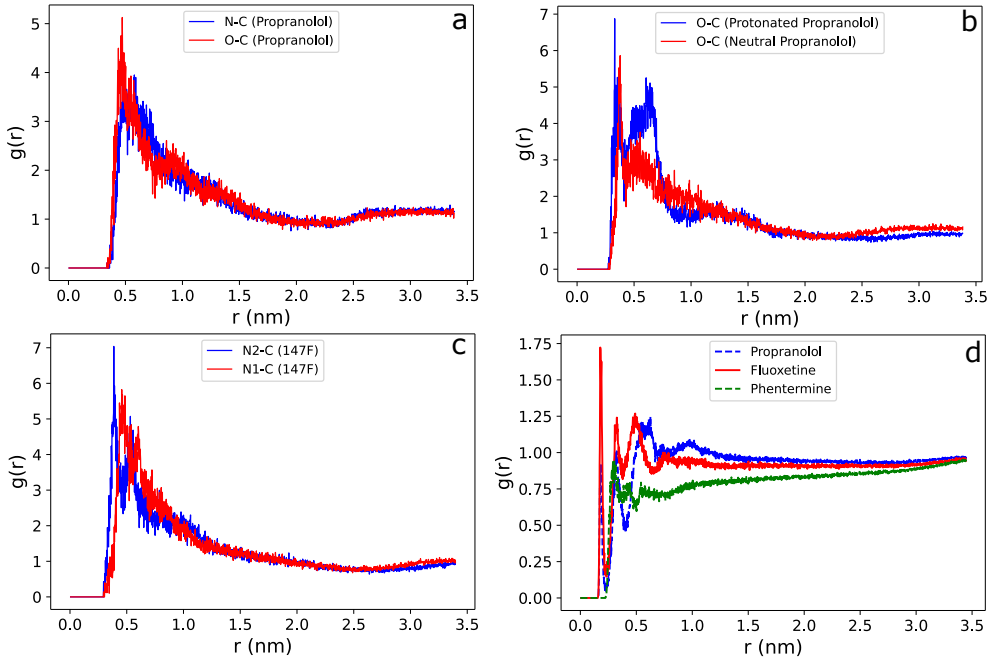


Figure 4.9: Radial distribution functions for (a) O-C and N-C bond distance in propranolol, (b) O-C bond distance in neutral and protonated propranolol, (c) N-C bond distance in different nitrogenic sites of 2-aminomethylbenzimidazole, (d) O (water) - H (protonated site) bond distance in fluoxetine, propranolol and phentermine

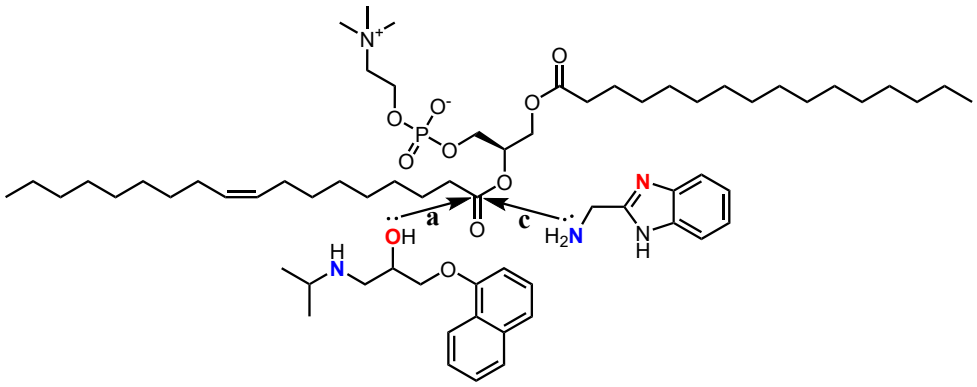


Figure 4.10: Chemical structure of the most likely interaction of propranolol (a) and 2-aminomethylbenzimidazole (c) with the POPC lipid

For 2-aminomethylbenzimidazole, the first peak for the free amine nitrogen is at a shorter distance than for the benzimidazole nitrogen atom, thus implying the reactivity is likely to come from the amine nitrogen rather than the ring nitrogen (figure 4.10). The final plot shows the water oxygen to protonated hydrogen distance. As fluoxetine exhibits high hydrolysis behaviour, the plot correlates well with earlier results and shows the highest peak at the shortest distance for the deprotonated species, and thus this CAD molecule must be deprotonated first before being able to act as a phase transfer catalyst. Phentermine exhibits no reactivity, and has the lowest peak height in $g(r)$ for any CAD molecule in the plot.

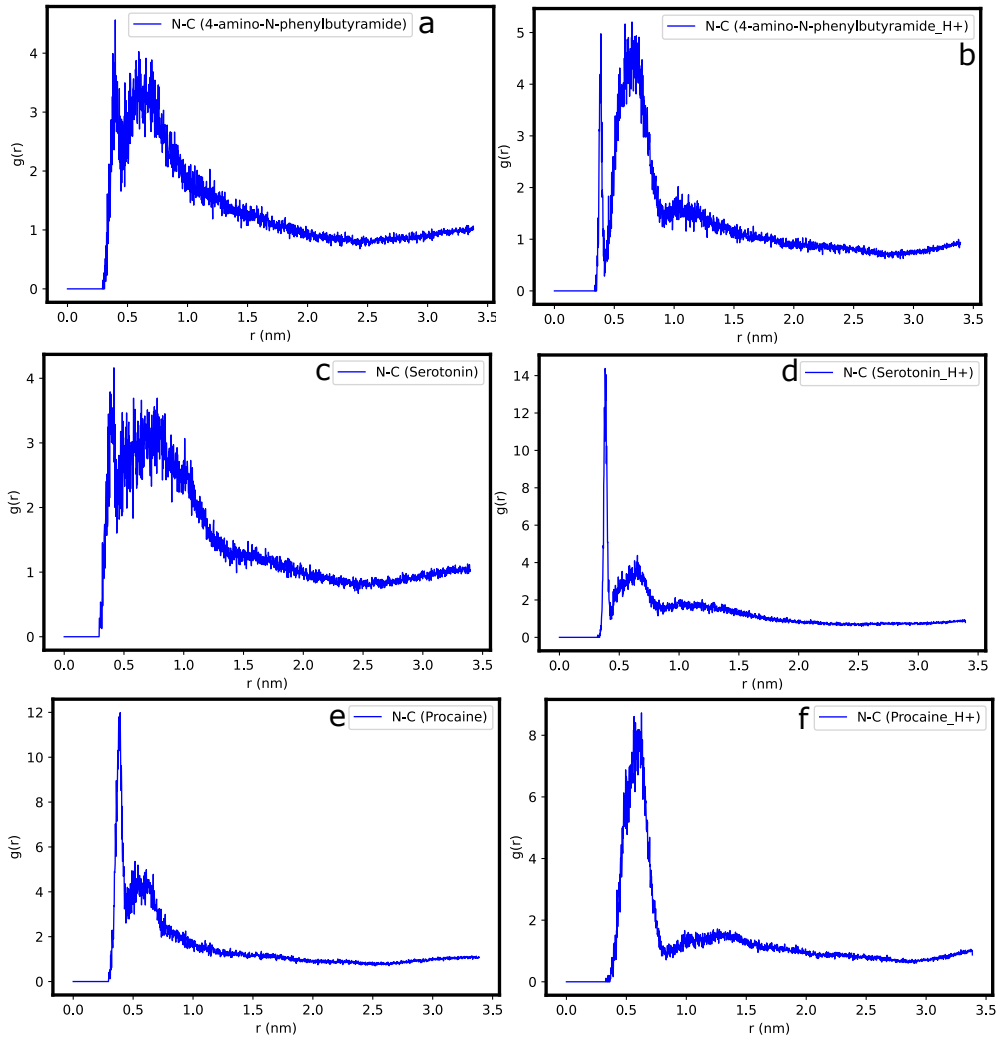


Figure 4.11: Radial distribution functions for (a) N-C bond distance in neutral 4-amino-*N*-phenylbutanamide, b) N-C bond distance in protonated 4-amino-*N*-phenylbutanamide, c) N-C bond distance in neutral serotonin, d) N-C bond distance in protonated serotonin, e) N-C bond distance in neutral procaine, f) N-C bond distance in protonated procaine

For 4-Amino-*N*-phenylbutanamide, only the free amine at the terminal of the aliphatic

chain has the potential to act as a nucleophile in reactivity with the lipid carbonyl carbon atom. The RDF of the distance between the nucleophilic nitrogen amine atom to the carbonyl carbon atom of the lipid headgroup is at a shorter distance, 0.4 nm, than its protonated analogue, 0.45 nm, and the first peak appears at a shorter distance with the neutral form (figure 4.11). Likewise with serotonin, the N-C distance between the nucleophilic amine and carbonyl carbon appears at a shorter distance for the first peak with the neutral vs the protonated form. The greatest difference in RDF peak distance can be seen in the drug procaine. Here, the first peak appears at a peak distance of 0.4 nm for the neutral form compared with its protonated form at 0.65 nm. In all these cases, the neutral form undergoes direct lipidation and the RDF predicts much closer contact for the nucleophilic nitrogen atom relative to all possible carbonyl carbon atoms of the lipid headgroup.

4.3.3 Orientation Behaviour

These set of calculations were run to predict the molecular orientation of each drug molecule, in the POPC membrane, along the principal (z) axis. This principal axis corresponds to the lowest eigenvalue of the moment of inertia tensor. The (z)-component of this vector was extracted, which provided the cosine of the principal axis relative to the membrane normal vector. This data was plotted as a function of simulation time, across 200 ns of the trajectory. This plot shows the preferred orientation of the molecule, along the z axis, as it traverses the bilayer. The second plot was a histogram showing the distribution of these cosine angles as a function of frequency along their trajectory. These plots predict whether there is preferred orientation of each molecule for reactivity.

As can be seen in figure 4.12, there is an orientational preference for protonated propranolol in the membrane. For all three systems, the cos θ fluctuates between -1.0 (minimum) and +1.0 (maximum). Whilst 2-aminomethylbenzimidazole and propranolol fluctuate evenly between these two positions, protonated propranolol favours an orientation with cos θ between 0 and -1.0. This implies that unlike protonated propranolol, neutral propranolol has no preferred orientation in the membrane. This could give credence to a lower reactivity with the membrane, compared

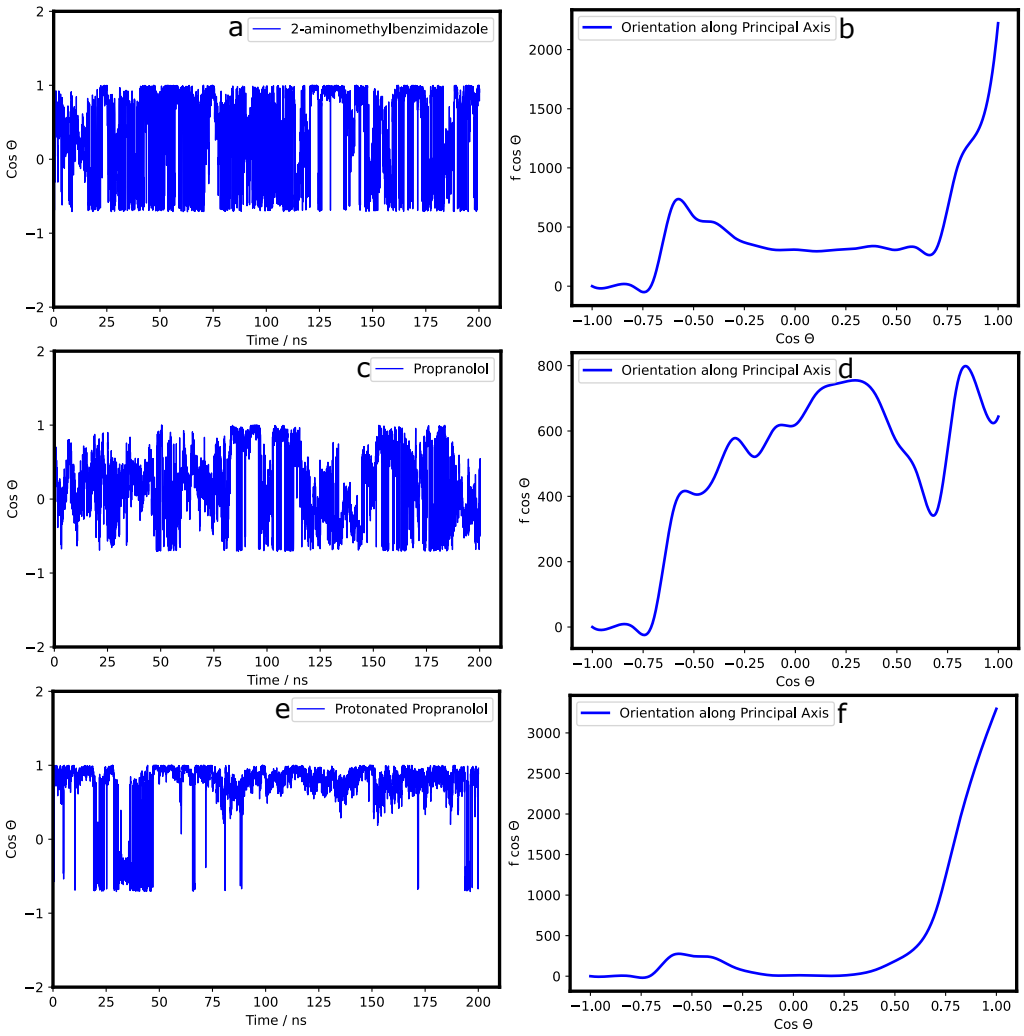


Figure 4.12: Orientation plots of cos θ vs simulation time for (a) 2-aminomethylbenzimidazole, (c) propranolol, (e) protonated propranolol. Histogram of cos θ plots for (b) 2-aminomethylbenzimidazole, (d) propranolol, (f) protonated propranolol

with its protonated analogue. This is further confirmed by histogram plots for cos θ. The orientation of 2-aminomethylbenzimidazole, along its principal axis, is most frequent at a cosine θ value of 1.0, confirming preferred angle of orientation. This behaviour is exhibited even more preferentially in protonated propranolol, with an even higher distribution of cos $\theta = 1$. Neutral propranolol evidently has no preferred cos θ from the histogram plot.

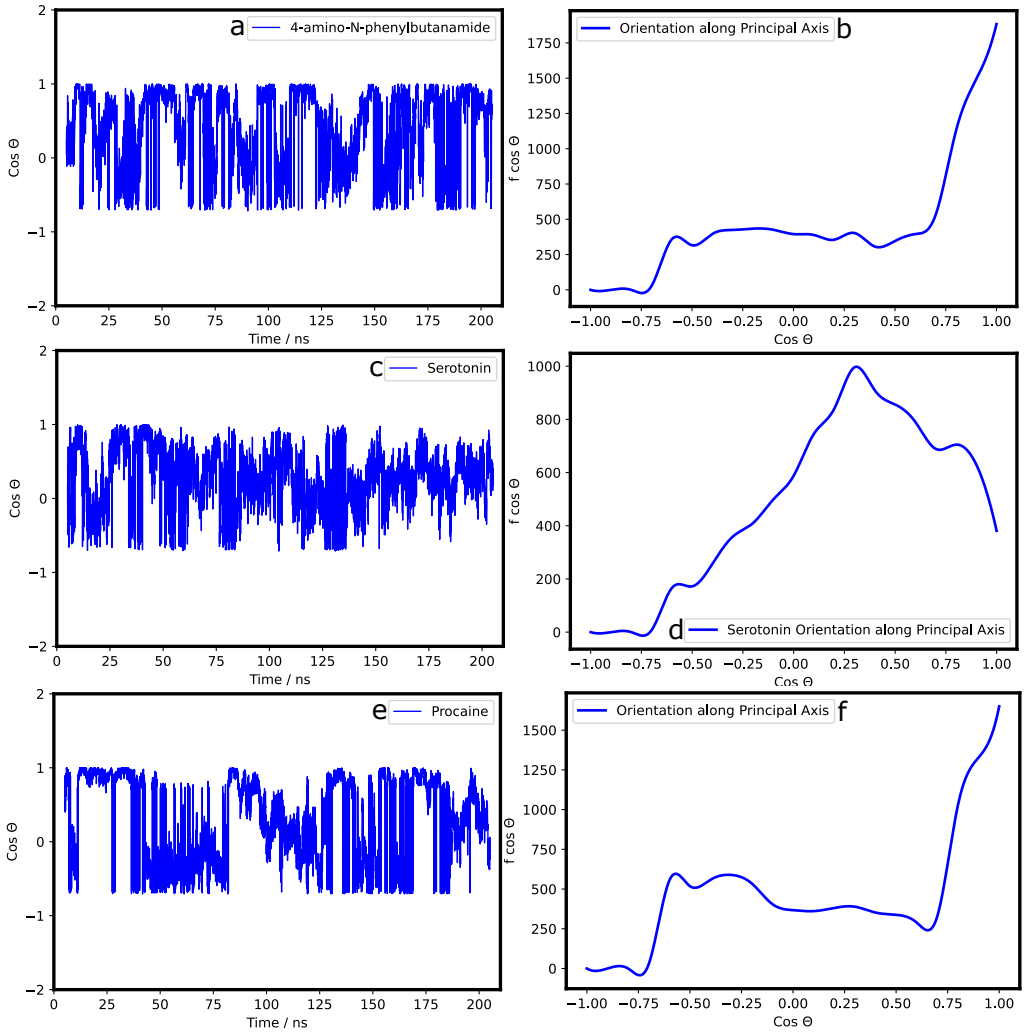


Figure 4.13: Orientation plots of cos θ vs simulation time for (a) 4-amino-*N*-phenylbutanamide, (c) serotonin, (e) procaine. Histogram of cos θ plots for (b) 4-amino-*N*-phenylbutanamide, (d) serotonin, (f) procaine

As can be seen in figure 4.13, the additional three reactive molecules were also subjected to the same orientation calculations. 4-Amino-*N*-phenylbutanamide, serotonin and procaine all exhibit a preferred orientation along their respective principal axes. All three drug molecules have a fluctuating cos θ value between -1.0 and 1.0 along the 200 ns trajectory. Both 4-Amino-*N*-phenylbutanamide and procaine have

a peak at cos $\theta = 1$, implying their preferred orientation. Serotonin has its highest distribution at cos $\theta = 0.3$. This perhaps points to favoured orientation of Serotonin relative to the other two molecules.

4.3.4 PMF Profiles

The potential of mean force (PMF) is a measure of the free energy change as the molecule is pulled along the membrane bilayer along a chosen reaction coordinate, in this case the z axis. The thermodynamic minimum in both neutral and protonated molecules indicates the positioning of the molecule at the membrane interface. From the work above we expect the location of the minima in the PMFs to be in the region of the lipid headgroups. The plots show that the neutral molecules prefer to be localised at the interface where the free energy is lowest. In this region there is the highest abundance of PC lipid headgroups. The protonated molecules prefer to be localised slightly further away from the lipid headgroups and closer to the aqueous region, as shown by the free energy minima. This is indicative of drug stability by the PC lipid headgroups for the neutral molecules, whereas the protonated forms are further out in the water region and must be deprotonated first (figure 4.14). These free energy differences therefore indicate preferred localisation of the molecule within the membrane (see table 4.8). As fluoxetine exhibits lipid hydrolysis behaviour, its free energy in the aqueous region is higher than for the other drug molecules. As phentermine exhibits no reactivity, its free energy is lower than both fluoxetine and propranolol in the aqueous region.

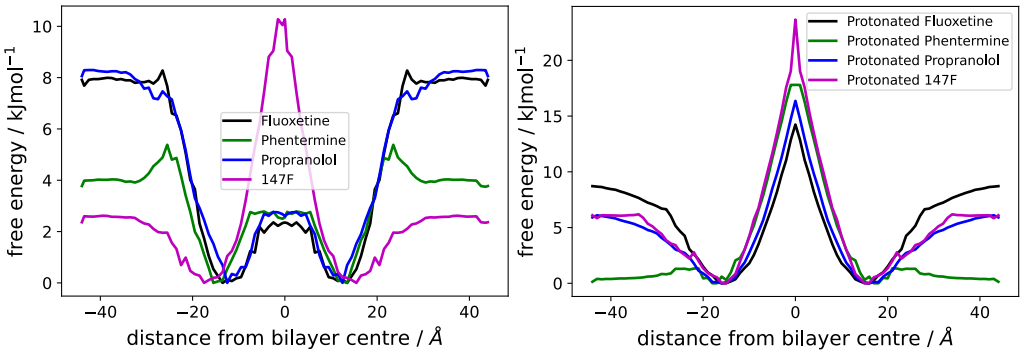


Figure 4.14: Potential of mean force for neutral (left) and protonated (right) CAD molecules

The stabilisation energy of a set of additional molecules in the membrane was also

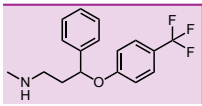
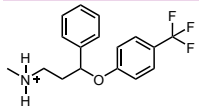
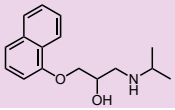
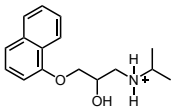
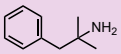
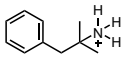
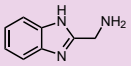
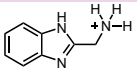
Chemical Structure	Drug Name	Δ*G* (Membrane) / kJ mol^{-1}	Δ*G* (Water) / kJ mol^{-1}	Drug Position (Minimum) / Å
	Fluoxetine	**2.35**	**8.28**	**13.5**
	Protonated fluoxetine	**14.2**	**8.72**	**16.0**
	Propranolol	**2.60**	**8.30**	**12.5**
	Protonated propranolol	**16.4**	**6.10**	**16.0**
	Phentermine	**2.50**	**5.38**	**14.5**
	Protonated phentermine	**17.8**	**1.34**	**18.0**
	2-aminomethylbenzimidazole	**10.1**	**2.62**	**17.0**
	Protonated 2-aminomethylbenzimidazole	**23.7**	**6.19**	**15.0**

Table 4.8: Comparison of drug position and free energy change from the membrane centre out to the water region for a set of neutral and protonated CAD molecules

simulated with biased molecular dynamics. Procaine and tetracaine have been found experimentally to undergo direct lipidation with membrane lipids[25]. 4-Amino-*N*-phenylbutanamide has unusual behaviour, in that its protonated form prefers to sit closer to the centre of the membrane than its neutral analogue (see figure 4.19b).

Biased simulations were then conducted to calculate the PMF for a wide series of drug molecules, which had been initially subjected to unbiased simulations. For the first series of indoles, indazoles and benzimidazoles, the PMF was calculated by applying a pulling force to gradually move the drug molecule from the centre of the membrane out to the aqueous region and measure the free energy change across the membrane in the z axis. As can be seen for 5-aminoindole, 5-aminoindazole, 5-aminobenzimidazole and 4-aminobenzotriazole, the free energy is lower in the aqueous region than in the centre of the membrane, implying greater stability in an aqueous environment (figure 4.15, 4.16). The thermodynamic minimum is evidently close to the PC head group lipids. The position of the minimum is similar for both neutral and protonated forms (table 4.9). The only drug capable of lipida-

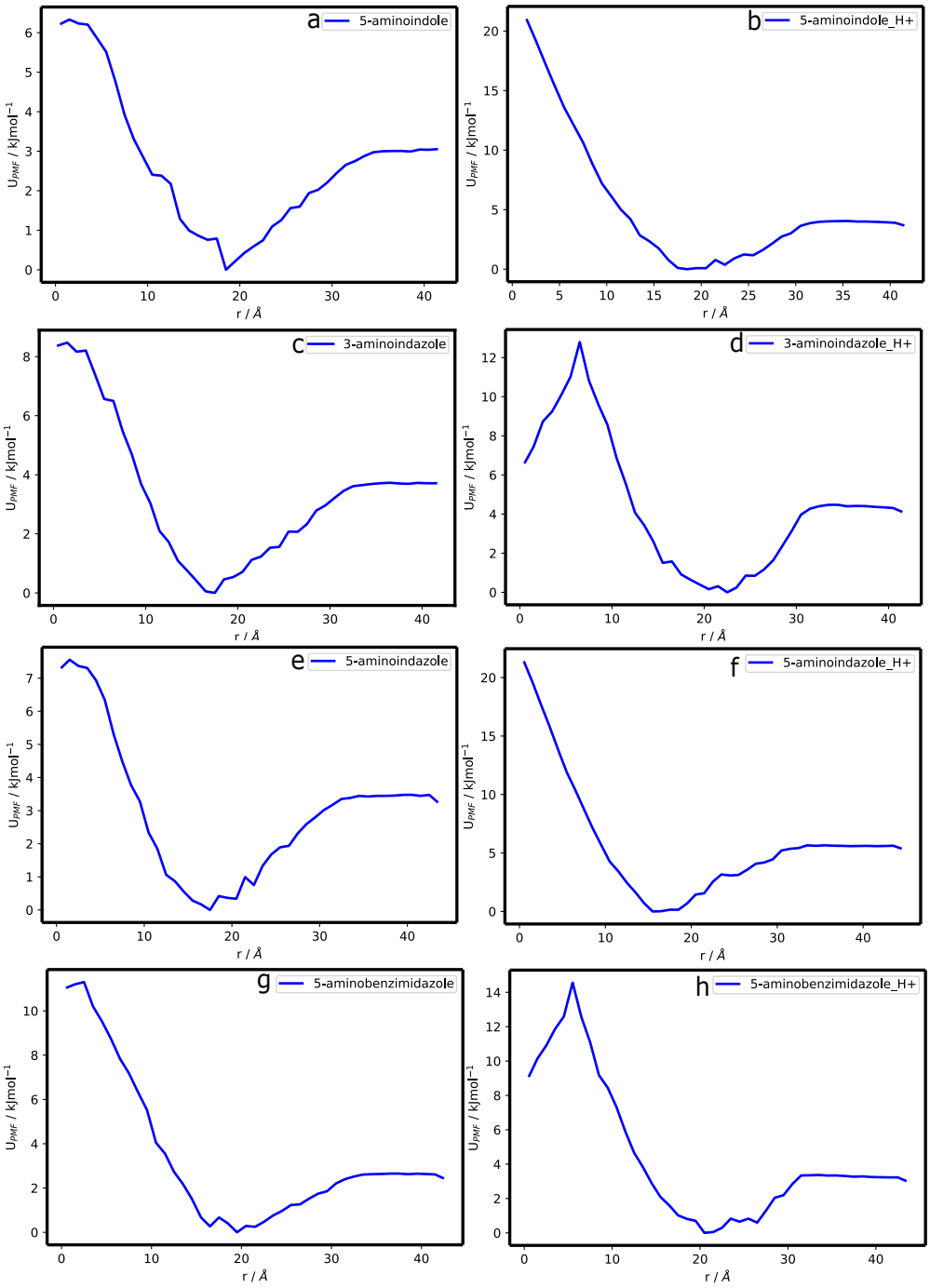


Figure 4.15: Potential of mean force for (a) 5-aminoindole, (b) protonated 5-aminoindole, (c) 3-aminoindazole, (d) protonated 3-aminoindazole, (e) 5-aminoindazole, (f) protonated 5-aminoindazole, (g) 5- aminobenzimidazole, (h) protonated 5-aminobenzimidazole

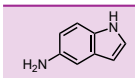
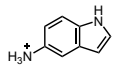
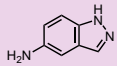
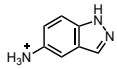
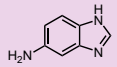
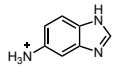
Chemical Structure	Drug Name	ΔG (Membrane) / kJ mol^{-1}	ΔG (Water) / kJ mol^{-1}	Drug Position (Minimum) / Å
	5-aminoindole	**6.33**	**3.05**	**18.5**
	Protonated 5-aminoindole	**21.02**	**4.05**	**19.5**
	3-aminoindazole	**8.38**	**3.72**	**17.5**
	Protonated 3-aminoindazole	**6.58**	**4.47**	**22.5**
	5-aminoindazole	**7.56**	**3.48**	**17.5**
	Protonated 5-aminoindazole	**21.4**	**5.66**	**15.5**
	5-aminobenzimidazole	**11.3**	**2.65**	**19.5**
	Protonated 5-aminobenzimidazole	**9.07**	**3.37**	**20.5**

Table 4.9: Comparison of drug position and free energy change from the membrane centre out to the water region for a set of neutral and protonated CAD molecules

tion, namely 3-aminoindazole, shows a minimum close to the PC headgroups (figure 4.15c). Its protonated form is further out in the aqueous region where deprotonation can occur and thus both the neutral and protonated forms of this molecule are more stable outside than inside the membrane.

The next set of PMF calculations were conducted for the substituted indazoles, with amine groups at varying positions on the aromatic ring (table 4.10, 4.11). For both neutral and protonated molecules, their free energy was lower when situated in the aqueous region relative to the centre of the membrane, and thus displaying greater stability outside of the membrane (figure 4.16, 4.17). The thermodynamic minimum exists at a smaller distance from the membrane centre for the neutral molecules rather than their protonated forms, implying reactivity with the PC lipid headgroups closer to the membrane-water interface (table 4.10). The only caveat in this trend appears for the protonated 5-(aminomethyl)-1*H*-indazole PMF. Here, a high free energy barrier (figure 4.16g) must be crossed first before undergoing lipidation at the membrane interface. This behaviour could be caused by protonation affecting localisation of the drug and causing it to be less stable and thus preferentially moving to the aqueous region.

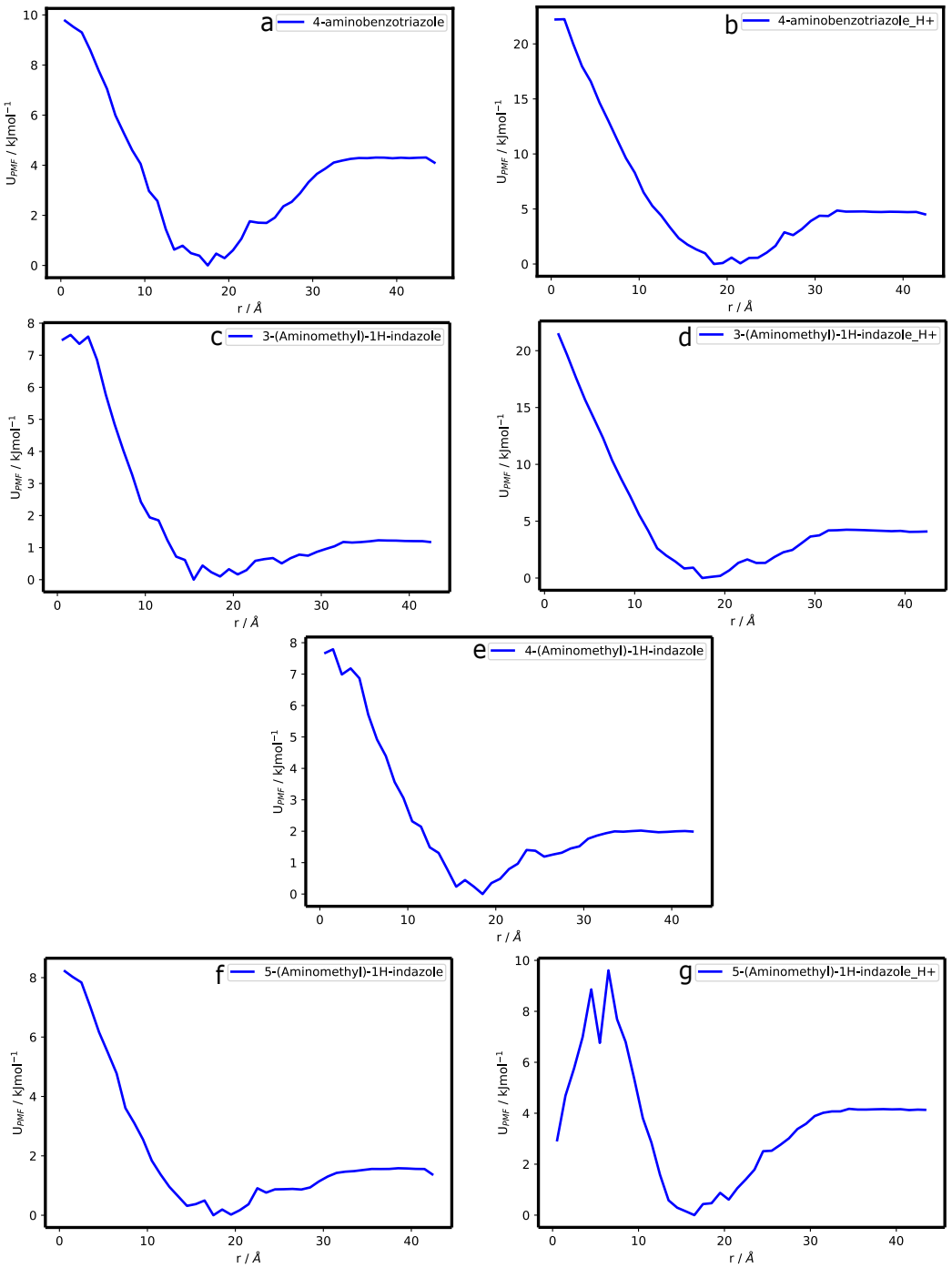


Figure 4.16: Potential of mean force for (a) 4-aminobenzotriazole, (b) protonated 4-aminobenzotriazole, (c) 3-(aminomethyl)-1*H*-indazole, (d) protonated 3-(aminomethyl)-1*H*-indazole, (e) 4-(aminomethyl)-1*H*-indazole, (f) 5-(aminomethyl)-1*H*-indazole, (g) protonated 5-(aminomethyl)-1*H*-indazole

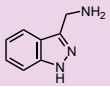
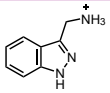
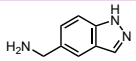
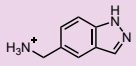
Chemical Structure	Drug Name	Δ*G* (Membrane) / kJ mol^{-1}	Δ*G* (Water) / kJ mol^{-1}	Drug Position (Minimum) / Å
	4-aminobenzotriazole	**9.77**	**4.31**	**17.5**
	Protonated 4-aminobenzotriazole	**22.2**	**4.86**	**18.5**
	3-(aminomethyl)-1H-indazole	**7.63**	**1.23**	**15.5**
	Protonated 3-(aminomethyl)-1H-indazole	**21.54**	**4.25**	**17.50**
	4-(aminomethyl)-1H-indazole	**7.79**	**2.02**	**18.50**
	5-(aminomethyl)-1H-indazole	**8.24**	**1.58**	**17.50**
	Protonated 5-(aminomethyl)-1H-indazole	**2.89**	**4.17**	**16.50**

Table 4.10: Comparison of drug position and free energy change from the membrane centre out to the water region for a set of neutral and protonated CAD molecules

The following set of PMF curves were obtained for increasingly substituted indoles (figure 4.17, 4.18). For these molecules, their neutral form localises closer to the membrane interface where lipidation can take place with the PC lipid headgroups, whilst their protonated forms position themselves closer to the aqueous region. Both 3-aminoindole and *N*-indol-3-ylmethylethylamine have quite a lower free energy, in the aqueous region in comparison to the membrane centre (table 4.11, 4.12). The free energy barrier between the minimum and aqueous region is lower for their protonated analogues compared with their neutral forms, perhaps implying easier diffusion across the bilayer. Protonated 3-(aminomethyl)indole has a large energy barrier to cross as it is pulled across the bilayer, perhaps in part due to again the protonation state of the molecule preferring to be out in the aqueous region (figure 4.17h). 3-(1*H*-Indol-3-yl)propan-1-amine seems to have an equal stability in the aqueous region and the membrane centre, as its free energy in both regions is comparably similar (table 4.12). Interestingly, its protonated form however still prefers to be localised closer to the head group region of the membrane (figure 4.18b).

The next set of PMF curves share the same trend, with their neutral forms being closer to the membrane interface than their protonated forms. The free energy is

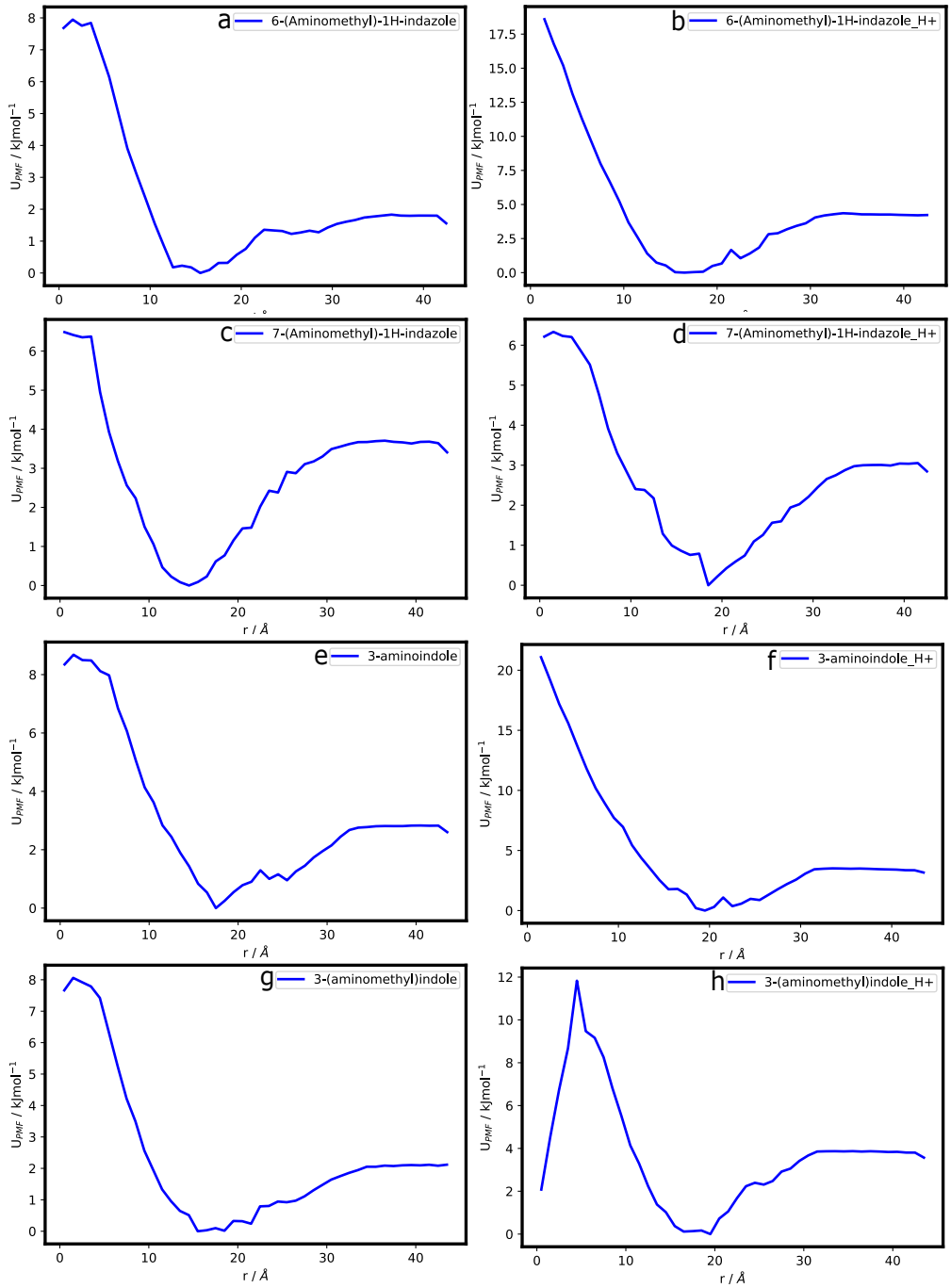


Figure 4.17: Potential of mean force for (a) 6-(aminomethyl)-1*H*-indazole, (b) protonated 6-(aminomethyl)-1*H*-indazole, (c) 7-(aminomethyl)-1*H*-indazole, (d) protonated 7-(aminomethyl)-1*H*-indazole, (e) 3-aminoindole, (f) protonated 3-aminoindole, (g) 3-(aminomethyl)-indole, (h) protonated 3-(aminomethyl)-indole

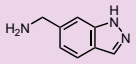
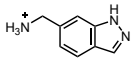
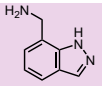
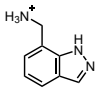
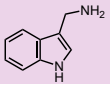
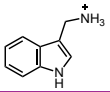
Chemical Structure	Drug Name	ΔG (Membrane) / kJ mol^{-1}	ΔG (Water) / kJ mol^{-1}	Drug Position (Minimum) / Å
	6-(aminomethyl)-1H-indazole	**7.95**	**1.83**	**15.5**
	Protonated 6-(aminomethyl)-1H-indazole	**18.59**	**4.36**	**16.5**
	7-(aminomethyl)-1H-indazole	**6.49**	**3.71**	**14.5**
	Protonated 7-(aminomethyl)-1H-indazole	**6.33**	**3.01**	**18.5**
	3-aminoindole	**8.68**	**2.83**	**17.5**
	Protonated 3-aminoindole	**21.09**	**3.51**	**19.50**
	3-(aminomethyl)-indole	**8.06**	**2.11**	**15.5**
	Protonated 3-(aminomethyl)-indole	**2.08**	**3.87**	**19.5**

Table 4.11: Comparison of drug position and free energy change from the membrane centre out to the water region for a set of neutral and protonated CAD molecules

lower in the aqueous region than at the membrane centre for most of these molecules (figure 4.19). The exceptions to this behaviour are protonated serotonin and protonated *N*-(4-aminophenyl)-1,1,1-trifluoromethanesulfonamide. There is literature precedence which indicates that both protonated serotonin and this other protonated molecule are not able to traverse the cell membrane[229]. These theoretical models predict that no thermodynamic minimum is reached (figure 4.19d,f) in these PMF plots and therefore these protonated molecules have no stability within the membrane. Therefore a comparison of the free energy difference is not possible as they are unable to cross the bilayer (see table 4.13).

The final set of PMF curves are for some biologically relevant drug molecules and others simulated in the POPC membrane. Procaine has a thermodynamic minimum closest to the bilayer centre than the other drugs in this series, but also close enough to the PC lipid headgroups for lipidation (figure 4.19g). Its protonated form has a bigger free energy difference in the aqueous region compared with the membrane

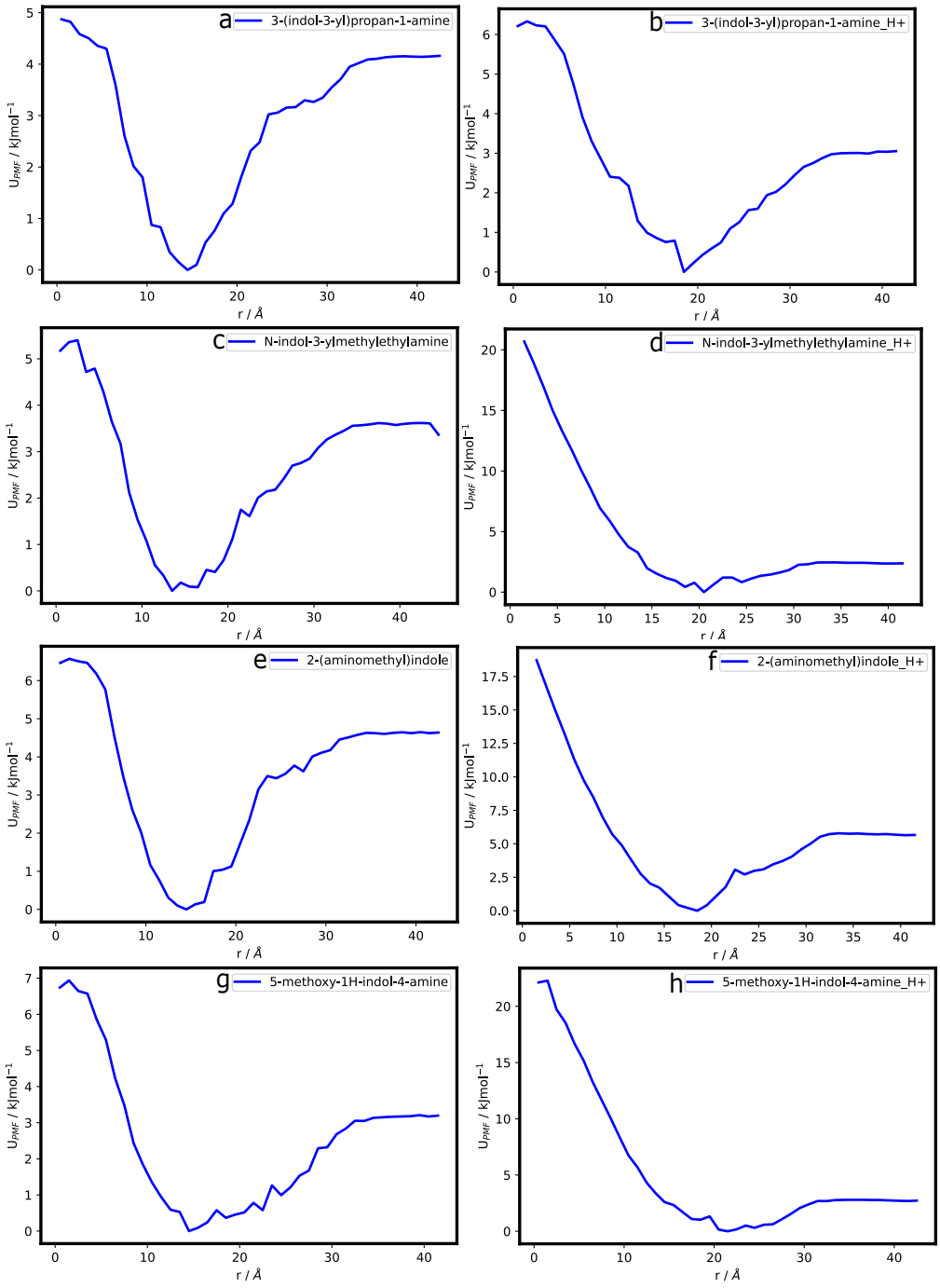


Figure 4.18: Potential of mean force for (a) 3-(1*H*-indol-3-yl)propan-1-amine, (b) protonated 3-(1*H*-indol-3-yl)propan-1-amine, (c) *N*-indol-3-ylmethylethylamine, (d) protonated *N*-indol-3-ylmethylethylamine, (e) 2-aminomethylindole, (f) protonated 2-aminomethylindole, (g) 5-methoxy-1*H*-indol-4-amine, (h) protonated 5-methoxy-1*H*-indol-4-amine

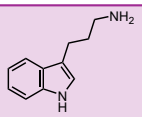
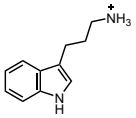
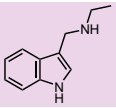
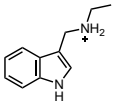
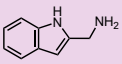
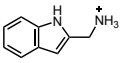
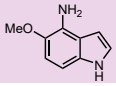
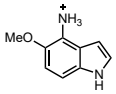
Chemical Structure	Drug Name	Δ*G* (Membrane) / kJ mol^{-1}	Δ*G* (Water) / kJ mol^{-1}	Drug Position (Minimum) / Å
	3-(1H-indol-3-yl)propan-1-amine	**4.87**	**4.15**	**14.5**
	Protonated 3-(1H-indol-3-yl)propan-1-amine	**6.33**	**3.01**	**18.5**
	N-indol-3-ylmethylethylamine	**5.40**	**3.62**	**13.5**
	Protonated N-indol-3-ylmethylethylamine	**20.69**	**2.45**	**20.5**
	2-aminomethylindole	**6.57**	**4.64**	**14.5**
	Protonated 2-aminomethylindole	**18.72**	**5.79**	**18.5**
	5-methoxy-1H-indol-4-amine	**6.94**	**3.21**	**14.5**
	Protonated 5-methoxy-1H-indol-4-amine	**22.28**	**2.80**	**21.50**

Table 4.12: Comparison of drug position and free energy change from the membrane centre out to the water region for a set of neutral and protonated CAD molecules.

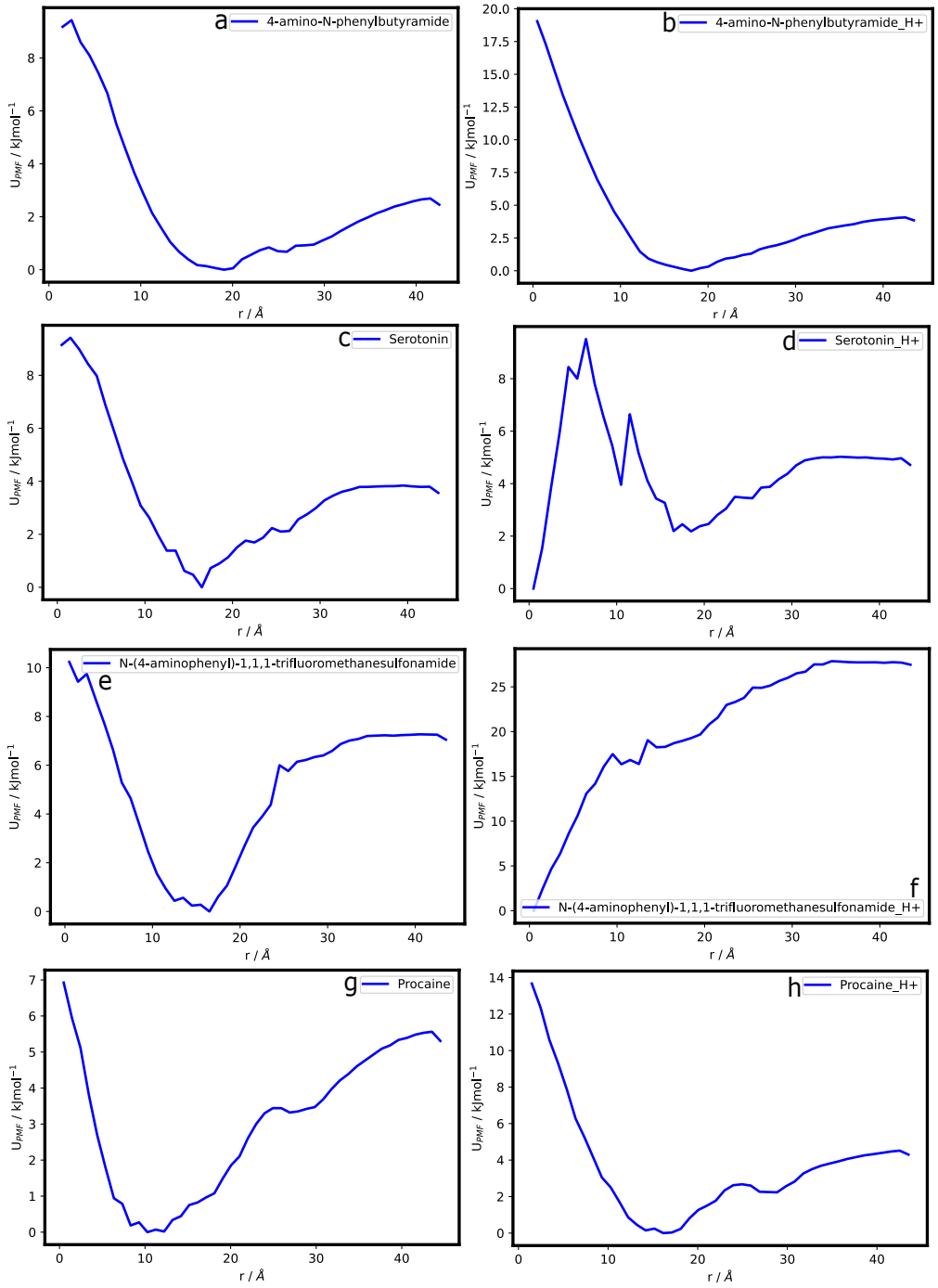


Figure 4.19: Potential of mean force for (a) 4-amino-*N*-phenylbutanamide, (b) protonated 4-amino-*N*-phenylbutanamide, (c) serotonin, (d) protonated serotonin, (e) *N*-(4-aminophenyl)-1,1,1-trifluoromethanesulfonamide, (f) protonated *N*-(4-aminophenyl)-1,1,1-trifluoromethanesulfonamide, (g) procaine, (h) protonated procaine

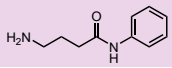
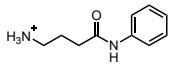
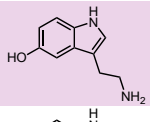
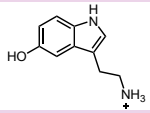
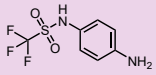
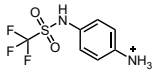
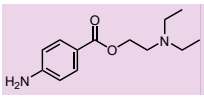
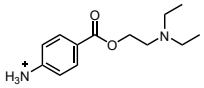
Chemical Structure	Drug Name	Δ*G* (Membrane) / kJ mol^{-1}	Δ*G* (Water) / kJ mol^{-1}	Drug Position (Minimum) / Å
	4-amino-*N*-phenybutanamide	**9.43**	**2.69**	**19.0**
	Protonated 4-amino-N-phenylbutanamide	**19.05**	**4.07**	**18.0**
	Serotonin	**9.42**	**3.84**	**16.50**
	Protonated Serotonin	**Outside Membrane**	**Outside Membrane**	**Outside Membrane**
	N-(4-aminophenyl)-1,1,1-trifluoromethanesulfonamide	**10.24**	**7.27**	**16.5**
	Protonated N-(4-aminophenyl)-1,1,1-trifluoromethanesulfonamide	**Outside Membrane**	**Outside Membrane**	**Outside Membrane**
	Procaine	**6.93**	**5.56**	**10.5**
	Protonated Procaine	**13.67**	**4.52**	**16.5**

Table 4.13: Comparison of drug position and free energy change from the membrane centre out to the water region for a set of neutral and protonated CAD molecules

centre versus its neutral form. Perhaps implicating its preferred ionised form when placed in a membrane. Once again, experimental literature has shown that neither protonated tetracaine[201], chloroquine[157], nor sertraline[168] are able to cross the membrane bilayer. These PMF plots (figure 4.20b,d,f) show no thermodynamic minimum in the energy and therefore no stability within the membrane. This correlates well with experimental findings to show that the no free energy difference across the membrane bilayer can be captured with these three protonated molecules (table 4.14). Both neutral chloroquine and sertraline have a much lower free energy in the membrane centre than in the aqueous region. They both still exhibit an energy minimum close to the PC lipid head groups, and thus prefer to be localised closer to the membrane-water interface (PC lipid head group abundance) for reactivity.

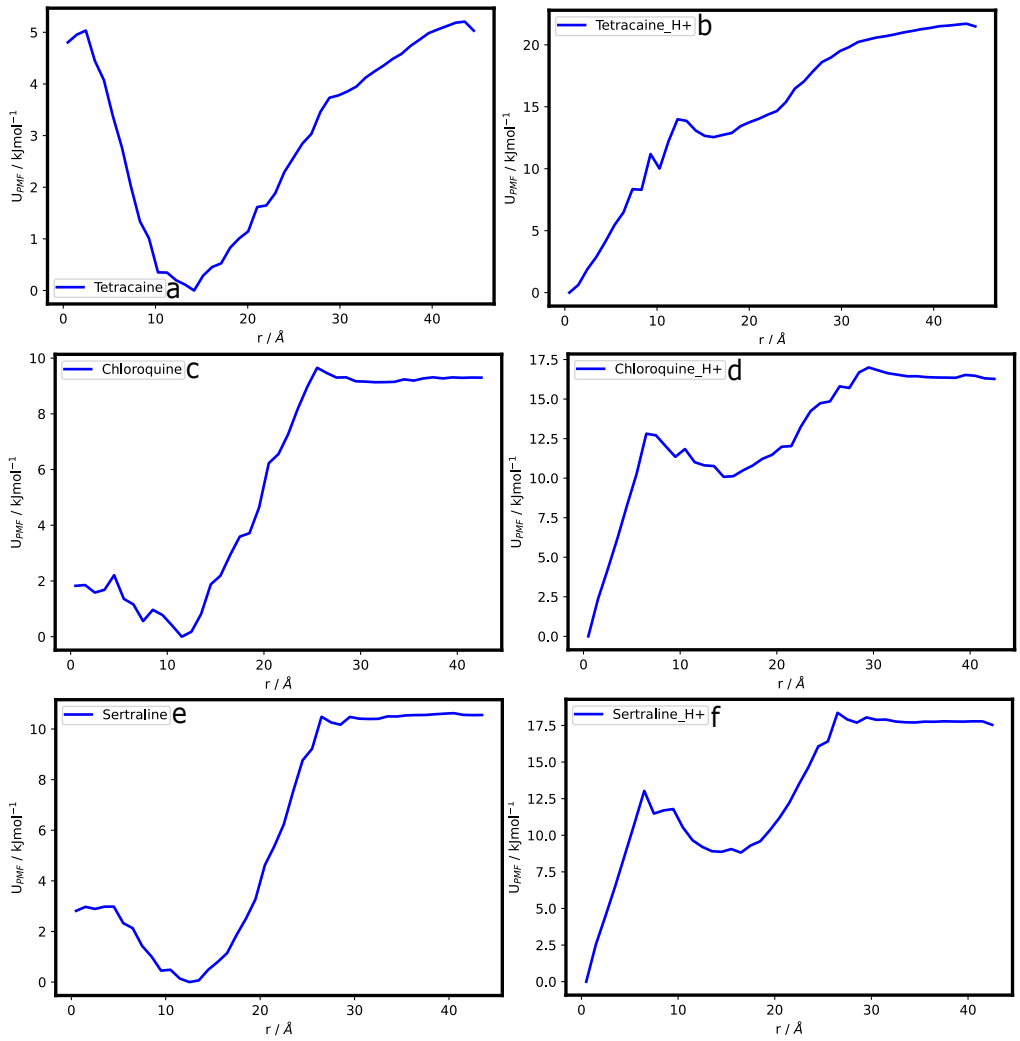


Figure 4.20: Potential of mean force for (a) tetracaine, (b) protonated tetracaine, (c) chloroquine, (d) protonated chloroquine, (e) sertraline, (f) protonated sertraline

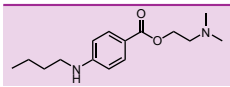
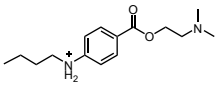
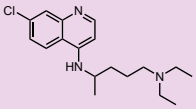
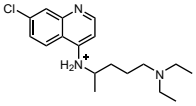
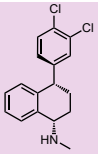
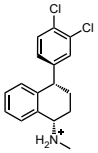
Chemical Structure	Drug Name	Δ*G* (Membrane) / kJ mol^{-1}	Δ*G* (Water) / kJ mol^{-1}	Drug Position (Minimum) / Å
	Tetracaine	**5.03**	**5.21**	**14.0**
	Protonated Tetracaine	**Outside Membrane**	**Outside Membrane**	**Outside Membrane**
	Chloroquine	**1.85**	**9.31**	**11.5**
	Protonated Chloroquine	**Outside Membrane**	**Outside Membrane**	**Outside Membrane**
	Sertraline	**2.97**	**10.63**	**12.5**
	Protonated Sertraline	**Outside Membrane**	**Outside Membrane**	**Outside Membrane**

Table 4.14: Comparison of drug position and free energy change from the membrane centre out to the water region for a set of neutral and protonated CAD molecules

4.4 Summary

In conclusion, this chapter presented results on atomistic simulations of drug-lipid interactions. Initial static MD simulations were run to establish preferred orientation and depth of partitioning of these molecules in the membrane. Classic CAD molecules, such as propranolol, procaine and 4-amino-*N*-phenylbutanamide, alongside other small organics, such as 2-aminomethylbenzimidazole, showed appropriate orientation and localisation close to the PC lipid headgroups, thus showing a high possibility of lipidation behaviour. Radial distribution functions showed close distance contacts between interacting atoms of the drug and lipid, thus again showing preferred orientation within the membrane. Orientation behaviour confirmed preferred orientation of protonated propranolol versus neutral propranolol, alongside 2-aminomethylbenzimidazole for lipidation reactivity. Furthermore, the CAD molecules 4-amino-*N*-phenylbutanamide, serotonin and procaine showed a preferred orientation in the membrane, thus confirming their potential lipidation behaviour as well. Finally, PMF calculations were run to demonstrate stability and localisation of each molecule in the membrane.

Experimentally, these molecules either undergo lipidation, hydrolysis or show no reactivity. A high likelihood of reactivity is confirmed by the drug molecule being close to the lipid headgroups and within a reasonable distance from the interacting carbonyls. Now that molecular orientations and positions needed for reactivity have been quantified, in the succeeding chapter, DFT calculations have been conducted to confirm stability of intermediates within the reaction pathway of a set of reactive molecules.

Chapter 5

QM/MM Simulations of Drug-Lipid Interactions

5.1 Background

Quantum mechanical/molecular mechanical (QM/MM) simulations have grown in popularity over the last few decades, culminating in the 2013 Nobel Prize for Chemistry, recognising the hugely diverse length and time scales possible in computer simulations of complex chemical systems. Historically, the first proponents of this hybrid technique were Warshel and Levitt in the late 1980s[64]. With this new discovery they were able to understand the catalytic properties in an enzymatic reaction, further verified in the 1990s by Karplus et al[43]. These researchers were the first to develop a method of combining semi-empirical quantum chemical methods with the CHARMM force field and thus providing a benchmark for comparison with not only full *ab initio* quantum chemical calculations but also against experimental data. In the modern age, QM/MM has found extensive use in studying biomolecular systems, organic/inorganic reactions as well as in organometallic chemistry[192].

The foundational concept of QM/MM is the combination of an *ab initio* quantum mechanical method, which treats a small site of chemical reactivity, with a classi-

cal molecular mechanical force field method, which treats the surrounding system including solvent. Both the QM and MM regions are in contact with each other, thus the potential energy of the system must include both the coupling term and boundary potential energy if any covalent bonds cross the two regions. Thus the total potential energy of the system encapsulates the potential energy of the QM, MM and QM-MM schemes and fully defines the QM/MM method. The potential issue of covalent bonds traversing the QM/MM region are dealt with by the common approach of 'link atoms'. Here, additional QM atoms are manually added to the terminal end of free bonds, usually a QM hydrogen atom. The major disadvantage however, is the possible overpolarisation of electron density by the close proximity of these QM atoms with the MM point charges. This is simply mitigated by assigning these point charges to more atoms closer to the boundary and thus spreading the charge over more atoms. Thus, the most important part of the QM/MM calculation is treating covalent bonds across the QM/MM boundary in the most computationally appropriate method[191].

Whilst MM methods have the advantage of simulating large systems to understand their dynamics, QM methods have the advantage of studying chemical reactions and the process of bond formation and breaking, which is of course pertinent to the chemical reactivity of organic molecules with membrane lipids. QM methods also have no requirement to be parameterised and are often of much higher accuracy than MM methods. Of course the major disadvantage with QM methods is the issue of appropriate scaling. The computing time of classical molecular mechanical simulations scale with n^2 (where n is the number of atoms), whilst QM simulations scale with n^4, and are thus far more computationally expensive. Therefore QM/MM utilises the accuracy of QM methods with the speed of MM force fields to achieve a fairly efficient computing time with respect to scaling of the system[57].

Another critical part of the QM/MM technique is the appropriate choice of QM/MM theory. Often this requires a high compatibility of the chosen QM level of theory, the MM force field at hand and the appropriate number of GPU/CPU cores from the available hardware. A balance of accuracy versus speed is always an important

consideration in QM/MM calculations. To ensure efficient QM/MM calculations, the number of atoms in the QM region should be as minimal as possible and ideally not exceed 80-100 atoms. Although semi-empirical methods, such as PM6 or AM1 yield faster simulations, their chemical accuracy is often less than full density functional theory (DFT) methods. DFT methods are diverse in their choice of density functionals and basis sets, and are the most widely used approach in QM/MM simulations[211].

The entire electronic structure of a molecule can be calculated from Density Functional Theory (DFT) by calculation of the total electron density of the molecule. This is computed by considering the distribution of electron density across the whole molecule, thus making it a good approximation to solving the Kohn-Sham equations (non interacting Schrödinger equation). The crucial benefit of DFT is the inclusion of a exchange correlation energy which acts as a stabilisation energy when calculating the Coulombic repulsion between interacting electrons. Thus the electron density is more accurately calculated from the sum of four distinct energy terms. Firstly, the kinetic energy of the independent electrons, second the Coulombic repulsions of the interacting electrons, thirdly the exchange-correlation energy and fourthly the electrostatic interactions between the nuclei and electrons[207]. The choice of DFT functional is vital in describing the exchange-correlation energy. Therefore, different levels of DFT functionals exist which vary in computational speed and accuracy. The least accurate but computationally fastest is the Local Density Approximation (LDA). The main failing of this functional for complex biomolecular systems is the assumption of homogeneity in the electron density[126]. A more accurate set of DFT functionals are based on Generalised Gradient Approximations (GGA). These offer an improvement on LDA by the addition of a gradient correction term that computes a non-uniform electron density, which is far more realistic for a heterogeneous biomolecular system. Even more accurate DFT functionals are provided by Hybrid Generalised Gradient Approximations (Hybrid GGA). This is of greatest value for QM/MM calculations where an additional Hartree-Fock correlation energy term is included along with an *ab initio* term. The classic B3LYP or PBE0 functionals provide the best accuracy to speed ratio when examining drug-lipid reactivity in

membranes[207].

There are many biomolecular applications of the QM/MM method which make it a suitable choice for comparison with experimental data. Firstly, in the area of structural refinement, QM/MM can be utilised to predict possible hydrogen bonds formed between potential substrates and the active site in enzyme catalysis. This can be directly compared with experimental crystallography data which refines the 3D structure and predicts any possible interactions. Furthermore, QM/MM has the ability to quantify the strength of potential interactions and calculate other physical properties such as electron density and electronic effects. This can inform on the potential site of reactivity and give useful information on choosing the appropriate size of the QM region within enzyme reactivity. Secondly, again within enzymatic reactions, QM/MM can accurately predict the most favourable reaction pathway in enzyme catalysis. QM/MM can predict the energy landscape of a potential enzyme reaction and calculate the free energy barriers to different intermediates. These computed quantities can be compared with experimental rate constants to either approve or disprove potential enzyme reaction pathways. Thus, this makes QM/MM a powerful diagnostic tool for the prediction of chemical reactions. Lastly, and applicable to this project, QM/MM has the capability of predicting transition states within an overall reaction pathway. QM/MM can accurately predict energy barriers between reactive intermediates and transition states, and thus help identify the rate-determining step within organic reaction pathways. This provides invaluable aid in drug discovery as the library of compounds required to search can be limited by discovering which intermediates are more likely to react with membrane lipids, thus which are required to chemically synthesise[132].

5.2 QM/MM Methodology

5.2.1 QM/MM Hamiltonian

The quantum mechanical analogue of classical force fields, in the computation of total potential energy of the system, is mathematically captured by the Hamiltonian. Classically, the Hamiltonian is the net sum of the potential and kinetic energy of the system, but quantum mechanically signifies the small reactive part of the overall system. The distinction between these two regions (QM and MM) is defined by a QM/MM boundary, which must allow the covalent bonding to be conserved across the boundary. The Hamiltonian (H) of the system is defined as the sum of the QM, MM and QM/MM regions. This then simplifies to the total energy (E) of all individual components of the QM/MM system[191]:

$$\hat{H} = \hat{H}_{QM} + \hat{H}_{QM/MM} + \hat{H}_{MM} \tag{5.1}$$

$$E = E_{QM} + E_{QM/MM} + E_{MM} \tag{5.2}$$

The QM and MM Hamiltonian are dependent on the choice of QM level of theory and the choice of compatible force field, respectively, whilst the QM/MM Hamiltonian accounts for interactions across the boundary between the small reactive QM part and the rest of the MM system. Often, this is treated by a so called embedding method (electrostatic or mechanical). From the perspective of electrostatic embedding, an additional term, namely the MM point charges close to the QM/MM boundary, are included with the QM Hamiltonian which accounts for electrostatics at the boundary. QM nuclei are assumed as van der Waals points and they, alongside the MM force field Hamiltonian, constitute the remaining non-covalent interactions. Within the overall equation, the terms a and i account for all QM nuclei and electrons respectively, whilst the term m contributes to all MM atoms (including full nuclear and electronic components) and r the distance. Z reflects the total nuclear charge of the QM atoms and q_m the charge of the MM atoms. Similarly to classical MM force fields, there is a short range repulsive ($C_{12}^{\frac{1}{2}}$) and long range attractive term ($C_6^{\frac{1}{2}}$) for the LJ potential in the total Hamiltonian for the QM/MM region[70]:

$$\hat{H}_{QM/MM} = -\sum_{i,m} \frac{q_m}{r_{im}} + \sum_{a,m} \frac{Z_a q_m}{r_{am}} + \sum_{a,m} \left(\frac{(C_{12,a})^{\frac{1}{2}}(C_{12,m})^{\frac{1}{2}}}{r_{am}^{12}} - \frac{(C_{6,a})^{\frac{1}{2}}(C_{6,m})^{\frac{1}{2}}}{r_{am}^{6}} \right) \quad (5.3)$$

Embedding Approaches

Embedding methods have been developed to account for the multiple covalent bonds which traverse the QM/MM boundary, thus must be well developed within a QM/MM scheme. Whilst van der Waals non-bonding interactions at this boundary are perfectly evaluated with MM force fields, the electrostatics can only be dealt with by a choice of two different embedding methods. The first, namely mechanical embedding, is the simpler approach as it involves only the need of MM force fields to describe the interaction at the QM/MM boundary. The Hamiltonian for the QM region is calculated in isolation (gas phase), and thus the electronic wavefunction need only account for this small subsystem rather than the boundary effects. Full atomic charges are assigned to each atom in the QM and MM region, and through Coulomb's law in the fully MM force field, electrostatic interactions are calculated between atoms across the QM/MM boundary. Therefore the major drawback of this scheme is the lack of polarisation of the MM region on the electron density in the QM region[150]. Only electronic point charges ($q_{A=QM}, q_{B=MM}$), not nuclear charges, are considered within mechanical embedding to provide the overall Hamiltonian of this method[235]. In this equation, r_A and r_B represent the positions of the QM and MM atoms respectively[154]:

$$E_{QM/MM} = \sum_A \sum_B q_A q_B \frac{1}{|r_A - r_B|} \quad (5.4)$$

Electrostatic embedding is the more advanced scheme and in contrast with mechanical, treats the electrostatic interactions across the QM/MM boundary at the QM level[42]. The point charges on all MM atoms in the electrostatic embedding scheme are now considered part of the Hamiltonian for the QM region. The only drawback of this scheme is the monopolarisation effect of the MM region onto the QM region, rather than both directions. The involvement of MM point charges in the

QM Hamiltonian is realised through an external potential term. This additional term again acts as point charges if their charges are considered as nuclear charges. This enables an electric field to be generated which results in polarisation of these additional point charges towards the QM electron density region. Another drawback with either this method or mechanical embedding is that only Coulomb's law is implied in the treatment of electrostatics, thus dispersion and potential Pauli repulsion terms are neglected[51]. Unlike mechanical embedding, there are now two terms in the overall Hamiltonian of electrostatic embedding. The QM region is handled with a nuclear charge calculation of all nuclei present in the system alongside the overall dispersed electron density across the QM/MM boundary. This distributed electron density $\rho(r)$ interacts with the so called electrostatic embedding potential $\phi(r)$ (generated from the MM region) as a function of the position of the QM atoms (r_A) to give the overall energy contribution[154]:

$$E_{QM/MM} = \sum_A Z_A \phi(r_A) - \int \rho(r)\phi(r)dr \tag{5.5}$$

Subtractive, Additive and ONIOM Method

When computing the Hamiltonian of any QM/MM system, there are a variety of approaches that can be utilised dependent upon the type of system. The subtractive method is the simplest and often the most viable approach within QM/MM calculations. It is versatile and can be compiled with any QM or MM software whilst also ensuring that each interaction is easily calculated without needing to alter the code. The 'Our own N-layered Integrated molecular Orbital and Molecular mechanics' ('ONIOM') scheme is a subset of this method and thus can utilise many different computational methods including an array of DFT methods within the QM part of the calculation[191]. Thus in the subtractive scheme, the Hamiltonian is expressed as:

$$\hat{H}_{QM/MM} = \hat{H}_{MM}(MM + QM) + \hat{H}_{QM}(MM) - \hat{H}_{MM}(QM) \tag{5.6}$$

The additive approach in contrast subsumes the QM region within the MM region and thus is additive in that the total Hamiltonian of the system includes the Hamiltonians of the QM and MM regions respectively plus the interaction energies between the two systems. To account for this summation, an additional coupling parameter is included in the computation so that the total Hamiltonian is represented by all of these three terms. Another critical difference is that the Hamiltonians of the QM and MM region are calculated explicitly, unlike the subtractive method. Both these methods make use of embedding strategies to properly deal with non-covalent interactions and the covalent bonds which cross the QM/MM boundary[217]. Thus, in the additive scheme, the Hamiltonian is thus expressed:

$$\hat{H}_{QM/MM} = \hat{H}_{QM}(QM) + \hat{H}_{MM}(MM) - \hat{H}_{QM-MM}(QM + MM) \tag{5.7}$$

The ONIOM method is one of the most widely applied techniques for studying biomolecular systems within a QM/MM approach. The ONIOM calculation is a hybrid, multi-layered approach which uses different levels of MM and QM theory for different layers of computation within the overall system. This method was first proposed in 1995 by Maseras et al[37]. They reported an 'Integrated molecular orbital (IMO) and Molecular Mechanics (MM)' called IMOMM, which served as the primary source of ONIOM methods. The initial calculation was one where the total Hamiltonian of the system, in a subtractive manner, was applied for a two layered system, one with an inner QM region and an outer MM region. The Hamiltonian of the system was calculated as the product of the QM Hamiltonian of the model system and the MM Hamiltonian of the real system subtracted from the MM region of the model system. This subtraction simply negates any possible double MM interactions not included in the Hamiltonian. Absence of the additional Hamiltonian coupling parameter in this scheme ensures the QM-MM interactions are not over or under represented[37]. For context, the real system reflects all atoms within the MM region and the model system reflects all atoms treated with QM theory. Thus the overall equation for a 2-layered ONIOM calculation, assuming the QM region is the high and the MM region the low level would be[216]:

$$\hat{H} = \hat{H}_{\text{high,model}} + \hat{H}_{\text{real}} - \hat{H}_{\text{low,model}} \tag{5.8}$$

Once the two-layered system is known, where the inner layer reflects the small reactive QM region and the outer layer reflects the surroundings and solvent, the total potential energy of the system (QM and MM regions) must be well defined; particularly in light of link atoms which circumvent the dangling atoms across the boundary problem. This geometry optimisation essentially calculates the energy gradient of the ONIOM calculation and thus has wide applicability in the calculation of transition states, reactive intermediates and the overall potential energy surface for a chemical process. Within the equation, there includes the Jacobian matrix (J), this mirrors the link atoms onto the link-connected atoms within the system. These additional link atoms are merely a function of the overall geometry and thus are captured well by the energy optimisation. The term q represents a vector which fully defines all atomic coordinates and the $q - QM$ and $q - MM$ reflects the Cartesian coordinates of all atoms within the QM and MM regions respectively[45]:

$$\frac{E^{\text{ONIOM}}}{\partial q} = \frac{E^{\text{real,MM}}}{\partial q} + \frac{E^{\text{model,QM}}}{\partial q} J_{QM} - \frac{E^{\text{model,MM}}}{\partial q} J_{MM} \tag{5.9}$$

Finally, the three-layered ONIOM scheme is an example of a more complex multi-layered approach compared to the simpler two-layered ONIOM scheme. This variation on the two-layered scheme now includes an additional 'intermediate' level, where the 'real' term encapsulates both the real and intermediate compartments. Furthermore, the levels of theory are now extended to low, medium and high, with an increase in DFT theory from low (MM) to medium (semi-empirical) to high (*ab initio* DFT). These three layers are versatile in being able to control further the number of atoms which can be accounted for in different levels of QM theory. This ensures that the computational expense is minimal as it limits the number of atoms contained within the full *ab initio* DFT region[203]. In my project, I have employed both a two-layered and three-layered ONIOM scheme (figure 5.1). In the former scheme the drug and the interacting lipid are placed in the inner layer at PM6 level

of theory, whilst the surrounding lipids and water are treated with classical MM force field. In the latter scheme the interacting nucleophile of the drug and the carbonyl group of the lipid are placed in the inner layer at the highest level of theory (B3LYP/PBE0), the middle layer contains the connecting parts of the single drug and lipid, placed in the medium level of theory (PM6) and the outer layer constitutes the surrounding membrane including solvent (MM). Mathematically, this three-layered calculation is expressed as:

$$E_{\text{ONIOM}} = E_{\text{high,model}} + E_{\text{medium,intermediate}} - E_{\text{medium,model}} + E_{\text{low,real}} - E_{\text{low,intermediate}} \tag{5.10}$$

Figure 5.1: 2-layered ONIOM method (left), 3-layered ONIOM method (right)

5.2.2 Quantum Chemistry

The aim of quantum chemistry/electronic structure theory is to be able to solve the time-independent Schrödinger equation (TISE) for any molecular system. This equation encompasses the dynamics of all nuclei and electrons present within the system, exclusive of any physical perturbations such as the effect of external electric or magnetic fields[107]. In equation 5.11, the Hamiltonian operator H is applied to the eigenfunction ψ, being identical to the product of the eigenvalue E and eigenfunction ψ, thus making it a solvable eigenvalue problem. For this project, where the electronic structure is in the ground state, we are interested in computing this equation for the reactive part of the system. With the aid of the Born-Oppenheimer approximation, one can take advantage of the nuclei being far larger in mass and makes the approximation that they can be considered as fixed/clamped in space. The kinetic energy of the nuclear component is zero and the nuclear-nuclear repulsions are also negated. This simplifies the equation to the so called 'electronic

Schrödinger equation', where the total Hamiltonian of the system is now dependent upon fewer terms[31]. The Hamiltonian then of the TISE would be:

$$\hat{H}\psi = E\psi \tag{5.11}$$

$$\hat{H}_{elec} = \hat{T} + \hat{V}_{ee} + \hat{V}_{ne} \tag{5.12}$$

$$\hat{T} = \frac{\hbar^2}{2m}\nabla^2 \tag{5.13}$$

$$\hat{V} = V(\boldsymbol{r}) \tag{5.14}$$

where T represents the kinetic energy operator, V the potential energy operator and ee and ne the electron-electron and nuclear-electron interactions respectively. Thus the T operator is defined from the Laplacian operator ∇^2 and the V operator is defined from the position operator $\boldsymbol{r}$[40].

With this in mind, the Born-Oppenheimer approximation can assume that the relaxation time of electronic motion relative to nuclear motion is instantaneous. Thus, these two components can be segregated and only the electronic energies of the clamped nuclei must be considered. As the nuclei are moving far slower than the electrons, their kinetic energy can be separated from the electrons, alongside the negation of any potential energy attractive term between the electrons and nuclei and lastly the repulsive potential energy term between the nuclei is simplified to a constant in the equation. This beautifully results in the complete electronic Schrödinger equation, where i reflects the electronic coordinates, j the nuclear coordinates, V_n is the nuclear-nuclear repulsion term (constant) and the eigenvalue of the equation is now the electronic energy (E_{el})[41]:

$$\hat{H}_{\mathrm{el}}\psi_{\mathrm{el}}(q_{\mathrm{i}}q_{\mathrm{j}}) = E_{\mathrm{el}}\psi_{\mathrm{el}}(q_{\mathrm{i}}q_{\mathrm{j}}) \tag{5.15}$$

Density Functional Theory

The exact solution to the electronic structure for a 1-electron hydrogen atom can be easily solved by the time-independent Schrödinger equation. Once the system increases in complexity to an N-body problem where N reflects any electronic structure higher than a single electron, the computation of this becomes significantly harder to solve. The associated electronic wavefunction becomes ever increasingly difficult to solve as the molecular system increases in size[205]. Density Functional Theory (DFT) is an approximation which calculates the electron density rather than deal with the overly complex electronic wavefunction. Fortunately, this greatly reduces the complexity of the system, and multiple DFT theorems have been proposed over the last few decades. The first was in 1964 where Hohenberg and Kohn suggested that a direct functional could be related to the electron density. They proposed allowing a generic box with a given electronic structure to be subjected to an external potential and the evident repulsion between the electrons, this external potential $v(\boldsymbol{r})$ would then be the functional of the electron density $n(\boldsymbol{r})$. The theorem was subjected to further testing later and found to be suitable for considering the electronic structure in both degenerate and non-degenerate systems[12]. As before the total Hamiltonian of the system is calculated as:

$$H = T + U + V \tag{5.16}$$

$$V = \int v(\boldsymbol{r})\psi^*(\boldsymbol{r})\psi(\boldsymbol{r})d\boldsymbol{r} \tag{5.17}$$

The potential energy V is calculated as an integral of the sum of complex wavefunction ψ^* with the other terms. They found that the product of T (kinetic energy) and U (electron-electron interaction operator) was merely a universal functional $F[n]$ of the charge density, similarly to the wavefunction ($\psi(r)$) being a functional of the electron density $n(r)$:

$$F[n(\boldsymbol{r})] = (\psi(T + U)\psi) \tag{5.18}$$

$$E_v[n] = \int v(\boldsymbol{r})n(\boldsymbol{r})d\boldsymbol{r} + F[n] \tag{5.19}$$

$$N[n] = \int n(\boldsymbol{r})d\boldsymbol{r} = N \tag{5.20}$$

Equation 5.20 is groundbreaking in that it accurately predicts that the functional will reach an integral minimum, for the electron density in the ground state $n(\boldsymbol{r})$, if the total number of particles N are kept constant. Thus this functional is entirely dependent upon the electronic density $(n(\boldsymbol{r}))$[67].

A year later, Kohn and Sham discovered an approximation to the Hohenberg-Kohn equations, where the energy functional could be re-written to include an exchange correlation term known as the Local Density Approximation (LDA). This additional term, $E_{xc}[n]$, accounts for both the exchange and correlation energies of the many-electron interactions present in the system[108]. This new approach was applied to a uniform homogeneous gas to obtain the electron exchange correlation energy, using the so called LDA approximation[163]:

$$E = \int v(\boldsymbol{r})n(\boldsymbol{r})d\boldsymbol{r} + \frac{1}{2}\iint \frac{n(\boldsymbol{r})n(\boldsymbol{r}')}{|\boldsymbol{r} - \boldsymbol{r}'|}d\boldsymbol{r}d\boldsymbol{r}' + T_s[n] + E_{xc}[n] \tag{5.21}$$

$$E_{xc}[n] = \int n(\boldsymbol{r})\epsilon_{xc}(n(\boldsymbol{r}))d\boldsymbol{r} \tag{5.22}$$

where $T_s[n]$ is defined as the kinetic energy of the non-interacting electrons with electron density $n(\boldsymbol{r})$, $E_{xc}[n]$ is the exchange correlation energy of the system of interacting electrons also with electron density $n(\boldsymbol{r})$. In equation 5.22, the $E_{xc}[n]$ is computed from both the exchange and correlation energy ($\epsilon_{xc}(n(\boldsymbol{r}))$ per electron of density n.

DFT Functionals and Semi-Empirical Methods

The usefulness of the Kohn-Sham DFT equations rest exclusively on the performance of the exchange-correlation functionals. Each approximation to $E_{xc}[n]$ is categorised based upon its characteristics. The simplest is the LDA, whose functionals are only

computed for the electron density. The next are the GGA functionals, which applies an additional density gradient to the LDA functional. Meta-GGA functionals, which are an improvement on GGA by the inclusion of a kinetic energy density term. Hybrid-GGA functionals which include a fraction of additional Hartree-Fock exchange in an equal balance with the remaining functional[99]. The final class are the double-hybrid-GGA functionals. Perdew cleverly devised a scheme based upon 'Jacob's Ladder' from the Biblical Old Testament[33], where instead of angels ascending and descending the ladder between Earth and Heaven, he placed DFT functionals at each rung of the ladder with increasing chemical accuracy as the functionals progressed from a lower to an upper rung. This is helpful in assessing the numerical efficiency of each functional. A pattern emerges such that the addition of a Hartree-Fock exchange correlation functional improves the functional performance significantly, but with added computational expense. Whilst all hybrid and double-hybrid functionals possess this term, all functionals of lower chemical accuracy do not. Perhaps the functionals with the best balance of accuracy versus speed are the semi-local meta-GGA functionals[32].

The two main sources of error with any DFT functional calculation are the self interaction error (SIE) and the lack of long-range dispersion. The latter is mitigated by the addition of a dispersion correction term to the functional, typically this would be the Grimme's D3 dispersion correction for B3LYP or for PBE0 functionals[78]. The SIE error is harder to mitigate as this error is inevitable due to the self-Coulombic interactions present between all electrons in the system. The SIE causes an over-delocalisation of electron density and thus affects semi-local functionals. Hybrid and double-hybrid functionals have the advantage of additional Hartree-Fock exchange correlation energy, thus can reduce the effect of this error[159].

For my project, I have compared the computational and numerical efficiency of the B3LYP and PBE0 hybrid DFT functionals, alongside the semi-empirical PM6 calculation. As a hybrid functional, the B3LYP functional is calculated from a semi-empirically derived Hartree-Fock exchange energy (accounting for 20%) mixed with a standard GGA exchange correlation energy. Parameters for these approaches are

obtained in a semi-empirical way by optimising calculations against experimental data such as ionisation potentials for a small group of molecules[198]. The term B3LYP derives from the 1988 Becke GGA functional with Lee-Yang-Parr correlation function[167]. This exchange energy comprises of an LDA exchange term (E_X^{LDA}), a Becke88 gradient correction exchange term ($\Delta E_X^{\mathrm{B88}}$) and of course the Hartree-Fock exchange (E_X^{HF}):

$$E_X^{\mathrm{B3LYP}} = 0.8E_X^{\mathrm{LDA}} + 0.2E_X^{\mathrm{HF}} + 0.72\Delta E_X^{\mathrm{B88}} \tag{5.23}$$

The additional correlation energy again includes the Lee-Yang-Parr correlation function (E_C^{LYP}) with the added Vosko–Wilk–Nusair III correlation energy (E_C^{VWN3}) to transform the functional from BLYP to B3LYP. This B3LYP is thus semi-empirical in nature, and all coefficients of the exchange correlation energy are fitted with experimental data obtained from atomisation energies and ionisation potentials of the molecular species in question. The exchange energy is summed with this correlation energy to give the final B3LYP functional:

$$E_C^{\mathrm{B3LYP}} = 0.19E_C^{\mathrm{VWN3}} + 0.81E_C^{\mathrm{LYP}} \tag{5.24}$$

The PBE0 functional, which is also a hybrid functional, is somewhat semi-empirical in its derivation. It is formed by 25% mixing of the Hartree-Fock exchange energy with 75% of the PBE exchange (GGA functional). The term PBE0 derives from the Perdew, Burke and Ernzerhof correlation functional (GGA functional) alongside again a Hartree-Fock exchange term. All the terms in the PBE exchange term are basic constants with pre-determined values (a=0.00449, b=0.00336, and x capturing the electron density and its associated gradient)[1]. This ensures the PBE functional is versatile and of good chemical accuracy to be compared with other GGA functionals, thus yielding the exchange term:

$$E_X^{\mathrm{PBE}} = \frac{bx^2}{1+ax^2} \tag{5.25}$$

Once this exchange energy is accounted for, the correlation energy must be computed to obtain the PBE0 functional exchange-correlation energy. Alongside the PBE exhange energy, the HF exchange and PBE exchange-correlation energy must be calculated. These latter two terms are combined and mixed, along with an empirical parameter, to obtain a so called exchange-correlation term known as the ACM1 method[181]. This yields the ACM1 exchange correlation energy with quite high accuracy:

$$E_{XC}^{\mathrm{ACM1}} = E_{XC}^{\mathrm{GGA}} + a1(E_X^{\mathrm{HF}} - E_X^{\mathrm{GGA}}) \tag{5.26}$$

Perdew et al. then proposed that the a1 empirical parameter could be approximated to 1/4 based upon a fourth order adiabatic term which ensures accurate results and became the ACM0 method. The 3:1 mixing ratio of Hartree-Fock to PBE exchange energy is derived from this perturbation. This PBE0 functional is fully non-empirical and a result of three terms; the Hartree-Fock exchange energy (E_X^{HF}), the PBE correlation energy (E_C^{PBE}) and the PBE exchange energy (E_X^{PBE})[136]. Thus the exchange-correlation for the ACM0 is fed directly into the PBE0 exchange correlation energy term:

$$E_{XC}^{\mathrm{ACM0}} = E_{XC}^{\mathrm{GGA}} + \frac{1}{4}(E_X^{\mathrm{HF}} - E_X^{\mathrm{GGA}}) \tag{5.27}$$

$$E_{XC}^{\mathrm{PBE0}} = \frac{1}{4}E_X^{\mathrm{HF}} + \frac{3}{4}E_X^{\mathrm{PBE}} + E_C^{\mathrm{PBE}} \tag{5.28}$$

The final comparison of these two hybrid-DFT methods is employed in this project along with the quite successful semi-empirical PM6 method. Historically, the classic AM1 (Austin Model 1) and PM3 (Parametric Method 3) have been used to model simple inorganic and organic systems with relatively good results. However recently, the so called PM6 (Parametric Method 6) has been developed to compute much

larger biochemical systems, where former methods have resulted in poor representation of the molecular geometry and resulted in poor computational efficiency[200]. This state-of-the-art method is a type of NDDO semi-empirical method which is a modification of the typical NDDO type by an additional core-core diatomic interaction parameter, as well as a d-orbital correction term[199]. This method had been applied to the parameterisation scheme of more than 80 chemical elements including many metals. Its wide applicability in large scale simulation makes it an effective alternative to standard DFT methods. Its main problem is the lack of well-defined non-covalent interactions in the system, or dispersion interactions. However, this issue is easily solved with an additional dispersion correction term, usually represented as PM6-D in a calculation[200]. This parameter is essentially a pairwise interatomic potential shown mathematically as:

$$\sum E_{\text{dispersion}} = -\sum f_{\text{damp}}(r_{ij}, R_{ij}^{0}) C_{6ij} \frac{1}{r_{ij}^{6}} \tag{5.29}$$

where $f_{\text{damp}}(r_{ij}, R_{ij}^{0})$ is a damping function between two atoms i and j, thus enabling no singularities at short distances, C_{6ij} are the dispersion coefficient between atoms i and j, $\frac{1}{r_{ij}^{6}}$ is the short-range attraction term as a function of interatomic distance between atoms i and j.

5.2.3 Simulation Details

QM/MM Simulations were performed with the QUICK program in AMBER23 and AmberTools24 software. This QUICK program was able to compute DFT functionals with the chosen basis set and CPU dispersion correction. The semi-empirical PM6 method was employed within Gaussian 16 software running on Hamilton8 whilst *ab initio* DFT calculations were computed using QUICK GPU software running on Bede. Initial starting structures were extracted from unbiased MD simulations in the *NPT* ensemble at 200 ns. A python script was written to calculate the distance and angle between each interacting nucleophile and the carbonyl carbon of the lipid headgroup, calculated across 10,000 frames of the trajectory. For 2-aminomethylbenzimiazole, this was the distance and angle between the nucleophilic

nitrogen atom and the carbonyl carbon of the lipid (see figure 5.2). For the drug propranolol, this was the distance and angle between the nucleophilic oxygen atom and the carbonyl carbon of the lipid (see figure 5.3). A python script was written to calculate all distances and angles between these interacting atoms. The normalised distance and angle distributions were then plotted as 1D histograms. This helped identify an initial starting structure from a single frame of the trajectory. These MD snapshots are extracted from the entire membrane-water system for visual simplicity (see figures 5.2-5.7). This frame would be close to the optimal Bürgi-Dunitz distance and angle for a typical nucleophilic attack onto a carbonyl.

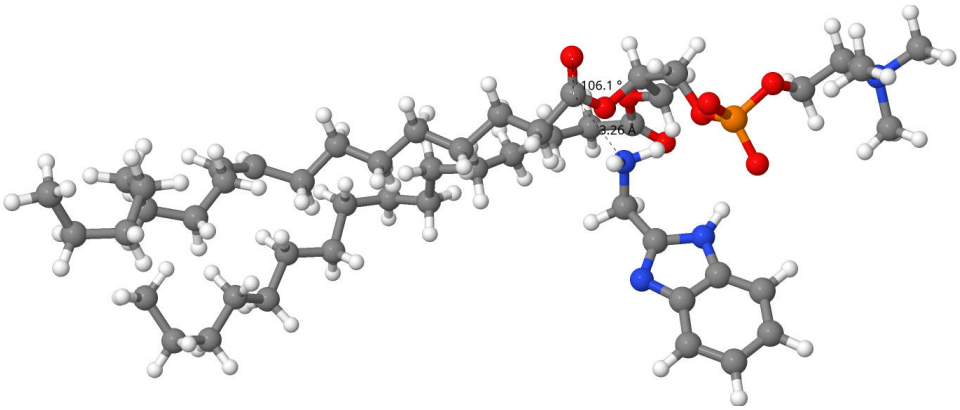


Figure 5.2: MD snapshot of the interaction of the nitrogen atom (blue, 2-aminomethylbenzimidazole) with the carbonyl carbon atom (grey, POPC lipid)

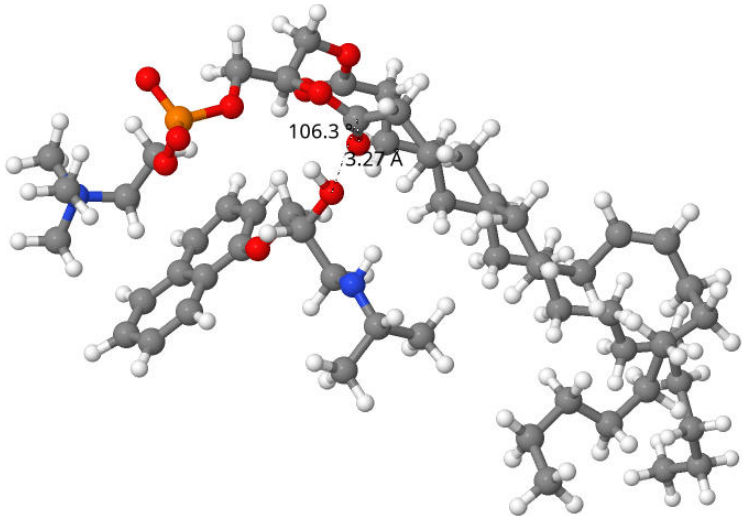


Figure 5.3: MD snapshot of the interaction of the oxygen atom (red, propranolol) with the carbonyl carbon atom (grey, POPC lipid)

From the initial structure of the drug-membrane system, a series of reaction coordinates were produced from a bash script, which were a function of a pre-defined restraint set up between the two reactive atoms. A force constant of 100 kJ mol^{-1}

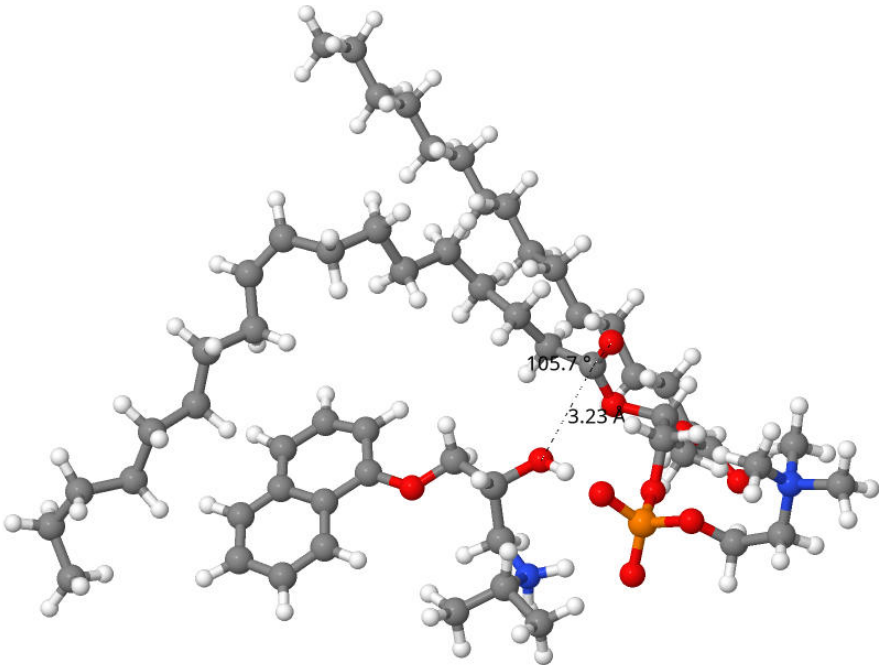


Figure 5.4: MD snapshot of the interaction of the oxygen atom (red, protonated propranolol) with the carbonyl carbon atom (grey, POPC lipid)

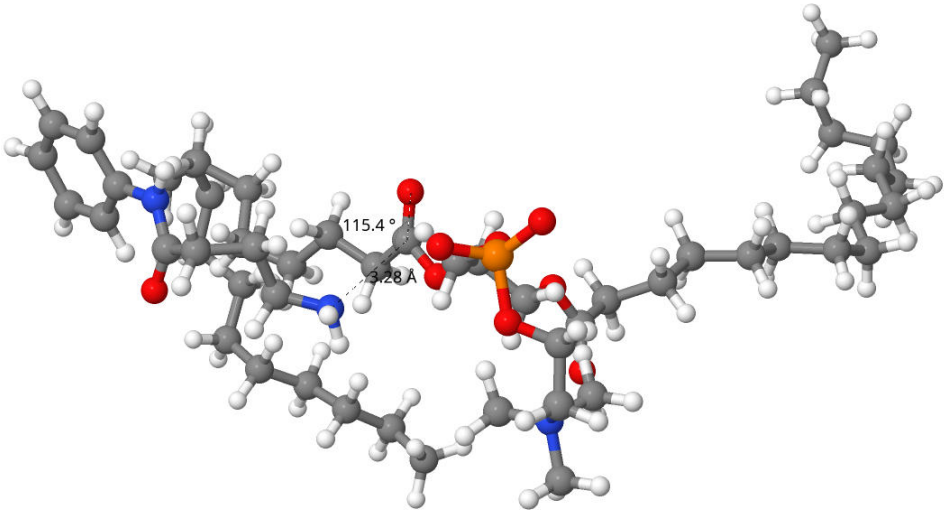


Figure 5.5: MD snapshot of the interaction of the nitrogen atom (blue, 4-amino-*N*-phenylbutanamide) with the carbonyl carbon atom (grey, POPC lipid)

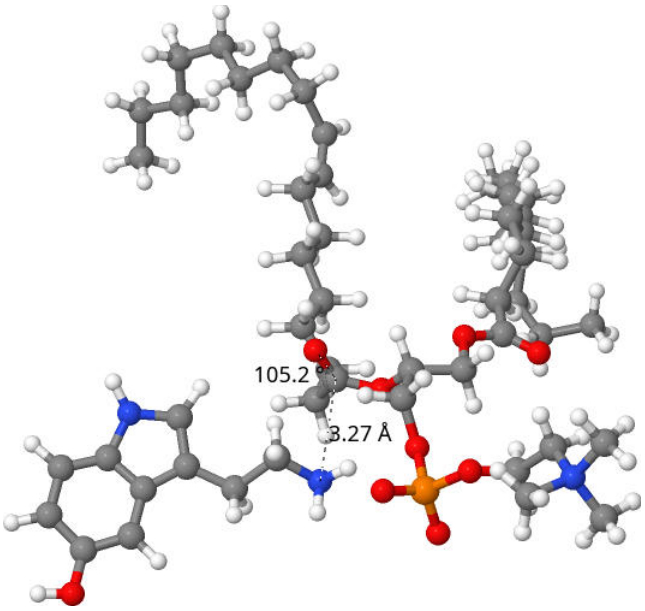


Figure 5.6: MD snapshot of the interaction of the nitrogen atom (blue, serotonin) with the carbonyl carbon atom (grey, POPC lipid)

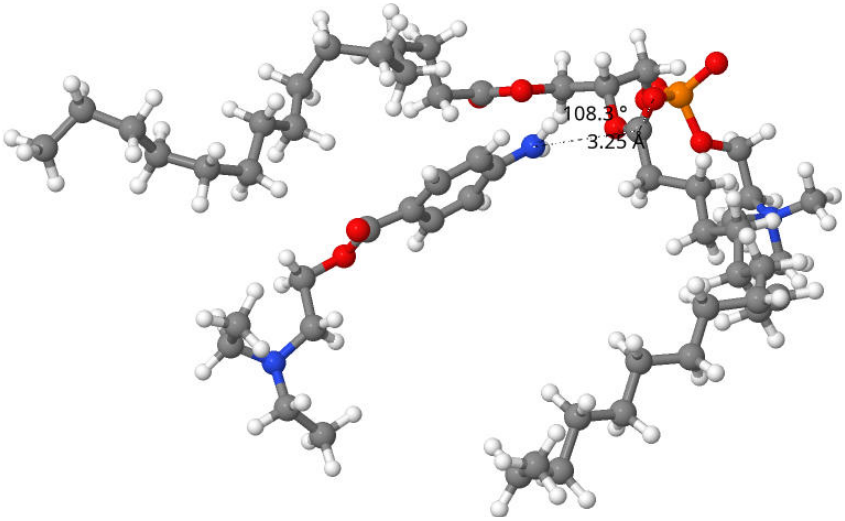


Figure 5.7: MD snapshot of the interaction of the nitrogen atom (blue, procaine) with the carbonyl carbon atom (grey, POPC lipid)

nm^2 was used to bring the atoms close enough together to form a chemical bond. Preliminary QM/MM dynamics were run at semi-empirical PM6 level. From each of these starting structures, 150-170 atoms were placed in the QM region (see figures 5.2-5.7) and the remaining membrane and solvent were placed in the MM region with AMBER force fields. This system was simulated for 1, 10 and 30 ps at 310 K in the *NPT* ensemble with and without a dispersion correction term (PM6-D). The corresponding DFT calculations were run at the B3LYP/def2-SVP and at the PBE0/def2-SVP level, where around 10 atoms were placed in the QM region with electrostatic embedding to account for the QM/MM boundary, and the remaining atoms were placed in the MM region handled with AMBER force fields. These systems were simulated for 10 and 30 ps at 310 K in the *NPT* ensemble with a Grimme-D3 dispersion correction term (B3LYP-D3 and PBE0-D3).

Once the simulations were complete, the so called 'WHAM' method was utilised to construct the free energy profile and its associated histogram[79] (as a measure of sampling quality). The free energy barriers were calculated and the corresponding reactive intermediate structures were obtained from the barrier heights. In each system, the free energy profile corresponded to the nucleophilic reaction of the nitrogen/oxygen atom of the drug with the carbonyl carbon of the lipid across a reaction coordinate. The minimum in each energy profile was then extracted and subjected to energy minimisation at PM6+water in SCIGRESS. This intermediate was then energy minimised again with PBE0 and B3LYP functionals in implicit solvent. In accordance with each reaction pathway (aminolysis or transesterification), subsequent T^0 (neutral), T^- (deprotonated), and acylated forms of each drug were energy minimised (figure 5.8) in an identical manner to the $T^{\pm}$ intermediate (minimum). These energies were then tabulated and the relative free energies between different reactive intermediates were calculated to establish the overall drug-lipid reactivity.

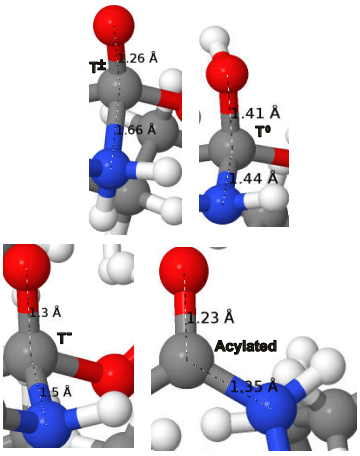


Figure 5.8: Structures obtained from DFT energy minimisation at the PBE0 level. Only the reactive site and surrounding atoms are shown.

5.3 Results and Discussion

5.3.1 Histograms of distances and angles

To perform accurate QM/MM calculations on an initial starting structure with good potential for chemical reactivity, the distance and angle between the nucleophile (drug) and electrophile (lipid) had to be calculated. The 'Bürgi-Dunitz' angle for organic reaction mechanisms is a good measure of potential reactivity, as it takes crystallographic values for bond distances, derived ultimately from van der Waals radii of overlapping atoms, coupled with bond angles to predict the likelihood of reactivity between a nucleophile and electrophile[174]. A python script was written to calculate the individual distance and angle between each interacting nucleophile of the drug molecule and the carbonyl carbon electrophile of the lipid. Having assumed that reactivity would preferably take place on the carbonyl carbon connected to the unsaturated, rather than the saturated, fatty acyl chain, the script calculated the distance between either the interacting nitrogen or oxygen atom of the nucleophilic drug with the C21 residues of the PC lipid headgroups. The same script calculated individual angles between the nucleophile, electrophile and carbonyl oxygen connected to the nucleophile. Once all distances and angles had been calculated across the 10,000 frames of the trajectory, the results were plotted as a normalised histogram. The histogram was useful in identifying the distribution of distances and angles across the whole molecular dynamics trajectory and thus concluding whether the respective drug molecule had the potential for chemical reactivity.

The typical Bürgi-Dunitz (BD) angle for a nucleophilic addition reaction of an S_N2 type, from a nucleophile to a carbonyl, is around 107° ± 5°. The typical distance for nucleophilic attack depends on the summed van der Waals radii for the interaction atoms[160] (figure 5.9). For a nitrogen nucleophilic drug, the interaction distance would be 3.25 Å, for an oxygen nucleophilic drug, the distance would be 3.22 Å. Thus, these histograms are indicative of the frequency of distances and angles closest to these optimal values for reactivity.

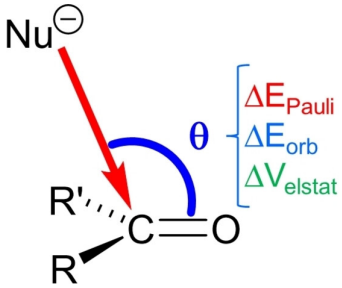


Figure 5.9: Diagram of the BD distance and angle for reactivity[174]

As is evident in figure 5.10 for propranolol, its highest probability is found between 3.5-4 Å and θ = 20-30°. Whilst the O-C distance is plausible for reactivity, the angle range is quite far from the optimal. Compared with protonated propranolol in figure 5.11, with the nitrogen atom being protonated, the highest probability of distances is still 3.5-4 Å, but now the angle range of 100-110° is at a higher probability (0.0075) than its neutral analogue (0.003). This distance and angle range are both very close to the optimal BD conditions and thus the protonated form has a higher potential for reactivity. We know from experimental literature[54] that protonated propranolol is the major form and neutral propranolol is the minor form. Protonation thus provides greater reactivity as the interacting oxygen is positioned closer to the lipid head group region for lipidation.

As can be seen in figure 5.11, protonated 2-aminomethylbenzimidazole exhibits no reactivity with the membrane. The histogram plot shows that the highest frequency of distances is around 5.5-6 Å and the highest frequency of angles is around 130-140°. Both the distance and angle range are vastly different from the BD conditions and thus its protonated form is unlikely to act as a nucleophile. Neutral 2-aminomethylbenzimidazole does act as a strongly interacting

nucleophile. Its highest distance distribution is between 3.5-4 Å, whilst its highest angle distribution is between 95-105°. Combined, these distances and angles place 2-aminomethylbenzimidazole in a good range which correlates well with the BD conditions required for reactivity.

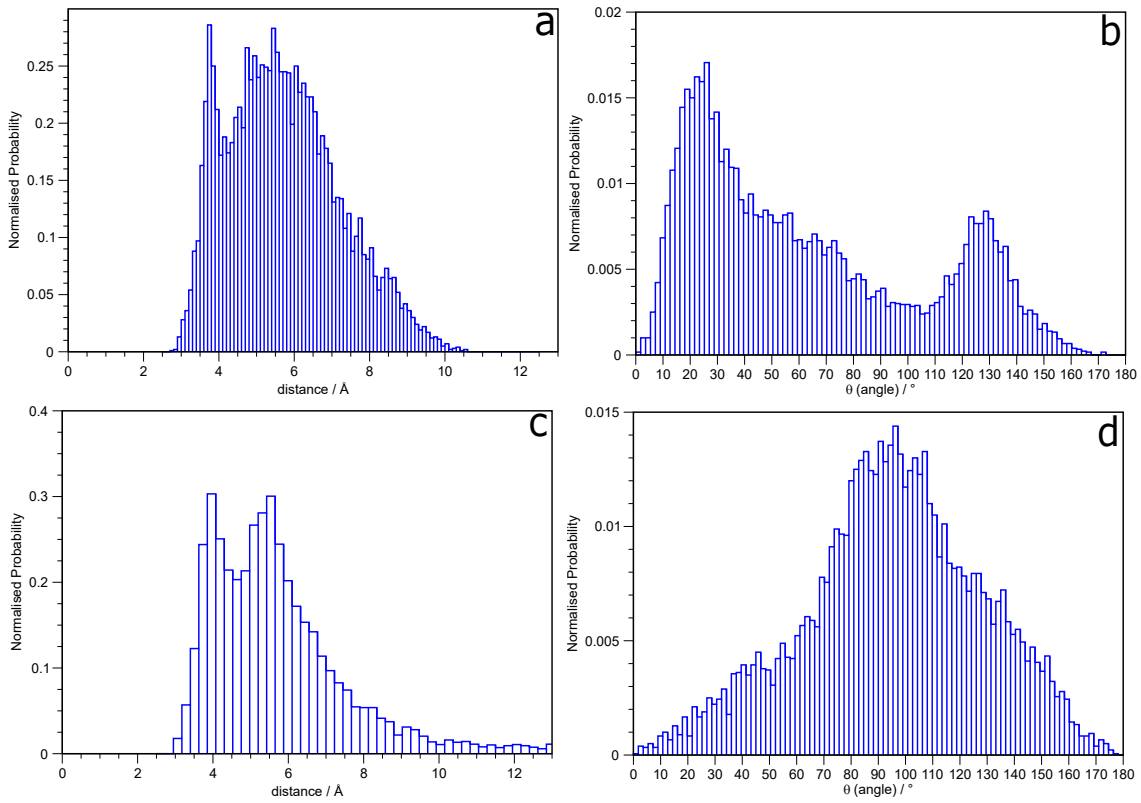


Figure 5.10: Normalised histograms of distances and angles for a,b) propranolol, c,d) 2-aminomethylbenzimidazole

As can be seen in figure 5.12 for fluoxetine, the highest distance distribution is in the range 5-5.5 Å and the highest angle distribution in the range 150-160°. These values suggest limited reactivity, and correlate with experimental data which confirms lipid hydrolysis behaviour once in the membrane. Similarly its protonated form (figure 5.13) confirms no lipidation behaviour. Furthermore, as phentermine exhibits no reactivity with the membrane, both the neutral and protonated forms (figures 5.12, 5.13) on the histogram plots indicate the farthest distance and angle from the optimal BD conditions of all drug-membrane systems.

4-amino-*N*-phenylbutanamide, serotonin and procaine have all been experimentally verified as molecules which undergo direct lipidation with DOPC membranes[25].

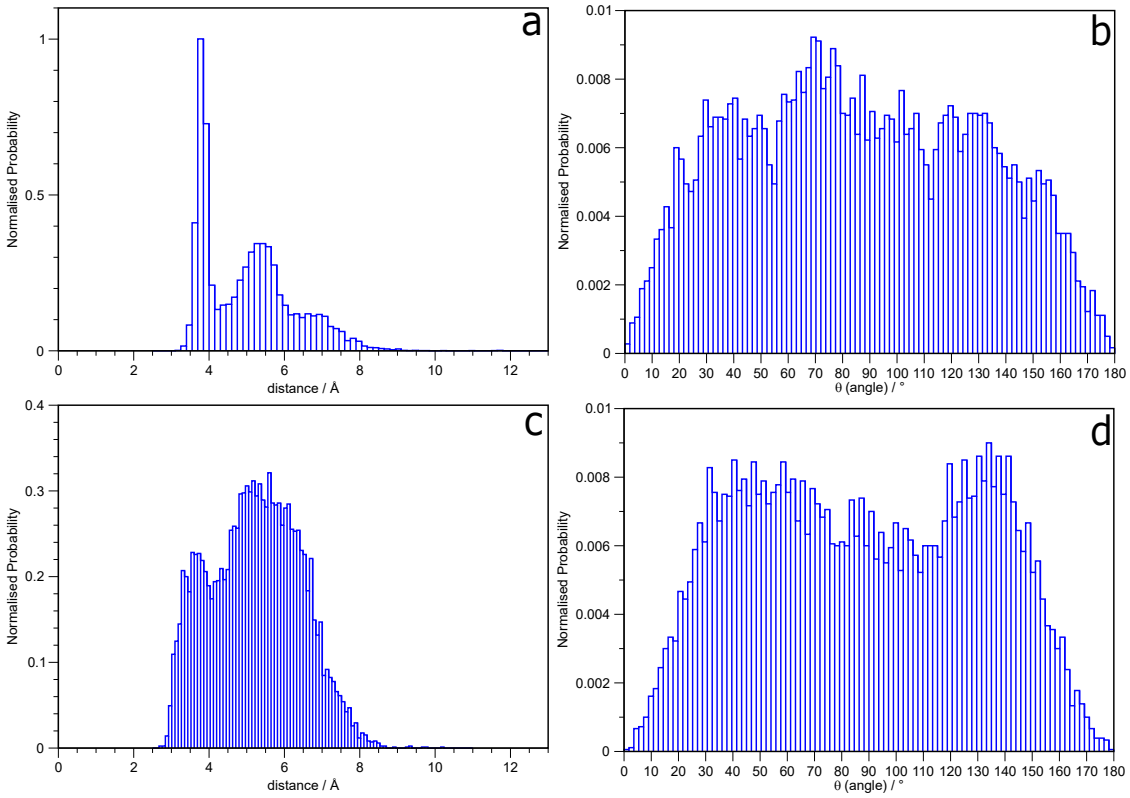


Figure 5.11: Normalised histograms of distances and angles for a,b) protonated propranolol, c,d) protonated 2-aminomethylbenzimidazole

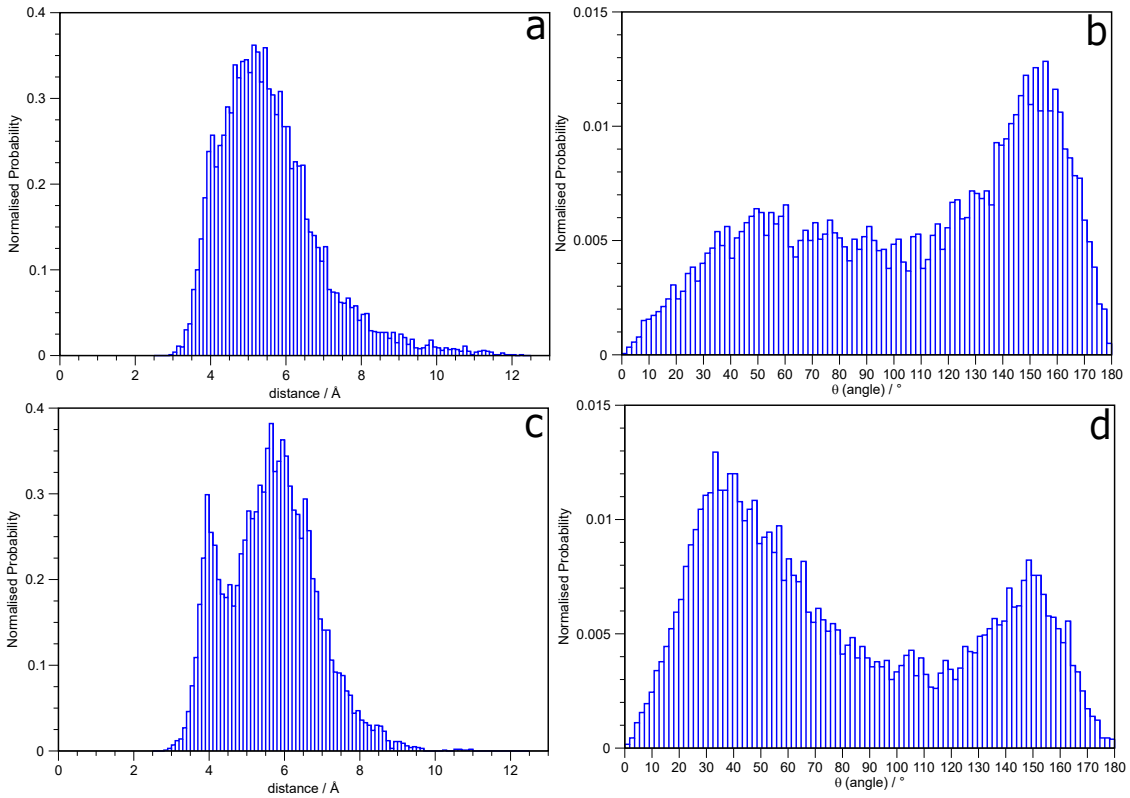


Figure 5.12: Normalised histograms of distances and angles for a,b) fluoxetine, c,d) phentermine

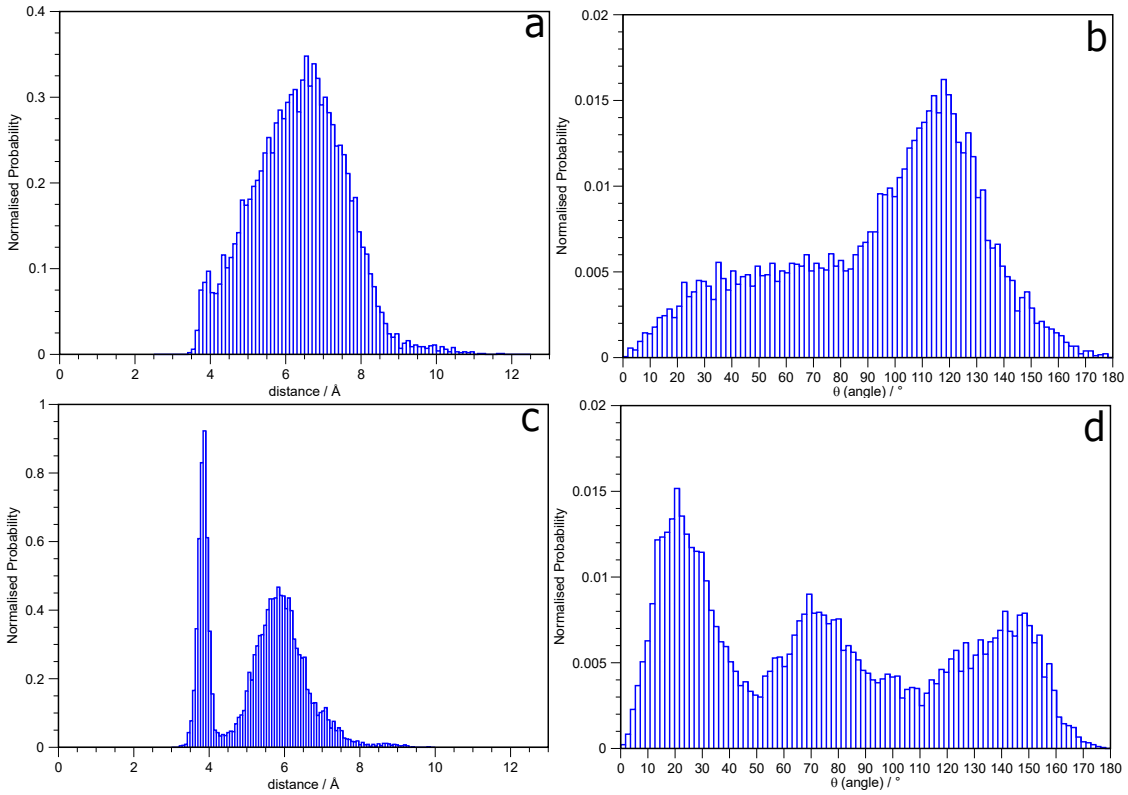


Figure 5.13: Normalised histograms of distances and angles for a,b) protonated fluoxetine, c,d) protonated phentermine

Neutral 4-amino-*N*-phenylbutyramide has its highest frequency distance between 5.5-6 Å and highest angle range between 130-150° (figure 5.14). Thus, although one can extract appropriate trajectories, the optimal BD conditions are less frequent in this trajectory than for protonated propranolol and 2-aminomethylbenzimidazole molecules. This potentially suggests that its reactivity from the nitrogen nucleophile may be lower than 2-aminomethylbenzimidazole. Its protonated form (figure 5.15) likewise has its highest distance and angle distribution far away from the optimal BD conditions.

Likewise serotonin has the potential to act as a nucleophile in the reaction. Although there is a good distribution of distances between 3.5-4.0 Å for reactivity, the highest probability of angles are found in the range 40-50 and 140-150°. Thus, although its distance is close to the optimal BD distance, its highest angle distribution is quite far off. Evidently from figure 5.15, its protonated form has completely the wrong geometrical orientation for nucleophilic attack, and subsequently there were no frames with the optimal angle for reactivity.

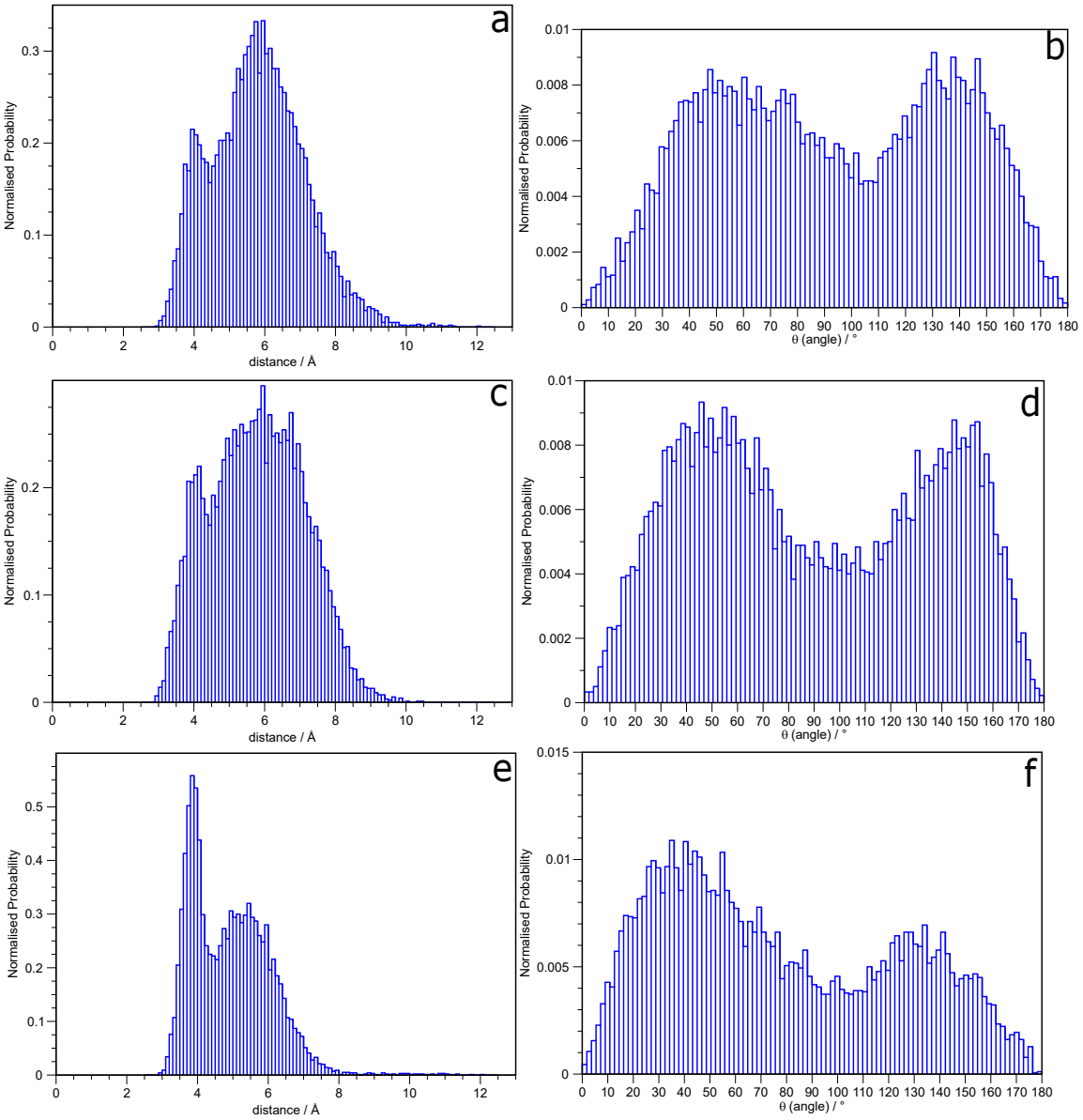


Figure 5.14: Normalised histograms of distances and angles for a,b) 4-amino-*N*-phenylbutanamide, c,d) serotonin, e,f) procaine

Finally, as can be seen in figure 5.14, procaine had the highest probability of distances in the range 3.5-4 Å of the three systems presented here. However, its angle distribution around 100-110° was slightly lower than both serotonin and 4-amino-*N*-phenylbutanamide. This implies that the geometrical orientation of the molecule is less favoured for nucleophilic attack than 4-amino-*N*-phenylbutanamide. Whilst the angle range was quite far from the BD conditions, an optimal distance and angle was still found from a configuration. Protonated procaine in figure 5.15, shows a distance distribution quite far away from the optimal BD conditions, with an incredibly varied angle distribution. This would be expected for a protonated molecule.

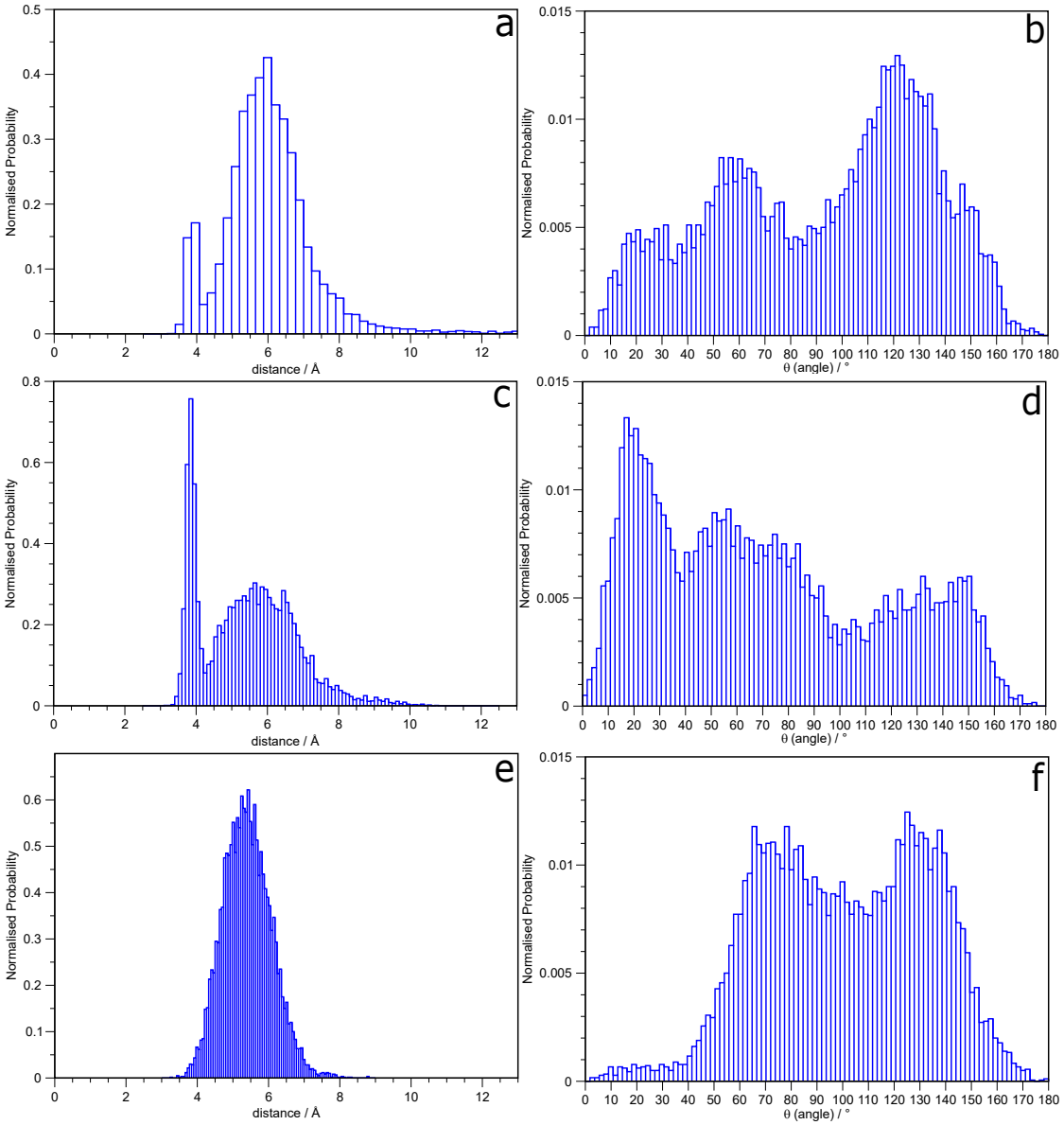


Figure 5.15: Normalised histograms of distances and angles for a,b) protonated 4-amino-*N*-phenylbutanamide, c,d) protonated serotonin, e,f) protonated procaine

5.3.2 QM/MM Systems: General Protocol

The initial distance and angle between the interacting nucleophilic atom of the drug and lipid carbonyl atom were obtained for each system. Initial starting structures were obtained from the 1D histogram structures, which correlated with bond distances and angles close to the BD optimal distance and angle of attack. The initial set of QM/MM simulations were all calculated with PM6 level of QM theory. Each simulation was run for 30 ps and for 10 ps with a Grimme-D3 dispersion correction term. The atomic coordinates of each drug molecule and the interacting lipid (PA, PC, OL lipids) were all placed within the QM region (see figures 5.2-5.7), and the rest of the lipids and system were all placed within the MM region (table 5.1). Thus, this was analogous to a 2-layer ONIOM method. A force constant of 100 kcal mol^{-1} nm^2 per 0.1 Å interval was applied along a well defined reaction coordinate. This reaction coordinate was defined as the difference in distance between bond formation of the drug nucleophile and lipid carbonyl carbon and the bond breaking between the lipid carbonyl carbon and lipid carbonyl oxygen. The electronic energy was then calculated for these two processes and a reaction profile was plotted with WHAM[79].

Drug System	N or O-C bond distance / nm	C=O bond length / nm	N or O-C=O bond angle / °	No. of Atoms in QM region (PM6)	No. of Atoms in MM region (AMBER)
2-aminomethylbenzimidazole	**0.326**	0.123	**106.1**	154	39271
Propranolol	**0.327**	0.125	**106.3**	174	38860
Protonated propranolol	**0.323**	0.124	**105.7**	175	38859
4-amino-N-phenylbutanamide	**0.328**	0.126	**115.4**	161	38860
Serotonin	**0.327**	0.122	**105.2**	159	39408
Procaine	**0.325**	0.119	**108.3**	171	38997

Table 5.1: QM/MM starting conditions for drug-membrane system

The second set of QM/MM calculations retained PM6 theory, but these simulations just involved a single lipid interaction with drug in the absence of all surrounding membrane lipids. This was used as a control to understand possible membrane effects on stabilising the $T^{\pm}$ intermediate. The single lipid-drug system was extracted from the .pdb file and solvated with TIP3P water. The drug-lipid system was frozen and the explicit solvent model underwent relaxation. The water molecules were energy minimised for 10,000 steps, heated from 0 to 310 K for 50,000 steps under

constant NVT conditions, with the restraint constant reduced from 500 down to 10 kJ mol^{-1} nm^2. The whole system was then equilibrated for 1 ns to ensure system stability. The single lipid and drug system were placed within the QM region and the remaining water molecules were placed within the MM region (table 5.2), again similarly to a 2-layer ONIOM scheme. The charge on the QM region was +1 for protonated propranolol but neutral for all other systems. A force constant of 100 kJ mol^{-1} nm^2 per 0.1 Å interval was applied along a reaction coordinate. The same difference in C-Nucleophile bond forming and C-O (carbonyl) bond breaking was calculated and the electronic energy of these processes was plotted with WHAM[79].

Drug System	N or O-C bond distance / nm	C=O bond length / nm	N or O-C=O bond angle / °	No. of Atoms in QM region (PM6)	No. of Atoms in MM region (AMBER)
2-aminomethylbenzimidazole	**0.435**	0.125	**134.1**	154	7668
Propranolol	**0.324**	0.123	**107.2**	174	7854
Protonated Propranolol	**0.288**	0.124	**113.8**	175	8771
4-amino-N-phenylbutanamide	**0.393**	0.126	**125.9**	161	8094
Serotonin	**0.313**	0.124	**117.1**	159	8249
Procaine	**0.380**	0.121	**126.5**	171	7851

Table 5.2: QM/MM starting conditions for single solvated lipid-drug system

The final set of calculations again retained the same PM6 method, but this simulation was for the interaction of the drug molecule with the same lipid molecule, omitting both a phosphate headgroup and fatty acyl chain. This control was calculated to predict whether the presence of the phosphate headgroup in the membrane contributes to stabilisation of the drug-lipid interaction. In a similar vein to the last set of calculations, the drug-lipid system was extracted from the whole pdb file. Manually in SCIGRESS, the lipid headgroup and fatty acyl chain was removed and terminated with a methyl. This new system was then solvated with TIP3P water. The new system underwent relaxation of its water molecules. The water molecules were energy minimised for 10,000 timesteps, heated from 0 to 310 K for 50,000 timesteps, with the restraint constant reduced from 500 to 10 kJ mol^{-1} nm^2 for the single lipid-drug system. The whole system was then equilibrated for 1 ns to ensure system stability. The simpler lipid-drug system was placed in the QM region, and the remaining water molecules were treated with MM force fields (table 5.3).

The charge on the QM region was +1 for protonated propranolol but neutral for all other systems. Once again the force constant of 100 kJ mol^{-1} nm^2 per 0.1 Å interval was applied along a reaction coordinate. The same difference in C-Nucleophile bond forming and C-O bond breaking was calculated and the electronic energy of these processes was plotted with WHAM[79].

Drug System	N or O-C bond distance / nm	C=O bond length / nm	N or O-C=O bond angle / º	No. of Atoms in QM region (PM6)	No. of Atoms in MM region (AMBER)
2-aminomethylbenzimidazole	**0.375**	0.123	**147.4**	58	4920
Propranolol	**0.326**	0.122	**104.7**	78	4665
Protonated Propranolol	**0.315**	0.121	**104.7**	73	4896
4-amino-N-phenylbutanamide	**0.405**	0.122	**153.1**	65	4938
Serotonin	**0.313**	0.123	**127.1**	65	5274
Procaine	**0.502**	0.122	**144.2**	75	5259

Table 5.3: QM/MM starting conditions for simpler solvated lipid-drug system without phosphocholine headgroup and lipid acyl chain

5.3.3 QM/MM System I: 2-Aminomethylbenzimidazole

2-Aminomethylbenzimidazole was the first small organic molecule candidate for QM/MM simulations. As can be seen in figure 5.16, the energy of the maximum peak corresponded to the transition state and the minimum corresponded to the $T^{\pm}$ intermediate, the first stable intermediate in the reaction pathway. Evidently from table 5.4, there is a reduction in the C-N bond distance for the $T^{\pm}$ intermediate. A reduction from 0.326 nm to 0.168 nm shows bond formation between the drug and lipid. The free energy change between these two intermediates was 0.73 kcal mol^{-1}. However, once this simulation was compared to the same PM6 curve with an added dispersion correction term, the ΔG was now 2.49 kcal mol^{-1}. Figure 5.16 clearly shows the same intermediates formed at a similar position along the reaction coordinate but with a greater than 3-fold increase in energy change. The minimum was lower for the dispersion correction calculation, despite the C-N bond distance being larger at 1.74 nm. Thus the $T^{\pm}$ intermediate could subsequently be used as a starting structure for DFT minimisations of other possible intermediates.

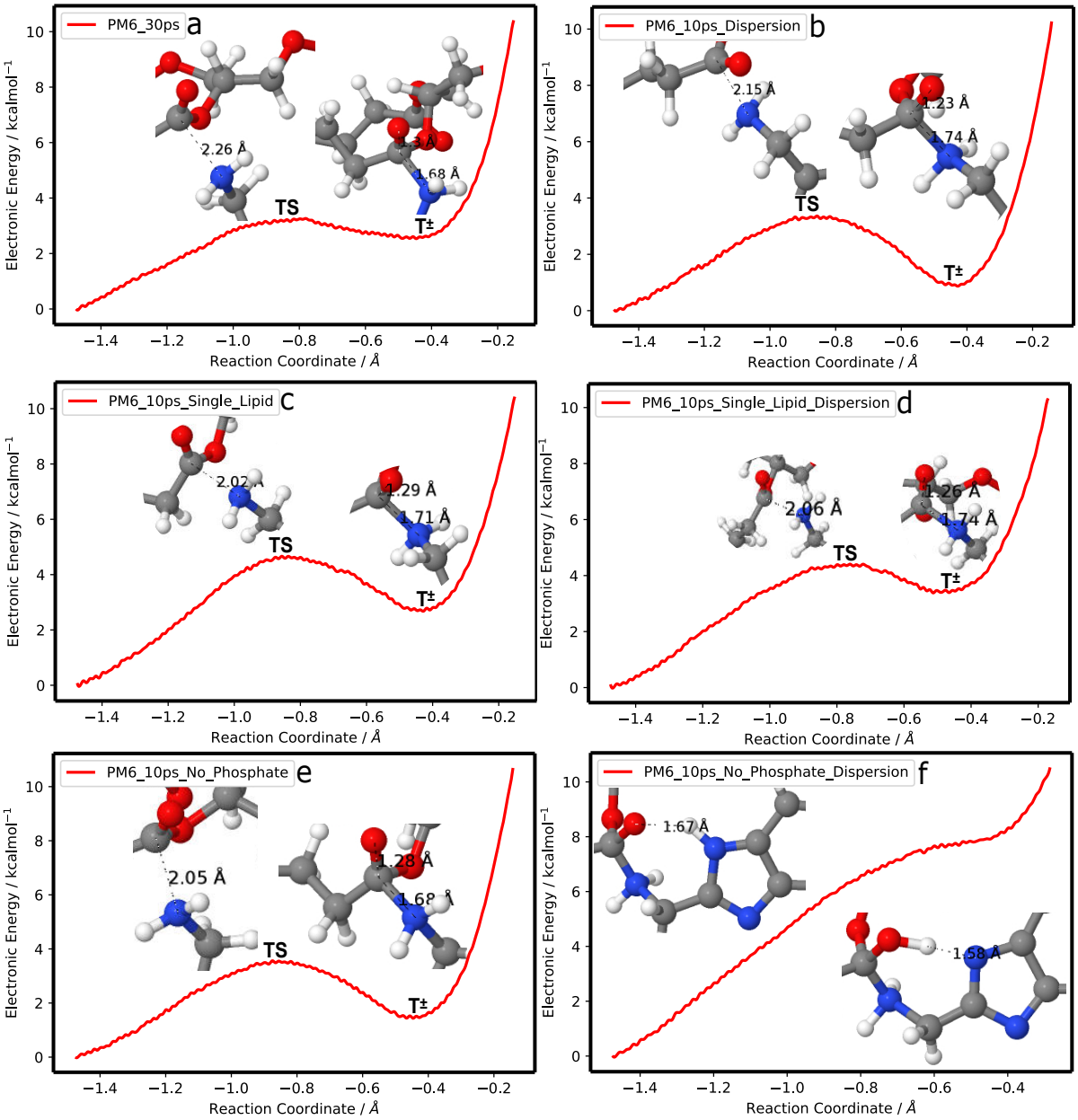


Figure 5.16: 2-aminomethylbenzimidazole reactivity profiles (PM6): (a) 30 ps, (b) 10 ps (+D3 dispersion), (c) 10 ps (single lipid-drug), (d) 10 ps (single lipid-drug+D3 dispersion), (e) 10 ps (no phosphate system), (f) 10 ps (no phosphate system+D3 dispersion)

Reaction Profile	Drug Intermediate	C-N Bond length / nm	*G* / kcal mol^{-1}	Δ*G* / kcal mol^{-1}
PM6_30ps	TS	**0.226**	3.26	**0.73**
	T$^{\pm}$	**0.168**	2.53	
PM6_10ps_Dispersion	TS	**0.215**	3.36	**2.49**
	T$^{\pm}$	**0.174**	0.87	
PM6_10ps_Single_Lipid	TS	**0.202**	4.67	**1.95**
	T$^{\pm}$	**0.171**	2.72	
PM6_10ps_Single_Lipid_Dispersion	TS	**0.206**	4.41	**1.03**
	T$^{\pm}$	**0.174**	3.38	
PM6_10ps_No_Phosphate	TS	**0.205**	3.57	**2.14**
	T$^{\pm}$	**0.168**	1.43	
PM6_10ps_No_Phosphate_Dispersion	N/A	H^{+} transfer	N/A	N/A
	N/A	H^{+} transfer	N/A	

Table 5.4: PM6 C-N bond lengths and free energy of intermediates

As is evident in table 5.4, the energy barrier to reach the transition state, in the single lipid drug interaction system, is higher (4.67 kcal mol^{-1}) than in the membrane (3.26 kcal mol^{-1}). This would imply additional stabilising effects in the membrane which lower the energy barrier for reactivity. The initial C-N bond distance at the beginning was 0.435 nm, and drastically reduced to 0.202 nm (TS) and 0.171 (T$^{\pm}$ intermediate). Furthermore, the free energy change between the transition state and T$^{\pm}$ intermediate was larger at 1.95 kcal mol^{-1} with the T$^{\pm}$ minimum having a higher energy than in the membrane. This suggests that the intermediate is also stabilised by membrane effects. The addition of a dispersion correction term lowers the energy barrier to 4.41 kcal mol^{-1} and also results in a lower free energy change between the transition state and T$^{\pm}$ intermediate (1.03 kcal mol^{-1}). It is expected that this dispersion term will improve the accuracy of the electronic energy calculation.

As can be seen in figure 5.16, the energy barrier to reach the transition state is higher (3.57 kcal mol^{-1}) in this simpler system (without lipid headgroup and acyl chain) than in the membrane. This shows the important role of the oxyanion part of the phosphate in the membrane which stabilises the chemical bonding and thus lowers the energy for bond formation in the membrane. This is further confirmed by the largest ΔG between the TS and T$^{\pm}$ intermediate (2.14 kcal mol^{-1}) of the

three systems, thus showing more energy is required to reach the stable $T^{\pm}$ intermediate without the phosphate oxyanion stabilisation effect. The addition of a dispersion correction term negates a free energy barrier and instead shows a proton transfer from the NH benzimidazole ring to the carbonyl atom to form a different intermediate along the reaction pathway in figure 5.16f.

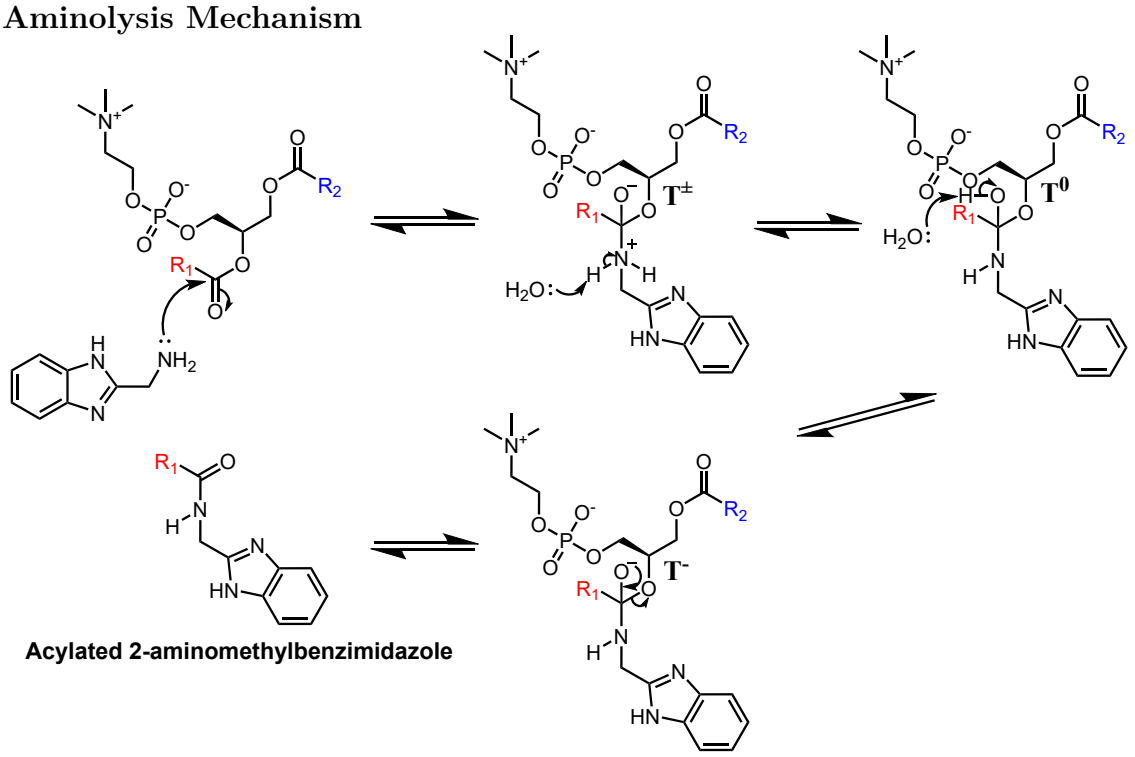


Figure 5.17: Reaction pathway mechanism for 2-aminomethylbenzimidazole

The trajectory frame for the minimum ($T^{\pm}$) was then extracted from the energy profile and energy minimised with PM6. In accordance with experimental literature, this molecule undergoes an aminolysis mechanism to form a stable acylated drug (figure 5.17). Nucleophilic attack of the drug to the lipid carbonyl forms a $T^{\pm}$ intermediate, deprotonation and protonation simultaneously forms a T^{0} intermediate, deprotonation of the cationic amine results in a T^{-} intermediate, and lastly aminolysis to form the amide. DFT minimisations of these intermediates were carried out with the B3LYP and PBE0 (+D3 Dispersion) functionals in implicit solvent.

As can be seen in table 5.5, both the C-N and C-O bond lengths after DFT minimisation are almost identical for the B3LYP and PBE0 functionals. A reduction in C-N bond length (0.166 to 0.145 nm) and elongation in C-O bond length (0.126

to 0.143 nm) from $T^{\pm}$ to T^{0} for both DFT functionals shows protonation of the oxyanion and deprotonation of the cationic amine to form a stable T^{0} intermediate. Furthermore, deprotonation of the alcoholic proton from T^{0} to T^{-} intermediate, resulting in a C-O bond reduction from 0.143 to 0.130 nm, causes a stable T^{-} intermediate to form. Lastly, re-formation of the carbonyl takes place as aminolysis ensures formation of the stable acylated drug. This is reflected in a reduced C-O distance of 0.124 nm and a C-N bond distance of 0.136 nm, both correlating well with amide bond distances.

DFT Functional	Intermediate	C-N Bond length / nm	C-O Bond length / nm
B3LYP	$T^{\pm}$	**0.166**	**0.126**
B3LYP	T^{0}	**0.145**	**0.143**
B3LYP	T^{-}	**0.151**	**0.130**
B3LYP	Acylated Product	**0.136**	**0.124**
PBE0 (+ D3)	$T^{\pm}$	**0.166**	**0.126**
PBE0 (+ D3)	T^{0}	**0.144**	**0.141**
PBE0 (+ D3)	T^{-}	**0.150**	**0.130**
PBE0 (+ D3)	Acylated Product	**0.135**	**0.123**

Table 5.5: Bond lengths post DFT minimisation of reaction intermediates

Figure 5.6 tabulates the energy changes between the different intermediates in the reaction pathway. A small relative energy change of -3.02 kcal mol^{-1} for the $T^{\pm}$ to the T^{0} intermediate (B3LYP functional) compares well with the experimental enthalpy required for a synchronous protonation, to from the cationic amine attached to the lipid carbon, and deprotonation, to form the oxyanion attached to the lipid carbon. The relative energy change from the T^{-} to acylated product, namely -85.9 kcal mol^{-1} (-359.4 kJ mol^{-1}), correlates well with the enthalpy change required to re-form the carbonyl in the formation of an amide in aminolysis. Furthermore, a negative free energy change of -28.5 kcal mol^{-1} from the $T^{\pm}$ straight to the acylated drug product shows a favourable aminolysis mechanism, as this reaction can happen spontaneously. The relative free energies are slightly lower for the PBE0 functional. The inclusion of the Grimme-D3 dispersion correction would slightly lower the energy, but all values are consistent with an aminolysis mechanism.

Functional	Basis Set	Drug Intermediate	*G* / Hartrees	Δ*G* / kcal mol^{-1}
B3LYP	6-311+g(d,p)	$T^{\pm}$ (154 atoms)	-3123.25893734	
B3LYP	6-311+g(d,p)	T^{0} (154 atoms)	-3123.26374793	**-3.02**
B3LYP	6-311+g(d,p)	T^{-} (153 atoms)	-3122.77882699	**60.4**
B3LYP	6-311+g(d,p)	Acylated Drug (154 atoms)	-3123.30430086	**-85.9**

Functional	Basis Set	Water	*G* / Hartrees	Δ*G* / kcal mol^{-1} **$T^{\pm}$ to Acylated Drug**
B3LYP	6-311+g(d,p)	Neutral	-76.4664127203	**-28.5**
B3LYP	6-311+g(d,p)	Protonated	-76.8550364477	

Functional	Basis Set	Drug Intermediate	*G* / Hartrees	Δ*G* / kcal mol^{-1}
PBE0 (+D3)	6-311+g(d,p)	$T^{\pm}$ (154 atoms)	-3119.93118357	
PBE0 (+D3)	6-311+g(d,p)	T^{0} (154 atoms)	-3119.93707788	**-3.70**
PBE0 (+D3)	6-311+g(d,p)	T^{-} (153 atoms)	-3119.45090495	**52.8**
PBE0 (+D3)	6-311+g(d,p)	Acylated Drug (154 atoms)	-3119.97630116	**-77.4**

Functional	Basis Set	Water	*G* / Hartrees	Δ*G* / kcal mol^{-1} **$T^{\pm}$ to Acylated Drug**
PBE0 (+D3)	6-311+g(d,p)	Neutral	-76.3694557125	**-28.3**
PBE0 (+D3)	6-311+g(d,p)	Protonated	-76.7715468686	

Table 5.6: DFT minimisation energies of reaction intermediates

5.3.4 QM/MM System II: Neutral/Protonated Propranolol

Propranolol

A comparison of propranolol vs protonated propranolol was then conducted at PM6 theory. As has been alluded to previously, the reactive nucleophilic part of both neutral and protonated propranolol is the oxygen atom. A major difference in the reaction profile between propranolol and 2-aminomethylbenzimidazole is that both the TS and acylated drug (minimum) are much higher in energy than 2-aminomethylbenzimidazole. This implies that the reactive intermediates in the overall energy pathway are less stable than 2-aminomethylbenzimidazole. This would give credence to formation of the acylated drug rather than a stable $T^{\pm}$ intermediate. As the literature has confirmed that neutral propranolol, compared with protonated, is the minor product which undergoes lipidation, the relative difficulty in reaching a transition state and subsequent reactive intermediates correlates well with experimental data. Initially at the TS, the C-O bond distance was 0.186 nm, whereas the C-O bond distance had reduced to 0.145 nm in the stable acylated drug (table 5.7). Also, there had been a proton transfer from the oxygen to the nitrogen to form the protonated drug, thus implying lower stability for the neutral form. Furthermore, with the inclusion of a dispersion correction, the maximum corresponds to a higher energy $T^{\pm}$ intermediate, rather than the TS, confirmed by a shorter C-O bond distance of 0.156 nm (see figure 5.18b). Although the energy barrier to reach the transition state was higher with dispersion (13.34 kcal mol^{-1}), than without (12.27 kcal mol^{-1}), the free energy change between the maximum and stable acylated drug was lower. This would indicate reaching the stable acylated drug is easier from the $T^{\pm}$ intermediate than from a transition state.

Based on figure 5.18, the energy barrier to reach the transition state is higher for the single lipid-drug system (17.99 kcal mol^{-1}) than in the membrane (12.27 kcal mol^{-1}). Similar to 2-aminomethylbenzimidazole, additional stabilising interactions in the membrane help lower the energy barrier for reactivity. This model is vastly improved with a dispersion correction. The dispersion calculation interestingly shows a T^{0} intermediate at the minimum, whereas lysolipid formation takes place during energy minimisation of the acylated drug (minimum) without the dispersion term.

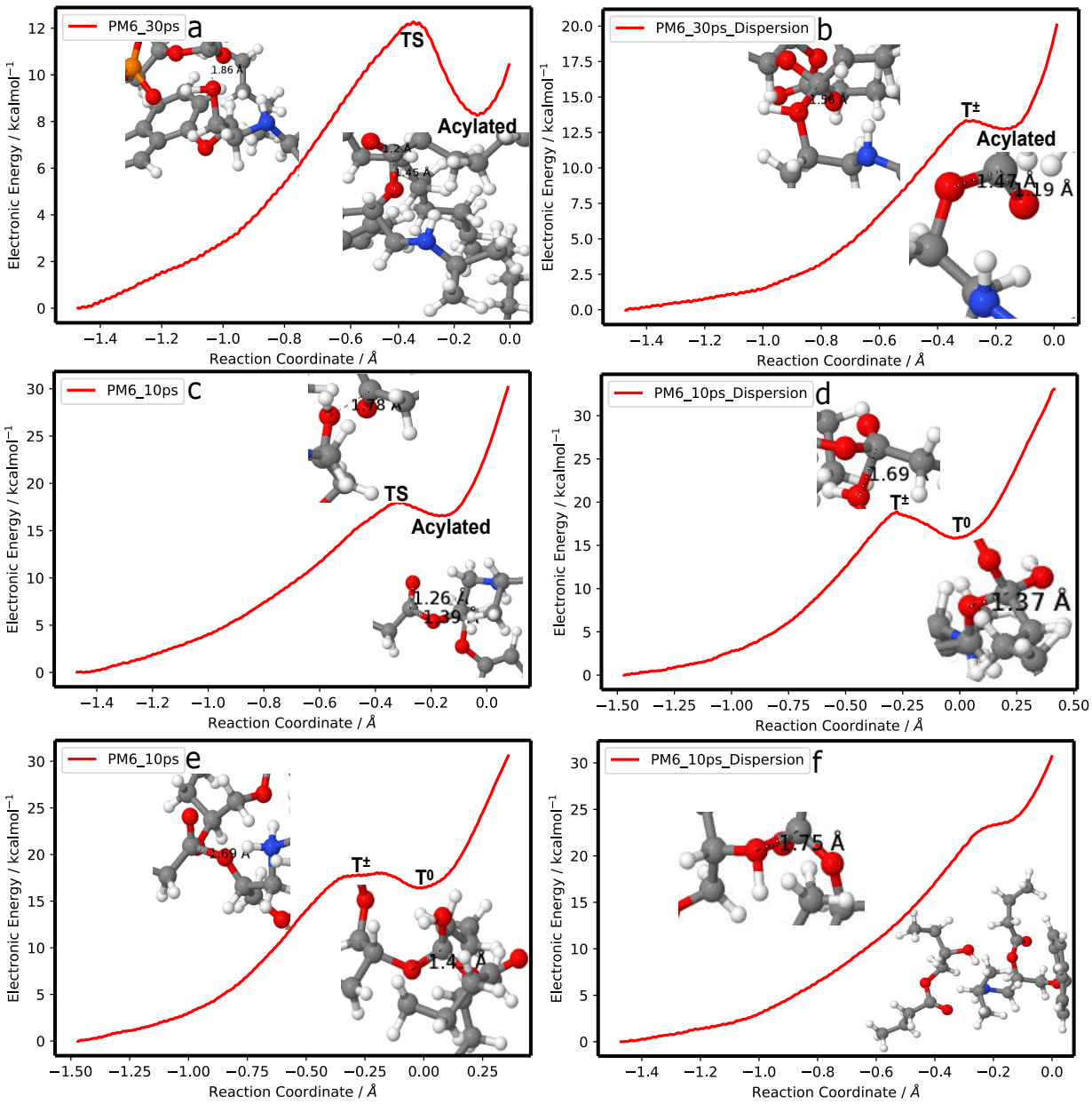


Figure 5.18: Propranolol reactivity profiles (PM6): (a) 30 ps, (b) 10 ps (+D3 dispersion), (c) 10 ps (single lipid-drug), (d) 10 ps (single lipid-drug+D3 dispersion), (e) 10 ps (no phosphate system), (f) 10 ps (no phosphate system+D3 dispersion)

Reaction Profile	Drug Intermediate	C-O Bond length / nm	*G* / kcal mol^{-1}	Δ*G* / kcal mol^{-1}
PM6_30ps	TS	**0.186**	12.27	**4.01**
	Acylated	**0.145**	8.26	
PM6_10ps_Dispersion	T$^{\pm}$	**0.156**	13.34	**0.61**
	Acylated	**0.147**	12.73	
PM6_10ps_Single_Lipid	TS	**0.178**	17.99	**0.61**
	Acylated	**0.139**	16.54	
PM6_10ps_Single_Lipid_Dispersion	T$^{\pm}$	**0.169**	18.90	**3.08**
	T^{0}	**0.137**	15.92	
PM6_10ps_No_Phosphate	T$^{\pm}$	**0.169**	18.06	**1.67**
	T^{0}	**0.143**	16.39	
PM6_10ps_No_Phosphate_Dispersion	N/A	Inflection	N/A	N/A
	N/A	Inflection	N/A	

Table 5.7: PM6 C-O bond lengths and free energy of intermediates

Experimental evidence shows that lysolipid formation only happens with protonated propranolol. The model calculates that the dispersion calculation is the more accurate one and clearly predicts the T^0 being lower in energy (15.92 kcal mol^{-1}) and thus less likely to form lysolipids. Also, the free energy difference between the maximum and minimum is smaller for the same calculation without the dispersion correction. The C-O bond distance of 0.169 nm for the $T^{\pm}$ intermediate (maximum) and its reduction to 0.137 nm for the T^0 intermediate are predicted well with the dispersion term.

Lastly for the simplest (without phosphate headgroup and lipid acyl chain) system, from table 5.7 we can see that the energy barrier to reach the transition state is higher (18.06 kcal mol^{-1}) compared with the membrane system (12.27 mol^{-1}). Thus, implicating the important role of the oxyanion phosphate in mediating the drug-lipid stability and thus lowering the energy barrier for reactivity. Once more, the addition of a dispersion correction term results in an inflection point on the curve, an indication of the sensitivity of the system to small changes. In fact during the course of the reaction, the proton on the -OH of the drug molecule is transferred to the ester oxygen attached to the carbonyl of the lipid, and then the ester bond is hydrolysed to form a lipidated drug system.

Protonated Propranolol

In protonated propranolol, the nitrogen atom is protonated and thus any potential proton transfer from the oxygen to nitrogen was negated. Somewhat unexpectedly, the energy barrier to reach the transition state was higher in protonated (14.17 kcal mol^{-1}) compared with neutral (12.27 kcal mol^{-1}) propranolol. When sampling the potential energy surface, there may have been alternative starting structures which result in lower TS energies. The C-O bond distance reduced from 0.186 nm ($T^{\pm}$) to 0.145 nm (acylated drug) (table 5.8). The minimum corresponding to the acylated drug, has a lower energy at 7.29 kcal mol^{-1} compared with 8.26 kcal mol^{-1} for neutral propranolol. This suggests that the protonated drug is more stable once it has reacted with the membrane lipid than its neutral analogue. The phosphate oxyanion evidently catalysed the proton transfer from the hydroxyl and thus favoured the acylated drug being formed. Addition of a dispersion term results in a far smaller energy difference between the $T^{\pm}$ and acylated drug in the reaction pathway, but with the final structure being less energetically stable (18.03 kcal mol^{-1} compared with 7.29 kcal mol^{-1}) than the non-dispersion calculation. Evidently dispersion here results in more energetically favourable lysolipid formation.

The same simulation at PM6 level, but with a simpler system, i.e. the drug reacting with a single lipid, was conducted in implicit solvent. Without a dispersion term there were no obvious intermediates. However, the inclusion of a dispersion term resulted in lysolipid formation (figure 5.19d). In contrast to the drug-membrane system, the C-O bond distance was shorter at 0.169 nm and thus implied formation of a higher energy intermediate TS. The energy minimum (15.11 kcal mol^{-1}) remained far higher than the same minimum in the drug-membrane system (7.29 kcal/mol), further confirming the role of additional stabilisation of the intermediates when in the membrane compared with implicit solvent. Also, this simpler system formed a T^{0} intermediate and not the final acylated drug, suggesting the critical role of not only membrane but also phosphate stabilisation effects in directly leading to lysolipid formation. The phosphate oxyanion plays the largest role in stabilisation, as the minimum corresponds to the acylated drug only in the drug-membrane system. The absence of any intermediates when the reaction takes place between the

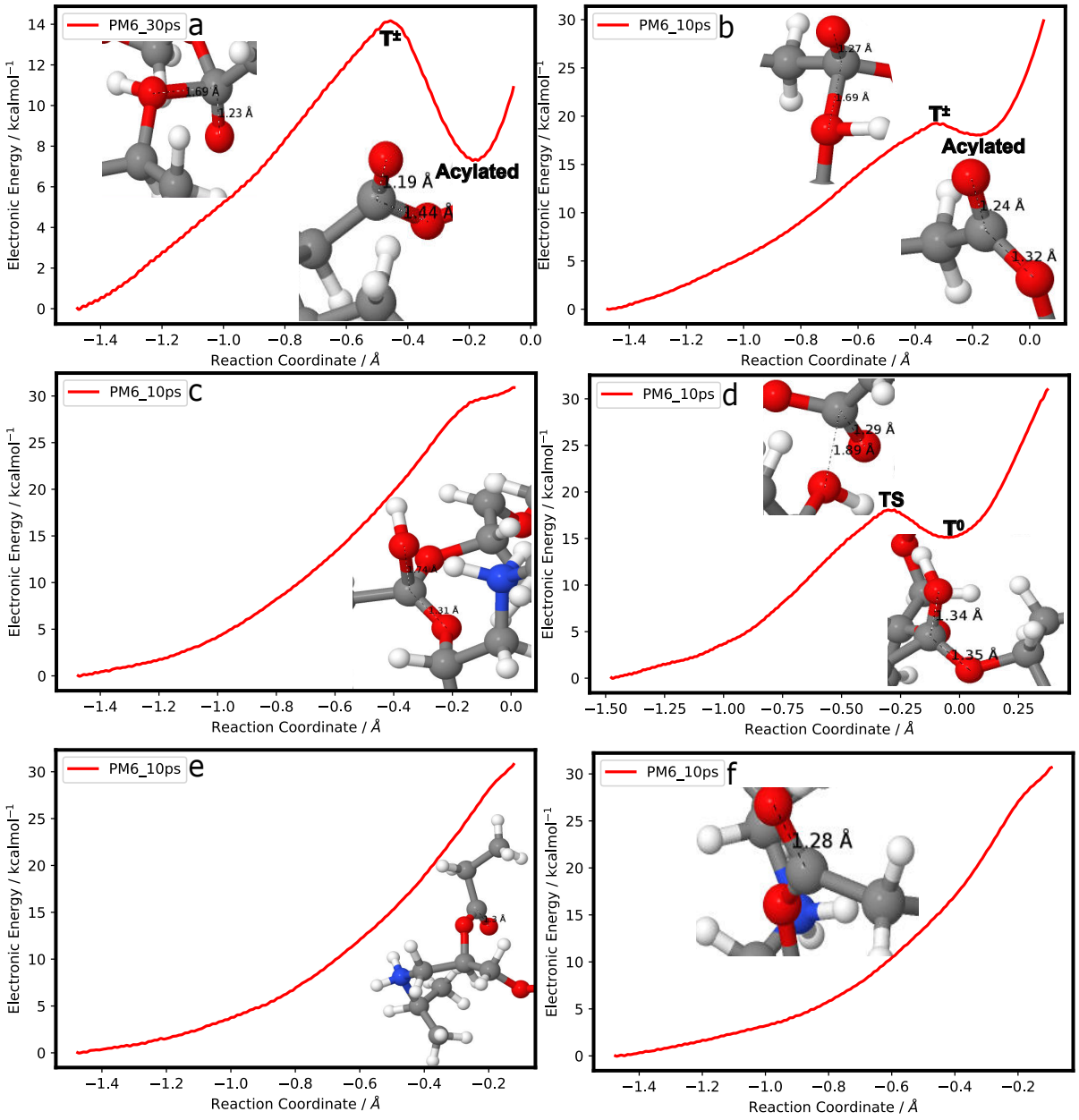


Figure 5.19: Protonated propranolol reactivity profiles (PM6): (a) 30 ps, (b) 10 ps (+D3 dispersion), (c) 10 ps (single lipid-drug), (d) 10 ps (single lipid-drug+D3 dispersion), (e) 10 ps (no phosphate system), (f) 10 ps (no phosphate system+D3 dispersion)

Reaction Profile	Drug Intermediate	C-O Bond Distance / nm	*G* / kcal mol^{-1}	Δ*G* / kcal mol^{-1}
PM6_30ps	T$^{\pm}$	**0.186**	14.17	**6.88**
	Acylated	**0.145**	7.29	
PM6_10ps_Dispersion	T$^{\pm}$	**0.156**	19.25	**1.22**
	Acylated	**0.147**	18.03	
PM6_10ps_Single_Lipid	N/A	Inflection	N/A	N/A
	N/A	Inflection	N/A	
PM6_10ps_Single_Lipid_Dispersion	TS	**0.169**	18.08	**2.97**
	T^{0}	**0.137**	15.11	
PM6_10ps_No_Phosphate	N/A	Inflection	N/A	N/A
	N/A	Inflection	N/A	
PM6_10ps_No_Phosphate_Dispersion	N/A	Inflection	N/A	N/A
	N/A	Inflection	N/A	

Table 5.8: PM6 C-O Bond lengths and free energy of intermediates

drug and a simpler lipid system (without a phosphate or long acyl chains) indicates no stabilisation and thus no acylation product.

Transesterification Mechanism

The trajectory frame for the minimum was then extracted and energy minimised with PM6. Unlike with the other systems which display aminolysis behaviour, both protonated and neutral propranolol undergo transesterification to form the stable acylated drug (figure 5.20). Furthermore, from the energy profiles, nucleophilic attack of each neutral and protonated drug forms a stable acylated drug as the energy minimum, without any other intermediates. As protonated propranolol results in increased lysolipid formation, this theoretical model may be consistent with the reaction going straight to acylated form. DFT minimisations were carried out on the final acylated form with both the B3LYP and PBE0 (+D3 Dispersion) DFT functionals in implicit solvent.

Table 5.9 shows both the energies and bond length changes during DFT minimisation. As the final acylated drug was the only intermediate in the reaction pathway, only energies in Hartrees could be provided without relative free energy comparison. Consistent with a D3 dispersion correction, the energy of the final acylated product

is lower with the PBE0 than the B3LYP minimisations. Importantly, all minimisations demonstrate a final C-O bond distance of 0.121 nm which correlate well with re-formation of the carbonyl during transesterification to form the acylated product.

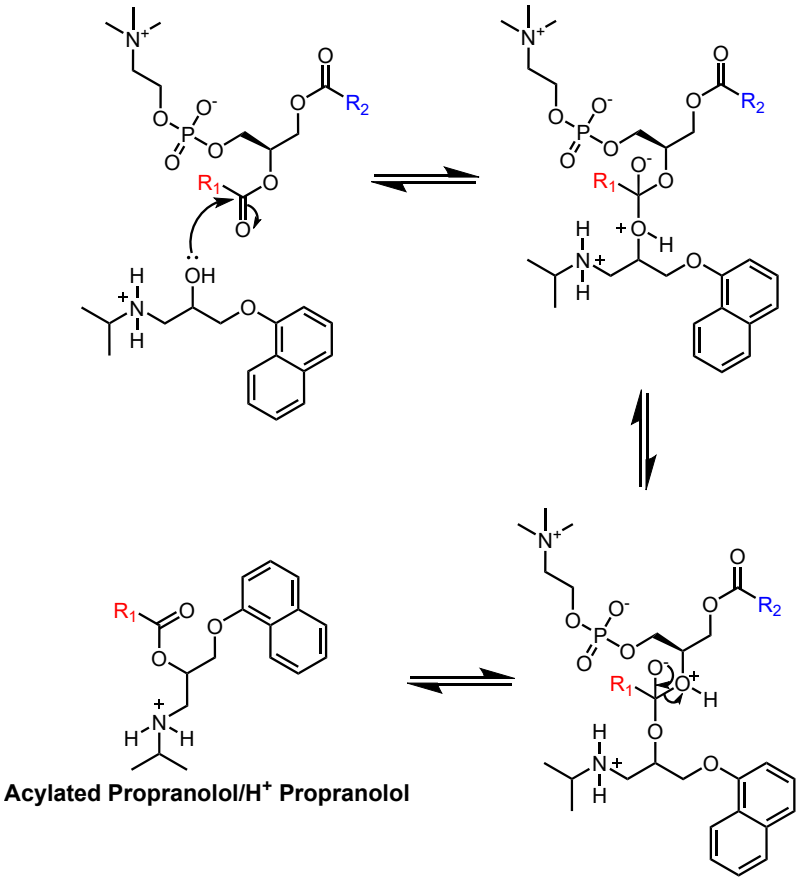


Figure 5.20: Reaction pathway mechanism for propranolol/protonated propranolol

Functional	Basis Set	Drug Intermediate	*G* / Hartrees	**C-O Bond length / nm**	**C=O Bond length / nm**
B3LYP	6-311+g(d,p)	Neutral Acylated Drug (174 atoms)	-3476.34971857	**0.135**	**0.121**
PBE0 (+D3)	6-311+g(d,p)	Neutral Acylated Drug (174 atoms)	-3472.61589399	**0.134**	**0.121**
B3LYP	6-311+g(d,p)	Protonated Acylated Drug (175 atoms)	-3476.83472652	**0.137**	**0.121**
PBE0 (+D3)	6-311+g(d,p)	Protonated Acylated Drug (175 atoms)	-3473.10968841	**0.135**	**0.121**

Table 5.9: Bond lengths and energies post DFT minimisation of final acylated product

5.3.5 QM/MM System III: 4-Amino-*N*-phenylbutanamide

The drug candidate 4-amino-*N*-phenylbutanamide had similar behaviour to 2-amino methylbenzimidazole. The nucleophilic attack of the nitrogen nucleophile (drug) onto the carbonyl carbon (lipid) resulted in a stable $T^{\pm}$ intermediate (minimum). At the TS, the C-N bond distance was 0.196 nm and this decreased to 0.167 nm at the stable $T^{\pm}$ intermediate (table 5.10). The $T^{\pm}$ intermediate was lower in energy for 4-amino-*N*-phenylbutanamide (1.82 kcal mol^{-1}) than 2-aminomethylbenzimidazole (2.53 kcal mol^{-1}). Also, the energy difference between the TS and $T^{\pm}$ intermediate was lower (2.26 kcal mol^{-1}) than 2-aminomethylbenzimidazole. This suggests that the $T^{\pm}$ intermediate is more stable and requires a smaller energy barrier to reach. Adding a dispersion correction term here only slightly changes the energy barrier and is not significant (figure 5.21b). Unlike with 2-aminomethylbenzimidazole, the phosphate oxyanion is much further away from the nucleophilic part of the drug molecule and thus unlikely to play a pivotal role in stabilisation.

When directly compared with a simpler system, whereby the drug molecule interacts with a single lipid in implicit solvent, the $T^{\pm}$ intermediate is actually more stable in solvent than in the membrane. Although the energy barrier between the TS and $T^{\pm}$ is larger (6.65 kcal mol^{-1}) compared with in the membrane (2.26 kcal mol^{-1}) (table 5.10), the $T^{\pm}$ intermediate reaches an energy value of exactly 0 kcal mol^{-1}. This implies a fully minimised intermediate. As the initial C-N bond distance is 0.201 nm and decreases to 0.165 nm for the $T^{\pm}$ intermediate, these values may be closer to an absolute minimum in the potential energy surface. The inclusion of the dispersion term merely reduces the energy barrier between the TS and $T^{\pm}$ intermediate. However, although the $T^{\pm}$ intermediate is lower in energy, the energy barrier to reach the transition state is still higher in implicit solvent (6.65 kcal mol^{-1}) compared with in the membrane (4.08 kcal mol^{-1}). Thus, the reactivity is more stable in the membrane than in water (figure 5.21). When compared with the reactivity of the same drug molecule to the lipid carbonyl carbon in the absence of both a phosphate headgroup and acyl chain, both the energy barrier required to reach the transition state (4.26 kcal mol^{-1}) and the $T^{\pm}$ intermediate (2.74 kcal mol^{-1}) are higher than in the drug-membrane system. Thus, confirming the importance of

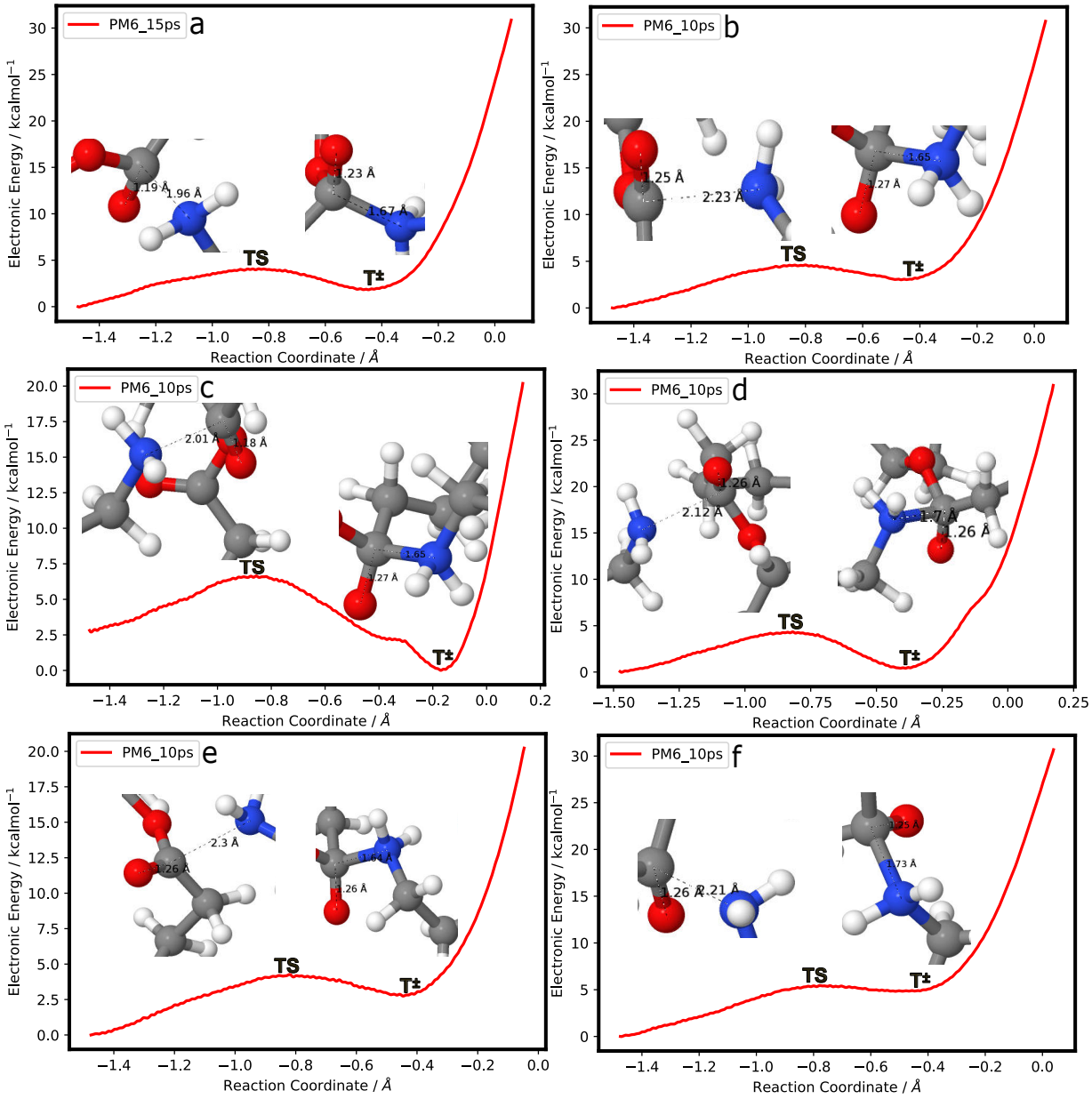


Figure 5.21: 4-amino-*N*-phenylbutanamide reactivity profiles (PM6): (a) 30 ps, (b) 10 ps (+D3 dispersion), (c) 10 ps (single lipid-drug), (d) 10 ps (single lipid-drug+D3 dispersion), (e) 10 ps (no phosphate system), (f) 10 ps (no phosphate system+D3 dispersion)

Reaction Profile	Drug Intermediate	C-N Bond length / nm	*G* / kcal mol^{-1}	Δ*G* / kcal mol^{-1}
PM6_30ps	TS	**0.196**	4.08	**2.26**
	T±	**0.167**	1.82	
PM6_10ps_Dispersion	TS	**0.223**	4.60	**1.59**
	T±	**0.165**	3.01	
PM6_10ps_Single_Lipid	TS	**0.201**	6.65	**6.65**
	T±	**0.165**	0.00	
PM6_10ps_Single_Lipid_Dispersion	TS	**0.212**	4.34	**3.94**
	T±	**0.170**	0.40	
PM6_10ps_No_Phosphate	TS	**0.230**	4.26	**1.52**
	T±	**0.164**	2.74	
PM6_10ps_No_Phosphate_Dispersion	TS	**0.221**	5.44	**0.62**
	T±	**0.173**	4.82	

Table 5.10: PM6 C-N bond lengths and free energy of intermediates

the lipid acyl chains in stabilisation of the reaction intermediates.

Aminolysis Mechanism

The trajectory frame for the minimum ($T^{\pm}$) was then extracted from the energy profile and energy minimised with PM6. In accordance with experimental literature[24], this molecule undergoes an aminolysis mechanism to form a stable acylated drug (figure 5.22). DFT minimisations of these intermediates were carried out with the B3LYP and PBE0 (+D3 Dispersion) functionals in implicit solvent.

As is evident in table 5.11, the C-N and C-O bond lengths, across the DFT functionals, remain almost identical. A reduction in C-N bond length from 0.165 to 0.146 nm, as well as an elongation in bond length from 0.126 to 0.141 nm, from $T^{\pm}$ to T^{0} intermediates, show the protonated amine becoming neutral and the oxyanion being protonated to form an hydroxyl. Further, a reduction in bond length from 0.141 to 0.129 nm from T^{0} to T^{-} intermediates implies deprotonation to reform the oxyanion attached to the lipid carbon atom. Finally, formation of the amide in the final step is confirmed by a reduction in both C-N and C-O bond lengths from the T^{-} intermediate to acylated product. Re-formation of the carbonyl and C-N amide bond is consistent with 0.124 and 0.135 nm respectively for the acylated product.

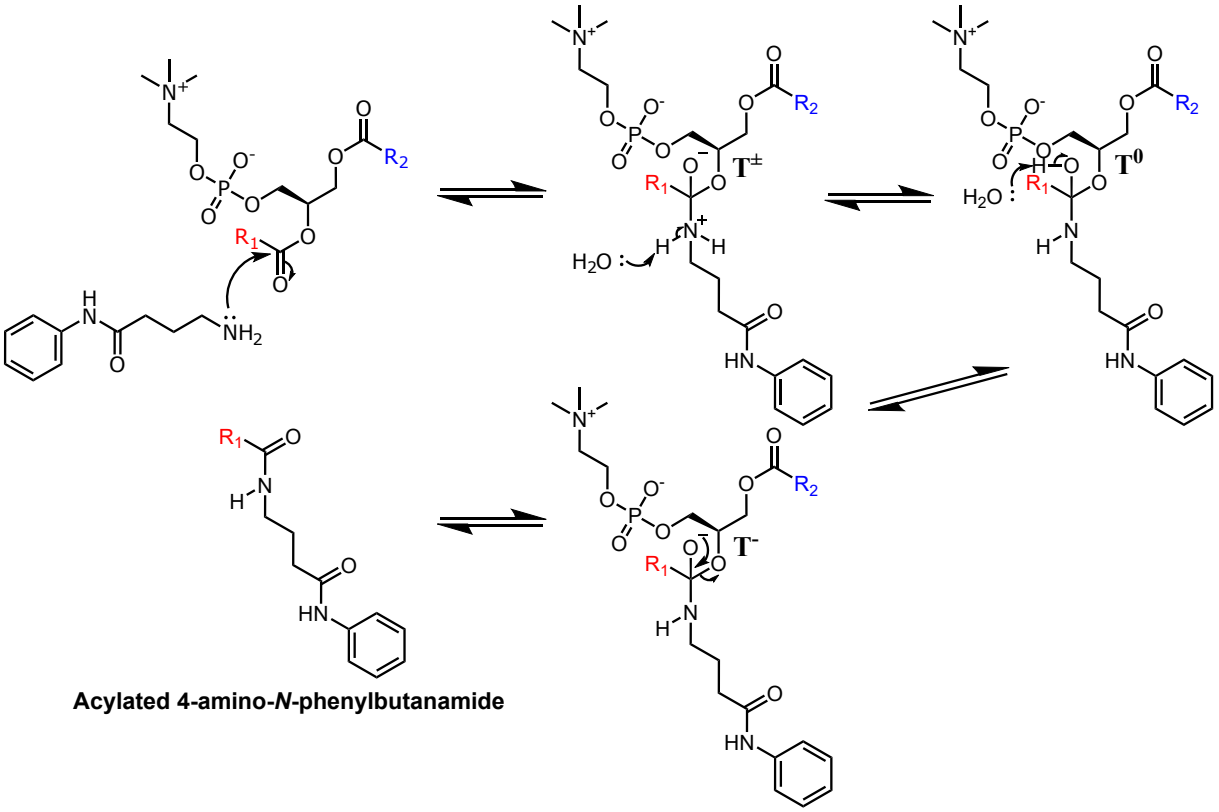


Figure 5.22: Reaction pathway mechanism for 4-amino-*N*-phenylbutanamide

Table 5.12 summarises energy changes between the different intermediates in the reaction pathway. When examining the B3LYP data, a small relative energy change of -1.18 kcal mol^{-1} between the $T^{\pm}$ and T^{0} intermediate would be experimentally consistent with both deprotonation and protonation occurring synchronously, particularly as the net charge would be equivalent. The relative energy change from T^{-} intermediate to acylated product is -92.6 kcal mol^{-1} (-387 kJ mol^{-1}). This reaction energy, which reforms the carbonyl and creates an C-N amide bond, is consistent with experimental enthalpy change from a single C-O to a double C=O bond. Also, as the free energy change is more negative from the T^{-} intermediate (-92.6 kcal mol^{-1}) to acylated drug (-32.0 kcal mol^{-1}) than the $T^{\pm}$ intermediate going straight to the acylated drug, this four step reaction pathway is energetically favoured. When directly compared with PBE0 data, all relative free energy changes between the different intermediates are of the same order of magnitude, but slightly lower in energy. The inclusion of a D3 dispersion term almost certainly lowers the relative energy differences for the PBE0 data. Importantly though, the same trend as for the B3LYP data is seen for the PBE0 data.

DFT Functional	Intermediate	C-N Bond length / nm	C-O Bond length / nm
B3LYP	$T^{\pm}$	**0.165**	**0.126**
B3LYP	T^{0}	**0.146**	**0.141**
B3LYP	T^{-}	**0.151**	**0.129**
B3LYP	Acylated Product	**0.135**	**0.124**
PBE0 (+D3)	$T^{\pm}$	**0.165**	**0.126**
PBE0 (+D3)	T^{0}	**0.146**	**0.140**
PBE0 (+D3)	T^{-}	**0.151**	**0.129**
PBE0 (+D3)	Acylated Product	**0.135**	**0.123**

Table 5.11: Bond lengths post DFT minimisation of reaction intermediates

Functional	Basis Set	Drug Intermediate	*G* / Hartrees	ΔG / kcal mol^{-1}
B3LYP	6-311+g(d,p)	$T^{\pm}$ (161 atoms)	-3223.17392159	
B3LYP	6-311+g(d,p)	T^{0} (161 atoms)	-3223.17580365	**-1.18**
B3LYP	6-311+g(d,p)	T^{-} (160 atoms)	-3222.68865948	**61.8**
B3LYP	6-311+g(d,p)	Acylated Drug (161 atoms)	-3223.22486275	**-92.6**

Functional	Basis Set	Water	*G* / Hartrees	ΔG / kcal mol^{-1} **$T^{\pm}$ to Acylated Drug**
B3LYP	6-311+g(d,p)	Neutral	-76.4664127203	**-32.0**
B3LYP	6-311+g(d,p)	Protonated	-76.8550364477	

Functional	Basis Set	Drug Intermediate	*G* / Hartrees	ΔG / kcal mol^{-1}
PBE0 (+D3)	6-311+g(d,p)	$T^{\pm}$ (161 atoms)	-3219.52788923	
PBE0 (+D3)	6-311+g(d,p)	T^{0} (161 atoms)	-3219.52714261	**0.469**
PBE0 (+D3)	6-311+g(d,p)	T^{-} (160 atoms)	-3219.03528646	**56.3**
PBE0 (+D3)	6-311+g(d,p)	Acylated Drug (161 atoms)	-3219.56808513	**-82.0**

Functional	Basis Set	Water	*G* / Hartrees	ΔG / kcal mol^{-1} **$T^{\pm}$ to Acylated Drug**
PBE0 (+D3)	6-311+g(d,p)	Neutral	-76.3694557125	**-25.2**
PBE0 (+D3)	6-311+g(d,p)	Protonated	-76.7715468686	

Table 5.12: DFT minimisation energies of reaction intermediates

5.3.6 QM/MM System IV: Serotonin

The drug candidate serotonin had similar behaviour to 2-aminomethylbenzimidazole. The nucleophilic attack of the nitrogen nucleophile (drug) onto the carbonyl carbon (lipid) resulted in a stable $T^{\pm}$ intermediate (energy minimum). The initial C-N bond distance for the TS was 0.198 nm and reduced to 0.169 nm for $T^{\pm}$, showing a stable intermediate (table 5.13). When compared with 2-aminomethylbenzimidazole and 4-amino-*N*-phenylbutanamide, the energy barrier to reach the transition state (4.87 kcal mol^{-1}) was higher than both of the other drug molecules, implying less nucleophilic attack onto the carbonyl carbon and a harder to reach TS. Also, the energy of the $T^{\pm}$ intermediate was higher than the other nitrogenous systems. This correlates well with experimental data which shows that Serotonin only has partial nucleophilic potential in the membrane. The energy barrier between the TS and $T^{\pm}$ intermediate (1.31 kcal mol^{-1}) is almost double that of 2-aminomethyl benzimidazole (0.73 kcal mol^{-1}) confirming further the lower nucleophilic potential of the nitrogen atom to the carbonyl lipid. The inclusion of a dispersion term mitigates this somewhat and results in a lower energy barrier to reach the TS (3.36 kcal mol^{-1}) and lower energy of the $T^{\pm}$ intermediate (0 kcal mol^{-1}). Similarly to 4-amino-*N*-phenylbutanamide, the role of the phosphate oxyanion in stabilising this molecule is very minor as the distance between this group and the nucleophilic drug is much larger than needed for stabilisation.

When subjected to the same calculation but with a simpler system where the drug is interacting with a single lipid in implicit solvent, different behaviour is observed. Although the initial C-N bond distance (0.23 nm) is similar to the same system in the membrane, the energy barrier required to reach the TS is lower (3.90 kcal mol^{-1}) and the energy of the $T^{\pm}$ intermediate (0 kcal mol^{-1}) is also lower (figure 5.23). This implies greater stabilisation of this drug molecule in a fully aqueous environment than in the lipid membrane. This phenomenon would explain the poorer nucleophilic ability of serotonin in the membrane. With the dispersion term, an identical energy of the $T^{\pm}$ intermediate is observed with only a slightly larger energy barrier to reach the transition state (see figure 5.23). Without the phosphate and acyl chain, the reaction between the nucleophilic drug and lipid carbonyl carbon

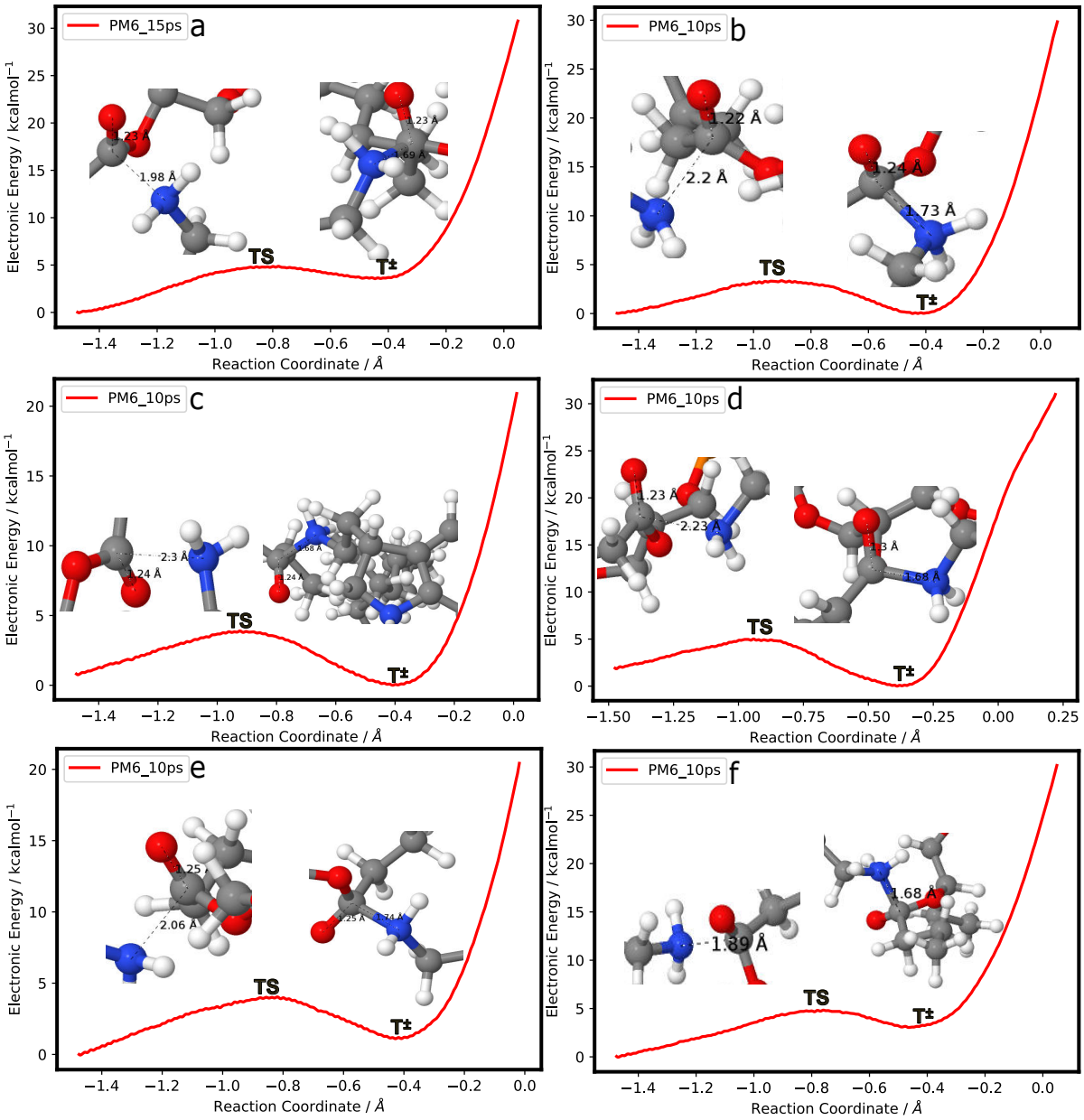


Figure 5.23: Serotonin reactivity profiles (PM6): (a) 30 ps, (b) 10 ps (+D3 dispersion), (c) 10 ps (single lipid-drug), (d) 10 ps (single lipid-drug+D3 dispersion), (e) 10 ps (no phosphate system), (f) 10 ps (no phosphate system+D3 dispersion)

Reaction Profile	Drug Intermediate	C-N Bond length / nm	*G* / kcal mol^{-1}	Δ*G* / kcal mol^{-1}
PM6_30ps	TS	**0.198**	4.87	**1.31**
	T±	**0.169**	3.56	
PM6_10ps_Dispersion	TS	**0.220**	3.36	**3.36**
	T±	**0.173**	0.00	
PM6_10ps_Single_Lipid	TS	**0.230**	3.90	**3.90**
	T±	**0.168**	0.00	
PM6_10ps_Single_Lipid_Dispersion	TS	**0.223**	4.97	**4.97**
	T±	**0.168**	0.00	
PM6_10ps_No_Phosphate	TS	**0.206**	4.04	**2.94**
	T±	**0.174**	1.10	
PM6_10ps_No_Phosphate_Dispersion	TS	**0.189**	4.83	**1.79**
	T±	**0.168**	3.04	

Table 5.13: PM6 C-N bond lengths and free energy of intermediates

results in a $T^{\pm}$ intermediate with higher energy (1.1 kcal mol^{-1}), and therefore less stable than with the single lipid system. The corresponding dispersion calculation reduces the energy barrier between the TS and $T^{\pm}$ intermediate (1.79 kcal mol^{-1}). This simulation also points to the phosphate oxyanion having virtually no effect on additional stabilisation in the membrane, as the $T^{\pm}$ intermediate is actually lower in energy for this simpler system without the phosphate (3.04 kcal mol^{-1}) than in the membrane.

Aminolysis Mechanism

The trajectory frame for the minimum ($T^{\pm}$) was then extracted from the energy profile and energy minimised with PM6. In the experimental literature, this molecule undergoes an aminolysis mechanism to form a stable acylated drug. A similar process occurred in the energy minimisation process. Nucleophilic attack of the drug to the lipid carbonyl forms a $T^{\pm}$ intermediate, deprotonation and protonation simultaneously forms a T^{0} intermediate, deprotonation of the cationic amine results in a T^{-} intermediate, and lastly aminolysis to form the amide (figure 5.24). DFT minimisations of these intermediates were carried out with the B3LYP and PBE0 (+D3 Dispersion) functionals in implicit solvent.

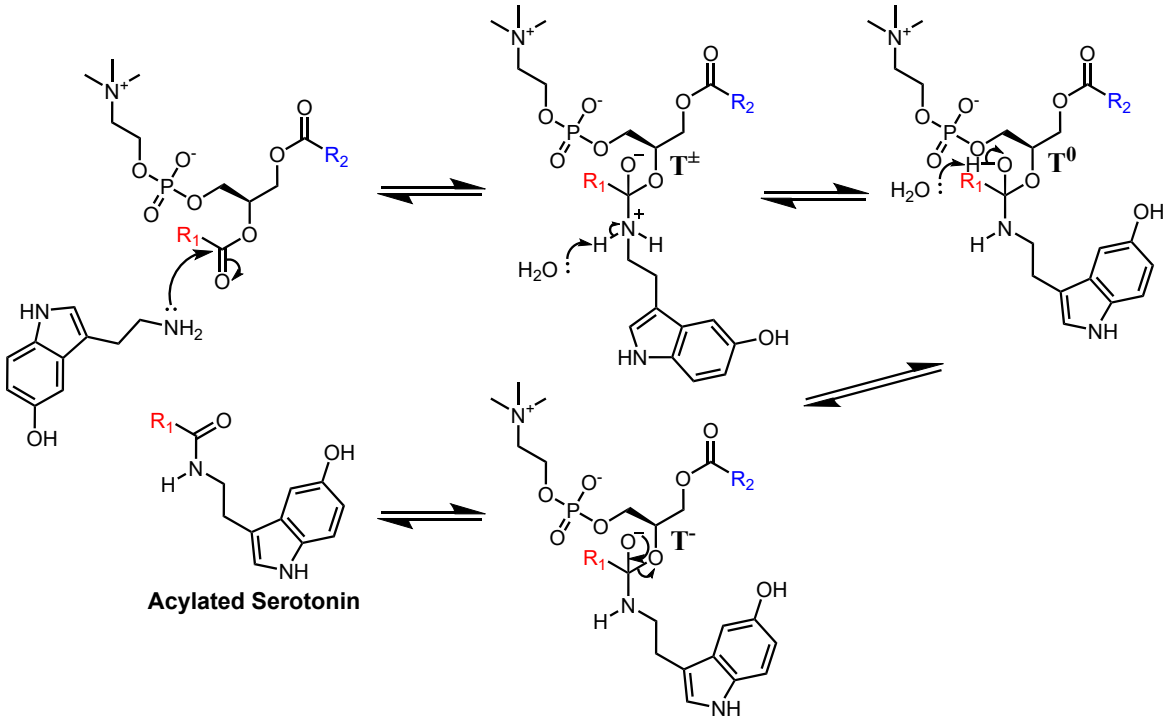


Figure 5.24: Reaction pathway mechanism for serotonin

Similarly to 2-aminomethylbenzimidazole, a decrease in C-N bond length (0.165 to 0.143 nm) and an increase in C-O bond length (0.127 to 0.142 nm) are consistent with simultaneous protonation of the oxyanion and deprotonation of the cationic amine in $T^{\pm}$ to form the stable T^{0} intermediate (table 5.14). Also, a reduction in C-O bond length (0.142 to 0.13 nm) represents deprotonation of the alcoholic proton to form T^{-} as the penultimate intermediate in the reaction profile. Lastly, reduction in both C-N and C-O bond length shows synchronous re-formation of the carbonyl and amide in the aminolysis reaction pathway to form the stable acylated drug. There is very little variation in C-N and C-O bond lengths between the B3LYP and PBE0 (+D3) DFT functionals, testifying to their respective accuracies in modelling changes in bond length.

Table 5.15 summarises energy changes between different intermediates in the reaction profile. When examining the B3LYP data, the relative free energy between the $T^{\pm}$ and T^{0} intermediates is prohibitively high (-122.7 kcal mol^{-1}) but the PBE0 (+D3) accurately shows the energy change as -5.34 kcal mol^{-1}. This value is higher than the other drug systems which undergo aminolysis, suggesting that more energy is required to form this intermediate than 2-aminomethylbenzimidazole or 4-amino-

DFT Functional	Intermediate	C-N Bond length / nm	C-O Bond length / nm
B3LYP	$T^{\pm}$	**0.165**	**0.127**
B3LYP	T^{0}	**0.143**	**0.142**
B3LYP	T^{-}	**0.147**	**0.13**
B3LYP	Acylated Product	**0.135**	**0.124**
PBE0 (+ D3)	$T^{\pm}$	**0.165**	**0.127**
PBE0 (+ D3)	T^{0}	**0.142**	**0.141**
PBE0 (+ D3)	T^{-}	**0.146**	**0.13**
PBE0 (+ D3)	Acylated Product	**0.134**	**0.123**

Table 5.14: Bond lengths post DFT minimisation of reaction intermediates

Functional	Basis Set	Drug Intermediate	*G* / Hartrees	ΔG / kcal mol^{-1}
B3LYP	6-311+g(d,p)	$T^{\pm}$ (159 atoms)	-3221.77361637	
B3LYP	6-311+g(d,p)	T^{0} (159 atoms)	-3221.96916946	**-122.7**
B3LYP	6-311+g(d,p)	T^{-} (158 atoms)	-3221.45984235	**75.7**
B3LYP	6-311+g(d,p)	Acylated Drug (159 atoms)	-3221.82387249	**15.4**

Functional	Basis Set	Water	*G* / Hartrees	ΔG / kcal mol^{-1} **$T^{\pm}$ to Acylated Drug**
B3LYP	6-311+g(d,p)	Neutral	-76.4664127203	**-31.5**
B3LYP	6-311+g(d,p)	Protonated	-76.8550364477	

Functional	Basis Set	Drug Intermediate	*G* / Hartrees	ΔG / kcal mol^{-1}
PBE0 (+D3)	6-311+g(d,p)	$T^{\pm}$ (159 atoms)	-3218.33385106	
PBE0 (+D3)	6-311+g(d,p)	T^{0} (159 atoms)	-3218.34236534	**-5.34**
PBE0 (+D3)	6-311+g(d,p)	T^{-} (158 atoms)	-3217.83696376	**64.8**
PBE0 (+D3)	6-311+g(d,p)	Acylated Drug (159 atoms)	-3218.37863408	**-87.6**

Functional	Basis Set	Water	*G* / Hartrees	ΔG / kcal mol^{-1} **$T^{\pm}$ to Acylated Drug**
PBE0 (+D3)	6-311+g(d,p)	Neutral	-76.3694557125	**-28.1**
PBE0 (+D3)	6-311+g(d,p)	Protonated	-76.7715468686	

Table 5.15: DFT minimisation energies of reaction intermediates

N-phenylbutanamide. Further, the relative free energy change between the T^- and acylated product is positive with the B3LYP functional but negative with the PBE0 functional. This might suggest greater accuracy in predicting relative free energies, coupled with the dispersion correction. A relative free energy change of -87.6 kcal mol^{-1} (-362.8 kJ mol^{-1}) correlates well with the experimental enthalpy of amide bond formation to form the stable acylated drug molecule. Lastly, the negative relative free energy change between the $T^{\pm}$ and acylated product intermediates is -28.1 kcal mol^{-1} (table 5.15). This implies a spontaneous reaction with favourable conditions for an aminolysis mechanism.

5.3.7 QM/MM System V: Procaine

The final system studied with QM/MM is the drug molecule procaine. Contrary to the previous examples, this system shows no stable intermediates in the reaction profile (figure 5.25). In the initial nucleophilic attack in the membrane, the final structure had an C-N bond length of 0.145 nm. Interestingly, during the reaction, a proton transfer from the nucleophilic nitrogen onto the oxyanion of the carbonyl carbon took place to form a stable T^0 intermediate. This behaviour is quite different from previous examples. Unlike with the former nitrogenous systems, a $T^{\pm}$ rather than a T^0 intermediate is formed. Virtually identical behaviour is exhibited with the dispersion term, apart from the absence of a proton transfer in the reaction profile. Although the starting conditions for this system were good and very close to the optimal BD conditions, there might have been alternative starting conditions on the potential energy surface which might have been more accurate.

For the same reaction but direct nucleophilic attack from the drug molecule onto the carbonyl carbon without a membrane and with implicit solvent, a very similar outcome was present. Rather curiously, addition of a dispersion term caused there to be a maximum in the profile, corresponding to the $T^{\pm}$ intermediate, and a minimum corresponding to the T^0 intermediate (figure 5.25). As the energy difference between these two states was very small (1.82 kcal mol^{-1}), this implied that the proton transfer to form the T^0 intermediate took place rather quickly and easily (table 5.16). However, for the system without a phosphate headgroup and fatty

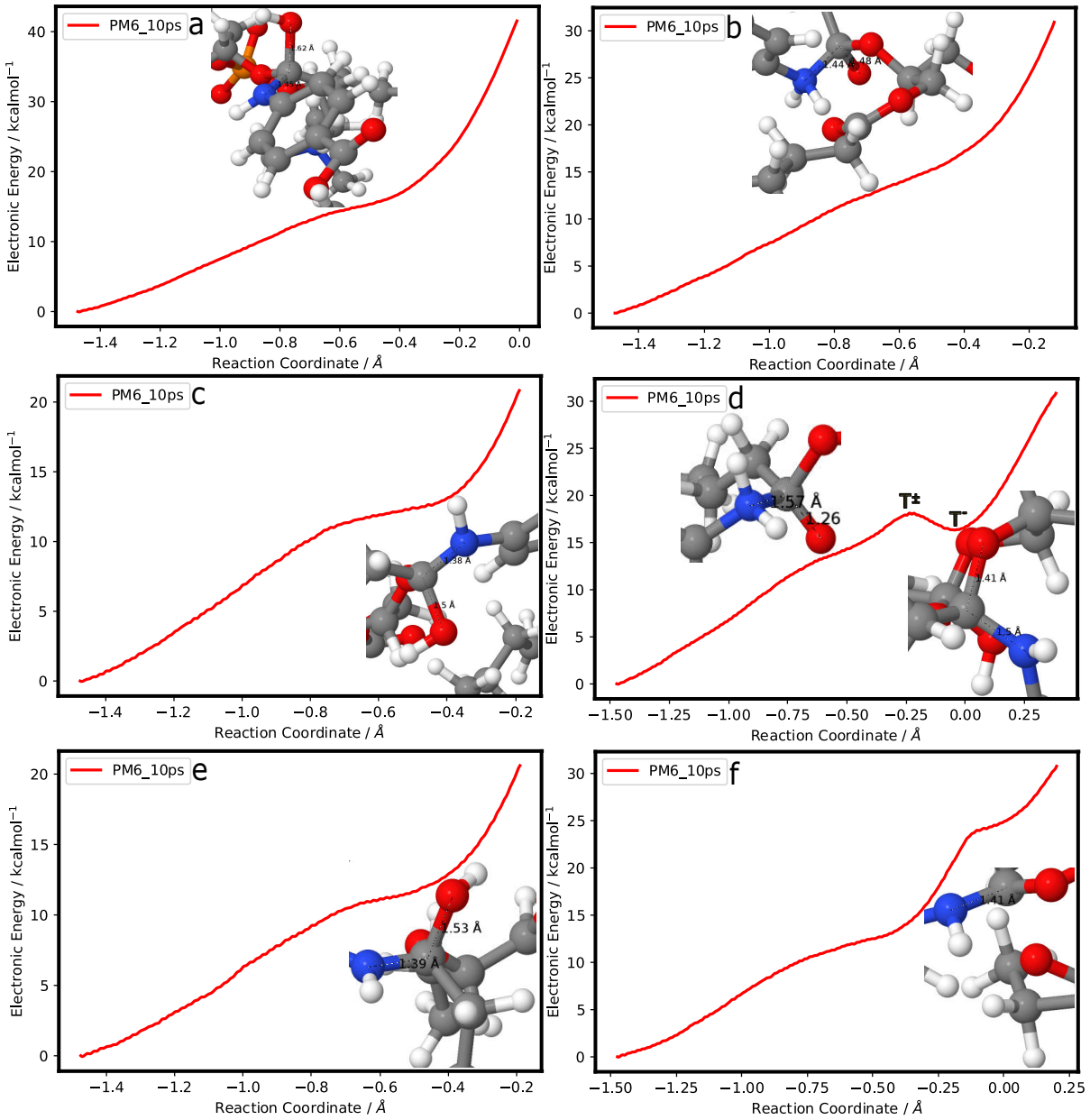


Figure 5.25: Procaine reactivity profiles (PM6): (a) 30 ps, (b) 10 ps (+D3 dispersion), (c) 10 ps (single lipid-drug), (d) 10 ps (single lipid-drug+D3 dispersion), (e) 10 ps (no phosphate system), (f) 10 ps (no phosphate system+D3 dispersion)

Reaction Profile	Drug Intermediate	C-N Bond length / nm	*G* / kcal mol^{-1}	Δ*G* / kcal mol^{-1}
PM6_30ps	N/A	Inflection	N/A	N/A
	N/A	Inflection	N/A	
PM6_10ps_Dispersion	N/A	Inflection	N/A	N/A
	N/A	Inflection	N/A	
PM6_10ps_Single_Lipid	N/A	Inflection	N/A	N/A
	N/A	Inflection	N/A	
PM6_10ps_Single_Lipid_Dispersion	T±	**0.157**	18.15	**1.82**
	T$^-$	**0.150**	16.33	
PM6_10ps_No_Phosphate	N/A	Inflection	N/A	N/A
	N/A	Inflection	N/A	
PM6_10ps_No_Phosphate_Dispersion	N/A	Inflection	N/A	N/A
	N/A	Inflection	N/A	

Table 5.16: PM6 C-N bond lengths and free energy of intermediates

acyl chain, no intermediates were observed and the final structure was once again the T^0 intermediate. It would thus be prudent to obtain alternative starting structures from the potential energy surface to ensure formation of intermediates in the reaction pathway.

Aminolysis Mechanism

Unlike other systems, no intermediates were found along this reaction pathway; despite having good initial starting conditions for reactivity. The trajectory of the last frame of this simulation was extracted and energy minimised with PM6. Experimental literature confirms reactivity of this molecule via an aminolysis mechanism. Thus, this structure, which was the T^0 intermediate was deprotonated to form the T^- intermediate and then aminolysis took place to form the stable acylated drug (figure 5.26). DFT minimisations of these intermediates were carried out with the B3LYP and PBE0 (+D3 Dispersion) functionals in implicit solvent.

Despite an initial starting intermediate of T^0, similar trends for former aminolysis reactions were present for this system. As can be seen in table 5.17, reduction of the C-O bond length (0.141 to 0.128 nm) from the T^0 to T^- intermediate clearly shows deprotonation of the alcoholic proton to a stable T^- intermediate. A further

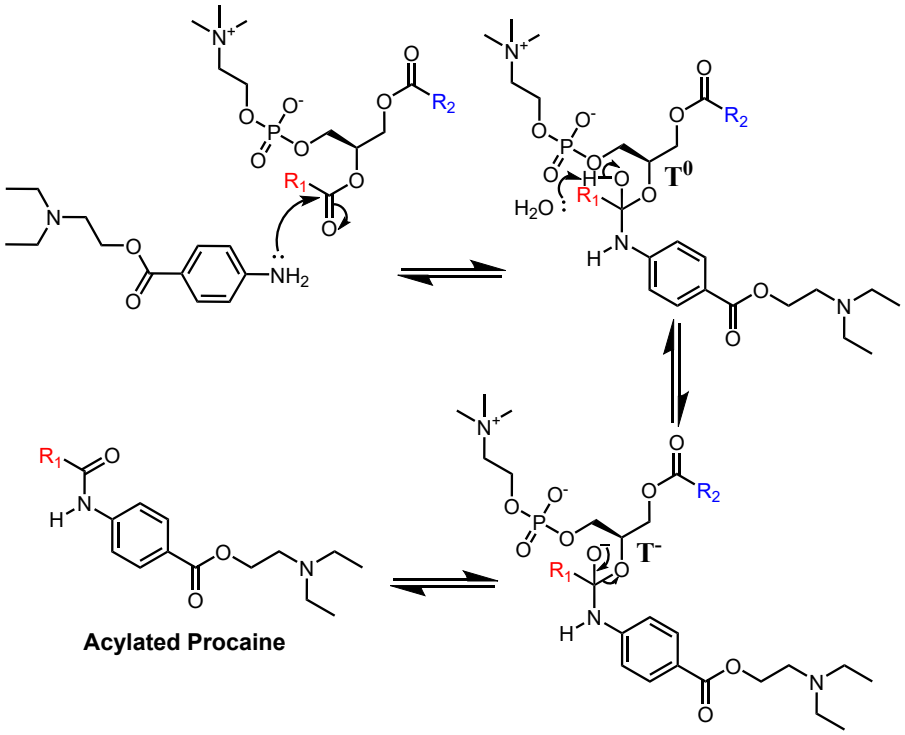


Figure 5.26: Reaction pathway mechanism for procaine

reduction in C-N bond length (0.153 to 0.137 nm) and reduction in C-O bond length (0.128 to 0.123 nm) points to re-formation of the carbonyl group and subsequent amide formation via an aminolysis mechanism. When compared directly with PBE0 data, respective C-N and C-O bond lengths for the different reactive intermediates remain almost identical.

Lastly, table 5.18 compares different relative free energies for the intermediates in the reaction pathway. Unlike with former drug-lipid systems, no direct comparison can be made between the $T^{\pm}$ and acylated drug intermediates, as this energy profile failed to reach an intermediate in the energy landscape. The relative free energy change between the T^0 and T^- intermediates is similar between the B3LYP (57.8 kcal mol^{-1}) and the PBE0 (+D3) functional (50.9 kcal mol^{-1}). A lower free energy change between the T^- and final acylated drug for the PBE0 functional (-27.3 kcal mol^{-1}) than for the B3LYP functional (-82.0 mol^{-1}) suggests that the PBE0, coupled with a dispersion term, more accurately predicts the energy change required for formation of the amide in aminolysis.

DFT Functional	Intermediate	C-N Bond length / nm	C-O Bond length / nm
B3LYP	T^0	**0.146**	**0.141**
B3LYP	T^-	**0.153**	**0.128**
B3LYP	Acylated Product	**0.137**	**0.123**
PBE0 (+ D3)	T^0	**0.145**	**0.139**
PBE0 (+ D3)	T^-	**0.151**	**0.128**
PBE0 (+ D3)	Acylated Product	**0.137**	**0.122**

Table 5.17: Bond lengths post DFT minimisation of reaction intermediates

Functional	Basis Set	Drug Intermediate	*G* / Hartrees	ΔG / kcal mol^{-1}
B3LYP	6-311+g(d,p)	T^0 (171 atoms)	-3416.18900182	
B3LYP	6-311+g(d,p)	T^- (170 atoms)	-3415.70830441	**57.8**
B3LYP	6-311+g(d,p)	Acylated Drug (171 atoms)	-3416.22764592	**-82.0**

Functional	Basis Set	Water	*G* / Hartrees
B3LYP	6-311+g(d,p)	Neutral	-76.4664127203
B3LYP	6-311+g(d,p)	Protonated	-76.8550364477

Functional	Basis Set	Drug Intermediate	*G* / Hartrees	ΔG / kcal mol^{-1}
PBE0 (+D3)	6-311+g(d,p)	T^0 (171 atoms)	-3412.53413167	
PBE0 (+D3)	6-311+g(d,p)	T^- (170 atoms)	-3412.05096878	**50.9**
PBE0 (+D3)	6-311+g(d,p)	Acylated Drug (171 atoms)	-3412.49664444	**-27.3**

Functional	Basis Set	Water	*G* / Hartrees
PBE0 (+D3)	6-311+g(d,p)	Neutral	-76.3694557125
PBE0 (+D3)	6-311+g(d,p)	Protonated	-76.7715468686

Table 5.18: DFT minimisation energies of reaction intermediates

5.3.8 Summary

In conclusion, this chapter initially presented results on the optimisation of distances and angles for nucleophilic attack. This data was presented as 1D histograms to show the distribution of distances and angles across 10,000 frames of the trajectory. In accordance with literature, protonated propranolol and 2-aminomethylbenzimidazole were found to have the highest probability of both distances and angles found close to the BD conditions required for nucleophilic attack. Both fluoxetine and phentermine exhibited a range of distances and angles quite far from the optimal BD conditions. 4-amino-*N*-phenylbutanamide, serotonin and procaine all displayed nucleophilic potential as demonstrated by initial starting conditions close to the BD optimal conditions for reactivity.

Secondly, a general QM/MM methodology was proposed to simulate reactivity of each molecule with the membrane lipids as a function of reaction coordinate. Initial N-C or O-C and C=O bond distances were calculated, alongside the N or O-C=O bond angle. These bond distances and angles were used as initial starting structures for QM/MM reactions. A well defined number of atoms were placed in the QM region and MM region and electrostatic embedding was used in all system to account for atoms across the QM/MM boundary. To understand stabilising effects in the membrane, these energy profiles were also compared with a single drug-lipid system and a simpler drug-lipid system without the phosphate headgroup, all in implicit solvent.

2-Aminomethylbenzimidazole showed a stable $T^{\pm}$ intermediate as the minimum in the energy profile. DFT minimisations of the reaction intermediates showed an aminolysis mechanism of reactivity. The formation of a stable acylated drug was energetically favourable. Also, the energy required to reach the initial TS was higher in the systems without membrane, implying the stabilising effects within the membrane. Propranolol/protonated propranolol was found to undergo transesterification and the minimum in the energy profile corresponded to the final stable acylated drug. Although no stable intermediates were found along the reaction pathway, a change in C-O and C=O bond distances as well as the trajectory confirmed formation of

the stable transesterified product.

Both 4-amino-*N*-phenylbutanamide and Serotonin displayed similar aminolysis behaviour to 2-aminomethylbenzimidazole. As previously seen with 2-aminomethyl benzimidazole, the minimum on the energy profile corresponded with the $T^{\pm}$ intermediate and the energy barrier required to reach the TS was lower in the membrane. DFT minimisations of all reaction intermediates were performed to show the favourable formation of the final acylated drug. Lastly, although procaine is thought to undergo aminolysis in the membrane, no intermediates were found along the reaction pathway. However, the final structure in the trajectory was extracted and subjected to DFT minimisations. Post minimisation, a stable acylated form of the drug was found. Future work is required to perform DFT minimisations of these drug-lipid systems in the POPC membrane. This would enable a comparison of relative free energy changes between the different intermediates in the reaction pathway in a membrane compared with outside a membrane.

Chapter 6

Experimental

6.1 General Methods

6.1.1 Materials

^{15}N isotopically enriched molecules were purchased from either Merck Life Science UK Ltd. (Dorset, UK) or Goss Scientific Instruments Ltd. (Crewe, UK). Phospholipids (POPC) were purchased from Avanti Polar Lipids (Amsterdam, Netherlands). Solvents were all purchased from either Fisher Scientific (Loughborough, UK) or Merck Life Science UK Ltd. (Dorset, UK). Deuterated solvents were all purchased from Apollo Scientific (Stockport, UK). All other chemicals required for synthetic work, including acids and bases, were all purchased from Merck Life Science UK Ltd. (Dorset, UK).

6.1.2 Solution State NMR

^{1}H, ^{13}C, ^{19}F NMR were performed at 400, 100 and 376 MHz respectively, using a Bruker spectrometer. All samples were prepared in either $CDCl_3$, DMSO-d_6, D_2O-d_2 or CD_3OD-d_4. J values are all reported in Hz. Chemical shifts (δ) are all reported in ppm, referenced according to the residual ^{1}H NMR solvent peak signal: 7.26 ppm for $CDCl_3$, 2.50 ppm for DMSO-d_6, 4.79 ppm for D_2O-d_2 and 3.31 ppm for CD_3OD-d_4. Also, referenced according to the residual ^{13}C NMR solvent peak signal: 77.16 ppm for $CDCl_3$, 39.52 ppm for DMSO-d_6 and 49.00 ppm for CD_3OD-d_4 [74].

6.1.3 Mass Spectrometry

All mass characterisation was conducted on a Waters Acquity SQD and a TQD with ESI (EI^+ mode) Mass Spectrometer. The UPLC was equipped with a 1.7 μm ACQUITY UPLC BEH C18 column, with an inner diameter of 2.1 mm and length of 50 mm. The mobile phase was composed of water with formic acid (0.1% w/v). The methanol mobile phase was set at a gradient concentration of 0.6 mL min^{-1} (1% MeOH: 99% H_2O to 99% MeOH: 1% H_2O over 5 min).

6.1.4 Drug Buffer Preparation

0.1 M citric acid and 0.2 M Na_2HPO_4 were fully dissolved in D_2O. pH was measured and buffered to a value of 7.4 using NaOD and DCl, according to the pH value.

6.1.5 Drug Buffer Stock Solutions

Each respective ^{15}N labelled compound for ssNMR experiments was dissolved in the buffer solution (D_2O) and UV-Vis measurements were conducted to calculate an accurate concentration of the drug in buffer solution. These calculations were conduced between 180-400 nm on a CARY100 UV-Visible spectrophotometer with Cary WinUV Scan Software 3.00[97]. For each drug buffer solution, the accurate concentration was calculated from the difference in absorbance between the drug in buffer solution and the buffer only solution (control). Absorbance was measured at 7.5 μL, 15 μL and 30 μL aliquot volumes made up to a total volume of 1.5 mL in a 2 mL quartz cuvette. Application of the Beer-Lambert Law provided accurate concentrations of the drug in buffer solution (table 6.1).

^{15}N labelled compounds	Absorbance λ / nm	Extinction Coefficient / M^{-1} cm^{-1}	Accurate Concentration / mg ml^{-1}
2-Aminomethylbenzimidazole	209	3715	**0.0586**
Propranolol	209	4880	**0.588**
4-Amino-*N*-phenylbutanamide	206	65000	**0.168**

Table 6.1: Accurate concentration calculations for ^{15}N labelled compounds

6.1.6 Liposome Preparation

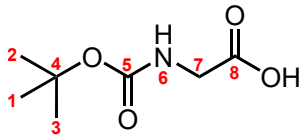
Stock solutions of POPC liposomes in $CHCl_3$ were prepared. An accurate amount of POPC was weighed out, dissolved in $CHCl_3$ to obtain a 10x lipid concentration relative to drug, and the solvent removed *in vacuo* to produce a thin-film layer. This was then stored in the desiccator overnight. The POPC lipid film was rehydrated with 150 μL of drug-buffer stock solution and underwent agitation with a vortex mixer[114]. To ensure good lipid dispersion, the mixture underwent five freeze-thaw cycles from -196°C (liquid N_2) to 37°C (RT). These samples were stored in the freezer (-20°C) until required for solid-state NMR experiments.

6.1.7 Solid State NMR

The maximally dispersed lipid-drug solution, at a concentration ratio of 10:1 respectively, was then subjected to ssNMR experiments. ^{1}H spectra were recorded using a Bruker Avance III HD spectrometer operating at 400.2 MHz and a 4 mm (rotor o.d.) magic-angle spinning (MAS) probe. They were obtained with a 4 s recycle delay. The sample spin rate was approximately 10 kHz. Spectral referencing is with respect to neat tetramethylsilane, carried out by setting the high-frequency signal from tetrakis(trimethylsilyl)silane to -9.9 ppm. ^{13}C magic-angle spinning measurements were carried out at 100.63 MHz using a Bruker Avance III HD spectrometer and 4 mm (rotor o.d.) probe. Spectra were acquired at a spin rate of 10 kHz. Carbon spectral referencing is relative to neat tetramethylsilane, carried out by setting the high-frequency signal from an external sample of adamantane to 38.5 ppm[238].

6.2 Chemical Synthesis

Synthesis of *N*-*tert*-butoxycarbonyl-glycine 16

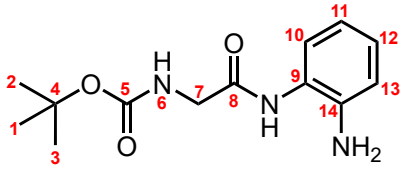


16

Di-*tert*-butyl dicarbonate (BOC anhydride) (3.05 g, 13.99 mmol) was added drop-

wise to a stirred solution of glycine (0.75 g, 9.99 mmol) in 1 M NaOH (10 mL) at 0°C over 30 min. Temperature was raised to 35 °C over 10 min, and stirring was continued for an additional 2.5 h. Hexane (20 mL) was then added to the solution and extracted with H_2O (2×25 mL) and saturated $NaHCO_3$ (2×25 mL). The combined aqueous layers were adjusted to pH = 1 and extracted with EtOAc (25 mL). This organic layer was washed once again with H_2O (2×25 mL). The organic layer was dried ($MgSO_4$), filtered and the solvent removed *in vacuo* to yield **16** as a pure white solid (0.561 g, 32%). **^{1}H NMR** ($CDCl_3$, 400 MHz) δ (ppm) 1.45 (s, 9H, $H_{1,2,3}$), 3.93 (d, J = 5.2 Hz, 2H, H_7), 6.73 (d, J = 5.2 Hz, 1H, H_6). **^{13}C NMR** ($CDCl_3$, 100 MHz) δ (ppm) 28.76 ($C_{1,2,3}$), 42.70 (C_7), 80.93 (C_4), 156.43 (C_5), 175.27 (C_8). **m.p.** = 86-87°C.[66]

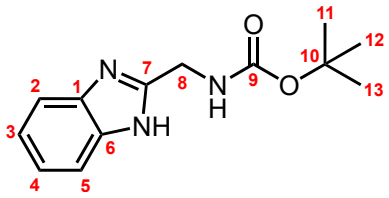
Synthesis of *tert*-butyl *N*-[2-(2-aminoanilino)-2-oxoethyl]carbamate 17



17

To a solution of *N*-*tert*-butoxycarbonyl-glycine **16** (0.28 g, 1.06 mmol) in THF (7 mL) was added 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide hydrochloride (0.22 g, 1.17 mmol), hydroxybenzotriazole (0.18 g, 1.17 mmol), o-phenylenediamine (0.13 g, 1.17 mmol) and Et_3N (0.22 mL, 1.59 mmol) and stirred for 12 h at room temperature. The solution was concentrated *in vacuo* and redissolved in DCM (20 mL). The organic layer was washed with H_2O (2×10 mL), brine (5 mL) and dried over $MgSO_4$. After removal of the drying agent by filtration, the solvent was removed *in vacuo* to yield **17** as a pure white powder (0.139 g, 49%). **^{1}H NMR** (DMSO, 400 MHz) δ (ppm) 1.39 (s, 9H, $H_{1,2,3}$), 3.73 (d, J = 6.0 Hz, 2H, H_7), 6.36 (d, J = 6.0 Hz, 1H, H_6), 6.70 (dd, J = 8.0, 1.4 Hz, 1H, H_{11}), 6.90 (td, J = 7.6, 1.6 Hz, 1H, H_{13}), 7.02 (t, J = 6.0 Hz, 1H, H_{12}), 7.12 (dd, J = 7.8, 1.5 Hz, 1H, H_{10}). **^{13}C NMR** (DMSO, 100 MHz) δ (ppm) 28.23 ($C_{1,2,3}$), 43.63 (C_7), 78.11 (C_4), 115.62 (C_{13}), 116.04 (C_{10}), 125.65 (C_{11}), 126.06 (C_{12}), 134.94 (C_{14}), 142.32 (C_9), 155.99 (C_5), 168.27 (C_8). **LRMS (ESI)** *m/z* 266.26 $[M+H]^+$; **HRMS (ASAP)** calculated for $C_{13}H_{19}N_3O_3$ $[M+H]^+$ 266.1510, found 266.1509. **m.p.** = 109-110°C.[71]

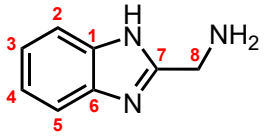
Synthesis of *tert*-butyl *N*-(1H-benzimidazol-2-ylmethyl)carbamate 18



18

To a solution of *tert*-butyl *N*-[2-(2-aminoanilino)-2-oxoethyl]carbamate **17** (0.12 g, 0.45 mmol) was added AcOH (3 mL) and stirred for 12 h under reflux (72°C). The solvent was removed *in vacuo* to yield the crude product. Excess acetic acid was removed under vacuum. This yielded **18** as an orange oil (0.10 g, 90%). **^{1}H NMR** ($CDCl_3$, 400 MHz) δ (ppm) 1.39 (s, 9H, $H_{11,12,13}$), 4.61 (d, J = 5.0 Hz, 2H, H_8), 6.85 (d, J = 5.0 Hz, 1H, NH), 7.19 – 7.32 (m, 2H, $H_{3,4}$), 7.53 (tt, J = 7.1, 4.2 Hz, 2H, $H_{2,5}$). **^{13}C NMR** ($CDCl_3$, 100 MHz) δ (ppm) 28.62 ($C_{11,12,13}$), 37.86 (C_8), 80.94 (C_{10}), 114.18 (C_2), 115.00 (C_5), 121.45 (C_3), 124.09 (C_4), 124.99 (C_6), 135.77 (C_1), 152.34 (C_7), 157.46 (C_9). **LRMS (ESI)** *m/z* 248.25 $[M+H]^+$; **HRMS (ASAP)** calculated for $C_{13}H_{17}N_3O_2$ $[M+H]^+$ 248.1376, found 248.1378.[82]

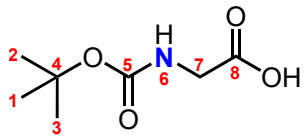
Synthesis of 2-(aminomethyl)benzimidazole 19



19

To a solution of *tert*-butyl *N*-(1H-benzimidazol-2-ylmethyl)carbamate **18** (0.020 g, 0.08 mmol) in MeOH (1 mL) was added 2 M HCl (aq., 1 mL) and stirred at room temperature for 1 h. The solvent was removed *in vacuo*. The residue was triturated with Et_2O (1 mL) and centrifugation was applied to obtain the pure pink solid precipitate **19** (0.010 g, 90%). **^{1}H NMR** (MeOD, 400 MHz) δ (ppm) 4.84 (s, 2H, H_8), 7.66 – 7.79 (m, 2H, $H_{3,4}$), 7.90 – 7.99 (m, 2H, $H_{2,5}$). **^{13}C NMR** (MeOD, 100 MHz) δ (ppm) 35.79 (C_8), 115.80 ($C_{2,5}$), 128.78 ($C_{3,4}$), 132.79 ($C_{1,6}$), 146.35 (C_7). **LRMS (ESI)** *m/z* 148.14 $[M+H]^+$; **HRMS (ASAP)** calculated for $C_8H_9N_3$ $[M+H]^+$ 148.0880, found 148.0879. **m.p.** = 266-267°C.[224]

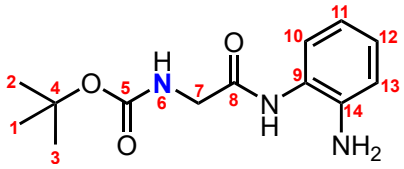
Synthesis of ^{15}N-*tert*-butoxycarbonyl-glycine 16



^{15}N **16**

BOC anhydride (3.05 g, 13.99 mmol) was added dropwise to a stirred solution of ^{15}N glycine (0.75 g, 9.99 mmol) in 1 M NaOH (10 mL) at 0°C over 30 min. The temperature was raised to 35°C over 10 min, and stirring was continued for an additional 2.5 h. Hexane (20 mL) was then added to the solution and extracted with H_2O (2×25 mL) and saturated $NaHCO_3$ (2×25 mL). The combined aqueous layers were adjusted to pH = 1 and extracted with EtOAc (25 mL). This organic layer was washed once again with H_2O (2×25 mL). The organic layer was dried ($MgSO_4$), filtered and the solvent removed *in vacuo* to yield ^{15}N **16** as a pure white solid (0.52 g, 30%). **^{1}H NMR** ($CDCl_3$, 400 MHz) δ (ppm) 1.45 (s, 9H, $H_{1,2,3}$), 3.93 (d, J = 5.2 Hz, 2H, H_7), 6.70 (d, J = 5.2 Hz, 1H, H_6). **^{13}C NMR** (100 MHz, $CDCl_3$) δ (ppm) 28.44 ($C_{1,2,3}$), 42.44 (C_7), 80.62 (C_4), 156.24 (C_5), 174.86 (C_8). **m.p.** = 86-87°C.[66]

Synthesis of *tert*-butyl ^{15}N–[2-(2-aminoanilino)-2-oxoethyl]carbamate 17

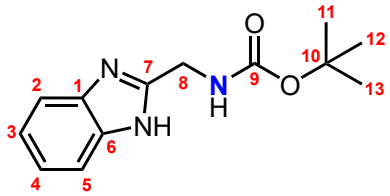


^{15}N **17**

To a solution of ^{15}N-*tert*-butoxycarbonyl-glycine **16** (0.25 g, 1.43 mmol) in THF (7 mL) was added 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide hydrochloride (0.3 g, 1.57 mmol), hydroxybenzotriazole (0.24 g, 1.57 mmol), o-phenylenediamine (0.17 g, 1.57 mmol) and Et_3N (0.3 mL, 2.15 mmol) and stirred for 12 h at room temperature. The solution was concentrated *in vacuo* and redissolved in DCM (20 mL). The organic layer was washed with H_2O (2×10 mL), brine (5 mL) and dried over $MgSO_4$. After removal of the drying agent by filtration, the solvent was removed *in vacuo* to yield ^{15}N **17** as a pure brown solid (0.224 g, 59%). **^{1}H NMR** (DMSO,

400 MHz) δ (ppm) 1.40 (s, 9H, $H_{1,2,3}$), 3.73 (d, J = 6.0 Hz, 2H, H_7), 6.36 (d, J = 6.0 Hz, 1H, H_6), 6.53 (td, J = 7.6, 1.4 Hz, 1H, H_{11}), 6.70 (dd, J = 8.0, 1.4 Hz, 1H, H_{13}), 6.85 – 6.97 (m, 1H, H_{12}), 7.09 – 7.20 (m, 1H, H_{10}). **^{13}C NMR** (DMSO, 100 MHz) δ (ppm) 28.39 ($C_{1,2,3}$), 43.84 (C_7), 78.24 (C_4), 115.77 (C_{13}), 116.18 (C_{10}), 125.79 (C_{11}), 126.20 (C_{12}), 136.23 (C_{14}), 142.47 (C_9), 156.26 (C_5), 168.42 (C_8). **LRMS (ESI)** *m/z* 267.29 $[M+H]^+$; **HRMS (ASAP)** calculated for $C_{13}H_{19}{}^{15}N_3O_3$ $[M+H]^+$ 267.1479, found 267.1479. **m.p.** = 109-110°C.[71]

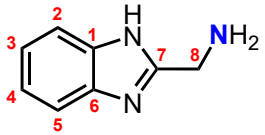
Synthesis of *tert*-butyl ^{15}N-(1H-benzimidazol-2-ylmethyl)carbamate 18



^{15}N **18**

To a solution of ^{15}N-[2-(2-aminoanilino)-2-oxoethyl]carbamate **17** (0.12 g, 0.45 mmol) was added AcOH (3 mL) and stirred for 12 h under reflux (72°C). The solvent was removed *in vacuo* to yield the crude product. Excess acetic acid was removed under vacuum. This yielded ^{15}N **18** as an orange oil (0.11 g, 98%). **^{1}H NMR** ($CDCl_3$, 400 MHz) δ (ppm) 1.40 (s, 9H, $H_{11,12,13}$), 4.59 (d, J = 4.6 Hz, 2H, H_8), 6.88 (d, J = 4.6 Hz, 1H, NH) 7.22 (dt, J = 6.2, 3.3 Hz, 2H, $H_{3,4}$), 7.51 (tt, J = 5.5, 3.3 Hz, 2H, $H_{2,5}$). **^{13}C NMR** ($CDCl_3$, 100 MHz) δ (ppm) 28.24 ($C_{11,12,13}$), 37.33 (C_8), 80.69 (C_{10}), 114.58 (C_2), 115.11 (C_5), 123.98 (C_3), 124.09 (C_4), 125.43 (C_6), 134.68 (C_1), 151.79 (C_7), 157.19 (C_9). **LRMS (ESI)** *m/z* 249.28 $[M+H]^+$; **HRMS** (ASAP) calculated for $C_{13}H_{17}{}^{15}N_3O_2$ $[M+H]^+$ 249.1373, found 249.1369.[82]

Synthesis of ^{15}N 2-(aminomethyl)benzimidazole 19

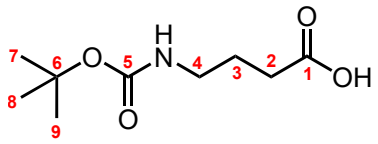


^{15}N **19**

To a solution of ^{15}N-(1H-benzimidazol-2-ylmethyl)carbamate **18** (0.13 g, 0.53 mmol)

in MeOH (3 mL) was added 2 M HCl (aq., 3 mL) and the solution stirred at room temperature for 1 h. The solvent was removed *in vacuo*. The residue was triturated with Et_2O (3 mL) and centrifugation was applied to obtain ^{15}N **19** as the pure pink solid precipitate (0.070 g, 89%). **^{1}H NMR** (MeOD, 400 MHz) δ (ppm) 3.35 (s, 2H, H_8), 7.58 – 7.69 (m, 2H, $H_{3,4}$), 7.81 – 7.90 (m, 2H, $H_{2,5}$). **^{13}C NMR** (MeOD, 100 MHz) δ (ppm) 35.52 (C_8), 115.51 ($C_{2,5}$), 128.45 ($C_{3,4}$), 132.59 ($C_{1,6}$), 146.07 (C_7). **^{15}N NMR** (MeOD, 400 MHz) δ (ppm) 30.77 (s, NH_2). **LRMS (ESI)** *m/z* 149.17 $[M+H]^+$; **HRMS (ASAP)** calculated for $C_8H_9{}^{15}N_3$ $[M+H]^+$ 149.0848, found 149.0845. **m.p.** = 266-267°C.[224]

Synthesis of 4-[(2-methylpropan-2-yl)oxycarbonylamino]butanoic acid 21

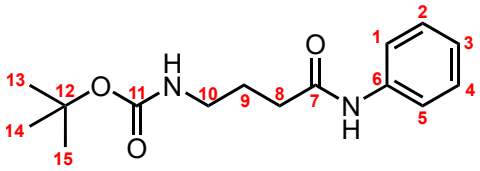


21

BOC anhydride (1.27 g, 5.82 mmol) in anhydrous THF (2.5 mL) was added dropwise to a solution of 4-aminobutanoic acid (0.5 g, 4.85 mmol) in H_2O (5 mL), THF (5 mL) and 1 M NaOH (0.39 g, 9.7 mmol) over 30 min. The solution was then allowed to stir at room temperature for 5 h. The solution was diluted with H_2O (10 mL) and washed with petroleum ether (40–60) (3×5 mL). The aqueous phase was acidified with 1 M HCl (pH 3) in an ice bath, and extracted with EtOAc (3×10 mL). The organic phase was dried over $MgSO_4$, filtered, and the solvent removed *in vacuo* to yield **21** a pure grey solid (0.877 g, 89%). **^{1}H NMR** (MeOD, 400 MHz) δ (ppm) 1.43 (s, 9H, $H_{7,8,9}$), 1.77 (p, J = 7.2 Hz, 2H, H_3), 2.33 (t, J = 7.4 Hz, 2H, H_2), 3.10 (t, J = 6.9 Hz, 2H, H_4). **^{13}C NMR** (MeOD, 100 MHz) δ (ppm) 25.80 (C_3), 28.56 ($C_{7,8,9}$), 31.74 (C_2), 40.25 (C_4), 79.49 (C_6), 157.80 (C_5), 176.55 (C_1). **m.p.** = 59-60°C.[28]

Synthesis of *tert*-butyl *N*-(4-oxo-4-phenylaminobutyl)carbamate 22

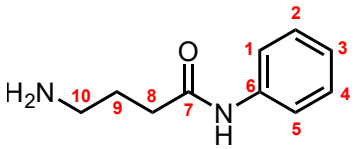
Carbonyldiimidazole (0.77 g, 4.75 mmol) was added to a solution of 4-[(2-methylpropan-2-yl)oxycarbonylamino]butanoic acid **21** (0.877 g, 4.32 mmol) in DCM (10 mL) and



22

stirred at room temperature for 1 h. Aniline (0.43 mL, 4.75 mmol) in DCM (10 mL) was then added to the solution and stirred at room temperature for 16 h. The reaction was then quenched with 5 mL NaOH (1 M) for 15 min. The organic layer was washed with 1 M HCl (3×10 mL), dried over $MgSO_4$, filtered, and concentrated to afford **22** as a pure light brown solid (0.281 g, 23%). **^{1}H NMR** (DMSO, 400 MHz) δ (ppm) 1.37 (s, 9H, $H_{13,14,15}$), 1.56 – 1.74 (m, 2H, H_9), 2.28 (t, J = 7.5 Hz, 2H, H_8), 2.87 – 3.00 (m, 2H, H_{10}), 7.01 (ddt, J = 8.6, 7.1, 1.2 Hz, 1H, H_3), 7.23 – 7.32 (m, 2H, $H_{2,4}$), 7.54 – 7.62 (m, 2H, $H_{1,5}$). **^{13}C NMR** (DMSO, 100 MHz) δ (ppm) 25.32 (C_9), 28.05 ($C_{13,14,15}$), 33.56 (C_8), 33.77 (C_{10}), 77.24 (C_{12}), 118.70 ($C_{1,5}$), 122.71 (C_3), 128.43 ($C_{2,4}$), 138.99 (C_6), 155.4 (C_{11}), 170.58 (C_7). **m.p.** = 92-93°C.[227]

Synthesis of 4-amino-*N*-phenylbutanamide 23

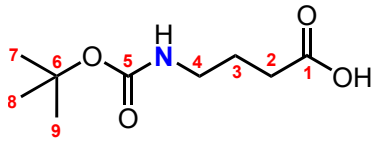


23

To a solution of *tert*-butyl *N*-(4-oxo-4-phenylaminobutyl)carbamate **22** (0.281 g, 0.1 M) in DCM (2.81 mL) was added an equal volume of 6-7 M iPrOH (2.81 mL, 0.05 M), bringing the total concentration to 0.05 M, and the reaction was stirred for 12 h at room temperature. The solvent was removed *in vacuo* and re-dissolved in DCM (5 mL), the residue was then triturated with Et_2O (5 mL), filtered and dried to obtain **23** as a pure pink solid (0.170 g, 96%). **^{1}H NMR** (MeOD, 400 MHz) δ (ppm) 1.83 (d, J = 7.4 Hz, 2H, H_9), 2.45 (t, J = 7.2 Hz, 2H, H_8), 2.72 – 2.88 (m, 2H, H_{10}), 7.02 (tt, J = 7.4, 1.2 Hz, 1H, H_3), 7.23 – 7.32 (m, 2H, $H_{2,4}$), 7.58 – 7.66 (m, 2H, $H_{1,5}$). **^{13}C NMR** (MeOD, 100 MHz) δ (ppm) 23.07 (C_9), 33.03 (C_8), 38.33 (C_{10}), 119.10 ($C_{1,5}$), 123.07 (C_3), 128.65 ($C_{2,4}$), 139.24 (C_6), 170.35 (C_7). **LRMS**

(ESI) *m/z* 179.16 $[M+H]^+$; **HRMS (ASAP)** calculated for $C_{10}H_{14}N_2O$ $[M+H]^+$ 179.1190, found 179.1184. **m.p.** = 115-116°C.[227]

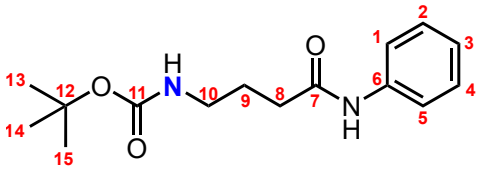
Synthesis of ^{15}N 4-[(2-methylpropan-2-yl)oxycarbonylamino]butanoic acid 21



^{15}N **21**

BOC anhydride (1.27 g, 5.82 mmol) in anhydrous THF (2.5 mL) was added dropwise to a solution of ^{15}N 4-aminobutanoic acid (0.5 g, 4.85 mmol) in H_2O (5 mL), THF (5 mL) and 1 M NaOH (0.39 g, 9.7 mmol) over 30 min. The solution was then allowed to stir at room temperature for 5 h. The mixture was diluted with H_2O (10 mL) and washed with petroleum ether (40–60) (3×5 mL). The aqueous phase was acidified with 1 M HCl (pH 3) in an ice bath, and extracted with EtOAc (3×10 mL). The organic phase was dried over $MgSO_4$, filtered, and the solvent removed *in vacuo* to yield ^{15}N **21** as a pure white solid (0.912 g, 92%). **^{1}H NMR** (MeOD, 400 MHz) δ (ppm) 1.43 (s, 9H, $H_{7,8,9}$), 1.77 (pd, J = 7.1, 2.6 Hz, 2H, H_3), 2.33 (t, J = 7.4 Hz, 2H, H_2), 3.10 (td, J = 6.9, 1.2 Hz, 2H, H_4). **^{13}C NMR** (MeOD, 100 MHz) δ (ppm) 25.81 (C_3), 28.61 ($C_{7,8,9}$), 31.73 (C_2), 40.30 (C_4), 79.48 (C_6), 157.88 (C_5), 176.56 (C_1). **m.p.** = 59-60°C.[28]

Synthesis of *tert*-butyl ^{15}N-(4-oxo-4-phenylaminobutyl)carbamate 22

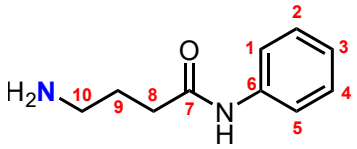


^{15}N **22**

Carbonyldiimidazole (0.80 g, 4.94 mmol) was added to a solution of ^{15}N 4-[(2-methylpropan-2-yl)oxycarbonylamino]butanoic acid **21** (0.912 g, 4.49 mmol) in DCM (10 mL) and stirred at room temperature for 1 h. Aniline (0.45 mL, 4.94 mmol) in

DCM (10 mL) was added and the reaction subsequently stirred at room temperature for 16 h. The reaction was then quenched with 5 mL NaOH (1 M) for 15 min. The organic layer was washed with 1 M HCl (3×10 mL), dried over $MgSO_4$, filtered, and concentrated to afford ^{15}N **22** as a pure light brown solid (0.389 g, 31%). **^{1}H NMR** (DMSO, 400 MHz) δ (ppm) 1.37 (s, 9H, $H_{13,14,15}$), 1.54 – 1.74 (m, 2H, H_9), 2.28 (td, J = 7.5, 2.0 Hz, 2H, H_8), 2.87 – 3.00 (m, 2H, H_{10}), 6.97 – 7.05 (m, 1H, H_3), 7.23 – 7.32 (m, 2H, $H_{2,4}$), 7.54 – 7.61 (m, 2H, $H_{1,5}$). **^{13}C NMR** (DMSO, 100 MHz) δ (ppm) 25.54 (C_9), 28.26 ($C_{13,14,15}$), 30.62 (C_8), 33.82 (C_{10}), 77.44 (C_{12}), 119.01 ($C_{1,5}$), 122.92 (C_3), 128.62 ($C_{2,4}$), 139.31 (C_6), 155.47 (C_{11}), 170.88 (C_7). **LRMS (ESI)** m/z 280.29 $[M+H]^+$; **HRMS (ASAP)** calculated for $C_{15}H_{22}{}^{15}N_2O_3$ $[M+H]^+$ 280.1740, found 280.1681. **m.p.** = 92-93°C.[227]

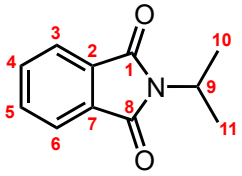
Synthesis of 4-amino-^{15}N-phenylbutanamide 23



^{15}N **23**

To a solution of *tert*-butyl ^{15}N-(4-oxo-4-phenylaminobutyl)carbamate **22** (0.281 g, 0.1 M) in DCM (2.81 mL) was added an equal volume of 6-7 M iPrOH (2.81 mL, 0.05 M), bringing the total concentration to 0.05 M, and the reaction was stirred for 12 h at room temperature. The solvent was removed *in vacuo* and re-dissolved in DCM (5 mL), the residue was then triturated with Et_2O (5 mL), filtered and dried to obtain ^{15}N **23** as a pure grey solid (0.161 g, 90%). **^{1}H NMR** (MeOD, 400 MHz) δ (ppm) 1.83 (qd, J = 7.4, 2.2 Hz, 2H, H_9), 2.45 (t, J = 7.2 Hz, 2H, H_8), 2.80 (tp, J = 11.7, 6.0 Hz, 2H, H_{10}), 7.02 (tt, J = 7.4, 1.2 Hz, 1H, H_3), 7.23 – 7.34 (m, 2H, $H_{2,4}$), 7.58 – 7.65 (m, 2H, $H_{1,5}$). **^{13}C NMR** (MeOD, 100 MHz) δ (ppm) 23.05 (C_9), 33.02 (C_8), 38.30 (C_{10}), 119.09 ($C_{1,5}$), 123.08 (C_3), 128.64 ($C_{2,4}$), 139.22 (C_6), 170.33 (C_7). **^{15}N NMR** (MeOD, 400 MHz) δ (ppm) 29.51 (s, NH_2). **LRMS (ESI)** m/z 180.18 $[M+H]^+$; **HRMS (ASAP)** calculated for $C_{10}H_{14}{}^{15}N_2O$ $[M+H]^+$ 180.1215, found 180.1154. **m.p.** = 115-116°C.[227]

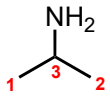
Synthesis of *N*-Isopropylphthalimide 10



10

Potassium phthalimide (1.1 g, 5.94 mmol) was added to a solution of 2-bromopropane (0.56 mL, 5.94 mmol) in anhydrous DMF (12.5 mL) and stirred under reflux (80°C) for 4 h. The solution was cooled to 0°C, added to water (10 mL) and extracted in DCM (15 mL). The organic layer was washed with saturated $NaHCO_3$ (3×5 mL) and dried over $MgSO_4$. After filtration, the solvent was removed *in vacuo* to yield **10** as a pure orange solid (0.59 g, 53%). **^{1}H NMR** ($CDCl_3$, 400 MHz) δ (ppm) 1.51 (d, J = 6.9 Hz, 6H, $H_{10,11}$), 4.55 (d, J = 6.9 Hz, 1H, H_9), 7.67 – 7.81 (m, 2H, $H_{4,5}$), 7.77 – 7.85 (m, 1H, H_3), 7.81 – 7.93 (m, 1H, H_6). **^{13}C NMR** ($CDCl_3$, 100 MHz) δ (ppm) 20.15 ($C_{10,11}$), 43.00 (C_9), 123.13 ($C_{3,6}$), 132.26 ($C_{2,7}$), 133.90 ($C_{4,5}$), 168.35 ($C_{1,8}$). **LRMS (ESI)** *m/z* 190.18 $[M+H]^+$; **HRMS (ASAP)** calculated for $C_{11}H_{11}NO_2$ $[M+H]^+$ 190.0868, found 190.0822. **m.p.** = 83-84°C.[102]

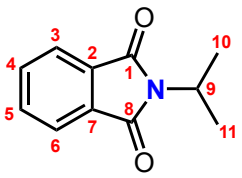
Synthesis of *N*-Isopropylamine 11



11

Hydrazine hydrate (0.165 mL, 3.3 mmol) was added to a solution of *N*-Isopropyl phthalimide **10** (0.25 g, 1.32 mmol) in anhydrous MeOH (5 mL) and heated under reflux (80°C) for 1 h, during which time a white solid precipitated out of solution. Water (1 mL) and 37% HCl (1 mL) were added to the mixture and continued heating for an additional 1 h. The solution was cooled to room temperature, the mixture was filtered, and the solvent removed from the filtrate to yield **11** as a pure yellow salt (0.215 g, 86%). **^{1}H NMR** (D_2O, 400 MHz) δ (ppm) 1.25 (d, J = 6.6 Hz, 6H, $H_{1,2}$), 3.44 (d, J = 6.7 Hz, 1H H_3). **^{13}C NMR** (D_2O, 100 MHz) δ (ppm) 19.74 ($C_{1,2}$), 44.00 (C_3). **m.p.** = 150-151°C.[202]

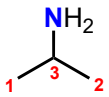
Synthesis of ^{15}N-Isopropylphthalimide 10



^{15}N **10**

^{15}N potassium phthalimide (1.765 g, 9.33 mmol) was added to a solution of 2-bromopropane (0.88 mL, 9.33 mmol) in anhydrous DMF (17.33 mL) and stirred under reflux (80°C) for 4 h. The solution was cooled to 0°C, added to water (10 mL) and extracted in DCM (15 mL). The organic layer was washed with saturated $NaHCO_3$ (3×5 mL) and dried over $MgSO_4$. After filtration, the solvent was removed *in vacuo* to yield ^{15}N **10** as a pure white solid (1.90 g, 75%). **^{1}H NMR** ($CDCl_3$, 400 MHz) δ (ppm) 1.49 (dd, J = 6.9, 2.6 Hz, 6H, $H_{10,11}$), 4.53 (dd, J = 7.0, 1.3 Hz, 1H, H_9), 7.64 – 7.92 (m, 4H, $H_{3,4,5,6}$). **^{13}C NMR** ($CDCl_3$, 100 MHz) δ (ppm) 20.28 ($C_{10,11}$), 43.17 (C_9), 123.12 (C_3), 123.14 (C_6), 132.29 (C_2), 133.92 (C_7), 134.49 ($C_{4,5}$), 168.62 ($C_{1,8}$). **LRMS (ESI)** *m/z* 191.09 $[M+H]^+$; **HRMS (ASAP)** calculated for $C_{11}H_{11}{}^{15}NO_2$ $[M+H]^+$ 191.0900, found 191.0839. **m.p.** = 83-84°C.[102]

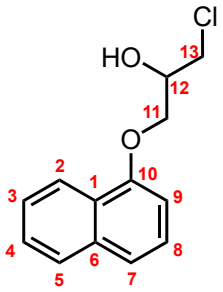
Synthesis of ^{15}N-Isopropylamine 11



^{15}N **11**

Hydrazine hydrate (0.32 mL, 6.6 mmol) was added to a solution of ^{15}N-Isopropyl phthalimide **10** (0.5 g, 2.64 mmol) in anhydrous MeOH (10 mL) and heated under reflux (80°C) for 1 h, during which time a white solid precipitated out of solution. Water (2 mL) and 37% HCl (2 mL) were added to the mixture and continued heating for an additional 1 h. The solution was cooled to room temperature, the mixture was filtered, and the solvent removed from the filtrate to yield ^{15}N **11** as a pure white salt (0.828 g, 83%). **^{1}H NMR** (D_2O, 400 MHz) δ (ppm) 1.23 (dd, J = 6.6, 3.2 Hz, 6H, $H_{1,2}$), 3.43 (dd, J = 6.6 Hz, 3.2 Hz, 1H, H_3).**^{13}C NMR** (D_2O, 100 MHz) δ (ppm) 22.22 ($C_{1,2}$), 46.49 (C_3). **m.p.** = 150-151°C.[202]

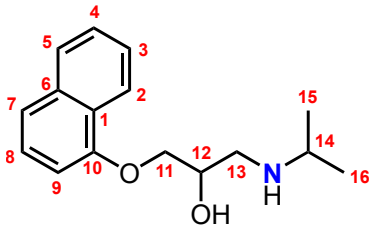
Synthesis of 1-Chloro-3-naphthalen-1-yloxypropan-2-ol 13



13

Epichlorohydrin (2.69 mL, 34.68 mmol) was added to a solution of 1-napthol (1 g, 6.94 mmol) in pyridine (60 μL) and stirred at 35°C for 24 h. Excess epichlorohydrin was removed *in vacuo* and the remaining solution re-dissolved in $CHCl_3$ (0.75 mL) and 37% HCl (0.375 mL). This solution was stirred for 1 h at 0°C. The mixture was diluted with $CHCl_3$ (25 mL), washed with H_2O (3×5mL), dried over $MgSO_4$ and after filtration the solvent was removed *in vacuo* to yield **13** as a dark red oil (1.15 g, 70%). **^{1}H NMR** ($CDCl_3$, 400 MHz) δ (ppm) 3.86 (d, J = 5.3 Hz, 2H, H_{13}), 4.12 – 4.27 (m, 2H, H_{11}), 4.35 (p, J = 5.2 Hz, 1H, H_{12}), 6.78 (dd, J = 7.7, 1.0 Hz, 1H, H_9), 7.43 (dd, J = 8.3, 7.6 Hz, 1H, H_4), 7.52 – 7.65 (m, 3H, $H_{2,7,8}$), 7.85 – 7.94 (m, 1H, H_3), 8.31 – 8.40 (m, 1H, H_5). **^{13}C NMR** ($CDCl_3$, 100 MHz) δ (ppm) 45.93 (C_{13}), 68.35 (C_{11}), 70.40 (C_{12}), 104.73 (C_9), 120.56 (C_2), 121.32 (C_1), 124.99 (C_5), 125.07 (C_7), 125.48 (C_3), 126.20 (C_4), 127.24 (C_8), 134.09 (C_6), 153.49 (C_{10}). **LRMS (ESI)** m/z 237.19 $[M+H]^+$; **HRMS (ASAP)** calculated for $C_{13}H_{13}{}^{35}ClO_2$ $[M+H]^+$ 237.0692, found 237.0682.[20]

Synthesis of ^{15}N Propranolol 14

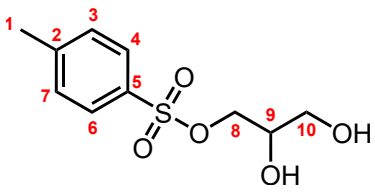


^{15}N **14**

To a solution of 1-Chloro-3-(1-naphthoxy)-2-propanol **13** (0.253 mL, 1.07 mmol) in isopropylamine (4.6 mL), was added ^{15}N isopropylamine hydrochloride (2.05 g, 21.4

mmol) in 20% NaOH sol. (4.6 mL). The mixture was stirred at RT for 16 h. Excess IPA was then removed *in vacuo*. The remaining aqueous mixture was washed with DCM (20 mL), dried over $MgSO_4$, and after filtration the solvent was removed *in vacuo* to yield ^{15}N **14** as a pink solid (0.103 g, 37%). **^{1}H NMR** ($CDCl_3$, 400 MHz) δ (ppm) 1.11 (d, J = 6.3 Hz, 6H, $H_{15,16}$), 2.76 – 2.94 (m, 2H, H_{13}), 3.01 (dd, J = 12.1, 3.3 Hz, 1H, H_{14}), 4.10 – 4.24 (m, 3H, $H_{11,12}$), 6.83 (dd, J = 7.6, 1.0 Hz, 1H, H_9), 7.37 (t, J = 7.9 Hz, 1H, H_4), 7.41 – 7.54 (m, 3H, $H_{2,7,8}$), 7.76 – 7.85 (m, 1H, H_3), 8.21 – 8.28 (m, 1H, H_5). **^{13}C NMR** ($CDCl_3$, 100 MHz) δ (ppm) 22.86 (C_{15}), 23.00 (C_{16}), 48.79 (C_{14}), 49.30 (C_{13}), 68.37 (C_{12}), 70.54 (C_{11}), 104.75 (C_9), 120.44 (C_2), 121.66 (C_1), 125.07 (C_5), 125.40 (C_7), 125.65 (C_3), 126.26 (C_4), 127.36 (C_8), 134.33 (C_6), 154.19 (C_{10}). **^{15}N NMR** (MeOD, 400 MHz) δ (ppm) 51.93 (s, NH). **LRMS (ESI)** *m/z* 260.26 $[M+H]^+$; **HRMS (ASAP)** calculated for $C_{16}H_{21}{}^{15}NO_2$ $[M+H]^+$ 260.1622, found 260.0921. **m.p.** = 90-91°C.[20]

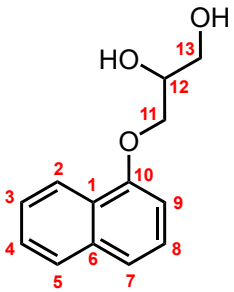
Synthesis of 2,3-dihydroxypropyl,4-methylbenzenesulfonate 6



6

To glycerol on molecular sieves (0.25 g, 2.71 mmol) was added p-toluenesulfonyl chloride (0.52 g, 2.71 mmol), dimethylaminopyridine (0.033 g, 0.271 mmol) and Et_3N (0.42 mL, 2.98 mmol) in anhydrous DCM (5 mL). The solution was stirred for 12 h at room temperature under Ar. The solvent was removed *in vacuo* to yield a crude product. Purification by silica gel chromatography using MeOH:DCM (5:95) yielded the compound **6** as a pure yellow oil (0.195 g, 29%). **^{1}H NMR** ($CDCl_3$, 400 MHz) δ (ppm) 2.01 (s, 3H, H_1), 3.12 - 3.62 (m, 5H, $H_{8,9,10}$), 6.93 (d, J = 7.6 Hz, 2H, $H_{3,7}$), 7.37 (d, J = 8.2 Hz, 2H, $H_{4,6}$). **^{13}C NMR** ($CDCl_3$, 100 MHz) δ (ppm) 21.30 (C_1), 62.44 (C_{10}), 69.35 (C_9), 70.54 (C_8), 127.63 ($C_{4,6}$), 129.72 ($C_{3,7}$), 131.96 (C_5), 144.92 (C_2). **LRMS (ESI)** *m/z* 247.15 $[M+H]^+$; **HRMS (ASAP)** calculated for $C_{10}H_{14}O_5S$ $[M+H]^+$ 247.0652, found 247.0654.[209]

Synthesis of 3-naphthalen-1-yloxypropane-1,2-diol 7

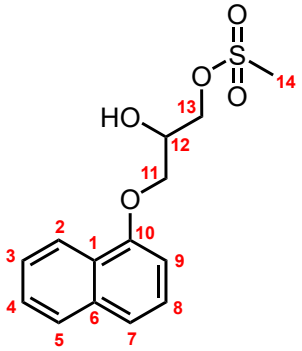


7

To 2,3-dihydroxypropyl,4-methylbenzenesulfonate **6** (0.194 g, 0.788 mmol) was added 1-napthol (0.136 g, 0.946 mmol), sodium iodide (0.0118 g, 0.0788 mmol) and potassium carbonate (0.436 g, 3.152 mmol) in MeCN (5 mL), and the solution was stirred under reflux (72°C) for 3.5 h. The solution was then diluted with H_2O (10 mL) and extracted with Et_2O (20 mL). The organic layer was washed with H_2O (10 mL), brine (10 mL), dried over $MgSO_4$, and after filtration the solvent was removed *in vacuo* to yield the crude product. Purification by silica gel chromatography using EtOAc yielded compound **7** as a pure reddish brown oil (0.025 g, 15%). **^{1}H NMR** ($CDCl_3$, 400 MHz) δ (ppm) 3.83 (dd, J = 11.5, 5.5 Hz, 2H, H_{13}), 3.91 (dd, J = 11.4, 3.7 Hz, 1H, H_{12}), 4.14 - 4.28 (m, 2H, H_{11}), 6.81 (dd, J = 7.6, 1.0 Hz, 1H, H_9), 7.32 - 7.38 (m, 1H, H_4), 7.43 - 7.52 (m, 3H, $H_{2,7,8}$), 7.76 - 7.84 (m, 1H, H_3), 8.18 - 8.27 (m, 1H, H_5). **^{13}C NMR** ($CDCl_3$, 100 MHz) δ (ppm) 63.65 (C_{13}), 69.11 (C_{12}), 70.40 (C_{11}), 104.87 (C_9), 120.77 (C_2), 121.41 (C_1), 125.23 (C_7), 125.26 (C_5) 125.61 (C_3), 126.35 (C_4), 127.45 (C_8), 134.34 (C_6), 153.87 (C_{10}). **LRMS (ESI)** *m/z* 219.25 $[M+H]^+$; **HRMS (ASAP)** calculated for $C_{13}H_{14}O_3$ $[M+H]^+$ 219.1007, found 219.1021.[116]

Synthesis of 2-hydroxy-3-(naphthalen-1-yloxy)propyl methanesulfonate 8

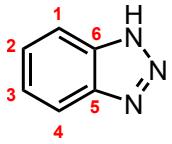
To 3-naphthalen-1-yloxypropane-1,2-diol **7** (0.025 g, 0.115 mmol) was added methanesulfonyl chloride (0.0089 mL, 0.115 mmol) and Et_3N (0.016 mL, 0.115 mmol) in anhydrous DCM (1 mL). The solution was stirred in an ice-bath for 2 h. The solution was diluted with H_2O (10 mL) and extracted with DCM (15 mL). The organic phase was washed with H_2O (5 mL), brine (5 mL), dried over $MgSO_4$, and after filtration



8

the solvent was removed *in vacuo* to yield a crude product. Purification by silica gel chromatography using MeOH:DCM (5:95) yielded compound **8** as a pure yellow oil (0.032 g, 94%). **^{1}H NMR** ($CDCl_3$, 400 MHz) δ (ppm) 3.09 (s, 3H, H_{14}), 4.17 - 4.23 (m, 2H, H_{13}), 4.48 - 4.56 (m, 3H, $H_{11,12}$), 6.85 (dd, J = 7.6, 1.0 Hz, 1H, H_9), 7.39 (dd, J = 8.3, 7.6 Hz, 1H, H_4), 7.45 - 7.59 (m, 3H, $H_{2,7,8}$), 7.77 - 7.88 (m, 1H, H_3), 8.19 - 8.28 (m, 1H, H_5). **^{13}C NMR** ($CDCl_3$, 100 MHz) δ (ppm) 37.79 (C_{14}), 68.31 (C_{13}), 68.63 (C_{12}), 70.55 (C_{11}), 105.26 (C_9), 121.46 (C_2), 121.61 (C_1), 125.48 (C_7), 125.75 (C_5), 125.90 (C_3), 126.80 (C_4), 127.86 (C_8), 134.70 (C_6), 153.84 (C_{10}). **LRMS (ESI)** *m/z* 297.24 $[M+H]^+$; **HRMS (ASAP)** calculated for $C_{14}H_{16}O_5S$ $[M+H]^+$ 297.0795, found 297.0797.[239]

Synthesis of 1*H*-Benzotriazole 24

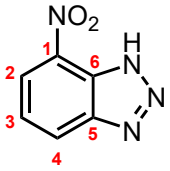


24

Tert-butyl nitrite (0.714 mL, 6 mmol) was added to a solution of o-phenylenediamine (0.324 g, 3 mmol) in MeCN (15 mL) and stirred at RT for 3 h. MeCN was then removed *in vacuo*, and the residue diluted with H_2O (30 mL) and extracted with EtOAc (40 mL). The organic layer was dried ($MgSO_4$), and after filtration the solvent was removed *in vacuo* to yield a crude product. Purification by silica gel chromatography using EtOAc yielded compound **24** as a pure red oil (0.268 g, 75%). **^{1}H NMR** ($CDCl_3$, 400 MHz) δ (ppm) 7.37 – 7.48 (m, 2H, $H_{2,3}$), 7.87 –

7.96 (m, 2H, $H_{1,4}$). **^{13}C NMR** ($CDCl_3$, 100 MHz) δ (ppm) 115.09 (C_1), 117.18 (C_4), 120.68 ($C_{2,3}$), 134.62 (C_6), 138.98 (C_5). **LRMS (ESI)** *m/z* 120.06 $[M+H]^+$; **HRMS (ASAP)** calculated for $C_6H_5N_3$ $[M+H]^+$ 120.0545, found 120.0562.[10]

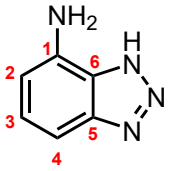
Synthesis of 4-nitro-1*H*-benzotriazole 25



25

To a solution of 1*H*-Benzotriazole **24** (0.268 g, 2.25 mmol) in conc. H_2SO_4 (0.81 mL) was added conc. HNO_3 (0.34 mL) dropwise at 0°C. The solution was stirred at RT for 12 h. The solution was then poured onto ice, and the precipitate collected by filtration. This was washed with H_2O (5 mL) and re-crystallised from EtOH (3 mL) to yield compound **25** as a pure brown solid (0.1 g, 27%). **^{1}H NMR** ($CDCl_3$, 400 MHz) δ (ppm) 7.64 (t, J = 8.0 Hz, 1H, H_3), 8.47 (dd, J = 7.9, 0.8 Hz, 1H, H_4), 8.59 (dd, J = 8.2, 0.8 Hz, 1H, H_2). **^{13}C NMR** ($CDCl_3$, 100 MHz) δ (ppm) 110.21 (C_4), 123.89 (C_1), 123.93 ($C_{2,3}$), 126.97 (C_5), 133.33 (C_6). **LRMS (ESI)** *m/z* 164.94 $[M+H]^+$; **HRMS (ASAP)** calculated for $C_6H_4N_4O_2$ $[M+H]^+$ 165.0408, found 165.0413. **m.p.** = 166-167°C.[223]

Synthesis of 4-amino-1*H*-benzotriazole 26

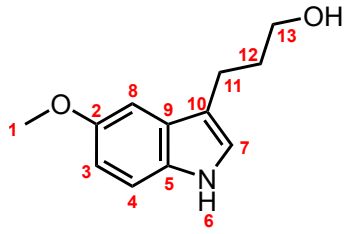


26

To a solution of 4-Nitro-1*H*-benzotriazole **25** in MeOH (4 mL) was added carbon (0.008 g, 10 mol%), K_2CO_3 (0.00677 g, 0.049 mmol), iron (III) chloride hexahydrate (0.00397 g, 0.0245 mmol) and the mixture was heated at reflux. Hydrazine hydrate (42.8 μL, 0.88 mmol) was added dropwise to the mixture and stirred for 2 h, then the mixture was cooled to RT and stirred for 12 h. The mixture was extracted with

DCM (5×3 mL) and dried ($MgSO_4$). After filtration, the solvent was evaporated *in vacuo* to yield the crude product. Upon recrystallisation with diisopropylether (5 mL) compound **26** was obtained as a pure red solid (0.103 g, 77%). **^{1}H NMR** ($CDCl_3$, 400 MHz) δ (ppm) 6.91 (dd, J = 8.2, 0.7 Hz, 1H, H_3), 7.06 – 7.21 (m, 1H, H_4) 7.58 (t, J = 8.0 Hz, 1H, H_2). **^{13}C NMR** ($CDCl_3$, 100 MHz) δ (ppm) 107.16 (C_4), 117.18 (C_1), 125.11 ($C_{2,3}$), 127.50 (C_5), 129.79 (C_6). **LRMS (ESI)** m/z 134.92 $[M+H]^+$; **HRMS (ASAP)** calculated for $C_6H_6N_4$ $[M+H]^+$ 135.0567, found 135.0571. **m.p.** = 81-82°C.[219]

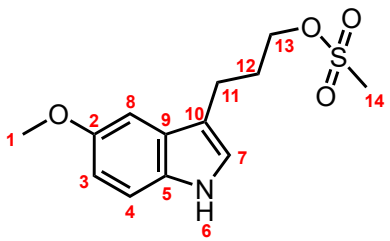
Synthesis of 3-(5-methoxy-1*H*-indol-3-yl)propanol 27



27

Dihydrofuran (0.63 mL, 6.92 mmol) was added dropwise to a solution of 4-methoxy phenyl hydrazine hydrochloride (1 g, 6.92 mmol) in 4% sulphuric acid (10 mL) and DMA (10 mL) at 60°C over 2 min. The solution was stirred for 3 h at reflux and then cooled to room temperature. The mixture was extracted with EtOAc (20 mL) and washed with H_2O (3×20 mL). The solution was dried over $MgSO_4$, filtered and solvent removed *in vacuo* to yield the crude product. Purification by silica gel chromatography using EtOAc yielded compound **27** as a pure yellow oil (0.854 g, 60%). **^{1}H NMR** ($CDCl_3$, 400 MHz) δ (ppm) 1.98 (dq, J = 7.4, 6.5 Hz, 2H, H_{12}), 2.78 – 2.88 (m, 2H, H_{11}), 3.72 (t, J = 6.4 Hz, 2H, H_{13}), 3.88 (s, 3H, H_1), 6.85 – 6.93 (m, 2H, $H_{3,4}$), 7.08 (d, J = 2.4 Hz, 1H, H_8), 7.21 (dd, J = 8.8, 0.6 Hz, 1H, H_7). **^{13}C NMR** ($CDCl_3$, 100 MHz) δ (ppm) 21.05 (C_{11}), 32.41 (C_{12}), 55.73 (C_1), 62.23 (C_{13}), 100.57 (C_8), 111.65 (C_{10}), 115.12 (C_4), 122.08 (C_3), 127.50 (C_7), 129.62 (C_9), 132.24 (C_5), 153.41 (C_2). **LRMS (ESI)** m/z 206.14 $[M+H]^+$; **HRMS (ASAP)** calculated for $C_{12}H_{15}NO_2$ $[M+H]^+$ 206.1183, found 206.1181.[210]

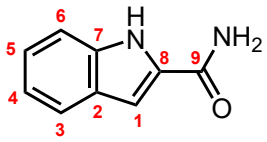
Synthesis of 3-(5-methoxy-1*H*-indol-3-yl)propyl methanesulfonate 28



28

To a solution of 3-(5-methoxy-1*H*-indol-3-yl)propanol **27** (0.5 g, 2.44 mmol) in anhydrous DCM (10 mL) was added Et_3N (0.34 mL, 2.44 mmol) and methanesulfonyl chloride (0.19 mL, 2.44 mmol) and stirred in an ice-water bath for 2 h. The solution was diluted with H_2O (10 mL) and extracted with DCM (15 mL). The organic phase was washed with H_2O (5 mL), brine (5 mL) and dried over $MgSO_4$. After filtration, the solvent was removed *in vacuo* to yield the crude product. Purification by silica gel chromatography using EtOAc yielded compound **28** as a pure colourless oil (0.57 g, 82%). **^{1}H NMR** ($CDCl_3$, 400 MHz) δ (ppm) 2.04 – 2.18 (m, 2H, H_{12}), 2.74 – 2.91 (m, 2H, H_{11}), 2.91 – 3.04 (m, 3H, H_{14}), 3.88 (d, J = 2.1 Hz, 3H, H_1), 4.21 – 4.30 (m, 2H, H_{13}), 6.86 (dd, J = 8.8, 2.4 Hz, 1H, H_8), 6.93 – 7.06 (m, 2H, $H_{3,4}$), 7.19 – 7.27 (m, 1H, H_7). **^{13}C NMR** ($CDCl_3$, 100 MHz) δ (ppm) 20.43 (C_{11}), 28.86 (C_{12}), 36.70 (C_{14}), 55.55 (C_1), 69.43 (C_{13}), 100.16 (C_8), 111.61 (C_{10}), 111.70 (C_4), 113.37 (C_3), 122.45 (C_7), 127.17 (C_9), 131.21 (C_5), 153.41 (C_2). **LRMS (ESI)** *m/z* 284.28 $[M+H]^+$; **HRMS (ASAP)** calculated for $C_{13}H_{17}NO_4S$ $[M+H]^+$ 284.0940, found 284.0941.[197]

Synthesis of 1*H*-indole-2-carboxamide 29

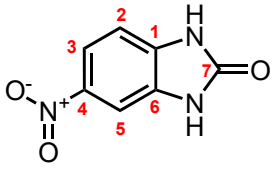


29

Carbonyldiimidazole (1.3 g, 8 mmol) was added to a solution of indole-2-carboxylic acid (0.64 g, 4 mmol) in dry THF (20 mL). The solution was stirred for 2 h at room temperature, after which ammonium hydroxide (3 mL) was added. The reaction was stirred for an additional 2 h before H_2O (6 mL) was added. The phases were

separated and the aqueous layer was extracted with EtOAc (2×15 mL). The organic extracts were washed with brine (15 mL), dried over $MgSO_4$, and after filtration, the solvent was removed *in vacuo* to yield **29** as a pure off-white solid (1.07 g, 83%). **^{1}H NMR** (DMSO, 400 MHz) δ (ppm) 7.02 (ddd, J = 8.0, 6.9, 1.0 Hz, 1H, H_4), 7.11 (dd, J = 2.2, 0.9 Hz, 1H, H_1), 7.17 (ddd, J = 8.2, 7.0, 1.1 Hz, 1H, H_5), 7.59 (dt, J = 8.0, 1.0 Hz, 1H, H_6), 7.64 (d, J = 1.1 Hz, 1H, H_3). **^{13}C NMR** (DMSO, 100 MHz) δ (ppm) 103.04 (C_1), 112.25 (C_6), 119.60 (C_3), 121.47 ($C_{4,5}$), 127.14 (C_8), 131.78 (C_2), 135.14 (C_7), 162.81 (C_9). **LRMS (ESI)** *m/z* 161.14 $[M+H]^+$; **HRMS (ASAP)** calculated for $C_9H_8N_2O$ $[M+H]^+$ 161.0717, found 161.0715. **m.p.** = 164-165°C.[122]

Synthesis of 5-nitro-1,3-dihydrobenzimidazol-2-one 30

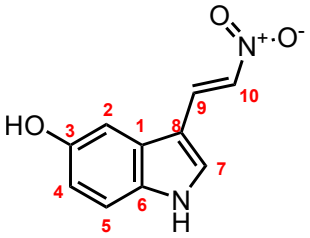


30

Carbonyldiimidazole (1.59 g, 9.78 mmol) was added to a solution of 4-nitro-o-phenylenediamine (1.0 g, 6.52 mmol) in anhydrous DMF (20 mL) at 0°C. The solution was then stirred at room temperature for 6 h. The mixture was filtered, and the solid filtrate was washed with H_2O (3×5 mL) and subsequently dried to give **30** as a pure yellow solid (1.42 g, 81%). **^{1}H NMR** (DMSO, 400 MHz) δ (ppm) 7.09 (dd, J = 8.6, 0.4 Hz, 1H, H_2), 7.70 (dd, J = 2.3, 0.3 Hz, 1H, H_3), 7.93 (dd, J = 8.6, 2.3 Hz, 1H, H_5). **^{13}C NMR** (DMSO, 100 MHz) δ (ppm) 103.63 (C_5), 108.07 (C_2), 117.77 (C_3), 129.72 (C_1), 135.70 (C_6), 141.26 (C_4), 155.45 (C_7). **LRMS (ESI)** *m/z* 180.14 $[M+H]^+$; **HRMS (ASAP)** calculated for $C_7H_5N_3O_3$ $[M+H]^+$ 180.0418, found 180.0409. **m.p.** = 304-305°C.[237]

Synthesis of 3-[(*E*)-2-nitroethenyl]-1*H*-indole 31

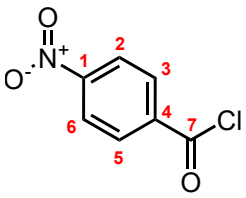
5-Hydroxyindole (0.375 g, 2.82 mmol) was added to a solution of 1-dimethylamino-2-nitroethylene (0.30 mL, 2.82 mmol) in TFA (1.5 mL) and stirred under Ar at 0°C for 30 min. The solution was then allowed to heat up to RT and poured onto



31

ice-cold H_2O, from which a dark red solid precipitated out. The aqueous layer was extracted with EtOAc (3×10 mL). The combined organic layers were washed with sat. $NaHCO_3$ (15 mL), brine (15 mL) and dried over $MgSO_4$. After filtration, the solvent was then removed *in vacuo* to yield a crude product. Purification by silica gel chromatography using EtOAc yielded compound **31** as a pure red solid (0.405 g, 70%). **^{1}H NMR** (MeOD, 400 MHz) δ (ppm) 6.82 (dd, J = 8.7, 2.3 Hz, 1H, H_4), 7.12 (dd, J = 2.3, 0.6 Hz, 1H, H_{10}), 7.30 (dd, J = 8.7, 0.6 Hz, 1H, H_5), 7.71 – 7.83 (m, 2H, $H_{2,7}$), 8.32 (d, J = 2.3 Hz, 1H, H_9). **^{13}C NMR** (MeOD, 100 MHz) δ (ppm) 105.69 (C_2), 109.39 (C_8), 114.09 (C_5), 114.04 (C_4), 126.98 (C_7), 131.36 (C_1), 133.65 (C_9), 135.86 (C_{10}), 136.22 (C_6), 154.19 (C_3). **LRMS (ESI)** m/z 205.19 $[M+H]^+$; **HRMS (ASAP)** calculated for $C_{10}H_8N_2O_3$ $[M+H]^+$ 205.0616, found 205.0613. **m.p.** = 304-305°C.[156]

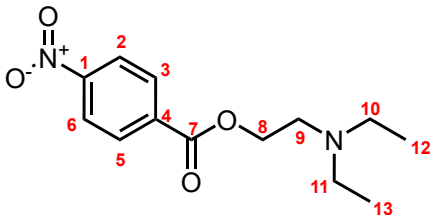
Synthesis of 4-nitrobenzoyl chloride 32



32

Oxalyl chloride (0.46 mL, 5.31 mmol) was added to a solution of 4-nitro-benzoic acid (0.59 g, 3.54 mmol) in a small amount of DMF (20 μL) in DCM (20 mL). The solution was then stirred at room temperature for 2 h. The solvent was removed *in vacuo* to yield **32** as a pure white-yellow solid (1.30 g, 99%). **^{1}H NMR** ($CDCl_3$, 400 MHz) δ (ppm) 8.23 - 8.41 (m, 4H, $H_{2,3,5,6}$). **^{13}C NMR** ($CDCl_3$, 100 MHz) δ (ppm) 124.19 ($C_{2,6}$), 132.38 ($C_{3,5}$), 138.16 (C_4), 151.73 (C_1), 167.18 (C_7). **m.p.** = 73-74°C.[55]

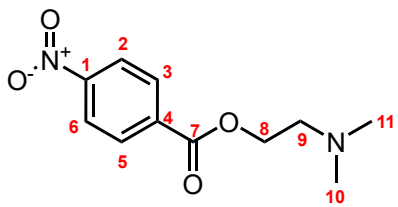
Synthesis of 2-(diethylamino)ethyl 4-nitrobenzoate 33



33

To a solution of 4-nitrobenzoyl chloride **32** (1 g, 5.39 mmol) and *N,N*-dimethyl aminopyridine (0.0658 g, 0.539 mmol) in anhydrous DCM (20 mL) was added diethylethanolamine (0.6 mL, 4.49 mmol). To the stirred solution was added Et_3N (0.94 mL, 6.74 mmol) dropwise at 0°C. The reaction was allowed to warm up to room temperature and stirred for an additional 12 h. 1 M HCl (10 mL) was then added to the reaction flask and the solution was extracted with DCM (3×10 mL) and the phases were separated. The organic phases were combined, washed with brine (10 mL), dried over $MgSO_4$, and after filtration, the solvent was removed *in vacuo* to yield **33** as a yellow solid (0.826 g, 57%). **^{1}H NMR** ($CDCl_3$, 400 MHz) δ (ppm) 1.46 (t, J = 7.3 Hz, 6H, $H_{12,13}$), 3.27 (q, J = 7.3 Hz, 4H, $H_{10,11}$), 3.46 (t, J = 5.5 Hz, 2H, H_9), 4.88 – 4.95 (m, 2H, H_8), 8.15 – 8.34 (m, 4H, $H_{2,3,5,6}$). **^{13}C NMR** ($CDCl_3$, 100 MHz) δ (ppm) 8.66 ($C_{12,13}$), 47.38 ($C_{10,11}$), 49.99 (C_9), 59.62 (C_8), 123.68 ($C_{2,6}$), 130.85 (C_4), 131.22 ($C_{3,5}$), 151.07 (C_1), 164.39 (C_7). **LRMS (ESI)** m/z 267.25 $[M+H]^+$; **HRMS (ASAP)** calculated for $C_{13}H_{18}N_2O_4$ $[M+H]^+$ 267.1364, found 267.1385. **m.p.** = 63-64°C.[221]

Synthesis of 2-(dimethylamino)ethyl 4-nitrobenzoate 34



34

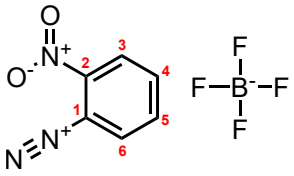
To a solution of 4-nitrobenzoyl chloride **32** (1.0 g, 5.39 mmol) in anhydrous DCM (40 mL) was added *N,N*-dimethylaminoethanol (0.54 mL, 5.39 mmol) dropwise and cooled to -5°C over 15 min. The reaction was then left to stir at RT for 48 h. The

precipitate was filtered, washed with Et_2O (20 mL) and dried for 1 h *in vacuo*. The solid was then dissolved in 1 M NaOH, extracted with Et_2O, dried over $MgSO_4$, filtered, and the solvent was evaporated *in vacuo* to yield compound **34** as a white solid (0.438 g, 34%). **^{1}H NMR** ($CDCl_3$, 400 MHz) δ (ppm) 2.22 (s, 6H, $H_{10,11}$), 2.65 (t, J = 5.6 Hz, 2H, H_9), 4.38 – 4.45 (m, 2H, H_8), 8.15 – 8.22 (m, 2H, $H_{3,5}$), 8.32 – 8.40 (m, 2H, $H_{2,6}$). **^{13}C NMR** ($CDCl_3$, 100 MHz) δ (ppm) 45.30 ($C_{10,11}$), 57.11 (C_9), 63.27 (C_8), 123.94 ($C_{2,6}$), 130.63 (C_4), 135.16 ($C_{3,5}$), 150.26 (C_1), 164.39 (C_7). **LRMS (ESI)** *m/z* 239.17 $[M+H]^+$; **HRMS (ASAP)** calculated for $C_{11}H_{14}N_2O_4$ $[M+H]^+$ 239.1067, found 239.1032. **m.p.** = 59-60°C.[109]

General Procedure for the Synthesis of Diazonium Tetrafluoroborates

To a solution of substituted aniline in EtOH (1.8 mL) was added HBF_4 (12 mmol, 0.75 mL) and stirred at 0°C. To this solution was added *tert*-butyl nitrite (12 mmol, 1.43 mL) dropwise and stirred at RT for 1 h. Diethyl ether was added and a precipitate formed. The mixture was filtered and further ether added. This process was repeated three times to ensure maximal yield of the substituted diazonium tetrafluoroborate.

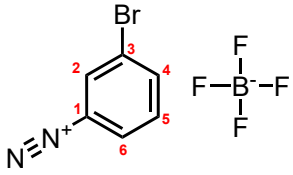
Synthesis of 2-nitrobenzenediazonium tetrafluoroborate 35



35

The generic method was applied to 2-nitroaniline (6 mmol, 0.82 g) to give compound **35** as a yellow solid (1.06 g, 75%). **^{1}H NMR** (DMSO, 400 MHz) δ (ppm) 8.41 (td, J = 7.9, 1.2 Hz, 1H, H_4), 8.49 – 8.58 (m, 1H, H_5), 8.79 (dd, J = 8.3, 1.2 Hz, 1H, H_6), 9.11 (dd, J = 8.1, 1.4 Hz, 1H, H_3). **^{13}C NMR** (DMSO, 100 MHz) δ (ppm) 111.13 (C_1), 128.02 (C_3), 136.51 (C_6), 136.54 (C_4), 142.25 (C_5), 144.50 (C_2). **LRMS (ESI)** *m/z* 149.82 $[M+H]^+$; **HRMS (ASAP)** calculated for $C_6H_4NO_2N_2$ $[M+H]^+$ 150.0307, found 150.0304.[129]

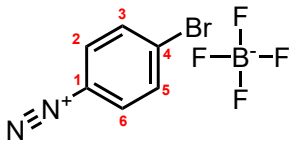
Synthesis of 3-bromobenzenediazonium tetrafluoroborate 36



36

The generic method was applied to 3-bromoaniline (8 mmol, 1.38 g) to give compound **36** as a brown solid (1.44 g, 67%). **^{1}H NMR** (DMSO, 400 MHz) δ (ppm) 7.92 (t, J = 8.3 Hz, 1H, H_5), 8.48 (ddd, J = 8.3, 2.0, 1.0 Hz, 1H, H_6), 8.69 (ddd, J = 8.3, 2.0, 1.0 Hz, 1H, H_4), 8.97 (t, J = 2.0 Hz, 1H, H_2). **^{13}C NMR** (DMSO, 100 MHz) δ (ppm) 117.84 (C_1), 122.27 (C_2), 131.89 (C_3), 132.82 (C_6), 134.25 (C_5), 143.76 (C_4). **LRMS (ESI)** m/z 184.78 $[M+H]^+$; **HRMS (ASAP)** calculated for $C_6H_4BrN_2$ $[M+H]^+$ 184.4583, found 184.4558.[129]

Synthesis of 4-bromobenzenediazonium tetrafluoroborate 37

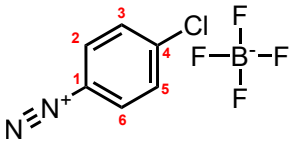


37

The generic method was applied to 4-bromoaniline (8 mmol, 1.38 g) to give compound **37** as a brown solid (1.64 g, 76%). **^{1}H NMR** (DMSO, 400 MHz) δ (ppm) 8.21 – 8.31 (m, 2H, $H_{3,5}$), 8.53 – 8.63 (m, 2H, $H_{2,6}$). **^{13}C NMR** (DMSO, 100 MHz) δ (ppm) 115.17 (C_1), 133.96 ($C_{3,5}$), 134.52 ($C_{2,6}$), 136.52 (C_4). **LRMS (ESI)** m/z 184.78 $[M+H]^+$; **HRMS (ASAP)** calculated for $C_6H_4BrN_2$ $[M+H]^+$ 184.9574, found 184.9558.[129]

Synthesis of 4-chlorobenzenediazonium tetrafluoroborate 38

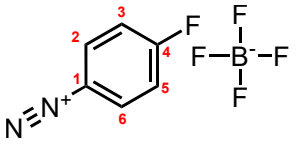
The generic method was applied to 4-chloroaniline (8 mmol, 1.02 g) to give compound **38** as a white solid (1.5 g, 83%). **^{1}H NMR** (DMSO, 400 MHz) δ (ppm) 8.06 – 8.17 (m, 2H, $H_{3,5}$), 8.63 – 8.75 (m, 2H, $H_{2,6}$). **^{13}C NMR** (DMSO, 100 MHz)



38

δ (ppm) 114.74 (C_1), 131.60 ($C_{3,5}$), 134.39 ($C_{2,6}$), 146.50 (C_4). **LRMS (ESI)** *m/z* 138.87 $[M+H]^+$; **HRMS (ASAP)** calculated for $C_6H_4{}^{35}ClN_2$ $[M+H]^+$ 139.0070, found 139.0063.[129]

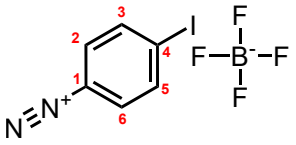
Synthesis of 4-fluorobenzenediazonium tetrafluoroborate 39



39

The generic method was applied to 4-fluoroaniline (8 mmol, 0.768 mL) to give compound **39** as a white solid (1.42 g, 85%). **^{1}H NMR** (DMSO, 400 MHz) δ (ppm) 7.84 – 7.94 (m, 2H, $H_{3,5}$), 8.75 – 8.85 (m, 2H, $H_{2,6}$). **^{13}C NMR** (DMSO, 100 MHz) 111.83 (C_1), 119.51 ($C_{3,5}$), 137.03 ($C_{2,6}$), 167.07 (C_4). **^{19}F NMR** (DMSO, 376 MHz) δ (ppm) -87.18, -148.23. **LRMS (ESI)** *m/z* 122.91 $[M+H]^+$; **HRMS (ASAP)** calculated for $C_6H_4FN_2$ $[M+H]^+$ 123.0357, found 123.0359.[129]

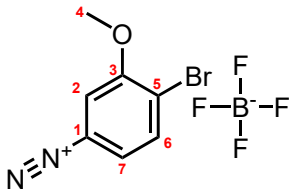
Synthesis of 4-iodobenzenediazonium tetrafluoroborate 40



40

The generic method was applied to 4-iodoaniline (8 mmol, 1.75 g) to give compound **40** as a white solid (1.93 g, 76%). **^{1}H NMR** (DMSO, 400 MHz) δ (ppm) 8.31 – 8.39 (m, 2H, $H_{3,5}$), 8.39 – 8.48 (m, 2H, $H_{2,6}$). **^{13}C NMR** (DMSO, 100 MHz) δ (ppm) 113.67 (C_4), 115.17 (C_1), 132.86 ($C_{3,5}$), 140.24 ($C_{2,6}$). **^{19}F NMR** (DMSO, 376 MHz) δ (ppm) -148.23. **LRMS (ESI)** *m/z* 230.81 $[M+H]^+$; **HRMS (ASAP)** calculated for $C_6H_4IN_2$ $[M+H]^+$ 230.9431, found 230.9419.[129]

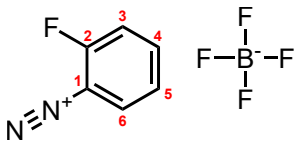
Synthesis of 4-bromo-3-methoxybenzenediazonium tetrafluoroborate 41



41

The generic method was applied to 4-bromo-3-methoxyaniline (8 mmol, 1.06 mL) to give compound **41** as a green solid (1.72 g, 71%). **^{1}H NMR** (DMSO, 400 MHz) δ (ppm) 4.00 (s, 3H, H_4), 8.20 – 8.31 (m, 2H, $H_{6,7}$), 8.42 (d, J = 2.2 Hz, 1H, H_2). **^{13}C NMR** (DMSO, 100 MHz) δ (ppm) 57.77 (C_4), 114.18 (C_1), 115.40 (C_5), 126.28 (C_2), 126.60 (C_7), 135.55 (C_6), 156.48 (C_3). **^{19}F NMR** (DMSO, 376 MHz) δ (ppm) -148.20. **LRMS (ESI)** *m/z* 212.90 $[M+H]^+$.[129]

Synthesis of 2-fluorobenzenediazonium tetrafluoroborate 42

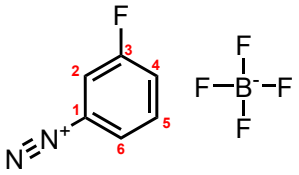


42

The generic method was applied to 2-fluoroaniline (6 mmol, 0.58 mL) to give compound **42** as a grey solid (0.338 g, 27%). **^{1}H NMR** (DMSO, 400 MHz) δ (ppm) 7.64 (ddd, J = 7.9, 5.6, 3.0 Hz, 1H, H_4), 7.81 (t, J = 8.0 Hz, 1H, H_3), 8.01 (q, J = 9.2 Hz, 1H, H_6), 8.32 – 8.42 (m, 1H, H_5). **^{13}C NMR** (DMSO, 100 MHz) δ (ppm) 116.40 (C_1), 118.67 (C_3), 125.10 (C_6), 127.52 (C_5), 133.49 (C_4), 144.71 (C_2). **^{19}F NMR** (DMSO, 376 MHz) δ (ppm) -102.57, -148.20. **LRMS (ESI)** *m/z* 122.83 $[M+H]^+$; **HRMS (ASAP)** calculated for $C_6H_4FN_2$ $[M+H]^+$ 123.0356, found 123.0359.[129]

Synthesis of 3-fluorobenzenediazonium tetrafluoroborate 43

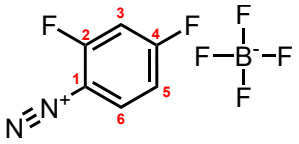
The generic method was applied to 3-fluoroaniline (6 mmol, 0.58 mL) to give compound **43** as a white solid (0.91 g, 72%). **^{1}H NMR** (DMSO, 400 MHz) δ (ppm) 8.06 (td, J = 8.4, 5.1 Hz, 1H, H_4), 8.22 (td, J = 8.8, 2.9 Hz, 1H, H_5), 8.53 – 8.73



43

(m, 2H, $H_{2,6}$). **^{13}C NMR** (DMSO, 100 MHz) δ (ppm) 119.32 (C_1), 119.62 (C_4), 129.01 (C_2), 129.22 (C_6), 129.81 (C_5), 133.58 (C_3). **^{19}F NMR** (DMSO, 376 MHz) δ (ppm) -105.70, -148.21. **LRMS (ESI)** *m/z* 122.95 $[M+H]^+$; **HRMS (ASAP)** calculated for $C_6H_4FN_2$ $[M+H]^+$ 123.0367, found 123.0370.[129]

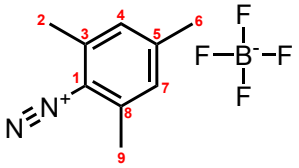
Synthesis of 2,4-difluorobenzenediazonium tetrafluoroborate 44



44

The generic method was applied to 2,4-difluoroaniline (6 mmol, 0.61 mL) to give compound **44** as a brown solid (1.19 g, 87%). **^{1}H NMR** (DMSO, 400 MHz) δ (ppm) 7.70 – 7.83 (m, 1H, H_3), 8.18 – 8.33 (m, 1H, H_5), 8.87 (ddd, J = 9.4, 7.0, 5.2 Hz, 1H, H_6). **^{13}C NMR** (DMSO, 100 MHz) δ (ppm) 105.78 (C_1), 108.48 (C_3), 116.22 (C_5), 136.46 (C_6), 163.63 (C_2), 168.95 (C_4). **^{19}F NMR** (DMSO, 376 MHz) δ (ppm) -79.90, -96.43, -148.34. **LRMS (ESI)** *m/z* 140.96 $[M+H]^+$; **HRMS (ASAP)** calculated for $C_6H_3F_2N_2$ $[M+H]^+$ 141.0264, found 141.0235.[129]

Synthesis of 2,4,6-trimethylbenzenediazonium tetrafluoroborate 45

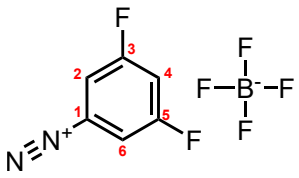


45

The generic method was applied to 2,4,6-trimethylaniline (6 mmol, 0.81 g) to give compound **45** as a purple solid (0.87 g, 62%). **^{1}H NMR** (DMSO, 400 MHz) δ (ppm) 2.31 (s, 3H, H_6), 2.50 (s, 6H, $H_{2,9}$), 7.33 (s, 2H, $H_{4,7}$) . **^{13}C NMR** (DMSO,

100 MHz) δ (ppm) 18.14 (C_6), 22.11 ($C_{2,9}$), 112.03 ($C_{4,7}$), 130.73 ($C_{3,8}$), 143.71 (C_1), 153.41 (C_5). **^{19}F NMR** (DMSO, 376 MHz) δ (ppm) -146.60. **LRMS (ESI)** *m/z* 147.095 $[M+H]^+$; **HRMS (ASAP)** calculated for $C_9H_{11}N_2$ $[M+H]^+$ 147.0561, found 147.0522.[129]

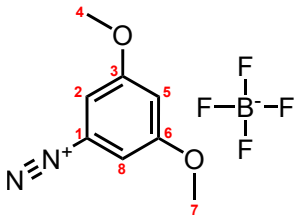
Synthesis of 3,5-difluorobenzenediazonium tetrafluoroborate 46



46

The generic method was applied to 3,5-difluoroaniline (6 mmol, 0.61 mL) to give compound **46** as a white solid (1.19 g, 87%). **^{1}H NMR** (DMSO, 400 MHz) δ (ppm) 8.43 (tt, J = 9.1, 2.4 Hz, 1H, H_4), 8.58 – 8.70 (m, 2H, $H_{2,6}$). **^{13}C NMR** (DMSO, 100 MHz) δ (ppm) 117.08 (C_1), 117.43 (C_3), 118.56 (C_5), 118.82 (C_6), 159.92 (C_2), 162.46 (C_4). **^{19}F NMR** (DMSO, 376 MHz) δ (ppm) -101.76, -148.14. **LRMS (ESI)** *m/z* 140.92 $[M+H]^+$; **HRMS (ASAP)** calculated for $C_6H_3F_2N_2$ $[M+H]^+$ 141.0279, found 141.0276.[129]

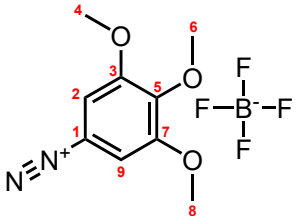
Synthesis of 3,5-dimethoxybenzenediazonium tetrafluoroborate 47



47

The generic method was applied to 3,5-dimethoxyaniline (6 mmol, 0.8 mL) to give compound **47** as a brown solid (1.03 g, 68%). **^{1}H NMR** (DMSO, 400 MHz) δ (ppm) 3.90 (s, 6H, $H_{4,7}$), 7.41 (t, J = 2.4 Hz, 1H, H_5), 7.95 (d, J = 2.4 Hz, 2H, $H_{2,8}$). **^{13}C NMR** (DMSO, 100 MHz) δ (ppm) 56.89 ($C_{4,7}$), 92.76 (C_5), 110.07 ($C_{2,8}$), 113.71 (C_1), 160.76 ($C_{3,6}$). **^{19}F NMR** (DMSO, 376 MHz) δ (ppm) -148.20.[129]

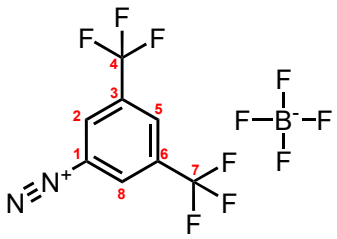
Synthesis of 3,4,5-trimethoxybenzenediazonium tetrafluoroborate 48



48

The generic method was applied to 3,4,5-trimethoxyaniline (6 mmol, 1.10 g) to give compound **48** as a brown solid (1.22 g, 72%). **^{1}H NMR** (DMSO, 400 MHz) δ (ppm) 3.90 (s, 6H, $H_{4,8}$), 4.03 (s, 3H, H_6), 8.18 (s, 2H, $H_{2,9}$). **^{13}C NMR** (DMSO, 100 MHz) δ (ppm) 57.21 ($C_{4,8}$), 61.66 (C_6), 106.46 ($C_{2,9}$), 111.00 (C_1), 149.26 (C_5), 152.53 ($C_{3,7}$). **^{19}F NMR** (DMSO, 376 MHz) δ (ppm) -148.25. **LRMS (ESI)** *m/z* 194.97 $[M+H]^+$; **HRMS (ASAP)** calculated for $C_9H_{11}N_2O_3$ $[M+H]^+$ 195.0823, found 195.0870.[129]

Synthesis of 3,5-bis(trifluoromethyl)benzenediazonium tetrafluoroborate 49

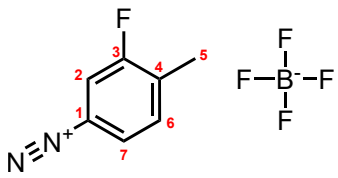


49

The generic method was applied to 3,5-bis(trifluoromethyl)aniline (6 mmol, 0.94 mL) to give compound **49** as a white solid (1.38 g, 70%). **^{1}H NMR** (DMSO, 400 MHz) δ (ppm) 7.12 (s, 1H, H_5), 7.58 (s, 2H, $H_{2,8}$). **^{13}C NMR** (DMSO, 100 MHz) δ (ppm) 113.60 ($C_{2,8}$), 116.69 (C_1), 120.23 (C_5), 121.81 ($C_{4,7}$), 134.31 ($C_{3,6}$). **^{19}F NMR** (DMSO, 376 MHz) δ (ppm) -61.65, -148.19. **LRMS (ESI)** *m/z* 240.88 $[M+H]^+$; **HRMS (ASAP)** calculated for $C_8H_3F_6N_2$ $[M+H]^+$ 241.0210, found 241.0213.[129]

Synthesis of 3-fluoro-4-methylbenzenediazonium tetrafluoroborate 50

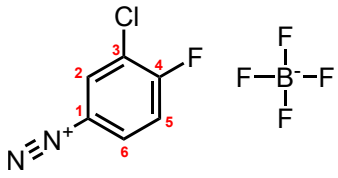
The generic method was applied to 3-fluoro-4-methylaniline (6 mmol, 0.68 mL) to



50

give compound **50** as a white solid (1.17 g, 87%). **^{1}H NMR** (DMSO, 400 MHz) δ (ppm) 2.33 (s, 3H, H_5), 7.94 (ddd, J = 8.4, 7.4, 0.9 Hz, 1H, H_6), 8.45 – 8.53 (m, 1H, H_7), 8.59 (dd, J = 8.4, 2.1 Hz, 1H, H_2). **^{13}C NMR** (DMSO, 100 MHz) δ (ppm) 15.64 (C_5), 113.47 (C_1), 118.65 (C_2), 129.59 (C_4), 134.24 (C_6), 141.94 (C_7), 160.30 (C_3). **^{19}F NMR** (DMSO, 376 MHz) δ (ppm) -109.27, -148.21. **LRMS (ESI)** m/z 136.93 $[M+H]^+$; **HRMS (ASAP)** calculated for $C_7H_6FN_2$ $[M+H]^+$ 137.0519, found 137.0515.[129]

Synthesis of 3-chloro-4-fluorobenzenediazonium tetrafluoroborate 51

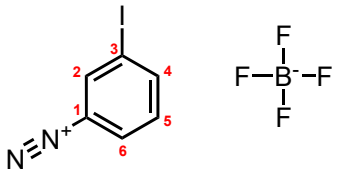


51

The generic method was applied to 3-chloro-4-fluoroaniline (6 mmol, 0.55 mL) to give compound **51** as a white solid (1.26 g, 86%). **^{1}H NMR** (DMSO, 400 MHz) δ (ppm) 7.97 – 8.11 (m, 1H, H_5), 8.73 – 8.86 (m, 1H, H_6), 9.09 (dd, J = 6.2, 2.6 Hz, 1H, H_2). **^{13}C NMR** (DMSO, 100 MHz) δ (ppm) 104.20 (C_1), 120.40 (C_5), 122.88 (C_3), 133.71 (C_6), 135.43 (C_2), 163.05 (C_4). **^{19}F NMR** (DMSO, 376 MHz) δ (ppm) -90.42, -148.22. **LRMS (ESI)** m/z 156.89 $[M+H]^+$; **HRMS (ASAP)** calculated for $C_6H_3{}^{35}ClFN_2$ $[M+H]^+$ 156.9899, found 156.9907.[129]

Synthesis of 3-iodobenzenediazonium tetrafluoroborate 52

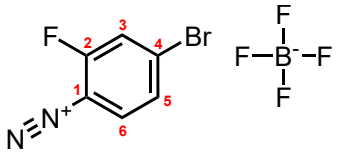
The generic method was applied to 3-iodoaniline (6 mmol, 1.31 g) to give compound **52** as a purple solid (1.53 g, 80%). **^{1}H NMR** (DMSO, 400 MHz) δ (ppm) 7.27 – 7.38 (m, 1H, H_5), 7.72 – 7.76 (m, 1H, H_6), 8.02 – 8.22 (m, 1H, H_4), 8.55 – 8.70



52

(m, 1H, H_2). **^{13}C NMR** (DMSO, 100 MHz) δ (ppm) 95.52 (C_1), 123.10 (C_3), 131.44 (C_2), 131.87 (C_6), 139.48 (C_5), 149.07 (C_4). **^{19}F NMR** (DMSO, 376 MHz) δ (ppm) -148.16. **LRMS (ESI)** *m/z* 230.88 $[M+H]^+$; **HRMS (ASAP)** calculated for $C_6H_4IN_2$ $[M+H]^+$ 230.9434, found 230.9419.[129]

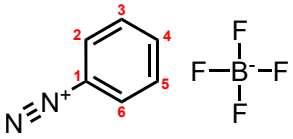
Synthesis of 4-bromo-2-fluorobenzenediazonium tetrafluoroborate 53



53

The generic method was applied to 4-bromo-2-fluoroaniline (6 mmol, 0.71 mL) to give compound **53** as a brown solid (0.825 g, 48%). **^{1}H NMR** (DMSO, 400 MHz) 7.64 – 7.70 (m, 1H, H_5), 7.75 (ddd, J = 8.5, 2.1, 1.1 Hz, 1H, H_6), 8.18 (d, J = 8.8 Hz, 1H, H_3). **^{13}C NMR** (DMSO, 100 MHz) δ (ppm) 105.02 (C_1), 121.40 (C_3), 122.49 (C_4), 131.00 (C_5), 133.83 (C_6), 158.24 (C_2). **^{19}F NMR** (DMSO, 376 MHz) δ (ppm) -100.38, -148.26. **LRMS (ESI)** *m/z* 200.78 $[M+H]^+$; **HRMS (ASAP)** calculated for $C_6H_3BrFN_2$ $[M+H]^+$ 200.9473, found 200.9475.[129]

Synthesis of benzenediazonium tetrafluoroborate 54

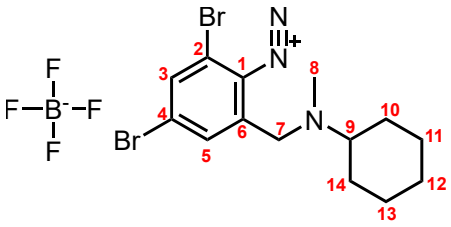


54

The generic method was applied to aniline (6 mmol, 0.55 mL) to give compound **54** as a brown solid (0.98 g, 85%). **^{1}H NMR** (DMSO, 400 MHz) 7.97 (s, 2H, $H_{3,5}$), 8.24 (s, 1H, H_4), 8.66 (s, 2H, $H_{2,6}$). **^{13}C NMR** (DMSO, 100 MHz) δ (ppm) 115.99

(C_1), 131.31 ($C_{3,5}$), 132.72 ($C_{2,6}$), 140.93 (C_4). **^{19}F NMR** (DMSO, 376 MHz) δ (ppm) -148.07. **LRMS (ESI)** m/z 104.93 $[M+H]^+$; **HRMS (ASAP)** calculated for $[M+H]^+$ 105.0447, found 105.0453.[129]

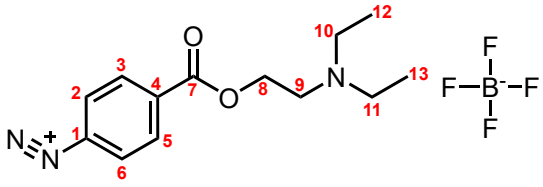
Synthesis of bromhexinediazonium tetrafluoroborate 55



55

The generic method was applied to bromhexine hydrochloride (2.5 mmol, 1.03 g) to give compound **55** as a white solid (0.94 g, 79%). **^{1}H NMR** (DMSO, 400 MHz) 1.29 (p, J = 13.4 Hz, 2H, H_{12}), 1.41 – 1.68 (m, 4H, $H_{11,13}$), 1.70 – 1.91 (m, 4H, $H_{10,14}$), 2.55 (s, 3H, H_8), 2.98 (s, 2H, H_7), 7.50 (d, J = 2.3 Hz, 1H, H_3), 7.72 (dd, J = 18.5, 1.9 Hz, 1H, H_5). **^{13}C NMR** (DMSO, 100 MHz) δ (ppm) 24.52 (C_{11}), 24.66 (C_{13}), 25.25 (C_{12}), 28.77 (C_{10}), 32.36 (C_{14}), 34.79 (C_8), 49.92 (C_7), 52.07 (C_9), 116.49 (C_2), 116.75 (C_4), 131.98 (C_5), 133.95 (C_6), 135.04 (C_3), 135.47 (C_1). **^{19}F NMR** (DMSO, 376 MHz) δ (ppm) -148.24. **LRMS (ESI)** m/z 387.80 $[M+H]^+$; **HRMS (ASAP)** calculated for $C_{14}H_{18}Br_2N_3$ $[M+H]^+$ 387.9868, found 387.9867.[129]

Synthesis of procainediazonium tetrafluoroborate 56

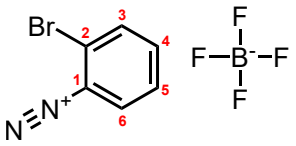


56

The generic method was applied to procaine (3 mmol, 0.71 g) to give compound **56** as a white solid (0.76 g, 76%). **^{1}H NMR** (DMSO, 400 MHz) 1.24 (t, J = 7.2 Hz, 6H, $H_{12,13}$), 3.27 (qdd, J = 7.1, 4.4, 3.1 Hz, 4H, $H_{10,11}$), 3.58 (q, J = 5.0 Hz, 2H, H_9), 4.62 – 4.69 (m, 2H, H_8), 8.43 – 8.51 (m, 2H, $H_{3,5}$), 8.79 – 8.87 (m, 2H, $H_{2,6}$). **^{13}C NMR** (DMSO, 100 MHz) δ (ppm) 8.58 ($C_{12,13}$), 47.14 ($C_{10,11}$), 49.54 (C_9), 60.46

(C_8), 120.51 (C_1), 131.37 (C_4), 133.11 ($C_{3,5}$), 138.84 ($C_{2,6}$), 162.93 (C_7). **^{19}F NMR** (DMSO, 376 MHz) δ (ppm) -148.27. **LRMS (ESI)** *m/z* 248.17 $[M+H]^+$; **HRMS (ASAP)** calculated for $C_{13}H_{18}N_3O_2$ $[M+H]^+$ 248.1419, found 248.1399.[129]

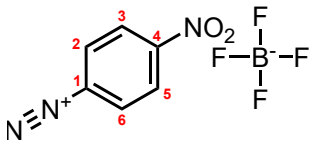
Synthesis of 2-bromobenzenediazonium tetrafluoroborate 57



57

The generic method was applied to 2-bromoaniline (4 mmol, 0.69 g) to give compound **57** as a white solid (0.86 g, 80%). **^{1}H NMR** (DMSO, 400 MHz) 7.99 (ddd, J = 8.6, 7.6, 1.1 Hz, 1H, H_5), 8.17 (ddd, J = 8.2, 7.6, 1.6 Hz, 1H, H_4), 8.31 (dd, J = 8.3, 1.1 Hz, 1H, H_3), 8.84 (dd, J = 8.3, 1.6 Hz, 1H, H_6). **^{13}C NMR** (DMSO, 100 MHz) δ (ppm) 118.72 (C_1), 124.44 (C_2), 130.42 (C_6), 135.07 (C_5), 135.30 (C_4), 141.87 (C_3). **^{19}F NMR** (DMSO, 376 MHz) δ (ppm) -148.22. **LRMS (ESI)** *m/z* 184.78 $[M+H]^+$; **HRMS (ASAP)** calculated for $C_6H_4BrN_2$ $[M+H]^+$ 184.9569, found 184.9558.[129]

Synthesis of 4-nitrobenzenediazonium tetrafluoroborate 58



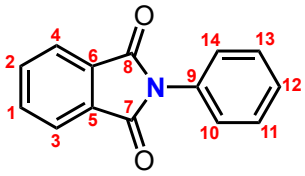
58

The generic method was applied to 4-nitroaniline (4 mmol, 0.55 g) to give compound **58** as a yellow solid (0.75 g, 79%). **^{1}H NMR** (DMSO, 400 MHz) 8.72 (td, J = 9.3, 5.1 Hz, 2H, $H_{2,6}$), 8.93 (td, J = 9.0, 4.9 Hz, 2H, $H_{3,5}$). **^{13}C NMR** (DMSO, 100 MHz) δ (ppm) 121.88 (C_1), 126.03 ($C_{3,5}$), 134.50 ($C_{2,6}$), 153.21 (C_4). **^{19}F NMR** (DMSO, 376 MHz) δ (ppm) -148.20. **LRMS (ESI)** *m/z* 149.78 $[M+H]^+$; **HRMS (ASAP)** calculated for $C_6H_4N_3O_2$ $[M+H]^+$ 150.0313, found 150.0304.[129]

General Procedure for the Synthesis of ^{15}N coupled phthalimides

To a stirred solution of ^{15}N potassium phthalimide (1 mmol, 0.186 g) and CuCl (1 mmol, 0.099 g) in DMSO (5 mL), under compressed air, was added the diazonium tetrafluoroborate in DMSO (3 mL) slowly dropwise over 10-15 min. The resulting mixture was stirred under ice cooling and compressed air for 12 h. The mixture was filtered to remove the catalyst and H_2O was added (15 mL). The aqueous phase was washed with EtOAc (3×15 mL). The organic phase was then washed with H_2O (3×10 mL). The solvent was evaporated *in vacuo* to yield the crude product. Purification by silica gel chromatography using petroleum ether:ethyl acetate (4:1), followed by a second purification with petroleum ether:ethyl acetate (8:1) yielded the compound as a pure ^{15}N coupled phthalimide derivative.

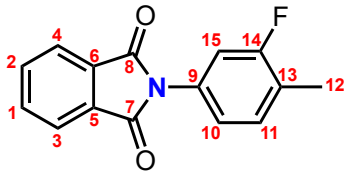
Synthesis of ^{15}N 2-phenylisoindole-1,3-dione 59



^{15}N **59**

The generic method was applied to benzenediazonium tetrafluoroborate **54** (3 mmol, 0.56 g) to give compound ^{15}N **59** as a pure white solid (0.126 g, 57%). **^{1}H NMR** ($CDCl_3$, 400 MHz) δ (ppm) 7.37 – 7.57 (m, 5H, $H_{10,11,12,13,14}$), 7.64 – 7.90 (m, 2H, $H_{1,2}$), 7.92 – 8.04 (m, 2H, $H_{3,4}$). **^{13}C NMR** ($CDCl_3$, 100 MHz) 123.90 ($C_{3,4}$), 126.72 ($C_{10,14}$), 128.25 ($C_{1,2}$), 129.27 (C_{12}), 131.73 ($C_{11,13}$), 131.95 ($C_{5,6}$), 134.54 (C_9), 167.49 ($C_{7,8}$). **LRMS (ESI)** *m/z* 225.03 $[M+H]^+$; **HRMS (ASAP)** calculated for $C_{14}H_9{}^{15}NO_2$ $[M+H]^+$ 225.0686, found 225.0682. **m.p.** = 205-206°C.[151]

Synthesis of ^{15}N 2-(3-fluoro-4-methylphenyl)isoindoline-1,3-dione 60

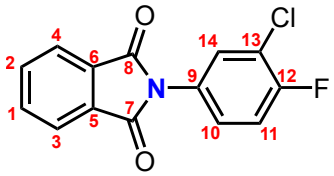


^{15}N **60**

The generic method was applied to 3-fluoro,4-methylbenzenediazonium tetrafluo-

roborate **50** (3 mmol, 0.67 g) to give compound ^{15}N **60** as a pure yellow solid (0.036 g, 14%). **^{1}H NMR** ($CDCl_3$, 400 MHz) δ (ppm) 2.33 (d, J = 2.1 Hz, 3H, H_{12}), 7.16 (dt, J = 9.4, 1.7 Hz, 2H, $H_{10,11}$), 7.25 – 7.36 (m, 1H, H_{15}), 7.75 - 7.86 (m, 2H, $H_{1,2}$), 7.91 – 8.00 (m, 2H, $H_{3,4}$). **^{13}C NMR** ($CDCl_3$, 100 MHz) 14.57 (C_{12}), 113.87 (C_{15}), 121.99 (C_{10}), 123.97 ($C_{3,4}$), 125.09 (C_{13}), 125.26 ($C_{1,2}$), 130.35 (C_{11}), 131.71 ($C_{5,6}$), 134.65 (C_9), 159.82 (C_{14}), 167.24 ($C_{7,8}$). **^{19}F NMR** (DMSO, 376 MHz) δ (ppm) -114.96. **LRMS (ESI)** m/z 257.10 $[M+H]^+$; **HRMS (ASAP)** calculated for $C_{15}H_{10}F^{15}NO_2$ $[M+H]^+$ 257.0739, found 257.0743. **m.p.** = 198-199°C.[151]

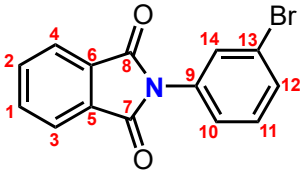
Synthesis of ^{15}N 2-(3-chloro-4-fluorophenyl)isoindoline-1,3-dione 61



^{15}N **61**

The generic method was applied to 3-chloro,4-fluorobenzenediazonium tetrafluoroborate **51** (3 mmol, 0.73 g) to yield compound ^{15}N **61** as a pure yellow solid (0.028 g, 10%). **^{1}H NMR** ($CDCl_3$, 400 MHz) δ (ppm) 7.23 – 7.52 (m, 2H, $H_{10,11}$), 7.53 – 7.63 (m, 1H, H_{14}), 7.79 - 7.85 (m, 2H, $H_{1,2}$), 7.92 – 8.02 (m, 2H, $H_{3,4}$). **^{13}C NMR** ($CDCl_3$, 100 MHz) 116.90 (C_{14}), 117.15 (C_{11}), 123.46 (C_{13}), 123.76 (C_{10}), 124.12 ($C_{3,4}$), 131.60 ($C_{5,6}$), 134.42 (C_9), 134.86 ($C_{1,2}$), 157.58 (C_{12}), 166.89 ($C_{7,8}$). **^{19}F NMR** (DMSO, 376 MHz) δ (ppm) -115.27. **LRMS (ESI)** m/z 276.90 $[M+H]^+$; **HRMS (ASAP)** calculated for $C_{14}H_7ClF^{15}NO_2$ $[M+H]^+$ 277.0194, found 277.0196. **m.p.** = 211-212°C.[151]

Synthesis of ^{15}N 2-(3-bromophenyl)isoindoline-1,3-dione 62

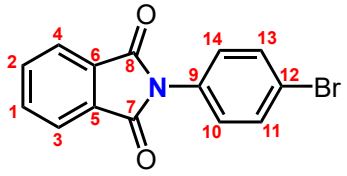


^{15}N **62**

The generic method was applied to 3-bromobenzenediazonium tetrafluoroborate **36**

(3 mmol, 0.81 g) to give compound ^{15}N **62** as a pure red solid (0.030 g, 9.9%). **^{1}H NMR** ($CDCl_3$, 400 MHz) δ (ppm) 7.37 – 7.43 (m, 2H, $H_{12,14}$), 7.54 (dt, J = 7.7, 1.6 Hz, 1H, H_{11}), 7.65 (q, J = 1.8 Hz, 1H, H_{10}), 7.82 (dd, J = 5.5, 3.0 Hz, 2H, $H_{1,2}$), 7.94 – 8.00 (m, 2H, $H_{3,4}$). **^{13}C NMR** ($CDCl_3$, 100 MHz) δ (ppm) 122.53 (C_{12}), 124.07 ($C_{10,14}$), 125.24 ($C_{3,4}$), 129.62 ($C_{1,2}$), 130.43 (C_9), 131.29 ($C_{5,6}$), 134.77 ($C_{11,13}$), 166.37 ($C_{7,8}$). **LRMS (ESI)** m/z 304.84 $[M+H]^+$; **HRMS (ASAP)** calculated for $C_{14}H_8Br^{15}NO_2$ $[M+H]^+$ 304.9829, found 304.9771. **m.p.** = 204-205°C.[151]

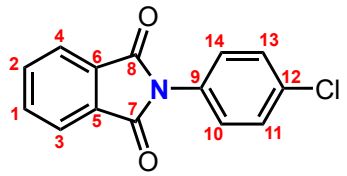
Synthesis of ^{15}N 2-(4-bromophenyl)isoindoline-1,3-dione 63



^{15}N **63**

The generic method was applied to 4-bromobenzenediazonium tetrafluoroborate **37** (3 mmol, 0.81 g) to give compound ^{15}N **63** as a pure yellow solid (0.030 g, 9.9%). **^{1}H NMR** ($CDCl_3$, 400 MHz) δ (ppm) 7.29 – 7.47 (m, 2H, $H_{11,13}$), 7.55 – 7.68 (m, 2H, $H_{10,14}$), 7.78 – 7.85 (m, 2H, $H_{1,2}$), 7.92 – 7.99 (m, 2H, $H_{3,4}$). **^{13}C NMR** ($CDCl_3$, 100 MHz) δ (ppm) 121.97 (C_{12}), 124.01 ($C_{10,14}$), 128.08 ($C_{3,4}$), 130.94 ($C_{1,2}$), 131.70 (C_9), 132.42 ($C_{5,6}$), 134.72 ($C_{11,13}$), 166.99 ($C_{7,8}$). **LRMS (ESI)** m/z 304.84 $[M+H]^+$; **HRMS (ASAP)** calculated for $C_{14}H_8Br^{15}NO_2$ $[M+H]^+$ 304.9829, found 304.9756. **m.p.** = 204-205°C.[151]

Synthesis of ^{15}N 2-(4-chlorophenyl)isoindoline-1,3-dione 64

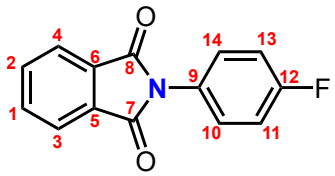


^{15}N **64**

The generic method was applied to 4-chlorobenzenediazonium tetrafluoroborate **38** (3 mmol, 0.68 g) to give compound ^{15}N **64** as a pure orange solid (0.049 g, 19%). **^{1}H NMR** ($CDCl_3$, 400 MHz) δ (ppm) 7.35 – 7.45 (m, 4H, $H_{10,11,13,14}$), 7.73 – 7.87

(m, 2H, $H_{1,2}$), 7.96 (dd, J = 5.5, 3.0 Hz, 2H, $H_{3,4}$). **^{13}C NMR** ($CDCl_3$, 100 MHz) δ (ppm) 124.00 (C_{12}), 127.81 ($C_{10,14}$), 129.46 ($C_{3,4}$), 130.41 ($C_{1,2}$), 131.79 (C_9), 133.95 ($C_{5,6}$), 134.71 ($C_{11,13}$), 167.06 ($C_{7,8}$). **LRMS (ESI)** m/z 258.97 $[M+H]^+$; **HRMS (ASAP)** calculated for $C_{14}H_8Cl^{15}NO_2$ $[M+H]^+$ 259.0322, found 259.0280. **m.p.** = 192-193°C.[151]

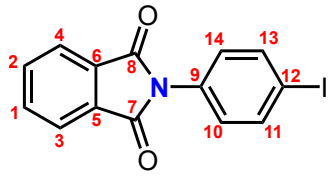
Synthesis of ^{15}N 2-(4-fluorophenyl)isoindoline-1,3-dione 65



^{15}N **65**

The generic method was applied to 4-fluorobenzenediazonium tetrafluoroborate **39** (3 mmol, 0.63 g) to give compound ^{15}N **65** as a pure yellow solid (0.024 g, 10%). **^{1}H NMR** ($CDCl_3$, 400 MHz) δ (ppm) 7.10 – 7.25 (m, 2H, $H_{11,13}$), 7.34 – 7.49 (m, 2H, $H_{10,14}$), 7.73 – 7.87 (m, 2H, $H_{1,2}$), 7.92 – 8.02 (m, 2H, $H_{3,4}$). **^{13}C NMR** ($CDCl_3$, 100 MHz) δ (ppm) 116.17 ($C_{11,13}$), 116.41 ($C_{10,14}$), 123.97 ($C_{3,4}$), 128.57 ($C_{1,2}$), 131.77 (C_9), 134.66 ($C_{5,6}$), 160.87 (C_{12}), 167.41 ($C_{7,8}$). **^{19}F NMR** ($CDCl_3$, 376 MHz) δ (ppm) -113.02. **LRMS (ESI)** m/z 242.97 $[M+H]^+$; **HRMS (ASAP)** calculated for $C_{14}H_8F^{15}NO_2$ $[M+H]^+$ 243.0650, found 243.0577. **m.p.** = 186-187°C.[151]

Synthesis of ^{15}N 2-(4-iodophenyl)isoindoline-1,3-dione 66

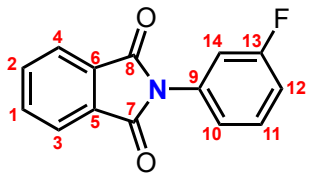


^{15}N **66**

The generic method was applied to 4-iodobenzenediazonium tetrafluoroborate **40** (3 mmol, 0.95 g) to give compound ^{15}N **66** as a pure yellow solid (0.030 g, 8.6%). **^{1}H NMR** ($CDCl_3$, 400 MHz) δ (ppm) 7.35 – 7.55 (m, 2H, $H_{11,13}$), 7.68 – 7.87 (m, 4H, $H_{1,2,10,14}$), 7.91 - 8.00 (m, 2H, $H_{3,4}$). **^{13}C NMR** ($CDCl_3$, 100 MHz) δ (ppm) 93.43 (C_{12}), 124.02 ($C_{10,14}$), 128.27 ($C_{3,4}$), 131.71 ($C_{1,2}$), 131.79 (C_9), 134.73

($C_{5,6}$), 138.41 ($C_{11,13}$), 167.08 ($C_{7,8}$). **LRMS (ESI)** *m/z* 351.71 $[M+H]^+$; **HRMS (ASAP)** calculated for $C_{14}H_8I^{15}NO_2$ $[M+H]^+$ 351.9737, found 351.9705. **m.p.** = 209-210°C.[151]

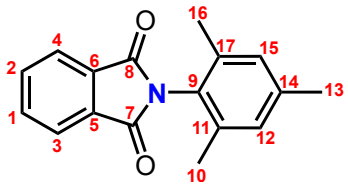
Synthesis of ^{15}N 2-(3-fluorophenyl)isoindoline-1,3-dione 67



^{15}N **67**

The generic method was applied to 3-fluorobenzenediazonium tetrafluoroborate **43** (3 mmol, 0.63 g) to give compound ^{15}N **67** as a pure red solid (0.036 g, 15%). **^{1}H NMR** ($CDCl_3$, 400 MHz) δ (ppm) 7.12 (tdd, J = 8.4, 2.5, 0.9 Hz, 1H, H_{12}), 7.19 – 7.45 (m, 2H, $H_{11,10}$), 7.48 (td, J = 8.2, 6.2 Hz, 1H, H_{14}), 7.75 – 7.90 (m, 2H, $H_{1,2}$), 7.93 – 8.11 (m, 2H, $H_{3,4}$). **^{13}C NMR** ($CDCl_3$, 100 MHz) δ (ppm) 115.27 (C_{12}), 122.14 (C_{14}), 122.17 ($C_{3,4}$), 124.06 (C_{10}), 130.31 ($C_{1,2}$), 131.75 (C_{11}), 133.21 ($C_{5,6}$), 134.76 (C_9), 166.94 (C_{13}), 171.31 ($C_{7,8}$). **^{19}F NMR** ($CDCl_3$, 376 MHz) -111.15. **LRMS (ESI)** *m/z* 242.93 $[M+H]^+$; **HRMS (ASAP)** calculated for $C_{14}H_8F^{15}NO_2$ $[M+H]^+$ 243.0650, found 243.05795. **m.p.** = 186-187°C.[151]

Synthesis of ^{15}N 2-mesitylisoindoline-1,3-dione 68

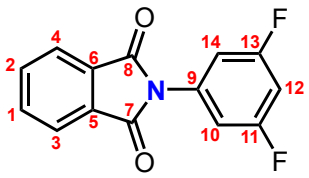


^{15}N **68**

The generic method was applied to 2,4,6-trimethylbenzenediazonium tetrafluoroborate **45** (3 mmol, 0.70 g) to give compound ^{15}N **68** as a pure red solid (0.026 g, 9.8%). **^{1}H NMR** ($CDCl_3$, 400 MHz) δ (ppm) 2.12 (s, 6H, $H_{10,16}$), 2.33 (s, 3H, H_{13}), 6.91 – 7.03 (m, 2H, $H_{12,15}$), 7.80 (dd, J = 5.5, 3.1 Hz, 2H, $H_{1,2}$), 7.91 – 8.01 (m, 2H, $H_{3,4}$). **^{13}C NMR** ($CDCl_3$, 100 MHz) δ (ppm) 18.12 ($C_{10,16}$), 21.28 (C_{13}), 123.89 ($C_{3,4}$), 127.24 ($C_{12,15}$), 129.45 ($C_{5,6}$), 132.23 ($C_{1,2}$), 134.41 ($C_{11,17}$), 136.60

(C_9), 139.54 (C_{14}), 167.63 ($C_{7,8}$). **LRMS (ESI)** *m/z* 267.21 $[M+H]^+$; **HRMS (ASAP)** calculated for $C_{17}H_{15}{}^{15}NO_2$ $[M+H]^+$ 267.1214, found 267.1160. **m.p.** = 177-178°C.[151]

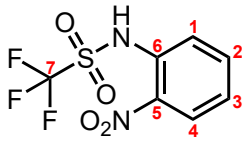
Synthesis of ^{15}N 2-(3,5-difluorophenyl)isoindoline-1,3-dione 69



^{15}N **69**

The generic method was applied to 3,5-difluorobenzenediazonium tetrafluoroborate **46** (3 mmol, 0.68 g) to give compound ^{15}N **69** as a pure yellow solid (0.027 g, 10%). **^{1}H NMR** ($CDCl_3$, 400 MHz) δ (ppm) 6.86 (tt, J = 8.9, 2.3 Hz, 1H, H_{12}), 7.06 – 7.17 (m, 2H, $H_{10,14}$), 7.80 – 7.86 (m, 2H, $H_{1,2}$), 7.95 – 8.02 (m, 2H, $H_{3,4}$). **^{13}C NMR** ($CDCl_3$, 100 MHz) δ (ppm) 103.61 (C_{12}), 109.59 ($C_{10,14}$), 123.26 (C_9), 124.22 ($C_{3,4}$), 131.50 ($C_{5,6}$), 134.98 ($C_{1,2}$), 161.78 ($C_{11,13}$), 166.78 ($C_{7,8}$). **^{19}F NMR** ($CDCl_3$, 376 MHz) δ (ppm) -108.29. **LRMS (ESI)** *m/z* 260.91 $[M+H]^+$; **HRMS (ASAP)** calculated for $C_{14}H_7F_2{}^{15}NO_2$ $[M+H]^+$ 261.0498, found 261.0496. **m.p.** = 202-203°C.[151]

Synthesis of of 1,1,1-trifluoro-*N*-(2-nitrophenyl)methanesulfonamide 70

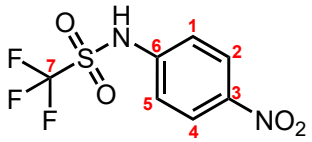


70

Trifluoromethanesulfonic anhydride (0.25 mL, 1.5 mmol) in DCM (10 mL) was added dropwise to a stirred solution of 2-nitroaniline (0.21 g, 1.5 mmol) and Et_3N (0.32 mL, 2.25 mmol) in anhydrous DCM (10 mL) over 1.5 h at 0°C under N_2. The solution was then warmed to RT and stirred for 18 h. The solution was then diluted with DCM (20 mL), washed with sat. $NaHCO_3$ (15 mL) and dried over $MgSO_4$. After filtration, the solvent was removed *in vacuo* to yield a crude red oil. Purification by silica gel chromatography using EtOAc yielded compound **70** as a pure orange oil

(0.29 g, 72%). **^{1}H NMR** ($CDCl_3$, 400 MHz) δ (ppm) 7.29 (ddd, J = 8.5, 7.3, 1.3 Hz, 1H, H_3), 7.64 (ddd, J = 8.4, 7.3, 1.6 Hz, 1H, H_2), 7.82 (dd, J = 8.4, 1.3 Hz, 1H, H_4), 8.13 (dd, J = 8.4, 1.6 Hz, 1H, H_1). **^{13}C NMR** ($CDCl_3$, 100 MHz) δ (ppm) 118.17 (C_3), 122.68 (C_7), 124.81 (C_1), 125.76 (C_4), 133.04 (C_2), 135.33 (C_5), 139.48 (C_6).[76]

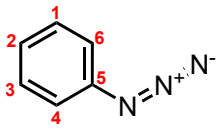
Synthesis of of 1,1,1-trifluoro-*N*-(4-nitrophenyl)methanesulfonamide 71



71

Trifluoromethanesulfonic anhydride (0.25 mL, 1.5 mmol) in DCM (10 mL) was added dropwise to a stirred solution of 4-nitroaniline (0.21 g, 1.5 mmol) and Et_3N (0.32 mL, 2.25 mmol) in anhydrous DCM (10 mL) over 1.5 h at 0°C under N_2. The solution was then warmed to RT and stirred for 18 h. The solution was diluted with DCM (20 mL), washed with sat. $NaHCO_3$ (15 mL) and dried over $MgSO_4$. After filtration, the solvent was removed *in vacuo* to yield a crude red oil. Purification by silica gel chromatography using EtOAc yielded compound **71** as a pure orange oil (0.32 g, 79%). **^{1}H NMR** ($CDCl_3$, 400 MHz) δ (ppm) 7.39 – 7.48 (m, 2H, $H_{2,4}$), 8.17 – 8.31 (m, 2H, $H_{1,5}$). **^{13}C NMR** ($CDCl_3$, 100 MHz) δ (ppm) 113.25 (C_7), 121.01 ($C_{1,5}$), 125.26 ($C_{2,4}$), 126.29 (C_6), 140.21 (C_3).[76]

Synthesis of Phenyl Azide 72



72

Sodium nitrite (0.51 g, 7.5 mmol) in water (3 mL) was added to a stirred solution of aniline (0.46 mL, 5 mmol) in 2 M HCl (aq., 25 mL) at 0°C. After 15 min of stirring at 0°C, sodium azide (0.32 g, 7.5 mmol) in water (3 mL) was added to the solution and left to stir for 1 h. The solution was then extracted with EtOAc (3×30 mL). The combined organic layers were dried over $MgSO_4$, filtered, and solvent removed

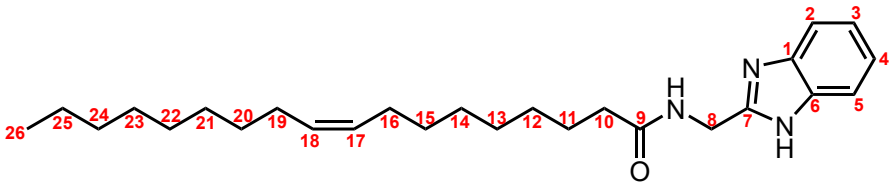
in vacuo to yield compound **72** as an orange oil (0.185 g, 31%). **^{1}H NMR** ($CDCl_3$, 400 MHz) δ (ppm) 7.00 – 7.10 (m, 2H, $H_{4,6}$). 7.11 – 7.23 (m, 1H, H_2), 7.32 – 7.44 (m, 2H, $H_{1,3}$). **^{13}C NMR** ($CDCl_3$, 100 MHz) δ (ppm) 118.77 ($C_{4,6}$), 124.61 (C_2), 129.49 ($C_{1,3}$), 139.75 (C_5).[195]

Synthesis of Aniline 73

73

Sodium borohydride (0.147 g, 3.88 mmol) was added to a solution of phenyl azide **72** in nickel(II) chloride hexahydrate (0.00184 g, 0.00775 mmol) and anhydrous MeOH (5 mL). The solution was stirred for 3 h whilst protected from light. The solvent was removed *in vacuo*, re-dissolved in water (10 mL) and extracted with Et_2O. The organic layer was dried over $MgSO_4$, filtered and solvent removed *in vacuo* to yield compound **73** as an orange oil (0.004 g, 2.8%). **^{1}H NMR** ($CDCl_3$, 400 MHz) δ (ppm) 6.66 – 6.72 (m, 2H, $H_{4,6}$), 6.80 – 6.86 (m, 1H, H_2), 7.12 – 7.19 (m, 2H, $H_{1,3}$). **^{13}C NMR** ($CDCl_3$, 100 MHz) δ (ppm) 115.59 ($C_{4,6}$), 119.06 (C_2), 129.75 ($C_{1,3}$), 146.2 (C_5).[165]

Synthesis of Acylated 2-aminomethylbenzimidazole 74



74

Oleoyl chloride (0.37 mL, 1.12 mmol) was added to a stirred solution of 2-aminomethyl benzimidazole (0.15 g, 1.02 mmol) and pyridine (90 μL, 1.12 mmol) in anhydrous DCM (10 mL) under Ar gas at RT for 12 h. The organic phase was washed with saturated NH_4Cl (3×5 mL), dried over $MgSO_4$, and after filtration, the solvent was removed *in vacuo* to yield compound **74** as a light pink oil. **^{1}H NMR** (DMSO, 400 MHz) δ (ppm) 0.83 (ddd, J = 7.1, 4.8, 1.9 Hz, 3H, H_{26}), 1.09 - 1.32 (m, 20H,

H_{alkyl}), 1.41 - 1.53 (m, 2H, H_{11}), 1.85 - 2.05 (m, 4H, $H_{16,19}$), 2.24 (t, J = 7.4 Hz, 2H, H_{10}), 3.55 (s, 2H, H_8), 5.24 - 5.36 (m, 2H, $H_{17,18}$), 7.11 (s, 1H, H_3), 7.46 (d, J = 5.8 Hz, 1H, H_4), 7.79 (t, J = 5.8 Hz, 1H, H_2), 8.05 (dt, J = 11.6, 5.8 Hz, 1H, H_5). **^{13}C NMR** (DMSO, 100 MHz) δ (ppm) 13.81 (C_{26}), 22.13 (C_{alkyl}), 24.43 (C_{alkyl}), 26.61 (C_{alkyl}), 28.53 (C_{alkyl}), 28.68 (C_{alkyl}), 28.77 (C_{alkyl}), 28.93 (C_{alkyl}), 29.17 (C_{alkyl}), 31.35 (C_{alkyl}), 33.26 (C_{10}), 35.07 (C_8), 127.68 ($C_{2,5}$), 129.46 ($C_{3,4}$), 129.49 ($C_{1,6}$), 129.55 ($C_{17,18}$), 155.38 (C_7), 173.07 (C_9).[56]

Synthesis of Acylated 4-amino-*N*-phenylbutanamide 75

75

Oleoyl chloride (0.37 mL, 1.12 mmol) was added to a stirred solution of 4-amino-*N*-phenylbutanamide (0.18 g, 1.02 mmol) and pyridine (90 μL, 1.12 mmol) in anhydrous DCM (10 mL) and stirred under Ar gas at RT for 12 h. The organic phase was washed with saturated NH_4Cl (3×5 mL), dried over $MgSO_4$, and after filtration, the solvent was removed *in vacuo* to yield compound **75** as an orange oil. **^{1}H NMR** (DMSO, 400 MHz) δ (ppm) 0.79 (t, J = 6.6 Hz, 3H, H_{28}), 1.15 - 1.27 (m, 20H, H_{alkyl}), 1.43 (p, J = 7.1 Hz, 2H, H_9), 1.90 (d, J = 6.4 Hz, 2H, H_8), 1.92 (s, 2H, H_{12}), 2.10 (t, J = 7.4 Hz, 2H, H_{10}), 5.18 - 5.30 (m, 2H, $H_{19,20}$), 6.84 - 7.01 (m, 1H, H_4), 7.13 - 7.25 (m, 1H, H_3), 7.56 (dd, J = 7.9, 3.9 Hz, 1H, H_5), 7.77 (t, J = 5.7 Hz, 1H, H_2), 8.08 (t, J = 7.6 Hz, 1H, H_6). **^{13}C NMR** (DMSO, 100 MHz) δ (ppm) 13.81 (C_{28}), 22.22 (C_{27}), 24.59 (C_{13}), 25.42 (C_9), 26.70 (C_{alkyl}), 28.72 (C_{alkyl}), 28.80 (C_{alkyl}), 28.84 (C_{alkyl}), 28.88 (C_{alkyl}), 29.05 (C_{alkyl}), 29.27 (C_{alkyl}), 31.46 (C_8), 33.72 (C_{12}), 38.13 (C_{10}), 119.04 ($C_{2,6}$), 128.43 (C_4), 128.49 ($C_{3,5}$), 129.45 ($C_{19,20}$), 139.46 (C_1), 170.89 (C_7), 172.18 (C_{11}).[56]

Chapter 7

Conclusion and Outlook

This thesis has explored and uncovered, both experimentally and computationally, the factors which contribute towards reactivity of molecules with membrane lipids. Whilst lipophilicity and the solubility of molecules within the membrane has been historically well studied, recent literature and the findings in this thesis have shown that orientation and depth of partitioning of molecules in the membrane dictate their reactivity.

Chapter 2 gave a literature overview of the historical synthetic routes to CAD molecules, followed by the proposed synthetic targets in this thesis. Various synthetic pathways for the incorporation of an ^{15}N isotope into 2-aminomethyl benzimidazole, propranolol, and 4-amino-*N*-phenylbutanamide were explored and tested to find the optimal route. The synthesis of ^{15}N propranolol, after much testing of different synthetic routes, was rather simple. *In situ* formation of the epoxide intermediate was formed and only a simple 2-step synthesis was needed to form ^{15}N propranolol. Having explored different protecting groups in the synthesis of ^{15}N 2-aminomethylbenzimidazole, attachment of BOC to the labelled glycine molecule, and subsequent synthetic steps, ensured successful deprotection in the final step. This same protecting group chemistry was then applied in the synthesis of ^{15}N 4-amino-*N*-phenylbutanamide. Having explored the economic cost of purchasing ^{15}N

precursor molecules for the synthesis of ^{15}N molecules, it became increasingly apparent that a better method was needed for this ^{15}N incorporation. The second part of this chapter gave an overview of the C-N bond formation in the literature, and then gave the proposed ^{15}N methodology development. The copper catalysed cross-coupling reaction of the substituted diazonium tetrafluoroborate with ^{15}N potassium phthalimide was successful. However, due to multiple purification steps, the yield suffered quite badly. Thus unfortunately, it was impractical to try deprotection of these labelled compounds as this would have resulted in insufficient quantities to analyse. Future work needs to be conducted into optimising the reaction conditions and finding better purification conditions to improve the yield. Once the yield can be improved and the final synthetic step is confirmed, this methodology can then be applied to a wide substrate scope of any drug molecule with aniline functionality.

Chapter 3 firstly gave a literature review of solid-state NMR in aiding the quantitative understanding of drug lipid reactivity. A viable methodology was proposed in this chapter to analyse close contacts/distances between protons on the drug molecule and protons in the lipid. 1D and 2D NOESY experiments confirmed reactivity as the distance between the aromatic protons on the drug and aliphatic protons on the lipid chain decreased as a function of time. To ensure a valid comparison of chemical shift changes, both acylated 2-aminomethylbenzimidazole and acylated 4-amino-*N*-phenylbutanamide were used as controls to detect the amine protons bonded to the carbonyl carbon of the palmitic chain. For the 1D NMR experiments, the data is presented with the drug only, lipid only, then the drug+liposome at low concentration followed by drug+liposome at higher concentration. These initial 1D experiments provided confidence that the proton chemical shifts confirmed close contact, and the emergence of multiple cross-peaks between the drug and lipid in 2D NOESY experiments, showed precisely the chemical species interacting with each other. Additional ^{1}H-^{15}N HMBC measurements showed a proton-nitrogen chemical shift correlating with formation of an amide for 2-aminomethylbenzimidazole and 4-amino-*N*-phenylbutanamide. Finally, distance measurements were provided by analysis of the build-up curves in the linear regime. In all cases the r value, correlating with distance, decreased between the drug and

lipid chemical species, confirming reactivity. In corroboration with molecular simulations, 2-aminomethylbenzimidazole exhibited intramolecular interactions and both propranolol and 4-amino-*N*-phenylbutanamide exhibited intermolecular interactions respectively. The former showed interactions between the aromatic protons on the molecule, whilst the latter two demonstrated interactions between polar parts of propranolol with the lipid head group protons and the aliphatic protons of 4-amino-*N*-phenylbutanamide with the lipid acyl chain protons close to the carbonyl site of reactivity. Future work would include synthesis of acylated propranolol and additional oxygen-carbon HMBC experiments to further confirm transesterification mechanisms present in propranolol. Also, more work needs to be conducted into quantifying the rate of lysolipid formation of each of the three drug-lipid systems, this would enable quantitative kinetic studies to be undertaken for each drug-lipid system.

Chapter 4 initially gave an overview of the literature and the usefulness of simulations in shedding light on drug-lipid reactivity, as well as utilised methodology in this project. The results were predominantly presented as partial density plots, radial distribution function plots, orientation plots and potential of mean force plots. Ubiquitously, all neutral molecules were found to localise at the membrane interface, close to the lipid heads where reactivity can take place. Protonated molecules were found to localise closer to the aqueous region where water can be found in abundance. Partial density plots confirmed this experimental observation and for each molecule-lipid system, distance of the drug molecule relative to the PC lipid headgroup was tabularly presented and plotted to confirm localisation behaviour. Radial distribution functions, for known reactive compounds, were presented to show which atomic species of the molecule would be more likely to undergo reactivity with the PC lipid headgroups. For propranolol, oxygen, rather than nitrogen was closer in space to the lipid headgroups and thus more likely to react, as confirmed with experiments. For 2-aminomethylbenzimidazole, the amine nitrogen was more likely to react than the imidazole nucleophilic nitrogen. Orientation calculations were presented as a function of cos of the angle relative to the principal angle of rotation. This confirmed both the preferred orientation of the molecule along the principal axis

and provided information about preferred interaction angles. Lastly, PMF calculations confirmed the preferred orientation and localisation of the molecule. Neutral molecules were again found to localise closer to the membrane interface, as confirmed by a thermodynamic minimum, whereas the energy minimum for protonated molecules was found further out in the aqueous region. This trend was found to be largely consistent for most molecules. The last few protonated molecules seemed to exhibit strange behaviour, where no energy minimum was found in the PMF profiles. However, this was consistent with experimental findings that these protonated molecules (serotonin, tetracaine, sertraline, chloroquine) are unable to cross the membrane bilayer.

Chapter 5 initially gave a literature perspective on the wide application of QM/MM in the study of chemical systems. It proposed that the accuracy of QM coupled with the speed of MM provided a good methodology for the study of drug-lipid interactions. The required methodology was outlined alongside different approaches used in QM/MM. From the earlier work in Chapter 4, 1D histograms were generated to show the frequency distribution of both the distance, between the nucleophilic atom and lipid, and the angle between the nucleophilic atom, and the lipid carbonyl oxygen. These results showed that 2-aminomethylbenzimidazole and protonated propranolol had structures close to the optimal BD conditions required for reactivity. Conversely, these plots showed that both fluoxetine and phentermine exhibited very few conformations close to the optimal BD conditions, affirming their hydrolytic and no reactivity respectively. 4-amino-*N*-phenylbutanamide, serotonin and procaine all showed some activity with varying degrees based upon their nucleophilic potential as ascertained from experimental literature. Secondly, tables were presented with initial conditions for QM/MM energy profiles for each drug-lipid system. This data was presented with bond distances, angles and the number of atoms placed in the QM and MM regions. The first was the drug-membrane system, the second was the drug-lipid system in implicit solvent and the last was the drug-lipid (without lipid headgroup) system in implicit solvent. The general trend seen across all drug-lipid systems was that the energy required to reach the TS was lower in the membrane than the other two systems, implying additional stabilisation from the membrane

lipids. 2-aminomethylbenzimidazole, 4-amino-*N*-phenylbutanamide, serotonin and procaine displayed aminolysis behaviour whilst propranolol/protonated propranolol showed transesterification behaviour. Rather curiously, propranolol/protonated propranolol proceeded to a stable transesterified product with no stable intermediates as the minimum in the energy profile. DFT energy minimisations were performed and showed that in all cases, a stable acylated drug was formed after a series of protonation/deprotonation steps of the stable minimum in each energy profile. Future work needs to be conducted in performing DFT minimisations on these drug-lipid systems in the membrane itself. This would enable a comparison of energy differences between intermediates minimised in implicit solvent compared with the membrane.

Final Summary and Outlook

A combination of experimental and theoretical work has been performed to study drug-lipid interactions. The narrative of this thesis has shown the complementary aspect of theoretical work required to verify experimental findings. At the start of this work, a key goal was to understand the factors that control the reactivity of small organic molecules in lipid membranes.

To this end, I was successful in showing that computational methods can successfully predict binding orientations of molecules that adopt conformations that are reactive towards lipidation. Radial distribution functions for 2-aminomethylbenzimidazole, propranolol and 4-amino-*N*-phenylbutanamide predicted the favourable interaction between atomic species on the drug relative to the lipid carbonyl. A combination of partial density and potential of mean force calculations showed how these neutral molecules localise to the membrane-water interface close to the lipid head groups, and are thus favourably orientated for reactivity. QM/MM simulations quantified the intermediate energies and bond lengths to show the most favoured reaction pathway of the drug with lipid.

Experimentally, NOESY data has provided qualitative evidence that the models are reproducing binding orientations. The nOe build-up curves in particular show favourable intermolecular interactions between the drug and lipid and multiple cross

peaks in the NOESY show close proximity between the protons on the drug relative to protons on the lipid. Importantly, these experimental findings are corroborated well with theoretical models which also confirm preferred orientation and depth of partitioning once induced with POPC lipids.

To expand this work further, additional experimental and theoretical investigations would need to be conducted. In the former case, more results would need to be obtained from ssNMR measurements. nOe difference spectra would help to identify spatial links between atoms on the drug relative to the lipid, this would further confirm close proximity for reactivity. Isotopic enrichment of the oxygen atom in propranolol would allow ^{17}O-^{13}C HMBC experiments to be conducted to characterise through-bond connectivity of the esterified product. Furthermore, ^{1}H-^{13}C HMBC data had poor resolution, and thus another aim would be to optimise the liposome-drug ratio to ensure high quality HMBC data. From a synthetic perspective, another goal would be to optimise the ^{15}N labelling method to ensure successful deprotection of the labelled coupled product to re-form the labelled material. If it were possible to improve the yield in the final synthetic step, this method could be applied to any aniline derivatised compound, including biologically active molecules. To obtain quantitative data from ssNMR experiments, rate constants of the binding of the molecule with the lipid would need to be calculated and plotted to ensure validated kinetic data.

In the latter case, it would be interesting to ascertain whether a coarse-grained approach might be possible to probe drug-lipid reactivity. The enormous computational cost of *ab-initio* calculations mean that one has to be disciplined in defining the number of atoms in the QM region, relative to the MM region. Thus, another goal would be to simulate reactivity with full DFT functionals, and compare its performance to the already obtained PM6 reactivity profiles. Furthermore, undertaking DFT minimisations of the reaction intermediates in the membrane, rather than just implicit solvent, would enable a further comparison of the stability of the reactive intermediates and whether membrane effects stabilise the rate-determining step in the membrane.

An important biological application of this research would be to incubate these isotopically enriched molecules in HepG2 cells. Molecular and microbiology techniques could be used to probe reactivity of these molecules with "real" membranes, rather than just model liposomes. These results would give even more credence and confidence that orientation and depth of partitioning are factors which affect reactivity *in vitro*, and thus naturally occur in biological systems.

Finally, this work will be useful for predicting the reactivity of other molecules with membrane lipids. The data presented in this thesis can be extended to include reactivity of both peptides and proteins with membrane lipids. As lipidation may underpin some diseases, an extension of this work may be to predict certain cellular diseases, such as DIPL. Overall, the project aims in this thesis have been successfully met in both quantitatively (theoretical) and qualitatively (experimental) demonstrating the important roles of orientation and depth of partitioning of small organic molecules for their reactivity at the membrane interface.

Bibliography

1 C. Adamo and V. Barone. Toward reliable density functional methods without adjustable parameters: The pbe0 model. *The Journal of chemical physics*, 110 (13):6158–6170, 1999.

2 S. A. Adcock and J. A. McCammon. Molecular dynamics: survey of methods for simulating the activity of proteins. *Chemical reviews*, 106(5):1589–1615, 2006.

3 R. Akhtar, A. F. Zahoor, M. Irfan, T. H. Bokhari, and A. ul Haq. Recent green synthetic approaches toward ullmann reaction: A review. *Chemical Papers*, 76 (12):7275–7293, 2022.

4 J. M. Albano, E. de Paula, and M. Pickholz. Molecular dynamics simulations to study drug delivery systems. *Molecular Dynamics*, 73:73–90, 2018.

5 S. E. Allen, R. R. Walvoord, R. Padilla-Salinas, and M. C. Kozlowski. Aerobic copper-catalyzed organic reactions. *Chemical reviews*, 113(8):6234–6458, 2013.

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

7 N. Anderson and J. Borlak. Drug-induced phospholipidosis. *FEBS letters*, 580 (23):5533–5540, 2006.

8 S. E. Ashbrook and P. Hodgkinson. Perspective: Current advances in solid-state nmr spectroscopy. *The Journal of Chemical Physics*, 149(4):040901–1–040901–14, 2018.

9 A. Avdeef, K. Box, J. Comer, C. Hibbert, and K. Tam. ph-metric logp 10. determination of liposomal membrane-water partition coefficients of lonizable drugs. *Pharmaceutical research*, 15:209–215, 1998.

10 S. Azeez, P. Chaudhary, P. Sureshbabu, S. Sabiah, and J. Kandasamy. tert-

butyl nitrite mediated nitrogen transfer reactions: synthesis of benzotriazoles and azides at room temperature. *Organic & Biomolecular Chemistry*, 16(37): 8280–8285, 2018.

11 M. Baciu, S. C. Sebai, O. Ces, X. Mulet, J. A. Clarke, G. C. Shearman, R. V. Law, R. H. Templer, C. Plisson, C. A. Parker, et al. Degradative transport of cationic amphiphilic drugs across phospholipid bilayers. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 364(1847):2597–2614, 2006.

12 D. Bagayoko. Understanding density functional theory (dft) and completing it in practice. *AIP Advances*, 4(12):127104–1–127104–12, 2014.

13 J. Baldwin, A. Raab, K. Mensler, B. Arison, and D. McClure. Synthesis of (r)- and (s)-epichlorohydrin. *The Journal of Organic Chemistry*, 43(25):4876–4878, 1978.

14 M. Baumann and I. R. Baxendale. An overview of the synthetic routes to the best selling drugs containing 6-membered heterocycles. *Beilstein journal of organic chemistry*, 9(1):2265–2319, 2013.

15 A. Bax, R. H. Griffey, and B. L. Hawkins. Correlation of proton and nitrogen-15 chemical shifts by multiple quantum nmr. *Journal of Magnetic Resonance (1969)*, 55(2):301–315, 1983.

16 A. A. Behroozmand, K. Keating, and E. Auken. A review of the principles and applications of the nmr technique for near-surface characterization. *Surveys in geophysics*, 36:27–85, 2015.

17 I. P. Beletskaya and A. V. Cheprakov. Copper in cross-coupling reactions: The post-ullmann chemistry. *Coordination Chemistry Reviews*, 248(21-24):2337–2364, 2004.

18 R. Bell and J. Saunders. Correlation of the intramolecular nuclear overhauser effect with internuclear distance. *Canadian Journal of Chemistry*, 48(7):1114–1122, 1970.

19 D. Bemporad, J. W. Essex, and C. Luttmann. Permeation of small molecules through a lipid bilayer: a computer simulation study. *The Journal of Physical Chemistry B*, 108(15):4875–4884, 2004.

20 H. Bevinakatti and A. Banerji. Practical chemoenzymic synthesis of both enan-

tiomers of propranolol. *The Journal of Organic Chemistry*, 56(18):5372–5375, 1991.

21 D. C. Blakemore, L. Castro, I. Churcher, D. C. Rees, A. W. Thomas, D. M. Wilson, and A. Wood. Organic synthesis provides opportunities to transform drug discovery. *Nature chemistry*, 10(4):383–394, 2018.

22 B. Breiden and K. Sandhoff. Emerging mechanisms of drug-induced phospholipidosis. *Biological chemistry*, 401(1):31–46, 2019.

23 M. S. Bretscher. The molecules of the cell membrane. *Scientific American*, 253 (4):100–109, 1985.

24 H. M. Britt, C. A. García-Herrero, P. W. Denny, J. A. Mosely, and J. M. Sanderson. Lytic reactions of drugs with lipid membranes. *Chemical Science*, 10(3):674–680, 2019.

25 H. M. Britt, A. S. Prakash, S. Appleby, J. A. Mosely, and J. M. Sanderson. Lysis of membrane lipids promoted by small organic molecules: Reactivity depends on structure but not lipophilicity. *Science Advances*, 6(17):eaaz8598, 2020.

26 K. R. Brownstein and C. Tarr. Importance of classical diffusion in nmr studies of water in biological cells. *Physical review A*, 19(6):2446, 1979.

27 B. M. Bruininks, T. A. Wassenaar, and I. Vattulainen. Unbreaking assemblies in molecular simulations with periodic boundaries. *Journal of chemical information and modeling*, 63(11):3448–3452, 2023.

28 S. Buchini, A. Buschiazzo, and S. G. Withers. A new generation of specific trypanosoma cruzi trans-sialidase inhibitors. *Angewandte Chemie*, 120(14):2740–2743, 2008.

29 D. E. Bugay. Solid-state nuclear magnetic resonance spectroscopy: theory and pharmaceutical applications. *Pharmaceutical research*, 10:317–327, 1993.

30 A. Bunker and T. Róg. Mechanistic understanding from molecular dynamics simulation in pharmaceutical research 1: drug delivery. *Frontiers in Molecular Biosciences*, 7:604770, 2020.

31 K. Burke and L. O. Wagner. Dft in a nutshell. *International Journal of Quantum Chemistry*, 113(2):96–101, 2013.

32 M. Bursch, J.-M. Mewes, A. Hansen, and S. Grimme. Best-practice dft protocols

for basic molecular computational chemistry. *Angewandte Chemie International Edition*, 61(42):e202205735, 2022.

33 R. Car. Fixing jacob's ladder. *Nature chemistry*, 8(9):820–821, 2016.

34 G. Carona, G. Steyaerta, A. Pagliaraa, F. Reymondb, P. Crivoria, P. Gaillarda, P.-A. Carrupta, A. Avdeefb, J. Comerc, K. J. Boxd, et al. Structure-lipophilicity relationships of neutral and protonated b-blockers. *Helvetica Chimica Acta*, 82: 1211, 1999.

35 D. Casey, K. Charalambous, A. Gee, R. V. Law, and O. Ces. Amphiphilic drug interactions with model cellular membranes are influenced by lipid chain-melting temperature. *Journal of The Royal Society Interface*, 11(94):20131062, 2014.

36 C. Chen, H. Liu, B. Zhang, Y. Wang, K. Cai, Y. Tan, C. Gao, H. Liu, C. Tan, and Y. Jiang. A simple benzimidazole quinoline-conjugate fluorescent chemosensor for highly selective detection of ag+. *Tetrahedron*, 72(27-28): 3980–3985, 2016.

37 L. W. Chung, W. Sameera, R. Ramozzi, A. J. Page, M. Hatanaka, G. P. Petrova, T. V. Harris, X. Li, Z. Ke, F. Liu, et al. The oniom method and its applications. *Chemical reviews*, 115(12):5678–5796, 2015.

38 T. D. Claridge. *High-resolution NMR techniques in organic chemistry*, volume 27. Elsevier, 2016.

39 S. Cockcroft. Mammalian lipids: structure, synthesis and function. *Essays in biochemistry*, 65(5):813–845, 2021.

40 C. J. Cramer. *Essentials of computational chemistry: theories and models*. John Wiley & Sons, 2013.

41 C. J. Cramer, F. Bickelhaupt, et al. Essentials of computational chemistry. *Angewandte Chemie International Edition in English*, 42(4):381–381, 2003.

42 K.-S. Csizi and M. Reiher. Universal qm/mm approaches for general nanoscale applications. *Wiley Interdisciplinary Reviews: Computational Molecular Science*, 13(4):e1656, 2023.

43 Q. Cui and M. Karplus. Catalysis and specificity in enzymes: a study of triosephosphate isomerase and comparison with methyl glyoxal synthase. *Advances in protein chemistry*, 66:315–372, 2003.

44 D. L. Daleke. Regulation of transbilayer plasma membrane phospholipid asymmetry. *Journal of lipid research*, 44(2):233–242, 2003.

45 S. Dapprich, I. Komáromi, K. S. Byun, K. Morokuma, and M. J. Frisch. A new oniom implementation in gaussian98. part i. the calculation of energies, gradients, vibrational frequencies and electric field derivatives. *Journal of Molecular Structure: THEOCHEM*, 461:1–21, 1999.

46 P. M. Derlet and R. Maass. Emergent structural length scales in a model binary glass-the micro-second molecular dynamics time-scale regime. *Journal of Alloys and Compounds*, 821:153209, 2020.

47 P. F. Devaux. Static and dynamic lipid asymmetry in cell membranes. *Biochemistry*, 30(5):1163–1173, 1991.

48 F. Di Meo, G. Fabre, K. Berka, T. Ossman, B. Chantemargue, M. Paloncỳová, P. Marquet, M. Otyepka, and P. Trouillas. In silico pharmacology: Drug membrane partitioning and crossing. *Pharmacological Research*, 111:471–486, 2016.

49 C. J. Dickson, V. Hornak, R. A. Pearlstein, and J. S. Duca. Structure–kinetic relationships of passive membrane permeation from multiscale modeling. *Journal of the American Chemical Society*, 139(1):442–452, 2017.

50 C. J. Dickson, R. C. Walker, and I. R. Gould. Lipid21: complex lipid membrane simulations with amber. *Journal of chemical theory and computation*, 18(3): 1726–1736, 2022.

51 A. O. Dohn. Multiscale electrostatic embedding simulations for modeling structure and dynamics of molecules in solution: a tutorial review. *International Journal of Quantum Chemistry*, 120(21):e26343, 2020.

52 V. Dotsch, G. Wider, and K. Wuthrich. Phase-sensitive spectra in a single scan with coherence selection by pulsed field gradients. *Journal of Magnetic Resonance*, 109(2):263–264, 1994.

53 W. Dowhan and M. Bogdanov. Functional roles of lipids in membranes. In *New comprehensive biochemistry*, volume 36, pages 1–35. Elsevier, 2002.

54 G. Duché and J. M. Sanderson. The chemical reactivity of membrane lipids. *Chemical Reviews*, 124(6):3284–3330, 2024.

55 M. R. Edelmann, T. Hartung, R. Trussardi, H. Iding, G. Galley, P. Pflieger, and R. D. Norcross. Synthesis of enantiomerically pure [14c]-labelled morpholine

derivatives for a class of trace amine-associate receptor 1 agonists. *Journal of Labelled Compounds and Radiopharmaceuticals*, 59(14):635–639, 2016.

56 H. Elshihawy, M. A. Helal, M. Said, and M. A. Hammad. Design, synthesis, and enzyme kinetics of novel benzimidazole and quinoxaline derivatives as methionine synthase inhibitors. *Bioorganic & Medicinal Chemistry*, 22(1):550–558, 2014.

57 M. Elstner, T. Frauenheim, and S. Suhai. An approximate dft method for qm/mm simulations of biological structures and processes. *Journal of Molecular Structure: THEOCHEM*, 632(1-3):29–41, 2003.

58 G. Enkavi, M. Javanainen, W. Kulig, T. Róg, and I. Vattulainen. Multiscale simulations of biological membranes: the challenge to understand biological phenomena in a living substance. *Chemical reviews*, 119(9):5607–5774, 2019.

59 U. Essmann, L. Perera, M. L. Berkowitz, T. Darden, H. Lee, and L. G. Pedersen. A smooth particle mesh ewald method. *The Journal of chemical physics*, 103 (19):8577–8593, 1995.

60 D. J. Evans and B. L. Holian. The nose–hoover thermostat. *The Journal of chemical physics*, 83(8):4069–4074, 1985.

61 J. C. Facelli. Chemical shift tensors: Theory and application to molecular structural problems. *Progress in nuclear magnetic resonance spectroscopy*, 58 (3-4):176, 2010.

62 E. Fahy, S. Subramaniam, H. A. Brown, C. K. Glass, A. H. Merrill, R. C. Murphy, C. R. Raetz, D. W. Russell, Y. Seyama, W. Shaw, et al. A comprehensive classification system for lipids1. *Journal of lipid research*, 46(5):839–861, 2005.

63 B. Farmer Ii, S. Macura, and L. Brown. The effect of molecular motion on cross relaxation in the laboratory and rotating frames. *Journal of Magnetic Resonance (1969)*, 80(1):1–22, 1988.

64 A. R. Fersht. Profile of martin karplus, michael levitt, and arieh warshel, 2013 nobel laureates in chemistry. *Proceedings of the National Academy of Sciences*, 110(49):19656–19657, 2013.

65 J. D. Firth and I. J. Fairlamb. A need for caution in the preparation and application of synthetically versatile aryl diazonium tetrafluoroborate salts. *Organic Letters*, 22(18):7057–7059, 2020.

66 R. R. Flavell, C. Von Morze, J. E. Blecha, D. E. Korenchan, M. Van Criekinge, R. Sriram, J. W. Gordon, H.-Y. Chen, S. Subramaniam, R. A. Bok, et al. Application of good's buffers to ph imaging using hyperpolarized 13 c mri. *Chemical Communications*, 51(74):14119–14122, 2015.

67 C. Fonseca Guerra, J. Snijders, G. t. Te Velde, and E. J. Baerends. Towards an order-n dft method. *Theoretical Chemistry Accounts*, 99:391–403, 1998.

68 L. F. Galiullina, H. A. Scheidt, D. Huster, A. Aganov, and V. Klochkov. Interaction of statins with phospholipid bilayers studied by solid-state nmr spectroscopy. *Biochimica et Biophysica Acta (BBA)-Biomembranes*, 1861(3):584–593, 2019.

69 C. Garcia, C. J. Andersen, and C. N. Blesso. The role of lipids in the regulation of immune responses. *Nutrients*, 15(18):3899, 2023.

70 D. P. Geerke, S. Thiel, W. Thiel, and W. F. van Gunsteren. Qm–mm interactions in simulations of liquid water using combined semi-empirical/classical hamiltonians. *Physical Chemistry Chemical Physics*, 10(2):297–302, 2008.

71 K. Ghosh and T. Sarkar. Selective sensing of cu (ii) by a simple anthracene-based tripodal chemosensor. *Supramolecular Chemistry*, 23(6):435–440, 2011.

72 A. M. Gomes, P. J. Costa, and M. Machuqueiro. Recent advances on molecular dynamics-based techniques to address drug membrane permeability with atomistic detail. *BBA advances*, page 100099, 2023.

73 F. M. Goñi. The basic structure and dynamics of cell membranes: An update of the singer–nicolson model. *Biochimica et Biophysica Acta (BBA)-Biomembranes*, 1838(6):1467–1476, 2014.

74 H. E. Gottlieb, V. Kotlyar, A. Nudelman, et al. Nmr chemical shifts of common laboratory solvents as trace impurities. *Journal of organic chemistry*, 62(21): 7512–7515, 1997.

75 M. E. Greenberg, X.-M. Li, B. G. Gugiu, X. Gu, J. Qin, R. G. Salomon, and S. L. Hazen. The lipid whisker model of the structure of oxidized cell membranes. *Journal of biological chemistry*, 283(4):2385–2396, 2008.

76 A. Greenfield and C. Grosanu. Convenient synthesis of primary sulfonamides. *Tetrahedron Letters*, 49(44):6300–6303, 2008.

77 R. B. Griffiths. Microcanonical ensemble in quantum statistical mechanics. *Journal of Mathematical Physics*, 6(10):1447–1461, 1965.

78 S. Grimme, A. Hansen, J. G. Brandenburg, and C. Bannwarth. Dispersion-corrected mean-field electronic structure methods. *Chemical reviews*, 116(9): 5105–5154, 2016.

79 A. Grossfield. An implementation of wham: The weighted histogram analysis method, version 2.0. 9. 18. *"WHAM: the weighted histogram analysis method"*, 2014.

80 E. Grunewald and R. Knight. A laboratory study of nmr relaxation times in unconsolidated heterogeneous sediments. *Geophysics*, 76(4):G73–G83, 2011.

81 J. R. Gullingsrud, R. Braun, and K. Schulten. Reconstructing potentials of mean force through time series analysis of steered molecular dynamics simulations. *Journal of Computational Physics*, 151(1):190–211, 1999.

82 S. Guzelj, M. Gobec, D. Urbančič, I. Mlinarič-Raščan, E. Corsini, and Ž. Jakopin. Structural features and functional activities of benzimidazoles as nod2 antagonists. *European journal of medicinal chemistry*, 190:112089, 2020.

83 X. Han. Cross coupling of 3-bromopyridine and sulfonamides (r1nhso2r2 · r1= h, me, alkyl; r2= alkyl and aryl) catalyzed by cui/1, 3-di (pyridin-2-yl) propane-1, 3-dione. *Tetrahedron Letters*, 51(2):360–362, 2010.

84 C. Hannesschlaeger, A. Horner, and P. Pohl. Intrinsic membrane permeability to small molecules. *Chemical reviews*, 119(9):5922–5953, 2019.

85 T. Hansson, C. Oostenbrink, and W. van Gunsteren. Molecular dynamics simulations. *Current opinion in structural biology*, 12(2):190–196, 2002.

86 M. D. Harmony, V. W. Laurie, R. L. Kuczkowski, R. Schwendeman, D. Ramsay, F. J. Lovas, W. J. Lafferty, and A. G. Maki. Molecular structures of gas-phase polyatomic molecules determined by spectroscopic methods. *J. Phys. Chem. Ref. Data*, 8:619, 1979.

87 F. A. Heberle and G. W. Feigenson. Phase separation in lipid membranes. *Cold Spring Harbor perspectives in biology*, 3(4):a004630, 2011.

88 H. Heinz, R. Vaia, B. Farmer, and R. Naik. Accurate simulation of surfaces and interfaces of face-centered cubic metals using 12- 6 and 9- 6 lennard-jones potentials. *The Journal of Physical Chemistry C*, 112(44):17281–17290, 2008.

89 J. W. Hennel and J. Klinowski. Magic-angle spinning: a historical perspective. *Topics in Current Chemistry*, pages 1–14, 2005.

90 L. Herbette, A. Katz, and J. Sturtevant. Comparisons of the interaction of propranolol and timolol with model and biological membrane systems. *Molecular pharmacology*, 24(2):259–269, 1983.

91 M. Hertrich. Imaging of groundwater with nuclear magnetic resonance. *Progress in Nuclear Magnetic Resonance Spectroscopy*, 53(4):227–248, 2008.

92 H. H. Hodgson. The sandmeyer reaction. *Chemical reviews*, 40(2):251–277, 1947.

93 V. Hornak, R. Abel, A. Okur, B. Strockbine, A. Roitberg, and C. Simmerling. Comparison of multiple amber force fields and development of improved protein backbone parameters. *Proteins: Structure, Function, and Bioinformatics*, 65 (3):712–725, 2006.

94 Y.-W. Hsueh, M.-T. Chen, P. J. Patty, C. Code, J. Cheng, B. J. Frisken, M. Zuckermann, and J. Thewalt. Ergosterol in popc membranes: physical properties and comparison with structurally similar sterols. *Biophysical journal*, 92 (5):1606–1615, 2007.

95 T. Hu, I. R. Baxendale, and M. Baumann. Exploring flow procedures for diazonium formation. *Molecules*, 21(7):918, 2016.

96 C.-h. Huang. Structural organization and properties of membrane lipids. In *Cell Physiology Source Book*, pages 43–63. Elsevier, 2001.

97 V. Hughes, P. Ellis, T. Burt, and N. Langlois. The practical application of reflectance spectrophotometry for the demonstration of haemoglobin and its degradation in bruises. *Journal of clinical pathology*, 57(4):355–359, 2004.

98 G. Hummer. Fast-growth thermodynamic integration: Error and efficiency analysis. *The Journal of Chemical Physics*, 114(17):7330–7337, 2001.

99 Y. Imamura, T. Otsuka, and H. Nakai. Description of core excitations by time-dependent density functional theory with local density approximation, generalized gradient approximation, meta-generalized gradient approximation, and hybrid functionals. *Journal of computational chemistry*, 28(12):2067–2074, 2007.

100 J. P. Jambeck and A. P. Lyubartsev. Exploring the free energy landscape of

solutes embedded in lipid bilayers. *The Journal of Physical Chemistry Letters*, 4(11):1781–1787, 2013.

101 P. Janmey and P. K. Kinnunen. Biophysical properties of lipids and dynamic membranes. *Trends in cell biology*, 16(10):538–546, 2006.

102 I.-Y. Jeong, W. S. Lee, S. Goto, S. Sano, M. Shiro, and Y. Nagao. Novel heterocyclic ring-expansion and/or dehydration-hydration reactions of propargylic and allenylic hydroxy γ-lactams in the presence of strong base or lewis acid. *Tetrahedron*, 54(48):14437–14454, 1998.

103 J. Karpińska. Derivative spectrophotometry—recent applications and directions of developments. *Talanta*, 64(4):801–822, 2004.

104 L. Kay, P. Keifer, and T. Saarinen. Pure absorption gradient enhanced heteronuclear single quantum correlation spectroscopy with improved sensitivity. *Journal of the American Chemical Society*, 114(26):10663–10665, 1992.

105 Q. Ke, X. Gong, S. Liao, C. Duan, and L. Li. Effects of thermostats/barostats on physical properties of liquids by molecular dynamics simulations. *Journal of Molecular Liquids*, 365:120116, 2022.

106 V. Kiessling, J. M. Crane, and L. K. Tamm. Transbilayer effects of raft-like lipid domains in asymmetric planar bilayers measured by single molecule tracking. *Biophysical journal*, 91(9):3313–3326, 2006.

107 W. Koch and M. C. Holthausen. *A chemist's guide to density functional theory*. John Wiley & Sons, 2015.

108 W. Kohn and L. J. Sham. Self-consistent equations including exchange and correlation effects. *Physical review*, 140(4A):A1133, 1965.

109 Z. Kokan, M. Dušková-Smrčková, and V. Šindelář. Supramolecular hydrogelation via host-guest anion recognition: Lamellar hydrogel materials for the release of cationic cargo. *Chem*, 7(9):2473–2490, 2021.

110 W. Kopeć, J. Telenius, and H. Khandelia. Molecular dynamics simulations of the interactions of medicinal plant extracts and drugs with lipid bilayer membranes. *The FEBS journal*, 280(12):2785–2805, 2013.

111 S. D. Krämer, A. Braun, C. Jakits-Deiser, and H. Wunderli-Allenspach. Towards the predictability of drug-lipid membrane interactions: the ph-dependent

affinity of propranolol to phosphatidylinositol containing liposomes. *Pharmaceutical research*, 15:739–744, 1998.

112 A. Kukol et al. *Molecular modeling of proteins*, volume 443. Springer, 2008.

113 E. C. Lai. Lipid rafts make for slippery platforms. *The Journal of cell biology*, 162(3):365, 2003.

114 A. Laouini, C. Jaafar-Maalej, I. Limayem-Blouza, S. Sfar, C. Charcosset, and H. Fessi. Preparation, characterization and applications of liposomes: state of the art. *Journal of colloid Science and Biotechnology*, 1(2):147–168, 2012.

115 D. D. Laws, H.-M. L. Bitter, and A. Jerschow. Solid-state nmr spectroscopic methods in chemistry. *Angewandte Chemie International Edition*, 41(17):3096–3129, 2002.

116 K. Leftheris and M. Goodman. Synthesis and. beta.-adrenergic antagonist activity of stereoisomeric practolol and propranolol derivatives. *Journal of medicinal chemistry*, 33(1):216–223, 1990.

117 J. W. Lehmann, D. J. Blair, and M. D. Burke. Towards the generalized iterative synthesis of small molecules. *Nature Reviews Chemistry*, 2(2):0115, 2018.

118 A. N. Leonard, E. Wang, V. Monje-Galvan, and J. B. Klauda. Developing and testing of lipid force fields with applications to modeling cellular membranes. *Chemical reviews*, 119(9):6227–6269, 2019.

119 L. Lerner and A. Bax. Sensitivity-enhanced two-dimensional heteronuclear relayed coherence transfer nmr spectroscopy. *Journal of Magnetic Resonance (1969)*, 69(2):375–380, 1986.

120 M. H. Levitt. *Spin dynamics: basics of nuclear magnetic resonance*. John Wiley & Sons, 2008.

121 H. Li, T. Zhao, and Z. Sun. Analytical techniques and methods for study of drug-lipid membrane interactions. *Reviews in Analytical Chemistry*, 37(1): 20170012, 2018.

122 X. Li, J. Han, S. Bujaranipalli, J. He, E. Y. Kim, H. Kim, J. H. Im, and W.-J. Cho. Structure-based discovery and development of novel o-glcnacase inhibitors for the treatment of alzheimer's disease. *European Journal of Medicinal Chemistry*, 238:114444, 2022.

123 T.-C. Lim. Alignment of buckingham parameters to generalized lennard-jones potential functions. *Zeitschrift für Naturforschung A*, 64(3-4):200–204, 2009.

124 H. Lin and D. Sun. Recent synthetic developments and applications of the ullmann reaction. a review. *Organic preparations and procedures international*, 45(5):341–394, 2013.

125 X. Liu, C. Wu, P. Bai, Y. Miao, Y. Hu, and Y. Xie. Recyclable silver-catalyzed selective hydrogenation of imides to primary amines via dual c–n bond cleavage. *Organic Letters*, 25(17):3066–3071, 2023.

126 A. Lodola and M. De Vivo. The increasing role of qm/mm in drug discovery. *Advances in protein chemistry and structural biology*, 87:337–362, 2012.

127 J. Lombard. Once upon a time the cell membranes: 175 years of cell boundary research. *Biology direct*, 9:1–35, 2014.

128 D. Lopes, S. Jakobtorweihen, C. Nunes, B. Sarmento, and S. Reis. Shedding light on the puzzle of drug-membrane interactions: Experimental techniques and molecular dynamics simulations. *Progress in lipid research*, 65:24–44, 2017.

129 D. Louvel, A. Chelagha, J. Rouillon, P.-A. Payard, L. Khrouz, C. Monnereau, and A. Tlili. Metal-free visible-light synthesis of arylsulfonyl fluorides: Scope and mechanism. *Chemistry–A European Journal*, 27(34):8704–8708, 2021.

130 G. Mackin and A. Shaka. Phase-sensitive two-dimensional hmqc and hmqc-tocsy spectra obtained using double pulsed-field-gradient spin echoes. *Journal of Magnetic Resonance, Series A*, 118(2):247–255, 1996.

131 S. Macura, Y. Huang, D. Suter, and R. Ernst. Two-dimensional chemical exchange and cross-relaxation spectroscopy of coupled nuclear spins. *Journal of Magnetic Resonance (1969)*, 43(2):259–281, 1981.

132 R. P. Magalhães, H. S. Fernandes, and S. F. Sousa. Modelling enzymatic mechanisms with qm/mm approaches: current status and future challenges. *Israel Journal of Chemistry*, 60(7):655–666, 2020.

133 A. Magarkar, V. Dhawan, P. Kallinteri, T. Viitala, M. Elmowafy, T. Róg, and A. Bunker. Cholesterol level affects surface charge of lipid membranes in saline solution. *Scientific reports*, 4(1):5005, 2014.

134 L. Malet-Sanz, J. Madrzak, S. V. Ley, and I. R. Baxendale. Preparation of arylsulfonyl chlorides by chlorosulfonylation of in situ generated diazonium salts

using a continuous flow reactor. *Organic & Biomolecular Chemistry*, 8(23): 5324–5332, 2010.

135 S. J. Marrink, V. Corradi, P. C. Souza, H. I. Ingolfsson, D. P. Tieleman, and M. S. Sansom. Computational modeling of realistic cell membranes. *Chemical reviews*, 119(9):6184–6226, 2019.

136 M. Marsman, J. Paier, A. Stroppa, and G. Kresse. Hybrid functionals applied to extended systems. *Journal of Physics: Condensed Matter*, 20(6):064201, 2008.

137 R. Menichetti, K. Kremer, and T. Bereau. Efficient potential of mean force calculation from multiscale simulations: solute insertion in a lipid membrane. *Biochemical and biophysical research communications*, 498(2):282–287, 2018.

138 S. Milne, P. Ivanova, J. Forrester, and H. A. Brown. Lipidomics: an analysis of cellular lipids by esi-ms. *Methods*, 39(2):92–103, 2006.

139 A. Missner and P. Pohl. 110 years of the meyer–overton rule: predicting membrane permeability of gases and other small compounds. *ChemPhysChem*, 10 (9-10):1405–1414, 2009.

140 T. R. Molugu, S. Lee, and M. F. Brown. Concepts and methods of solid-state nmr spectroscopy applied to biomembranes. *Chemical reviews*, 117(19):12087–12132, 2017.

141 S. Moradi, A. Nowroozi, and M. Shahlaei. Shedding light on the structural properties of lipid bilayers using molecular dynamics simulation: a review study. *RSC advances*, 9(8):4644–4658, 2019.

142 C. R. Morcombe and K. W. Zilm. Chemical shift referencing in mas solid state nmr. *Journal of Magnetic Resonance*, 162(2):479–486, 2003.

143 A. Mudi and C. Chakravarty. Effect of the berendsen thermostat on the dynamical properties of water. *Molecular Physics*, 102(7):681–685, 2004.

144 M. P. Muller, T. Jiang, C. Sun, M. Lihan, S. Pant, P. Mahinthichaichan, A. Trifan, and E. Tajkhorshid. Characterization of lipid–protein interactions and lipid-mediated modulation of membrane protein function through molecular simulation. *Chemical reviews*, 119(9):6086–6161, 2019.

145 F. Nardelli, S. Borsacchi, L. Calucci, E. Carignani, F. Martini, and M. Geppi. Anisotropy and nmr spectroscopy. *Rendiconti Lincei. Scienze Fisiche e Naturali*, 31(4):999–1010, 2020.

146 G. L. Nicolson. Update of the 1972 singer-nicolson fluid-mosaic model of membrane structure. *Discoveries*, 1(1):500–550, 2013.

147 I. Ø. Nielsen, K. K. B. Clemmensen, D. L. Fogde, T. N. Dietrich, J. D. Giacobini, M. Bilgin, M. Jäättelä, and K. Maeda. Cationic amphiphilic drugs induce accumulation of cytolytic lysoglycerophospholipids in the lysosomes of cancer cells and block their recycling into common membrane glycerophospholipids. *Molecular biology of the cell*, 35(3):ar25, 2024.

148 T. Nonoyama and R. Fukuda. Drug-induced phospholipidosis-pathological aspects and its prediction. *Journal of toxicologic pathology*, 21(1):9–24, 2008.

149 E. Okamura. Solution nmr to quantify mobility in membranes: diffusion, protrusion, and drug transport processes. *Chemical and Pharmaceutical Bulletin*, 67(4):308–315, 2019.

150 J. M. H. Olsen, C. Steinmann, K. Ruud, and J. Kongsted. Polarizable density embedding: A new qm/qm/mm-based computational strategy. *The Journal of Physical Chemistry A*, 119(21):5344–5355, 2015.

151 Y.-Q. Ouyang, Z.-H. Yang, Z.-H. Chen, and S.-Y. Zhao. Efficient copper-catalyzed n-arylation of nh-containing heterocycles and sulfonamides with arenediazonium tetrafluoroborates. *Synthetic Communications*, 47(8):771–778, 2017.

152 M. Paloncyova, K. Berka, and M. Otyepka. Molecular insight into affinities of drugs and their metabolites to lipid bilayers. *The Journal of Physical Chemistry B*, 117(8):2403–2410, 2013.

153 A. C. Pan, D. W. Borhani, R. O. Dror, and D. E. Shaw. Molecular determinants of drug–receptor binding kinetics. *Drug discovery today*, 18(13-14):667–673, 2013.

154 X. Pan, E. Rosta, and Y. Shao. Representation of the qm subsystem for long-range electrostatic interaction in non-periodic ab initio qm/mm calculations. *Molecules*, 23(10):2500, 2018.

155 R. Pastor and A. MacKerell Jr. Development of the charmm force field for lipids. *The journal of physical chemistry letters*, 2(13):1526–1532, 2011.

156 A. Paul, D. Chatterjee, S. Banerjee, and S. Yadav. Synthesis of 3-alkenylindoles

through regioselective c–h alkenylation of indoles by a ruthenium nanocatalyst. *Beilstein Journal of Organic Chemistry*, 16(1):140–148, 2020.

157 M. Paulikat, G. Piccini, E. Ippoliti, G. Rossetti, F. Arnesano, and P. Carloni. Physical chemistry of chloroquine permeation through the cell membrane with atomistic detail. *Journal of Chemical Information and Modeling*, 63(22):7124–7132, 2023.

158 C. Peetla, A. Stine, and V. Labhasetwar. Biophysical interactions with model lipid membranes: applications in drug discovery and drug delivery. *Molecular pharmaceutics*, 6(5):1264–1276, 2009.

159 J. P. Perdew and A. Zunger. Self-interaction correction to density-functional approximations for many-electron systems. *Physical review B*, 23(10):5048, 1981.

160 J. P. Perdew, A. Ruzsinszky, J. Sun, S. Glindmeyer, and G. I. Csonka. van der waals interaction as a summable asymptotic series. *Physical Review A—Atomic, Molecular, and Optical Physics*, 86(6):062714, 2012.

161 H. M. Petrassi, K. B. Sharpless, and J. W. Kelly. The copper-mediated cross-coupling of phenylboronic acids and n-hydroxyphthalimide at room temperature: synthesis of aryloxyamines. *Organic Letters*, 3(1):139–142, 2001.

162 R. R. Pissurlenkar, M. S. Shaikh, R. P. Iyer, and E. C. Coutinho. Molecular mechanics force fields and their applications in drug design. *Anti-Infective Agents in Medicinal Chemistry (Formerly Current Medicinal Chemistry-Anti-Infective Agents)*, 8(2):128–150, 2009.

163 J. A. Pople, P. M. Gill, and B. G. Johnson. Kohn—sham density-functional theory within a finite basis set. *Chemical physics letters*, 199(6):557–560, 1992.

164 M. Porcs-Makkay, T. Mezei, and G. Simig. New practical synthesis of the key intermediate of candesartan. *Organic process research & development*, 11(3): 490–493, 2007.

165 G. Proietti, K. J. Prathap, X. Ye, R. T. Olsson, and P. Dinér. Nickel boride catalyzed reductions of nitro compounds and azides: nanocellulose-supported catalysts in tandem reactions. *Synthesis*, 54(01):133–146, 2022.

166 J. X. Qiao and P. Y. Lam. Copper-promoted carbon-heteroatom bond cross-coupling with boronic acids and derivatives. *Synthesis*, 2011(06):829–856, 2011.

167 M. M. Quintal, A. Karton, M. A. Iron, A. D. Boese, and J. M. Martin. Bench-

mark study of dft functionals for late-transition-metal reactions. *The Journal of Physical Chemistry A*, 110(2):709–716, 2006.

168 M. M. Rainey, D. Korostyshevsky, S. Lee, and E. O. Perlstein. The antidepressant sertraline targets intracellular vesiculogenic membranes in yeast. *Genetics*, 185(4):1221–1233, 2010.

169 F. Ramos-Martín and N. D'Amelio. Biomembrane lipids: When physics and chemistry join to shape biological activity. *Biochimie*, 203:118–138, 2022.

170 J. Ramsden. Partition coefficients of drugs in bilayer lipid membranes. *Experientia*, 49:688–692, 1993.

171 K. S. Rao and T.-S. Wu. Chan–lam coupling reactions: synthesis of heterocycles. *Tetrahedron*, 68(38):7735–7754, 2012.

172 B. Reif, S. E. Ashbrook, L. Emsley, and M. Hong. Solid-state nmr spectroscopy. *Nature Reviews Methods Primers*, 1(1):2, 2021.

173 H. M. Roche. Unsaturated fatty acids. *Proceedings of the Nutrition Society*, 58 (2):397–401, 1999.

174 H. A. Rodríguez, F. M. Bickelhaupt, and I. Fernández. Origin of the bürgi-dunitz angle. *ChemPhysChem*, 24(17):e202300379, 2023.

175 T. Róg, M. Pasenkiewicz-Gierula, I. Vattulainen, and M. Karttunen. Ordering effects of cholesterol and its analogues. *Biochimica et Biophysica Acta (BBA)-Biomembranes*, 1788(1):97–121, 2009.

176 B. Roux. The calculation of the potential of mean force using computer simulations. *Computer physics communications*, 91(1-3):275–282, 1995.

177 A. C. Rustan and C. A. Drevon. Fatty acids: structures and properties. *Encyclopedia of life sciences*, 1:1–7, 2005.

178 C. Salata, A. Calistri, C. Parolin, A. Baritussio, and G. Palù. Antiviral activity of cationic amphiphilic drugs. *Expert review of anti-infective therapy*, 15(5): 483–492, 2017.

179 O. M. Salo-Ahen, I. Alanko, R. Bhadane, A. M. Bonvin, R. V. Honorato, S. Hossain, A. H. Juffer, A. Kabedev, M. Lahtela-Kakkonen, A. S. Larsen, et al. Molecular dynamics simulations in drug discovery and pharmaceutical development. *Processes*, 9(1):71, 2020.

180 C. Sambiagio, S. P. Marsden, A. J. Blacker, and P. C. McGowan. Copper catal-

ysed ullmann type chemistry: from mechanistic aspects to modern development. *Chemical Society Reviews*, 43(10):3525–3550, 2014.

181 J.-C. Sancho-Garcia, E. Brémond, A. Pérez-Jiménez, I. Ciofini, and C. Adamo. Non-empirical double-hybrid density functionals as reliable tools for electronic structure calculations. *Electronic Structure*, 4(4):043001, 2022.

182 J. M. Sanderson. Resolving the kinetics of lipid, protein and peptide diffusion in membranes. *Molecular membrane biology*, 29(5):118–143, 2012.

183 J. M. Sanderson. Far from inert: membrane lipids possess intrinsic reactivity that has consequences for cell biology. *BioEssays*, 42(3):1900147, 2020.

184 M. B. Santos, B. d. S. Costa, and E. E. G. Rojas. Calorimetric techniques applied to the thermodynamic study of interactions between proteins and polysaccharides. *Ciência Rural*, 46(8):1491–1497, 2016.

185 N. C. Santos, M. Prieto, and M. A. Castanho. Quantifying molecular partition into model systems of biomembranes: an emphasis on optical spectroscopic methods. *Biochimica et Biophysica Acta (BBA)-Biomembranes*, 1612(2):123–135, 2003.

186 H. N. Sarode, G. E. Lindberg, Y. Yang, L. E. Felberg, G. A. Voth, and A. M. Herring. Insights into the transport of aqueous quaternary ammonium cations: a combined experimental and computational study. *The Journal of Physical Chemistry B*, 118(5):1363–1372, 2014.

187 H. A. Scheidt and D. Huster. The interaction of small molecules with phospholipid membranes studied by 1h noesy nmr under magic-angle spinning. *Acta Pharmacologica Sinica*, 29(1):35–49, 2008.

188 S. L. Schreiber. Organic synthesis toward small-molecule probes and drugs. *Proceedings of the National Academy of Sciences*, 108(17):6699–6702, 2011.

189 H. L. Scott. Modeling the lipid component of membranes. *Current Opinion in Structural Biology*, 12(4):495–502, 2002.

190 A. M. Seddon, D. Casey, R. V. Law, A. Gee, R. H. Templer, and O. Ces. Drug interactions with lipid membranes. *Chemical Society Reviews*, 38(9):2509–2519, 2009.

191 H. M. Senn and W. Thiel. Qm/mm studies of enzymes. *Current opinion in chemical biology*, 11(2):182–187, 2007.

192 H. M. Senn and W. Thiel. Qm/mm methods for biomolecular systems. *Angewandte Chemie International Edition*, 48(7):1198–1229, 2009.

193 F. Shahidi and Y. Zhong. Lipid oxidation and improving the oxidative stability. *Chemical society reviews*, 39(11):4067–4079, 2010.

194 I. G. Shenderovich and H.-H. Limbach. Solid state nmr for nonexperts: an overview of simple but general practical methods. *Solids*, 2(2):139–154, 2021.

195 J. T. Simmons, J. R. Allen, D. R. Morris, R. J. Clark, C. W. Levenson, M. W. Davidson, and L. Zhu. Integrated and passive 1, 2, 3-triazolyl groups in fluorescent indicators for zinc (ii) ions: thermodynamic and kinetic evaluations. *Inorganic chemistry*, 52(10):5838–5850, 2013.

196 S. J. Singer and G. L. Nicolson. The fluid mosaic model of the structure of cell membranes: Cell membranes are viewed as two-dimensional solutions of oriented globular proteins and lipids. *Science*, 175(4023):720–731, 1972.

197 P. Smid, H. K. Coolen, H. G. Keizer, R. van Hes, J.-P. de Moes, A. P. den Hartog, B. Stork, R. H. Plekkenpol, L. C. Niemann, C. N. Stroomer, et al. Synthesis, structure- activity relationships, and biological properties of 1-heteroaryl-4-[ω-(1 h-indol-3-yl) alkyl] piperazines, novel potential antipsychotics combining potent dopamine d2 receptor antagonism with potent serotonin reuptake inhibition. *Journal of medicinal chemistry*, 48(22):6855–6869, 2005.

198 S. F. Sousa, P. A. Fernandes, and M. J. Ramos. General performance of density functionals. *The Journal of Physical Chemistry A*, 111(42):10439–10452, 2007.

199 J. J. Stewart. Application of the pm6 method to modeling the solid state. *Journal of molecular modeling*, 14:499–535, 2008.

200 J. J. Stewart. Application of the pm6 method to modeling proteins. *Journal of molecular modeling*, 15(7):765–805, 2009.

201 G. Strichartz. Kinetics of local anesthetic interactions with model membranes: Adsorption and permeation. *Open Journal of Anesthesiology*, 12(1):55–75, 2021.

202 K. Sun, C. Fang, W. Yang, Z. Xu, H. Wang, W. Sun, Y. Luo, and Y. Xu. Efficient synthesis of d6-clenproperol and d6-cimaterol using deuterium isopropylamine as labelled precursor. *Journal of Labelled Compounds and Radiopharmaceuticals*, 59(13):552–556, 2016.

203 M. Svensson, S. Humbel, R. D. Froese, T. Matsubara, S. Sieber, and K. Mo-

rokuma. Oniom: a multilayered integrated mo+ mm method for geometry optimizations and single point energy predictions. a test for diels- alder reactions and pt (p (t-bu) 3) 2+ h2 oxidative addition. *The Journal of Physical Chemistry*, 100(50):19357–19363, 1996.

204 M. Taillefer and D. Ma. *Amination and Formation of sp2 CN Bonds*, volume 46. Springer, 2013.

205 A. M. Teale, T. Helgaker, A. Savin, C. Adamo, B. Aradi, A. V. Arbuznikov, P. W. Ayers, E. J. Baerends, V. Barone, P. Calaminici, et al. Dft exchange: sharing perspectives on the workhorse of quantum chemistry and materials science. *Physical chemistry chemical physics*, 24(47):28700–28781, 2022.

206 C. H. Tse, J. Comer, Y. Wang, and C. Chipot. Link between membrane composition and permeability to drugs. *Journal of chemical theory and computation*, 14(6):2895–2909, 2018.

207 C. E. Tzeliou, M. A. Mermigki, and D. Tzeli. Review on the qm/mm methodologies and their application to metalloproteins. *Molecules*, 27(9):2660, 2022.

208 M. Utsintong, A. Massarotti, A. Caldarelli, and S. Theeramunkong. Parallel synthesis of "click" chalcones as antitubulin agents. *Medicinal Chemistry*, 9(4): 510–516, 2013.

209 Y. Utsugi, H. Kobuchi, Y. Kawamura, A. S. A. Atito, M. Nagao, H. Isoda, and Y. Miyamae. Importance of the proximity and orientation of ligand-linkage to the design of cinnamate-gw9662 hybrid compounds as covalent pparγ agonists. *Molecules*, 24(10):2019, 2019.

210 M. Uyanik, H. Tanaka, and K. Ishihara. Hypoiodite-catalyzed chemoselective tandem oxidation of homotryptamines to peroxy-and epoxytetrahydropyridoindolenines. *Organic Letters*, 22(20):8049–8054, 2020.

211 M. W. van der Kamp and A. J. Mulholland. Combined quantum mechanics/molecular mechanics (qm/mm) methods in computational enzymology. *Biochemistry*, 52(16):2708–2728, 2013.

212 G. Van Meer, D. R. Voelker, and G. W. Feigenson. Membrane lipids: where they are and how they behave. *Nature reviews Molecular cell biology*, 9(2): 112–124, 2008.

213 K. Vanommeslaeghe, O. Guvench, et al. Molecular mechanics. *Current pharmaceutical design*, 20(20):3281–3292, 2014.

214 M. Vater, L. Möckl, V. Gormanns, C. Schultz Fademrecht, A. M. Mallmann, K. Ziegart-Sadowska, M. Zaba, M. L. Frevert, C. Bräuchle, F. Holsboer, et al. New insights into the intracellular distribution pattern of cationic amphiphilic drugs. *Scientific reports*, 7(1):44277, 2017.

215 I. Vattulainen and T. Rog. Lipid simulations: a perspective on lipids in action. *Cold Spring Harbor perspectives in biology*, 3(4):a004655, 2011.

216 T. Vreven, K. Morokuma, Ö. Farkas, H. B. Schlegel, and M. J. Frisch. Geometry optimization with qm/mm, oniom, and other combined methods. i. microiterations and constraints. *Journal of computational chemistry*, 24(6):760–769, 2003.

217 T. Vreven, K. S. Byun, I. Komáromi, S. Dapprich, J. A. Montgomery Jr, K. Morokuma, and M. J. Frisch. Combining quantum mechanics methods with molecular mechanics methods in oniom. *Journal of chemical theory and computation*, 2(3):815–826, 2006.

218 D. O. Walsh, E. Grunewald, P. Turner, A. Hinnell, and P. Ferre. Practical limitations and applications of short dead time surface nmr. *Near Surface Geophysics*, 9(2):103–113, 2011.

219 F. Wang, C. Wang, G. Sun, and G. Zou. Highly efficient palladium-catalyzed cross-coupling of diarylborinic acids with arenediazoniums for practical diaryl synthesis. *Tetrahedron Letters*, 61(7):151491, 2020.

220 J. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman, and D. A. Case. Development and testing of a general amber force field. *Journal of computational chemistry*, 25(9):1157–1174, 2004.

221 L.-C. Wang, B. Chen, and X.-F. Wu. Cobalt-catalyzed direct aminocarbonylation of ethers: Efficient access to α-amide substituted ether derivatives. *Angewandte Chemie International Edition*, 61(23):e202203797, 2022.

222 T.-Y. Wang and J. R. Silvius. Cholesterol does not induce segregation of liquid-ordered domains in bilayers modeling the inner leaflet of the plasma membrane. *Biophysical journal*, 81(5):2762–2773, 2001.

223 Z. Wang, X. Ye, M. Jin, Q. Tang, S. Fan, Z. Song, and X. Shi. 4-aminobenzotriazole (abta) as a removable directing group for palladium-

catalyzed aerobic oxidative c–h olefination. *Organic letters*, 24(17):3107–3112, 2022.

224 Z.-K. Wang, X.-Q. Hong, J. Hu, Y.-Y. Xing, and W.-H. Chen. Synthesis and biological activity of squaramido-tethered bisbenzimidazoles as synthetic anion transporters. *RSC advances*, 11(7):3972–3980, 2021.

225 H. Watson. Biological membranes. *Essays in biochemistry*, 59:43–69, 2015.

226 N. Weizenmann, D. Huster, and H. A. Scheidt. Interaction of local anesthetics with lipid bilayers investigated by 1h mas nmr spectroscopy. *Biochimica et Biophysica Acta (BBA)-Biomembranes*, 1818(12):3010–3018, 2012.

227 P. A. Wiget, L. A. Manzano, J. M. Pruet, G. Gao, R. Saito, A. F. Monzingo, K. R. Jasheway, J. D. Robertus, and E. V. Anslyn. Sulfur incorporation generally improves ricin inhibition in pterin-appended glycine-phenylalanine dipeptide mimics. *Bioorganic & medicinal chemistry letters*, 23(24):6799–6804, 2013.

228 M. P. Williamson and D. Neuhaus. Symmetry in noe spectra. *Journal of Magnetic Resonance (1969)*, 72(2):369–375, 1987.

229 I. Wood, M. F. Martini, and M. Pickholz. Similarities and differences of serotonin and its precursors in their interactions with model membranes studied by molecular dynamics simulation. *Journal of Molecular Structure*, 1045:124–130, 2013.

230 C. Wu, A. Ramamoorthy, and S. Opella. High-resolution heteronuclear dipolar solid-state nmr spectroscopy. *Journal of Magnetic Resonance Series A*, 109: 270–270, 1994.

231 L. Yang, C.-h. Tan, M.-J. Hsieh, J. Wang, Y. Duan, P. Cieplak, J. Caldwell, P. A. Kollman, and R. Luo. New-generation amber united-atom force field. *The journal of physical chemistry B*, 110(26):13166–13176, 2006.

232 Z.-H. Yang, Z.-H. Chen, Y.-L. An, and S.-Y. Zhao. Selective and tunable synthesis of 3-arylsuccinimides and 3-arylmaleimides from arenediazonium tetrafluoroborates and maleimides. *RSC advances*, 6(28):23438–23447, 2016.

233 Z. Yue, C. Li, G. A. Voth, and J. M. Swanson. Dynamic protonation dramatically affects the membrane permeability of drug-like molecules. *Journal of the American Chemical Society*, 141(34):13421–13433, 2019.

234 J. Zhang, M. Jiang, C.-S. Wang, K. Guo, Q.-X. Li, C. Ma, S.-F. Ni, G.-Q. Chen,

Y. Zong, H. Lu, et al. Transition-metal free c–n bond formation from alkyl iodides and diazonium salts via halogen-atom transfer. *Nature Communications*, 13(1):7961, 2022.

235 Y. Zhang. Electrostatic interaction of the electrostatic-embedding and mechanical-embedding schemes for qm/mm calculations. *Communications in Computational Chemistry*, 1(2):109–117, 2013.

236 Y.-N. Zheng, H. Zheng, T. Li, and W.-T. Wei. Recent advances in copper-catalyzed c- n bond formation involving n-centered radicals. *ChemSusChem*, 14(24):5340–5358, 2021.

237 H. Zhu, R. Mishra, L. Yuan, S. F. Abdul Salam, J. Liu, G. Gray, A. D. Sterling, M. Wunderlich, J. Landero-Figueroa, J. T. Garrett, et al. Oxidative cyclization-induced activation of a phosphoinositide 3-kinase inhibitor for enhanced selectivity of cancer chemotherapeutics. *ChemMedChem*, 14(22):1933–1939, 2019.

238 S. Zhurenko, A. Tkachev, A. Gunbin, and A. Gippius. Upgrade of a bruker nmr spectrometers using a modern digital base. *Instruments and Experimental Techniques*, 64(3):427–433, 2021.

239 S. C. Zimmerman and R. Breslow. Asymmetric synthesis of amino acids by pyridoxamine enzyme analogs utilizing general base-acid catalysis. *Journal of the American Chemical Society*, 106(5):1490–1491, 1984.