Predicting Absolute Binding Free Energies for Lipid Exposed Binding Sites



Charlie Cook

University College

University of Oxford

A thesis submitted in partial fulfilment of

the requirements for the degree of

Doctor of Philosophy

Michaelmas 2024

Acknowledgements

First, I would like to express my gratitude to my supervisor Prof. Phil Biggin, who has always been there to provide advice on matters both scientific and non-scientific. His patience and academic insights have allowed this project to keep progressing. He has given me the space to explore and develop my scientific thinking, whilst keeping me working towards producing this thesis. My growth as a scientist over the last 4 years is largely down to his mentorship.

I would also like to thank Ewa Chudyk and Ronald Knegtel, my industrial supervisors at Vertex. Their enthusiasm and ideas led to my project exploring many avenues I otherwise would not have considered. It was a fantastic opportunity to work alongside them and see how industrial research can lead to transformative therapeutics. I would also like to thank Georgia McGaughey for facilitating my involvement with Vertex and her insights.

I must of course also thank the many members of the Structural Bioinformatics & Computational Biology (SBCB) group. I have been very lucky to work alongside such brilliant scientists and people. I will fondly remember all of the scientific discussions and fantastic seminars, but I shall even more fondly remember the group lunches, coffee breaks, and pub trips. Particular thanks must go to Dr. Irfan Alibay and Dr Aphroditi Zaki for answering my multitude of questions, particularly in the first few months of my DPhil. There are many SBCB members past and present who have made my experience memorable, but I would particularly like to thank Harold, Michael, Simon, Bertie, Linda, Jodie, Eddie, Johnny, and Matt. My time in Oxford has been enriched by so many people outside of the lab but Alessandro and Lucy and my housemates Rob, Bradley, and Danny all deserve special

mention. Will and Bjarne deserve special mention for managing to put up with me both at home and in the lab.

My parents have given me incredible support throughout my life, and the last 4 years have been no different. Their curiosity, love of knowledge, and passion for a good debate have shaped the scientist and person who I am today. Thank you also to Paula and Christian who have allowed me to stay at their home in Cornwall and escape Oxford and DPhil life and for generally showing me exceptional kindness. Finally, I must thank Molly. First, thank you for proofreading this thesis – I'm sure you now know my DPhil thesis in excruciating detail. You have been there for me every time I lost confidence and doubted myself. You put up with being locked in a tiny flat during a series of national lockdowns with all the pressures of us both doing PhD's. Your enthusiasm and general joy in life has made my DPhil experience infinitely better.

Table of Contents

TABLE OF CONTENTS ..... **IV**

ABSTRACT ..... **VIII**

TABLE OF FIGURES ..... **IX**

LIST OF TABLES ..... **XII**

LIST OF ABBREVIATIONS ..... **XIII**

1. INTRODUCTION **1**

1.1. TABLE OF CONTENTS 2

1.2. TABLE OF FIGURES 3

1.3. STRUCTURE-BASED DRUG DESIGN 4

1.3.1. Introduction to Structure-Based Drug Design *4*

1.3.2. The role of simulation *7*

1.3.3. Free energy calculations *10*

1.4. MEMBRANE PROTEINS 22

1.4.1. Introduction to Membrane Proteins *22*

1.4.2. Membrane Structure and Dynamics *24*

1.4.3. GPCRs *27*

1.5. LIPID EXPOSED BINDING SITES AS NOVEL DRUG TARGETS 36

1.5.1. The Microkinetic Model *39*

1.5.2. Experimental Affinity Measurement *42*

1.5.3. Lipid-Exposed Binding Sites as a Challenge for ABFEs *45*

1.6. THE SYSTEMS 49

1.6.1. P2Y1R *49*

1.6.2. Glucagon Receptor *55*

1.7. AIMS AND MOTIVATIONS 61

2. THEORY **63**

2.1. TABLE OF CONTENTS 64
2.2. TABLE OF FIGURES 65
2.3. COMPUTATION AT SCALE 66
2.3.1. Quantum Mechanics 66
2.4. MOLECULAR DYNAMICS & FORCE FIELDS 69
2.4.1. Molecular Mechanics Force Fields 69
2.4.2. Energy Minimisation 77
2.4.3. Molecular Dynamics 78
2.4.4. Periodic Boundary Conditions 85
2.4.5. Thermodynamic Ensembles and Pressure/Temperature coupling 86
2.4.6. Thermostats 87
2.4.7. Barostats 90
2.4.8. Long Range Interactions 92
2.4.9. Replica Exchange 98
2.5. FREE ENERGY CALCULATIONS 100
2.5.1. Defining Free Energy 100
2.5.2. Computational Methods for Free Energy Prediction 102
2.5.3. Alchemical Methods 103
3. FREE ENERGY CALCULATIONS FOR P2Y$_1$R 110
3.1. TABLE OF CONTENTS 111
3.2. TABLE OF FIGURES 112
3.3. LIST OF TABLES 113
3.4. INTRODUCTION 114
3.5. METHODS 117
3.6. INVESTIGATING LIPID DYNAMICS 121
3.7. P2Y$_1$R BPTU ABFE VALIDATION 127
3.7.1. ABFEs are accurate and precise 127
3.7.2. Multiple embeddings provide a more realistic estimate of the error 128
3.7.3. Results can be reproduced using multiple forcefields 130
3.7.4. Lipid sampling in the apo state 133

3.7.5. Lipid sampling in the partially interacting state *135*

3.7.6. Can we improve lipid sampling? *137*

3.8. FREE ENERGY CALCULATIONS OF $P2Y_1R$ SAR 140

3.8.1. Pseudo-symmetric groups require consideration *142*

3.8.2. ABFEs Accurately Rank $P2Y_1R$ SAR *144*

3.9. ABFES FOR BPTU SELECTIVITY 148

3.10. DISCUSSION 156

3.11. CONCLUSION 162

4. GLUCAGON RECEPTOR 164

4.1. TABLE OF CONTENTS 165

4.2. FIGURES 166

4.3. INTRODUCTION 167

4.4. METHODS 170

4.5. INITIAL SIMULATIONS 174

4.5.1. Apo Conformations *179*

4.5.2. Binding Site Occupancy *183*

4.6. RBFE CALCULATIONS 186

4.7. ABFE CALCULATIONS 190

4.8. SAR ABFES 199

4.8.1. ABFEs from 'true' apo conformation *204*

4.9. DISCUSSION 209

4.10. CONCLUSION 214

5. THE APO STATE IN ABFES 215

5.1. TABLE OF CONTENTS 216

5.2. FIGURES 217

5.3. INTRODUCTION 218

5.4. METHODS 222

5.5. GCGR 224

5.6. T4 AND $\boldsymbol{p38\alpha}$ 228

5.7. Discussion 237
5.8. Conclusion 241
6. ABFES IN COMPLEX BILAYERS 243
6.1. Table of Contents 244
6.2. Table of Figures 245
6.3. List of Tables 245
6.4. Introduction 246
6.5. Methods 250
6.6. GCGR & PIP_2 253
6.7. $P2Y_1R$ and Cholesterol 257
6.7.1. Atomistic simulations of crystallographic cholesterols 259
6.7.2. Doped Cholesterol Bilayer: Coarse-grained Simulations 261
6.8. ABFEs and the Microkinetic Model 265
6.9. Discussion 271
6.10. Conclusion 278
7. SUMMARY 281
7.1. Contextualization of this thesis' findings 282
7.2. Conclusion 286
8. SUPPLEMENTARY MATERIAL: CHAPTER 3 288
9. SUPPLEMENTARY MATERIAL: CHAPTER 4 292
10. SUPPLEMENTARY MATERIAL: CHAPTER 6 298
11. REFERENCES 300

Abstract

Absolute binding free energy (ABFE) calculations are finding increasing use within drug-design programs. To date, most work has focused on soluble proteins, with relatively little work done on membrane proteins. Structural studies are increasingly revealing the presence of lipid-exposed allosteric binding sites on membrane proteins. These sites provide an exciting new way to target GPCRs. Ligands binding at these sites are exposed to lipids rather than water and this is presumed to make ABFE calculations significantly more challenging. This thesis focuses on two GPCRs with lipid-exposed binding sites - the $P2Y_1$ receptor and the glucagon receptor (GCGR). I showed that it is challenging to sample the lipids' interactions on the timescale of ABFE calculations and make suggestions for monitoring and overcoming this. I showed that ABFE calculations can accurately rank SAR for these sites at a comparable level to many soluble systems ($r_P = 0.68$ and $r_P = 0.60$ respectively). ABFEs on the GCGR system show a consistent and large overestimation of the $\Delta G^{\theta}_{binding}$ on the order of 13 $kcal/mol$. I identified this as being a result of being unable to properly sample the apo state in the non-interacting windows of our ABFE calculations. I make suggestions for how to consider the apo conformation of the protein. This leads to improvements of $\sim$4 $kcal/mol$ relative to experimental values. I also show that for GCGR a proper description of the apo state must include the lipid $PI(4,5)P_2$ at a site overlapping the ligand binding site. I show that ABFEs in a mixed bilayer are not currently plausible due to sampling challenges. I therefore suggest an additional stage of the alchemical cycle to consider the binding of $PI(4,5)P_2$. For the GCGR system, I show $PI(4,5)P_2$ binds preferentially to POPC with $\Delta\Delta G_{binding} = 4.9 \pm 0.2\ kcal/mol$. These two corrections significantly improve the prediction relative to experimental data.

Table of Figures

FIGURE 1.1: ALCHEMICAL RBFE CYCLE 14

FIGURE 1.2: ALCHEMICAL ABFE CYCLE. 16

FIGURE 1.3: EQUILIBRIUM VS NON-EQUILIBRIUM ALCHEMICAL CALCULATIONS 18

FIGURE 1.4: A PROTEIN EMBEDDED IN A POPC:PIP_2:CHOLESTEROL BILAYER 26

FIGURE 1.5: ACTIVE AND INACTIVE STRUCTURES OF $\beta 2AR$. 30

FIGURE 1.6: THE DIFFERENT TYPES OF LIPID-EXPOSED SITES 38

FIGURE 1.7: THE STRUCTURE OF BPTU BOUND TO $P2Y_1R$ 54

FIGURE 1.8: THE STRUCTURE OF MK-0893 BOUND TO GCGR 60

FIGURE 2.1: SCHEMATICS OF BONDED INTERACTIONS FOUND IN TYPICAL FORCE FIELDS 72

FIGURE 2.2: QUALITATIVE PLOT OF THE LENNARD-JONES POTENTIAL 75

FIGURE 2.3: QUALITATIVE PLOT OF THE COULOMBIC POTENTIAL OF TWO PARTICLES WITH EQUAL SIGNED CHARGE 76

FIGURE 2.4: A SCHEMATIC OF THE LEAPFROG INTEGRATOR. 84

FIGURE 2.5: AN ILLUSTRATION OF PERIODIC BOUNDARY CONDITIONS IN 2D. 86

FIGURE 2.6: A QUALITATIVE REPRESENTATION OF A LENNARD-JONES POTENTIAL WITH A SWITCHING FUNCTION APPLIED. 94

FIGURE 2.7: A SCHEMATIC OF THE VERLET NEIGHBOUR LIST SCHEME. 95

FIGURE 2.8: 1D ILLUSTRATION OF THE EWALD SUM. 97

FIGURE 2.9: A SIMPLE SCHEMATIC OF THE REPLICA EXCHANGE METHOD. 100

FIGURE 2.10: THE ALCHEMICAL CYCLE FOR AN ABSOLUTE BINDING FREE ENERGY CALCULATION. 103

FIGURE 3.1: BPTU $P2Y_1R$ INTERACTIONS 122

FIGURE 3.2: BPTU/POPC INTERACTIONS 124

FIGURE 3.3 PROTEIN/POPC INTERACTIONS 125

FIGURE 3.4: CRYSTALLOGRAPHIC ABFES 128

FIGURE 3.5: ABFES FROM MULTIPLE EMBEDDINGS 130

FIGURE 3.6: ABFES WITH MULTIPLE FORCEFIELDS 132

FIGURE 3.7: ABFE EQUILIBRATION 135

FIGURE 3.8: $Lipid\ Sampling\ at\ \lambda VdW = 0.7$ 136

FIGURE 3.9 LIPID HREX....138
FIGURE 3.11: P2 SAR STRUCTURES....141
FIGURE 3.12: 6B CONFORMATIONS....142
FIGURE 3.13: 6B TERMINAL DIHEDRAL....144
FIGURE 3.14: $P2Y_1R$ SAR ABFE RESULTS....147
FIGURE 3.15: $P2Y_1R$ MUTATIONS....154
FIGURE 4.1: MK-0893 INTERACTIONS....177
FIGURE 4.2: POPC/PROTEIN INTERACTIONS....179
FIGURE 4.3: APO CONFORMATIONAL CHANGES....182
FIGURE 4.4: APO PROTRUDE DISTANCE....185
FIGURE 4.5: RBFE CYCLE AND SAR....189
FIGURE 4.6: RBFE RESULTS....189
FIGURE 4.7: GCGR/MK-0893 FORCEFIELDS....192
FIGURE 4.8: VDW 20 PROTEIN/POPC INTERACTIONS....195
FIGURE 4.9: VDW20 PROTRUDE DISTANCES....198
FIGURE 4.10 GCGR SAR....200
FIGURE 4.11: GCGR SAR ABFE RESULTS....203
FIGURE 4.12: APO/HOLO ABFE CALCULATIONS....208
FIGURE 5.1: SCHEMATIC OF OCCUPANCY DEPENDENCY....220
FIGURE 5.2: APO/HOLO ABFES WITH HREX....225
FIGURE 5.3: GCGR APO/HOLO ABFE OCCUPANCIES....227
FIGURE 5.4: T4 AND P38α STRUCTURES....230
FIGURE 5.5: T4 AND P38α APO/HOLO ABFES....234
FIGURE 5.6: GCGR SAR APO ABFE RESULTS....236
FIGURE 6.1: GCGR PIP_2 STABILITY....254
FIGURE 6.2: POPC/PIP_2 RBFE CYCLE....256
FIGURE 6.3: CHOLESTEROL SAMPLING....258
FIGURE 6.4: BPTU ABFES WITH CHOLESTEROL....260
FIGURE 6.5: THE MICROKINETIC MODEL AND ALCHEMICAL CYCLES....266
FIGURE 6.6: SELECTED SAR STRUCTURES.....267

FIGURE 6.7: PARTITION SIMULATION ENDPOINTS 269
FIGURE 6.8: P2Y$_1$R SAR MEMBRANE PARTITIONING RESULTS 270
SUPPLEMENTARY FIGURE 8.1: UNBIASED SIMULATIONS RMSD 288
SUPPLEMENTARY FIGURE 8.2: NON-EQUILIBRIUM WORK DISTRIBUTION. 289
SUPPLEMENTARY FIGURE 8.3: P2Y$_1$R BPTU SAR GRID 290
SUPPLEMENTARY FIGURE 8.4: MM/GBSA 6B CONFORMER PREDICTION 291
SUPPLEMENTARY FIGURE 9.1 GCGR SYMMETRY MATE. 292
SUPPLEMENTARY FIGURE 9.2: UNBIASED SIMULATION RMSDS 292
SUPPLEMENTARY FIGURE 9.3: STABILITY OF WATER BRIDGES 293
SUPPLEMENTARY FIGURE 9.4: BESPOKEFIT TORSIONS 294
SUPPLEMENTARY FIGURE 9.5: GCGR SAR ABFE RESULTS & STRUCTURES GRID. 295
SUPPLEMENTARY FIGURE 9.6: DUMMY ATOM VALIDATION 296
SUPPLEMENTARY FIGURE 9.7: MK-0893 APO/HOLO ABFE OVERLAP MATRICES 297
SUPPLEMENTARY FIGURE 9.8: GCGR WATER BRIDGES 297
SUPPLEMENTARY FIGURE 10.1: VDW20 R346$^{6.37B}$/PIP$_2$ CONTACTS 298
SUPPLEMENTARY FIGURE 10.2: RBFE ORIENTATIONAL DEPENDENCE 298
SUPPLEMENTARY FIGURE 10.3: PARTITION METADYNAMICS 299

List of Tables

TABLE 2.1: COMMON THERMODYNAMIC ENSEMBLES 87

TABLE 2.2: THE CHARGE OF A SYSTEM CAN BE KEPT NEUTRAL WHEN ANNIHILATING A CHARGED LIGAND BY USING A CO-ALCHEMICAL ION. AN OPPOSITELY CHARGED ION CAN BE 'CO-ANNIHILATED', WHILST A SIMILARLY CHARGED ION CAN BE USED FOR 'CHARGE-TRANSFER'. 98

TABLE 3.1: LIPID HREX RESULTS 139

TABLE 3.2:MUTATION RESULTS 151

TABLE 6.1: ABFES WITH CHOLESTEROL RESULTS 260

List of Abbreviations

ABFE	Absolute binding free energy.
CHS	Cholesteryl hemisuccinate
CV	Collective variable
Cryo-EM	Cryogenic electron microscopy
FEP	Free energy perturbation.
GCGR	Glucagon Receptor
GPCR	G-protein coupled receptor
HREX	Hamiltonian Replica Exchange
ITC	Isothermal titration calorimetry.
MAE	Mean absolute error
MBAR	Multistate Bennett's acceptance ratio.
MD	Molecular dynamics.
MM	Molecular mechanics
MMPBSA	Molecular mechanics Poisson-Boltzmann surface area
$P2Y_1R$	Purinergic Receptor P2Y1
PDB	Protein Data Bank
PIP2	Phosphatidylinositol 4,5-bisphosphate
PMF	Position of mean force
POPC	1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine
QM	Quantum mechanics
RBFE	Relative binding free energy.
RMSD	Root mean squared deviation.
TMD	Transmembrane domain

vdW	van der Waals.
WT	Wild type

1. Introduction

1.1. Table of Contents

TABLE OF FIGURES **3**

STRUCTURE-BASED DRUG DESIGN **4**

Introduction to Structure-Based Drug Design 4

The role of simulation 7

Free energy calculations 10

MEMBRANE PROTEINS **22**

Introduction to Membrane Proteins 22

Membrane Structure and Dynamics 24

Lipid Composition *24*

Protein-Lipid Interactions *26*

GPCRs 27

Structure, Function, and Activation *28*

Simulation of GPCRs *33*

LIPID EXPOSED BINDING SITES AS NOVEL DRUG TARGETS **36**

The Microkinetic Model 39

Experimental Affinity Measurement 42

Lipid-Exposed Binding Sites as a Challenge for ABFEs 45

THE SYSTEMS **49**

P2Y1R 49

Glucagon Receptor 55

AIMS AND MOTIVATIONS **61**

1.2. Table of Figures

FIGURE 1.1: ALCHEMICAL RBFE CYCLE 14

FIGURE 1.2: ALCHEMICAL ABFE CYCLE. 16

FIGURE 1.3: EQUILIBRIUM VS NON-EQUILIBRIUM ALCHEMICAL CALCULATIONS 18

FIGURE 1.4: A PROTEIN EMBEDDED IN A POPC:PIP_2:CHOLESTEROL BILAYER 26

FIGURE 1.5: ACTIVE AND INACTIVE STRUCTURES OF $\beta 2AR$. 30

FIGURE 1.6: THE DIFFERENT TYPES OF LIPID-EXPOSED SITES 38

FIGURE 1.7: THE STRUCTURE OF BPTU BOUND TO $P2Y_1R$ 54

FIGURE 1.8: THE STRUCTURE OF MK-0893 BOUND TO GCGR 60

1.3.Structure-Based Drug Design

1.3.1. Introduction to Structure-Based Drug Design

Drug discovery is the process in which a therapeutic agent is identified, engineered, and optimised with the purpose of treating a pathological phenotype in a safe, efficacious, and specific way. On a molecular level, this describes identifying the underlying cause of the pathology, and then modulating the behaviour of the causal biomolecule (often a protein) via interactions with the designed therapeutic.

Drug design is a multifactorial problem, with factors such as pharmacokinetics, affinity, and toxicity all needing to be optimised. Each of these individual factors remains challenging to solve, meaning the cost of getting a drug to market, accounting for the cost of failed schemes, was estimated to be $2.6 billion in 2016[1], with even a more conservative 2020 estimate placing the mean cost at approximately $1.3 billion [2].

There is therefore a great interest in streamlining the drug discovery process, to reduce the costs associated with drug discovery and increase the range of diseases for which we have efficacious drugs. The significant demographic shifts seen in most developed nations drives a need to develop novel drugs to allow these ageing populations to live longer, healthier lives [3]. The need for a pharmaceutical industry that can quickly produce novel therapeutic agents was highlighted by the COVID-19 pandemic, with the development of numerous vaccines by companies including Pfizer, AstraZeneca, and Moderna in collaboration with academic institutions being both a technological and regulatory success [4,5]. There has also been intense research into anti-viral compounds [6], with the COVID Moonshot project providing a new paradigm as an open-sourced, non-profit venture for drug discovery leading to a consortium

focused on pandemic preparedness [7,8]. Other motivating factors for a more efficient drug discovery process include concerns around drug resistance [9] and a desire for the 'personalisation' of medicine [10].

Once the cause of an underlying pathology has been discovered, a non-trivial process in itself [11], the challenge is to find a therapeutic that modulates its function. This is done by designing either a small-molecule or a biologic that interacts with the pathological protein. Historically, drug discovery relied on trying to screen huge libraries of drug-like molecules to identify which molecules would impact the function of the target protein [12]. Indeed, the major limitation of high throughput screening is the sheer scale of possible chemical space, which has been estimated at 10^{23} to 10^{60} molecules [13]. Current pharmaceutical libraries are typically on the order 10^{6} molecules [14], and therefore cover only a tiny fraction of this. The libraries used in high throughput screening typically contain molecules of 20 – 100 heavy atoms, although the utilisation of smaller molecules, so called 'fragments' is increasingly popular as it allows a greater exploration of chemical space[15].

Structure based-drug discovery aims to use structural knowledge to improve the efficiency of the drug discovery process. It uses knowledge about either the structure of the protein or physicochemical properties of known binders to engineer the ligand. The interactions between a protein and ligand, either covalent or non-covalent in nature, are determined by the 3-dimensional structure, dynamics, and charge distribution of the protein and ligand. This is the key underlying principle of structure-based drug discovery.

Structural information about the target is typically obtained via X-ray crystallography [16], nuclear magnetic resonance (NMR) [17], or, as is becoming increasingly common, single-

particle cryogenic electron microscopy (cryoEM) [18]. Recent advances in these fields, particularly the 'resolution revolution' in cryoEM [19] and the utilisation of X-ray crystallography for library screening [20] are driving structural biology and structure-based drug discovery.

Despite these advances in structural biology, obtaining a structure remains an expensive, time-consuming, and technically challenging problem. This is particularly true for membrane proteins or those that show significant conformational heterogeneity [21]. There has therefore been a long-standing interest in the computational chemistry field to be able to predict the structure of a protein from its amino acid sequence. Extensive research has been necessary to lay the groundwork for recent advances in protein structure prediction, including the development and maintenance of the Protein Data Bank [22] and numerous rounds of Critical Assessment of Structure of Proteins (CASP) [23]. AlphaFold2 [24] and RoseTTAFold [25] are the most popular current methods for structure prediction and show remarkable accuracy (RMSD is often < 1 Å). The utility of predicted structures for structure-based drug discovery remains an open question [26].

As discussed above, an initial hit can be obtained from a high-throughput screen. If there is an available structure, a virtual screen (VS) can be run. The most common way to do this is to use docking. Docking works by proposing a pose of the ligand in the binding site and then ranking it using either a physics based or statistical scoring function [27,28]. Docking is computationally inexpensive but is typically unable to account for the role of water, entropy, or dynamics, meaning results are often inaccurate [29]. Ligand binding often induces changes in the binding site, making docking based on apo structures less likely to be accurate [30,31]. It remains unclear how well computationally predicted structures will work for VS for similar

reasons. This is largely due to the sensitivity of a ligand's affinity to small changes in the binding site. Indeed, recent work suggests that VS on predicted structures return less accurate results than docking runs on experimental structures, and the authors attribute this to a non-holo like binding site [32].

Once a hit has been identified, it must be elaborated. This is because, as discussed previously, drug discovery is a multifactorial problem, and it is likely the hit will not have sufficient properties across all of these factors. The information encoded in the protein-ligand structure is invaluable for guiding the elaboration process. This ranges from visual inspection of the structure to better understand the binding site and allow medicinal chemists to suggest potential elaborations, to molecular dynamics-based modelling to predict the free energy of binding.

1.3.2. The role of simulation

Molecular dynamics (MD) simulations use classical mechanics to evolve the system over time. They are a theoretically rigorous way to study a system in atomistic detail. Recent years have seen significant growth in access to computational resources, both in terms of the number and quality of CPUs and GPUs. This, coupled with improvements to the underlying theory and implementation, means methods that would have been seen as too computationally expensive in the past are becoming increasingly feasible. MD remains a computationally expensive method, but these computational advances mean we are able to simulate larger systems for increasingly long times, ultimately allowing us to ask ever more complex questions.

One application of MD simulations is to better understand the dynamics of a protein. Proteins do not exist in a single conformation, but instead exist as a dynamic ensemble across a conformational landscape. These distinct conformations are key for the function of a protein, but also provide novel opportunities for drug discovery, for instance by the discovery of cryptic sites [33], understanding the mechanisms of allostery [34,35], or for designing small molecules which exhibit ligand bias [36]. The conformational landscape is often incredibly complex, and it remains challenging to identify the different conformations of a protein experimentally. Experimental methods, such as NMR, X-ray crystallography, and Förster resonance energy transfer (FRET), typically report a single average value for many observables or a distribution for a small number of variables.

MD simulations can, in theory, provide an insight into the full conformational ensemble, including information about which values occur simultaneously. The complexity of the free energy landscape of a protein means it is challenging to sample the full landscape by brute force alone. However, access to significant computational resources does allow certain insights to be gained. For instance, the *folding@home* distributed computing platform recently simulated the SARS-CoV2 proteome for a total of 0.1 seconds of simulation, identifying 50 cryptic pockets in the Spike protein [37]. A combination of unbiased and enhanced sampling methods identified a cryptic pocket at the HIV-1 protease dimerization interface that was not evident from the crystal structure of the monomer [38].

Simulation can be utilised to improve the efficacy of virtual screens. One way in which this can be done is to generate multiple conformers of the protein to use as the input to virtual screens. For instance, Kapla *et al.* [39] used MD simulations to refine models of the D_3 dopamine receptor (D_3R). They showed that virtual screens on the refined receptors led to

improved ligand binding pose prediction, with some models better distinguishing known binders and decoys. Once docking has been run, MD can be used to refine and validate the poses. This can vary from simple methods to investigate the stability of binding, to more complex methods to investigate the properties of binding. An instance of the former is work undertaken by Kacker *et al.*, where they utilised MD simulations to assess the stability of proposed docks, distinguish stable and transient hydrogen bonds, and understand the role of bridging waters [40]. Rastelli *et al.* developed a pipeline to rescore docks using MM-PB/GBSA calculations, enhancing the hit enrichment rate [41].

Simulation can be used to investigate the mechanism of ligand binding. This is starting to become plausible via unbiased simulation due to improvements in computer architecture, for instance, the development of Anton 3, which can simulate a 1 million atom system at over 100 $\mu s/day$ [42].

Previous studies have long timescale unbiased simulations to identify the binding pathway to proteins, including kinases [43] and GPCRs [44]. Access to such specialist hardware is limited, and often we wish to see multiple binding/unbinding events, particularly when we wish to understand the binding kinetics. This is particularly relevant for GPCRs, where studies indicate that k_{off} better correlates with *in vitro* efficacy than $\Delta G_{binding}$ due to the increased metabolism of molecules with low residency time [45–48]. The k_{off} for many protein ligand complexes is on the order of days, well beyond the accessible timescales of MD [49]. Thus, only a few cases exist where unbiased MD has been used successfully for simple systems [50]. Multiple techniques have been developed to try and overcome this, with particularly popular methods being based around Markov State Models [51], path sampling methods such as weighted ensemble [52], and biased simulation methods [53]. These will not be discussed further

here (for more information, please see Ahmad *et al.* [54]), but a key takeaway is that the estimation of binding kinetics remains computationally expensive and challenging for most systems.

1.3.3. Free energy calculations

Despite a growing interest in kinetics, thermodynamic properties, particularly the free energy of binding, remain central in structure-based drug discovery. The Gibbs free energy change (ΔG) is related to the enthalpic change (ΔH) and entropic change (ΔS):

$$\Delta G = \Delta H - T\Delta S \qquad \textit{Equation 1}$$

The enthalpic change is a balance of the new interactions formed between the protein/ligand and those lost between the protein/water and ligand/water. The entropic change is driven by the removal of ordered waters around the protein.

Simulation is essential for understanding the thermodynamics of drug binding because molecular recognition is inherently a dynamic process. Factors such as conformational changes of the protein and ligand, entropy-enthalpy compensation, and solvent effects must be considered when predicting $\Delta G_{binding}$ [55,56].

Docking scoring functions, the cheapest methods for predicting $\Delta G_{binding}$, consider only a single pose and therefore do not explicitly account for these factors. This is an underlying cause for the continued poor performance of scoring functions, and a reason for the continued interest in machine-learning based scoring functions over physics-based scoring functions

[57,58]. Despite the fact that a vast number of deep learning architectures have been developed for scoring functions and that many of these publications report Pearson correlation coefficients in the range of 0.80–0.85, the typical performance on external test sets is far worse [59]. The inability of docking scoring functions to generalise highlights the need to validate the results of virtual screens with more rigorous and generalisable methods.

Simulation based methods are theoretically rigorous, inherently incorporate dynamic factors, and are generalisable. Numerous methods exist, varying in computational expense and accuracy. Endpoint methods involve the post-processing of explicit solvent simulations using implicit solvent. A popular method is molecular mechanics with Poisson-Boltzmann or generalised Born surface area (MM-PB/GBSA) [60]. MM-PB/GBSA is of intermediate expense and accuracy and is therefore often used as an initial validation of virtual screens [61]. The use of implicit solvent and the need to estimate the entropic contribution means that MM-PB/GBSA often performs worse than more expensive methods [62]. Steered or pulling method approaches apply a force to a ligand to induce unbinding. Harmonically restrained umbrella sampling windows can then be run along this physical pathway to compute a potential of mean force [63–66].

Alchemical methods for free energy prediction utilise the fact that free energy is a state function and therefore is independent of the path. We can therefore define a non-physical pathway between the two endpoints. Defining a non-physical pathway allows us to avoid defining a physical pathway between the two endpoints, which is complex, unknown, and hard to sample.

The non-physical (or 'alchemical') pathway involves decoupling or annihilating the ligand – that is transforming the ligand from 'interacting' with its environment to 'non-interacting'. This is done by introducing a coupling parameter λ. λ is used to modify the Hamiltonian of the system and indicates your position on the cycle, with $\lambda = 0$ indicating one endpoint and $\lambda = 1$ indicating the other. Intermediate λ values represent a partially interacting ligand and we are therefore linking two physical endpoints with a series of non-physical intermediates. In equilibrium simulations, discrete λ values are picked and short simulations ('windows') are run. The free energy between the two endpoints can then be calculated using a variety of estimators.

Alchemical free energy methods can be broadly split into relative binding free energy calculations (RBFEs), which calculate $\Delta\Delta G_{binding}$, and absolute binding free energy calculations (ABFEs), which calculate $\Delta G_{binding}$.

RBFEs calculate $\Delta\Delta G_{binding}^{A \to B}$. $\Delta\Delta G_{binding}^{A \to B}$ is the free energy difference between two ligands – ligand A and ligand B. RBFEs alchemically transform ligand A into ligand B (Figure 1.1). As such, RBFEs are reliant on the two molecules being compared being highly similar (although recent work by Baumann *et al.* on separated topologies looks to overcome this [67]).

ABFEs work by removing the molecule's interactions with the environment (decoupling) or removing the molecule's inter *and* intra-molecular interactions (annihilating) (Figure 1.2). This is done in solution and in complex, with the difference giving $\Delta G_{binding}^{\theta}$. As the entire molecule is being annihilated/decoupled, the weakly interacting molecule may leave the binding site, making sampling significantly more challenging.

Unlike RBFEs, ABFEs require the use of restraints. The most commonly used restraints were proposed by Boresch *et al.* and are often referred to as Boresch restraints [68]. They essentially restrain the ligand in a particular orientation in the binding site. This is achieved via the introduction of a distance restraint, 2 angular restraints, and 3 dihedral restraints. These restraints all work intermolecularly between the ligand and the protein, maintaining the ligand in the binding site. The energetic impact of these restraints must be accounted for via simulation for the complex leg but can be calculated analytically in solution.

As ABFEs and RBFEs calculate different properties, they typically occupy different niches in the drug discovery process, with ABFEs offering promise for hit validation and RBFEs used for hit optimisation. However, there are situations where RBFEs are not appropriate for ranking ligands – for instance in the case where the binding mode may change (i.e. fragments), linker elaboration, or situations where rings are opened/closed [69]. In these situations, it is clear that the $\Delta\Delta G_{binding}^{A\rightarrow B}$ can be calculated from $\Delta G_{binding}^{B} - \Delta G_{binding}^{A}$. RBFEs are typically considered to be easier to run than ABFE calculations as they do not require restraints, there is a cancellation of errors, and there is a smaller alchemical region. Consequently, RBFEs are already commonly used in computational pipelines in pharmaceutical companies and as a research tool in academia [70–72], with various automated workflows making implementation relatively easy [73].

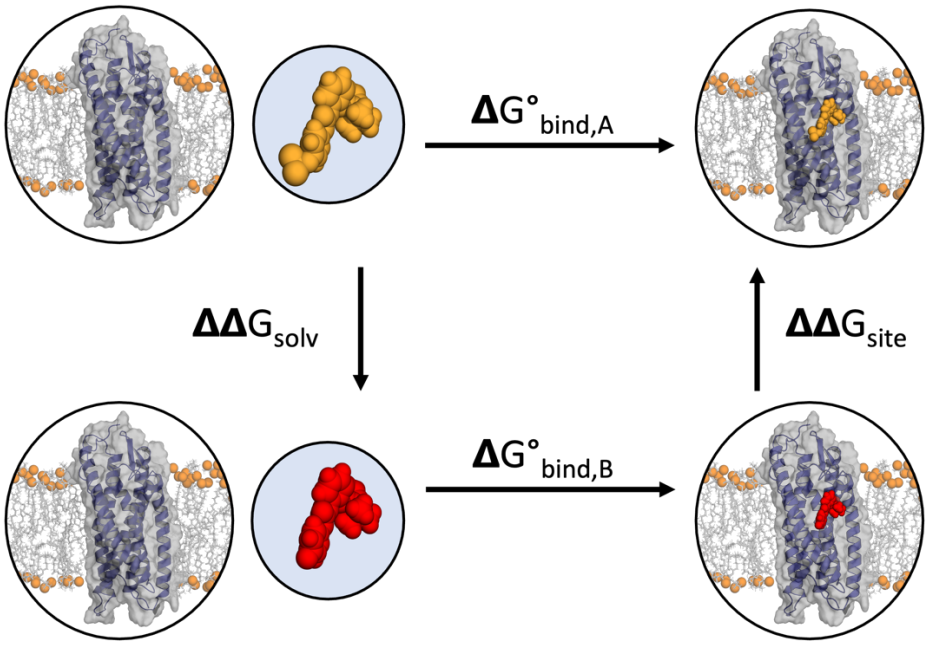


Figure 1.1: Alchemical RBFE Cycle. The alchemical cycle for calculating the relative binding free energy of binding ($\Delta\Delta G_{binding}$) of two ligands. Ligand A (yellow) is alchemically modified into Ligand B (red) in solution (lefthand side of cycle) and bound to the receptor (righthand side of cycle). The difference in free energy of these two legs gives $\Delta\Delta G_{binding}$.

Absolute alchemical binding free energy (ABFE) calculations are typically considered the most accurate way of calculating $\Delta G_{binding}$ as they have been successfully applied across multiple systems.

Aldeghi *et al.* applied ABFE calculations to investigate the binding of a series of chemically diverse small molecules to bromodomain-containing protein 4 (BRD4) [74]. Using the ligand bound crystal structures as a starting point they achieved a mean absolute error (MAE) of 0.6 kcal/mol with a Pearson's correlation coefficient (r_P) of 0.84 ± 0.05. Remarkably, they repeated the process in a pseudo-prospective manner by generating starting poses using docking and obtained a MAE of 1.0 kcal/mol and r of 0.77 ± 0.04. More recent work has shown that for three targets (BACE, CDK2, and thrombin) ABFEs can be used to improve

the hit enrichment rate from a set of known binders and decoys better than docking alone [75] and that ABFEs can be used to guide fragment elaboration [76].

One common way to improve the performance of ABFEs is Hamiltonian replica exchange (HREX). HREX involves exchanging Hamiltonians between adjacent windows, with the goal of improving sampling. Several papers report the improvement of performance of ABFEs when using HREX[74,77], although it is not clear that it provides benefits in all cases and it also requires simulations to be run simultaneously which may not be applicable depending on your available hardware.

However, despite these success cases and technical advances, the blind SAMPL challenges indicate that ABFEs still remain challenging [78]. This raises the question – what accuracy is required for ABFEs to be useful in a typical drug discovery program? This was discussed by Mobley *et al.* who theorised that calculations with a root mean square error (RMSE) of 0.2 kcal/mol and a throughput on the order of 10s of molecules a week would reduce the amount of synthesis required by a factor of 3 [79]. ABFEs have typically only been applied retrospectively to soluble proteins and it remains an ongoing to challenge to identify when ABFEs would be sufficiently accurate to reach the threshold for usefulness or how to identify when they have not achieved accurate results.

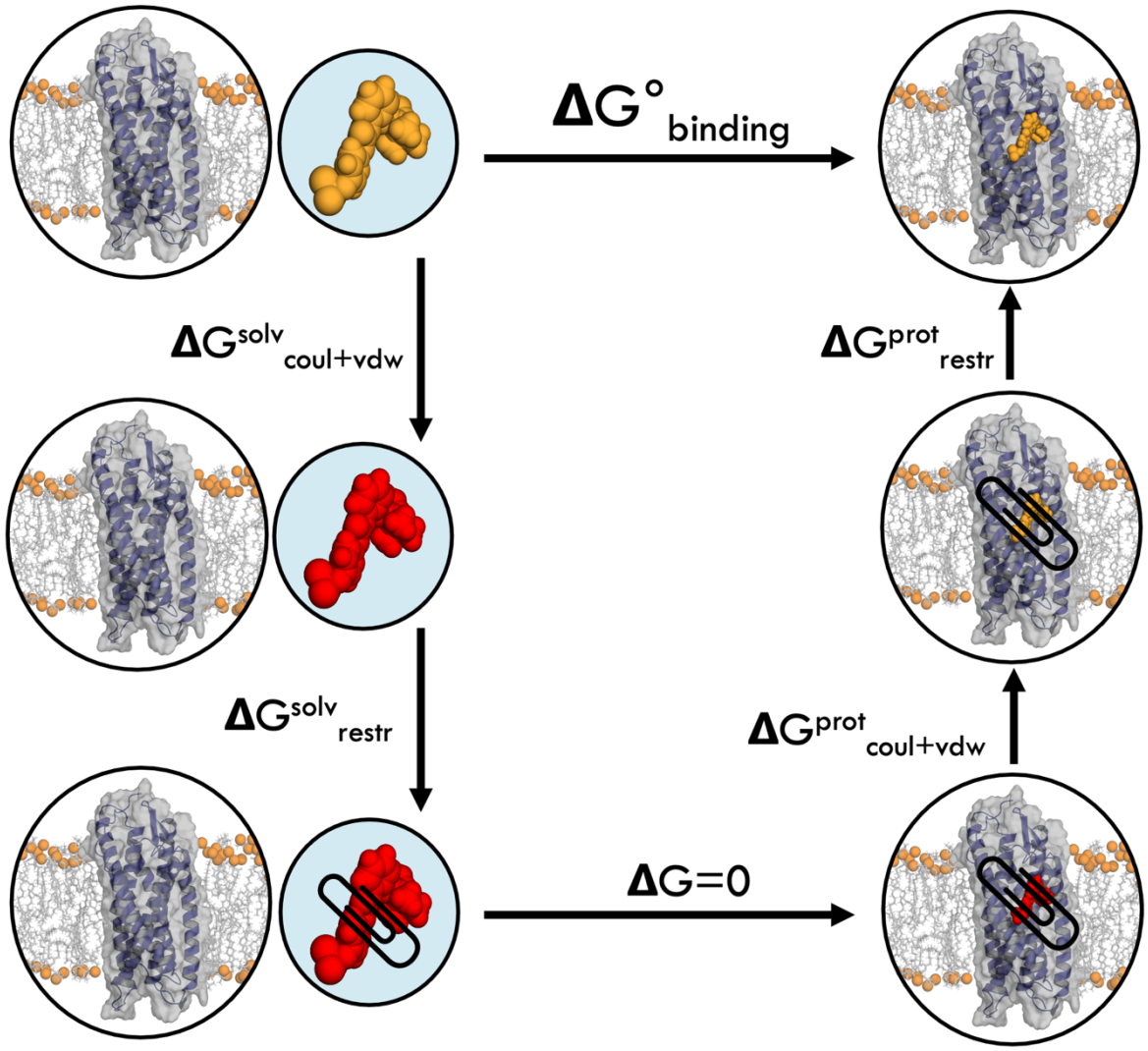


Figure 1.2: Alchemical ABFE Cycle. The alchemical cycle for calculating the free energy of binding - $\Delta G^{\theta}_{binding}$. The interacting ligand (yellow) is annihilated or decoupled to give a non-interacting ligand (red). This is done in solution and in complex, with the difference giving $\Delta G^{\theta}_{binding}$. The application of restraints is necessary to stop the weakly interacting ligand leaving the binding site. Restraints consisting of a single distance, two angles, and three dihedrals, proposed by Boresch et al., were used in the simulations in this thesis unless stated otherwise [68]. The energetic effect of the restraints is calculated from simulations in complex and analytically in solution.

The work described above is based on equilibrium ABFEs. Equilibrium ABFEs remain the most commonly used method, and benefit from extensive research to validate them, understand their sources of error, and make their implementation easier [80,81]. However, of growing popularity are non-equilibrium ABFEs.

Non-equilibrium ABFEs run two equilibrium simulations – one corresponding to the holo state, the other to the apo state. Snapshots from these equilibrium simulations are then used to seed non-equilibrium simulations (Figure 1.3). These non-equilibrium simulations drive the ligand from interacting to non-interacting, and vice-versa. Exponential averaging of the non-equilibrium work is carried out based on the Jarzynski equality (for uni-directional non-equilibrium simulations) [82,83] or Crooks fluctuation theorem (for bi-directional non-equilibrium simulations) [84].

The main proposed advantage of non-equilibrium free energy calculations is the ability to increase the sampling time at the endpoints, which are ultimately what we are interested in. The Jarzynski equality, applied when running non-equilibrium simulations in only a single direction, is sensitive to the tails of the work distribution due to the exponential averaging, which are often poorly sampled. This is why running in two directions and using the Crooks fluctuation theorem is preferable. The Crooks fluctuation theorem is reliant on overlap between the forward and reverse work distributions for good estimations of the free energy. In the case of poor overlap, the free energy prediction will likely be incorrect. However, despite growing interest in non-equilibrium ABFEs, they remain less popular than equilibrium ABFEs, largely due to issues around getting overlap between the forward and reverse simulations. It is not clear how this method will behave in more complex systems and in dealing with larger transitions.

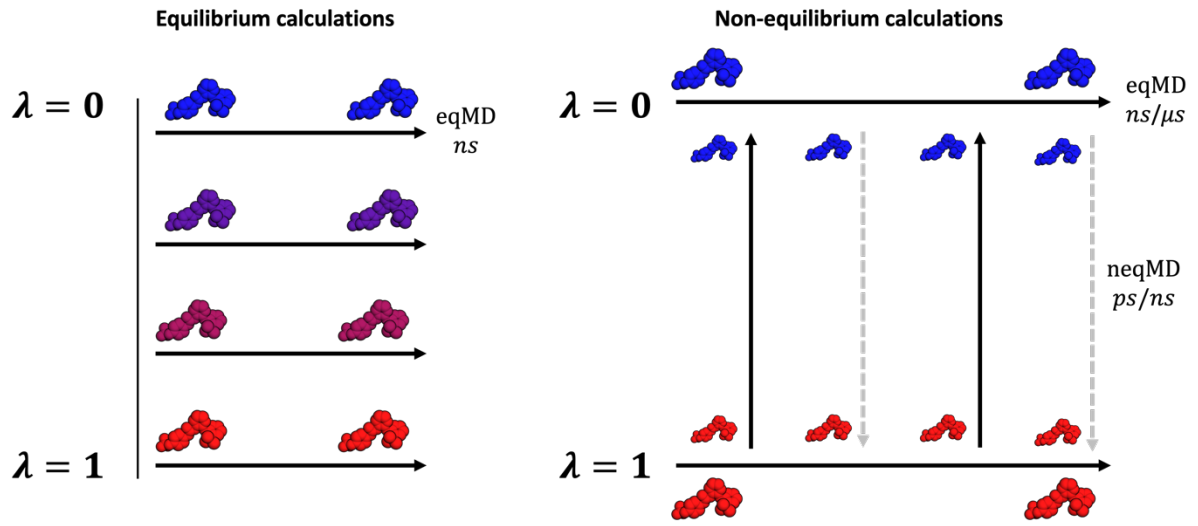


Figure 1.3: Equilibrium vs Non-Equilibrium Alchemical Calculations. Equilibrium ABFEs involve multiple equilibrium simulations at different values of λ on the order of 10s of ns. Non-equilibrium calculations involve two long equilibrium simulations at $\lambda = 0$ and $\lambda = 1$. Snapshots are extracted from these simulations and non-equilibrium simulations (typically on the order of ps or a few ns) are used to drive the system from $\lambda = 0 \rightarrow \lambda = 1$, and vice-versa.

One of the fundamental challenges of ABFEs is the requirement to properly sample the apo state of the protein. Whether one subscribes to the induced fit or conformational selection theory of ligand binding, it is clear that the major conformation of the apo protein often differs from the crystallographically solved holo structure. This introduces a problem as these changes often happen on timescales that are too long to be sampled in free energy calculations. There is therefore an overestimation of the magnitude of the free energy of binding, as the energetic cost of transitioning from the holo to apo state is not considered. This can be quite substantial, even for subtle conformational changes. Mobley *et al.* showed that missing the rearrangement of a valine between the apo and holo structures in T4 lysozyme can lead to an error of several kcal/mol [85]. Khalak *et al.* showed that a threonine rearrangement in $p38\alpha$ not being sampled leads to an error of 9 kJ/mol [86]. Chen *et al.* attributed a 10 kcal/mol error they observed in their predictions for PTP1B down to a structural rearrangement of the WPD-loop not being sampled [87].

Failing to sample the apo state properly clearly causes significant errors in the absolute estimation across a large number of systems. One can only expect this problem to be far more significant for more dynamic systems, such as GPCRs, many of which are of major pharmacological interest.

The energetics of any change in conformation will likely differ between proteins, making the use of ABFEs to explain selectivity challenging. The one study to use ABFEs to do this (Aldeghi *et al.* looking at the promiscuity of bromosporine [88]) showed moderate accuracy. This could potentially be explained by the different energetic penalties from poor sampling of the apo states. The energetics of conformational changes will also differ for different sites within a protein, thus making it more difficult to use ABFEs to identify the correct binding site on a protein.

Despite the prevalence of this problem , there is no generally accepted solution. Khalak *et al.* propose a method using non-equilibrium alchemical free energy calculations [86]. Essentially, they run the equilibrium simulation for the apo-state from an apo crystal structure or the endpoint of a long simulation. For the $p38\alpha$ system they see a non-reversible transition in the threonine rotamer and improved free energy results seeding the non-equilibrium simulations from these structures. However, it is not clear how well work distributions will overlap in the cases of larger transitions.

Therefore, an approach based on the more popular equilibrium ABFEs would be of great use. Previous work has looked at combing Monte Carlo simulations with molecular dynamics simulations, with notable examples being *BLUES* [89] from the Mobley lab and *grand* [90] from

the Essex lab. These Monte Carlo based methods work by proposing moves which are accepted or rejected based on a Metropolis-Hastings criterium. These methods have been shown to improve sampling of side chain movements and buried waters [91–93]. These methods are designed to try and negate the major challenge in Monte Carlo simulations – the acceptance rate. Proposed moves typically introduce steric clashes and are therefore often rejected. BLUES, for example, uses non-equilibrium candidate Monte Carlo to relax the system after the proposed move, thus increasing the acceptance rate.

These Monte Carlo based methods remain computationally expensive and often require the use of a specific ensemble (for instance, the Grand method [90] requires the use of the Grand Canonical ensemble). The acceptance rate remains a challenge and is likely to be far worse for systems where the moves are being proposed in a lipid bilayer, due to the fact lipids are larger than water molecules and exhibit slower relaxation. It is also not clear whether Monte Carlo would be sufficient to sample large changes in the conformation of a protein, or concerted sidechain movements.

Recent work has tried to assess how useful AlphaFold2 and RoseTTAFold predicted structures will be for free energy calculations. Work by Beuming *et al.* looked at using predicted structures to run relative binding free energy calculations (RBFEs) for a series of proteins [94]. To mimic more realistic conditions of a drug discovery project, sequences with more than 30% sequence identity were removed. Their results were mixed, with some predicted structures showing comparable results to experimental, and others showing significantly worse results.

1.4. Membrane Proteins

1.4.1. Introduction to Membrane Proteins

Membrane proteins are key drug targets because of their diverse functions and central role in propagating signals from the exterior of the cell to the interior. A few of the functions include transport (both of ions and larger molecules), structural regulation of the membrane, and cellular recognition. Membrane proteins play key roles in many of the key processes of life including respiration, photosynthesis, and the immune system.

Whilst peripheral membrane proteins associate with the exterior of the membrane, integral membrane proteins embed themselves into the bilayer. The transmembrane region is typically either α-helical or forms a β-barrel. It is estimated that for most species 20-30% of proteins are α-helical membrane proteins [95]. The residues in the centre of the bilayer are typically hydrophobic as they interact with the non-polar tails of phospholipids, whilst there is more scope for polar and charged residues in the headgroup and water facing regions of the structure. Consideration of the amino acid composition of different regions of both the membrane and membrane proteins is key for drug discovery as the environment will change the properties of the small molecules that are likely to bind there. Arginine and lysine, when positioned towards the termini of transmembrane helices, are able to 'snorkel' - that is orient their positive charges to the water interface region and pull waters into the hydrocarbon core. The complex and dynamic interactions of membrane proteins, lipids, and water are one of the many challenges of modelling membrane proteins.

Membrane proteins provide a gateway between the exterior and interior of the cell and cellular compartments. There are three major classes of membrane proteins: G-protein coupled receptors (GPCRs), transporters, and ion channels. They propagate signals, allow the

passive diffusion of ions, or transport small organic molecules. Their balanced function is necessary for the proper function of the cell in general, and therefore they are tightly regulated by a range of stimuli including voltage, ligands, and solute concentration gradients.

The dysfunction of membrane proteins is therefore a common cause of disease, with membrane proteins being linked with all of the most common cancers [096], cardiovascular disease [97], and rare diseases such as cystic fibrosis [98], to name but a few. These diseases can be caused by misfolding [99,100], incorrect trafficking [101], loss of enzymatic function [102], and incorrect protein-protein [103] or protein-lipid [104] interactions. Membrane proteins also play a key role in pathogen life cycles [105–107], the development of drug resistance [108], and autoimmune diseases [109,110]. This underlying role in pathology is why membrane proteins are such a key target for drug discovery programs. This is reflected in the fact that despite accounting for approximately 20% of the human proteome, proteins located at the cell surface account for 60% of drug targets [111].

One of the major challenges of structure-based drug discovery campaigns is the difficulty of obtaining high quality structures of membrane proteins. Historically, X-ray crystallography was at the forefront of elucidating membrane structures. Recent advances in the expression, purification, addition of mutations and insertions for solubility and stability, the automation of procedures, and advances to synchrotron facilities have driven the increase in the number of membrane protein structures being solved. However, the generation of crystals remains challenging, and the entire process is highly dependent on the use of the correct detergents [112]. Recent advances in cryo-electron microscopy (cryo-EM) are driving a surge in the number of available membrane protein structures. As of 15/10/2023, there are 3,220 published structures of 1,640 unique integral membrane proteins in the RCSB PDB, with just

900 unique proteins represented in 2018 [113]. This revolution in membrane protein structural biology opens up a plethora of opportunities for structure-based drug discovery.

Key families of membrane proteins (including transporters, ion channels, and G-protein coupled receptors (GPCRs)) are best considered as inherently dynamic entities. The different structures that membrane proteins can take are essential for their function. The structural challenges of working with membrane proteins means that MD simulations are a natural tool for understanding the conformational landscape [114].

Simulations have been widely used to investigate ion channels, for example, to understand the gating mechanism[115], transfer rates [116], and the role of mutations in causing disease [117]. Transporters undergo significant conformational changes to facilitate the transport of molecules across a membrane. The scale of this transformation means that steered MD, a commonly used method for identifying transition paths, displays significant hysteresis, that is there is a difference in the computed values of $A \rightarrow B$ vs $B \rightarrow A$. Recent work has tried to overcome this by generating realistic, history independent intermediates that can be used to seed umbrella sampling windows [118].

1.4.2. Membrane Structure and Dynamics

1.4.2.1. Lipid Composition

Membranes play a key role in the function of cells, acting as a barrier between the interior and exterior of cells and organelles. These membranes are formed of a bilayer of lipids, the most abundant being phospholipids. Phospholipids consist of long acyl tails and polar/charged headgroups. These amphiphilic molecules spontaneously aggregate into a

bilayer, burying the hydrophobic tails and leaving the headgroups exposed to solvent. The width of the hydrophobic region of a bilayer is typically on the order of 30 – 40 Å.

Membranes are a complex and dynamic mixture of gylcerophospholipids, sphingolipids, sterols, and membrane proteins. The exact composition of each membrane depends on the cell type and the cellular location of the membrane to allow the membrane to perform its specific function [119–122]. The variable composition leads to significant varieties in the structure and biophysical properties of the bilayer, with plasma membranes being anionic. For instance, cholesterol changes the fluidity and thickness of the bilayer [123], whilst the extreme curvature seen in cristae has been proposed to be caused by the specific lipid composition of the inner mitochondrial bilayer and the dimerization of ATP synthase [124].

The composition of individual leaflets also varies. For instance, the extracellular leaflet of the plasma membrane is enriched with phosphatidylcholine, sphingolipids (such as sphingomyelin), and glycosphingolipids. The intracellular leaflet is enriched in phosphatidylethanolamine, phosphatidylserine, and phosphatidylinositol [121]. It is also worth noting that lipid distribution is not consistent within a leaflet, with lipids often being localised in 'rafts' [125].

The complexity of membranes and slow dynamics of lipids makes it highly challenging to computationally model. As a result of this, it is typical to use a bilayer composed of a single lipid type when running atomistic simulations. When it is believed that this simplifying assumption is not applicable for the given system, coarse-grained simulations provide a way to sample the dynamics of a complex bilayer [126,127]. The curvature of membranes is often

typically ignored, but there is a growing interest in incorporating curvature into MD simulations [128].

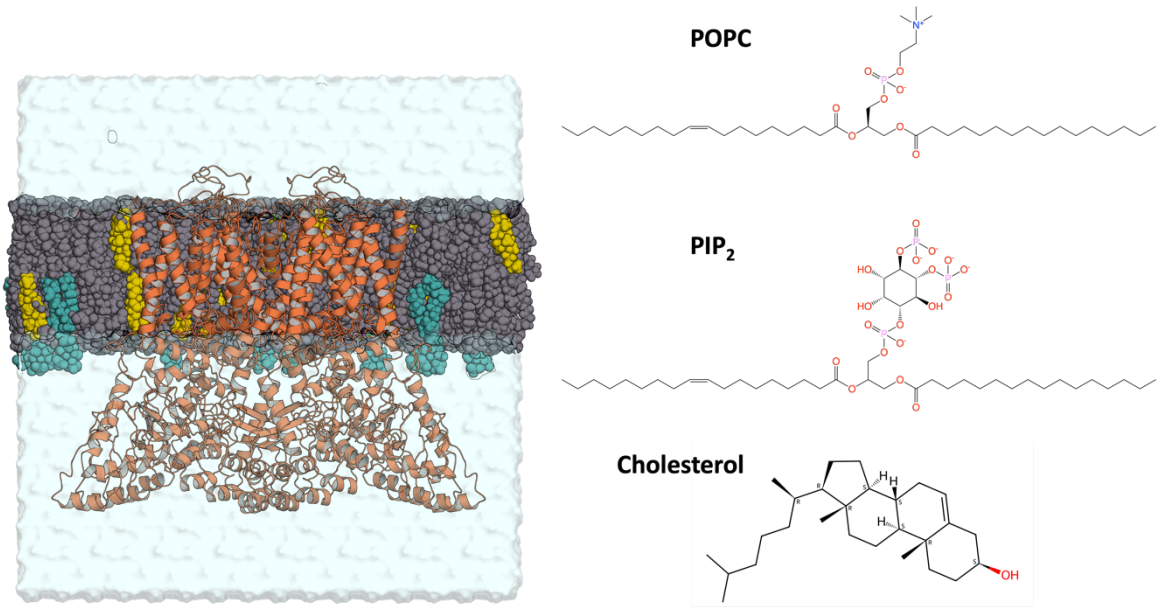


Figure 1.4: A protein embedded in a POPC:PIP_2:Cholesterol bilayer. The TRPV6 channel (PDB: 7K4A) embedded in a bilayer containing POPC (grey), cholesterol (yellow), and PIP_2 (teal) in a ratio of 10:2:1. The structures of the three key lipids used in this thesis – POPC, $PI(4,5)P_2$, and cholesterol. $PI(4,5)P_2$ is shown carrying a charge of -5e.

1.4.2.2. Protein-Lipid Interactions

Membrane proteins form specific interactions with the lipid components of the bilayer. These lipids often have a key functional role in regulating the function of the membrane protein. This can include the stabilisation of specific conformations [129] or the stabilisation of protein-protein interactions [130]. These lipids exert their effects allosterically.

Bound lipids are increasingly observed in experimental structures of membrane proteins [131,132]. Interpretation of structural cholesterols must be done carefully as the exact lipid species is often difficult to identify and the bound lipids observed may not reflect the

interactions seen in the native bilayer environment [133]. Purification can remove bound lipids and the presence of a lipid does not necessarily imply structural or functional significance. Numerous computational studies have looked at using coarse-grained simulations to identify lipid interactions with membrane proteins [134–136], identify the species of ambiguous lipid-like densities [137], and calculate the free energy of binding [138]. Lipids modulate the function of all the major classes of membrane proteins. This shows that it is possible to allosterically modulate the function of a protein via molecules that bind within the bilayer in a potentially near ubiquitous fashion.

1.4.3. GPCRs

There are over 800 G-protein coupled receptors (GPCRs), making them the largest family of proteins in the human genome. GPCRs are broadly split into six classes – class A, class B, class C, class D, class E, and class F. Mammalian GPCRs are often categorised by the GRAFS system (in which class A is **R**hodopsin, class B is split into **S**ecretin and **A**dhesion, class C is **G**lutamate, and class F is **F**razzled).

The classical role of GPCRs is to transduce extra-cellular signals across the cell membrane causing the recruitment and activation of a cytoplasmic G-protein that then activates downstream effector proteins. GPCRs respond to a huge number of signals including photons, ions, small molecules (including neurotransmitters), and peptides.

GPCR's central role in many signalling pathways makes them attractive drug targets for a broad spectrum of diseases. GPCRs make up 50% of drug targets [139]. As of 2021, there are 527 FDA approved drugs that target GPCRs, and 60 further candidates in clinical trials [140]. The vast majority of these drugs target class A GPCRs. As of 2022, FDA approved drugs

targeted only ~110 unique GPCRs, meaning there is great therapeutic scope in targeting the remaining GPCRs [140–142]. Research into GPCRs is extensive and broad, with some key areas of interest being improving our ability to solve the structure of GPCRs [143], understanding the dynamics of GPCRs (particularly with regard to activation [144], allosteric modulation [145], and biased signalling [146]), identifying GPCRs with unknown physiological agonists and functions (so called 'orphan' receptors) [147], and improving specificity of drug binding [148–150]. Advances by Kobilka and Lefkowitz into the structure and function of GPCRs led to them receiving the 2012 Nobel Prize in Chemistry [151].

1.4.3.1. GPCR Structure, Function, and Activation

GPCRs are integral membrane proteins and have 7 transmembrane α-helices with an extracellular N-terminus and intra-cellular C-terminus. The most simplistic description of the function of GPCRs is that the endogenous ligand binds to the orthosteric site within the 7-TM bundle on the extracellular side of the receptor. This leads to a conformational change on the intracellular side of the protein, causing association with and activation of a trimeric G-protein by the exchange of GDP for GTP. This then causes a signalling cascade that ultimately leads to some change in the physiology of the cell.

This section will particularly focus on the structure and activation of GPCRs, with the downstream signalling largely being beyond the scope of this thesis. I will cover the early stages of the structural elucidation of GPCRs, the current research in understanding the dynamics of GPCRs, and the emerging concepts in GPCR structural biology. Throughout this thesis, I will utilise the Ballesteros-Weinstein notation for GPCRs, where the most conserved residue in a TM helix is assigned the number 50 and other residues are numbered relative to this [152].

Class A GPCRs are the most well studied, with a particular focus being placed on the β_2 adrenergic receptor ($\beta_2 AR$). GPCRs proved difficult to solve structurally because of their conformational flexibility and large unstructured loops, as well as the other general problems associated with the crystallisation of membrane proteins.

The structural work around $\beta_2 AR$ by the Kobilka lab utilised a wide range of the latest crystallographic technologies, including stabilising mutations, nanobodies, state-of-the-art beamlines, novel lipids, and fusion proteins. The first structure of $\beta_2 AR$ was solved in an inactive state with a high-affinity inverse agonist in 2007 [153], with a nanobody bound active state following in 2011 [154] , and a G_s-bound active state also in 2011 [155].

These structures showed relatively small differences in the orthosteric site between active and inactive structures, with the main difference being a 2 Å inward movement of TM5 around S207 (Figure Figure 1.5). The changes on the intracellular side are drastic, with TM6 moving 14 Å outwards upon agonist binding. These structures also provided insight into how the activation process occurs. The inactive conformation is stabilised by the packing of P211 in TM5, F282 in TM6, I121 in TM3, and N318 in TM7. This packing cannot be sustained following the movement of TM5. There is therefore a rearrangement to facilitate the agonist bound state. This rearrangement involves the rotation of TM6 around F282, causing the large outwards movement of the cytoplasmic end of TM6 [156]. This large movement of TM6 has become the hallmark of GPCR activation [142,157–159].

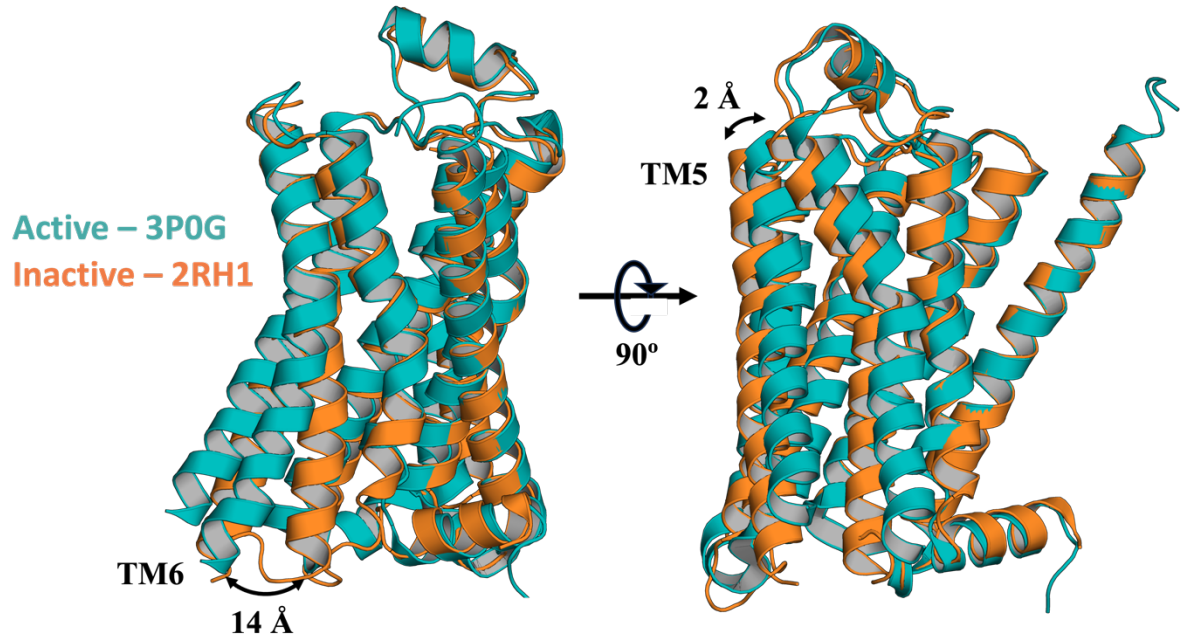


Figure 1.5: Active and inactive structures of β_2AR. Structure of β_2AR in the inactive conformation (orange, PDB 2RH1) and active conformation (teal, PDB 3P0G). The canonical activation movement of TM6 outwards is shown, as is the movement of TM5 on the extracellular side of the receptor.

The structural rearrangement around F282 (or in BW notation $F^{6.44}$) is highly conserved and is part of the PIF motif [160]. The PIF motif shows coupled movement with the CWxP motif ($C^{6.47}$, $W^{6.48}$, and $P^{6.50}$) [161]. The combination of these two motifs is key for linking conformational changes in the binding site to the cytoplasmic side of the receptor and is sometimes termed the 'transmission switch' [162]. The 'ionic lock' at the intracellular end of TM3 and TM6 is another important microswitch that stabilises the inactive state [163]. Mutations to these switches causes the stabilisation of different conformations of the receptor and changed function of the receptor [164,165].

The activation process is, however, more complicated than this. First, GPCRs show a baseline level of GTP exchange activity in the absence of agonist. This indicates the active conformation exists in the equilibrium population of the receptor in the absence of agonist.

Agonists increase the frequency of this active conformation and therefore the level of GTP exchange. Second, the simple picture of agonists stabilising an active conformation is not fully correct. Agonist binding actually appears to increase the dynamic properties of TM6 [166]. It is only upon binding of a G-protein (or something mimicking one) that the active state is stabilised [167]. The binding of the G-protein also increases the affinity of the small molecule for the receptor [168].

The signalling of GPCRs is also complicated by the fact that they also signal through arrestins. Arrestin recruitment was initially described upon the phosphorylation of the C-terminal tail of the receptor. This recruitment blocked the association with G-proteins and promoted receptor internalisation [169]. However, it has become clear that some ligands can cause arrestin binding and subsequently activate distinct signalling pathways from G-protein mediated signalling [170,171]. Different ligands can also induce interactions with Gα-protein subtypes [172].

These signalling phenomena arise as a function of the complex dynamics of GPCRs and their dynamic nature is the cause of some of the difficulties in drug discovery for GPCRs, but also many of the emerging opportunities. There are a number of experimental techniques applied to the study of GPCR dynamics. NMR studies for instance showed that the $\beta_2 AR$ exists as multiple functionally important conformations [173], the mechanisms of receptor activation [174,175], ligand specific dynamics [176], and to understand biased signalling [177]. Manglik *et al.* [178] showed that the inactive form of $\beta_2 AR$ exists in at least two distinct conformations. These conformations exchange on the timescale of microseconds to milliseconds. These NMR studies were supplemented with double electron-electron resonance (DEER [179]) studies that showed these conformations corresponded to the presence or absence of the ionic lock.

NMR studies have also been used to look at the mode of action of different types of inhibitors. Two such studies looked at the impact of inverse agonists and agonists on the conformational population and dynamics for β_2AR [180] and A_2AR [181]. For the A_2AR, NMR in the absence of ligand identified four minima – two inactive, one active, and one intermediate. Upon the addition of an inverse agonist, the relative populations of the inactive states increased, and the active states decreased. This indicates the inverse agonist lowers the free energy of the inactive states relative to the active state. For β_2AR, the addition of saturating levels of agonist leads to the sampling of a third conformational state and the reduction of the populations of the 2 inactive states. However, the addition of agonist alone means that the inactive states still account for 40-60% of the population. It is only upon addition of a G-protein mimetic nanobody that a fourth state, corresponding to the fully active state seen in the crystal structure, is seen.

It is worth noting that this creates a challenge for calculating the binding affinity of agonists. The difference in occupation of the active and inactive ensembles indicates that, in the absence of ligand, the free energy of the inactive state is lower. There is therefore an energetic penalty of switching from the inactive to active state that the binding of the small molecule compensates for. If one fails to account for this, the magnitude of the free energy of agonist binding will be overpredicted.

NMR has also been used to probe the structural causes of partial agonism and ligand bias. Ligand bias has been studied by multiple experimental techniques recently as interest in its therapeutic potential grows. The studies suggest that ligand bias is caused by stabilising different conformations to balanced agonists. These studies suggest a greater structural

change in TM7 in arrestin-biased drugs, leading to a more occluded intracellular G-protein binding pocket [182,183]. G-protein biased ligands on the other hand favour the canonical active state more than balanced agonists [184]. However, these results are often contradictory – the mechanism of partial agonism for instance is proposed to be due to stabilisation of an intermediate state [185,186] and also by only partially shifting the equilibrium population [187–189].

These experiments highlight the key role of dynamics in understanding the behaviour of GPCRs but the presence of contradictory results between them highlights their limitations. Research has shown that the type of atom used as the label, the labelling site [190], and the detergent environment all impact the quality of results [191]. These experiments also utilised non-WT sequences which can greatly impact the dynamics of proteins [192,193].

1.4.3.2. Simulation of GPCRs

These experimental methods have been extensively supplemented by MD simulations. Whereas experiments such as DEER and FRET provide only information about the distance of two probes, MD simulations provide an all-atomistic picture.

Of particular interest has been the activation/deactivation process. Early work on this utilised long unbiased MD simulations to look at the deactivation of $\beta_2 AR$ [194]. Active state structures were stripped of the G-protein / G-protein mimetic and tens of repeats were run for between 2 and 50 μs each. These simulations identified intermediate states for the ligand binding site, the G-protein site, and the orthosteric site, but could not characterise the free energy profile for the entire process. Later work studied the activation process of the same protein using 2.15 ms of simulation time and Markov State Models (MSM) [195]. The simulation time consisted of 10,000 – 20,000 short simulations and identified several activation pathways and

intermediate states. The MSM showed that ligands can be rationalised as agonists or inverse agonists by how they alter the kinetics of the interconversion of states.

The sheer amount of simulation time utilised by these studies highlights the need for enhanced sampling methods to better probe these systems. The activation of the adenosine A2A receptor was studied using metadynamics [196]. Two collective variables (CVs) described the rotameric state of the toggle switch residue. The transition of this rotamer is import for activation. Agonist binding increased the free energy difference and lowered the energy barrier between the active and inactive rotameric states. The change in the rotamer led to an opening of the G-protein state. Other work has used adiabatic biasing MD simulations (that is where a time-dependent bias is applied to drive the system from state A to state B) [197,198], replica exchange with solute tempering (REST2) [199], string with swarms MD [200], and adaptively-biased MD [201].

MD simulations can also be used to better understand ligand binding. The atomistic level of detail MD simulations provide is vital for understanding ligand binding but due to the timescales required, enhanced sampling methods are often used. Typical examples are the identification of binding modes/sites, understanding the allosteric effect of ligand binding, or to study the binding/unbinding pathways.

Fleetwood *et al.* used the string of swarms MD to study 6 $\beta_2 AR$ with different efficacies [202]. They saw a stabilisation of inactive and active states for antagonists and agonists respectively, but also show that each ligand stabilises a different conformation and that there is a strong correlation between the states of the microswitches and experimentally measured cAMP responses. Soldner *et al.* used metadynamics to identify and cluster ligand binding modes to

three GPCR systems and correctly reproduced the experimental binding mode [203]. Xu *et al.* used funnel metadynamics to investigate the 10-fold selectivity of norepinephrine to B_1AR over β_2AR [204]. Structural studies had indicated the binding sites for β_1AR and β_2AR are identical. Metadynamics showed that the different properties of the extracellular vestibule led to different binding pathways, which effected the k_{on} and affinities. De Waal *et al.* used enhanced sampling methods to understand the undesirable arrestin-biased binding of fentanyl to the μ-opiod receptor and design G-protein biased small molecules [205].

1.5. Lipid Exposed Binding Sites as Novel Drug Targets

Drug discovery remains challenging for GPCRs. This is for numerous reasons, but two of the more fundamental problems centre around the challenge of achieving specificity, both for the receptor of interest and the specific pathway of interest.

GPCR families are further divided into subclasses. It is often desirable to specifically target one of these subtypes as targeting of related proteins can lead to severe side effects. For instance, $\beta_2 AR$ is a target for asthma and pulmonary diseases, whilst off-target effects at $\beta_1 AR$ lead to serious cardiac side effects. This led to the development of $\beta_2 AR$ specific drugs, such as salbutamol and albuterol [206]. Subtype specificity is difficult to achieve due to the high similarity of orthosteric sites.

Whilst biologics, particularly monoclonal antibodies, offer one route to increased specificity [207], I will focus on small molecules that exert their effect allosterically. Allosteric sites are typically less conserved than orthosteric sites, making specificity easier to achieve. The muscarinic acetylcholine receptors (mAChR) provide a good example of this. mAChR contains five subtypes, termed M_1, M_2, M_3, M_4, and M_5. Specific targeting of M_1 and M_4 subtypes is important for treatment of central nervous system disorders such as schizophrenia, whilst activation of M_2 and M_3 subtypes leads to serious adverse effects. Specificity at the orthosteric site proved incredibly challenging due to the high structural similarity of the subtypes. Allosteric modulators have however been developed that show remarkable specificity [208,209].

The other inherent challenge of designing molecules that target GPCRs is their complex and diverse signalling profile. The signalling pathways activated by GPCR signalling (via either

G-proteins or β-arrestin) lead to diverse effects, many of which are non-desirable. There is therefore a desire to design small molecules that are "biased" toward a specific pathway. An example of a biased allosteric modulator is SBI-553 binding to NTSR1 [210]. SBI-553 is a positive allosteric modulator that is β-arrestin biased. The β-arrestin bias leads to reduced addictive behaviours [211].

One particular type of allosteric site sits at the protein-lipid interface. They are buried in the bilayer and are formed by the transmembrane region of membrane proteins. Ligands that bind here are exposed to the protein and lipid. We therefore term these lipid-exposed binding sites. Small molecules that bind here exhibit their effects allosterically. Numerous structures of small molecules binding at lipid exposed binding sites have been solved for GPCRs, transporters, and ion channels. The exact nature of the binding site varies significantly. For instance, molecules have been found to bind at different depths in the membrane, be almost fully enveloped by the protein or be more lipid exposed, and some exist entirely extra-helically whilst others protrude through the helices (Figure 1.6). These binding sites offer a new paradigm for drug discovery, offering new opportunities, but also presenting novel challenges.

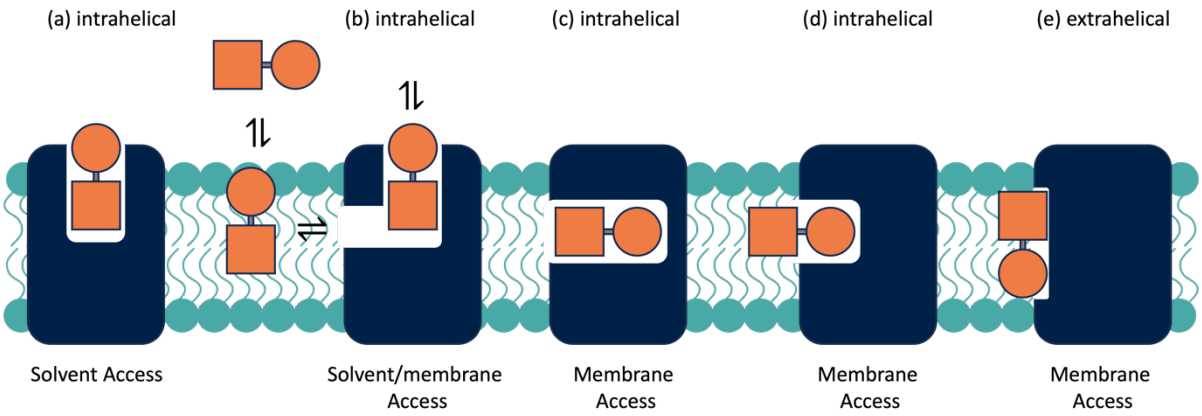


Figure 1.6: The different types of lipid-exposed sites. The membrane can facilitate access to binding sites that sit within the transmembrane helices (b,c), protrude through the helices into the bilayer (d), or are entirely exposed to the bilayer (e).

The Lipinski rule of five is a set of guidelines that suggest how likely a small molecule is to have acceptable oral bioavailability [212]. It states there should be no more than 5 hydrogen bond donors, no more than 10 acceptors, have a mass less than 500 Daltons, and have an octanol/water partition coefficient (clogP) ≤ 5. The key concern for adherents to the Lipinski rule of five would be that small molecules that would be largely exposed to the bilayer must be highly lipophilic, breaking the 'rule' that $clogP \leq 5$. For instance, the small molecule BPTU, that binds to $P2Y_1R$ has $clogP = 5.7$ [212]. Highly lipophilic molecules have been linked to increased toxicity [213]. However, the Lipinski rule of 5 is controversial, with many studies recommending consideration, but not blind following of the terms [214,215]. Recent analyses suggest that marketed drugs actually cover a wide range of physicochemical space, and that drug discovery projects may benefit from consideration of a wider range of lipophilicity [216]. It is also worth considering that these rules are largely derived with the primary purpose of targeting soluble proteins or the solvent accessible orthosteric sites of membrane proteins. In these cases, interactions with the membrane are non-desirable, whilst for lipid-exposed sites this may not be the case. In short, for lipid-exposed sites, despite

limited bioavailability potentially necessitating alternative delivery mechanisms, it may be an exploitable feature, not a bug [217].

1.5.1. The Microkinetic Model

Ligands that bind at lipid-exposed binding sites must necessarily partition into the bilayer before binding at the lipid-exposed site. Ligand binding can therefore be considered a two-step binding process; the ligand first partitions into the bilayer and then binds to the protein. This is termed the microkinetic model. Each step has an associated ΔG, the first step being called $\Delta G_{membrane}$ and the second being termed $\Delta G_{intrinsic}$. These sum to give $\Delta G_{observed}$ which is what is experimentally measured. The bilayer can act as a sink for the ligand, increasing the local concentration of the drug [218]. The bilayer reduces the dimensionality of ligand binding to a two-dimensional problem [217]. It has also been shown that the bilayer may pre-orientate a drug at a particular depth, conformation, or orientation [219].

The microkinetic model has significant implications for drug discovery [220]. It has been hypothesised that small changes in the molecule that impact $\Delta G_{membrane}$ may be the cause of steep SARs commonly seen in membrane proteins [220]. Partitioning into the bilayer can also significantly impact the pharmacokinetics of a small molecules. Increasing the propensity of molecule to partition into the membrane has been shown to increase the half-life of a drug [221]. It has also been proposed that cytochrome P450 directly recruits substrates from the bilayer, meaning a more lipophilic molecule may be metabolised more readily [222]. It has also been shown that there is a link between increased membrane partitioning and xenobiotic detoxification mechanisms [223].

The fact that the small molecule interacts with the bilayer can be used to improve the observed affinity. For instance, it is possible to add lipophilic groups to increase the affinity for the membrane. This can be achieved via the addition of lipophilic tails or basic amines. For instance, the phenylalkoxyalkyl tail of salmeterol makes extensive contacts with β_2AR but also effects the observed potency via increasing the on-rate of membrane partitioning [224,225].

Intra-molecular hydrogen bonding is also important for membrane partitioning, as it prevents hydrogen bond donors being exposed to the bilayer. This is seen in the SAR for ivacaftor, as substitutions that impact the intra-molecular hydrogen bonding lead to a decrease in potency, possibly by inhibiting its membrane-partitioning capacity [226]. Halogenation of a molecule is also a popular method in drug discovery for membrane proteins. Trifluoromethyl groups are particularly popular for increasing lipophilicity and blocking metabolism [221]. The hydrophobic hydrogen-bond acceptor sulfonyl group is also commonly used. A significant limitation of this approach is that lipophilic molecules bring an increased risk of toxicity [227,228].

Clinical toxicity is often caused by a few ion channels that are susceptible to highly lipophilic molecules. For instance, basic amines possess an increased risk of hERG channel inhibition [229] and phospholipidosis [230]. This highlights the need to preferentially increase $\Delta G_{observed}$ via increases to $\Delta G_{intrinsic}$ over $\Delta G_{membrane}$. Understanding the breakdown of $\Delta G_{observed}$ into $\Delta G_{intrinsic}$ and $\Delta G_{membrane}$ can be useful for designing drugs with reduced risk of side effects.

There remains relatively little information about the energetics of membrane partitioning. This is because experiments either struggle to replicate the complex nature of a bilayer or have limited scalability. Using liposomes can provide a realistic model of a membrane but current issues around liposome stability and preparation limit the throughput [231]. A less realistic but more high-throughput option is HPLC using immobilised artificial membrane (IAM) columns [232].

Another problem is the flawed idea that membrane partitioning is driven by exclusively by bulk lipohillicity. This ignores the variety of interactions that a small molecule can make with a bilayer and the role of the complex mixture of lipids. For instance, the electrostatic interactions of basic amines and phosphate groups are known to be important, cholesterol has been shown to reduce the partitioning of small molecules [233], and certain molecules are known to preferably partition into cells/membrane microdomains with particular lipid contents [234].

Work by Sykes *et al.* looked to disentangle the role of membrane partitioning from the observed binding process [225]. They used HPLC on IAM columns to calculate $\log(K_{IAM})$, the membrane partition coefficient for these immobilised bilayers. They showed that $\log(K_{IAM})$ correlates with k_{on} and not k_{off}. They argued that this indicated the concentration of ligand near the receptor is different from that in solution. They use $\log(K_{IAM})$ to correct for the local concentration at the bilayer and then correct k_{on} and K_d for this higher local concentration. They find this 'corrected' k_{on} deviates a lot less between the ligands than the uncorrected k_{on}, suggesting the difference in lipophilicity drives the difference in observed on rates. They also show that the 'true' K_d can be calculated as:

$$pK_d^{true} = pK_i^{obs} + pK_{IAM} \quad \textit{Equation 2}$$

This provides a simpler way for correcting the K_d without the need for large amounts of kinetic data.

1.5.2. Experimental Affinity Measurement

One of the striking observations of ligands binding at GPCRs, and particularly those that bind at lipid-exposed sites, is the lack of high-quality experimental data. Isothermal titration calorimetry (ITC) is typically considered the gold standard for measuring the binding affinity. ITC involves measuring the amount of energy needed to maintain a cell in which the reaction is taking place at a constant temperature relative to a reference cell. At the start of the experiment only species A is present in the cell. A small amount of species B is then injected into the cell. This will result in the binding of B to A which is detected as heat. As the experiment progresses, more of B is iteratively injected into the cell. This progresses until all of A is occupied, meaning the addition of more B leads to no change in temperature.

ITC is technically challenging and time-consuming. An inter-lab study also showed significant deviation between predictions for a simple system, well beyond each individual lab's predicted error [235]. This was later found to be due to concentration errors, something that ITC is highly sensitive to [236]. Another general challenge of ITC is the enthalpy of mixing, caused by mismatches in the buffer compositions.

ITC of membrane proteins is significantly more challenging, with the first ITC study of any membrane protein only being published in 2008 [237]. I will provide a quick overview of the challenges of ITC for membrane proteins, but for a more detailed review please see

Rajarathnam *et al.* [238]. For soluble proteins, mixing errors are corrected for by dialysing the protein, ensuring the buffers are identical. This is far more challenging in membrane proteins due to the presence of detergent. This is because a fraction of the lipid will exist as a micelle, making it too large for efficient dialysis. Detergent can also lead to strangely shaped or double peaks, further complicating interpretation. In most modern ITC experiments of soluble proteins, the ligand is present in the cell and the protein is injected in. This is not possible for membrane proteins. We expect non-specific binding will also be a major problem for ITC at lipid-exposed sites due to the hydrophobicity of the small molecules. This will likely lead to non-specific binding which will cause concentration errors. The requirement for keeping the membrane protein stable throughout the experiment also remains non-trivial.

The predominant experimental method we have observed for determining binding affinities of small molecules binding at lipid-exposed sites is radioligand binding assays. These are typically considered less accurate than ITC and other methods but are significantly easier to run. For a detailed review of radioligand binding assays, please see Flanagan [239]. In essence, a known concentration of a radioactive known binder is added to a known concentration of the protein. This is left to reach equilibrium before the ligand of interest is added. The system is then left to reach equilibration before the system is washed to remove unbound radioligand. The only radioligand left is therefore that which remains bound to the protein. The radioactivity is then measured. This is repeated for a range of ligand concentrations to give a binding curve. From this, the IC_{50} can be determined. This can be converted to K_i using the Cheng-Prusoff equation, provided information about the mode of binding is known:

$$K_i = \frac{IC_{50}}{(1 + \frac{K_M}{[S]})} \qquad \textit{Equation 3}$$

Where $[S]$ is the concentration of the radioligand is used and K_M is the Michaelis constant of the radioligand for the receptor. This form of the Cheng-Prusoff only applies to competitive inhibitors. From this we can see that $IC_{50} \geq K_i$. K_i is related to the $\Delta G_{binding}$ as such:

$$\Delta G^{\theta}_{binding} = -\mathrm{RT} \ln (\mathrm{K_i}) \qquad \textit{Equation 4}$$

When calculating $\Delta G_{binding}$ it is essential to have K_i or K_D. However, IC_{50} is commonly used in place of K_i. K_i can be approximated as IC_{50} for competitive inhibitors when $[S] \gg K_M$. When calculating the $\Delta\Delta G_{binding}$ the two $\frac{K_M}{[S]}$ terms cancel, allowing determination of the relative free energy difference to be calculated directly from IC_{50}.

One factor of importance is non-specific binding. All radioligands will partition into the bilayer or bind to other proteins and glassware used in the assay [240]. The measured radioactivity will be a function of specific and non-specific binding. We must therefore disentangle the role of non-specific binding from the observed binding to get the specific binding. This will particularly be an issue if using radioligands that bind in the bilayer, due to their propensity to partition into the bilayer. The amount of non-specific binding can be reduced by selecting less hydrophobic ligands, reducing the partitioning into the bilayer. Non-specific binding is usually measured by adding an excess of a non-labelled competitive ligand to the radioligand. This will prevent all binding of the radioligand. Therefore, any residual radioactivity is from non-specific binding. Careful consideration must therefore be paid to the conditions of the ligand binding assay.

1.5.3. Lipid-Exposed Binding Sites as a Challenge for ABFEs

Computational tools are utilised throughout the modern drug discovery process. To maximise the therapeutic potential of lipid-exposed binding sites, existing computational tools must be validated and extended. The central theme of this project will focus on assessing how well alchemical free energy calculations work for lipid exposed binding sites. Research into free energy calculations has focused on soluble proteins. The occasional paper that has looked at membrane proteins has focused on relative binding free energy calculations. There is therefore a great need to:

- Systematically assess how well absolute binding free energy calculations work for a variety of lipid exposed sites.
- Identify challenges for free energy calculations for lipid-exposed binding sites. This involves finding both novel challenges and known challenges from soluble proteins that become greater for lipid-exposed sites.
- Suggest practical ways to identify these problems. Where possible, this should be done without knowledge of whether predictions align with experimental data to allow calculations to be applied in a prospective manner.
- Suggest practical solutions to these problems.

As previously discussed, free energy calculations for GPCRs are made significantly more difficult by the presence of multiple protein conformations that are often separated by large energetic barriers. This makes the sampling of the apo state more challenging. For instance, let us imagine a situation where the binding of an antagonist induces a change in the orientation of one of the side chains in the binding site. Previous studies have shown that this is non-trivial to identify *a priori*, is difficult to sample, and can make a difference on the

order of several kcal/mol [85]. It is possible that the ligand binding may favour a different inactive conformation of the GPCR. This would be incredibly difficult to sample fully in simulation and it is often not known what conformation(s) are present in the absence of ligand and their relative populations.

Let us finally consider an agonist that works by stabilising the active state of the GPCR. The presence of basal activity of GPCRs indicates this conformation exists in the ensemble, but it is less energetically favourable. The transitions between active and inactive states are incredibly difficult to sample. Failing to sample this transition means we will not account for the energetic penalty of selecting the active conformation of the GPCR, leading to an overestimation in the magnitude of the free energy of binding.

The correct description of lipid exposed sites is made significantly more challenging by the presence of multiple different lipid types. There are several examples of apo structures of systems showing lipids to be present where the ligand is in the holo structure [241]. The displacement of these lipids must be considered to generate the correct free energy of binding. The correct description of the apo state of the system is further complicated by the fact that in many cases even if a lipid does bind at the ligand binding site in the absence of ligand:

1. There may not be an apo structure.
2. If there is an apo structure, the lipid may not be resolved in the binding site.
3. It may not be possible to identify the lipid type.
4. It may not be possible to identify the correct orientation of the lipid. This is a particular problem for cholesterol.

5. Structural and binding affinity work may be carried out in different membrane compositions.
6. Membrane composition is often not explicitly stated and there is often a lack of reliable information on the composition of the bilayer of the specific cell-line used.

As discussed previously, it is possible to identify lipid binding sites to proteins using coarse-grained simulations. These simulations are relatively cheap and should be run for any novel system to understand what the true apo state looks like in terms of lipid occupancies. Coarse-grained simulations have identified numerous lipid binding sites but practitioners should seek other methods to further validate these sites as coarse-graining smoothens the energy landscape, leading to an imperfect representation of the system.

Lipids can also stabilise the holo state by forming direct interactions with the ligand. This is the case for double-knot toxin binding to TRPV1 [242]. Double-knot toxin is a bivalent peptide consisting of two ICK motifs joined by a short linker. Hydrophobic fingers from each knot extend into the bilayer and clear densities corresponding to phospholipids can be observed. These lipids form interactions with both the toxin and the channel. The authors hypothesise that this stabilises the open state of the protein.

In a similar way to the challenges described for finding apo lipid interactions, it may be difficult to know *a priori* for a system whether there are important ligand-lipid stabilising interactions. The recent release of Martini 3 and its improved description of small molecules offers the chance to investigate the interactions between lipids and bound small molecules [243,244]. This is particularly possible given the recent release of cholesterol parameters [245].

Atomistic simulations can be used to identify possible interactions but are likely not long enough to give information on whether these interactions are transient or meaningful.

ABFE calculations can theoretically identify whether a specific lipid-ligand interaction is favourable (i.e. it stabilises the holo state). However, a practical concern again is the potential to generate an incorrect apo state. Let us consider a scenario where a lipid species, in this case cholesterol, binds to the ligand-protein complex but does not occupy the binding site in the apo state. In the non-interacting windows, there is a vacuum where the ligand was. In this case we know that this space should not be occupied by cholesterol, but it is possible that the proximity of the cholesterol to the vacuum will lead to it occupying the space. This would artificially destabilise the apo state, giving the impression the holo state is more favourable than it is. This would make interpretation of these calculations difficult.

For simplicity, most atomistic simulations use a single species bilayer, often POPC. As described above, this loses interactions involving specific lipid types that may stabilise either the holo or apo state. This assumption also removes any lipids that change the properties of the bilayer, such as fluidity and curvature. Specific knowledge of the system will decide whether these simplifications are appropriate.

1.6. The Systems

We picked systems to study that cover as much of the diversity of lipid-exposed sites in terms of their location, the properties of the bilayer, and the properties of the ligand as possible. In general, it is preferable to use a diverse set of ligands when assessing absolute binding free energy calculations. However, the challenges of pose prediction would likely reduce our ability to isolate other issues that are of more interest. Using a SAR series allows us to place the molecules in the binding site using the shared scaffold as a guide.

A SAR series would typically be elaborated via RBFEs but serve as an appropriate test set for ABFEs in lipid-exposed sites. The affinities of the SAR series should also ideally be measured using a high-quality method such as ITC or SPR. However, as discussed in 'Experimental Affinity Measurement', high quality experimental data is challenging to acquire for membrane protein systems, particularly lipid-exposed sites. This makes validation more challenging, but also highlights the potential utility of alchemical free energy calculations for this system. In the absence of high-quality affinity data, significant preference should be given to data reported in the same paper as hopefully any errors or artefacts will be consistent across the series.

1.6.1. P2Y1R

Human purinergic receptors, a family of class A GPCRs, are split into two subfamilies – $P2Y_1$ and $P2Y_{12}$ like receptors. Both are activated by ADP and trigger glutamate release. The primary physiological role of the $P2Y_1$ and $P2Y_{12}$ receptors are to maintain the integrity of blood vessels via thrombus formation. They are also responsible for the formation of pathogenic thrombi at atherosclerotic plaques. These pathogenic thrombi are the major causes

of myocardial infarction and 80% of strokes [246]. Therefore, there is pharmacological interest in purinergic receptors for their potential as anti-thrombotic targets.

Numerous drugs, including clopidogrel, ticlopidine, prasugrel, and ticareglor, target the $P2Y_{12}R$ [247]. These drugs act as competitive antagonists to the $P2Y_{12}R$, preventing thrombus formation. However, loss of activity of the $P2Y_{12}R$ means platelets no longer fulfil their role in primary haemostasis. This manifests as excessive bleeding and can be seen in $P2Y_{12}$ knockout mice or human patients with $P2Y_{12}$ deficiency [248–250]. Importantly from a drug discovery point of view, an increased bleed liability is also seen for patients treated with clopidogrel and other $P2Y_{12}$ antagonists [249]. This increased bleed liability is not seen when using selective $P2Y_1$ antagonists [251].

Initial $P2Y_1$ selective molecules showed limited bioavailability, leading to a concerted effort to find novel selective $P2Y_1$ antagonists. Chao *et al.* found the small molecule 1-(2-(2-tert-butylphenoxy)pyridin-3-yl)-3–4-(trifluoromethoxy)phenylurea (BPTU) to be an effective and selective $P2Y_1$ antagonist [252]. BPTU consists of a tert-butyl-phenoxy-pyridine group and trifluoromethoxyphenyl connected by a urea linker (Figure 1.7). It binds $P2Y_1R$ with $K_i = 6 \pm 0.6\ nM$ and $P2Y_{12}R$ with $K_i > 70{,}000\ nM$. Chao *et al.* report an extensive SAR series with 36 molecules. The affinity for these molecules was determined by a radioligand binding assay and many of the ligands have ≥ 2 repeats.

The structure of BPTU bound to the $P2Y_1R$ was solved in 2015 (Figure 1.7) [253]. This was the first example of a structure of a small molecule binding at a lipid-exposed site. BPTU sits outside of helices I, II, and III at a depth of 22.7 Å in a relatively shallow binding pocket. The shallow nature of the pocket means 40% of the solvent accessible surface area is actually

exposed to the bilayer [219]. BPTU forms predominantly hydrophobic interactions with the protein, with the only polar interactions being between the urea group of the ligand and the backbone carbonyl of L102$^{2.55}$. This backbone carbonyl is available for hydrogen bonding due to the presence of P105$^{2.58}$ one helical turn away precludes intra-helical hydrogen bonding. The pyridyl group forms hydrophobic interactions with A106$^{2.59}$ and F119. The benzene ring within the phenoxy group wedges in between helices II and III and forms interactions with T103$^{2.56}$, M123$^{3.24}$, L126$^{3.27}$, and Q127$^{3.28}$. The ureido phenyl ring forms two aromatic edge-to-face interactions with F62$^{1.43}$ and F66$^{1.47}$. The structure was further validated using mutagenesis. A106$^{2.59}$W/F/L mutants lost the ability to bind BPTU, as did the T103$^{2.56}$W mutant. A106$^{2.59}$ is unique to $P2Y_1R$ amongst purinergic receptors, explaining the selectivity of BPTU. It is believed BPTU likely exerts its effect by blocking the conformational change of TM1-TM3, thus locking the receptor in the inactive conformation [255].

Several aspects of the $P2Y_1R$/BPTU system have been studied via MD simulations. One study proposed that BPTU binding stabilises an ionic lock and observes an increase in lipid order upon the binding of BPTU [256].

Work by Yuan *et al.* investigated the binding pathway of BPTU [257]. They ran umbrella sampling simulations to calculate the potential of mean force (PMF) across the lipid bilayer. Interestingly the free energy difference is just under 12 kcal/mol, very similar to the experimental affinity of binding to $P2Y_1R$. The results are symmetrised which can often hide issues from sampling or the pulling simulation used to generate starting configurations, significantly limiting the quantitative interpretation of their PMF results. They used well-tempered metadynamics to examine the binding pathway. They found BPTU partitions into

the bilayer before interacting with extracellular loop 2 before finally reaching the binding pocket. They also utilised funnel metadynamics to measure the $\Delta G^{\circ}_{binding}$. They achieve remarkable accuracy with experimental results - $\Delta G^{\circ predicted}_{binding} = -11.5\ kcal/mol$ and $\Delta G^{\circ experimental}_{binding} = -11.7\ kcal/mol$. This result required 750 ns of simulation time with the restraints implemented using PLUMED.

It is worth noting that the performance of funnel metadynamics code is slow. Our initial tests for this system in 2020 were obtaining 25 ns / day on a single GPU meaning 750 ns of simulation time would take approximately a month. Recent improvements in the implementation in the PLUMED funnel metadynamics code and the use of multiple walkers may significantly improve this. They did not run repeats of this process to test the reproducibility of the procedure. They also did not test other molecules in the SAR series to check their ability to rank this data. Positioning of the funnel is key in these situations and experience within our lab indicates that this is non-trivial and significantly impacts results, limiting the utility of funnel metadynamics simulations.

Recent work by Dickson *et al.* used alchemical RBFE calculations on the $P2Y_1R$ [258]. They ran 43 pairwise transformations with an RMSE of 1.37 kcal/mol. They converted these into 30 compound $\Delta G_{binding}$ scores which had an $r^2 = 0.58$. This shows alchemical calculations can work for these sites. However, as previously discussed, RBFEs benefit from many factors that make them easier than ABFEs such as the cancellation of errors and a smaller alchemical space. This is particularly true for GPCRs due to the dynamics of the receptor and lipid-exposed sites where slow lipid dynamics may exacerbate problems around equilibration and sampling. This work therefore does not provide much insight into how effective ABFEs would be on this system. Dickson *et al.* also used RBFEs and the microkinetic model to

decompose $\Delta G_{observed}$ into $\Delta G_{membrane}$ and $\Delta G_{intrinsic}$. They found that substituents on the ortho position of the benzene of the phenoxy-pyridine group improve $\Delta G_{intrinsic}$ whilst those at the meta position increase $\Delta G_{membrane}$. This informs medicinal chemists to focus on the ortho position, despite ortho and meta substituents showing similar improvements in $\Delta G_{observed}$. They propose that the substituents in the ortho position improve interactions with the sidechain of L102$^{2.55}$ and claim that this forms an important binding hot spot.

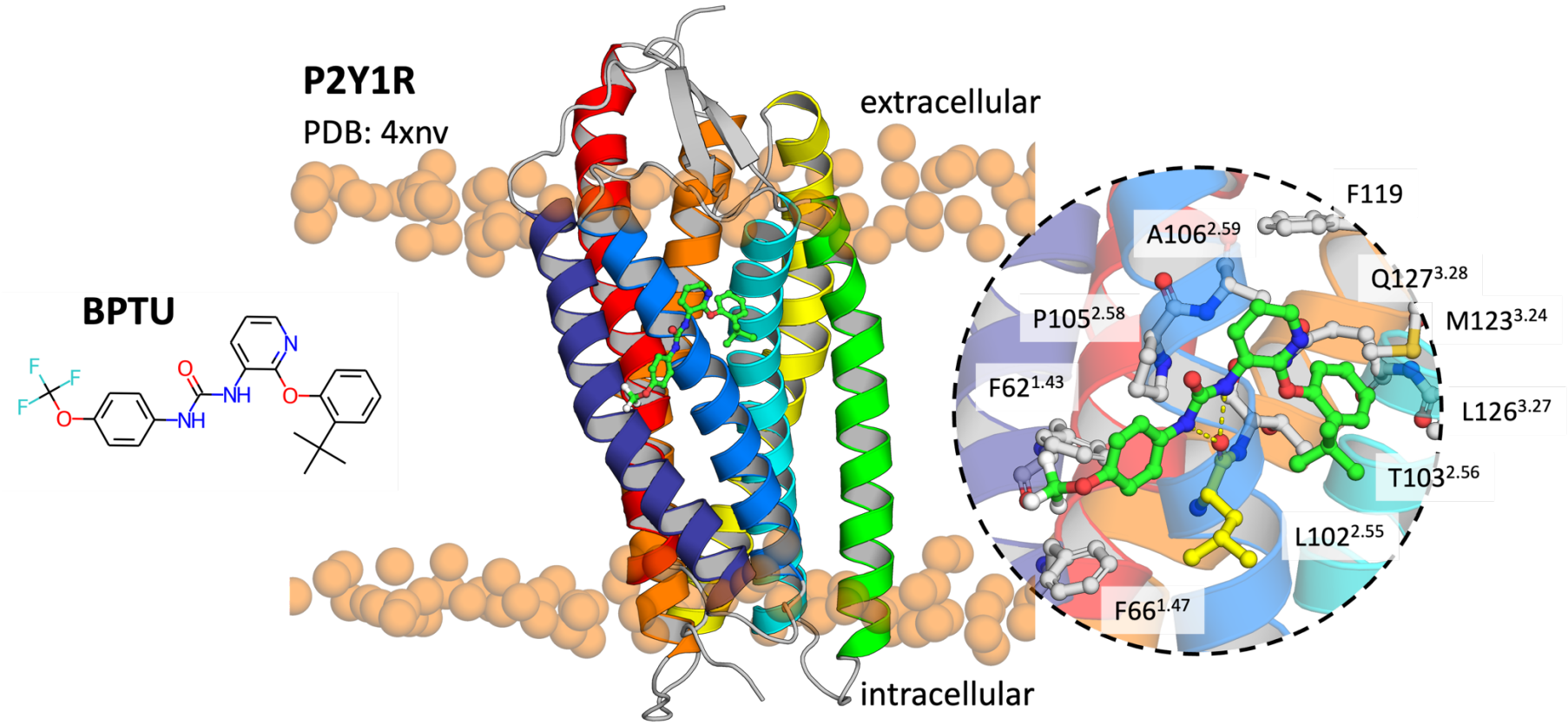


Figure 1.7: The structure of BPTU bound to $P2Y_1R$. The small molecule BPTU binds $P2Y_1R$ outside of helices I (dark blue cartoon), II (royal blue cartoon), and III (cyan cartoon). The only polar interactions are between the urea group of the ligand and the backbone carbonyl of L102$^{2.55}$ (yellow sticks). BPTU forms many hydrophobic interactions with $P2Y_1R$ sidechains (grey sticks). Phosphorous atoms in the POPC headgroups are shown as translucent orange spheres. PDB:4XNV

1.6.2. Glucagon Receptor

Pancreatic alpha cell secretion of glucagon, a 29 residue peptide, is stimulated in response to hypoglycaemia and stimulates an increase in circulating blood glucose [259]. Exogenous administration induces hyperglycaemia in healthy men and can recover diabetic patients in hypoglycaemic comas [260,261]. The glucagon receptor (GCGR) is the endogenous receptor for glucagon. It is a class B GPCR primarily expressed in the liver that is responsible for responding to glucagon levels and stimulating the downstream signalling that leads to the increase in blood glucose levels. Type II diabetes is characterised by chronically elevated blood glucose levels which over time leads to insulin desensitisation. This is particularly relevant in a number of developed countries as type II diabetes is a comorbidity of obesity [262].

There is therefore a need to find ways to reduce blood sugar levels. Antagonism of GCGR is an attractive anti-hyperglycaemic therapeutic target. Knockouts and pharmacological inhibition in animal models and studies of healthy humans have shown GCGR antagonism can prevent glucagon-induced glycaemic increases without severe side effects [263–265].

The first candidate advanced to clinical trials was MK-0893. It successfully reduced fasted blood glucose levels by 34% [266]. However, they also noted that there was a 15% increase in circulating LDL-C and elevated liver alanine transaminase. These side effects led to the discontinuation of MK-0893 in clinical trials. These side effects have since been seen in other GCGR antagonists [267]. Interestingly, current interest in the GCGR for treatment of type II diabetes is via agonism as this has been shown to reduce food intake [268]. Antagonism may still have a role in the treatment of type II diabetes as the side effects, whilst not well understood in humans, are not seen in mice and non-human primates [269]. This difference could be caused by reduced off-target effects in these species. Structural investigations into

understanding any off-target effects in humans and specificity seen in other species may allow the design of GCGR antagonists with reduced side effects.

N-[(4-{(1S)-1-[3-(3, 5-Dichlorophenyl)-5-(6-methoxynaphthalen-2-yl)-1H-pyrazol-1-yl]ethyl}phenyl)carbonyl]-β-alanine (MK-0893) was discovered as a competitive and reversible antagonist of GCGR in 2012 [270]. It was discovered as part of a SAR series containing 31 molecules. MK-0893 was found to have an $IC_{50} = 6.6 \pm 3.5\ nM$ in radioligand binding assays of the human receptor embedded in a membrane preparation of the CHO cell line. ^{125}I-glucagon was used as the radioligand. Standard deviations (when reported) are the mean of at least two technical repeats.

Conversion of this into K_i is a little tricky due to K_M not being explicitly reported for glucagon. However, $K_M \geq K_i$ and K_i is a well approximated by EC_{50}, which is known ($0.22 \pm 0.02\ nM$). Therefore the reported EC_{50} can be used to provide a lower bound for the K_i and assuming $K_i = IC_{50}$ can provide an upper bound. The reported radioligand concentration was $0.050\ nM$. Therefore, the lower bound is $K_i = \frac{IC_{50}}{1+\frac{[S]}{K_M}} = \frac{6.6}{1+\frac{0.050}{0.220}} = 5.4\ nM$ and the upper bound is $6.6\ nM$. This difference is less than the error of the experimental value. Schild plot analysis indicated $K_B = 8.5\ nM$.

Functional assays looking at cAMP production reported an $IC_{50} = 15.7 \pm 0.54\ nM$ in humans, $104.5 \pm 62.8\ nM$ in dogs, and $727 \pm 350\ nM$ in rats. These assays also showed that MK-0893 binds another class B GPCR GIPR with an $IC_{50} = 1020\ nM$. It binds other tested class B GPCRs either with a high IC_{50} (on the order of $10{,}000\ nM$) or one that is beyond the sensitivity of the experiment.

The structure of MK-0893 bound to GCGR was solved in 2016 [271]. MK-0893 consists of a central pyrazole with three prongs attached. The prongs are: ethyl-phenyl-carbonyl-β-alanine (prong 1), methoxynaphthalene (prong 2), and dicholorophenyl (prong 3) (Figure 1.8). MK-0893 carries a charge of -1 at pH 7.4 due to the carboxylate group on prong 1. MK-0893 straddles helix VI of GCGR, with prong 1 occupying a polar cleft between helices VI and VII, prong 2 occupying a hydrophobic cleft between helices V and VI, and prong 3 protruding into the bilayer. It is not evident from the structure what interactions prong 3 will form with the bilayer or water molecules. Prong 2 forms hydrophobic interactions L329$^{5.61b}$, F345$^{6.36b}$, L352$^{6.43b}$, T353$^{6.44b}$, and K349$^{6.40b}$. The amide group of prong 1 forms hydrogen bonds with K349$^{6.40b}$ and S350$^{6.41b}$. The carboxylate group forms a salt bridge with R346$^{6.37b}$ and polar interactions with Asn404$^{7.61b}$ and the backbone of Lys405 (found on the loop between helices VII and VIII). It also forms water-mediated hydrogen bonds with Ser350$^{6.41b}$ and Leu399$^{7.56b}$.

This is an extremely challenging system for ABFEs. The position of the binding site at the protein/lipid/water interface creates numerous challenges. First, as the ligand will be interacting with the headgroups there will be more polar interactions than buried sites. The poor sampling of these interactions has the potential to substantially impact ABFE calculations. Additionally, it is not clear what will or even should occupy the binding site in the absence of ligand; will R346 ‘snorkel’ and draw in water or will the space be occupied by lipid headgroups? The hydrogen bonding network consisting of waters and multiple residues must be properly modelled and sampled. The charge of the ligand adds additional complexity.

It is also worth noting that there are two concerns around the quality of this structure. The first is the R173$^{2.46b}$A mutation. This is on the intracellular side of helix II and directly

opposite the carboxylate group of prong 1. It is possible the mutation may disrupt the water network, the hydrogen bond network between residues, or directly interact with prong 1. The other is the interaction of prong 3 of helix IV of a symmetry mate. This unphysiological interaction observed in the crystallographic packing may have stabilised an unnatural pose.

The dynamics of GCGR, particularly the activation process, has been studied by MD simulations. Li *et al.* used aMD and PMF calculations to study GCGR activation [272]. Their simulations with glucagon bound identified three energy minima – corresponding to inactive, intermediate, and active states. The class B GPCR GCGR activation process shows the same characteristic movement of helix VI as seen in class A GPCRs. L253$^{3.58}$ and Y343$^{6.34}$ $C\alpha$ distance and RMSD of the RxxK motif are identified as useful activation reaction coordinates. The same structure in the absence of glucagon did not see a transition to the active state.

Bai *et al.* extended this work to look at the effect of the binding of the antagonist MK-0893 on receptor dynamics. They use aMD simulations of the same structure in the presence or absence of MK-0893. They see that in the absence of ligand, some of their simulations see a transition to the active intracellular conformation, and the presence of the antagonist prevents this. MD simulations of agonist binding shows that the agonist stabilises a pre-activated conformation, not the active state. The active state is stabilised in the receptor-agonist-G-protein ternary complex [273].

GCGR has also been studied in the context of receptor-lipid interactions. Coarse-grained MD simulations identified that glycosphingolipid GM3 stabilises the open state of the extracellular region of GCGR [274]. The same study also identified four PIP_2 lipid binding sites

– TM1/ICL1/TM2/TM4, TM5/ICL3, TM6/TM7, and H8. No binding was observed at TM3/ICL2 as is seen in class A GPCRs [275]. The TM6/TM7 site overlaps with the MK-0893 binding site. $R/K^{6.37}$ and $R/K^{6.40}$ residues at the intracellular side of helix VI are conserved across class B1 GPCRs. Recent work using coarse-grained simulations and native mass spec has supported the existence of these sites and shown PIP_2 binding is dependent on tail length and receptor conformation [276].

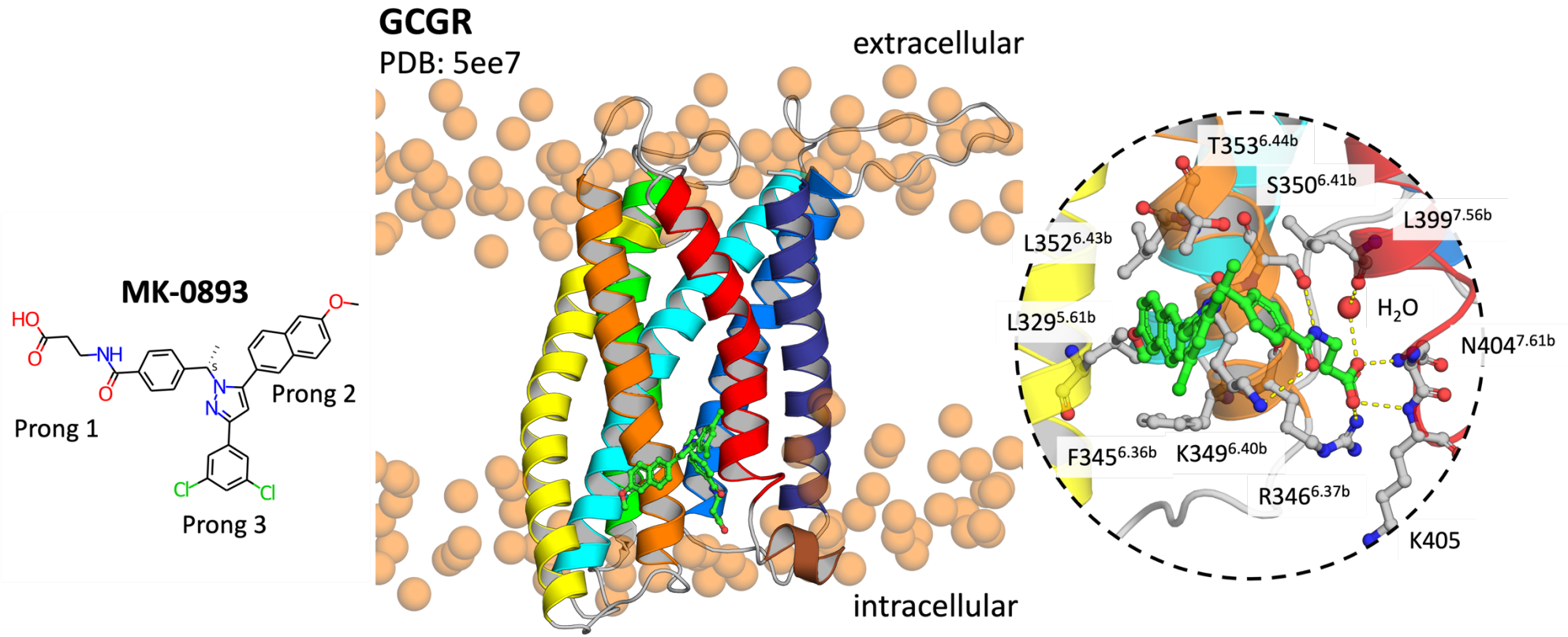


Figure 1.8: The structure of MK-0893 bound to GCGR. The small molecule MK-0893 binds at an extrahelical site on the intracellular side of helices V (yellow), VI (orange), and VII (red). Prong 2 forms hydrophobic interactions L329$^{5.61b}$, F345$^{6.36b}$, L352$^{6.43b}$, T353$^{6.44b}$, and K349$^{6.40b}$. The amide of prong 1 forms hydrogen bonds with K349$^{6.40b}$ and S350$^{6.41b}$. The carboxylate group forms a salt bridge with R346$^{6.37b}$, polar interactions with N404$^{7.61b}$ and the backbone of K405. It also forms water-mediated hydrogen bonds with S350$^{6.41b}$ and L399$^{7.56b}$. Prong 3 protrudes into the bilayer. Phosphorous atoms of POPC headgroups are shown as translucent orange spheres. PDB: 5EE7

1.7. Aims and Motivations

This thesis will focus on the application of alchemical free energy calculations to small molecules binding at lipid-exposed sites on GPCRs. These sites hold significant potential for drug discovery. Alchemical free energy calculations are a popular tool to understand the energetics of ligand binding. Lipid-exposed sites present a significant and novel challenge to alchemical free energy calculations. The presence of lipids and the highly dynamic nature of GPCRs are factors that are expected to cause significant challenges. The ultimate aim for this research is to contribute towards ABFEs being a useful tool in prospective drug discovery pipelines. I will therefore endeavour to not only assess how well ABFEs work, but also to work out how to identify when they do not work, why they do not work, and suggest potential solutions. This thesis will look primarily at two different systems $P2Y_1R$ (Chapter 3) and GCGR (Chapter 4). I will then discuss the challenges and importance of finding the apo state in alchemical free energy calculations (Chapter 5) and how to handle free energy calculations in complex bilayers (Chapter 6).

2. Theory

2.1. Table of Contents

2. THEORY 63

2.1. TABLE OF CONTENTS 64

2.2. TABLE OF FIGURES 65

2.3. COMPUTATION AT SCALE 66

2.3.1. Quantum Mechanics 66

2.4. MOLECULAR DYNAMICS & FORCE FIELDS 69

2.4.1. Molecular Mechanics Force Fields 69

2.4.2. Energy Minimisation 77

2.4.3. Molecular Dynamics 78

2.4.4. Periodic Boundary Conditions 85

2.4.5. Thermodynamic Ensembles and Pressure/Temperature coupling 86

2.4.6. Thermostats 87

2.4.7. Barostats 90

2.4.8. Long Range Interactions 92

2.4.9. Replica Exchange 98

2.5. FREE ENERGY CALCULATIONS 100

2.5.1. Defining Free Energy 100

2.5.2. Computational Methods for Free Energy Prediction 102

2.5.3. Alchemical Methods 103

2.2. Table of Figures

FIGURE 2.1: SCHEMATICS OF BONDED INTERACTIONS FOUND IN TYPICAL FORCE FIELDS.....72

FIGURE 2.2: QUALITATIVE PLOT OF THE LENNARD-JONES POTENTIAL.....75

FIGURE 2.3: QUALITATIVE PLOT OF THE COULOMBIC POTENTIAL OF TWO PARTICLES WITH EQUAL SIGNED CHARGE 76

FIGURE 2.4: A SCHEMATIC OF THE LEAPFROG INTEGRATOR.....84

FIGURE 2.5: AN ILLUSTRATION OF PERIODIC BOUNDARY CONDITIONS IN 2D. 86

FIGURE 2.6: A QUALITATIVE REPRESENTATION OF A LENNARD-JONES POTENTIAL WITH A SWITCHING FUNCTION APPLIED. 94

FIGURE 2.7: A SCHEMATIC OF THE VERLET NEIGHBOUR LIST SCHEME. 95

FIGURE 2.8: 1D ILLUSTRATION OF THE EWALD SUM. 97

FIGURE 2.9: A SIMPLE SCHEMATIC OF THE REPLICA EXCHANGE METHOD. 100

FIGURE 2.10: THE ALCHEMICAL CYCLE FOR AN ABSOLUTE BINDING FREE ENERGY CALCULATION. 103

2.3. Computation at scale

Modern computational chemists consider problems that occur across a wide range of physical scale, from questions at the atomic scale to recent attempts to model entire cells [277]. In this chapter I will discuss the theoretical underpinnings of many tools used by computational chemists and discuss what makes them (in)appropriate for specific research questions. An in-depth description is beyond the scope of this thesis, but further information can be found in several textbooks, notably by Leach [278].

2.3.1. Quantum Mechanics

Quantum mechanics (QM) provides an unrivalled insight into molecular mechanisms at the smallest level. QM methods utilise the time-independent Schrödinger equation to describe the exact electronic structure of the system:

$$\widehat{H}|\Psi\rangle = E|\Psi\rangle \qquad \textit{Equation } 5$$

Where $\widehat{H}$ is the quantum mechanical Hamiltonian, $|\Psi\rangle$ is the state of the system, and E is the energy of the system. The full Hamiltonian for a molecule (of n electrons, and N nuclei of charge Z_I) can be written as:

$$\widehat{H} = \left[\frac{-\hbar^2}{2\mu}\nabla^2 + V(r)\right] \qquad \textit{Equation } 6$$

$$= -\frac{\hbar^2}{2\mu}\sum_{i=1}^{n}\nabla_i^2 - \sum_{i=1}^{n}\sum_{I=1}^{N}\frac{(Z_I e)e}{4\pi\epsilon_0 \boldsymbol{r}_{iI}} + \sum_{i<j}^{n}\frac{e^2}{4\pi\epsilon_0 \boldsymbol{r}_{ij}}$$

Where $\hbar$ is the reduced Planck's constant, μ is the reduced mass of the electron, ∇^2 is the Laplace operator, e is the charge of the electron, $\boldsymbol{r}_{iI}$ is the distance between the ith electron

and Ith nucleus, ϵ_0 is vacuum permittivity, and $\boldsymbol{r}_{ij}$ is the distance between the ith electron and jth electron. The three terms in $Equation\ 6$ correspond to the kinetic energy, electron-nuclear, and interelectronic repulsion respectively.

The complexity of $Equation\ 6$ can be reduced by the Born-Oppenheimer approximation [279]. Under the Born-Oppenheimer approximation, the wavefunction of a molecule can be split into terms relating to the wavefunction of the nuclei and of the electrons ($Equation\ 7$). This separation is motivated by the significant difference between the mass of an electron (e_M) and that of a nucleus (Z_M), with the ratio $Z_M : e_M \approx 2000$. The Born-Oppenheimer approximation makes numerical solutions to the Schrödinger equation more viable to obtain. This allows us to look at the movement of electrons in a system. Methods of note include Hartree-Fock [280,281] and Density Functional Theory [282,283].

$$\Psi_{tot}(nuclei, electrons) = \Psi(electrons) + \Psi(nuclei) \qquad Equation\ 7$$

Even with the Born-Oppenheimer approximation, evaluating the Schrödinger equation for a system remains highly computationally expensive. CCSD(T), that is coupled clusters with singlets and doublets, is often considered the gold standard QM method. Coupled cluster is a numerical technique that takes the Hartree–Fock molecular orbital method and constructs wavefunctions using the exponential cluster operator to account for electron correlation. CCSD(T) approximately scales $\propto N^7$, where N is the number of atoms [285]. This renders the use of QM methods impractical for a lot of biological questions. However, QM methods are widely used in studying chemical reactions as they explicitly represent the electrons, allowing the breaking and formation of bonds. This differentiates it from force field based approaches.

However, the full time-evolved QM description of an entire protein remains computationally impractical. Work by Levitt, Karplus, and Warshel, which ultimately led to them being awarded the 2013 Nobel Prize, worked to 'bridge the gap' by combining QM methods with molecular mechanics (MM)[286]. In essence, the region of interest in a system is treated via QM whilst the rest of the system is treated using classical mechanics, which is far less computationally expensive. In computational biochemistry, QM/MM is most widely used to study enzyme-catalysed reactions, where the catalytic site and substrate are described via QM and the rest of the protein by classical mechanics[286].

2.4. Molecular Dynamics & Force Fields

The previous section provided a brief introduction to the application of quantum mechanical (QM) methods to answer specific biological questions. QM methods offer a very high level of theoretical rigour and accuracy, but are limited to being useful for only a small subset of biological questions due their high computational cost. In this section, I will discuss the theory and application molecular dynamics (MD) simulations. Briefly, these utilise a molecular mechanics (MM) force field to integrate Newton's equations of motion, thus providing a classical mechanics description of the system. MD simulations therefore act as 'computational microscope', allowing us to investigate the dynamics of systems at a resolution beyond what can be done via experimental techniques. Whilst MD simulations lack the accuracy of QM calculations, they allow the exploration of much larger systems (such as complex lipid bilayers and even recent attempts to model an entire cell) and timescales. The ongoing improvement in MD implementation, investment in high performance computing clusters, and improvements in GPU design mean that larger systems can be simulated for longer. This ultimately means that the utility of MD is increasing, allowing us to answer a broader array of biochemical questions.

2.4.1. Molecular Mechanics Force Fields

Molecular mechanics models are commonly referred to as force fields. They describe the potential energy (U) of a system as a function of N nuclear coordinates ($\boldsymbol{r}^N$), via bonded and non-bonded interactions:

$$U_{total} = U(\boldsymbol{r}^N) = \sum U_{bonded} + \sum U_{non-bonded}$$ *Equation 8*

These bonded and non-bonded terms can be further decomposed into various functional forms; the functional form used by AMBER is shown in Equation 8 and will be further discussed in the next sections. The terms in $Equation\ 9$ correspond to the contributions of: bond stretching, angle bending, torsions, van der Waal's interactions, and electrostatic interactions, respectively.

$$\begin{aligned} U_{total} = {} & \textstyle\sum_{bonds} k_r (r - r_0)^2 \\ & + \sum_{angles} k_\theta (\theta - \theta_0)^2 \\ & + \sum_{torsion} \sum_{n} \frac{1}{2} V_n [1 + cos(n\phi - \delta)] \\ & + \sum_{i} \sum_{i \neq j} 4\epsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^{6} \right] \\ & + \sum_{i} \sum_{i \neq j} \frac{q_i q_j}{\epsilon_0 r_{ij}} \end{aligned}$$ *Equation 9*

The success of an MD simulation is dependent on the accuracy of the force field. This has been a key area of research for numerous years, with numerous protein force fields being produced (such as AMBER [288], GROMOS [289], CHARMM [290] etc.). These are all 'empirical' force fields as they are parameterised to reproduce various types of experimental data, including crystallography, spectroscopy, and QM calculations. One important feature of a forcefield is its transferability, meaning that there is an expectation that it can reasonably

model a series of molecules. It is worth noting that different force fields utilise different parameterisation processes and consequently force fields should not be mixed.

Force field development remains an active area of research. One area of focus is machine learning force fields [284]. In essence, ML-forcefields try and learn the underlying chemistry directly from the input data, typically high quality QM data. They are therefore theoretically bounded by the QM method of choice, and in terms of computational performance, the fastest ML-force fields remain typically 1-3 orders of magnitude slower than conventional force fields due to their more complex functional forms.

Other recent areas of interest include looking to streamline the force field parameterisation process (allowing new data to be incorporated and validated more easily) [291] and avoiding atom-typing. Atom-typing is the process of assigning an atom a particular atomtype based on its environment. The atom then derives nonbonded and bonded properties from these atomtypes. This causes bloating of a forcefield as novel chemistry requires novel atomtypes, and also increases the chance of inaccurate parameters being included. Recent work by OpenFF has looked to use 'direct chemical perception' to assign parameters based directly on the chemical environment, leading to SMIRNOFF, a more streamlined force field[293] [292].

2.4.1.1. Bonded Interactions

Typical forcefields consider three bonded contributions: bond stretching, angle bending, and bond torsions.

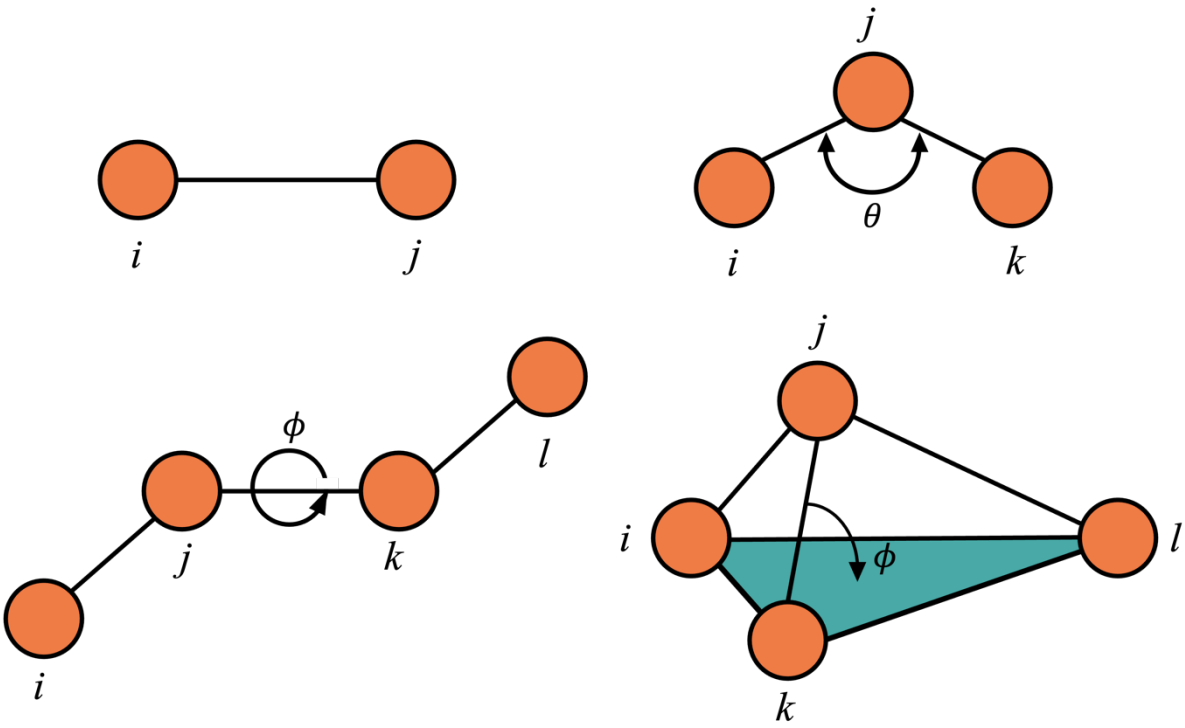


Figure 2.1: Schematics of bonded interactions found in typical force fields. *Clockwise from top-left is: bonded, angles, improper dihedrals, proper dihedrals*

The potential energy of the bonded interactions of a system can be written as:

$$U_{bonded} = U_{bonds} + U_{angles} + U_{dihedrals} \qquad \textit{Equation } 10$$

The potential energy terms for bond stretching and angle bending are most typically modelled as a simple harmonic using Hooke's law formula:

$$U_{bond}(r_{ij}) = \frac{k_{ij}^{r}}{2}\left(r_{ij} - r_{ij}^{0}\right)^{2} \qquad \textit{Equation } 11$$

r_{ij} is the instantaneous distance between particles i and j, k_{ij}^{r} is the force constant ('stiffness') of the bond that controls the energetic cost of deviation from the reference value (r_{ij}^{0}).

$$U_{angle} = \frac{k_{ijk}^{\theta}}{2}\left(\theta_{ijk} - \theta_{ijk}^{0}\right)^{2} \qquad \textit{Equation } 12$$

θ_{ijk} is the instantaneous angle between particles i, j, and k, k_{ijk}^{θ} is the force constant ('stiffness') of the bond that controls the energetic cost of deviation from the reference value (θ_{ijk}^{0}).

The force constants for bond stretching and angle bending are typically high, meaning deviations are unlikely. For instance, a deviation of 0.2Å from the reference value for a Csp^3-Csp^3 bond with a typical k_{ij}^{r} of 300 kcal/mol/Å^2 would result in an increase of potential energy of 12 kcal/mol. The energy is typically less for angle bending, but still high enough for bond stretching and angle bending to sometimes be referred to as 'hard' degrees of freedom.

Dihedral angles, which are the angles between two planes that contain points i, j, k and j, k, l, often show more flexibility and contribute significantly to conformational diversity of a system. Theoretically, non-bonded interactions between the 1,4 atoms in a dihedral could be sufficient to accurately determine the torsional profile, but most force fields use an explicit torsional term. A typical function form for a dihedral is:

$$U_{torsions} = \sum_{n=0}^{N} \frac{V_n}{2}[1 + \cos(n\phi - \delta] \qquad \textit{Equation } 13$$

Where ϕ is the torsion, n is the multiplicity (which corresponds to the number of minima in the profile), and V_n is the barrier to rotation.

Improper torsions are often used to maintain the planarity or stereochemistry of a molecule. These differ from proper dihedrals in that improper dihedrals are not bonded in the sequence 1-2-3-4. They can have the same functional form as proper dihedrals, or have a simple harmonic form.

In a number of systems, the majority of conformational changes are due to changes in the dihedral angles. Accurate modelling of torsional energy profiles is therefore of significant importance, but remains highly challenging. Utilising QM calculations to generate more accurate torsional profiles is an attractive proposition. However, a number of molecules of interest are too large to have their torsions routinely parameterised using QM due to the scaling of QM calculations with the number of particles. This includes larger small molecules and macromolecules. Recent research has looked at streamlining this process, including by utilising fragmentation methods that maintain the chemical environment of the dihedral of interest[294,295].

2.4.1.2. Non-Bonded Interactions

The two types of non-bonded interaction in a classical force field are van der Waal's (vdW) forces and electrostatic interactions. Non-bonded interactions measure interactions both *inter* and *intra* molecularly. This is worth noting as it will be relevant in the discussion of how to carry out free energy calculations.

vdW interactions are typically modelled by a pairwise potential between a pair of atoms (i and j), which acts as a function of their internuclear distances. This often takes the form of a 6-12 Lennard Jones potential [296]:

$$U_{vdW}(r_{ij}) = U_{LJ}(r_{ij}) = 4\epsilon_{ij}\left[\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6}\right]$$ *Equation* 14

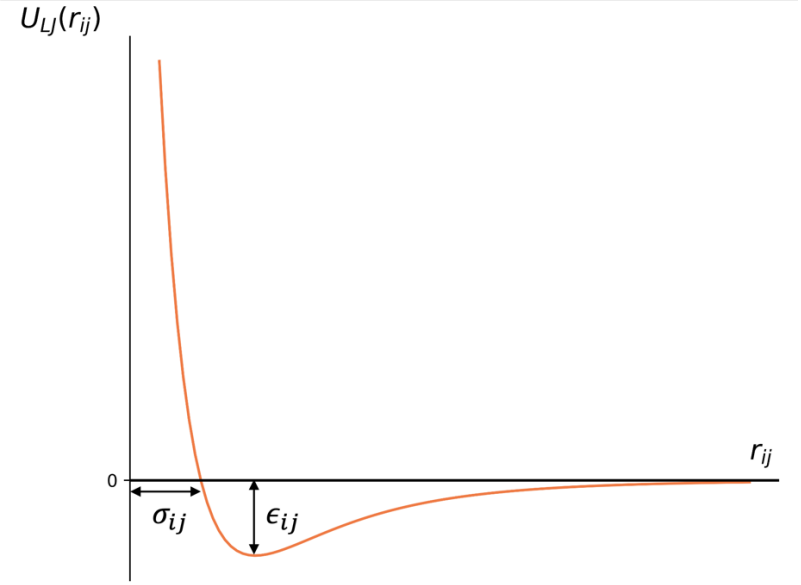


Figure 2.2: Qualitative plot of the Lennard-Jones potential

Where r_{ij} is the internuclear distance, ϵ_{ij} is the well depth, and σ_{ij} is the distance at which $U_{vdW}(r_{ij}) = 0$, that is the vdW radius or collision diameter. The r^{-12} term describes the Pauli repulsion at short ranges and the r^{-6} term describes the long-range attraction due to dispersion forces. σ_{ij} and ϵ_{ij} are normally derived from the atom types of the two interacting atoms using the Lorentz-Berthelot combination rules:

$$\sigma_{ij} = \frac{1}{2}(\sigma_i + \sigma_j)$$ *Equation 15*

$$\epsilon_{ij} = \sqrt{\epsilon_i + \epsilon_j}$$ *Equation 16*

In a classical force field, every particle in the system is assigned a partial charge. Electrostatic interactions therefore occur between any two atoms carrying a partial charge. Coulomb's law is used to model the potential energy of these electrostatic interactions:

$$U_{electrostatic} = \sum_{i} \sum_{i \neq j} \frac{q_i q_j}{4\pi\epsilon_0 r_{ij}}$$ *Equation 17*

Where q_i and q_j are the partial charges of particles i and j, respectively, ϵ_0 is the permittivity of a vacuum and r_{ij} is the distance between particles i and j.

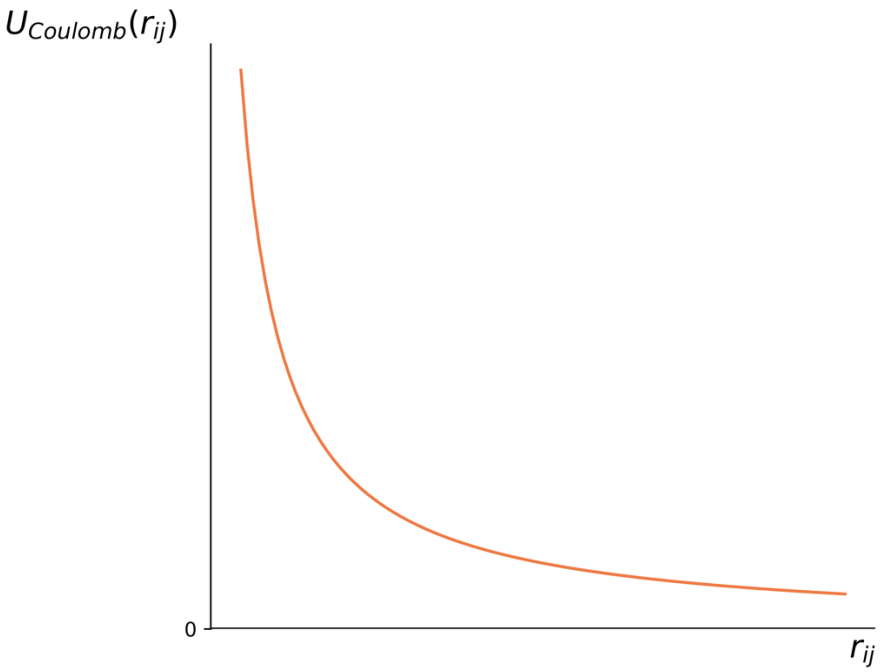


Figure 2.3: Qualitative plot of the Coulombic potential of two particles with equal signed charge

2.4.2. Energy Minimisation

The initial atomic coordinates of a system are typically determined from an experimental structure, a homology model, or, in the case of protein-ligand complexes, by docking. Consequently, unfavourable contacts and geometries are common, leading to large forces and instability issues in the simulation. Prior to simulation, it is therefore normal to run an energy minimisation process. Energy minimisation is a process that aims to optimise the geometry of a system by finding a minima in the potential energy surface. A minima is a set of coordinates where the first derivative of the potential energy is negative, and the second derivative is positive:

$$\frac{\partial U}{\partial \boldsymbol{r}} = 0 \text{ and } \frac{\partial^2 U}{\partial \boldsymbol{r}^2} > 0 \qquad \textit{Equation } 18$$

In practice this involves removing steric clashes and finding acceptable bond lengths, angles, and dihedrals. Due to the complexity of a typical potential energy surface, this is a complicated process and typically only a local minima is found. The complexity of the potential energy surface means analytical solutions are not possible, meaning a numerical method is required. Minimisation algorithms can be broadly split into two categories: those that use the derivative of the potential energy with respect to the atomic coordinates, and those that do not. These are said to be derivative and non-derivative methods respectively. Examples of non-derivative minimisation methods include the simplex and the sequential univariate model. These will not be discussed further as derivative minimisation methods are more common in molecular modelling. The first derivative provides the 'slope', guiding the system towards the energy minima, with the magnitude of the first derivative informing how steep the local slope is. Derivative methods can be further divided into those that use only the first derivative ('first order minimisation methods') and those that also use the second derivative ('second order minimisation methods'). Second order minimisation methods include the *Newton-Raphson* Method [297]; first order minimisation methods include *steepest*

descent and *conjugate gradients* minimisation [298]. Steepest descent provides an efficient method for finding minima, and was used in this work. Briefly, it works by iteratively changing the atom coordinates in the negative direction of the gradient of the potential energy.

2.4.3. Molecular Dynamics

Molecular dynamics (MD) is a useful tool for inspecting the dynamics of a system over time. Provided a starting three-dimensional structure of a protein, MD simulations utilise the potential energy function provided via a force field (discussed above) to evaluate the forces acting on the atoms. Newton's second law of motion can then be integrated to give the positions of the atoms as the simulation progresses. This section will discuss these ideas in more detail, as well as looking at other theoretical aspects of MD simulations.

Newton's second law of motion states that force is equal to the rate of change of momentum. The force acting on a particle i can be written as:

$$\boldsymbol{F}_i = m\boldsymbol{a}_i \qquad \textit{Equation } 19$$

$$-\nabla_{\mathbf{r}_i} U(\boldsymbol{r}_i) = m\ddot{\boldsymbol{r}}_i(t) \qquad \textit{Equation } 20$$

$$\iint -\nabla_{\mathbf{r}_i} U(\boldsymbol{r}_i) dtdt = m\boldsymbol{r}_i(t) \qquad \textit{Equation } 21$$

$$\boldsymbol{r}_i(t) = \frac{1}{m} \iint -\nabla_{\mathbf{r}_i} U(\boldsymbol{r}_i)\, dtdt \qquad \textit{Equation } 22$$

This shows how the equation of motions can yield a time resolved trajectory, in which $\boldsymbol{r}_i(t)$, that is the co-ordinates of the system, is dependent on $U(\boldsymbol{r}_i)$. However, due to the many-body nature of this problem, $Equation\ 22$ cannot be solved analytically, and instead must be discretised and solved numerically. In MD simulations, integrators are responsible for this process.

MD simulations are typically used to calculate ensemble properties of the system. As a result of the *ergodic hypothesis* [299], an ensemble property A can be calculated as a time average of the simulation:

$$\langle A \rangle = \frac{1}{M}\sum_{i=1}^{M} A(\boldsymbol{p}^N, \boldsymbol{r}^N) \qquad \textit{Equation 23}$$

Where $\boldsymbol{p}^N$ is the momenta of the particles in the system.

When simulating the system, it is important the simulation maintains the conservation of energy. The total energy of a system (E) is the sum of the kinetic and potential energy. Differentiating the energy with respect to time shows that Newton's equations of motions respect the conservation of energy:

$$E = \sum_{i=1}^{N} \frac{1}{2} m_i \dot{\boldsymbol{r}}_i^2 + U(\boldsymbol{r}_1, \ldots, \boldsymbol{r}_N) \qquad \textit{Equation 24}$$

$$\frac{dE}{dt} = \sum_{i=1}^{N} \frac{1}{2} m_i \dot{\boldsymbol{r}}_\iota \cdot \ddot{\boldsymbol{r}}_\iota + \sum_{i=1}^{N} \frac{\partial U}{\partial \boldsymbol{r}_i} \cdot \dot{\boldsymbol{r}}_i \qquad \textit{Equation 25}$$

$$= \sum_{i=1}^{N} \dot{\boldsymbol{r}}_i \cdot \left[m_i \ddot{\boldsymbol{r}}_i + \frac{\partial U}{\partial \boldsymbol{r}_i} \right] \qquad \textit{Equation 26}$$

$$= \sum_{i=1}^{N} \dot{\boldsymbol{r}}_i \cdot [m_i \ddot{\boldsymbol{r}}_i - \boldsymbol{F}_i] \qquad \textit{Equation 27}$$

$$= 0$$

2.4.3.1. Finite Difference Methods

Under the influence of a continuous potential, the motions of particles in a system are coupled, giving rise to a many-body problem. This means $Equation\ 22$ cannot be solved analytically and must be discretised and solved numerically. Integration of the equations of motion are integrated using a *finite difference method* under these circumstances. Essentially, the integration is broken down into many small stages of length δt. The total force on a particle at time t is calculated. The acceleration is then calculated from the force and from the acceleration, the positions at time t, and the velocities at time t, the positions and velocities at time $t + \delta t$ are calculated. This is done iteratively to work out the positions and velocities at time $t + 2\delta t,\ t + 3\delta t$ etc. This is how the system evolves over time and our simulation progresses. There are many different methods for integrating the system. Integration is not typically a large part of the computation (calculation of forces normally dominates). Therefore, the main considerations when picking an integration method is that the integration algorithm should conserve energy and momentum, be time-reversible, and permit a long δt. Time-reversibility (as well as completely reproducible simulation) is not possible practically due to limited machine precision.

The selection of an appropriate δt is vital. If the timestep is too small, the system will evolve very slowly and cover little phase space, meaning excessive simulation would be required to answer your question. If the timestep is too large, the forces and velocities will be updated too infrequently, leading to high energy overlaps between atoms. These clashes can cause instabilities and program failure due to the high energies. A guide for chosing δt is that it

should be less than a tenth of the time of the shortest period of motion. This is typically bond stretching, particularly those involving H atoms, which happens on the order of 10 fs. These high frequency modes are typically of little interest in MD simulations. One way to extend δt is to use constraints for these bonds, essentially freezing them at their equilibrium value. SETTLE [300] is the constraint algorithm used by GROMACS for water molecules, with P-LINCS [301] being used for other biomolecules. These constraints allow a timestep of 2 fs to be used. Hydrogen mass repartitioning, a process that involves transferring some of the mass from heavy atoms to hydrogens, allows a timestep of 4 fs [302]. HMR is appropriate for free energy calculations as it effects the kinetics of the system but not the thermodynamics.

It is also important to discuss how to get the initial velocities. The initial structure is typically obtained from an experimental structure that contains no information on the velocities of the atoms. I therefore instead randomly assign velocities based on the Maxwell-Boltzmann distribution:

$$p(\boldsymbol{v}_{ix}) = \left(\frac{m_i}{2\pi k_B T}\right)^{1/2} exp\left[-\frac{1}{2}\frac{m_i \boldsymbol{v}_{ix}^2}{k_B T}\right] \qquad \textit{Equation 28}$$

This describes the probability of an atom i of mass m_i having a velocity v_{ix} in the x direction at a temperature T.

2.4.3.2. Verlet Integrator

The Verlet integrator is a very commonly used integrator [303]. The Verlet integrator, and other integrators, all utilise the underlaying assumption that the positions of a system and the dynamic properties of the system, such as velocity and acceleration, can be approximated as Taylor series expansions:

$$\boldsymbol{r}(t+\delta t) = \boldsymbol{r}(t) + \delta t\boldsymbol{v}(t) + \frac{1}{2}\delta t^2\boldsymbol{a}(t) + \frac{1}{6}\delta t^3\boldsymbol{b}(t) + \frac{1}{24}\delta t^4\boldsymbol{c}(t) + \ldots \quad \textit{Equation 29}$$

$$\boldsymbol{v}(t+\delta t) = \boldsymbol{v}(t) + \delta t\boldsymbol{a}(t) + \frac{1}{2}\delta t^2\boldsymbol{b}(t) + \frac{1}{6}\delta t^3\boldsymbol{c}(t) + \ldots \quad \textit{Equation 30}$$

$$\boldsymbol{a}(t+\delta t) = \boldsymbol{a}(t) + \delta t\boldsymbol{b}(t) + \frac{1}{2}\delta t^2\boldsymbol{c}(t) + \ldots \quad \textit{Equation 31}$$

$$\boldsymbol{b}(t+\delta t) = \boldsymbol{b}(t) + \delta t\boldsymbol{c}(t) + \ldots \quad \textit{Equation 32}$$

The Verlet algorithm uses the positions and accelerations at time t, and the positions from the previous step $\boldsymbol{r}(t-\delta t)$, to calculate the new positions $\boldsymbol{r}(t+\delta t)$ at $t+\delta t$. The relationships between $\boldsymbol{r}(t+\delta t)$ and $\boldsymbol{r}(t-\delta t)$ can be written down in terms of the positions and velocities at time t.

$$\boldsymbol{r}(t+\delta t) = \boldsymbol{r}(t) + \delta t\boldsymbol{v}(t) + \frac{1}{2}\delta t^2\boldsymbol{a}(t) + \ldots \quad \textit{Equation 33}$$

$$\boldsymbol{r}(t-\delta t) = \boldsymbol{r}(t) - \delta t\boldsymbol{v}(t) + \frac{1}{2}\delta t^2\boldsymbol{a}(t) - \ldots \quad \textit{Equation 34}$$

These can be summed to give (if we ignore higher order terms):

$$\boldsymbol{r}(t+\delta t) = 2\boldsymbol{r}(t) - \boldsymbol{r}(t-\delta t) + \delta t\boldsymbol{v}(t) + \delta t^2\boldsymbol{a}(t) \quad \textit{Equation 35}$$

The Verlet algorithm is incredibly popular for its ease of implementation and limited memory requirements. One drawback of the Verlet algorithm is that velocities do not explicitly appear in the Verlet algorithm, although these can be easily calculated:

$$\boldsymbol{v}(t) = \frac{[\boldsymbol{r}(t+\delta t) - \boldsymbol{r}(t-\delta t)]}{2\delta t}$$ *Equation 36*

This equation, however, requires that the positions at the next step have been calculated.

2.4.3.3. Leap-frog Integrator

An equivalent algorithm to the Verlet integrator is the leap-frog integrator [304]. This calculates velocities on the half time step:

$$\boldsymbol{r}(t+\delta t) = \boldsymbol{r}(t) + \delta t \boldsymbol{v}\left(t + \frac{1}{2}\delta t\right)$$ *Equation 37*

$$\boldsymbol{v}\left(t + \frac{1}{2}\delta t\right) = \boldsymbol{v}\left(t - \frac{1}{2}\delta t\right) + \delta t \boldsymbol{a}(t)$$ *Equation 38*

This leads the velocities 'leap-frogging' over the positions at t to give velocities at $t + \frac{1}{2}\delta t$ (Figure 2.4). The two main advantages of the leap-frog algorithm over the Verlet algorithm are that it explicitly includes the velocity, and does not require the calculation of the differences of large numbers, which can lead to a loss of precision. The main disadvantage is that the positions and velocities are not synchronised, meaning that it is not possible to calculate the contribution of the kinetic energy (dependent on velocities) to the total energy at the same time as the potential energy (dependent on positions).

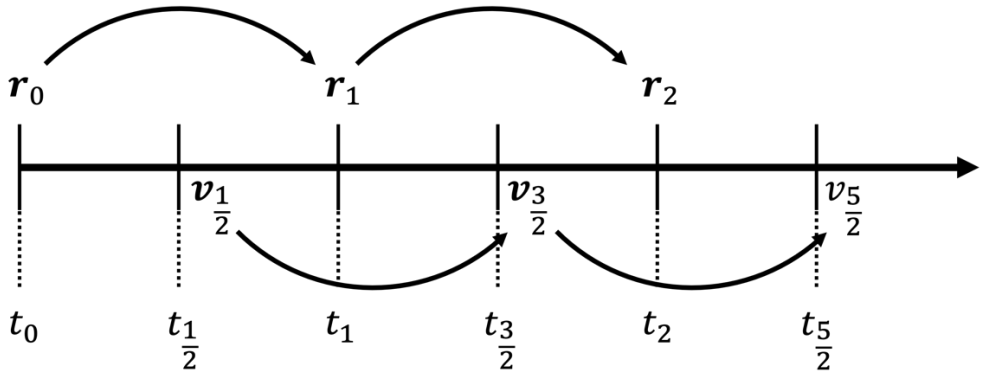


Figure 2.4: A schematic of the Leapfrog integrator. The positions ($\boldsymbol{r}$) are calculated at the full timestep, whilst the velocities ($\boldsymbol{v}$) are calculated at the half timestep

Other examples of Verlet-related integrators include the *velocity Verlet* [303] algorithm and Beeman's algorithm.

2.4.3.4. Stochastic Dynamics

Stochastic dynamics (also known as Langevin dynamics) adds a friction and a noise term to Newton's second equation of motion [305], so that for a particle i:

$$m\ddot{\boldsymbol{r}}_{\boldsymbol{i}} = \boldsymbol{F}_i(\boldsymbol{r}) - m_i\gamma\dot{\boldsymbol{r}} + \boldsymbol{R}_i \qquad \textit{Equation 39}$$

Where $\boldsymbol{F}_i(\boldsymbol{r})$ is the force deriving from the particles interactions with other particles; $m_i\gamma\dot{\boldsymbol{r}}$ is the force from the motion of the particle through the solvent (typically referred to as a frictional term, with γ being the friction coefficient); and $\boldsymbol{R}_i$ being the force deriving from the fluctuations caused by random interactions with solvent particles. $\boldsymbol{R}_i$ has a Gaussian distribution with a mean of zero.

Stochastic dynamics can be thought of as molecular dynamics with stochastic temperature-coupling. This results from the dependence of $\boldsymbol{R}_i$ on temperature. It can be shown that the resulting velocities converge to a Maxwell-Boltzmann distribution at the reference temperature. The main advantage of stochastic dynamics over other temperature-coupling

methods is that it guarantees the generation of a Maxwell-Boltzmann distribution. This is particularly relevant for free energy calculations as the decoupled atoms can have incorrect velocity distributions as they are decoupled from the system. In the work presented in this thesis, Stochastic Dynamics was used for the free energy calculations.

2.4.4. Periodic Boundary Conditions

When simulating a system, one must consider what happens at the boundary of the system. The absence of atoms beyond the edge of our simulation box would indicate the presence of a vacuum, which would lead to unphysical behaviour. The solution for this is to apply periodic boundary conditions (PBC). This involves defining the system as a unit cell in an infinite lattice, where all other unit cells contain a replica of the system (Figure 2.5). If a particle leaves the unit cell, it will be replaced by an equivalent particle from one of the periodic images. Under PBC, only interactions with the nearest copy of a particle are considered. This is called the minimum image convention.

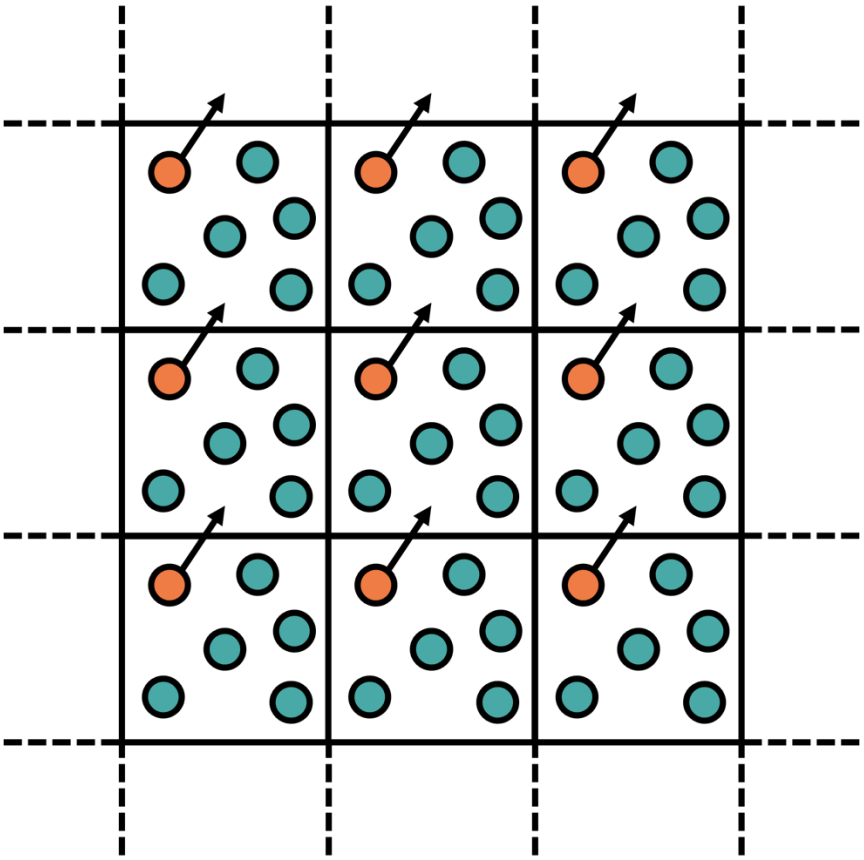


Figure 2.5: An illustration of periodic boundary conditions in 2D. The central system is surrounded by periodic images of itself. If the orange particle passes out of the simulation cell, an equivalent particle from a periodic image will replace it.

2.4.5. Thermodynamic Ensembles and Pressure/Temperature coupling

2.4.5.1. Ensembles

A macroscopic observables (A_{obs}) of a system is determined by the underlying microstates of the system. An ensemble is the collection of all the unique microstates in a multi-dimensional space called phase space ($\boldsymbol{\Gamma}$), where a microstate is described by the positions ($\boldsymbol{q}$) and momenta ($\boldsymbol{p}$) of the system. Common ensembles are shown in Table 2.1. As time evolves, the system will pass through all possible microstates in $\boldsymbol{\Gamma}$. $\rho(\boldsymbol{\Gamma})$ is the density of the microstates for a particular ensemble and represents the probability of finding a system at $\boldsymbol{p}, \boldsymbol{q}$. The macroscopic observable is the time average of the system, and tends towards the 'true' A_{obs} as time tends towards infinity. However, infinite time cannot be achieved in MD simulations

so the *ergodic hypothesis* is applied; that is that the time average of the system is equal to the ensemble average. Molecular dynamics allows us to sample states from the ensemble, allowing us to estimate the macroscopic observable. Sufficiently sampling the full phase space remains a central challenge in MD. The production simulations were run in isothermal-isobaric ensemble as this most closely resembles cellular conditions.

$$A_{obs} = \langle A_{time} \rangle = \langle A(\mathbf{\Gamma}(t)) \rangle_{time} \quad \textit{Equation 40}$$

$$= \lim_{t_{obs} \to \infty} \frac{1}{t_{obs}} \int_0^{t_{obs}} A\left(\mathbf{\Gamma}(t)\right) dt \quad \textit{Equation 41}$$

$$= \langle A \rangle_{ensemble} \quad \textit{Equation 42}$$

Ensemble	Abbreviation	Fixed Parameters
Micro-canonical	NVE	Number of particles, volume, energy
Canonical	NVT	Number of particles, volume, temperature
Isothermal-isobaric	NPT	Number of particles, pressure, temperature
Grand Canonical	μVT	Chemical potential, volume, temperature

Table 2.1: Common thermodynamic ensembles

2.4.6. Thermostats

The temperature (T) of a system with N particles is related to the kinetic energy (K) of the particles:

$$K = \frac{1}{2}\sum_{i=1}^{N} m_i \boldsymbol{v}_i^2 = \frac{3}{2} N k_B T \quad \textit{Equation 43}$$

$$T = \frac{2}{3}\frac{K}{N k_b} = \frac{1}{3 N k_b}\sum_{i=1}^{N} m_i \boldsymbol{v}_i^2 \quad \textit{Equation 44}$$

Therefore, the simplest way to correct the instantaneous temperature of the system ($T(t)$) to the desired temperature (T_{ref}) is to scale the velocities according to the factor λ, where $\lambda = \sqrt{T_{ref}/T(t)}$.

$$\Delta T = \frac{1}{3}\sum_{i=1}^{N}\frac{m_i(\lambda \boldsymbol{v}_i)^2}{Nk_B} - \frac{1}{3}\sum_{i=1}^{N}\frac{m_i \boldsymbol{v}_i^2}{Nk_B} \quad \textit{Equation 45}$$

$$= \frac{1}{3Nk_b}\left(\sum_{i=1}^{N} m_i(\lambda \boldsymbol{v}_i)^2 - \sum_{i=1}^{N} m_i \boldsymbol{v}_i^2\right) \quad \textit{Equation 46}$$

$$= \frac{1}{3Nk_b}\sum_{i=1}^{N} m_i \boldsymbol{v}_i^2\,(\lambda^2 - 1) \quad \textit{Equation 47}$$

$$= T(t)(\lambda^2 - 1) \quad \textit{Equation 48}$$

$$\lambda = \sqrt{T_{ref}/T(t)} \quad \textit{Equation 49}$$

This method is not appropriate, however, as it does not produce the correct thermodynamic behaviour in the canonical ensemble. If one applies this method at the limit where the velocities are scaled every time step, the kinetic energy is constant, with no fluctuations. In statistical mechanics the kinetic energy is expected to fluctuate around the mean, as described by the below equation:

$$\sigma_K^2 = \frac{3}{2}\mathrm{Nk_B T^2} \quad \textit{Equation 50}$$

The *Berendsen thermostat* [306] overcomes this by coupling the system to an external heat bath. The external heat bath is held constant at T_{ref} and adds/removes heat from the system by scaling the velocities, such that the rate of change in temperature is proportional to the difference between T_{ref} and $T(t)$.

$$\frac{dT(t)}{dt} = \frac{1}{\tau}(T_{ref} - T(t)) \qquad \textit{Equation } 51$$

Where τ is a coupling parameter that defines how tightly the system and bath are coupled, essentially controlling the rate at which the system exponentially decays to T_{ref}. The difference in temperature between successive steps is:

$$\Delta T = \frac{\Delta t}{\tau}(\mathrm{T_{ref}} - \mathrm{T(t)}) \qquad \textit{Equation } 52$$

Meaning the scaling factor for the velocities is:

$$\lambda^2 = 1 + \frac{\Delta t}{\tau}\left(\frac{T_{ref}}{T(t)} - 1\right) \qquad \textit{Equation } 53$$

Whilst the Berendsen thermostat does allow fluctuations in kinetic energy, it has been shown that it does not produce the correct canonical ensemble and is not appropriate for use in production simulations.

Other thermostats do reproduce the correct ensemble – for instance the *Anderson thermostat* [307] achieves this by *stochastic collisons*. In this process, a particle is selected at random and its velocity is reassigned by random selection from the Maxwell-Boltzmann distribution. This is equivalent to the system being in contact with a heat bath that emits random 'thermal particles' which collide at random with particles in the system. The correct collision rate is essential – too low and we do not reproduce the correct distribution of energies, too high and we do not see the expected fluctuations in kinetic energy. The *Nosé-Hoover* thermostat [305] is an example of an *extended system* thermostat. It essentially works by treating the external

reservoir as in integral part of the system (s), which has its own mass (Q) and velocity ($\dot{s}$). The rewritten Hamiltonian can be seen below:

$$H(\boldsymbol{q},\boldsymbol{p},s,p_s) = \sum_{i=1}^{N} \frac{\boldsymbol{p}_i^2}{2m_i} + V(\boldsymbol{q}) + \frac{\boldsymbol{p}_s^2}{2Q} + gkTln(s) \qquad \textit{Equation } 54$$

Where g is the number of degrees of freedom in the system. The first two terms describe the kinetic and potential energy of the system, whilst the third and fourth terms describe the kinetic energy of the virtual system s and the potential of s.

2.4.6.1. Langevin Dynamics

As previously discussed, Langevin dynamics can be thought of as molecular dynamics with inbuilt stochastic temperature coupling. It applies friction and a noise term, so that the change in velocity between adjacent time steps is described by the following stochastic differential equation:

$$\Delta\boldsymbol{v}_i = a_i\Delta t - \gamma_i\boldsymbol{v}_i\Delta \mathrm{t} + \xi\sqrt{2k_BT_{ref}\gamma_i\Delta t/m_i} \qquad Equation\ 55$$

The second term is a frictional term and the third is the noise term. The friction coefficient γ dampens the velocities, and ξ is a normally distributed number with random variance. The advantage of Langevin dynamics is that it works on degrees of freedom individually and therefore returns the correct distribution of kinetic energies for decoupled states. This is important for free energy calculations because in a number of windows the ligand will be decoupled from the system.

2.4.7. Barostats

Experiments are commonly run under conditions of constant temperature and pressure and so to replicate these experiments, it is often desirable to run simulations in the *isothermal-*

isobaric ensemble. A simulation in the isothermal-isobaric ensemble maintains constant pressure by changing the volume of the cell. The amount of fluctuation is related to the isothermal compressibility, κ:

$$\kappa = -\frac{1}{V}\left(\frac{\partial V}{\partial P}\right)_{\mathrm{T}} \qquad \textit{Equation 56}$$

One method for maintaining pressure, the *Berendsen barostat* [306], is analogous to the Berendsen thermostat. That is, the system is coupled to an external 'pressure bath'. The rate of change of pressure is given by:

$$\frac{dP}{dt} = \frac{1}{\tau_P}(\mathrm{P}_{ref} - \mathrm{P(t)}) \qquad \textit{Equation 57}$$

Where τ_P is the coupling constant, P_{ref} is the reference pressure that the pressure bath is at, and $P(t)$ is the pressure at time t. P(t) can be calculated via the virial theorem. The volume of the box is scaled by a factor λ and consequently, atomic coordinates are scaled by a factor of $\lambda^{1/3}$.

$$\lambda = 1 - \kappa\frac{\delta t}{\tau_P}(\mathrm{P(t)} - \mathrm{P_{ref}}) \qquad \textit{Equation 58}$$

The new positions are given by:

$$\boldsymbol{r}_i' = \lambda^{1/3}\mathbf{r}_{\mathrm{i}} \qquad \textit{Equation 59}$$

The Berendsen barostat yields the correct pressure of the system, but not the exact isothermal-isobaric ensemble. It is therefore typically only used in equilibration simulations. For production simulations, the *Parrinello-Rahman* [308] barostat was used. The Parrinello-Rahman barostat is an extended system method, similar to the Nosé-Hoover method for temperature coupling. It is an extension to the *Anderson barostat* that allows the system to scale anisotropically. The Anderson barostat adds the volume of the box as an additional

degree of freedom to the system. This can be thought of as adding an external piston to the system which is capable of changing the volume of the box. The piston has associated kinetic energy and potential terms which are added to give a modified Hamiltonian.

2.4.8. Long Range Interactions

In MD simulations, the calculation of non-bonded interactions is the most time consuming part of the simulation. This results from the fact that all particles interact with all other particles in the system, meaning the number of pairwise interactions calculated scales on the order of N^2. This section will discuss this problem for Lennard-Jones and Coulombic interactions, and how modern simulations deal with these problems.

2.4.8.1. Long-Range Lennard-Jones Interactions

The Lennard-Jones potential decays with an r^{-6} dependence on distance. This rapid decay means the contribution is relatively negligible at longer distances. For instance, at a distance of 2.5σ, the Lennard-Jones potential has a value of 1% compared to the value at σ. It is therefore possible to apply a cut-off distance (r_c) and assign values of 0 to interactions beyond this cut-off distance. This is equivalent to truncating the potential $v_c(r_{ij})$.

$$v_c(r_{ij}) = \begin{cases} v(r_{ij}) & r_{ij} \leq r_c \\ 0 & r_{ij} > r_c \end{cases}$$ *Equation 60*

Particles at the boundary may suffer from unphysical forces. A shifted force potential can be used to overcome this. Essentially, the potential is shifted by v_c and a linear term added, so that at r_c the derivative is zero:

$$v_{SF}(r_{ij}) = \begin{cases} v(r_{ij}) - v_c - (r_{ij} - r_c)\left(\dfrac{dv(r_{ij})}{dr_{ij}}\right) & r_{ij} \leq r_c \\ 0 & r_{ij} > r_c \end{cases} \quad \textit{Equation 61}$$

A switching potential function modifies the potential as such:

$$v'(r) = v(r)S(r) \quad \textit{Equation 62}$$

These can be applied across the entire potential function, or to gradually taper the value to 0 between two cutoffs, r_l and r_u. A simple linear form of the latter can be expressed as:

$$S = 1.0 \qquad r_{ij} < r_l \quad \textit{Equation 63}$$

$$S = (r_u - r_{ij})/(r_u - r_l) \qquad r_l < r_{ij} < r_u \quad \textit{Equation 64}$$

$$S = 0.0 \qquad r_u < r_{ij} \quad \textit{Equation 65}$$

S being 1 means the potential is normal, whilst a value of 0 means the potential is 0. However, this linear form suffers from discontinuities in energy and force at the two cutoff values. Therefore, a switching function should switch smoothly from 1 below r_l to 0 at r_u and satisfy the following requirements:

$$S_{r=r_l} = 1; \qquad \left(\frac{dS}{dr}\right)_{r=r_l} = 0; \qquad \left(\frac{d^2S}{dr^2}\right)_{r=r_l} = 0 \quad \textit{Equation 66}$$

$$S_{r=r_u} = 1; \qquad \left(\frac{dS}{dr}\right)_{r=r_u} = 0; \qquad \left(\frac{d^2S}{dr^2}\right)_{r=r_u} = 0 \quad \textit{Equation 67}$$

The subsequent Lennard-Jones potential is shown below.

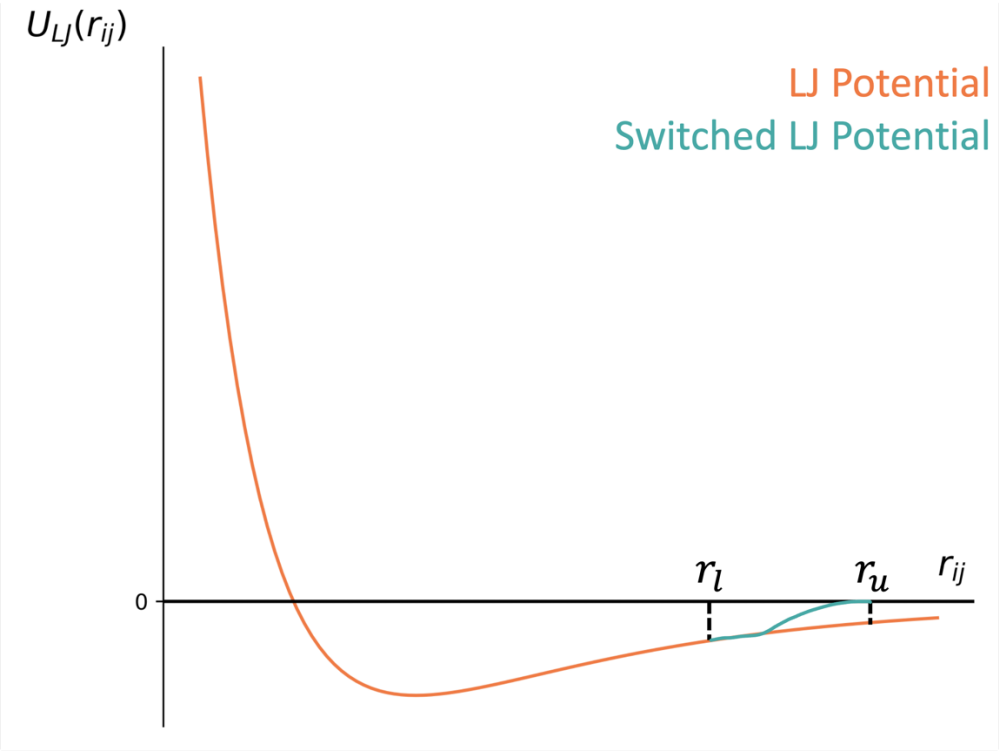


Figure 2.6: A qualitative representation of a Lennard-Jones potential with a switching function applied.

It is possible to go further than just applying a cutoff (r_c) for which at greater distances we do not calculate the interaction. This is because calculating the distances for all particles, and therefore which ones are closer than r_l, is costly. The solution is to use a neighbour list, as first proposed by Verlet. For a particular particle, the particles which fall within a cutoff distance (r_l) from particle i are tracked using a neighbour list. For particle i, only the Lennard-Jones interactions between itself and the particles in its neighbour list will be calculated. The neighbour list must be updated periodically. Too frequently is inefficient, but too frequently will lead to atoms not in the neighbour list becoming closer to i than the r_l threshold, leading to incorrectly calculated forces. Thompson proposed a solution to this; track the displacement of the particles since the last neighbour list update, and when one particle in the system is displaced by more than $r_l - r_c$, we update the neighbour list.

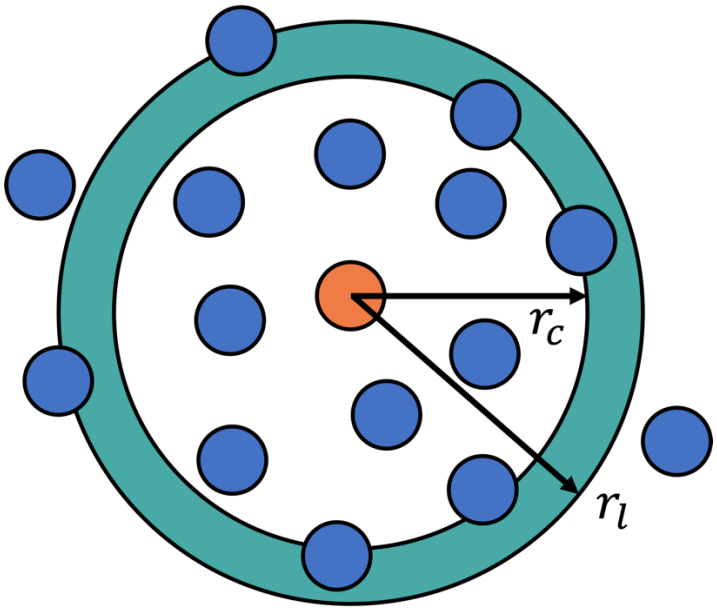


Figure 2.7: A schematic of the Verlet neighbour list scheme. A particle i (orange) only interacts with the particles that are within the cutoff radius r_c. Only particles that are within the radius r_l are tracked in the Verlet neighbour list. When a particle is displaced by more than $r_l - r_c$ the neighbour lists are updated. (Image adapted from Frenkel and Smit)

The utilisation of a cut-off can impact the energy of free energy calculations; specifically there is a dependence on the size of cut-off used. Gromacs tries to remove artefacts caused by the cut-off by applying an analytical correction to correct for energy and pressure. This method assumes isotropic distribution of solvent outside of the cut-off distance, which does not hold when simulating large biomolecules. A numerical correction for non-isotropic systems called 'EXP-LR' was proposed by Shirts *et al.* [309]. This involves including an additional free energy term in the alchemical cycle - ΔG_{EXP-LR}. The term corrects for the difference in free energy between using a cut-off and not by recalculating the energies of the existing system with no cut-off. For soluble systems, the value of this cut-off is typically on the order of 0.1 kcal/mol.

2.4.8.2. Long-Range Electrostatic Interactions

The electrostatic potential decays with r^{-1}, meaning a hard cut-off scheme is not appropriate. The interactions of a particle with particles in other unit cells, not just the unit cell it is in, must be considered. Methods based around Ewald summation, particularly particle mesh Ewald (PME), are commonly used [310,311]. The insight of Ewald was that the slowly converging term could be converted into two quickly converging terms:

$$\varphi(\boldsymbol{r}) = \varphi_{sr}(\boldsymbol{r}) + \varphi_{lr}(\boldsymbol{r}) \qquad \textit{Equation 68}$$

The first term handles short range interactions, and quickly decays in real space. The second term handles the long-range interactions and decays quickly in reciprocal space. In the Ewald method, a Gaussian charge distribution of opposite sign is added to every charge. The 'real space' summation over all the point charges of the system is now converted to a sum of the charges *plus* the neutralising distributions. A second charge distribution is added to the system of the same shape and size as the first distribution, but of opposite sign (therefore the same sign as the original charge), thus counteracting the first distribution (Figure 2.8). The summation of this second distribution is carried out in reciprocal space.

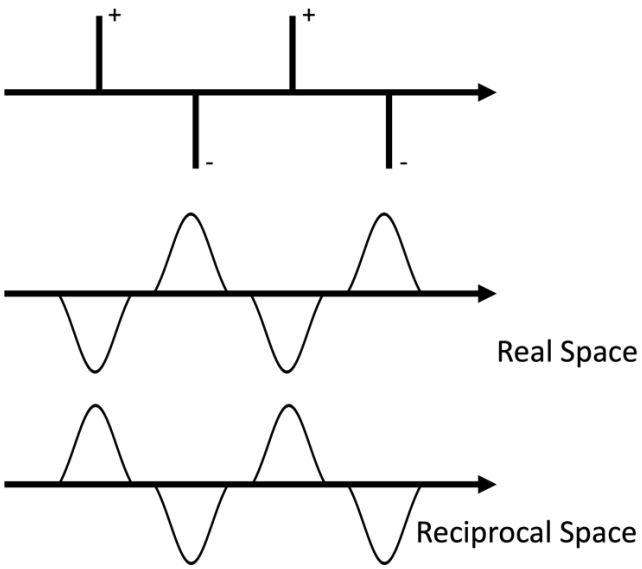


Figure 2.8: 1D illustration of the Ewald sum. Neutralising distributions are added to point charges in real space, whilst equivalent distributions of the same sign as the point charge are added in real space.

Ewald summation scales as $\mathcal{O}N^{\frac{3}{2}}$, limiting its utility to most systems of interest. PME extends on Ewald summation by changing the continuous representation of point charges to grid-based charges. This allows the use of Fast Fourier Transform calculations to be used, improving the speed and scaling (Nln(N)).

The use of PME and periodic boundary conditions requires the average of the electric potential over the box to be set to zero. This causes issues when the charge of a molecule changes during the course of a free energy calculation – for instance when annihilating a charged small molecule. This is because as the molecule is annihilated, the simulation box ceases to be neutral. To rectify this, a neutralising background charge is added to the entire simulation box. The magnitude of this neutralising background charge at a given point is dependent on the box size, and we therefore see a dependence on box size for free energy calculations of charged small molecules. Rocklin *et al.* proposed an analytical correction for this, plus other artefacts introduced by the periodic boundary conditions [312]. This correction can be applied post-hoc to calculations and is therefore widely used. However, this works on

the assumption that the system is isotropic, which is not true for membrane systems. Recent work by Wu *et al.* expanded on the Rocklin correction with the aim of making it applicable to membrane-bound systems [313]. The recommendation of this work, is to utilise a '*co-alchemical ion'* when running free energy calculations with a change in the total charge of the system in membrane systems. Essentially, one can either *co-annihilate* an ion of the opposite charge to the ligand, or charge an ion of the same charge (*charge-transfer*) (Table 2.2). The alchemical ion should be restrained sufficiently far from the charged ligand to prevent interactions between the two.

Charge of ligand	**Co-annihilation**	**Charge-transfer**
-1	Ligand: -1 → 0 Co-alchemical ion: +1 → 0	Ligand: -1 → 0 Co-alchemical ion: 0 → -1
+1	Ligand: +1 → 0 Co-alchemical ion: -1 → 0	Ligand: +1 → 0 Co-alchemical ion: 0 → +1

Table 2.2: The charge of a system can be kept neutral when annihilating a charged ligand by using a co-alchemical ion. An oppositely charged ion can be 'co-annihilated', whilst a similarly charged ion can be used for 'charge-transfer'.

2.4.9. Replica Exchange

Replica exchange simulation methods enhance sampling of the system by simultaneously running multiple simulations (replicas) of the system at different thermodynamic states. At regular points, replicas are then allowed to exchange some property, depending on an acceptance criteria. The first application of replica exchange was by Swendsen and Wang [311]. They ran multiple simulations at different temperatures $T_0 < T_1 < T_2$, where T_0 is the temperature of interest. At frequent intervals as the parallel simulations are progressing,

exchange moves are proposed between replicas. The moves are accepted on the basis of a Metropolis criterion:

$$p(1 \leftrightarrow 2) = \min\left(1, \exp\left(\left[\left(\frac{1}{k_B T_1} - \frac{1}{k_B T_2}\right)(U_1 - U_2)\right]\right)\right) \qquad \textit{Equation 69}$$

This states a move will be accepted with a probability dependent on the potential energy difference of the two systems. The higher temperatures are better able to explore phase space as they are more able to cross energy barriers because of the increased kinetic energy at higher temperatures. The simulation at T_0 remains thermodynamically accurate, but benefits from the increased search of phase space. As $Equation\ 69$ tells us, exchanges will only be accepted frequently if there is sufficient energy overlap between the two states. It is worth noting that as the potential energy of the system must be evaluated for the calculation of forces, the computational overhead remains low. The main limitation for using replica exchange is the requirement to run all replicas concurrently, which is not always feasible, depending on the compute resources available to the researcher.

In *Hamiltonian replica exchange* (HREX), the Hamiltonian of the system differs between repeats instead of temperature [314]. In equilibrium free energy calculations, the interactions of a ligand with the environment are scaled by a parameter λ. In the HREX simulations exchanges are proposed along the λ vector. These systems must show a level of phase space overlap to allow the accurate calculation of the final free energy. This means there is energy overlap between the adjacent states, so HREX can be used to enhance mixing of the different thermodynamic states. Gibb's sampling replica exchange is an extension that allows sampling across all states, not just adjacent ones. The exchange probability for HREX is:

$$p(1 \leftrightarrow 2) = \min\left(1, \exp\left(\left[\left(\frac{1}{k_B T}\right)(U_1(\boldsymbol{r}_2) - U_1(\boldsymbol{r}_1)) + (U_2(\boldsymbol{r}_1) - U_2(\boldsymbol{r}_2))\right]\right)\right) \qquad \textit{Equation 70}$$

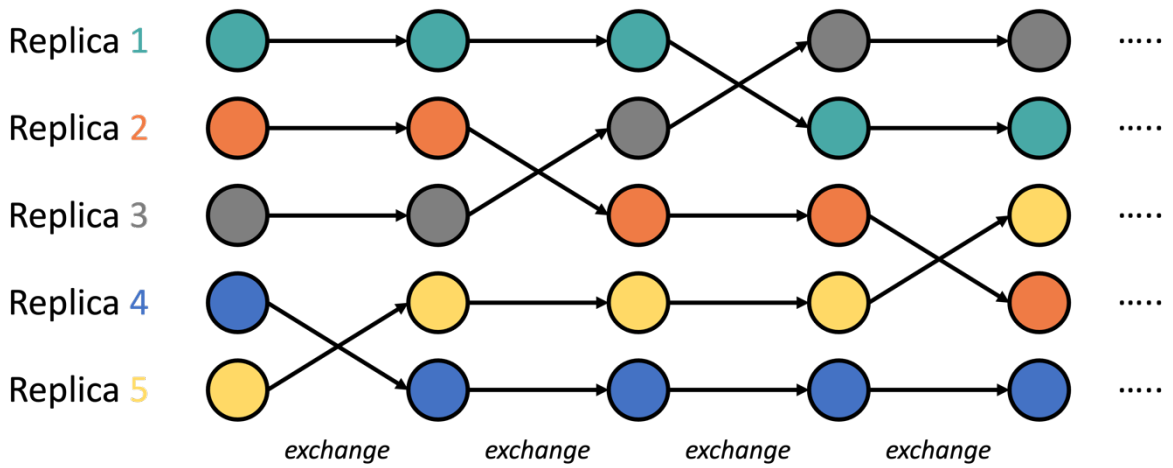


Figure 2.9: A simple schematic of the replica exchange method.

2.5. Free Energy Calculations

Free energy is probably the most important thermodynamic property, as the difference in free energy defines the likelihood of finding the system in either state. It determines whether a process is favourable, meaning it indicates whether work can be obtained from the process, or whether work needs to be put in to make the process happen. Properties such as ligand solvation, binding, stability, and phase transitions can all be described by their respective free energy differences. The accurate computational prediction of free energy remains something of a 'holy grail' in computational chemistry and it remains highly challenging for all but the simplest of systems. This section will define free energy, discuss some commonly used computational methods, and discuss the pros and cons associated with each method.

2.5.1. Defining Free Energy

The free energy difference of a system is directly related to the partition function of the system. In the canonical (NVT) ensemble, the Helmholtz free energy (A) can be written as:

$$A = k_B T ln(Q_{NVT}) = U - TS \qquad \textit{Equation 71}$$

Where Q_{NVT} is the partition function (that is the sum over all possible microstates in the system) in the NVT ensemble, U is the internal energy of the system, T is the temperature, and S is the entropy of the system. In the isothermal-isobaric (NPT) ensemble, the Gibb's free energy is defined as:

$$G = A + PV = U - T + PV = H - TS \qquad \textit{Equation 72}$$

Where P is the pressure, V is the volume of the system, and $U + PV = H$, the enthalpy of the system. Gibb's free energy differs from the Helmholtz free energy by including a term that incorporates the energetics of volume changes. We cannot calculate A or G of the system directly from MD simulations as we would require the system to sample the entirety of phase space, or at least the parts that make a significant contribution for a reasonable estimate. However, as the below equation for the Helmholtz free energy shows, points at high energy make an import contribution to A (because of the $\exp\left(\frac{\mathcal{H}(\boldsymbol{p}^N,\boldsymbol{r}^N)}{k_B T}\right)$ term), whilst MD simulations preferentially sample the lower energy regions of the system.

$$A = k_B T ln\left(\iint d\boldsymbol{P}^N d\boldsymbol{r}^N \exp\left(\frac{\mathcal{H}(\boldsymbol{p}^N,\boldsymbol{r}^N)}{k_B T}\right) p(\boldsymbol{p}^N,\boldsymbol{r}^N)\right) \qquad \textit{Equation 73}$$

However, we are often interested in calculating the difference in free energy between two states – that is the ΔA or ΔG, depending on the ensemble. The isothermal-isobaric ensemble is typically used in experiments, and ΔG is the relevant thermodynamic property to drug discovery. In this thesis, the isothermal-isobaric ensemble was used, so from now on, any reference to free energy is to the Gibb's free energy unless stated otherwise. In the NPT

ensemble, the Gibb's free energy of a ligand binding to a protein is related to the dissociation constant K_d by:

$$\Delta G^{\theta}{}_{binding} = k_B T ln\left(\frac{K_d}{c^{\ominus}}\right) \qquad \textit{Equation 74}$$

Where $c^{\ominus}$ is the standard concentration, which is 1 M by convention. $c^{\ominus}$ converts the free energy to a standard state, allowing free energies to be compared. $Equation\ 74$ refers to the free energy of binding for a single molecule. It is typical to refer to the binding free energy of a mole of ligand, so the equation is multiplied by the Avogadro constant N_A. $k_B N_A$ is the gas constant R.

2.5.2. Computational Methods for Free Energy Prediction

Many computational methods exist for attempting to predict the affinity of a small molecule for its target. These methods vary significantly in their accuracy and computational demands, and subsequently occupy different spaces in the drug discovery process. The least computationally demanding methods are based on *scoring functions*. These employ a physical or statistical potential to score a protein-ligand complex. These methods are fast but have consistently been shown to be inaccurate. Occupying the middle ground in terms of accuracy and computational cost are *end-point* methods. These involve the post-processing of snapshots from MD or Monte Carlo simulations. End-point methods typically utilise implicit solvent models for computational efficiency. It is worth noting that scoring functions and end-point methods do not return a free energy difference but instead they return a value designed to be a proxy for the free energy. *Pathway* methods, however, are theoretically rigorous from a thermodynamic point of view and therefore provide an estimate of the free energy of binding. Pathway methods are the most computationally expensive and accurate. They are based on molecular dynamics simulations. The two major classes of pathway methods are *alchemical* and *geometric* methods. Alchemical methods involve the removal of

an atom from its environment in a non-physical thermodynamic cycle. Geometric pathway methods instead use a physical pathway that is generated by applying a force to pull the molecule from the binding site and calculating the work required to do this.

2.5.3. Alchemical Methods

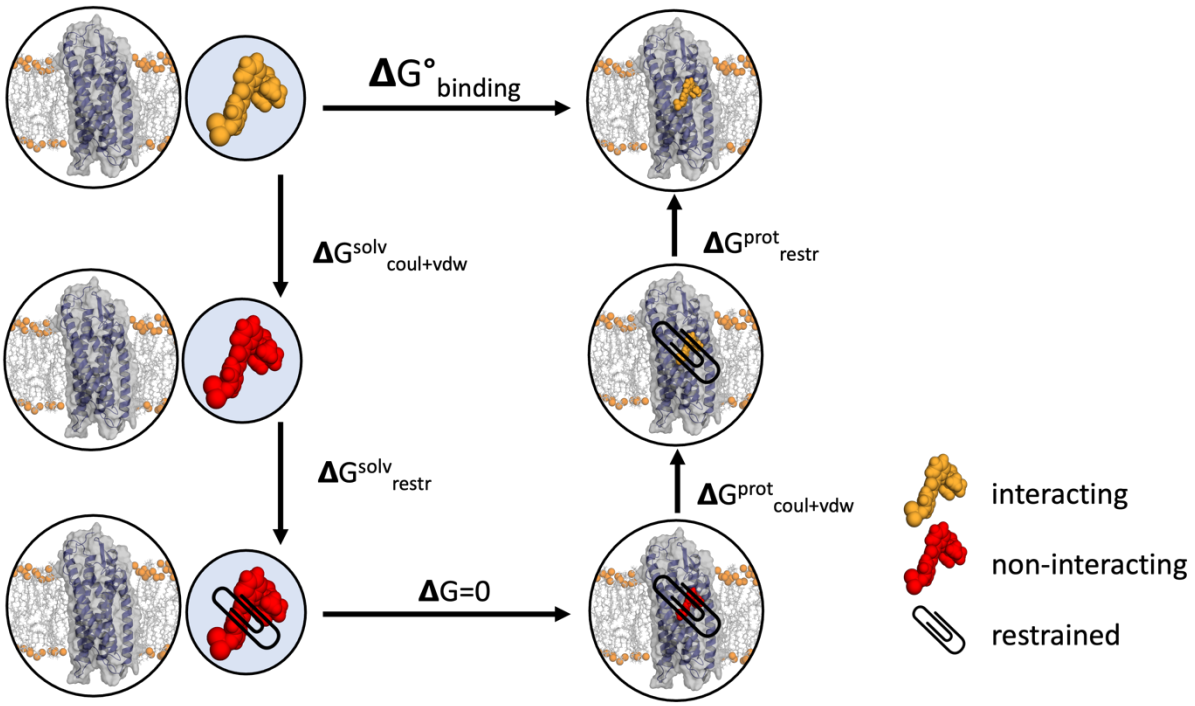


Figure 2.10: The alchemical cycle for an absolute binding free energy calculation.

As mentioned previously, alchemical free energy methods involve defining a non-physical pathway between two end-points (Figure 2.10). This non-physical pathway utilises the malleable nature of the functional form of modern force fields to scale the ligand's interactions with the rest of the system. The scaling of the interactions is controlled by coupling the Hamiltonian of the system to the variable λ. λ varies between zero and one, with the two end states reflecting physically meaningful states. This is all possible because free energy is a state function and is therefore independent of the path taken.

Alchemical methods typically outperform other methods for several reasons. First, they explicitly represent all atoms in the system. This allows them to better capture key interactions (for instance bridging waters) than methods that rely on implicit solvent. The explicit representation of atoms and correct sampling from the ensemble means that entropic contributions are automatically considered.

Equilibrium free energy methods involve running simulations at a constant λ for each simulation. These simulations are typically referred to as 'windows'. As we are decoupling an entire ligand, when the ligand is weakly/non-interacting it is free to leave the binding site. This would make sampling these states prohibitively hard, so instead we apply a restraint scheme. Many restraints have been discussed, but I utilised a restraint system proposed by Boresch *et al*. [68]. Boresch restraints maintain the ligand in a particular pose in the binding site by using one distance restraint, two angular restraints, and three dihedral restraints. The impact of these restraints can be accounted for analytically for the ligand in solution.

$$\Delta G_{rest}^{solv} = -k_B T ln \left[\frac{8\pi 2V^0}{r_0^2 sin\theta_{A,0} sin\theta_{B,0}} \cdot \frac{\left(K_r K_{\theta_A} K_{\theta_B} K_{\phi_A} K_{\phi_B} K_{\phi_C}\right)^{\frac{1}{2}}}{(2\pi k_B T)^3} \right] \quad \textit{Equation 75}$$

Where K is the force constant of a harmonic potential, r_0 is the equilibrium value of the distance restraint, θ_A and θ_B are equilibrium values of the angles, and $\phi_A, \phi_B,$ and ϕ_C are the equilibrium values of the dihedral restraints. Boresch restraints can introduce some instabilities, leading to failure of the simulation. To overcome this, restraints were derived using MDRestraintsGenerator [76]. In essence, this takes an unrestrained vanilla MD simulation and looks for protein-ligand distances, angles, and dihedrals that remain stable over the

course of the simulation. It then returns suggested atoms, equilibrium values, and force constants.

Boresch restraints require that the system is isotropic for the analytical correction to be applied as they include a standard state correction. In the case of calculating the free energy of partitioning from the bilayer to the protein, an alternative restraints scheme is required. I used the restraints scheme from Salari *et al.* [315]. The implementation of these restraints was done in a similar manner to Corey *et al.* [316].

$$\frac{Z_b}{Z_f} = \left(\frac{Z_f^{=}}{Z_f}\right)\left(\frac{Z_f^{-} Z_g^{=}}{Z_f^{=}}\right)\left(\frac{Z^{\circ}{}_g}{Z_g^{=}}\right)\left(\frac{Z^{\circ\hat{}}_g}{Z^{\circ}{}_g}\right)\left(\frac{Z_b^{\hat{}}}{Z_f Z_g^{\hat{}}}\right) \qquad \textit{Equation 76}$$

Where the configuration partition function for each state is shown as Z with the restraint scheme imposed. =, -, o and ^ correspond to the restraint schemes. The lipid states are f=free, b=bound, g=gas phase. Terms 1 (the cost of imposing bulk-phase restraints) and 3 can be calculated analytically:

$$Term\ 1: \quad \frac{Z_f^{=}}{Z_f} = \mathrm{n}_L \alpha \qquad \textit{Equation 77}$$

Where n_L is the number of ligands (1) and α is the fraction of the simulation box enclosed by $=$ restraints. Term 3 can be calculated as such:

$$Term\ 3: \quad \frac{Z^{\circ}{}_g}{Z_g^{=}} = \frac{V^o}{V^{=}}\frac{\Omega^o}{\Omega^{=}} = \frac{2\pi r_R^3}{3 Z_R A^{=}}\frac{1}{(1 - cos\theta_R)} \qquad \textit{Equation 78}$$

$Z^{\circ}{}_g$ and $Z^{=}{}_g$ are the partition functions under the o and = restraints systems respectively. Where the 2 terms correspond to the ratio of volumes available in the gas phase (V^o and $V^{=}$)

and orientational aspect of the restraints ($\Omega^o \; and \; \Omega^=$). r_R and θ_R are the values used for the distance and angle terms in the bulk restraints. z_R is the distance used in the gas phase. Term 2 is calculated from the annihilation of the molecule in bulk membrane, term 4 is the cost of restraining the molecule in the gas phase, and term 5 is calculated from annihilating the molecule when bound to the receptor.

We will limit our discussion to equilibrium free energy methods, as they are the ones primarily used in this thesis, but non-equilibrium methods based on the Jarzynski equality or the Crook's fluctuation theorem can be used [82–84].

2.5.3.1. The Zwanzig Relationship

Let us consider two well-defined states X and Y, where X and Y could be an apo protein and protein-ligand complex, or two different small molecules in solution. For this chapter, I will be focusing on absolute binding free energy calculations. Therefore, states X and Y can be considered the apo protein and ligand-bound complex.

The difference in free energy between the two states can be written as:

$$\Delta G = G_Y - G_X = k_B T ln\left(\frac{Z_Y}{Z_X}\right) \qquad \textit{Equation 79}$$

$$\Delta \mathrm{G} = \beta^{-1} ln\langle e^{-\beta[-(\mathcal{H}_X - \mathcal{H}_Y)]} \rangle_0 \qquad \textit{Equation 80}$$

Where Z are the partition functions for particular states, $\beta \; is \; 1/k_B T$, and $\mathcal{H}$ is the Hamiltonian. This is the *Zwanzig relationship* and forms the basis for *exponential averaging* [317]. Exponential averaging involves averaging the exponential of the energy differences. Whilst this is sometimes called free energy perturbation (FEP), this term is often used to refer

to alchemical methods in general. I will therefore refer to this as exponential averaging to avoid confusion.

Simulations are run at constant λ values. The Zwanzig relationship requires phase space overlap to return accurate free energy estimates. For real systems, using only two lambda values will return insufficient phase space overlap. We therefore additionally run simulations at intermediate λ values to improve overlap.

$$\Delta G = G_Y - \mathrm{G_X} = (G_Y - \mathrm{G_N}) + (G_N - \mathrm{G_{N-1}}) \ldots + (G_1 - \mathrm{G}_X) \qquad \textit{Equation 81}$$

$$\Delta G = -k_B T ln \left[\frac{Z(Y)}{Z(N)} \cdot \frac{Z(N)}{Z(N-1)} \ldots \cdot \frac{Z(1)}{Z(X)}\right] \qquad \textit{Equation 82}$$

Where N is the number of intermediate simulations. Picking an appropriate value for N is vital for ensuring enough phase space overlap. There are commonly used N but one should check there is sufficient overlap in the simulations. The important thing to note is that these intermediate terms cancel, meaning the ΔG is only dependent on the end-points and is independent of the number of intermediates. Exponential Averaging is exact but has been shown to converge slowly with regards to the amount of data collected. Therefore, other estimators based on exponential averaging are more commonly used.

2.5.3.2. Bennet's Acceptance Ratio

Exponential averaging will, provided there are sufficient samples, provide the same answer in forward and reverse directions. Therefore, one could simply average the results of the two directions to improve upon exponential averaging. However, Bennett instead noticed that due

to the relationship between the distributions of potential energy in forward and reverse directions, a more robust method could be developed [318,319]. Bennett's first observation was:

$$\Delta G = k_B T \, ln\left(\frac{\langle a(\boldsymbol{x}) \exp[-\beta(H_0 - H_1)]\rangle_1}{\langle a(\boldsymbol{x}) \exp[-\beta(H_1 - H_0)]\rangle_0}\right) \qquad \textit{Equation 83}$$

This holds for any function $a(\boldsymbol{x})$. Bennett then utilised variational calculus to find a function $a(\boldsymbol{x})$ that minimises the variance of the free energy. This gave the implicit formula for ΔG that can be solved numerically.

$$\sum_{i=1}^{n_1} \frac{1}{1 + exp\left[\ln\left(\frac{n_1}{n_0}\right) + \beta\Delta H_{10} - \beta\Delta G\right]} - \sum_{j=1}^{n_0} \frac{1}{1 + exp\left[\ln\left(\frac{n_0}{n_1}\right) + \beta\Delta H_{01} - \beta\Delta G\right]} = 0 \qquad \textit{Equation 84}$$

Where n_1 and n_0 are the number of samples from each state. It returns the maximum likelihood estimate of the free energy, given the samples from states X and Y.

2.5.3.3. Multistate Bennet's Acceptance Ratio

An extension to BAR, the *multistate Bennet's acceptance ratio (MBAR)*, was proposed by Shirt's *et al.* [319]. In a calculation with K states, $K \times K$ weighting functions $a(\boldsymbol{x})$ are derived to minimise the uncertainty of the free energy estimate. This means MBAR considers all K states simultaneously. If K=2, MBAR and BAR are identical. MBAR can be considered a zero-width Weighted Histogram Analysis Method (WHAM). The Python implementation of MBAR was utilised throughout this project. MBAR has been shown to be the lowest variance estimator.

2.5.3.4. Thermodynamic Integration

Thermodynamic integration (TI) is a gradient based approach to calculating free energy differences. The derivative of the Hamiltonian is taken with respect to λ.

$$\frac{dG}{d\lambda} = \frac{d}{d\lambda}\int e^{-\frac{H(\lambda,\boldsymbol{x})}{k_B T}}\,\mathrm{d}\mathbf{x} = \langle\frac{dH(\lambda,\boldsymbol{x})}{d\lambda}\rangle_\lambda \qquad \textit{Equation 85}$$

$$\Delta G = \int_0^1 \langle\frac{dH(\lambda,\boldsymbol{x})}{d\lambda}\rangle_\lambda \qquad \textit{Equation 86}$$

In the canonical ensemble, the Hamiltonian reduces to the potential energy U as the velocities are independent of λ (assuming λ does not scale the particle masses). In the NPT ensemble, one must also consider the contribution of the change in volume to the ΔG. However, this is typically negligible. Numerical integration is required to convert the$\langle\frac{dH(\lambda,\boldsymbol{x})}{d\lambda}\rangle$ values to ΔG values. Any numerical integration method can be applied, but the trapezoid rule is typically used due to its simplicity [320]. Unlike the other estimators discussed, TI is not based on the Zwanzig relationship and so can provide a useful comparison to estimates from Zwanzig relationship based methods.

3. Free Energy Calculations for $P2Y_1R$

3.1. Table of Contents

3. FREE ENERGY CALCULATIONS FOR P2Y$_1$R 110

3.1. TABLE OF CONTENTS 111

3.2. TABLE OF FIGURES 112

3.3. LIST OF TABLES 113

3.4. INTRODUCTION 114

3.5. METHODS 117

3.6. INVESTIGATING LIPID DYNAMICS 121

3.7. P2Y$_1$R BPTU ABFE VALIDATION 127

3.7.1. ABFEs are accurate and precise 127

3.7.2. Multiple embeddings provide a more realistic estimate of the error 128

3.7.3. Results can be reproduced using multiple forcefields 130

3.7.4. Lipid sampling in the apo state 133

3.7.5. Lipid sampling in the partially interacting state 135

3.7.6. Can we improve lipid sampling? 137

3.8. FREE ENERGY CALCULATIONS OF P2Y$_1$R SAR 140

3.8.1. Pseudo-symmetric groups require consideration 142

3.8.2. ABFEs Accurately Rank P2Y$_1$R SAR 144

3.9. ABFES FOR BPTU SELECTIVITY 148

3.10. DISCUSSION 156

3.11. CONCLUSION 162

9. SUPPLEMENTARY MATERIAL: CHAPTER 3 288

3.2. Table of Figures

FIGURE 3.1: BPTU P2Y$_1$R INTERACTIONS....122

FIGURE 3.2: BPTU/POPC INTERACTIONS....124

FIGURE 3.3 PROTEIN/POPC INTERACTIONS....125

FIGURE 3.4: CRYSTALLOGRAPHIC ABFES....128

FIGURE 3.5: ABFES FROM MULTIPLE EMBEDDINGS....130

FIGURE 3.6: ABFES WITH MULTIPLE FORCEFIELDS....132

FIGURE 3.7: ABFE EQUILIBRATION....135

FIGURE 3.8: $Lipid\ Sampling\ at\ \lambda VdW = 0.7$....136

FIGURE 3.9 LIPID HREX....138

FIGURE 3.11: P2 SAR STRUCTURES....141

FIGURE 3.12: 6B CONFORMATIONS....142

FIGURE 3.13: 6B TERMINAL DIHEDRAL....144

FIGURE 3.14: P2Y$_1$R SAR ABFE RESULTS....147

FIGURE 3.15: P2Y$_1$R MUTATIONS....154

SUPPLEMENTARY FIGURE 8.1: UNBIASED SIMULATIONS RMSD....288

SUPPLEMENTARY FIGURE 8.2: NON-EQUILIBRIUM WORK DISTRIBUTION.....289

SUPPLEMENTARY FIGURE 8.3: P2Y$_1$R BPTU SAR GRID....290

SUPPLEMENTARY FIGURE 8.4: MM/GBSA 6B CONFORMER PREDICTION....291

3.3. List of Tables

TABLE 3.1: LIPID HREX RESULTS 139

TABLE 3.2:MUTATION RESULTS 151

3.4. Introduction

Prior to this work, alchemical absolute binding free energy (ABFE) calculations had largely focused on soluble proteins [321–325]. Recently, there has been a growth in the application of relative free energy (RBFE) calculations to membrane proteins [326].

The application of ABFEs remains largely retrospective or pseudo-prospective [74], with few examples in prospective projects as the application of ABFEs remains challenging [327]. The sources of errors for soluble proteins are reasonably well understood [328]. They typically come under forcefield problems and sampling limitations.

Forcefield problems can be assessed by using multiple forcefields for the ligand or the representation can be improved via the use of polarisable force fields [329] or additional consideration of the torsional parameters [294]. These methods add expense in both running and setting up the simulations, so their implementation is typically only done when previous simulations do not work. This decision requires knowledge of the experimental value and/or significant knowledge of the system or chemical intuition. Knowing whether more expensive representations is needed *a priori* remains challenging.

Furthermore, lipid-exposed sites present additional difficulties for absolute binding free energy calculations. The first major difficulty is the role of the lipid in 'solvating' the binding site. This provides a challenge for free energy calculations as it is not clear how well we will be able to sample all of these interactions given the expected slow dynamics of lipids. This is a concern even for single species bilayers but is even more problematic for more realistic multiple species bilayers. There remains a concern that for the short simulations used in free

energy calculations, the phase space sampled will be dependent on the starting configuration of the lipids.

Another concern centres around the fact that we start non-interacting windows (which should resemble the apo state) from the holo state. This causes multiple issues around the sampling of the apo state of the receptor and generating realistic apo lipid configurations. GPCRs are inherently dynamic entities, and the sampling of even simple protein dynamics causes significant issues. It is not clear at this stage how much the dynamics of GPCRs will hinder our ability to get accurate free energy results.

The experimental data for lipid-exposed sites is often competition radioligand binding, where typically an IC_{50} is reported. Although K_d can be derived by the Chun-Prussof equation, great care should be taken when comparing different experimental protocols, particularly when the solute concentrations differ. Interpretation of experimental values is further complicated by the fact that it is often not clear what the exact membrane composition is. This poses a problem in terms of validating our calculations, but the challenges of generating high-quality experimental data also highlights a potential niche for ABFEs for lipid-exposed sites.

There is a need to run a comprehensive set of absolute binding free energy calculations on a series of molecules binding at a lipid-exposed binding site to assess how great these challenges are.

BPTU bound to $P2Y_1R$ is a good system to test the performance of ABFE calculations to work at lipid-exposed sites. It has the essential components for running free energy

calculations; there is a crystal structure of BPTU bound to $P2Y_1R$ [330] and there is an extensive reported SAR series to allow us to test our predictions [331].

The SAR series provides a few advantages. First, docking scoring algorithms have not trained or validated on lipid-exposed sites. The shared scaffold allows us to place the molecules in the binding site. Second, when using docked structures, uncertainties in the binding pose may increase the uncertainty in predictions. This is important to be considered before running ABFEs at lipid-exposed sites prospectively but may make it more difficult to assess what is caused by the factors we are more interested in, such as the role of lipids and the dynamic nature of GPCRs. BPTU is a neutral ligand and the performance of neutral ligands is typically better than charged ones [76]. The binding site is also deep enough in the bilayer that water penetration into the bilayer is unlikely. BPTU is also an antagonist with the BPTU/$P2Y_1R$ structure being in an inactive state and therefore the complication of modelling two states (apo and holo) is not present. Previous research has suggested that metadynamics may obtain a good estimate of the binding free energy [332]. Work published during the undertaking of this project also suggests RBFE calculations can accurately rank BPTU SAR [258]. However, ABFEs are significantly more challenging to implement and there remains a need to assess how well ABFEs can work at these kinds of binding sites.

In this chapter, I ran ABFEs of the BPTU SAR compounds binding to the $P2Y_1R$ utilising current best practices [333]. I aim to properly understand the causes of errors and describe methods on how to identify and overcome them. This is with the aim of advancing the utility of ABFEs at lipid-exposed such that they can be used in a more prospective manner.

3.5. Methods

The crystal structure of BPTU bound to $P2Y_1R$ (PDB ID: 4xnv) was used as the starting point for these simulations [271]. This was selected as it contains the ligand of interest. 2 other structures of 4xnv exist but neither have BPTU bound. 4xnw shows an inactive structure of $P2Y_1R$ bound to MRS2500. MRS2500 is an antagonist that binds at the endogenous orthosteric site. Interestingly, there are no major sidechain differences around the BPTU binding site and the global conformations are highly similar. These similarities suggest these may be similar to the inactive state in the absence of ligand. 7xxh is a cryo-EM active state structure which shows minor sidechain torsional differences from the inactive structures but no major conformational changes. Running simulations from these structures would be interesting to see if they show structural differences.

Missing atoms, residues, and loops were modelled using MODELLER [334]. Of the 50 generated models, the top 10 models were rescored using the QMEAN scoring function [335,336]. The best QMEAN scoring model was selected. The protein was capped with N-methyl and acetyl caps using PyMol. Mutations were introduced using MODELLER. For the $P2Y_{12}R$ system, the AlphaFold model of the inactive state as available on GPCRdb was used [24,337].

Simulations were performed using Gromacs 2019 [338]. The forcefield used was AMBER99SB-ILDN [288]. Unless otherwise stated, the ligand was parameterised with GAFF2 [339] using antechamber. Gromacs topologies were generated using ParmEd [340]. The protonation state of the ligand was the major state in solution with Marvin Sketch used for the pKa calculations. In some simulations, the ligand was represented with OpenFF v2.0.0 (Sage) [341] or with optimised torsional parameters from BespokeFit [294]. The protein was embedded into the

bilayer using InflateGro [342]. This consisted of approximately 170 lipids total. The use of CHARMM-GUI [343] for embedding was also validated. The lipids were represented using SLipids 2016 [344] or SLipids 2020 [345]. The systems were solvated in a box of approximately $8 \times 8 \times 10\ nm$. The box was solvated (~ 21,000 water molecules) and sodium and chloride ions (~140 ions total) were added to neutralise the system and to create a concentration of 0.15 M. The final system size was approximately 70,000 atoms.

The water model used was TIP3P [346]. Electrostatic simulations were handled using particle mesh Ewald method (PME) [347,348]. A cut-off of 1.2 nm was used. A cut-off of 1.2 nm was used for vdW interactions. Post embedding, the systems were equilibrated for 100 ns with the protein restrained to facilitate equilibration of the lipids.

For ABFE calculations, the vdW and electrostatic interactions were annihilated (using the mdp flag $couple - intramol = yes$). The vdW and electrostatic interactions were annihilated using a linear pathway of $\Delta\lambda = 0.05$ and 0.1 respectively. The restraints were added using a non-linear pathway (0.0, 0.01, 0.025, 0.05, 0.075, 0.1, 0.15, 0.2, 0.3, 0.5, 0.75, 1.0). These values represent the asymmetric energy profile of adding restraints. The coulombic and vdW were separated into distinct vectors due to differences in the free energy kernel's different abilities to offload work to the GPU. This resulted in 43 simulations total for the complex and 32 for the ligand in solution. A soft-core potential for decoupled vdW interactions was used [349].

For each window, the system was equilibrated before the production run. The production runs, unless specified otherwise, were 20 ns for the complex and 15 ns for the ligand in solution. These are standard length simulations and longer simulations did not yield improved

results. We also show the degrees of freedom we need to sample are too slow to be sampled in realistic ABFE windows. The equilibration schemes began with a 10,000 step energy minimisation using a steepest descent algorithm. For the ligand in solution, this was followed by 100 ps in the canonical ensemble with harmonic position restraints with a force constant of 1,000 kJ mol^{-1} nm^{-2} on the solute heavy atoms. The temperature was coupled to the reference temperature of 298.15 K using Langevin dynamics. It was then further equilibrated in the isothermal-isobaric (NPT) ensemble for 100 ps using the Berendsen weak coupling algorithm [306]. This was then followed by a further 500 ns using the Parinnello-Rahman barostat [350] before restraints were removed and the system was equilibrated for a final 5 ns. The initial equilibration scheme for the complex simulations follow the same process. Simulation quality was monitored by checking the temperature, pressure, potential energy, RMSD of the protein and the ligand, and lipid order parameters.

Analysis was carried out using the Python implementation of pymbar v3.0.3 [319] via alchemlyb. Uncorrelated samples were generated by calculating the autocorrelation time. The error is reported as the standard deviation of at least three repeats. These repeats were run from separate embedding processes. During the complex simulations, the ligand was restrained in the binding site using Boresch restraints [68]. These involve restraining an intermolecular distance, 2 angles, and 3 dihedrals with force constants of 41.8 kJ mol^{-1} nm^{-2}. The contribution of these restraints to the free energy is calculated analytically for the ligand in solution. The optimal atoms to use for these restraints was picked from a 20 ns simulation using MDRestraintsGenerator [76]. Hamiltonian replica exchange (HREX) was utilised as implemented in Gromacs. 1,000,000 exchanges were attempted every 1,000 steps with exchanges attempted between all λ windows.

Trajectory analysis was carried out using MDAnalysis [351] and ProLiF [352]. Ligand structures were plotted using ChemDraw and RDKit 2023.03.1. Visualisation was carried out using PyMol [353]. RDKit was used for calculating protrude distances. Throughout the text 95% CI are reported in square brackets.

3.6. Investigating Lipid Dynamics

To better understand the $P2Y_1R$ system, I began my investigation with unbiased MD simulations. I was particularly interested in understanding the lipid sampling in both the presence and absence of ligand. This is essential as the lipid constitutes a major part of the ligand's environment and is therefore expected to contribute significantly to the free energy of binding.

To assess the lipid sampling, I ran 5 x 1 μs unbiased MD simulations in both the holo and apo states embedded in a pure POPC bilayer. POPC was selected as it is a neutral lipid, it is commonly found in a variety of membranes, and it is widely used allowing for easier comparison with other studies. The crystal structure of BPTU bound to $P2Y_1R$ (PDB: 4XNV) was used as a starting point for the holo and apo simulations. The simulations were run from separately embedded complexes to generate different starting configurations. The protein remained stable throughout the simulations, as did the ligand in the holo simulations (Supplementary Figure 8.1).

I aimed to characterise the interactions the ligand forms with the protein and POPC, and our ability to sample them (Figure 3.1A). The percentage of frames exhibiting specific interactions was tracked using the ProLiF Python library [352]. ProLif defines the interactions via the properties of the atoms involved and the distance and angles between them. We note they define a hydrophobic interaction as any interaction where the atoms are <4.5 Å from one another, and vdW interactions are defined as those less than the sum of the vdW radii plus a tolerance factor of 0.6 Å.

The ligand forms exclusively π stacking, hydrogen bond donor, vdW contacts, and hydrophobic interactions with the protein (Figure 3.1B). The π stacking interactions occur between F62$^{1.43}$ and F66$^{1.47}$ and the ureido phenyl ring of the ligand only. The hydrogen bond between the ligand's urea and the backbone of carbonyl L102$^{2.55}$. The hydrophobic and vdW contacts are the most prevalent form of interactions. The average number of contacts is consistent between repeats and this low standard deviation is present from early in the simulation (Figure 3.1C). The observed interactions are consistent with the crystal structure. This suggests that sampling the ligand protein interactions may not be problematic.

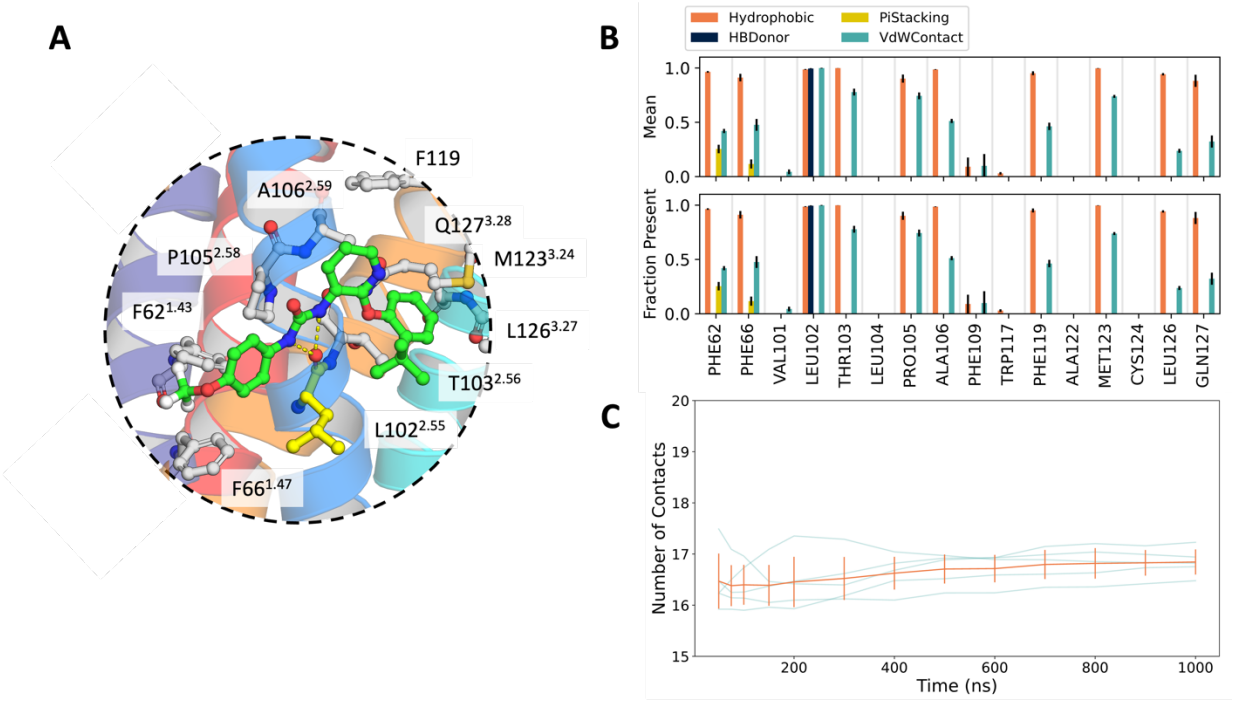


Figure 3.1: BPTU P2Y$_1$R Interactions. (A) The BPTU (green sticks) binding site on P2Y$_1$R. Residues that form interactions with the ligand are shown in sticks (grey for hydrophobic interactions and yellow for residues that form hydrogen bonds) (B) The interactions formed between BPTU and the residues of the binding site as defined by ProLIF. Hydrophobic and vdWContact are most prevalent. The ligand acts as a hydrogen bond donor with L102$^{2.55}$ and forms π stacking interactions with F62 and F66. Error bars are the standard deviation across 5 repeats. (C) The evolution of the total number of contacts formed across the simulation. The average results across 5 simulations are shown in orange. Error bars are the standard deviation across the 5 repeats. Individual repeats are shown in cyan. The number of contacts remains stable across the 5 repeats with a slight decrease in the standard deviation as the simulations progress (orange line).

We see that POPC forms exclusively vdW contacts and hydrophobic interactions with the ligand (Figure 3.2A). BPTU forms on average 4.8 ± 0.3 hydrophobic interactions and 3.4 ± 0.1 vdW interactions with the bilayer. This is consistent with the buried position of the ligand in the bilayer. It is worth noting the standard deviation between repeats remains high for a significant fraction of the simulation (Figure 3.2B). For instance, after 400 ns $\sigma_{hydrophobic} = 0.54$ and after 50 ns $\sigma_{hydrophobic} = 0.71$. This indicates we are unable to sample the ligand-lipid interactions fully on these timescales.

Interestingly, error estimates for some of the repeats from block averaging of single repeats are low. This indicates individual repeats may appear well converged whilst only sampling a fraction of the relevant phase space. This highlights the need for multiple simulations to get a more accurate picture of the error. I broke down the 1 μs simulations into 20 ns blocks. Here we can see significant spread amongst the 20 ns chunks, for instance for repeat 1 $\sigma_{hydrophobic} = 0.93$ (Figure 3.2C). This is relevant as this is the length of a typical window in an ABFE calculation – indicating the lipid sampling issues are significant on this timescale. Increasing the length of the blocks to 100 ns reduces the standard deviation but it remains high (for repeat 1 $\sigma_{hydrophobic} = 0.79$).

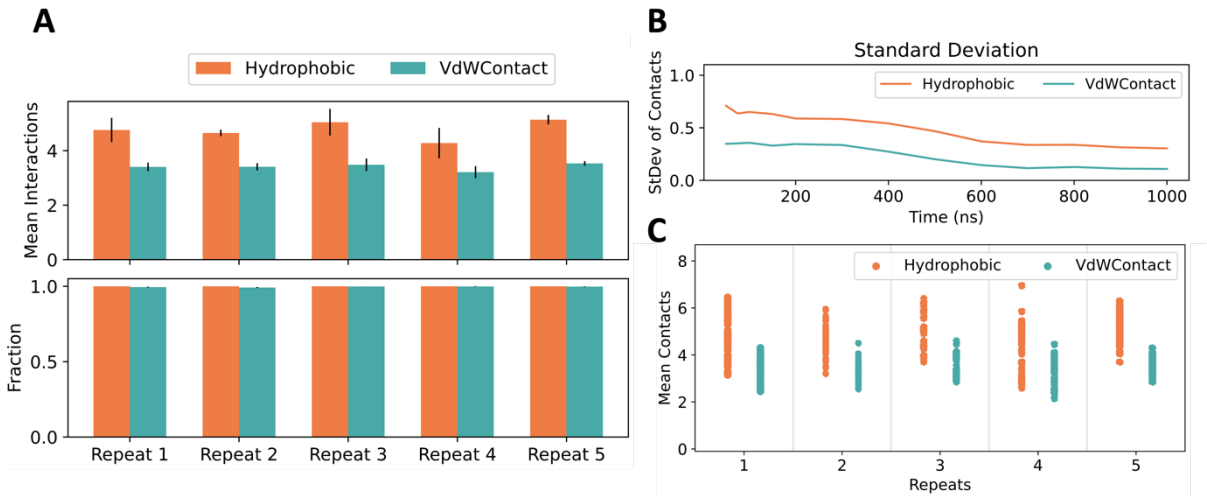


Figure 3.2: BPTU/POPC Interactions. (A) The number of contacts (upper) and the fractions of frames which contain at least one contact (lower) from $5 \times 1\ \mu s$ *unbiased simulations. Error bars are from block averaging. (B) The standard deviation of the number of contacts across the simulations (C) The number of contacts formed across* $20\ ns$ *chunks of the* $1\ \mu s$ *simulations*

We also assessed our ability to sample the lipid's interactions with the protein on a per residue basis using ProLIF. I began by looking at the protein as a whole. Figure 3.3A shows how the standard deviation of the number of interactions formed for each residue between repeats decreases as the simulation time increases. This shows that we cannot realistically sample the whole lipid ensemble across the order of a typical free energy window ($20\ ns$), as one may have expected.

Looking at the interactions formed with the residues that form the binding site, we see that there are only two interaction types – vdW contact and hydrophobic (Figure 3.3B). The average number of hydrophobic interactions for all binding site residues is 23.4 and the average number of vdW contacts is 9.2. The standard deviation of the number of contacts is low across the 5 simulations. However, the standard deviation remains relatively high for a significant proportion of the simulation (Figure 3.3C). For instance, for L126, the standard

deviation of the number of hydrophobic contacts after 100 ns is 0.44. This number drops to about 0.1 after 1 μs of simulation. This raises questions about the ability to sample the relevant protein/POPC interactions on the timescale of an ABFE window. To get a better idea of this, I broke the simulation down into 20 ns blocks. Using L126 as a representative residue, we see the number of contacts varies significantly within repeats (Figure 3.3D). The distributions show a similar pattern between repeats but there is some deviation in the width of the repeats.

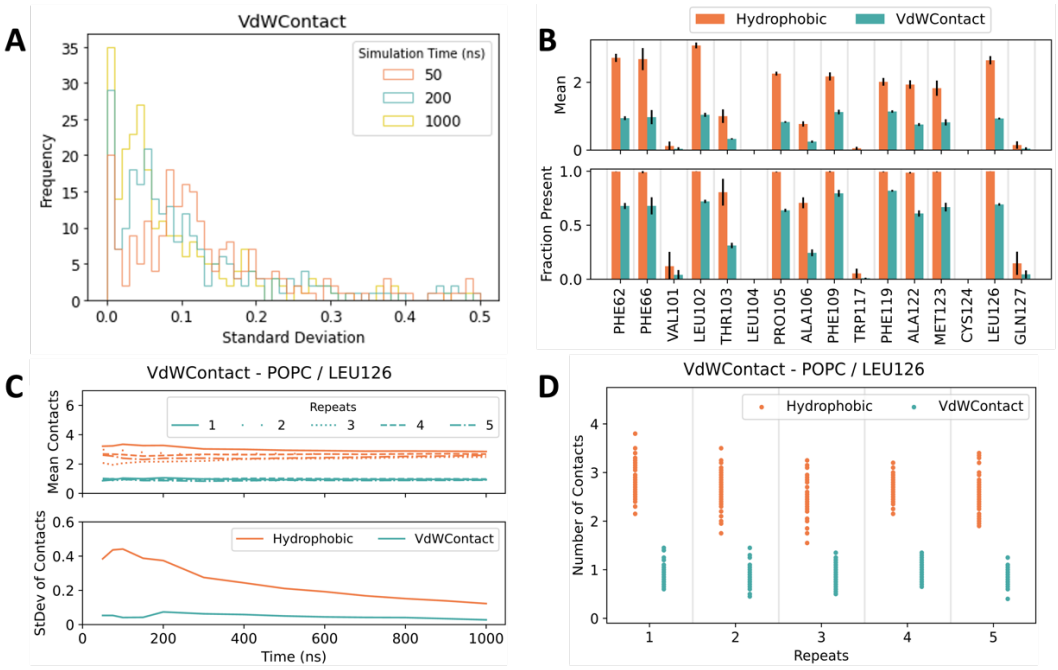


Figure 3.3 Protein/POPC Interactions. *(A) The standard deviation of vdW contacts for all residues across 5 × 1 μs simulations at 50 ns (orange), 200 ns (teal), and 1000 ns (yellow). (B) The number of contacts (upper) and the number of frames which contain at least 1 interaction (lower) between the binding site residues and POPC. (C) The average number of contacts formed between L126 and POPC as a function of time (upper) and how the standard deviation of this between repeats evolves over the simulations (lower) (D) The number of contacts formed in 20 ns sections of the simulations.*

Here I have shown that sampling the lipid is challenging even in a simple pure POPC bilayer and even on timescales significantly longer than current ABFE windows. Consistent with the

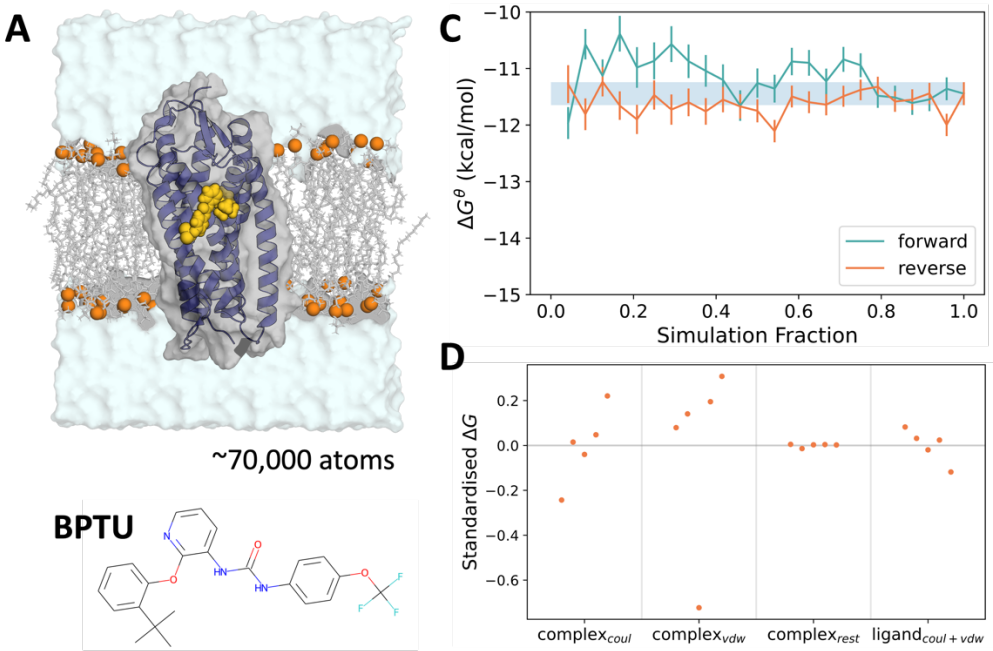
buried position of the binding site, we see that the majority of interactions are hydrophobic. I have shown that we can expect to sample only a fraction of the protein/POPC interactions over a 20 ns simulation. We also see some inconsistencies in our ability to sample the BPTU/POPC interactions at this timescale. This contact analysis makes no comment on the strength of interactions formed and so may not translate into significant errors in ABFEs.

3.7. $P2Y_1R$ BPTU ABFE Validation

To date, no ABFEs have been run on ligands binding at lipid-exposed sites. It is therefore worth carefully analysing the calculations on the crystallographic ligand to assess the accuracy and reliability of the calculations, address existing concerns, and identify novel ones. These concerns centre around the ability to generate converged results, to sample the lipid, and to generate the correct apo state. I developed a workflow that builds upon the current best practices [76,333] and adapted and extended them to be applicable with membrane proteins. Where possible, I consider multiple approaches to processes to identify artefacts and assess the general applicability of the process.

3.7.1. ABFEs are accurate and precise

We began by running 5 ABFEs of the crystallographic ligand BPTU. These were run from a single embedding of the protein (Figure 3.4A). The calculations predicted $\Delta G^{\theta}_{binding} = -11.4 \pm 0.4\ kcal/mol$. This is in remarkable agreement with the experimental value of $-11.7 \pm 0.2\ kcal/mol$ (Figure 3.4B). The individual repeats do not indicate significant issues around convergence (Figure 3.4C). However, I report the error as the standard deviation of 5 repeats, which is more rigorous and greater in magnitude than error estimates from bootstrapping or the uncertainty of the estimator as it includes information about incomplete sampling. I broke down the contributions to the total free energy to see where this error might be originating from (Figure 3.4D). The $complex_{rest}$ and $ligand_{coul+vdw}$ legs have very little error. $\sigma_{complex_{coul}} = 0.1\ kcal/mol$ with no extreme points. $\sigma_{complex_{vdw}} = 0.4\ kcal/mol$ which is largely driven by a single value that differs by ~0.8 kcal/mol from the other data points. $\sigma_{total} = 0.4\ kcal/mol$ is remarkably low and comparable to what one might expect to see for a soluble protein.



B

	P2Y$_1$R / BPTU
Experimental (kcal/mol)	-11.7 ± 0.2
Predicted (kcal/mol)	-11.4 ± 0.4

Figure 3.4: Crystallographic ABFEs. (A) The P2Y$_1$R/BPTU system embedded in a pure POPC bilayer and the chemical structure of BPTU. (B) The experimental $\Delta G^{\theta}_{binding}$ [331] and the predicted value from ABFE calculations. The error for the prediction is the standard deviation for 5 repeats from a single embedding. (C) A forward/reverse convergence plot (repeat 1). The blue bar shows the MBAR estimator error using all available data (D) The components of the free energy standardised to the leg mean.

3.7.2. Multiple embeddings provide a more realistic estimate of the error

One major concern is the ability sample the contribution of the lipid. I have thus far shown that even in a pure POPC bilayer, it is challenging to sample the lipid interactions on the timescale of 1 μs. I have shown that this is exacerbated on the timescale of a typical ABFE window (~ 20 ns). However, this analysis is assessing the number of interactions formed and makes no comment on the strength of these interactions. I should note that the differences are relatively small in absolute terms and are for typically weak interactions. It therefore remains an unanswered question what impact this will have on the accuracy and precision of ABFEs.

To assess this, I ran 4 more embedding processes and ran 5 repeats from each embedded protein (Figure 3.5A). This resulted in 5 different lipid configurations and 25 repeats total.

The results for the 25 repeats is $-11.6 \pm 1.5\ kcal/mol$ (Figure 3.5B). The mean is in good agreement with the experimental value and the single embedding results. However, there is a significant increase in the σ_{total}. For each of the five starting conformations, we see strong agreements between repeats (standard deviations of: 0.4, 0.4, 0.7, 0.3, and 0.4). The increase in σ_{total} is largely driven by repeat 3, specifically the $complex_{vdw}$ leg (Figure 3.5C). I assessed the starting structure of the repeats and could not identify any obvious differences in ligand or protein conformation. A plausible explanation therefore is a contribution from the lipid configuration. It is expected the phase space sampled by the lipid will be largely dependent on the starting configuration. This would explain the internal consistency for repeats from individual embedding processes.

I therefore decided to assess the interactions the lipid forms with the protein in the non-interacting window from the complex side of the cycle. The mean number of interactions formed between individual binding site residues and POPC can be seen in Figure 3.5D. These were then coloured by how many standard deviations the mean for each window differs from the mean number of interactions observed in the long simulations (Figure 3.5E). We see that the number of interactions can differ significantly between repeats and that there can be a significant starting state dependence. For instance, let us consider the interactions F109 forms with POPC. The number of interactions formed in embedding 3 is 2.3 ± 0.2, whilst for the other 3 repeats it is 3.2 ± 0.3. This is statistically significant ($\rho < 0.01$). The number of interactions formed by F109 is abnormal when compared to long simulations (all 5 repeats are more than 3 standard deviations from the mean of the long simulations).

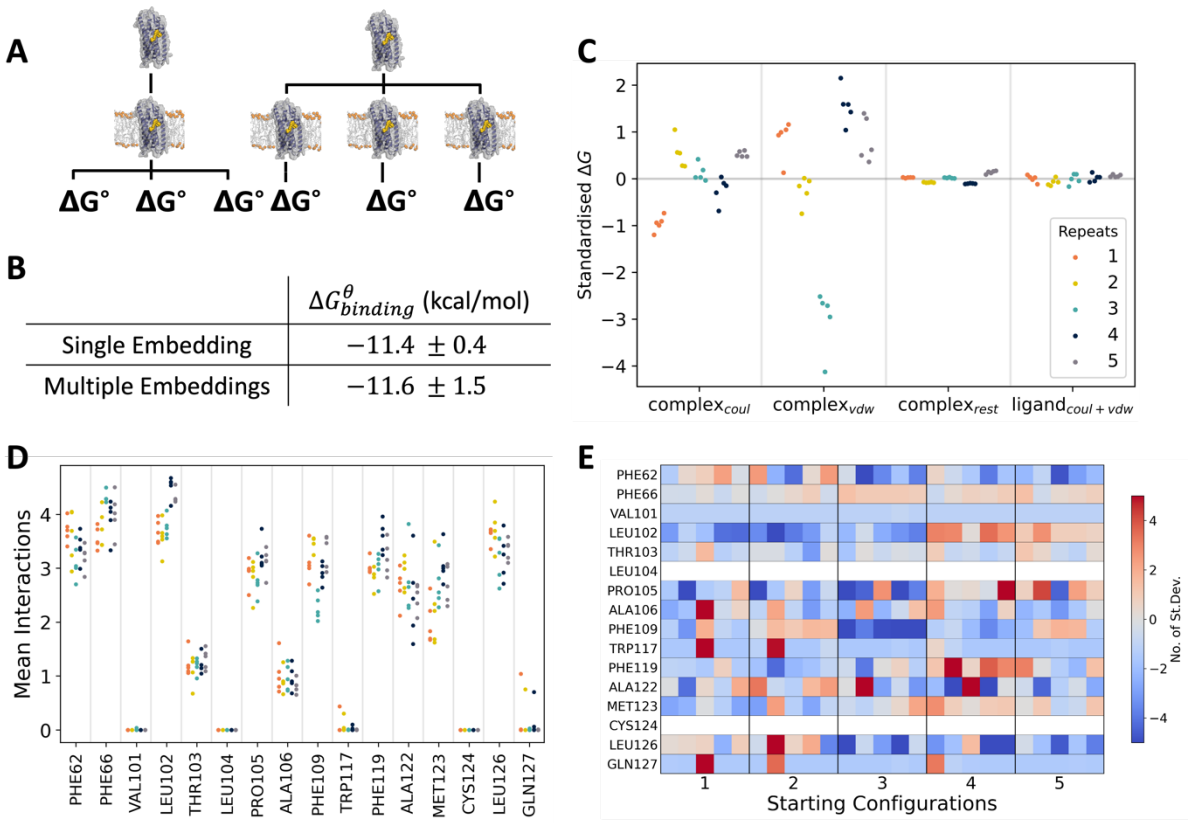


Figure 3.5: ABFEs from multiple embeddings. (A) Schematic of using a single embedding process for repeats (left) or using unique embeddings for each repeat (right) (B) The results of ABFE calculations from 5 repeats from a single embedding and 25 repeats from 5 different embeddings. (C) ΔG for individual legs standardised. Embedding 3 shows significant deviation from the other embedding in the complex$_{vdw}$ leg (D) The average number of interactions formed between binding site residues and POPC in the non-interacting window (vdw20) (E) Heatmap of binding site /POPC interactions across the vdw20 windows from the ABFE repeats. The mean and standard deviation for each residue is calculated from $5 \times 1\ \mu s$ *unbiased simulations. Each cell in the heatmap is then coloured by the number of standard deviations the simulation mean is from the long simulation mean.*

3.7.3. Results can be reproduced using multiple forcefields

The ability of ABFEs to generate accurate results should be independent of the forcefield used. The ABFE calculations described thus far were carried out using the GAFF2 representation for ligands and SLipids (2016). This is because GAFF2 and the previous generation GAFF have been used widely in previous absolute binding free energy

calculations [74,76]. It provides a reasonably accurate representation and utilising the same parameters where possible allows for better comparison between studies. SLipids has been validated as the most accurate AMBER compatible representation of lipids [354].

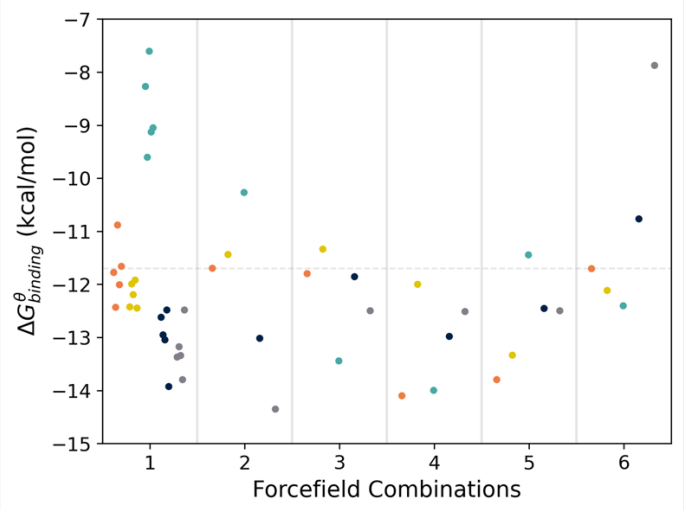
I assessed whether sage (OpenFF v2.0), sage with BespokeFit optimised torsional parameters, and the recently released SLipids (2020) provide comparable results with GAFF2 and SLipids (2016). Here I tested 6 forcefield and simulation protocol combinations (Figure 3.6). This is not meant to be exhaustive and I note this is not an ideal system for comparing forcefields as the large internal errors for forcefields makes comparisons challenging. We can however check if a particular forcefield combination gives substantial qualitatively different results.

To see if we could significantly reduce the standard deviation between embeddings, I ran 1 repeat from each of the 5 previously discussed embedding using HREX (with GAFF2 and SLipids 2020). This produced comparable results to the simulations without HREX ($-12.2 \pm 1.4\ kcal/mol$ and $-11.6 \pm 1.5\ kcal/mol$ respectively). For the rest of the simulations, separate embedding processes were run from these 2 previously discussed.

The predicted affinities should be independent of the embedding process. Thus far I have used the inflategro script [342] but other embedding methods exist [355,356]. Here I assess one of these methods, the CHARMM-GUI server [343]. The server was used 5 separate times and simulations run from each repeat. The results are in broad agreement with the inflategro results ($-12.2 \pm 0.7\ kcal/mol$).

Simulations run with sage (OpenFF v2.0) parameters for the ligand see a slight increase in the magnitude of the $\Delta G^{\theta}_{binding}$ prediction ($-13.1 \pm 0.8\ kcal/mol$). The subtle difference from the GAFF2 results is not significant and is largely driven by the GAFF2 outlier. The same is true for the sage ligand with torsions optimised with BespokeFit ($-12.7 \pm 0.8\ kcal/mol$). The results using SLipids 2020 with GAFF2 are highly similar to those using SLipids 2016 ($-11.0 \pm 1.6\ kcal/mol$). Interestingly, it also shows the pattern of 4 embeddings in very good agreement with one significantly underestimating the magnitude of $\Delta G^{\theta}_{binding}$. Again, I could not identify any obvious conformational differences in the protein or ligand.

Ultimately, I have shown that, for these forcefields, the ABFE results are largely independent of the forcefield. In general, they all produce results that are in good agreement with experimental, but the presence of particular configurations that seem to lead to significant deviation make comparison hard.



	Forcefields	Predicted ΔG° (kcal/mol)
1	AMBER99SB-ILDN/GAFF2/SLipids(2016)	-11.6 ± 1.5
2	AMBER99SB-ILDN/GAFF2/SLipids(2016) HREX	-12.2 ± 1.4
3	AMBER99SB-ILDN/GAFF2/SLipids(2016) CHARMM-GUI	-12.2 ± 0.7
4	AMBER99SB-ILDN/Sage/SLipids(2016)	-13.1 ± 0.8
5	AMBER99SB-ILDN/BespokeFit/SLipids(2016)	-12.7 ± 0.8
6	AMBER99SB-ILDN/GAFF2/SLipids(2020)	-11.0 ± 1.6

Figure 3.6: ABFEs with multiple forcefields. ABFEs were run using multiple forcefields/simulation protocols. The colours correspond to repeats from the same starting configuration. These are only internally meaningful (i.e. configuration 1 is

different in combinations 3 and 4 despite being the same colour). The only exception to this is forcefield combinations (1) and (2) in which the starting configurations are the same.

3.7.4. Lipid sampling in the apo state

3.7.4.1. Lipid Rapidly Enters the Binding Site in the Absence of Ligand

Another concern is the ability of the system to equilibrate around the non-interacting and partially interacting ligand. In the windows where the ligand is non-interacting there is a vacuum from the perspective of the protein and the lipid. It is not exactly clear how long the system needs to be equilibrated for to allow the lipids to occupy this space. The slow dynamics of lipids could exacerbate this problem relative to soluble systems. The vacuum could also lead to the deformation of the protein.

The progress of the equilibration could be tracked by a number of metrics, for instance looking at the number of POPC atoms within a threshold distance of a key residue. This leads to questions around what threshold distance is relevant and leads to discrete changes in value. A more complex description of the system using multiple centres and distances could provide a more realistic description of the volume of interest but is more complex to define and would require active user input. I instead tracked the protrude distance. The protrude distance is 1 – the volumetric overlap, where the volumetric overlap of POPC with the ligand, for example, is $\frac{V_{ligand} \cap V_{POPC}}{V_{ligand}}$ (Figure 3.7A). This is often utilised in cheminformatics [357], but to the best of our knowledge has not been used to analyse MD simulations before. It has the advantage of being a continuous variable, the volume of the ligand is explicitly defined, and it is easy to define for a novel system.

To investigate our ability to equilibrate in the non-interacting window, I ran 30 simulations of 0.5 ns with the protein restrained, mimicking a typical equilibration step (Figure 3.7B). These 30 simulations were started from 3 different starting configurations, with 10 repeats per starting configuration.

I tracked the protrude distances for the ligand paired with POPC, water, and protein. The protrude distance for the ligand with all species is initially 1 as the starting configuration is based upon the interacting ligand, meaning there is no overlap. This then falls sharply for POPC, indicating POPC is entering the space previously occupied by ligand. The protrude distance stabilises at a value of $0.60\ [0.57, 0.63] \pm 0.10\ [0.08, 0.12]$ (95% CI) typically after approximately 100 ps. The equilibration mimicking simulations can be compared to the protrude distance values for the previously discussed 5 x 1 μs simulations. Snapshots were taken every 100 ns from the apo simulations. The mean of these snapshots was $0.64\ [0.62, 0.67] \pm 0.08\ [0.07, 0.10]$. This is in good agreement with the short simulations mimicking equilibration. This indicates the equilibration is generating configurations that, at least in terms of the species occupying the ligand volume, are realistic.

The fact our equilibration mimicking simulations are run from 3 different starting configurations allows us to investigate whether there is a dependence on the starting configuration. Breaking down the protrude distance by starting configuration shows that there is a clear dependence on the starting configuration (Figure 3.7C). The means of the equilibration endpoints are $0.52\ [0.48,0.56] \pm 0.06\ [0.04, 0.09]$, $0.69\ [0.67,0.73] \pm 0.05\ [0.03,0.08]$, and $0.59\ [0.54,0.64] \pm 0.07\ [0.08,0.12]$ for configurations 1, 2, and 3 respectively. This starting dependence can also be seen in the interactions formed between the binding site residues and POPC (Figure 3.7D). The endpoints of the equilibration mimicking

processes were tested for the presence of interactions with each residue. The distributions for each starting configuration were then compared. Multiple residues were found to have a dependence on the initial configuration. This dependence is not surprising given the previously highlighted challenges with sampling the lipid interactions.

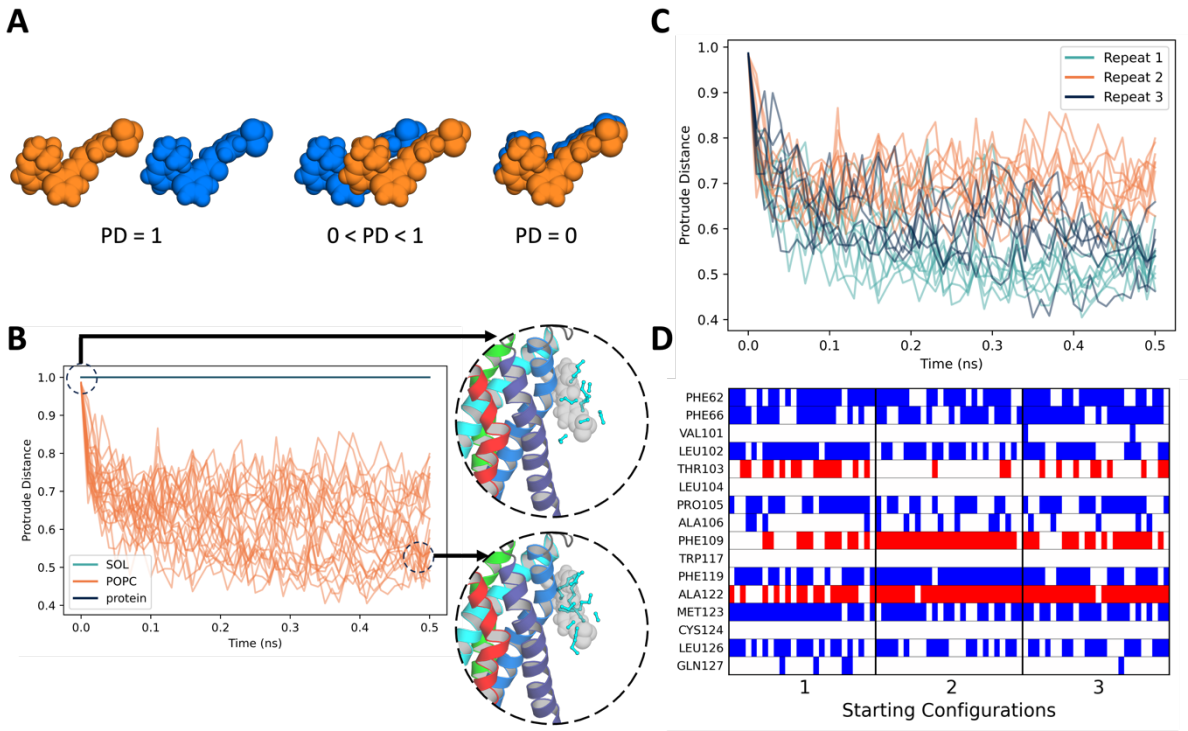


Figure 3.7: ABFE Equilibration. (A) Schematic of the protrude distance shown with BPTU as an example. (B) Protrude distance over equilibration mimicking simulations for BPTU and protein (dark blue), water (teal), and POPC (orange). (C) Protrude distance for BPTU and POPC over equilibration mimicking simulations coloured by starting configuration (D) The presence of interactions with binding site residues at the end of the equilibration mimicking simulations. The residue is coloured red if there is a significant difference between the distributions for different configurations.

3.7.5. Lipid sampling in the partially interacting state

As we progress along the λ vector, the ligand goes from completely blocking access to the binding site for other species, to completely allowing it. At intermediate vdW windows, the ligand partially blocks entry into the binding site. This typically occurs around $\lambda = 0.60$-

0.70. This makes sampling in these windows challenging. This has previously been shown to be a general challenge for ABFEs. In a study looking at HSP90 the limited sampling of water in the intermediate van der Waal's windows was strongly correlated with $dH/d\lambda$ and led to errors in the ΔG prediction [77]. I was concerned that this could be exacerbated by the slow dynamics of the lipid. I looked at the protrude distance of POPC and the ligand volume and plotted it against $dH/d\lambda$ for each of the 5 repeats of the 5 starting configurations. Figure 3.8A shows a representative window at $\lambda_{VdW} = 0.70$. $dH/d\lambda$ is strongly correlated with the volumetric overlap, with $\rho < 0.001$ for 20 of the 25 repeats. These intermediate λ_{VdW} have amongst the highest $\sigma_{overlap}$ of the windows. Interestingly, if we look at the variation in ΔG between adjacent windows as predicted by TI, we see that these windows are the largest source of error in the van der Waal's leg of the cycle (Figure 3.8B).

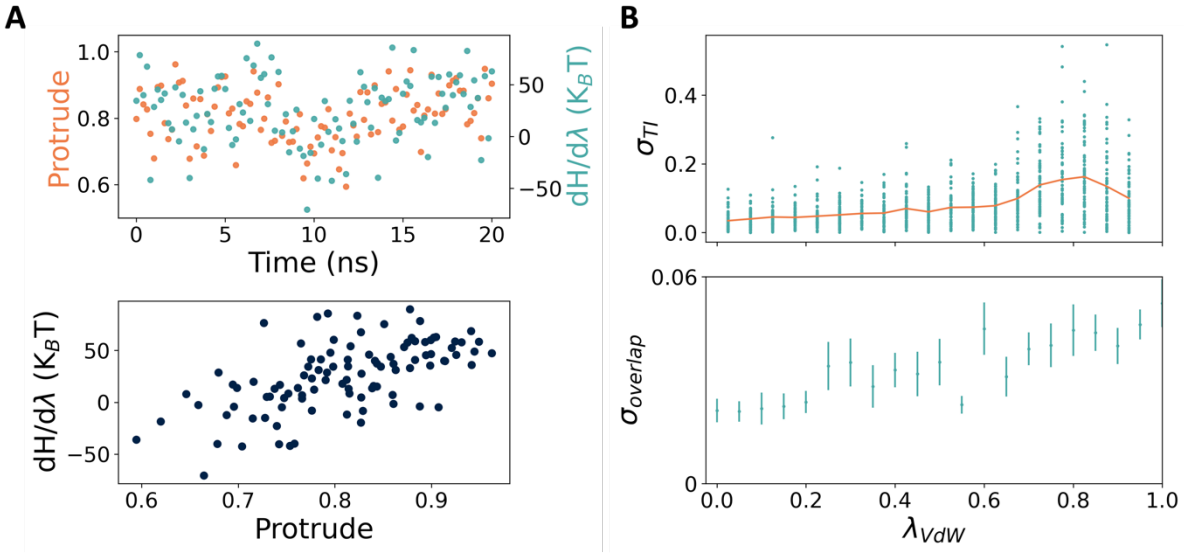


Figure 3.8: Lipid Sampling at $\lambda_{VdW} = 0.7$. (A) Protrude distance and dH/dλ plotted against time (upper) and each other (lower) (B) The standard deviation in ΔG across repeats between adjacent windows as predicted by TI for all ligand in the $P2Y_1R$ SAR series (upper) and the standard deviation for volumetric overlap of in each vdW window for the 25 BPTU repeats (lower)

3.7.6. Can we improve lipid sampling?

We have shown that lipid sampling is inconsistent across the repeats and that this can lead to significant deviation. This deviation significantly hinders our ability to interpret the results and therefore limits the utility of ABFEs. It would therefore be useful to develop a method to incorporate the variable lipid configurations into a single calculation. On the surface, non-equilibrium ABFEs would seem well suited to this problem as they would allow extended sampling of the lipid's interactions at the endpoints of the free energy cycle. However, even when using significantly longer non-equilibrium windows than currently commonly utilised (10 ns non-equilibrium windows), we were unable to obtain overlapping forward and reverse work distributions (Supplementary Figure 8.2).

I therefore developed an approach based on equilibrium ABFEs. We start windows from different configurations and utilised HREX to facilitate improved lipid sampling. HREX proposes moves and accepts the moves with a probability based on the potential energy of the two states. We used starting configurations 1, 2, and 3 described above to represent the different lipid configurations. These were selected as the free energies derived from configuration 3 ($-8.7 \pm 0.7\ kcal/mol$) deviates significantly from the other configurations ($-11.4 \pm 0.4\ kcal/mol$ and $-12.0 \pm 0.4\ kcal/mol$). This difference in free energy was largely driven by differences in $\Delta G_{VdW}^{complex}$. I therefore hoped that running the vdW windows using multiple lipid configurations might lead to improved consistency. There was good mixing between states (Figure 3.9B) but, interestingly, I saw the trajectories partition into different portions of λ space (Figure 3.9C). For instance, I saw configuration 3 occupy higher λ values (where the ligand is weakly interacting) and configurations 1 and 2 showing a preference for lower λ windows. This is likely due to the differences in $U(r)$ between systems biasing the HREX moves in a particular direction. It is worth noting that the

movement of trajectories is not unidirectional and that we do see diffusive behaviour of the trajectories across λ space.

5 ABFEs using the different starting configurations were run and resulted in $\Delta G^{\theta} = -10.5 \pm 0.8\ kcal/mol$ (Figure 3.9A & Table 3.1). The original results for the three starting configurations was $-10.9 \pm 1.6\ kcal/mol$. I note that there is no significant change in the mean, but there is a marked improvement in the standard deviation. This could therefore provide a useful method to reduce the standard deviations between repeats caused by poor lipid sampling.

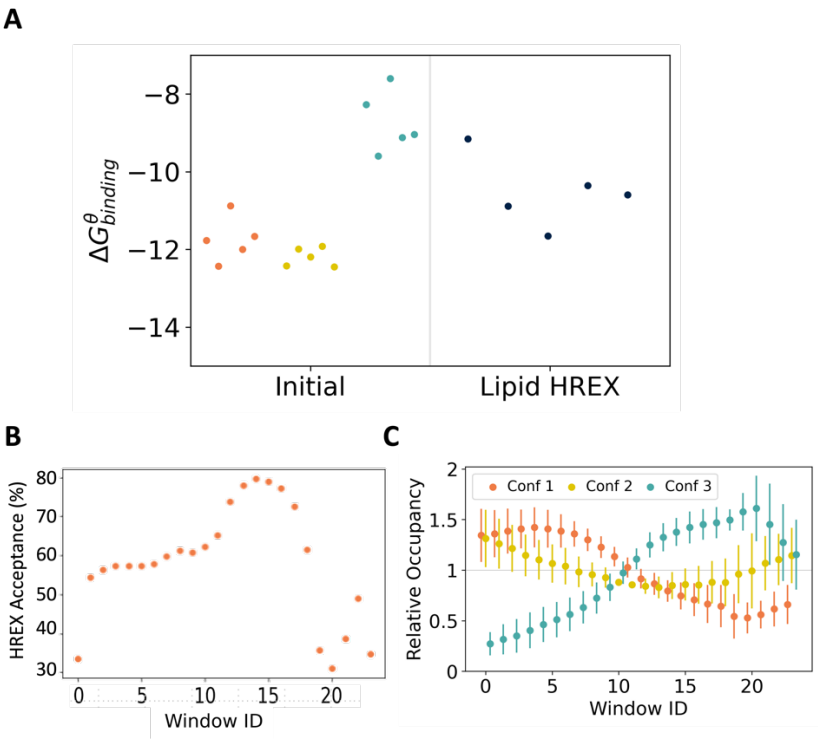


Figure 3.9 Lipid HREX. *(A) Graphical representation of the $\Delta G^{\theta}_{binding}$ for 3 conformers and the results from using a mixture of those starting configurations in the same simulation with HREX (B) HREX acceptance rate is defined as when a simulation ends up at a new λ window after a set of proposed moves. This shows that exchanges occur between states. (C) The standardised proportion each configuration spends in each window.*

	$P2Y_1R$ / BPTU (kcal/mol)
Experimental	-11.7 ± 0.2
$\text{Initial}_{\text{Conf 1}}$	-11.4 ± 0.4
$\text{Initial}_{\text{Conf 2}}$	-12.0 ± 0.4
$\text{Initial}_{\text{Conf 2}}$	-8.7 ± 0.7
Lipid HREX	-10.5 ± 0.8

Table 3.1: Lipid HREX Results. The $\Delta G^{\theta}_{binding}$ for 3 conformers and the results from using a mixture of those starting configurations in the same simulation with HREX.

3.8. Free Energy Calculations of $P2Y_1R$ SAR

We next decided to test our methodology on the SAR series described in the discovery of BPTU (Figure 3.10). This series of molecules provides a useful challenge as they are numerous, cover a decent range of affinities ($\sim 4\ kcal/mol$), and cover elaborations of a number of different vectors. The reported values all come from a single paper as inconsistencies in methodology make comparisons between data more challenging.

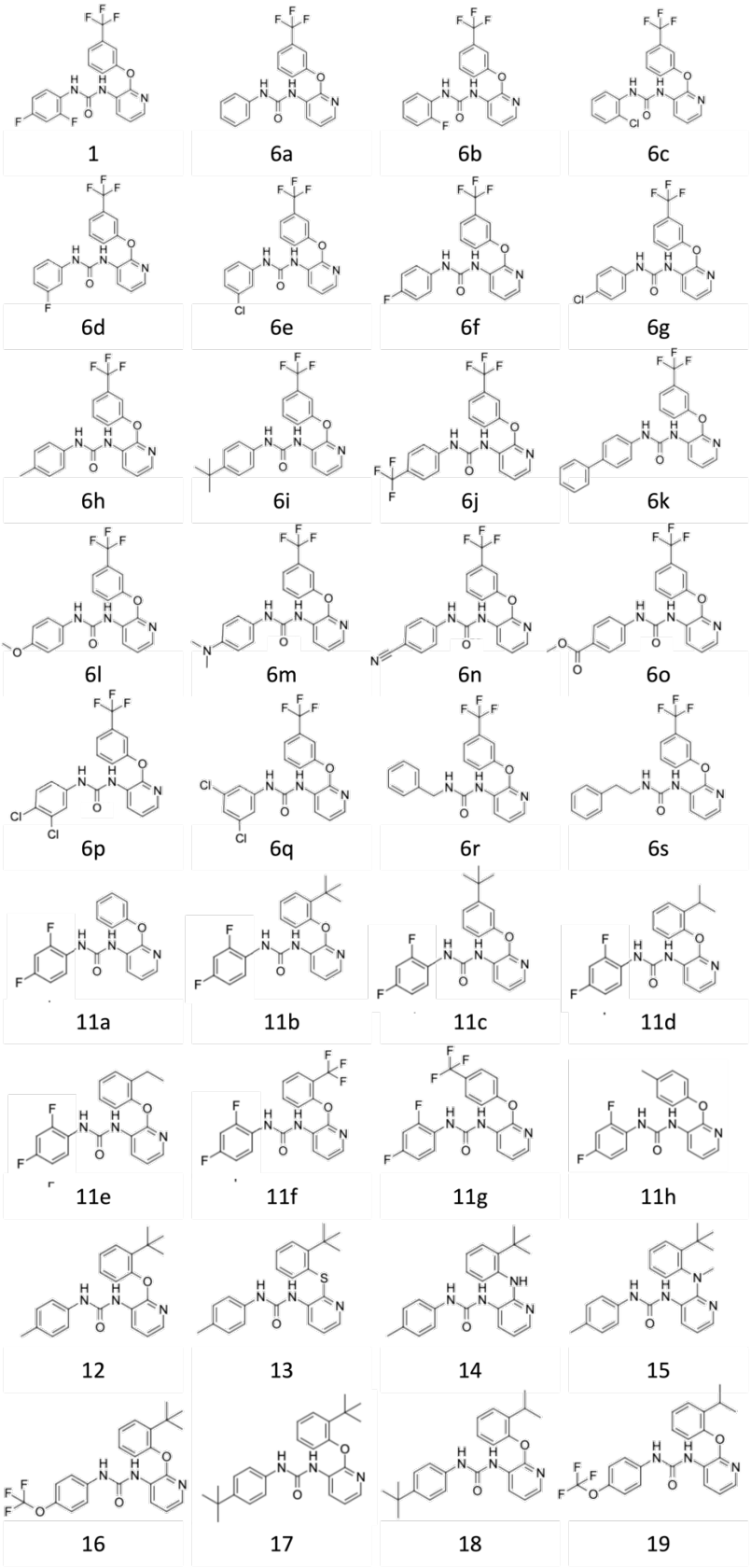


Figure 3.10: P2 SAR Structures

3.8.1. Pseudo-symmetric groups require consideration

In a prospective study, one would likely utilise docking to generate a starting pose. Here, that is not useful as we are interested in understanding the systematic errors of free energy calculations and errors from starting pose could hinder the identification of other issues. Docking algorithms are not trained for these systems so significant experimental data, methodological development, and validation would be needed before this were to be approached.

One of the practical advantages of this system is that the similarity of scaffold allows the placement of molecules into the binding site. One of the issues of this approach is how to deal with pseudo-symmetric groups. To highlight this problem, I will discuss the molecule *6b* from the SAR series (Figure 3.11, Figure 3.12A). 6b has a trifluoromethyl group in the meta position on the terminal benzene. This terminal benzene can adopt two conformations - the trifluoromethyl group could point towards the protein ('in') or it could point away from the protein ('out') (Figure 3.12B). The preferred orientation is not obvious *a priori.*

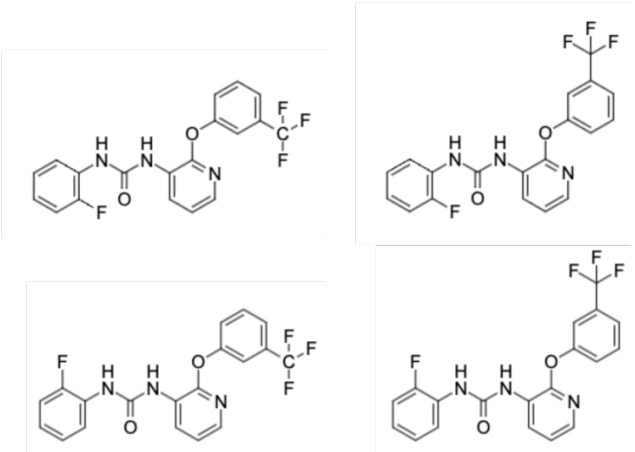


Figure 3.11: 6b conformations. Conformations of 6b with the two terminal susbtituents 'in' and 'out'

I ran ABFE calculations for both poses and saw a significant difference between the 'in' and 'out' conformations. The 'in' conformation had a predicted $\Delta G^{\theta}_{binding} = -7.0 \pm 0.7\ kcal/mol$ whilst the 'out' conformation had a value of $\Delta G^{\theta}_{binding} = -10.5 \pm 0.6\ kcal/mol$. The 'out' value is in very good agreement with the experimental value of $\Delta G^{\theta}_{binding} = -10.3\ kcal/mol$. This highlights the importance of generating the correct starting conformation of the ligand.

The ligand binding site of 6b is relatively shallow so were surprised to not see interconversion between rotamers (Figure 3.12C). I hypothesised that the presence of lipid could slow the sampling of internal degrees of freedom of the ligand. To assess this, I ran $100\ ns$ simulations of the ligand free in solution (Figure 3.12E) and in a pure POPC bilayer (Figure 3.12D). We saw that the dihedral readily interconverts in water whilst showing limited mixing of states in the bilayer. This is relevant for ABFE calculations as if the sampling of rotamers is slowed by lipid, then careful consideration of rotamers may be required at a large number of lipid-exposed sites.

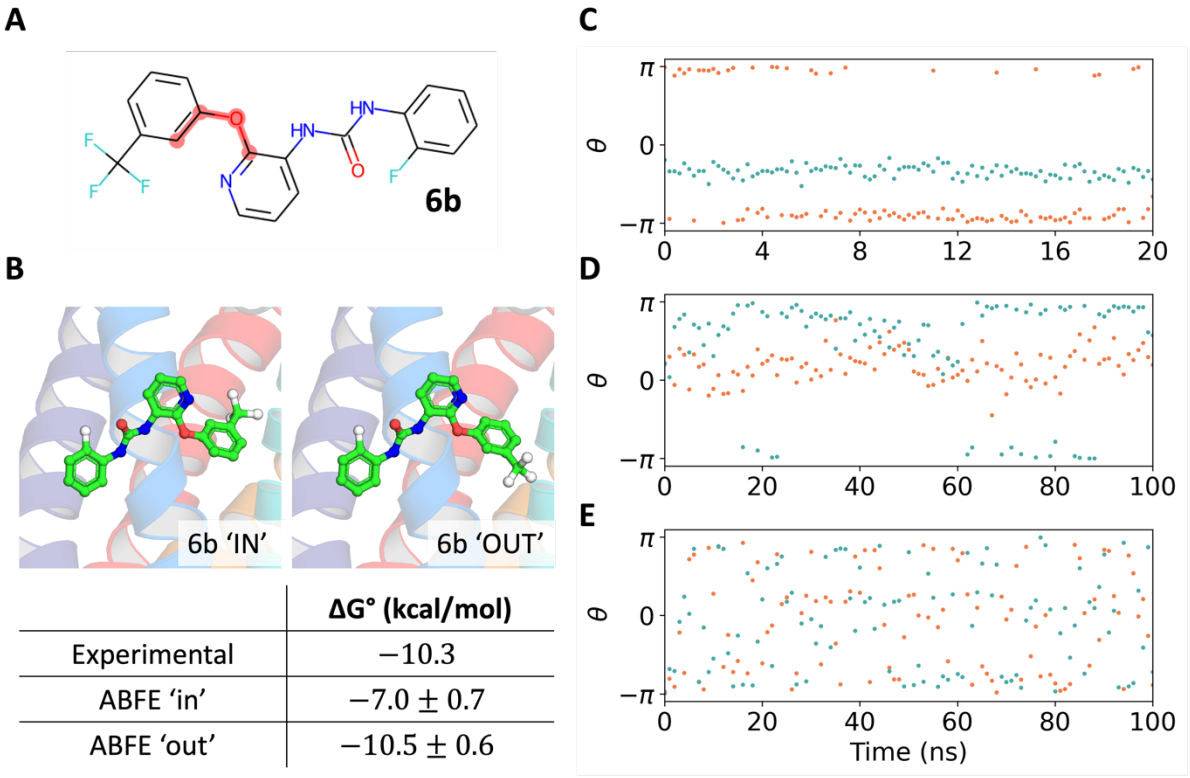


Figure 3.12: 6b terminal dihedral. (A) The chemical structure of 6b with the dihedral linking the trifluoromethyl-benzene and pyridine highlighted (B) The 'in' and 'out' conformations of 6b with the respective $\Delta G^{\theta}_{binding}$ (C) θ in the 'in' and 'out' configurations (teal and orange) do not interconvert on the timescale of an ABFE window (D) There is limited interconversion when run in the bilayer (E) the dihedral freely interconverts in water

3.8.2. ABFEs Accurately Rank P2Y$_1$R SAR

We ran ABFEs on the SAR series. To address the issue with pseudo-symmetric groups described above, I ran the calculations for each of the 'in' and 'out' conformations of the terminal dihedrals. This could result in up to 4 conformations per ligand, drastically increasing the required simulation time.

To address this, I developed a simple workflow to reduce the computational cost of the simulations. I started with running 20 ns unbiased simulations. These simulations have a low computational overhead and are already part of our standard workflow as they are used to

generate suitable Boresch restraints. The dihedrals were tracked over the course of the simulation. I defined a dihedral as needing treatment as two separate states if we did not see a transition from 'in' to 'out' and 'out' to 'in' in the respective trajectories. A single ABFE repeat was then run for the remaining states of the ligand. I then removed conformations that had a high $\Delta G^{\theta}_{binding}$. This was defined as ligands that were more than 3 $kcal/mol$ more positive than the best scoring conformation.

The 3 $kcal/mol$ threshold was selected as when doing Boltzmann weighted averages, as the difference between two values increases, the contribution of the lesser value decreases exponentially. It is typically ignored once the values are 2 $kcal/mol$. The value for two equal free energies is $-\beta^{-1} \ln 2$, which is $\approx 0.41\ kcal/mol$ at 300K [358]. I then added an additional 1 $kcal/mol$ to the threshold to consider the impacts of our uncertainty of the true value of the conformation for a particular conformation from a single repeat. This threshold is a balance between reducing computational overheads and trying to prevent the spurious removal of relevant conformations.

The selected conformations were then run for an additional 2 repeats, each with separate embeddings and equilibration processes. The lowest scoring ligand was taken as representative. Ideally, one would like to treat the conformations as separate states, but instead I used only a single conformer. The requirement for the separation of states is that the ligands sample different parts of phase space. However, as our ability to sample the rotamers is dependent on the ligand's environment, we may see transitions between rotamers in the weakly interacting states. Also, due to the stochastic nature of simulations, we may see a rotation occur even in the fully interacting state. Separation of states is not therefore technically applicable in this scenario [358].

The results can be seen in Figure 3.13. The error bars are the standard deviation of the three repeats. The results are strong and comparable to those seen in soluble systems. I consistently identify strong and weak binders correctly. I have a good r_P of 0.68 [0.39, 0.84], MAE of 1.39 [1.02, 1.86], and RMSE of 1.81 [1.39, 2.35].

Whilst the majority of ligands are predicted accurately, there a few for whom the magnitude of their affinity is underpredicted. The 6 most problematic ligands were: 6a, 6f, 6h, 6l, 6m, and 13. These 6 molecules do not share any obvious features that would explain why they are poorly predicted. Molecule 13 contains a thioether instead of the ether. Sulfurs are known to be problematic in parameterising small molecules [292,359]. Interestingly, the 5 molecules in the '6' section of the SAR series all contain smaller substituents at the para position of the terminal benzene than BPTU's trifluoromethyl group. One could apply a symmetry correction which would improve the results by $\sim 0.41\ kcal/mol$, but it is not clear how appropriate this is. If one does apply a symmetry correction to the applicable molecules, the results would be: $r_P = 0.71\ [0.41,0.85],\ MAE = 1.26[0.95,1.68],\ RMSE = 1.60\ [1.24,2.08]$.

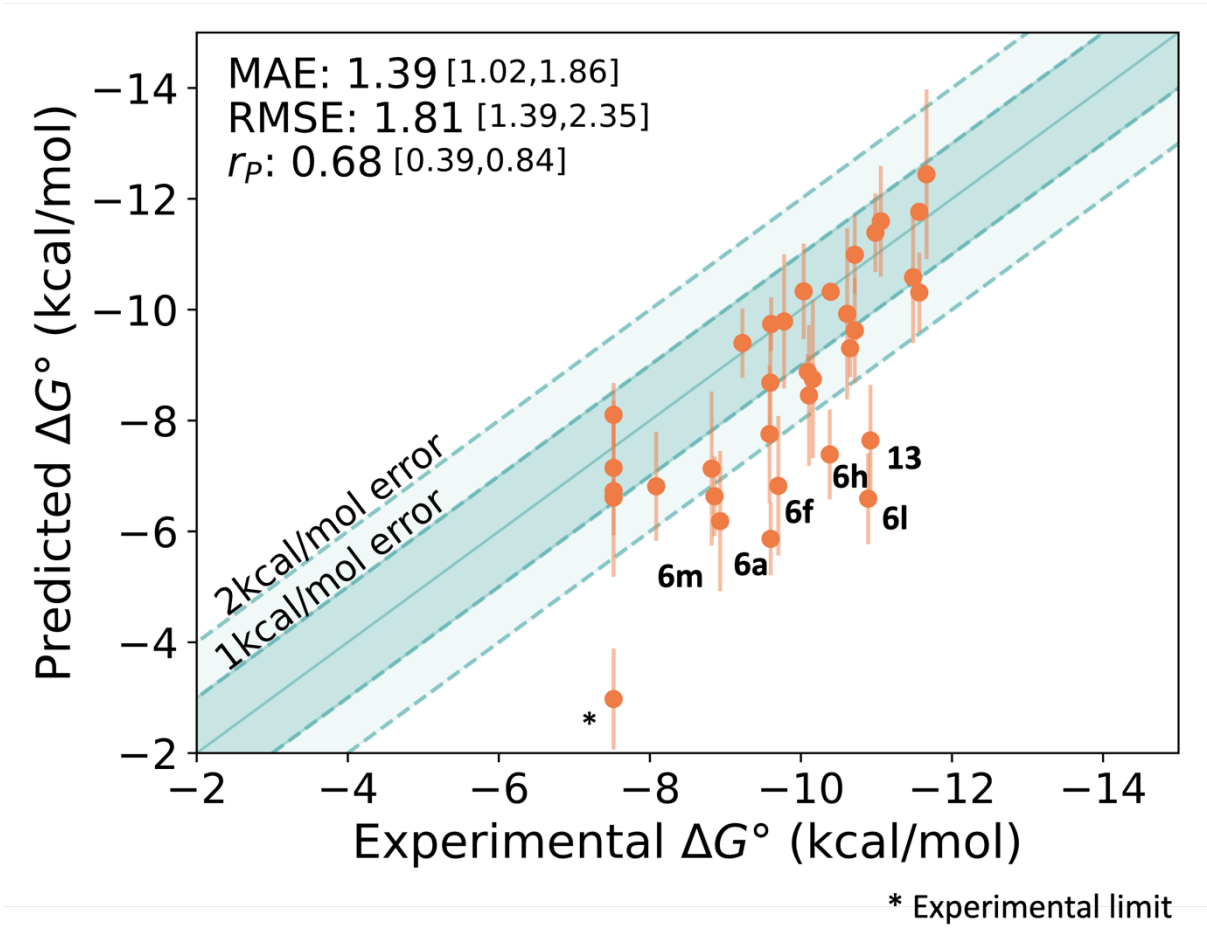


Figure 3.13: P2Y$_1$R SAR ABFE Results. The results for ABFE calculations of the BPTU SAR binding to P2Y$_1$R. The points with experimental values equal to -7.6 kcal/mol are beyond the sensitivity of the experiment and are simply reported as non-binding. These points are not included in the reported statistics. The 6 ligands for which the predicted values most disagrees with the experimental value are labelled. The line of experimental = predicted is shown with shaded areas either side representing 1 and 2 kcal/mol error.

3.9. ABFEs for BPTU Selectivity

One of the interesting properties of BPTU is that it is selective for $P2Y_1R$ over other $P2Y_X$ receptors. RBFES can be used to work out the free energy difference of mutating one residue to another [360,361]. RBFEs are limited to small changes in the binding site and could not handle changes in binding pose or significant conformational changes. ABFEs provide a useful alternative method for calculating selectivity without making assumptions on similarities on binding site and binding mode. The dynamic nature of GPCRs means that assumptions around similar protein structure and binding mode are less likely to hold. It is therefore interesting to assess whether ABFEs can predict selectivity.

The BPTU/$P2Y_1R$ system is a good system to investigate this. There is a clear hypothesis for the cause of specificity – the sidechain of $A106^{2.59}$ in $P2Y_1R$ is small enough to accommodate the BPTU phenyl ring whilst the larger sidechains found in other $P2Y_X$ receptors prevent binding [253]. This is supported by mutagenesis. shows selected mutagenesis data for two molecules from the BPTU SAR series [253]. These mutations should not introduce significant structural changes and provide an interesting starting to point to validate the utility of ABFEs for selectivity prediction.

Another interesting *in silico* mutation to investigate is $P105^{2.58}A$. The presence of $P105^{2.58}$ prevents $L102^{2.55}$ forming intra-helical hydrogen bonds, allowing the main chain carbonyl to form the key hydrogen bonds with BPTU. One would therefore expect the $P105^{2.58}A$ to reduce binding by allowing $L102^{2.55}$ to form intrahelical bonds. It is worth noting that there is no experimental data to support this hypothesis. This is a significantly more challenging mutation as the mutation would require structural rearrangements as the intrahelical bonds form. This would make RBFEs unsuitable for this problem. Finally, there is also binding data

for selected molecules in the BPTU SAR series binding to other $P2Y_X$ receptors [252]. This provides a direct test of our ability to predict selectivity.

shows binding data for selected SAR molecules binding to selected $P2Y_X$ receptors. Ligands 12 and 16 bind to $P2Y_1R$ with highly similar affinities. They both also show no ability to bind $P2Y_{12}R$. Visualisation of an AlphaFold [24] model of the inactive $P2Y_{12}R$ shows that $F79^{2.59}$ is involved in a steric clash with the ligand. This is in agreement with the previously mentioned hypothesis that $A106^{2.59}$ in $P2Y_1R$ is important for selectivity. Other rotamers of $F79^{2.59}$ do not appear to resolve this in a satisfactory manner. In this case, structural observations are enough to explain the lack of observed binding. More interestingly, for the $P2Y_2R$, ligand 16 (BPTU) exhibits no binding ($K_i > 15\ \mu M$) whilst ligand 12 does ($K_i = 2.7\ \mu M$). The two ligands are highly similar, differing only in that 16 has a trifluoromethoxy group where 12 has a methyl group. There are multiple changes between the $P2Y_2R$ and $P2Y_1R$. $L88^{2.59}$ is in the equivalent position to $A106^{2.59}$. Interestingly, this change is not proximal to where the two ligands differ and there are no obvious structural clues as to why one would be tolerated and the other not.

We began by running the four mutations in the BPTU binding site described in Zhang *et al.*[330]. These are $T103^{2.56}W$, $A106^{2.59}L$, $A106^{2.59}F$, and $A106^{2.59}W$ (Figure 3.14A). These mutations all introduce much bulkier mutations that are believed to introduce steric clashes. All of them, with the exception of $A106^{2.59}L$, reduce binding beyond the sensitivity of the radioligand binding assay (Figure 3.14). $A106^{2.59}L$ has a $\Delta G^{\theta}_{binding} = -7.6 \pm 0.1\ kcal/mol$. I ran 3 ABFEs for each mutation. All mutations led to significant reductions in $\Delta G^{\theta}_{binding}$. $A106^{2.59}L$ reduced binding to $-9.1 \pm 0.5\ kcal/mol$. This is the smallest reduction and is in good agreement with the experimental maintenance of binding. $A106^{2.59}F$ and $A106^{2.59}W$

reduce binding to $-4.7 \pm 0.3\ kcal/mol$ and $-5.7 \pm 0.7\ kcal/mol$, respectively. Again, this matches the experimental observation of a loss of binding.

I then decided to directly assess how well I can predict selectivity. To do this I ran ABFEs for BPTU and 12 binding to $P2Y_2R$ (Figure 3.14C). This is a significant challenge due to the large number of mutations and the lack of a crystal structure. I used the AlphaFold model of $P2Y_2R$ [24]. I manually checked for issues with the binding poses and residue sidechains. I then ran a 100 ns equilibration with the ligand restrained to allow the sidechains and find a more holo-like conformation. ABFE calculations were then run from these structures. For both BPTU and 12, we saw significant decreases from their predicted $P2Y_1R$ values ($-11.6 \pm 1.5\ kcal/mol$ to $-9.2 \pm 1.3\ kcal/mol$ and $-10.3 \pm 0.7\ kcal/mol$ to $-6.9 \pm 1.5\ kcal/mol$, respectively). Whilst this is in qualitative agreement with the experimental decreases, experimental data indicates BPTU ceases to bind whilst 12 has an affinity of $-7.9\ kcal/mol$. However, we observe a greater decrease in 12 than BPTU. This could be a result of the use of a predicted protein structure and the experimental results are also from a single repeat so may be less reliable.

	Experimental ΔG° (kcal/mol)	**Predicted ΔG° (kcal/mol)**
$P2Y_1R$ WT	-11.7 ± 0.2	-11.6 ± 1.5
$P2Y_1R$ $T103^{2.56}W$	> -7.1	-8.6 ± 1.0
$P2Y_1R$ $A106^{2.59}L$	-7.6 ± 0.1	-9.1 ± 0.5
$P2Y_1R$ $A106^{2.59}F$	> -7.1	-4.7 ± 0.3
$P2Y_1R$ $A106^{2.59}W$	> -7.1	-5.7 ± 0.7
$P2Y_1R$ $P105^{2.58}A$	N/A	-10.5 ± 0.4
$P2Y_1R$ $P105^{2.58}A_{modeller}$	N/A	-8.0 ± 2.1
$P2Y2R_{BPTU}$	> -7.1	-9.2 ± 1.3
$P2Y2R_{12}$	-7.9	-6.9 ± 1.5

Table 3.2:Mutation Results. Experimental and predicted binding affinities for a range of mutations, receptors, and ligands.

Finally, I looked at the $P105^{2.58}A$ mutation (Figure 3.14B). There is no experimental data for how this mutation impacts BPTU binding but I hypothesised it would reduce binding affinity by preventing $L102^{2.55}$ hydrogen bonding with BPTU as it is forming intrahelical hydrogen bonds with $A105^{2.58}$. The structural rearrangement you may see on this mutation to facilitate the intrahelical bonding would make RBFEs inappropriate.

I modelled the mutation in two ways. The first was simply to change the residue in question. This did not result in intra-helical hydrogen bonding and BPTU continued to hydrogen bond with $L102^{2.55}$. As a result, the ABFE predictions were similar to WT ($-10.5 \pm 0.4\ kcal/mol$). I then decided to model the mutation with Modeller whilst allowing a large portion of the helix to rearrange. This resulted in polar interactions forming between $L102^{2.55}$ and $A105^{2.58}$. I then ran ABFEs from these simulations. The predicted affinity was $-8.0 \pm$

2.1 $kcal/mol$. This is a large decrease in the mean but is accompanied by a large standard deviation.

To investigate what is causing this standard deviation, I decided to more closely investigate the stability of the interactions we introduced between L102$^{2.55}$ and A105$^{2.58}$. To do this, I ran unbiased simulations of the protein in the presence and absence of BPTU. Interestingly, we saw that the interaction was maintained in the absence of ligand but the presence of BPTU caused it to break (Figure 3.14D).

I then retrospectively analysed the windows from our ABFE calculations and saw that the average distance between the carbonyl oxygen and the amide nitrogen is large for the fully interacting ligand, but decreases as the coulombic interactions are perturbed (Figure 3.14E). This corresponds with an increase in the distance between the nitrogen of the urea as the coulombic interactions are 'turned off' - that is the coulombic interactions are scaled to 0 by a coupling parameter λ such that at $\lambda = 1$ ligand does not form interaction interactions with the environment (Figure 3.14F).

This creates a picture of a subtle system where the preferred conformation of the protein is different in the apo and holo states. The ligand still binds and forms interactions with the carbonyl of L102$^{2.55}$ but we would expect to see a decrease in affinity as it must pay an energetic cost for the breaking of the intrahelical hydrogen bonds. Poor sampling of the transition between these two states could explain the large standard deviation between repeats.

I looked back at the calculations and indeed saw inconsistent sampling of the states. Figure 3.14G shows the intrahelical distance and angle for the fully interacting (teal) and fully non-interacting (yellow) windows for selected repeats. Repeat 1 showed the distribution I expected – the apo state maintains the modelled distance whilst the holo state sees the distance increase. This returns an intermediate affinity of $-7.1\ kcal/mol$. Repeats 2 and 4 however find only the apo like state and only the holo like state respectively. They therefore return extreme affinities of $-4.0\ kcal/mol$ and $-10.1\ kcal/mol$.

It is worth noting of course, that the modelled system may not be sufficient to accurately model this system. It could be that the real system could have more optimal geometry for the intrahelical hydrogen bonding, meaning the carbonyl cannot rotate in the presence of BPTU. The mutation could also introduce other larger scale changes to the structure that I have not accurately modelled here. However, what this does highlight is how subtle rearrangements can make a large difference to the predicted affinity and how poor sampling of this can lead to large deviations between repeats.

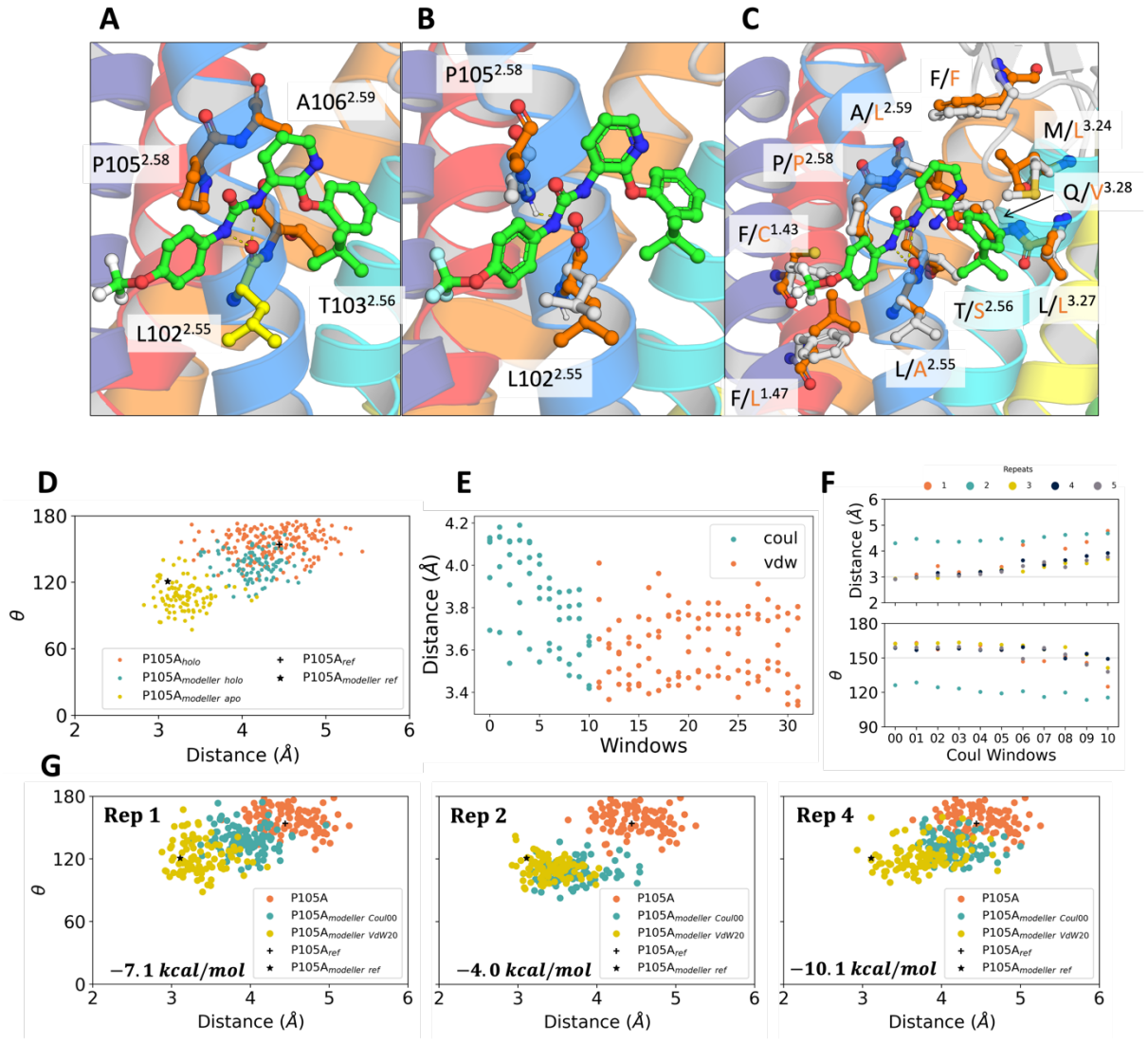


Figure 3.14: P2Y$_1$R Mutations. (A) BPTU binding to P2Y$_1$R (PDB: 4XNV) Sidechains shown in orange are those that were investigated with in silico mutations. In yellow is L102 which forms the key hydrogen bonds with BPTU. (B) P105A mutation with no structural rearrangement (orange) and with structural refinement by Modeller [362] to create intrahelical hydrogen bonding (grey). Intra-helical bonding between L102 and A105 is shown in yellow. (C) The P2Y$_2$R binding site as predicted by AlphaFold[24] (orange sticks) overlaid with P2Y$_1$R (grey sticks) (D) The distance (Å) between the O of the backbone carbonyl of L102 and the backbone N of A105 and the O--H--N bond angle. These values are plotted for a holo simulation in the unrefined structure (orange), the refined structure both holo (teal) and apo (yellow). The reference values for the unrefined structure and Modeller structure are shown as a black plus and star respectively. (E) The mean O--N distance for each window in an ABFE across 5 repeats. Windows where the coulombic interactions are perturbed are shown in teal, and vdW in orange. (F) An N in the urea group of BPTU to backbone carbonyl O distance and corresponding N--H--O angle for coulombic windows across 5 repeats. (G) The distance and angle described in (D) in the fully interacting window and fully non-interacting windows of select repeats and the corresponding predicted affinity.

3.10. Discussion

This chapter has investigated ABFE calculations for a lipid-exposed site at $P2Y_1R$. This is a novel binding site for ABFE calculations so I have assessed multiple factors including our ability to sample the lipid's contribution, the dynamics of the ligand, and the equilibration methodology. I have shown that despite some of the challenges I identified, I have been able to generate accurate results, comparable with what has been seen previously for soluble proteins.

The lipid constitutes a major component of the ligand's environment. This poses a number of challenges for molecular simulations, so I began by assessing our ability to sample the lipid's contribution. Our main concern was our ability to sample the contribution of the lipid due its slow dynamics. I assessed this with a series of μs simulations, both of the apo and holo states. I showed that the number of contacts formed varies significantly over the course of a simulation and can differ even over the course of long simulations. Our analysis here was based on the number of contacts formed, so to assess what impact this would have in practice, I ran multiple ABFE calculations of the crystallographic ligand.

I began by running 5 repeats from a single embedding process. The results ($-11.4 \pm 0.4\ kcal/mol$) showed very good agreement with the experimental values ($-11.7 \pm 0.2\ kcal/mol$). The standard deviation of these repeats was very low but, as I have shown there is a dependence of lipid interactions based on starting position, I decided to run 5 repeats each for 4 more embedding process. This resulted in a total of 25 repeats. The results ($-11.6 \pm 1.5\ kcal/mol$) maintain the strong agreement with the experimental prediction.

However, I see a marked increase in the standard deviation. This standard deviation is much greater than the internal standard deviation for each of the starting configurations. A large amount of the variation is driven by repeat 3, with a predicted affinity of $-8.7 \pm 0.7\ kcal/mol$. I could see no structural clues in the protein and ligand as to the cause. I then assessed the protein/lipid interactions in the non-interacting window. I saw a strong dependence on the starting configuration for the interactions that were formed. This provides a reasonable hypothesis as to the cause of this deviation.

I therefore ran our ABFE repeats from a unique embedding process to generate novel lipid configurations as I believe this gives a more realistic estimate in the uncertainty of our predictions. I also note that one could try and run ever longer windows to better sample the lipid's contribution. I have shown that poor sampling of the lipid can persist well beyond the scope of typical ABFE windows. For instance, I took the 5 non-interacting windows for repeat 3 and extended them to $100\ ns$ each. These results were highly similar to the original results ($-8.9 \pm 0.5\ kcal/mol$). I proposed a method to use multiple starting points for ABFE windows. I showed that this reduced the standard deviation between repeats whilst maintaining a similar mean. I also showed that the ABFE calculations are robust for several forcefield combinations, embedding processes, and simulation parameters.

As part of my methodology, I ran equilibration simulations prior to each window, totalling $5.6\ ns$ of simulation per window. I wanted to assess whether this was appropriate given the slow dynamics of the lipids. I saw a dependence of the protein lipid interactions on the starting configuration. This is to be expected based on previous simulations and cannot easily be corrected. I therefore wanted to see if these starting configurations were at least plausible. I was concerned about how long the lipid would take to occupy the binding site and whether

the vacuum would cause deformation of the protein. To assess this, I developed a pipeline to measure the protrude distance over the course of equilibration mimicking simulations. I saw that lipid quickly occupies the binding site at a volume which is in good agreement with the values observed in long simulations. The exact volume filled shows a starting dependence, again highlighting the importance of multiple starting lipid configurations. Reassuringly, I saw no protein entering the ligand's volume, indicating we do not see a hydrophobic collapse. This suggests that our equilibration process is as good as we can expect on realistic timescales.

We also saw that the dynamics of ligands are slowed in the bilayer relative to water. I assessed this by tracking the terminal dihedral of SAR molecule 6b. This freely interconverted in water but did not in the pure POPC bilayer. This dihedral was unable to freely interconvert between rotamers in the binding site when fully interacting. It is not obvious *a priori* which rotamer is preferred. This creates a problem as ABFE calculations showed that the 'out' conformation is in good agreement with the experimental value, whilst the 'in' conformation is not.

Assessing each conformer for ABFEs is therefore necessary but leads to a massive growth in the computational requirements. To overcome this, I developed a workflow to assess whether dihedrals can freely interconvert and then ran single repeats of ABFE calculations. I then applied a threshold of $-3.0\ kcal/mol$ for conformations to be taken forward to additional repeats. This threshold is selected to balance keeping important conformations (the contribution decays exponentially as $\Delta\Delta G$ between conformers increases) and not using excessive computational resources. This threshold could of course be adapted to fit the needs of the system and the design project. In my case, I set a relatively high threshold of 3

kcal/mol due to the observed standard deviation between repeats and our desire to ensure we do not miss key conformations.

Typically, one would use a separation of states methodology and treat the conformations as separate binding poses. However, this is not applicable as I see the dihedral interconvert in the latter windows of the cycle. For smaller molecules, the Boresch restraints can be used to ensure only one conformation is sampled. This allows symmetry corrections and separation of states to be applied. However, for a large molecule this is not trivial and although it could feasibly be done, would significantly impact the dynamics of the ligand and reduce the stability of the simulations. Instead, one could use a post-hoc corrections of frames that are in the wrong conformer or directly apply restraints to the dihedral of interest [358,363]. We assessed the post-hoc correction and found it increased the magnitude of the affinity by approximately 1 $kcal/mol$. Directly restraining the dihedrals of interest to a single conformer was attempted but it proved challenging to find combinations for the Boresch and dihedral restraints that did not exert an influence on one another.

Using alchemical free energy calculations to rank poses is computationally expensive. One can envision a situation whereby MM-GBSA calculations are used to predict the preferred pose as an additional filtering step. The 20 ns simulations used to derive the Boresch restraints could be used to make the computational overhead fairly negligible. MM-PB/GBSA suffers from issues around predicting entropic contributions and the use of implicit solvent ignores structurally important waters. It would be interesting to see how these issues manifest for lipid-exposed sites. Initial results (using the Python package gmx_MMPBSA [364] to run the calculations) suggest some ability to distinguish the 6b conformations (Supplementary

Figure 8.4). Further work would be needed to validate these results, including determining an appropriate scheme for modelling the dielectric constant across the membrane [365].

The ABFE calculations for the SAR series show promising results. The r_P of 0.68 [0.39, 0.84] is relatively strong and our absolute values are in reasonably good agreement with the experimental data (MAE of 1.39 [1.02, 1.86] and RMSE of 1.81 [1.39, 2.35]). This is a better correlation than observed in using RBFEs which obtained $r_P = 0.58$ [258]. This is likely due to the fact I use quite a thorough ABFE scheme and that RBFEs are dependent on a number of factors including the design of the network. These results also perform well when compared to soluble proteins. For instance, if we look at the 8 series of ABFEs on soluble systems carried out by Chen *et al.* [87], in terms of r_P I outperformed all but 1 of them. I identified a few outlier ligands. For ligand 13, I believe the discrepancy could be as a result of parameterisation issues caused by the presence of a sulfur. For the other ligands, I noted that they all had symmetric terminal groups which could have been improved upon by the use of symmetry corrections.

Predicting selectivity is a key potential use of ABFEs but to date few studies have successfully demonstrated this is possible [88]. This is particularly important for GPCRs due to the similarity of family members and how targeting multiple family members leads to serious side effects. Here I aimed to assess whether predicting selectivity is likely to be possible for GPCRs. I began by assessing our ability to recreate mutagenesis data. I looked at 4 mutations - T103$^{2.56}$W, A106$^{2.59}$L, A106$^{2.59}$F, and A106$^{2.59}$W. These all introduce bulkier sidechains and either make binding beyond the sensitivity of the experiment (T103$^{2.56}$W, A106$^{2.59}$F, and A106$^{2.59}$W) or significantly reduce it (A106$^{2.59}$L $\Delta G^{\theta}_{binding} = -7.6 \pm 0.1\ kcal/mol$). I reproduce the data well, with all 4 mutations leading to significant reductions and A106$^{2.59}$L

having the smallest magnitude reduction (A106$^{2.59}$L $-7.6 \pm 0.1\ kcal/mol$; A106$^{2.59}$F $-4.7 \pm 0.3\ kcal/mol$; A106$^{2.59}$W $-5.7 \pm 0.7\ kcal/mol$; T103W $-8.6\ \pm 1.0\ kcal/mol$).

I then attempt to directly predict selectivity by predicting the binding affinities for 12 and 16 binding to P2Y$_{12}$R. I was able to recreate the 2 molecules' preferences for P2Y$_1$R over P2Y$_{12}$R. However, I was unable to recreate 12's relative preference for P2Y$_{12}$R compared to 16. This could be a function of our use of a predicted structure for P2Y$_{12}$R. I finally looked at the P105$^{2.58}$A mutation. This mutation is interesting as it likely requires a rearrangement relative to the WT structure to facilitate intra-helical hydrogen bonding. I see using different modelling techniques yields significantly different results. I also see large standard deviations in our free energy predictions. I linked both of these factors to the variable presence and absence of hydrogen bonds between the backbone carbonyl of L102$^{2.55}$ and the ligand or A105$^{2.58}$. It appears in the absence of the ligand the intra-helical bonding is favoured, whilst the presence of the ligand leads to the breaking of the intra-helical bonding. This is interesting as it highlights how even subtle structural changes can lead to significant errors in prediction.

3.11. Conclusion

In this chapter I assessed how well ABFEs work for predicting the binding of BPTU and its related SAR series to a lipid-exposed site on $P2Y_1R$. I began by showing that fully sampling the contribution of lipids may be challenging on the timescale of ABFE windows. I showed several ways to assess the quality of equilibration and whether I generate realistic protein lipid interactions in ABFE windows. I showed that ABFEs run from a single embedding process appear to underestimate the error and I therefore suggest people should run unique embedding processes for each repeat. I showed that ABFEs can predict the absolute value of BPTU binding accurately and robustly across multiple forcefield combinations. I also showed that ABFEs can be used to accurately rank the BPTU SAR at a comparable level to recent results for soluble systems ($r_P = 0.68$). Finally, I showed that ABFEs can be used to recreate mutagenesis data for $P2Y_1R$ and can predict BPTU's preference for $P2Y_1R$ over the closely related $P2Y_{12}R$. These results show that ABFEs can work well in a novel and challenging system.

4. Glucagon Receptor

4.1. Table of Contents

4. GLUCAGON RECEPTOR 164

4.1. TABLE OF CONTENTS 165

4.2. FIGURES 166

4.3. INTRODUCTION 167

4.4. METHODS 170

4.5. INITIAL SIMULATIONS 174

4.5.1. Apo Conformations 179

4.5.2. Binding Site Occupancy 183

4.6. RBFE CALCULATIONS 186

4.7. ABFE CALCULATIONS 190

4.8. SAR ABFES 199

4.8.1. ABFEs from 'true' apo conformation 204

4.9. DISCUSSION 209

4.10. CONCLUSION 214

9. SUPPLEMENTARY MATERIAL: CHAPTER 4 292

4.2. Figures

FIGURE 4.1: MK-0893 INTERACTIONS 177

FIGURE 4.2: POPC/PROTEIN INTERACTIONS 179

FIGURE 4.3: APO CONFORMATIONAL CHANGES 182

FIGURE 4.4: APO PROTRUDE DISTANCE 185

FIGURE 4.5: RBFE CYCLE AND SAR 189

FIGURE 4.6: RBFE RESULTS 189

FIGURE 4.7: GCGR/MK-0893 FORCEFIELDS 192

FIGURE 4.8: VDW 20 PROTEIN/POPC INTERACTIONS 195

FIGURE 4.9: VDW20 PROTRUDE DISTANCES 198

FIGURE 4.10 GCGR SAR 200

FIGURE 4.11: GCGR SAR ABFE RESULTS 203

FIGURE 4.12: APO/HOLO ABFE CALCULATIONS 208

SUPPLEMENTARY FIGURE 9.1 GCGR SYMMETRY MATE. 292

SUPPLEMENTARY FIGURE 9.2: UNBIASED SIMULATION RMSDS 292

SUPPLEMENTARY FIGURE 9.3: STABILITY OF WATER BRIDGES 293

SUPPLEMENTARY FIGURE 9.4: BESPOKEFIT TORSIONS 294

SUPPLEMENTARY FIGURE 9.5: GCGR SAR ABFE RESULTS & STRUCTURES GRID. 295

SUPPLEMENTARY FIGURE 9.6: DUMMY ATOM VALIDATION 296

SUPPLEMENTARY FIGURE 9.7: MK-0893 APO/HOLO ABFE OVERLAP MATRICES 297

SUPPLEMENTARY FIGURE 9.8: GCGR WATER BRIDGES 297

4.3. Introduction

In Chapter 3, I have shown that ABFE calculations can work effectively for BPTU (and related analogues) binding at a lipid-exposed site on $P2Y_1R$. I therefore decided to extend our calculations to an even more ambitious system - MK-0893 binding to the glucagon receptor (GCGR) [270,271].

MK-0893 consists of 3 prongs: ethyl-phenyl-carbonyl-β-alanine (prong 1), methoxynaphthalene (prong 2), and dicholorophenyl (prong 3) (Figure 1.8). A SAR series focusing predominantly on prongs 2 and 3 was reported in the discovery of MK-0893 [270]. MK-0893 sits outside of helices V, VI, and VII (Figure 1.8). The binding site consists of a polar cleft between helices VI and VII (occupied by prong 1) and a hydrophobic cleft between helices V and VI (occupied by prong 2). Prong 3 protrudes into the bilayer.

MK-0893 as a ligand provides several challenges. First, it consists of 67 atoms, making it larger than many ligands used in free energy calculations. This means we need to transform a large alchemical region. The ligand is also flexible with 9 rotatable bonds. The ligand also carries a formal charge of -1 at physiological pH (-7.2). Due to our use of PME, we must account for the change in the system's total charge as the ligand is perturbed to prevent a finite-size effect leading to a dependence on system size. In soluble systems this is typically corrected in a *post hoc* manner based on work by Rocklin *et al.* which is commonly referred to as the Rocklin correction [312]. The Rocklin correction has been shown to be effective for soluble systems [366,367] but recent work has shown it is not applicable for membrane systems and instead suggests the use of a co-alchemical ion [313]. This involves the (de)charging of an ion at the same time as the ligand to maintain a neutral box. I use the co-alchemical ion method throughout the simulations.

The nature of the binding site also provides many additional challenges. The binding site is located at the protein/bilayer/water interface. This places the ligand amongst the lipid headgroups increasing the chance for stronger and more specific interactions between the protein/lipid and ligand/lipid. This could exacerbate the sampling challenges I observed in $P2Y_1R$.

The presence of structured water also complicates the system. The crystal structure provides evidence of a crystallographic water forming interactions with the backbone carbonyl of $L399^{7.56b}$ and the carboxylate of MK-0893. The presence and stability of this water should be closely monitored. WaterDock [368] predicts the existence of water at this site (data not shown) whilst docking of MK-0893 into the site predicts qualitatively different conformations in the presence and absence of the water (data not shown). It is not clear what would occupy the binding site in the absence of ligand. It could be that the binding site is occupied by lipid, water, or a dynamic combination of the two.

The ligand is also bound at a highly dynamic part of the protein structure. The outward movement of helix VI is associated with activation in class B GPCRs and is more extensive than the rearrangements seen in class A GPCRs [369,370]. This activation is facilitated by numerous subtle rearrangements of interaction networks and motifs [371]. MK-0893 is an antagonist so whilst we do not have to consider the energetics of activation, it may induce subtle changes in the bonding networks. This is particularly relevant for the intracellular hydrogen bonding network consisting of $R173^{2.54b}$, $R346^{6.37b}$, $N404^{8.47b}$, and $E406^{8.49b}$ as these residues constitute a key part of the MK-0893 binding site. Any rearrangements that MK-0893 has induced would need to be fully sampled, which may not be trivial.

I also note that the crystal structure has the R173$^{2.54b}$A mutation. I simulated the WT protein but it is important to ensure we are sampling a realistic conformation of the system, both in terms of the sidechain interactions and the water network. The other concern with the structure is the presence of interactions between prong 3 of MK-0893 with a symmetry mate. I therefore first aim to understand the system before assessing how well ABFEs work for this system. Despite these limitations, I chose this structure out of the 13 available (4L6R, 5EE7, 5XEZ, 5XF1, 5YQZ, 6LMK, 6LML, 6WHC, 6WPW, 7V35, 8FU6, 8JRU, 8JRV) as it contains the ligand of choice, ensuring the best chance of the protein being in an appropriate conformation.

4.4. Methods

The crystal structure of MK-0893 bound to GCGR (PDB ID: 5ee7) was used as the starting point for these simulations [271]. Missing atoms, residues, and loops were modelled using MODELLER [334]. Of the 50 generated models, the top 10 models were rescored using the QMEAN scoring function [335,336]. The best QMEAN scoring model was selected. The protein was capped with N-methyl and acetyl caps using PyMol.

Simulations were performed using Gromacs 2021 [338]. The forcefield used was AMBER99SB-ILDN [288]. The ligand was parameterised in the following ways: unless stated otherwise OpenFF v2.0.0 (Sage) [341] and GAFF2 [339] using antechamber. Gromacs topologies were generated using ParmEd [340]. The protonation state of the ligand was the major state in solution with Marvin Sketch used for the pKa calculations. The protein was embedded into the bilayer using InflateGro [342]. The lipids were represented using SLipids 2016 [344]. The systems were solvated in a box of approximately (8,8,10 nm). The box was solvated and sodium and chloride ions were added to neutralise the system and to create a concentration of 0.15 M. The water model used was TIP3P [346]. Electrostatic simulations were handled using particle mesh Ewald method (PME) [347,348]. A cut-off of 1.2 nm was used. A cut-off of 1.2 nm was used for vdW interactions. Post embedding, the systems were equilibrated for 100 ns with the protein restrained to facilitate equilibration of the lipids.

For ABFE calculations, the vdW and electrostatic interactions were either decoupled (if not stated) or annihilated. Annihilation is the removal of inter and intra-molecular interactions, whilst decoupling is just the removal of inter-molecular interactions. Annihilation was done using the mdp flag $couple - intramol = yes$ whilst decoupling was done by converting the atoms to custom atomtypes that behaved identically except that they did not interact with the

environment. This was done to avoid known issues involved in the handling of nonbonded exclusions in the Gromacs 2021 code [372,373]. The vdW and electrostatic interactions were decoupled using a linear pathway of $\Delta\lambda = 0.05$ and 0.1 respectively. The restraints were added using a non-linear pathway (0.0, 0.01, 0.025, 0.05, 0.075, 0.1, 0.15, 0.2, 0.3, 0.5, 0.75, 1.0). The coulombic and vdW were separated into distinct vectors due to differences in the free energy kernel's different abilities to offload work to the GPU. This resulted in 43 simulations total for the complex and 32 for the ligand in solution. On certain occasions, to optimise utilisation of the available clusters, additional λ windows were added (λ_{vdW}=0.62,0.67,0.72 and $\lambda_{coul} = 0.65$). This did not impact the results. A soft-core potential for decoupled vdW interactions was used [349].

For each window, the system was equilibrated before the production run. The production runs, unless specified otherwise, were 20 ns for the complex and 15 ns for the ligand in solution. The equilibration schemes began with a 10,000 step energy minimisation using a steepest descent algorithm. For the ligand in solution, this was followed by 100 ps in the canonical ensemble with harmonic position restraints with a force constant of 1,000 kJ mol^{-1} nm^{-2} on the solute heavy atoms. The temperature was coupled to the reference temperature of 298.15 K using Langevin dynamics. It was then further equilibrated in the isothermal-isobaric (NPT) ensemble for 100 ps using the Berendsen weak coupling algorithm [306]. This was then followed by a further 500 ns using the Parinnello-Rahman barostat [350] before restraints were removed and the system was equilibrated for a final 5 ns. The initial equilibration scheme for the complex simulations follow the same process. I also use an adapted equilibration scheme which is the same except up until the final 3 ns of equilibration restraints are applied to the water molecules.

Analysis was carried out using the Python implementation of pymbar v3.0.3 [319] via alchemlyb. During the complex simulations, the ligand was restrained in the binding site using so-called Boresch restraints [68]. These involve restraining an intermolecular distance, 2 angles, and 3 dihedrals with force constants of 41.8 kJ mol^{-1} nm^{-2}. The contribution of these restraints to the free energy is calculated analytically for the ligand in solution. The optimal atoms to use for these restraints was picked from a 20 ns simulation using MDRestraintsGenerator [76]. The ligands carry a formal charge of -1 at pH 7.2. To maintain charge neutrality as the ligand is decoupled, a co-alchemical ion was annihilated [313]. Hamiltonian replica exchange (HREX) was utilised as implemented in Gromacs. 1,000,000 exchanges were attempted every 1,000 steps.

The GCGR 'true' apo state was taken from the endpoints of 3 independent unbiased simulations. The ligand was added to these endpoints using alchembed [355]. The alchembed process was run prior to the λ window being equilibrated. The interactions of the ligand were turned on only to the extent of the λ window to prevent distortion. This was done by manipulating the 'delta-lambda' value in the *mdp files. To prevent lipid getting stuck between the prongs, dummy atoms were added to the volume between the two. These dummy atoms were parameterised so that they would interact only with lipid. They therefore prevent lipid being stuck between the prongs without distorting the rest of the system. These atoms were turned on concurrently with the ligand and at the same rate to prevent over disturbance of the lipid.

For RBFE calculations, mapping was performed using an in-house script. Merged topology and gro structure files were created using alchemicalitp (https://alchemicalitp.readthedocs.io). These were carried out using a sequential transformation. This is a three-step process where

partial charges of the atoms only present in ligand A are annihilated, non-charge related transformations are run, and the atoms that are only present in ligand B are charged. Linear λ spacing schemes are used with $\Delta\lambda = 0.1$ for charged related stages and $\Delta\lambda = 0.05$ for non-charge related stages.

For each window, equilibration was run for production. This consisted of a 10,000 step energy minimisation using a steepest descent algrotithm. This was followed by 50 ps in the canonical ensemble with harmonic position restraints with a force constant of 1,000 kJ mol^{-1} nm^{-2} on the solute heavy atoms. The temperature was coupled to the reference temperature of 298.15 K using Langevin dynamics. It was then further equilibrated in the isothermal-isobaric (NPT) ensemble for 50 ps using the Berendsen weak coupling algorithm [306]. This was then followed by a further 500 ns using the Parinnello-Rahman barostat [350] before restraints were removed and the system was equilibrated for a final 1 ns. Production runs were 5 ns for the ligand in solution and 10 for the ligand in complex.

Trajectory analysis was carried out using MDAnalysis [351] and ProLiF [352]. Ligand structures were plotted using ChemDraw and RDKit 2023.03.1. Visualisation was carried out using PyMol [353].

4.5. Initial Simulations

The model of the WT protein was generated from PDB ID 5ee7 with crystallographic mutations reverted to WT, the T4-lysozyme insert removed (to aid crystallisation), and missing atoms modelled. I then ran $5 \times 2\ \mu s$ unbiased simulations in the absence of MK-0893 and $5 \times 2\ \mu s$ unbiased simulations in the presence of MK-0893. The ligand remains stable across the holo simulations and the backbone RMSD indicates the overall fold remains stable (Supplementary Figure 9.2). I note that prong 3 of the ligand appears to be stable, reducing our concerns about the significance of the contact with the symmetry mate. In a similar fashion to $P2Y_1R$, I decided to quantify the protein/lipid/ligand/water interactions to assess how well I can identify interactions, whether I can realistically sample the interactions, and whether there is a difference between the apo and holo states. Again, I used ProLiF to assess these interactions.

We began by looking at the interactions the ligand forms with the protein and the lipid across the $5 \times 2\ \mu s$ unbiased simulations. MK-0893 forms a mixture of hydrophobic (5.9 ± 0.15), hydrogen bond as the donor (0.8 ± 0.11), hydrogen bond as the acceptor (2.3 ± 0.2), anionic (1.0 ± 0.16), π-cation (0.2 ± 0.08), and vdW (6.8 ± 0.32) interactions with the protein (Figure 4.1A).

The contacts formed are broadly in agreement with the crystal structure. The residues forming the hydrophobic cleft between helices V and VI ($L329^{5.61b}$, $F345^{6.36b}$, $L352^{6.43b}$, $T353^{6.44b}$) form hydrophobic and vdW interactions. The interactions with $L329^{5.61b}$ are only present in 0.13 ± 0.03 of the frames. This is not due to a significant conformational change in the ligand, but likely results from the crystal distance (3.5 Å) being near the limit of the definition of the ProLiF interactions. In the crystal, prong 1 forms hydrogen bonds with

K349$^{6.40b}$, S350$^{6.41b}$ and N404$^{7.61b}$, a salt bridge with R346$^{6.37b}$, and water-mediated hydrogen bonds with S350$^{6.41b}$ and L399$^{7.56b}$. The ionic interaction between R346$^{6.37b}$ and the ligand are maintained across the simulations. The hydrogen bond with S350$^{6.41b}$ is maintained. The hydrogen bond with K349$^{6.40b}$ is not maintained throughout the simulations although the K349 does form π-cation interactions with the central ring of MK-0893.

I do not see any interactions formed by the ligand with R173$^{2.54b}$. This is reassuring as the residue is mutated to alanine in the crystal. The interactions that are formed between the ligand and the bilayer are primarily hydrophobic (4.9 ± 0.34) and vdW interactions (3.7 ± 0.17). Rarely, I observed π-cation interactions and ionic interactions. These anionic interactions are primarily seen in repeat 5 when the carboxylate on prong 1of the ligand briefly moves out slightly from between helices VI and VII, allowing it to form these interactions, before moving back in between the helices. The interactions cannot be well sampled across short simulations – after 40 ns $\sigma_{hydrophobic} = 0.84$, after 1 μs $\sigma_{hydrophobic} = 0.46$.

We also tracked the water mediated hydrogen bonds between the ligand and S350$^{6.41b}$ and L399$^{7.56b}$ that are present in the crystal structure, as well as those that are formed with R173$^{2.54b}$ (Supplementary Figure 9.3). Examples of these water bridges are shown in Supplementary Figure 9.8). Here we use the term first order interactions to refer to a single water bridging two residues, with a second order network containing two waters. The first order water mediated interactions between the ligand and S350$^{6.41b}$ and L399$^{7.56b}$ are maintained throughout the simulations, with the exception of a period between 400 ns and 900 ns in repeat 5. This corresponds to the period described above in which we see the ligand move out from between helices VI and VII and form a salt bridge with a lipid.

Defining the potential water network between MK-0893 and R173$^{2.54b}$ is not present in the crystal structure as it is mutated to alanine in the crystal structure. I therefore looked at considering first order networks only and considering both first and second order networks. When considering second order networks, the only time we see a consistent absence of water network hydrogen bonds between the ligand R173$^{2.54b}$ corresponds to the outward movement of the carboxylate of MK-0893 in repeat 5. Considering only first order networks gives a more complicated picture. Repeat 3 sees consistent first order networks across the entire simulation whilst repeat 4 sees no first order networks. Repeats 1, 2, and 5 all see networks formed, broken, and reformed.

Another interesting question is if the water network is the same in the presence and absence of ligand. To assess this, I tracked the presence of a first order network between S350$^{6.41b}$ and L399$^{7.56b}$. We saw that a first order water network is present in 0.91 ± 0.09 of frames. If we look at the $5 \times 2\mu s$ apo simulations, we see the network is only present in 0.37 ± 0.32 of frames. This indicates the network is stabilised by the presence of the ligand. The high standard deviation suggests this could be a challenging degree of freedom to adequately sample. This could provide an additional challenge in our ABFE calculations as we must sample the λ-dependent rearrangement of this network.

I decided to investigate what behaviour we might expect from intermediate windows so ran $3 \times 250\ ns$ with the ligand's coulombic interactions fully decoupled and the vdW interactions fully intact. The network is present in 0.74 ± 0.17 of frames. This is not significantly different from the WT ($p = 0.16$) but this could be a result of the small number of repeats or the difference in simulation lengths. I was also interested to see if the network

was changed by the R173$^{2.54b}$A mutation. I ran 5 × 200 ns holo simulations and observed the water bridge between S350$^{6.41b}$ and L399$^{7.56b}$ was maintained in 0.86 ± 0.05 of frames – highly similar to that observed in the WT simulations.

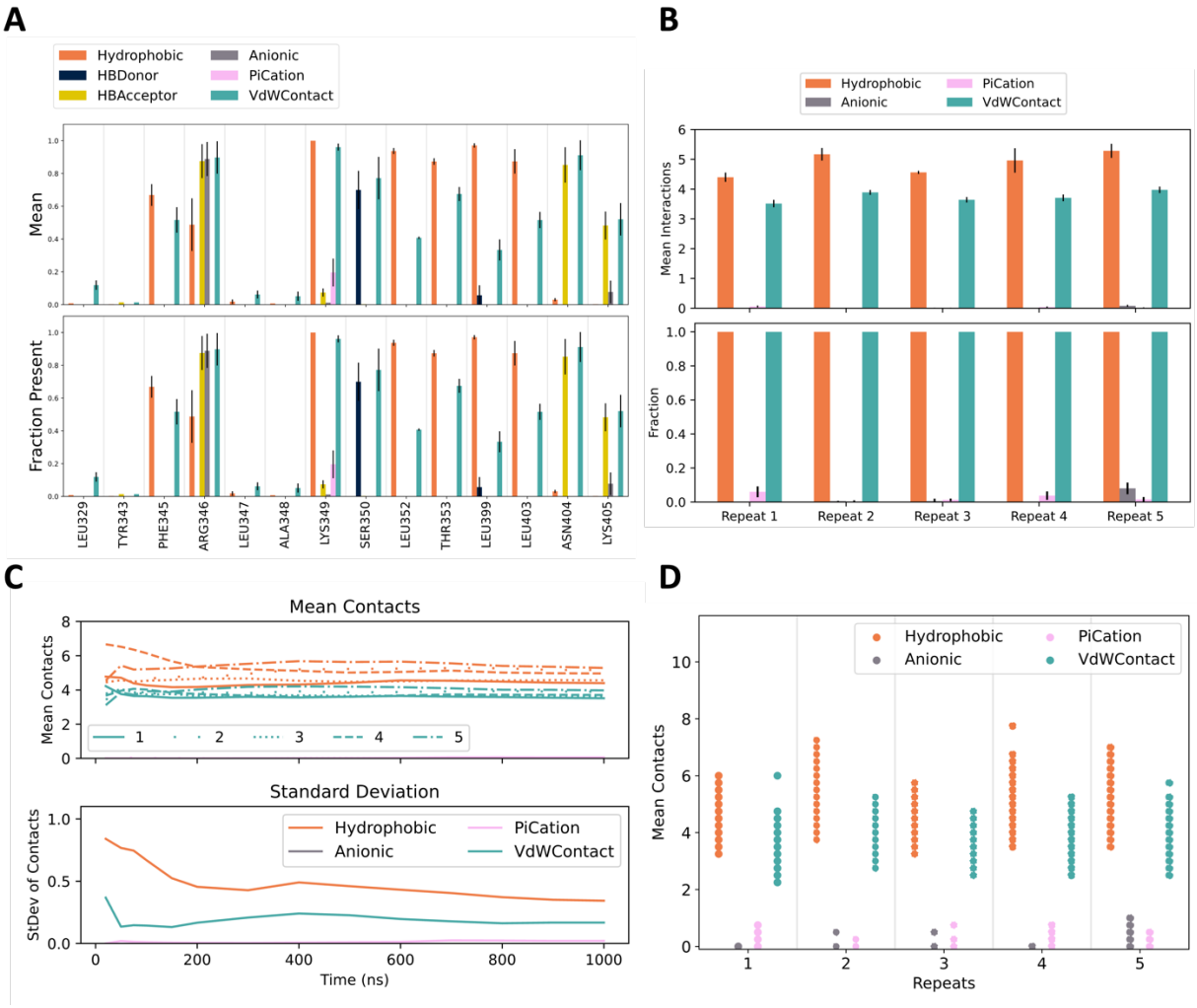


Figure 4.1: MK-0893 Interactions. (A) The interactions MK-0893 forms with the binding site residues across 5 × 2μs simulations. The error bars are the standard deviation of the 5 repeats (B) The interactions between MK-0893 and POPC. The error bars are from block analysis. (C) The number of MK-0893 POPC interactions by type as a function of simulation time (upper) and the standard deviation between repeats (lower) (D) The mean MK-0893/POPC interactions across 20 ns chunks of the simulation.

Next, I looked at the interactions POPC forms with the protein in the absence of the ligand across the 5 × 2 μs unbiased simulations (Figure 4.2A). The major interactions formed are hydrophobic (15.1 ± 1.45), vdW contacts (10.1 ± 1.73), hydrogen bond acceptor (2.5 ±

0.46), and anionic (2.5 ± 0.4). These high standard deviations, even after 2 μs, indicate significant sampling problems. Looking at K349$^{6.40b}$ as an example residue, we see $\sigma_{hydrophobic} = 0.39$ after 20 ns whilst after 500 ns $\sigma_{hydrophobic} = 0.22$ (Figure 4.2B).

It is worth noting that certain interactions are harder to sample adequately. For instance, K405 forms on average 0.11 vdW contacts with POPC in repeat 1 whilst in repeat 3 it makes 1.1 contacts. 20 ns chunks show significant deviations (Figure 4.2C).Interestingly, the standard deviation between repeats increases between 20 ns ($\sigma_{hydrophobic} = 0.33$) whilst after 2 μs $\sigma_{hydrophobic} = 0.42$. A similar increase is seen for R346$^{6.37b}$ (500 ns $\sigma_{hydrophobic} = 0.15$ whilst 2 μs $\sigma_{hydrophobic} = 0.26$). The cause of these discrepancies is not slow lipid dynamics, but slow rearrangements of the sidechains upon removal of the ligand.

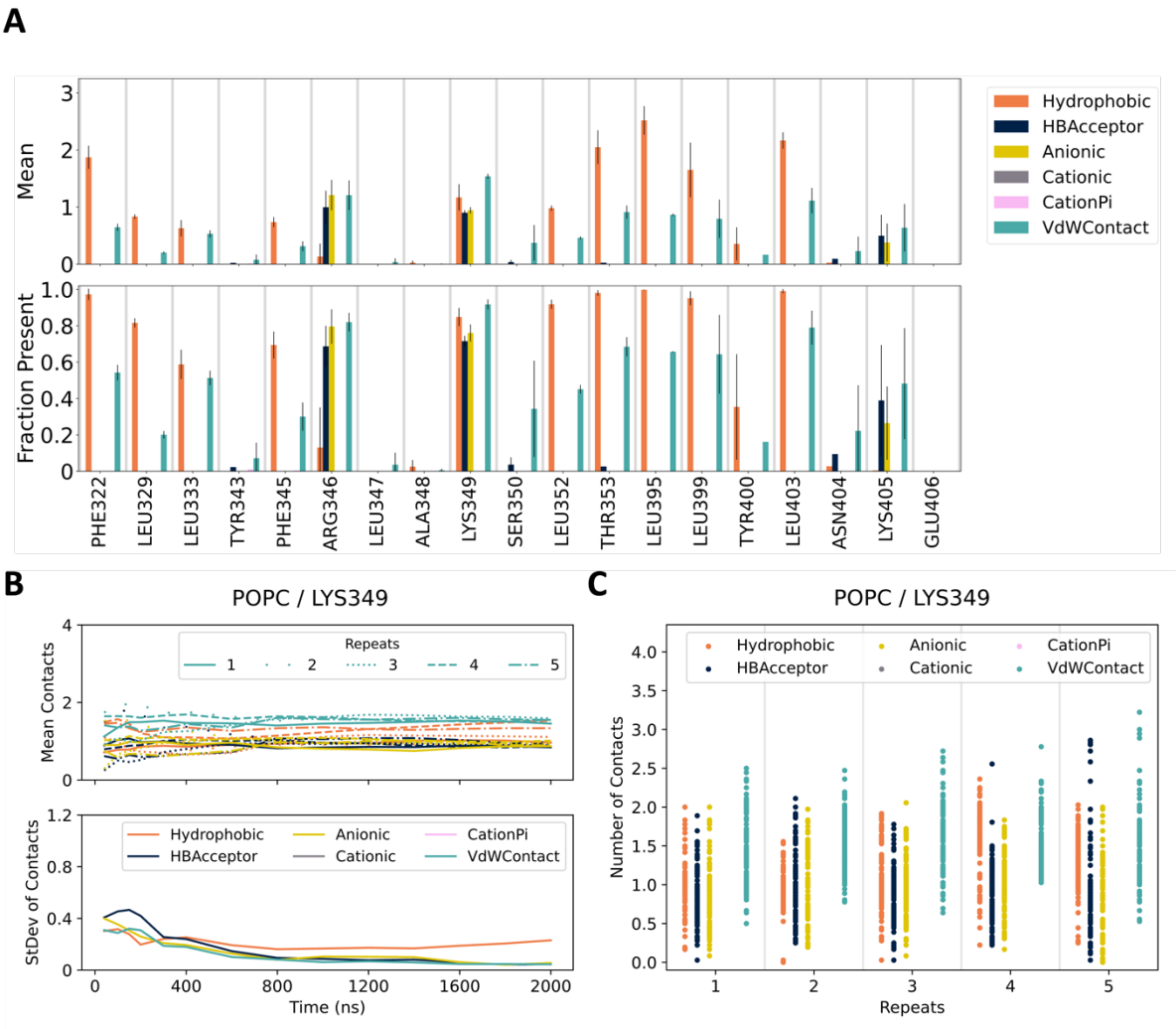


Figure 4.2: POPC/Protein Interactions. (A) The interactions between binding site residues and POPC across 5 × 2 μs apo simulations. Error bars are standard deviation of the 5 repeats. (B) The number (upper) and standard deviation (lower) of interactions that K349 forms with the bilayer. (C) The interactions formed between K349 and POPC in 20 ns chunks of the 5 × 2 μs apo simulations

4.5.1. Apo Conformations

One of the major challenges of simulating GPCRs is that their conformational ensemble is highly sensitive to environmental factors. One must therefore be careful assuming structural features observed in a holo structure are preserved in the apo structures. The MK-0893 binding site sits adjacent to an intracellular hydrogen bonding network consisting of

R173$^{2.54b}$, R346$^{6.37b}$, N404$^{8.47b}$, and E406$^{8.49b}$. The breaking of this intracellular network upon activation allows the canonical outward movement of TM6 [371,374].

It is not clear what this network should look like in the apo state, but it seems feasible that MK-0893 could stabilise a state other than that seen in the absence of MK-0893. I have already shown, for example, that the S350$^{6.41b}$ and L399$^{7.56b}$ water mediated hydrogen bond network is disrupted in the absence of ligand. If we look at the 4 other structures labelled as 'inactive' on GPCRdb [337], 3 are solved with NNC-0640 a negative allosteric modulator that binds in a similar manner to MK-0893 in which I observe similar sidechain orientations (PDB IDs: 5XF1[375], 5XEZ[375], and 4LR6[376]) and the final one is solved with NNC-1702, a partial agonist (PDB ID: 5YQZ[377]). This lack of structural information on the apo state makes it difficult to assess whether or not we are sampling the apo state in the non-interacting ABFE windows.

I therefore looked at the $5 \times 2\ \mu s$ apo simulations to assess the stability of the binding site residues in the absence of ligand. I observed an outward movement of R346$^{6.37b}$ into the bilayer, a rotation of Y343$^{6.34b}$ towards helix VII, and the movement of E406 away from helix VI (Figure 4.3A). This can be seen by tracking the distances of these residues from the midpoint of helices VI and VII (Figure 4.3B). R346$^{6.37b}$, Y343$^{6.34b}$, and E406 stay at the crystallographic distance across the holo simulations. However, in the apo simulations, we see the distances increase for R346$^{6.37b}$ and E406 and decrease for Y343$^{6.34b}$. This indicates that the ligand stabilises the crystallographic state, and in the absence of it we see a rearrangement in the sidechains of the binding site.

I tracked the interactions of these residues over the course of the simulations. We see the crystallographic interaction between R346$^{6.37b}$ and E406 breaks in the absence of ligand (Figure 4.3C) and R346$^{6.37b}$ subsequently forms additional ionic interactions with POPC (Figure 4.3D). This is further followed by the formation of interactions Y343$^{6.34b}$ with K405 (Figure 4.3E). The failure of repeat 5 to form this interaction could explain the reformation of R346$^{6.37b}$ and E406 interactions in this repeat alone. E406 forms stable interactions with R173$^{2.54b}$ in repeats 1 and 4, whilst failing to stably form this interaction in repeats 2, 3, and 4 (Figure 4.3F).

This highlights the difficulty of sampling the complex interaction networks in GPCRs. The changes in the interaction network indicate the crystal structure may not be highly represented in the conformational ensemble in the absence of ligand. The structures at the endpoints of these simulations may more closely resemble a more populated state. I will therefore refer to these as the 'true' apo states for convenience. However, I should note a single state is not the correct way to think about the conformational ensemble of a GPCR, and that further rearrangements could occur that would identify a more favourable conformation but generating a converged conformational landscape is beyond the scope of this thesis.

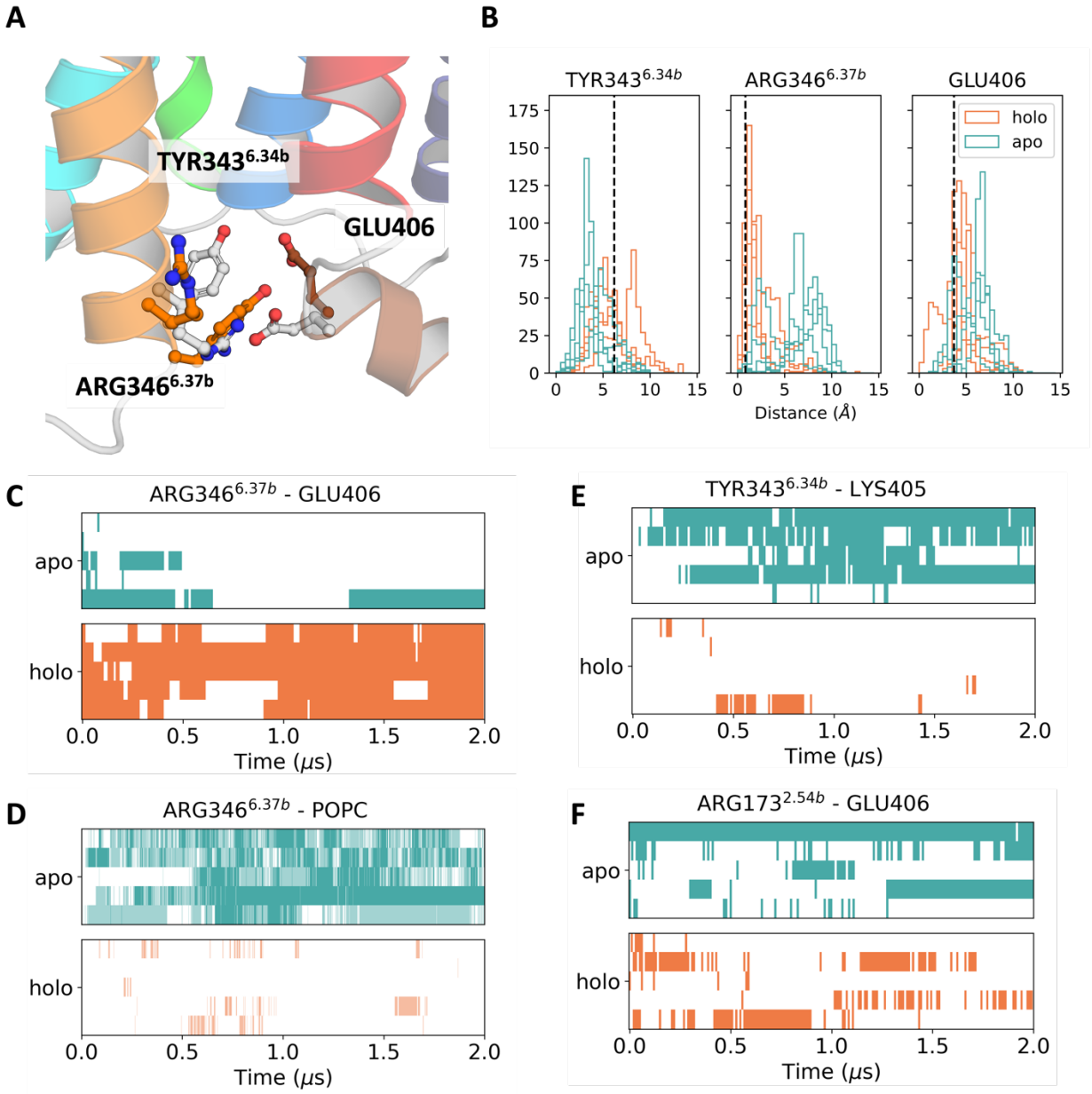


Figure 4.3: Apo conformational changes. (A) Structure of selected binding site residues (Y343$^{6.34b}$, R346$^{6.37b}$, and E406). The crystal orientation is shown as grey sticks and a representative snapshot of the 'true' apo state from repeat 4 is shown as coloured sticks. (B) Histograms of distance from the midpoint of helices 6 and 7. The crystallographic distance is shown as a black dashed line. (C) Interactions between R346$^{6.37b}$ and E406 across 5 × 2 μs apo simulations (upper) and 5 × 2 μs holo simulations (lower) (D) Interactions between R346$^{6.37b}$ and POPC across 5 × 2 μs apo simulations (upper) and 5 × 2 μs holo simulations (lower) (E) Interactions between Y343$^{6.34b}$ and K405 across 5 × 2 μs apo simulations (upper) and 5 × 2 μs holo simulations (lower) (F) Interactions between R173$^{2.54b}$ and E406 across 5 × 2 μs apo simulations (upper) and 5 × 2 μs holo simulations (lower)

4.5.2. Binding Site Occupancy

We have shown that the protein/POPC interactions are complicated in the binding site and should be monitored carefully during ABFE calculations. Another consideration to monitor is the potential for water to occupy the binding site in the absence of ligand. It is not clear ad hoc whether water or lipid should be present here.

To investigate this I calculated the protrude distances of protein, POPC, and water across the $5 \times 2\ \mu s$ apo simulations (Figure 4.4A). This reveals a complicated picture. In general, we see an amount of water and protein in the binding site, but it is predominantly occupied by POPC. However, the exact amounts of POPC and water in the binding sites fluctuate within simulations and are different across simulations. Breaking down the protrude distance analysis into the 3 prongs, we see even less consistency. Looking at prong 1, we see repeat 1 averages protrude distances of 0.74 and 0.84 for water and POPC respectively, whilst for repeat 5 we see protrude distances of 0.89 and 0.50. As one would expect due to their competing for the same space, the protrude distances for POPC and water with prong 1 are negatively correlated (Figure 4.4D). Figure 4.4B shows the protrude distances for repeat 3 and how much they fluctuate over the course of the simulation whilst Figure 4.4C shows the oxygens of waters in the binding sites at high and low protrude distance frames.

These differences are likely caused by changes in the protein conformation, as well as slow lipid dynamics. For instance, repeats 1 and 4 appear to show the greatest transition towards the 'true' apo state and also see the least POPC occupying prong 1. To assess this further, I defined frames as being 'apo like' or 'holo like' based on whether the $R346^{6.37b}$ to helix VI/VII midpoint is more or less than a 5 Å cutoff respectively (Figure 4.4E). This is not a perfect definition as the apo like conformation appears more complex than this. However, we

see the ‘apo like’ state has a mean POPC protrude distance of 0.74 whilst the ‘holo like’ had a mean of 0.58 and they have water protrude distances of 0.82 and 0.87 respectively. This supports the idea that there may be different water compositions in the apo and holo states. The dynamic nature of the occupancies makes definitive conclusions challenging, but it is clear that we will not fully sample the volumetric overlaps in ABFE windows, and we should carefully monitor what is entering the binding site.

This section has shown that MK-0893 binding to GCGR is an extremely complicated system. There is an array of polar and charged interactions between the ligand/POPC and the protein making sampling these interactions particularly important. There also appears to be slow transitions between the crystal structure and the ‘true apo’ state. This transition not only appears to involve side chain transitions, but also seems to involve the disruption of the water mediated hydrogen bond network, novel interactions with the lipid, and potentially water entering the binding site. All of these factors could seriously hinder our ability to generate precise and accurate results from free energy calculations.

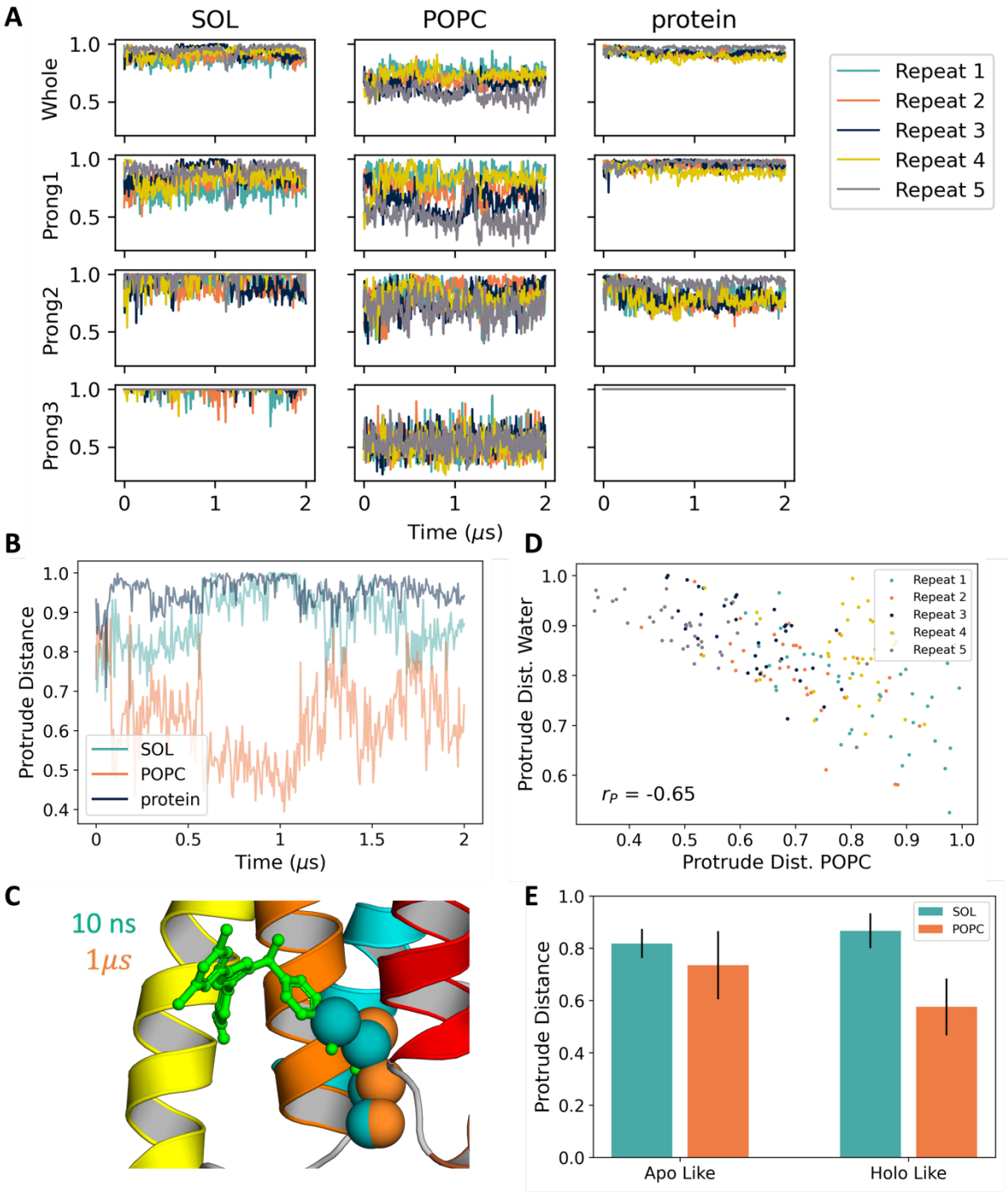


Figure 4.4: Apo Protrude Distance. (A) The protrude distance of MK-0893 and its substituent prongs with POPC, protein, and water over 5 × 2 μs apo simulations. (B) The protrude distance of water, POPC, and protein against prong 1 for repeat 3. (C) Binding site of selected frames from repeat 3(10 ns teal and 1 μs orange). Waters in the binding site are shown with crystallographic ligand overlaid (green sticks). (D) Protrude distance of water plotted against the protrude distance of POPC (E) Frames from the 2 μs trajectories are defined as 'apo like' or 'holo like' based on whether the R346$^{6.37b}$ to helix VI/VII midpoint is more or less than 5 Å cutoff respectively. Error bars are the standard deviation of the 5 × 2 μs repeats.

4.6. RBFE Calculations

Due to the complexity of the system, I first wanted to assess how well RBFE calculations would work for this system. This is a novel challenge for RBFEs as the first RBFEs at a lipid exposed binding site was only published in 2021 for the simpler $P2Y_1R$ system [258]. RBFEs are used in academia and industry to suggest whether elaborations are likely to improve affinity. I used a standard alchemical cycle for RBFEs (Figure 4.5A). To try and mimic a drug discovery program, I ran 3 RBFE calculations per transformation on two series reported by Xiong *et al.* (Figure 4.5B) [270].

The 2 vectors I assessed are equivalent to prongs 2 (referred to from now as vector 1) and 3 (vector 2) in MK-0893. Vector 1 interacts with the protein whilst vector 2 purely interacts with the bilayer. Modifications in areas that interact with lipid may provide additional challenges. For instance, it may be challenging to sample novel lipid interactions that are introduced by the new modification or the lipid environment may be slow to equilibrate around transformations. I use a simple 'hub and spoke' map (where the proposed ligands are connected to the lead compound only). For vector 1 the hub molecule is ***2*** from Xiong *et al.* [270] whilst for vector 2 I use both ***2*** and ***9m*** (MK-0893).

The RBFEs returned positive results. We see good agreement between repeats for ligands, with $\sigma < 0.5\ kcal/mol$ for 29/34 of the transformations. The overall results indicate a strong correlation – $r_P = 0.85\ [0.69,0.92]$. The magnitude of the difference is generally in agreement with experimental ($MAE = 1.21\ [0.96,1.48]$ and $RMSE = 1.44\ [1.21,1.74]$) with 29/34 out of the transformations within 2 $kcal/mol$of the experimental value.

We see a tendency to overestimate the magnitude of $\Delta\Delta G^{\theta}_{binding}$ with the gradient of the line of best fit = 1.8. For 4 of the 5 transformations where the predicted differs from experimental by more than 2 $kcal/mol$, $\Delta\Delta G^{\theta}_{binding}$ is more positive than it should be. Of these, the two furthest from experimental both involve ***15a***. The transformations $\mathbf{9m} \rightarrow \mathbf{15a}$ and $\mathbf{2} \rightarrow \mathbf{15a}$ differ from experimental by 2.9 $kcal/mol$ and 2.8 $kcal/mol$ respectively. The transformations involve going from 3,5-dichlorophenyl substituents to a trifluoromethyl. This is a relatively large transformation. I note the next smallest substituent in vector 2 (***15b*** with a tert-butyl group) also sees a discrepancy of 1.5 $kcal/mol$ for $\mathbf{9m} \rightarrow \mathbf{15b}$.

The individual subseries (vector 1 with ***2*** as hub, vector 2 with ***2*** as hub, and vector 2 with ***9m*** as hub) also perform well with $r_P = 0.84$ [0.52,0.95], $r_P = 0.71$ [−0.58,0.95], and $r_P = 0.69$ [−0.17,0.95] respectively. Vector 1 performs better than both vector 2 subseries. Vector 2 protrudes into the bilayer so this could indicate a challenge in sampling these transformations. However, it is worth noting there is significant uncertainty around the true correlation coefficients due to the small number of repeats and the low range of affinities. Combining the 2 vector 2 series of transformations yields $r_P = 0.86$ [0.63,0.95]. It would therefore appear that we can account for differences in the directly lipid facing substituents. The SAR series also contains 2 sets of stereoisomers: **9m** and **9n,** and **9p** and **9q**. The experimental $\Delta\Delta G^{\theta} =$ 2.9 $kcal/mol$ and −1.2 $kcal/mol$ respectively. In both cases, the preferred stereoisomer is clearly picked out ($\mathbf{2} \rightarrow \mathbf{9m} = -3.3 \pm 0.38, \mathbf{2} \rightarrow \mathbf{9n} = 3.7 \pm 0.44, \mathbf{2} \rightarrow \mathbf{9p} = -3.3 \pm 0.55, and\ \mathbf{2} \rightarrow \mathbf{9q} = -0.9 \pm 0.08$).

This is a complicated system for RBFEs and the results are positive. The good correlation shows that we can capture trends in the SAR despite the complicated position in the bilayer. I show that the effect of proposed elaborations which purely interact with the bilayer can be

predicted with comparable to those that directly interact with the protein.The presence of numerous charged groups and water does not lead to prohibitively large standard deviations. Although further validation is needed, this result supports the idea that RBFEs can be reliably used at lipid-exposed sites.

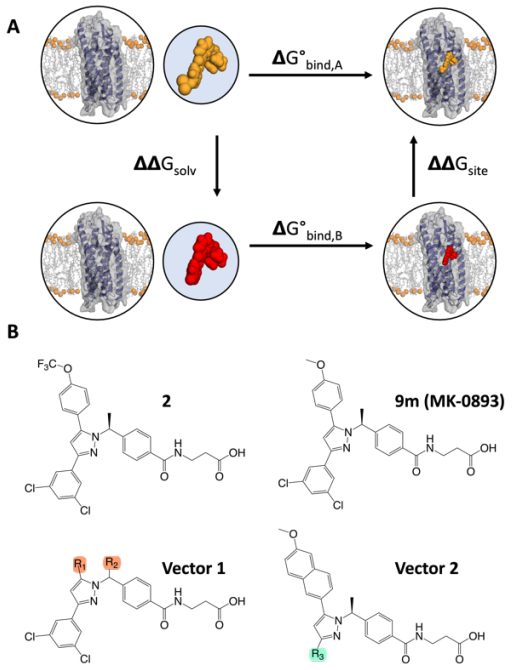


*Figure 4.5: RBFE Cycle and SAR. (A) Schematic of the alchemical RBFE cycle (B) Structures of hub molecules (**2** and **9m**) and the vectors of the SAR series.*

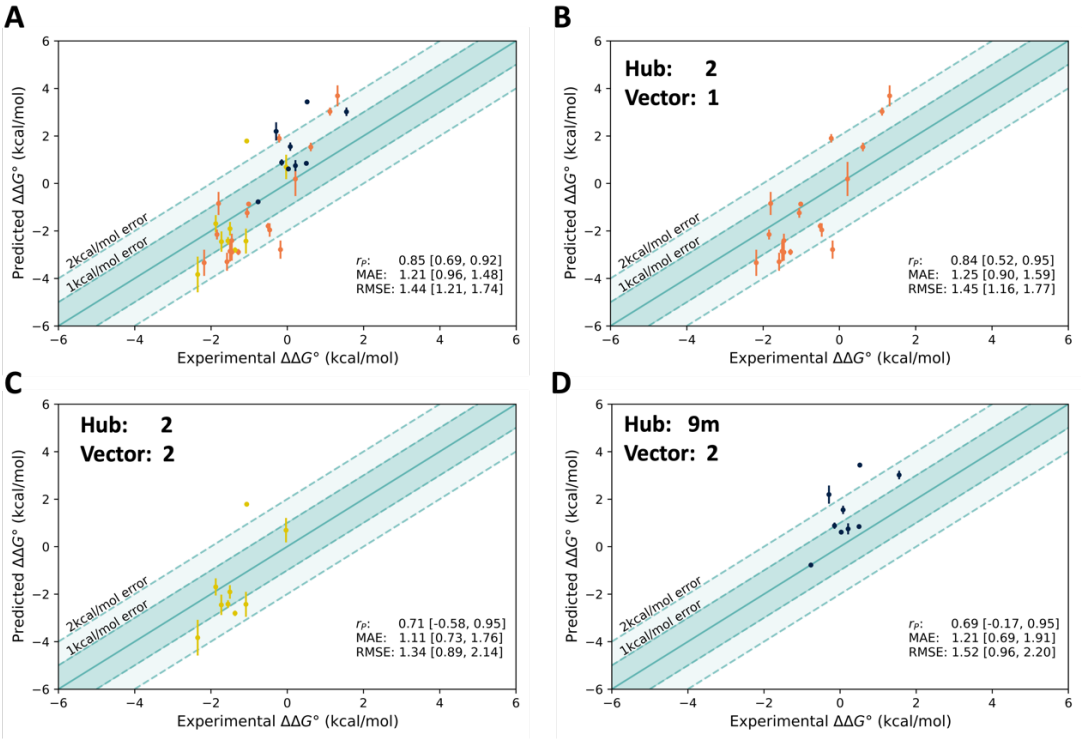


*Figure 4.6: RBFE Results. (A) RBFE results of all 3 vectors. Error bars are the standard deviation of 3 repeats (B) RBFE results of vector 1 with **2** as the hub molecule (C) RBFE results of vector 2 with **2** as the hub molecule (D) RBFE results of vector 2 with **9m** as the hub molecule.*

4.7. ABFE Calculations

Having shown that RBFE calculations can rank the MK-0893 SAR, I next looked to assess ABFE calculations. I suspected these would be significantly more challenging than RBFEs. Additional challenges include the larger alchemical region, the lack of cancellation of errors in lipid sampling, the change in charge, and regenerating the correct apo state.

We first needed to adapt the workflow to work for the GCGR/MK-0893 system. A major difference was the presence of a charged ligand. Following current best practices [313], I used a co-alchemical ion to maintain charge neutrality. Issues around the handling of nonbonded exclusions observed in the free energy code in GROMACS 2021 [372,373] meant we needed to perform annihilation (via the use of *'couple-intramol = yes'*) or to perform decoupling we should morph the atoms into atomtypes that did not interact with the non-ligand atoms (referred to as the 'dummy' method). They should theoretically provide identical answers. The decoupling scheme was deemed preferrable as the solvation free energy is a more physically intuitive concept and the maintenance of intramolecular interactions could prevent the formation of unrealistic conformations. I validated the use of both decoupling and annihilation schemes using Sage.

I ran our initial calculations using GAFF2, the dummy atom free energy method, and the co-annihilation of a sodium ion to maintain neutral charge. These initial simulations, 3 repeats from 3 separate embeddings, returned $\Delta G^{\theta}_{binding} = -22.5 \pm 1.8\ kcal/mol$. These repeats are in good agreement with one another and appeared internally converged. However, they significantly disagree with the experimental value of $\Delta G^{\theta}_{binding} = -11.6 \pm 0.4\ kcal/mol$ (Figure 4.7A) [270].

GAFF2 is known to struggle parameterising ligands with heterocyclic rings. I therefore used sage to run 5 repeats per embedding for 3 embeddings (totalling 15 repeats). I have shown this was an issue for $P2Y_1R$ but the complexity of the binding site and presence of lipid headgroups meant I suspected the potential for systematic biases was greater. The overall results ($-24.0 \pm 1.3\ kcal/mol$) are similar to that observed for GAFF2.

We see a dependence on the starting conformations - -25.2 ± 0.8,-24.2 ± 0.7, and-22.6 ± 0.7 respectively. The breakdown of the standard deviations for the coulombic and vdW legs were very similar ($\sigma_{complex\ coul} = 1.7$ and $\sigma_{complex\ vdw} = 1.5$). The high standard deviation could be explained by the variable positions of lipid headgroups proximal to the ligand.

I hypothesized that these interactions could be very slow to equilibrate so I extended one of the repeats to 200 ns per window for the vdW leg and 100 ns per window for coulombic legs (Figure 4.7B). This is far longer than a typical ABFE window. Breaking this down into 20 ns chunks we see that the predicted ΔG° does not tend towards a different value, but instead oscillates around the initial value. The 20 ns chunks vary by as much as 1.8 $kcal/mol$ for the vdW leg and 2.0 $kcal/mol$ for the coulombic leg. This lack of internal consistency indicates a slow degree of freedom we are struggling to sample which I hypothesise could be the ligand's interactions with the bilayer. I also ran 1 repeat from each of the 3 starting configurations using HREX. This did not improve the results ($-24.0 \pm 0.9\ kcal/mol$).

The utilisation of Sage did not improve the results, but I suspected the 9 rotatable bonds of MK-0893 could be challenging to parameterise accurately, so I optimised the torsions using

BespokeFit (Supplementary Figure 9.4). The torsional profiles better fit the QM data but the ABFE results ($-22.2 \pm 0.9\ kcal/mol$) does not bring us substantially closer to the experimental value.

Unless specified otherwise, subsequent simulations were run using sage. I also assessed whether charging a similarly charged ion ('charge transfer') proved to be beneficial over co-annihilation of an oppositely charged ion. This also did not significantly change the results ($-24.4 \pm 1.2\ kcal/mol$). The independence of this discrepancy from forcefield, simulation length, and methodology indicated there was a systematic issue with our system. The next sections will discuss our attempts to identify the sources of this discrepancy.

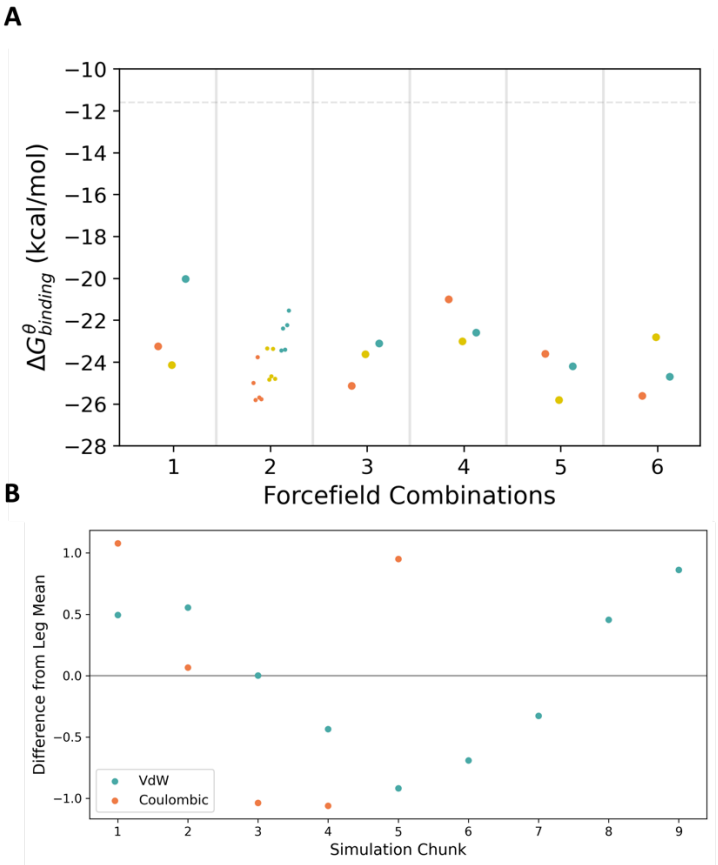


Figure 4.7: GCGR/MK-0893 Forcefields. (A) ABFEs run with different forcefields and simulation protocols. Unless specified, these calculations used sage, dummy decoupling, co-annihilation of an alchemical ion, and no HREX (1) GAFF2 (2) Sage dummy (3) HREX (4) BespokeFit (5) Annihilation (6) charge transfer of an alchemical ion. The experimental value is shown as a grey dashed line (B) 20 ns chunks of 1 repeat of 100 ns coulombic windows and 200 ns vdW windows.

Our first interest was in looking at the interactions between POPC and the protein in the non-interacting window. We see significant variation across the repeats (Figure 4.8A). For instance, K349$^{6.40b}$ forms on average 6.6 interactions with POPC in one repeat and as few as 1.7 in another. There are some instances of significant differences between different configurations (for instance F345$^{6.36b}$ differs between repeats 2 and 3 ($p < 0.01$)) but there are large differences within configurations. This may contribute to the relatively large standard deviation between repeats.

This data can be compared to data obtained from the $5 \times 2\ \mu s$ simulations (Figure 4.8B). This shows that there are several residues that differ substantially from the long simulations. L352$^{6.43b}$, for instance, forms significantly more interactions in the repeat 1 windows than in the long simulations. Repeat 3 shows fewer interactions between POPC and the protein than the other repeats and fewer than the long simulations. This is interesting as repeat 3 has a lower magnitude predicted $\Delta G^{\theta}_{binding}$ than the other embeddings. This could be caused by a more stabilised apo state, which one may have suspected beforehand involved an increased number of contacts. This either acts as a reminder that the number of contacts alone is not a direct proxy for interaction strength, or that there is a different cause of the difference between repeats.

I was also interested to see whether our equilibration process produces similar interaction profiles. To do this I ran 25 equilibration mimicking simulations of $5\ ns$ for each configuration. I plotted the presence of an interaction between the residue and POPC in the last frame and coloured the starting configuration based on whether there is a dependence on the starting configuration (Figure 4.8C) and whether the starting configuration distribution

differs from the long simulations (Figure 4.8D). We see that there is no dependence on the starting configuration in terms of the contacts formed. This is likely a combination of the fact that we are running a relatively long equilibration and our use of the Bonferroni correction [378].

We do see some of the residues form significantly different interactions than the long simulations. However, this may at least partly be a function of the dynamic nature of the protein. For instance, we see R346$^{6.37b}$ forms significantly fewer interactions in the equilibration endpoint distribution than in the long simulations. I have shown earlier that a change in sidechain orientation on the order of 100s of ns facilitates these additional interactions. It is therefore not realistic to expect the equilibration mimicking simulations, which are comparatively very short and apply backbone restraints, to see these conformational changes. To offset this as best we could, I drew our distributions from the first $500\ ns$ of the long simulations.

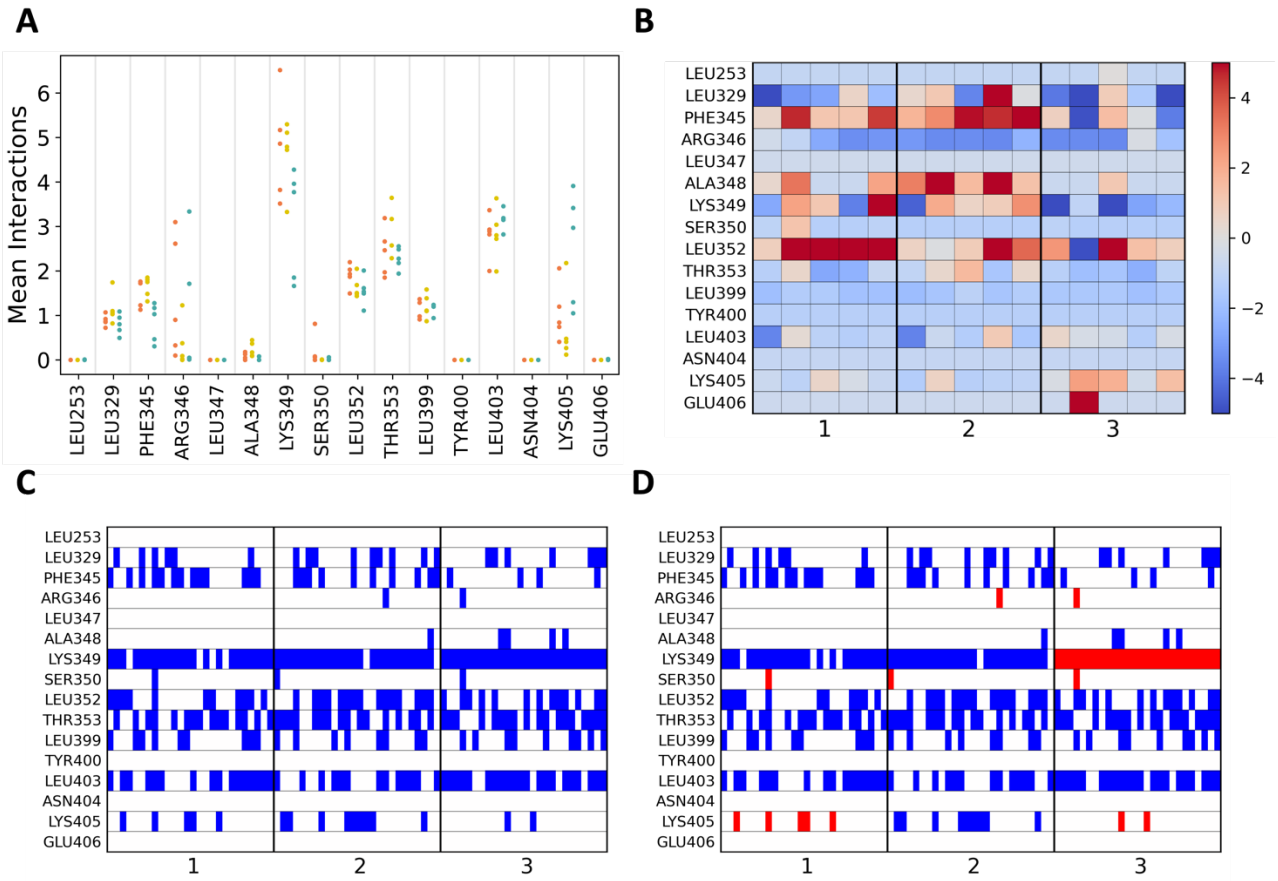


Figure 4.8: vdW 20 Protein/POPC Interactions. (A) The mean number of interactions formed between POPC and the binding site residues across the non-interacting window. (B) The average POPC/protein interactions for the non-interacting windows coloured by the number of standard deviations (from 5 × 2 μs simulations) the average is from the long simulation mean. (C) The presence of interactions at the endpoint of 5 ns equilibration mimicking simulations. If there is a starting dependence, it is coloured red and if not it is coloured blue (D) The presence of interactions at the endpoint of 5 ns equilibration mimicking simulations, coloured red if the starting configuration is significantly different from the long apo simulations.

Whilst the variable POPC interactions may contribute to the standard deviation, it does not provide a clear explanation for the discrepancy between experimental and predicted. As I have shown before, it is not inherently clear what should be present in the binding site in the absence of ligand. Long simulations show a complicated picture where the protrude distance for POPC and water fluctuates significantly across the simulations (Figure 4.9A) and that there may be a link to the conformation of the binding site residues. I hypothesised the presence of a vacuum in the absence of ligand may lead to an artificial state where the water

over occupies the ligand's volume, particularly with regard to prong 1 as it fills the space more rapidly. To assess this, I looked at 90 repeats (30 repeats for each starting configuration) of 5 ns equilibration mimicking simulations run *without* restraints on the water (Figure 4.9B). Water consistently occupies prong 1 more than POPC. Looking at the endpoints, I see the mean protrude distances for water and POPC are 0.69 ± 0.10 and 0.86 ± 0.12 respectively. Whilst the picture is complicated for the long simulations, I tend to see more POPC occupying the space than water (0.70 ± 0.13 and 0.83 ± 0.06).

I hypothesised this was an artefact caused by the faster dynamics of water allowing it to occupy the ligand vacuum. I ran 30 repeats (10 repeats per starting configuration) of 5 ns equilibration mimicking simulations run *with* restraints on the water (Figure 4.9C). As we would expect the amount of water in the binding site significantly reduces (0.94 ± 0.03). This also leads to the amount of POPC in the binding site increasing (0.73 ± 0.13). The large standard deviations make interpretation difficult, but it would appear these restraints create a different binding site environment which may better resemble the true apo state. If we look at the equilibration mimicking simulations with restraints, we see a dependence on the starting configuration. Looking at the POPC protrude distance with prong 1 for the final frame, the averages for conformation 1, 2, and 3 are 0.67 ± 0.11, 0.67 ± 0.14, and 0.84 ± 0.06.

Conformations 1 and 2 are significantly different from conformer 3 (T-test $p < 0.005$ for both conformations). This is interesting as conformer 3 provided the least negative ABFE predictions and in our analysis of the long simulations we see water occupancy as a potential indicator of transitioning to a more apo like state. To see if this would translate into improved results, I adapted our equilibration process to position restrain the waters for 2.6 ns of simulation before doing a further 3 ns with the waters unrestrained. This was with the aim of

allowing POPC the time to enter the binding site, whilst reducing the chance of additional artefacts from introducing the restraints. To see if this improved results, I compared ABFEs from the two equilibration schemes. The altered equilibration process did not improve our ABFE results ($-24.0 \pm 1.3\ kcal/mol$) compared to the unrestrained results ($-23.8 \pm 1.5\ kcal/mol$).

In this section I have looked at the water occupancy in the binding site in the non-interacting windows. I have shown that leaving the waters unrestrained throughout the equilibration process leads to an over representation of water in the binding site. This can be prevented by applying position restraints to the water. This however risks introducing new artefacts, so I propose an intermediate equilibration process whereby the waters are initially restrained to allow POPC to enter the binding site before being released. Ultimately, this does not lead to a change in ABFE results, indicating there are other factors at play.

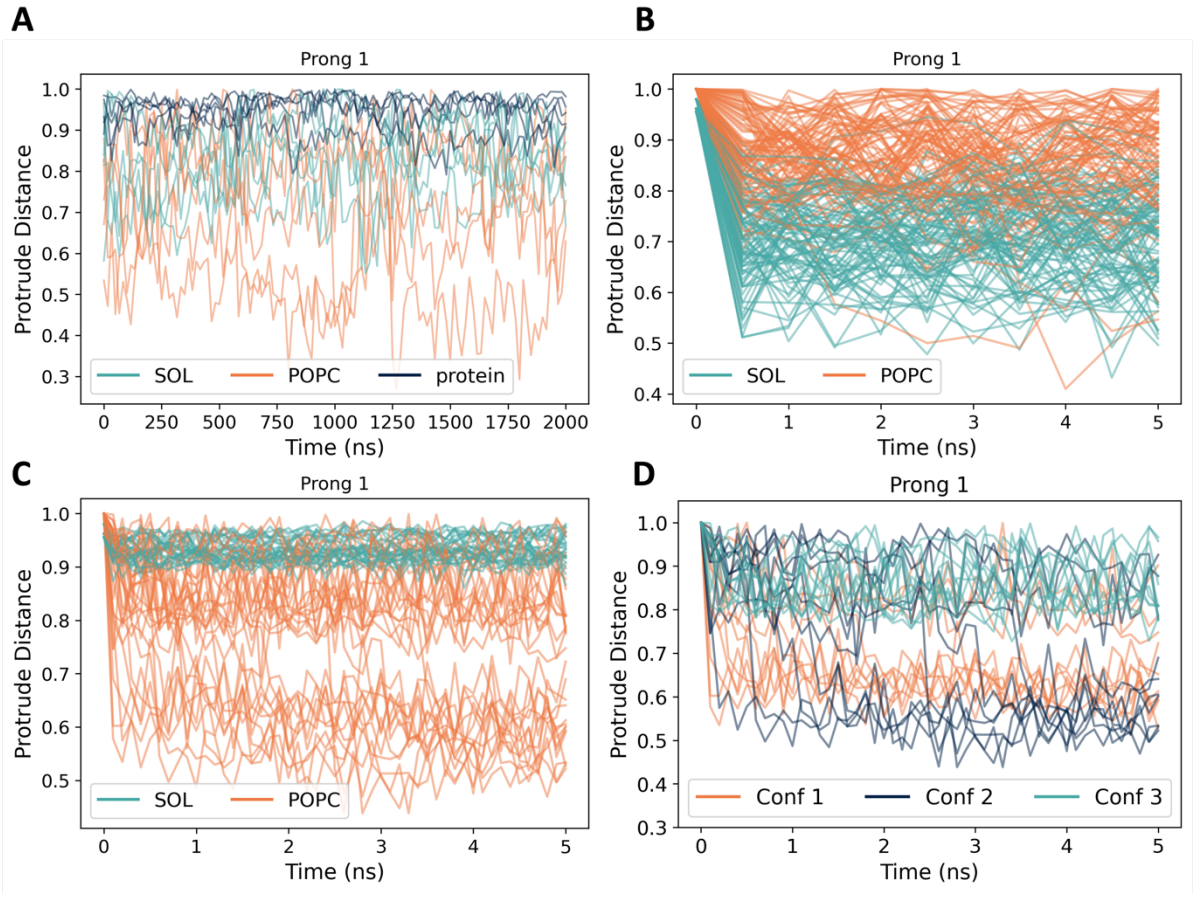


Figure 4.9: vdW20 Protrude Distances. (A) The protrude distances for prong 1 with water (teal), POPC (orange), and protein (dark blue) over 5×2 *μs unbiased apo simulations. (B) Protrude distance for prong 1 with water (teal) and POPC (orange) in 5 ns equilibration mimicking simulations without water restraints (C) Protrude distance for prong 1 with water (teal) and POPC (orange) in 5 ns equilibration mimicking simulations with water restraints (D) Protrude distance for prong 1 with. POPC in 5 ns equilibration mimicking simulations with water restraints. The lines are coloured by starting configuration.*

4.8.SAR ABFEs

As I expected we were facing a systematic problem with the ABFE calculations, I decided to proceed with analysing the SAR data for MK-0893 [270]. The structures of these SAR molecules, their predicted affinity, and experimental affinity can be seen in Supplementary Figure 9.5 or the structures alone can be seen in Figure 4.10. This SAR series primarily focuses on two vectors that test different groups at prongs 2 and 3 (Figure 4.5B). All molecules share the β-alanine group that sits between helices VI and VII and that I suspected may be causing some structural rearrangements which may lead to the disparity. Looking at multiple ligands will allow us to test whether this is a systematic issue caused by all of these ligands and whether we are still able to accurately rank the molecules.

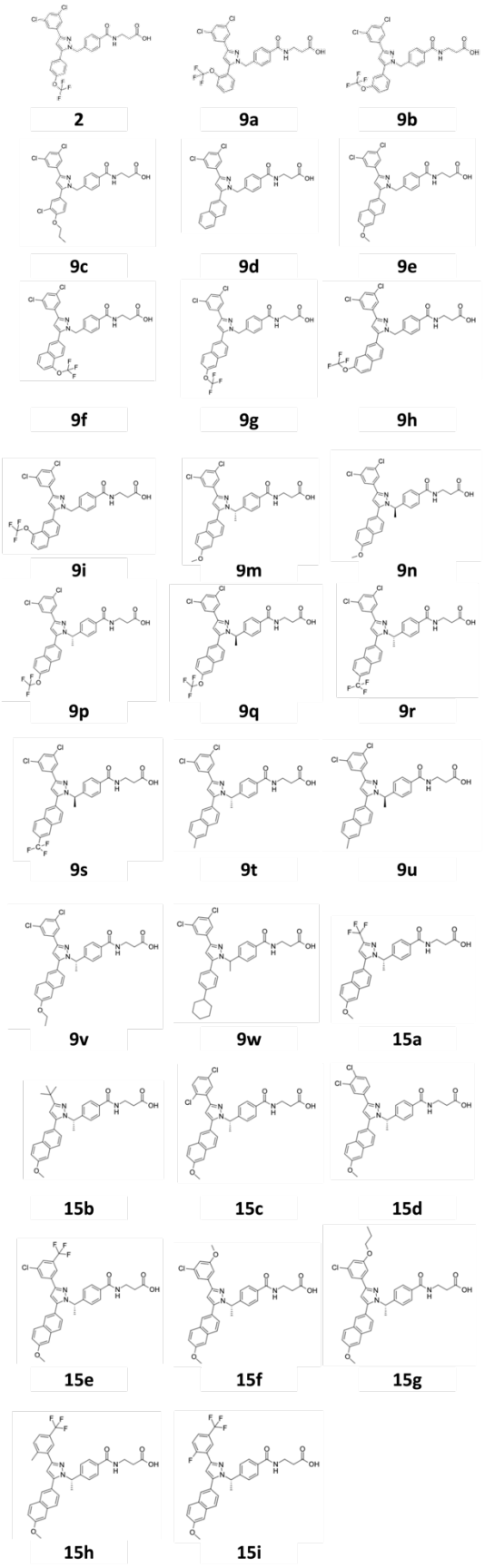


Figure 4.10 GCGR SAR

I ran ABFEs on the SAR molecules using our initial equilibration process ($5.6\ ns$ of equilibration per window with no water restraints) (Figure 4.11A), our adapted equilibration process ($5.6\ ns$ of equilibration per window, $2.6\ ns$ with water restraints followed by $3\ ns$ unrestrained) (Figure 4.11B), and our adapted equilibration process with HREX (Figure 4.11C). Considering the complexity of the system, the results are positive. We see relatively strong correlation coefficients ($0.46\ [0.10,0.77]$, $0.45\ [0.13,0.69]$, and $0.60\ [0.23,0.80]$) and consistently predict the affinities of the tightest binders of being of the largest magnitude. We see no major differences between the statistics of the two equilibration schemes, although the adapted scheme appears to have a tighter grouping of molecules as evidenced by the smaller difference between MAE confidence intervals. All 3 setups have similar MAE ($12.55\ [11.43,13.27]$, $12.05\ [11.33,12.70]$, and $11.64\ [10.90,12.23]$). The standard deviations between repeats are relatively high across all the different equilibration schemes (Figure 4.11E).

Breaking down the standard deviations into coulombic and vdW components shows that the coulombic contribution tends to be higher than the vdW (Figure 4.11F). This is in contrast to $P2Y_1R$ where we saw the predominant contribution came from the vdW legs. I suspect this is due to the fact the binding site is proximal to the lipid headgroups and that we are perturbing a charged ligand.

The use of HREX leads to improved correlation but does not improve the MAE. I looked at our ability to predict the SAR from the different vectors (Figure 4.11D) and both have strong internal Pearson correlation coefficients ($0.74\ [0.31,0.93]$ for prong 2/vector 1 and $0.74\ [-0.04, 0.97]$ for prong 4 / vector 2). This is interesting as it indicates lipid facing

modifications are as well predicted as those facing the protein. However, I should note the confidence intervals are much broader for prong 3 as there are few molecules at the lower magnitude end of the dynamic range. If we look at the HREX outliers, we see **15a** is lower than the rest of the SAR with a MAE of 6.9 compared to the average of 11.51 for the SAR series. It also has a comparatively low MAE in the other simulation setups. A potential explanation for this is a challenge in parameterising the molecule. It has a trifluoromethyl at prong 3. Although fluorine is not typically considered as a strong halogen bond donor due to the limited size of the σ-hole, halogen bonds have been shown to be important in small molecule membrane interactions [379] It could therefore be interesting to see if modelling the σ-hole would lead to improved results for the halogenated molecules binding at lipid-exposed sites.

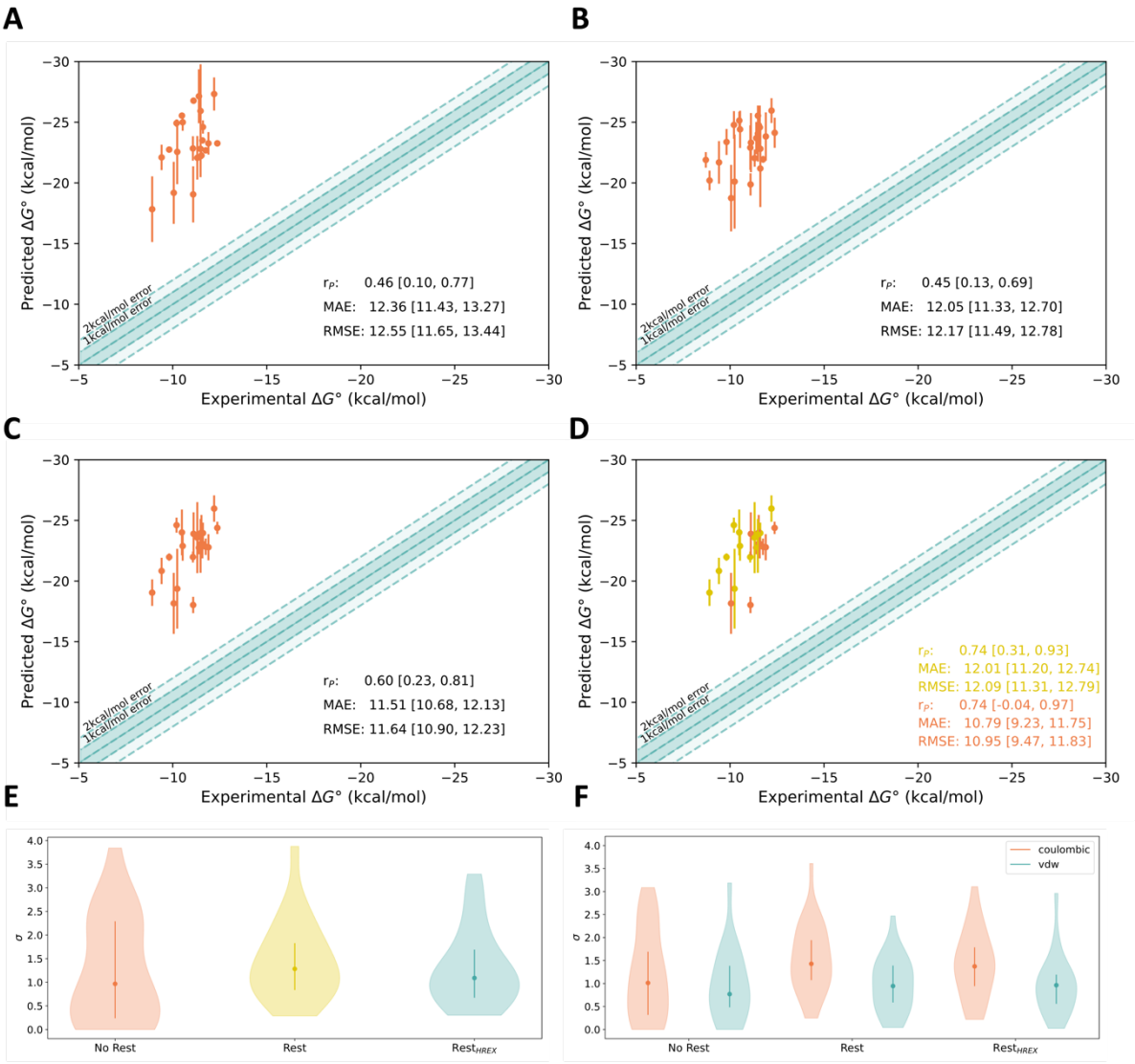


Figure 4.11: GCGR SAR ABFE Results. (A) ABFE results for GCGR SAR series using an equilibration process entirely with unrestrained waters (B) ABFE results for GCGR SAR series using an equilibration process initially with restrained waters and subsequently with unrestrained waters (C) ABFE results for GCGR SAR series using an equilibration process initially with restrained waters and subsequently with unrestrained waters. The production simulation was run using HREX. (D) The data in subplot C coloured by whether the SAR is elaborating vector 1/prong 2 (yellow) or vector 2/prong 3 (orange). (E) The standard deviations of the repeats for individual ligands across the 3 simulation setups. 'No rest' is simulations run without water restraints on during the equilibration, 'rest' is with water restraints during the equilibration, and 'RestHREX' equilibration is run with water restraints and production with HREX. The points are the median values and the whiskers extend to the first and third quartiles (F) The standard deviation broken down into coulombic and vdW contributions

4.8.1. ABFEs from 'true' apo conformation

I have shown that across 2 μs simulations we see the formation of a relatively stable conformation that I term the 'true' apo state. This involves a series of sidechain movements, rearrangements of the water network, and the formation of novel interactions between the protein and POPC. These interactions take on the order of 100s of ns to form and if we do not see a full transition, they can revert to the holo like structure. Sampling this transition is vital if we are to get accurate ABFE results but due to the short timescales of ABFE windows I expected we would likely not see this transition. I tracked 4 indicators of the apo transition: the breaking of R346$^{6.37b}$ and E406 interaction (Figure 4.12A), the formation of ionic interactions R346$^{6.37b}$ and POPC (Figure 4.12B), the formation of interactions between Y343$^{6.34b}$ and K405 (Figure 4.12C), and the loss of stability of the crystal water. We see the R346$^{6.37b}$ and E406 interaction is broken in most of the simulations but several repeats see it present for the entire simulation and it is maintained for significant chunks of other simulations. The R346$^{6.37b}$ and POPC ionic interactions are only formed for more than 50% of the simulation in 4/15 of the simulations and 2 ionic interactions are only formed in 7% of frames, whilst for the long apo simulations it is present in 37% of frames.

In the long simulations, we saw water mediated hydrogen bonds between S350$^{6.41b}$ and L399$^{7.56b}$ in 0.91 ± 0.09 of the frames in the holo simulations and 0.37 ± 0.32 in the apo simulations. For the vdW 20 simulations we see the water bridge is overly stable, being present in 0.61 ± 0.22 of frames. This is another indicator that we are not sampling the 'true' apo state in our simulations.

We therefore decided to assess what the impact of including the apo state explicitly was. To do this, I took the endpoints of 3×1 μs simulations and added the ligand to the structure. I

note that these are not the simulations previously discussed. Adding the ligand to the structures will likely introduce clashes between the ligand and POPC so I used alchembed [355] to remove these clashes. Alchembed turns on the ligand's interactions with the environment to facilitate embedding. This was done for each λ window. To prevent excessive distortion of the apo state, alchembed was only used to turn on the atoms to the same λ as the ligand. An additional complicating factor is the prong like nature of MK-0893. I worried the alchembed method would lead to lipids being trapped. To resolve this, I placed additional dummy atoms within the ligand between prongs 1 and 2 (Figure 4.12D). These dummy atoms were parameterised so that they would not interact with protein, the ligand, themselves, or water. Their interactions were turned on concurrently with the ligand's interactions, thus allowing lipid to enter the inter-prong space in the weakly interacting windows but not the more strongly interacting windows. Validation of these atoms can be seen in Supplementary Figure 9.6.

The motivation for this is the belief that we are getting an incorrect $\Delta G^{\theta}_{binding}$ due to not sampling the correct apo state in the endpoint. There is therefore a strong argument for running the last window from the apo state. At the opposite end of the spectrum, the fully interacting window must be run from the holo state. Running either endpoint from an unfavourable conformation will lead to an energetic penalty. In the case of the holo window, it will lead to an underestimation of the magnitude of the free energy, whilst in the apo case it will lead to an overestimation of the magnitude of the free energy.

Justifying a change to the holo state that reduces the predicted $\Delta G^{\theta}_{binding}$ is problematic, whilst finding a more favourable apo state is probably a more justifiable way of reducing the predicted $\Delta G^{\theta}_{binding}$. At some intermediate λ value, there will be no preference for either of

the states. Running windows from the apo state at λ values nearer the holo state than this λ will result in a reduction of $\Delta G^{\theta}_{binding}$ by penalising the holo state. It is therefore important to prevent running excessive windows from the apo state. The 50:50 value of λ is unknown and is expected to be system specific.

If we were to run an increasing number of windows from the apo state (assuming there is an energetic difference between the apo and holo states), I hypothesise that one would initially expect to see a sharp decrease in $\Delta G^{\theta}_{binding}$ as the apo state is stabilised. We would then expect to see a plateauing in $\Delta G^{\theta}_{binding}$ as we approach the 50:50 λ value as the difference between energetics in the two states becomes less. We would then expect to begin destabilising the holo state, leading to an increasingly steep increase in $\Delta G^{\theta}_{binding}$.

I ran the final 10 windows from the 'true' apo states for all the ligands in the SAR series. I then substituted in the data from these simulations into increasing numbers of the windows (Figure 4.12E). The difference between adjacent windows can be extracted from TI. I can therefore look on a per window basis at the impact of adding additional windows from the 'apo' like state. Figure 4.12F shows the difference between $(x-1)_{holo} \rightarrow x_{apo}$ and $(x-1)_{holo} \rightarrow x_{holo}$ on a per window bases, whilst Figure 4.12G shows the cumulative impact. This shows a significant change across the windows.

Interestingly, we see substantial deviations from using all holo. Windows run from apo conformations 2 and 3 show significant deviations from the holo values on the order of 1.5 $kcal/mol$. Conformation 1 is a far larger difference, with the average deviation being 5 $kcal/mol$. This discrepancy is substantial. I assessed the starting apo structures for each configuration and note they all exhibit signs of resembling the 'true' apo state. They exhibit

the broken R346$^{6.37b}$ to E406 bond, R173$^{2.54b}$ interactions, substantial rearrangements of Y343$^{6.34b}$, and loss of the S350$^{6.41b}$ and L399$^{7.56b}$ water mediated hydrogen bond. Conformations 2 and 3 show 1 and 2 ionic interactions respectively with POPC, but interestingly, conformation 1 does not and it instead protrudes into the water. This could also be a favourable interaction given arginine's propensity to snorkel.

Another possible explanation for the reduced differential in conformations 2 and 3 is that we see movement of the extracellular ends of helices 6 and 7 respectively when compared to the crystal structure, whilst conformation 1 closely resembles the crystallographic structure in these areas. It would be interesting to repeat these simulations with more apo like structures to further assess the consistency of these observations, as well as to try and identify the contribution of specific conformational changes. I was also concerned that I may be introducing discrete breaks in the phase space overlap between windows. This poor phase space overlap could result in incorrect answers. Two ways to identify this are the overlap matrix and the use of multiple estimators (as MBAR and BAR rely on phase space overlap but TI instead depends on a smooth integrand). We see good agreement between estimators and good phase space overlap across the simulations (Supplementary Figure 9.7).

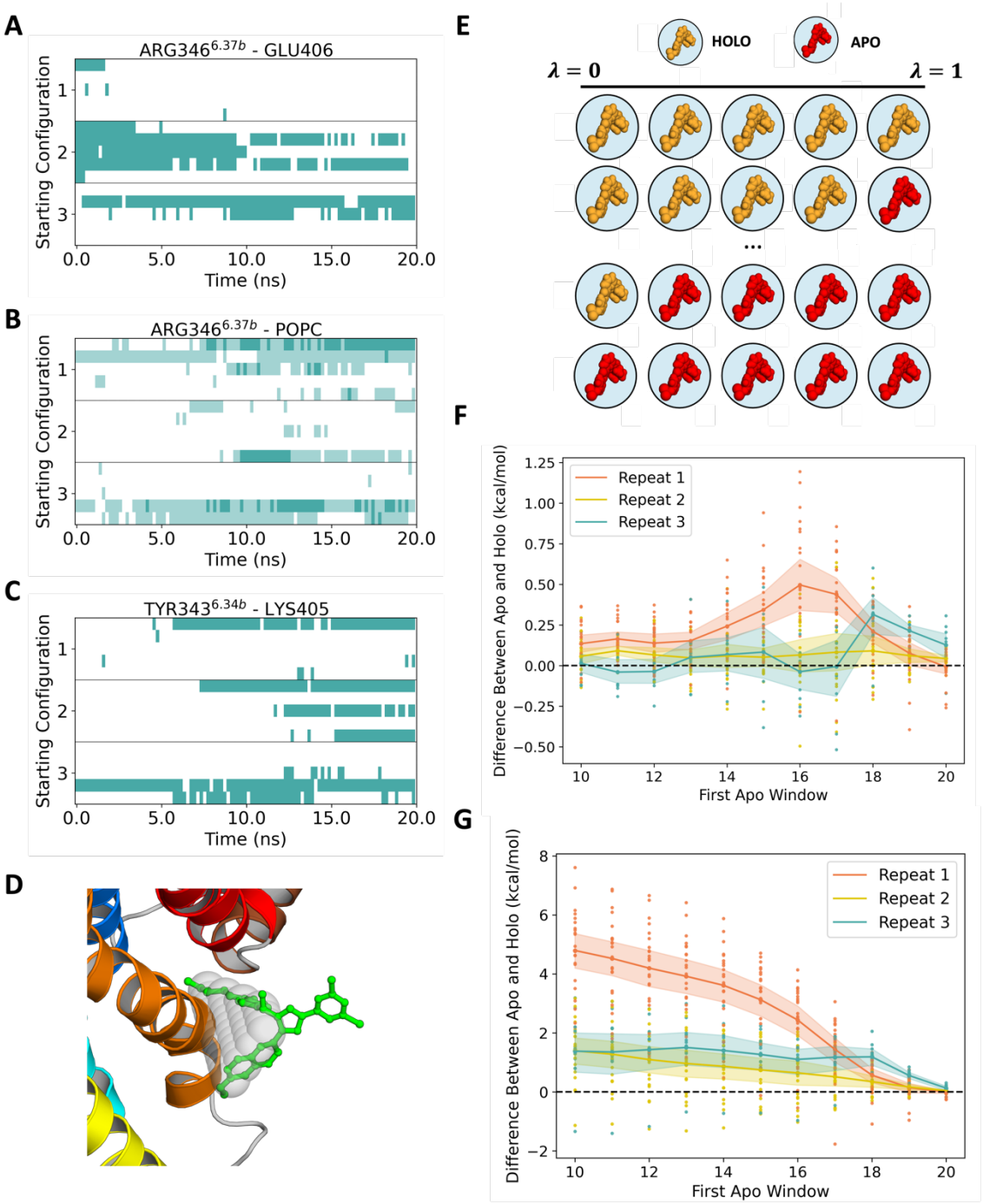


Figure 4.12: Apo/Holo ABFE Calculations. (A) The presence of R346$^{6.37b}$ and E406 interaction across the non-interacting ABFE window (B) The presence of ionic interactions R346$^{6.37b}$ and POPC across the non-interacting ABFE window (C), the presence of interactions between Y343$^{6.34b}$ and K405 across the non-interacting ABFE window (D) Dummy atoms added to prevent lipid occupying the space between prongs 1 and 2 of MK-0893. (E) Schematic of running ABFE windows using variable numbers of windows starting from the apo/holo conformation. The top row corresponds to using all holo states (as is common current practice) whilst the bottom row corresponds to using all apo. (F) The difference between adjacent windows (from TI) is calculated. The plot shows the difference between the 2 values when going from holo → apo vs holo → holo at a given window (G) The cumulative effect of these differences.

4.9. Discussion

In this chapter I have investigated the binding of MK-0893 and the corresponding SAR series to GCGR. This is an extremely challenging system. The presence of the binding site at the protein/lipid/water interface means there are numerous groups interacting with one another, meaning I need to sample a large number of interactions. These interactions are often polar/ionic in nature making sufficient sampling even more important. The binding site sits at the bottom of helix VI, which is key in GCGR activation. The outward movement of helix VI is caused by the rearrangement of several sidechain hydrogen-bond networks. The proximity of the MK-0893 binding site to these networks made us particularly wary about conformational changes in the binding site. The ligand is also large, with 9 rotatable bonds, and carries a formal charge. These factors combined makes GCGR/MK-0893 extremely challenging with a number of factors that have not been addressed in ABFE calculations before. I therefore decided to take an in-depth look at the system to better understand the complexities described above. I then preliminarily assessed the complexity by running RBFE calculations before looking at ABFE calculations. I aimed to assess the accuracy of these calculations, but I particularly focused on identifying systematic issues in the system and challenging degrees of freedom and proposing potential solutions.

I began by assessing the stability of the binding site in the presence and absence of ligand. I found the binding site remained stable throughout the long simulations. I assessed this in terms of key ligand protein interactions, protein-protein interactions, and the stability of water-mediated hydrogen bond networks. Sampling of the lipid interactions with the ligand is poor across short simulations but does improve with simulation length. The apo state, however, is significantly more challenging. We see the destabilisation of the R346$^{6.37b}$ and E406 interaction, the movement of R346$^{6.37b}$ outwards to face into the bilayer and form ionic

interactions with POPC, the formation of interactions between Y343$^{6.34b}$ with K405, and the destabilisation of the water mediated hydrogen bond between S350$^{6.41b}$ and L399$^{7.56b}$.

The dynamic nature of these interactions and the reformation of some more 'holo' like interactions in some of the simulations makes it difficult to define an explicit apo state. GPCRs exist as a conformational ensemble, meaning the concept of a single 'true' apo state is not strictly correct. However, these endpoints likely provide a state which is more populated in the absence of ligand than the crystal conformation. It would be interesting to assess this further. I ran unbiased simulations of 2 μs. Extending the simulations would provide more insight into the stability of the interactions we identified and potentially identify novel transformations. Enhanced sampling methods would allow us to explore more of the phase space, but defining meaningful CVs would be challenging [380]. Methods which do not require an explicit definition of a CV (such as GaMD [381]) could be useful in this scenario. Enhanced sampling methods would likely be necessary to get a converged picture of the energy landscape, but this would be challenging due to the number of orthogonal degrees of freedom.

Given the challenges we expected from ABFE calculations, I decided to run a series of RBFE calculations. This is a challenging system for RBFEs given the limited validation of RBFEs at lipid-exposed sites and this being a particularly challenging lipid-exposed site. Vector 2 also explores elaborations of prong 3 which protrudes into the bilayer. Again, I was unsure of our ability to sample changes at these points due to the slow dynamics of lipids. I obtained strong results with r_P = 0.85 [0.69,0.92]. This is comparable to soluble systems. I was able to distinguish between stereoisomers and saw strong results for vector 2 which protrudes into the bilayer (r_P = 0.86 [0.63,0.95]). It would be interesting to look at

molecules where modify prong 1 is modified. The strength of interactions, complexity of the bonding networks, and the potential timescales required for rearrangements to occur could make this extremely challenging for RBFEs and methodological advances may be needed to account for these slow dynamics.

Nevertheless, given the apparent success of RBFEs, I decided to run a series of ABFE calculations. We see a significant discrepancy between the experimental value ($-11.6 \pm 0.4\ kcal/mol$) and our initial results with GAFF2 ($-22.5 \pm 1.8\ kcal/mol$). Due to known issues with GAFF2, I used sage to parameterise the ligand but saw no improvement in the results ($-24.0 \pm 1.3\ kcal/mol$). I tried several other schemes (including improving the ligand parameters using BespokeFit, utilising HREX, and changing the method for perturbing the ligand) but the discrepancy with experimental remained consistent. We saw relatively large differences between repeats. Our repeats were run from unique embeddings but when I ran repeats from the same embedding I saw a starting configuration dependence (-25.2 ± 0.8, -24.2 ± 0.7, and -22.6 ± 0.7). This highlights that we are struggling to sample the lipid and that running from starting configurations is important to get a better idea of the true error. I also note that when I extended the windows there is a significant difference between 20 ns chunks. HREX did not improve the discrepancy but did reduce the deviation between repeats ($-24.0 \pm 0.9\ kcal/mol$). Additional repeats would be needed to identify whether or not this is significant.

We looked at potential causes of this. I looked at protein/POPC interactions in the apo state and whilst there are some differences between the ABFE windows and the long simulations, these appear to be linked to slow conformational changes. I also looked at what

was entering the binding site in the absence of ligand. We saw more water entering the binding site than in the long simulations. I hypothesised this was a function of the water more quickly filling the vacuum in the absence of ligand and the absence of lipid in the binding site could contribute to the discrepancy. I prevented this by adapting our equilibration scheme to position restrain the waters during the initial stages, but this did not change our results significantly ($-23.8 \pm 1.5\ kcal/mol$). During the long simulations, we see both high and low presence of water and this could explain the comparatively similar free energies.

We ran a series of ABFE calculations on the SAR series. I used 3 different methodologies – unrestrained water during the equilibration, restrained water during the first part of the equilibration, and restrained water during the first part of the equilibration with HREX during production runs. All 3 methodologies showed large MAE errors. However, we saw reasonable correlation coefficients. HREX returned the best correlation coefficient (0.60 [0.23,0.80]). This is comparable to a number of recently published soluble systems[321,327]. I showed that ABFEs can accurately rank SAR when a significant proportion of the ligand is exposed to lipid headgroups.

The long apo simulations indicated that there is a transition from the holo state to the apo state on the order of 100s of ns. This means it was unlikely we were fully sampling (or potentially even finding) the 'true' apo state in our short ABFE simulations. We must sample the apo state in the non-interacting window, but we also hypothesised that the latter vdW windows would also have a preference for the apo state. Failure to sample the correct apo state has been shown to lead to overestimation of the magnitude of the affinity [325,382]. Methods including 'confine-and-release' and using non-equilibrium simulations with apo equilibrium simulations have been proposed previously. However, the large numbers of

degrees of freedom makes confine-and-release unsuitable and I have been unable to reproduce the non-equilibrium results (as we do not generate overlapping work distributions).

Consequently, there is a niche for a method that uses the well validated equilibrium ABFE methodology. I therefore ran the final 11 windows (corresponding to $\lambda_{vdW} = 0.5$ to $\lambda_{vdW} = 1.0$ – the completely non-interacting window) from the apo protein conformations. I tracked how including this data for variable numbers of the terminal windows changed results. I saw the magnitude of $\Delta G^{\theta}_{binding}$ decrease towards the experimental as the number of apo windows increased. This indicates there is a difference between the apo and holo states, with the latter windows preferring the apo conformation. We saw good agreement between estimators and good overlap in the overlap matrix indicating we are not introducing large discrete energy gaps. This shows consideration of the apo state can be used to improve results. However, I am entirely biasing the states at which the windows are being run. Let us consider the λ value at which there is no preference for either state. The window should sample both states equally. However, we do not see transitions between states in our ABFE windows so I am biasing it away from its preferred distribution. We also do not know at which point to stop adding λ windows. Initially, adding apo windows is preferable as we are stabilising the apo state, but beyond a certain (unknown) point, we will begin destabilising the more interacting windows. This will also lead to a decrease in the magnitude of $\Delta G^{\theta}_{binding}$ but this is not justifiable.

4.10. Conclusion

We have completed an in-depth assessment of our ability to carry out ABFE calculations for an extremely challenging system. I showed that RBFEs and ABFEs can both rank the molecules. I provided a detailed description of our attempts to define the 'true' apo state. I have shown the importance of considering the true apo state when running ABFEs. This is vital for achieving an accurate absolute value. This is vital for hit validation, selectivity prediction, and binding site prediction. I therefore attempted to incorporate this knowledge of the apo state into our simulations. I saw the magnitude of the predicted free energy decreased as increasing number of windows were considered from the apo state. This showed a dependence on the particular 'true' apo state used. I note that the method used in this chapter may introduce some artefacts, and I will discuss potential solutions to this in Chapter 5.

5. The Apo State in ABFEs

5.1. Table of Contents

5. THE APO STATE IN ABFES 215

5.1. TABLE OF CONTENTS 216

5.2. FIGURES 217

5.3. INTRODUCTION 218

5.4. METHODS 222

5.5. GCGR 224

5.6. T4 AND $p38\alpha$ 228

5.7. DISCUSSION 237

5.8. CONCLUSION 241

5.2. Figures

FIGURE 5.1: SCHEMATIC OF OCCUPANCY DEPENDENCY 220

FIGURE 5.2: APO/HOLO ABFES WITH HREX 225

FIGURE 5.3: GCGR APO/HOLO ABFE OCCUPANCIES 227

FIGURE 5.4: T4 AND P38α STRUCTURES 230

FIGURE 5.5: T4 AND P38α APO/HOLO ABFES 234

FIGURE 5.6: GCGR SAR APO ABFE RESULTS 236

5.3. Introduction

The accurate calculation of ABFE calculations must include consideration of the apo state. Failure to generate the apo state will lead to an overprediction in the magnitude of the free energy. This will hinder our ability to use ABFEs to validate hits, predict binding sites, and predict selectivity. Inadequate sampling of the apo state has been shown to be a problem in multiple soluble systems where even simple conformational changes can cause serious errors [325,327,382]. I showed in Chapter 4 that we are unable to properly sample the apo state in our ABFE calculations in the GCGR system. I expect that this will be a general problem for simulations of GPCRs due to their flexibility and sensitivity to changes in the environment. Existing proposed solutions include confine-and-release [382] and using non-equilibrium ABFEs with holo and apo equilibrium simulations [325]. These are either inappropriate for the large number of degrees of freedom in a GPCR conformational change or need significant further validation. There is therefore a niche for a method based on equilibrium ABFEs that accounts for the apo transition.

In Chapter 4, I showed that starting the latter λ windows from the apo state can improve predictions. However, I identified 2 major concerns. First, I am biasing which state a simulation is in. This leads to a chance we will introduce discrete breaks in the phase space overlap. Simulations at a given λ window should be able to sample the states at a ratio according to their preference for that state. By this I mean the non-interacting window will sample almost exclusively the apo state whilst the fully interacting should sample almost exclusively the holo state. However, at intermediate λ windows it is not clear at which ratio of apo:holo states the window would prefer to sample. At a certain unknown λ value, there will be no preference between the states and the preferred sampling ratio would be 50:50.

Due to the short simulations in ABFE windows I would not be able to sample this. The second concern is around determining how many windows to start from the apo state. The number of apo windows included will directly influence the free energy prediction by changing what a given window samples. Inclusion of too few windows would lead to the apo state being underrepresented and the inclusion of too many would lead to the underrepresentation of the holo state.

I propose running windows starting with variable numbers of apo and holo windows with HREX could help solve both problems. First, we allow the windows to exchange meaning we can have a window which samples apo and holo conformations equally (and more generally windows can theoretically sample at their preferred ratio). We can track the ratio of time each window spends in a trajectory started from the apo or holo state. I will refer to this as the 'apo occupancy', that is the amount of time a window spends sampling the apo state of the protein. By looking at an individual ABFE calculation we can see how the apo occupancy depends on λ (Figure 5.1A). If we run multiple ABFE calculations which differ by the number of windows started from the apo state, we can see how the occupancy of a particular λ window varies as a function of the number of apo windows (Figure 5.1C) and how the apo occupancy as a function of λ varies as a function of the number of apo windows (Figure 5.1B).

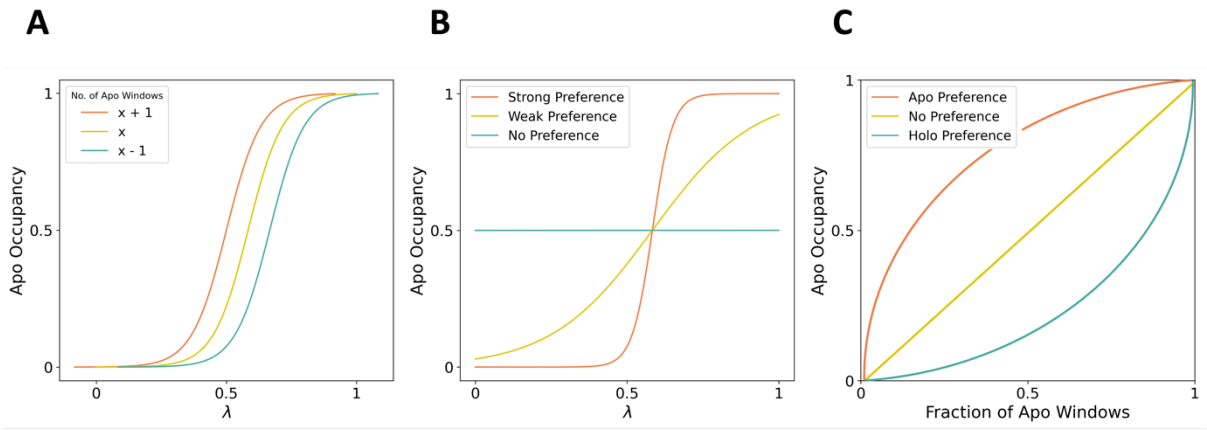


Figure 5.1: Schematic of Occupancy Dependency. (A) Apo occupancy against λ as the number of apo windows is increased (orange) or decreased (teal) from a reference (yellow). $\lambda = 0$ is holo and $\lambda = 1$ is apo. (B) How the occupancy changes as a function of λ for regions with no relative preference (teal), intermediate preference (yellow) and strong preference (orange). (C) How the apo occupancy changes as the fraction of apo windows is varied for a window with an apo preference (orange), no preference (yellow) and a holo preference (teal).

Let us consider adding the first apo window at the fully non-interacting window. If there is no difference between the preferences for the states then the window will diffuse across λ space evenly, changing the occupancies for all windows relative to the all holo simulation by the same amount ($\Delta_{occ} = \frac{1}{N_{windows}}$) . In this situation, as we add more windows, the occupancies for all windows will increase linearly and at the same rate. Now let us consider the situation where there is a strong preference for the apo state at the non-interacting window. For the final λ window $\Delta_{occ} \gg \frac{1}{N_{windows}}$, whilst for many of the more interacting windows $\Delta_{occ} \ll \frac{1}{N_{windows}}$. Upon addition of a second window, the penultimate window will see a large increase in occupancy. Repeating this for all of the possible ratios of apo:holo starting conformations we can see how the occupancy of each λ increases as a function of the ratio of starting conformations. Windows with a strong preference for the apo state will have an

initial gradient larger than $\frac{1}{N_{windows}}$ whilst those with a preference for the holo state will have a lesser gradient.

I therefore propose selecting the λ value at which most closely tracks the line $y = x$. This will provide something of an intermediate preference and therefore reduces the chance of spuriously forcing windows into the apo or holo state. However, I note that this is a relative preference where strictly speaking we are looking for an absolute preference. If we know the endpoints prefer different conformations then the relative preference will be linked to the absolute preference. This information can be gauged from crystallographic data where available. If it is not available, simulation can be used to find possible apo states and free energy calculations can help confirm these findings. In this discussion, I have also assumed the correct occupancy will be found by HREX. This may not be the case and it is important to check that we see suitable numbers of exchanges and convergence of occupancies across our simulations.

5.4. Methods

The apo and holo GCGR systems were prepared as described in Chapter 4.4. For the T4 system, the holo structure bound to p-xylene was used (PDB ID: 187L) and the apo structure was also used (PDB ID: 1LGU). The holo structure of $p38\alpha$ bound to 6-(2,4-difluorophenoxy)-8-methyl-2-(propan-2-ylamino)pyrido[6,5-d]pyrimidin-7-one was taken from PDB ID: 3FLY. The apo structure was taken from PDB ID: 1wfc. The proteins were capped with N-methyl and acetyl using PyMol [353]. The ligand was parameterised using OpenFF v2.0.0 (Sage) [341]. The box was solvated and sodium and chloride ions were added to neutralise the system and to create a concentration of 0.15 M. The water model used was TIP3P [346]. Equilibration simulations were run to equilibrate the system and to find appropriate values for the Boresch restraints using MDRestraintsGenerator [68,76]. Electrostatic simulations were handled using particle mesh Ewald method (PME) [347,348]. A cut-off of 1.2 nm was used. A cut-off of 1.2 nm was used for vdW interactions. The equilibration schemes began with a 10,000 step energy minimisation using a steepest descent algorithm. This was followed by 100 ps in the canonical ensemble with harmonic position restraints with a force constant of 1,000 kJ mol^{-1} nm^{-2} on the solute heavy atoms. The temperature was coupled to the reference temperature of 298.15 K using Langevin dynamics. It was then further equilibrated in the isothermal-isobaric (NPT) ensemble for 250 ps using the Berendsen weak coupling algorithm [306]. This was then followed by a further 500 ns using the Parinnello-Rahman barostat [350] before restraints were removed and the system was equilibrated for a final 10 ns.

ABFEs were run starting from either the holo state, the apo state, or with a mixture of the two conformations. The ratio of this mixture was varied from all holo to all apo. The production

runs were 20 ns for the complex and 15 ns for the ligand in solution. The equilibration schemes began with a 10,000 step energy minimisation using a steepest descent algorithm. This was followed by 100 ps in the canonical ensemble with harmonic position restraints with a force constant of 1,000 kJ mol^{-1} nm^{-2} on the solute heavy atoms. The temperature was coupled to the reference temperature of 298.15 K using Langevin dynamics. It was then further equilibrated in the isothermal-isobaric (NPT) ensemble for 100 ps using the Berendsen weak coupling algorithm [306]. This was then followed by a further 500 ns using the Parinnello-Rahman barostat [350] before restraints were removed and the system was equilibrated for a final 5 ns. Hamiltonian replica exchange (HREX) was utilised as implemented in Gromacs. 1,000,000 exchanges were attempted every 1,000 steps. In the 'restrained' ABFE simulations, the $N - C_\alpha - C_B - C_\gamma$ dihedral in V111 (T4) or T106 ($p38\alpha$) was restrained using a flat-bottomed restraint with a force constant of 41.8 kJ mol^{-1} nm^{-2}.

Trajectory analysis was carried out using MDAnalysis [351] and ProLiF [352]. Ligand structures were plotted using ChemDraw and RDKit 2023.03.1. Visualisation was carried out using PyMol [353]. Apo occupancy was calculated from information in the GROMACS log file.

5.5.GCGR

We began by looking at the GCGR. I ran ABFEs using HREX and a mixture of apo and holo states. I changed the number of windows starting from the apo and holo states (ranging from all holo to all apo inclusive). I ran increasing numbers of windows at higher λ values starting with the apo conformation. I track this progress via the lowest initial λ value which corresponds to the lowest window started from an apo state. A higher value indicates fewer windows from the apo state. The apo and holo windows should be able to interconvert depending on the preference for the particular conformation at a given λ. We see accepted transitions between states (Figure 5.2A). We again see the same trend whereby as more apo windows are introduced the magnitude of the leg decreases (Figure 5.2B). We see that conformation 1 also provides the largest difference. The magnitude of the difference in repeat 1 (5 $kcal/mol$) is similar to that seen without HREX.

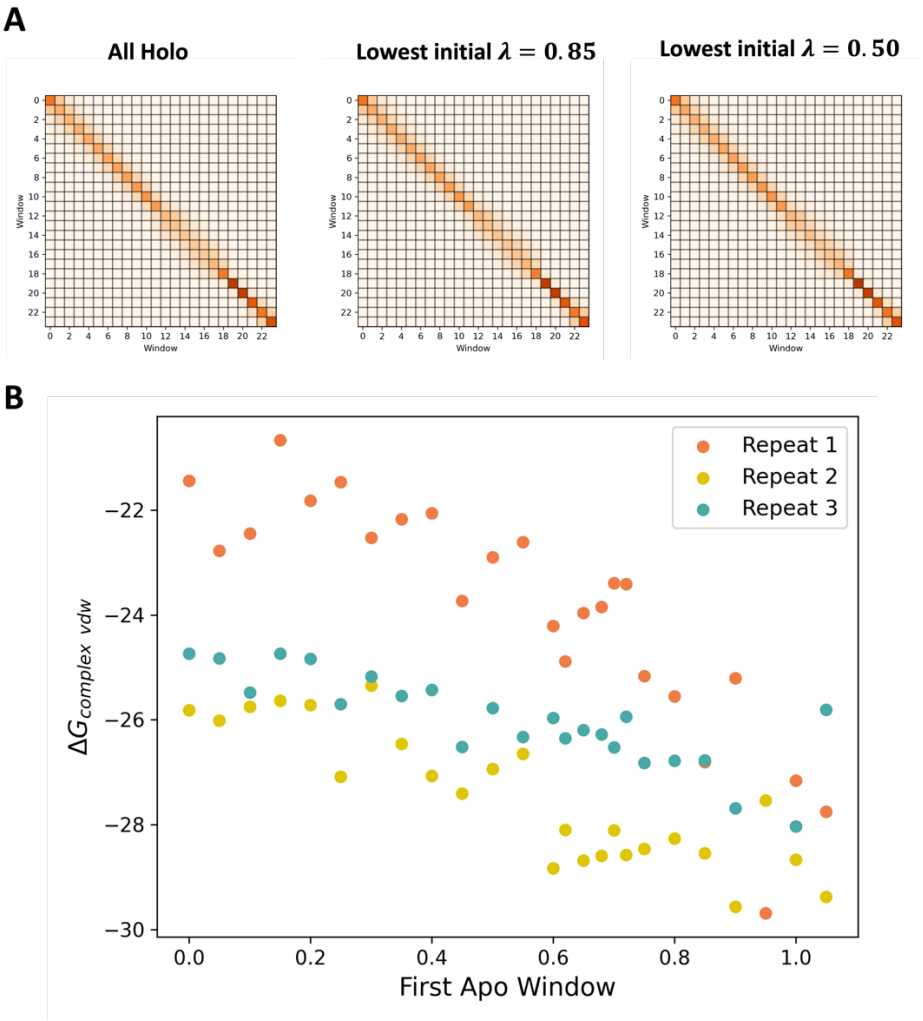


Figure 5.2: Apo/Holo ABFEs with HREX. (A) The HREX transition matrices for selected numbers of apo windows (B) The results of the $complex_{vdw}$ leg of ABFEs run with HREX using variable numbers of windows starting from apo and holo conformations. Each point is the result of a single ABFE calculation. The points at first apo window = 1.05 correspond to using all holo windows.

We plotted the occupancies of specific λ windows across these simulations (Figure 5.3A,B,C,D,E). As more holo windows are included, the curve left shifts. Figure 5.3B and Figure 5.3E show the inclusion of 1 holo state and 1 apo state respectively. Interestingly, we see a flatter curve in B than E. This indicates less preference for the holo state $\lambda = 0$ than for the apo state at $\lambda = 1$. I propose that the shape of the curve can help guide how many windows should be included from each state. Steeper gradients indicate greater differences in

preferences between states whilst shallower gradients indicate a more similar preference. Looking at the derivatives of the fit data, we see that the first few apo windows lead to significant gradients at the midpoint of occupancy, whilst the latter windows show a much smoother transition in preference (Figure 5.3F). For instance, the gradient when the initial λ window = 0.8 the gradient is 2.2, whilst when the initial λ window = 0.6 the gradient is 0.7. The gradients plateau at lower initial λ values, indicating a lesser preference for the λ windows to be in the holo state. The HREX simulations should provide a converged picture of the occupancy of windows. To assess this, I broke the simulations into 5 ns chunks and assessed whether they show similar occupancy profiles. Whilst in general the 5 ns chunks display similar trends, there is some difference between the values (Figure 5.3G & H) indicating the simulations are not fully converged.

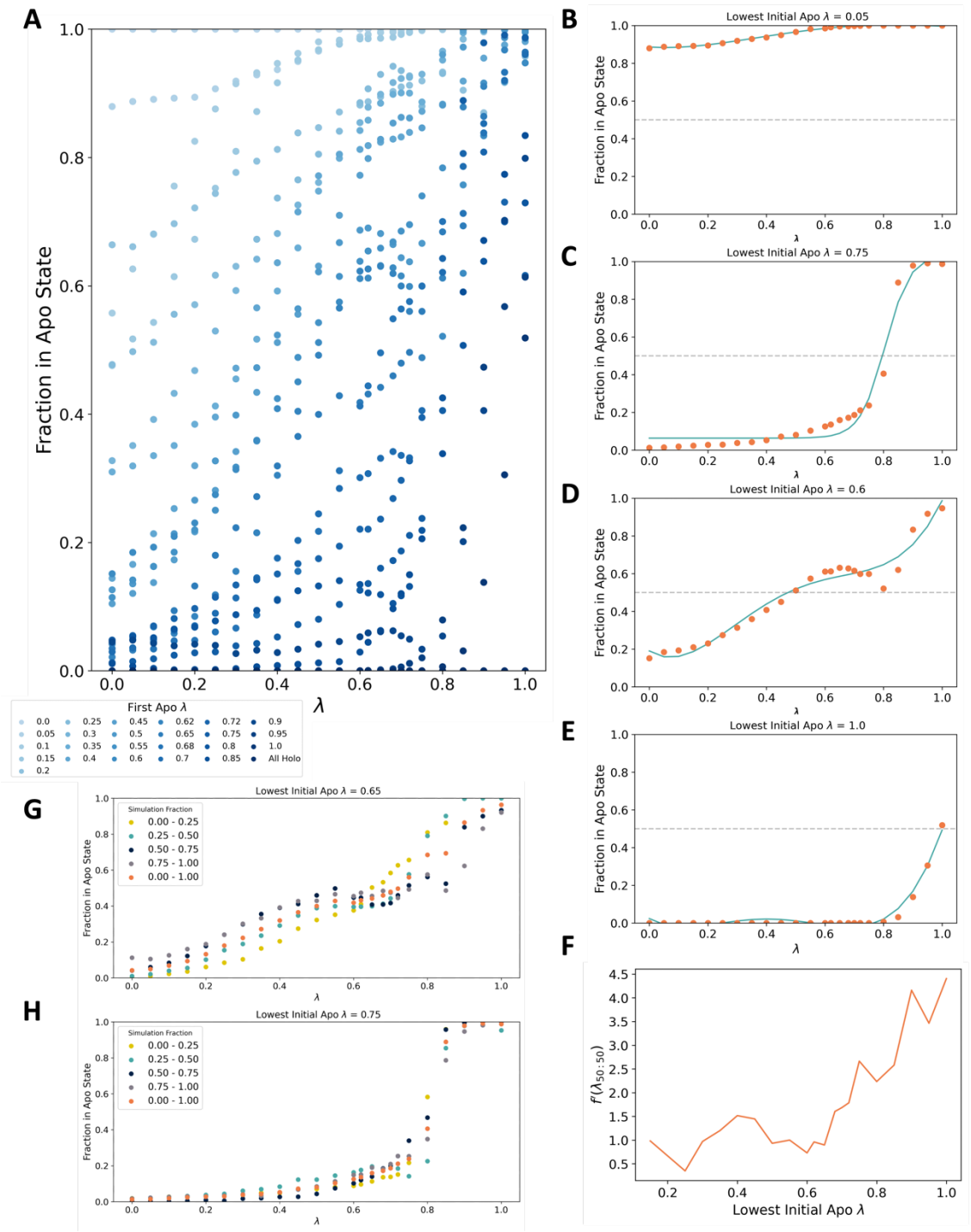


Figure 5.3: GCGR apo/holo ABFE occupancies. (A) The fraction of frames in an apo originating conformation at various λ values across different ratios of apo:holo starting conformations. (B) The occupancies of different windows in which the initial λ window is 0.05 (C) 0.75 (D), 0.60 (E), 1.0. (F) The gradients as the fraction in the apo state is 50:50 at various λ values (G) the occupancies across 5 ns chunks of simulation at $\lambda = 0.65$ *(H) and at* $\lambda = 0.75$.

5.6. T4 and $p38\alpha$

I decided that to better understand this problem, we needed to look at a simple system where there is less error between repeats (allowing us to better distinguish trends and noise) and where there is a clearer and simpler definition of the apo and holo states. I therefore looked at two systems in which a simple sidechain reorientation is known to cause issues – p38α and the T4 lysozyme (Figure 5.4).

Mobley *et al.* studied the impact of the reorientation of V111 in T4 lysozyme [85]. They showed ABFE calculations of *p-xylene* (Figure 5.5A) run from the holo state (PDB ID: 187l [383]) returned $-7.27 \pm 0.09\ kcal/mol$, whilst the apo state (PDB ID: 1lgu [384]) returned values of $-2.96 \pm 0.06\ kcal/mol$. The experimental value lies in between at $-4.67 \pm 0.06\ kcal/mol$. The only difference between these starting structures is the orientation of V111 [383]. Their studies show that this V111 does not interconvert in their simulations, but I should note their production simulations were only $1.11\ ns$ at each λ.

They propose the confine-and-release method for delaing with protein conformational change. This essentially involves calculating the energetic cost of confining the protein to the holo state, running an ABFE from the confined holo state, and then calculating the energetic cost of releasing the protein. Using umbrella sampling to calculate these values, they suggest the corrected value to be $-3.3 \pm 0.2\ kcal/mol$ and the values calculated from the apo and holo states to be statistically indistinguishable. However, they also state that a requirement for this method to be effective is that there are relatively few degrees of freedom to allow sufficient sampling of them. This means this approach is likely not appropriate for GPCRs

where there are a number of degrees of freedom and generating an accurate free energy landscape remains challenging [385].

I will also look at the $p38\alpha$ system. Work by Khalak *et al.* [325] showed that a single rotameric change in T106 from *trans* to *gauche-* between holo (PDB ID: 3fly [383]) and apo (PDBID: 1wfc [386]) states led to a change in prediction of 9 kJ/mol. They also showed a large loop motion observed in the crystals does not impact the binding affinity. In 5 repeats from the holo state without ligand they saw the V111 *trans* to *gauche-* transition stably occur within 200 ns on each occasion. It is therefore possible that we will see this transition in our ABFE windows. I looked at the binding of 6-(2,4-difluorophenoxy)-8-methyl-2-(propan-2-ylamino)pyrido[6,5-d]pyrimidin-7-one (the 3fly crystallographic ligand) (Figure 5.5B).

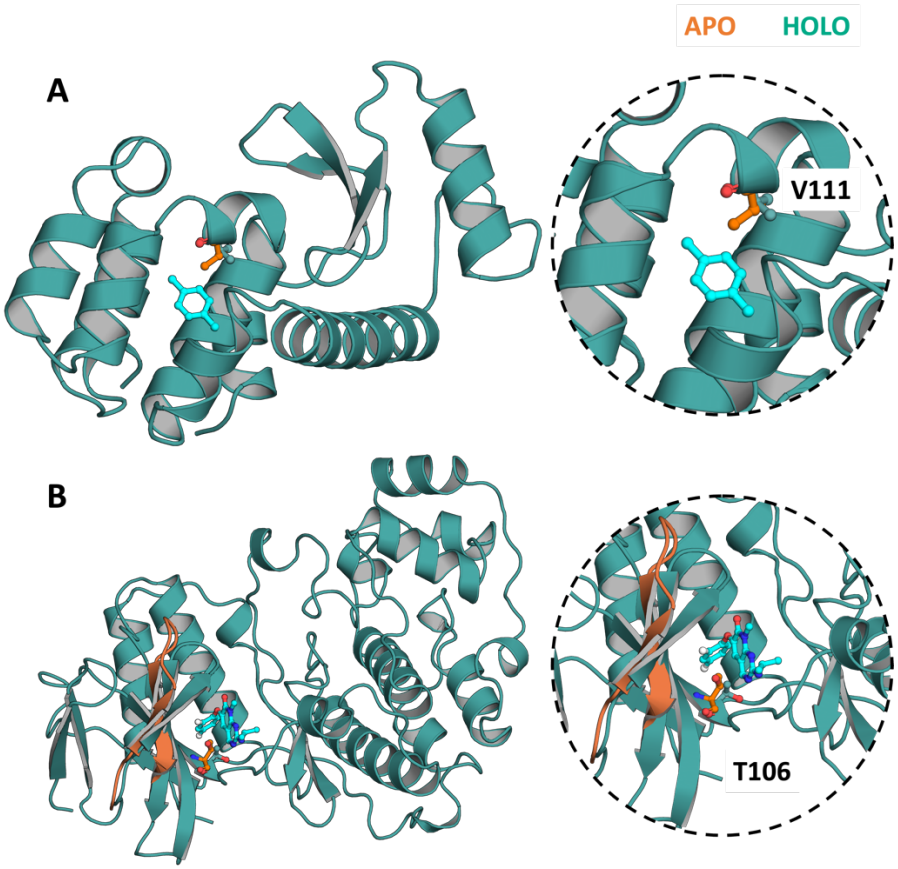


Figure 5.4: T4 and p38α structures. (A) The holo structure of T4 lysozyme (teal, PDB ID: 187L) bound to p-xylene (cyan sticks). The apo orientation of V111 is shown in orange sticks (PDB ID: 1LGU) (B) The holo structure of p38α *(teal, PDB ID: 3FLY) bound to 6-(2,4-difluorophenoxy)-8-methyl-2-(propan-2-ylamino)pyrido[6,5-d]pyrimidin-7-one (cyan sticks). The apo orientation of T106 and a loop rearrangement are shown in orange (PDB ID: 1wfc)*

Initital T4 simulations in the holo state returned values of $-4.82 \pm -0.05\ kcal/mol$. This in good agreement with the experimental value of $-4.67\ \pm\ 0.06\ kcal/mol$. The value from the apo state was -3.4 ± 0.25. I note this is less than the range observed by Mobley et al. $-7.27\ \pm\ 0.09\ kcal/mol$ and $-2.96\ \pm\ 0.06\ kcal/mol$ for holo and apo respectively. This could be a function of different protocols, forcefields, or our extended windows seeing apo/holo transitions on the timescale of the simulation. I suspected this could be a result of observed transitions in the V111 torsion so I ran ABFEs with the torsion restrained to the respective value with a flat-bottom restraint. We saw values of $-5.34\ \pm\ 0.09\ kcal/mol$ and $-3.0\ \pm\ 0.14\ kcal/mol$ for holo and apo respectively. This supports the idea that at least

some of the discrepancy due to poor sampling of the V111 can be resolved by increased simulation time.

However, I was interested to see if we could see an improvement in using the HREX method described previously. First, I ran simulations using all holo and all apo starting points. Interestingly, the holo simulation was significantly different from without HREX ($-4.19 \pm -0.06\ kcal/mol$) and the value from the apo state was -3.90 ± 0.14 I suspect this could be due to HREX facilitating more holo to apo transitions and *vice versa* which then propagate through λ space. I also ran ABFEs using various numbers of apo and holo starting configurations in the intermediate windows (Figure 5.5B). In these windows, we are perturbing both the vdW and the coulombic interactions. I therefore refer to $\lambda = 0$ as fully interacting, $\lambda = 1$ as just the vdW interactions turned on, and $\lambda = 2$ as fully non-interacting. There is a trend with the magnitude decreasing as we tend from all holo to all apo, but further comment is challenging due to the small difference observed between the two states and the standard deviation between repeats. We also note that we did not track the exchanges for these simulations due to time constraints. Due to the small energy differences we would expect exchanges to occur but it would be interesting to confirm this and check the frequency.

I therefore decided to look at the p38α system. Our ABFE calculations returned $-11.7 \pm 1.3\ kcal/mol$ and $-7.6 \pm 0.4\ kcal/mol$ for holo and apo respectively. I again ran repeats with varying number of intermediate simulations starting from apo and holo using HREX. The range between all holo and all apo was again less than observed when not using HREX ($-10.9 \pm 0.1\ kcal/mol$ and $-9.2 \pm 0.5\ kcal/mol$), indicating that again HREX is potentially facilitating more apo to holo / holo to apo transitions. We see a clearer trend that

as more apo windows are introduced, the magnitude of $\Delta G^{\theta}_{binding}$ decreases. We would expect this behaviour but the problem remains as to which λ do we decide as a cutoff.

I tracked the occupation of apo states for each window as the fraction of apo simulations increased for T4 (Figure 5.5D) and p38α (Figure 5.5E). For both plots, we see the high λ values that correspond to the non-interacting state (orange) are over-represented by apo conformations, whilst those at lower λ windows (yellow) are under-represented. This shows there is a relative preference difference between the interacting and non-interacting windows. By minimising the area between the curve and the line $y = x$, we can find the λ value that best corresponds to zero relative preference. In both cases this corresponds to $\lambda = 1.25$.

Taking the values from this λ, we get predicted values of ($-3.9 \pm 0.2\ kcal/mol$ and -10.5 ± 0.7), which remain in good agreement with the experimental ($-4.67 \pm 0.06\ kcal/mol$ and $-11.5\ kcal/mol$). I note these values do not significantly improve the results relative to experimental, but this prior agreement could have been due to different errors cancelling.

Looking at the individual plots for T4 (Figure 5.5F) and $p38\alpha$ (Figure 5.5G), we see smooth transitions at these values. Interestingly, there is still some occupancy of the lower λ values in T4. This indicates a weak preference for the holo state in the interacting windows. This fits with the small change in affinity as more apo windows are added and the tighter grouping of relative occupancies as a function of the fraction of apo windows around the line of $y = x$.

The T4 curve stays flat until $\lambda > 1$, whilst $p38\alpha$ begins increasing at $\lambda < 1$. This indicates the T4 windows at $\lambda < 1$ have the same state preferences, whilst the $p38\alpha$ do not. This is

interesting as this suggests there is a coulombic component to the p38α preferences whilst there is not for T4. This is supported by the proposed structural differences – whilst both are caused by changer in rotamer, T4 is caused by a non-polar side-chain rotating (V111) whilst $p38\alpha$ is caused by the rotation of a polar sidechain (T106). I note that several of the windows show significant discrepancies between occupancies as a function of the simulation. This is often the initial quarter of the windows so this affect could be lessened by improved equilibration. It would be interesting to start the simulations with the apo and holo simulations distributed randomly along the λ vector to further demonstrate the occupancies are not a function of the starting conformation. When observing the occupancies as a function of λ, we often see the endpoints exhibit slight discrepancies from what one might expect – that is the non-interacting window being less occupied by the apo state and the holo being more occupied by the apo state. I suspect this may be caused by the flux of exchanges being impacted due their end position in λ space.

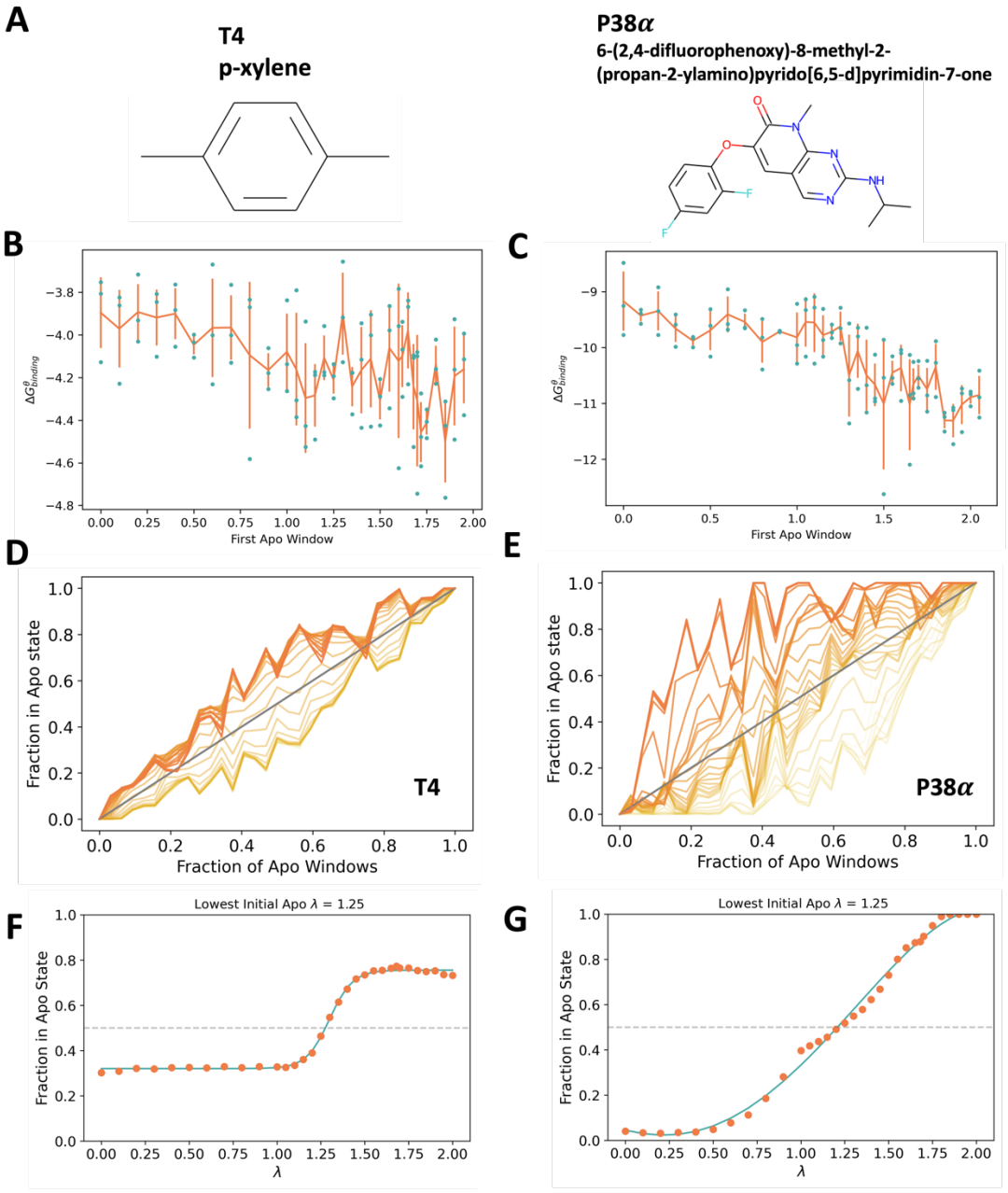


Figure 5.5: T4 and p38α apo/holo ABFEs. (A) Structure of the ligands used in the ABFE calculations (B)$\Delta G^{\theta}_{binding}$ as a function of the number of apo starting conformations used for the T4 system. Simulations were run with HREX to allow mixing of states. (C)$\Delta G^{\theta}_{binding}$ as a function of the number of apo starting conformations used for the p38α system. Simulations were run with HREX to allow mixing of states. (D) The apo occupancy of individual λ values as a function of varying the fraction of windows started from the apo conformation for the T4 system. As the ligand becomes more interacting, the colour transitions from orange (non-interacting) to yellow (interacting). (E) The apo occupancy of individual λ values as a function of varying the fraction of windows started from the apo conformation for the p38α system (F) The occupancies as a function of λ at the proposed lowest preference number of λ windows for the T4 system (G) The occupancies as a function of λ at the proposed lowest preference number of λ windows for the p38α system

Whilst more work is required to validate this methodology, I used it to retrospectively analyse our GCGR results. I looked at the HREX results with variable numbers of starting conformations and plotted the fraction of starting conformations dependence of specific λ values occupancy (Figure 5.6A). We see a strong preference for the apo conformation in the non-interacting windows (orange). The λ value which minimises the area between its line and $y = x$ is 0.5. I utilised the existing simulations from the apo windows for the SAR series and plotted the values from this, using the λ cutoff of 0.5 (Figure 5.6B). The calculations produce a very similar correlation coefficient to the all holo results (0.48 $[0.10,0.76]$ and 0.46 $[0.08,0.77]$). However, we see a significant improvement in the MAE (9.64 $[8.52,10.57]$ vs 12.36 $[11.42,13.25]$). As we would expect given the dependence on the particular apo starting conformation used to generate the apo state, we see large standard deviations.

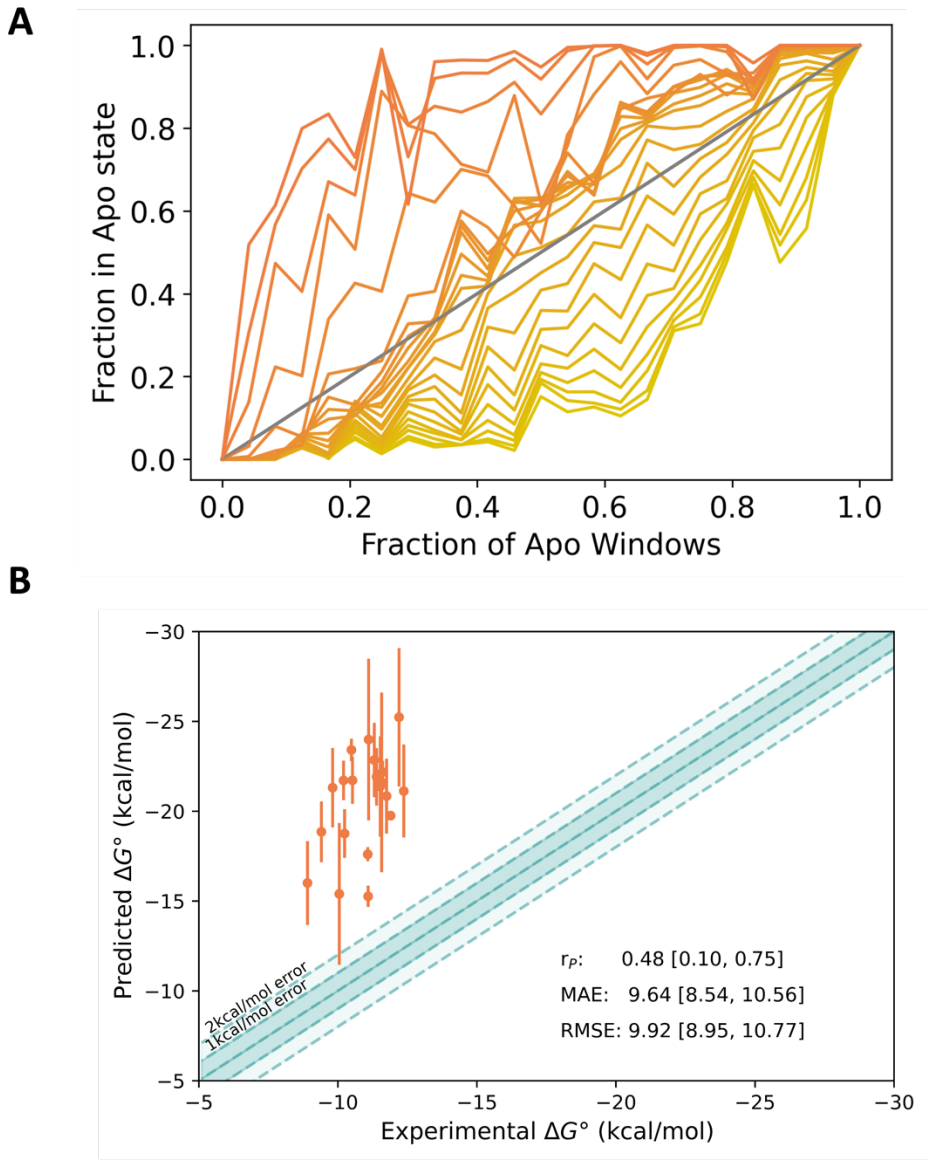


Figure 5.6: GCGR SAR Apo ABFE Results. (A) The apo occupancy of individual λ values as a function of varying the fraction of windows started from the apo conformation for the GCGR system. As the ligand becomes more interacting, the colour transitions from orange (non-interacting) to yellow (interacting). (B) The results of ABFEs with the last 11 windows run from the apo state. The error bars are the standard deviation of 3 repeats.

5.7. Discussion

In Chapter 4 I showed ABFE calculations for GCGR can be improved by directly considering the apo state in the latter windows. This methodology does not allow transitions between states in our ABFE windows, so I am biasing windows away from their preferred distribution. We also do not know at which point to stop adding λ windows. Initially, addition of apo windows is preferable as we are stabilising the apo state, leading to a decrease in the magnitude of $\Delta G^{\theta}_{binding}$. Beyond a certain (unknown) point, we will begin destabilising the more interacting windows. This will also lead to a decrease in the magnitude of $\Delta G^{\theta}_{binding}$ but this is not justifiable. I propose running windows starting with variable numbers of apo and holo windows with HREX could help solve both problems.

We applied this methodology to p-xylene binding to T4 [382] and 6-(2,4-difluorophenoxy)-8-methyl-2-(propan-2-ylamino)pyrido[6,5-d]pyrimidin-7-one binding to $p38\alpha$ [325]. Both systems are known to differ between apo and holo states by a sidechain conformational change – V111 and T106 respectively. I reproduced previous data that showed that both systems overpredict the magnitude of binding from the holo state and underpredict from the apo state. I showed that in the case of T4, longer simulations and HREX allow this discrepancy to be decreased due to improved sampling. Restraining the dihedral to the apo and holo conformations increased the discrepancies in both cases. I then ran ABFEs with variable numbers of apo windows and saw a slow transition from the apo to holo predicted values. However, the discrepancy between apo and holo across our simulations was only 0.3 $kcal/mol$ when run with HREX so analysis was challenging. In the case of p38α, we see a discrepancy of 2 $kcal/mol$ between the all holo and all apo HREX simulations ($-10.9 \pm$

0.1 $kcal/mol$ and -9.2 ± 0.5 $kcal/mol$). The predicted affinity decreases in magnitude as more λ windows are added. Using the proposed 'optimal' λ, we see -10.5 ± 0.7 $kcal/mol$ which is in good agreement with the experimental -11.5 $kcal/mol$. I used this concept of finding an optimum λ to revisit the exisiting GCGR SAR simulations. Using this process led to the MAE of simulations improving from 12.05 to 9.64. This indicates that the poor apo sampling is a contributor to the discrepancy but not the whole picture.

Strictly speaking, the occupancy should be determined by the absolute preference for a state. However, we are determining the relative preference for a state. One can imagine a situation whereby one has generated a conformation that does not represent the 'true' apo state. This incorrect state may be variably preferred by different λ windows. One should therefore have some belief *a priori* that the two states are meaningful. This information can be gauged from structural data where available. If it is not available, simulation can be used to find possible apo states. In the case of GCGR, we know the holo state is likely correct because of the crystal structure and the stability of simulations. We also believe the 'apo' state is meaningfully preferred in the absence of ligand from simulations. However, we only consider the vdW component of the λ vector, making our endpoint the ligand with coulombic interactions turned off.

I ran 3×500 ns simulations with the ligand's coulombic interactions turned off to see if this intermediate state tended towards the apo or holo state. It remained in the holo like state (data not shown) throughout, but this could be a function of the shorter simulations or the vdW interactions could be kinetically trapping the conformation in the holo state. It would be good to repeat this analysis using variable HREX windows across the full λ space and using HREX for the SAR series to generate better sampling of apo and holo states. I suspect the coulombic

stage is where we would start seeing a strong preference for the holo state due to the importance of charged interactions in this system. This is supported by the different shapes of the T4 and $p38\alpha$ λ vs apo occupancy curves. T4 is flat at λ values corresponding to coulombic perturbations whilst for p38α the higher coulombic λ values show a preference relative to the lower ones. This interesting as the T4 system is dependent on the rotation of a non-polar sidechain (V111) whilst $p38\alpha$ is dependent on a polar sidechain (T106).

When looking at occupancies across different 5 ns sections of the simulations, we see some discrepancies between sections. This could be improved by longer simulations or more windows to facilitate better mixing. I also noticed a quirk whereby the endpoints of the HEX simulations showed different occupancies than you would expect compared to their neighbours. For instance, the fully interacting window often has higher apo occupancy than its neighbours. I propose that this is a function of the flux of conformations leading to a build-up of the unfavourable state at the endpoint.

For the GCGR simulations, we see a strong dependence on the particular 'true' apo conformation used. It would be interesting to look further into what causes this discrepancy and what information this provides about how 'true' the apo states are. The motivation for this work was to improve the absolute values of the affinities of a similar SAR series. The fact all ligands have the same carboxylate group indicates they all stabilise the holo state in a similar way. It would be interesting to look at a set of molecules where they stabilise different states. One would expect the differences in the reorganisation energetics would make ranking the molecules challenging without explicitly considering the reorganisation. A key challenge in GPCR drug discovery is predicting selectivity for specific family members. Getting an

accurate absolute value is key for this but the reorganisation energy will likely differ between family members making selectivity determination challenging.

The reorganisation energy will also differ for different sites on the same protein. This makes comparison between scaffolds that bind at different sites difficult. For instance, I performed ABFE calculations on two SAR series on the FFAR receptor (data not shown). These both individually showed a consistent offset within the series and reasonable correlation coefficients, but comparisons could not be made due to the difference in discrepancies which we suspect were caused by different rearrangement energies. Another possible utility for ABFEs is binding site identification and different reorganisation energies will hinder this in all but the most stable binding sites.

5.8. Conclusion

Accounting for apo conformational transitions remains a challenging problem. I investigated the use of different numbers of apo and holo starting conformations with HREX. I demonstrated that this could produce mixing between the apo and holo conformations and showed this reproduced the transition in binding affinities. I then tried to address the major problem: what ratio of apo:holo conformations should be used? I attempted to address this by finding the λ window which showed the least preference between apo and holo states. I proposed a method for describing the window which shows the least preference for either state. I note that this is a relative preference, and not an absolute preference as is strictly required. Further work is needed to better assess the validity and rigour of this method. I applied the method to T4 and $p38\alpha$ and generated results in good agreement with experimental. For GCGR, I saw an improvement on the results but there is still an overestimation in $\Delta G^{\theta}_{binding}$. This indicates poor sampling of the apo conformation of the receptor is part of the picture but that there are other factors at play. In Chapter 6 I will discuss the importance of considering multiple lipid types as a contributing factor.

6. ABFEs in Complex Bilayers

6.1. Table of Contents

6. ABFES IN COMPLEX BILAYERS 243

6.1. TABLE OF CONTENTS 244

6.2. TABLE OF FIGURES 245

6.3. LIST OF TABLES 245

6.4. INTRODUCTION 246

6.5. METHODS 250

6.6. GCGR & PIP_2 253

6.7. $P2Y_1R$ AND CHOLESTEROL 257

6.7.1. Atomistic simulations of crystallographic cholesterols 259

6.7.2. Doped Cholesterol Bilayer: Coarse-grained Simulations 261

6.8. ABFES AND THE MICROKINETIC MODEL 265

6.9. DISCUSSION 271

6.10. CONCLUSION 278

11. SUPPLEMENTARY MATERIAL: CHAPTER 6 298

6.2. Table of Figures

FIGURE 6.1: GCGR PIP_2 STABILITY 254
FIGURE 6.2: POPC/PIP_2 RBFE CYCLE 256
FIGURE 6.3: CHOLESTEROL SAMPLING 258
FIGURE 6.4: BPTU ABFES WITH CHOLESTEROL 260
FIGURE 6.5: THE MICROKINETIC MODEL AND ALCHEMICAL CYCLES 266
FIGURE 6.6: SELECTED SAR STRUCTURES. 267
FIGURE 6.7: PARTITION SIMULATION ENDPOINTS 269
FIGURE 6.8: $P2Y_1R$ SAR MEMBRANE PARTITIONING RESULTS 270
SUPPLEMENTARY FIGURE 11.1: VDW20 $R346^{6.37B}$/PIP_2 CONTACTS 298
SUPPLEMENTARY FIGURE 11.2: RBFE ORIENTATIONAL DEPENDENCE 298
SUPPLEMENTARY FIGURE 11.3: PARTITION METADYNAMICS 299

6.3. List of Tables

TABLE 6.1: ABFES WITH CHOLESTEROL RESULTS 260

6.4. Introduction

Our previous calculations have made simplifying assumptions around the composition and structure of the bilayer. Of note is the simplification that the bilayer can be adequately modelled by a single species. In our case, I have been simulating in a pure POPC bilayer. I make this simplification to reduce the amount of simulation required to get reasonable sampling. However, this environment is significantly different from a realistic membrane. Membranes are a complex and dynamic mixture of gylcerophospholipids, sphingolipids, sterols, and membrane proteins. Varying the bilayer's lipid composition can significantly change the physicochemical properties of the bilayer [123]. The composition of bilayers varies across cell-type and for different membrane types. The exact composition can be determined by mass spectrometry [387]. The lipid composition is often linked to the function of that specific bilayer [119,121,122,124]. Lipids have a significant regulatory role for membrane proteins [129,130] and are increasingly present in structural data [131,132].

In the case of drug binding, specific lipid species may make a significant contribution to the binding free energy. They may occupy the binding site in the absence of ligand, interact directly with the ligand, or change the physicochemical properties of the bilayer. Indeed, the variable lipid compositions of different bilayers could provide a means to achieve more specific tissue or organelle targeting [388]. The role of lipids in regulating membrane protein behaviour leads to the possibility that small molecules can utilise the same regulatory sites as the regulatory lipids. This means one might expect a significant number of sites targeted by small molecules to have a lipid bound in the absence of ligand.

It is possible an apo structure may reveal the presence of a lipid bound at the same site. However, interpretation of this structural data needs to be done carefully as the identity and

orientation of lipids can be difficult to determine, they may be an artefact of the structure resolving process, they may bind in a conformation dependent manner [129,275], or they may not be present in the cell line used to generate the experimental data or in the particular membrane we wish to target. The absence of a lipid from the structure may not necessarily indicate the absence of binding there as it may not be resolved or it may have been stripped during the experimental protocol.

There are several papers that demonstrate the utility of coarse-grained MD to find lipid-exposed binding sites [134–136,389] and pipelines to identify ambiguous densities [390]. Work by Corey *et al.* [316] showed how coarse-grained free energy calculations could be used to further validate the existence of these proposed binding sites. The identification of direct lipid contacts with the ligand of interest is challenging. Atomistic simulations may struggle to reach the timescales required to adequately sample any interactions, whilst coarse-grained simulations may lack the resolution to adequately capture the subtle interactions. The improved representation of small molecules in Martini 3 may provide an interesting way to directly probe ligand/lipid interactions [244,245].

In Chapter 5, I discussed the importance of generating the correct description of the apo state for ABFE calculations. In the case of lipid-exposed binding sites, generating the correct apo state must also include finding all relevant protein/lipid interactions. I have shown how it is important to monitor the equilibration scheme to ensure lipids relax into the vacuum of the non-interacting ligand. This can be taken further to modify the restraint system to prevent water over-occupying the binding site in preference to lipid. It is important to find any key protein lipid interactions. In our pure POPC simulations, I assessed the quality of our sampling in the non-interacting window of our ABFEs by comparing it to long unbiased

simulations. In the case of simulations in complex bilayers, prohibitively long atomistic simulations may be required to define the apo state adequately. It is also highly unlikely that one would be able to routinely find the apo-like interactions in the current typical ABFE window timescales.

Another way in which lipids could impact binding is by altering the small molecule's affinity for the bilayer [234,235]. The microkinetic model suggests:

$$\Delta G_{observed} = \Delta G_{memb} + \Delta G_{intrinsic} \quad \textit{Equation 87}$$

Where ΔG_{memb} is the free energy of partitioning into the bilayer, and $\Delta G_{intrinisc}$ is the free energy of partitioning from the bilayer to the binding site (Figure 6.5A). $\Delta G_{intrinsic}$ is challenging to calculate experimentally [232,233] and simple models assuming partitioning is driven by bulk lipophillicity ignore the specific interactions small molecules form with specific lipids. Molecular simulations provide a means to calculate ΔG_{memb} and to probe its dependence on the membrane composition.

Umbrella sampling to calculate PMFs is the most common methodology for doing this. Typically, a simulation in which the small molecule is pulled through the bilayer is used to generate the starting conformations of the small molecule. This means the starting windows are not independent of one another, leading to hysteresis and an underestimation of the error [65]. I predict PMFs in a mixed bilayer will exacerbate this effect as the proximity of the small molecule to particular lipid types will show a dependence on the starting conformation. PMFs will often see the ΔG differing in water on either side of the membrane. Symmetrisation of the PMF profile is often used to correct for this. I am therefore interested in other methodologies that could generate a more accurate value for ΔG_{memb}.

To date, no ABFE calculations have assessed how well ABFEs perform in a complex bilayer. In this chapter, I will discuss our attempts to incorporate knowledge of multiple lipid species into an ABFE pipelines, the challenges I faced, and propose potential solutions. I will consider how to handle lipid occupying the binding site in the absence of ligand, direct protein lipid interactions, and the physicochemical properties of the bilayer changing as a result of the lipid composition. I will consider two of the most prominent regulatory lipids: cholesterol and phosphatidylinositol 4,5-bisphosphate (PIP_2). I will then also briefly discuss methodologies for determining membrane partitioning in the context of the microkinetic model.

6.5. Methods

Atomistic simulations were carried out in a similar manner to that previously described. Protein structures and ligand parameters were the same as used in Chapters 2 & 3. Cholesterol parameters were used from SLipids (2016) [344] whilst phosphatidylinositol (4,5) bisphosphate (PIP_2) was parameterised using OpenFF v2.0 (Sage) [341]. To prevent over aggregation of sodium ions to the negative chargers of PIP_2, exclusions were applied using Plumed [391]. Symmetric 4:1 POPC:cholesterol and asymmetric POPC:PIP_2 bilayers were construced (extracellular leaflet POPC, intracellular leaflet 90:10 POPC:PIP_2). The system was equilibrated for 50 ns post embedding with positions restraints (1,000 kJ mol^{-1} nm^{-2}) on the protein/ligand heavy atoms.

Coarse-grained simulations were carried out using MARTINI 2.2[392–394]. The structures were converted to coarse grained using the *martinize.py* script [395] and embedded into the bilayer using *insane.py* [396]. The ElNeDyn elastic network was used with a force constant of 500 kJ mol^{-1} nm^{-2} and a cutoff of 0.9 nm [397]. The protein was centred in the middle of a 15 x 15 x 17 nm box. For GCGR, extracellular leaflet 70:30 POPC:cholesterol was used and for the intracellular leaflet 60:30:10 POPC:PIP_2 was used to mimic the asymmetry of a plasma membrane. For $P2Y_1R$, 80:20 POPC:cholesterol was used for both leaflets. The latest *.itp files compatible with MARTINI 2.2 and the virtual sites cholesterol model were used [398]. The system was solvated using MARTINI water and 150 mM NaCl. Each repeat was setup and equilibrated individually. Simulations were run with a 20 fs timestep. Temperature was maintained at 323 K using the V-rescale thermostat [399] with a coupling constant $\tau_t = 1.0\ ps$. Pressure was maintained at 1 bar using the Parrinello-Rahman barostat ($\tau_p = 12.0\ ps$ and compressibility = 3×10^{-4} bar $^{-1}$). Coulombic interactions were handled with the reaction field method and a cutoff of 1.1 nm. vdW interactions used the potential-shift Verlet and a

cutoff of 1.1 nm. Bonds were constrained using the LINCS algorithm [400]. Protein/lipid interactions were assessed using PyLipID [389].

Coarse grained RBFEs were run using a similar protocol to Corey *et al.* [316]. The transformation involved the removal of beads PO1, PO2, RP1, RP2 and the conversion of bead RP3 from type SP1 to Q0 (changing the charge from 0 to +1). Charge neutrality of the system was maintained using the charge removal of 5 sodium ions. Coulombic and vdW interactions were annihilated separately. Both coulombic and vdW legs used 11 windows and a linear λ vector ([0.00 0.10 0.20 0.30 0.40 0.50 0.60 0.70 0.80 0.90 1.00).

Atomistic well-tempered metadynamics simulations were used to investigate the partitioning of $P2Y_1R$ SAR compunds into the bilayer. The 2 CVs biased were the z-distance from the centre of mass of the bilayer to the centre of mass of the ligand and the angle a pre-defined vector through the ligand makes with the z-axis. The CVs were biased using Plumed using a bias-factor of 10, a height of 0.8 kcal mol^{-1}, Gaussian widths of 0.1 nm and 0.3 radians, and a 1 ps deposition rate. Plotting and analysis was done using metadynminer (https://www.metadynamics.cz/metadynminer/).

Cholesterol free energy calculations were run using the alchemical cycle described by Salari *et al.* [315]. $Equation$ 88 describes gives an overview of the cycle. f,g, and b refer to the cholesterol 'free' in the bilayer, in the 'gas' phase, and 'bound' respectively. =, -, °, and ^ refer to different restraints schemes. Z is the partition function for a particular state. $Equation$ 88 is comprised of 5 terms. Terms 1 and 3 can be solved analytically. For more detail of how to calculate the analytical corrections, please see the Theory section.

$$\frac{Z_b}{Z_f} = \left(\frac{Z_f^{=}}{Z_f}\right)\left(\frac{Z_f^{-}Z_g^{=}}{Z_f^{=}}\right)\left(\frac{Z^{\circ}{}_g}{Z_g^{=}}\right)\left(\frac{Z^{\circ\hat{}}_g}{Z^{\circ}{}_g}\right)\left(\frac{Z_b^{\hat{}}}{Z_f Z_g^{\hat{}}}\right) \qquad \textit{Equation } 88$$

Term 2 corresponds to the FEP in the bulk membrane. Coulombic and vdW interactions were annihilated separately using linear λ vectors of 11 and 21 windows respectively. Restraints maintaining the cholesterols orientation relative to the bilayer normal and distance in z to the bilayer centre of mass were implemented using Plumed. I followed the implementation protocol in Corey *et al.* [316]. Please refer to this for more detailed discussion. Term 4 is the restraint FEP in the gas phase. The restraints are perturbed over 8 windows. Term 5 is the FEP of the lipid bound to the receptor. Term 5 uses position restraints to maintain the lipid's bound pose. Coulombic and vdW interactions were annihilated separately using linear λ vectors of 11 and 21 windows respectively. All simulated windows consisted of 20 ns of production simulation.

Free energy calculations to calculate ΔG_{memb} of 10 P2 SAR (1, 6a, 6b, 6d, 6j, 11a, 11c, 11g, 12, and 16 (BPTU)) molecules were run. Starting points were identified as the endpoints of 200 ns unbiased simulations starting from solution. The ligand was decoupled separately from the bilayer and solution. Both legs used linear λ vectors of 11 and 21 windows respectively for coulombic and vdW transformations. The ligand was restrained to the bilayer using Plumed. Volume restraints were selected to restrain the ligand to the lipid volume, meaning no correction was needed on the bilayer side of the cycle (as the ligand in solution is assumed to make a negligible contribution to the partition function). A standard state correction was applied to the ligand in solution.

6.6. GCGR & PIP_2

The binding site of MK-0893 sits outside of helices V, VI, VII, and VIII. The pocket formed by the residues on helices VI, VII, and VIII contain a number of positively charged residues ($R346^{6.37b}$ and K405). These charged residues are conserved across class B1 GPCRs. This, coupled with the importance of helix VI in activation, indicate that this could be a lipid regulatory site. In particular, PIP_2 seems a good candidate because of the large negative charge it carries. Indeed, coarse-grained simulations have suggested PIP_2 does bind at this site [274]. Further coarse-grained simulations coupled with native mass spec suggested binding is conformation dependent [276]. If PIP_2 does indeed occupy the binding site in preference to uncharged lipids, we must account for this in our ABFE calculations.

As the other studies used different conformations of GCGR, I began by confirming that we see PIP_2 binding for the conformation of GCGR in 5ee7. The other inactive GCGR crystal structures provide no evidence of PIP_2 binding at this site (PDB IDs: 5XF1[375], 5XEZ[375], and 4LR6[376], and 5YQZ[377]). I ran $5 \times 10\ \mu s$ simulations in an asymmetric bilayer. The upper leaflet consisted of 70:30 POPC:cholesterol and the lower leaflet consisted of 60:30:10 POPC:cholesterol:PIP_2. I observed PIP_2 binding at the MK-0893 binding site (Figure 6.1A). Analysis via PyLipID identified this site was occupied by PIP_2 for 96% of the frames (binding site 2) (Figure 6.1B). I verified this binding site by converting the system to atomistic and running $3 \times 200\ ns$ atomistic simulations. The presence of PIP_2 was measured by the presence of interactions between the 2 key charged residues - $R346^{6.37b}$ and K405 (Figure 6.1C).

Before running ABFE calculations, I wanted to assess whether we would be able to identify these interactions using atomistic simulations without prior knowledge of their existence. To

do this, I embedded the protein in an asymmetric atomistic bilayer (upper all POPC, lower 90:10 POPC:PIP_2). I ran $8 \times 500\ ns$ simulations of the protein in this bilayer (Figure 6.1D). 3 embedding processes were used to see if we saw a dependence on the interactions based on the embedding, which we did not. All 8 simulation do show interactions between the binding site residues and the protein. The interactions are stably maintained once they are made. I saw that one simulation found the binding site during the equilibration process, whilst the other 7 took between 225 and 475 ns to identify the binding site. This shows that whilst we can confirm a binding site and identify it during an atomistic simulation, we are unlikely to be able to find it in the non-interacting windows of an ABFE simulation.

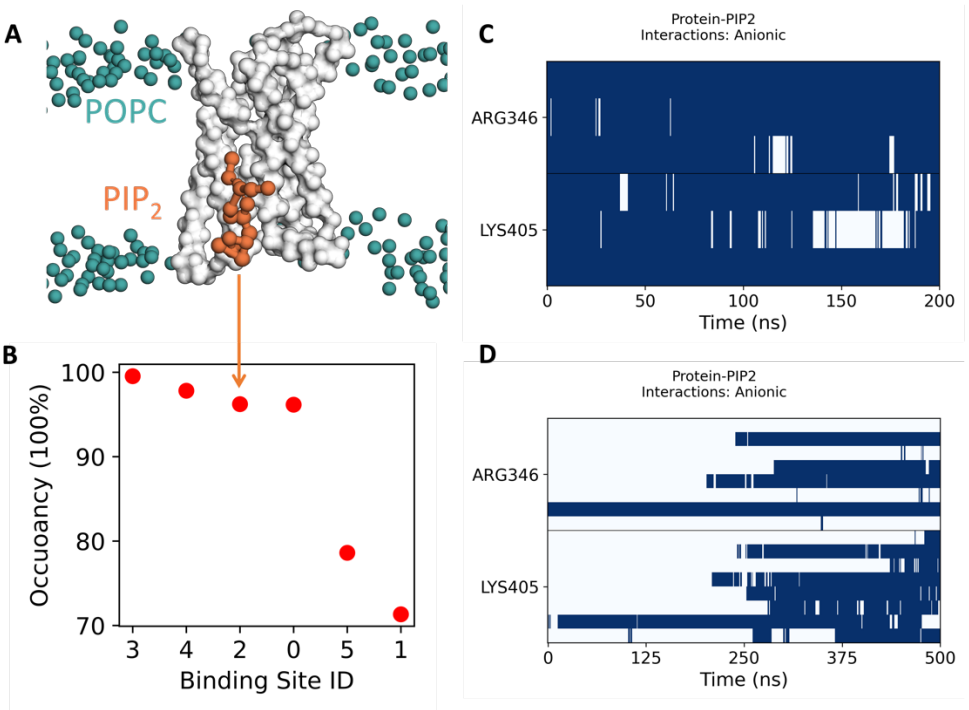


Figure 6.1: GCGR PIP_2 Stability. (A) A representative pose of PIP_2 (orange sticks) bound to GCGR between helices VI and VII. (B) The occupancy of PIP_2 at the various binding sites, as calculated by PyLipID [389]*. Binding site 2 corresponds to the site between helices VI and VII. (C) The presence of ionic interactions between PIP_2 and the key positively charged binding site residues (* $R346^{6.37b}$ *and K405) during simulations started from an atomistic representation of the coarse-grain identified binding pose. The presence of interactions was tracked using ProLiF* [352]*. (D) The presence of PIP_2 and charged binding site residues across* 8×500 *ns simulations started from embedding into a complex bilayer.*

We wanted to assess the performance of ABFEs in a PIP_2 doped bilayer. I suspected finding specific interactions would be challenging and could lead to increased deviation between repeats. The presence of a large numbers of negative charges could also have impacted the calculations. I tested this hypothesis by running 3 ABFE calculations of MK-0893 in a doped PIP_2 bilayer. The results ($-24.8 \pm 1.7\ kcal/mol$) were not significantly different from that observed in a pure POPC bilayer ($-24.0 \pm 1.3\ kcal/mol$). We saw no contacts between PIP_2 and the key residues in the non-interacting window (Supplementary Figure 10.1). This indicates that despite it being an intuitive way to include multiple lipid species, running ABFEs in a doped bilayer would require simulations far beyond currently accessible timescales.

We therefore propose an extension to the alchemical cycle where the ABFE is run in a pure POPC bilayer before an additional stage accounting for the PIP_2 binding. This takes the form of an RBFE calculation mutating PIP_2 to POPC (Figure 6.2). I use a methodology based upon that reported by Corey *et al.* [316]. This methodology has several advantages. As $\Delta\Delta G_{bind\ (PIP_2-POPC)}$ is independent of the ligand, a single calculation can be run and applied to all of the SAR series. Whilst atomistic simulations could be used they would struggle to sample the flexible motions of the lipid tails. RBFEs would also transforming the ring structures of the headgroups which would make the calculations significantly more challenging. Coarse-grained simulations however will be unable to capture the subtle interactions in the binding site so the results should not be over interpreted.

I ran 5 repeats of the RBFE calculations and obtained $\Delta\Delta G_{bind\ (PIP_2-POPC)} = 4.9 \pm 0.2\ kcal/mol$. This further validates that this is a PIP_2 binding site. The MK-0893 SAR

ABFE results had a MAE of 9.64 $kcal/mol$ after using the apo conformation for the latter ABFE windows. Applying this correction would bring the MAE down to 4.74 $kcal/mol$, a significant improvement on the original MAE of 12.05 $kcal/mol$.

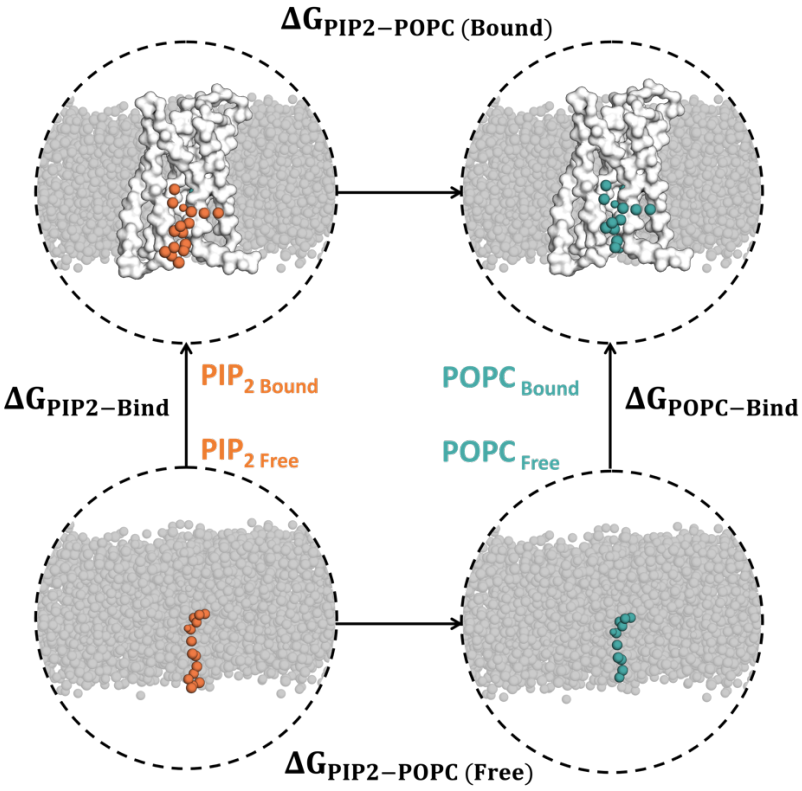


Figure 6.2: POPC/PIP_2 RBFE Cycle. $\Delta\Delta G_{bind\ (PIP_2-POPC)}$ can be calculated by transforming PIP_2 (orange) to POPC (teal) both in complex with the protein and free in the bilayer.

6.7. $P2Y_1R$ and Cholesterol

Cholesterol is a fundamental component of mammalian membranes. This work was motivated by the observation that different inhibition constants are reported for BPTU binding to $P2Y_1R$. Chao *et al.* reported a value of $K_i = 6 \pm 0.6\ nM$ [331] whilst Zhang *et al.* reported a value of $K_i = 173 \pm 45\ nM$ [330]. The reported methodologies are highly similar with the major difference being the cell line used to create the membrane preparations – HEK-293 and Sf9 respectively. Interestingly, HEK-293 does not contain any cholesterol, whilst Sf9 cells do [401]. The idea that cholesterol may influence binding is supported by the possibility of aromatic interactions between cholesterol and BPTU or the possibility of the hydroxyl of cholesterol forming interactions with the backbone carbonyl of $L102^{2.55}$ in the absence of ligand.

The crystallographic structure of BPTU bound to $P2Y_1R$ contains 1 cholesterol and 3 cholesteryl hemisuccinates (CHS) (Figure 6.3A). The CHS at BS3 is proximal to the ligand. There is no further evidence for the role of cholesterol in binding, but this would likely be the case in a prospective drug discovery campaign. This system therefore provides a useful case to investigate the potential roles of complex bilayers in free energy calculations and how to identify systematic challenges or errors. Here I assess the ability of atomistic and coarse-grained molecular dynamics simulations to identify and sample key lipid interactions in a doped bilayer. I then discuss methods to run atomistic free energy calculations in a doped bilayer.

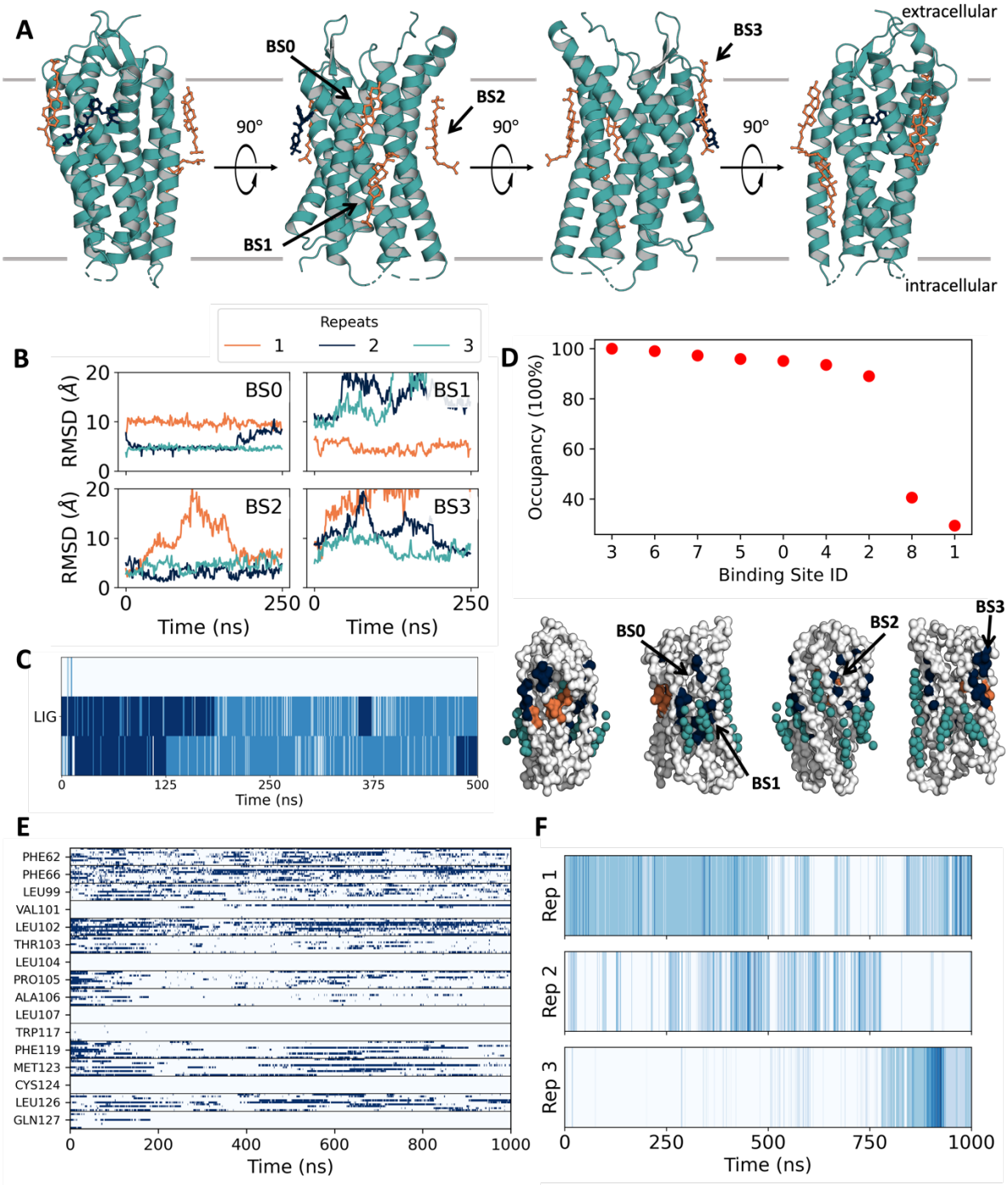


Figure 6.3: Cholesterol Sampling. (A) The cholesterol (BS0) and cholesteryl hemisuccinates (BS1, BS2, and BS3) resolved in the structure of BPTU bound to P2Y$_1$R (PDB ID: 4XNV) (B) The RMSD of cholesterols occupying the crystallographic binding sites. (C) The interactions of the cholesterol at BS3 with BPTU (D) The binding sites identified from coarse-grained simulations by PyLipID. Representative poses for high occupancy cholesterols (teal) do not overlap with the BPTU binding site (orange) or crystallographic sites (BS0, BS2, and BS3, shown in dark blue). PyLipID site 6 overlaps with crystallographic BS1. (E) The interactions between cholesterol and the binding site residues from 9 × 1 μs apo simulations. (F) The interactions between cholesterol and BPTU from 3 × 1 μs simulations.

6.7.1. Atomistic simulations of crystallographic cholesterols

Lipid-like densities are commonly found around the transmembrane region in structures of membrane proteins [402,403] but identification of the particular lipid species remains challenging [390]. Identification of the correct orientation of cholesterol can also lead to errors. Studies have also suggested that although they are chemically similar, the less hydrophobic nature of CHS may lead to different binding poses relative to cholesterol [390].

I therefore wanted to test the stability of the bound cholesterols to assess whether they are appropriate to be included in free energy calculations. To test this, I ran $3 \times 250\ ns$ simulations of the BPTU/$P2Y_1R$ system with the crystallographic cholesterols. The 3 CHS molecules were replaced with cholesterols to mimic a more realistic membrane environment. The cholesterols are generally unstable (Figure 6.3B). We see multiple unbinding events over the course of the simulations. This could be explained by the different properties of CHS. For instance, BS2 is occupied by a CHS which interacts with R301 via its carboxylate group. BS0, the only site in which a cholesterol is modelled, is stable in 2/3 simulations. Interestingly, despite unbinding in repeat 1, the cholesterol in BS3 interacts with the ligand over the course of repeats 2 and 3 (Figure 6.3C). I therefore decided to run 3 ABFEs including the crystallographic cholesterols (Figure 6.4A & Table 6.1). The results of these simulations ($-11.9 \pm 0.9\ kcal/mol$) are not significantly different from previous simulations without the presence of the crystallographic cholesterols.

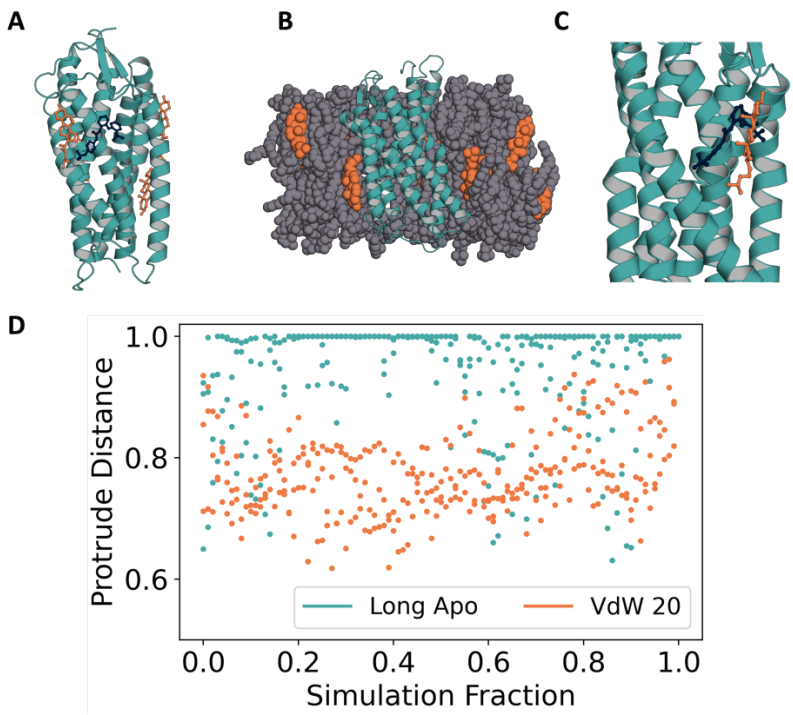


Figure 6.4: BPTU ABFEs with cholesterol. ABFEs of BPTU (dark blue) bound to $P2Y_1R$ (teal cartoon) were run in a number of different lipid environments (A) with the crystal cholesterols (orange) (B) a 4:1 POPC:cholesterol bilayer and (C) from a snapshot in which a cholesterol is in contact with BPTU (D) the protrude distance of cholesterol and the ligand for $3 \times 1\ \mu s$ *cholesterol repeats (teal) and the* $3 \times 20\ ns$ *windows from the non-interacting windows from ABFEs run from (C)*

	ΔG° (kcal/mol)
Experimental 1 [331]	$\mathbf{-11.7 \pm 0.2}$
Experimental 2 [330]	$\mathbf{-9.6 \pm 0.1}$
POPC	$\mathbf{-11.6 \pm 1.5}$
Crystallographic Cholesterol	$\mathbf{-11.9 \pm 0.9}$
POPC:Cholesterol (4:1) Equil 50 ns	$\mathbf{-12.6 \pm 1.4}$ -11.1 ± 0.4 -12.5 ± 0.3 -14.2 ± 0.7
POPC:Cholesterol (4:1) Equil1000 ns	$\mathbf{-12.2 \pm 1.3}$ -13.9 ± 0.3 -10.9 ± 0.1 -11.6 ± 0.4
POPC:Cholesterol (4:1) Selected	**-14.1 ± 0.4**

Table 6.1: ABFEs with cholesterol results. The results of ABFE calculations run with cholesterol via different methods. Bold values are the values for all repeats, non-bold text are the values for individual starting configurations.

6.7.2. Doped Cholesterol Bilayer: Coarse-grained Simulations

Another way the binding free energy could be impacted by cholesterol is if it occupies the ligand binding site in the absence of ligand. It would therefore need to be displaced by the ligand, reducing the observed binding affinity. To identify if cholesterol occupies the binding site in the absence of ligand, I ran $10 \times 10\ \mu s$ coarse-grained simulations of $P2Y_1R$ in a doped 4:1 POPC:cholesterol bilayer. PyLipID [389] was used to identify binding sites (Figure 6.3D).

7 binding sites show high cholesterol occupancy and these are shown in teal. BS0, BS2, and BS4 (coloured dark blue) do not overlap well with any of the predicted high occupancy sites by PyLipID. PyLipID site 6 does overlap well with BS1. The poor identification of crystallographic cholesterols is interesting, but it does agree with the poor stability observed in atomistic simulations.

We do not identify any cholesterol binding at the ligand binding site. To further validate this hypothesis, I ran $9 \times 1\ \mu s$ atomistic simulations in a 4:1 POPC:cholesterol doped bilayer and tracked the interactions between the binding site residues and cholesterol (Figure 6.3E). Here we see a number of interactions being formed but these tend to be relatively short lived. This supports the coarse-grained data and suggests cholesterol does not bind at this site. This indicates that if cholesterol is influencing binding, it is doing so via another route.

Cholesterol could influence the binding indirectly via the changes it induces in the structure of the membrane. To assess this, I embedded the protein and ligand in a 4:1 POPC:cholesterol bilayer (Figure 6.4B). Separate embedding process were used to ensure different cholesterol positions. 3 ABFE calculations were run from each of the 3 starting configurations, resulting

in 9 total repeats. The overall results ($-12.6 \pm 1.4\ kcal/mol$) are similar to the calculations in a pure POPC bilayer.

The repeats from configuration 3, however, represent the lowest ΔG^{θ} ($-14.2 \pm 0.7\ kcal/mol$). This could be a function of favourable positioning of cholesterols generated by the embedding process so I decided to equilibrate the bilayers further to see if we could get more consistency across configurations. To do this I ran $3 \times 1\ \mu s$ atomistic simulations in a 4:1 POPC:cholesterol bilayer (Figure 6.3F). We see multiple interactions with BPTU and cholesterol. However, these interactions are not consistently present and they are formed with different parts of the ligand. For each of the endpoints of these simulations I ran 3 ABFE calculations, resulting in 9 repeats in total. We see a similar pattern to the unequilibrated simulations – the general result ($-12.2 \pm 1.3\ kcal/mol$) is similar to the calculations carried out in a pure POPC bilayer. The endpoints of the $1\ \mu s$ simulations, however, continue to show differences between them in terms of the presence of cholesterol and BPTU interactions. I therefore decided to pick a frame that represented the most consistent BPTU and cholesterol interactions across our simulations. This was a representative frame (Figure 6.4C) from the first 500 ns of repeat 1 (Figure 6.3F). The 3 ABFEs run from this configuration returned results of $-14.1 \pm 0.4\ kcal/mol$. This suggests there could be some significant changes from considering a pure BPTU bilayer.

However, it is possible this is an artefact rather than a genuine result. Figure 6.4E shows the protrude distance for BPTU and cholesterol across $3 \times 1\ \mu s$ simulations and the non-interacting windows from these ABFEs. The long simulations indicate that an absence of overlap is favourable, and any overlap is transient. The ABFE windows, however, show a consistent overlap (Figure 6.4D). This could artificially increase the apparent magnitude of

$\Delta G^{\theta}_{binding}$ by making the apo state unduly unfavourable. This is likely due to be a general issue with ABFE simulations where a lipid is suspected of stabilising the ligand binding.

As with GCGR, I propose an additional stage to the cycle. This would calculate $\Delta G^{\theta}_{binding}$ of cholesterol to the binding site defined by the protein and the ligand. To the best of our knowledge, ABFE calculations where a small molecule defines part of the binding site have not been carried out previously. I am also interested in calculating the free energy of binding from the bilayer to the binding site. This cannot be done using Boresch restraints [68] as the analytical correction assumes the system is isotropic. I instead use restraints as defined by Salari *et al.* [315] and applied in Corey *et al.* [316] (Figure 6.5C). For discussion of these restraints, please see the Theory Section (Section 2.5) or the Methods section. The calculation essentially consists of 5 steps: 2 analytical corrections (terms 1 & 3), perturbing the ligand both in complex with the protein and free in the bilayer (terms 5 and 2), and restraining it in the gas phase (term 4).

Applying this approach yields a cholesterol $\Delta G_{binding}$ value of $+4.1 \pm 0.7\ kcal/mol$. This suggests the binding is unfavourable, which is to be expected given the transient nature of the interaction in the simulation. The restraints here consider the non-isotropic nature of the water/bilayer system. When calculating the analytical correction for restraining cholesterol in the bilayer, it is possible to consider the penalty of restraining it from the volume of the box or the volume of the bilayer. I consider the volume of the bilayer, but the difference is on the order of $0.4\ kcal/mol$ which is quite small given the other errors observed in the calculations. Using the bilayer's volume assumes the bilayer is isotropic which is also not strictly correct. The scheme described here is meant to be descriptive of a potential route to

assess the stability of observed ligand/lipid interactions. Further optimisation and validation against known experimental values is needed.

Thus, for $P2Y_1R$, ABFE calculations suggest that cholesterol has little, if any, influence on $\Delta G^{\theta}_{binding}$ of BPTU and the related SAR compounds.

6.8. ABFEs and the Microkinetic Model

The microkinetic model states that $\Delta G_{observed} = \Delta G_{memb} + \Delta G_{intrinsic}$. This is relevant for drugs binding at lipid-exposed binding sites as it suggests the observed affinity can be increased by increasing the lipophilicity of the molecule [218]. Increasing the intrinsic affinity is important as highly lipophilic molecules will likely show less specificity and therefore potentially greater side effects [221]. There is therefore an unmet need to characterise the role of ligand/membrane energetic in observed SAR [220]. MD based techniques are well placed to play a role in disentangling the SAR due to the difficulties of generating experimental data [233] and because MD naturally considers the complex array of ligand/membrane interactions. Any computational methods should be flexible enough to consider the role of different membrane compositions [234,235]. The restraints used for cholesterol above could facilitate these calculations. In this section, I will discuss using alchemical calculations to calculate ΔG_{memb} for selected molecules of the $P2Y_1R$ SAR.

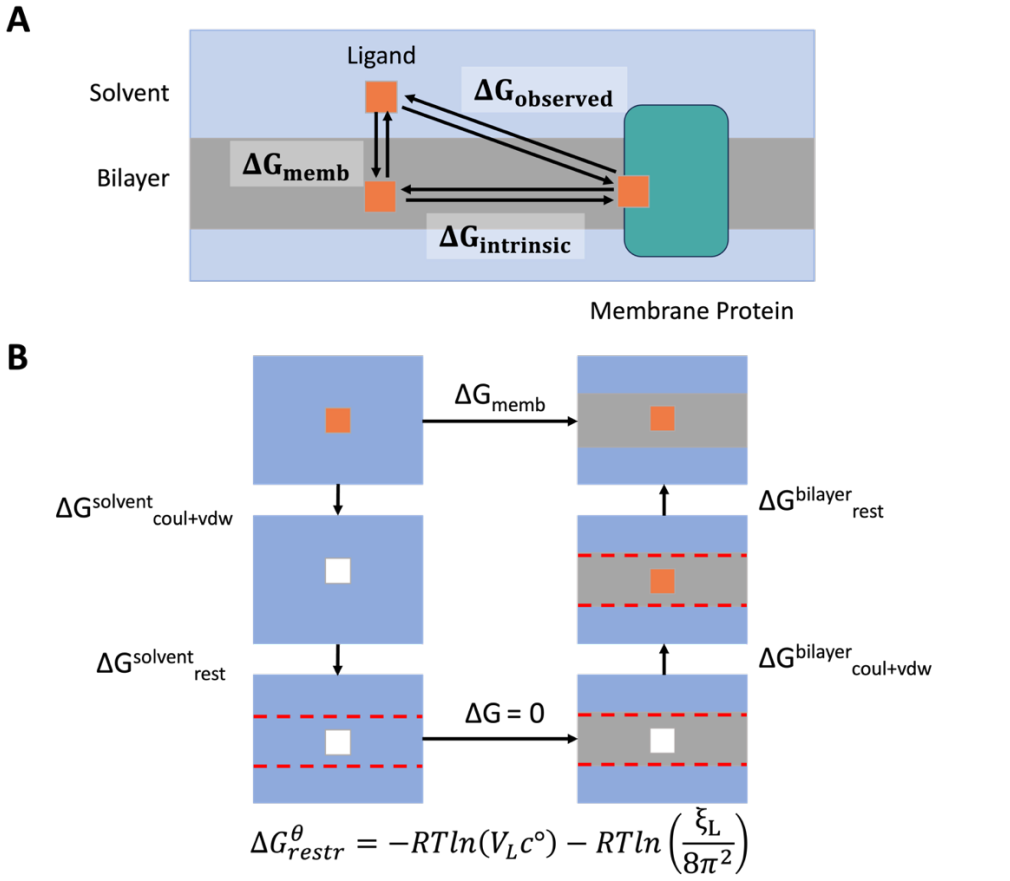


Figure 6.5: The microkinetic model and Alchemical Cycles. (A)The microkinetic model. The observed $\Delta G_{observed}$ can be broken into the affinity for the bilayer relative to solution (ΔG_{memb}) and the affinity from the bilayer to the bound pose ($\Delta G_{intrinsic}$) (B) The alchemical cycle for membrane partitioning ABFEs. The ligand is decoupled from solution and the bilayer. Analytical corrections for the restraints are applied. For more details, please see the Methods section of this chapter. (C) The alchemical cycle for calculating $\Delta G_{intrinsic}$. Based on Corey et al. [316]. The ligand in the bilayer (f) is restrained relative to the centre of mass of the protein and the z-axis. It is then decoupled from the bilayer. Whilst in the gas phase the restraints are adapted to a volume-based restraint and then a position based restraint. The ligand is then placed in the binding site before the interactions are turned back on.

We selected 10 SAR molecules that provide a diverse array of the ligand structures and $\Delta G_{observed}$ values. These were: 1, 6a, 6b, 6d, 6j, 11a, 11c, 11g, 12, and 16 (BPTU) (Figure 6.6).

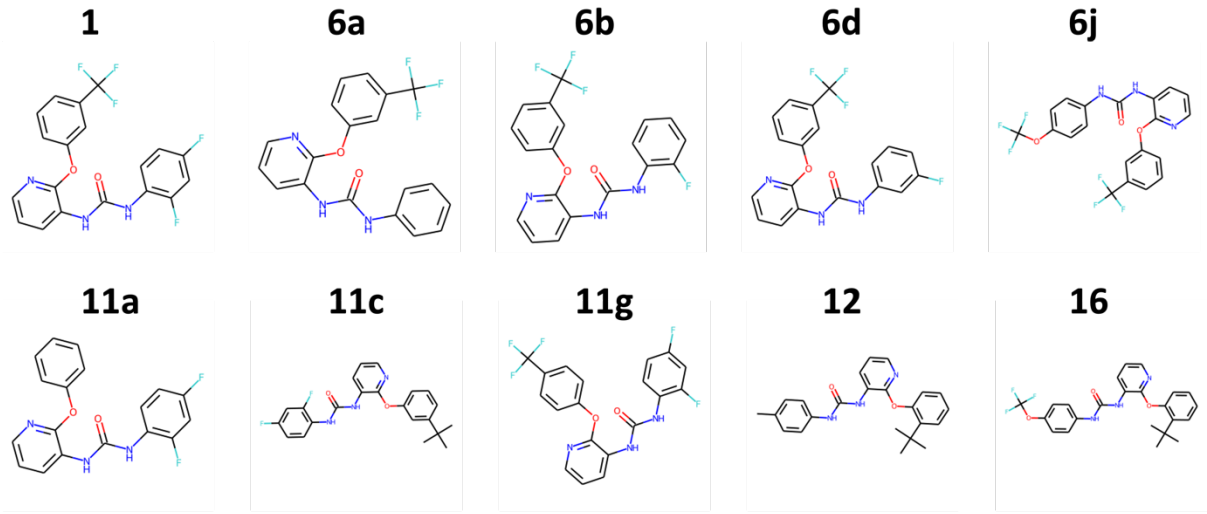


Figure 6.6: Selected SAR Structures.

For calculating both ΔG_{memb} and ΔG_{calc} it is important to generate realistic starting systems for our calculations, both in terms of bilayer depth (z) and the orientation of the molecule. I define a vector through the scaffold of the SAR molecules and track the angle this makes relative to the bilayer normal to track the orientation (θ) (Figure 6.7A). For other ligands, particularly flexible ones, more degrees of freedom may need to be considered. One of the advantages of ABFEs for comparing $\Delta G_{intrinisc}/\Delta G_{memb}$ of different SAR is that it allows different molecules to find their preferred orientation in the bilayer. RBFEs on the other hand make the assumption that they occupy the same preferred orientation [258]. This will likely not hold when a charged or polar group is introduced to replace a non-polar group. I demonstrated that RBFEs show an orientation dependence (Supplementary Figure 10.2).

To generate realistic conformations, I used unbiased MD simulations. Selecting an appropriate simulation length involves balancing computational expediency and ensuring enough time is given for potentially large energy barriers to be overcome. To assess the reproducibility of this, I ran 3 repeats and tracked z and θ (Figure 6.7A & C). The ligands are broadly in agreement with one another, which is likely due to their highly similar scaffolds. The z values are typically between $0.8\ nm$ and $1.5\ nm$. This is in agreement with the minima seen in the PMF of BPTU from Yuan *et al.* [332]. This area corresponds to the interface region between the hydrophobic core and the headgroups, potentially suggesting the ligand could be orientating to form polar interactions with the headgroups and bury the hydrophobic groups. θ shows a slight preference for the angle to lie between 45° and 90° but there is more variability than with z.

The simulation endpoints are in agreement with $500\ ns$ metadynamics simulations using z and θ as CVs (Supplementary Figure 10.3). These metadynamics simulations indicate a strong z dependence with a preference for the ligand centre of mass to be between 0.5 and 1.5 nm from the bilayer centre but no clear dependence on θ. I note that our metadynamics simulations were not converged, as evidenced by the discrepancy in the free energy between leaflets. Some of the free energy surfaces indicate an orientational preference whilst others do not. This could be a function of the limited sampling or the properties of the individual ligand. This would support the idea that the assumption of shared bilayer orientations does not always hold, even in the case of similar SAR molecules. Converged metadynamics simulations could provide an estimate of ΔG_{memb} but care would need to be taken to consider what reaction co-ordinates are important.

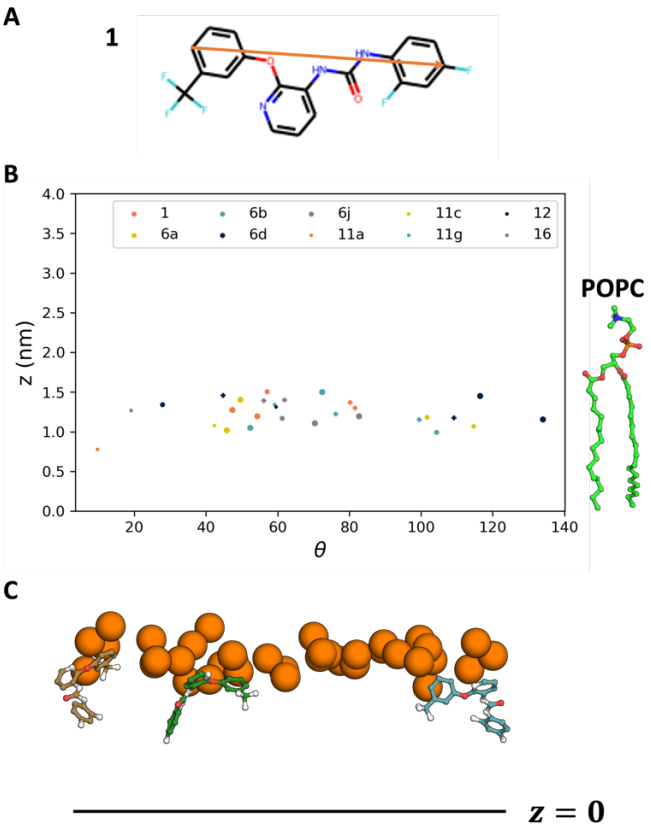


Figure 6.7: Partition Simulation Endpoints. (A) The vector through the scaffold used to define the ligand's orientation relative to z, shown on SAR molecule 1 (B) The distance from the ligand centre of mass to the bilayer centre of mass and the angle the vector in A forms relative to the z-axis (θ) at the endpoint of the partition simulations. A POPC molecule is shown to indicate scale. (C) Snapshots of the 3 simulations for SAR molecule 1. The z-component centre of mass of the bilayer is shown as a black line. Orange spheres are the phosphorous atoms in POPC.

We used these endpoints as the starting points for free energy calculations to calculate ΔG_{memb}. I began by looking at the 10 SAR molecules reported for $P2Y_1R$. I wanted to understand how uncertainty in the exacting starting orientation would impact results, so I ran 3 repeats per ligand, each starting from the endpoint of a different partitioning simulation (Figure 6.8). We saw the repeats were in good agreement (the average standard deviation was 0.63 $kcal/mol$). The ΔG_{memb} predictions were on average only 3.6 $kcal/mol$ from the experimental $\Delta G_{binding}$. This indicates partitioning into the bilayer is a significant portion of the observed affinity. A major motivation for predicting ΔG_{memb} is to better interpret SAR

and understand whether changes in $\Delta G_{observed}$ are caused by changes in the bulk partitioning or increasing the specificity. In this context, it is interesting to observe that the predictions of ΔG_{memb} correlated with the experimental $\Delta G_{observed}$ ($r_P = 0.43$ [0.18,0.86]).

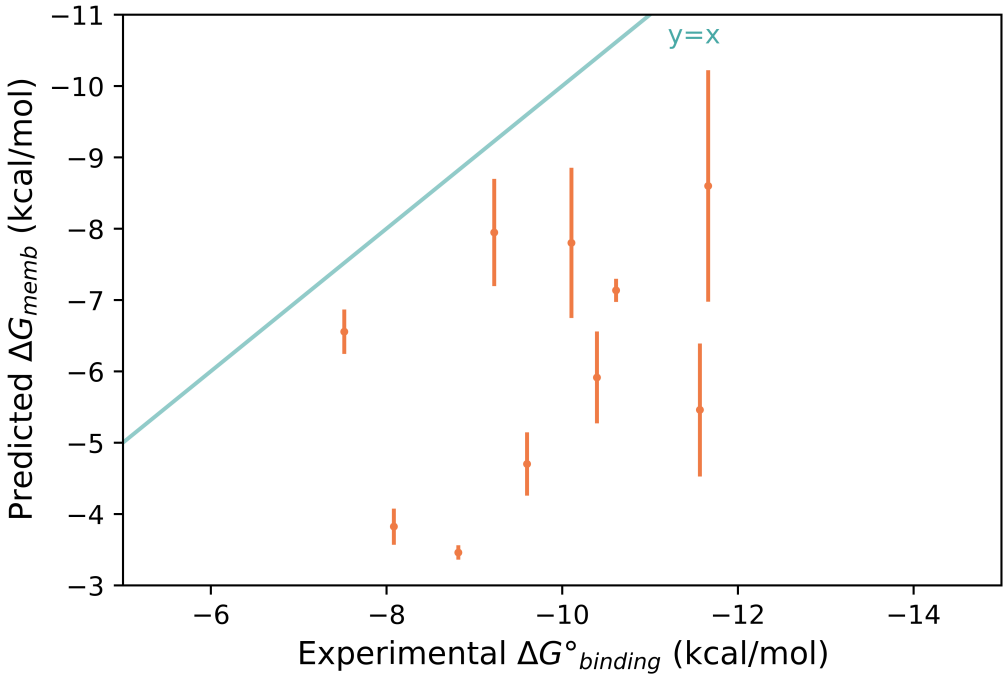


Figure 6.8: P2Y$_1$R SAR Membrane Partitioning Results. Alchemical calculations were run on selected P2Y$_1$R SAR compounds to calculate ΔG_{memb}. These predictions are plotted against the experimental values for $\Delta G_{observed}$. The error bars are the standard deviation of three repeats. The line of $y = x$ is shown in teal.

6.9. Discussion

A majority of atomistic membrane protein simulations make the simplifying assumption that the bilayer can be modelled by a single lipid species. This assumption is made to ease the sampling of lipids. However, multiple studies have shown the lipid composition can impact the dynamics of proteins in multiple ways [404]. It can therefore be expected that, as the lipid environment plays such a central role in defining the ligand's environment in the binding site, that ligand binding could be impacted by the lipid composition. This could occur via interactions between the ligand and drug, changes in the structure of the bilayer, or via lipid occupation of the binding site in the absence of ligand.

We began by looking at how PIP_2 could influence the binding of MK-0893. Previous studies have suggested that PIP_2 can bind between helices VI and VII. This site overlaps with the MK-0893 binding site, specifically where prong 1 binds [274,276]. It is therefore expected that the ligand may have to displace PIP_2, reducing the observed $\Delta G_{binding}$. I confirmed PIP_2 binds to our particular conformation of GCGR with coarse-grained MD and confirmed the stability of the bound pose with atomistic simulations. Ionic interactions between PIP_2 and charged residues - $R346^{6.37b}$ and K405 – are maintained throughout the atomistic simulations. I was interested to understand how we could incorporate this information into ABFE calculations. I decided to assess whether it was likely these contacts would be found in a non-interacting window of an ABFE calculation. To do this, I ran $8 \times 500\ ns$ simulations of GCGR in a PIP_2 containing bilayer. Whilst all 8 simulations eventually formed these interactions, 7/8 took between 250 and 500 ns to do so. This indicates we will be unable to find these interactions in an ABFE window. I confirmed this by running ABFE windows in a PIP_2 doped bilayer and did not observe any binding site occupancy or altered affinity predictions. I therefore decided that given the current accessible timescales, free energy

calculations in a doped bilayer are not currently a realistic way for dealing with mixed bilayers.

I therefore proposed an alternative method in which the SAR molecule is decoupled from a pure POPC bilayer, before an additional corrective term corresponding to $\Delta\Delta G_{POPC \rightarrow PIP_2}$ is calculated in an additional stage of the alchemical cycle. $\Delta\Delta G_{POPC \rightarrow PIP_2}$ is independent of the SAR molecule being decoupled so only a single correction is needed for the entire series. I ran coarse-grained RBFEs to calculate $\Delta\Delta G_{POPC \rightarrow PIP_2}$. Using coarse-grained simulations makes the correction significantly less expensive and simpler as the atomistic perturbation would involve ring modification. I ran the calculations following the protocol validated in Corey *et al.* [138]. These calculations returned values of $\Delta\Delta G_{POPC \rightarrow PIP_2} = -4.9 \pm 0.2\ kcal/mol$. Whilst we must be careful interpreting coarse-grained free energy predictions (particularly when comparing to atomistic values), these results support the existence of this being a PIP_2 binding site. Whilst further validation of coarse-grained free energy calculations is needed, generating experimental data lipid binding data against which to validate it remains extremely challenging [405]. If we were to add this correction to the correction observed for considering the 'true' apo state in a pure POPC bilayer, the MAE would decrease from an initial value of $12.05\ kcal/mol$ to 4.74 $kcal/mol$. Ultimately, these results show ABFEs at lipid-exposed sites must consider the apo state both in terms of protein conformation and ligand presence.

We then moved on to assess the role of cholesterol in BPTU binding to $P2Y_1R$, partly motivated by differences in experimental measurements in cholesterol free and cholesterol containing systems [254,331]. I was carrying out this investigation in the absence of structural data indicating cholesterol binding or a clear theory of what, if any, impact cholesterol would

have. I began by assessing our ability to identify cholesterol binding sites and sample these interactions. I found the crystallographic cholesterol and CHSs were not consistently stable in atomistic simulations. ABFEs including these crystallographic lipids made no difference to the predicted $\Delta G^{\theta}_{binding}$ despite interactions between the ligand and adjacent CHS. We could not identify any cholesterol binding sites either at the ligand binding site or proximal to it in either atomistic or coarse-grained simulations.

I therefore looked at the presence of direct interactions between cholesterol and the ligand. Coarse-grained simulations would allow us to probe these interactions at longer timescales than atomistic simulations. Martini3 provides an improved representation of small molecules so could provide a means to probe specific ligand and lipid interactions at longer timescales [244,245]. However, we were unable to build a coarse-grained model of BPTU in Martini3 which remained bound and, at the time of this research being undertaken, there were no validated Martini3 parameters for cholesterol. The recent publication of validated cholesterol parameters [246] and an automated Martini3 ligand parameterisation protocol [406] could allow a more detailed insight into the presence of interactions. We however relied upon $3 \times 1\ \mu s$ atomistic simulations. We identified several instances in which cholesterol interacted with the ligand. We did not see enough binding and unbinding events to get a clear picture on preferable interactions or the existence of stable interactions. We did see one interaction which was stable for ~ 500 ns. Other interactions were formed but these interactions were not consistent in terms of which part of the ligand was interacted with.

The limited timescales of atomistic simulations meant we were unable to get a converged picture of the interactions, so I decided to see if a cholesterol containing bilayer made an impact on affinity predictions. I began by running 9 ABFE calculations from 3 different

configurations of $P2Y_1R$ and BPTU in 4:1 POPC:cholesterol bilayers. These systems had undergone minimal equilibration. The overall results ($-12.6 \pm 1.4\ kcal/mol$) are similar to a pure POPC bilayer, but we do see the repeats from configuration three ($-14.2 \pm 0.7\ kcal/mol$) are lower than the results observed for the pure POPC simulations. I suspected the variance could be down to the variable cholesterol positions caused by the separate embedding processes, so I ran ABFEs from the endpoints of the long simulations. Again, we saw variable results ($-12.2 \pm 1.3\ kcal/mol$). An explanation for this is the free energy surface being relatively flat, leading to an absence of a single metastable state. It is also possible that this state does exist but we have not found it due to the timescale of our simulations.

I selected a frame from the longest stably maintained cholesterol/ligand interaction. The ABFEs run from this frame yielded the largest magnitude of free energy we observed for the $P2Y_1R$ system ($-14.1 \pm 0.4\ kcal/mol$). However, we noted the cholesterol/ligand protrude distance was significantly lower than observed in long apo simulations. This may be an artefact from the cholesterol being proximal to the binding site and then entering the non-interacting ligand vacuum. This means the observed difference could either be noise or function of this artefact. This will likely be a general problem when considering adjacent lipids. As a result of this and the general challenge of sampling mixed bilayers, I proposed an additional stage to alchemical cycle in which the energy of the cholesterol partitioning from the bilayer is calculated. This requires the use of an alternative alchemical cycle and restraints scheme [315]. Our simulations returned results of $4\ kcal/mol$. This appears reasonable due to the transient nature of the interactions, and I therefore propose that this particular interaction can be discounted. However, there is no experimental evidence to support this making assessment of the methodology challenging.

The microkinetic model is a key aspect of understanding the binding of small molecules to lipid-exposed sites. Decomposing $\Delta G_{observed}$ into ΔG_{memb} and $\Delta G_{intrinsic}$ will allow the design of more specific drugs with fewer side effects. Computational methods could be useful for disentangling this information. Recent work has showed the utility of RBFEs for predicting whether the contribution is to ΔG_{memb} or $\Delta G_{intrinsic}$ [258]. I showed that RBFEs require the ligands have similar orientational preferences, which may not be the case, particularly upon the addition of charged groups. RBFEs also do not provide information for comparing different scaffolds or assessing potential hit molecules. I therefore propose using ABFE calculations to predict $\Delta G_{intrinsic}$ and ΔG_{memb}. This requires careful consideration of a standard state correction and appropriate restraints.

I proceeded to look at ΔG_{memb} as a more tractable way of decomposing the binding energetics. Currently, umbrella sampling is the most common MD based method but often leads to asymmetric PMF profiles (especially if initial coordinates for windows were derived from steered MD). Our aim was to find an alternative to umbrella sampling that could provide more accurate results. I ran alchemical free energy calculations to calculate ΔG_{memb}. Each repeat was started from the endpoint of different unbiased simulations. The simulations returned results that were precise despite the different starting points (the average standard deviation was 0.63 $kcal/mol$). The results were relatively close in magnitude to the experimental $\Delta G_{observed}$ (the average difference was 3.6 $kcal/mol$) indicating ΔG_{memb} is a significant contributor to the observed affinity.

Interestingly, we see our predictions of ΔG_{memb} correlated with the experimental $\Delta G_{observed}$ ($r_P = 0.43$ [0.18,0.86]). This agrees with previous research using RBFEs that suggests the

changes in SAR are largely driven by increasing ΔG_{memb} [258]. Validating this with more molecules and breaking it down to different SAR vectors could be interesting. Identification of ΔG_{memb} driven $\Delta G_{observed}$ could help medicinal chemists propose better SAR modifications and select more promising hit molecules. To more directly assess our ability to predict ΔG_{memb}, future work could look at β_2AR ligands which have had their membrane affinity measured via immobolised artificial membrane (IAM) chromatography [226]. It would also be interesting to assess the correlation of these results with logP which is currently often used as a proxy for membrane partitioning.

To calculate $\Delta G_{intrinsic}$, the restraints and alchemical cycle proposed by Salari *et al.* could be used [315]. This restraints scheme has not been widely used and detailed implementation information is lacking. Implementation would likely require further optimisation of the workflow and how restraints depend on the implementation of the restraints and corrections. As an alternative to calculating $\Delta G_{intrinsic}$ directly, using the microkinetic model ($\Delta G_{intrinsic} = \Delta G_{observed} - \Delta G_{memb}$) to calculate $\Delta G_{observed}$ and ΔG_{memb} could prove more tractable.

Calculations of both ΔG_{memb} and $\Delta G_{intrinsic}$ require an accurate representation of the ligand in the bilayer. I ran 200 ns simulations to achieve this. All 3 repeats of each ligand partitioned to between 0.8 nm and 1.5 nm. I decided to further investigate this via metadynamics. These results did not converge enough to provide accurate free energy results, but they did identify a free energy minimum between 0.5 nm and 1.5 nm. The free energy surfaces also suggested some ligands show an orientational preference, whilst others do not. These orientational preference differences are caused by subtle differences in the ligand structure introducing novel lipid interactions. Consideration of other factors, such as changes in the formal/partial

charge of the ligand upon membrane partitioning [408,409], could lead to improved results. The challenge of sampling interactions will be increased when trying to consider multiple lipid species. These can significantly change the partition properties and should ideally be considered. Enhanced sampling methods would likely be needed for this. Care should be taken to identify the case where orthogonal degrees of freedom (such as specific lipid interactions) are not sampled and thus lead to an underestimation of the error.

6.10. Conclusion

Membrane proteins exist in a complex and dynamic bilayer which modulates their structures and functions. To date, most atomistic simulations are carried out in a single species bilayer (often POPC) to ease the challenges of sampling multiple lipid species and the myriad of interactions they could form with the protein and or ligand. However, this assumption is not strictly appropriate for ligands binding at lipid-exposed sites as their binding could be impacted by direct lipid interactions, the changed properties of the bilayer, or specific lipid occupying the binding site in the absence of ligand.

In this chapter, I show that the results of MK-0893 binding to GCGR can be significantly improved by considering the binding of PIP_2 in the absence of ligand. I show that this cannot easily be accounted for in atomistic ABFE windows due to the current accessible timescales. I therefore propose running calculations in a pure POPC bilayer before explicitly accounting for the PIP_2 binding by running an additional stage in which POPC is transformed to PIP_2.

I then explore the role of cholesterol in BPTU binding to $P2Y_1R$. Although no binding is observed in the absence of ligand, transient cholesterol ligand interactions were observed. I showed that consideration of this proximal cholesterol is challenging as in the non-interacting window we see occupancy of the BPTU space by cholesterol that we do not see in the long simulations. I therefore propose a method in which the binding affinity of cholesterol to the binding site, defined by protein and ligand, is calculated.

Finally, I briefly discuss membrane partitioning calculations in the context of the microkinetic model. I show that we can generate reasonable membrane bound conformations and use ABFEs to calculate ΔG_{memb}. The absolute values for ΔG_{memb} indicate binding is

largely driven by the molecule's lipophilicity. Interestingly, we see a correlation between $\Delta G_{memb\ (predicted)}$ and $\Delta G_{observed\ (experimental)}$, indicating affinity changes in the SAR series could be a function of increasing the lipophilicity of the molecule.

7. Summary

7.1. Contextualization of this thesis' findings

In this thesis, I have focused on 2 GPCR systems – $P2Y_1R$ and GCGR. These GPCRs, as with the GPCR family as a whole, are of significant pharmacological interest [410–412]. Their key role in signalling pathways means their dysfunction is associated with a plethora of diseases. The drug discovery process for GPCRs, from validating a protein as a valid target to creating a small molecule that restores normal function, remains challenging. I am particularly interested in the role structure-based drug discovery can play in facilitating the efficient discovery of efficacious drugs with minimal side effects. Computational tools have a key role to play in making a more efficient, reactive, and robust drug discovery pipeline for GPCRs. Modern computational chemists use an array of tools, but due to the dynamic nature of GPCRs, their environment, and the drug binding process, MD can play a key role. Key trends in drug discovery for GPCRs (including allosteric molecules, ligand bias, and selectivity) all also require detailed understanding of the dynamics of the protein. MD, in concert with experimental methods, is driving forward our understanding of these processes, showing how integrating computational and experimental work can facilitate significant advances [44,142,194,225,385,413,414].

One of the historic challenges for GPCR drug discovery has been obtaining structures. Recent advances in X-ray crystallography and particularly cryo-EM mean there is an ever-increasing amount of structural data [16,415,416]. The challenge is how do we maximise the conversion of this structural data into tangible improvements in our understanding of biological process and improved outcomes for patients. One area which structural data has highlighted is the presence of allosteric binding sites for small molecules outside of the 7TM helices. These sites hold great promise for drug discovery and give rise to several questions: Are they less conserved between subfamily members and does this allow the development of more specific

drugs? Can their dependence on the bilayer environment be utilised to develop molecules that target specific tissues? These molecules are likely to be highly lipophilic, therefore is it possible to develop drugs which do not have significant off-target effects? Can we decompose $\Delta G_{observed}$ into ΔG_{memb} and $\Delta G_{intrinsic}$? These challenges require investigation, and this thesis attempts to contribute towards answering some of them.

One of the fundamental challenges of developing computational techniques to predict binding affinities at lipid exposed sites is the lack of high-quality experimental data. Most experimental data for lipid-exposed sites are radioligand assays. These often lack important details required to reproduce them computationally, for instance the exact lipid composition. This makes it difficult to validate ABFE calculations against individual systems. The lack of consistency in the conditions in which these experiments are undertaken and the general concerns around the quality of the data mean the generation of large, reliable, and directly comparable datasets is challenging. This is important for developing empirical / ML based scoring functions for docking. Inconsistencies in the quality and quantity of experimental data is a factor that hinders development of improved docking scoring functions, and we expect this to be exacerbated for lipid exposed sites. Lack of high-quality experimental data also hinders our ability to develop and validate robust methods for predicting membrane partitioning. However, the challenges around generating high quality experimental data, whilst making validation more challenging, creates a niche which can be occupied by computational methods.

We have demonstrated that ABFE calculations can work well at lipid-exposed binding sites on some systems, whilst on others it can be incredibly challenging. I have identified numerous challenges, suggested how to monitor for these problems in the case of prospective

drug discovery case, and proposed solutions to them. Two of the major issues I identified in this thesis are based around sampling. I showed that sampling lipid interactions can be challenging even in a single species bilayer. As the importance of the role of lipids continues to be increasingly recognised, the need to incorporate them into our systems will also grow. The other major challenge is the conformational heterogeneity of many systems of interest. This conformational flexibility is key to the function of many proteins and processes of interest. For instance, probing a GPCR agonist would require sampling large conformational changes associated with GPCR activation. Massive performance boosts from improved hardware and implementation of MD code have allowed larger systems to be simulated for longer but as hardware improves, so does the scope of the questions we ask. Membrane protein systems are incredibly important and complex systems and therefore often drive advances in theory and application. Understanding membrane protein activation and allostery is driving numerous enhanced sampling methods to better understand protein dynamics [385] and understanding the role of protein/lipid interactions has led to numerous advances in coarse-grained representations of molecules [244].

From the perspective of free energy calculations, equilibrium alchemical calculations are still considered the most accurate. They also have the advantages of being shown to be effective on a number of systems [68,74,76,88,321–323], having various pieces of software facilitate the setup and analysis [81], and the likely sources of error are reasonably well understood. Ultimately absolute binding free energy calculations should aim to become a tool that is widely used in drug discovery programs. To achieve this, beyond being easy to setup and interpret, free energy calculations must provide results quickly and economically. The average repeat of a single ligand required $1.36\ \mu s$ of simulation time ($0.88\ \mu s$ for the complex simulations and $0.48\ \mu s$ for the ligand in solution simulations). We note that this consists of 76 individual

windows, each of which take less than 12 hours. The time to run individual simulations is expected to continue to decrease with time and we expect chemists to have access to larger clusters. One can therefore imagine a medicinal chemist wanting to try a new scaffold being able to submit simulations as they leave work and having completed simulations for when they return to work the next day.

It is important to keep pushing the frontiers of what is possible to better understand the capabilities and limitations of the method. The worked discussed in this thesis is one of the most ambitious applications of free energy calculations. However, as a more diverse set of problems are considered, equilibrium ABFEs will not be the most suitable tool available for every research question. Recent years have seen the development of methods such as AWH [417] and the increased popularity of non-equilibrium free energy calculations. Potential benefits of these systems include the sampling of slow degrees of freedom [325]. This is particularly interesting for GPCRs where both lipid sampling and conformational changes represent slow and potentially important degrees of freedom. It will be interesting to see how these methods evolve and perform under continued evaluation. In general, continued hardware improvements, from specialised supercomputers [42] and to a potential role for quantum computing [418], coupled with investment in infrastructure and further advances in enhanced sampling methods will allow us to ask ever more pertinent questions and answer them in ever more robust ways.

Another fundamental area seeing significant improvements is how molecules are represented. Classical MM forcefields, whilst a good compromise of accuracy and efficiency in many cases, continues to be shown to be unsuitable in many circumstances. This is particularly the case when simulating membranes and ligands. Methods which allow the charges of a

molecule to change over the course of a simulation, polarisable forcefields [419], and those which do not treat an atom as a single point charge will facilitate better representations of the interactions seen with bilayers [407,420]. Improvements in forcefield design pipelines are facilitating a more efficient forcefield design process [341]. Torsional parameters for ligands remain challenging, but increasingly efficient and user-friendly packages will likely facilitate further improved representation of ligands [294].

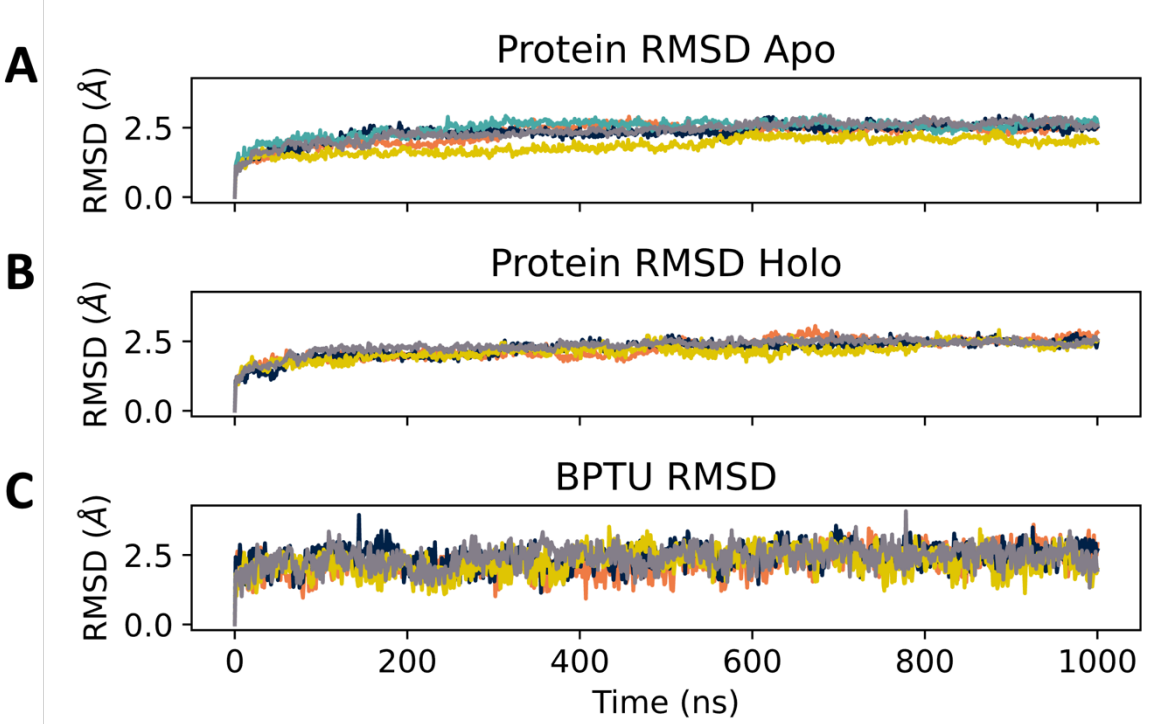
7.2. Conclusion

In this thesis, I aimed to validate the use of ABFE calculations on small molecules that bind at lipid-exposed binding sites. In particular, I attempted to identify and address challenges that are unique to or exacerbated in these sites. In Chapter 3, I show that ABFE calculations can be applied successfully to predict binding affinities for BPTU (and related analogues) to $P2Y_1R$ and reproduce mutagenesis data. I also show that lipid sampling is challenging on the timescale of ABFE windows. I propose methods to monitor equilibration and lipid sampling. In Chapter 4, I looked at MK-0893 (and related analogues) binding to GCGR. This is an exceptionally challenging system. However, I show that both ABFEs and RBFEs can accurately rank the MK-0893 SAR series. I observed a large overestimation of the magnitude of binding and showed how explicit treatment of the apo state can improve results. In Chapter 5, I investigated how to include information about the apo state in a robust manner. I proposed a method to improve sampling of the apo and holo states and how to select an appropriate ratio of apo:holo windows. In Chapter 6, I showed how consideration of multiple lipid species can be vital for accurate affinity predictions. I showed that for the GCGR system $PI(4,5)P_2$ binding in the absence of ligand must be considered. I showed that the effect of complex bilayers cannot be accounted for by simply running ABFEs in doped bilayers due to

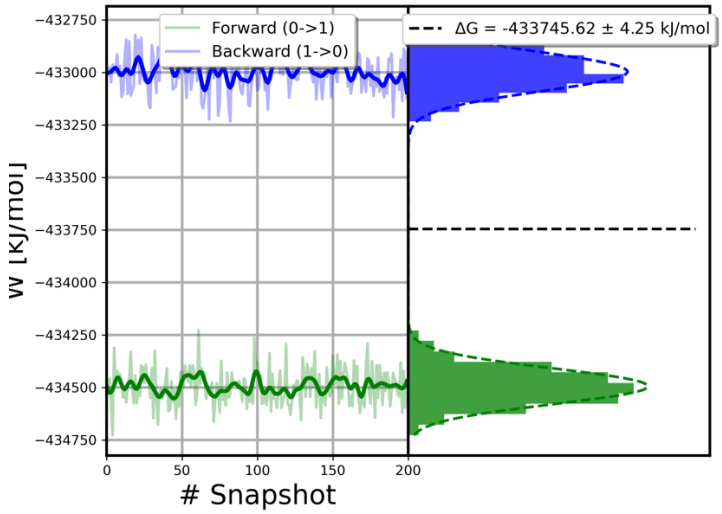
the currently accessible timescales. I therefore propose methods involving an additional stage of the alchemical cycle. The data discussed in this thesis is available on Zenodo.

Future work in this area should be aiming to facilitate the use of ABFEs in a prospective manner. I must therefore overcome known challenges and develop better ways to identify errors when we cannot compare to experimental data. Of particular importance for ABFE calculations on GPCRs and at lipid-exposed sites are conformational changes in the protein and the complex nature of the lipid bilayer.

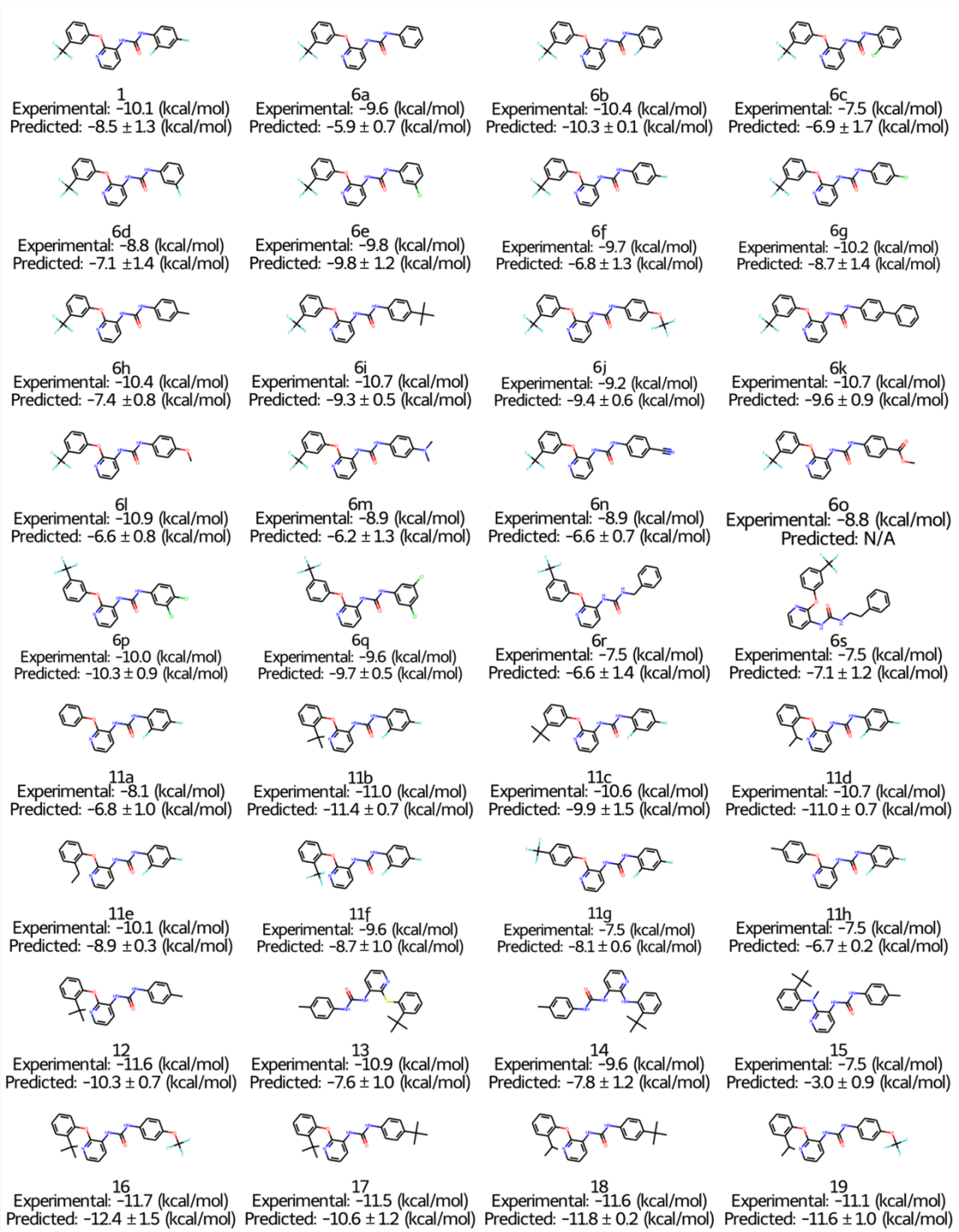
8. Supplementary Material: Chapter 3



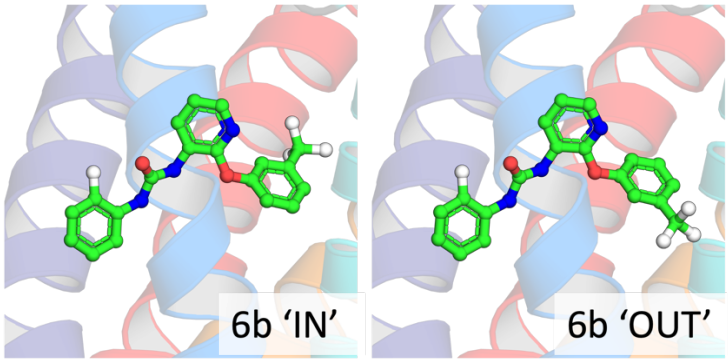
Supplementary Figure 8.1: Unbiased Simulations RMSD. (A) RMSD of $P2Y_1R$ in the absence of BPTU (B) RMSD of $P2Y_1R$ in the presence of BPTU (C) RMSD of BPTU bound to $P2Y_1R$



Supplementary Figure 8.2: Non-equilibrium work distribution.



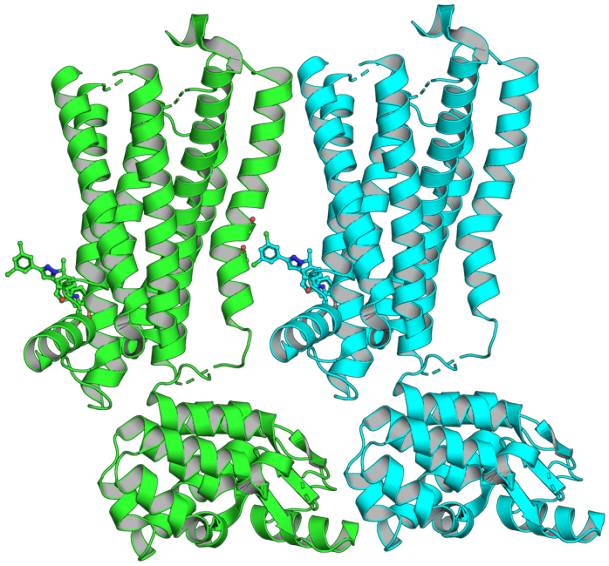
Supplementary Figure 8.3: P2Y$_1$R BPTU SAR Grid. The BPTU SAR series with experimental [331] *and predicted* $\Delta G^{\theta}_{binding}$.



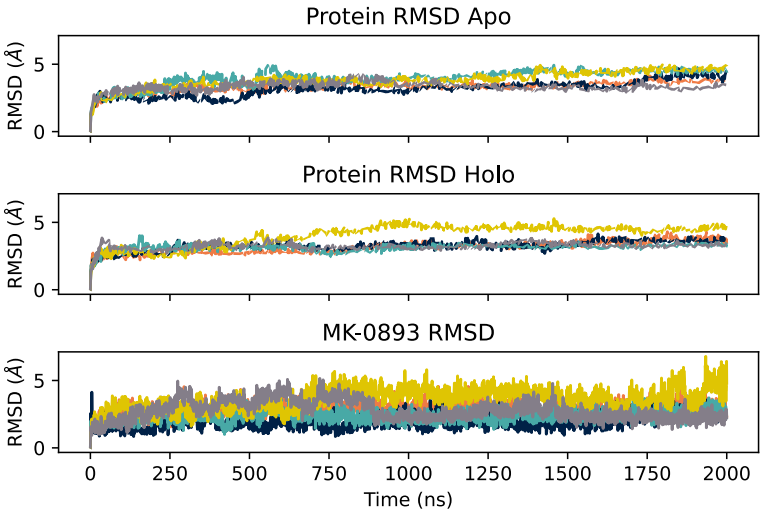
	ΔG° (kcal/mol)
Experimental	-10.3
ABFE 'in'	-7.0 ± 0.7
ABFE 'out'	-10.5 ± 0.6
MM-GBSA 'in'	-21.5
MM-GBSA 'out'	-17.1

Supplementary Figure 8.4: MM/GBSA 6b conformer prediction.

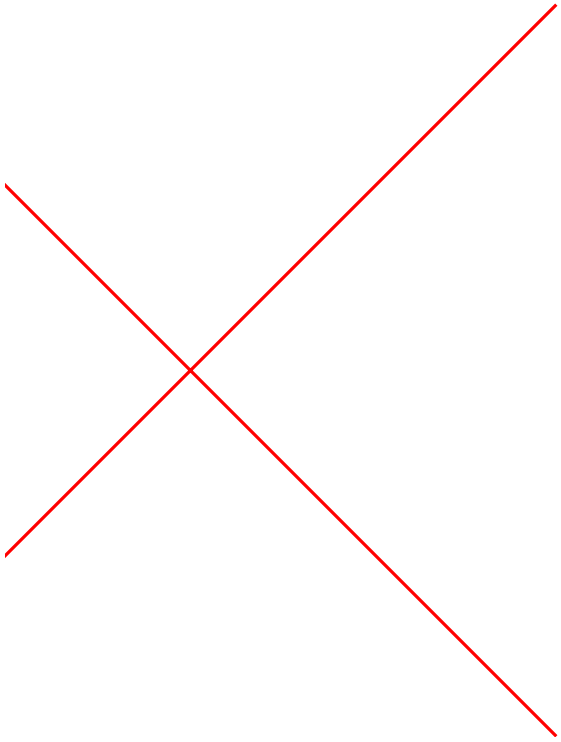
9. Supplementary Material: Chapter 4



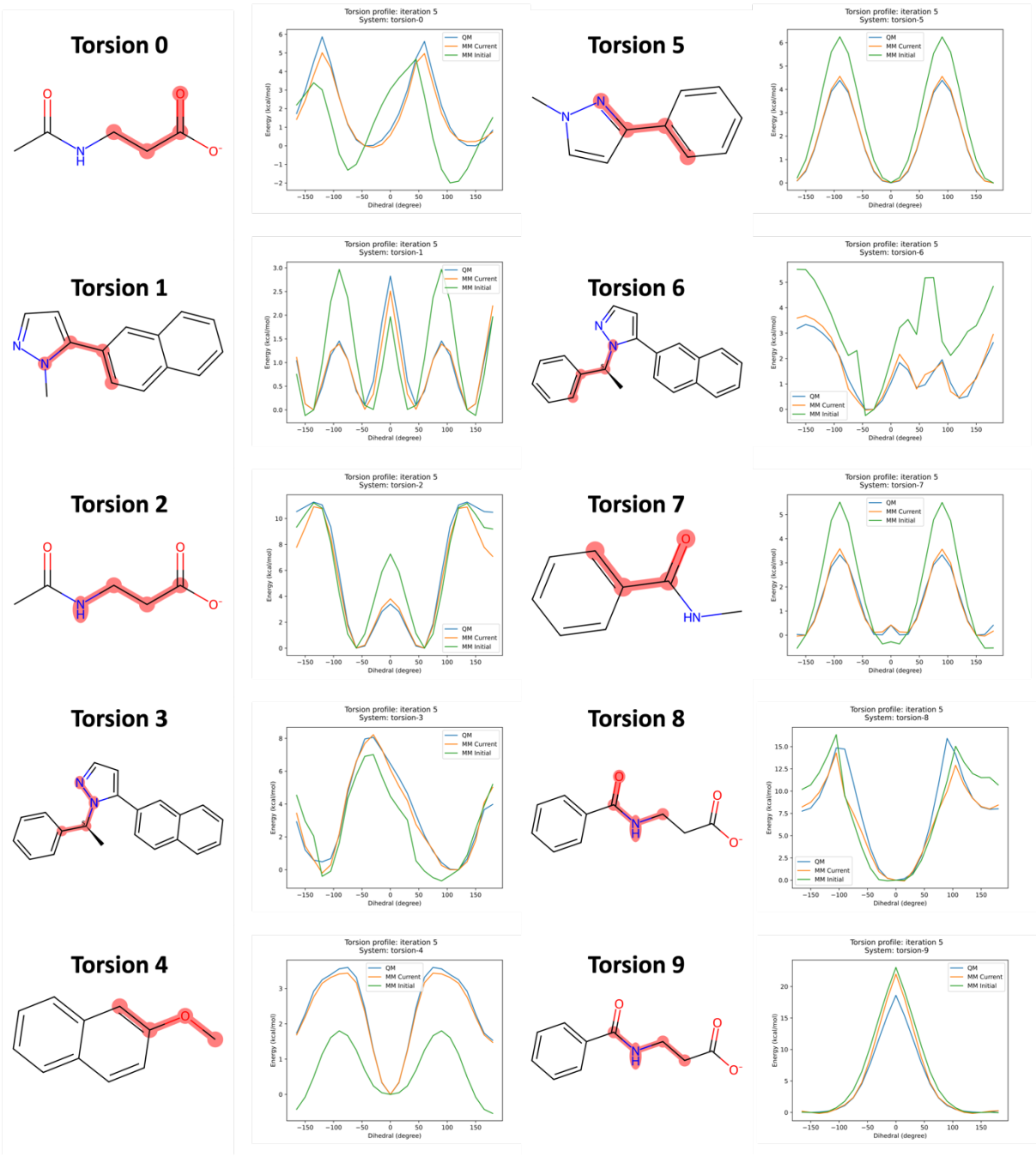
Supplementary Figure 9.1 GCGR symmetry mate. MK-0893 is shown in sticks bound to GCGR (PDB: 5ee7) and making contact with the symmetry mate in the adjacent crystal cell



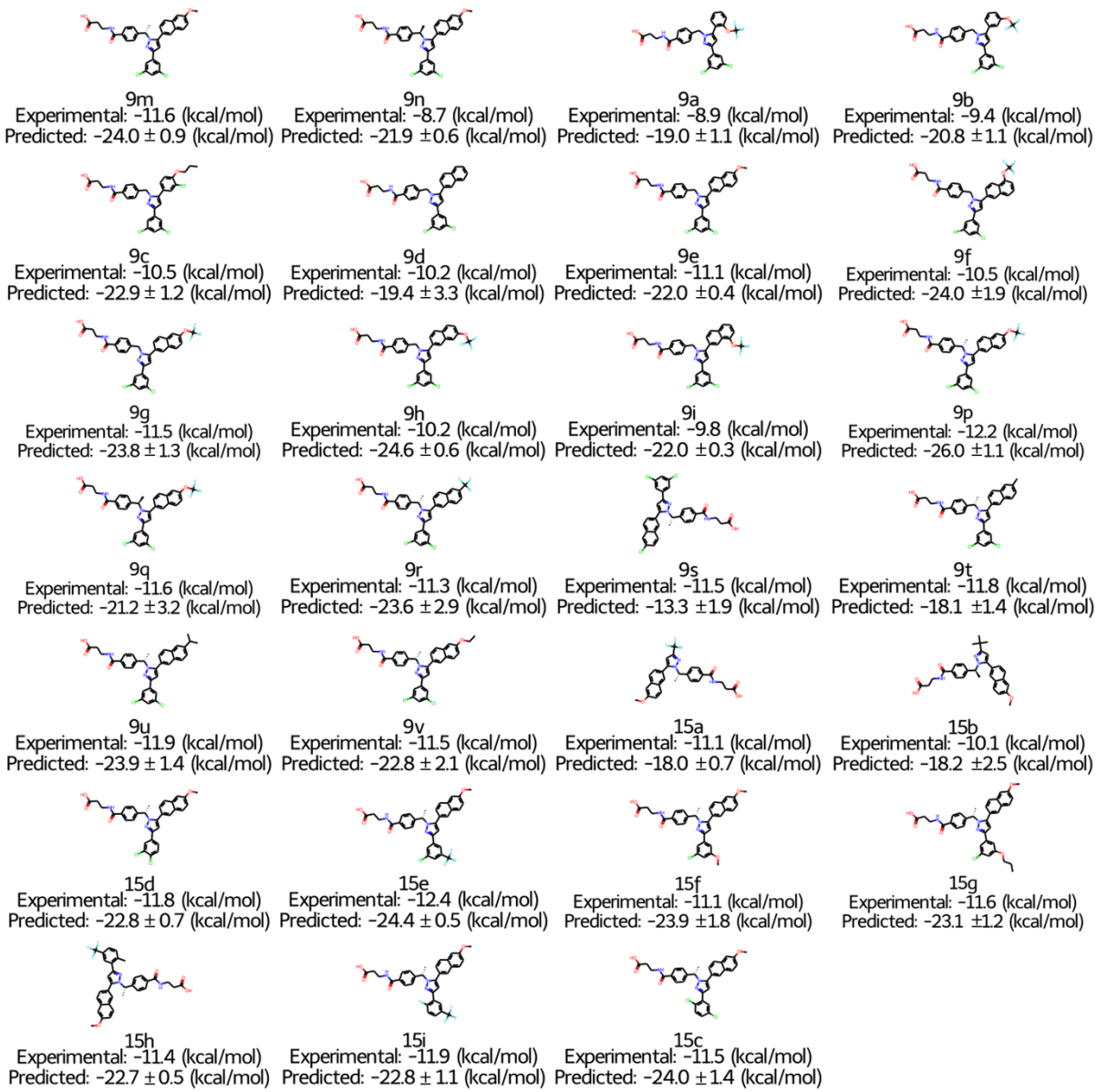
Supplementary Figure 9.2: Unbiased simulation RMSDs. Top: RMSD of GCGR in the absence of MK-0893 Middle: RMSD of GCGR in presence of MK-0893 Bottom: RMSD of MK-0893



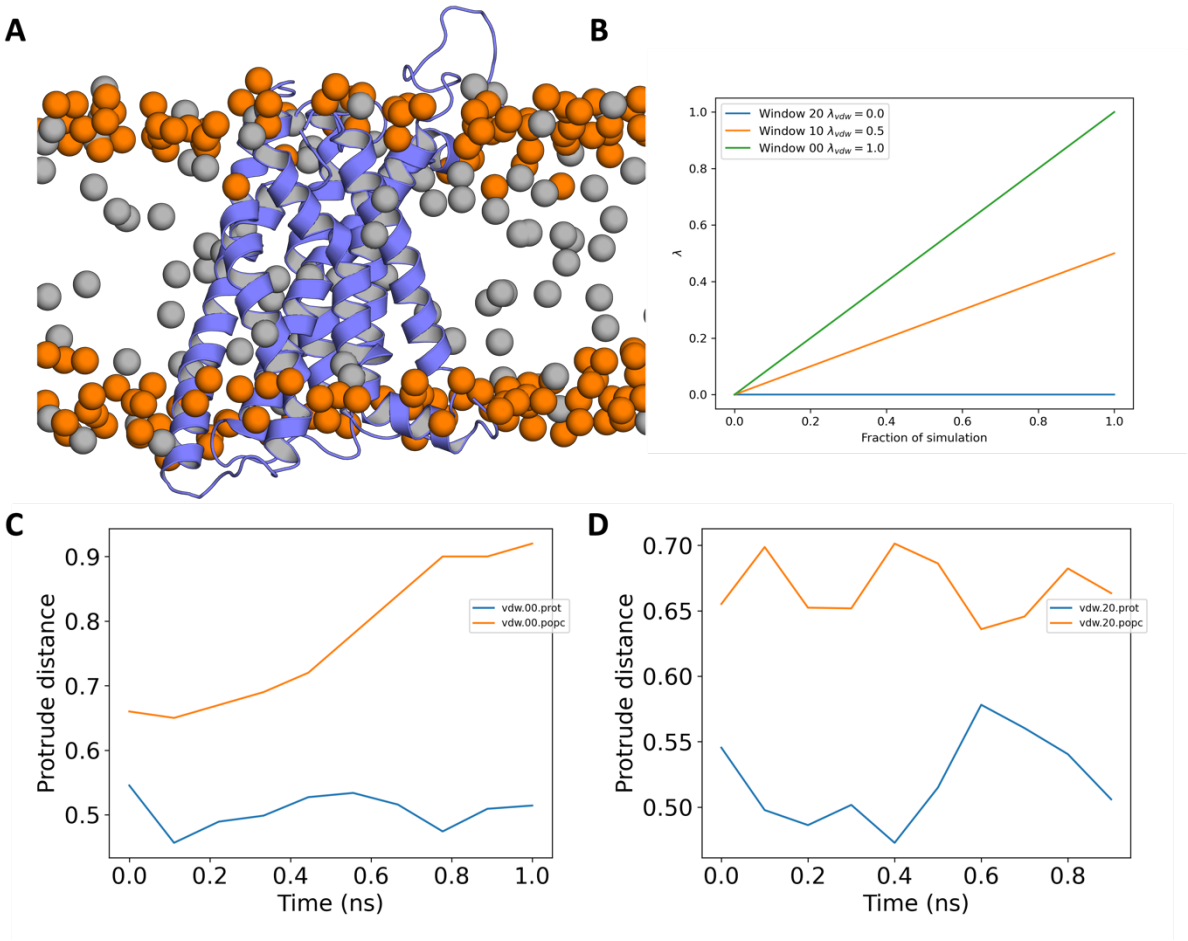
Supplementary Figure 9.3: Stability of water bridges. The presence of water bridges between the ligand and selected binding site residues across 5×2 *μs simulations. The top 3 rows are first order networks for S350, L399, and R173 whilst the final row is the presence of first and second order networks for R173.*



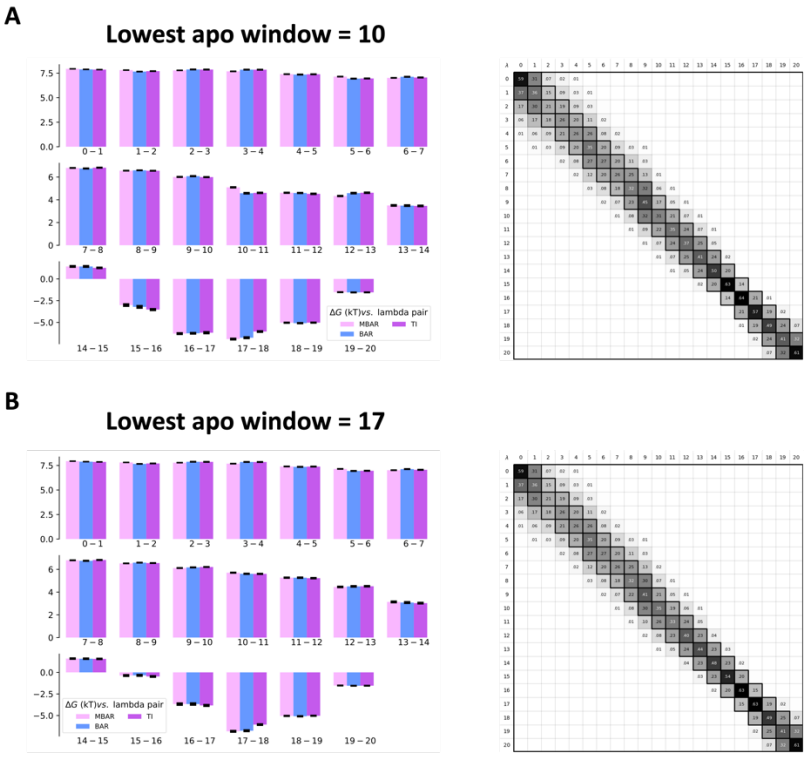
Supplementary Figure 9.4: BespokeFit Torsions. The fragments and the initial sage torsion landscape (green), the QM profile (blue), and the optimised MM parameters (orange)



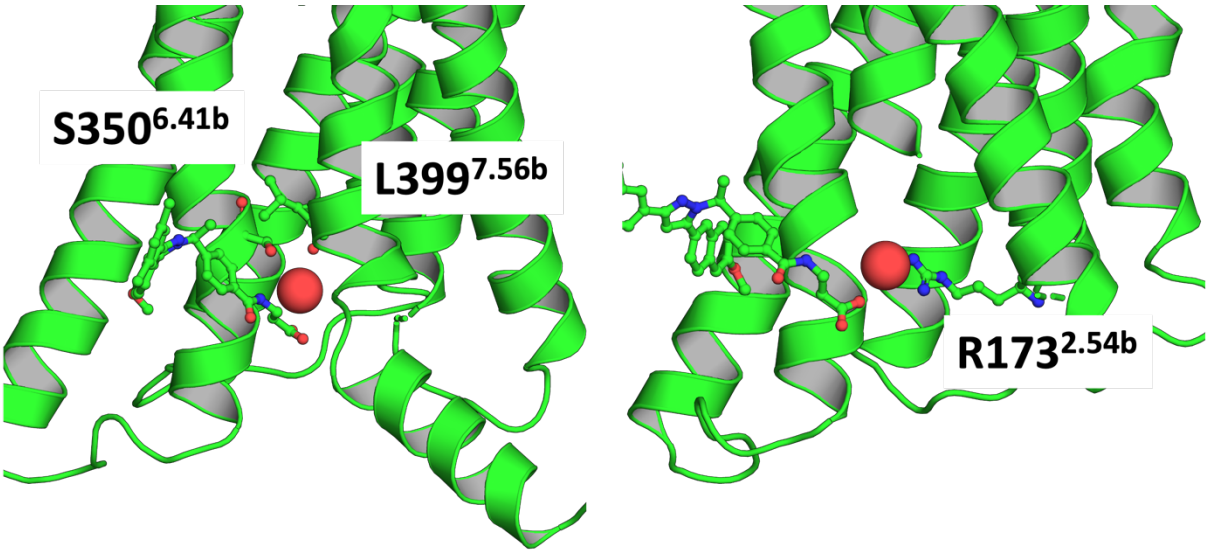
Supplementary Figure 9.5: GCGR SAR ABFE Results & Structures Grid. Structures of the MK-0893 SAR with their experimentally determined $\Delta G^{\theta}_{binding}$ and the predicted values with HREX



Supplementary Figure 9.6: Dummy atom validation. (A) Structure of GCGR (purple cartoon) in a POPC bilayer (headgroup P shown as orange spheres). Non-interacting dummy atoms (grey spheres) are placed randomly throughout the bilayer and protein. (B) The value of λ across different alchembed equilibration simulations. λ increases linearly across the simulations at different gradients. (C) The protrude distance for non-interacting dummy atoms for the vdw.00 windows. The protrude distance for the dummy atoms with the protein stays the same across the simulation as the dummy atoms do not interact with protein (blue) whilst for the POPC it increases over the simulation (orange). (D) In the vdw.20 mimicking simulation both protein and POPC protrude distances remain the same as $\lambda = 0$ throughout.

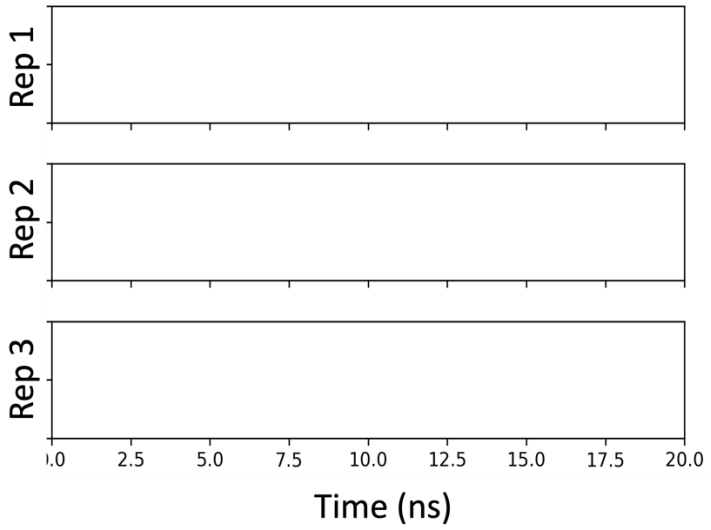


Supplementary Figure 9.7: MK-0893 Apo/Holo ABFE Overlap Matrices. (A) The window by window predictions for different estimators (left) and the MBAR overlap matrix for results where windows 0-9 are run from the holo conformation, whilst 10-20 are run from the apo and (B) 0-16 are run from the holo conformation, whilst 17-20 are run from the apo

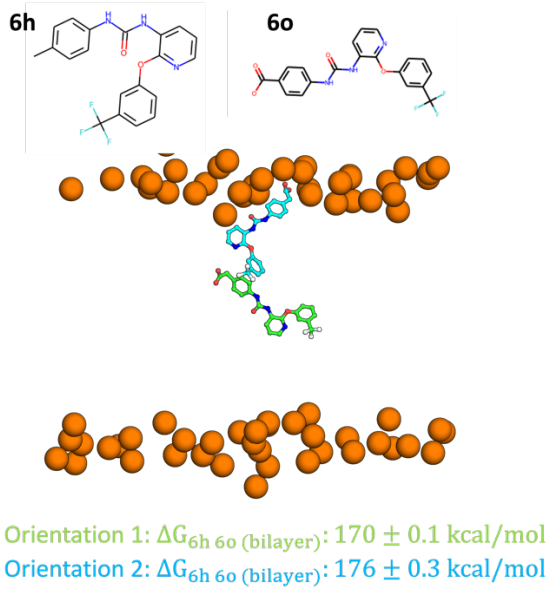


Supplementary Figure 9.8: GCGR water bridges. MK-0893 (shown in green sticks) forming interactions via a bridging water to S350$^{6.41b}$, L399$^{7.56b}$, and R173$^{2.54b}$.

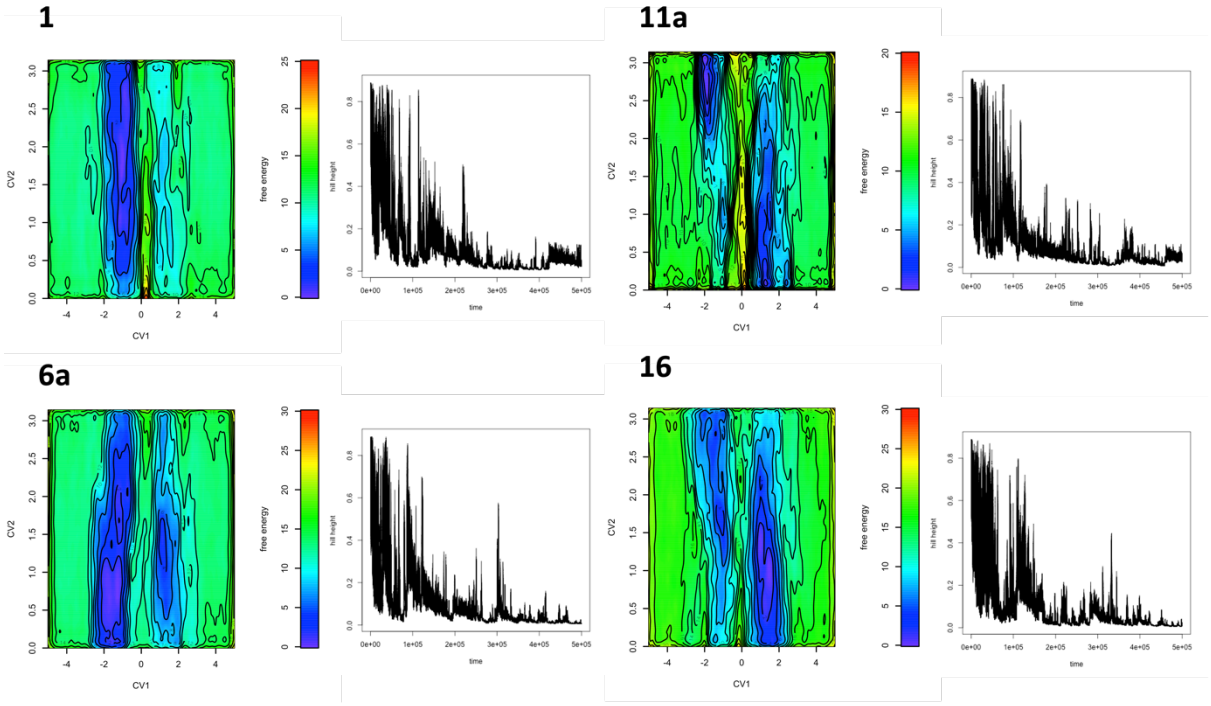
10. Supplementary Material: Chapter 6



Supplementary Figure 10.1: vdW20 R346$^{6.37b}$/PIP$_2$ contacts. There were no contacts observed between R346$^{6.37b}$ and PIP$_2$ in the non-interacting windows of ABFEs of MK-0893 in a doped PIP$_2$ bilayer



Supplementary Figure 10.2: RBFE Orientational Dependence. The bilayer phase of an RBFE was run converting 6h to 6o with 2 different orientations of the ligand. The carboxylate of 6o was either placed in the bilayer core or in the headgroups. The results differ by about 6 kcal/mol.



Supplementary Figure 10.3: Partition Metadynamics. Free energy surfaces from metadynamics with their corresponding hills plots. The deposition of Gaussians and the asymmetry in the values predicted for the leaflets indicates a lack of convergence.

11. References

1. DiMasi, J. A., Grabowski, H. G. & Hansen, R. W. Innovation in the pharmaceutical industry: New estimates of R&D costs. *J Health Econ* **47**, 20–33 (2016).
2. Wouters, O. J., McKee, M. & Luyten, J. Estimated Research and Development Investment Needed to Bring a New Medicine to Market, 2009-2018. *JAMA* **323**, 844 (2020).
3. Gao, L., Maidment, I., Matthews, F. E., Robinson, L. & Brayne, C. Medication usage change in older people (65+) in England over 20 years: findings from CFAS I and CFAS II. *Age Ageing* **47**, 220–225 (2018).
4. Baden, L. R. *et al.* Efficacy and Safety of the mRNA-1273 SARS-CoV-2 Vaccine. *New England Journal of Medicine* **384**, 403–416 (2021).
5. Polack, F. P. *et al.* Safety and Efficacy of the BNT162b2 mRNA Covid-19 Vaccine. *New England Journal of Medicine* **383**, 2603–2615 (2020).
6. Drożdżal, S. *et al.* An update on drugs with therapeutic potential for SARS-CoV-2 (COVID-19) treatment. *Drug Resistance Updates* **59**, 100794 (2021).
7. Boby, M. L. *et al.* Open Science Discovery of Potent Non-Covalent SARS-CoV-2 Main Protease Inhibitors. *bioRxiv* 2020.10.29.339317 (2023) doi:10.1101/2020.10.29.339317.
8. Crowdsourced Covid Antiviral Project Gets $11 Million in Funding - Bloomberg. https://www.bloomberg.com/news/articles/2021-09-27/crowdsourced-covid-antiviral-project-gets-11-million-in-funding?leadSource=uverify%20wall.
9. Miethke, M. *et al.* Towards the sustainable discovery and development of new antibiotics. *Nature Reviews Chemistry 2021 5:10* **5**, 726–749 (2021).

10. Dugger, S. A., Platt, A. & Goldstein, D. B. Drug development in the era of precision medicine. *Nature Reviews Drug Discovery 2017 17:3* **17**, 183–196 (2017).

11. Hu, Y., Zhao, T., Zhang, N., Zhang, Y. & Cheng, L. A Review of Recent Advances and Research on Drug Target Identification Methods. *Curr Drug Metab* **20**, 209–216 (2019).

12. Mayr, L. M. & Bojanic, D. Novel trends in high-throughput screening. *Curr Opin Pharmacol* **9**, 580–588 (2009).

13. Polishchuk, P. G., Madzhidov, T. I. & Varnek, A. Estimation of the size of drug-like chemical space based on GDB-17 data. *J Comput Aided Mol Des* **27**, 675–679 (2013).

14. Follmann, M. *et al.* An approach towards enhancement of a screening library: The Next Generation Library Initiative (NGLI) at Bayer — against all odds? *Drug Discov Today* **24**, 668–672 (2019).

15. Chilingaryan, Z., Yin, Z. & Oakley, A. J. Fragment-Based Screening by Protein Crystallography: Successes and Pitfalls. *International Journal of Molecular Sciences 2012, Vol. 13, Pages 12857-12879* **13**, 12857–12879 (2012).

16. Maveyraud, L. & Mourey, L. Protein X-ray Crystallography and Drug Discovery. *Molecules* **25**, (2020).

17. Sugiki, T., Furuita, K., Fujiwara, T. & Kojima, C. Current NMR Techniques for Structure-Based Drug Discovery. *Molecules : A Journal of Synthetic Chemistry and Natural Product Chemistry* **23**, (2018).

18. Zhu, K. F. *et al.* Applications and prospects of cryo-EM in drug discovery. *Military Medical Research 2023 10:1* **10**, 1–14 (2023).

19. de Oliveira, T. M., van Beek, L., Shilliday, F., Debreczeni, J. & Phillips, C. Cryo-EM: The Resolution Revolution and Drug Discovery. *SLAS Discovery* **26**, 17–31 (2021).

20. Thomas, S. E. *et al.* Structure-guided fragment-based drug discovery at the synchrotron: screening binding sites and correlations with hotspot mapping. *Philos Trans A Math Phys Eng Sci* **377**, (2019).

21. Subramaniam, S. & Kleywegt, G. J. A paradigm shift in structural biology. *Nature Methods 2022 19:1* **19**, 20–23 (2022).

22. Berman, H. M. *et al.* The Protein Data Bank. *Nucleic Acids Res* **28**, 235–242 (2000).

23. Kryshtafovych, A., Schwede, T., Topf, M., Fidelis, K. & Moult, J. Critical Assessment of Methods of Protein Structure Prediction (CASP) – Round XIV. *Proteins* **89**, 1607 (2021).

24. Jumper, J. *et al.* Highly accurate protein structure prediction with AlphaFold. *Nature 2021 596:7873* **596**, 583–589 (2021).

25. Baek, M. *et al.* Accurate prediction of protein structures and interactions using a three-track neural network. *Science (1979)* **373**, 871–876 (2021).

26. Nussinov, R., Zhang, M., Liu, Y. & Jang, H. AlphaFold, allosteric, and orthosteric drug discovery: Ways forward. *Drug Discov Today* **28**, 103551 (2023).

27. Trott, O. & Olson, A. J. AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J Comput Chem* **31**, 455–461 (2010).

28. Kumar, A. & Zhang, K. Y. J. Hierarchical virtual screening approaches in small molecule drug discovery. *Methods* **71**, 26–37 (2015).

29. Fischer, M., Coleman, R. G., Fraser, J. S. & Shoichet, B. K. Incorporation of protein flexibility and conformational energy penalties in docking screens to improve ligand discovery. *Nat Chem* **6**, 575–583 (2014).

30. McGovern, S. L. & Shoichet, B. K. Information decay in molecular docking screens against holo, apo, and modeled conformations of enzymes. *J Med Chem* **46**, 2895–2907 (2003).

31. Bender, B. J. *et al.* A practical guide to large-scale docking. *Nature Protocols 2021 16:10* **16**, 4799–4832 (2021).

32. Díaz-Rovira, A. M. *et al.* Are Deep Learning Structural Models Sufficiently Accurate for Virtual Screening? Application of Docking Algorithms to AlphaFold2 Predicted Structures. *J Chem Inf Model* **63**, 1668–1674 (2023).

33. Vajda, S., Beglov, D., Wakefield, A. E., Egbert, M. & Whitty, A. Cryptic binding sites on proteins: definition, detection, and druggability. *Curr Opin Chem Biol* **44**, 1 (2018).

34. Colombo, G. Computing allostery: from the understanding of biomolecular regulation and the discovery of cryptic sites to molecular design. *Curr Opin Struct Biol* **83**, 102702 (2023).

35. Hu, G., Doruker, P., Li, H. & Demet Akten, E. Editorial: Understanding Protein Dynamics, Binding and Allostery for Drug Design. *Front Mol Biosci* **8**, 681364 (2021).

36. Karl, K., Paul, M. D., Pasquale, E. B. & Hristova, K. Ligand bias in receptor tyrosine kinase signaling. *J Biol Chem* **295**, 18494–18507 (2020).

37. Zimmerman, M. I. *et al.* SARS-CoV-2 Simulations Go Exascale to Capture Spike Opening and Reveal Cryptic Pockets Across the Proteome. *Nat Chem* **13**, 651 (2021).

38. Pietrucci, F., Vargiu, A. V. & Kranjc, A. HIV-1 Protease Dimerization Dynamics Reveals a Transient Druggable Binding Pocket at the Interface. *Scientific Reports 2015 5:1* **5**, 1–12 (2015).

39. Kapla, J., Espigares, I. R., Ballante, F., Selent, J. & Carlsson, J. Can molecular dynamics simulations improve the structural accuracy and virtual screening performance of GPCR models? *PLoS Comput Biol* **17**, e1008936 (2021).

40. Kacker, P., Masetti, M., Mangold, M., Bottegoni, G. & Cavalli, A. Combining dyad protonation and active site plasticity in BACE-1 structure-based drug design. *J Chem Inf Model* **52**, 1079–1085 (2012).

41. Rastelli, G., Degliesposti, G., Del Rio, A. & Sgobba, M. Binding estimation after refinement, a new automated procedure for the refinement and rescoring of docked ligands in virtual screening. *Chem Biol Drug Des* **73**, 283–286 (2009).

42. Shaw, D. E. *et al.* Anton 3: Twenty Microseconds of Molecular Dynamics Simulation before Lunch. *International Conference for High Performance Computing, Networking, Storage and Analysis, SC* (2021) doi:10.1145/3458817.3487397.

43. Shan, Y. *et al.* How does a drug molecule find its target binding site? *J Am Chem Soc* **133**, 9181–9183 (2011).

44. Dror, R. O. *et al.* Pathway and mechanism of drug binding to G-protein-coupled receptors. *Proc Natl Acad Sci U S A* **108**, 13118–13123 (2011).

45. Zhang, R. & Monsma, F. Binding kinetics and mechanism of action: Toward the discovery and development of better and best in class drugs. *Expert Opin Drug Discov* **5**, 1023–1029 (2010).

46. Copeland, R. A. Evolution of the drug-target residence time model. *Expert Opin Drug Discov* **16**, 1441–1451 (2021).

47. Hothersall, J. D., Brown, A. J., Dale, I. & Rawlins, P. Can residence time offer a useful strategy to target agonist drugs for sustained GPCR responses? *Drug Discov Today* **21**, 90–96 (2016).

48. Copeland, R. A. The drug–target residence time model: a 10-year retrospective. *Nature Reviews Drug Discovery 2015 15:2* **15**, 87–95 (2015).

49. Tummino, P. J. & Copeland, R. A. Residence time of receptor-ligand complexes and its effect on biological function. *Biochemistry* **47**, 5481–5492 (2008).

50. Pan, A. C., Xu, H., Palpant, T. & Shaw, D. E. Quantitative Characterization of the Binding and Unbinding of Millimolar Drug Fragments with Molecular Dynamics Simulations. *J Chem Theory Comput* **13**, 3372–3377 (2017).

51. Singhal, N., Snow, C. D. & Pande, V. S. Using path sampling to build better Markovian state models: predicting the folding rate and mechanism of a tryptophan zipper beta hairpin. *J Chem Phys* **121**, 415–425 (2004).

52. Huber, G. A. & Kim, S. Weighted-Ensemble Brownian Dynamics Simulations for Protein Association Reactions. *Biophys J* **70**, 97–110 (1996).

53. Tiwary, P. & Parrinello, M. From metadynamics to dynamics. *Phys Rev Lett* **111**, 230602 (2013).

54. Ahmad, K. *et al.* Enhanced-Sampling Simulations for the Estimation of Ligand Binding Kinetics: Current Status and Perspective. *Front Mol Biosci* **9**, 899805 (2022).

55. Bissantz, C., Kuhn, B. & Stahl, M. A medicinal chemist's guide to molecular interactions. *J Med Chem* **53**, 5061–5084 (2010).

56. Chodera, J. D. & Mobley, D. L. Entropy-enthalpy compensation: Role and ramifications in biomolecular ligand recognition and design. *Annu Rev Biophys* **42**, 121 (2013).

57. Huang, S. Y., Grinter, S. Z. & Zou, X. Scoring functions and their evaluation methods for protein–ligand docking: recent advances and future directions. *Physical Chemistry Chemical Physics* **12**, 12899–12908 (2010).

58. Li, J., Fu, A. & Zhang, L. An Overview of Scoring Functions Used for Protein–Ligand Interactions in Molecular Docking. *Interdisciplinary Sciences: Computational Life Sciences 2019 11:2* **11**, 320–328 (2019).

59. Volkov, M. *et al.* On the Frustration to Predict Binding Affinities from Protein-Ligand Structures with Deep Neural Networks. *J Med Chem* **65**, 7946–7958 (2022).

60. Wang, C., Greene, D., Xiao, L., Qi, R. & Luo, R. Recent developments and applications of the MMPBSA method. *Front Mol Biosci* **4**, 312193 (2018).

61. Poli, G., Granchi, C., Rizzolio, F. & Tuccinardi, T. Application of MM-PBSA Methods in Virtual Screening. *Molecules* **25**, (2020).

62. Aldeghi, M., Bodkin, M. J., Knapp, S. & Biggin, P. C. Statistical Analysis on the Performance of Molecular Mechanics Poisson-Boltzmann Surface Area versus Absolute Binding Free Energy Calculations: Bromodomains as a Case Study. *J Chem Inf Model* **57**, 2203–2221 (2017).

63. Johannes K. Umbrella sampling. doi:10.1002/wcms.66.

64. Roux, B., Andersen, O. S. & Allen, T. W. Comment on ‘free energy simulations of single and double ion occupancy in gramicidin A’ (The Journal of Chemical Physics (2007) 126 (105103)). *Journal of Chemical Physics* **128**, (2008).

65. Markthaler, D., Jakobtorweihen, S. & Hansen, N. Lessons Learned from the Calculation of One-Dimensional Potentials of Mean Force [Article v1.0]. *Living J Comput Mol Sci* **1**, 11073–11073 (2019).

66. You, W., Tang, Z. & Chang, C. E. A. Potential Mean Force from Umbrella Sampling Simulations: What Can We Learn and What Is Missed? *J Chem Theory Comput* **15**, 2433–2443 (2019).

67. Baumann, H. M. *et al.* Broadening the scope of binding free energy calculations using a Separated Topologies approach. (2023) doi:10.26434/CHEMRXIV-2023-9678K.

68. Boresch, S., Tettinger, F., Leitgeb, M. & Karplus, M. Absolute Binding Free Energies: A Quantitative Approach for Their Calculation. (2003) doi:10.1021/jp0217839.

69. Liu, S., Wang, L. & Mobley, D. L. Is Ring breaking feasible in relative binding free energy calculations? *J Chem Inf Model* **55**, 727–735 (2015).

70. Abel, R., Wang, L., Harder, E. D., Berne, B. J. & Friesner, R. A. Advancing Drug Discovery through Enhanced Free Energy Calculations. *Acc Chem Res* **50**, 1625–1632 (2017).

71. Armacost, K. A., Riniker, S. & Cournia, Z. Novel Directions in Free Energy Methods and Applications. *J Chem Inf Model* **60**, 1–5 (2020).

72. Cournia, Z., Allen, B. & Sherman, W. Relative Binding Free Energy Calculations in Drug Discovery: Recent Advances and Practical Considerations. *J Chem Inf Model* **57**, 2911–2937 (2017).

73. Moore, J. H. *et al.* Automated relative binding free energy calculations from SMILES to ΔΔG. *Communications Chemistry 2023 6:1* **6**, 1–12 (2023).

74. Aldeghi, M., Heifetz, A., Bodkin, M. J., Knapp, S. & Biggin, P. C. Accurate calculation of the absolute free energy of binding for drug molecules. *Chem Sci* **7**, 207–218 (2015).

75. Feng, M., Heinzelmann, G. & Gilson, M. K. Absolute binding free energy calculations improve enrichment of actives in virtual compound screening. *Scientific Reports 2022 12:1* **12**, 1–11 (2022).

76. Alibay, I., Magarkar, A., Seeliger, D. & Biggin, P. C. Evaluating the use of absolute binding free energy in the fragment optimisation process. *Communications Chemistry 2022 5:1* **5**, 1–13 (2022).

77. Baumann, H. M., Gapsys, V., De Groot, B. L. & Mobley, D. L. Challenges Encountered Applying Equilibrium and Nonequilibrium Binding Free Energy Calculations. *Journal of Physical Chemistry B* **125**, 4241–4261 (2021).

78. Işık, M. *et al.* Overview of the SAMPL6 pKa Challenge: Evaluating small molecule microscopic and macroscopic pKa predictions. *J Comput Aided Mol Des* **35**, 131 (2021).

79. Mobley, D. L. & Klimovich, P. V. Perspective: Alchemical free energy calculations for drug discovery. *J Chem Phys* **137**, 230901 (2012).

80. Deng, N. *et al.* Large scale free energy calculations for blind predictions of protein–ligand binding: the D3R Grand Challenge 2015. *J Comput Aided Mol Des* **30**, 743–751 (2016).

81. Heinzelmann, G. & Gilson, M. K. Automation of absolute protein-ligand binding free energy calculations for docking refinement and compound evaluation. *Scientific Reports 2021 11:1* **11**, 1–18 (2021).

82. Jarzynski, C. Equilibrium free energy differences from nonequilibrium measurements: a master equation approach. *Phys Rev E Stat Phys Plasmas Fluids Relat Interdiscip Topics* **56**, 5018–5035 (1997).

83. Jarzynski, C. A nonequilibrium equality for free energy differences. *Phys Rev Lett* **78**, 2690–2693 (1996).

84. Crooks, G. E. Entropy production fluctuation theorem and the nonequilibrium work relation for free energy differences. *Phys Rev E* **60**, 2721 (1999).

85. Mobley, D. L., Chodera, J. D. & Dill, K. A. Confine-and-release method: Obtaining correct binding free energies in the presence of protein conformational change. *J Chem Theory Comput* **3**, 1231–1235 (2007).

86. Khalak, Y. *et al.* Alchemical absolute protein–ligand binding free energies for drug design. *Chem Sci* **12**, 13958–13971 (2021).

87. Chen, W. *et al.* Enhancing Hit Discovery in Virtual Screening through Absolute Protein-Ligand Binding Free-Energy Calculations. *J Chem Inf Model* **63**, 3171–3185 (2023).

88. Aldeghi, M., Heifetz, A., Bodkin, M. J., Knapp, S. & Biggin, P. C. Predictions of ligand selectivity from absolute binding free energy calculations. *J Am Chem Soc* **139**, 946–957 (2017).

89. Gill, S. C. *et al.* Binding Modes of Ligands Using Enhanced Sampling (BLUES): Rapid Decorrelation of Ligand Binding Modes via Nonequilibrium Candidate Monte Carlo. *Journal of Physical Chemistry B* **122**, 5579–5598 (2018).

90. Samways, M. L., Bruce Macdonald, H. E. & Essex, J. W. Grand: A python module for grand canonical water sampling in OpenMM. *J Chem Inf Model* **60**, 4436–4441 (2020).

91. Burley, K. H., Gill, S. C., Lim, N. M. & Mobley, D. L. Enhancing Sidechain Rotamer Sampling Using Non-Equilibrium Candidate Monte Carlo. *J Chem Theory Comput* **15**, 1848 (2019).

92. Melling, O. J., Samways, M. L., Ge, Y., Mobley, D. L. & Essex, J. W. Enhanced Grand Canonical Sampling of Occluded Water Sites Using Nonequilibrium Candidate Monte Carlo. *J Chem Theory Comput* **19**, 1050–1062 (2023).

93. Bergazin, T. D. *et al.* Enhancing Water Sampling of Buried Binding Sites Using Nonequilibrium Candidate Monte Carlo. *J Comput Aided Mol Des* **35**, 167 (2021).

94. Beuming, T. *et al.* Are Deep Learning Structural Models Sufficiently Accurate for Free-Energy Calculations? Application of FEP+ to AlphaFold2-Predicted Structures. *J Chem Inf Model* **62**, 4351–4360 (2022).

95. Almén, M. S., Nordström, K. J. V., Fredriksson, R. & Schiöth, H. B. Mapping the human membrane proteome: a majority of the human membrane proteins can be classified according to function and evolutionary origin. *BMC Biol* **7**, 50 (2009).

96. Kampen, K. R. Membrane proteins: the key players of a cancer cell. *J Membr Biol* **242**, 69–74 (2011).

97. Lee, S. H., Kim, D. H., Kuzmanov, U. & Gramolini, A. O. Membrane proteomic profiling of the heart: past, present, and future. *Am J Physiol Heart Circ Physiol* **320**, H417–H423 (2021).

98. Moore, J. M., Bell, E. L., Hughes, R. O. & Garfield, A. S. ABC transporters: human disease and pharmacotherapeutic potential. *Trends Mol Med* **29**, 152–172 (2023).

99. Ng, D. P., Poulsen, B. E. & Deber, C. M. Membrane protein misassembly in disease. *Biochim Biophys Acta* **1818**, 1115–1122 (2012).

100. Nagy, J. K. & Sanders, C. R. Destabilizing mutations promote membrane protein misfolding. *Biochemistry* **43**, 19–25 (2004).

101. Junne, T., Schwede, T., Goder, V. & Spiess, M. Mutations in the Sec61p channel affecting signal sequence recognition and membrane protein topology. *J Biol Chem* **282**, 33201–33209 (2007).

102. Costin, G. E. *et al.* Mutations in dopachrome tautomerase (Dct) affect eumelanin/pheomelanin synthesis, but do not affect intracellular trafficking of the mutant protein. *Biochem J* **391**, 249–259 (2005).

103. Yin, H. & Flynn, A. D. Drugging Membrane Protein Interactions. *Annu Rev Biomed Eng* **18**, 51 (2016).

104. Mandelker, D. *et al.* A frequent kinase domain mutation that changes the interaction between PI3Kalpha and the membrane. *Proc Natl Acad Sci U S A* **106**, 16996–17001 (2009).

105. Candido, K. L. *et al.* Spike protein of SARS-CoV-2 variants: a brief review and practical implications. *Brazilian Journal of Microbiology* **53**, 1133 (2022).

106. Chlanda, P. *et al.* Structural Analysis of the Roles of Influenza A Virus Membrane-Associated Proteins in Assembly and Morphology. *J Virol* **89**, 8957 (2015).

107. Brandsma, A. M., Hilmer, C., Rauch, M., Matuschewski, K. & Montagna, G. N. Plasmodium early transcribed membrane proteins appear tailored to the host range of malaria parasites. *Int J Parasitol* **52**, 135–143 (2022).

108. Blanco, P. *et al.* Bacterial Multidrug Efflux Pumps: Much More Than Antibiotic Resistance Determinants. *Microorganisms* **4**, (2016).

109. Z, V., P, H., G, P., R, G. & Z, K. Involvement of membrane channels in autoimmune disorders. *Curr Pharm Des* **13**, 2456–2468 (2007).

110. Delporte, C. & Soyfoo, M. Aquaporins: Unexpected actors in autoimmune diseases. *Autoimmun Rev* **21**, 103131 (2022).

111. Overington, J. P., Al-Lazikani, B. & Hopkins, A. L. How many drug targets are there? *Nature Reviews Drug Discovery 2006 5:12* **5**, 993–996 (2006).

112. Birch, J. *et al.* The fine art of integral membrane protein crystallisation. *Methods* **147**, 150–162 (2018).

113. White, S. Known Structures of Membrane Proteins. *https://blanco.biomol.uci.edu/mpstruc/*.

114. Harpole, T. J. & Delemotte, L. Conformational landscapes of membrane proteins delineated by enhanced sampling molecular dynamics simulations. *Biochimica et Biophysica Acta (BBA) - Biomembranes* **1860**, 909–926 (2018).

115. Rao, S., Klesse, G., Lynch, C. I., Tucker, S. J. & Sansom, M. S. P. Molecular Simulations of Hydrophobic Gating of Pentameric Ligand Gated Ion Channels: Insights into Water and Ions. *Journal of Physical Chemistry B* **125**, 981–994 (2021).

116. Kutzner, C., Grubmüller, H., De Groot, B. L. & Zachariae, U. Computational Electrophysiology: The Molecular Dynamics of Ion Channel Permeation and Selectivity in Atomistic Detail. *Biophys J* **101**, 809 (2011).

117. Musgaard, M., Paramo, T., Domicevica, L., Andersen, O. J. & Biggin, P. C. Insights into channel dysfunction from modelling and molecular dynamics simulations. *Neuropharmacology* **132**, 20–30 (2018).

118. Lichtinger, S. M. & Biggin, P. C. Tackling Hysteresis in Conformational Sampling: How to Be Forgetful with MEMENTO. *J Chem Theory Comput* **19**, 3705–3720 (2023).

119. Harayama, T. & Riezman, H. Understanding the diversity of membrane lipid composition. *Nature Reviews Molecular Cell Biology 2018 19:5* **19**, 281–296 (2018).

120. Van Meer, G., Voelker, D. R. & Feigenson, G. W. Membrane lipids: where they are and how they behave. *Nature Reviews Molecular Cell Biology 2008 9:2* **9**, 112–124 (2008).

121. Van Meer, G. & De Kroon, A. I. P. M. Lipid map of the mammalian cell. *J Cell Sci* **124**, 5–8 (2011).

122. Coskun, Ü. & Simons, K. Cell membranes: the lipid perspective. *Structure* **19**, 1543–1548 (2011).

123. Boughter, C. T., Monje-Galvan, V., Im, W. & Klauda, J. B. Influence of cholesterol on phospholipid bilayer structure and dynamics. *Journal of Physical Chemistry B* **120**, 11761–11772 (2016).

124. Klecker, T. & Westermann, B. Pathways shaping the mitochondrial inner membrane. *Open Biol* **11**, (2021).

125. Lingwood, D. & Simons, K. Lipid rafts as a membrane-organizing principle. *Science* **327**, 46–50 (2010).

126. Stansfeld, P. J. & Sansom, M. S. P. From coarse grained to atomistic: A serial multiscale approach to membrane protein simulations. *J Chem Theory Comput* **7**, 1157–1166 (2011).

127. Khalid, S., Piggot, T. J. & Samsudin, F. Atomistic and Coarse Grain Simulations of the Cell Envelope of Gram-Negative Bacteria: What Have We Learned? *Acc Chem Res* **52**, 180–188 (2019).

128. Larsen, A. H. Molecular Dynamics Simulations of Curved Lipid Membranes. *Int J Mol Sci* **23**, (2022).

129. Dämgen, M. A. & Biggin, P. C. State-dependent protein-lipid interactions of a pentameric ligand-gated ion channel in a neuronal membrane. *PLoS Comput Biol* **17**, (2021).

130. Gupta, K. *et al.* The role of interfacial lipids in stabilizing membrane protein oligomers. *Nature 2017 541:7637* **541**, 421–424 (2017).

131. Whorton, M. R. & MacKinnon, R. Crystal structure of the mammalian GIRK2 K+ channel and gating regulation by G proteins, PIP2, and sodium. *Cell* **147**, 199–208 (2011).

132. Norimatsu, Y., Hasegawa, K., Shimizu, N. & Toyoshima, C. Protein-phospholipid interplay revealed with crystals of a calcium pump. *Nature* **545**, 193–198 (2017).

133. Corey, R. A., Sansom, M. S. P. & Stansfeld, P. J. Identification and Characterization of Specific Protein–Lipid Interactions Using Molecular Simulation. *Methods in Molecular Biology* **2315**, 121–139 (2021).

134. Schmidt, M. R., Stansfeld, P. J., Tucker, S. J. & Sansom, M. S. P. Simulation-based prediction of phosphatidylinositol 4,5-bisphosphate binding to an ion channel. *Biochemistry* **52**, 279–281 (2013).

135. Lee, J. Y. & Lyman, E. Predictions for cholesterol interaction sites on the A2A adenosine receptor. *J Am Chem Soc* **134**, 16512–16515 (2012).

136. Stansfeld, P. J., Hopkinson, R., Ashcroft, F. M. & Sansom, M. S. P. PIP2-binding site in Kir channels: Definition by multiscale biomolecular simulations. *Biochemistry* **48**, 10926–10933 (2009).

137. Ansell, T. B. *et al.* LipIDens: Simulation assisted interpretation of lipid densities in cryo-EM structures of membrane proteins. *bioRxiv* 2022.06.30.498233 (2022) doi:10.1101/2022.06.30.498233.

138. Corey, R. A., Vickery, O. N., Sansom, M. S. P. & Stansfeld, P. J. Insights into Membrane Protein-Lipid Interactions from Free Energy Calculations. *J Chem Theory Comput* **15**, 5727–5736 (2019).

139. Santos, R. *et al.* A comprehensive map of molecular drug targets. *Nature Reviews Drug Discovery 2016 16:1* **16**, 19–34 (2016).

140. Yang, D. *et al.* G protein-coupled receptors: structure- and function-based drug discovery. *Signal Transduction and Targeted Therapy 2020 6:1* **6**, 1–27 (2021).

141. Hauser, A. S., Attwood, M. M., Rask-Andersen, M., Schiöth, H. B. & Gloriam, D. E. Trends in GPCR drug discovery: new agents, targets and indications. *Nature Reviews Drug Discovery 2017 16:12* **16**, 829–842 (2017).

142. Huang, W. *et al.* Structural insights into μ-opioid receptor activation. *Nature* **524**, 315–321 (2015).

143. Thal, D. M. *et al.* Recent advances in the determination of G protein-coupled receptor structures. *Curr Opin Struct Biol* **51**, 28–34 (2018).

144. Weis, W. I. & Kobilka, B. K. The Molecular Basis of G Protein–Coupled Receptor Activation. *Annu Rev Biochem* **87**, 897 (2018).

145. Wold, E. A., Chen, J., Cunningham, K. A. & Zhou, J. Allosteric Modulation of Class A GPCRs: Targets, Agents, and Emerging Concepts. *J Med Chem* **62**, 88 (2019).

146. Gurevich, V. V. & Gurevich, E. V. Biased GPCR signaling: Possible mechanisms and inherent limitations. *Pharmacol Ther* **211**, 107540 (2020).

147. Lu, S., Jang, W., Inoue, A. & Lambert, N. A. Constitutive G protein coupling profiles of understudied orphan GPCRs. *PLoS One* **16**, e0247743 (2021).

148. Nusse, R. & Clevers, H. Wnt/β-Catenin Signaling, Disease, and Emerging Therapeutic Modalities. *Cell* **169**, 985–999 (2017).

149. Hutchings, C. J. Mini-review: antibody therapeutics targeting G protein-coupled receptors and ion channels. *Antib Ther* **3**, 257–264 (2020).

150. Fan, L. *et al.* Haloperidol bound D2 dopamine receptor structure inspired the discovery of subtype selective ligands. *Nat Commun* **11**, (2020).

151. The Nobel Prize in Chemistry 2012 - Advanced information - NobelPrize.org. https://www.nobelprize.org/prizes/chemistry/2012/advanced-information/.

152. Ballesteros, J. A. & Weinstein, H. [19] Integrated methods for the construction of three-dimensional models and computational probing of structure-function relations in G protein-coupled receptors. *Methods in Neurosciences* **25**, 366–428 (1995).

153. Rasmussen, S. G. F. *et al.* Crystal structure of the human β2 adrenergic G-protein-coupled receptor. *Nature 2007 450:7168* **450**, 383–387 (2007).

154. Rasmussen, S. G. F. *et al.* Structure of a nanobody-stabilized active state of the β2 adrenoceptor. *Nature 2011 469:7329* **469**, 175–180 (2011).

155. Rasmussen, S. G. F. *et al.* Crystal structure of the β2 adrenergic receptor-Gs protein complex. *Nature* **477**, 549–557 (2011).

156. Kobilka, B. The Structural Basis of G-Protein-Coupled Receptor Signaling (Nobel Lecture). *Angew Chem Int Ed Engl* **52**, 6380 (2013).

157. Standfuss, J. *et al.* Crystal structure of constitutively active rhodopsin: How an agonist can activate its GPCR. *Nature* **471**, 656 (2011).

158. Kruse, A. C. *et al.* Activation and allosteric modulation of a muscarinic acetylcholine receptor. *Nature* **504**, 101 (2013).

159. Carpenter, B., Nehmé, R., Warne, T., Leslie, A. G. W. & Tate, C. G. Structure of the adenosine A2A receptor bound to an engineered G protein. *Nature* **536**, 104 (2016).

160. Venkatakrishnan, A. J. *et al.* Molecular signatures of G-protein-coupled receptors. *Nature* **494**, 185–194 (2013).

161. Deupi, X. & Kobilka, B. Activation of G protein-coupled receptors. *Adv Protein Chem* **74**, 137–166 (2007).

162. Trzaskowski, B. *et al.* Action of molecular switches in GPCRs--theoretical and experimental studies. *Curr Med Chem* **19**, 1090–1109 (2012).

163. Rovati, G. E. *et al.* The DRY motif and the four corners of the cubic ternary complex model. *Cell Signal* **35**, 16–23 (2017).

164. Gabilondo, A. M., Krasel, C. & Lohse, M. J. Mutations of Tyr326 in the beta 2-adrenoceptor disrupt multiple receptor functions. *Eur J Pharmacol* **307**, 243–250 (1996).

165. Ragnarsson, L., Andersson, Å., Thomas, W. G. & Lewis, R. J. Mutations in the NPxxY motif stabilize pharmacologically distinct conformational states of the α1B- and β2-adrenoceptors. *Sci Signal* **12**, (2019).

166. Nussinov, R., Ma, B. & Tsai, C. J. Multiple conformational selection and induced fit events take place in allosteric propagation. *Biophys Chem* **186**, 22 (2014).

167. Xiao, J. Y. *et al.* The effect of ligand efficacy on the formation and stability of a GPCR-G protein complex. *Proc Natl Acad Sci U S A* **106**, 9501–9506 (2009).

168. De Lean, A., Stadel, J. M. & Lefkowitz, R. J. A ternary complex model explains the agonist-specific binding properties of the adenylate cyclase-coupled beta-adrenergic receptor. *Journal of Biological Chemistry* **255**, 7108–7117 (1980).

169. Gurevich, V. V. & Gurevich, E. V. GPCR signaling regulation: The role of GRKs and arrestins. *Front Pharmacol* **10**, 437285 (2019).

170. Luttrell, L. M. & Gesty-Palmer, D. Beyond Desensitization: Physiological Relevance of Arrestin-Dependent Signaling. *Pharmacol Rev* **62**, 305–330 (2010).

171. Luttrell, L. M. *et al.* β-arrestin-dependent formation of β2 adrenergic receptor-src protein kinase complexes. *Science (1979)* **283**, 655–661 (1999).

172. Wu, N. *et al.* High-mass MALDI-MS unravels ligand-mediated G protein-coupling selectivity to GPCRs. *Proc Natl Acad Sci U S A* **118**, e2024146118 (2021).

173. Manglik, A. *et al.* Structural Insights into the Dynamic Process of β2-Adrenergic Receptor Signaling. *Cell* **161**, 1101–1111 (2015).

174. Nygaard, R. *et al.* The dynamic process of β(2)-adrenergic receptor activation. *Cell* **152**, 532–542 (2013).

175. Staus, D. P. *et al.* Allosteric nanobodies reveal the dynamic range and diverse mechanisms of G-protein-coupled receptor activation. *Nature* **535**, 448–452 (2016).

176. Bokoch, M. P. *et al.* Ligand-specific regulation of the extracellular surface of a G-protein-coupled receptor. *Nature* **463**, 108–112 (2010).

177. Liu, J. J., Horst, R., Katritch, V., Stevens, R. C. & Wüthrich, K. Biased signaling pathways in β2-adrenergic receptor characterized by 19F-NMR. *Science* **335**, 1106–1110 (2012).

178. Manglik, A. *et al.* Structural Insights into the Dynamic Process of β2-Adrenergic Receptor Signaling. *Cell* **161**, 1101–1111 (2015).

179. Jeschke, G. DEER distance measurements on proteins. *Annu Rev Phys Chem* **63**, 419–446 (2012).

180. Manglik, A. *et al.* Structural Insights into the Dynamic Process of β2-Adrenergic Receptor Signaling. *Cell* **161**, 1101–1111 (2015).

181. Ye, L., Van Eps, N., Zimmer, M., Ernst, O. P. & Scott Prosser, R. Activation of the A2A adenosine G-protein-coupled receptor by conformational selection. *Nature* **533**, 265–268 (2016).

182. Wingler, L. M. *et al.* Angiotensin Analogs with Divergent Bias Stabilize Distinct Receptor Conformations. *Cell* **176**, 468-478.e11 (2019).

183. Liu, J. J., Horst, R., Katritch, V., Stevens, R. C. & Wüthrich, K. Biased signaling pathways in β2-adrenergic receptor characterized by 19F-NMR. *Science* **335**, 1106–1110 (2012).

184. Lamichhane, R. *et al.* Single-molecule view of basal activity and activation mechanisms of the G protein-coupled receptor β2AR. *Proc Natl Acad Sci U S A* **112**, 14254–14259 (2015).

185. Imai, S. *et al.* Structural equilibrium underlying ligand-dependent activation of β2-adrenoreceptor. *Nat Chem Biol* **16**, 430–439 (2020).

186. Ye, L., Van Eps, N., Zimmer, M., Ernst, O. P. & Scott Prosser, R. Activation of the A2A adenosine G-protein-coupled receptor by conformational selection. *Nature* **533**, 265–268 (2016).

187. Okude, J. *et al.* Identification of a Conformational Equilibrium That Determines the Efficacy and Functional Selectivity of the μ-Opioid Receptor. *Angew Chem Int Ed Engl* **54**, 15771–15776 (2015).

188. Wu, F. J. *et al.* Probing the correlation between ligand efficacy and conformational diversity at the α1A-adrenoreceptor reveals allosteric coupling of its microswitches. *J Biol Chem* **295**, 7404–7417 (2020).

189. Liu, J. J., Horst, R., Katritch, V., Stevens, R. C. & Wüthrich, K. Biased signaling pathways in β2-adrenergic receptor characterized by 19F-NMR. *Science* **335**, 1106–1110 (2012).

190. Ye, L., Larda, S. T., Frank Li, Y. F., Manglik, A. & Prosser, R. S. A comparison of chemical shift sensitivity of trifluoromethyl tags: optimizing resolution in ^{19}F NMR studies of proteins. *J Biomol NMR* **62**, 97–103 (2015).

191. Chung, K. Y. *et al.* Role of detergents in conformational exchange of a G protein-coupled receptor. *J Biol Chem* **287**, 36305–36311 (2012).

192. Grahl, A., Abiko, L. A., Isogai, S., Sharpe, T. & Grzesiek, S. A high-resolution description of β1-adrenergic receptor functional dynamics and allosteric coupling from backbone NMR. *Nature Communications 2020 11:1* **11**, 1–13 (2020).

193. Isogai, S. *et al.* Backbone NMR reveals allosteric signal transduction networks in the β1-adrenergic receptor. *Nature* **530**, 237–241 (2016).

194. Dror, R. O. *et al.* Activation mechanism of the β 2-adrenergic receptor. *Proc Natl Acad Sci U S A* **108**, 18684–18689 (2011).

195. Kohlhoff, K. J. *et al.* Cloud-based simulations on Google Exacycle reveal ligand modulation of GPCR activation pathways. *Nature Chemistry 2013 6:1* **6**, 15–21 (2013).

196. Li, J., Jonsson, A. L., Beuming, T., Shelley, J. C. & Voth, G. A. Ligand-dependent activation and deactivation of the human adenosine A 2A receptor. *J Am Chem Soc* **135**, 8749–8759 (2013).

197. Provasi, D., Artacho, M. C., Negri, A., Mobarec, J. C. & Filizola, M. Ligand-Induced Modulation of the Free-Energy Landscape of G Protein-Coupled Receptors Explored by Adaptive Biasing Techniques. *PLoS Comput Biol* **7**, e1002193 (2011).

198. Provasi, D. & Filizola, M. Putative Active States of a Prototypic G-Protein-Coupled Receptor from Biased Molecular Dynamics. *Biophys J* **98**, 2347–2355 (2010).

199. Sena, D. M., Cong, X., Giorgetti, A., Kless, A. & Carloni, P. Structural heterogeneity of the μ-opioid receptor's conformational ensemble in the apo state. *Scientific Reports 2017 7:1* **7**, 1–7 (2017).

200. Fleetwood, O., Matricon, P., Carlsson, J. & Delemotte, L. Energy Landscapes Reveal Agonist Control of G Protein-Coupled Receptor Activation via Microswitches. *Biochemistry* **59**, 880–891 (2020).

201. de Waal, P. W. *et al.* Molecular mechanisms of fentanyl mediated β-arrestin biased signaling. *PLoS Comput Biol* **16**, e1007394 (2020).

202. Fleetwood, O., Carlsson, J. & Delemotte, L. Identification of ligand-specific g-protein coupled receptor states and prediction of downstream efficacy via data-driven modeling. *Elife* **10**, 1–46 (2021).

203. Söldner, C. A., Horn, A. H. C. & Sticht, H. A Metadynamics-Based Protocol for the Determination of GPCR-Ligand Binding Modes. *International Journal of Molecular Sciences 2019, Vol. 20, Page 1970* **20**, 1970 (2019).

204. Xu, X. *et al.* Binding pathway determines norepinephrine selectivity for the human β1AR over β2AR. *Cell Research 2020 31:5* **31**, 569–579 (2020).

205. de Waal, P. W. *et al.* Molecular mechanisms of fentanyl mediated β-arrestin biased signaling. *PLoS Comput Biol* **16**, e1007394 (2020).

206. Cullum, V. A., Farmer, J. B., Jack, D. & Levy, G. P. Salbutamol: a new, selective β-adrenoceptive receptor stimulant. *Br J Pharmacol* **35**, 141–151 (1969).

207. Hutchings, C. J. A review of antibody-based therapeutics targeting G protein-coupled receptors: an update. *Expert Opin Biol Ther* **20**, 925–935 (2020).

208. Leach, K. *et al.* Molecular mechanisms of action and in vivo validation of an M4 muscarinic acetylcholine receptor allosteric modulator with potential antipsychotic properties. *Neuropsychopharmacology* **35**, 855–869 (2010).

209. Conn, P. J., Jones, C. K. & Lindsley, C. W. Subtype-selective allosteric modulators of muscarinic receptors for the treatment of CNS disorders. *Trends Pharmacol Sci* **30**, 148–155 (2009).

210. Pinkerton, A. B. *et al.* Discovery of β-Arrestin Biased, Orally Bioavailable and CNS Penetrant Neurotensin Receptor 1 (NTR1) Allosteric Modulators. *J Med Chem* **62**, 8357 (2019).

211. Slosky, L. M. *et al.* β-Arrestin-Biased Allosteric Modulator of NTSR1 Selectively Attenuates Addictive Behaviors. *Cell* **181**, 1364-1379.e14 (2020).

212. Lipinski, C. A., Lombardo, F., Dominy, B. W. & Feeney, P. J. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv Drug Deliv Rev* **46**, 3–26 (2001).

213. Wang, T. C. *et al.* Discovery of diarylurea P2Y1 antagonists with improved aqueous solubility. *Bioorg Med Chem Lett* **23**, 3239–3243 (2013).

214. Yukawa, T. & Naven, R. Utility of Physicochemical Properties for the Prediction of Toxicological Outcomes: Takeda Perspective. *ACS Med Chem Lett* **11**, 203–209 (2020).

215. Leeson, P. D. & Springthorpe, B. The influence of drug-like concepts on decision-making in medicinal chemistry. *Nature Reviews Drug Discovery 2007 6:11* **6**, 881–890 (2007).

216. Young, R. J. Today's drug discovery and the shadow of the rule of 5. *Expert Opin Drug Discov* **18**, 965–972 (2023).

217. Miller, R. R. *et al.* Integrating the Impact of Lipophilicity on Potency and Pharmacokinetic Parameters Enables the Use of Diverse Chemical Space during Small Molecule Drug Optimization. *J Med Chem* **63**, 12156–12170 (2020).

218. Vauquelin, G. & Packeu, A. Ligands, their receptors and … plasma membranes. *Mol Cell Endocrinol* **311**, 1–10 (2009).

219. Vauquelin, G. On the 'micro'-pharmacodynamic and pharmacokinetic mechanisms that contribute to long-lasting drug action. *Expert Opin Drug Discov* **10**, 1085–1098 (2015).

220. Szlenk, C. T., Gc, J. B. & Natesan, S. Does the Lipid Bilayer Orchestrate Access and Binding of Ligands to Transmembrane Orthosteric/Allosteric Sites of G Protein-Coupled Receptors? *Mol Pharmacol* **96**, 527 (2019).

221. Payandeh, J. & Volgraf, M. Ligand binding at the protein–lipid interface: strategic considerations for drug design. *Nature Reviews Drug Discovery 2021 20:9* **20**, 710–722 (2021).

222. Gunaydin, H. *et al.* Strategy for Extending Half-life in Drug Design and Its Significance. *ACS Med Chem Lett* **9**, 528–533 (2018).

223. Baylon, J. L., Lenov, I. L., Sligar, S. G. & Tajkhorshid, E. Characterizing the membrane-bound state of cytochrome P450 3A4: Structure, depth of insertion, and orientation. *J Am Chem Soc* **135**, 8542–8551 (2013).

224. Shukla, S. J. *et al.* Identification of Clinically Used Drugs That Activate Pregnane X Receptors. *Drug Metabolism and Disposition* **39**, 151–159 (2011).

225. Masureel, M. *et al.* Structural insights into binding specificity, efficacy and bias of a β2AR partial agonist. *Nature Chemical Biology 2018 14:11* **14**, 1059–1066 (2018).

226. Sykes, D. A. *et al.* Observed drug-receptor association rates are governed by membrane affinity: The importance of establishing 'micro-pharmacokinetic/pharmacodynamic relationships' at the β2-adrenoceptor. *Mol Pharmacol* **85**, 608–617 (2014).

227. Hadida, S. *et al.* Discovery of N -(2,4-Di- tert -butyl-5-hydroxyphenyl)-4-oxo-1,4-dihydroquinoline-3-carboxamide (VX-770, Ivacaftor), a potent and orally bioavailable CFTR potentiator. *J Med Chem* **57**, 9776–9795 (2014).

228. Barton, P. & Riley, R. J. A new paradigm for navigating compound property related drug attrition. *Drug Discov Today* **21**, 72–81 (2016).

229. Yukawa, T. & Naven, R. Utility of Physicochemical Properties for the Prediction of Toxicological Outcomes: Takeda Perspective. *ACS Med Chem Lett* **11**, 203–209 (2020).

230. Jamieson, C., Moir, E. M., Rankovic, Z. & Wishart, G. Medicinal chemistry of hERG optimizations: Highlights and hang-ups. *J Med Chem* **49**, 5029–5046 (2006).

231. Ploemen, J.-P. H. T. M. *et al.* Use of physicochemical calculation of pKa and CLogP to predict phospholipidosis-inducing potential: a case study with structurally related piperazines. *Exp Toxicol Pathol* **55**, 347–355 (2004).

232. Tsopelas, F., Giaginis, C. & Tsantili-Kakoulidou, A. Lipophilicity and biomimetic properties to support drug discovery. *Expert Opin Drug Discov* **12**, 885–896 (2017).

233. Tsopelas, F., Vallianatou, T. & Tsantili-Kakoulidou, A. Advances in immobilized artificial membrane (IAM) chromatography for novel drug discovery. *Expert Opin Drug Discov* **11**, 473–488 (2016).

234. Mason, R. P., Moisey, D. M. & Shajenko, L. Cholesterol alters the binding of Ca2+ channel blockers to the membrane lipid bilayer. *Mol Pharmacol* **41**, (1992).

235. Pavel, M. A., Petersen, E. N., Wang, H., Lerner, R. A. & Hansen, S. B. Studies on the mechanism of general anesthesia. *Proc Natl Acad Sci U S A* **117**, 13757–13766 (2020).

236. Myszka, D. G. *et al.* The ABRF-MIRG'02 study: Assembly state, thermodynamic, and kinetic analysis of an enzyme/inhibitor interaction. *Journal of Biomolecular Techniques* **14**, 247–269 (2003).

237. Tellinghuisen, J. & Chodera, J. D. Systematic errors in isothermal titration calorimetry: Concentrations and baselines. *Anal Biochem* **414**, 297–299 (2011).

238. Nisius, L., Rogowski, M., Vangelista, L. & Grzesiek, S. Large-scale expression and purification of the major HIV-1 coreceptor CCR5 and characterization of its interaction with RANTES. *Protein Expr Purif* **61**, 155–162 (2008).

239. Rajarathnam, K. & Rösgen, J. Isothermal Titration Calorimetry of Membrane Proteins – Progress and Challenges. *Biochim Biophys Acta* **1838**, 69–77 (2014).

240. Flanagan, C. A. GPCR-radioligand binding assays. *Methods Cell Biol* **132**, 191–215 (2016).

241. Mendel, C. M. & Mendel, D. B. 'Non-specific' binding. The problem, and a solution. *Biochemical Journal* **228**, 269–272 (1985).

242. Tang, L. *et al.* Structural basis for inhibition of a voltage-gated Ca2+ channel by Ca2+ antagonist drugs. *Nature 2016 537:7618* **537**, 117–121 (2016).

243. Gao, Y., Cao, E., Julius, D. & Cheng, Y. TRPV1 structures in nanodiscs reveal mechanisms of ligand and lipid action. *Nature 2016 534:7607* **534**, 347–351 (2016).

244. Souza, P. C. T. *et al.* Martini 3: a general purpose force field for coarse-grained molecular dynamics. *Nature Methods 2021 18:4* **18**, 382–388 (2021).

245. Souza, P. C. T. *et al.* Protein–ligand binding with the coarse-grained Martini model. *Nature Communications 2020 11:1* **11**, 1–11 (2020).

246. Borges-Araújo, L. *et al.* Martini 3 Coarse-Grained Force Field for cholesterol. (2023) doi:10.26434/CHEMRXIV-2023-LH7BQ.

247. Mackman, N. Triggers, targets and treatments for thrombosis. *Nature 2008 451:7181* **451**, 914–918 (2008).

248. Cattaneo, M. P2Y12 receptors: structure and function. *Journal of Thrombosis and Haemostasis* **13**, S10–S16 (2015).

249. Foster, C. J. *et al.* Molecular identification and characterization of the platelet ADP receptor targeted by thienopyridine antithrombotic drugs. *Journal of Clinical Investigation* **107**, 1591 (2001).

250. Cattaneo, M. The platelet $P2Y_{12}$ receptor for adenosine diphosphate: congenital and drug-induced defects. *Blood* **117**, 2102–2112 (2011).

251. Léon, C. *et al.* Defective platelet aggregation and increased resistance to thrombosis in purinergic P2Y1 receptor-null mice. *Journal of Clinical Investigation* **104**, 1731–1737 (1999).

252. Hechler, B. *et al.* MRS2500 [2-iodo-N6-methyl-(N)-methanocarba-2'-deoxyadenosine-3',5'-bisphosphate], a potent, selective, and stable antagonist of the platelet P2Y1 receptor with strong antithrombotic activity in mice. *J Pharmacol Exp Ther* **316**, 556–563 (2006).

253. Chao, H. *et al.* Discovery of 2-(phenoxypyridine)-3-phenylureas as small molecule P2Y 1 antagonists. *J Med Chem* **56**, 1704–1714 (2013).

254. Zhang, D. *et al.* Two disparate ligand binding sites in the human P2Y1 receptor. *Nature* **520**, 317 (2015).

255. Jacobson, K. A. *et al.* Update of P2Y receptor pharmacology: IUPHAR Review 27. *Br J Pharmacol* **177**, 2413–2433 (2020).

256. Yuan, S. *et al.* The Molecular Mechanism of P2Y1 Receptor Activation. *Angew Chem Int Ed Engl* **55**, 10331 (2016).

257. Yuan, X., Raniolo, S., Limongelli, V. & Xu, Y. The Molecular Mechanism Underlying Ligand Binding to the Membrane-Embedded Site of a G-Protein-Coupled Receptor. *J Chem Theory Comput* **14**, 2761–2770 (2018).

258. Dickson, C. J., Hornak, V. & Duca, J. S. Relative Binding Free-Energy Calculations at Lipid-Exposed Sites: Deciphering Hot Spots. *J Chem Inf Model* **61**, 5923–5930 (2021).

259. ELRICK, H., HLAD, C. J., ARAI, Y. & SMITH, A. THE INTERACTION OF GLUCAGON AND INSULIN ON BLOOD GLUCOSE. *Journal of Clinical Investigation* **35**, 757 (1956).

260. GREBEN, S. E. & SCHULMAN, J. L. The Effect of Glucagon on the Blood Glucose Level and the Clinical State in the Presence of Marked Insulin Hypoglycemia. *Journal of Clinical Investigation* **36**, 74 (1957).

261. Maccuish, A. C., Munro, J. F. & Duncan, L. J. P. Treatment of hypoglycaemic coma with glucagon, intravenous dextrose, and mannitol infusion in a hundred diabetics. *Lancet* **2**, 946–949 (1970).

262. Stern, J. H. *et al.* Obesity dysregulates fasting-induced changes in glucagon secretion. *J Endocrinol* **243**, 149 (2019).

263. Petersen, K. F. & Sullivan, J. T. Effects of a novel glucagon receptor antagonist (Bay 27-9955) on glucagon-stimulated glucose production in humans. *Diabetologia* **44**, 2018–2024 (2001).

264. Christensen, M., Bagger, J. I., Vilsbøll, T. & Knop, F. K. The Alpha-Cell as Target for Type 2 Diabetes Therapy. *Rev Diabet Stud* **8**, 369 (2011).

265. Lee, Y., Wang, M. Y., Du, X. Q., Charron, M. J. & Unger, R. H. Glucagon receptor knockout prevents insulin-deficient type 1 diabetes in mice. *Diabetes* **60**, 391–397 (2011).

266. Guan, H. P. *et al.* Glucagon receptor antagonism induces increased cholesterol absorption. *J Lipid Res* **56**, 2183 (2015).

267. Cheng, C., Jabri, S., Taoka, B. M. & Sinz, C. J. Small molecule glucagon receptor antagonists: an updated patent review (2015-2019). *Expert Opin Ther Pat* **30**, 509–526 (2020).

268. Al-Massadi, O., Fernø, J., Diéguez, C., Nogueiras, R. & Quiñones, M. Glucagon Control on Food Intake and Energy Balance. *Int J Mol Sci* **20**, (2019).

269. Okamoto, H. *et al.* Glucagon Receptor Blockade With a Human Antibody Normalizes Blood Glucose in Diabetic Mice and Monkeys. *Endocrinology* **156**, 2781–2794 (2015).

270. Xiong, Y. *et al.* Discovery of a novel glucagon receptor antagonist N-[(4-{(1 S)-1-[3-(3, 5-dichlorophenyl)-5-(6-methoxynaphthalen-2-yl)-1 H -pyrazol-1-yl]ethyl}phenyl) carbonyl]-β-alanine (MK-0893) for the treatment of type II diabetes. *J Med Chem* **55**, 6137–6148 (2012).

271. Jazayeri, A. *et al.* Extra-helical binding site of a glucagon receptor antagonist. *Nature 2016 533:7602* **533**, 274–277 (2016).

272. Li, Y., Sun, J., Li, D. & Lin, J. Activation and conformational dynamics of a class B G-protein-coupled glucagon receptor. *Physical Chemistry Chemical Physics* **18**, 12642–12650 (2016).

273. Mattedi, G., Acosta-Gutí Errez, S., Clark, T. & Gervasio, F. L. A combined activation mechanism for the glucagon receptor. **117**, 15414–15422 (2020).

274. Ansell, T. B., Song, W. & Sansom, M. S. P. The Glycosphingolipid GM3 Modulates Conformational Dynamics of the Glucagon Receptor. *Biophys J* **119**, 300 (2020).

275. Song, W., Yen, H. Y., Robinson, C. V. & Sansom, M. S. P. State-dependent Lipid Interactions with the A2a Receptor Revealed by MD Simulations Using In Vivo-Mimetic Membranes. *Structure(London, England:1993)* **27**, 392 (2019).

276. Kjølbye, L. R. *et al.* Lipid Modulation of a Class B GPCR: Elucidating the Modulatory Role of PI(4,5)P2Lipids. *J Chem Inf Model* **62**, 6788–6802 (2022).

277. Stevens, J. A. *et al.* Molecular dynamics simulation of an entire cell. *Front Chem* **11**, 1106495 (2023).

278. Leach, A. R. *Molecular Modelling: Principles and Applications.* (2001).

279. Born, M. & Oppenheimer, J. R. On the Quantum Theory of Molecules. (1927).

280. Artree, D. R. H. & Artree (, W. H. Self-consistent field, with exchange, for beryllium. *Proc R Soc Lond A Math Phys Sci* **150**, 9–33 (1935).

281. Hartree, D. R. The Wave Mechanics of an Atom with a Non-Coulomb Central Field. Part II. Some Results and Discussion. *Mathematical Proceedings of the Cambridge Philosophical Society* **24**, 111–132 (1928).

282. Hohenberg, P. & Kohn, W. Inhomogeneous electron gas. *Physical Review* **136**, B864 (1964).

283. Kohn, W. & Sham, L. J. Self-consistent equations including exchange and correlation effects. *Physical Review* **140**, A1133 (1965).

284. Unke, O. T. *et al.* Machine Learning Force Fields. *Chem Rev* **121**, 10142–10186 (2021).

285. Stanton, J. F. Why CCSD(T) works: a different perspective. *Chem Phys Lett* **281**, 130–134 (1997).

286. Warshel, A. & Levitt, M. Theoretical studies of enzymic reactions: Dielectric, electrostatic and steric stabilization of the carbonium ion in the reaction of lysozyme. *J Mol Biol* **103**, 227–249 (1976).

287. Van Der Kamp, M. W. & Mulholland, A. J. Combined quantum mechanics/molecular mechanics (QM/MM) methods in computational enzymology. *Biochemistry* **52**, 2708–2728 (2013).

288. Lindorff-Larsen, K. *et al.* Improved side-chain torsion potentials for the Amber ff99SB protein force field. *Proteins* **78**, 1950 (2010).

289. Oostenbrink, C., Villa, A., Mark, A. E. & Van Gunsteren, W. F. A biomolecular force field based on the free enthalpy of hydration and solvation: the GROMOS force-field parameter sets 53A5 and 53A6. *J Comput Chem* **25**, 1656–1676 (2004).

290. Huang, J. *et al.* CHARMM36m: An Improved Force Field for Folded and Intrinsically Disordered Proteins. *Nat Methods* **14**, 71 (2017).

291. Wang, Y. *et al.* End-to-end differentiable construction of molecular mechanics force fields. *Chem Sci* **13**, 12016–12033 (2022).

292. Mobley, D. L. *et al.* Escaping atom types in force fields using direct chemical perception. *J Chem Theory Comput* **14**, 6076 (2018).

293. Mobley, D. L. *et al.* Escaping Atom Types in Force Fields Using Direct Chemical Perception. *J Chem Theory Comput* **14**, 6076–6092 (2018).

294. Horton, J. T. *et al.* Open Force Field BespokeFit: Automating Bespoke Torsion Parametrization at Scale. *J Chem Inf Model* **62**, 5622–5633 (2022).

295. Qiu, Y. *et al.* Driving torsion scans with wavefront propagation. *J Chem Phys* **152**, 244116 (2020).

296. Lennard-Jones, J. E. Cohesion. *Proceedings of the Physical Society* **43**, 461 (1931).

297. Galántai, A. The theory of Newton's method. *J Comput Appl Math* **124**, 25–44 (2000).

298. Nakano, A. Parallel multilevel preconditioned conjugate-gradient approach to variable-charge molecular dynamics. *Comput Phys Commun* **104**, 59–69 (1997).

299. Reichert, P. The Ergodic Hypothesis: A Typicality Statement. (2023).

300. Miyamoto, S. & Kollman, P. A. Settle: An analytical version of the SHAKE and RATTLE algorithm for rigid water models. *J Comput Chem* **13**, 952–962 (1992).

301. Hess, B. P-LINCS: A Parallel Linear Constraint Solver for Molecular Simulation. *J Chem Theory Comput* **4**, 116–122 (2007).

302. Balusek, C. *et al.* Accelerating Membrane Simulations with Hydrogen Mass Repartitioning. *J Chem Theory Comput* (2019) doi:10.1021/acs.jctc.9b00160.

303. Swope, W. C., Andersen, H. C., Berens, P. H. & Wilson, K. R. A computer simulation method for the calculation of equilibrium constants for the formation of physical clusters of molecules: Application to small water clusters. *J Chem Phys* **76**, 637–649 (1982).

304. Skeel, R. D. Variable step size destabilizes the Störmer/leapfrog/verlet method. *BIT* **33**, 172–175 (1993).

305. Evans, D. J. & Holian, B. L. The Nose–Hoover thermostat. *J Chem Phys* **83**, 4069–4074 (1985).

306. Berendsen, H. J. C., Postma, J. P. M., Van Gunsteren, W. F., Dinola, A. & Haak, J. R. Molecular dynamics with coupling to an external bath. *J Chem Phys* (1984) doi:10.1063/1.448118.

307. Andersen, H. C. Molecular dynamics simulations at constant pressure and/or temperature. *J Chem Phys* **72**, 2384–2393 (1980).

308. Parrinello, M. & Rahman, A. Crystal Structure and Pair Potentials: A Molecular-Dynamics Study. *Phys Rev Lett* **45**, 1196 (1980).

309. Shirts, M. R., Mobley, D. L., Chodera, J. D. & Pande, V. S. Accurate and efficient corrections for missing dispersion interactions in molecular simulations. *Journal of Physical Chemistry B* **111**, 13052–13063 (2007).

310. Essmann, U. *et al.* A Smooth Particle Mesh Ewald Method. *Article in The Journal of Chemical Physics* (1995) doi:10.1063/1.470117.

311. Swendsen, R. H. & Wang, J. S. Replica Monte Carlo Simulation of Spin-Glasses. *Phys Rev Lett* **57**, 2607 (1986).

312. Rocklin, G. J., Mobley, D. L., Dill, K. A. & Hünenberger, P. H. Calculating the binding free energies of charged species based on explicit-solvent simulations employing lattice-sum methods: an accurate correction scheme for electrostatic finite-size effects. *J Chem Phys* **139**, (2013).

313. Wu, Z. & Biggin, P. C. Correction Schemes for Absolute Binding Free Energies Involving Lipid Bilayers. *J Chem Theory Comput* **18**, 2657–2672 (2022).

314. Bussi, G. Molecular Physics An International Journal at the Interface Between Chemistry and Physics Hamiltonian replica exchange in GROMACS: a flexible implementation INVITED ARTICLE Hamiltonian replica exchange in GROMACS: a flexible implementation. *Mol Phys* **112**, (2014).

315. Salari, R., Joseph, T., Lohia, R., Hénin, J. & Brannigan, G. A streamlined, general approach for computing ligand binding free energies and its application to GPCR-bound cholesterol. *J Chem Theory Comput* **14**, 6560 (2018).

316. Corey, R. A., Vickery, O. N., Sansom, M. S. P. & Stansfeld, P. J. Insights into Membrane Protein-Lipid Interactions from Free Energy Calculations. *J Chem Theory Comput* **15**, 5727–5736 (2019).

317. Zwanzig, R. W. High-Temperature Equation of State by a Perturbation Method. I. Nonpolar Gases. *J Chem Phys* **22**, 1420–1426 (1954).

318. Bennett, C. H. Efficient estimation of free energy differences from Monte Carlo data. *J Comput Phys* **22**, 245–268 (1976).

319. Shirts, M. R. & Chodera, J. D. Statistically optimal analysis of samples from multiple equilibrium states. *Journal of Chemical Physics* (2008) doi:10.1063/1.2978177.

320. Jorge, M., Garrido, N. M., Queimada, A. J., Economou, I. G. & MacEdo, E. A. Effect of the integration method on the accuracy and computational efficiency of free energy calculations using thermodynamic integration. *J Chem Theory Comput* **6**, 1018–1027 (2010).

321. Feng, M., Heinzelmann, G. & Gilson, M. K. Absolute binding free energy calculations improve enrichment of actives in virtual compound screening. *Scientific Reports 2022 12:1* **12**, 1–11 (2022).

322. Fu, H., Zhou, Y., Jing, X., Shao, X. & Cai, W. Meta-Analysis Reveals That Absolute Binding Free-Energy Calculations Approach Chemical Accuracy. *J Med Chem* **65**, 12970–12978 (2022).

323. Singh, N. & Li, W. Absolute Binding Free Energy Calculations for Highly Flexible Protein MDM2 and Its Inhibitors. *Int J Mol Sci* **21**, 1–15 (2020).

324. Gapsys, V. *et al.* Accurate absolute free energies for ligand–protein binding based on non-equilibrium approaches. *Communications Chemistry 2021 4:1* **4**, 1–13 (2021).

325. Khalak, Y. *et al.* Alchemical absolute protein–ligand binding free energies for drug design. *Chem Sci* **12**, 13958–13971 (2021).

326. Papadourakis, M. *et al.* Alchemical Free Energy Calculations on Membrane-Associated Proteins. *J Chem Theory Comput* (2023) doi:10.1021/ACS.JCTC.3C00365.

327. Chen, W. *et al.* Enhancing Hit Discovery in Virtual Screening through Absolute Protein-Ligand Binding Free-Energy Calculations. *J Chem Inf Model* **63**, 3171–3185 (2023).

328. Baumann, H. M., Gapsys, V., De Groot, B. L. & Mobley, D. L. Challenges Encountered Applying Equilibrium and Nonequilibrium Binding Free Energy Calculations. *Journal of Physical Chemistry B* **125**, 4241–4261 (2021).

329. Jing, Z. *et al.* Polarizable force fields for biomolecular simulations: Recent advances and applications. *Annu Rev Biophys* **48**, 371 (2019).

330. Zhang, D. *et al.* Two disparate ligand-binding sites in the human P2Y1 receptor. *Nature 2015 520:7547* **520**, 317–321 (2015).

331. Chao, H. *et al.* Discovery of 2-(phenoxypyridine)-3-phenylureas as small molecule P2Y 1 antagonists. *J Med Chem* **56**, 1704–1714 (2013).

332. Yuan, X., Raniolo, S., Limongelli, V. & Xu, Y. The Molecular Mechanism Underlying Ligand Binding to the Membrane-Embedded Site of a G-Protein-Coupled Receptor. *J Chem Theory Comput* **14**, 2761–2770 (2018).

333. Mey, A. S. J. S. *et al.* Best Practices for Alchemical Free Energy Calculations [Article v1.0]. *Living J Comput Mol Sci* **2**, (2020).

334. Webb, B. & Sali, A. Comparative protein structure modeling using MODELLER. *Curr Protoc Bioinformatics* (2016) doi:10.1002/cpbi.3.

335. Waterhouse, A. *et al.* SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Res* **46**, W296–W303 (2018).

336. Benkert, P., Tosatto, S. C. E. & Schomburg, D. QMEAN: A comprehensive scoring function for model quality assessment. *Proteins: Structure, Function, and Bioinformatics* **71**, 261–277 (2008).

337. Pándy-Szekeres, G. *et al.* GPCRdb in 2023: state-specific structure models using AlphaFold2 and new ligand resources. *Nucleic Acids Res* **51**, D395–D402 (2023).

338. Abraham, M. J. *et al.* GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **1–2**, 19–25 (2015).

339. He, X., Man, V. H., Yang, W., Lee, T. S. & Wang, J. A fast and high-quality charge model for the next generation general AMBER force field. *Journal of Chemical Physics* **153**, 114502 (2020).

340. Shirts, M. R. *et al.* Lessons learned from comparing molecular dynamics engines on the SAMPL5 dataset. *J Comput Aided Mol Des* **31**, 147–161 (2017).

341. Boothroyd, S. *et al.* Development and Benchmarking of Open Force Field 2.0.0: The Sage Small Molecule Force Field. *J Chem Theory Comput* **19**, 3251–3275 (2023).

342. Schmidt, T. H. & Kandt, C. LAMBADA and InflateGRO2: Efficient membrane alignment and insertion of membrane proteins for molecular dynamics simulations. *J Chem Inf Model* (2012) doi:10.1021/ci3000453.

343. Jo, S., Kim, T., Iyer, V. G. & Im, W. CHARMM-GUI: A web-based graphical user interface for CHARMM. *J Comput Chem* **29**, 1859–1865 (2008).

344. Ermilova, I. & Lyubartsev, A. P. Extension of the slipids force field to polyunsaturated lipids. *Journal of Physical Chemistry B* (2016) doi:10.1021/acs.jpcb.6b05422.

345. Grote, F. & Lyubartsev, A. P. Optimization of Slipids Force Field Parameters Describing Headgroups of Phospholipids. *Journal of Physical Chemistry B* (2020) doi:10.1021/acs.jpcb.0c06386.

346. Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W. & Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J Chem Phys* (1983) doi:10.1063/1.445869.

347. Darden, T., York, D. & Pedersen, L. Particle mesh Ewald: An N·log(N) method for Ewald sums in large systems. *J Chem Phys* **98**, 10089–10092 (1993).

348. Essmann, U. *et al.* A smooth particle mesh Ewald method. *JChPh* **103**, 8577–8593 (1995).

349. Beutler, T. C., Mark, A. E., van Schaik, R. C., Gerber, P. R. & van Gunsteren, W. F. Avoiding singularities and numerical instabilities in free energy calculations based on molecular simulations. *Chem Phys Lett* **222**, 529–539 (1994).

350. Parrinello, M. & Rahman, A. Polymorphic transitions in single crystals: A new molecular dynamics method. *J Appl Phys* (1981) doi:10.1063/1.328693.

351. Gowers, R. *et al.* MDAnalysis: A Python Package for the Rapid Analysis of Molecular Dynamics Simulations. in *Proceedings of the 15th Python in Science Conference* (2016). doi:10.25080/majora-629e541a-00e.

352. Bouysset, C. & Fiorucci, S. ProLIF: a library to encode molecular interactions as fingerprints. *J Cheminform* **13**, 1–9 (2021).

353. DeLano, W. L. Pymol: An open-source molecular graphics tool. *CCP4 Newsletter On Protein Crystallography* (2002).

354. Antila, H. S., M. Ferreira, T., Ollila, O. H. S. & Miettinen, M. S. Using Open Data to Rapidly Benchmark Biomolecular Simulations: Phospholipid Conformational Dynamics. *J Chem Inf Model* (2021) doi:10.1021/acs.jcim.0c01299.

355. Jefferys, E., Sands, Z. A., Shi, J., Sansom, M. S. P. & Fowler, P. W. Alchembed: A Computational Method for Incorporating Multiple Proteins into Complex Lipid Geometries. *J Chem Theory Comput* **11**, 2743–2754 (2015).

356. Wolf, M. G., Hoefling, M., Aponte-SantamaríA, C., Grubmüller, H. & Groenhof, G. g_membed: Efficient insertion of a membrane protein into an equilibrated lipid bilayer with minimal perturbation. *J Comput Chem* **31**, 2169–2174 (2010).

357. Hua, Y. *et al.* FSDscore: An Effective Target-focused Scoring Criterion for Virtual Screening. *Mol Inform* **42**, 2200039 (2023).

358. Mobley, D. L., Chodera, J. D. & Dill, K. A. On the use of orientational restraints and symmetry corrections in alchemical free energy calculations. *J Chem Phys* **125**, 084902 (2006).

359. Ehrman, J. N. *et al.* Improving Small Molecule Force Fields by Identifying and Characterizing Small Molecules with Inconsistent Parameters. *J Comput Aided Mol Des* **35**, 271 (2021).

360. Aldeghi, M., de Groot, B. L. & Gapsys, V. Accurate Calculation of Free Energy Changes upon Amino Acid Mutation. *Methods Mol Biol* **1851**, 19–47 (2019).

361. Brankin, A. E. & Fowler, P. W. Predicting antibiotic resistance in complex protein targets using alchemical free energy methods. *J Comput Chem* **43**, 1771 (2022).

362. Eswar, N. *et al.* Comparative Protein Structure Modeling Using Modeller. *Current protocols in bioinformatics / editoral board, Andreas D. Baxevanis ... [et al.]* **0 5**, Unit (2006).

363. Wang, J., Deng, Y. & Roux, B. B. Absolute Binding Free Energy Calculations Using Molecular Dynamics Simulations with Restraining Potentials. (2006) doi:10.1529/biophysj.106.084301.

364. Valdés-Tresanco, M. S., Valdés-Tresanco, M. E., Valiente, P. A. & Moreno, E. Gmx_MMPBSA: A New Tool to Perform End-State Free Energy Calculations with GROMACS. *J Chem Theory Comput* **17**, 6281–6291 (2021).

365. Greene, D., Qi, R., Nguyen, R., Qiu, T. & Luo, R. Heterogeneous Dielectric Implicit Membrane Model for the Calculation of MMPBSA Binding Free Energies. *J Chem Inf Model* **59**, 3041–3056 (2019).

366. Rizzi, A. *et al.* The SAMPL6 SAMPLing challenge: assessing the reliability and efficiency of binding free energy calculations. *J Comput Aided Mol Des* **34**, 601–633 (2020).

367. Huai, Z., Yang, H., Li, X. & Sun, Z. SAMPL7 TrimerTrip host–guest binding affinities from extensive alchemical and end-point free energy calculations. *J Comput Aided Mol Des* **35**, 117–129 (2021).

368. Sridhar, A., Ross, G. A. & Biggin, P. C. Waterdock 2.0: Water placement prediction for Holo-structures with a pymol plugin. *PLoS One* **12**, e0172743 (2017).

369. de Graaf, C. *et al.* Extending the Structural View of Class B GPCRs. *Trends Biochem Sci* **42**, 946–960 (2017).

370. Liang, Y. L. *et al.* Cryo-EM structure of the active, Gs-protein complexed, human CGRP receptor. *Nature* **561**, 492–497 (2018).

371. Mattedi, G., Acosta-Gutiérrez, S., Clark, T. & Gervasio, F. L. A combined activation mechanism for the glucagon receptor. *Proc Natl Acad Sci U S A* **117**, 15414–15422 (2020).

372. solvation free energy with couple-intramol set to no - Redmine #3403 (#3403) · Issues · GROMACS / GROMACS · GitLab. https://gitlab.com/gromacs/gromacs/-/issues/3403.

373. Exclusions not treated correctly for the interactions beyond rcoulomb (#4051) · Issues · GROMACS / GROMACS · GitLab. https://gitlab.com/gromacs/gromacs/-/issues/4051.

374. Liang, Y. L. *et al.* Phase-plate cryo-EM structure of a class B GPCR–G-protein complex. *Nature 2017 546:7656* **546**, 118–123 (2017).

375. Zhang, H. *et al.* Structure of the full-length glucagon class B G-protein-coupled receptor. *Nature 2017 546:7657* **546**, 259–264 (2017).

376. Siu, F. Y. *et al.* Structure of the human glucagon class B G-protein-coupled receptor. *Nature 2013 499:7459* **499**, 444–449 (2013).

377. Zhang, H. *et al.* Structure of the glucagon receptor in complex with a glucagon analogue. *Nature 2018 553:7686* **553**, 106–110 (2018).

378. Noble, W. S. How does multiple testing correction work? *Nat Biotechnol* **27**, 1135 (2009).

379. Nunes, R. S., Vila-Viçosa, D. & Costa, P. J. Halogen Bonding: An Underestimated Player in Membrane-Ligand Interactions. *J Am Chem Soc* **143**, 4253–4267 (2021).

380. Valsson, O., Tiwary, P. & Parrinello, M. Enhancing Important Fluctuations: Rare Events and Metadynamics from a Conceptual Viewpoint. *https://doi.org/10.1146/annurev-physchem-040215-112229* **67**, 159–184 (2016).

381. Miao, Y., Feher, V. A. & McCammon, J. A. Gaussian Accelerated Molecular Dynamics: Unconstrained Enhanced Sampling and Free Energy Calculation. *J Chem Theory Comput* **11**, 3584–3595 (2015).

382. Mobley, D. L., Chodera, J. D. & Dill, K. A. Confine-and-release method: Obtaining correct binding free energies in the presence of protein conformational change. *J Chem Theory Comput* **3**, 1231–1235 (2007).

383. Morton, A. & Matthews, B. W. Advance ACS Abstracts. *Biochemistry* **34**, 8576–8588 (1995).

384. Wei, B. Q., Baase, W. A., Weaver, L. H., Matthews, B. W. & Shoichet, B. K. A Model Binding Site for Testing Scoring Functions in Molecular Docking. *J Mol Biol* **322**, 339–355 (2002).

385. Harpole, T. J. & Delemotte, L. Conformational landscapes of membrane proteins delineated by enhanced sampling molecular dynamics simulations. *Biochimica et Biophysica Acta (BBA) - Biomembranes* **1860**, 909–926 (2018).

386. Wilson, K. P. *et al.* Crystal Structure of p38 Mitogen-activated Protein Kinase. *Journal of Biological Chemistry* **271**, 27696–27700 (1996).

387. Kostidis, S., Sánchez-López, E. & Giera, M. Lipidomics analysis in drug discovery and development. *Curr Opin Chem Biol* **72**, 102256 (2023).

388. Peetla, C., Stine, A. & Labhasetwar, V. Biophysical interactions with model lipid membranes: applications in drug discovery and drug delivery. *Mol Pharm* **6**, 1264 (2009).

389. Song, W. *et al.* PyLipID: A Python Package for Analysis of Protein-Lipid Interactions from Molecular Dynamics Simulations. *J Chem Theory Comput* **18**, 1188–1201 (2022).

390. Ansell, T. B. *et al.* LipIDens: simulation assisted interpretation of lipid densities in cryo-EM structures of membrane proteins. *Nat Commun* **14**, 7774 (2023).

391. Bonomi, M. *et al.* PLUMED: a portable plugin for free-energy calculations with molecular dynamics. *Comput Phys Commun* **180**, 1961–1972 (2009).

392. De Jong, D. H., Baoukina, S., Ingólfsson, H. I. & Marrink, S. J. Martini straight: Boosting performance using a shorter cutoff and GPUs. *Comput Phys Commun* **199**, 1–7 (2016).

393. Monticelli, L. *et al.* The MARTINI Coarse-Grained Force Field: Extension to Proteins. *J Chem Theory Comput* **4**, 819–834 (2008).

394. Marrink, S. J., Risselada, H. J., Yefimov, S., Tieleman, D. P. & De Vries, A. H. The MARTINI force field: coarse grained model for biomolecular simulations. *J Phys Chem B* **111**, 7812–7824 (2007).

395. De Jong, D. H. *et al.* Improved Parameters for the Martini Coarse-Grained Protein Force Field. *J Chem Theory Comput* **9**, 687–697 (2013).

396. Wassenaar, T. A., Ingólfsson, H. I., Böckmann, R. A., Tieleman, D. P. & Marrink, S. J. Computational Lipidomics with insane: A Versatile Tool for Generating Custom Membranes for Molecular Simulations. *J Chem Theory Comput* **11**, 2144–2155 (2015).

397. Periole, X., Cavalli, M., Marrink, S. J. & Ceruso, M. A. Combining an Elastic Network With a Coarse-Grained Molecular Force Field: Structure, Dynamics, and Intermolecular Recognition. *J Chem Theory Comput* **5**, 2531–2543 (2009).

398. Melo, M. N., Ingólfsson, H. I. & Marrink, S. J. Parameters for Martini sterols and hopanoids based on a virtual-site description. *J Chem Phys* **143**, (2015).

399. Bussi, G., Donadio, D. & Parrinello, M. Canonical sampling through velocity rescaling. *J Chem Phys* **126**, (2007).

400. Hess, B., Bekker, H., Berendsen, H. J. C. & Fraaije, J. G. E. M. LINCS: A linear constraint solver for molecular simulations. *Wiley Online LibraryB Hess, H Bekker, HJC Berendsen, JGEM FraaijeJournal of computational chemistry, 1997•Wiley Online Library* **18**, 14631472 (1997).

401. Dawaliby, R. *et al.* Phosphatidylethanolamine Is a Key Regulator of Membrane Fluidity in Eukaryotic Cells. *Journal of Biological Chemistry* **291**, 3658–3667 (2016).

402. Xu, P. *et al.* Structural insights into the lipid and ligand regulation of serotonin receptors. *Nature* **592**, 469–473 (2021).

403. Qi, C., Minin, G. Di, Vercellino, I., Wutz, A. & Korkhov, V. M. Structural basis of sterol recognition by human hedgehog receptor PTCH1. *Sci Adv* **5**, (2019).

404. Tieleman, D. P. *et al.* Insights into lipid-protein interactions from computer simulations. *Biophys Rev* **13**, 1019 (2021).

405. Patrick, J. W. *et al.* Allostery revealed within lipid binding events to membrane proteins. *Proc Natl Acad Sci U S A* **115**, 2976–2981 (2018).

406. Stroh, K. S., Souza, P. C. T., Monticelli, L. & Risselada, H. J. CGCompiler: Automated Coarse-Grained Molecule Parametrization via Noise-Resistant Mixed-Variable Optimization. *J Chem Theory Comput* (2023)

doi:10.1021/ACS.JCTC.3C00637/ASSET/IMAGES/LARGE/CT3C00637_0008.JPEG
.

407. Nunes, R., Vila Viçosa, D. & Costa, P. J. Halogen Bonding: An Underestimated Player in Membrane–drug Interactions. (2020) doi:10.26434/CHEMRXIV.11894547.V1.

408. Huang, Y., Henderson, J. A. & Shen, J. Continuous constant pH Molecular Dynamics Simulations of Transmembrane Proteins. *Methods Mol Biol* **2302**, 275 (2021).

409. Bignucolo, O., Chipot, C., Kellenberger, S. & Roux, B. Galvani Offset Potential and Constant-pH Simulations of Membrane Proteins. *Journal of Physical Chemistry B* **126**, 6868–6877 (2022).

410. Wootten, D., Christopoulos, A. & Sexton, P. M. Emerging paradigms in GPCR allostery: Implications for drug discovery. *Nature Reviews Drug Discovery* Preprint at https://doi.org/10.1038/nrd4052 (2013).

411. Congreve, M., Oswald, C. & Marshall, F. H. Applying Structure-Based Drug Design Approaches to Allosteric Modulators of GPCRs. *Trends in Pharmacological Sciences* vol. 38 837–847 Preprint at https://doi.org/10.1016/j.tips.2017.05.010 (2017).

412. Hauser, A. S., Attwood, M. M., Rask-Andersen, M., Schiöth, H. B. & Gloriam, D. E. Trends in GPCR drug discovery: New agents, targets and indications. *Nat Rev Drug Discov* (2017) doi:10.1038/nrd.2017.178.

413. Dror, R. O. *et al.* Structural basis for modulation of a G-protein-coupled receptor by allosteric drugs. *Nature* (2013) doi:10.1038/nature12595.

414. Dror, R. O., Dirks, R. M., Grossman, J. P., Xu, H. & Shaw, D. E. Biomolecular simulation: A computational microscope for molecular biology. *Annual Review of Biophysics* Preprint at https://doi.org/10.1146/annurev-biophys-042910-155245 (2012).

415. Cheng, Y. Single-particle Cryo-EM at crystallographic resolution. *Cell* Preprint at https://doi.org/10.1016/j.cell.2015.03.049 (2015).

416. Lau, C., Hunter, M. J., Stewart, A., Perozo, E. & Vandenberg, J. I. Never at rest: insights into the conformational dynamics of ion channels from cryo-electron microscopy. *J Physiol* **596**, 1107 (2018).

417. Lundborg, M., Lidmar, J. & Hess, B. The accelerated weight histogram method for alchemical free energy calculations. *Journal of Chemical Physics* **154**, 204103 (2021).

418. Gaidai, I. *et al.* Molecular dynamics on quantum annealers. *Scientific Reports 2022 12:1* **12**, 1–10 (2022).

419. Ponder, J. W. *et al.* Current status of the AMOEBA polarizable force field. *Journal of Physical Chemistry B* **114**, 2549–2564 (2010).

420. Ford, M. C. & Ho, P. S. Computational Tools to Model Halogen Bonds in Medicinal Chemistry. *J Med Chem* (2016) doi:10.1021/acs.jmedchem.5b00997.