Computational approaches to fragment based screening

Kevin John Smith

A thesis submitted for the degree of Doctor of Philosophy (Ph.D.)

School of Biological Sciences

University of Essex

September 2016

Abstract

Polarization is an often-neglected term in molecular modelling, and this is particularly the case in docking. However, the growing interest in fragment-based drug design, coupled with the small size of fragments that makes them amenable to quantum mechanical treatment, has created new opportunities for including polarization, anisotropic electrostatics and realistic repulsion potentials in docking. We have shown that polarization implemented as induced charges can offer in the region of a 10-15% improvement in native docking results, as judged by the percentage of poses within a rather tight threshold of 0.5 or 1.0 Å, where accurate prediction of binding interactions, are more likely. This is a significant improvement given the quality of current commercial docking programs (such as Glide used here). This improvement is most apparent when the correct pose is known a priori, so that the extent of polarization is correctly modelled, and scoring is based on force-fields that do not scale the electrostatics. The introduction of specific active-site water molecules was shown to have a far greater effect than the polarization, probably because of the introduction of 3 additional full charges, rather than introduction of smaller charge perturbations. With active site waters, polarization is more likely to improve the docking when the water molecule is carefully orientated using quantum mechanical/molecular mechanics (QM/MM) methods. The placement of such water molecules is a matter of great current interest; we have shown that the water molecule can be placed with some degree of reliability simply by docking with the ligand present, provided that the water makes good hydrogen bonding interactions (these are the very conditions under which it is desirable to include the specific active-site water). Anisotropic electrostatics and exponential repulsion for rigid fragments was investigated using Orient and compared to QM/MM methods, all methods merited further research. The general hierarchy is that native docking using Glide (with polarization) > QM/MM (with MM polarization) > Orient-based methods. Thus, we expanded the Glide (with polarization) dataset to include more realistic crossdocking experiments on over 5000 structures. RMSD analysis resulted in many examples of clear improvement for including polarization.

Acknowledgements

First and foremost, I would like to thank my supervisor Prof Chris Reynolds for providing me with this amazing opportunity. This work would not have been possible without his guidance, encouragement, and incredible patience, and funding. Indeed, he also provided an excellent role model with his comprehensive knowledge; good nature; keen intelligence; integrity; professionalism; and ability to help with matters beyond the realm of this research. I would also like to thank his lovely wife Ruth Reynolds. Special thanks also go to both my internal and external examiners Dr Mike Hough, and Dr Peter Winn respectively, who gave up their time to apply a level of rigour to my thesis. Dr David Reha is another person who made this work possible, with his Perl method implementations, technical guidance, training and good humour. He assisted in my development in learning Perl, Linux, and computer algorithms for this research, and as a computational scientist.

I also thank all of the past and present members of the Biocomputing Laboratory at Essex, including: Dr Bruck Taddese, Dr Hani Mohammedali, Dr James McElhinney, Dr David Crompton, Dr David Woodlock, Dr Taniya Kayawatra, Dr Juan Carlos Mobarec, Gamze Karakullukcu, Joanna Farr, Sian Jordan, Nuradin Abdulla, Andreea Manole, Meenakshi Pardamwar, Vandana Desikan, and Himanshu Sharma for their help, support, discussions, and humour during my research. There are many other people at the University of Essex who have supported me throughout my PhD, particular thanks go to administrators Julie Snell, Emma Revill, and Sheila Pill. Then academics Prof Glyn Stanway, Dr Anthony Vickers, Prof Mike Wilson, Prof Nelson Fernandez, Dr Renate Leithauser, and Dr Selwa Alsam.

I'd like to thank previous inspirational teachers Mrs White, Mrs Bennet, Mr Morgan, Mr Arnold, Mrs Weston, Miss Shields, Mrs West, Andy Wright and Dr Rick Gosling. Saving the best for last, I would like to thank all my family including my amazing wife Hannah Smith for putting up with me throughout this PhD; my Parents for their love, help, and support; my brother Colin, my sister Fiona; and also my amazing children Harley, Evie, Poppy, and Willow, and my wonderful Nephews and Nieces for making my world a better place.

Contents

1 An introduction to Fragment based Drug Design

1.1 Background to Fragment Screening

The drug design process involves a number of stages, can take many years and can cost up to 5 billion dollars (ACS drug discovery symposium, 2014, www.acs.org). The initial stage involves a strategic decision on the target disease, followed by target identification and validation. To develop a drug, it is then necessary to develop appropriate assays, which may be target-based or phenotypic. Once appropriate assays are in place, it is possible to employ various forms of screening, either high-throughput screening or fragment-based screening, which may be complemented by virtual screening. From this screening, it is hoped to identify a lead from the various hits, that can be optimized for preclinical development, where issues such as toxicology and formulation are addressed. The final stage involves clinical trial. Many practitioners are wary of the pitfalls of high throughput screening and so have turned to fragment-based screening; the work in this thesis is most relevant to fragment-based screening.

1.1.1 Combinatorial Chemistry and HTS, a paradigm shift in library design

Innovations in medicinal chemistry such as Merrifield's method for the linear solid-phase synthesis of peptides, and the development of automated peptide synthesizers and eventually parallel and then split and pool synthesizers lead to ever more complex automation of the synthesis of peptides, with similar developments occurring in small non-peptide molecule synthesis. These advances in chemical synthesis and new High

throughput screening (HTS) technologies went hand in hand, each driving advances in the other field and together they constituted the industrialization of the pharmaceutical industry. The first medicinal Combinatorial Chemistry (CC) library, featured benzodiazepines and was reported by Bunin and Ellman in 1992 (Gershell and Atkins, 2003).

The industrialization of the drug industry has not been without its hiccups, and the resulting improvements in quality of compounds and data generated is analogous to the improvements in car quality after the introduction of assembly line automation (Macarron et al., 2011).

The focus in early CC libraries was on the impressive number of compounds that could be generated, but early libraries often used relatively simple chemistries. Solid-phase chemistry was in itself soon considered limiting with respect to the diversity of chemistries that could be performed (Mario Geysen et al., 2003). Additionally, the technical challenges of making traditional drug-like molecules would typically involve more than 20 chemical operations and only a few less purifications. Therefore medicinal chemists largely experimented with screening of many simple compounds, often sharing wells in the hope that sheer numbers would generate hits (Gershell and Atkins, 2003). These early hits, although numerous, could often not be validated in duplicate screens due to artefacts such as metabolite instability. The subsequent purification of hit compounds also led to largely biologically inactive substances (Rydzewski, 2008). This approach referred to as the 'massive library' model has now been largely abandoned in favour of small focused non-peptide libraries with elaborations on a small number of promising scaffolds (Gershell and Atkins, 2003).

These libraries have been given names such as 'lead-like' and 'scaffold' libraries which along with their lead optimization derivatives are perhaps where the highest value of CC is found. Retrospective studies of physicochemical properties that predict good oral bioavailability through consideration of Absorption, Distribution, Metabolism, Excretion and Toxicity (ADMET) and their structural activity relationships (SAR) have become common in the industry and are used to support the development of new design strategies (Rydzewski, 2008). For example, a study of the Comprehensive medicinal chemistry database (CMCD) at Vertex pharmaceuticals identified 5120 suitable drug compounds for SAR investigations. Analysis of these 5120 suitable drug compounds indicated that just 32 'SHAPES' or scaffolds consisting of 5 and 6 member rings described over 50% of them. All of the 32 'SHAPES'

comprised of either 1 ring system or 2 ring systems joined by a linker (Bemis and Murcko, 1996). This and other studies concluded that the significant structural contributions determining the oral bioavailability and other desirable physicochemical properties of a drug are overwhelmingly found in the scaffolds or core structure, and not the contributions of the side groups (Vieth et al., 2004, Boehm et al., 2000). Elaboration of scaffolds such as SHAPES have therefore become a key philosophy in both fragment and CC libraries (Fejzo et al., 1999).

Small lead-like and indeed fragment-like molecules are more likely to bind to a receptor because they have a better chance of finding a binding mode than larger drug-like molecules (Hann et al., 2001). Once a small polar molecule has been found with µM affinity, more focused libraries can rapidly improve on it. For example, the addition of a lipophilic side group can improve affinity and pharmacokinetic properties. In addition, as this strategy only requires one or two step elaboration from a series of promising templates it therefore does not require the use of complicated chemistries and split and pool techniques (Teague et al., 1999). High architectural diversity can also be achieved without a significant impact on molecular weight (MW) by having a high proportion of sp^3 saturated atoms within the core structures of molecules. This is an important consideration in the design of all libraries as high MW has a negative impact on what are deemed good 'drug-like' properties; this has been identified through retrospective studies on predicting good oral-bioavailability through ADMET (see 1.1.2). There is also a link between size (usually achieved predominantly through hydrophobic contributions) and permeability, solubility and promiscuity (Lovering et al., 2009, Lipinski et al., 1997).

Early robotic HTS screening systems were designed to operate on 96 well plates, often using fluorescence as readout. Steady improvements in speed, reliability, accuracy and miniaturization soon led to 384 and 1536 well plates respectively (Rydzewski, 2008). In demonstration of the new HTS technology, the first paper available on PubMed citing 'HTS' in the title appeared 1991; it took 6 years before ten such papers were published in a single

year (Macarron et al., 2011). Early HTS conference discussions and excitement focused on the technology innovations required to reach ultra HTS (100,000 compounds screened a day). In this scenario low affinity, low MW fragments were not considered an attractive starting point for screening as the binding events of fragments were largely undetectable using the new HTS screening technologies (Rydzewski, 2008). When ultra HTS was achieved attention instead focused on what else HTS could offer to the pre-clinical process other than lead identification, such as High Throughput target validation. The relatively low levels of initial success in finding desirable 'drug-like' compounds with good ADMET properties led to studies of the desirable physicochemical properties required, and the main focus of industry discussion turned to 'what targets should be screened?' (Walters and Namchuk, 2003). It was within this receptive forum that FBDD emerged as an important complimentary drug design strategy (Congreve et al., 2008).

1.1.2 From what makes a good drug to Fragment based drug Design

A retrospective study of the United States adopted name (USAN) library, consisting predominantly of drug candidates that have undergone preclinical and phase 1 safety evaluation, revealed that the physicochemical properties of HTS hits were different from pre HTS era hits. These changes emerged as an artefact of the screening processes used to detect hits. HTS compounds had higher molecular weight (MW) and higher LogP (log of the octanol-water partition coefficient measuring a drug's lipophilicity) and lower turbidimetric solubility than pre HTS leads (Lipinski et al., 1997).

Lipinski et al., (1997) studied four parameters universally associated with solubility and permeability, which along with potency form a triad of important drug absorption properties. The drug-like property cut-offs for these parameters are all close to or multiples of 5, so they were named the 'rule of 5' guidelines. These drug-like parameter limits are MW <500 Daltons (Da), logP <5, (or Moriguchi (M) logP <4.15), and <10 H-bond acceptors (such as O and N), and <5 H-bond donors (such as OH and NH). Approximately 10% of USAN compounds exceed one or more of the cut-offs, and therefore exceeding one or more 'rules of 5' cut-offs

reduces the chance of a drug candidate making it to clinical trials to less than 10% (Lipinski et al., 1997).

A more recent study of patent application drug candidates showed a positive correlation between promiscuity and increased logP. Excessively lipophilic molecules are more likely to bind to multiple targets (be promiscuous); such molecules are associated with toxicity and therefore increased attrition (Leeson and Springthorpe, 2007). We note that binding to multiple targets is required for the action of some drugs such as central nervous system (CNS) agents

A study of over 1100 drug candidates at GlaxoSmithKline for oral bioavailability measurements in rats also revealed that MW > 500 Da largely results in poor bioavailability. It was also found that having a high polar surface area (PSA - which is from the sum of polar atoms in a molecule) and a large number of rotatable bonds has a negative impact on permeation rate. In addition PSA<140 $Å^2$ was found to correlate better with increased permeation rate than cLogP (calculated LogP) (Veber et al., 2002). Cross-correlation studies showed a 0.96 R^2 between PSA and the number of O and N atoms (Vieth et al., 2004).

 The 'rule of 5' guidelines are widely accepted guidelines as to what constitutes drug-like physicochemical properties. Although useful, the 'rule of 5' guidelines have been reported as mostly unsuitable for identifying lead compounds from screening campaigns.   This is because most drug candidates complete a physicochemical property-changing lead optimization process necessary to improve target affinity to biologically functional levels. HTS screen studies revealed potent ligands (IC_{50}<0.1 μM) occur with small probability (<1:10^6), where most ligands (83%) occur in the low affinity range of (IC_{50}1- 10 μM) (Teague et al., 1999). The lead optimization process can yield a 100-1000 fold improvement in target affinity equating to 2.8 - 4.2 kcal mol^{-1} of binding energy (Erlanson et al., 2004). Studying the lead optimization of a set of small polar low affinity binders revealed increases in MW of 1-

200 Da and in clogP of 0.5-4.0. Therefore, optimizing drug-like lead molecules is likely to result in physicochemical properties that exceed the 'rule of 5' cut-offs and hence is likely to yield poor bioavailability. Instead, scaffolds with MW 100-350 Da and clogP = 1-3 are a highly superior starting points for lead-like libraries that can be elaborated on (Teague et al., 1999).

Fragment libraries have smaller sizes still (typically MW 120-250 Da or 8-18 non-hydrogen atoms). The properties of 40 diverse fragment hits were identified against a range of targets using High Throughput X-ray crystallographic screening technology; the results of this study indicated on average that a new mnemonic called the 'rule of 3' would form appropriate guidelines as to what is Fragment-like. These parameters were MW <300, ≤ 3 H-bond donors, ≤ 3 H-bond acceptors, cLogP≤3 and in addition the same study suggested that keeping the number of rotatable bonds ≤ 3 and the average PSA ≤ 60 $Å^2$, might provide further guidance on fragment selection (Rees et al., 2004).

The 'rule of 3' is a widely accepted view of what is fragment-like, so by definition fragments have good physicochemical properties (Congreve et al., 2008). The technical difficulties of working with fragments have been largely overcome in the last 10-15 years, although FBDD is still heavily reliant on biophysical methods for detection. The idea that smaller is better when designing screening libraries has become increasingly popular drug design strategy in a large number of both academic and industry groups (Congreve et al., 2008).

1.2 The concepts of FBDD

1.2.1 Fragments are more than the sum of their parts

William Jencks was an early pioneer of the theoretical basis of FBDD. Jencks, when studying rates of reactions between compounds, showed that ligand interactions, can be viewed as a combination of two or more individual binding epitopes. This view of the combination of individual binding epitopes later became synonymous with the term and concept of 'fragments' (Rees et al., 2004). Small fragments often bind weakly or negligibly in their

individual moieties, but when linked together they can become high affinity binders. This is because the intrinsic binding energy of connected fragments can be approximated as additive once cost in binding entropy has been considered. Linked fragments effectively pay only one conformational entropy loss penalty when they bind to their individual protein target, (Jencks, 1981). A similar conformational entropy loss penalty phenomenon has been observed for example in the binding of chelating agents to metals and in the non-polar interactions and hydrogen bonding that holds a protein in its native structure when the groups are connected in a peptide chain. However, it is also possible to introduce conformational strain from suboptimal binding geometries. Therefore, this approximation is most accurate when the binding epitopes are small molecules and they match their moieties optimally (Rees et al., 2004).

1.2.2 Receptor Match

Andrews attempted to quantify the average strength of individual functional group contributions of drug binding. Data was collected on the binding constants and structural components of 200 drugs and enzyme inhibitors. The data was used to calculate the average binding energy contributions of the ten most commonly occurring functional groups. Each of the 200 drug compounds, in which the functional groups occurred, was assessed for the number of these present. An equation summed each of a drug's present functional groups average binding energy contributions, then entropy penalties were added to these additive totals giving an additive average expected binding energy for each drug. The contributions from each functional group in diminishing order were charged, polar, and non-polar groups. The binding energy penalties incurred by each drug were -14 kcal mol^{-1} for rotational/translational entropy loss and -0.7 kcal mol^{-1} entropy loss per degree of internal conformational freedom. The observed binding energies for the 200 drugs were also calculated from their binding constants. The average binding energy expected from each of these drugs' functional group contributions were then compared to the calculated observed binding energies to give an indication of how well each of the drugs matched it's receptor

energetically; regression analysis and substitution of entropy values were used to test the appropriateness of the method (Andrews et al., 1984).

Receptor match was a precursor to ligand efficiency (discussed in 1.2.4) and relates to how many of a ligand's substituents are involved directly in binding. Exceptionally well matching drugs for example were identified such as biotin and camphor, that have binding energies that are significantly above the calculated average functional group binding energy contribution, and poorly matching drugs were identified such as methotrexate and ouabain having binding energies significantly below the calculated average expectations. methotrexate's excellent K_i was noted, even though it produced the worst energy fit. The reason suggested was that the bound form was likely a high energy conformation, with a portion of the binding energy being used to achieve this form. The energy deviations between the calculated averages and the observed averages of the 200 compounds were found to be up to +/- 16-17 kcal mol^{-1}. These results were as a whole in keeping with the expected observable energies as when optimally bound, a drug's geometry should be relatively fixed, resulting in negligible rotational and translational entropy loss (Andrews et al., 1984).

The 'Andrews analysis' method was again used to assess the quality of a ligands match to its receptor in 10^6 HTS screening studies using pKi. 22% of the compounds of average molecular complexity screened were expected to have an average pKi < 10 nM. An observed IC_{50} < 10 nM was seen in none of the compounds, indicating poor receptor matches. This supported their hypothesis that drug-like leads such as those found in HTS campaigns are likely to achieve µM affinity using many poorly optimized interactions (Teague et al., 1999). Optimization of these HTS hits is deemed 'difficult' as without detailed structural information, about the protein-ligand interaction a medicinal chemist does not know where to focus improvements see Figure 1.1 (Rees et al., 2004) .
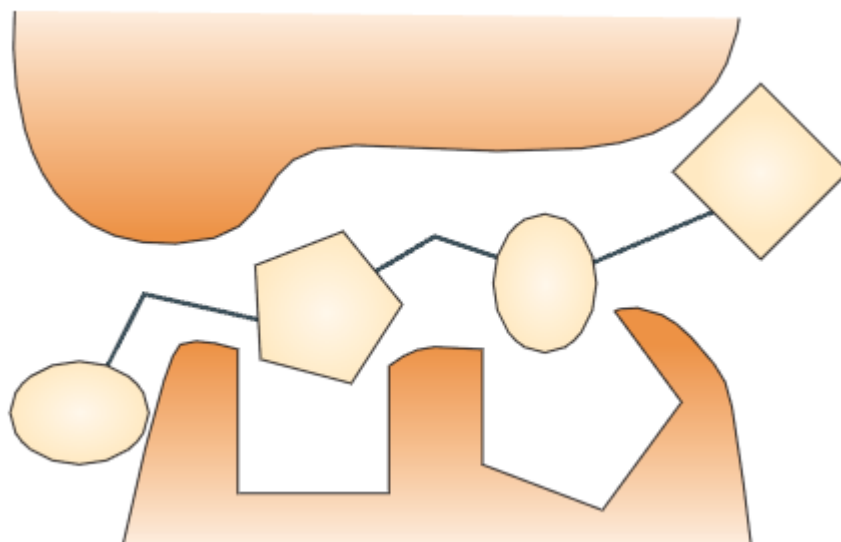
Figure 1.1. Schematic representing the many poorly optimized interactions of a drug-like lead**.** Taken from (Rees et al., 2004)

1.2.3   Fragments are efficient binders

Fragments may have a better chance of finding a binding mode, but must be extremely

efficient binders for their size. This is to counter the entropy loss involved in binding, which

has only a small relationship to MW.   A large proportion of a fragments atoms must be

involved in the binding interactions, thus requiring high binding energies per unit of mass,

and therefore a good receptor match is also required (see Figure 1.2). This high per atom

interaction ratio for the fragment, must be carefully maintained, throughout the lead

optimization process until high affinity is reached (Chessari and Woodhead, 2009).
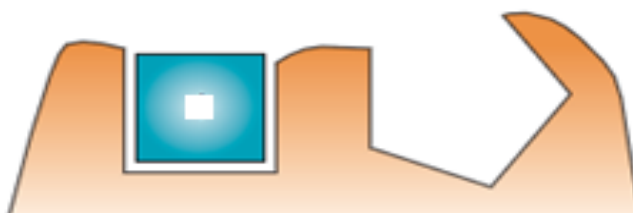
Figure 1.2. Schematic representing the good receptor match of a fragment. Adapted from (Rees et al., 2004)

The maximal affinity of ligands was studied by Kuntz et al., (1999) indicating that an increase

in potency of up to -1.5 kcal mol^{-1} per heavy (i.e. non-hydrogen) atom could be achieved in

the strongest non-metallic complexes from natural or synthetic ligands. The contributions per heavy atom followed a sharp slope until about 15 heavy atoms where contributions became increasingly negligible. In addition, a maximal affinity plateau for the tightest binding ligands picomolar (pM) was achieved at $\approx$-15 kcal mol^{-1} regardless of the heavy atom count (HAC) (Kuntz et al., 1999).

In a related study, Hadjuk (2006) looked at 18 highly optimized inhibitors from 15 lead discovery programs at Abbott Laboratories. Findings suggested successful lead optimization yielded an approximate uniform potency increase of -0.3 kcal mol^{-1} per heavy atom. Additionally, a series limit in affinity was approached at $\approx$-12 kcal mol^{-1}. These results were found to be in good agreement with Kuntz et al., (1999), as only subjective differences appeared in interpretation. If the Kuntz et al., (1999) series was viewed between 5 and 25 heavy atoms, a uniform slope of -0.27 kcal mol^{-1} per heavy atom emerged. In addition, apart from the tightest picomolar binders, the series also approached a maximal affinity limit at $\approx$-12 kcal mol^{-1} (see Figure 1.3). The results were also in agreement with Hopkins et al.'s, (2004) work on a concept called Ligand efficiency (*LE*) where a value of 0.3 kcal mol^{-1}kcal mol^{-1} per heavy atom or LE >0.3 had been suggested as the minimum binding efficiency per heavy atom for lead optimization. It was also found that to improve pK_D by 1 log$_{10}$ unit required on average an increase in MW of 64 Da and of clogP of 1 unit (Hajduk, 2006).
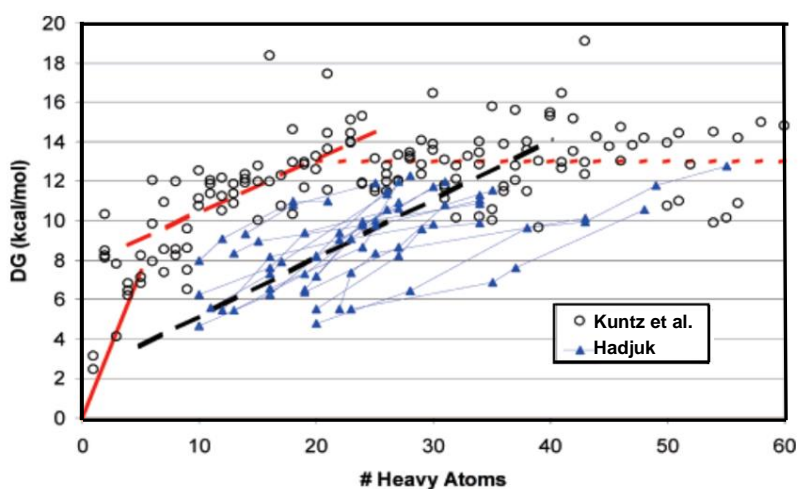
Figure 1.3. Plots of the HAC against the calculated Gibbs free energy of binding ($\varDelta G$ = $-RT\ In\ K_D$ at 300 K). Sharp slope through the zero intercept refers to the initial 1.5 kcal mol[-1] slope in Kuntz et al., (1999). Parallel slopes: Red 5-25 atom trend line using Kuntz et al., (1999) data; Black trend line from Hadjuk (2006) data. Taken from Hadjuk (2006)

1.2.4 Binding efficiency indices

There have been many binding efficiency indices developed in the last few years in response to the work of Kuntz et al., (1999) and the development of FBDD; some are mentioned below.

Ligand Efficiency *(LE)*

LE=-$\varDelta G$/ HAC ≈-RTln(IC50) / HAC

LE suggests a value ≥0.3 per heavy atom for lead optimization (Units of *LE* are kcal mol[-1]) (Congreve et al., 2008)

Group Efficiency *(GE)*

GE = -$\varDelta G$/HAC

GE looks at the group contributions to overall Gibbs free energy of binding (ΔG) by using a matched pair of compounds compared with a Free-Wilson analysis. The conversion of ΔG into *GE* is done identically to *LE* so that *GE* ≥ 0.3 indicates an acceptable group contribution. This metric can be used to quickly identify active site 'hot spots' were contributions from groups on an inhibitor are most significant (Congreve et al., 2008).

Ligand-Lipophilicity Efficiency *(LLE)*

LLE = pIC50 (or pKi) - cLogP (or LogD)

LLE can be used to guide increases in lipophilicity during optimization and is based on the relationship between high lipophilicity, promiscuity and increased toxicity identified by Leeson and Springthorpe (2007) (see 2.2), who recommend an LLE for a low nM potency lead of ~5-7 or greater. Astex, a world leader in FBDD, identified that their patent molecules had lower

average cLogP (2.4) when compared with 4 Big Pharma companies that have more established HTS influenced strategies, namely Astra Zeneca (3.7); Glaxosmithkline (4.2); Merck (4.0); Pfizer (3.5), suggesting another advantage of using fragments (Leeson and Springthorpe, 2007); (Congreve et al., 2008).

1.2.5   Complexity and Chemical Space in Fragment screening

Hann et al., (2001) identified that with regards to molecular complexity there is an inverse relationship between the chances of detecting a binding event and the probability of that binding event having a unique mode. The distribution under the opposing trend lines is where a useful binding event can be detected, see Figure 1.4.   A fragments small size offers them a higher chance of finding a binding mode, yet low affinity binding, at least with conventional HTS methods, is difficult to detect. Although FBDD is heavily reliant on biophysical methods to detect the weaker binding events, when detection occurs the area of useful event distribution and therefore the probability of detecting a useful binding event is greatly increased.
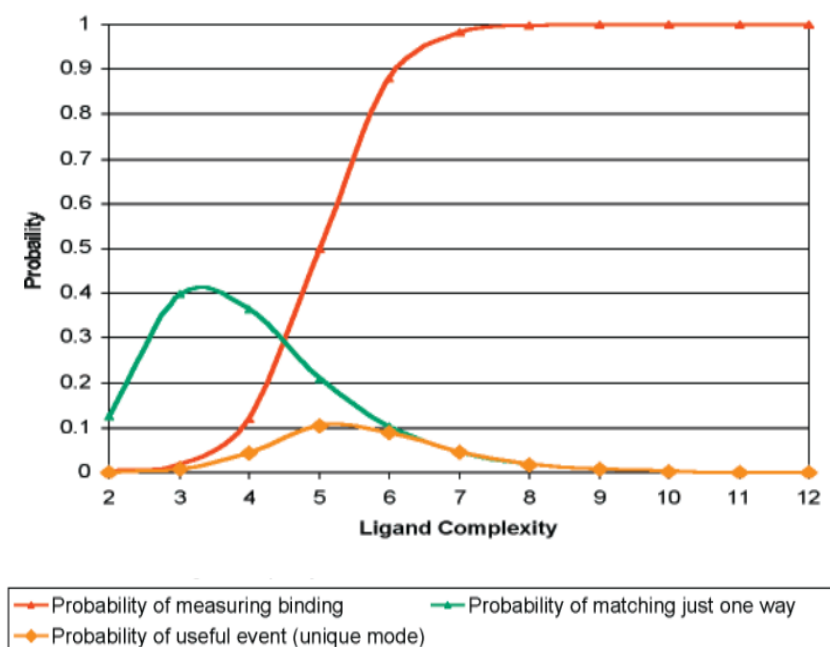
Figure 1.4. The impact of molecular complexity on the probability of detecting a useful binding event. The Green line gives the chance of binding with unique mode, if the red detection line is moved to the left as with fragment screening, the distribution for the orange useful event line is increased accordingly. Taken from (Hann et al., 2001).

An additional advantage of fragment screening is the ability to better explore the useful chemical space, when compared with established HTS screening methods of larger drug-like and even lead-like libraries. Estimates of the number of synthesisable organic molecules range between 10^{18}-10^{200} compounds (Fink et al., 2005). This phenomenal number is put in perspective when we consider the universe is only approximately 10^{18} seconds old (Rydzewski, 2008). The chemical universe at typical drug-like screening size ≈30 heavy atoms is estimated to be more than 10^{60} compounds. The significant exploration of typical drug-like chemical space even with 10^6 screening techniques is therefore negligible (Congreve et al., 2008). In contrast, if the chemical space up to 11 heavy atoms is explored there are only 13.9 million estimated drug-like compounds (Fink et al., 2005) and for up to 13 heavy atoms 970 million drug-like compounds are estimated (Blum and Reymond, 2009). Therefore fragment-like screening of chemical space typically ≤ 18 heavy atoms is likely to cover larger diverse chemical space than HTS even using 10^3-10^4 screening techniques (Congreve et al., 2008).

1.3 Screening of Fragments

NMR. There were some early *in silico* successes in the fledgling field of FBDD (Verlinde et al., 1992), but only when a biophysical experimental technique called SAR by NMR (A 2-D isotope edited NMR method) emerged to detect low MW low affinity binders (Shuker et al., 1996), did the field of FBDD really begin to be valued as an alternate or complementary design strategy (Congreve et al., 2008). SAR by NMR is a linked fragment approach patented by the Abbott group that can detect the weak binding of fragments and was able to make SAR determinations of a fragment library demonstrated against the FK506 binding protein (FKBP).

The protein target (e.g., FKBP) is radio isotope labelled with ^{15}N before screening occurs. The observable chemical shift changes in two- dimensional ^{15}N-Heteronuclear single

quantum correlation (^{15}H-HSQC) spectra then indicate binding in the presence of a ligand. A second ligand interacting with a nearby site was also sought, using detected binding observed in a different set of amide chemical shift changes, which was then analysed to give the approximate location of the second bound ligand. The two fragments once selected had their location and orientation in the protein-ligand complex established by either NMR spectroscopy or X-ray crystallography. The structural data was used to guide the synthesis of compounds where the two fragments were linked in the hope of producing a high affinity ligand (Shuker et al., 1996).

An advantage of SAR by NMR is that no signal from the ligand is observable when N^{15} spectral editing is used, making the binding event detectable at high compound concentrations. In contrast, conventional HTS screening assays, such as fluorimetric and colorimetric assays often give large background signals at high compound concentrations.

Despite the success of SAR by NMR in beginning a revolution in applications of biophysical screening techniques, it was patented, required radiolabelled protein and was only applicable to small molecules (protein) <30 kD that can be obtained in quantities of >200 mg (Shuker et al., 1996). A plethora of alternative screening approaches soon began to emerge in the drive to explore the new science of FBDD. By the year 2000 a glut of papers on these techniques appeared including the SHAPES, NEEDLES and TETHERING strategies used in FBDD (Chessari and Woodhead, 2009).

Hoffman-La Roche used in silico 'NEEDLE' screening to identify hits that predicted good physicochemical properties and solubility at high concentrations, then using programs such as LUDI and CATALYST performed an early High Content Screening (HCS) method to validate hits, before lead optimization (Boehm et al., 2000). A collaboration between Sunesis Pharmaceuticals and University of California saw that the presence of a disulphide linker between two weak binders could be TETHERED by a native or engineered cysteine in an active site to increase affinity 10^3 fold. In addition, cysteine-captured ligands have highly

stable conformations and are detectable by Mass Spectroscopy (MS) at typical drug screening concentrations (10-200μM) (Erlanson et al., 2000). The SHAPES method used elaborations on Bemis and Murcko (1996) SHAPES library using Nuclear Overhauser Effect (NOE) ligand detection NMR techniques for screening.   This NMR technique uses a qualitative but extremely reliable line broadening of ^{1}H spectra of a ligand-protein mixture to detect the binding event. The SHAPES method doesn't require the use of radioisotope labelled ligand and has no limitations in MW, which is in contrast to SAR by NMR (Fejzo et al., 1999). Many NMR techniques now exist and NMR has become a widely used tool with particularly successful applications in the FBDD process (Pellecchia et al., 2008). NMR and X-ray crystallography have become very popular FBDD screening techniques.

X-ray crystallography. In parallel to the evolution of FBDD, new High-throughput crystallography techniques have emerged showing improvements in speed at all stages of the crystallization process and in the subsequent structure solving from X-ray diffraction data (Blundell et al., 2002).   This has proven a massive boon to FBDD and SBDD in general, giving structural determinations for a large number of protein targets and making practical X-ray crystallography screening of small fragment libraries (typically 10^3 compounds) possible. X-ray crystallography comes into its own when a target is amenable to structural determination early in the drug design process. The information helps kick start the lead optimization process (Chessari and Woodhead, 2009).   The advantages and disadvantages of common screening techniques are reviewed by Carr et al., (2005), see Table 1.1.

Surface Plasmon Resonance. More recently Surface Plasmon Resonance (SPR) has emerged as an increasingly popular method for fragment screening as information on stoichiometry is provided. Surface plasmons are collective oscillations of electrons at a metal surface. The oscillations manifest as electromagnetic waves confined to the interfacial plane between a metal and a dielectric. In a phenomenon called resonance these oscillations occur with greater amplitude at some frequencies than they do at others.   The waves can be excited to resonate using a variety of optical setups (e.g., grating or prism

coupling) that result in the coupling of polarized light to the surface plasmon modes of the metal (e.g., gold or silver). At resonance conditions, photon energy is transferred to the surface plasmon mode and hence the reflected light exhibits a sharp attenuation (SPR minimum). The angle and wavelength position of this reflectance minimum, has a strong dependence on the refractive index of the medium adjacent to the metal surface. Processes that alter the local refractive index such as absorption of biomolecules on the sensor layer result in a shift in resonance position and consequently the SPR minimum. This shift can then be used to deduce kinetic constants and hence thermodynamic equilibrium binding data, namely K_d, k_{on} and k_{off} (Neumann et al., 2007)

Isothermal titration calorimetry. Isothermal titration calorimetry (ITC) is another widely used screening technique to acquire thermodynamic and stoichiometric data. ITC involves a single titration experiment where stepwise changes in enthalpy of the ligand-receptor interaction are measured. An appropriate model to describe the interaction makes the thermodynamic parameters readable from a non-linear least squares curve fitting. Difficulties arise when ITC is applied to fragments as accepted models (such as the Wiseman isotherm) are unable to read the thermodynamic parameters of low affinity binders. This is because

Table 1.1 Comparison of fragment screening methods arranged in order of decreasing throughput and increasing information content.

Approach	Typical throughput per screen (compounds)	Quality of information about ligand binding mode	Resource and instrumentation requirements	Protein structure required	Key technical considerations
HTS	100-1000K	None	Specialized infrastructure required	No	Not suitable for fragments
High concentration bioassay	10-50K	None	Very straightforward method	No	High false-positive rates can often hinder interpretation of data
Surface Plasmon resonance	10-50K	None	Straightforward method, but requires costly instrumentation	No	Protein or compounds must be immobilized; false positives possible

Affinity mass spectrometry	10-50K	None	Straightforward method, but requires costly instrumentation	No	Limited applications reported
Covalent attachment and MS	10-50K	None	Specialized infrastructure required	No	Requires cysteine residue close to active site
Dynamic combinatorial chemistry and LC/MS	1-10K	None	Straightforward method	No	Limited range of chemistry is suitable
Ligand-detected NMR (1D/2D)	1-10K	Can distinguish active site vs non-active site binders	Straightforward methods using ^{1}H or ^{19}F, but requires costly instrumentation. Well suited to screening of mixtures	No	Protein typically <20 kDa in size; moderate protein requirements
Protein-detected NMR (2D)	1-10K	Information on principle interactions between ligand and protein	Straightforward methods using ^{1}H/^{15}N NMR resonance assignments for amide groups. Requires costly instrumentation	Usually	Protein typically <30 kDa in size; high protein requirements
X-ray crystallography	500 – 1000	Detailed binding mode elucidated	Specialized infrastructure required	Yes	Limited to ~35% drug targets where the structure can be solved; moderate protein requirements

Adapted from (Carr et al., 2005).

the reading of a desired sigmoidal curve is not possible if the c value (the product of receptor concentration (M) and association constant (K_a)) is too low. The c value may be viewed as the ratio between the receptor concentration and dissociation constant (K_d) and influences the shape of the isotherm. A c value between 10-500 is required for the Wiseman isotherm to work where fragments often have c value <10 or even <1.   Alternative models and careful experimental design have made ITC screening of fragments possible (Turnbull and Daranas, 2003).

1.4 The Optimization of Fragments

There are considered to be two approaches to drug design, these are blind and rational. HTS is typically a blind technique that can produce leads even when structural information isn't amenable. Many drug discovery programmes now take an orthogonal approach to research and development, utilising the advantages and minimizing the impact of the disadvantages of a number of strategies. FBBD is a rational or semi-rational approach often synonymous with structure based drug design (SBDD). '*In silico*' rational methods require extensive structural information to progress through a lead optimization phase and produce a promising drug candidate. Ligand-protein complexes offer many opportunities to a skilled computational chemist, such as identifying key interactions that guide 3D optimization of fragments and the creation of '*de novo*' ligands and acting as a test bed for assessing the accuracy of docking simulations, and in the identification of new binding sites and validation of optimized leads (Congreve et al., 2008).

'*De novo*' design involves the '*in silico'* creation of a New Chemical Entity (NCE) that is distinct from its constituent parts. Notable overlaps exist in the optimization strategies of FBDD and '*de novo*' drug design and at least to some extent they have become interchangeable (Loving et al., 2010). *In silico* methods used to describe fragment-like molecules existed at least in concept before SAR by NMR. The GRID program for instance used small chemical probes to assess energetically favourable 'hotspots' in the active site with compact significant contributions to the binding event; the results of these probes correlated with what was observed in known enzyme inhibitors (Goodford, 1985). Fragments can be used to identify new binding sites by blocking an active site hotspot with a binding fragment and then to probe the resulting protein-ligand complex for a second binding event, potentially allowing linked fragments if the binding sites are adjacent.

The early 1990's saw the first linked fragment approaches pioneered computationally (Verlinde et al., 1992, Böhm, 1992). In an ensuing report, the lead optimization approaches that would later become most typical in FBDD were described; these approaches were

referred to as building (later called growing or evolving) and linking, the fragments were referred to both as 'fragments' and 'seeds' (Verlinde and Hol, 1994).

1.4.1 Evolving of Fragments

The binding of fragments can usually be improved by substituted additional functionality at one or more vectors, see Figure 1.5**.** Therefore, a fragment with at least one chemical handle (which has suitable functionality for further chemical elaboration) should be selected in the library design process (Siegal et al., 2007). The fragment evolution will only succeed rationally if the binding mode is maintained during optimization (Congreve et al., 2008). Deconstruction of a known ß-lactamase inhibitor showed that this is not always the case. (Babaoglu and Shoichet, 2006). If however the fragment is an optimal inhibitor, or close to it, there is no convincing evidence that the binding mode will not be maintained (Ciulli and Abell, 2007). Many examples of evolved fragments now exist (e.g., BACE-1 inhibitors that show promise in treating Alzheimer's disease, see Figure 1.6). Evolving is the most popular and successful lead optimization strategy in FBDD (Congreve et al., 2008). The reason for the popularity of this approach over linking is that a multistep optimization of LE and MW can be used to give a more flexible exploration of the binding site. In contrast, linking is constrained by the size of the original fragments and the size of the linker, resulting in a rapid single step build-up of atoms. Additionally, the conformational strain and flexibility of linked fragments results in an entropy loss during linkage (Hung et al., 2009).
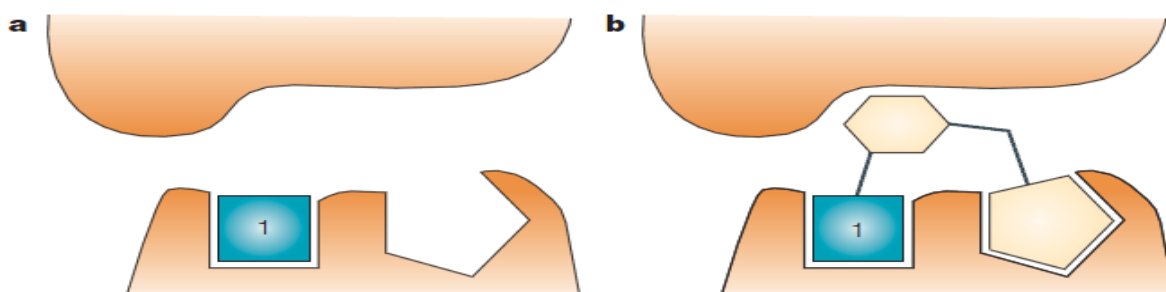
Figure 1.5 Schematic of Fragment Evolution. a Tight binding fragment/inhibitor in a protein active site moiety. b The Fragment is evolved across the active site, utilizing other potential interactions in the active site. Taken from (Rees et al., 2004)
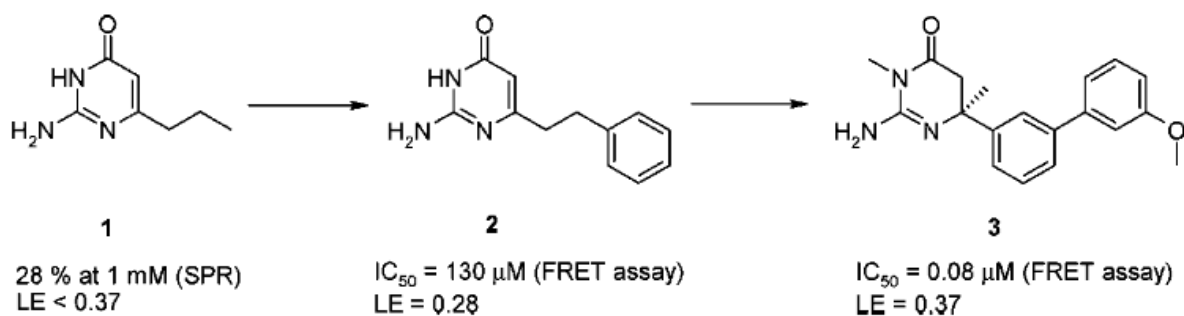
Figure 1.6 Evolved lead optimization resulting in compound [3] a potent BACE-1 inhibitor. Taken from (Congreve et al., 2008)

1.4.2 Linked Fragments

The rationale behind the linking of fragments is discussed in 1.2.1 and appropriate screening methods to identify a second fragment for linking such as 'SAR by NMR' is discussed in 1.3. Fragment linking requires the often challenging identification of a linker that allows both fragments to obtain their original binding mode when combined into a new ligand with acceptable affinity and LE, see Figure 1.7. An example of a successfully linked approach leading to potent non-amidine containing thrombin inhibitors is given in Figure 1. 8 (Congreve et al., 2008).
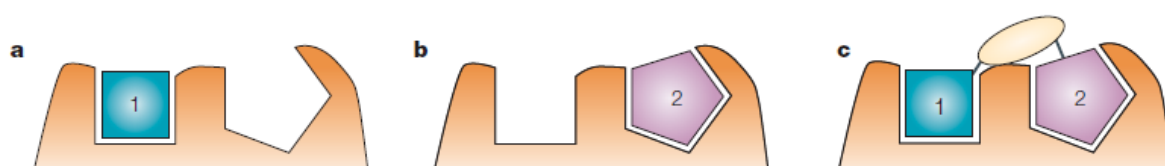
Figure 1.7. Schematic of Fragment Linking. a Fragment binds to a receptor at one site. b fragment binds to a receptor at adjacent site. c the addition of a linker group allows the lead molecule to span both sites. Taken from (Rees et al., 2004)
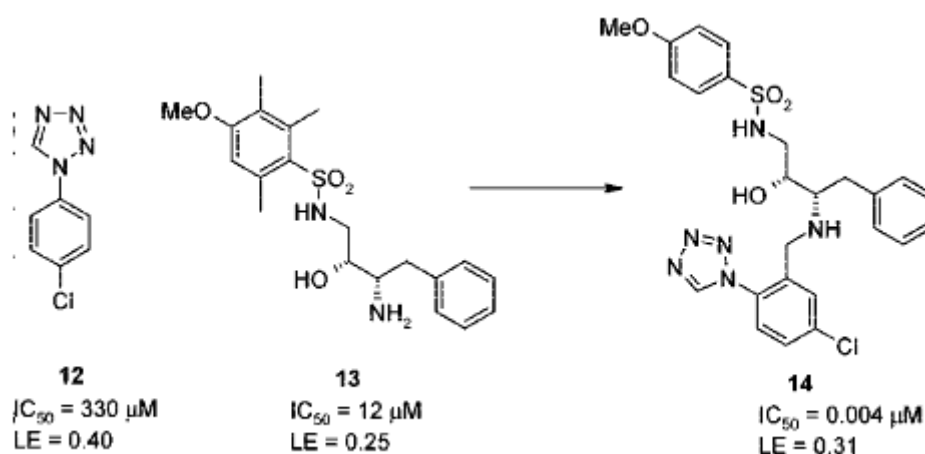
Figure 1.8. Fragment linking leading to potent thrombin inhibitors. Taken from (Congreve et al., 2008)

1.5 Relevance to project

Computational methods lend themselves to many areas of the FBDD process yet are not considered accurate enough to predict reliably the binding mode of fragments, which is still heavily dependent on biophysical methods for hit detection and validation (Congreve et al., 2008). Fragment docking is particularly sensitive to the accurate description of charges and other parameters that govern their interaction with their protein target and it is in the context that this research was initiated. Because structural information is so important in FBDD, we are seeking to introduce more accurate docking methods to aid in fragment growth. We will therefore seek to improve the charge description through the inclusion of polarization via induced charges, since this approach is compatible with state of the art docking programs such as GLIDE and Autodock. However, this is only part of the problem so we will also consider the use of anisotropic charge distributions through the use of distributed multipole analysis (DMA) and the use of more realistic repulsion than the commonly used r^{12} term. General computational approaches are discussed in the next chapter.

1.6 Conclusion

FBDD is proving a successful design strategy with many compounds either in or entering clinical trials (de Kloe et al., 2009). In a scenario where resources are often limited, design strategy often comes down to what experience and resources are available for use from previous campaigns such as the available equipment, stored compound collections and philosophies and expertise of staff in employment (Rydzewski, 2008). For many large organizations the significant up-front investment and timely commitment to structural biology to establish binding modes and eliminate false positives is difficult to achieve. Incentive to add to HTS strategies is further reduced by need for specialist equipment /staff and in that only a proportion of targets are readily available to 3-D structural determinations. In contrast, in an academic setting, assembling a small library of fragments and their screening using biophysical techniques such as NMR and X-ray crystallography followed by SBDD is much more achievable than the assembly and screening of a large compound library in a HTS bioassay. This, coupled with its intellectual appeal, makes FBDD a promising approach to drug design. The requirement for structural biology and the need to use improved computational chemistry methods to exploit this structure is one of the driving forces behind this project.

1.7 References

ANDREWS, P., CRAIK, D. & MARTIN, J. 1984. Functional group contributions to drug-receptor interactions. *Journal of Medicinal Chemistry,* 27**,** 1648-1657.

BABAOGLU, K. & SHOICHET, B. K. 2006. Deconstructing fragment-based inhibitor discovery. *Nature chemical biology,* 2**,** 720-723.

BEMIS, G. W. & MURCKO, M. A. 1996. The properties of known drugs. 1. Molecular frameworks. *Journal of medicinal chemistry,* 39**,** 2887-2893.

BLUM, L. C. & REYMOND, J. L. 2009. 970 million druglike small molecules for virtual screening in the chemical universe database GDB-13. *Journal of the American Chemical Society,* 131**,** 8732-8733.

BLUNDELL, T. L., JHOTI, H. & ABELL, C. 2002. High-throughput crystallography for lead discovery in drug design. *Nat Rev Drug Discov,* 1**,** 45-54.

BOEHM, H.-J., BOEHRINGER, M., BUR, D., GMUENDER, H., HUBER, W., KLAUS, W., KOSTREWA, D., KUEHNE, H., LUEBBERS, T., MEUNIER-KELLER, N. & MUELLER, F. 2000. Novel Inhibitors of DNA Gyrase: 3D Structure Based Biased Needle Screening, Hit Validation by Biophysical Methods, and 3D Guided Optimization. A Promising Alternative to Random Screening. *Journal of Medicinal Chemistry,* 43**,** 2664-2674.

BÖHM, H.-J. 1992. The computer program LUDI: A new method for the de novo design of enzyme inhibitors. *Journal of Computer-Aided Molecular Design,* 6**,** 61-78.

CARR, R. A. E., CONGREVE, M., MURRAY, C. W. & REES, D. C. 2005. Fragment-based lead discovery: leads by design. *Drug Discovery Today,* 10**,** 987-992.

CHESSARI, G. & WOODHEAD, A. J. 2009. From fragment to clinical candidate--a historical perspective. *Drug discovery today,* 14**,** 668-675.

CIULLI, A. & ABELL, C. 2007. Fragment-based approaches to enzyme inhibition. *Current opinion in biotechnology,* 18**,** 489-496.

CONGREVE, M., CHESSARI, G., TISI, D. & WOODHEAD, A. J. 2008. Recent Developments in Fragment-Based Drug Discovery. *Journal of Medicinal Chemistry,* 51**,** 3661-3680.

DE KLOE, G. E., BAILEY, D., LEURS, R. & DE ESCH, I. J. P. 2009. Transforming fragments into candidates: small becomes big in medicinal chemistry. *Drug Discovery Today,* 14**,** 630-646.

ERLANSON, D. A., BRAISTED, A. C., RAPHAEL, D. R., RANDAL, M., STROUD, R. M., GORDON, E. M. & WELLS, J. A. 2000. Site-directed ligand discovery. *Proceedings of the National Academy of Sciences of the United States of America,* 97**,** 9367.

ERLANSON, D. A., MCDOWELL, R. S. & O'BRIEN, T. 2004. Fragment-based drug discovery. *J. Med. Chem,* 47**,** 3463-3482.

FEJZO, J., LEPRE, C. A., PENG, J. W., BEMIS, G. W., AJAY, MURCKO, M. A. & MOORE, J. M. 1999. The SHAPES strategy: an NMR-based approach for lead generation in drug discovery. *Chemistry & Biology,* 6**,** 755-769.

FINK, T., BRUGGESSER, H. & REYMOND, J. L. 2005. Virtual Exploration of the Small Molecule Chemical Universe below 160 Daltons. *Angewandte Chemie International Edition,* 44**,** 1504-1508.

GERSHELL, L. J. & ATKINS, J. H. 2003. A brief history of novel drug discovery technologies. *Nat Rev Drug Discov,* 2**,** 321-327.

GOODFORD, P. J. 1985. A computational procedure for determining energetically favorable binding sites on biologically important macromolecules. *Journal of Medicinal Chemistry,* 28**,** 849-857.

HAJDUK, P. J. 2006. Fragment-based drug design: how big is too big? *Journal of Medicinal Chemistry,* 49**,** 6972-6976.

HANN, M. M., LEACH, A. R. & HARPER, G. 2001. Molecular complexity and its impact on the probability of finding leads for drug discovery. *Journal of chemical information and computer sciences,* 41**,** 856-864.

HUNG, A. W., SILVESTRE, H. L., WEN, S., CIULLI, A., BLUNDELL, T. L. & ABELL, C. 2009. Application of fragment growing and fragment linking to the discovery of inhibitors of Mycobacterium tuberculosis pantothenate synthetase. *Angewandte Chemie,* 121**,** 8604-8608.

JENCKS, W. P. 1981. On the attribution and additivity of binding energies. *Proceedings of the National Academy of Sciences of the United States of America,* 78**,** 4046.

KUNTZ, I., CHEN, K., SHARP, K. & KOLLMAN, P. 1999. The maximal affinity of ligands. *Proceedings of the National Academy of Sciences of the United States of America,* 96**,** 9997.

LEESON, P. D. & SPRINGTHORPE, B. 2007. The influence of drug-like concepts on decision-making in medicinal chemistry. *Nat Rev Drug Discov,* 6**,** 881-890.

LIPINSKI, C. A., LOMBARDO, F., DOMINY, B. W. & FEENEY, P. J. 1997. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Advanced drug delivery reviews,* 23**,** 3-25.

LOVERING, F., BIKKER, J. & HUMBLET, C. 2009. Escape from flatland: increasing saturation as an approach to improving clinical success. *Journal of Medicinal Chemistry,* 52**,** 6752-6756.

LOVING, K., ALBERTS, I. & SHERMAN, W. 2010. Computational Approaches for Fragment-Based and De Novo Design. *Current Topics in Medicinal Chemistry,* 10**,** 14-32.

MACARRON, R., BANKS, M. N., BOJANIC, D., BURNS, D. J., CIROVIC, D. A., GARYANTES, T., GREEN, D. V. S., HERTZBERG, R. P., JANZEN, W. P. & PASLAY, J. W. 2011. Impact of high-throughput screening in biomedical research. *Nature Reviews Drug Discovery,* 10**,** 188-195.

MARIO GEYSEN, H., SCHOENEN, F., WAGNER, D. & WAGNER, R. 2003. Combinatorial compound libraries for drug discovery: an ongoing challenge. *Nat Rev Drug Discov,* 2**,** 222-230.

NEUMANN, T., JUNKER, H., SCHMIDT, K. & SEKUL, R. 2007. SPR-based fragment screening: advantages and applications. *Current topics in medicinal chemistry,* 7**,** 1630-1642.

PELLECCHIA, M., BERTINI, I., COWBURN, D., DALVIT, C., GIRALT, E., JAHNKE, W., JAMES, T. L., HOMANS, S. W., KESSLER, H., LUCHINAT, C., MEYER, B., OSCHKINAT, H., PENG, J., SCHWALBE, H. & SIEGAL, G. 2008. Perspectives on NMR in drug discovery: a technique comes of age. *Nat Rev Drug Discov,* 7**,** 738-745.

REES, D. C., CONGREVE, M., MURRAY, C. W. & CARR, R. 2004. Fragment-based lead discovery. *Nat Rev Drug Discov,* 3**,** 660-672.

RYDZEWSKI, R. M. 2008. *Real world drug discovery: a chemist's guide to biotech and pharmaceutical research*, Elsevier Science.

SHUKER, S. B., HAJDUK, P. J., MEADOWS, R. P. & FESIK, S. W. 1996. Discovering high-affinity ligands for proteins: SAR by NMR. *Science,* 274**,** 1531.

SIEGAL, G., AB, E. & SCHULTZ, J. 2007. Integration of fragment screening and library design. *Drug Discovery Today,* 12**,** 1032-1039.

TEAGUE, S. J., DAVIS, A. M., LEESON, P. D. & OPREA, T. 1999. The Design of Leadlike Combinatorial Libraries. *Angewandte Chemie International Edition,* 38**,** 3743-3748.

TURNBULL, W. B. & DARANAS, A. H. 2003. On the value of c: can low affinity systems be studied by isothermal titration calorimetry? *Journal of the American Chemical Society,* 125**,** 14859-14866.

VEBER, D. F., JOHNSON, S. R., CHENG, H. Y., SMITH, B. R., WARD, K. W. & KOPPLE, K. D. 2002. Molecular properties that influence the oral bioavailability of drug candidates. *Journal of Medicinal Chemistry,* 45**,** 2615-2623.

VERLINDE, C. L. & HOL, W. G. 1994. Structure-based drug design: progress, results and challenges. *Structure,* 2**,** 577-587.

VERLINDE, C. M. J., RUDENKO, G. & HOL, W. J. 1992. In search of new lead compounds for trypanosomiasis drug design: A protein structure-based linked-fragment approach. *Journal of Computer-Aided Molecular Design,* 6**,** 131-147.

VIETH, M., SIEGEL, M. G., RICHARD, E., WATSON, I. A., ROBERTSON, D. H., SAVIN, K. A., DURST, G. L. & HIPSKIND, P. A. 2004. Characteristic physical properties and structural fragments of marketed oral drugs. *Journal of Medicinal Chemistry,* 47**,** 224-232.

WALTERS, W. P. & NAMCHUK, M. 2003. Designing screens: how to make your hits a hit. *Nature Reviews Drug Discovery,* 2**,** 259-266.

2 Methods Overview

2.1 Introduction to Virtual Screening

Despite ultra HTS's power to perform individual biochemical assays on a huge number of compounds, it still has limitations in hit identification. The technical complexity and high cost and attrition rates as discussed in chapter 1 encourage the continued exploration of alternative and supplementary techniques (Tuccinardi, 2009). These techniques include *in silico* methods, which are not yet considered accurate or reliable enough to rival biophysical screening techniques and so cannot be used independently. They can however be used in an orthogonal fashion with biophysical techniques, and when successful provide insights that drive forward the science and reduce costs in the research and development stage (Hajduk and Greer, 2007, Congreve et al., 2008).

Virtual screening (VS) integrates biophysical principles with computer science and largely attempts to predict likely 'hits' for subsequent bioassay or HTS evaluation. VS techniques can be classified as ligand or receptor-based. Ligand based methods use known binders (compound or active series) of a particular biomolecular target to extract other molecules with similar properties from a compound library. Examples methods include Similarity searching, quantitative structure-activity relationship (QSAR) models, fingerprint and pharmacophore searching.

Docking-based Virtual Screening (DBVS), our focus here, is an important receptor based method in the drug design process where 3D structural knowledge of a receptor target is used to search for molecules with favourable interactions (Tuccinardi, 2009, Jorgensen, 2004). Until recently, Structure based Drug Design (SBDD) 'in silico' methods had contributed to approximately fifty compounds that have entered clinical trials (Jorgensen, 2004). However progress in crystallography, homology modelling, and a range of computational techniques and the hardware they use has resulted in receptor based *in silico*

methods in general being increasingly recognized in speeding up and reducing the costs involved with hit identification and lead optimization (Zoete et al., 2009, Tuccinardi, 2009).

2.1.1 Docking-based Virtual Screening

DBVS is a strong area of active research that has three main applications at different stages of the drug discovery process. Firstly, there is in the prediction of binding mode for a known active, secondly identifying potential 'hits' using virtual screening, and thirdly predicting binding affinity of compounds from within a known active series. The area where arguably the most success has been reported (and the main area of focus here) is in successful prediction of a ligand binding mode (Huang et al., 2010).

Docking studies attempt to address the complex issues involved in identifying features of specific biomolecular recognition or ligand binding affinity prediction. These challenges can be difficult to understand and harder to simulate. Therefore Docking Programs adopt multi-step approaches, which filter the number of potential binding solutions with incremental increases in complexity (Kitchen et al., 2004). Docking, although complex, can be roughly summarized as a search algorithm that attempts to suggest the most likely small molecule binding geometries (poses) and a scoring function that will endeavour to rank these to identify the true or native binding mode (Zoete et al., 2009).

2.1.2 Docking as a complex problem

The ideal DBVS tool needs to combine a docking algorithm capable of finding the correct ligand orientation and conformation (pose) with a fast reliable scoring function, while considering the hardware limitations of the average computer workstation (Kellenberger et al., 2004). The current state of the art in DBVS programs can usually provide a selection of poses that largely includes the pose associated with the X-ray experimental binding, across a diverse range of protein classes. The problem with reliability is therefore more often attributed to the scoring functions, which in many cases are not yet considered sophisticated

enough or versatile enough to rank poses accurately across these same diverse range of protein classes (Warren et al., 2006).

Successful and reliable DBVS is technically challenging with many potential pitfalls to consider across a range of scientific disciplines. Many will be mentioned below, but in brief regard these include, but are by no means exclusive to:

- The accuracy and relevance of the crystallographic model, i.e. the resolution of x-ray diffraction data, protonation states and the assignment of the hydrogen bond network;
- the induced fit problem associated with identifying novel binders i.e. ligand/protein flexibility;
- Single binding pocket focus, with poor consideration or validation against allosteric or multiple binding pockets.
- The difficulties in modelling of the condensed phase i.e. solvation effects, the treatment of ions and consideration for water mediated binding interactions.
- The selection of potentials and schemes in the modelling of bonded and non-bonding terms (e.g., the pair-wise approximation).
- The exploration of phase space both when considering the scheme and the philosophy e.g., energy minimization methods only explore local minima over pragmatic time steps;
- The exploration of chemical space and the diversity/focus of the training set e.g., adoption of 'Lipinski' or 'fragment space', potentially defines the usefulness against future docking runs. the expectation to find high affinity binders without consideration of the optimization potential, and reliable and consistent methods of docking performance evaluation when considering the possibilities with different training sets and protein classes. (Scior et al., 2012, Jorgensen and Tirado-Rives, 2005, Kitchen et al., 2004).

2.2   The Molecular Representation

Molecular representations used for docking in order of increasing complexity and computational expense are the surface, grid and atomic representations. The surface representation is primarily for describing protein-protein interactions. While the potential energy grid representation pioneered by (Goodford, 1985), stores the receptor's energetic contributions e.g., electrostatic and van der Waals (VDW) terms on a number of grid points. Grid representations greatly reduce computational expense for the evaluation and treatment of the potential energy surface when posing and scoring the ligand, but complicate the ability to model conformational flexibility of the receptor. The majority of the work described here uses a grid since this is the approach used by Glide. Atomic representations generally use a potential energy function that rank the pair-wise interactions at the final stage of ranking by the scoring function (Kitchen et al., 2004).

2.2.1   The Docking Algorithm

The Docking algorithm, through a fuzzy process known as 'Posing' (see Figure 2.1), attempts to explore the chemically relevant conformational space and orientation of a ligand's binding within its target receptor. It does this through a process of first placing and then scoring the ligand to determine the most likely available solutions (Kitchen et al., 2004). Even simple rigid body posing (lock and key model) without consideration of protein or ligand flexibility is a complex computational task and must model the six degrees of freedom (DOF), i.e. translational and rotational of one body relative to another (Leach et al., 2006).   However, during formation of the protein-ligand complex, a process known as induced fit occurs where both the ligand and protein adapt conformations to each other.
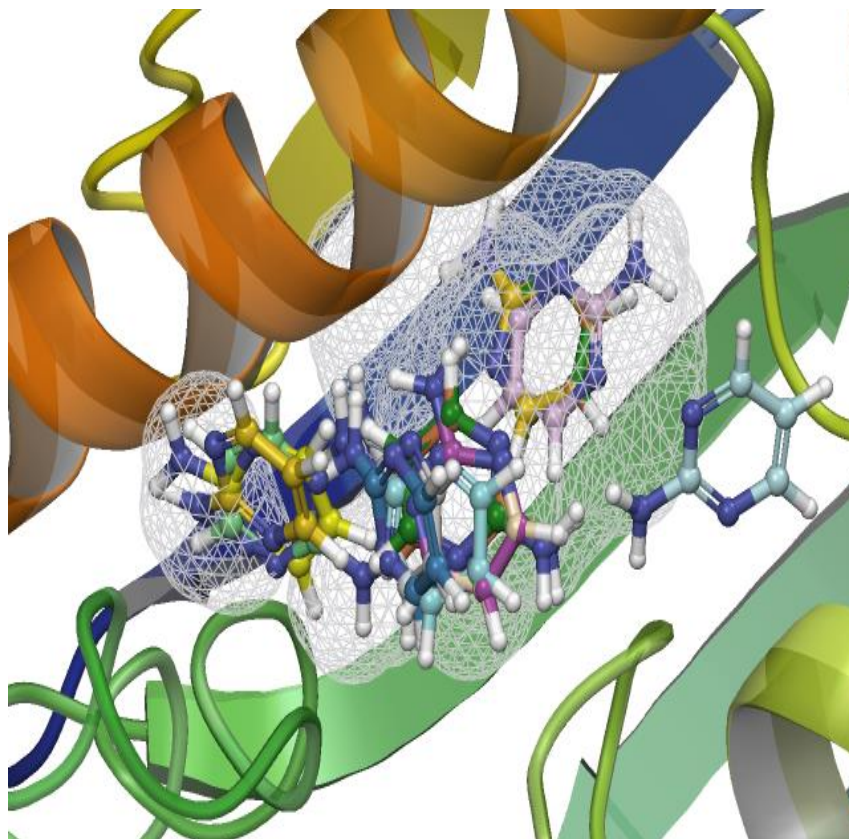
Figure 2.1 Example of Glide Posing solutions within active (wireframe) pocket. e.g., PDB 2jjc shows several ligand poses at the bottom of the pocket including 2 that are closer to the left, 2 at the top and 1 on the right. Each ligand is distinguished by different coloured carbon atoms.

Sometimes a protein may have multiple or allosteric pockets; this becomes a difficult problem to consider and validate using docking algorithms, which are designed to dock ligands with single binding pockets in mind (Scior et al., 2012). For example, in tandem with the Induced fit model is the Population shift protein-ligand complex model. In this model, the inhibitor (ligand) changes the enzyme (protein) conformation to a less active or an inactive conformation which (in terms of statistical mechanics) occurs throughout the population (Teague, 2003). While the binding pocket is preformed to facilitate this conformation change, it is also possible to destabilize the folded structure of the protein through allosteric binding (Horn and Shoichet, 2004).

2.2.2 The Induced fit model

The induced fit phenomena in the context of enzyme activity, can be described by enzyme activation requiring the specific substrate, in finding its binding conformation and orientation, to make appreciable changes in the three dimensional structure/relationship of the amino acids at the active site to precisely orient the catalytic groups for reaction (Koshland Jr, 1958). The 'induced fit' conformational changes can be subtle or linked to refolding and orientating of entire protein domains. X-ray crystallography of protein-ligand complexes typically show the ligand to have ~70%-100% of their surface area buried, giving evidence of protein encapsulation of the ligand (Teague, 2003). Therefore, modelling of protein/ligand flexibility should ideally be considered in the docking algorithm. However these additional conformational degrees of freedom (DOF) lead to a combinatorial explosion of the sampling space that, when unrestricted, lead to a near infinite number of putative pose possibilities (Zoete et al., 2009, Durrant and McCammon, 2010).

$$N_{Conformations} = \prod_{i=1}^{N} \prod_{j=1}^{n_{inc}} \frac{360}{\theta_{ij}}$$
(2.1)

This can be represented in Eq. 2.1, where N is the number of rotatable bonds and θ_{ij} is the size of the incremental rotational angle j for bond i, where the total number of increments for each bond, $n_{inc,}$ i can be calculated using $n_{inc=} \frac{360}{\theta_{ij}}$ (Kitchen et al., 2004). There are three broad categories of sampling algorithm used to treat mainly ligand flexibility these are systematic, stochastic (or random), and simulation search methods. However, within this work, protein flexibility is generally not included due to limitations within Glide.

2.2.3 Ligand Flexibility

2.2.3.1 Systematic Sampling Algorithms

The Systematic approach ideally requires fast exploration of all the conformational and rigid body degrees of freedom (DOF) to find the native binding mode. Strategies for dealing with

the DOF combinatorial explosion for systematic searches involve various methods for the incremental build of the ligand in the binding site (Zoete et al., 2009). The ligand is treated as a series of fragments and in this way these approaches are equivalent to the linking and growing of fragments in FBDD. The linking strategy docks molecular fragments in the active site during posing, linking them covalently. Alternatively, ligands are docked as core fragments, with the flexible parts treated as side chains. These side chains are grown incrementally from the docked fragment during posing based on favourable scoring. Sometimes additional energy minimization occurs between each accepted increment, as in the Hammerhead Program (Welch et al., 1996). The Glide docking program (Friesner et al., 2004) chosen for this research is considered among the state of art docking programs and uses an exhaustive systematic search algorithm made possible by introducing a number of hierarchal filters to incrementally reduce the search space.

2.2.3.2 Random Sampling Algorithms

Random (or stochastic) sampling algorithms consider the ligand as a whole. Stepwise changes are made to the starting pose or population of poses. These steps are kept or rejected based on the favourability of scoring while posing. Monte Carlo or evolutionary algorithms (EA) are considered stochastic sampling methods. EA mimic the process of Darwinian evolution; the stepwise changes called operators mimic biological (gene) mutations that introduce binding mode perturbations such as rotation of a dihedral angle and crossovers (mimicking gene recombination) that combine two poses. Operators (for posing) are applied to selected poses to identify the fittest elements of the population and attempt to generate fitter solutions (Zoete et al., 2009). Monte Carlo after generation of a random starting pose, typically keeps the stepwise changes which are further randomly generated poses based on the 'Metropolis criterion'. These criteria are to keep the new solution immediately if a new minimum is obtained. If not a Boltzmann-based probability function test is applied where if more probable, the new pose is also kept; otherwise it is rejected. These methods as with 'tabu search' algorithms that relax the strictness of conditions for finding

local search minima to check for similar solutions that may offer improvement, need to keep track of the favourability of already explored areas of conformational space. Ideally, these methods lead to convergence into a relatively small pose population of potential solutions (Kitchen et al., 2004). The Glide Program also uses Monte Carlo sampling in the final stages of hierarchal pose evaluation (Friesner et al., 2004).

2.2.4 Molecular Dynamics Methods

Molecular dynamics (MD) is the most popular simulation method and is capable of giving many insights unavailable to other more conventional docking algorithms. However, in consideration of docking, a local minima problem arises as MD is often unable to cross high energy barriers within pragmatic simulation time scales. Therefore, a common usage strategy is to start molecular dynamics calculations from different ligand positions perhaps from the final pose population reached using other docking methods. The use of alternative conformations can be handled using the locally enhanced sampling (LES) method (Simmerling et al., 1998), which simulates multiple non-interacting copies of the 'enhanced' part of the system, with modifications to the force field that enhance sampling. Another approach to MD simulation attempts the simulation of different parts of the protein-ligand complex at different temperatures. Replica exchange molecular dynamics (REMD) (Sugita and Okamoto, 1999) uses multiple copies of the system, with each copy at a different temperature. The copies at high temperature can sample more widely, and so can find new regions of conformational space; configurations can exchange according to the Metropolis criteria and so the main room temperature configuration can effectively sample wider regions. These methods can be used to provide a better homology model (see 2.2.5.1). Alternatively to MD, energy minimization methods although largely incapable of exploring beyond local energy minima are often used in conjunction with other search methods to refine geometries during posing to hopefully reach a few well defined minima (Kitchen et al., 2004, Zoete et al., 2009).

2.2.5   Protein Flexibility

Protein flexibility can be simulated by molecular dynamics, subject to the limitations discussed in 2.2.4, but is often not included in docking for the reasons discussed in section 2.2.2. Glide does offer a solution to limited flexibility of a small number of protein side chains through the use of Glide in combination with the Prime homology modelling program, but Prime was not available for this work and so a rigid protein model was used. This could be a problem when a pose is generated that presents say a strongly positive atomic charge on the ligand in the vicinity of a strongly positive charge on the protein. Normally this clash would be resolved by movement of the two groups away from each other. In our studies, this situation can be partially ameliorated by polarization of charges on both the ligand and the protein so that electrostatic polarization reduces the clash. This type of flexibility (in the electronic distribution) is not available in standard docking methods.

2.2.5.1   Global Flexibility

MD is commonly used as an alternative to NMR, or X-ray crystallography, or Monte Carlo simulation with regards to the induced fit problem to generate multiple conformationally diverse structures that can be used in modelling of global protein flexibility, even for homology models. Computational methods are attractive because they generate a full continuum of structures, although concerns over adequate sampling of conformational space at largely nanosecond scales exist. An additional method is the ensemble average energy grid which uses a composite energy grid generated by averaging grids calculated over multiple structures for use during DBVS. Dynamic pharmacophore modelling typically uses multiple MD structures to characterize the active binding site regions in building a composite model, which can then be screened against a ligand database for compounds with complementary chemical features. Normal mode analysis has also been incorporated into ligand-identification protocols to account for protein flexibility (Durrant and McCammon, 2010). These methods have generally not been explored in this thesis, though flexibility has

been explored in cases where multiple structures (including multiple protein chains within an asymmetric unit) are available.

2.2.5.2 Local Receptor Flexibility

General treatments that incorporate local receptor flexibility in increasing order of flexibility are soft docking, side chain flexibility, and side chain and backbone flexibility.   Soft docking exchanges the commonly used Lennard-Jones potential which rapidly increases to infinity as interatomic distances approach zero with a more forgiving often exponential potential for describing the VDW force.   The soft potential does not approach infinity or carry massive energy penalties at short interatomic distances, effectively allowing for minor steric clashes (Durrant and McCammon, 2010). Soft docking is explored in this research, both through scaling of the van der Waals' potential in a 1.0/0.8 scaling configuration where polar ligand radii remain unscaled and nonpolar ligand radii are scaled by 0.8 as a standard Glide option (Friesner et al., 2004) and through the use of an exponential repulsion, here the Williams potential (Williams, 1999, Williams, 2001a, Williams, 2001b) in Chapter 4.

Rotamer libraries containing a number of experimentally observed discrete rotameric states can explore side chain flexibility amenable to binding. These are commonly optimized by recursively removing dead-end rotation steps, that do not contribute to a minimum-energy structure (Kitchen et al., 2004). A fuller spectrum of backbone and sidechain motions can also be explored by using energy refinement techniques involving relaxed geometry optimization following docking; these still suffer from only being able to explore local energy minima (Durrant and McCammon, 2010).

2.2.5.3 Difficulties with Induced fit modelling and application to this research

The treatment of protein flexibility is currently considered less advanced than that of ligand flexibility (Kitchen et al., 2004). In a ligand similarity study of the Protein data bank (PDB) it was shown that the binding sites of paired proteins with structurally similar ligands differ in 83% of cases, where the most common difference was water molecule architecture, side

chain movements (altered receptor conformations) were observed in half the pairs, while backbones were largely unmoving. It was also noted that small changes in ligand or binding site character can often lead to unpredictable changes in activity, e.g. an incorrect (specificity) of side chain rotamer which is difficult to account for using molecular similarity techniques (Boström et al., 2006). Like docking itself, the difficulties involved in homology and induced fit modelling of the receptor make reliable universally applicable techniques hard to establish (Seifert and Lang, 2008, Kairys et al., 2005, Alvesalo et al., 2006).

Solving steric problems caused by lack of motion / flexibility in protein models used for docking, have previously been shown to outweigh the contribution from using more sophisticated MM non bonded force-field related terms in providing docking success (Kitchen et al., 2004). For instance, using now aging force-field based scoring methods in a number of native case studies, a docking success rate of ~80% was obtained (RMSD < 2 Å of the X-ray pose). This success rate dropped to only 56% during a cross docking study of the same ligands that used a different version of the protein structure with which the ligand was not co-crystallized/soaked (Pérez and Ortiz, 2001). Here we largely consider just the problem of successful prediction of an X-ray experimental ligand binding mode within its receptor, where the need to model for receptor induced fit is greatly reduced. While ligand flexibility is dealt with by a filtered exhaustive systematic flexible ligand search carried out by Glide (Friesner et al., 2004). For the native docking experiments in chapters 3-5 the active pockets have been primed through protein-ligand complex energy minimization using the IMPACT Program (Banks et al., 2005) to expect the native ligand and some modelling of the water molecule architecture has also been applied (see 3.2.4). We address these biases through investigating the cross-docking problem with and without polarization in Chapter 6.

2.3 Generation of the Protein-ligand Model

X-ray Crystallography is the main source of structural information for protein-ligand complexes generally and in this research (Davis et al., 2003). The assumption (which is

usually justified) is that the protein-ligand complex crystallizes with the ligand in its most thermodynamically stable or ideal geometry (Kleywegt, 2006). In a crystallographic (X-ray diffraction) experiment, the positions and intensities of reflections are given by the diffraction pattern of the crystal. Structure-factor amplitudes can be approximately calculated as the square root of these intensities. Fourier transforms on the (post phase established) structure-factors provide a local electron density map in the form of a three dimensional matrix of numbers. The crystallographer (or sometimes a computer program) uses this map through an iterative process to build a discrete atomic model in a piece-meal fashion. A refinement program is used at each stage to ensure the model is chemically reasonable and best describes the experimental data, the crystallographer then evaluates the model for improvement (Davis et al., 2003).   Molecular mechanics (MM) simulation (see 2.5.4), is typically used by the refinement program to perform a restrained energy minimization which is parameterized based on experimental data e.g., favourably observed bond rotations angles etc. (Leszczynski, 2012).

If the resolution of the X-ray diffraction data is higher than 1.5 angstroms (<1.5 Å) then it is probable that above 95% of the atomic model is directly a result of observed data. However, at lower resolutions below (>2.5 Å), the atomic model likely becomes undesirablly subjective. For example, at ~2 Å the isoelectronic nature of the sidechain atoms of asparagine (e.g., $N\delta$ and $O\delta$) and glutamine (e.g., $N\gamma$ and $O\gamma$) typically make their relative positions undeterminable from observation of the electron density because they are isoelectronic. Atom assignment is then based on inspecting the local hydrogen-bonding networks, these uncertainties add to the model-associated pitfalls during docking and scoring efforts (Davis et al., 2003).   X-ray diffraction data resolution has been given consideration in compound selection for this research - see 3.2.3.

2.4 Scoring Functions (applied to this research)

Scoring functions, may be classified as physical, or statistical (empirical/ knowledge based) potentials (Tuccinardi, 2009). The physical based scoring functions use force-field methods (see 2.5) and are the main target for improvement in this research. Empirical scoring functions use weighted energy terms where $\Delta G = \sum_i W_i \cdot \Delta G_i$. ΔG represents the change in Gibbs energy in the formation of the protein-ligand complex (see 2.5.1), and ΔG_i represents the energy terms such as hydrogen bonding, hydrophobicity, entropy, and desolvation energy in addition to standard VDW and electrostatic terms (Huang et al., 2010). The weighting coefficients, W_i are typically determined by fitting experimental activity values for known protein-ligand X-ray structures (Huang et al., 2010, Tuccinardi, 2009). The empirical scoring function used by the Glide Program in this research is known as Glide Score (Friesner et al., 2004) and is explained elsewhere (see 3.2.6.1). In keeping with previous work (Kontoyianni et al., 2003), although Glide is considered to have an empirical scoring function (see Table 2.1) we have also treated the EMODEL (Heavily Force-field weighted) and CVDW (Just Force-field) terms as separate scoring functions (see 3.2.6.2 and 3.2.6.3). Additionally, exploring the efficacy of an explicit polarization treatment (Ferenczy and Reynolds, 2001) upon them (see 3.2.7 and 3.2.8) while docking fragments. Knowledge based scoring functions are largely overlooked by this research. They borrow the inverse Boltzmann method from statistical mechanics (Physics) to obtain pair-wise potentials based on their occurrence frequency within an experimentally determined structural database (Huang et al., 2010).

2.4.1 The current state of Scoring Functions

The accuracy of the scoring function is a primary factor in the success of the docking experiment (hence the purpose of these studies on the effect of polarization). Study results can mislead when the final rankings are unable to identify true poses over decoys. Therefore, docking decoys can be useful for testing and improving docking algorithms. Hit-list decoys occur when a non-binder is predicted over a true binder for a target, and

geometric decoys occur when an incorrect ligand pose is predicted over a true one. The latter is more simple to address through evaluation of whether the highest ranking pose is also closest to the experimental binding mode (Graves et al., 2005), this is discussed further in subsequent chapters.

Table 2.1 A variety of docking programs and the strategies they use.

Program	Ligand flexibility	Protein flexibility	Scoring function
Autodock 4.0 (Morris et al., 1998, Huey et al., 2007).	Evolutionary algorithm	Flexible side chains	Force field
Gold (Jones et al., 1997, Verdonk et al., 2003).	Evolutionary algorithm	Protein side chain and backbone flexibility	Empirical score
FlexX (Rarey et al., 1996), FlexE (Claußen et al., 2001).	Incremental build	Ensemble of protein structures	Empirical score
Dock 6.2 (Kuntz et al., 1982, Ewing et al., 2001).	Incremental build	Protein side chain and backbone flexibility	Force field or contact score
Glide (Friesner et al., 2004, Friesner et al., 2006).	Exhaustive search	-	Empirical score
ICM (Abagyan et al., 1994), IFREDA (Cavasotto and Abagyan, 2004).	Pseudo-Brownian sampling and internal coordinate local minimization	Flexible side chains	Force field and Empirical score
QXP (McMartin and Bohacek, 1997).	Monte Carlo	-	Force field
Hammerhead (Welch et al., 1996).	Incremental build	-	Empirical score

EA dock (Zoete et al., 2009)	Evolutionary algorithm	Protein side chain and backbone flexibility	Force field

Adapted from (Zoete et al., 2009).

Many older scoring functions consider just the geometric fit of the ligand within the protein structure and make relatively simple calculations of electrostatic potential (ESP) and VDW radius in the treatment of energetics with little consideration for entropic and solvation effects (Kitchen et al., 2004).   However, this situation has improved in recent years but progress is ongoing. For instance, in discussion of entropy treatment, where empirical scoring functions commonly account for the favourable entropy of ligand desolvation, this is largely done through pair-wise atom-atom terms or buried surface area terms. These terms are parameterized by averaging over many different molecular environments to represent a measure of the hydrophobic contacts between protein and ligand, while this is often a valid treatment, it works best when different molecular environments do not vary significantly (Young et al., 2007).

MD simulations (Young et al., 2007), analysed largely using inhomogeneous solvation theory (Lazaridis, 1998), revealed two protein active site molecular recognition motifs where a 'gross underestimation' of binding affinity occurred. The first motif occurred when multiple water molecules were surrounding the ligand in an enclosure containing multiple hydrophobic side chains. The second motif occurred in the formation of multiple protein-ligand hydrogen bond interactions (stated between one and three), when the rest of the local environment was hydrophobically enclosed (Young et al., 2007). Results using an advanced treatment of hydrophobic enclosure, implemented in Glide XP (Friesner et al., 2006) are shown in Appendix A.

Another commonly used, and often useful scoring function metric, is to give torsional entropy (energy) penalties proportional to a ligand's rotatable bond number. However, there is a lack of compelling evidence that ligands lose the same amount of rotational and

translational energy on binding. Contradictions arise based on the level of ligand/protein conformation change required in binding. Therefore, favourable starting conditions can greatly improve a scoring function's ability to cope with the more difficult to quantify aspects of treating entropy (Chang et al., 2007).

These conditions arise when the initial conformation is highly comparable to the bound one (i.e., largely sampled with 6 DOF). Ligands with lower degrees of torsional freedom will tend to be planar, and have fewer rotatable bonds. Ligands achieve further reductions in torsional freedom through high ligand-pocket structural complementarity or 'snugness of its fit' (Chang et al., 2007), see ligand binding indices,1.2.4. Fragment molecules have largely planar cores and relatively few rotatable bonds, and efficient active site interactions making them a good starting point for DBVS, by potentially reducing the need of a scoring function to describe more difficult entropy based terms, see 1.2 and 2.5.2.

2.4.2 Scoring Fragments

An early fragment docking study (Verdonk et al., 2004) using the GOLD docking program (Jones et al., 1995, Jones et al., 1997) focused on CDK2 inhibitors. It showed that fragment size (<250 Da) actives were harder to retrieve than larger actives (>250 Da) from a database seeded with decoys. Despite this, docking was still considered useful, offering approximately a 14-fold enrichment of the top 1% of the database.

Most fragment libraries consist of between 1000 and 10000 focused compounds. It is unlikely that they represent all the chemo types available in the 300 000 commercially available fragments. This large gap in chemical space can potentially be explored through fragment SBVS. However there are additional concerns over the reliability of docking fragments, over and above those associated with drug-like molecules (Marcou and Rognan, 2007, Chen and Shoichet, 2009). These concerns are mainly the appropriateness of the scoring functions, which have typically been empirically parameterized using Lipinski-space molecules and the perceived promiscuity of fragments in their binding mode that is difficult to

predict (Babaoglu and Shoichet, 2006). The lack of studies comparing X-ray experiment binding to docked fragment results also contributed to this concern (Chen and Shoichet, 2009).

Where fragments make relatively few protein interactions they are efficient ones, and there is now evidence from several studies (Chen and Shoichet, 2009, Gleeson and Gleeson, 2009, Sándor et al., 2010, Verdonk et al., 2011), that native docking experiments can correctly predict the binding poses of fragment inhibitors and the interactions are largely accurate when compared with X-ray experiment structures. A reason attributed to this success, despite a lack of empirical parameterization with fragment training sets, is that the physics-based elements of the scoring functions have no specific bias against fragment compounds (Chen and Shoichet, 2009). Fragments (by virtue of their small size) are more sensitive to change in enthalpy ΔH when binding; they are typically very polar and water soluble relative to previous drug-like compounds. They also tend to be more rigid than leads, with relatively few rotatable bonds (Congreve et al., 2008). This reduces the difficulty in accounting for changes in entropy ΔS (Freire, 2008), for which the empirical elements of the scoring function are largely developed, making them easier to treat by docking (Chen and Shoichet, 2009), e.g., Glide performs better at posing and scoring molecules with fewer rotatable bonds (Friesner et al., 2004).

Even the promiscuity of fragments may be a virtue and perhaps arises from their ability to adapt to different environments within proteins, forming different hydrogen bonds with the receptor (Chen and Shoichet, 2009). The highly polar nature of fragments has been exploited in the evaluation of an explicit polarization treatment (Ferenczy and Reynolds, 2001) upon the OPLS 2005 force-field (Banks et al., 2005) and is a main theme of this research and is outlined in sections 3.2.7 and 3.2.8, which has been applied to the rescoring of the Glide Docking Program (Friesner et al., 2004), considered an industry standard.

2.5 Force-Field based Methods

2.5.1 Gibbs free energy calculation at chemical equilibrium

The reversible interaction between a protein (P) and a ligand (L) that forms the protein ligand-complex is $P + L \leftrightarrow PL$. When considering the Gibbs energy change (ΔG) from forming this complex, it is useful to think of the relationship between the total free energy and the reaction mixture as it approaches equilibrium. When experimental data is available this potentially allows the calculation of K_d from $\Delta G = \Delta G^0 + RT\ln Q$. In terms of $\Delta G = \Delta G^0 + RT\ln K$, ΔG^0 is the standard free energy of the interaction. Standard refers to the conditions where partial pressure of any gases involved in the reaction is 0.1 MPa (mega pascals), and any aqueous solution involved has a concentration of 1 M (Molar). R is the gas constant (R=1.99 cal mol^{-1}K^{-1} / 8.184 J mol^{-1} K^{-1}), T is the absolute temperature K, is the equilibrium constant for a specific equilibrium mixture, which is equal to the reaction quotient Q at thermodynamic equilibrium. If an equilibrium mixture is mainly reactants, K<1 and ΔG^0 >0, if the mixture is mainly products K>1 and ΔG^0 <0, but the mixture contains comparable amounts of both reactants and products when K=1, lnK =0 and ΔG^0=0 (Fay and McMurry, 2012). Simplifying $\Delta G^0 = -RT\ln(K_{eq})$ now allows direct calculation of ΔG^0 for the protein-ligand complex from experimental values of K. or the reciprocal of K (i.e. K_d). Substituting therefore gives $\Delta G^0 = RT\ln(K_d)$ which in principle allows direct calculation of K_d (de Azevedo et al., 2008).

2.5.2 Entropy and Enthalpy

ΔG in forming of the protein-ligand complex is given by $\Delta G^0 = \Delta H^0 - T\Delta S^0$. It can be considered in terms of variations in standard enthalpy (ΔH^0) and entropy (ΔS^0). Ideally, both should contribute favourably to binding in order to achieve `extremely high or super affinity` in the low nano-molar to pico-molar range (Freire, 2008, Young et al., 2007). The balance between ΔH^0 and ΔS^0 of a compound during binding is the thermodynamic signature. The easiest and most popular lead optimization strategy has historically been to target favourable

ΔS^0 gains in ΔG. The two major contributors to ΔS in ΔG are desolvation ΔS^0 and conformational ΔS^0 (Freire, 2008). The former is commonly cited to be the primarily driven by the hydrophobic effect (Young et al., 2007). However, this is considered only to be true if all the surfaces are large enough As the force behind the assembly of clusters of hydrophobic particles in water, although involving entropy contributions is also considered to be driven by the difference in solvation free energy between entropically dominated small molecules contribution with linear growth in their free energies relating to their excluded volume, and the enthalpically dominated large surface contributions with linear growth relating instead to their solvated hydrophobic surface area (Chandler, 2005). The latter conformational contributor almost always incurs an ΔS^0 energy penalty to ΔG, caused by loss of degrees of freedom for the protein-ligand complex. Typically the interaction energy is entropically optimized by the addition of hydrophobic groups to the ligand within the framework of 'Lipinski's rule of 5' and conformational restraints are engineered to maintain the similarity between the bound and free ligand conformations e.g., more planar, fewer rotatable bonds (Freire, 2008). These are typical fragment-like considerations too (see 1.2).

The complete thermodynamic data for two classes of drug from the first in marketed drugs to the best in class compounds, emerging after over a decade of development for HIV-1 protease inhibitors and statins, showed higher selectivity, better potency, pharmokinetics and/or drug resistance profile as improvements were made.   It has been noted that the best in class drugs also had a favourable enthalpy contribution to the thermodynamic signature (Freire, 2008). Improving affinity through the primary mechanism of adding hydrophobic groups decreases drug solubility, permeability and consequently leads to a poorer **absorption, distribution, metabolism, and excretion – toxicity** (ADMET) profile. The improvement in the type of interactions determining binding when ΔH^0 is also optimized has been shown to lead to a better drug profile. Although now an important consideration is that optimizing ΔH^0 interactions is a notoriously harder strategy (Freire, 2008).

The difficulty arises from a conflict in the contributions to ΔH^0. Firstly, favourable ΔH^0 comes largely from formation of hydrogen bonds and VDW contacts (where perfect geometric fit between protein and ligand optimizes ΔH^0). While unfavourable ΔH^0 is largely a result of the energy penalty associated with the desolvation of polar groups ~8-9 kcal/mol at 25°C. Hydrogen bond strength is maximal when the distance and angle between acceptors and donors are optimal. A well placed hydrogen bond that is stronger than the one formed with water by either the ligand or protein can have a favourable ΔH^0 of ~4-5 kcal/mol. (Freire, 2008)It is estimated that for every 1.4 kcal mol^{-1} of ΔH^0, binding affinity is increased/decreased by an order of magnitude when $\Delta S^0 = 0$. Therefore the maximal increase in binding affinity is 1000-5000 fold for a well placed hydrogen bond (Freire, 2008).

However, this favourable ΔH^0 can be negated by compensating entropy loss primarily due to conformational structuring induced by the hydrogen bond. This too can be managed by directing several hydrogen bonds against the same donor/acceptor so the first one pays the entropy penalty and subsequent ones bind to the structured region. SBDD can be used to visualize the general location of hydrogen bond donors and acceptors which can then be more precisely defined by experiment e.g., effect of hydrogen bond functionality placement can be assessed by isothermic calorimetry (ITC) (Freire, 2008). ΔH^0 contributions such as VDW forces and electrostatics can largely be calculated using force-field based methods (de Azevedo et al., 2008) e.g., MM (see 2.5.4). However, the accuracy required for modelling of H-bonds to the necessary tenths of an angstrom using SBDD alone is beyond the current state of the field (Freire, 2008). To address this somewhat, we explore the potential of an improved description of electrostatics and hydrogen bonding called Distributed Multipole Analysis (DMA) (Stone, 1981) in Chapter 4.

2.5.3 The Quantum Mechanical view of Intermolecular forces.

Molecular Dynamics (MD) simulation has principally aided in solving chemical problems that involve the 3D structure of the macromolecule such as protein folding pathway prediction

(using restricted experimental data) and refinement of structures derived from X-ray diffraction data (where resolution is largely unachievable by experiment alone). But MD, despite its semi-empirical nature, is also usefully applied to the problems of protein-ligand complexation and molecular recognition (Leszczynski, 2012). Molecular mechanics (MM) and MD use classical (physics) potential energy functions, that can reproduce the basic atomic level features of the protein energy landscape (Ponder and Case, 2003). We discuss amongst the simplest and widely used potential energy function below (see 2.5.4).

However, it is the quantum mechanics (QM) intermolecular interaction theories, that best describe the physical phenomena responsible for the attraction and repulsion between molecules.   These phenomena, which arise most importantly from the electrostatic interactions between all particles (e.g. electrons and nuclei) within two interacting molecules (Stone, 1996), should not be overlooked here.

London (London, 1937) proposed that interaction energy between molecules is the sum of four basic components. The first three, electrostatic, induction, dispersion are considered 'long range' effects, where the approximate behaviour of their interaction energy follows an inverse power function of separation distance r (Stone, 1996, Cieplak et al., 2009). The remaining 'short range' component to the interaction energy is exchange or exchange-repulsion.   At short distance molecular wave-functions overlap significantly, and electron exchange becomes possible between molecules.   The exchange contributions have displayed an exponential dependence on r (Cieplak et al., 2009). Two other effects can arise at 'long range', these are resonance and magnetic effects, but as they have little importance in the context of intermolecular forces, we can avoid discussion of these for now (Stone, 1996).

The 'long range' contributions survive at large separation, yet remain present at short distance even when there is a strong molecular overlap. Thus, when looking at the physical phenomena of the electrostatic interaction, it remains finite until the nuclei come into contact

and is well defined at any distance in an unperturbed molecular charge distribution. The approximation for modelling long range interaction with inverse powers of r introduces two immediate errors. The first is a 'penetration error' where these 'long-range' series converge to a point instead of extend into space at insufficient distances of r>0. The second is a common 'truncation error' where in practice it is required to have a finite number of points or terms in the series (Stone, 1996).

The four basic interaction energy components each have a specific interpretation in terms of molecular properties. The electrostatic effects involve the static charge distributions, including permanent electric multipole moments of the classical interaction between two molecules. Electrostatic effects can be either attractive or repulsive, and are pairwise additive (Stone, 1996). Induction is an important part of polarization energy and is defined as, 'one molecules permanent multipoles interacting with the multipole moments induced in another molecule' (Cieplak et al., 2009). Induction is always attractive, and strongly non-additive, as the fields of other neighbouring molecules may either reinforce or cancel each other (Stone, 1996). Dispersion involves the mutually polarized fluctuations of the charge distributions of two molecules (Cieplak et al., 2009), as the electrons move. The electrons' motions become correlated in a way that favours low energy configurations, with the electron correlations becoming stronger as the molecules approach. This gives rise to an attractive interaction, with the average effect being a lowering of energy (Stone, 1996).

The 'short range' energy exchange or exchange-repulsion component, is thought to comprise two effects, the weaker effect is the attractive one arising from exchange-induction and exchange-dispersion. The other effect which dominates overall at short range is a repulsive effect. This arises from the energy cost of the Pauli anti-symmetry requirement, that same-spin electrons must not occupy the same orbital. The above energy contributions are summarized in table 2.2, it is recommended to think of the charge-transfer term, as part of the induction energy rather than a separate component, at least in terms of perturbation theory (Stone, 1996, Stone and Misquitta, 2009).

There are two general quantum mechanical approaches to dealing with the calculation of intermolecular interactions: the perturbation theory, and the supermolecular theory. In brief, the perturbation approach, in its simplest form can be referred to as 'the polarization

Table 2.2. Contribution to the interaction energy between molecules. Taken from (Stone, 1996).

Contribution	Additive	Sign	Comment
Long-range ($U \sim R^{-n}$)			
Electrostatic	Yes	$\pm$	Strong orientational dependence
Induction	No	-	
Dispersion	Approx.	-	Always present
Resonance	No	$\pm$	Degenerate states only
Magnetic	Yes	$\pm$	Very small
Short-range ($\sim e^{-\alpha R}$)			
Exchange-repulsion	Approx.	+	Dominates at very short distance
Exchange-induction	Approx.	-	
Exchange-dispersion	Approx.	-	
Charge transfer	No	-	Donor-acceptor interaction

approximation', where the perturbation expansion parameter (ξ) is introduced to the

Schrodinger equation. It is perhaps enough to say here, that the ξ in terms of the total

interaction energy can be viewed as an infinite series of corrections called polarization

energies. Where the largest first order polarization energy is the electrostatic energy, which

incorporates the most important exchange term, accounting for ~90% of the total exchange

energy (Jeziorski et al., 1994). The second and third order terms are sums of induction and

dispersion energies. It is at this point multipole expansions of the operator V can be applied

(V is intermolecular interaction operator). These expansions lead to energy components

dependent on the separation distance r (Cieplak et al., 2009).

The supermolecular approach is used to define the difference between the energy of the dimer and the energies of the two monomers as the total interaction energy. An important consideration here is that interaction energy between two molecules is frequently in the order of several kcal mol^{-1}. However, this may be five or six orders of magnitude less than energy of the dimer obtained quantum mechanically from using the supermolecular approach. Therefore the calculation level, in terms of theory sophistication and basis set quality effects the results (Cieplak et al., 2009). Approaches to quantify supermolecular energies into clearly defined energy contributions, highlight the presence of the basis set superposition error (BSSE), (Boys and Bernardi, 1970). BSSE entails the non-physical lowering of the monomers' energy in a dimer calculation, since each monomer uses its partner's basis set to lower its own energy.

While in supermolecular theory, the charge-transfer term is commonly accepted in partitioning interaction energies, but suffers from being contaminated by the BSSE. It is harder to define charge-transfer within perturbation theory. However, when extracted from the induction energy component (which contains a polarization term), it is less dependent on basis set (Stone, 1996). The term also becomes small to negligible when proper handling of exchange repulsion is taken into account (Stone and Misquitta, 2009).

Hybrid QM/MM approaches contain features of both perturbation theory and the supermolecule approach. However, in the traditional approach, there is no charge transfer or BSSE and only the QM part is polarized. Empirical force fields typically consist of electrostatic, repulsion and attraction terms. They all employ various levels of approximation and error compensation. After general discussion in the next section, we can further explore their effectiveness as potentials (see 2.5.4 to 2.5.9).

2.5.4 Molecular Mechanics Force Fields

A force-field consists of classical potential energy expressions and their associated adjustable parameters (Jorgensen and Tirado-Rives, 2005). For a system of N interacting

particles, the total potential energy of that system can be written as a series of n-body

potentials (Elrod and Saykally, 1994, Hodges et al., 1997, Dahlke and Truhlar, 2006), where

V_n is the *n*-body interaction term:

$$V(r_1, r_2, \dots, r_N) = \sum_{i<j}^{N} V_2\left(r_i, r_j\right) + \sum_{i<j<k}^{N} V_3\left(r_i, r_j, r_k\right) + \dots + \sum_{i<j<k,\dots,z}^{N} V_n\left(r_i, r_j, r_k, \dots, r_z\right)$$

(2.2)

Truncating of (Eq. 2.2) after the first term in the series however leads to the pair-wise

additive approximation that is used in most practical molecular mechanics-based

applications: $E(\vec{X}) = \sum_{a<b} E_{ab} + \sum_a E_a^{int}$. When $E(\vec{X})$ refers to the total potential energy of

the above system with coordinates $\vec{X}$. $E(\vec{X})$ is equal to the sum of the intramolecular

interaction energies between the components $\sum_{a<b} E_{ab}$ plus the sum of the intermolecular

energies of the components $\sum_a E_a^{int}$ (Jorgensen and Tirado-Rives, 2005). The subscript a<b

ensures that interactions are not double-counted. In principle, inclusion of polarization using

induced dipoles involves truncating Eq. 2.2 after the second term, even though our

implementation of polarization as a set of induced charges means that the energy only has to

be evaluated with pair-wise additive terms (see 2.5.9).

Condensed-phase simulations largely adopt the pair-wise additive approximation

(Jorgensen and Tirado-Rives, 2005), although considered less quantitatively accurate than

the many-body potential description, it is considered valid for many measurable properties,

(Elrod and Saykally, 1994, Dahlke and Truhlar, 2006). Nevertheless, this is an approximation

that many consider should be avoided (see 2.5.8 to 2.5.9).

Molecular mechanics (MM) simulation as eluded to (2.5.3) uses classical physics.

MM applies Newtonian mechanics to the system without considering electrons which are too

small to obey the rules of classical mechanics. The convention is to treat the system

involving the protein and ligand as a series of hard spheres (atoms) connected by springs

(covalent bonds). The simplest system for this is described in Figure 2.2 (de Azevedo et al., 2008).
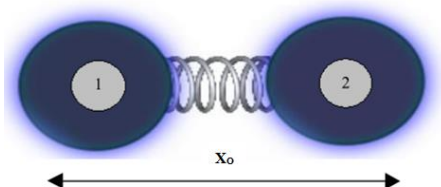
Figure 2.2. The simplest system of two covalently bound atoms represented as hard balls. Adapted from (de Azevedo et al., 2008).

In Figure 2.2, atom 1 is fixed at position A, atom 2 is free to move to a new position x along the x-axis from position X_0, which is the equilibrium point of the spring. Where potential energy, U, is 0 at $x = x_0$ (Figure 2.2). When atom 2 in this system is moved the spring is compressed or stretched at the x position.

This is best described by Hooke's law for simple (classical) harmonic oscillators, namely F = -kx. When an atom is moved from its equilibrium position the spring acts with a conservative (path independent) force as in Newton's 2nd law, namely $\vec{F} = m\vec{a}$. Here a force, F, on the atomic mass pulls it back towards its origin, where k is the spring constant (Hrabovsky and Susskind, 2012, de Azevedo et al., 2008), this gives:

$$U(x) = - \int_{x_0}^{x} F.dx = - \int_{x_0}^{x} (-kx)\, dx = \frac{1}{2}(x - x_0)^2$$

(2.3)

The force-field energy potential U(**r**) is given by a sum of terms, e.g., $U(r) = E_{bond} + E_{angle} + E_{tors} + E_{vdw} + E_{el}$, with each describing the energy required to distort a molecule in a

specific fashion (Jensen, 2007). The most common interaction potential to describe (U(***r***))

(Weiner and Kollman, 1981) can be expressed as (de Azevedo et al., 2008):

$$U(r) = \frac{1}{2} \sum_{bond} k_b(b_i - b_{i,o})^2 + \frac{1}{2} \sum_{angle} k_\theta(\theta_i - \theta_{i,o})^2 + \frac{1}{2} \sum_{torsion} V_n[1 + \cos(n\omega - y)]$$

$$+ \sum_{i=1}^{N} \sum_{j=i+1}^{N} k_{ij} \left\{ 4\varepsilon \left[\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6} \right] + \frac{q_i q_j}{4\pi\varepsilon_o r_{ij}} \right\}$$

$$(2.4)$$

This potential is common to the AMBER type or 'Class 1' force-fields. It can describe the

protein energy landscape in an effective but somewhat limited way (Ponder and Case, 2003).

The bonded interactions (Figure 2.3) are those over bonds (1-2 interactions), angles (1-3

interactions), and torsion (1-4 interactions). The term 'improper' torsion applies when the

four angle-defining atoms are not all connected covalently. These improper torsions are a

functional form for planarity enforcement around sp^2 central atoms. The final pairwise sum of

atoms with subscripts i and j, excludes the 1-2 and 1-3 interactions and often uses separate

parameters for 1-4 interactions as compared with atoms separated by three or more bonds

(Ponder and Case 2003).
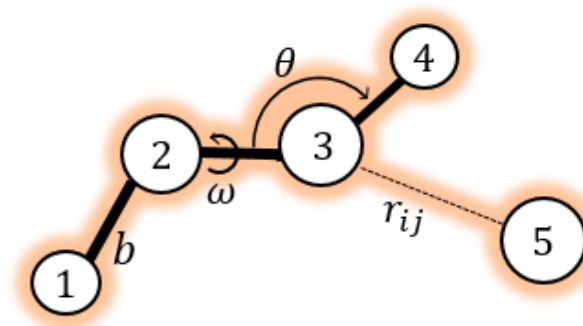
Figure 2.3. Schematic showing the terms used to calculate potential energy. Adapted from (Ponder

and Case, 2003). Solid lines apply to bonded terms and the dashed line represents non-bonded terms.

The mathematical forms of the bonded terms b, θ and ϖ and non-bonded terms r_{ij} are given in Eq. 2.3.

Dispersion and exchange forces, often referred to as the 'van der Waals' terms, are typically modelled by the Lennard-Jones 6-12 potential, which is considered rather simplistic (Cieplak et al., 2009), (see 2.5.5). The bonded terms (see Figure 2.2, 2.3) that form the first half of (Eq. 2.3) applied to the intramolecular potential, where the bond stretching (b) and angle bending (θ) terms usually follow the 'simple Hooke's law' dependencies from (Eq. 2.3), but modern force-fields, typically include cubic or quartic terms in a Taylor series to reflect anharmonicity, while a Fourier series (Eq. 2.5) is applied to each torsional angle (ω), (Jorgensen and Tirado-Rives, 2005, Maple et al., 2005). Additionally, descriptions for improper torsions or out of plane bending and a number of cross terms most importantly bond/angle are more commonly added to the force field, blurring older 'class 1' and 'class 2' distinctions (Jensen, 2007, Leszczynski, 2012).

$$E_{torsion} = \frac{1}{2}\Sigma\{V_1[1 + \cos(\varphi)] + V_2[1 - \cos(2\varphi)] + V_3[1 + \cos(3\varphi)] + V_4[1 - \cos(4\varphi)]\} \qquad (2.5)$$

The non-bonded terms given in the second half of (Eq. 2.3) are the Coulomb $E_{el} = \frac{q_i q_j}{4\pi\varepsilon_o r_{ij}}$ and the afore mentioned Lennard-Jones ,$E_{vdw} = 4\varepsilon\left[\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6}\right]$, potentials, which involve the atoms separated by 3 or more bonds (Jorgensen and Tirado-Rives, 2005).

2.5.5 van der Waals non-bonded interactions

The dispersive and exchange-repulsion interactions between atoms can be described in a highly non-trivial way using quantum mechanics calculation methods that require electron correlation and large basis sets. However, force-fields typically model the interatomic potential curve of the van der Waals non-bonded energy (E_{vdw}) using only a rapid calculation method using a simple empirical expression (Leach, 2001). The Lennard-

Jones potential (Jones, 1924), written here as $E_{vdw} = 4\varepsilon \left[\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^{6} \right]$, has become a

popular and fairly straight forward way to satisfy the functional conditions of the VDW non-

bonded energy (E_{vdw}). The E_{vdw} describes the repulsion (Fierz and Pauli, 1939) and

attraction (London, 1937) between atoms that are not directly bonded (Eq. 2.6):

$$E_{vdw}(R^{AB}) = E_{repulsive}(R^{AB}) - \frac{C^{AB}}{(R^{AB})^6}$$

(2.6)

$E_{vdw}(R^{AB})$ is very positive over small distances, reaches it minimum when two atoms are

slightly touching and it heads towards zero as distance increases. Although the repulsive

interaction $E_{repulsive}(R^{AB})$ is not directly derivable, it is a sufficient approximation that it

heads towards zero faster than the $(R^{AB})^6$ term as distance goes to infinity (Jensen, 2007).

The mean of atomic radii (R^{AB}) is modelled using constants C_{AB} (or σ_{ij}, Eq. 2.4; and

C_{AB} Eq. 2.6) that act as fitting parameters (Jensen, 2007). Indeed, the majority of the

constants in equation 2.4 are fitting parameters that are largely derived from experiment. An

approximation inherent in this work is that these parameters are reasonable and that they are

transferable between similar systems (Jensen, 2007).

The Lennard-Jones potential contains just two adjustable or fitting parameters – the

collision diameter σ_{ij} at which the separation gives an energy of zero and the well depth ε,

which is the minimum of the potential (Leach, 2001). The Lennard-Jones potential uses

$\left(r_{ij} \right)^{12}$ from divisor above for the repulsive part. Theoretically this has no strong basis, where

more accurate quantum mechanical exchange-repulsion energy depends on overlapping of

electronic densities, which displays an exponential dependence on r, therefore suggesting

an exponential power form (Cieplak et al., 2009). In contrast, the twelfth power form

produces a potential curve considered rather steep for modelling systems such as

hydrocarbons, its popularity instead arises from rapid calculation properties in reducing computational expense (Leach, 2001).

The short-range repulsive $\left(r_{ij}\right)^{12}$ term is calculated directly from squaring the $\left(r_{ij}\right)^{6}$ term. The relatively fast calculation of the $\left(r_{ij}\right)^{6}$ can also be achieved from the square of the distance, which does not require an otherwise typical square root calculation (Leach, 2001). As $\left(r_{ij}\right)^{12}$ has no physical basis, different powers have been used to produce a less steep curve such as 9, 10 to improve modelling of E_{vdw} against experimental results (Leach, 2001, Cieplak et al., 2009)   Some attempts at better correlation with experimental data, also propose alternative calculation schemes. These attempt to recreate performance with comparable rapid calculation properties, such as a buffered 14-7 potential (Halgren, 1992, Halgren, 1996a, Halgren, 1996b).

The longer range attractive $(R^{AB})^{6}$ term has the same dependence as the (dipole-dipole) London dispersion energy, which for two particles with polarizabilities α is proportional to $-\alpha^2/r^6$. Lennard Jones parameters typically don't use known values of α, but this interaction provides a way to parameterize to include polarization in an average way in non-polarizable models (Rick and Stuart, 2003). This same power law relationship, used in theoretical treatments of dispersion energy (e.g., The Drude model), gives $\left(r_{ij}\right)^{6}$ a theoretical basis (Leach 2002). Attractive power terms greater than $\left(r_{ij}\right)^{6}$ are rarely used, but they have theoretical basis in the multipole expansion of induction and dispersion energy components. It is noteworthy, that popular terms employed to describe intermolecular interactions in classical force fields such as r^{12}, r^6 and r, absorb many of the energy components (such as induction) in a way that is effective, yet not always correct (Cieplak et al., 2009).

2.5.6   Electrostatic potential and Coulomb's Law.

The van der Waals interactions and the electrostatic interactions are considered 'through-space' non-bonded interactions as they are not dependent upon a specific bonding relationship between atoms. They are instead normally modelled in the general form of a function giving some inverse power of distance (Leach, 2001) see Figure 2.4 and 2.5.5. Applied here, this means that the electrostatic force/energy between two particles gets smaller as they get further apart see Figure 2.4.
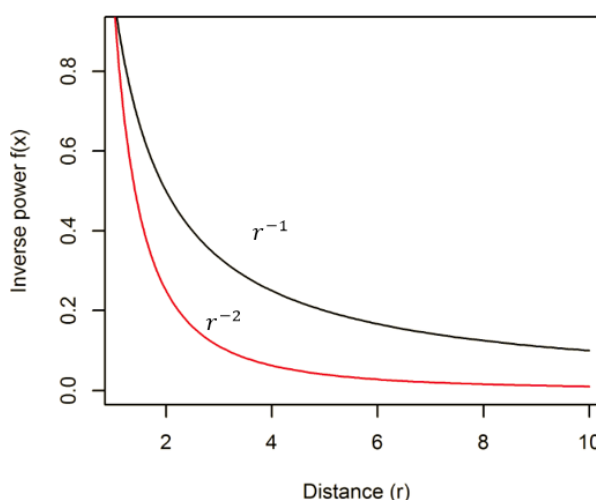
Figure 2.4. The variation of potential energy V=f(r) as a function of r. f(r) is typically used in describing the interaction between two molecules, r^{-1} black line for electrostatic potential energy interaction, and r^{-2} red line the electrostatic force. Higher powers can be used in describing a multipole series e.g., r^{-3} describes the charge-dipole interaction.

The electrostatic (or Coulomb) potential energy (E_{el}) is the potential or stored energy of a charged object due to its position in a force field. When considered as just an interaction energy between two charges it can be given simply as $E_{el} = \frac{q_i q_j}{4\pi\varepsilon_o r_{ij}}$, e.g., it varies inversely with distance. The value $q_i q_j$ is the product of the charges on the two particles identifiable from their individual subscripts i and j.   The term r_{ij} gives the distance between point charges q_i and q_j. The term $4\pi\varepsilon_o$, refers to where 4π is the isotropic spherical nature of the charge density cancelled by the field strength (both having r^2 in the divisor).   The E_{el} can be

described by Coulomb's law in terms of an electrostatic potential (v). Where the coulomb potential for two interacting molecules or different parts of the same molecule is pair-wise additive giving, $v = \sum_{i=1}^{N} \sum_{j=i+1}^{N} \left\{ \frac{q_i q_j}{4\pi\varepsilon_o r_{ij}} \right\}$ from Eq. 2.4.

There is an attraction between oppositely charged objects, and that gives rise to a force between them. Forces are typically proportional to the inverse square of the distance between two objects, and when here those objects are charges we call this the electrostatic or Coulomb interaction (Keeler and Wothers, 2003). The electrostatic force (F) between two particles carrying a charge, acting upon each other is given by $F = k_e \frac{q_i q_j}{r^2}$. This formulation because of the 'inverse square rule' closely mimics Newton's law of gravitation. This 'inverse square rule' can be visualized as a cone projecting from a point, with a circular cross section where the vertices of the distance proportional square fit perfectly. The one important difference from the law of gravitation is that the two types of charges commonly called positive and negative affect the action of the force. When making an arithmetic comparison of the product of charge $q_i q_j$, ++ or -- charges give a positive force which is repelling, and unlike charges +- or -+ give a negative attractive force. The convention is to assign the positive charge to protons and negative charge to electrons. The term $k_e = 1/4\pi\varepsilon_o$ relates to the permittivity of free space see 2.5.7 (Leach, 2001).

The electrostatic potential is defined as the force acting on a unit positive charge, placed at that point. We can use Figure 2.5 to demonstrate the electrostatic potential $v(r)$ in action at point P. Forces derived from a potential are conservative forces; the negative gradient of energy is the force. The work done (W) by a conservative force is $W = -\Delta U$, where ΔU is the change in potential energy associate by force, negative indicating that work done against and force field increases potential energy. Transferring the work done into point P, we get $v(r) = \frac{1}{4\pi\varepsilon_0} \left(\frac{q_1}{r_1} + \frac{q_2}{r_2} + \frac{q_3}{r_3} + \frac{q_4}{r_4} + \frac{q_5}{r_5} \right)$. Where the $k_e = \frac{1}{4\pi\varepsilon_0}$ term described previously is multiplied by the sum of individual charges divided by their distances from point P.
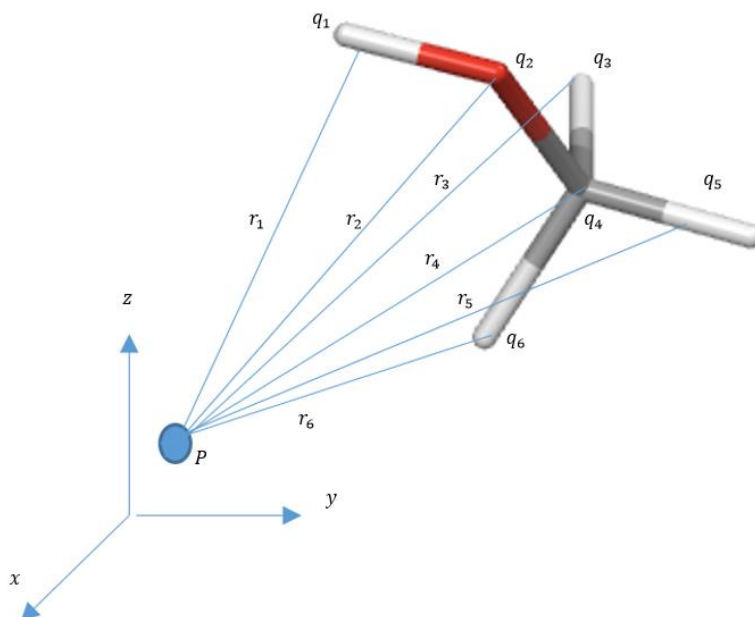
Figure 2.5. Hypothetical Methoxyl system demonstrating electrostatic potential $\phi(r)$ in action at point P.

The electrostatic potential is directly observable from the quantum mechanical wavefunction, it is considered to be the sum of contributions from nuclei and electrons. It has continuous properties, that is hard to represent using an analytical function, and therefore requires discrete representation for analysis numerically. The charges are typically constrained to their nuclear centres, in the lowest and most widespread 'point charge' approximation for the consideration of electron density. This is limited to the assignment of positive and negative partial charges, q_i, to an interaction site on each atom, which along with their LJ fitting parameters result in typical class 1 all atom (AA) force-field models (Leach, 2001), using the non-bonded terms $\sum_{i=1}^{N} \sum_{j=i+1}^{N} \left\{ 4\varepsilon \left[\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6} \right] + \frac{q_i q_j}{4\pi\varepsilon_o r_{ij}} \right\}$ from Eq. 2.4 (de Azevedo et al., 2008). We explore the limitation of these approximation further below (see 2.5.8 to 2.5.9).

2.5.7 Dielectric Modelling

The formula for electrostatic energies, forces and potentials include a term relating to the permittivity of free space (ε_o), or vacuum permittivity (8.854 x 10^{-12} F m^{-1}), in the gas phase it's unit is farads meter^{-1} (F m^{-1}) (Keeler and Wothers, 2003). However, when it is desired to mimic solvent effects which dampen electrostatic interactions (without the inclusion of explicit solvent molecules), a different dielectric model can be used in the equations for electrostatic interactions. The simplest dampening model applies a dielectric constant, also called the relative permittivity (of free space), which is a relative scaling factor when compared to vacuum. This relative permittivity (ε_r) models the different properties of an insulating material to increase capacitance or stored charge, when placed between the plates of a capacitor (Fischer-Cripps, 2014). In context here, it is more appropriately the factor that the electric field between charges is dampened relative to a vacuum by using an appropriate value for ε_r in Coulomb's law equation (i.e. $\varepsilon = \varepsilon_o\varepsilon_r$). As ε_o appears in the divisor the greater the value of ε_r the lower the magnitude of the interaction between two charges, thus *in vacuo* charges have the greatest electrostatic effect (Leach, 2001).

For instance, paper has an ε_r value of 2-3, and glass an ε_r value of 8 (Fischer-Cripps, 2014). It may be set to a value or magnitude of 80 to model properties of water, allowing for an implicit treatment of solvation (Jensen, 2007, Leszczynski, 2012), but ε_r is only set/scaled to 1, the value for a vacuum, when all atoms including water are treated explicitly. In the condensed phase (where we find solids and liquids together), when we look at a large solute (i.e. protein) in solvent, the shortest distance between two charges could pass through solute rather than the solvent. This makes it difficult to choose an appropriate value for ε_r, one solution is to choose an arbitrary intermediate value between water and vacuum, then adjust it to best fit the desired modelling properties required from the force-field, this is also a very simplistic way to include an averaged treatment of polarization implicitly (Leach 2001).

Another approach introduces a 'distance-dependent dielectric', where the dielectric is dependent upon the distance of charged species separation. The simplest model is to make the relative permittivity ε_r (i.e. $\varepsilon = \varepsilon_o \varepsilon_r$) proportional to the distance r. It is worth noting that this electrostatic potential $v(r) = \frac{1}{4\pi\varepsilon_0} \frac{q_i q_j}{r^2}$ with r and ε_r added to the divisor, now looks quantitatively identical to the inverse square rule applied to the electrostatic force between two particles. The introduction of the invisible ε_r in the divisor product is the difference. This simplest distance dielectric, although considered a slight improvement on the previous electrostatic dampening model (without the extra divisor r), is not normally recommended as it is considered without physical basis (Leach, 2001). An improvement at this level of solvation continuum complexity is to add a sigmoidal (dielectric) function for distance, where at short range relative permittivity is low and then approaches bulk values at longer distances (Smith and Pettitt, 1994). There are many solvent continuum, and explicit solvation schemes to model the free energy of solvation of a solute. The polarization term is often considered of major importance. It adds sophistication at each level of the modelling complexity hierarchy, particularly when treated in an explicitly (Leach, 2001). We look at the influence of polarization below (see 2.5.8 and 2.5.9).

2.5.8 The form and limitations of the typical force fields

The analytic formula used to calculate energy as the sum of bonded and non-bonded terms, can be traditionally separated into two main groups Class 1 and Class 2. The diagonal or Class 1 force fields, can include amongst the most popular of their number: OPLS (used heavily in this research) (Jorgensen et al., 1996, Kaminski et al., 2001, Banks et al., 2005, Friesner et al., 2004), CHARMM (MacKerell et al., 1998), ECEPP (Zimmerman et al., 1977), GROMOS (Ott and Meyer, 1996, Schuler et al., 2001), and AMBER (Weiner and Kollman, 1981, Cornell et al., 1995, Wang et al., 2001, Duan et al., 2003). The last to be mentioned AMBER (Assisted Model Building with Energy Refinement) has also become synonymous with the naming of this type of Class 1 force field (Cieplak et al., 2009). We

have discussed the functional form of the AMBER type potential (see 2.5.4 to 2.5.7) above, and can consider this a general start point for this group of force fields. Force field research has been around for many decades now, with each new generation borrowing heavily from the last (Leszczynski, 2012). It has proven technically difficult to judge a 'gold standard' for direct force field comparison in the condensed phase environment. It has been observed that condensed phase experimental measurements involve substantial thermal averaging of conformers. However, the conformational sampling of protein (or peptide) simulations have until recently lacked meaningful convergence to a level where deviations can be attributed to force field deficits alone (Ponder and Case, 2003).

The Class 1 force-fields that apply a fixed point charge model have displayed a high level of convergence in their electrostatic models (Ponder and Case, 2003). It is apparent that they have just minor adjustments to the number of energy terms and treat van der Waals and electrostatic separately which are scaled in magnitude (Cieplak et al., 2009) (Eq.2.4). Class 1 force fields have evolved to treat large systems (e.g., DNA or protein), and have made high levels of approximation and kept their functional from relatively simple (Ponder and Case, 2003, Jensen, 2007, Cieplak et al., 2009). It has been proved that the qualitatively 'correct' behaviour of these force fields is strongly linked to the type of simulation and desired results. Historically force-fields and the computer codes that implement them have shown a less than ideal, yet important interdependency; this has become less true with time (Ponder and Case, 2003). There has also been a blurring of the traditional force field groups, with Class 1 bonded terms now including cross terms, and expansions to cubic and quartic level terms that reflect anharmonicity, as these force fields have evolved (Banks et al., 2005, Leszczynski, 2012).

We now briefly consider Class 2 force fields, which typically have more complicated and varied functional forms for their interaction energy potentials. A traditional distinction is that Class 2 force fields contain higher order bonded terms, and cross terms, as we have just seen this is no longer a straight forward distinction. The class 2 force fields offer varied

functional forms of their electrostatic models, e.g., employing an exponential type potential (Buckingham and Corner, 1947) for E_{vdw} (see 2.5.5), or point dipoles located at chemical bonds (Allinger et al., 1989). Class 2 force fields aim to reproduce with high accuracy the interactions of small to medium size molecules. The descriptive level of interaction is also expanded beyond the calculating of relative energies and geometries, to also perform vibrational analyses (Cieplak et al., 2009). Some of the popular Class 2 force fields are: CFF (Niketic and Rasmussen, 1977), UFF (Rappé et al., 1992), MMFF (Halgren, 1999), and MM4 (Allinger et al., 2003). The moniker 'Class 3' has been offered to force-fields allowing parameters that depend on neighbouring atom types (e.g., hyperconjugation) and include polarization effects (Jensen, 2007). This may just be an evolution in the Class 2 approach e.g., just as Class 1 force-fields have evolved in their parameterizations and bonded terms. Perhaps a clearer way of thinking of these newer developments is just going beyond the 'minimalist approach' (Leszczynski, 2012).

The emphasis on this 'minimalist approach' is to achieve a good enough level of approximation, while cleverly limiting computational expense, based on current performance vs technical capabilities at the everyday workstation level (Leszczynski, 2012). We eluded to (see 2.5.5) the electrostatic potential as a continuous property that requires discrete representation for numerical analysis. The typical All Atom (AA) model charge distribution involves the reproduction of a molecules electrostatic properties, through careful assignment of fractional point charges restricted to nuclear centres (Leach, 2001). The partitioning of these fractional charges can be done in huge number of ways that reflect the reproduction of the same desired experimental properties. Some considerations in the force field and parameterization design are: the emphasis on which experimental data to reproduce, the computational efficiency, the transferability of parameters outside the 'fit set' between force-fields with same or different experimental data emphasis, the environment of the molecular interaction (e.g. solvent, vacuum), and indeed the physical meaning of the parameters used.

Despite their relative convergence over time, typical force fields still exhibit significant differences (Ponder and Case, 2003).

The fixed point-charge models approximate hydrogen bonding through electronegativity. Linus Pauling defined electronegativity as 'the power of an atom to attract electrons to itself' (Leach, 2001). So for instance hydrogen partial charges are partitioned relative to the electronegativity of their bonded neighbours (Veenstra et al., 1992). The attraction between an electron deficient hydrogen increases with highly electronegative elements such as oxygen or nitrogen. This gives rise to unequal distribution of charge density and this inequality is reflected in the partitioning of related partial charge parameters (Leach, 2001, Jorgensen and Tirado-Rives, 2005, Jensen, 2007). Typically, after fitting, the partial charges, and 'stiff' internal angle and bond parameters, only a small number of LJ parameters for the E_{vdw} term are typically required for a level of agreement with experiment. These can be largely referenced from densities and heats of vaporization in liquid-state simulation (Ponder and Case, 2003).

A number of protein simulation subjective features e.g., the average folded structure and the fluctuations about this average structure are deemed insensitive to force field parameterization. The non-bonded interactions between moving group ends are intricately linked to the energy profile concerning rotation about torsion angles and their related potentials. The partitioning of torsional parameters into their bonded and non-bonded contributions is a grey area with less theoretical underpinning and typically adopts an empirical approach. The 'soft' torsional parameters around single bonds, are usually the last to be fitted when generating a new force field. The torsional parameters therefore serve as an empirical 'error function' for fitting the force field to a target extracted from QM or experiment (Ponder and Case, 2003). Here the energy expression can be considered as entirely empirical parameters and are adjusted until the closest coincidence to experiment occurs (e.g., via the least squares method) (Leszczynski, 2012). As the torsion parameters

are empirically linked to their non-bonded potentials, they typically lack transferability between force fields (Ponder and Case, 2003).

During the 1980s and 1990s there was considerable refinement in Class 1 force field partial charge values, but more recently emphasis has turned more to refining protein torsional potentials (Cieplak et al., 2009). Improvements in key backbone torsional parameters have arisen from the ability to carry out converged peptide or protein simulations then compare the conformational populations with experimental results (Damm and van Gunsteren, 2000). There has also been a desire to better model protein dynamics through the prediction of side chain rotamer preferences. This has led to extensions that fit torsional parameters to quantum mechanical results (Kaminski et al., 2001). Modelling or fitting of partial charges based on QM calculation results is subject to the limitations of the fixed atom-centred charge approach (Ponder, 2004).

When carrying out an energy minimization or MD simulation, it is important to know the forces acting directly on the nuclei. An advantage of the fixed atom-centred partial charge approach is it simplifies this problem (Leach, 2001). However, this approach lacks the mathematical flexibility to describe many features of a molecules charge distribution. Additionally the fixed charges cannot respond directly to the molecular environment e.g., the charge of a carbonyl oxygen is the same whether exposed to water on the protein surface or buried within a folded protein (Ponder and Case, 2003). The first round of AMBER type force fields used an electrostatic potential fitting (ESP) fitting method using a series of molecular surfaces. Each surface would have points at which the potential was fitted to the van der Waals radii, which gradually increased over the series. The basis set used to derive the wave function heavily influences the charge sets (Leach, 2001).

However, larger basis sets do not necessarily mean better charges. It is also important to treat the QM calculation with the same basis set. The 6-31G* basis set is largely considered to provide a good level of agreement to experimental results relevant to

protein modelling in the condensed phase (Leach, 2001). It is also possible to use scale factors to obtain comparable results using smaller basis sets STO-3G, or lower levels of approximation e.g., AM1 (Ferenczy et al., 1990). The first generation of the AMBER AA force field (Cornell et al., 1995) recognised the problem of buried charges having artificially high charges, and introduced the RESP or Restrained ESP fitting model (Bayly et al., 1993). RESP uses a hyperbolic restraints on non-hydrogen atoms that serves best to reduce the charges on buried carbon atoms (Leach, 2001). MULFIT charges do not suffer this problem and can be derived without the use of restraints (Winn 1997) see 2.5.9 and 2.7.

Despite improvements, a problem with ESP fitting, even with RESP, is that the fit charges are undesirably dependent on the conformation during the QM calculation (Williams, 1990). This can in part be addressed statistically by fitting charges weighting by a Boltzmann factor for the conformation population, derived from the calculated Boltzmann distributions of multiple conformations (Reynolds et al., 1992a). Alternatively continuous variation in charges based on conformation can be achieved through a charge equilibration model (Rappe and Goddard III, 1991). Ultimately, any amount of parameterization cannot fully address the inflexibility of the fixed charge model. Therefore, better electrostatic models (Stone, 1981, Winn et al., 1997, Ferenczy and Reynolds, 2001) are required to describe polar molecules within 'chemical accuracy' (Ponder and Case, 2003), which is the main theme of this research.

The OPLS (optimized potentials for liquid simulation) potential (Jorgensen et al., 1996), and its earlier incarnations highly influenced the first generation of AA force fields to describe solute-solvent versus solvent-solvent interactions in a balanced way. These new potentials would be more suited to condensed phase simulation, as previous emphasis was mainly on gas phase behaviour (Ponder and Case, 2003). OPLS achieved this balance by parameterization to reproduce heats of vaporization and densities for neat organic liquids (Jorgensen et al., 1996). A force field fitted to reproduce experiment and QM calculated results for isolated (gas phase) molecules, will likely translate poorly when applied to liquid

state-simulation. This applies to a much greater extent when polarization is ignored in the electrostatic model (Ponder and Case, 2003, Cieplak et al., 2009).

It had been known for some time that Hartree-Fock QM calculations at the 6-31G* basis level have a tendency to overestimate the experimental gas phase values for bond-dipoles for the 20 protein amino acids. Fortuitously there is also an over estimation of the water dimerization energy in the TIP3P water model. The CHARMM (Chemistry at HARvard Macromolecular Mechanics) force field (MacKerell et al., 1998) utilized these overestimates to provide a relatively easy way of achieving balanced water-solvent, solvent-solvent interactions for the 20 amino acids. In practice, a single water was fitted to several interaction sites around the amino acids or model compounds, and supermolecular QM calculations were performed with the interaction energies weighted with a 1.16 scaling factor that mirrored the ratio of water dimerization energy of TIP3P compared to the amino acid calculations at HF/6-31G* (Ponder and Case, 2003). This overestimation of charges can be viewed as an implicitly included 'over polarization', this only improves the flexibility of the charge distribution in a very average way. To further improve the charge distributions ability to adapt to the molecular environment it is necessary to go beyond the fixed point charge approximation (Leach, 2001, Ponder and Case, 2003), which we shall explore in the next section (see 2.5.9).

2.5.9 Beyond fixed point charges

Many-body effects describe the motion of every atom influenced by the motion of all of its atomic neighbours. These effects are non-additive, and cannot be solved easily, requiring coupled equations to describe the dynamics of the system (Illingworth and Domene, 2009). If we consider the interaction of two molecules, a dipole induced on one molecule will affect the electric field and subsequent charge distribution on the other. Additionally, moments can be many-body effects, as if a dipole is induced on a polarizable molecule it can be affected by the presence of a second dipole and so on (Leach, 2001). Modelling effects such as

polarization and the multipole expansion series, are important steps in improving the description of electrostatics (Leach, 2001, Cieplak et al., 2009).

The historical united atom (UA) force fields made a charge distribution and force centre approximation to reduce computational expense. This approximation was the inclusion of only polar hydrogens and heavy atom force centres (Leach, 2001). Computational power improvements and the limitations from this charge distribution model gave impetus for a next generation of AA force fields (Cornell et al., 1995). The UA approach with heavy atom only force centres made it difficult to describe 'pucker' of five member aliphatic rings or pseudo rotation between conformations or the electrostatics of aromatic rings (Ponder and Case, 2003). The situation is analogous to the current situation with AA force fields where better electrostatic models are now required (Ponder and Case, 2003, Cieplak et al., 2009).

The complete description of the interaction energy between two molecular charge distributions is an infinite series of interacting non-zero moments. Therefore in practice truncation or approximation is required at some point and as contribution to total energy can tail off quickly, the pair-wise approximation with charges placed at atomic nuclei has largely sufficed in large atomic models until now (Leach, 2001). However, the omission of higher order multipoles and explicit polarization can have large impacts on the description of the interaction and total energies (Winn et al., 1997). The induction energy and related polarization term alone is thought to account for 10%-20% of the interaction energy (Illingworth et al., 2006b, Cieplak et al., 2009). It is also thought that similar levels of error can occur from ignoring the multipole expansion (Winn et al., 1997). The fixed point charge AA force-field models, still struggle with describing the electrostatic interactions of ions with π electron system including aromatic rings, and charge concentrated atomic ions. This is largely due to a lack of consideration for many body non-additive contributions to the electrostatics (Jorgensen and Tirado-Rives, 2005).

The atom-centred charge approach without explicit extensions models a spherical or isotropic charge density (Leach, 2001). This is most applicable where a molecule's electron density typically contains a high degree of anisotropy. This is particular true when we consider phenomena such as atoms with lone pairs (Leach, 2001) and π electron systems (typical to aromatic rings), where each atom has a large quadrupole moment. The standard charge density ESP fitting method (e.g., not MULFIT), described in the last section, are the CHELPG (for Charges from Electrostatic Potentials using a Grid based) methods (Breneman and Wiberg, 1990). Fitting charges this way allows some consideration for the omission of the higher moments, and using a constrained least square fitting method, charge sets that model the molecular dipole moment can be produced. However, point charges cannot account effectively for the π systems where there are no neighbouring atoms above or below the plane to adjust (Stone, 2013).

One simple yet effective solution was able to reproduce the major geometric features of aromatic dimers (Hunter and Sanders, 1990) and later apply the principals to larger systems (e.g., DNA base-pairs). This model used dummy atoms with negative charges placed above and below the plane, combined with modified positive ring carbon charges (Stone, 2013). This inclusion of dummy atoms highlights some of the limitations of representing higher moments as point charges that are dependent on the plane of neighbouring atoms, and the importance of anisotropy in the charge distribution. The error in the electrostatic potential due to exclusion of higher moments is greater in proximity to polar and quadrupolar atoms, that are well represented in the active site of protein enzymes (Stone, 2013).

The high anisotropy of a molecule's charge distribution, effectively makes a single point charge set inadequate regardless of the fitting scheme (Wiberg and Rablen, 1993). One problem with introducing additional sites (e.g., to mimic lone pairs) to the fixed point charge model, is the arbitrary nature of such an approach. A more systematic way to improve a fixed electrostatic model is to include higher multipole terms such as dipole terms

or above that automatically include anisotropy (Wiberg and Rablen, 1993, Stone, 2013). The model we use in this work is the Distributed Multipole Analysis (DMA) method (Stone, 1981), which we have used up to rank 4 , i.e. hexadecapole moments (Stone, 2013).   In this method, distributed multipoles are derived directly from the properties of the Gaussian basis functions of the QM wave function. This is where the overlap of two of these functions, produces a Gaussian which represents their product along the connecting line. At any point P on the line of the product Gaussian there is a corresponding charge density. The multipole expansion of this charge density at P can be expressed through a Taylor series, allowing higher order moments that are dependent on the basis set used (Leach, 2001). If the ESP fitting of point charges is desired, the MULFIT procedure (Winn et al., 1997), can use a DMA as the source of the fitting potential. Doing ESP fitting from a DMA is faster than a charge density fitting and the potential avoids the penetration error (Stone, 2013) (see 2.5.3), and as we have access to DMA we also use MULFIT in this research as part of the QM/MM polarization methodology.

It is widely recognised that one set of parameters is inadequate to describe both the gas-phase and condensed-phase without at least the inclusion of explicit polarization (Ponder and Case, 2003). Work on polarisable force-fields has provided promising results using hybrid QM/MM methodologies (Cho et al., 2005, Maple et al., 2005). Polarization can be defined as, 'The redistribution of a particle's electron density due to an electric field' (Rick and Stuart, 2003). In a protein's condensed phase-like environment, any particle can be considered to be in the electric field of all its neighbours (Illingworth and Domene, 2009). The induction energy can be modelled in MM force fields somewhat through a polarization term (Cieplak et al., 2009, Stone, 2013). In addition to the above definition, we can refer to polarization in terms of the resulting changes in a molecule's conformation and charge density as a result of electron density redistribution (Winn et al., 1999).

Polarization improves the description of a number of modelled interactions, such as, ions with π electron system (e.g., potassium-benzene) (Jorgensen and Tirado-Rives, 2005,

Cieplak et al., 2009), atoms with lone pairs (e.g., N-methyl-acetamide-water dimers) (Cieplak et al., 2009), and water dimers (Stone, 2013), and consequently more generally hydration effects (Ren et al., 2012). For instance, water is an example of a highly polar molecule, the electronegative oxygen atom draws shared electron density from the two covalently bonded hydrogen atoms. The uneven sharing of electron density creates a concentrated negative charge (δ^-) on the oxygen atom, and in relation to the oxygen a distal area of positive charge(δ^+) on the hydrogens. This in turn leads to dipoles directed from the two hydrogen atoms towards the oxygen atom in the plane of the covalent bonds, the resultant dipole is shown in Figure 2.6 (Schwartz et al., 1997).

The mutual polarization of water molecules alone in bulk (see Figure 2.7a) or in the solvent-solute hydration interaction (see Figure 2.7b) can be thought of in terms of the tendency of molecules to align their permanent moments, and the combinations of their additive and non additive contributions. Mutual polarization during the solvent-solute hydration interaction explains the dampening of electrostatics between two hydrated molecules, and therefore the high dielectric constant for water (see 2.5.7) (Stillinger, 1975). This highlights the importance of inclusion of particularly an explicit polarization term in solvation models (Ponder and Case, 2003, Cieplak et al., 2009), and therefore we research the explicit inclusion of polarization here both with and without consideration for water.

There are several approaches to including polarization in MM force fields. The Drude oscillator model employs an extra point charge called a Drude particle, which is connected on a spring at shell distances to each atom, this particle is moves in response to the electrostatic forces to reproduce polarization effects (Cieplak et al., 2009). The fluctuating charge model (Rappé et al., 1992) is based on the electronegativity of
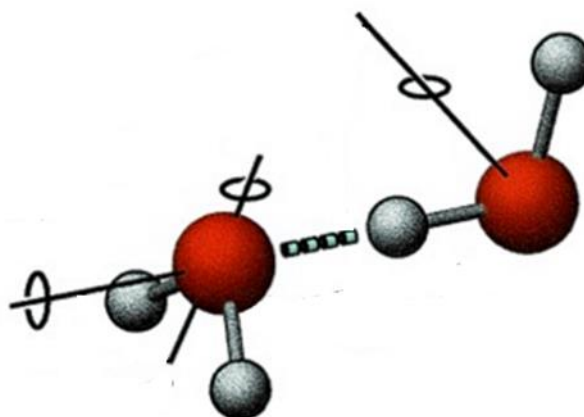
Figure 2.6 Example of a hydrogen bond formed by a water dimer, lines with rings represent dipole moment including the two permanent moments, through the two oxygen atoms. Adapted from (Harker et al., 2007).
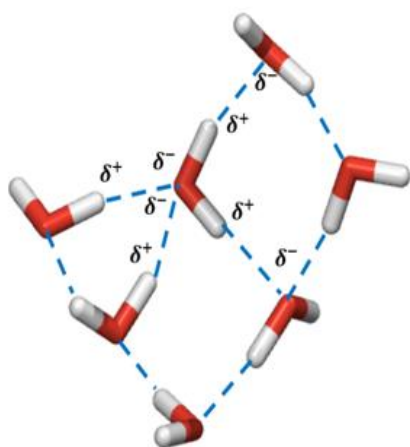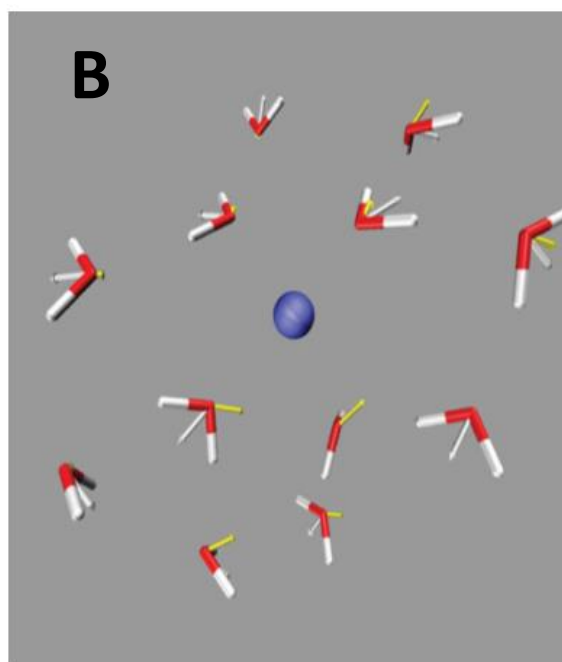
Figure 2.7 (A) A water molecule can form 4 hydrogen bonds, two at its δ^- oxygen atom, and one at each of its δ^+. This figure was generated using PyMOL using QSITE calculations from JAGUAR 7.8 and GLIDE 5.9 **(B)** Mutual polarization response of insertion of cation (K^+) into bulk water calculated using AMOEBA Force Field(Ponder, 2004). Figure 2.7 B Taken from (Ren et al., 2012).

atoms becoming equalized, where charges are coupled to their molecular environment and adjusted to mimic polarization (Cieplak et al., 2009, Illingworth and Domene, 2009). The widely researched induced dipole method, where polarization is included through the addition of the induced dipole moments arising from the point charges (Ponder and Case, 2003).Finally, there is the QM model, where polarization is derived at the electronic level from the wavefunction calculation (Illingworth and Domene, 2009).

A 'polarization catastrophe' can arise when nearby centres undergo an infinite mutual polarization, increasing the interaction energies absolute value (Ferenczy and Reynolds, 2001). One way this can occur is through using point polarizabilities instead of more accurate diffuse charge approaches (Cieplak et al., 2009). Related to point polarizabilities, is a close range infinite catastrophe produced by the interaction of two particles with polarizabilities α, which is proportional to $-\alpha^2/r^6$ (Cieplak et al., 2009). The Drude model avoids this situation naturally, by keeping interactions at E_{vdw} shell distances typical of the Lennard Jones Potential (Ponder and Case, 2003). The MM scheme, where 1-2 and 1-3 bonded interactions are turned off is also a cause of the 'polarization catastrophe' when using point polarizabilties typically observed in the induced dipole method (Cieplak et al., 2009).

Approaches to avoiding the polarization catastrophe include using a 1-2 and 1-3 interaction damping function (Thole, 1981), or turning them off completely within the polarization scheme (Cieplak et al., 2009). Alternatively, Gaussian functions applied to higher moments are able to interact at short distances, and can thus be used to describe the charge distributions without a polarization catastrophe (Elking et al., 2007). Another way is to avoid interaction within induced moment models (Ferenczy and Reynolds, 2001).

The induced charge method used in this research is an approximation of the induced dipole model, but offers ~2-4 fold time saving in respect to the iterated calculation for the induced dipole method. This approximation is effect with errors at close range between 10%-40% (with water-water interaction at the low end), with better approximation at

increased separation (Ferenczy and Reynolds, 2001). The induced charge Method (Winn et al., 1999, Ferenczy and Reynolds, 2001) adopts a dispersed charge approach, with each atom giving and receiving charge from its atomic neighbour, it is related to the fluctuating charge model in this way. However, it is the induced dipole that is approximated on each atom from the dipoles of its bonded neighbours, in an iterative fashion (Ferenczy and Reynolds, 2001). While the combination of charge-charge, charge-dipole, and dipole-dipole interactions are many body effects (Leach, 2001), the product of the groups method is an induced point charge that can be evaluated in the pair-wise MM scheme in which the induced charges were created. We now explore the platform for this research further.

2.6 Previous Work relevant to this research

2.6.1 Induced charges

The most significant in-house development work relevant to this research is the induced charge model for explicit polarization of an MM force field (Winn et al., 1999, Ferenczy and Reynolds, 2001). The use of a distributed multipole analysis (DMA) (Stone, 1981), has also been explored within the group, both to facilitate the induced charge work (Winn et al 1997, Illingworth et al., 2008a) and in its own right for improving interaction energies (Ferenczy et al., 1997). The point charges derived from the purely MM induced charge polarization scheme were assessed against QM Hartree-Fock calculations at the HF/6-31G* level (Boys et al., 1956, Hehre et al., 1972) on the water-water interaction of the dimer and trimer, and for small molecules (e.g., DNA base pairs) (Winn et al., 1999). The induced charge method was applied within a QM/MM scheme calculated with the BLYP functional (Becke, 1988, Lee et al., 1988) and a DZVP double zeta basis set with a polarization function for the study of iodine-based halogen bonding (Gooding et al., 2000). Further, water-water interactions and small molecule studies (albeit also with larger polypeptides) were explored under a QM/MM scheme using double and triple zeta quality basis sets, calculated with the B3LYP functional (Illingworth et al., 2006a). The QM/MM

scheme for induced charge polarization was then explored with Morokuma energy decomposition analysis (Umeyama et al., 1975, Kitaura and Morokuma, 1976), with various model chemistries on water dimers at the Hartree-Fock level, where medium sized basis sets (e.g. 6-31G*) were found to give the best agreement with the induced charge method, and larger basis sets gave the best agreement with the induced dipole method (Illingworth et al., 2008b). While these studies showed that small molecule interaction energies could be improved through inclusion of polarization relative to the use of fixed charges calculated on isolated molecules, none of these studies addressed protein-ligand docking.

2.6.2 Balanced parameterization in QM/MM calculations

For the QM/MM study of protein-ligand systems, a balance is usually struck between computational expense and model accuracy, which affects the degree of electron correlation or the size of the basis set (Illingworth et al., 2008b). QM/MM schemes generally have the potential problem of MM parameterization being inconsistent with the level of QM theory used for the calculation, but the 6-31G* (Hehre et al., 1972) family of basis sets is consistent with the treatment of polarization via induced charges (Illingworth et al., 2008b). Moreover, it has been proposed that density functional theory (DFT) approaches (e.g., B3LYP) can avoid many of the problems related to the parameterization of ligands since the ligand can be fully treated at the QM level (Cho et al., 2005). Together, these two sets of observations pave the way of a QM/MM treatment of docking that incorporates polarization.

2.6.3 QM/MM methods in docking: inclusion of ligand polarization

The investigation of the accuracy of the point charges including polarization of the QM ligand, by the MM protein derived at a QM/MM B3LYP/6-31G* level, has been applied to protein-ligand docking (Cho et al., 2005). In that instance, the QM ligand charges were derived from the QM/MM calculated wave function using the more accurate electrostatic potential (ESP) fitting method (which yields more accurate charges than Mulliken population analysis), and carried out as single point calculations (e.g., without iteration for optimization). The hybrid QM/MM calculation was carried out with an MM protein using OPLS 2005 (Banks

et al., 2005) and a QM ligand. The QM/MM interface was simplified by treating the non-covalent ligand and enzyme at the QM and MM levels respectively, and as the QM region was calculated in the presence of the MM environment; this naturally allowed the response to the external field and thus the polarization of the ligand (Cho et al., 2005). It additionally removed the complication of ligand parameterization (Illingworth et al., 2006a). QM/MM calculations treat the MM region point charges as an additional set of nuclear centres without basis functions; these charges are treated as a perturbation in the core Hamiltonian (Illingworth et al., 2008a). Thus, avoiding the problem of basis set superimposition errors (BSSE), for QM/MM calculations (Gooding et al., 2000).

The resulting ligand QM point charges replaced the fixed charges from the OPLS 2005 force field parameterization (Banks et al., 2005) and were evaluated within the MM framework of Glide (Friesner et al., 2004), which is considered amongst the most accurate docking programs (Cho et al., 2005, Warren et al., 2006). This method gave overall improved RMSD analysis results for Glide standard precision docking for 40 diverse co-crystallized or native protein complexes. The usage of this method was currently suggested in a lead optimization context, where the effort of producing higher quality charges could be balanced against computational expense when screening hundreds or thousands of ligands, compared to potentially millions of HTS ligands. Cho et al. hoped that they would address full explicit complex polarization in future studies (Cho et al., 2005). Our work fulfils this aspiration of Cho et al., as we additionally include MM polarization via the induced charge method.

2.6.4   Inclusion of polarization in Autodock

The groups own investigations to address the problem of docking through the accuracy of the point charges and attempting to expand on previous work by other groups (Cho et al., 2005), included the implementation of the induced charge method for the enzyme to supplement the ligand polarization that arises through use of the QM/MM method (Illingworth et al., 2008a). In that work, six distinct experiments using differing point charges were carried out on 12 co-crystallized protein structures. The resulting charges used in the

docking experiments were evaluated using a beta version of Autodock 4.0 (Morris et al., 2009). The Autodock framework was selected for the experiments, as in addition to direct energy evaluation, the size of the cluster pose populations for accessible state energy minima gives an indication of the quality of a docking pose (Illingworth et al., 2008a). Like the Glide docking experiments mentioned above, regardless of the source of the charges, the docking force field evaluation was carried out at the MM level (Illingworth et al., 2008a).

The six distinct charge experiments are now listed, the first used default Gasteiger charges (Gasteiger and Marsili, 1980) for ligand and receptor. The second used default Gasteiger charges for the ligand and AMBER potential derived charges (Wang et al., 2004), for the receptor. The third through to sixth charge experiments used QM (e.g., third experiment) and QM/MM calculations at the B3LYP/6-31G* (Hehre et al., 1972, Becke, 1993, Lee et al., 1988) carried out by the Gaussian 03 program (Frisch et al., 2004). The third and fourth experiments were used as benchmarks for QM/MM charge treatment without the induced charge method. The third experiment also used AMBER charges for the receptor, but the ligand charges were calculated in vacuo, a DMA was obtained using GDMA (Stone, 2005), and then MULFIT (Winn et al., 1997) was used to convert to point charges (e.g., without penetration errors). The fourth experiment derived point charges using QM/MM calculations at the X-ray crystallographic positions of the QM ligand and the (AMBER) MM protein with the caveat the MM region was defined as the residues with at least one atom within 5.5 Å of the ligand. This cut off was deemed apt as induction effects are short-range and largely only significant within this distance (Illingworth et al., 2008a). The AMBER MM region polarized the QM region of the ligand, and MULFIT was used to assign charge to the ligand based on the converged wave function. The fifth and sixth experiment used the induced charge method within the QM/MM scheme with and without iteration (sixth and fifth experiments respectively). The induced charge method allowed for the MM region that polarized the ligand to be mutually polarized, with and without convergence. Scalar isotropic atomic polarizabilities were taken from (Miller and Savchik, 1979). MULFIT again assigned

charges to the ligand. The docking results from the induced charge experiments did not show overall clear improvement for the 12 complexes studied using the induced charge method. However, there were some encouraging examples of improvement that merited further investigation (Illingworth et al., 2008a), hence this current investigation. The induced charge method is described briefly in 3.2.8 for a full description see (Ferenczy and Reynolds, 2001).

2.6.5   Aims of the current work: inclusion of polarization

In this work we intend to return to the original Glide SP platform used for ligand only polarization, where results showed clear overall improvement (Cho et al., 2005), but to extend the method to include MM polarization. The intention is to apply a more rigorous evaluation of explicitly polarized charges within the IMPACT MM framework using the OPLS 2005 force field, and use the Glide docking engine (Friesner et al., 2004) for posing and rescoring (Banks et al., 2005). We also intend to investigate the effects of polarization beyond the co-crystallized native position since that is not known in a de novo drug design situation. Thus, the effects on docking of polarization at different pose geometries within the protein will be investigated along with studies of how to polarize when the answer is not known. For example, one approach will be to polarize the optimized ligand geometries from QM/MM optimization using QSITE in the JAGUAR 7.9 (Murphy et al., 2000) and IMPACT (Banks et al., 2005) environments.

2.6.6   Aims of the current work: beyond the point charge approximation

We also intend to build on previous work using DMA (Ferenczy et al., 1997) and use the MM framework of ORIENT 4.6 (Stone et al., 2006) that is capable of handling rank 4 multipoles, and therefore go beyond the point charge approximation. Despite this sophistication, ORIENT it is not designed for docking and therefore we will use the default Glide standard precision flexible ligand docked poses (Friesner et al., 2004) presented from the docking experiments and then recalculate their energies in ORIENT using the eigenvector following method with geometries optimized through reorientation (Stone et al.,

2006). Thus, as Glide presents a number of rank ordered poses for each molecule the calculations will have to be carried out for each specific pose and for its pose specific QM region within the protein environment to avoid symmetry problems with the interface between the ligand and the protein environment DMA. In the Orient framework the QM/MM calculations and subsequent DMAs will be defined as the residues with at least one atom within a revised 4 Å of the ligand, as substantial QM calculations on neighbouring residues are required that will also be subject to DMAs, and some consideration must be given for the treatment of charge transfer (e.g., cut offs, ions, capping). Through these calculations we hope to address the question, as to whether going beyond the point charge approximation can offer improved geometries and give a better description of hydrogen bonding.

We hope from our investigation into high quality charges we will be able to address the docking problem. However, these QM/MM calculations can be computationally expensive so in these initial studies we have used fragments as representative ligands both because of their smaller size resulting in potential large computational time savings (relative to typical drug-like molecules), and their largely polar nature making them amenable to QM calculations. In addition to charge quality, neglect of explicit water molecules is another source of potential error (Cho et al., 2005) that will also be explored.

2.6.7   Aims of the current work: cross-docking

Additionally, although for native docking of a ligand in its co-crystallized protein structure the error is non-trivial, the situation for cross-docking is far worse (Cho et al., 2005, Warren et al., 2006). Cross-docking refers to when a ligand is docked in a protein with which it wasn't co-crystallized, the problem occurs as the ligand is allowed to move yet the protein remains rigid (e.g. grid representations), and therefore potential steric problems are more pronounced as the protein environment has not been primed for the ligand in the cross docked structures. Therefore, we will try to address this problem with the caveat that are initial native docking results are encouraging enough for further research. We will also be

careful to maintain a low level of docking bias by not priming the protein with the non-native ligand prior to docking.

2.7 Additional Computational Aspects of this Research

The calculation of the DMA is relatively inexpensive in computational terms when compared to some other multipole expansion methods (Stone, 2013). However, QM calculations are still largely limited to small molecules, so as suggested in 2.7 we propose to use QM/MM methods (Illingworth et al., 2008a) applied to fragments. Regarding the DMA calculations, we shall use the protein ligand modelling strategy of including a thin shell of protein residues in the QM region. For now, the DMA residue(s) and ligand calculations shall be done separately from each other and the rest of the protein *in-vacuo.* We will therefore attempt to address the problem of evaluating one charge set arising from different environments, with a different treatment scalar isotropic treatment of polarization, as implemented in ORIENT 4.6 (Stone et al.). As the Lennard-Jones potential (See 2.5.4 Eq. 2.4) is less quantitatively accurate than the three parameter models using an exponential component for exchange (Buckingham and Corner, 1947, Williams, 1999), we will also explore the soft Williams potential (Williams, 1999) within the DMA framework.

The current QM/MM implementation of induced charges was scripted in house. In this implementation, the MM environment is included in the QM calculation in Gaussian 03 as a set of point charges and additional force centres (Frisch et al., 2004), which allows fast approximation of the electrostatic potential (and electric field) at each atomic centre, which is extracted to compute the MM induced dipoles, and subsequently reduced to induced charges and added to the MM point charge framework of OPLS 2005 (Banks et al., 2005). The MM polarization implementation and iterations are controlled through a series of in-house Perl scripts. Docking was largely carried out using Glide 5.8 SP (Friesner et al., 2004).

Rescoring of docked poses was carried out largely using Glide 5.8 SP protocols, namely GSCORE, CVDW and EMODEL. Atom Typing / Parameterization was carried out by IMPACT 5.8 (Banks et al., 2005) using the parameters from OPLS 2005 force-field (Banks et al., 2005). When considering atomic charges other than OPLS 2005, ligand atomic charges were determined using Gaussian ESP (Frisch et al., 2004) and/or Gaussian in conjunction with MULFIT 2.1 to remove penetration errors (Reynolds et al., 1992b, Winn et al., 1997).

Charges, as done previously (Cho et al., 2005, Illingworth et al., 2008a), were determined using density functional theory at the B3LYP/6-31G* level rather than at the Hartree-Fock level. This is because the Hartree-Fock method overestimates the polarity of the wavefunction, which is useful for studies in which the effects of polarization are implicit in the parameterization, but would result in double-counting if used with explicit polarization. The density functional method, because it includes electron correlation does not suffer this problem (Winn et al., 1997). Distributed multipole analysis was carried out on the QM region using Gaussian 03 in combination with GDMA 2.2 (Stone, 2005) and glide docked poses were re-orientated, with pose specific distributed multipoles using ORIENT 4.6 (Stone, 1981, Stone et al., 2006). Visualization was carried out using Maestro 9.3 (Banks et al., 2005); pictures were generated using PyMOL. Analysis of the results, unless otherwise mentioned was facilitated by Perl Scripts. As the biggest difficulty in docking is not the force fields, but rather the treatment of hydration, our polarization approach is also applied to the inclusion of specific water molecules.

2.8 References

ABAGYAN, R., TOTROV, M. & KUZNETSOV, D. 1994. ICM—a new method for protein modeling and design: applications to docking and structure prediction from the distorted native conformation. *Journal of computational chemistry,* 15**,** 488-506.

ALLINGER, N. L., CHEN, K. H., LII, J. H. & DURKIN, K. A. 2003. Alcohols, ethers, carbohydrates, and related compounds. I. The MM4 force field for simple compounds. *Journal of computational chemistry,* 24**,** 1447-1472.

ALLINGER, N. L., YUH, Y. H. & LII, J. H. 1989. Molecular mechanics. The MM3 force field for hydrocarbons. 1. *Journal of the American Chemical Society,* 111**,** 8551-8566.

ALVESALO, J. K. O., SIISKONEN, A., VAINIO, M. J., TAMMELA, P. S. M. & VUORELA, P. M. 2006. Similarity Based Virtual Screening: A Tool for Targeted Library Design. *Journal of medicinal chemistry,* 49**,** 2353-2356.

BABAOGLU, K. & SHOICHET, B. K. 2006. Deconstructing fragment-based inhibitor discovery. *Nature chemical biology,* 2**,** 720-723.

BANKS, J. L., BEARD, H. S., CAO, Y., CHO, A. E., DAMM, W., FARID, R., FELTS, A. K., HALGREN, T. A., MAINZ, D. T. & MAPLE, J. R. 2005. Integrated modeling program, applied chemical theory (IMPACT). *Journal of computational chemistry,* 26**,** 1752-1780.

BAYLY, C. I., CIEPLAK, P., CORNELL, W. & KOLLMAN, P. A. 1993. A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges: the RESP model. *The Journal of Physical Chemistry,* 97**,** 10269-10280.

BECKE, A. D. 1988. Density-functional exchange-energy approximation with correct asymptotic behavior. *Physical review A,* 38**,** 3098.

BECKE, A. D. 1993. Density-functional thermochemistry. III. The role of exact exchange. *The Journal of Chemical Physics,* 98**,** 5648-5652.

BOSTRÖM, J., HOGNER, A. & SCHMITT, S. 2006. Do Structurally Similar Ligands Bind in a Similar Fashion? *Journal of medicinal chemistry,* 49**,** 6716-6725.

BOYS, S. F. & BERNARDI, F. D. 1970. The calculation of small molecular interactions by the differences of separate total energies. Some procedures with reduced errors. *Molecular Physics,* 19**,** 553-566.

BOYS, S. F., COOK, G. B., REEVES, C. M. & SHAVITT, I. 1956. Automatic Fundamental Calculations of Molecular Structure. *Nature,* 178**,** 1207-1209.

BRENEMAN, C. M. & WIBERG, K. B. 1990. Determining atom-centered monopoles from molecular electrostatic potentials. The need for high sampling density in formamide conformational analysis. *Journal of Computational Chemistry,* 11**,** 361-373.

BUCKINGHAM, R. & CORNER, J. 1947. Tables of second virial and low-pressure Joule-Thomson coefficients for intermolecular potentials with exponential repulsion. *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences,* 189**,** 118-129.

CAVASOTTO, C. N. & ABAGYAN, R. A. 2004. Protein flexibility in ligand docking and virtual screening to protein kinases. *Journal of molecular biology,* 337**,** 209-225.

CHANDLER, D. 2005. Interfaces and the driving force of hydrophobic assembly. *Nature,* 437**,** 640-647.

CHANG, C. A., CHEN, W. & GILSON, M. K. 2007. Ligand configurational entropy and protein binding. *Proceedings of the National Academy of Sciences,* 104**,** 1534.

CHEN, Y. & SHOICHET, B. K. 2009. Molecular docking and ligand specificity in fragment-based inhibitor discovery. *Nature chemical biology,* 5**,** 358-364.

CHO, A. E., GUALLAR, V., BERNE, B. J. & FRIESNER, R. 2005. Importance of accurate charges in molecular docking: Quantum mechanical/molecular mechanical (QM/MM) approach. *Journal of Computational Chemistry,* 26**,** 915-931.

CIEPLAK, P., DUPRADEAU, F.-Y., DUAN, Y. & WANG, J. 2009. Polarization effects in molecular mechanical force fields. *Journal of Physics: Condensed Matter,* 21**,** 333102.

CLAUßEN, H., BUNING, C., RAREY, M. & LENGAUER, T. 2001. FlexE: efficient molecular docking considering protein structure variations. *Journal of molecular biology,* 308**,** 377-395.

CONGREVE, M., CHESSARI, G., TISI, D. & WOODHEAD, A. J. 2008. Recent Developments in Fragment-Based Drug Discovery. *Journal of Medicinal Chemistry,* 51**,** 3661-3680.

CORNELL, W. D., CIEPLAK, P., BAYLY, C. I., GOULD, I. R., MERZ, K. M., FERGUSON, D. M., SPELLMEYER, D. C., FOX, T., CALDWELL, J. W. & KOLLMAN, P. A. 1995. A second generation force field for the simulation of proteins, nucleic acids, and organic molecules. *Journal of the American Chemical Society,* 117**,** 5179-5197.

DAHLKE, E. E. & TRUHLAR, D. G. 2006. Assessment of the pairwise additive approximation and evaluation of many-body terms for water clusters. *The Journal of Physical Chemistry B,* 110**,** 10595-10601.

DAMM, W. & VAN GUNSTEREN, W. F. 2000. Reversible peptide folding: dependence on molecular force field used. *Journal of Computational Chemistry,* 21**,** 774-787.

DAVIS, A. M., TEAGUE, S. J. & KLEYWEGT, G. J. 2003. Application and Limitations of X-ray Crystallographic Data in Structure-Based Ligand and Drug Design. *Angewandte Chemie International Edition,* 42**,** 2718-2736.

DE AZEVEDO, J., WALTER, F. & DIAS, R. 2008. Computational methods for calculation of ligand-binding affinity. *Current Drug Targets,* 9**,** 1031-1039.

DUAN, Y., WU, C., CHOWDHURY, S., LEE, M. C., XIONG, G., ZHANG, W., YANG, R., CIEPLAK, P., LUO, R. & LEE, T. 2003. A point-charge force field for molecular mechanics simulations of proteins based on condensed-phase quantum mechanical calculations. *Journal of computational chemistry,* 24**,** 1999-2012.

DURRANT, J. D. & MCCAMMON, J. A. 2010. Computer-aided drug-discovery techniques that account for receptor flexibility. *Current opinion in pharmacology,* 10**,** 770-774.

ELKING, D., DARDEN, T. & WOODS, R. J. 2007. Gaussian induced dipole polarization model. *Journal of computational chemistry,* 28**,** 1261-1274.

ELROD, M. J. & SAYKALLY, R. J. 1994. Many-body effects in intermolecular forces. *Chemical reviews,* 94**,** 1975-1997.

EWING, T. J., MAKINO, S., SKILLMAN, A. G. & KUNTZ, I. D. 2001. DOCK 4.0: search strategies for automated molecular docking of flexible molecule databases. *Journal of computer-aided molecular design,* 15**,** 411-428.

FAY, R. C. & MCMURRY, J. E. 2012. *Chemistry, 6th Ed,* New York, Pearson.

FERENCZY, G. G. & REYNOLDS, C. A. 2001. Modeling polarization through induced atomic charges. *The Journal of Physical Chemistry A,* 105**,** 11470-11479.

FERENCZY, G. G., REYNOLDS, C. A. & RICHARDS, W. G. 1990. Semiempirical AM1 electrostatic potentials and AM1 electrostatic potential derived charges: A comparison with ab initio values. *Journal of computational chemistry,* 11**,** 159-169.

FERENCZY, G. G., WINN, P. J. & REYNOLDS, C. A. 1997. Toward improved force fields. 2. Effective distributed multipoles. *The Journal of Physical Chemistry A,* 101**,** 5446-5455.

FIERZ, M. & PAULI, W. 1939. On relativistic wave equations for particles of arbitrary spin in an electromagnetic field. *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences,* 173**,** 211-232.

FISCHER-CRIPPS, A. C. 2014. *The physics companion*, CRC Press.

FREIRE, E. 2008. Do enthalpy and entropy distinguish first in class from best in class? *Drug Discovery Today,* 13**,** 869-874.

FRIESNER, R. A., BANKS, J. L., MURPHY, R. B., HALGREN, T. A., KLICIC, J. J., DANIEL, T., REPASKY, M. P., KNOLL, E. H., SHELLEY, M. & PERRY, J. K. 2004. Glide: a new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy. *Journal of medicinal chemistry,* 47**,** 1739-1749.

FRIESNER, R. A., MURPHY, R. B., REPASKY, M. P., FRYE, L. L., GREENWOOD, J. R., HALGREN, T. A., SANSCHAGRIN, P. C. & MAINZ, D. T. 2006. Extra precision glide: docking and scoring incorporating a model of hydrophobic enclosure for protein-ligand complexes. *Journal of medicinal chemistry,* 49**,** 6177-6196.

FRISCH, M., TRUCKS, G., SCHLEGEL, H., SCUSERIA, G., ROBB, M., CHEESEMAN, J., MONTGOMERY JR, J., VREVEN, T., KUDIN, K. & BURANT, J. 2004. Gaussian 03, revision c. 02; Gaussian. *Inc., Wallingford, CT,* 4.

GASTEIGER, J. & MARSILI, M. 1980. Iterative partial equalization of orbital electronegativity—a rapid access to atomic charges. *Tetrahedron,* 36**,** 3219-3228.

GLEESON, M. P. & GLEESON, D. 2009. QM/MM as a tool in fragment based drug discovery. A cross-docking, rescoring study of kinase inhibitors. *Journal of chemical information and modeling,* 49**,** 1437-1448.

GOODFORD, P. J. 1985. A computational procedure for determining energetically favorable binding sites on biologically important macromolecules. *Journal of Medicinal Chemistry,* 28**,** 849-857.

GOODING, S. R., WINN, P. J., MAURER, R. I., FERENCZY, G. G., MILLER, J. R., HARRIS, J. E., GRIFFITHS, D. V. & REYNOLDS, C. A. 2000. Fully polarizable QM/MM calculations: An application to the nonbonded iodine–oxygen interaction in dimethyl-2-iodobenzoylphosphonate. *Journal of Computational Chemistry,* 21**,** 478-482.

GRAVES, A. P., BRENK, R. & SHOICHET, B. K. 2005. Decoys for docking. *Journal of medicinal chemistry,* 48**,** 3714-3728.

HAJDUK, P. J. & GREER, J. 2007. A decade of fragment-based drug design: strategic advances and lessons learned. *Nat Rev Drug Discov,* 6**,** 211-219.

HALGREN, T. A. 1992. The representation of van der Waals (vdW) interactions in molecular mechanics force fields: potential form, combination rules, and vdW parameters. *Journal of the American Chemical Society,* 114**,** 7827-7843.

HALGREN, T. A. 1996a. Merck molecular force field. I. Basis, form, scope, parameterization, and performance of MMFF94. *Journal of computational chemistry,* 17**,** 490-519.

HALGREN, T. A. 1996b. Merck molecular force field. II. MMFF94 van der Waals and electrostatic parameters for intermolecular interactions. *Journal of Computational Chemistry,* 17**,** 520-552.

HALGREN, T. A. 1999. MMFF VI. MMFF94s option for energy minimization studies. *Journal of Computational Chemistry,* 20**,** 720-729.

HARKER, H., KEUTSCH, F., LEFORESTIER, C., SCRIBANO, Y., HAN, J.-X. & SAYKALLY, R. 2007. Toward a precise determination of the acceptor switching splitting in the water dimer. *Molecular Physics,* 105**,** 497-512.

HEHRE, W. J., DITCHFIELD, R. & POPLE, J. A. 1972. Self—consistent molecular orbital methods. XII. Further extensions of gaussian—type basis sets for use in molecular orbital studies of organic molecules. *The Journal of Chemical Physics,* 56**,** 2257-2261.

HODGES, M. P., STONE, A. J. & XANTHEAS, S. S. 1997. Contribution of Many-Body Terms to the Energy for Small Water Clusters: A Comparison of ab Initio Calculations and Accurate Model Potentials. *The Journal of Physical Chemistry A,* 101**,** 9163-9168.

HORN, J. R. & SHOICHET, B. K. 2004. Allosteric inhibition through core disruption. *Journal of molecular biology,* 336**,** 1283-1291.

HRABOVSKY, G. & SUSSKIND, L. 2012. *The theoretical minimum: What you need to know to start doing physics.* New York: Allen Lane.

HUANG, S.-Y., GRINTER, S. Z. & ZOU, X. 2010. Scoring functions and their evaluation methods for protein–ligand docking: recent advances and future directions. *Physical Chemistry Chemical Physics,* 12**,** 12899-12908.

HUEY, R., MORRIS, G. M., OLSON, A. J. & GOODSELL, D. S. 2007. A semiempirical free energy force field with charge-based desolvation. *Journal of computational chemistry,* 28**,** 1145-1152.

HUNTER, C. A. & SANDERS, J. K. 1990. The nature of. pi.-. pi. interactions. *Journal of the American Chemical Society,* 112**,** 5525-5534.

ILLINGWORTH, C. J. & DOMENE, C. Many-body effects and simulations of potassium channels. Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences, 2009. The Royal Society, rspa. 2009.0014.

ILLINGWORTH, C. J., GOODING, S. R., WINN, P. J., JONES, G. A., FERENCZY, G. G. & REYNOLDS, C. A. 2006a. Classical polarization in hybrid QM/MM methods. *The Journal of Physical Chemistry A,* 110**,** 6487-6497.

ILLINGWORTH, C. J., MORRIS, G. M., PARKES, K. E., SNELL, C. R. & REYNOLDS, C. A. 2008a. Assessing the role of polarization in docking. *The Journal of Physical Chemistry A,* 112**,** 12157-12163.

ILLINGWORTH, C. J., PARKES, K. E., SNELL, C. R., FERENCZY, G. R. G. & REYNOLDS, C. A. 2008b. Toward a consistent treatment of polarization in model QM/MM calculations. *The Journal of Physical Chemistry A,* 112**,** 12151-12156.

ILLINGWORTH, C. J. R., GOODING, S. R., WINN, P. J., JONES, G. A., FERENCZY, G. G. & REYNOLDS, C. A. 2006b. Classical polarization in hybrid QM/MM methods. *J. Phys. Chem. A,* 110**,** 6487-6497.

JENSEN, F. 2007. *Introduction to computational chemistry*, John Wiley & Sons.

JEZIORSKI, B., MOSZYNSKI, R. & SZALEWICZ, K. 1994. Perturbation theory approach to intermolecular potential energy surfaces of van der Waals complexes. *Chemical Reviews,* 94**,** 1887-1930.

JONES, G., WILLETT, P. & GLEN, R. C. 1995. Molecular recognition of receptor sites using a genetic algorithm with a description of desolvation. *Journal of Molecular Biology,* 245**,** 43-53.

JONES, G., WILLETT, P., GLEN, R. C., LEACH, A. R. & TAYLOR, R. 1997. Development and validation of a genetic algorithm for flexible docking. *Journal of Molecular Biology,* 267**,** 727-748.

JONES, J. E. 1924. On the Determination of Molecular Fields. II. From the Equation of State of a Gas. *Proceedings of the Royal Society of London. Series A,* 106**,** 463-477.

JORGENSEN, W. L. 2004. The Many Roles of Computation in Drug Discovery. *Science,* 303**,** 1813-1818.

JORGENSEN, W. L., MAXWELL, D. S. & TIRADO-RIVES, J. 1996. Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids. *Journal of the American Chemical Society,* 118**,** 11225-11236.

JORGENSEN, W. L. & TIRADO-RIVES, J. 2005. Potential energy functions for atomic-level simulations of water and organic and biomolecular systems. *Proceedings of the National Academy of Sciences of the United States of America,* 102**,** 6665-6670.

KAIRYS, V., FERNANDES, M. X. & GILSON, M. K. 2005. Screening Drug-Like Compounds by Docking to Homology Models: A Systematic Study. *Journal of chemical information and modeling,* 46**,** 365-379.

KAMINSKI, G. A., FRIESNER, R. A., TIRADO-RIVES, J. & JORGENSEN, W. L. 2001. Evaluation and Reparametrization of the OPLS-AA Force Field for Proteins via Comparison with Accurate Quantum Chemical Calculations on Peptides†. *The Journal of Physical Chemistry B,* 105**,** 6474-6487.

KEELER, J. & WOTHERS, P. 2003. *Why chemical reactions happen*, Oxford University Press.

KELLENBERGER, E., RODRIGO, J., MULLER, P. & ROGNAN, D. 2004. Comparative evaluation of eight docking tools for docking and virtual screening accuracy. *Proteins: Structure, Function, and Bioinformatics,* 57**,** 225-242.

KITAURA, K. & MOROKUMA, K. 1976. A new energy decomposition scheme for molecular interactions within the Hartree-Fock approximation. *International Journal of Quantum Chemistry,* 10**,** 325-340.

KITCHEN, D. B., DECORNEZ, H., FURR, J. R. & BAJORATH, J. 2004. Docking and scoring in virtual screening for drug discovery: methods and applications. *Nat Rev Drug Discov,* 3**,** 935-949.

KLEYWEGT, G. J. 2006. Crystallographic refinement of ligand complexes. *Acta Crystallographica Section D: Biological Crystallography,* 63**,** 94-100.

KONTOYIANNI, M., MCCLELLAN, L. M. & SOKOL, G. S. 2003. Evaluation of Docking Performance: Comparative Data on Docking Algorithms. *Journal of medicinal chemistry,* 47**,** 558-565.

KOSHLAND JR, D. 1958. Application of a theory of enzyme specificity to protein synthesis. *Proceedings of the National Academy of Sciences of the United States of America,* 44**,** 98.

KUNTZ, I. D., BLANEY, J. M., OATLEY, S. J., LANGRIDGE, R. & FERRIN, T. E. 1982. A geometric approach to macromolecule-ligand interactions. *Journal of molecular biology,* 161**,** 269-288.

LAZARIDIS, T. 1998. Inhomogeneous fluid approach to solvation thermodynamics. 1. Theory. *The Journal of Physical Chemistry B,* 102**,** 3531-3541.

LEACH, A. R. 2001. *Molecular modelling: principles and applications*, Pearson education.

LEACH, A. R., SHOICHET, B. K. & PEISHOFF, C. E. 2006. Docking and scoring. *J. Med. Chem,* 49**,** 5851-5855.

LEE, C., YANG, W. & PARR, R. G. 1988. Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. *Physical Review B,* 37**,** 785-789.

LESZCZYNSKI, J. 2012. Handbook of Computational Chemistry.

LONDON, F. 1937. The general theory of molecular forces. *Trans. Faraday Soc.,* 33**,** 8b-26.

MACKERELL, A. D., BROOKS, B., BROOKS, C. L., NILSSON, L., ROUX, B., WON, Y. & KARPLUS, M. 1998. CHARMM: the energy function and its parameterization. *Encyclopedia of computational chemistry*.

MAPLE, J. R., CAO, Y., DAMM, W., HALGREN, T. A., KAMINSKI, G. A., ZHANG, L. Y. & FRIESNER, R. A. 2005. A Polarizable Force Field and Continuum Solvation Methodology for Modeling of Protein–Ligand Interactions. *Journal of Chemical Theory and Computation,* 1**,** 694-715.

MARCOU, G. & ROGNAN, D. 2007. Optimizing fragment and scaffold docking by use of molecular interaction fingerprints. *Journal of chemical information and modeling,* 47**,** 195-207.

MCMARTIN, C. & BOHACEK, R. 1997. QXP: Powerful, rapid computer algorithms for structure-based drug design. *Journal of Computer-Aided Molecular Design,* 11**,** 333-344.

MILLER, K. J. & SAVCHIK, J. 1979. A new empirical method to calculate average molecular polarizabilities. *Journal of the American Chemical Society,* 101**,** 7206-7213.

MORRIS, G. M., GOODSELL, D. S., HALLIDAY, R. S., HUEY, R., HART, W. E., BELEW, R. K. & OLSON, A. J. 1998. Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function. *Journal of computational chemistry,* 19**,** 1639-1662.

MORRIS, G. M., HUEY, R., LINDSTROM, W., SANNER, M. F., BELEW, R. K., GOODSELL, D. S. & OLSON, A. J. 2009. AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. *Journal of computational chemistry,* 30**,** 2785-2791.

MURPHY, R., PHILIPP, D. & FRIESNER, R. 2000. Frozen orbital QM/MM methods for density functional theory. *Chemical Physics Letters,* 321**,** 113-120.

NIKETIC, S. & RASMUSSEN, K. 1977. The Consistant Force Field. *Lecture Notes in Chemistry,* 3.

OTT, K. H. & MEYER, B. 1996. Parametrization of GROMOS force field for oligosaccharides and assessment of efficiency of molecular dynamics simulations. *Journal of computational chemistry,* 17**,** 1068-1084.

PÉREZ, C. & ORTIZ, A. R. 2001. Evaluation of docking functions for protein-ligand docking. *Journal of medicinal chemistry,* 44**,** 3768-3785.

PONDER, J. W. 2004. TINKER: Software tools for molecular design. *Washington University School of Medicine, Saint Louis, MO,* 3.

PONDER, J. W. & CASE, D. A. 2003. Force fields for protein simulations. *Advances in protein chemistry,* 66**,** 27-86.

RAPPÉ, A. K., CASEWIT, C. J., COLWELL, K., GODDARD III, W. & SKIFF, W. 1992. UFF, a full periodic table force field for molecular mechanics and molecular dynamics simulations. *Journal of the American chemical society,* 114**,** 10024-10035.

RAPPE, A. K. & GODDARD III, W. A. 1991. Charge equilibration for molecular dynamics simulations. *The Journal of Physical Chemistry,* 95**,** 3358-3363.

RAREY, M., KRAMER, B., LENGAUER, T. & KLEBE, G. 1996. A fast flexible docking method using an incremental construction algorithm. *Journal of molecular biology,* 261**,** 470-489.

REN, P., CHUN, J., THOMAS, D. G., SCHNIEDERS, M. J., MARUCHO, M., ZHANG, J. & BAKER, N. A. 2012. Biomolecular electrostatics and solvation: a computational perspective. *Quarterly reviews of biophysics,* 45**,** 427-491.

REYNOLDS, C. A., ESSEX, J. W. & RICHARDS, W. G. 1992a. Atomic charges for variable molecular conformations. *Journal of the American Chemical Society,* 114**,** 9075-9079.

REYNOLDS, C. A., FERENCZY, G. G. & RICHARDS, W. G. 1992b. Methods for determining the reliability of semiempirical electrostatic potentials and potential derived charges. *Journal of Molecular Structure: THEOCHEM,* 256**,** 249-269.

RICK, S. W. & STUART, S. J. 2003. Potentials and Algorithms for Incorporating Polarizability in Computer Simulations. *Reviews in Computational Chemistry.* John Wiley & Sons, Inc.

SÁNDOR, M., KISS, R. & KESERŰ, G. R. M. 2010. Virtual fragment docking by Glide: A validation study on 190 protein– fragment complexes. *Journal of chemical information and modeling,* 50**,** 1165-1172.

SCHULER, L. D., DAURA, X. & VAN GUNSTEREN, W. F. 2001. An improved GROMOS96 force field for aliphatic hydrocarbons in the condensed phase. *Journal of Computational Chemistry,* 22**,** 1205-1218.

SCHWARTZ, A. T., BUNCE, D., SILBERMAN, R., STANITSKI, C., STRATTON, W. & ZIPP, A. 1997. *Chemistry in context: Applying chemistry to society*, Wm. C. Brown Dubuque, IA.

SCIOR, T., BENDER, A., TRESADERN, G., MEDINA-FRANCO, J. L., MARTÍNEZ-MAYORGA, K., LANGER, T., CUANALO-CONTRERAS, K. & AGRAFIOTIS, D. K. 2012. Recognizing Pitfalls in Virtual Screening: A Critical Review. *Journal of chemical information and modeling,* 52**,** 867-881.

SEIFERT, M. H. & LANG, M. 2008. Essential factors for successful virtual screening. *Mini Reviews in Medicinal Chemistry,* 8**,** 63-72.

SIMMERLING, C., FOX, T. & KOLLMAN, P. A. 1998. Use of locally enhanced sampling in free energy calculations: Testing and application to the α→ β anomerization of glucose. *Journal of the American Chemical Society,* 120**,** 5771-5782.

SMITH, P. E. & PETTITT, B. M. 1994. Modeling solvent in biomolecular systems. *The Journal of physical chemistry,* 98**,** 9700-9711.

STILLINGER, F. H. 1975. Theory and molecular models for water. *Adv. Chem. Phys,* 31**,** 10.

STONE, A. 1981. Distributed multipole analysis, or how to describe a molecular charge distribution. *Chemical Physics Letters,* 83**,** 233-239.

STONE, A. 1996. The Theory of Intermolecular Forces, 1996, Clarendon. Oxford.

STONE, A. 2005. Distributed Multipole Analysis of Gaussian wavefunctions GDMA version 2.2. 02.

STONE, A. 2013. *The theory of intermolecular forces*, Oxford University Press.

STONE, A., DULLWEBER, A., ENGKVIST, O., FRASCHINI, E., HODGES, M., MEREDITH, A., NUTT, D. & POPELIER, P. 2006. LA; Wales, DJ Orient: a program for studying interactions between molecules, Version 4.6; University of Cambridge: 2006. *Inquiries to AJ Stone, ajs1@ cam. ac. uk.*

STONE, A., DULLWEBER, A., ENGKVIST, O., FRASCHINI, E., HODGES, M., MEREDITH, A., NUTT, D., POPELIER, P. & WALES, D. ORIENT: a program for studying interactions between molecules, version 4.6, University of Cambridge, 2002. *There is no corresponding record for this reference*.

STONE, A. J. & MISQUITTA, A. J. 2009. Charge-transfer in symmetry-adapted perturbation theory. *Chemical Physics Letters,* 473**,** 201-205.

SUGITA, Y. & OKAMOTO, Y. 1999. Replica-exchange molecular dynamics method for protein folding. *Chemical physics letters,* 314**,** 141-151.

TEAGUE, S. J. 2003. Implications of protein flexibility for drug discovery. *Nature Reviews Drug Discovery,* 2**,** 527-541.

THOLE, B. T. 1981. Molecular polarizabilities calculated with a modified dipole interaction. *Chemical Physics,* 59**,** 341-350.

TUCCINARDI, T. 2009. Docking-based virtual screening: recent developments. *Combinatorial Chemistry &# 38; High Throughput Screening,* 12**,** 303-314.

UMEYAMA, H., KITAURA, K. & MOROKUMA, K. 1975. Energy decomposition analysis along the reaction coordinate. Theory and example: F−+ HF→[FHF]−. *Chemical Physics Letters,* 36**,** 11-15.

VEENSTRA, D. L., FERGUSON, D. M. & KOLLMAN, P. A. 1992. How transferable are hydrogen parameters in molecular mechanics calculations? *Journal of computational chemistry,* 13**,** 971-978.

VERDONK, M. L., BERDINI, V., HARTSHORN, M. J., MOOIJ, W. T. M., MURRAY, C. W., TAYLOR, R. D. & WATSON, P. 2004. Virtual screening using protein-ligand docking: avoiding artificial enrichment. *Journal of chemical information and computer sciences,* 44**,** 793-806.

VERDONK, M. L., COLE, J. C., HARTSHORN, M. J., MURRAY, C. W. & TAYLOR, R. D. 2003. Improved protein–ligand docking using GOLD. *Proteins: Structure, Function, and Bioinformatics,* 52**,** 609-623.

VERDONK, M. L., GIANGRECO, I., HALL, R. J., KORB, O., MORTENSON, P. N. & MURRAY, C. W. 2011. Docking performance of fragments and druglike compounds. *Journal of medicinal chemistry,* 54**,** 5422-5431.

WANG, J., WANG, W., KOLLMAN, P. A. & CASE, D. A. 2001. Antechamber: an accessory software package for molecular mechanical calculations. *J. Am. Chem. Soc,* 222**,** U403.

WANG, J., WOLF, R. M., CALDWELL, J. W., KOLLMAN, P. A. & CASE, D. A. 2004. Development and testing of a general amber force field. *Journal of computational chemistry,* 25**,** 1157-1174.

WARREN, G. L., ANDREWS, C. W., CAPELLI, A. M., CLARKE, B., LALONDE, J., LAMBERT, M. H., LINDVALL, M., NEVINS, N., SEMUS, S. F. & SENGER, S. 2006. A critical assessment of docking programs and scoring functions. *Journal of medicinal chemistry,* 49**,** 5912-5931.

WEINER, P. K. & KOLLMAN, P. A. 1981. AMBER: Assisted model building with energy refinement. A general program for modeling molecules and their interactions. *Journal of computational chemistry,* 2**,** 287-303.

WELCH, W., RUPPERT, J. & JAIN, A. N. 1996. Hammerhead: fast, fully automated docking of flexible ligands to protein binding sites. *Chemistry & biology,* 3**,** 449-462.

WIBERG, K. B. & RABLEN, P. R. 1993. Comparison of atomic charges derived via different procedures. *Journal of Computational Chemistry,* 14**,** 1504-1518.

WILLIAMS, D. E. 1990. Alanyl dipeptide potential-derived net atomic charges and bond dipoles, and their variation with molecular conformation. *Biopolymers,* 29**,** 1367-1386.

WILLIAMS, D. E. 1999. Improved intermolecular force field for crystalline hydrocarbons containing four- or three-coordinated carbon. *Journal of Molecular Structure,* 485–486**,** 321-347.

WILLIAMS, D. E. 2001a. Improved intermolecular force field for crystalline oxohydrocarbons including O⬚H⬚O hydrogen bonding. *Journal of computational chemistry,* 22**,** 1-20.

WILLIAMS, D. E. 2001b. Improved intermolecular force field for molecules containing H, C, N, and O atoms, with application to nucleoside and peptide crystals. *Journal of computational chemistry,* 22**,** 1154-1166.

WINN, P. J., FERENCZY, G. G. & REYNOLDS, C. A. 1997. Toward improved force fields. 1. Multipole-derived atomic charges. *The Journal of Physical Chemistry A,* 101**,** 5437-5445.

WINN, P. J., FERENCZY, G. G. & REYNOLDS, C. A. 1999. Towards improved force fields: III. Polarization through modified atomic charges. *Journal of Computational Chemistry,* 20**,** 704-712.

YOUNG, T., ABEL, R., KIM, B., BERNE, B. J. & FRIESNER, R. A. 2007. Motifs for molecular recognition exploiting hydrophobic enclosure in protein–ligand binding. *Proceedings of the National Academy of Sciences,* 104**,** 808.

ZIMMERMAN, S. S., POTTLE, M. S., NÉMETHY, G. & SCHERAGA, H. A. 1977. Conformational analysis of the 20 naturally occurring amino acid residues using ECEPP. *Macromolecules,* 10**,** 1-9.

ZOETE, V., GROSDIDIER, A. & MICHIELIN, O. 2009. Docking, virtual high throughput screening and in silico fragment-based drug design. *Journal of cellular and molecular medicine,* 13**,** 238-248.

3 Inclusion of polarization in docking

3.1 Introduction

Fragment based drug design is an exciting new approach to drug design that is highly reliant on biophysical methods, particularly those related to structure (Congreve et al., 2008). In concert with this, it is important to fully exploit computational methods, both to identify new lead fragments and once these have been identified to direct the way they are grown into potential lead compounds. Virtual screening therefore plays a significant part in this process. Moreover, for the growing process to be effective, it is important for the modelled interactions to be as accurate as possible. While current docking programs such as Autodock (Ewing et al., 2001) and Glide (Friesner et al., 2004) have been effective in many applications, there are still a number of known effects that influence docking that are not included in such programs, partly because the programs need to be fast so as to handle the huge number of molecules in commercial screening collections. Here we restrict ourselves to an investigation of the effect of polarization on docking. For the docking experiments, we have investigated polarization of the QM/MM calculated charges using the induce charge method as explained in 2.6 and 3.2.8. The three polarization treatments were 1. neglect, 2. polarization of the QM ligand only in the field of the MM protein, and 3. mutual polarization of MM protein by the MM polarized QM ligand, where iteration to convergence was allowed (e.g., typically 3 iterations).

Furthermore, to mimic the effects of applying polarization, when the answer is not known we decided to carry out the two active polarization treatments at a number of geometries. These geometries were divided into two groups: control experiments where the answer is known, and experiments to simulate the situation where the answer is not known but arrived at using various methods described in Table 3.1. In the spirit of the requirement for rapid docking, are induced charge polarization method produces a set of accurate point charges, so that the

docking program can work in its usual manner, albeit with modified (polarized) atomic charges.

Table 3.1 The geometries where polarization occurs in our charge experiments, and the abbreviations used to describe them in figure 3.1 and figure 3.3.

GEOMETRY POLARIZED	METHOD OF ACQUIRING GEOMETRY
GLIDE SP	Not polarized (neglect).
POSE	Polarized at the geometry of the top ranking glide pose (GSCORE protocol).
RMSD	Polarized at the lowest RMSD geometry to the x-ray experimental result, measured in angstroms (Å). This is a control.
ENERGY SP	Polarized at the lowest single point energy geometry using QSITE QM/MM electrostatic potential fit calculation at the B3LYP 6-31G* level using the IMPACT 5.8 and JAGUAR 7.9 environments.
ENERGY	Polarized at the lowest optimized energy geometry using QSITE QM/MM electrostatic potential fit calculation at the B3LYP 6-31G* level using the IMPACT 5.8 and JAGUAR 7.9 environments.
REF	Polarized at the X-ray experimental result after Maestro 9.3 protein preparation wizard applied. This is a control.
INDIV	Polarized at the individual geometries of each of the requested ~15 rank ordered poses presented from Glide SP docking for each of the 74 molecules.

Both Autodock and Glide use scaled atomic charges as part of their parameterization, which is part of the implicit treatment of hydration. However, Glide uses a number of scoring methods, namely GSCORE, CVDW and EMODEL, which place a differing emphasis on the electrostatics. Consequently, we will use all three methods so as to more fully explore the effects of polarization. It could be argued that induced dipoles would offer an improved approach to polarization. However, such an approach would be inconsistent for a number of reasons, not least because it would mean that the major electrostatic effects were treated at the monopole level while the minor polarization effects would be treated at a higher level. It is true that induced charges do not capture 100% of the polarization effect as the degree of

anisotropy possible is dependent on the molecular shape. However, it should be noted that there are also other major deficiencies in docking programs, as discussed in Chapters 4 and 5, and so an over-emphasis on polarization would be inappropriate as long as the other errors are ignored. Nevertheless, polarization is relevant to the case of docking to a rigid enzyme, because the polarization of repulsive charges can help to alleviate this effect (Illingworth et al., 2008b).

3.2 Methods

3.2.1 Glide Exhaustive Search Algorithm

Glide was designed to perform as close to exhaustive search as is feasible, of the ligand pose space, also while achieving sufficient computational speed to screen large libraries. Glide achieves this by using a novel multi-grid approach. Like more conventional grid representations, Glide treats the receptor conformation as a rigid object, thus reducing computational cost by allowing the pre-computing of the molecular mechanics potential for the receptor and mapping it onto a number of grid points. The novel multi-grid technology uses several resolutions of grid-boxes where the energy gradient attributable to a particular ligand atom in the field of the protein can be rapidly computed via standard interpolation techniques. The hierarchal search itself employs an increasing level of sophistication, with every round of ligand conformation eliminations. The flexible ligand conformations are explored as docked poses in a manner approximate to the incremental search (see **2.2.3.1**). To account for solvation, water molecules are also docked with the energetically favourable ligands, as isoelectronic 2.8 Å spheres on the energy grid. As mentioned in section **2.2.5.2**, soft docking is included in a 1.0 and 0.8 setup where non-polar ligand atom radii are scaled to 0.8 to allow for minor steric clashes (Banks et al., 2005).

The early stages eliminate conformations with poor steric matches, and the intermediate stages adopt 'rough scoring' and 'greedy scoring', where an experimentally parameterized heuristic screening function eliminates high energy conformers unsuitable for

binding using a truncated version of the OPLS Force-field. In the later stages of posing, the final typically 400 poses are assessed through energy minimization using the multi-grid approach. The final ~3-6 poses are subjected to Monte Carlo simulation (see 2.2.3.2); the poses are then re-scored by EMODEL (see 3.2.6.3) which serves to direct the final conformational-search algorithm (Friesner et al., 2004).

3.2.2   Glide /IMPACT Molecular Mechanics environment

3.2.2.1 Atom Typing/Parameter fitting

Atom typing and parameter fitting programs such as AMBER's Antechamber are used to: recognize the atom type and bond type, judge atomic equivalence, find reasonable substitutes for missing force field parameters and then generate a residue topology file (Wang et al., 2001).

The IMPACT – Integrated Modelling Program, Applied Chemical Theory, Molecular Mechanics environment used in this work by Glide employs atom types that are obtained by fitting the molecules' substituent groups to parameterized functional groups covered by the force-field.   These substituent groups are stored as character strings that that use similar notation to SMILES/SMARTS language (Weininger, 1988), which is used in cheminformatics. The SMARTS algorithm relies on the Lewis structure of the molecule, which is defined by the atomic number and formal charge of each atom as well as the bond orders of each covalent bond. The Lewis structure may be available from the topology file, but when it is inconsistent or unavailable it is calculated iteratively from the formal charge and valence data, computed as the sum of bond orders, for every atom (Banks et al., 2005).

3.2.2.2 OPLS 2005 Parameterization

The OPLS 2005 force-field was used in this work to describe the MM region; it evolved from the AMBER force-field. Initially, the stretching and angle bending parameters were largely taken from AMBER, while torsional and non-bonded parameters were derived computationally and predicted parameters compared until in good agreement with available

experimental results (Jorgensen et al., 1996). Further parameters were then added for amines that better described hydrogen bonding, as a result improving the ordering of free energies of hydration (Rizzo and Jorgensen, 1999). The AMBER parameters were later replaced using accurate 'ab initio' data derived using quantum mechanical LMP2/cc-pvTZ(-f)/HF/6-31G** calculations, where it was shown that a sufficiently large basis set such as cc-pvTZ(-f) could yield excellent results when compared to experiment using a Localized Moller-Plesset method- LMP2 (Kaminski et al., 2001, Murphy et al., 1995).

3.2.3   Source of Molecules used for Validation Dataset

The search for molecules suitable for this validation set was simplified by the existence of a purpose built fragment data set, namely the SERAPhiC set (Favia et al., 2011). These fragment complexes were extracted from the Protein Databank (Berman et al., 2003). To ensure high quality, the criteria for acceptance within the SERAPhiC dataset were firstly that: the resolution of the X-ray crystal structure was <2.5 Å; that X-ray diffraction data was available; that the structure was a recent submission (after the year 2000) and that it was subject to a scientific study resulting in publication. Specific considerations were that the protein contained no mutations, had a polymer type that was not DNA or RNA, and that the protein contained less than 200 residues. The considerations for the ligand were that the molecular weight (MW) fell into the fragment range >78 (MW for benzene) and <300, and that the ligand was not commonly found in crystallization buffers (e.g., sulphates, glycol etc) (Favia et al., 2011). In addition, the four distinct proteins where the active site was located at the interface between two chains were not used in our final dataset (e.g., 1yki, 2p1o, 2wx, 3dsx).   This gave a total of 54 complexes, which included some cases with more than one ligand (e.g., 1e2i, 2hdq, 2q6m). In addition to the SERAPhiC dataset, two more sources of molecules that when used alone formed the preliminary dataset and in conjunction with SERAPhiC formed the final dataset. These were firstly 12 molecules (namely 1eqg; 1fv9; 1gwq; 1n1m; 1qwc; 1sj9; 1wcc; 1yz3; 2adu; 2c90; 2jjc; and 2ohk) which are popular within fragment datasets (Sándor et al., 2010) and initially selected as case studies to cover a

representative varied range of targets where fragment screening has been successful, often resulting in molecules that have reached phase two clinical trials (Congreve et al., 2008). Then secondly, some fragments (e.g., 3imc; 3ime; 3img) were used as examples of molecules screened in Mycobacterium Tuberculosis Pantothenate Synthetase (Hung et al., 2009). This produced 8 molecules by considering the steric effects of docking the two fragment ligands in both chains. The final tally for this dataset was therefore 74 molecules.

3.2.4 Ligand and Protein Preparation

The crystal structures were imported in pdb format and then prepared using the Protein preparation Wizard in Maestro (Banks et al., 2005) as follows. The bond orders were assigned, hydrogen atoms were added, and metal atoms were treated using default parameters, and protein capping was applied where absent. Then at this point additional chains were removed from the structure. There were initially, two approaches to the subsequent assignment of the hydrogen bond network and MM minimization of the protein. In the first approach, the water molecules and ions and common crystallization buffer molecules not interacting with the ligand were deleted prior to hydrogen bond network assignment and MM minimization. In the second approach, the water molecules, ions and common crystallization buffer molecules were retained during hydrogen bond network assignment and MM minimization before being deleted.

Test results from the preliminary dataset of 20 molecules (Congreve et al., 2008, Hung et al., 2009) docked using Glide SP gave a docking success rate at 55% for the first approach and 75% for the second approach to protein preparation, as judged by docking to within a threshold of < 2 Å RMSD (results not shown). Consequently, the second approach was taken even though this left less scope for improvement when considering polarization, but gave Glide a better chance to get the right answer, which seemed the most appropriate approach. A possible reason for this improved model performance is that small binding pockets shrink less when minimized with the water molecules present, thus allowing for fewer steric concerns.

The Hydrogen bond assignment was optimized using exhaustive sampling to assist in selection of correct rotamer and ionization states of the amides (Asn and Gln), hydroxyls (Ser, Thr, and Tyr), thiol groups (Cys), and Histidines (His). The option for these decisions, partially informed by pKA calculations carried out within Maestro using PROPKA 3.0 (Olsson et al., 2011), was also set and the sensible option (Barillari et al., 2007) to remove waters with less than 3 non water Hydrogen bonds was temporarily disabled.

To complete formation of each reference complex, restrained energy minimization was carried out on both the protein and ligand to within 0.3 Å RMSD from the X-ray experimental structure. This was done within the IMPACT 5.8. Molecular Mechanics environment using the OPLS 2005 force-field parameters (Banks et al., 2005) –see sections 3.2.2 and 2.5.4.

The EPIK 2.3 Program (for assigning pKa's) was unavailable for this research, so for consistency default tautomerization states were used in this work, despite one or two alternatives being identified through visualization.

3.2.5   Docking Setup

3.2.5.1 Receptor Grid Generation

Based on the results from a few preliminary test molecules for each protein structure, the default outer grid box and the subsequent area that was considered by the molecular mechanics energy potential was increased from the default 30 x 30 x 30 $Å^3$ box to the maximum 46 x 46 x 46 $Å^3$ box. The inner grid box in which the posing through the exhaustive search algorithm see 3.2.1 was done was kept at the default 10 x 10 x 10 $Å^3$.

This large outer grid box although not recommended (Glide manual) because of a ~2-3 fold increase in computational expense for receptor grid generation, still kept grid generation times within ~5 minutes using an intel i7 860 cpu @2.8 GHz desktop PC with 16 GB of DDR3 non ECC memory @1066 MHz. Therefore, we persisted with this larger treatment area for consistency, despite retrospectively only slight changes in docking performance. The same grids were used to dock the explicitly polarized ligand geometries

(see 3.2.8), but when the newly polarized enzyme partial charges were used instead the 'use input partial charges' option was ticked in the receptor tab. In all other respects default parameters were used, and no constraints were included. Grid generation although largely created manually within the Maestro 9.3 GUI (Banks et al., 2005), was in one case for individual polarization of ~15 ligand pose complexes (see 3.2.8) generated by a perl script from a template ligand specific input file. In this instance, centroids were calculated from Cartesian co-ordinates and considering the large number of poses to generate grids for and the subsequent use of 'score in place' docking protocol, the outer box size was reduced to 15 x 15 x 15 Å.[3]

3.2.5.2 In Consideration of Docking Protocols

The Glide SP docking algorithm was used solely in this research into the effects of full explicit polarization. Glide SP was selected for its 'softer' more forgiving potential as it is considered flexible and still fast enough to screen large ligand libraries (Friesner et al., 2004).In addition, this algorithm provides a larger number of possible pose solutions, when compared with Glide XP. Thus, it was considered to have higher potential to recreate the experimental binding and explore low lying energy minima when the top ranked pose is wrongly scored. Glide XP, an alternative 'harder' more chemically aware docking algorithm (Banks et al., 2005), designed for lead optimization, and more focused studies on only a small number of reliable confirmed ligand candidates (Friesner et al., 2006), Glide XP was also explored for performance across the full dataset. These results, with the addition of several different docking protocols and sampling schemes some including polarization, can be found in Appendix A.

3.2.5.3 Docking Protocol

For Glide SP and all but one of the polarization schemes, default docking protocols and sampling schemes were used in Glide re-docking, i.e. we used the Glide SP flexible ligand docking option, with post processing (or pose docking) MM energy minimization of the pose candidates. 5000 ligands were considered in the initial docking phase, with the 400 best

poses kept for energy minimization after a rough score sorting phase, and the final ~3-6 poses subjected to Monte Carlo stochastic methods. Glide was set to write a report file that included the best 15 poses per ligand.

3.2.6   Scoring Methods

3.2.6.1   GSCORE

GSCORE or Glide scoring is designed to rank ligands with varying degrees of similarity e.g., different net charges, from within a virtual compound library. It uses the empirical based ChemScore function as a start point (Eldridge et al., 1997):

$$\Delta G^0 = C_0 + C_{lipo} \sum f(r_{1r}) + C_{hbond} \sum g(\Delta r)h(\Delta \alpha) + C_{metal} \sum f(r_{1m}) + C_{rotb}H_{rotb}$$

(3.1)

In Eq.3.1, the second term extends over all ligand-atom/receptor-atom pairs that are defined by ChemScore as lipophilic, the third term extends to ligand/receptor hydrogen-bonding interactions. Here f; g; and h serve as weighting nominals (scored between 1.00-0.00) that give a full score when optimal. Here, (g) is for distance, (h) is for angle and (f) is a factor or fractional term related to metal and lipophilic interactions e.g., $g(\Delta r)$ is 1.00 if the H···X hydrogen bond distance is <0.25 Å of a nominal value of 1.85 Å but approaches zero in a linear fashion if the distance lies between 2.10 and 2.50 Å. Similarly (carbonyl scenario), the $Z-H···X$ angle is within 30^0 of 180º and decreases to zero between 150^0 and 120^0 (Friesner et al., 2004).

GSCORE here refers to the softer Glide SP version of GSCORE, which is softer for more general screening use compared with Glide XP. The GSCORE extended version of the ChemScore function is shown as:

$$\Delta G^0 = C_{lipo-lipo} \sum f(r_{1r}) +$$

$$C_{hbond-neut-neut} \sum g(\Delta r)h(\Delta \alpha) +$$

$$C_{hbond-neut-charged} \sum g(\Delta r)h(\Delta \alpha) +$$

$$C_{hbond-charged-charged} \sum g(\Delta r)h(\Delta \alpha) +$$

$$C_{max-metal-ion} \sum f(r_{1m}) + C_{rotb}H_{rotb} +$$

$$C_{polar-phob}V_{polar-phob} + C_{coul}E_{coul} +$$

$$C_{vdW}E_{vdw} + solvation\ terms \tag{3.2}$$

In Eq.3.2, the hydrogen bonding and lipophilic-lipophilic terms are taken in the ChemScore form, but the hydrogen bonding term is separated into differently weighted components. These depend on whether both the donor/acceptor are neutral (n-n); or both charged (c-c); or one from each (n-c) e.g., here n-n is the most stabilizing and the c-c least important. The metal-ligand interaction again uses the same form as ChemScore, but three important adjustments to improve metal-ligand descriptions are made: (i) It only considers interactions with anionic acceptor atoms by preference (ii) but only if the net charge of the metal in the apo protein is positive, (iii) it only counts the single best interaction when two metal ligations are found.

In addition to these adjustments, the coulombic term is reduced by ~50% on formally charged groups e.g., carboxylates; and the van der Waals interaction energy is also scaled on the atoms directly involved, this is done to make the gas-phase non-bonding terms a better predictor of binding. The CVDW force-field scoring below also uses these adjustments to ionic charge interactions except in the case of anionic ligand–metal interactions. EMODEL does not reducing the weighting of these interactions, as it already in part incorporates GSCORE.

Finally solvation is taken into account by docking explicit waters into the binding site of each of the energetically competitive ligand pose, and the water scoring is made efficient by the use of grid algorithms (Friesner et al., 2004).

3.2.6.2 CVDW - molecular mechanics force-field scoring

CVDW uses the IMPACT environment OPLS 2005 force-field based MM (Banks et al., 2005); the setup of this environment is described in 3.2.2. It uses the MM methods described in 2.5.4, but, with the adjustments described at the end of 3.2.6.1above.

3.2.6.3 EMODEL

EMODEL uses heavily-weighted force-field scoring that is a specific combination of CVDW (without formal charge weighting - see 3.2.6.1), ligand strain energy and GSCORE. It is used to rank order poses of the same ligand, EMODEL is also considered the best at distinguishing the experimental pose from geometric decoys and as such is used to direct the final stages of the conformational search algorithm while posing (Friesner et al., 2004).

3.2.7 In Consideration of Explicit Polarization

Mixed QM/MM methods are currently considered to be amongst the most accurate methods applicable to molecular docking programs. These methods are particularly accurate when some consideration of polarization is made (Cho et al., 2005). The MM region usually polarizes the QM region and the MM region remains un-polarized (Nevertheless the MM region may contain implicit polarization by virtue of the fact that the atomic charges are sometimes set deliberately too high to include this effect in an average way), but polarization of the MM region is often considered to be the missing term in molecular docking programs (Jorgensen, 2007). Previous work within the research group has looked in detail at the accuracy and applications of polarization within QM/MM methods (Illingworth et al., 2006, Illingworth et al., 2008b, Illingworth et al., 2008c) using techniques originally developed by (Winn et al., 1999, Gooding et al., 2000, Ferenczy and Reynolds, 2001). Here we seek to undertake a more rigorous investigation using a state of art collection of commercial

programs namely GAUSSIAN 03 (Frisch et al., 2004) plus *Maestro 9.3, Glide 5.8, QSITE and JAGUAR 7.9 which are* part of the SCHRODINGER SUITE of molecular modelling programs (Murphy et al., 2000, Friesner et al., 2004, Halgren et al., 2004). These commercial programs will be combined with in-house software designed to implement the modelling of full induced charge polarization briefly described in the following methodology in 3.2.8.

3.2.8 Full Polarization using the Induced Charge method

MM treats the electrostatic energy as the sum of atomic charge interactions, which are evaluated using Coulomb's law (Eq.3.3).

$$E^{ele} = \frac{1}{2} \Sigma_{I,J} q_I^{per} q_J^{per} \frac{1}{r_{IJ}} \quad (3.3)$$

Here, E^{ele} is the Electrostatic energy, q^{per}_J is the Permanent atomic charge on atom J, which includes the average 'implicit polarization' by overestimating the permanent atomic charges. r_{IJ} is the distance between the interacting atomic charges on atoms I and J.

Quantum mechanics treats electrostatic interactions using an explicitly polarized electronic wave function. In typical Hybrid QM/MM methods, the QM region is polarized by the MM region and the MM region usually remains unpolarized.

The induced dipole method of polarization is an approach to polarization, with a large contribution from permanent charges (monopoles) and a small contribution from the induced dipole moments (Eq. 3.4) only make a small contribution to the method (Eq. 3.5).

$$\overrightarrow{\mu_A} = \alpha_A \vec{F}_A \quad (3.4)$$

Here, μ_A is the induced dipole, α_A is the atomic polarizability of atom A and F_A is the electrostatic field at atom A. The total energy from the interaction between the permanent atomic charges and the induced dipoles is given by equation (3.5).

$$E^{ele} = \frac{1}{2} \sum_{I,J} q_I^{per} q_J^{per} \frac{1}{r_{IJ}} + \sum_{I,J} q_I^{per} \left(\nabla \frac{1}{r_{IJ}} \right) \vec{\mu}_j^{ind} + \frac{1}{2} \sum_{I,J} \vec{\mu}_j^{ind} \left(\nabla\nabla \frac{1}{r_{IJ}} \right) \vec{\mu}_j^{ind} + E^{self}$$

(3.5)

Here, μ^{ind} is the induced dipole and E^{self} is self polarization energy (the energy to create the induced dipole). The induced charge method approximates the dipole from charges on neighbouring atoms (Eq. 3.6)

$$\vec{\mu}_A = \chi_A \cdot \vec{p}_A \quad (3.6)$$

Here, the vector $\vec{\mu}_A$ contains the induced dipoles, χ_A is a matrix based on the geometry and $\vec{p}_A$ is a matrix containing the partial induced charges.

The geometric term is based on (Eq. 3.7)

$$\chi_A = (\vec{r}_{B1-A}, \vec{r}_{B2-A}, \vec{r}_{B3-A}) \quad (3.7)$$

Here $\vec{r}_{Bn-A}$ is the monopole moment for each of the neighbouring atoms.

The partial induced charge is derived from (Eq. 3.8)

$$\vec{p}_A = \begin{pmatrix} p_{A(B1)} \\ p_{A(B2)} \\ p_{A(B3)} \end{pmatrix} \quad (3.8)$$

Here, Bn are the neighbouring atoms partial charges.

The induced dipole method can now be written as (Eq.3.9)

$$\vec{p}_A = \alpha_A \cdot (\chi_A^+ \chi_A)^{-1} \chi_A^+ \cdot \vec{F}_A \quad (3.9)$$

However, using the Taylor series expansion for electrostatic potential (ESP), the final formula can be truncated and generalized and written as (Eq.3.10)

$$\boxed{\vec{p}_A = \alpha_A \cdot (\chi_A^+ \chi_A)^{-1} \cdot \vec{\Delta}(\Phi_A)}$$

(3.10)

Here Φ_A is the electrostatic potential determined from the wave function. As long as derivatives are not required (as in this work), equation (3.10) avoids the calculation of the electrostatic field required by equation (3.4).

For a full and detailed explanation of this method see (Ferenczy and Reynolds, 2001).

3.2.9 Programs used in Induced Charge method and Comparisons

The *Maestro 9.3* all purpose molecular modelling environment (Banks et al., 2005) is used to retrieve and prepare the protein complex. *Glide 5.8* (Friesner et al., 2004, Halgren et al., 2004) is used to perform docking tasks within the *Maestro 9.3* environment. The *IMPACT 5.8* program is used for non QM minimizations which call upon the OPLS 2005 force-field (Banks et al., 2005, Jorgensen et al., 1996, Kaminski et al., 2003). The Qsite program is used within the Maestro 9.3 environment to prepare hybrid QM/MM calculations (see Chapter 4) that call on the JAGUAR 7.9 program for the QM component of calculations (Cho et al., 2005, Murphy et al., 2000). The *Gaussian 03* program (Frisch et al., 2003) is used to calculate the Gaussian wave function, and GDMA 2.2 (Stone, 2005) is then used to assign DMA multipole expansions. The *Mulfit* 2.1 program (Illingworth et al., 2006) is used to derive improved point charges from the DMA multipole expansions.

3.2.10 Evaluation Methods

3.2.10.1 RMSD evaluation

The success of a docking program in predicting a ligand binding pose is most popularly measured by root mean-square deviation (RMSD). The RMSD metric, is used to study the displacement of atoms in a docked pose compared with the experimental pose (Cole et al., 2005, Baber et al., 2009) as shown:

$$rmsd = \sqrt{\frac{1}{N}\sum_{i=1}^{N} d_i^2}$$

(3.11)

Where d$_i$ is the Euclidean distance between *N* pairs of equivalent *i* atoms (Baber et al., 2009). The most frequently used statistic to describe docking success rate is the number of test complexes that have a 1st ranked pose of RMSD of ≤ 2 Å from the experimental pose (Cole et al., 2005, Warren et al., 2006, Sándor et al., 2010). However, it has been noted that low RMSD's do not always result in poses that maintain the experimental pose interactions, this is particularly true of small ligands such as fragments (Cole et al., 2005). In response to this many other evaluation metrics have emerged (Kroemer et al., 2004, Yusuf et al., 2008, Baber et al., 2009), but despite these alternate metrics RMSD remains popular.

It is hard to judge the value of a particular RMSD result e.g., if a pose has an RMSD of 0.3 Å or 0.6 Å from the experimental pose they may be of equal use to a medicinal chemist (Cole et al., 2005). However in a study of the RMSD deviations, from a dataset of 69 ligands with 5 or more rotatable bonds it was shown through visualisation that Glide invariably achieved 'close' to experimental interactions at RMSD of ≤ 1.5 Å (Kontoyianni et al., 2003). Also in a previous work, an RMSD of ≤ 2 Å was considered to be a good indication of docking success, while an RMSD of ≤ 4 Å a rather rough one (Warren et al., 2006). Considering the small size of the fragments docked here and the thresholds reported in the literature (Kontoyianni et al., 2003, Warren et al., 2006), we have decided to use 4 discrete thresholds of RMSD of ≤ 0.5 Å; RMSD of ≤ 1 Å; RMSD of ≤ 2 Å; and RMSD of ≤ 4 Å respectively. Although still somewhat subjective the first two thresholds may be viewed when comparing the top ranked pose to the experimental pose as maintaining close interactions, while the latter two thresholds are more approximately within the active site, with the RMSD of ≤ 2 Å threshold being of greater value, both to the medicinal chemist and to our ability to evaluate docking performance.

3.2.10.2   Boltzmann probabilities and clustering

Native retrospective docking studies, of the type reported here in chapters 3 to 5, report reasonably high accuracy of about 70-80% success, with success defined as docking the best pose of the native ligand to within 2 Å RMSD of its X-ray structure (Cho et al., 2005,

Warren et al., 2006, Kolb and Irwin, 2009). Typically docking programs use a scoring function to rank order each molecule's best pose, often from a database containing millions of distinct poses (Kolb and Irwin, 2009).

The poses are typically subjected to a clustering algorithm, as is the case with GLIDE (Friesner et al., 2004). This, in addition to a distinct selection of ligand molecules, offers an investigator an increased likelihood that the final rank ordered poses per molecule show useful conformational alternatives, e.g. for consideration during ligand optimization. Post-docking protocols may apply to a given number of top-ranked poses. However, the final decision on the molecules acquired and tested from the top scoring hits lies with the investigator. It has been suggested, that the cherry picking of compounds from expert experience, introduces a subjective bias, influencing the number of successful screens attributed to docking (Kolb and Irwin, 2009).

In this research, for the retrospective native docking studies chapters 3 to 5, we have a number of energy calculation methods involving e.g., grid-based energy calculations, QM/MM energies. The posing search algorithm of Glide SP (e.g., an initial flexible Glide SP dock) has been combined with these methods, to produce ~15 poses to address the problem of energy calculation only exploring local energy minima. We have made the decision to explore a level of depth within the rank ordered poses, searching for the correct docked pose, when it is not presented as the top ranked or lowest energy pose. However, these poses still display a level of convergence with energy minimization, and we would like to investigate the distinct poses.

Therefore, the first step, (with the exception of the pre-clustered and relatively low energy weighted GLIDE GSCORE scoring function), involved applying a simple divisive hierarchal clustering algorithm, to all methods and protocols. The algorithm assessed the similarity of docked poses using the distance metric of RMSD (measured in angstroms, see 3.2.10.1),

and also the linkage criteria of minimum binding energy, which was measured by relative energy in kcal mol^{-1} (or equivalent, in the case of EMODEL).

In addition to the hierarchal clustering algorithm, we also used the Boltzmann probability, as an assessment tool. In brief, all poses were sorted lowest to highest by energy. The Boltzmann distribution of the population was approximated (e.g. 15 poses considered to include all accessible states), and relative population values p_i, were calculated, Eq. 3.12.

$$p_i = \frac{e^{-1000\varepsilon_i/RT}}{\sum_{j=1}^{M} e^{-1000\varepsilon_j/RT}}$$

$$(3.12)$$

Here i is the pose, ε is the energy of state, j is the sum of accessible states i, where M is the number of accessible states (typically 15), T is the absolute temperature, here 298 K, R is the Universal gas constant, here 8.314 J mol^{-1} K^{-1}., The relative population values p_i estimates the probability of pose i, being accessed (Illingworth et al., 2008a). The p_i values are summed within a cluster of poses. The poses were clustered (generously) so that each cluster had an RMSD of ≤ 1 Å and energy within 1.1 kcal mol^{-1} of the lowest energy member of that pose cluster; the choice of 1.1 kcal mol^{-1} is somewhat arbitrary but was found to give reasonable clustering, as explained below.

An explorative heuristic approach (although approximately quantitative), was adopted for the (generous) cluster criteria. We aimed to produce the scenario where many pose clusters could potentially exist, but just one distinct alternative low lying minima would be energetically accessible, judged by the somewhat arbitrary Boltzmann probability p_i>10%. The threshold of 1.1 kcal mol^{-1} also prevented formation of a third accessible cluster with p_i>10% for 73 of the 74 cases. Therefore, the distinct lowest energy pose within the second pose cluster could be tested for reproducing the experimental binding within the RMSD of ≤ 2

Å threshold. This approach was useful when the overall lowest energy / top ranking pose failed within the RMSD of ≤ 2 Å, because it provides an approximate way of taking into account the error in the Glide scoring functions and energy calculations. The same clustering method and thresholds were used to assess if the pose closest to experiment, measured by RMSD, was also within a cluster containing the overall lowest energy or top ranked pose (e.g., Fig. 3.1C and E white bars, discussed in 3.3.2).

3.3 Results

3.3.1 Analysis of top ranked poses

3.3.1.1 Ligand polarization

Glide SP flexible ligand docking is usually able to identify the experimental pose amongst the full set of returned poses, whether as the top ranked pose, or as a lower ranked pose, as will be discussed later (see 3.3.3 and 3.3.4). Figure 3.1A (first bar) shows that the top ranked pose (according to Glide SP, i.e., GSCORE) is also the pose with the smallest RMSD to the experimental pose in 36% of the 74 cases. This 36% is therefore one of the benchmark results that can be used to see if explicit polarization can improve docking. When the ligand partial charges are explicitly polarized using (i) the geometries of the 1st Ranked GLIDE SP (GSCORE) pose (2nd bar, denoted POSE(L)), (ii) the lowest Energy QM/MM (Jaguar/Impact) single-point pose (4th bar, ENERGY SP(L)), (iii) the lowest Energy QM/MM (Jaguar/Impact) optimized pose (5th bar, ENERGY(L)), then flexible ligand re-docking results in the top ranked pose being the one with the smallest RMSD to the experimental pose in 38%, 41% and 38% of cases respectively. Under these circumstances, polarization does not improve docking. The key issue with this analysis is whether the ligand is correctly polarized – if the top pose is not similar to the experimental pose then the ligand may not be polarized correctly. To test whether a correctly polarized ligand results in better results, we have (i) polarized the ligand

in the pose with the lowest RMSD to the experimental pose and (ii) polarized the experimental pose. Fig 3.1.A, bar 3 (RMSD(L)) and bar 6 (REF(L)) show that under these circumstances, the top ranked pose is the one with the smallest RMSD to the experimental pose in (i) 36% and (ii) 53% of cases respectively. Explicitly polarizing of the ligand partial charges at each of the ~15 individual ligand geometries of the pose population then re-docking using the 'Score in Place' (7th bar, INDIV(L)) results in the top ranked pose being the one with the smallest RMSD to the experimental pose in 35% of cases. It is disappointing that, in general, explicit polarization of the ligand at one of the docked poses does not give an improvement in the results as determined by this measure. However, the controls show that polarization clearly can improve docking results as polarization of the experimental pose improves the results from 36% to 53%. However, this is not strong evidence alone, as polarization at the smallest RMSD to experimental pose shows no improvement.
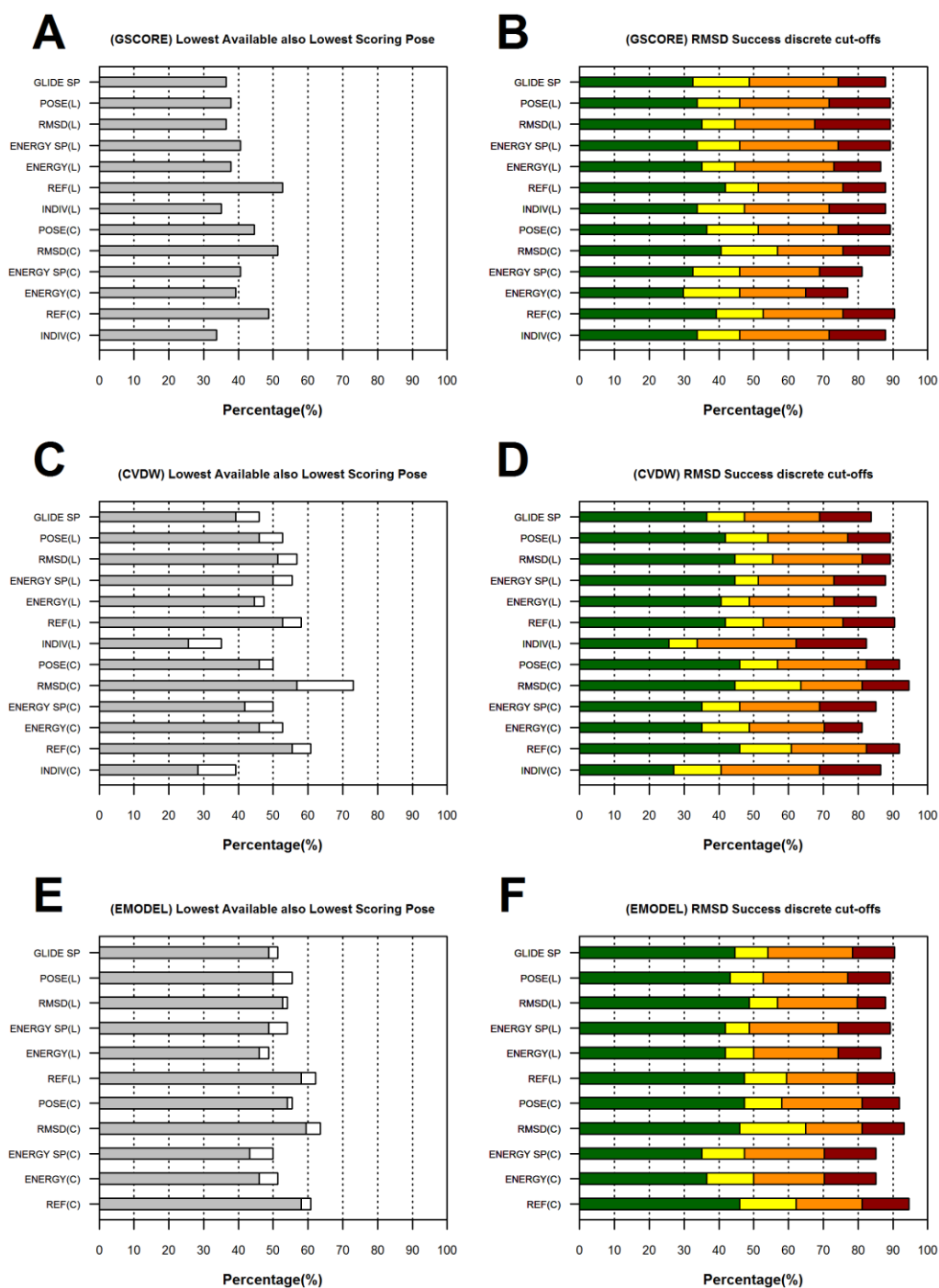
Figure 3.1. The effect of polarization on docking, for poses generated by Glide. For A and B, the poses (≤ 15) were generated and ranked using GSCORE. For C and D, the poses were generated according to GSCORE and ranked according to CVDW. For E and F, the poses were generated according to GSCORE and ranked according to EMODEL. The percentage of times that the top ranked pose is also the pose closest to the experimental pose, as measured by RMSD, is denoted in A, C and E by a grey bar; the percentage of times that a pose from the first cluster is also the pose closest to the experimental pose, as measured by RMSD, is denoted in A, C and E by a white bar. The percentage of cases that the top ranked pose is below the 0.5 Å, 1.0 Å, 2.0 Å and 4.0 Å thresholds is shown by green, yellow, orange and dark red bars respectively. In each panel, the first bar, denoted GLIDE SP, indicates docking with no polarization. In subsequent bars, (L) indicates that only the ligand was

polarized and (C) indicates that the ligand and protein were polarized. POSE indicates that the ligand was polarized using the geometry of the top ranked pose docked using glide, RMSD indicates that the ligand was polarized using the geometry of the pose closest to the experimental result, ENERGY SP indicates that the ligand was polarized using the geometry of the top ranked pose scored according to the single point QM/MM calculations, ENERGY indicates that the ligand was polarized using the geometry of the top ranked pose scored according to the QM/MM optimization calculations and REF indicates that the polarized charges for were determined at the experimental geometry. INDIV indicates that each of the initial 15 poses was polarized individually.

3.3.1.2 Ligand and protein polarization

In the previous section we considered polarization of the ligand; here we follow similar polarization strategies, but both the ligand and the protein are polarized. In Fig.3.1A, when the ligand and protein partial charges are explicitly polarized using (i) the geometries of the 1st Ranked GLIDE SP pose (8th bar, denoted POSE(C)), (ii) the lowest Energy QM/MM (Jaguar/Impact) single-point pose (10th bar, denoted ENERGY SP(C)), (iii) the lowest Energy QM/MM (Jaguar/Impact) optimized pose (11th bar, denoted ENERGY (C), then flexible ligand re-docking results in the top ranked pose being the one with the smallest RMSD to the experimental pose (according to Glide SP, i.e. GSCORE) in 44%, 41% and 39% of cases respectively (Fig. 3.1A). Under these circumstances, polarization alludes to improve docking, particularly for the first case with ~8% improvement over Glide SP benchmark of 36%, compared with ~2% improvement with ligand only polarization. Notably, this ~8% improvement is the best that comes without knowledge of the correct answer. Again, to test whether a correctly polarized ligand results in better results, we have (i) polarized the ligand and complex in the pose with the lowest RMSD to the experimental pose and (ii) polarized the experimental pose. Fig 3.1A, bar 9 (denoted RMSD(C)) and bar 12 (denoted REF(C)) shows that under these circumstances, the top ranked pose is the one with the smallest RMSD to the experimental pose in (i) 52% and (ii) 49% of cases respectively. Although, full polarization of the experimental pose, offers marginally less improvement than ligand only polarization. There is a case that polarization can improve docking. Indeed, when fully

polarized, both controls show clear improvement. Finally, success is only 34% when each individual pose complex partial charges is explicitly polarized before docking using 'score in place' occurs (13th bar, denoted INDIV(C)); this suggests there is no advantage in polarizing the ligand at each geometry.
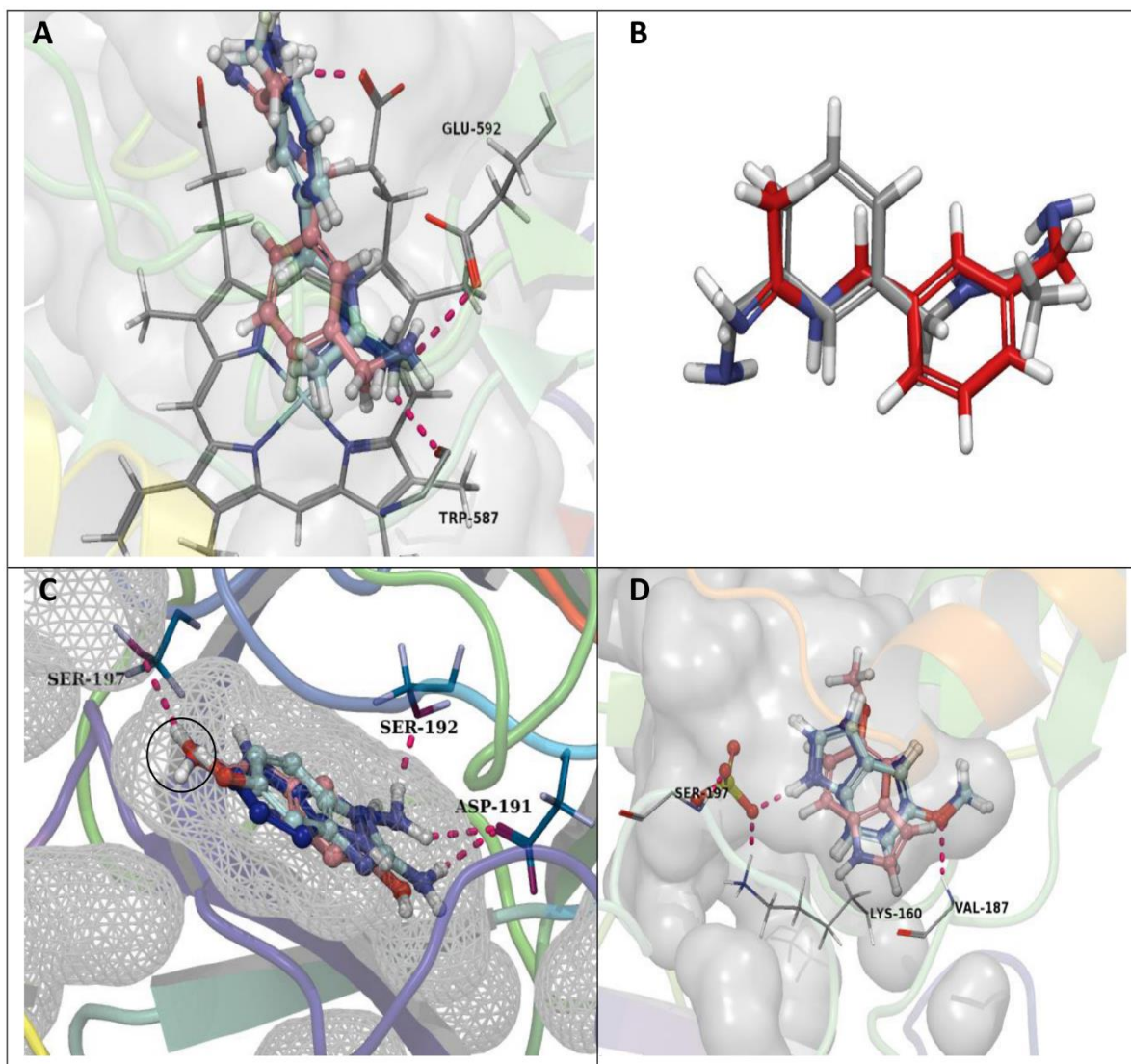

Figure 3.2 Examples of where polarized docking improves upon basic glide docking. In the three examples, basic glide docking causes the ligand to flip whereas fully polarized docking generates a native-like pose. A. Nitric oxide synthase oxygenase domain, pdb code 1qwc, showing the interactions of the heme, Trp[587] and Glu[592] with the ligand (14W). The reference ligand (X-ray pose) is in cyan, the fully polarized ligand is blue, and the basic Glide-docked ligand is pink. B. A Close up of the 1qwc reference ligand (grey) and Glide docked ligand (red), (shown in A) demonstrating a 180^0 rotation. C. Human microurokinase, pdb code 1fv9, showing key interaction between the ligand, 2-amino-5-hydroxy-benzimidazole, and Asp[191]. The reference ligand (X-ray pose) is in cyan, the fully polarized ligand is blue, and the basic Glide-docked ligand is pink. The fully polarized ligand undergoes $\sim40^0$ rotation to interact with Ser[192], while the basic Glide-docked ligand flips to interact with Ser[197], giving a much larger error. D. Mycobacterium Tuberculosis Pantothenate Synthetase, pdb code, 3img (chain

B) BZ3 showing the interaction of the ligand with Val[187] and a bridged sulphate ion. The reference ligand (X-ray pose) is in cyan, the QM/MM optimized ligand is pink. The basic Glide-docked ligand flips (results not shown).
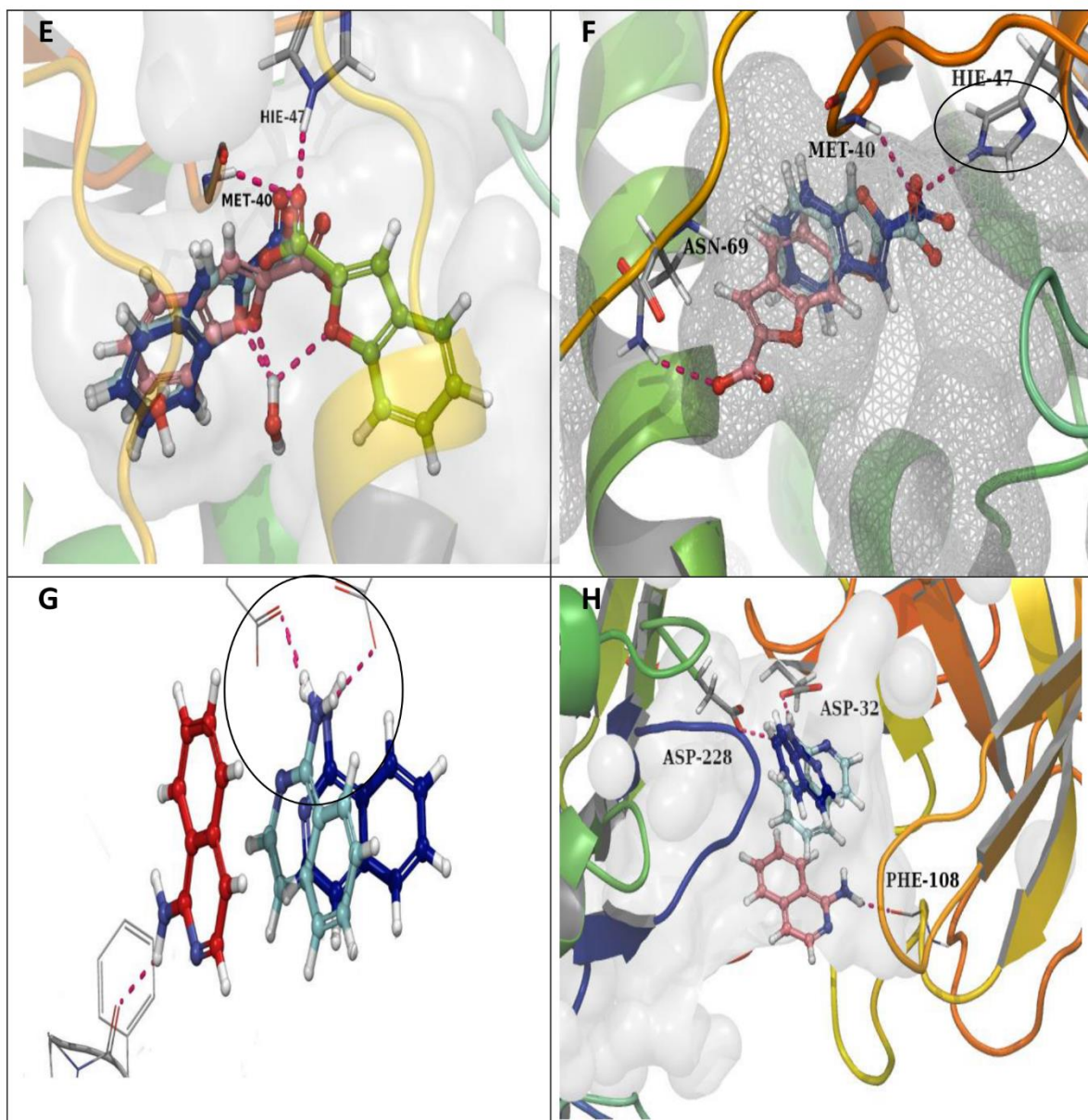

Figure 3.2 (continued). E. Mycobacterium Tuberculosis Pantothenate Synthetase, pdb code 3img, showing interactions of His[47], Met[40], and HOH[627] with ligand. The reference ligand (X-ray pose) is in cyan, the fully polarized ligand is blue, the basic Glide-docked ligand is pink and the optimized QM/MM method is in green. All the QM methods (e.g. QM/MM opt) cause a symmetrical pivot along the key interaction axis stabilized by the HOH[627] interaction. F. Mycobacterium Tuberculosis Pantothenate Synthetase, pdb code 3ime, showing interactions of His[47] and Met[40] with ligand. The reference ligand (X-ray pose) is in cyan, the QM/MM optimized ligand is blue, the basic Glide-docked ligand is pink. The basic Glide-docked ligand moves to interact with Asn[69]. beta secretase, pdb code 2ohk, The reference ligand (X-ray pose) is in cyan, the fully polarized ligand is blue, the basic Glide-docked ligand is red/pink. H. An alternative view of G, showing how the Glide docked ligand (1-amino-isoquinoline) moves to interact with Phe[108].

3.3.2 Alternative approaches to identifying the top pose

The results of Figure 3.1A are interesting as they suggest that some improvement can be obtained by polarizing just the top ranked pose, provided that both the ligand and protein are polarized. For this reason, we have considered alternative scoring protocols for ranking the poses, namely CVDW (Figure 3.1C) and EMODEL (Figure 3.1E).

3.3.2.1 CVDW rescoring: Ligand polarization

In Figure 3.1C (first grey bar, denoted GLIDE SP) we consider the top ranked pose (docked according to (GSCORE) but with the poses rescored according to CVDW); the figure shows that this pose (and for subsequent grey bars) is also the one with the smallest RMSD to the experimental pose in 39% of the 74 cases. This 39% is a slight improvement over the GSCORE ranking of Figure 3.1A (bar 1, denoted GLIDE SP). Moreover, the poses have been clustered (generously) so that each cluster has an RMSD of ≤ 1 Å and energy within 1.1 kcal mol^{-1} of the lowest energy member of that pose cluster (see 3.2.10.2).

When we set the criteria that the pose with the lowest RMSD to experiment could also be a member of the first cluster, then success increases to 46% (white bar(s) in Fig. 3.1C). When the ligand partial charges are explicitly polarized using (i) the geometries of the 1st Ranked GSCORE pose (2nd bar, denote POSE(L)), (ii) the lowest Energy QM/MM (Jaguar/Impact) single-point pose (4th bar, denoted ENERGY SP(L)), (iii) the lowest Energy QM/MM (Jaguar/Impact) optimized pose (5th bar, denoted ENERGY(L)), then flexible ligand re-docking results in the top ranked pose being the one with the smallest RMSD to the experimental pose in 46%, 50% and 44% of cases respectively; this rises to 52%, 55% and 47% when the cluster rather than the first pose is considered. Under these circumstances, ligand polarization does give improved results. The control results where the ligand is polarized at either the lowest RMSD structure (3rd bar, denoted RMSD(L)) or the experimental structure (6th bar, denoted REF(L)) are 52% and 53% respectively, rising to 57% and 58% when the second cluster is considered, giving clear evidence that ligand polarization can improve the docking results.

Explicit polarization of the ligand partial charges at the ~15 individual ligand geometries of the pose population then re-docking using the 'Score in Place' option (7th bar, denoted INDIV(C)) lowers the success to 26% for the first pose and 35% when the second pose cluster is considered, again indicating that there is no value in polarizing poses other than the top-ranked pose.

3.3.2.2 CVDW rescoring: Ligand and protein polarization

Because there is already improvement resulting from *ligand* polarization, the scope for further improvement from polarizing both the *ligand and the protein* is reduced. Nevertheless, polarizing the ligand and protein does give improvements over Glide SP. Thus, the control results where the ligand is polarized at either the lowest RMSD structure (9th bar, denoted RMSD(C)) or the experimental structure (12th bar, denoted REF(C)) are 57% and 56% respectively, rising to 73% and 61% when the first second pose cluster is considered. These figures are well above the initial results of 39% (top pose) and 46% (first cluster), and show that polarization does improve docking success. The practical approaches, relevant to when the experimental result is not known and where polarization is based on the top pose determined by the scoring methods, gives results in the range of 42%-46% (8th bar, 10th bar, 11th bar), which rise to 50-53% when the first cluster is considered. For this to work, it is essential that the top ranked pose has features of the correct pose; it seems that the CVDW score is more appropriate to this task than GSCORE. Finally, success is only 28% when the individual pose complex partial charges are explicitly polarized before re-docking with 'score in place' occurs (bar 13), this rises to 39% for the first pose cluster.

3.3.2.3 EMODEL rescoring: Summary

Figure 3.1E shows that EMODEL gives very similar results for explicit polarization to scoring with CVDW (especially when considering re-scoring by the lowest Energy QM - Jaguar/MM - Impact poses), and thus offers an alternative method of selecting the top ranked pose for polarizing the ligand and the enzyme. However, there is less scope for improvement as when we consider the top ranked pose (docked according to GLIDE SP/GSCORE but with the

poses rescored according to EMODEL); the figure shows that this pose (1st grey bar – all colour coding as Figure 3.1C) is also the one with the smallest RMSD to the experimental pose in 48% of the 74 cases. This 48% is a significant improvement over the 36% achieved by GSCORE or the 39% achieved by CVDW ranking of Figure 3.1A (bar 1, denoted GLIDE SP) and Figure 3.1C (bar 1, denoted GLIDE SP) respectively. Moreover, when we set the criteria that the pose with the lowest RMSD to experiment should also be a member of the cluster as with CVDW, then success increases slightly to 52% (1st white bar (denoted GLIDE SP) in Fig. 3.1E). This finding is consistent with EMODEL being Glide's best scoring function for ranking poses of the same ligand molecule (Friesner et al., 2004). Despite the reduced scope for improvement, we still see minor increases in performance particularly when considering the *ligand* (2nd bar, denoted POSE(L)) and then *complex* partial charges (7th bar, denoted POSE(C)) that are explicitly polarized when using the geometries of the 1st Ranked GSCORE pose then flexible ligand re-docked. This results in the top ranked pose being the one with the smallest RMSD to the experimental pose in 50%, and 54% of cases respectively; this rises to 56% and also 56%, when the first cluster rather than the first pose is considered.

The control results where the ligand (3rd bar) or complex (8th bar) is polarized at the lowest RMSD structure are 53% and 59% respectively; this rises to 54% and 63%, when the first second pose cluster rather than the first pose is considered. The control results where the ligand (6th bar) or complex (11th bar) is polarized at the experimental structure are 58% and also 58% respectively, rising to 63% and 61% when the first cluster is considered. So here all the non-clustered first pose controls slightly outperform CVDW scoring (e.g., grey bars 3, 6, 8 and 11), while the clustered pose results are similar (with the exception of Fig. 3.1C bar 9, denoted RMSD(C)).

3.3.3 RMSD analysis

In sections 3.3.1 and 3.3.2, Figure 3.1A, 3.1C and 3.1E we considered the proportion of cases where the top ranked pose was also the one with the lowest RMSD to the

experimental structure. However, in some of these cases the lowest RMSD might still be rather high. In figures 3.1B, 3.1D and 3.1E we consider the proportion of molecules where the RMSD is below 0.5 Å (green); RMSD of ≤ 1 Å (yellow); RMSD of ≤ 2 Å (orange); RMSD of ≤ 4 Å (dark red) respectively.

3.3.3.1 Benchmark results

Figure 3.1B, 3,1D and 3,1E (1st bar, denoted GLIDE SP) shows percentage of times that the top ranked pose (according to GSCORE, cVDW and EMODEL) is within a given threshold; this threshold is the distance measured by RMSD from the experimental pose. These benchmark results shown in table 3.2 can be used to see if explicit polarization can improve docking.

Table 3.2 The percentage of times that the top ranked pose is within a given threshold, evaluated over all 74 cases.

	Threshold			
Method	0.5 Å	1.0 Å	2.0 Å	4.0 Å
GSCORE	32%	48%	74%	88%
cVDW	37%	47%	68%	84%
EMODEL	44%	54%	78%	91%

In the section that follows, we focus on the 0.5 Å GSCORE results (Figure 3.1B); the remaining results will be summarized.

3.3.3.1.1 RMSD of ≤ 0.5 Å from experimental geometry GSCORE (green bars)

When the *ligand* partial charges are explicitly polarized using (i) the geometries of the 1st Ranked GSCORE pose (2nd bar, denoted POSE (L)), (ii) the lowest Energy QM/MM single-point pose (4th bar, denoted ENERGY SP(L)), (iii) the lowest Energy QM/MM optimized pose (5th bar, denoted ENERGY SP(L)), then flexible ligand re-docking results in the top ranked

pose having an RMSD of ≤ 0.5 Å 33%, 33% and 34 % of the time, respectively. When the *complex* partial charges are explicitly polarized in the same fashion using (i) the geometries of the 1st Ranked GSCORE pose (8th bar, denoted POSE(C)), (ii) the lowest Energy QM/MM single-point pose (10th bar, ENERGY SP(C)), (iii) the lowest Energy QM/MM optimized pose (11th bar, denoted ENERGY(C)), then flexible ligand re-docking results in the top ranked pose having an RMSD of ≤ 0.5 Å 36%, 32% and 29 % of the time respectively.

To explore the same key issue as in 3.3.1.1 as to whether the *ligand* is correctly polarized. We have (i) polarized the *ligand* in the pose with the lowest RMSD to the experimental pose and (ii) polarized the experimental pose. When the *ligand* partial charges are explicitly polarized then flexible ligand re-docked we find the top ranked pose having an RMSD of ≤ 0.5 Å (i) 35% (bar 3, denoted RMSD(L)) and (ii) 42% (bar 6, denoted REF(L)) of the time respectively. When the *ligand and protein* partial charges are explicitly polarized instead, these change to (i) 41% (bar 9, denoted RMSD(C)) and (ii) 39% (bar 12, denoted REF(C)) respectively.

Explicit polarization of the *ligand* partial charges (7th bar, denoted INDIV (L)) and the *ligand and protein* partial charges (13th bar, denoted INDIV(C)) at each of the ~15 individual ligand geometries of the pose population then re-docking using the 'Score in Place' results in an RMSD of ≤ 0.5 Å from the experimental pose 33% and also 33% of the time. The controls (where the ligand is polarized correctly, i.e. RMSD (L/C) and REF (L/C)) show that polarization can offer up to a 10% improvement to the docking results. The average increase is ~6.5% for ligand polarization and ~8% for the ligand and protein polarization. Significantly, polarization gives a 4% improvement even when the answer is not known (POSE(C)). Overall these results show that polarization can make a positive effect in generating accurate poses with an RMSD of less than 0.5 Å, but the superior results of the control results shows that it is important to polarize the ligands correctly.

3.3.3.1.2 RMSD of ≤ 1 Å from experimental geometry GSCORE (yellow bars)

The controls show that polarization can offer up to a 9% improvement to the docking results. The average increase is ~0% for *ligand* polarization and ~7% for the *ligand and protein* polarization at this 1.0 Å threshold.

Significantly, polarization can give up to a 4% improvement even when the answer is not known, but here the average over all methods is 0%. Overall these results show that polarization can make an effect in generating accurate poses with an RMSD of less than 1.0 Å, but the effect is less marked than at the 0.5 Å threshold.

3.3.3.1.3 RMSD of ≤ 2 Å from experimental GSCORE (orange bars)

Here the controls show little advantage from polarization which at best offered a 2 % increase from 74% to 76% at this threshold. There is even an overall minor negative effect to GSCORE when the answer is not known. It is noteworthy that polarization has a positive effect at the lower thresholds of 1.0 Å and particularly 0.5 Å, but little effect at the 2.0 Å threshold. It is likely that at this higher level of inaccuracy (2.0 Å threshold) that factors other than polarization play a role in determining the accuracy. A similar conclusion arises from analysis of the results at the 4 Å threshold (dark red bars).

3.3.3.1.4 RMSD of ≤ 0.5 Å from experimental geometry CVDW (green bars)

There is slightly less scope for improvement compared to GSCORE at this threshold as there is a rise from 32% (GSCORE) to 37% (CVDW) for the basic unpolarized results. The controls show that polarization does again offer up to 10% improvement to docking results. The average increase is ~6% for the *ligand* and ~9% for *ligand and protein* polarization. Significantly, *ligand* polarization gives up to a 7% improvement when the answer is not known, and *ligand and protein* polarization gives up to a 9% improvement.

3.3.3.1.5 RMSD of ≤ 1 Å from experimental geometry CVDW (yellow bars)

The controls show that polarization can offer up to a 16% improvement to the docking results at this 1.0 Å threshold. The average increase is 7.5% for ligand polarization and 15%

for the ligand and protein polarization at this 1.0 Å threshold. We also see an improvement in results up to 7% for ligand polarization, and up to 10% for the ligand and protein polarization, when the answer is not known. This occurs when polarizing by ligand or ligand and protein at the geometries of the 1st Ranked GSCORE pose.

3.3.3.1.6 RMSD of ≤ 2 Å from experimental geometry CVDW (orange bars)

When the *ligand* partial charges are explicitly polarized using (i) the geometries of the 1st Ranked GSCORE pose (2nd bar, denoted POSE(L)), (ii) the lowest Energy QM/MM single-point pose (4th bar, denoted ENERGY SP(L)), (iii) the lowest Energy QM/MM optimized pose (5th bar, denoted ENERGY(L)), then flexible ligand re-docking results in the top ranked pose having an RMSD of ≤ 2 Å 76%, 73% and 73 % of the time respectively, compared to the benchmark results of 68%. When the *complex* partial charges are explicitly polarized in the same fashion using (i) the geometries of the 1st Ranked GSCORE pose (8th bar, denoted POSE(C)), (ii) the lowest Energy QM/MM single-point pose (10th bar, denoted ENERGY SP(C)), (iii) the lowest Energy QM/MM optimized pose (11th bar, denoted ENERGY(C)), then flexible ligand re-docking results in the top ranked pose having an RMSD of ≤ 2 Å 83%, 69% and 71% of the time respectively.

For the controls where we have: (i) polarized the ligand in the pose with the lowest RMSD to the experimental pose and (ii) polarized the experimental pose. When the ligand partial charges are explicitly polarized then flexible ligand re-docked we find the top ranked pose having an RMSD of ≤ 2 Å (i) 82% (bar 3, denoted RMSD(L)) and (ii) 76% (bar 6, denoted REF(L)) of the time respectively. When the ligand and protein partial charges are explicitly polarized instead this changes slightly to (i) 82% (bar 9, RMSD(C)) and (ii) 83% (bar 12, REF(C)) respectively.

The controls again show that polarization clearly does offer improvement to the docking results at this threshold as polarization of the ligand and protein partial charges improves the result by up to 15% (average ~14.5%). Polarizing the *ligand* partial charges improves results by up to 14% (average 8%). When the answer is not known, polarizing the

geometries of the 1st Ranked GSCORE pose again gives the best results (15% improvement, giving 83% of poses within 2 Å) particularly when considering *ligand and protein* polarization. Some improvement is also seen at the 4 Å threshold (not discussed).

It is interesting to note that in this CVDW section 3.3.3.1.6, but not in the GSCORE section 3.3.3.1.3 , polarization gave improved results. This is linked to the use of CVDW rather than GSCORE for ordering the poses and probably arises because the electrostatics are not scaled in the CVDW method.

3.3.3.2 EMODEL rescoring: Summary

Figure 3.1F shows that EMODEL gives very similar results to scoring with CVDW, but as the initial unpolarized method gives a higher percentage of poses with the threshold, the scope for improvement is less (results not discussed).
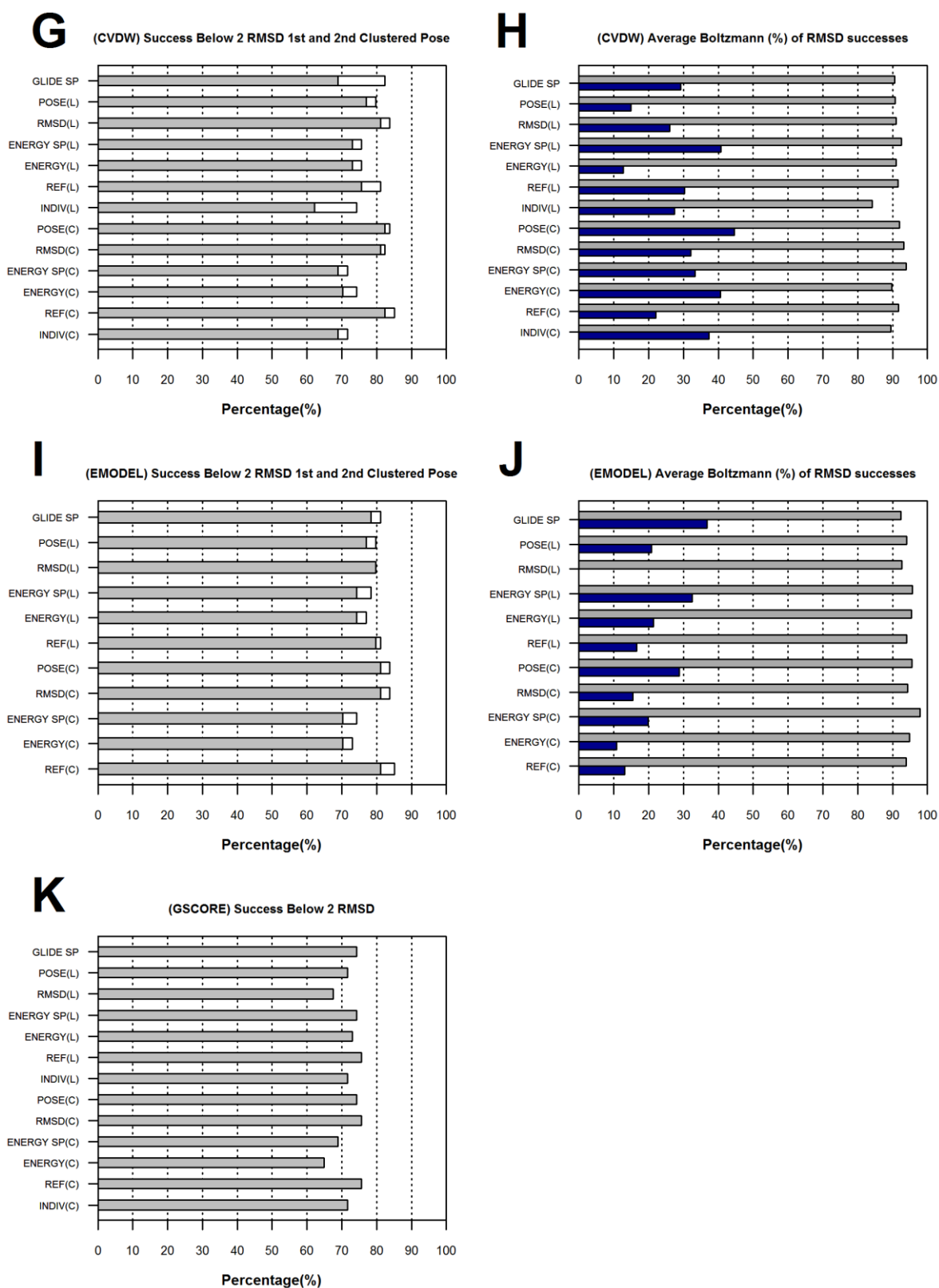
Figure 3.3. The influence of clustering on polarized docking performance. For G and H, the poses were generated according to GSCORE and ranked according to CVDW. For I and J, the poses were generated according to GSCORE and ranked according to EMODEL. K is given for reference and is equivalent to the orange bars in Fig 3.1B. In G and I, the percentage of cases that the top ranked pose

is below the 2.0 Å threshold is denoted by grey bars; the increase in the percentage of cases below the 2.0 Å threshold obtained by additionally considering the top ranked pose of a second (low-lying) cluster is denoted by white bars. For H and J, the mean sum of the Boltzmann probabilities of the poses in the first cluster for molecules with RMSD < 2.0 Å is denoted by grey bars; for the subset of molecules were there was no pose with an RMSD < 2.0 Å within the first cluster, the mean sum of the Boltzmann probabilities of the poses in the second cluster is denoted by the blue bars. GLIDE SP, indicates docking with no polarization. In subsequent bars, (L) indicates that only the ligand was polarized and (C) indicates that the ligand and protein were polarized. POSE indicates that the ligand was polarized using the geometry of the top ranked pose docked using glide, RMSD indicates that the ligand was polarized using the geometry of the pose closest to the experimental result, ENERGY SP indicates that the ligand was polarized using the geometry of the top ranked pose scored according to the single point QM/MM calculations, ENERGY indicates that the ligand was polarized using the geometry of the top ranked pose scored according to the QM/MM optimization calculations and REF indicates that the polarized charges for were determined at the experimental geometry. INDIV indicates that each of the initial 15 poses was polarized individually.

3.3.4   Cluster analysis.

In figures 3.1B, 3.1D and 3.1F we considered the proportion of molecules where the RMSD is below the four given thresholds, namely 0.5 Å (green), 1 Å (yellow), 2 Å (orange) and 4 Å(dark red) respectively. In some cases while the lowest energy pose may have a rather high RMSD, there may still be a pose with a low RMSD that is not too high in energy above the lowest energy pose; we consider this scenario in Figure 3.3 for the popular 2.0 Å threshold (Warren et al., 2006). As discussed in 3.2.10.2, the poses have been clustered so that each cluster has an RMSD of ≤ 1 Å and an energy within 1.1 kcal mol^{-1} of the lowest energy member of that pose cluster.   We then set the criteria that the lowest energy member of the 2nd pose cluster should also be considered for analysis at an RMSD of ≤ 2 Å, in the cases where the lowest energy member of the first pose cluster was docked unsuccessfully. This clustered RMSD analysis assumes that the lowest member of the 1st and 2nd pose clusters is representative of all members of that cluster and so the other members are not therefore considered. We evaluated the improvement in the docking results when both clusters were considered, as shown in Figures 3.3G and 3.3I, where the white bars indicate the higher percentage that comes by including the second cluster.

To ensure that poses in the second cluster are accessible, only poses that are members of a cluster with a Boltzmann probability (p_i) greater than 10% have been considered. The extent of this Boltzmann probability has been explored in Figures 3.3H and 3.3J and in section 3.3.4.4.1 (blue and grey bars). Each molecules Boltzmann probability was summed for all members of each of clusters. The mean (μ) Boltzmann probabilities for specific clusters was reported for two subsets of molecules from the above clustered RMSD analysis (that used lowest energy / top ranked poses that was representative of all members of that cluster). The first subset included the molecules with 1st clusters with the lowest energy/ top ranked pose was within RMSD of ≤ 2 Å. (grey bars). The second subset included molecules from the 2nd clusters with the lowest energy/ top ranked pose was within RMSD of ≤ 2 Å, and the best overall lowest energy/ top ranked pose from 1st cluster was not (blue bars). These are different subset of molecules so the percentages should not add up to 100%.

This probability evaluation explored high values which are indicative of alternative poses of similar energy, a situation that may arise, for example, when approximately symmetric molecules flip their binding mode.

3.3.4.1 Clustered CVDW Rescoring: Benchmark results

GSCORE was not used for clustering in Figure 3.3, mainly because of the heavily scaled nature of the force-field terms and also because of its default usage within Glide, but GSCORE results are shown at the RMSD of ≤ 2 Å threshold for reference in Figure 3.3K (identical to the orange bars in Fig 3.1B). As in Figure 3.1D, (orange, 1st bar denoted GLIDE SP), the results in Figure 3.3G (1st bar – grey, denoted GLIDE SP), are for the top ranked pose (docked according to GSCORE but with all the poses rescored according to CVDW).; Figure 3.3G (and Figure 3.1D) shows that the top ranked pose has an RMSD below 2 Å in 68% of cases. When we also consider the lowest energy member of the 2nd cluster at an RMSD of ≤ 2 Å (for the cases where first cluster did not yield an RMSD < 2 Å) the success rate rises to 83% (1st bar – white, denoted GLIDE SP). These two sets of results show that

there is considerable improvement of ~15% when clustering of the CVDW scored poses occurs. The energetic reason for considering such low lying clusters is (a) that the energy function is not sufficiently accurate to determine that the 1st cluster is indeed the lowest energy cluster and (b) low lying poses would normally be populated.

3.3.4.2 Clustered CVDW Rescoring: Ligand and Ligand and protein polarization

Considering the rest of the results in Figure 3.3G, the figure shows, when the *ligand* partial charges are explicitly polarized using (i) the geometries of the 1st Ranked GSCORE (re-scored by CVDW) pose (2nd bar, denoted POSE(L)), (ii) the lowest Energy QM/MM single-point pose (4th bar, denoted ENERGY SP(L)), (iii) the lowest Energy QM/MM optimized pose (5th bar, denoted ENERGY(SP)), then flexible ligand re-docking results in the top ranked pose having an RMSD of ≤ 2 Å 77%, 73% and also 73 % of the time respectively (grey bars). When we also consider the lowest energy member of the 2nd cluster at an RMSD of ≤ 2 Å (in the cases where the lowest energy member of the first cluster had an RMSD > 2, Å white bars), the success rate rises to 80%; 76% and also 76% respectively.

When the *ligand and protein* partial charges are explicitly polarized in the same fashion using (i) the geometries of the 1st Ranked GSCORE pose (8th bar, denoted POSE(C)), (ii) the lowest Energy QM/MM single-point pose (10th bar, denoted ENERGY SP(C)), (iii) the lowest Energy QM/MM optimized pose (11th bar, denoted ENERGY(C)), then flexible ligand re-docking results in the top ranked pose having an RMSD of ≤ 2 Å 83%, 69% and 71% of the time respectively (grey bars). When we also consider the lowest energy member of the 2nd cluster at an RMSD of ≤ 2 Å (in the cases where the lowest energy member of the first cluster is unsuccessful (white bars)) the success rate rises to 84%, 73% and 75% respectively.

For the controls where we have: (i) polarized the *ligand* in the pose with the lowest RMSD to the experimental pose and (ii) polarized the experimental pose. When the *ligand* partial charges are explicitly polarized then re-docked (flexible ligand) we find the top ranked pose having an RMSD of ≤ 2 Å (i) 82% (bar 3 – grey, denoted RMSD (L)) and (ii) 76% (bar 6

– grey, denoted REF(L)) of the time respectively. When we also consider the lowest energy member of the 2nd cluster at an RMSD of ≤ 2 Å, the success rate rises to (i) 84% (bar 3 - white) and (ii) 81% (bar 6 - white).

When the *ligand and protein* partial charges are explicitly polarized instead, this changes to (i) 82% (bar 9 – grey, denoted RMSD(C)) and (ii) 83% (bar 12 –grey, denoted REF(C)) respectively. When we also consider the lowest energy member of the 2nd pose cluster at an RMSD of ≤ 2 Å, the success rate rises to (i) 83% (bar 9 – white) and (ii) 85% (bar 12 - white). Explicit polarization of the ligand partial charges (7th bar, denoted INDIV (L)) and the ligand and protein partial charges (13th bar, denoted INDIV(C)) does not convey any advantage, as discussed above.

3.3.4.3 Clustered CVDW Rescoring: Summary remarks

The results here indicate that at the RMSD of ≤ 2 Å threshold, there seems to be a cap on docking success at ~85%, above which polarization or clustering and inclusion of the lowest energy member of the 2nd cluster cannot offer further improvements. We conclude that there are likely to be other issues that polarization alone cannot address in these cases, the most likely being protein flexibility and hydration. While the success rate of Glide SP at ~68% for the RMSD of ≤ 2 Å threshold, scored by CVDW, offers improvement over GSCORE. It is also significant that clustering and inclusion of the lowest energy member of the 2nd cluster improves the docking success to 83%, which is approximately the same success for including ligand and protein explicit polarization (Figure 3.3G, 3rd bar denoted RMSD(L), 7th bar denote POSE(C), 8th bar denoted RMSD(C) and 11th bar denoted REF(C)).

The margins for clustering improvement depend on there being a pose with a low RMSD that is not too high in energy above the lowest energy pose; such a scenario is quite common, e.g. for approximately symmetric molecules that have two similar binding modes separated by a rotation of 180°. These margins decrease when explicit polarization is included. This is

an indication that when polarization can help differentiate the correct pose from an energetically close geometric decoy.

3.3.4.4 Clustered EMODEL Rescoring: Summary Remarks

3.3.4.4.1 Analysis of the extent of the energetic viability of using a 2nd pose cluster.

Figure 3.3I shows that rescoring with EMODEL yields very good results as 78% of the poses have an RMSD ≤ 2 Å (1st grey bar, denoted GLIDE SP). Given these excellent results, there is little scope for improvement, either by clustering and considering the second pose, or by inclusion of polarization. which rises to 81% when we consider the second cluster. Indeed, examination of Figure 3.3I shows that clustering can improve the results by about 3% (e.g. 11th bar, denoted REF(C)) and that polarization can improve the results by about 3% (e.g. 11th bar), up to 85% (e.g. 11th white bar, denoted ref(C)).

As discussed in section 3.3.2.1, the poses have been clustered so that the each cluster has an RMSD of ≤ 1 Å and energy within 1.1 kcal mol^{-1} of the lowest energy member of that pose cluster. We also determined that to consider the lowest energy member of the 2nd cluster in the clustered docking success results at an RMSD of ≤ 2 Å, there was a need to ensure a significant probability of the pose adopting the 2nd cluster minima as opposed to the first cluster minima. Consequently, only poses that were members of a cluster with a sum greater than 10% Boltzmann probability were included in the results.

3.3.4.5 Clustered CVDW Rescoring: Summary

Figure 3.3H shows the percentage Boltzmann probability of the sum of the poses within the first cluster minima (all 74 molecules), and the second cluster minima (the subset of molecules with no pose < 2.0 Å in the first cluster). These Boltzmann probabilities have been averaged over the total number of molecules docked successfully with an RMSD of ≤ 2 Å, thus corresponding with the results in Figure 3.3G. This gives the average percentage

Boltzmann probability for a pose adopting the first minima (grey bars, associated with Figure 3.3G grey bars), and subsequently the second minima (dark blue bars, associated with Figure 3.3G white bars).   Consequently, for example for Glide SP Figure 3.3G (grey bar 1), there was a 68% docking success at an RMSD of ≤ 2 Å. The corresponding average of the sums of the Boltzmann probabilities for first cluster members within this minima across these successful molecules is shown in Figure 3.3H (grey bar 1) to be 91%. There is only once an average Boltzmann probability less than 90% at 84%, which applies to ~15 individually *ligand* polarized geometries (grey bar 7, denoted INDIV(L)). The highest average Boltzmann probability is 94% is for *ligand and protein* polarization of the lowest Energy QM/MM single-point pose (grey bar 10, denoted ENERGY SP(C)). These are sensible numbers considering the generous nature of the clustering criteria as seen again above in 3.3.4.4 and that they also refer to the lowest energy members of the first clusters, where these minima are always most likely to be populated. We next consider the extent to which the 2nd cluster could be also be populated.

We have seen for Glide SP Figure 3.3G (white bar 1, denoted GLIDE SP) that there was a lowest energy pose member from the 2nd pose cluster that was successful at the RMSD of ≤ 2 Å threshold, when the lowest energy member of the 1st pose cluster was not. When his was included, we saw an increase in success from 68% to 83%. This ~15% improvement in docking successes equates to ~11 molecules.

The corresponding average of the sums of the Boltzmann probability for second pose cluster members within this 2nd minima across these 11 successful molecules is shown in Figure 3.3H (dark blue bar 1, denoted GLIDE SP) to be ~29%. This is significantly high, when we consider the generous clustering criteria, the number of molecules involved and that the theoretical range based on our criteria gives an average Boltzmann probability between ~10% and ~49% for each molecule. So when we look at the next largest clustered docking success improvement in Figure 3.3G considering the lowest energy pose member from the 2nd cluster by our criteria, there was an ~11% improvement (~8 molecules). This was from explicit polarization of the ligand partial charges at each of the ~15 individual ligand

geometries of the pose population then re-docking using the 'Score in Place', where the

lowest energy pose by CVDW was originally 63% (bar 7 – grey, denoted INDIV(L)) and rose

to 74%(bar 7 – white). The corresponding average of the sums of the Boltzmann probability

for second cluster members within this 2nd minima across these 8 successful molecules is

shown in Figure 3.3H (dark blue bar 7, denoted INDIV(L)) to be 27%. This is very similar to

the Glide SP results shown above. It is therefore reasonable to say in both instances with an

average ~28% that with a middle of the road result when considering theoretical range of

~10% to ~49% that the 2nd pose cluster minima is also likely to be populated, with the rest of

the results in Figures 3.3H supporting this when clustering is done by the criteria in 3.2.10.2.

The most likely scenario for involvement of a low-lying second cluster is when the ligand is

pseudo-symmetrical and can bind in alternative modes with similar energies.

3.3.4.6 Clustered EMODEL Rescoring: Summary

Figure 3.3I shows that EMODEL again gives very similar results to scoring with CVDW, and

with the average Boltzmann probabilities for the 1st pose cluster minima understandably

slightly higher as there are smaller improvements from considering the 2nd cluster minima

compared with CVDW (Figure 3.3G). However, there is still a reasonable probability (> 10%)

that the 2nd pose cluster minima is also likely to be populated, as shown in Figure 3.3J.
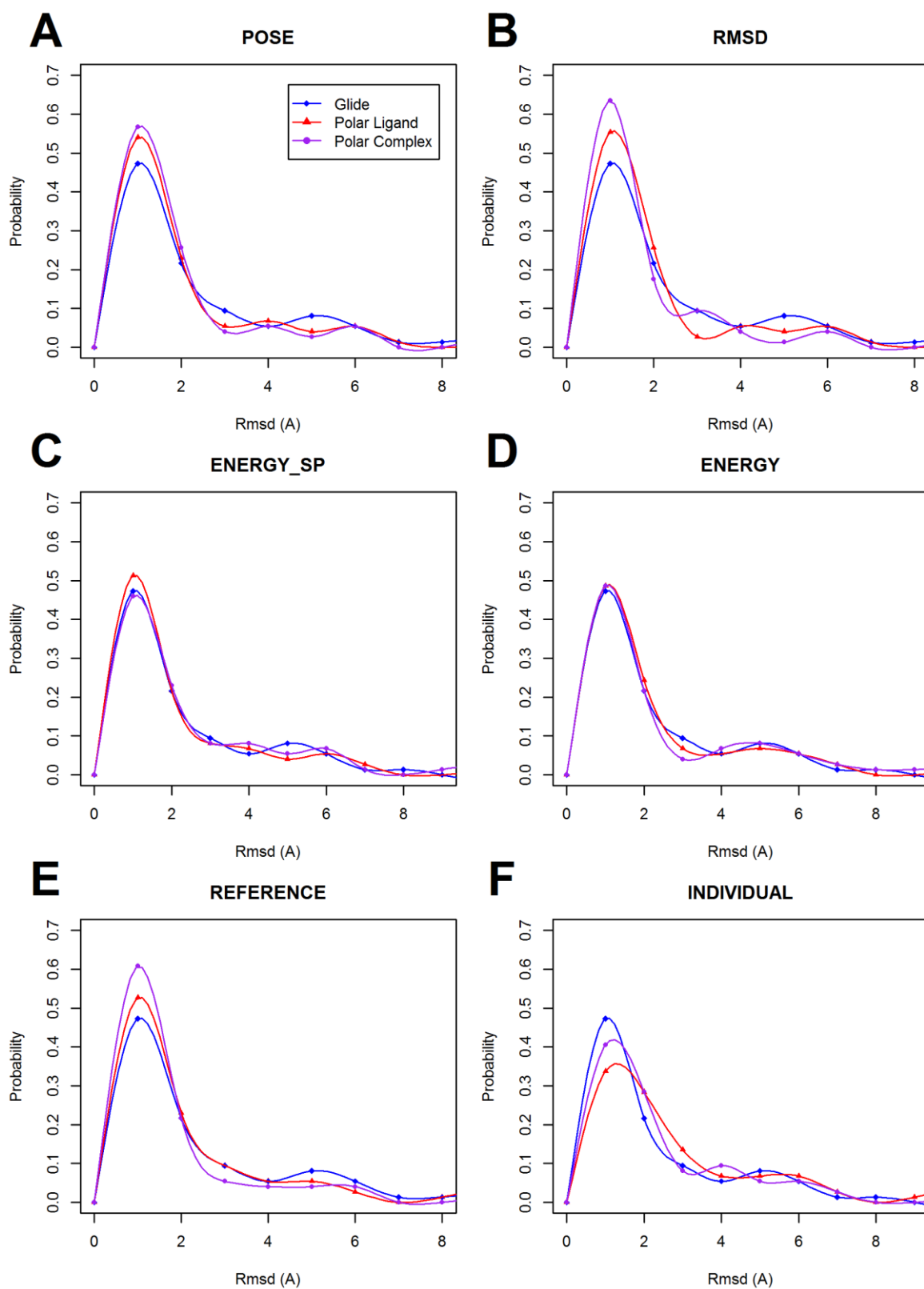
Figure 3.4 The probability that a pose is within a discrete 1 angstrom class bin, for a given value of RMSD. The probability was calculated using variables for RMSD from the top-ranked CVDW poses.

3.3.5 Probability plots

Given the improvements in the polarization results for CVDW compared to GSCORE, we have determined the probabilities of achieving an accuracy at a given RMSD class bin, as shown in Figure 3.4. This provides a different graphical representation of the probability data, other than a stacked percentage bar plot in Figure 3.1D. The results and cumulative probabilities ascertained from the area under the curve, otherwise remains the same. Figure 3.4E shows the improvement in docking results through polarization of the ligands at the control (experimental geometry, denoted REFERENCE). The purple line shows that the polarized ligand and protein is likely to have a higher probability of having a low RMSD of ~1 Å than the regular Glide results as it is higher than the blue line and more to the left. Similarly, polarization of ligand and protein is less likely to result in a large error of ~4 Å or more as the purple line is below the blue line at this point. Similar results are shown in Figure 3.4B for polarization according to the geometry of the pose with the lowest RMSD. Polarization of the ligand alone (red line) gives rise to intermediate results in Figure 3.4B and 3.4E. These two plots perhaps provide the strongest indication that polarization improves docking. Indeed, for example, Figure 3.4B shows the increased probability peaks at less or equal to one angstrom RMSD, which of course match the results of 3.3.5.1.1., where we see the incremental increase of probability as a decimal instead of a percentage from 0.473 for benchmark result, to 0.554 an ~8% improvement for ligand polarization, to 0.635 for ligand and protein polarization ~16% improvement for ligand and protein polarization respectively. Thus, the purple line (and the red line) show that polarization improves both the number and the quality of the good poses.

Figure 3.4 A, C and D show the effect of polarization when the correct answer is not known. The best results of these three are shown in Figure 3.4A, where the ligand and protein are polarized according to the best Glide pose. Again, the purple peak is higher than the blue peak at low RMSD class bins of 1 and 2 angstroms, and the higher RMSD class bins the remaining curve is within that of the blue line. In these results, ligand polarization

again gives rise to intermediate results when compared with Glide SP and ligand and protein polarization especially for the low RMSD class bins of 1 and 2 angstroms.

In principle quantum mechanics offers the promise of improved results. When we look at the 1 angstrom class bin peaks in Figure 3.4C, they show that polarizing the ligand at the best QM/MM single point result (red line), shows an improvement over the blue line CVDW benchmark. This improvement is lost in polarizing it at the QM/MM optimized result in Figure 3.4D. One possible interpretation of this is that the QM/MM optimization moves the ligand away from the experimental position, especially for poses with a high RMSD. This is presumably because GSCORE, through which the poses of Fig 3.4C were determined (although not ranked finally), includes solvent effects (in a parameterized way) while the QM/MM calculations ignore solvent effects. An alternative explanation is the imbalances in the QM/MM force field at the interface region. Polarization offers no change in the quality of the results over CVDW in Fig 3.4C, but polarization of the ligand only does offer a small improvement. See Appendix C for 0.5 angstrom RMSD binned probability plots.

3.4 Discussion

We have shown that polarization of fragments can offer in the region of a 10-15% improvement in docking results, as judged by the percentage of poses within a rather tight threshold of 0.5 or 1.0 Å. Clearly, such an improvement could make a significant difference to a fragment-based drug design program. These results are most apparent when the correct pose is known a priori, as under these circumstances the ligand and the protein can be polarized correctly, and polarizing the ligand and the protein gives better results than just polarizing the ligand. It does not seem to matter whether the 'correct pose' is the experimentally determined one or the docked pose with the lowest RMSD to this. Analysis of the results shows that the improvement is generally small and incremental, but the docking of some ligands did improve considerably (e.g., 2rdr; 2brt; 1s5n; 3ime (B); 1wcc; 1fv9), albeit the improvement was most marked when scored by CVDW scoring only.

The improvement is more apparent for the CVDW results as these are based on a molecular mechanics force field and so the electrostatics (and hence polarization effects) are not scaled down as much as they are in GSCORE; the improvement there continues up to the 2 Å threshold. The improvement is also apparent for EMODEL, which is a hybrid of GSCORE and CVDW, and so again the polarization effects are more apparent, as they are scaled down less than for GSCORE. Analysis of the results that are correct to within a given threshold suggests that polarization is more relevant to improving the good results (thresholds 0.5 – 1 Å, and possible 2 Å) than it is for improving the less good results (threshold 4 Å). In previous work on polarization in docking, the improvement was not so marked, and could only be taken into account by also checking the number of similar poses returned (Illingworth et al., 2008b). This is partly because electrostatics plays a more prominent role in docking than in Autodock.

There are many factors that contribute towards accurate docking. The Glide program is well optimized and so generally gives good results. This makes it difficult to improve these results. Nevertheless, there are several deficiencies in the Glide method that arise primarily from the requirement to be fast. One of these is the problem of the rigid enzyme/receptor. In some ways, the electronic flexibility introduced by polarization may alleviate this problem. Indeed, in fully flexible docking with a rigid charge distribution, rotamer and backbone changes make take place to minimize electrostatic clashes that should otherwise be reduced by polarization (Illingworth et al., 2008b). Hydration is another major problem in ligand design (Mason et al., 2013) that we will address in subsequent chapters. Crystal structure prediction competitions have shown that the semi-hard sphere repulsion of the standard 12-6 potentials is a weakness and that softer exponential repulsion can give improved results (Misquitta et al., 2008). In the next chapter we address this problem by using Orient for rigid body docking. This method also enables us to address the deficiencies of the fixed charge distribution.

3.5 References

BABER, J. C., THOMPSON, D. C., CROSS, J. B. & HUMBLET, C. 2009. GARD: a generally applicable replacement for RMSD. *Journal of chemical information and modeling,* 49**,** 1889-1900.

BANKS, J. L., BEARD, H. S., CAO, Y., CHO, A. E., DAMM, W., FARID, R., FELTS, A. K., HALGREN, T. A., MAINZ, D. T. & MAPLE, J. R. 2005. Integrated modeling program, applied chemical theory (IMPACT). *Journal of computational chemistry,* 26**,** 1752-1780.

BARILLARI, C., TAYLOR, J., VINER, R. & ESSEX, J. W. 2007. Classification of water molecules in protein binding sites. *Journal of the American Chemical Society,* 129**,** 2577-2587.

BERMAN, H., HENRICK, K. & NAKAMURA, H. 2003. Announcing the worldwide protein data bank. *Nature Structural & Molecular Biology,* 10**,** 980-980.

CHO, A. E., GUALLAR, V., BERNE, B. J. & FRIESNER, R. 2005. Importance of accurate charges in molecular docking: Quantum mechanical/molecular mechanical (QM/MM) approach. *Journal of Computational Chemistry,* 26**,** 915-931.

COLE, J. C., MURRAY, C. W., NISSINK, J. W. M., TAYLOR, R. D. & TAYLOR, R. 2005. Comparing protein–ligand docking programs is difficult. *Proteins: Structure, Function, and Bioinformatics,* 60**,** 325-332.

CONGREVE, M., CHESSARI, G., TISI, D. & WOODHEAD, A. J. 2008. Recent Developments in Fragment-Based Drug Discovery. *Journal of Medicinal Chemistry,* 51**,** 3661-3680.

ELDRIDGE, M. D., MURRAY, C. W., AUTON, T. R., PAOLINI, G. V. & MEE, R. P. 1997. Empirical scoring functions: I. The development of a fast empirical scoring function to estimate the binding affinity of ligands in receptor complexes. *Journal of computer-aided molecular design,* 11**,** 425-445.

EWING, T. J., MAKINO, S., SKILLMAN, A. G. & KUNTZ, I. D. 2001. DOCK 4.0: search strategies for automated molecular docking of flexible molecule databases. *Journal of computer-aided molecular design,* 15**,** 411-428.

FAVIA, A. D., BOTTEGONI, G., NOBELI, I., BISIGNANO, P. & CAVALLI, A. 2011. SERAPhiC: A benchmark for in silico fragment-based drug design. *Journal of chemical information and modeling,* 51**,** 2882-2896.

FERENCZY, G. G. & REYNOLDS, C. A. 2001. Modeling polarization through induced atomic charges. *The Journal of Physical Chemistry A,* 105**,** 11470-11479.

FRIESNER, R. A., BANKS, J. L., MURPHY, R. B., HALGREN, T. A., KLICIC, J. J., DANIEL, T., REPASKY, M. P., KNOLL, E. H., SHELLEY, M. & PERRY, J. K. 2004. Glide: a new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy. *Journal of medicinal chemistry,* 47**,** 1739-1749.

FRIESNER, R. A., MURPHY, R. B., REPASKY, M. P., FRYE, L. L., GREENWOOD, J. R., HALGREN, T. A., SANSCHAGRIN, P. C. & MAINZ, D. T. 2006. Extra precision glide: docking and scoring incorporating a model of hydrophobic enclosure for protein-ligand complexes. *Journal of medicinal chemistry,* 49**,** 6177-6196.

FRISCH, E., FRISCH, M. & TRUCKS, G. W. 2003. *Gaussian 03*, Gaussian.

FRISCH, M., TRUCKS, G., SCHLEGEL, H., SCUSERIA, G., ROBB, M., CHEESEMAN, J., MONTGOMERY JR, J., VREVEN, T., KUDIN, K. & BURANT, J. 2004. GAUSSIAN 03 program. *Gaussian Inc., Wallingford, CT.*

GOODING, S. R., WINN, P. J., MAURER, R. I., FERENCZY, G. G., MILLER, J. R., HARRIS, J. E., GRIFFITHS, D. V. & REYNOLDS, C. A. 2000. Fully polarizable QM/MM calculations: An application to the nonbonded iodine–oxygen interaction in dimethyl-2-iodobenzoylphosphonate. *Journal of Computational Chemistry,* 21**,** 478-482.

HALGREN, T. A., MURPHY, R. B., FRIESNER, R. A., BEARD, H. S., FRYE, L. L., POLLARD, W. T. & BANKS, J. L. 2004. Glide: a new approach for rapid, accurate docking and scoring. 2. Enrichment factors in database screening. *Journal of medicinal chemistry,* 47**,** 1750-1759.

HUNG, A. W., SILVESTRE, H. L., WEN, S., CIULLI, A., BLUNDELL, T. L. & ABELL, C. 2009. Application of fragment growing and fragment linking to the discovery of inhibitors of Mycobacterium tuberculosis pantothenate synthetase. *Angewandte Chemie,* 121**,** 8604-8608.

ILLINGWORTH, C. J., MORRIS, G. M., PARKES, K. E., SNELL, C. R. & REYNOLDS, C. A. 2008a. Assessing the role of polarization in docking. *The Journal of Physical Chemistry A,* 112**,** 12157-12163.

ILLINGWORTH, C. J. R., GOODING, S. R., WINN, P. J., JONES, G. A., FERENCZY, G. G. & REYNOLDS, C. A. 2006. Classical polarization in hybrid QM/MM methods. *J. Phys. Chem. A,* 110**,** 6487-6497.

ILLINGWORTH, C. J. R., MORRIS, G. M., PARKES, K. E. B., SNELL, C. R. & REYNOLDS, C. A. 2008b. Assessing the Role of Polarization in Docking. *The Journal of Physical Chemistry A,* 112**,** 12157-12163.

ILLINGWORTH, C. J. R., PARKES, K. E. B., SNELL, C. R., FERENCZY, G. R. G. & REYNOLDS, C. A. 2008c. Toward a Consistent Treatment of Polarization in Model QM/MM Calculations. *The Journal of Physical Chemistry A,* 112**,** 12151-12156.

JORGENSEN, W. L. 2007. Special issue on polarization. *J. Chem. Theory Comput,* 3**,** 1877.

JORGENSEN, W. L., MAXWELL, D. S. & TIRADO-RIVES, J. 1996. Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids. *Journal of the American Chemical Society,* 118**,** 11225-11236.

KAMINSKI, G. A., FRIESNER, R. A., TIRADO-RIVES, J. & JORGENSEN, W. L. 2001. Evaluation and Reparametrization of the OPLS-AA Force Field for Proteins via Comparison with Accurate Quantum Chemical Calculations on Peptides†. *The Journal of Physical Chemistry B,* 105**,** 6474-6487.

KAMINSKI, G. A., STERN, H. A., BERNE, B. J. & FRIESNER, R. A. 2003. Development of an Accurate and Robust Polarizable Molecular Mechanics Force Field from ab Initio Quantum Chemistry. *The Journal of Physical Chemistry A,* 108**,** 621-627.

KOLB, P. & IRWIN, J. J. 2009. Docking screens: right for the right reasons? *Current topics in medicinal chemistry,* 9**,** 755-770.

KONTOYIANNI, M., MCCLELLAN, L. M. & SOKOL, G. S. 2003. Evaluation of Docking Performance: Comparative Data on Docking Algorithms. *Journal of medicinal chemistry,* 47**,** 558-565.

KROEMER, R. T., VULPETTI, A., MCDONALD, J. J., ROHRER, D. C., TROSSET, J.-Y., GIORDANETTO, F., COTESTA, S., MCMARTIN, C., KIHLÉN, M. & STOUTEN, P. F. 2004. Assessment of docking poses: interactions-based accuracy classification (IBAC) versus crystal structure deviations. *Journal of chemical information and computer sciences,* 44**,** 871-881.

MASON, J., BORTOLATO, A., WEISS, D., DEFLORIAN, F., TEHAN, B. & MARSHALL, F. 2013. High end GPCR design: crafted ligand design and druggability analysis using protein structure, lipophilic hotspots and explicit water networks. *Silico Pharmacol,* 1**,** 23.

MISQUITTA, A. J., WELCH, G. W., STONE, A. J. & PRICE, S. L. 2008. A first principles prediction of the crystal structure of. *Chemical Physics Letters,* 456**,** 105-109.

MURPHY, R. B., BEACHY, M. D., FRIESNER, R. A. & RINGNALDA, M. N. 1995. Pseudospectral localized Mo/ller–Plesset methods: Theory and calculation of conformational energies. *The Journal of Chemical Physics,* 103**,** 1481-1490.

MURPHY, R. B., PHILIPP, D. M. & FRIESNER, R. A. 2000. A mixed quantum mechanics/molecular mechanics (QM/MM) method for large-scale modeling of chemistry in protein environments. *Journal of Computational Chemistry,* 21**,** 1442-1457.

OLSSON, M. H., SØNDERGAARD, C. R., ROSTKOWSKI, M. & JENSEN, J. H. 2011. PROPKA3: consistent treatment of internal and surface residues in empirical p K a predictions. *Journal of Chemical Theory and Computation,* 7**,** 525-537.

RIZZO, R. C. & JORGENSEN, W. L. 1999. OPLS all-atom model for amines: resolution of the amine hydration problem. *Journal of the American Chemical Society,* 121**,** 4827-4836.

SÁNDOR, M., KISS, R. & KESERŰ, G. R. M. 2010. Virtual fragment docking by Glide: A validation study on 190 protein– fragment complexes. *Journal of chemical information and modeling,* 50**,** 1165-1172.

STONE, A. J. 2005. Distributed Multipole Analysis of Gaussian wavefunctions GDMA version 2.2. 02.

WANG, J., WANG, W., KOLLMAN, P. A. & CASE, D. A. 2001. Antechamber: an accessory software package for molecular mechanical calculations. *J. Am. Chem. Soc,* 222**,** U403.

WARREN, G. L., ANDREWS, C. W., CAPELLI, A. M., CLARKE, B., LALONDE, J., LAMBERT, M. H., LINDVALL, M., NEVINS, N., SEMUS, S. F. & SENGER, S. 2006. A critical assessment of docking programs and scoring functions. *Journal of medicinal chemistry,* 49**,** 5912-5931.

WEININGER, D. 1988. SMILES, a chemical language and information system. 1. Introduction to methodology and encoding rules. *Journal of chemical information and computer sciences,* 28**,** 31-36.

WINN, P. J., FERENCZY, G. G. & REYNOLDS, C. A. 1999. Towards improved force fields: III. Polarization through modified atomic charges. *Journal of computational chemistry,* 20**,** 704-712.

YUSUF, D., DAVIS, A. M., KLEYWEGT, G. J. & SCHMITT, S. 2008. An alternative method for the evaluation of docking performance: RSR vs RMSD. *Journal of chemical information and modeling,* 48**,** 1411-1422.

4 Assessing the use of distributed multipoles and exponential repulsion in docking for FBDD

4.1 Introduction

In chapter 3 we saw that ligand docking using Glide can achieve reasonably high success rates of at best ~83%, as judged by the proportion of ligands where the RMSD of the top pose is < 2 Å. However, lack of polarization is not the only weakness in standard docking methods and so to increase docking success further we proposed that it is also important to address some of the other weaknesses. Solvation is clearly a major issue that is not considered fully in Glide and other similar docking programs, and we will address this to some degree in Chapter 5. However, here we propose to address two major weaknesses of traditional molecular mechanics (MM) force fields.

The first is the electrostatic element of the force field. Atomic charges are inherently isotropic and so can fail to describe the anisotropic nature of intermolecular interactions around an atom, e.g., as seen in halogen bonding (Metrangolo et al., 2005), where a linear interaction is seen between a carbon-halogen bond and an oxygen atom, and the anisotropic distribution of hydrogen bonds around a carbonyl group (Singh and Thornton, 1993). One way to go beyond the point charge distribution is to use quantum mechanics (QM) in QM/MM methods where the ligand is treated by QM and the enzyme target by MM (Lonsdale and Mulholland, 2014). This effectively improves the description of the ligand, but not the contact residues in the protein – unless key residues are also included in the QM region. Another way to go beyond the point charge approximation is to include a multipole series (i.e. charge, dipole, quadruple, hexadecapole...) on each atom, either through a distributed multipole analysis (Stone, 1981) or through effective multipoles (Ferenczy et al., 1997), for example as implemented in the Amoeba force field (Ponder and Case, 2003). The treatment of flexible molecules using a DMA is not well developed, except perhaps in the Amoeba force field, as implemented in Tinker (Ponder, 2004), but this is not a major problem for preliminary studies as a large proportion of fragments are rigid. Consequently, we propose to address the

docking of fragments using Orient, which was specifically developed for DMA calculations, since there is expertise in using Orient in the lab.

The second weakness is the r^{-12} repulsion term, which was introduced because of its ease of calculation, as it is trivially related to the r^{-6} dispersion term. This repulsion is rather hard and is not too dissimilar to a hard-sphere repulsion. An exponential repulsion term is more realistic (Buckingham and Corner, 1947, Cieplak et al., 2009), and can be readily incorporated into Orient. Consequently, we propose to use Orient to dock rigid ligands to their target and compare the results to Glide and QM/MM calculations. Where the ligand is not rigid, it will nevertheless be held in its experimental conformation.

4.2 Methods

4.2.1 Method Specific Technicalities

The protein structure PDB files were imported into Glide and written as single point truncated Newtonian compressed IMPACT files. The verbose pparam command was added to the IMPACT input file (next to species command), to create an IMPACT parameter file, that was parameterized as OPLS 2005 (Banks et al., 2005). To maintain consistency in atom numbering between the hydrogens and other atom types in the IMPACT parameter file and exported PDB co-ordinates, hydrogens and bond orders were left by prep-wizard to be assigned by IMPACT 5.8, see section 3.2.1 and 3.2.2, but waters were removed, (and disulphide bridges added) in Maestro 9.3. Then after IMPACT 5.8 was run, and file uncompressed, the maestro file was exported back into maestro 9.3 and exported again in PDB, and XYZ file formats for protein and ~15 ligand pose files from Glide SP 5.8 (Friesner et al., 2004) respectively.

Perl scripting was applied to acquire the shell of neighbouring residues with an atom within 4 Å of the ligand, then capping was applied at the C-terminus and N-terminus of the amino acid residues that were treated as continuous if their primary 'backbone' sequence was in within the RMSD range. This capping (replacing -NHR at the N-terminus, -COR at the

C-terminus and the side chain with H) was done by altering the Euclidean distances for the

three atoms attached to the tetrahedral backbone chiral (unless glycine) Cα atom beyond the

amide bonds to 1.09 angstroms e.g., hydrogen bond lengths, with cysteine bridge forming

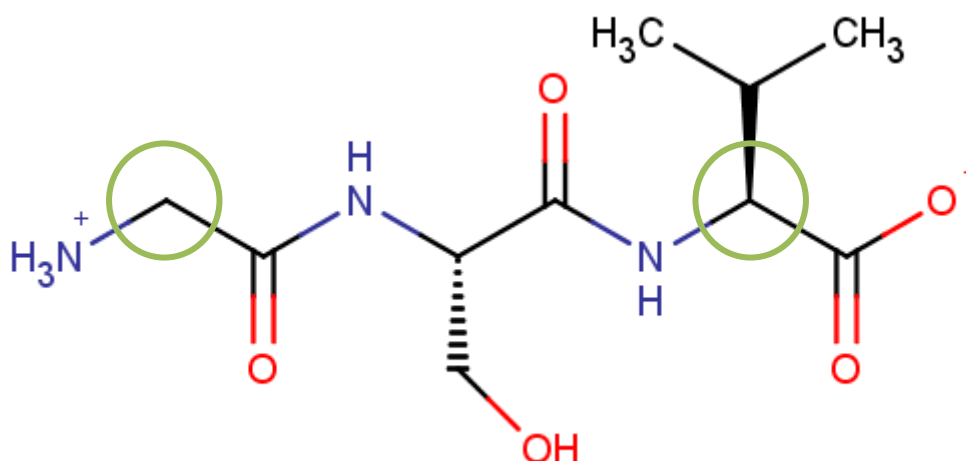atoms also being capped at the $C\beta$ atoms in the same fashion, see Figure 4.1.

Figure 4.1 An example of where protein caps were applied for ab initio calculations. Here, a Glycine-Serine-Valine tri-peptide was capped at Cα atoms, circled in green, to calculate the Serine residue in isolation.

This was achievable as the pdb atom types exported by glide, were standardized

(although the perl script required slightly different pdb atom names for replacing rules for

Glycine and Proline). Gaussian accepts input files that are in Cartesian XYZ file format with

formal charges and some input specifications, so once the capping of the pdb file was

achieved, it was then converted to xyz format, with the net charge being calculated from the

sum of the formal charge column conveniently provided by Maestro 9.3. That is unless ions

were involved where grouping was needed to allow for charge transfer and then the net

charges of the neighbourhood residues were also taken into account, which were grouped to

within 3 Å RMSD of the ion. This grouping provided neighbouring residues, allowing the ion

to realistically spread it charge. Once the net charge of the Gaussian file was accurate and

sufficient residues included, in either isolation or as grouped molecules accordingly, then the

Gaussian calculations were carried out using the Becke-3-Lee-Yang-Parr – B3LYP (Lee et

al., 1988, Becke, 1993) DFT method (Hohenberg and Kohn, 1964) and a Pople 6-31G* basis

set (Hehre et al., 1972), which was consistent with other parts of the thesis where the

QM/MM energy calculations prepared in the Maestro 9.3 using the QSITE module to call on

JAGUAR 7.9 also used these settings (see 3.2.9 for more detail). The Gaussian wave

functions were then used within the GDMA 2.2 program (Stone, 2005) to create a rank 4

atom-based distributed multipole series., Finally, for the neighbouring residues, the caps

were removed (including for cysteine bridges) and the point charges of the cap residues

belonging to the protein were added in equal proportion to the rest of the residue or group of

backbone attached residues; this was carried out through a perl script.

Meanwhile, the atom typing parameters in the IMPACT file were read, and Cartesian

co-ordinates from the pdb file were used to parameterize and convert to an ORIENT 4.6

(Stone et al., 2006) file template. ORIENT accepts user defined potentials, and was designed

to perform DMA minimizations; the Lennard Jones potential (Jones, 1924) was initially used,

then additionally the Williams 1999 Exp6 potential (Williams, 1999) was created using the

parameters from (Williams, 2001). Simple isotropic polarization without charge flow was

adding later to fit in with our polarization scheme using largely literature polarizabilities (Miller

and Savchik, 1979). However Orient is capable of handling tensor polarizabilities with charge

flow (Stone et al., 2006).

After GDMA 2.2 minimization, the residue groups were uncapped, with the point

charges spread from the caps to the remaining atoms in a 'quick fix' to avoid using

complicated Langragian multipliers. The now uncapped Rank 4 DMA (hexadecapole) atoms,

replaced the corresponding atoms in the ORIENT template file with ~matching Cartesian co-

ordinates. Finally, the ~15 Glide ligand pose XYZ files were separately run through

Gaussian 03 and GDMA 2.2 (no capping required), before being inserted into the template

file. Then the Orient program was run for a geometry optimization using eigenvector

following, with the final optimized XYZ co-ordinates being re-inserted into a template maestro file, for perl scripted RMSD geometry analysis. However, as the enzyme was moved in orient, a superimposition matrix using Jacobian rules and user defined symmetry awareness rules was incorporated (using in-house code). The results (energies, geometries) were extracted from the Orient output file for further analysis; the results were plotted in R.

4.3 Results and Discussion

4.3.1 Analysis of top ranked poses

4.3.1.1 Hybrid QM/MM and DMA Benchmarks

As discussed in 3.3.1.1, Glide SP flexible ligand docking is usually able to identify the experimental pose amongst the full set of returned poses, whether as the top ranked pose, or as a lower ranked pose. Additionally, as discussed in section 2.2.4, energy minimization methods are usually only able to explore local minima. To take advantage of the Glide's fuller coverage of the pose space, we have taken the full set of ~15 returned poses from Glide SP and implemented various QM/MM based energy minimizations to provide a benchmark dataset. A primary reason for this was to explore the efficacy of using a DMA and/or a softer exponential exchange repulsion term within high quality classical energy calculations. To make an effective quality comparison we have used high quality QM/MM benchmark results, namely results obtained from the QSITE module within the Jaguar QM and Impact MM environments (referred to here as QM/MM). To establish a control in Figure 4.3A (first bar and second bars), we consider the lowest energy pose according to QM/MM taken at (i) the single point geometry (ii) the QM/MM optimized geometry; the figure shows that these poses are also the ones with the smallest RMSD to the experimental pose in (i) 44% (grey bar 1, denoted Jaguar (S)) and (ii) 38% (grey bar 2, denoted Jaguar(O)) of the 74 cases. Using the same clustering criteria as in 3.2.10.2, when we set the criteria instead that the pose with the lowest RMSD to experiment should also be a member of the first pose

cluster, then the success rate increased slightly to (i) 45% (white bar 1, denoted Jaguar (S)) and (ii) 41% (white bar 2, denoted Jaguar (O)) respectively. These four QM/MM energy results can be considered as benchmarks by which to judge the implementation of DMA and/or a softer exponential exchange repulsion term. As an alternative benchmark, the Glide SP results (38% for the lowest energy pose) are given as the bar 10 at the end of Figure 4.3A (c.f. Figure 3.1A).

However, before comparing these results to other methods, it is instructive to compare the single point QM/MM results to the optimized results. All figures 4.2 to 4.6 in these results show, that the single point results are superior to the optimized results; this is particular evident from the probabilities as demonstrated below in Figure 4.2, with further examples in Figures 4.5 and 4.6. This observation is somewhat counter-intuitive, but presumably arises because of the lack of solvation treatment in these simple QM/MM calculations, meaning that the optimization may in some cases move the pose further away from experiment. On the positive side, single point calculations are easier and faster to carry out, and so this means that it is more feasible to apply QM/MM calculations to fragment screening.
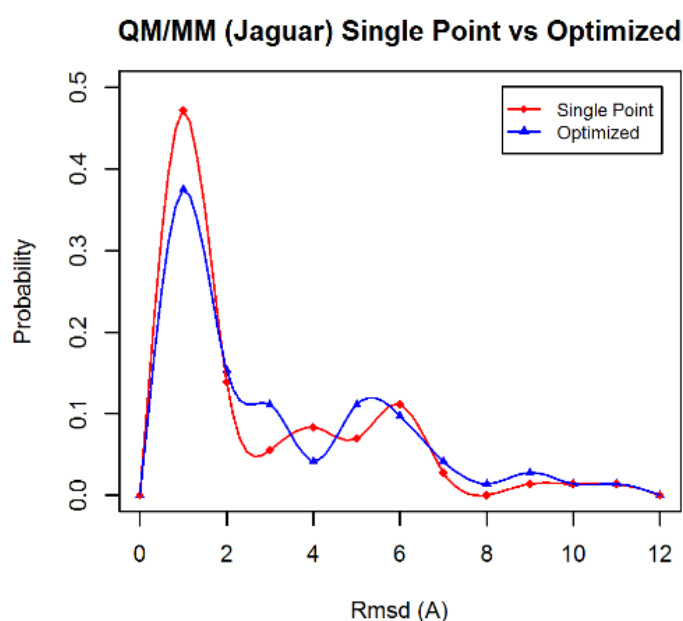

Figure 4.2 RMSD across the 74 molecules top-ranked poses as a binned (1 angstrom RMSD) probability plot for the QM/MM calculations.

4.3.1.2 Single Point Calculations

The Orient based single point energy calculations under consideration include the use of: (i) DMA with a Lennard Jones potential; (ii) DMA with an Exponential 6 Williams potential; (iii) DMA with a Lennard Jones potential and explicit ligand polarization (iv) DMA with an Exponential 6 Williams potential and explicit ligand polarization. Figure 4.3A shows that the lowest single point energy poses are also the ones with the smallest RMSD to experiment in (i) 36% (grey bar 3, denoted LJ (S)), (ii) 34% (grey bar 7, denoted EXP6 (S)), (iii) 32% (grey bar 5, denoted LJ (LS)) and (iv) 34 % (grey bar 9, denoted EXP6 (LS)) of the cases respectively. When we set the criteria instead that the pose with the lowest RMSD to experiment should also be a member of the first pose cluster, then success rate increases by 2-8% to (i) 41% (white bar 3, denoted LJ (S)), (ii) 42% (white bar 7, denoted EXP6 (S)), (iii) 34% (white bar 5, denoted LJ (LS)) and (iv) 39 % (white bar 9, denoted EXP6 (LS)) respectively.

When we assess whether the lowest single point energy pose also has the lowest RMSD to experiment, the benchmark result from using Jaguar to calculate the QM region single point energy has the highest success rate at 44% (Figure 4.3A, bar 1 and bar 2). This is about an 8% improvement on the DMA with a Lennard Jones potential. This is also true when we consider the members of the lowest pose cluster by the same criteria where the success rate is 45% (Jaguar, single point calculations, Fig 4.3A, 1st bar). Thus, it appears that DMA offers no improvement over QM/MM here, but is able to give comparable results especially when using clustering with either potential in cases (i) and (ii) (Figure 4.3A, bar 3 denoted LJ (S) and bar 7, denoted EXP6 (S)) where there is 41-42% success rate. Additionally, the explicit *ligand* polarization also offers no improvement to the DMA based Orient calculations but rather makes the results worse, possibly indicating that polarization is not implemented in a balanced way.

4.3.1.3 Optimized Energy Calculations

The Orient based optimized energy calculations as above in 4.3.1.2 include the use of the same four energy schemes as in the previous section ((i) DMA with a Lennard Jones potential; (ii) DMA with an Exponential 6 Williams potential; (iii) DMA with a Lennard Jones potential (the best DMA result) and explicit ligand polarization (iv) DMA with an Exponential 6 Williams potential and explicit ligand polarization). Figure 4.3A again shows that the lowest optimized energy poses are also the ones with the smallest RMSD to experiment in (i) 35% (grey bar 4, denoted LS (O)), (ii) 33% (grey bar 8, denoted EXP6 (O)), (iii) 29% (grey bar 6, denoted LJ (LO)) and (iv) 34 % (grey bar 10, denoted EXP6 (LO)) of the cases respectively. When we set the criteria that the pose with the lowest RMSD to experiment should also be a member of the first pose cluster, then success rate increases to (i) 38% (white bar 4, denoted LS (O)), (ii) 45% (white bar 8, denoted EXP6 (O)), (iii) 31% (white bar 6, denoted LJ(LO)) and (iv) 43% (white bar 10, denoted EXP6 (LO)) respectively.

      The general conclusions to emerge from Figure 4.3A are that optimization of the poses, either by QM/MM or by DMA confers no advantage. Moreover, inclusion of ligand polarization within the DMA treatment by Orient also conveys no advantage. As discussed in Chapter 3, this may be due to the lack of treatment of solvent in these approaches,

4.3.1.4 Assessment of DMA minimized geometries

After the results of Figure 4.3A showed largely no improvement over QM/MM or results in chapter 3, we now investigate whether the Glide geometries that had undergone energy minimization with DMA, offer any improvement over the standard Glide geometries. This idea is applied to the Glide scoring methods namely GSCORE and CVDW, using the 'score in place' option. To establish a performance benchmark in Figure 4.3C, we consider the 1st Ranked or lowest energy Glide pose (docked according to GSCORE) using flexible ligand docking and scored by (i) GSCORE and (ii) CVDW. The figure shows that these poses are also the ones with the smallest RMSD to the experimental pose in (i) 37% (grey bar 1,

denoted Glide(GS)) and (ii) 39% (grey bar 4, denoted Glide(CVDW)of the 74 cases. Using the same clustering criteria as in 3.2.10.2, when we set the criteria that the pose with the lowest RMSD to experiment should also be a member of the first pose cluster, for CVDW only then the success rate increase to 46% (white bar 4, denoted Glide(CVDW)). So for GSCORE, the benchmark success was 37% and for CVDW 39%, then 46% when the members of the first pose cluster were included, using the 'score in place' option. The Orient-based optimized geometries (as in 4.3.1.3) from (i) DMA with a Lennard Jones potential; and (ii) DMA with an Exponential 6 Williams potential were scored using GSCORE; both gave ~39% success rate in identifying the top ranking pose as also the lowest RMSD to experiment in the 74 cases. When CVDW was used instead for scoring, the success rates were (i) 46% (grey bar 6, denoted LJ (CVDW)) and (ii) 43% (grey bar 6, denoted EXP6(CVDW)) respectively. The addition of considering the members of the first pose cluster as also the lowest RMSD to experimental pose saw the success rates rise to (i) 51% and (ii) 52 % for the 74 cases respectively. So, there was a ~2% improvement using GSCORE for both DMA-minimized geometries, and up to ~7% (averaging ~5.5%) improvement using CVDW scoring for the DMA-minimized geometry (DMA and a Lennard Jones potential). Also, up to ~6% improvement over CVDW scoring (averaging again ~5.5%) was seen when considering members of the first pose cluster as also the lowest RMSD to experimental pose, indicating an overall Glide scoring preference for the DMA-minimized Geometries.

It is interesting therefore that although Glide is an excellent method, improved results can be obtained by optimizing the geometry and re-scoring with Glide – this offers the potential of determining the geometry using the more sophisticated DMA while retaining the advantages of the scoring within Glide that is able to include hydration effects. The fact that the impact (OPLS) Leonard Jones potentials are better than the exponential repulsion is contrary to accepted knowledge, but this probably arises because the OPLS force field is well parameterized while the exponential repulsion used is possibly not so well parameterized, and not so well balanced within the force field as a whole.
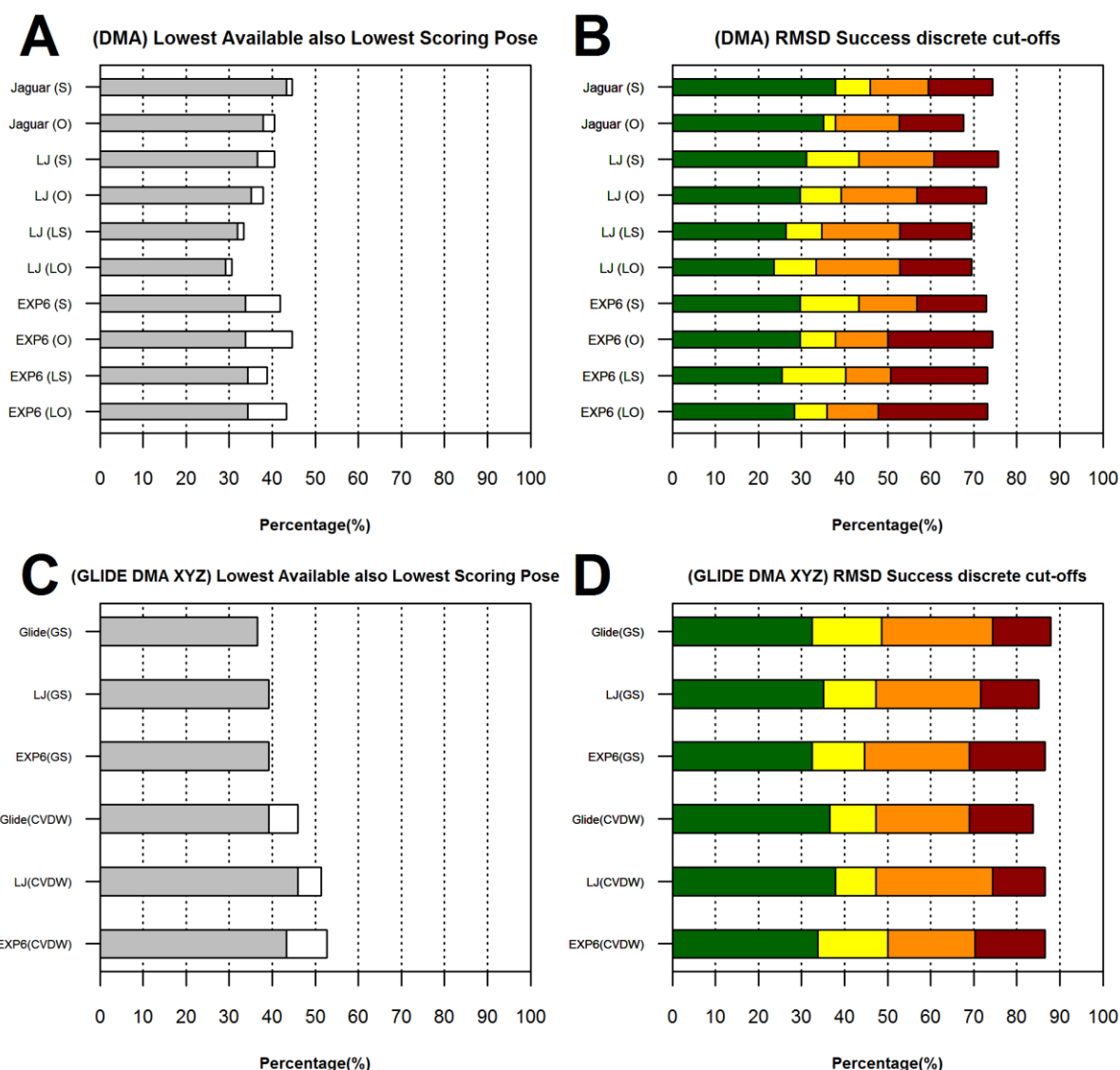
Figure 4.3 Assessment of Docking success using DMA. In A and B, Jaguar indicates - QM/MM energy calculations are done in Jaguar. DMA methods are indicated by (i) LJ and (ii) EXP6, where (i) and (ii) refer the specific type of exchange repulsion used for modelling van der Waals interactions. Here, Single point calculation (S) and Optimized calculation (O) denote the energy used for scoring, and (L) indicates that isotropic ligand polarization has been included using Orient. In C and D, Glide refers to Glide SP (see Chapter 3), LJ and EXP6 again refer to treatment of van der Waals forces in DMA methods, (GS) indicates GlideScore scoring was used, and (CVDW) indicates that the Glide MM Force-field only was used for 'score in place' of the optimized Orient (DMA) geometries. In plots A and C, the bars show the percentage (over 74 molecules) where the highest ranked/lowest scoring Pose (Grey), Pose cluster (White**) is also the lowest available pose. In B and D, the bars show the percentage (over 74* molecules) where highest ranked/lowest scoring pose is successfully docked at given discrete cut-offs: <0.5 RMSD (Green); <1 RMSD (Yellow); <2 RMSD (Orange); <4 RMSD (Red). * Isotropic ligand polarization was attempted in Orient for only 68 and 72 molecules for Exp6 and LJ respectively Exp6 due to convergence/ implementation issues.

4.3.2   RMSD analysis

In sections 4.3.1 and Figure 4.3A and 4.3C we considered the proportion of cases where the top ranked pose was also the one with the lowest RMSD to the experimental structure. However, in some of these cases the lowest RMSD, as in 3.3.3, might still be rather high. In Figures 4.3B and 4.3D we consider the proportion of molecules where the RMSD is below 0.5 Å (green); RMSD of ≤ 1 Å (yellow); RMSD of ≤ 2 Å (orange); RMSD of ≤ 4 Å(dark red) respectively.

4.3.3   QM/MM Benchmark Results

In Figure 4.3B (bar 1, denoted Jaguar (S)) we consider the lowest single point energy QM/MM pose; the figure shows that this pose has a distance measured by RMSD from the experimental pose of: (i) ≤ 0.5 Å; (ii) ≤ 1 Å; (iii) ≤ 2 Å; (iv) ≤ 4 Å in (i) 38%, (ii) 46%, (iii) 59% and (iv) 74% of the 74 cases respectively. When instead we consider the lowest optimized energy QM/MM pose, Figure 4.3B (bar 2, denoted Jaguar (O)) shows that this pose has an RMSD from experiment of: (i) ≤ 0.5 Å; (ii) ≤ 1 Å; (iii) ≤ 2 Å; (iv) ≤ 4 Å in (i) 35%, (ii) 38%, (iii) 53% and (iv) 67% of the 74 cases respectively. It is clear that QM/MM optimization reduces the proportion of successful docking by 3-8%, probably because Glide has a more appropriate approximation to hydration and entropy than QM/MM as implemented in QSITE and the optimization gives worse results because it is carried out in the absence of a hydration treatment. These benchmark results can be used as a guideline to see if energy calculations that include DMA can improve the docking results.

4.3.3.1   RMSD of ≤ 0.5 Å from experimental geometry (green bars)

The Orient based single point energy calculations include the same 4 energy functions as in section 4.3.1.2, (i) DMA with a Lennard Jones potential (bar 3, denoted LS (S)); (ii) DMA with an Exponential 6 Williams potential (bar 7, denoted EXP6 (S)); (iii) DMA with a Lennard Jones potential and explicit *ligand* polarization (bar 5, denoted LJ (LS)); and (iv) DMA with an Exponential 6 Williams potential and explicit *ligand* polarization (bar 9, denoted EXP6 (LS)).

Figure 4.3B shows that the lowest single point energy poses have an RMSD of ≤ 0.5 Å in (i) 31%, (ii) 27%, (iii) 30% and (iv) 26% of the cases respectively.

The Orient based optimized energy calculations again include the use of: (i) DMA with a Lennard Jones potential (bar 4, denoted LJ (O)); (ii) DMA with an Exponential 6 Williams potential (bar 8, denoted EXP6 (O)); (iii) DMA with a Lennard Jones potential and explicit *ligand* polarization (bar 6, denoted LS (LO)); and (iv) DMA with an Exponential 6 Williams potential and explicit ligand polarization (bar 10, denoted EXP6 (LO)). Figure 4.3B shows that the lowest optimized energy poses have an RMSD of ≤ 0.5 Å in (i) 30%, (ii) also 30%, (iii) 24% and (iv) 28% of the cases respectively.

When compared with the QM/MM benchmarks with an RMSD threshold of ≤ 0.5 Å, the (i) lowest single point energy and the (ii) lowest optimized energy are within this threshold in (i) 38% (bar 1, denoted Jaguar (S)) and (ii) 35% (bar 2, denoted Jaguar (O)) respectively, the Orient calculations with a DMA and explicit polarization, offer no improvements over the QM/MM results, and indeed deliver worse results. Similar results are seen at the 1.0 Å threshold (yellow bars in Fig 4.3B, not discussed)

While we tend to see a reduction in performance under optimization, this does not arise when exponential repulsion is used, possibly this is related to the softer repulsion potentials, but further investigation is likely to be difficult because the exponential force fields are probably not well balanced.

4.3.3.2 RMSD of ≤ 2 Å from experimental geometry (orange bars)

The Orient calculations with DMA only offer improvements over QM/MM for both single point and optimized calculations when DMA is implemented with a Lennard Jones potential without explicit ligand polarization. Specifically, there is a ~2% improvement for the single point calculation (bar 3, denoted LS (S)) and a ~4% improvement for the optimized calculation (bar 4, denoted LJ (O)) relative to QM/MM. The Orient calculations with explicit ligand polarization again are worse than the results without polarization.

4.3.3.3 RMSD of ≤ 4 Å from experimental geometry (dark red bars)

The 4 Å cut-off is rather crude and generally just measures whether the ligand is in the

correct binding site. The Orient calculations with DMA only offer improvements for the single

point calculations when DMA is implemented with a Lennard Jones potential without explicit

ligand polarization, where there is ~2% improvement. When the Orient calculations with DMA

are optimized the lowest energies for all methods improve on docking success over QM/MM

at this threshold. The explicit ligand polarization again however offers no improvement or

makes the results worse for Orient calculations with DMA.

Assessment of DMA energy minimized geometries

Figure 4.3D shows whether the Glide geometries that have undergone energy minimization

with DMA offer any improvement in docking success over the standard Glide geometries if

they are rescored. This idea is applied to the Glide scoring methods namely GSCORE and

CVDW, using the 'score in place' option. The benchmark results as in 4.3.1.4 are the 1st

Ranked or lowest energy Glide pose (docked according to GSCORE) using flexible ligand

docking and scored by GSCORE and CVDW.

Figure 4.3D (bar 1, denoted Glide(GS)) for the 1st Ranked GSCORE pose, shows that this

pose has a distance measured by RMSD from the experimental pose of: (i) ≤ 0.5 Å; (ii) ≤ 1 Å;

(iii) ≤ 2 Å; (iv) ≤ 4 Å in (i) 32%, (ii) 48%, (iii) 74% and (iv) 88% respectively of the 74 cases

respectively.

When we consider the lowest energy CVDW pose, Figure 4.3D (bar 4) shows a pose

distance measured by RMSD from the experimental pose of: (i) ≤ 0.5 Å; (ii) ≤ 1 Å; (iii) ≤ 2 Å;

(iv) ≤ 4 Å in (i) 37%, (ii) 47%, (iii) 68% and (iv) 84% of the 74 cases respectively. These

benchmark results can again be used as a guideline to see if energy calculations that include

DMA can improve docking.

4.3.3.4 Assessment RMSD from the experimental geometry

Figure 4.3D shows that the Orient - optimized geometries, optimized using DMA with a Lennard Jones Potential, offered slight improvements of ~3% for GSCORE scoring (bar 2, green denoted LJ(GS)) and ~1% for CVDW scoring (bar 5, green denoted LJ(CVDW)). At the 1.0 Å threshold, only CVDW scoring with both the Orient based optimized energy geometries (with DMA) offered some improvements in docking success at this threshold with at best ~3% improvement when the Exponential 6 Williams potential was used (bar 6, yellow, denoted EXP6 (CVDW), and ~1% improvement for the Lennard Jones potential geometry (bar 5, yellow, denoted LJ (CVDW). At the 2.0 Å threshold, only CVDW scoring with both the Orient based optimized energy geometries (with DMA) offers any comparability in docking success (bar 5, orange, denoted LJ (CVDW). All methods give similar results at the 4.0 Å threshold. The DMA results in Figure 4.3D are not hugely different to the non-DMA results, with some indication of improved results at lower thresholds, which is consistent with the ability of DMAs to treat anisotropic interactions. Given that the DMA-based docking methods have not been optimized in anyway, this is a positive result.

4.3.4 Hydrogen bond analysis.

In Table 4.1 below, we decided just to consider the 32 cases where RMSD analysis produces a pose below two angstroms for Glide SP (GSCORE) and both DMA methods. To see if the RMSD when the answer is approximately correct, improves with DMA-based docking methods. However, this analysis revealed no notable improvements in performance across the methods, with Glide and DMA with a Lennard Jones potential producing marginally better results than DMA with a William Exp6 potential. Therefore, we also looked at the number of hydrogen bonds between the ligand and protein, using the molecular graphic program - visual molecular dynamics (vmd) (Humphrey et al., 1996) and its *hbond* analysis tool (Caddigan et al., 2003).

In this first investigation of the quality of the hydrogen bonds, we looked at the lowest energy/top ranking poses for Glide SP (GSCORE), and both DMA methods. We set two

hydrogen bond thresholds for angle and distance, one tighter than the other. The angle

thresholds were set at 20° and 30°, from the ideal hydrogen bond angles. The 20° angle was

default, and the 30° was recommended 'looser criteria' in the vmd literature (Caddigan et al.,

2003). The distance thresholds from donor atom D with a hydrogen bond attached to it,

(e.g., D…H) were set at 2.8 Å and 3.2 Å from the acceptor atom respectively. The distances

were arbitrarily set at 0.2 Å either side of the default distance for hydrogen bond counting of

3 Å. The results of the numerical hydrogen bond analysis are shown in Table 4.2.

Table 4.1 Molecules with RMSD <2 A for both Glide SP and both DMA Methods.

PDB	RMSD Å Glide SP	RMSD Å LJ	RMSD Å Williams	ΔRMSD Å LJ	ΔRMSD Å Williams
1eqg	**1.1**	1.3	1.3	0.2	0.2
1s39	0.2	**0.1**	0.4	-0.1	0.2
1yz3	0.4	0.3	**0.2**	-0.1	-0.2
2adu	1.5	**1.2**	1.4	-0.3	-0.1
3imc A	**0.2**	**0.2**	**0.2**	0	0
3imc B	**0.2**	**0.2**	0.3	0	0.1
3img A BZ3	**0.1**	0.2	**0.1**	0.1	0
3img B BZ3	0.2	**0.1**	0.2	-0.1	0
1f5f	**0.2**	**0.2**	**0.2**	0	0
1f8e	**1**	1.1	1.1	0.1	0.1
1m3u	**0.9**	1.7	**0.3**	0.8	-0.6
1mlw	**1.3**	1.8	1.6	0.5	0.3
1ofz 1313	0.4	0.3	**0.2**	-0.1	-0.2
1ofz 1314	0.2	**0.1**	0.2	-0.1	0
1pwm	0.2	0.3	0.3	0.1	0.1
1r5y	**0.1**	**0.1**	0.2	0	0.1
1sqn	**0.1**	**0.1**	**0.1**	0	0
1uwc	0.2	**0.1**	0.7	-0.1	0.5
1w1a	**0.6**	0.7	0.7	0.1	0.1
1xfg	**0.6**	**0.6**	**0.6**	0	0
1ynh	**0.6**	**0.6**	0.8	0	0.2
2b0m	**0.6**	1.6	1.2	**1**	0.6
2bkx	**1.5**	2.1	2.1	0.6	0.6
2bl9	0.2	**0.1**	**0.1**	-0.1	-0.1
2f6x	1.7	1.8	**1.5**	0.1	-0.2
2fdv	1.3	**0.6**	1.2	-0.7	-0.1
2ff2	0.2	0.2	**0.1**	0	-0.1
2gvv	0.8	**0.5**	**0.5**	-0.3	-0.3
2i5x	1.5	1.3	**0.3**	-0.2	**-1.2**
2iba	**0.5**	0.8	0.6	0.3	0.1

2j5s	1.4	**0.3**	0.4	**-1.1**	**-1**
3c0z	**0.9**	1	1.9	0.1	**1**
μRMSD	**0.65**	**0.68**	**0.66**		
TOTAL BEST	16	16	13		

The table shows individual molecule RMSD Å columns 2 to 4, and mean **μ**RMSD Å from these results. **Bold** in these three columns indicates the lowest (or matching lowest) RMSD for that molecule. TOTAL BEST is a sum of these lowest RMSDs by molecule for each methods lowest energy/top ranking pose from Glide SP and the DMA methods. LJ is DMA with Lennard Jones potential, and Williams is DMA with the Williams potential. The last two columns show the change in RMSD (ΔRMSD Å), compared to Glide SP (GSCORE) by DMA method. For these last two columns **bold** indicates an improvement in RMSD >1 Å, **bold red** indicates a refractory case where Glide shows an improvement over DMA method >1 Å. Capital letters e.g., A or B after the pdb code in first column indicate the chain, also BZ3, 1313, and 1314 refer to specific ligands where more than one ligand was present in the pdb and used in our dataset.

Table 4.2 Numerical Hydrogen Bond Analysis using vmd *hbond* plug in and the cut-off criteria: (1) Acceptor atom distance from donor atom with hydrogen bond attached, and (2) angle within threshold from ideal hydrogen bond.

PDB	30º & 3.2Å Glide SP	20º & 2.8Å Glide SP	30º & 3.2Å LJ	20º & 2.8Å LJ	30º & 3.2Å Williams	20º & 2.8Å Williams
1eqg	3	3	3	3	3	1
1s39	4	0	4	1	3	**2**
1yz3	1	0	1	0	1	0
2adu	0	0	0	0	0	0
3imc A	1	0	1	0	1	0
3imc B	2	0	1	0	1	0
3img A BZ3	2	0	2	0	2	0
3img B BZ3	3	0	2	0	1	0
1f5f	1	0	2	0	2	0
1f8e	6	**2**	6	**2**	4	1
1m3u	3	1	2	1	3	**2**
1mlw	1	0	1	**1**	1	0
1ofz_1313	5	**2**	5	**2**	5	1
1ofz_1314	3	0	3	0	3	0
1pwm	2	0	2	0	2	0
1r5y	6	**1**	5	**1**	5	0
1sqn	0	0	0	0	0	0
1uwc	2	0	3	0	0	0
1w1a	4	1	5	1	5	**2**
1xfg	6	2	7	**3**	8	1
1ynh	7	**3**	8	**3**	9	2
2b0m	3	1	2	**2**	2	1

2bkx	6	**3**	6	**3**	3	0
2bl9	2	0	3	0	3	0
2f6x	5	2	7	**4**	6	2
2fdv	1	0	1	0	0	0
2ff2	5	3	6	**4**	5	3
2gvv	2	**2**	2	1	2	0
2i5x	4	2	5	**3**	5	0
2iba	3	0	3	**1**	3	0
2j5s	0	0	1	**1**	1	0
3c0z	2	0	2	0	0	0
TOTALS	95	28	101	37	89	18
TOTAL BEST		6		13		3

The table shows the amount of hydrogen bonds at two cut-off criteria for the lowest energy/ top ranking poses from Glide SP and the DMA methods. LJ is DMA with Lennard Jones potential, and Williams is DMA with the Williams potential. Total number of hydrogen bonds are shown at the bottom for each method and criteria. **Bold** indicates the highest per molecule number of hydrogen bonds at the tighter threshold across the methods (or a match above zero), a TOTAL BEST sum is provided at the bottom, for each method.

The results in Table 4.2 show, that DMA with a Williams exp6 potential does not improve over Glide SP (GSCORE), when the lowest energy/ top ranked geometries are approximately correct. However, using a Lennard Jones potential with DMA increases the number of hydrogen bonds identified at both cut-off criteria, with 6 more at the looser thresholds, and 9 more at the tighter thresholds. The tighter cut-off criteria indicate that the quality of these hydrogen bonds have improved. However, future research, visualization and measurements would be helpful, before being able to draw any clearer conclusions.

4.3.5 Cluster analysis.

The results are presented in figure 4.4 using the same analysis and criteria as 3.3.4 and 3.2.10.2 and are similar to the results presented above, namely that there are a few low-lying energy minima that have a low RMSD and where the pose would be thermally accessible. This emphasises that it is sometimes mis-leading to only consider the top ranked pose.

4.3.6 Probabilities.

The results presented in figures 4.5 and 4.6 are just a different method of presenting the previous discrete 'cut off' results, that may provide better visualisation of some trends. So we

again see there are no notable improvements using DMA methods, over QSITE Jaguar QM/IMPACT MM calculations. Again the figures show, single point calculations generally fare better than optimized calculations, and polarization does not improve results, but rather makes them worse.
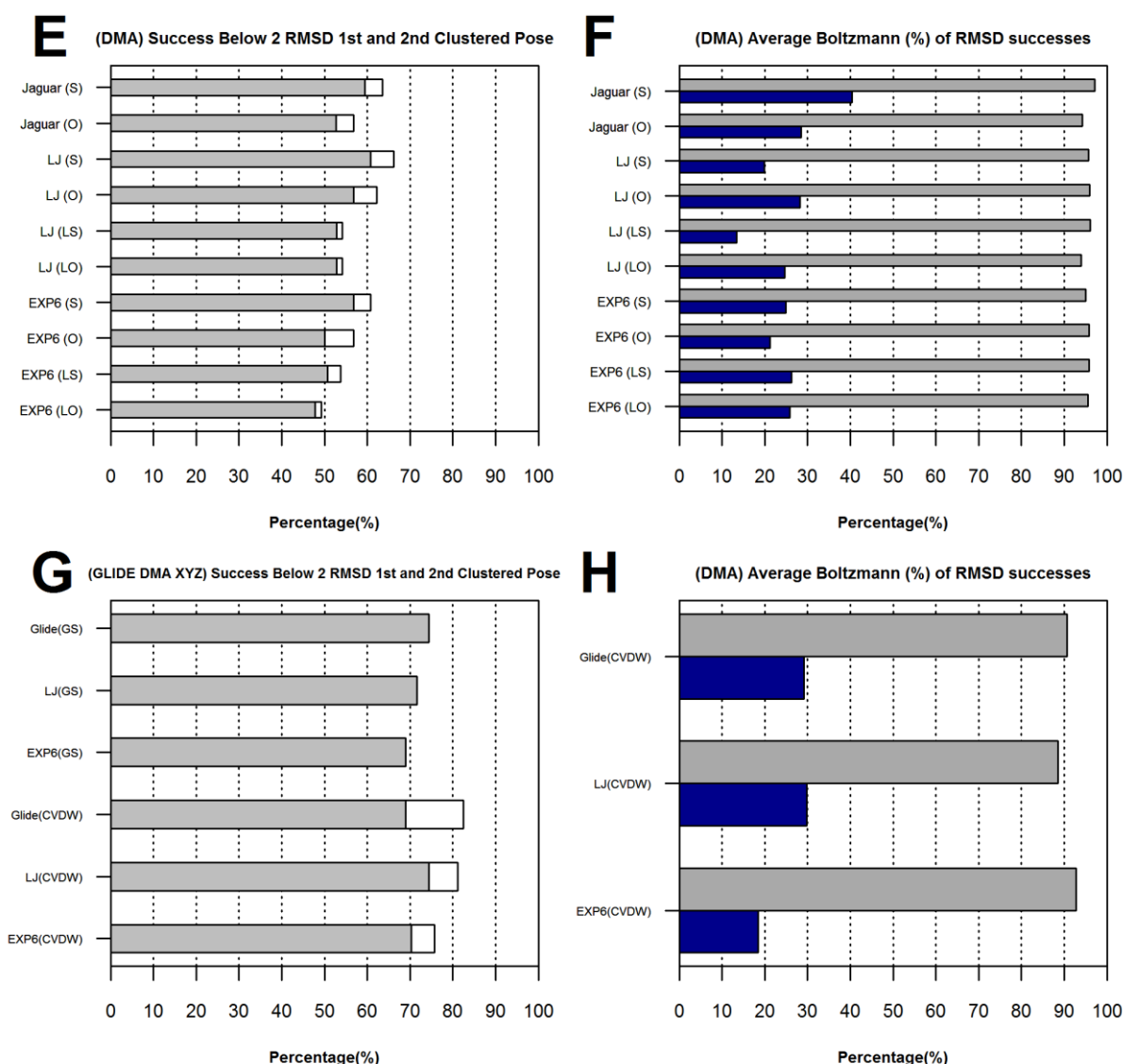

Figure 4.4. A Rough (<2 RMSD) Assessment of Additional Clustering on (15 pose) Docking Success (as Figure 4.3). In E and F, Jaguar indicates - QM/MM energy calculations are done in Jaguar. DMA methods are indicated by (i) LJ and (ii) EXP6 to denote the specific type of exchange repulsion used: Single point calculation (S) and Optimized calculation (O) denote the energy used for scoring, and (L) indicates that isotropic ligand polarization has been used in Orient. In G and H, Glide refers to Glide SP (also Chapter 3), LJ and EXP6 again refer to treatment of van der Waals forces in DMA methods, (GS) indicates GlideScore scoring used (Plot G), and (CVDW) indicates Glide MM Force-field only used for 'score in place' of the optimized Orient (DMA) geometries. Plots E and G, show the percentage over 74* molecules where the highest ranking/lowest scoring pose cluster (Grey) is successful (RMSD < 2 Å) and the % increase when additionally accounting for 2nd Pose cluster (White),

RMSD < 2 Å when 1st pose is not. Plots H and J show the Boltzmann probabilty (%) averaged over 74 molecules for the Pose Population adopting the highest ranked/lowest scoring pose minima (Grey) and of adopting the 2nd Pose Cluster minima when the first does not have an RMSD <2 Å (Dark Blue). *When isotropic ligand polarization was attempted in Orient only 72 molecules were used for LJ and 68 for Exp6 when calculating docking success percentages due to convergence/implementation issues. The basic unpolarized Glide results (scored using GSCORE) are given in Figure 3.1.
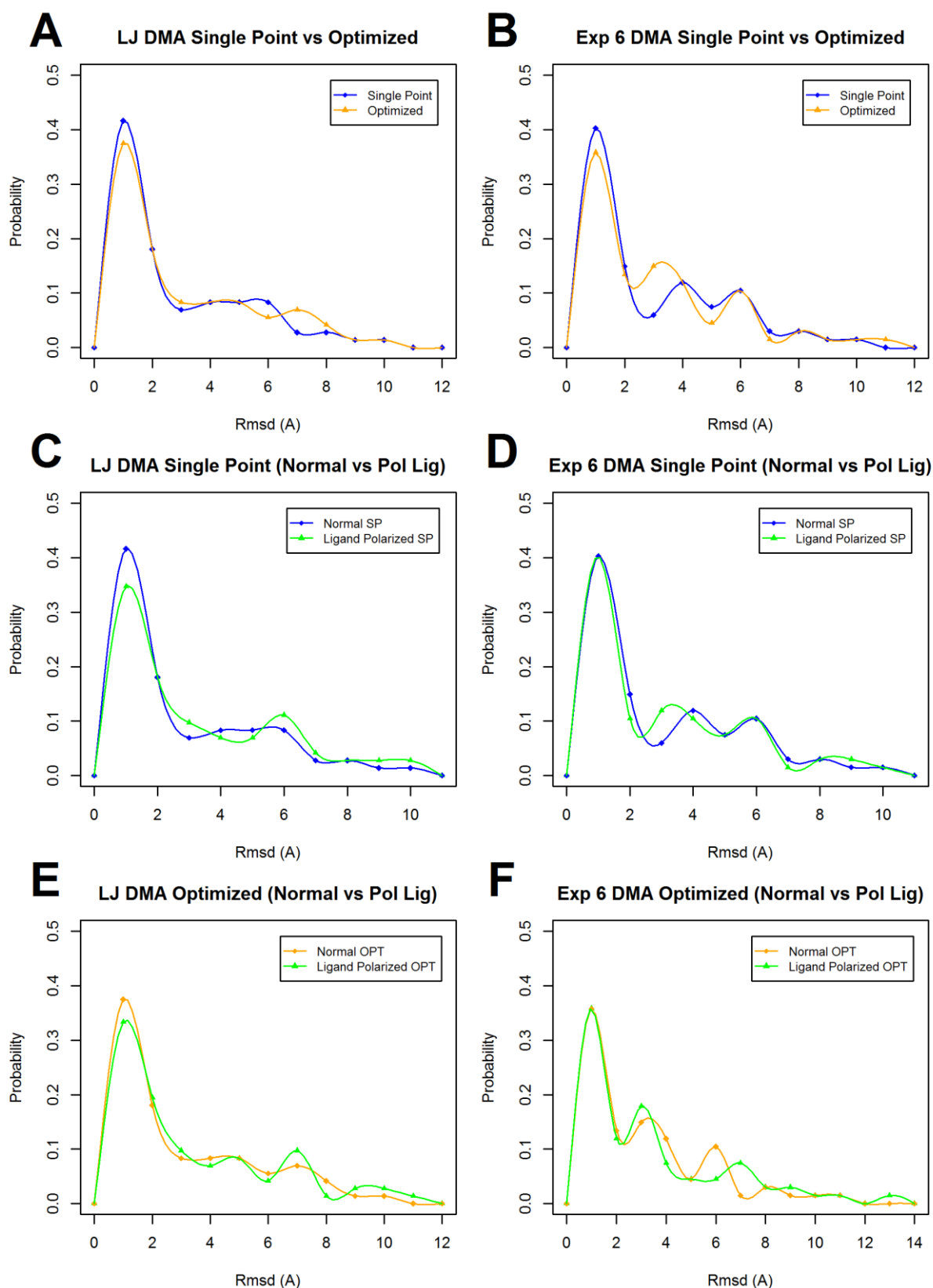
Figure 4.5. RMSD across the 74 molecules top-ranked poses as a 1 angstrom binned probability plot (decimal) for DMA/Orient-based methods.
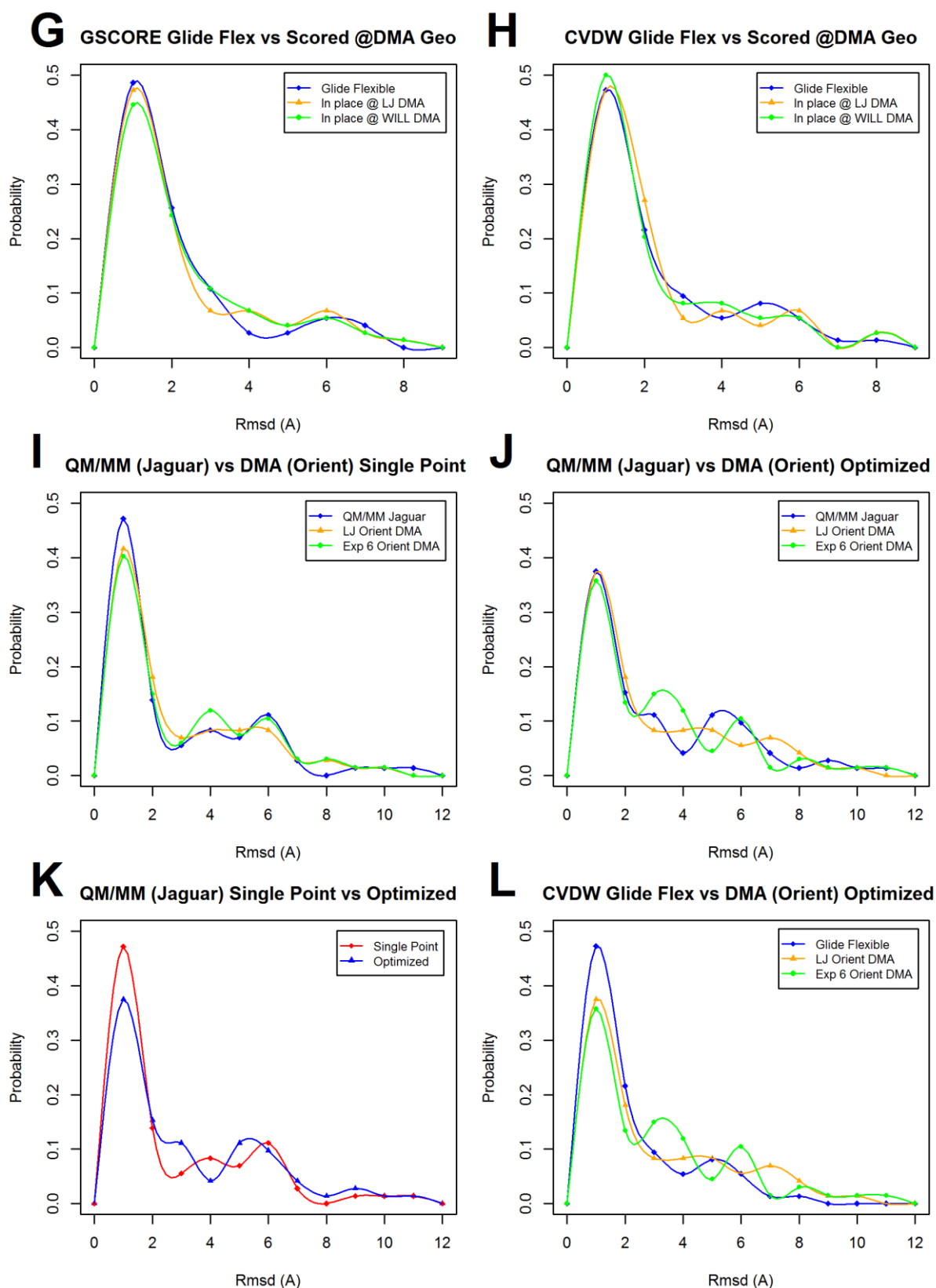
Figure 4.6. RMSD across the 74 molecules top-ranked poses as a 1 angstrom binned probability plot(decimal) for Jaguar QM/MM and DMA/Orient-based methods and Glide score in place. WILL DMA includes exponential-6 repulsion. The QM/MM results (K) are given for comparative purposes and is the same as Figure 4.2.

4.4 Conclusions

Generally, excluding the hydrogen bond analysis section 4.3.4, the DMA results are disappointing as the results are not as good as those given in Chapter 3, despite the potential to include important factors such as anisotropic charge distributions and exponential repulsion through Orient. However, before dismissing the relevance of these factors for drug design, it is important to recognize a number of factors. Firstly, Orient was not designed as a drug design tool, meaning that it was not possible to treat the whole enzyme, but rather the calculations were restricted to a small shell of residues around the active site. This process was rather laborious and so in its current implementation, would not be suitable for drug design applications. Secondly, the exponential terms of the force field were not optimized to be included with the anisotropic electrostatics, hence the Lennard-Jones repulsion performed better than exponential repulsion in Figure 4.6I. Thirdly, there was no treatment of solvation. This manifested itself in a rather subtle way, in that the results from optimization were not as good as the results for single point calculations – for both Orient (DMA) and Jaguar (QM) calculations. The most obvious explanation of this is that the 'gas phase' optimization moved the structures away from the solution structures, which can to some extent be modelled using Glide due to its implicit treatment of solvent effects. In addition, it is possible that a higher basis set, e.g. one containing diffuse basis functions (Illingworth et al., 2008) might have given a better DMA, and hence better results. Given a hierarchy of errors, it is important to address the most important errors first, hence in the next chapter we address the role of specific water molecules in the active site. However, the results in this chapter are encouraging as judged by comparing the performance of basic Glide SP with QM/MM and DMA/LJ and DMA/LJ poses rescored by CVDW, and with the polarized Glide results as shown in Table 4.3. The Table shows that in most respects the performance of the QM/MM and DMA methods is comparable to the Glide-based methods.

Table 4.3. A comparison of the new methods introduced in Chapter 4 with the basic Glide and polarized glide methods discussed in Chapter 3.

	Percentage		
Method	top pose is lowest RMSD pose	RMSD < 0.5 Å	RMSD < 2.0 Å
Glide	36%	32%	74%
Glide(CVDW)	39%	37%	68%
QM/MM(S)	44%	38%	59%
LJ(S)	36%	31%	61%
LS(CVDW)	46%	38%	74%
POSE(C)	44%	36%	74%
REF(C)	49%	39%	75%

This favourable comparison suggests that these issues should be explored more fully in the future, but ideally a more protein-related piece of software should be used. In this respect, it would probably be worthwhile exploring the use of the Amoeba force field in Tinker, because this uses anisotropic electrostatics, and a softer, more realistic repulsion potential (but not exponential repulsion). Moreover, the program is designed for protein simulations, and so would offer the benefit of a molecular dynamics treatment of hydration. As an initial approach, it would be possible to use Tinker/Amoeba in a similar fashion to the way we used Orient, without the inherent disadvantages of Orient. For subsequent refinement, it would be possible to simulate the fully hydrated protein-ligand complex to check whether it was stable. When this work commenced, Tinker was not parallelized and so was not such an obvious choice, but we understand progress has been made in this area. Moreover, the Amoeba force field has a well-defined parameterization protocol and so extension to new ligands would not be as difficult as in our approach. Of the methods studied here, the QM/MM results were generally the best as shown in Figure 4.3A, but Figure 4.6I and 4.6J do show that the DMA out-performed the QM/MM calculations in some respects, see also Table 4.1. However, while QM/MM has many benefits over classical approaches, Chapter 3 shows that Glide generally works better (by virtue of its specific parameterization). The occasional success of QM/MM methods in FBDD should therefore be viewed with caution (Gleeson and Gleeson,

2009), but again these are methods that could be improved. The improvement through

polarization was generally less than in Chapter 3 and this is probably because the effects are

being swamped by other errors.

4.5. References

BANKS, J. L., BEARD, H. S., CAO, Y., CHO, A. E., DAMM, W., FARID, R., FELTS, A. K., HALGREN, T. A., MAINZ, D. T. & MAPLE, J. R. 2005. Integrated modeling program, applied chemical theory (IMPACT). *Journal of computational chemistry,* 26**,** 1752-1780.

BECKE, A. D. 1993. Density-functional thermochemistry. III. The role of exact exchange. *The Journal of Chemical Physics,* 98**,** 5648-5652.

BUCKINGHAM, R. & CORNER, J. 1947. Tables of second virial and low-pressure Joule-Thomson coefficients for intermolecular potentials with exponential repulsion. *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences,* 189**,** 118-129.

CADDIGAN, E., COHEN, J., GULLINGSRUD, J. & STONE, J. 2003. Vmd user's guide. *Urbana,* 51**,** 61801.

CIEPLAK, P., DUPRADEAU, F.-Y., DUAN, Y. & WANG, J. 2009. Polarization effects in molecular mechanical force fields. *Journal of Physics: Condensed Matter,* 21**,** 333102.

FERENCZY, G. G., WINN, P. J. & REYNOLDS, C. A. 1997. Toward improved force fields. 2. Effective distributed multipoles. *The Journal of Physical Chemistry A,* 101**,** 5446-5455.

FRIESNER, R. A., BANKS, J. L., MURPHY, R. B., HALGREN, T. A., KLICIC, J. J., DANIEL, T., REPASKY, M. P., KNOLL, E. H., SHELLEY, M. & PERRY, J. K. 2004. Glide: a new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy. *Journal of medicinal chemistry,* 47**,** 1739-1749.

GLEESON, M. P. & GLEESON, D. 2009. QM/MM calculations in drug discovery: a useful method for studying binding phenomena? *Journal of chemical information and modeling,* 49**,** 670-677.

HEHRE, W. J., DITCHFIELD, R. & POPLE, J. A. 1972. Self—consistent molecular orbital methods. XII. Further extensions of gaussian—type basis sets for use in molecular orbital studies of organic molecules. *The Journal of Chemical Physics,* 56**,** 2257-2261.

HOHENBERG, P. & KOHN, W. 1964. Inhomogeneous electron gas. *Physical review,* 136**,** B864.

HUMPHREY, W., DALKE, A. & SCHULTEN, K. 1996. VMD: visual molecular dynamics. *Journal of molecular graphics,* 14**,** 33-38.

ILLINGWORTH, C. J. R., PARKES, K. E. B., SNELL, C. R., FERENCZY, G. G. & REYNOLDS, C. A. 2008. Toward a consistent treatment of polarization in model QM/MM calculations. *The Journal of Physical Chemistry A,* 112**,** 12151-12156.

JONES, J. E. 1924. On the Determination of Molecular Fields. II. From the Equation of State of a Gas. *Proceedings of the Royal Society of London. Series A,* 106**,** 463-477.

LEE, C., YANG, W. & PARR, R. G. 1988. Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. *Physical Review B,* 37**,** 785-789.

LONSDALE, R. & MULHOLLAND, A. J. 2014. QM/MM modelling of drug-metabolizing enzymes. *Current topics in medicinal chemistry,* 14**,** 1339-1347.

METRANGOLO, P., NEUKIRCH, H., PILATI, T. & RESNATI, G. 2005. Halogen Bonding Based Recognition Processes: A World Parallel to Hydrogen Bonding†. *Accounts of Chemical Research,* 38**,** 386-395.

MILLER, K. J. & SAVCHIK, J. 1979. A new empirical method to calculate average molecular polarizabilities. *Journal of the American Chemical Society,* 101**,** 7206-7213.

PONDER, J. W. 2004. TINKER: Software tools for molecular design. *Washington University School of Medicine, Saint Louis, MO,* 3.

PONDER, J. W. & CASE, D. A. 2003. Force fields for protein simulations. *Advances in protein chemistry,* 66**,** 27-86.

SINGH, J. & THORNTON, J. 1993. Atlas of protein side-chain interactions. Vols. I and II. *Acta Cryst,* 49**,** 355-356.

STONE, A. 1981. Distributed multipole analysis, or how to describe a molecular charge distribution. *Chemical Physics Letters,* 83**,** 233-239.

STONE, A. 2005. Distributed Multipole Analysis of Gaussian wavefunctions GDMA version 2.2. 02.

STONE, A., DULLWEBER, A., ENGKVIST, O., FRASCHINI, E., HODGES, M., MEREDITH, A., NUTT, D. & POPELIER, P. 2006. LA; Wales, DJ Orient: a program for studying interactions between molecules, Version 4.6; University of Cambridge: 2006. *Inquiries to AJ Stone, ajs1@ cam. ac. uk.*

WILLIAMS, D. E. 1999. Improved intermolecular force field for crystalline hydrocarbons containing four- or three-coordinated carbon. *Journal of Molecular Structure,* 485–486**,** 321-347.

WILLIAMS, D. E. 2001. Improved intermolecular force field for molecules containing H, C, N, and O atoms, with application to nucleoside and peptide crystals. *Journal of computational chemistry,* 22**,** 1154-1166.

5 Inclusion of specific water molecules

5.1 Introduction

It has been claimed that water is the new dimension in drug design (Congreve et al., 2011). While a full treatment of water in docking is beyond the focus of this thesis where we have largely concentrated on the effects of polarization, it has become apparent that water is an extremely important factor in drug-target interactions. Mason and other have discussed the idea that ideally, a drug should displace 'unhappy' (i.e. weakly bound) water molecules to gain the entropy increase upon their release, but that there is little to be gained by displacing 'happy' (i.e. strongly bound) water molecules (Congreve et al., 2011, Wang et al., 2011, Young et al., 2007). Barillari et al., have shown that the number of hydrogen bonds formed by a water identified in X-ray crystal structures is a key factor in determining how strongly it binds, as determined by free energy simulations (Barillari et al., 2007). We have therefore taken a simple approach to identifying 'happy' water molecules: we have simply counted the number of hydrogen bonds to other non-water molecules, i.e. to protein and to ligand. This is in line with Barillari et al., who found from Monte Carlo simulations that tightly bound water molecules tended to be in polar cavities and make 3 or 4 hydrogen bonds while weakly bound water molecules tended to be in polar cavities with less than 3 hydrogen bonds (Barillari et al., 2007). We have focused on water molecules that interact directly with the ligand and have classified them by the number of hydrogen bonds formed. In previous chapters we ignored the role of water molecules in docking, but given that 'happy' water molecules play a positive role in ligand binding, we have investigated whether incorporation of these water molecules (i.e. with multiple hydrogen bonds) in the target increases docking success. In addition, we have investigated whether docking can be used to predict the position of a subset of these water molecules, both in the presence of ligand (which will be useful in low resolution structures where water molecules are not well defined) and in the absence of ligand (which may be useful in allosteric binding sites or in cases such as HIV protease where the water pattern is ligand dependent).

5.2 Methods

All methods follow the same format as those described in section 3.2 with the exception of the following.

5.2.1 Source of Molecules used for Validation Dataset and Water preparation

We used the previous 74 molecule dataset that included the 54 molecules from the SERAPhiC set (Favia et al., 2011) and the 20 molecules from the preliminary set (Congreve et al., 2008, Hung et al., 2009) discussed in Chapter 3.2.3. In addition to the two 'no water' preliminary test set docking runs discussed in 3.2.4, we also did two initial preliminary test set docking runs 'with water', the first of these docking runs included just optimized waters from the X-ray structure. The second was achieved by using only structural homology source clustered waters within a 5 angstrom cut-off of the ligand, in these structures the water molecules are present at roughly the same position in the structural homologues. These were acquired by running a BLASTp search (Madden et al., 1996) of the PDB structural databank (Berman et al., 2003). This was done for all our preliminary models, with the rest of the Mycobacterium Tuberculosis Pantothenate Synthetase structures mention in 3.2.3 (Hung et al., 2009), sharing BLAST structural homology results with 3img.

The resulting structures between (5 and 42) were superimposed in Maestro 9.3. (Banks et al., 2005) and clustered visually (for all 20 molecules). This was done by displaying waters within 5 angstroms then showing the overlaid superimposed water result, and keeping the waters for the final structure when at least 5, or 60% of the waters (when less than 10 structures were available) clustered at the water site. This revealed a good number of potential conserved water sites within the protein structures. The results for both docking runs were very good (data not shown). But it became apparent that few (4) molecules from the preliminary dataset had conserved waters that made useful interactions with the ligand and the protein binding site, therefore docking success was achieved for these molecules simply through confining the ligand conformation possibilities when posing

(e.g., 3img (2), 2jjc, 1yz3). It was with this in mind a new approach was taken, inspired by (Barillari et al., 2007), where a single water molecule that had a near optimal number of non water-water hydrogen bond interactions would be used for the Docking studies. At this point, the rest of the structural homology conserved water data for the SERAPhiC set (Favia et al., 2011) was also gathered for later use (not used here). However, the new strategy was to re-import the dataset of 74 molecules into Maestro 9.3 (Banks et al., 2005) and carry out the ligand and protein preparation steps as in 3.2.4 without removing the water and then afterwards visualizing the hydrogen bonding interactions made with the ligand and protein, for cataloguing. These interactions were divided into three groups, the water molecules were divided by having: (i) *Strong interactions* - three or more hydrogen bond interactions (one or more at each water hydrogen), with at least one ligand and one protein interaction, (ii) *Moderate interactions* - two hydrogen bond interactions with at least one ligand and one protein interaction, (iii) *Weak interactions*: one hydrogen bond interaction with the ligand.

Waters were then all removed from the protein structure with the exception of the 'lone water' molecule that was categorized as above. All of these groups and structures were used in water docking experiments in Figure 5.3, see section 5.3.7 and 5.3.8, including an additional docking run where the ligand was removed for the *strong interactions*. However, the *strong interaction* and *moderate interaction* structures were only used for ligand docking runs, that included the particular 'lone water' molecule, the results for the *moderate interactions* are considered in Appendix B only. The new structures containing *strong interaction* 'lone water' molecules used in 5.3.1 to 5.3.6 are summarized in Table 5.1. below. Additionally, another set of structures where the hydrogens of these 'lone water' molecules from Table 5.1 were QM/MM orientated using methods described in 5.2.2 from the pre-prepared MM structure.

Table 5.1 Dataset structures containing the strongly interacting 'lone water' molecules. Here (brackets) indicate the ligand molecule used where more than one was available. 3IMG A and 3IMG B refer to chain A and B respectively.

PDB NAME	WATER ID	NO H-BONDS	PDB NAME	WATER ID	NO H-BONDS
1F8E	484	3	1WOG	1680	3
1F8E	732	3	2F6X	6	3
1MLW	629	3	2HDQ (501)	532	3
1Y2K	1016	3	2HDQ (501)	537	3
2BKX	2362	3	2V77	2365	3
2I5X	61	3	3EKO	902	4
2J5S	2247	3	3EKO	904	3
1E2I (APS)	2044	3	1YZ3	1040	3
1E2I (APS)	2123	3	2JJC	2166	3
1K0E	1049	4	3IMG A (BZ2)	627	3
1SD1	1337	4	3IMG B (BZ2)	633	3
1SD1	1361	3			

5.2.2   Additional tasks during Ligand and Protein Preparation

The crystal structures were imported and prepared for docking using the methods described in 3.2.4, with the additional water considerations discussed in 5.2.1. When polarization was used, this involved polarization of the 1st ranked GSCORE pose, which produced the best results in sections 3.3.1 to 3.3.5 (see also Appendix A) using the methods of section 3.2.7 and 3.2.8. In addition to the previous preparation, a new set of *strong interaction* structures where produced by the QM/MM orientation of the 'lone water' hydrogens, using the QSITE program within the Maestro 9.3, which called upon the Jaguar 7.9 program for the QM region geometry optimization calculation, and IMPACT 5.8 was used for the MM region single-point calculation, as section 3.2.9. The quantum mechanical calculations employed the B3LYP method with a 6-31G* basis set, denoted B3LYP/6-31G*. The ligand was removed in grid generation and QM/MM derived partial charges were not kept at this stage.

5.2.3   Docking Setup

Flexible ligand docking with the same grid and docking protocols used in section 3.2.5 was used, with the exception of the inclusion of the 'lone water' molecule within the grid. In addition to this amendment, the outer box size for the area considered when posing

by the Glide search algorithm was reduced to 15 x 15 x 15 Å ^{3} when carrying out the 'lone water' docking, but not when docking the ligand.

5.3 Results

5.3.1 Analysis of top ranked poses

5.3.1.1 Using a single strongly bound ligand attached water molecule

As described previously in 3.3.1.1, Glide SP flexible ligand docking is usually able to identify the experimental pose amongst the full set of returned poses, whether as the top ranked pose, or as a lower ranked pose. In Figure 5.1 generally we consider the 23 cases where a ligand attached water molecule with three or more non water hydrogen bonds exists within our Dataset and the effects of adding these 'lone waters' to the target receptor grid on subsequent ligand docking. Firstly, though, to establish benchmarks in Figure 5.1A (1st bar, denoted GLIDE 23)) we consider the 'no water' 23 molecules across the above range from the original docking run for the top ranked pose (according to GSCORE), the figure shows that this pose is also the one with the smallest RMSD to the experimental pose in 22% of the current 23 cases. When we consider these 23 molecules from the original 'no water' docking run where the (i) ligand partial charges (3rd bar, denoted POSE (L)) and (ii) ligand and protein partial charges (5th bar, denoted POSE(C)) were explicitly polarized using the geometries of the 1st Ranked GSCORE pose, then flexible ligand re-docking results in the top ranked pose being the one with the smallest RMSD to the experimental pose in 43% and 35% of cases respectively. These three results of 22%, 43% and 35% are therefore benchmark results that can be used to see if a strongly bonded ligand attached 'lone water' molecule can improve ligand docking.

When we now include the 'lone water' molecule in the receptor grid for the 23 cases using   (i) the default OPLS 2005 partial charges (2nd bar, denoted GLIDE(3H)); (ii) the polarized ligand partial charges (4th bar, denoted POSE(L3H)); and (iii) the polarized ligand and protein partial charges (6th bar, denoted POSE(C3H)) that are both explicitly polarized

using the geometries of the 1st Ranked GSCORE pose, then flexible ligand re-docking results in the top ranked pose being the one with the smallest RMSD to the experimental pose in (i) 52%, (ii) 43% and (iii) 43% of cases respectively. Under these circumstances, adding a 'lone water' offers a clear ~30% improvement upon default 'no water' Glide docking when scored by GSCORE, while polarization does not further improve the docking results where a 'lone water' is included. While having the luxury of including a 'lone water' clearly helps, again as in 3.3.1.1 a key issue is whether the ligand is correctly polarized - an issue that is largely affected by the initial geometry. To test whether further improvement to these lone water docking results could be made, the reference structures were docked again after the hydrogens from the water molecule had been orientated by a quantum mechanical (QM/MM) optimization. When we now include the 'lone water' molecule with QM/MM orientated hydrogens in the receptor grid for the 23 cases using (i) default partial charges (7th bar, denoted GLIDE(3HO)), then we also use the (ii) ligand partial charges (8th bar, denoted POSE(L3HO)) and (iii) ligand and protein partial charges (9th bar, denoted POSE(C3HO)) that are explicitly polarized using the geometries of the 1st Ranked GSCORE pose, then flexible ligand re-docking results in the top ranked pose being the one with the smallest RMSD to the experimental pose in (i) 65%, (ii) 48% and (iii) 65% of cases respectively. The 'lone water' with QM/MM orientated hydrogen results offer ~13% improvement over the previous 'lone water' with MM orientated hydrogen results and ~43% improvement over the results with 'no water' across the same range of 23 cases. Interestingly, while polarization still offers no further improvement, the result for the explicitly polarized ligand and protein partial charges is no worse, also at 65% success.
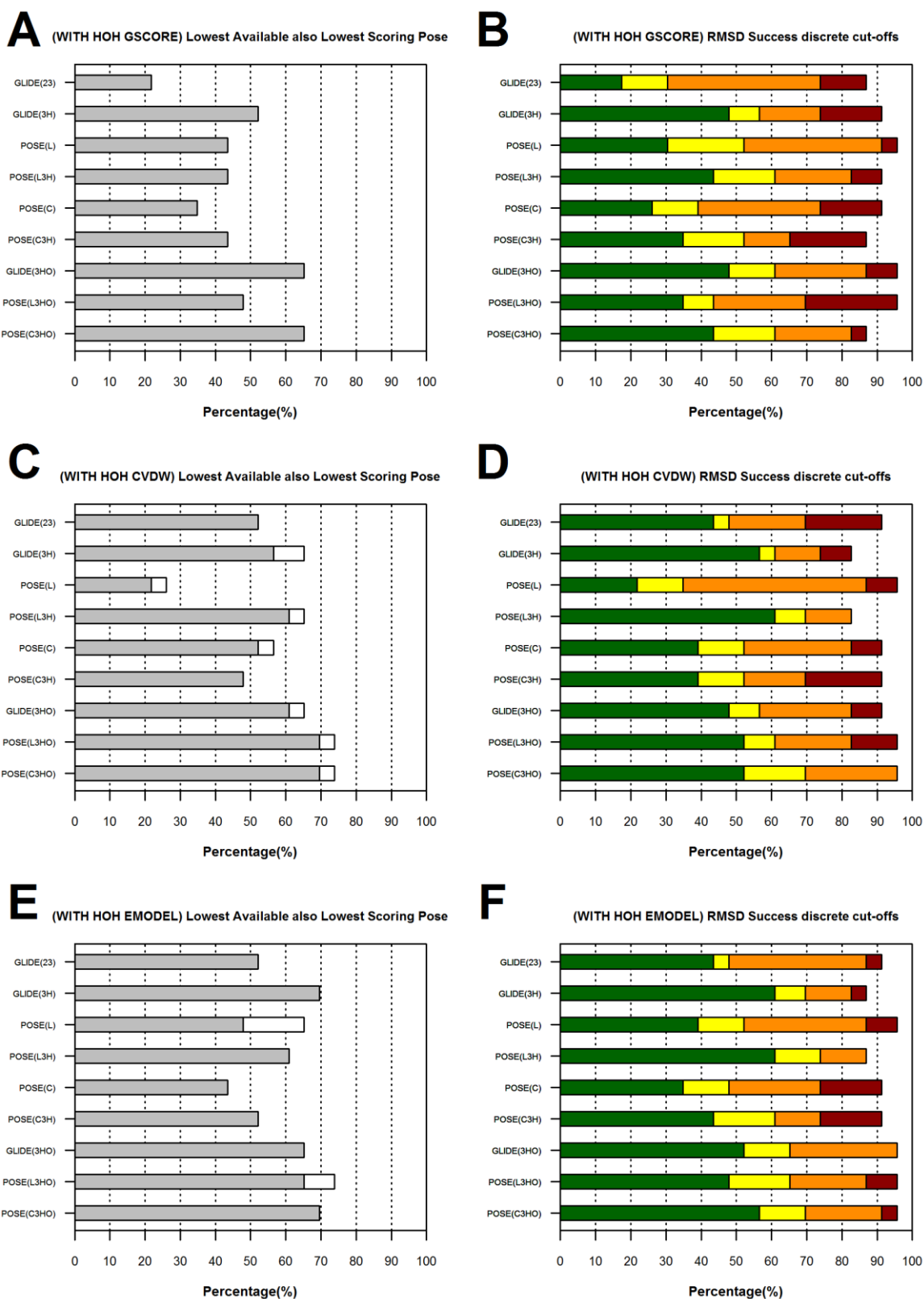
Figure 5.1 Assessment of Polarization of the Ligand/Complex and the a Ligand attached water Flexible re-docking success (15 poses) using Glide Standard Precision - GLIDE where (23) refers to number of molecules from original dataset without single water molecule, (3H) refers to the minimum number of protein/ligand hydrogen bonds formed namely 3, (L) is Re-Docked with Polarized Ligand partial charges or (C) is Complex partial charges (C); (3HO)

is the same as (3H) but the single water molecule hydrogens were orientated by QM/MM calculation before the docking run. POSE – is polarized at the Highest Ranking Glide GSCORE pose geometry. (CVDW), (EMODEL), (GSCORE) are the scoring functions used by Glide. Plots A, C, E – Percentage over 23 of the 74 molecules where highest ranked/lowest scoring Pose (Grey), Pose cluster (White*) is the lowest available pose. Plots B, D, F – Percentage over 23 of the 74 molecules where highest ranked/lowest scoring pose is at discrete cut-offs <0.5 Å RMSD (Green);<1 Å RMSD (Yellow); <2 Å RMSD (Orange); <4 Å RMSD (Red). *Clustering parameters set at <1.1 kcal/mol and <1 Å RMSD. The initial Docking strategy was to keep 15 poses.

5.3.2 Alternative approaches to identifying the top pose

The results of Figure 5.1A are interesting as they suggest a large improvement can be

obtained by including a lone water molecule with strong ligand and protein hydrogen bond

interactions. However, when considering polarization, in 3.3.2, we previously saw a marked

improvement when rescoring with alternative protocols. For this reason, as we have again

considered alternative scoring protocols for ranking the poses, namely CVDW (Figure 5.1C)

and EMODEL (Figure 5.1E).

5.3.2.1 CVDW rescoring: Using a strongly bound ligand attached water molecule

5.3.2.1.1 Top Ranked Pose (grey bars).

In Figure 5.1C (1st bar, denoted GLIDE(23)) we consider the 'no water' 23 molecules across

the above range from the original docking run for the top ranked pose (according to CVDW),

the figure shows that this pose is also the one with the smallest RMSD to the experimental

pose in 53% of the 23 cases. When we consider these 23 molecules from the original 'no

water' docking run where the (i) ligand partial charges (3rd bar, denoted POSE (L)) and (ii)

ligand and protein partial charges (5th bar, denoted POSE(C)) were explicitly polarized using

the geometries of the 1st Ranked GSCORE pose, then flexible ligand re-docking results in

the top ranked pose being the one with the smallest RMSD to the experimental pose in 22%

and 53% of cases respectively. These three results of 53%, 22% and 53% are therefore the

CVDW benchmark results that can be used to see if a strongly bonded ligand attached 'lone

water' molecule can improve ligand docking.

When we now include the 'lone water' molecule in the receptor grid for the 23 cases using   (i) the default OPLS 2005 partial charges (2nd bar, denoted GLIDE(3H)); (ii) the polarized ligand partial charges (4th bar, denoted POSE(L3H)); and (iii) the polarized ligand and protein partial charges (6th bar, denoted POSE(C3H)) that are both explicitly polarized using the geometries of the 1st Ranked GSCORE pose, then flexible ligand re-docking results in the top ranked pose being the one with the smallest RMSD to the experimental pose in (i) 57%, (ii) 61% and (iii) 48% of cases respectively. Under these circumstances, adding a 'lone water' offers some improvement ~4% improvement over the default 'no water' Glide docking when scored by CVDW, while polarization improves the docking results with a 'lone water' only when the ligand partial charges are polarized giving a further ~4% improvement.   This now ~8% improvement is also a large ~39% improvement over the 'no water' polarized ligand partial charges result across the 23 cases. This ~39% is also the same improvement over the 'no water' default partial charges result when scored by GSCORE above in 5.3.1.1.

When we now include the 'lone water' molecule with QM/MM orientated hydrogens in the receptor grid for the 23 cases using (i) default partial charges (7th bar, denoted GLIDE(3HO)), then we also use the (ii) ligand partial charges (8th bar, denoted POSE(L3HO)) and (iii) ligand and protein partial charges (9th bar, denoted POSE(C3HO)) that are explicitly polarized using the geometries of the 1st Ranked GSCORE pose, then flexible ligand re-docking results in the top ranked pose being the one with the smallest RMSD to the experimental pose in (i) 61%, (ii) 70% and (iii) also 70% of cases respectively. The 'lone water' with QM/MM orientated hydrogen results offer up to ~13% improvement over the previous 'lone water' with MM orientated hydrogen results when polarization of the ligand or ligand and protein partial charges is included, this is at best ~17% better than the 'no water' default partial charges glide docking result when scored by CVDW in the 23 cases.

5.3.2.1.2 Top Ranked Pose Cluster (white bars): Summary

In brief consideration of the cluster analysis in Figure 5.1C, when the first pose cluster measured by the criteria (see 3.2.10.2) that each cluster has an RMSD of ≤ 1 Å and energy

within 1.1 kcal mol^{-1} of the lowest energy member of that pose cluster and the pose with the lowest RMSD to experiment should also be a member of the first cluster, there are no particularly interesting patterns other than there is usually a slight improvement ~4% (white bars 3-5 and 7-9), however Glide SP with default charges and a 'lone water' (2nd white bar) shows ~8% improvement.

5.3.2.2   EMODEL rescoring: Using a strongly bound ligand attached water molecule

5.3.2.2.1 Top Ranked Pose (grey bars).

In Figure 5.1E (1st bar, denoted GLIDE(23)) we consider the 'no water' 23 molecules across the above range from the original docking run for the top ranked pose (according to EMODEL), the figure shows that this pose is also the one with the smallest RMSD to the experimental pose in 53% of the 23 cases. When we consider these 23 molecules from the original 'no water' docking run where the (i) ligand partial charges (3rd bar, denoted POSE(L)) and (ii) ligand and protein partial charges (5th bar, denoted POSE(C)) were explicitly polarized using the geometries of the 1st Ranked GSCORE pose, then flexible ligand re-docking results in the top ranked pose being the one with the smallest RMSD to the experimental pose in 48% and 44% of cases respectively. These three results of 53%, 48% and 43% are therefore the EMODEL benchmark results that can be used to see if a strongly bonded ligand attached 'lone water' molecule can improve ligand docking.

When we now include the 'lone water' molecule in the receptor grid for the 23 cases using   (i) the default OPLS 2005 partial charges (2nd bar, denoted GLIDE(3H)); (ii) the polarized ligand partial charges (4th bar, denoted POSE(L3H)); and (iii) the polarized ligand and protein partial charges (6th bar, denoted POSE(C3H)) that are both explicitly polarized using the geometries of the 1st Ranked GSCORE pose, then flexible ligand re-docking results in the top ranked pose being the one with the smallest RMSD to the experimental pose in (i) 70%, (ii) 61% and (iii) 52% of cases respectively. Under these circumstances, adding a 'lone

water' offers at best ~17% improvement over the default 'no water' Glide docking, when scored by EMODEL and polarization is not included, while polarization does not improve the docking results with a 'lone water', or the 'no water' results here.

When we now include the 'lone water' molecule with QM/MM orientated hydrogens in the receptor grid for the 23 cases using (i) default partial charges (7th bar, denoted GLIDE(3HO)), then we also use the (ii) ligand partial charges (8th bar, denoted POSE(L3HO)) and (iii) ligand and protein partial charges (9th bar, denoted POSE(C3HO)) that are explicitly polarized using the geometries of the 1st Ranked GSCORE pose, then flexible ligand re-docking results in the top ranked pose being the one with the smallest RMSD to the experimental pose in (i) 65%, (ii) 65% and (iii) 70% of cases respectively. The 'lone water' with QM/MM orientated hydrogen results match the previous 'lone water' with MM orientated hydrogen results when polarization of the ligand and protein partial charges is included, but are slightly worse otherwise.

The best results here are notably similar to the previous CVDW results, with the 'lone water' with QM/MM orientated hydrogen results performing the best on average, with polarization less important.

5.3.2.2.2 Top Ranked Pose Cluster (white bars): Summary

In brief consideration of the cluster analysis in Figure 5.1E when the first pose cluster measured by the criteria (see 3.2.10.2) that each cluster has an RMSD of ≤ 1 Å and energy within 1.1 kcal mol^{-1} of the lowest energy member of that pose cluster and the pose with the lowest RMSD to experiment should also be a member of the first cluster, there are only two cases with an improvement. These are using: (i) the 'no water' polarized ligand partial charges; and (ii) the 'lone water' with QM/MM orientated hydrogens and polarized ligand partial charges where success increases to (i) 65% and (ii) 74% for the 23 cases giving ~17% and ~9% improvement respectively.

As seen previously, the cluster analysis tends to offer less scope for improvement when alternative methods that also offer improvement have been incorporated.

5.3.3 RMSD analysis (GSCORE scoring)

In sections 5.3.1 and 5.3.2 and Figure 5.1A, 5.1C and 5.1E we considered the proportion of cases where the top ranked pose was also the one with the lowest RMSD to the experimental structure. However, in some of these cases the lowest RMSD might still be rather high. In Figures 5.1B, 5.1D and 5.1F we consider the proportion of molecules where the RMSD is below 0.5 Å (green); RMSD of ≤ 1 Å (yellow); RMSD of ≤ 2 Å (orange); RMSD of ≤ 4 Å(dark red) respectively.

Once again we consider the 23 cases where a ligand-attached water molecule with three or more non-water hydrogen bonds exists within our Dataset and the effects of adding these 'lone waters' to the target receptor grid on subsequent ligand docking by the above RMSD analysis criteria.

5.3.3.1 RMSD of ≤ 0.5 Å from experimental geometry (green bars)

In Figure 5.1B (1st bar, denoted GLIDE(23)) we consider the 'no water' 23 molecules across the above range from the original docking run for the top ranked pose (according to GSCORE), the figure shows that this pose has a distance measured by RMSD of ≤ 0.5 Å from the experimental pose in 17% of the 23 cases. When we also consider these 23 molecules from the original 'no water' docking run where the (i) ligand partial charges (3rd bar, denoted POSE(L)) and (ii) ligand and protein partial charges (5th bar, denoted POSE(C)) where explicitly polarized using the geometries of the 1st Ranked GSCORE pose, then flexible ligand re-docking results in the top ranked pose having an RMSD of ≤ 0.5 Å in 31% and 26 % of the cases, respectively. These three results of 17%, 31% and 26% are therefore the GSCORE benchmark results that can be used to see if a strongly bonded ligand attached 'lone water' molecule can improve ligand docking at the RMSD of ≤ 0.5 Å threshold.

When we now include the 'lone water' molecule in the receptor grid for the 23 cases using   (i) the default OPLS 2005 partial charges (2nd bar, denoted GLIDE(3H)); (ii) the polarized ligand partial charges (4th bar, denoted POSE(L3H)); and (iii) the polarized ligand

and protein partial charges (6th bar, denoted POSE(C3H)) that are both explicitly polarized using the geometries of the 1st Ranked GSCORE pose, then flexible ligand re-docking results in the top ranked pose having an RMSD of ≤ 0.5 Å in 48%, 43% and 34 % of the cases, respectively.

Under these circumstances, adding a 'lone water' offers at best ~31% improvement over the default 'no water' Glide docking when scored by GSCORE and polarization is not included, while polarization does not improve the docking results with a 'lone water', but improves over the rather poor 'no water' results here.

When we now include the 'lone water' molecule with QM/MM orientated hydrogens in the receptor grid for the 23 cases using (i) default partial charges (7th bar, denoted GLIDE(3HO)), then we also use the (ii) ligand partial charges (8th bar, denoted POSE(L3HO)) and (iii) ligand and protein partial charges (9th bar, denoted LOSE(C3HO)) that are explicitly polarized using the geometries of the 1st Ranked GSCORE pose then flexible ligand re-docking results in the top ranked pose having an RMSD of ≤ 0.5 Å in 48%, 34% and 43 % of the cases, respectively.

The 'lone water' with QM/MM orientated hydrogen results match the previous 'lone water' with MM orientated hydrogen results without polarization, giving again ~31% improvement over the 'no water' results at his threshold. However, on average there is no change for better or worse over the MM orientated hydrogen results.

Overall the 'lone water' results therefore show a significant effect in generating accurate poses with an RMSD of less than 0.5 Å. While polarization is only effective for the 'no water' result set at this threshold, when scored by GSCORE.

5.3.3.2 RMSD of ≤ 1 Å, 2 Å and 4 Å from experimental geometry

Overall the 'lone water' results in Figure 5,1show a significant effect in generating accurate poses with an RMSD of less than 1 Å, 2 Å and 4 Å. Polarization is generally effective for the 'no water' result set at these thresholds, particularly when using the ligand only partial charges.   However, consideration of the results at the 1 Å, 2 Å and 4 Å

thresholds shows there is no clear indication that polarization improves the results in the presence of a 'lone water'.

5.3.4   RMSD Analysis (CVDW rescoring)

The results of Figure 5.1B are again of interest as they suggest a large improvement can be obtained by including a lone water molecule with strong ligand and protein hydrogen bond interactions. When considering polarization previously, we have seen a marked improvement when rescoring with alternative protocols. For this reason, we have again considered alternative scoring protocols for ranking the poses, namely CVDW (Figure 5.1D) here, and EMODEL (Figure 5.1F) in section 5.3.5. Here CVDW rescoring is expected to provide the clearest effects of polarization as it uses a molecular mechanics force-field scoring function.

5.3.4.1   RMSD of ≤ 0.5 Å from experimental geometry (green bars)

In Figure 5.1D (1st bar, denoted GLIDE(23)) we consider the 23 'no water' molecules across the above range from the original docking run for the top ranked pose (according to CVDW), the figure shows that this pose has a distance measured by RMSD of ≤ 0.5 Å from the experimental pose in 43% of the 23 cases. When we also consider these 23 molecules from the original 'no water' docking run where the (i) ligand partial charges (3rd bar, denoted POSE(L)) and (ii) ligand and protein partial charges (5th bar, denoted POSE(C)) where explicitly polarized using the geometries of the 1st Ranked GSCORE pose, then flexible ligand re-docking results in the top ranked pose having an RMSD of ≤ 0.5 Å in 22% and 39 % of the cases, respectively. These three results of 43%, 22% and 39% are therefore the CVDW benchmark results that can be used to see if a strongly bonded ligand-attached 'lone water' molecule can improve ligand docking at the RMSD of ≤ 0.5 Å threshold.

When we now include the 'lone water' molecule in the receptor grid for the 23 cases using   (i) the default OPLS 2005 partial charges (2nd bar, denoted GLIDE(3H)); (ii) the polarized ligand partial charges (4th bar, denoted POSE(L3H)); and (iii) the polarized ligand and protein partial charges (6th bar, denoted POSE(C3H)) that are both explicitly polarized using the geometries of the 1st Ranked GSCORE pose, then flexible ligand re-docking results

in the top ranked pose having an RMSD of ≤ 0.5 Å in 57%, 61% and 39 % of the cases, respectively.

Under these circumstances, adding a 'lone water' offers at best ~18% improvement over the default 'no water' Glide docking when scored by CVDW and polarization of the ligand partial charge is included, while full polarization does not improve the docking results with a 'no water', polarization of the ligand partial charges gives ~4% improvement over the 'lone water' default charges result at this threshold.

When we now include the 'lone water' molecule with QM/MM orientated hydrogens in the receptor grid for the 23 cases using (i) default partial charges (7th bar, denoted GLIDE(3HO)), then we also use the (ii) ligand partial charges (8th bar, denoted POSE(L3HO)) and (iii) ligand and protein partial charges (9th bar, denoted POSE(C3HO)) that are explicitly polarized using the geometries of the 1st Ranked GSCORE pose then flexible ligand re-docking results in the top ranked pose having an RMSD of ≤ 0.5 Å in 48%, 52% and also 52% of the cases, respectively. The 'lone water' with QM/MM orientated hydrogens results offer no improvement over the 'lone water' with MM orientated hydrogen results, but adding polarization to either the ligand or ligand and protein partial charges gives ~4% improvement over the default charges when the 'lone water' hydrogens are QM/MM orientated. This takes the best improvement over 'no water' default glide partial charges to ~9 % improvement compared with the ~18% when using the 'lone water' hydrogens that are MM orientated at this threshold.

Overall the 'lone water' results therefore show a significant effect in generating accurate poses with an RMSD of less than 0.5 Å. Polarization in most also cases offers a slight improvement but only when applied to the 'lone water' results, with the exception of the MM orientated hydrogens using ligand partial charge result, when scored by CVDW.

5.3.4.2 RMSD of ≤ 1 Å from experimental geometry (yellow bars)

Overall the 'lone water' results therefore show a significant effect in generating accurate poses with an RMSD of less than 1 Å. Polarization also shows a significant effect

for the 'lone water' results, but only when the ligand partial charges are used for the MM orientated hydrogens result, and when the ligand and protein partial charges are used for the QM/MM orientated hydrogens result, when scored by CVDW. It appears that polarization of just the ligand is a relatively safe practice, but when QM/MM optimization of the water hydrogen atoms is included then polarization of the enzyme is more effective because the hydrogens are placed more appropriately.

5.3.4.3 RMSD of ≤ 2 Å from experimental geometry (orange bars)

Overall the 'lone water' results therefore show a significant effect in generating poses with an RMSD of less than 2 Å. Polarization shows a significant effect for the 'no water' results, but only when the ligand partial charges are used. Polarization also shows a significant effect for the 'lone water' results, but only when the ligand partial charges are used for the MM orientated hydrogens result, and when the ligand and protein partial charges are used for the QM/MM orientated hydrogens result, when scored by CVDW at this threshold.

5.3.4.4 Summary of RMSD of ≤ 4 Å from experimental geometry (dark red bars)

There are three best results that all achieve 96% success at this threshold these are the 'no water' result using polarized ligand partial charges (3rd bar, denoted POSE(L)), and both the QM/MM orientated hydrogen 'lone water' results where polarization is used for the ligand partial charges (8th bar, denote POSEL3HO)), and the ligand and protein partial charges (9th bar, denoted POSE(C3HO)). The last of these three results is notable as in these 96% of cases, all were also within the RMSD of ≤ 2 Å threshold.

 The two worst results at this threshold were both 83% for the 'no water' with default charges result (1st bar, denoted GLIDE(23)); and the 'lone water' MM orientated hydrogen result with polarization of the ligand partial charges (4th bar, denoted POSE(L3H)), the latter of which also had all of the 83% of cases within the RMSD of ≤ 2 Å threshold.

5.3.4.5 RMSD Analysis (EMODEL rescoring): Summary

Figure 5.1F shows that EMODEL gives similar results to scoring with CVDW (Figure 5.1D), and thus offers an alternative method of selecting the top ranked pose for polarizing the ligand and the enzyme. Again the results show distinct improvements when including the lone water. For some approaches at some thresholds, polarization offers some improvement, but the results are not consistent and the effect of including the additional water molecule is clearly dominant.
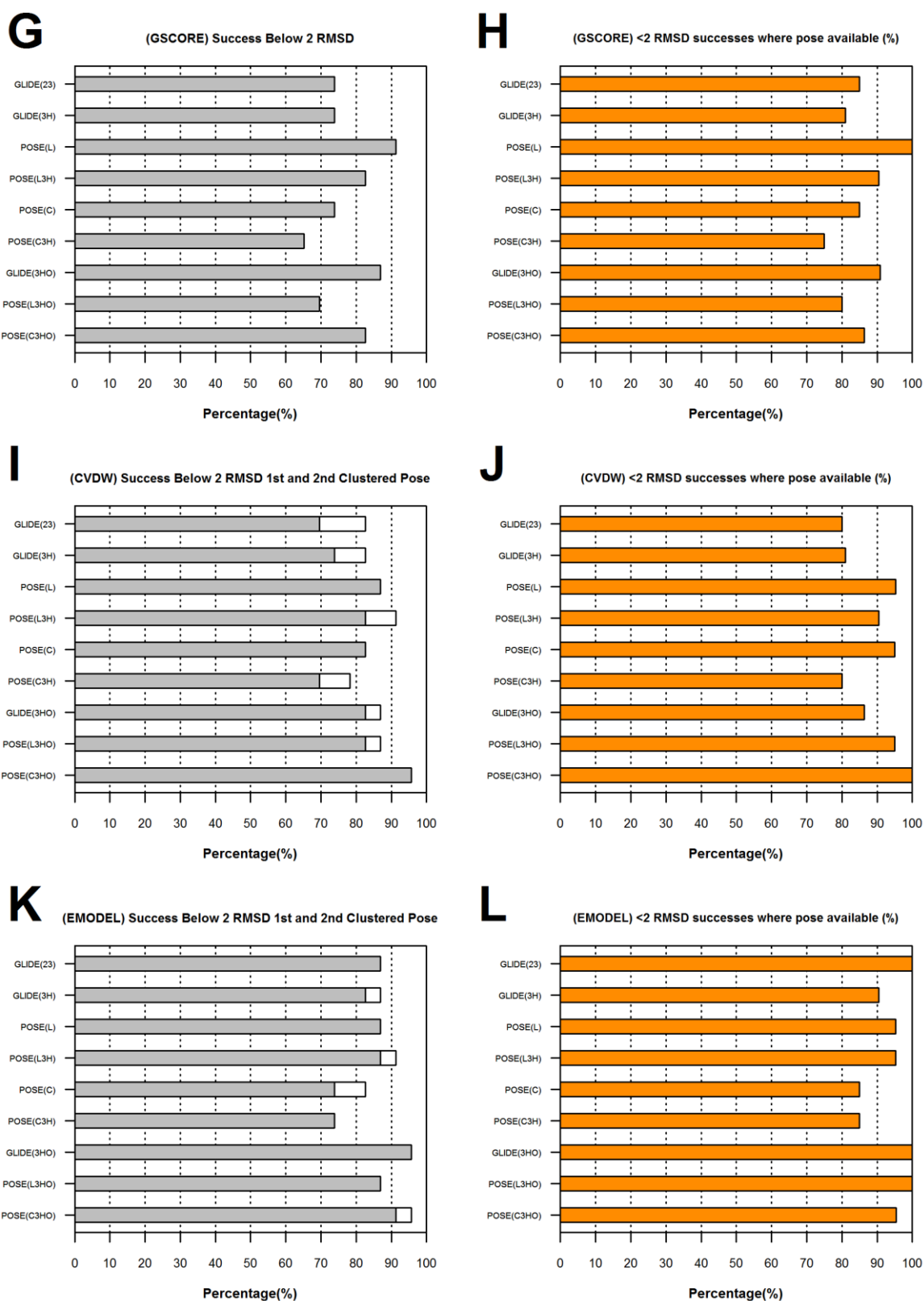
Figure 5.2. An (<2 Å RMSD) assessment of additional clustering on (requested 15 pose) docking success (as Figure 5.1) using Glide Standard Precision, denoted GLIDE, and where (23) refers to number of molecules from original dataset without single water molecule, (3H) refers to the minimum number of protein/ligand hydrogen bonds formed namely 3, (L) is Re-Docked with Polarized Ligand partial charges or (C) is Complex partial charges (C); (3HO)

is the same as (3H) but the single water molecule hydrogens were orientated by a QM/MM optimization before the docking run. POSE denotes that the pose was polarized at the Highest Ranking Glide GSCORE pose geometry. (CVDW), (EMODEL), (GSCORE) are the scoring functions used by Glide. Plots G, I, K – The percentage over 23 of the 74 molecules where the highest ranking/lowest scoring pose cluster (Grey) is a successful pose (below 2 Å RMSD). The % increase when including the 2nd Pose cluster (White) <2 Å RMSD (when 1st pose is > 2 Å) is also given. Plots H, J, L – The percentage over 23 of 74 molecules where the sampling error is ignored and a pose exists below 2 RMSD and the highest ranking/lowest scoring pose is a successful (below 2 Å RMSD - orange). Clustering parameters were set at <1.1 kcal mol^{-1} and <1 Å RMSD.

5.3.5   Cluster analysis.

In figures 5.1B, 5.1D and 5.1F we considered the proportion of molecules where the RMSD is below 0.5 Å (green); RMSD of ≤ 1 Å (yellow); RMSD of ≤ 2 Å (orange); RMSD of ≤ 4 Å (dark red) respectively.

In some cases, the lowest energy pose may have a rather high RMSD, but there may still be a pose with a low RMSD that is not too high in energy above the lowest energy pose. We consider this scenario in Figure 5.2. The RMSD of ≤ 2 Å threshold from 5.1B, 5.1D and 5.1F has been used as the benchmark threshold for testing due to its popularity (Cole et al., 2005, Warren et al., 2006, Sándor et al., 2010). These results can be seen again in Figures 5.2G, 5.2I and 5.1K (grey bars). As was first mentioned in 3.2.10.2, the poses have been clustered (generously) so that each cluster has an RMSD of ≤ 1 Å and energy within 1.1 kcal mol^{-1} of the lowest energy member of that pose cluster.   We now set the RMSD analysis criteria (see 3.3.4) that the lowest energy member of the 2nd pose cluster should also be considered for analysis at an RMSD of ≤ 2 Å, in the cases where the lowest energy member of the first pose cluster was docked unsuccessfully. Here we evaluate the improvement in the docking results when both clusters were considered by the above criteria, as shown in Figures 5.2I, 5.2K. The white bars indicate the higher percentage that comes from including the second cluster. To ensure that there is a significant probability (in terms of energy) of a pose adopting the 2nd pose cluster minima as opposed to the first pose minima, only poses that are members of a cluster with a sum greater than 10% Boltzmann probability have been considered.

5.3.5.1   Clustered CVDW Rescoring: Summary

GSCORE was not used for clustering in Figure 5.2, but GSCORE results are shown at the

RMSD of ≤ 2 Å threshold for reference in Figure 5.2G (Figure 5.2G is identical to Figure

5.1B, 2 Å threshold).

In 5.2I, when considering the 'no water' default partial charges result (1st grey bar, denoted

GLIDE(23)) for top ranked pose according to CVDW using Glide SP, the figure shows that

this pose has an RMSD below 2 Å in 70% of the 23 cases. When we also consider the

lowest energy member of the 2nd cluster at an RMSD of ≤ 2 Å, in the cases where first cluster

did not yield an RMSD < 2 Å, the success rate rises to 83% (1st bar – white, denoted

GLIDE(23)). This ~13% improvement is the largest shown from clustering, when scored by

CVDW in figure 5.2I. The next three largest ~9% improvements all come from the MM

orientated 'lone water' results using the default partial charges (2nd white bar, denoted

GLIDE(3H)); the polarized ligand partial charges (4th white bar, denoted POSE(L3H)); and

the polarized ligand and protein partial charges (6th white bar, denoted POSE(C3H)) taking

the success at the RMSD < 2 Å to 83%; 91%; and 78% respectively. Under these

circumstances, polarization offers no effect on clustering improvement and/or in the

distinguishing between two low lying pose clusters.   We see a reduced improvement of ~4%

from clustering the QM/MM orientated hydrogen 'lone water' results using the default partial

charges (7th bar, denoted GLIDE(3HO)); and the polarized ligand partial charges (8th green

ba, denoted POSE(L3HO)), which now offer ~87% success in both cases at this RMSD < 2 Å

threshold.   While both the 'no water' results using polarization for the ligand partial charges

(3rd bar, denoted POSE(L)); and the ligand and protein partial charges (5th bar, denoted

POSE(C)); and also the QM/MM orientated hydrogen 'lone water' result using the polarized

ligand and protein partial charges (9th bar, denoted POSE(C3HO)), offer no improvement

from clustering the results, remaining at 87%, 83%, and a rather high 96% success

respectively.

Overall again the margins for clustering improvement depend on there being a pose with a low RMSD that is not too high in energy above the lowest energy pose. These margins decrease in all cases when compared with the 'no water' default partial charges scored by CVDW, explicit polarization in both cases decrease these margins to zero for the 'no water' results. This is a good sign that polarization may be working properly in differentiating the correct pose from an energetically close geometric decoy. In the case of the 'lone water' results, the most significant factor in decreasing these margins appears to be the whether the hydrogens are orientated by QM/MM or MM energy minimization, with QM/MM energy minimization having the smallest margins from clustering of the two.

5.3.5.2 Clustered EMODEL rescoring: Summary

When EMODEL and clustering are used for scoring, Figure 5.2K shows there are three results that offer ~4% improvement for considering the 2nd pose cluster with a potential alternate low lying minima, two of which are from the MM orientated 'lone water' results, namely using the default partial charges (2nd white bar, denoted GILE(3H)) and the polarized ligand partial charges (4th white bar, denoted POSEL3H)). The third comes from the QM/MM orientated hydrogen 'lone water' result using the polarized ligand and protein partial charges (9th white bar, denoted POSE(C3HO)). The only other result that shows an improvement is the 'no water' using the polarized ligand and protein partial charges (5th white bar, denoted POSE(C)) which shows the largest margin of ~9% (~2 molecules) from considering the lowest energy member of the 2nd pose cluster, when the first pose cluster does not provide a correct answer within the below 2 Å RMSD threshold. This result is interesting in that in Chapter 3, when the 'no water' dataset is 74 cases instead of 23 cases there were no further low lying minima, that were energetically reasonable alternatives. Providing a cautionary message that size of the dataset has an effect on the patterns that the data reveals.

5.3.6 Analysis of RMSD of available ≤ 2 Å from experimental geometry

In Chapters 3 and 4 we considered the extent of the energetic viability of using a 2nd pose cluster. Here as this answer has been previously established as providing a reasonable

likelihood, we instead explore an alternative question.   As previously mentioned, docking search algorithms are considered to show a good level of reliability in producing a pose close to the experimental pose in a given number of pose solutions, but the scoring functions are generally considered to be not yet reliable or sophisticated enough to correctly rank the final poses on offer (Warren et al., 2006). With high levels of reliability ~70% in identifying the top ranked pose, and even higher success rates at the RMSD < 2 Å, threshold, we now re-assess the efficacy of the search algorithm versus the final ranking from the scoring protocols namely GSCORE, CVDW, EMODEL when the answer for a 'lone water' molecule with good hydrogen bond contacts is known, and polarization is included within the force-field. Although this is approaching an idealized situation this still does not consider the whole picture, e.g., there is still no protein flexibility and so forth.

5.3.6.1   RMSD of available poses ≤ 2 Å GSCORE: Summary

In Figure 5.2H, when we ignore the sampling error and consider just the cases where there is a pose with an RMSD < 2 Å within each of the requested ~15 pose solutions, and scored by GSCORE. The figure shows that there is one instance of the 'no water' result with polarized ligand partial charges (3rd bar, denoted POSE(L)) where the scoring function is successful in 100% at this threshold. The next best two results come from the MM orientated hydrogens 'lone water' result where the ligand partial charges have been polarized (4th bar, denoted POSE(L3H)); and the QM/MM orientated hydrogens 'lone water' result using the default partial charges (7th bar, denoted GLIDE(3HO)) these offer 90% and 91% success at this threshold respectively. The 'no water' result with default partial charges now offers 85% success (1st bar, denoted GLIDE(23), an ~11% increase over the initial docked result at this threshold. Overall the patterns although a few percent more pronounced, follow the patterns of the initial scoring results in that do not always provide an answer that approaches the experimental geometry. Nevertheless, improvements over the default GLIDE results come from either including a lone water or including polarization.

5.3.6.2 RMSD of available poses ≤ 2 Å CVDW: Summary

In Figure 5.2J, when we ignore the sampling error and consider just the cases where there is a pose with an RMSD < 2 Å within each of the requested ~15 pose solutions, and scored by CVDW. Then the figure shows, there is again one instance the QM/MM orientated hydrogens 'lone water' result with polarized ligand and protein partial charges (9th bar, denoted POSE(C3HO)) where the scoring function is successful in 100% at this threshold. The next three best results come from the 'no water' result with polarized ligand partial charges (3rd bar, pose(L)); the 'no water' result with polarized ligand and protein partial charges (5th bar, POSE(C)); and the QM/MM orientated hydrogens 'lone water' result with polarized ligand partial charges (8th bar, denoted POSE(L3HO)), these all offer ~95% success at this threshold.

 The MM orientated hydrogens 'lone water' result where the ligand partial charges have been polarized (4th bar, denoted POSE(L3H)) offers 90% success. The three worst results include the 'no water' result with default partial charges (1st bar, denoted GLIDE(23)); the 'no water' result with lone water (2nd bar, denoted GLIDE(3H)); and the MM orientated hydrogens 'lone water' result using the fully polarized partial charges (6th bar, denoted POSE(C3H)), these all offer ~80% success.

Overall, polarization offers the best results when a pose is available within the ~15 requested poses below 2 Å RMSD scored by CVDW. In these circumstances, polarization of the ligand partial charges fares better when the 'lone water' hydrogens are MM orientated, and polarization of the ligand and protein partial charges fares better when the 'lone water' hydrogens are QM/MM orientated, with both 'no water' polarization results also doing well.

5.3.6.3 RMSD of available poses ≤ 2 Å EMODEL: Summary

In Figure 5.2L, we ignore the sampling error and consider the cases where there is a pose with an RMSD < 2 Å within each of the requested ~15 pose solutions scored by EMODEL. The Figure shows that there are three instances where 100% success is achieved, these are the 'no water' result with default partial charges (1st bar, denoted GLIDE(23)); the QM/MM

orientated hydrogens 'lone water' results with default partial charges (7th bar, denoted GLIDE(3HO)), and with polarized ligand partial charges (8th bar, denoted POSE(L3HO)). There are also three results where EMODEL is successful in ~95% at this threshold. These are the 'no water' result with polarized ligand partial charges (3rd bar, denoted POSE(L)), the MM orientated hydrogens 'lone water' result with polarized ligand partial charges (4th bar, denoted POSE(L3H)) and the QM/MM orientated hydrogens 'lone water' result with polarized ligand and protein partial charges (9th bar, denoted POSE(C3HO)). Overall, EMODEL does best at scoring when a pose is available within this threshold, but the fact that the 'no water' default charge result gives 100% success is indicative that EMODEL is most directly related to posing and ranking of poses from the same ligand molecule (Banks et al., 2005). However, two of the QM/MM orientated hydrogens 'lone water' results with default partial charges (7th bar, denoted GLIDE(3HO)) and polarized ligand partial charges (8th bar, denoted POSE(L3HO)) were also able to achieve 100% success, while the QM/MM orientated hydrogens 'lone water' result with the polarized ligand and protein partial charges (9th bar, denoted POSE(C3HO)) achieved ~95% success. Also in Figure 5.2K the two best results were the QM/MM orientated hydrogens 'lone water' results regarding with default partial charges, and polarized ligand and protein partial charges (7th and 9th bar) outperforming EMODEL at the RMSD < 2 Å threshold, with the other QM/MM orientated hydrogens 'lone water' result using polarized ligand partial charges (8th bar) matching EMODEL. This indicates that any negative effects from tweaking the docking using EMODEL are cancelled by the improvements in docking success gained.

5.3.7 Docking of the 'Lone Water' molecule, top ranked pose analysis: Summary

In Figure 5.3A and Figure 5.3C, we consider the docking of the 'lone water' molecule with default partial charges, with the ligand and protein complex serving as the receptor grid unless otherwise stated. It has been shown elsewhere that the number of non water-water hydrogen bonds formed should be a major factor in producing the lowest binding energy (Barillari et al., 2007). In a simple test in Figure 5.3A, we consider the 'lone waters' with

three or more non water-water hydrogen bonds in the previous dataset range covered in 5.3.1 to 5.3.6, then dock the 'lone water' with and without the ligand in the receptor grid. This removal of the ligand obviously reduces the number of hydrogen bond interactions the 'lone water' can make at its current position within the receptor. When the ligand is present we can also assess the efficacy of a number of scoring protocols in recovering our 'lone water' pose. The figure shows that flexible 'lone water' docking results in the top ranked pose being the one with the smallest RMSD to the experimental pose more often when the ligand is present be it for GSCORE, CVDW, EMODEL, QM/MM single-point or the Optimized calculation than when the ligand is absent (Note that here the RMSD is for a single atom and so is equivalent to displacement – but RMSD is used for comparability to the other sections). The 'ligand present' highest success ~73% is for GSCORE (1st bar, denoted LIG(GS)); the next three best results of ~64% success come from CVDW (3rd bar, denoted LIG(CVDW)), EMODEL (5th bar, denoted LIG(EMOD)) and the QM/MM lowest optimized energy calculation (9th bar, denoted LIG, OQM)). The worst ligand present result of ~59% comes from the QM/MM lowest single point energy calculation (7th bar, denoted LIG(SQM)) when docking a 'lone water'. When the ligand is absent the best two results of ~50% success both come from QM/MM Energy calculations for the lowest single-point energy (8th bar, denoted NOLIG(SQM)) and the lowest optimized energy (10th bar, denoted noLIG(OQM)). The worst result of ~27% success in identifying the top ranked pose as the one with the smallest RMSD to experimental pose comes from GSCORE (2nd bar, denoted NOLIG(GS)).

Figure 5.3C gives further confirmation that the number of non water-water hydrogen bonds formed is the most significant factor in flexible 'lone water' docking resulting in the top ranked pose being the one with the smallest RMSD to the experimental pose. The figure shows that the best three results come from the 'lone waters' with two non water-water hydrogen bonds when scored by GSCORE (1st bar, denoted "H(GS)), CVDW (2nd bar, 2H(CVDW)) and EMODEL (3rd bar, denoted "H(EMOD)) offering 50%, 55% and 55% success respectively, with the worst result of ~23% coming from the single non water-water hydrogen bond result scored by GSCORE (6th bar, denoted 1H(GS)). The best single non

water-water hydrogen bond result is ~40% success from the QM/MM lowest optimized energy calculation (10th bar, denoted 1H(OQM)), the next three best single non water-water hydrogen bond results all of ~33% success come from CVDW (7th bar, denoted 1H(CVDW)), EMODEL (8th bar, denoted 1H(EMOD)) and the QM/MM lowest single-point energy calculation (9th bar, denoted 1H(OQM)).

Overall the best performance in flexible 'lone water' docking resulting in the top ranked pose being the one with the smallest RMSD to the experimental pose unsurprisingly come when there are more non water-water hydrogen bond interactions.

5.3.8 Docking of the 'Lone Water' molecule, RMSD analysis: Summary

In section 5.3.7 and Figures 5.3A and 5.3C, we considered the proportion of cases where the top ranked pose was also the one with the lowest RMSD to the experimental structure. However, in some of these cases the lowest RMSD might still be rather high. In figures 5.3B and 5.3D we consider the proportion of molecules where the RMSD is below 0.5 Å (green bars); RMSD of ≤ 1 Å (yellow bars); RMSD of ≤ 2 Å (orange bars); RMSD of ≤ 4 Å (dark red bars) respectively. However unlike ligand docking, the RMSD analysis is only meaningful up to the RMSD of ≤ 1 Å threshold as the docking of water molecules at larger distances would indicate alternate pose minima being adopted.

In Figure 5.3B when we consider the more significant RMSD of ≤ 0.5 Å, and RMSD of ≤ 1 Å thresholds, the Figure simply shows that flexible 'lone water' docking results in the top ranked pose being within these thresholds more often when the ligand is present, be it for GSCORE, CVDW, EMODEL or QM/MM single-point or the QM/MM Optimized calculations than when the ligand is absent. The three best results with the ligand present are for GSCORE (1st bar, denoted LIG(GS)); CVDW (3rd bar, denoted LIG(CVDW)) and EMODEL (5th bar, denoted LIG(EMOD)) at the RMSD of ≤ 1 Å threshold offering ~73%, ~73% and ~64% success respectively. Most of these 1st Ranked Poses are also within the RMSD of ≤ 0.5 Å threshold, which is to be expected considering the water molecule's small size offering ~64%; also ~64%; and 59% success respectively. Both the QM/MM calculation results with

the ligand present fare only slightly worse overall, with the lowest optimized energy result (9th bar, denoted LIG(OQM)) being the best of the two, particularly at the RMSD of ≤ 0.5 Å threshold with ~59% success. The best results when the ligand is absent come from the QM/MM calculation results for the lowest single-point energy (8th bar, denoted NOLIG(SQM)) and lowest optimized energy (10th bar, denoted NOLIG(OQM)); these offer ~27% and ~36% success respectively within the RMSD of ≤ 0.5 Å threshold, and offer ~45% and ~41% success respectively within the RMSD of ≤ 1 Å threshold. The worst result is for GSCORE when the ligand is absent with ~14% success at the RMSD of ≤ 0.5 Å and the RMSD of ≤ 1 Å thresholds.

In Figure 5.3D, for water molecules making fewer than 3 hydrogen bonds, we consider the more significant RMSD of ≤ 0.5 Å and RMSD of ≤ 1 Å thresholds, the figure simply shows that flexible 'lone water' docking results in the top ranked pose being within these thresholds more often when the 2 hydrogen bonds are formed, be it for GSCORE, CVDW, EMODEL, QM/MM single-point or the QM/MM optimized calculation, than when only 1 hydrogen bond is formed (where essentially there is no worthwhile docking success rate). In brief, the three best results arise in the 2 non water-water hydrogen bond results for GSCORE (1st bar, denoted 2H(GS)), CVDW (2nd bar, denoted 2H(CVDW)) and EMODEL (3rd bar, denoted 2H(EMOD)) at the RMSD of ≤ 0.5 Å threshold offering ~30%; also ~34%; and ~34% success respectively. There is no further success from increasing the threshold to an RMSD of ≤ 1Å.

Overall, by removing the ligand, the active site becomes less polar and there is a reduction in the number of hydrogen bonds that can be formed in the 'lone water' molecule's initial experimental minima (by ~1-2), regardless of hydrogen orientation. This in effect makes the lone water 'unhappy', making it far more likely to prefer a more energy favourable minima (from the newly returned ~15 poses) where more hydrogen bonds can be formed within the range of the posing space (Congreve et al., 2011). This 'unhappiness' also extends to the situation where the number of non water-water hydrogen bonds are reduced from three, to

two, and then to one (where the one is formed with the ligand only) respectively. While on the other hand, reasonable success can be achieved in 'lone water' water molecule docking, when the water can form at least 3 hydrogen bonds within its initial experimental minima. This means it is far more likely to return to its 'happy' place, regardless of scoring function and result in a lower RMSD to the experimental pose. So again the best performance in flexible 'lone water' docking resulting in the top ranked pose being the one with the smallest RMSD to the experimental pose (as with the previous section 5.3.7) unsurprisingly comes when there are more non water-water hydrogen bond interactions.
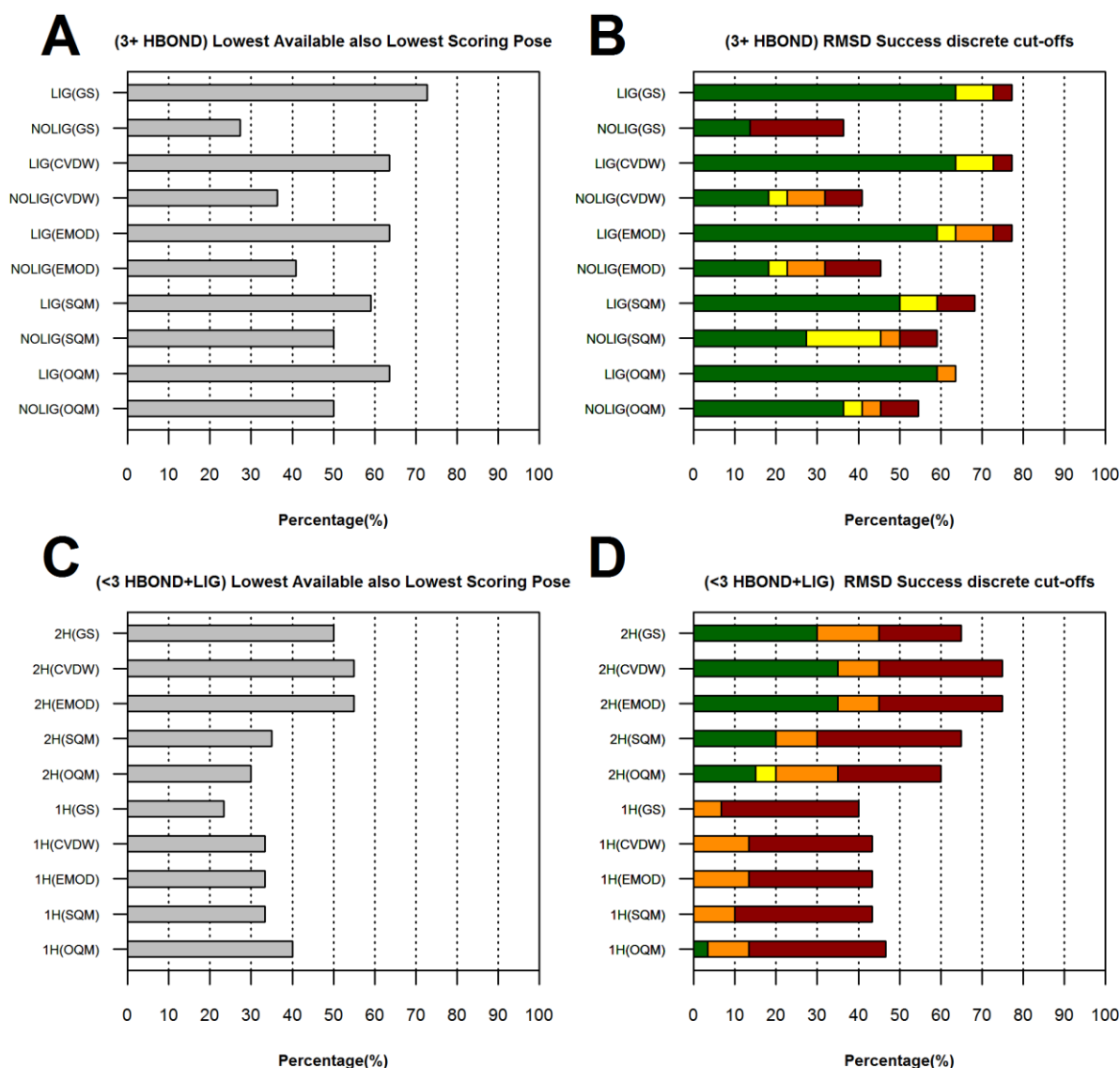
Figure 5.3 Assessment of 'Lone Water' flexible docking success (15 poses) using Glide with QM/MM Re-scoring. In plots A and B, the Water forms 3+ non water-water hydrogen bonds (at least 1 with protein and ligand), and LIG indicates the ligand is present; and NOLIG indicates the water is docked after the ligand is removed (thus reducing the number of hydrogen bonds accordingly). In plots C and D, 2H shows the water forms 2 non water-water hydrogen bonds (1 with protein and ligand); 1H refers to the water molecule only forming 1 hydrogen bond with the ligand prior to docking. In General brackets relate to the scoring method where (GS) is scored by GSCORE; (CVDW) is scored by CVDW; (EMOD) is scored by EMODEL; (SQM) is scored as lowest single point QM/MM Energy; and (OQM) is scored as lowest optimized QM/MM Energy. Plots A, C – Percentage from the 3H+; 2H; and 1H waters in the 22*; 21; and 30 of the original 74 cases respectively where the highest ranked/lowest scoring pose (Grey) is also lowest available pose. Plots B, D – Percentage over the same ranges of cases where the highest ranked/lowest scoring pose is at discrete cut-offs <0.5 Å RMSD (Green);<1 Å RMSD (Yellow); <2 Å RMSD (Orange); <4 Å RMSD (Red). *Omitted docking 1 of the water molecules from 3H+ waters dataset unintentionally.

5.4 Discussion

Active site water molecules clearly have an important role in ligand binding and so their treatment is an important factor in drug design (De Beer et al., 2010). Glide does take into account water to some extent (Banks et al., 2005, Friesner et al., 2006) but there are clear benefits for including explicit waters rather than treating water implicitly, as in Glide or continuum methods (Fogolari et al., 2002, Labute, 2008, Loving et al., 2010, Lie et al., 2011). The improvement in docking observed through inclusion of specific water molecules, as described in this chapter, is far more significant than that obtained through inclusion of polarization, as described in Chapter 3. Such observations on the importance of explicit water molecules in docking have been recorded by others (Corbeil and Moitessier, 2009, Thilagavathi and Mancera, 2010), but we understand this is the first report on the relative importance of explicit waters and polarization. Indeed, (Lie et al., 2011, Lemmon and Meiler, 2013) found improvements from the *simultaneous* docking of ligand and water, but implementation of this approach within our polarization framework would not be straightforward. For example, methods that simultaneously dock water and ligand (Verdonk et al., 2005) may have penalties for the annihilation of water and this could possibly be handled by a fully classical version of our method, but not by the current method that

requires QM calculations. The polarization effect introduced in this chapter is swamped by a far greater effect of the water, and so the improvement due to polarization, though generally present, is not always apparent. It is worth considering why inclusion of specific water molecules makes such a profound difference. The most obvious trivial reason is that the steric volume of the water reduces the search space, giving the docking process a greater chance of success by directing the poses more towards the correct binding site. This is particularly true for fragments as the water molecule can significantly increase the small number of hydrogen bond contacts between the fragment and its target. In addition, it provides a significant electrostatic steer to the docking process. Polarization generally only increases the magnitude of the interacting charges (and hence the resultant electrostatic energy) by about 10-15%, but the introduction of a polar water molecule has a far greater effect due to the large magnitude of the extra charges ($qH = 0.417$, $qO = -0.834$ for TIP3P water (Mark and Nilsson, 2001)). Since these charges are large, they will enhance the steric effect – as long as the water is oriented appropriately.

The positioning and orientation of water molecules is clearly a major problem since usually only a small proportion of the water molecules in a given binding site are visible. Water molecules may be absent in low resolution structures, but they might also be absent because of their dynamic nature. Moreover, the water pattern for one bound ligand may be different to that in other bound ligands. Thus, not all water molecules are the same. We have only considered significant water molecules that make at least 3 hydrogen bonds to the ligand and or protein, when docking ligands. An alternative approach may be to include water molecules that are conserved across a number of structures, since this is an alternative way to identify water molecules that can enhance docking success (Thilagavathi and Mancera, 2010). Indeed, it may be beneficial to investigate whether the lack of conservation in such structures indicates an unhappy water that should be omitted. However, such an investigation is outside the scope of this thesis. However, in the growing stage of a fragment-based drug design program, these water molecules may be under study for replacement. We

have shown that the water molecule can be placed with some degree of reliability simply by docking when the ligand is present, and the initially good hydrogen bonding conditions are met. This may offer a reasonable strategy as the X-ray crystal structure is indeed likely to have a ligand present and the generation of X-ray structures is often a key part of a FBDD program. The presence of such a water molecule could then enhance virtual screening programs for the purpose of finding other good fragment hits. We have shown that docking a water molecule in the absence of a ligand is unlikely to be successful.

Whether polarization helps to improve the docking in the presence of a water molecule is open to debate, as improvement is not reported in all cases studied. However, some general principles seem to emerge. If the water orientation is optimized by MM methods, then polarization of the ligand is generally more successful than polarization of the ligand and the protein. However, if the water orientation is optimized by QM/MM methods then polarization of the ligand and the protein is generally more successful. This is in contrast to the work of Roberts and Mancera, as they observed that the method of water optimization was not significant (Roberts and Mancera, 2008). One possible conclusion from this is that if the water is not correctly orientated then the protein may not be correctly polarized and so omission of protein polarization is preferable. It would seem that QM/MM orientation of the water molecule is preferable. This raises an important issue with regards to treatment of hydration and polarization within binding sites. One of the issues is that polarization in protein–ligand – specific water systems may be ligand dependent (Sahai and Biggin, 2011). Our results on the benefit of QM/MM orientation are in conflict with those of Roberts and Mancera (Roberts and Mancera, 2008), since they found that the water optimization method didn't have a large effect - but they did not include subtle polarization effects, or ensure the initial number of hydrogen bonds formed was at least three (done here by visualization after protein preparation steps). We will briefly discuss a number of methods for finding water binding sites but none of them rely on quantum mechanics, but it is clearly important to not only place the water molecule correctly but also to orientate it correctly.

A number of methods for placing water molecules in enzyme binding sites have been described. GRID was probably one of the earliest and most successful methods (Goodford, 1985). GRID is based on a set of molecular mechanics probes that interact with the enzyme over a 3D grid, it forms the basis of the waterflap method that is based on a rather sophisticated 'DRY' water probe (Mason et al., 2013) and has yielded good results on water networks (Cappel et al., 2011). The method can be used to find 'happy' and 'unhappy' water molecules and these can be exploited in drug design. Watermap from Schrödinger is an alternative to waterflap, and involves short molecular dynamics simulations to find water molecules that have a given residency at particular positions (Wang et al., 2011). Other methods based on MD or Monte Carlo methods can also be used (Woods et al., 2011, Bodnarchuk et al., 2014). Conservation of water sites across homologous proteins or across identical proteins containing different ligands is another approach to finding ligands that are significant for both ligand interaction and discrimination between ligands. Such methods offer great potential for identifying the small number of key water molecules that can profitably be included in docking experiments.

5.5 References

BANKS, J. L., BEARD, H. S., CAO, Y., CHO, A. E., DAMM, W., FARID, R., FELTS, A. K., HALGREN, T. A., MAINZ, D. T. & MAPLE, J. R. 2005. Integrated modeling program, applied chemical theory (IMPACT). *Journal of computational chemistry,* 26**,** 1752-1780.

BARILLARI, C., TAYLOR, J., VINER, R. & ESSEX, J. W. 2007. Classification of water molecules in protein binding sites. *Journal of the American Chemical Society,* 129**,** 2577-2587.

BERMAN, H., HENRICK, K. & NAKAMURA, H. 2003. Announcing the worldwide protein data bank. *Nature Structural & Molecular Biology,* 10**,** 980-980.

BODNARCHUK, M., VINER, R., MICHEL, J. & ESSEX, J. W. 2014. Strategies to calculate water binding free energies in protein-ligand complexes. *Journal of chemical information and modeling*.

CAPPEL, D., WAHLSTRÖM, R., BRENK, R. & SOTRIFFER, C. A. 2011. Probing the Dynamic Nature of Water Molecules and Their Influences on Ligand Binding in a Model Binding Site. *Journal of chemical information and modeling*.

COLE, J. C., MURRAY, C. W., NISSINK, J. W. M., TAYLOR, R. D. & TAYLOR, R. 2005. Comparing protein–ligand docking programs is difficult. *Proteins: Structure, Function, and Bioinformatics,* 60**,** 325-332.

CONGREVE, M., CHESSARI, G., TISI, D. & WOODHEAD, A. J. 2008. Recent Developments in Fragment-Based Drug Discovery. *Journal of Medicinal Chemistry,* 51**,** 3661-3680.

CONGREVE, M., LANGMEAD, C. J., MASON, J. S. & MARSHALL, F. H. 2011. Progress in structure based drug design for G protein-coupled receptors. *Journal of medicinal chemistry,* 54**,** 4283-4311.

CORBEIL, C. R. & MOITESSIER, N. 2009. Docking ligands into flexible and solvated macromolecules. 3. Impact of input ligand conformation, protein flexibility, and water molecules on the accuracy of docking programs. *Journal of chemical information and modeling,* 49**,** 997-1009.

DE BEER, S., VERMEULEN, N. P. E. & OOSTENBRINK, C. 2010. The role of water molecules in computational drug design. *Current topics in medicinal chemistry,* 10**,** 55-66.

FAVIA, A. D., BOTTEGONI, G., NOBELI, I., BISIGNANO, P. & CAVALLI, A. 2011. SERAPhiC: A benchmark for in silico fragment-based drug design. *Journal of chemical information and modeling,* 51**,** 2882-2896.

FOGOLARI, F., BRIGO, A. & MOLINARI, H. 2002. The Poisson–Boltzmann equation for biomolecular electrostatics: a tool for structural biology. *Journal of Molecular Recognition,* 15**,** 377-392.

FRIESNER, R. A., MURPHY, R. B., REPASKY, M. P., FRYE, L. L., GREENWOOD, J. R., HALGREN, T. A., SANSCHAGRIN, P. C. & MAINZ, D. T. 2006. Extra precision glide: docking and scoring incorporating a model of hydrophobic enclosure for protein-ligand complexes. *Journal of medicinal chemistry,* 49**,** 6177-6196.

GOODFORD, P. J. 1985. A computational procedure for determining energetically favorable binding sites on biologically important macromolecules. *Journal of Medicinal Chemistry,* 28**,** 849-857.

HUNG, A. W., SILVESTRE, H. L., WEN, S., CIULLI, A., BLUNDELL, T. L. & ABELL, C. 2009. Application of fragment growing and fragment linking to the discovery of inhibitors of Mycobacterium tuberculosis pantothenate synthetase. *Angewandte Chemie,* 121**,** 8604-8608.

LABUTE, P. 2008. The generalized Born/volume integral implicit solvent model: estimation of the free energy of hydration using London dispersion instead of atomic surface area. *Journal of computational chemistry,* 29**,** 1693-1698.

LEMMON, G. & MEILER, J. 2013. Towards ligand docking including explicit interface water molecules. *PloS one,* 8**,** e67536.

LIE, M. A., THOMSEN, R., PEDERSEN, C. N. S., SCHIØTT, B. & CHRISTENSEN, M. H. 2011. Molecular docking with ligand attached water molecules. *Journal of chemical information and modeling*.

LOVING, K., ALBERTS, I. & SHERMAN, W. 2010. Computational Approaches for Fragment-Based and De Novo Design. *Current Topics in Medicinal Chemistry,* 10**,** 14-32.

MADDEN, T. L., TATUSOV, R. L. & ZHANG, J. 1996. [9] Applications of network BLAST server. *Methods in enzymology,* 266**,** 131-141.

MARK, P. & NILSSON, L. 2001. Structure and dynamics of the TIP3P, SPC, and SPC/E water models at 298 K. *The Journal of Physical Chemistry A,* 105**,** 9954-9960.

MASON, J., BORTOLATO, A., WEISS, D., DEFLORIAN, F., TEHAN, B. & MARSHALL, F. 2013. High end GPCR design: crafted ligand design and druggability analysis using protein structure, lipophilic hotspots and explicit water networks. *Silico Pharmacol,* 1**,** 23.

ROBERTS, B. C. & MANCERA, R. L. 2008. Ligand-protein docking with water molecules. *Journal of chemical information and modeling,* 48**,** 397-408.

SAHAI, M. A. & BIGGIN, P. C. 2011. Quantifying Water-Mediated Protein–Ligand Interactions in a Glutamate Receptor: A DFT Study. *The Journal of Physical Chemistry B*.

SÁNDOR, M., KISS, R. & KESERŰ, G. R. M. 2010. Virtual fragment docking by Glide: A validation study on 190 protein– fragment complexes. *Journal of chemical information and modeling,* 50**,** 1165-1172.

THILAGAVATHI, R. & MANCERA, R. L. 2010. Ligand– Protein cross-docking with water molecules. *Journal of chemical information and modeling,* 50**,** 415-421.

VERDONK, M. L., CHESSARI, G., COLE, J. C., HARTSHORN, M. J., MURRAY, C. W., NISSINK, J. W. M., TAYLOR, R. D. & TAYLOR, R. 2005. Modeling water molecules in protein-ligand docking using GOLD. *Journal of medicinal chemistry,* 48**,** 6504-6515.

WANG, L., BERNE, B. & FRIESNER, R. 2011. Ligand binding to protein-binding pockets with wet and dry regions. *Proceedings of the National Academy of Sciences,* 108**,** 1326-1330.

WARREN, G. L., ANDREWS, C. W., CAPELLI, A. M., CLARKE, B., LALONDE, J., LAMBERT, M. H., LINDVALL, M., NEVINS, N., SEMUS, S. F. & SENGER, S. 2006. A critical assessment of docking programs and scoring functions. *Journal of medicinal chemistry,* 49**,** 5912-5931.

WOODS, C. J., MALAISREE, M., HANNONGBUA, S. & MULHOLLAND, A. J. 2011. A water-swap reaction coordinate for the calculation of absolute protein–ligand binding free energies. *The Journal of Chemical Physics,* 134**,** -.

YOUNG, T., ABEL, R., KIM, B., BERNE, B. J. & FRIESNER, R. A. 2007. Motifs for molecular recognition exploiting hydrophobic enclosure in protein–ligand binding. *Proceedings of the National Academy of Sciences,* 104**,** 808.

6 Addressing the Cross docking problem through MM polarization

6.1 Introduction

Virtual screening, an integral part of many drug design programmes, is driven by the conflicting requirements of being able to dock large databases of compounds in a short time while at the same time generating meaningful results; given the constant drive for improvements in this field (Yuriev et al., 2015), we have addressed the potential of polarization for delivering significant improvements. The docking problem can be broken down into one of (a) sampling, in which multiple ligand poses are generated within the enzyme or receptor binding site and (b) scoring, in which the poses are ranked. Much progress has been made in the quality of these docking programs (Sousa et al., 2013) such as GOLD (Jones et al., 1995), Glide (Friesner et al., 2004b), Autodock (Morris et al., 1996), FlexX (Kramer et al., 1999) and Dock (Allen et al., 2015) and impressive performances have been recorded by the latest versions. Thus, the self-docking problem is solved to some degree, with programs usually able to correctly dock a ligand back into its own protein X-ray structure reasonably well most of the time (Warren et al., 2006). Indeed, the results of Chapter 3 confirm this favourable situation for self-docking. However, the problem of docking a ligand back into its own protein where the structure was crystallized in the presence of a different ligand (cross-docking) remains considerably more challenging (Sandor et al., 2010, Morris et al., 2009, Liu et al., 2013b). This is partly because of the requirement of keeping the protein rigid or semi-rigid in order to process a huge number of (flexible) ligands in a short time, an approximation that may be ameliorated through the use of soft potentials. Often this rigidity can be partially addressed by using an ensemble of structures (Yuriev and Ramsland, 2013), either from modelling (Rueda et al., 2009, Moroy et al., 2015) or from crystallography where docking results can be improved by cross-docking into a range of structures and selecting the optimal X-ray structure for each target (Sandor et al., 2010). However, such approaches may be undesirable in a real-life structure-based drug design context (Gleeson

and Gleeson, 2009). What is generally not addressed in docking is the rigidity of the electronic distribution, though there are exceptions (Illingworth et al., 2008, Cho et al., 2005, Gleeson and Gleeson, 2009, Liu et al., 2013a). Thus, polarization of the ligand and protein is usually not included, and this may exacerbate the steric repulsion in cases where there is also electrostatic repulsion that cannot be fully alleviated by geometric rearrangements. Polarization of the ligand is however included in the quantum polarized docking method (Cho et al., 2005), and our approach (Illingworth et al., 2008) extends this by also including polarization of the protein. Indeed, in the absence of polarization, electrostatic repulsions could be alleviated by inappropriate geometric rearrangements. With this in mind, we have addressed the question as to whether inclusion of polarization can give rise to improved results in the cross-docking problem, since cross-docking, rather than self-docking, is more related to the real-world issue of virtual screening in drug design. This is particularly true in cases where resistant mutations can arise (Allen et al., 2015).

6.2 Methods

Protein choice. The initial search for proteins of identical or near-identical sequences with different ligands was based on the 12 complexes described in (Congreve et al., 2008b) and the 58 complexes described in (Favia et al., 2011) (excluding those with pdb codes 1YKI, 2P10, 2QWX and 3DSX where the binding site resides interfaced between 2 chains). Each sequence was BLASTED against the pdb sequence database using default parameters. Hits of a given sequence were accepted if there were (a) at least 7, (b) there was a ligand that roughly met Lipinski's rules (Lipinski et al., 2001), and the percentage identity was at least 99% and the coverage of the sequence was at least 97%. At the beginning of this investigation, drug like properties were not the most pressing consideration, as we looked to study a large number of unique proteins. Consequently, we allowed molecular weight (MW) <600 Da instead of <500 Da to increase the number of proteins with at least 7 ligands, this produced 196 ligands out of a total of 257 that strictly met Lipinski's rules of 5, and many of the remaining ligands only had one MW Lipinski rule violation (Lipinski et al., 2001), from the

initial list of 66 proteins, 17 proteins (15 unique) gave hits that met these criteria. Details on these 15 initial proteins are given in Table 6.1 and Appendix Table D.7. The 15 initial protein ligands and their associated hit ligands (from the total 257 proteins) are given in Appendix Table D.8. The folds of these proteins are shown in Figure 6.1.

Table 6.1. Overview of proteins used in cross docking.

Name	Pdb code	Group size	# cross-docking[a]	Mean resolution[b]	%age of proteins with clashes[c]	Mean # of C-C clashes[d]
Herpes simplex virus type 1	1e2i	9	72	2.2	51.4	1.8±1.2
Aldose reductase	1pwm	24	552(528)	1.5	55.6	9.4±6.4
Progesterone receptor	1sqn	7	42(29)	1.9	66.7	4.4±4.4
Isocitrate dehydrogenase	1t0l	8	56	2.6	51.8	8.1±13.8
Farnesyl diphosphate synthase	1yv5	10	90	1.9	18.9	1.2±2.8
methionine aminopeptidase	2gg7	10	90	1.7	63.3	4.4±3.7
microurokinase	1fv9	15	210	1.9	32.8	1.8±0.9
Oestrogen receptor	1gwq	36	1260 (1075)	2.3	53.3	5.2±5.7
Dipeptidyl peptidase IV	1n1m	14	182	2.5	37.9	3.3±3.8
nitric oxide synthase	1qwc	17	272	2.1	44.1	1.9±1.0
Cyclin-dependent kinase 2	1wcc	34	1122	2.2	73.0	4.6±3.4
tRNA-Guanine transglycosylase	1s39	15	210	1.7	16.7	1.2±1.9
Thrombin	2c90	8	56	2.0	23.3	2.0±2.0
HSP90	2jjc	30	870(866)	1.8	49.2	7.7±10.0
β-secretase 1	2ohk	20	380	2.2	70.3	6.3±7.0

ATP was included as a ligand for farnesyl diphosphate synthase.

a The number in parenthesis indicates the number of successful dockings in cases where there were unsuccessful dockings (226). b All proteins had a resolution of less than 3.0 Å; the majority had a resolution under 2.0 Å. c Percentage of proteins with C-C clashes (< 3 Å) between protein A and fitted ligand B. d Number of C-C clashes between protein A and fitted ligand B

Protein preparation. Commonly occurring cofactors and ions were retained in the complexes assessed on a protein by protein basis. The complexes were processed using the protein preparation wizard of Maestro in order to add hydrogens, assign bond orders, cap the protein ends with –NCOCH$_3$ and NHCH (except for short peptide ligands) and remove waters. Amino acid pKas were determined using PropKa 3.0 (Sondergaard et al., 2011, Olsson et al., 2011) at pH=7 and side chains adjusted accordingly. The proteins were then minimized using the OPLS 2005 force field (Kaminski et al., 2001, Banks et al., 2005a) (restrained to remain with 0.3 Å of the X-ray structure). Each set of the 15 groups of proteins were structurally aligned using the Maestro protein structure alignment facility; this enabled the coordinates of each ligand to be merged into the coordinates of every other protein structure in the set (see below). The initial protein complexes (Table 6.1) and an associated (pblast hit) complex, with an 'average' ligand is shown in Figure 6.2 and Figure 6.3. This 'average' ligand is identified using the Tanimoto index (Tanimoto, 1957, Bajusz et al., 2015), where the similarity of each ligand relative to the others in its subset was individually assessed. The structures forming the associated complexes are selected based on proximity to the mean Tanimoto index.

A docking grid input file was created for each aligned protein structure using its receptor text file, as the accompanying grid file template (exported without ligand and renamed from Maestro). An average Cartesian coordinates ligand centroid was determined for the structures of each of the 15 protein sets. These centroids were added to the input files, and the grids were generated in a batch (257 grids in total at this initial stage). The same grid sizes were used throughout. The inner grid was 14 Å × 14 Å × 14 Å and the outer grid was 34 Å × 34 Å × 34 Å. The ligands from each of the prepared structurally aligned co-crystallized (self) protein structures, were exported into a text file folder. For each ligand a docking input file was prepared for each of its protein's using (1) the ligand text file, (2) the native/ self ligand text file (for RMSD analysis reference), and (3) the native protein grid. This produced 257 self docking input files, and 5238 cross docking input files following a

'<pdb>_in_<pdb>' naming pattern for the ligand and receptor respectively. These were again run in batch generating report files for the 5495 docking experiments (226 cross docks failed from a possible 5721 docking experiments because Glide was not able to find any valid poses).

The merged structures that were used for polarizing the protein and the ligand at a given geometry, were generated after the ligands were docked. The merging took advantage of a Schrodinger script pv_convert.py that uses the docking output pose viewer text file from the 5495 experiments. The script creates a new complex (maestro) text file from the receptor, and by default, the first appearing/ top ranking pose (by Gscore). So, for each of the 5238 cross receptors and 257 self receptors, a temporary copy of the pose viewer file was created. The first ligand poses from each file were substituted (using text place markers e.g, :::) with each of its protein's ligands at the co-ordinates of their co-crystallized position. In the cross receptors, the protein-ligand complexes weren't primed together in protein preparation wizard, and they only shared approximately the same global co-ordinates from the protein structural alignment. The native complex structures were successfully tested against the maestro complexes for the appropriateness of using this pv_convert.py method. A Needleman Wunsch algorithm was used to compare the sequences of the ligand pockets (defined as a residue with at least one atom within 4 Å of the ligand) to check for mutations and to check the RMSD of the backbone superpositions of the residues in the ligand binding pockets (c.f. Table 6.1).

Initial Self-docking and cross-docking. The ligands were docked using flexible Glide SP (2015) (with a rigid protein) into each relevant protein structure, giving 5238 cross-docked structures (from a possible 5464, with 226 docking failures) and 257 self-docked structures. By comparison, the previous study by Sandor et al. used 8 complexes and 63 structures (Sandor et al., 2010). The OPLS 2005 force field was used for both the ligand and the protein, but for atoms with a partial charge smaller than 0.15 the van der Waals radii were reduced by 80% to ameliorate the approximation of rigid docking [Virtual Screening workflow,

Schrodinger LLC, New York 2007]. For each complex, 15 unique poses were determined,

which were then minimized within Glide to yield 15 poses (or fewer if Glide was unable to find

15 poses). The top-scoring pose was then selected using the GLIDE scoring functions,

namely Gscore, cVDW and Emodel. Gscore is most appropriate for comparing different

ligands, as in a virtual screen. Emodel is the best model for comparing the same ligand in the

different poses and is a combination of Gscore, the force field and internal ligand strain

energy. cVDW is essentially the force field (Banks et al., 2005b). All three rescoring methods

gave similar results and so only the cVDW results are reported below.
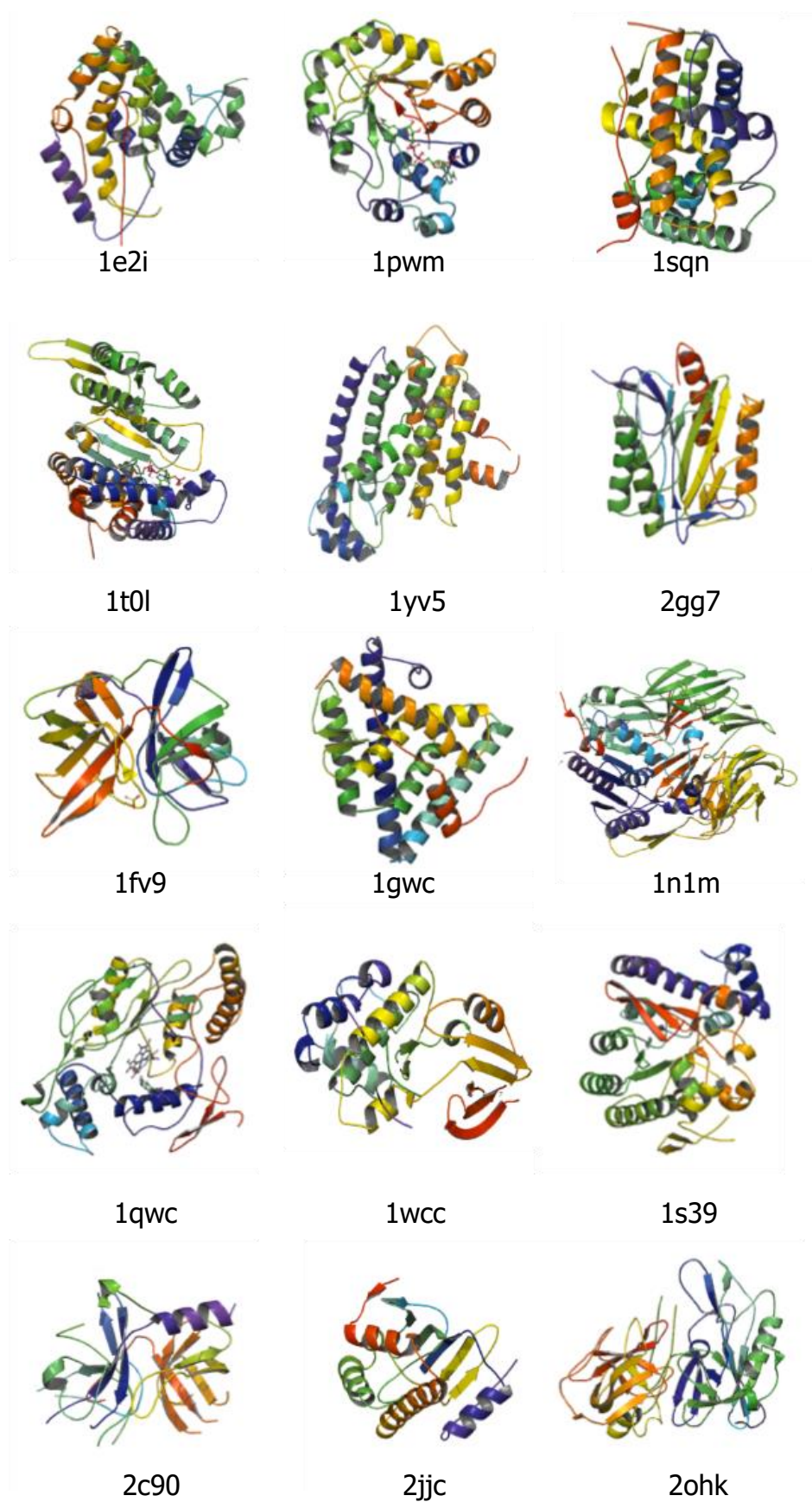
Figure 6.1. The folds of the proteins used in cross-docking.
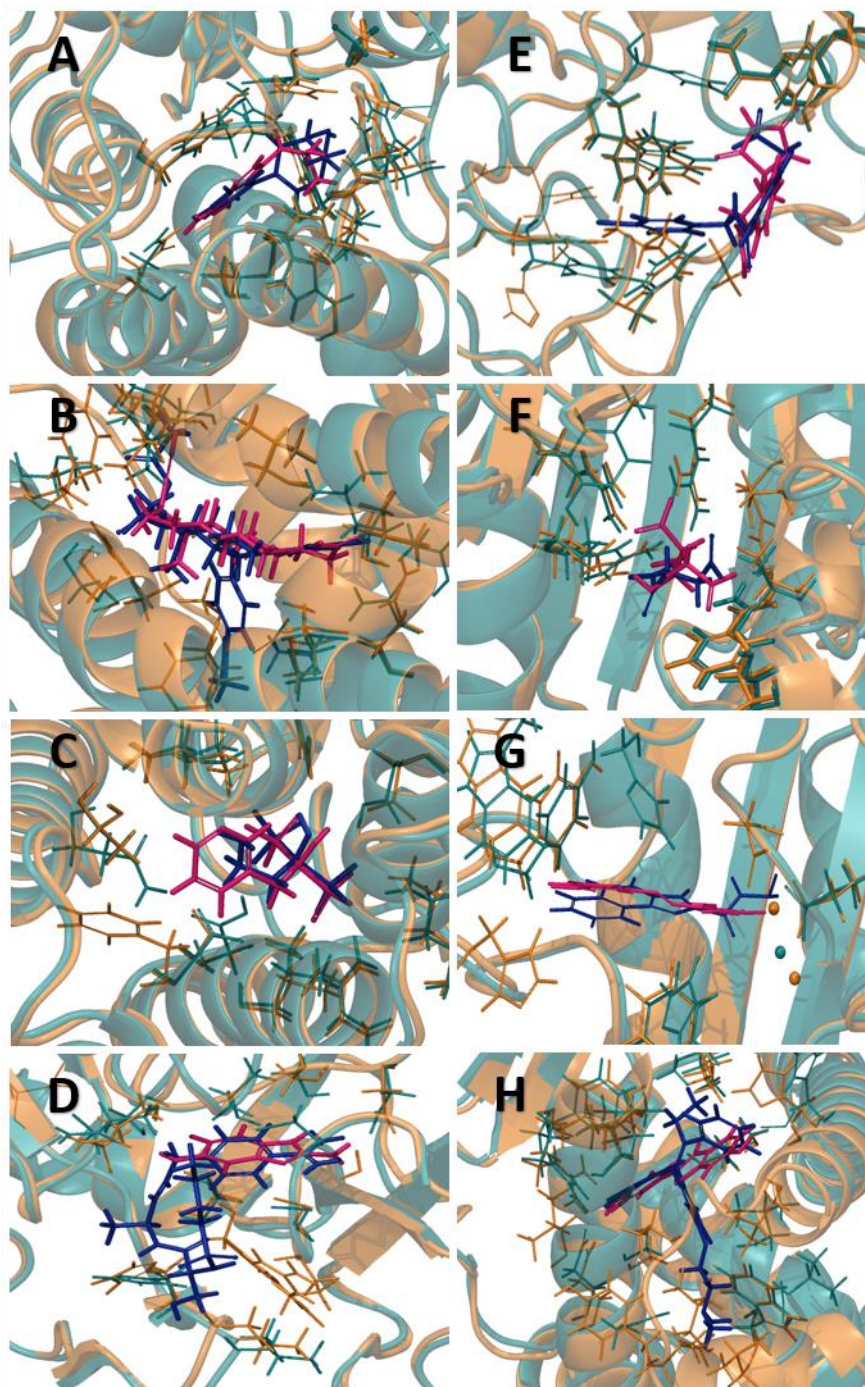
Figure 6.2. The initial protein complexes (Table 6.1, green protein with pink ligand) and an associated complex with the 'average' ligand (average defined by the Tanimoto index, brown protein with blue ligand). The initial protein complex pdb codes appear first in the following list of pairs: A. 1e2i and 1ki6 B. 1sqn and 4oar C. 1yv5 and 4kpj D. 1fv9 and 1w0z E. 1pwm and 3lz5 F. 1t0l and 4l04 G. 2jj7 and 2p98 H. 1gwq and 4q50
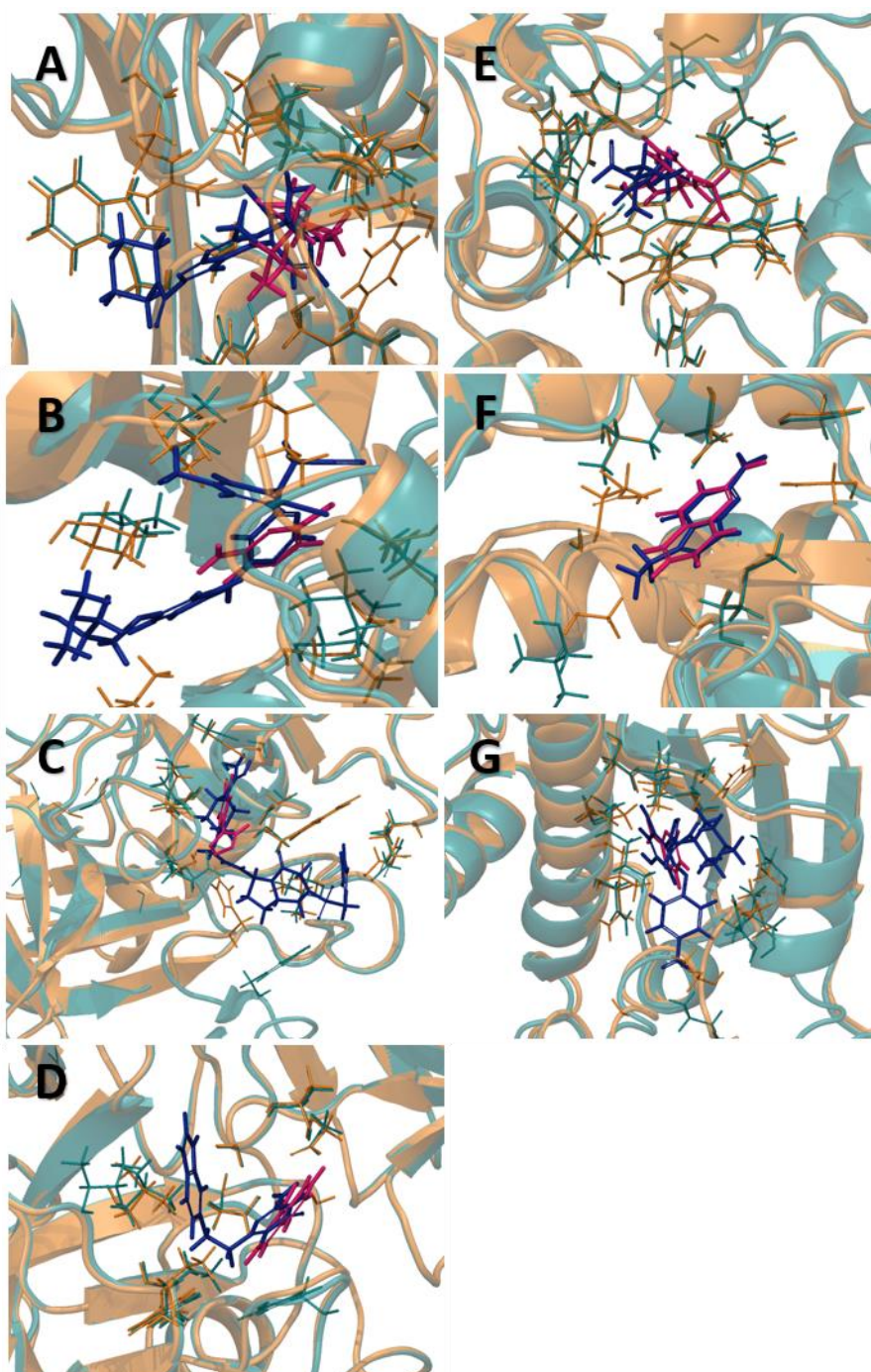
Figure 6.3. The initial protein complexes (Table 6.1, green protein with pink ligand) and an associated complex with the 'average' ligand (average defined by the Tanimoto index, brown protein with blue ligand). The initial protein complex pdb codes appear first in the following list of pairs: A. 1n1m and 3nox B. 1wcc and 2iw6 C. 2c90 and 1jwt D. 2ohk and 2va5 E. 1qwc and 1lzx F. 1s39 and 4gcx G. 2jjc and 4b7p

Polarized self-docking and polarized cross-docking. The docking was repeated using (a) polarized charges for the ligand in conjunction with the unpolarised charges for the enzyme and (b) using fully polarized ligand and enzyme charges, as in Chapter 3. Possible double counting of polarization effects that are included implicitly in the charges as a result of the parameterization process (Winn et al., 1999), were minimized as Glide scoring protocols reduce the coulombic term by ~50% on formally charged groups; and the van der Waals interaction energy is also scaled on the atoms directly involved during evaluation (Friesner et al., 2004a). Basis set superimposition effects are also avoiding using QM/MM methods (Gooding et al., 2000b). For each of the 5495 complexes, a QM/MM calculation was performed using Gaussian 03 (Frisch et al., 2004), where the ligand was QM and the whole protein was included in the MM region, which was modelled using the OPLS 2005 force field (Banks et al., 2005a), including OPLS charges (in the case of full polarization, these protein OPLS charges were augmented with our induced charges). The scalar isotropic atom polarizabilities were taken from (Miller, 1990) and assigned according to atom type (see Appendix D, Table D.1), as discussed previously. The B3LYP DFT method (Becke, 1993, Lee et al., 1988) was used with a 6-31G* basis set (Hehre et al., 1972). The 3-21G* basis set was used for a ligand that contained iodine, as the 6-31G* basis set is not available (Rassolov et al., 2001, Frisch et al., 2004). The Gaussian 03 QM/MM calculation determines the QM potential at each atom in the protein. While the induced dipole at these atoms requires determination of the field through the equation $\mu = \alpha E$, the determination of the induced charges only requires determination of the potential, as discussed in Chapter 2 (Ferenczy and Reynolds, 2001); the induced charges are then added to the unpolarised charges. The QM/MM calculation is repeated in the presence of the polarized charges and the process is repeated to convergence on the charges to 3DP (usually about 3 iterations). At the end of the calculation, a distributed multipole analysis (DMA) (Stone, 1981) is derived from the Gaussian checkpoint file using GDMA (Stone, 2005) and the potential-derived charges determined from this DMA using mulfit (Ferenczy, 1991, Ferenczy et al., 1997, Winn et al., 1997). These potential-derived charges are for the ligand in the field of the polarized

enzyme and correspond to the polarized charges for the ligand. The polarised charges for the ligand (where only the ligand is polarized) were determined in the same way from a QM calculation on the unpolarised enzyme. The protein preparation protocols can have a significant effect on docking quality (Sandor et al., 2010). Thus, Allen et al. were able to obtain impressive results for cross-docking when each ligand was minimized in each protein structure (Allen et al., 2015). Consequently, the same protocols were used for the self-docking and the cross-docking to ensure that the results provide a fair test of the effects of polarization.

Assessment of docking quality. The RMSD of the top docked pose for each ligand was determined against the native structure for self-docking and against the merged ligand coordinates for cross-docking. Four thresholds were used, namely 0.5 Å, 1.0 Å, 1.5 Å and 2.0 Å. In addition, for each pair of structures in a given set, we determined the number of instances where correct cross-docking of ligand A into protein B was associated with the correct docking of ligand B into protein A. In order assess the difficulty of cross-docking, ligand similarity was determined using the Tanimoto (Bajusz et al., 2015) and Molshacs (de Lima and Nascimento, 2013) similarity indices. Drug-like properties of the ligand were determined using the molinspiration webserver (molinspiration.com) based on SMILES strings exported from Maestro.

Bootstrapping. To assess statistical variance a Monte Carlo algorithm for bootstrapping (random resampling with replacement, implemented in R) was used to estimate the standard deviation σ in the percentage of hits with an RMSD below a given threshold (the test statistic). For each protein parent group (original data set), 1000 random samples were taken of the same size as the parent group (e.g. 72 for 1e2i); variation in each sample arose by permitting a given docking to be included multiple times (replacement). The test statistic was computed for each sample, producing a population parameter $\hat{\theta}$ of 1000 percentages, on which to estimate standard deviation (σ) (Manly, 2006). A similar procedure was followed for assessing the accuracy of the mutual pair docking. For each method, protocol, and

threshold, the test statistic was evaluated with the same random index list, before generating a new resample for each n in the $\hat{\theta}$ series n=1000. The use of pre-set random numbers ensured repeatability as required. Appendix D Tables D.4 to D.6 show population parameter statistics for the bootstrap in Table 6.2.

Properties. Ligand and protein volumes were calculated using Maestro.

6.3 Results

6.3.1 Sampling Errors

The biggest issue in cross-docking is the problem of sampling, since the protein is often held rigid and probably adopts a sub-optimum configuration for the ligand under study. Here the question under study is whether the various docking experiments are able to generate a pose with an RMSD of less than 2.0 Å, regardless of how well the pose is ranked. Figure 6.4A shows for sampling efficiency, that both polarization treatments give a modest increase for self-docking (~5%), and a larger increase for cross-docking (~10%). However, the improvement is not uniform across the board, with around a third of the groups e.g. the 1e2i group having minimal improvement, a third having modest improvements of about 5-10%, e.g. the 1gwq group and a third such as the 1yv5 group having significant improvements of 10-44%. Figure 6.4B shows that full polarization gives the best sampling for cross docking for each of the 15 groups; Ligand polarization does not give better results than full polarization, but does give equivalent results in two of the groups.
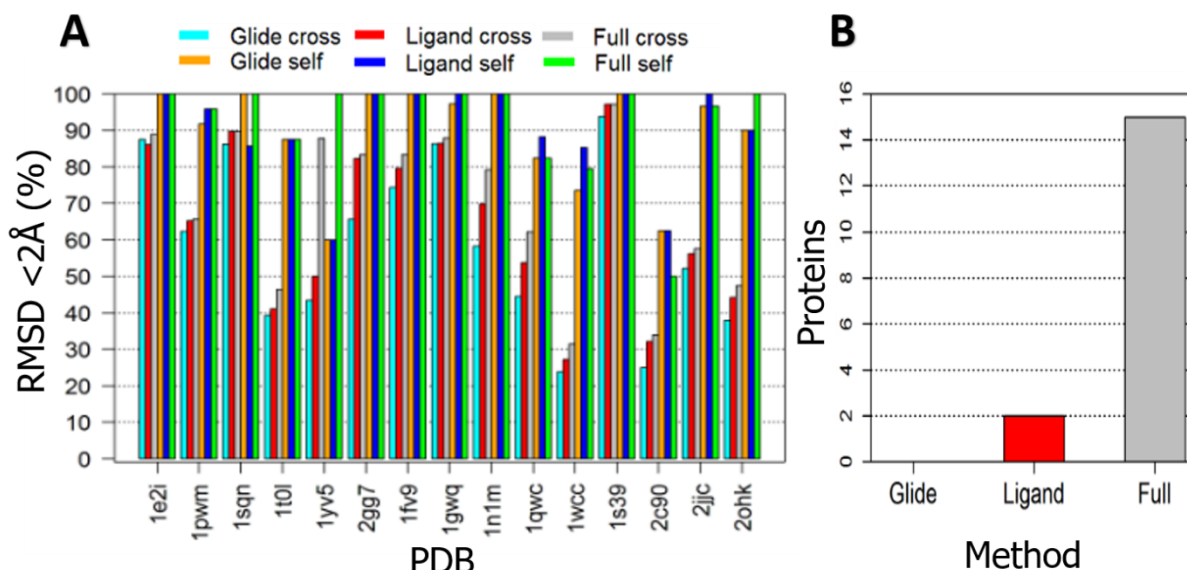
Figure 6.4. (A) A measure of sampling efficiency for self-docking and cross-docking, with and without polarization. For each complex family, the percentage of docking experiments that yielded a pose with an RMSD < 2 Å is given for both self-docking (denoted self) and cross-docking (denoted cross), for no polarization (denoted Glide), ligand polarization (denoted ligand) and full polarization (denoted full). (B). Counts of the method that gives (or matches) the best sampling. For the cross-docking case, the number of times that no polarization (Glide), ligand polarization and full polarization gives the highest percentage in Figure 6.4A is given.

RMSD analysis: self-docking. The basic RMSD results are given in Figure 6.5 for the 257 self-docking experiments and the 5238 cross-docking experiments. The results clearly show that self-docking is almost a solved problem as 95% of ligands docked to within 4.0 Å and 79% docked to within 2.0 Å in the absence of polarization. The ability to dock ligands to within an RMSD 2.0 Å is a standard measure of docking success and indeed similar results, of 66 – 91% (mean 76±8%, Glide 71%) were reported for a recent docking challenge involving multiple docking programs (Yuriev et al., 2015), but exact comparison is difficult as docking results can be system dependent (Warren et al., 2006, Yuriev et al., 2015). Because Glide is so effective in self-docking, the options to improve this process through either ligand polarization or full polarization are limited. Nevertheless, full polarization gives an improvement of 2%, 7%, 11% and 8% respectively at the 4 thresholds of 0.5 Å, 1.0 Å, 2.0 Å and 4.0 Å respectively.
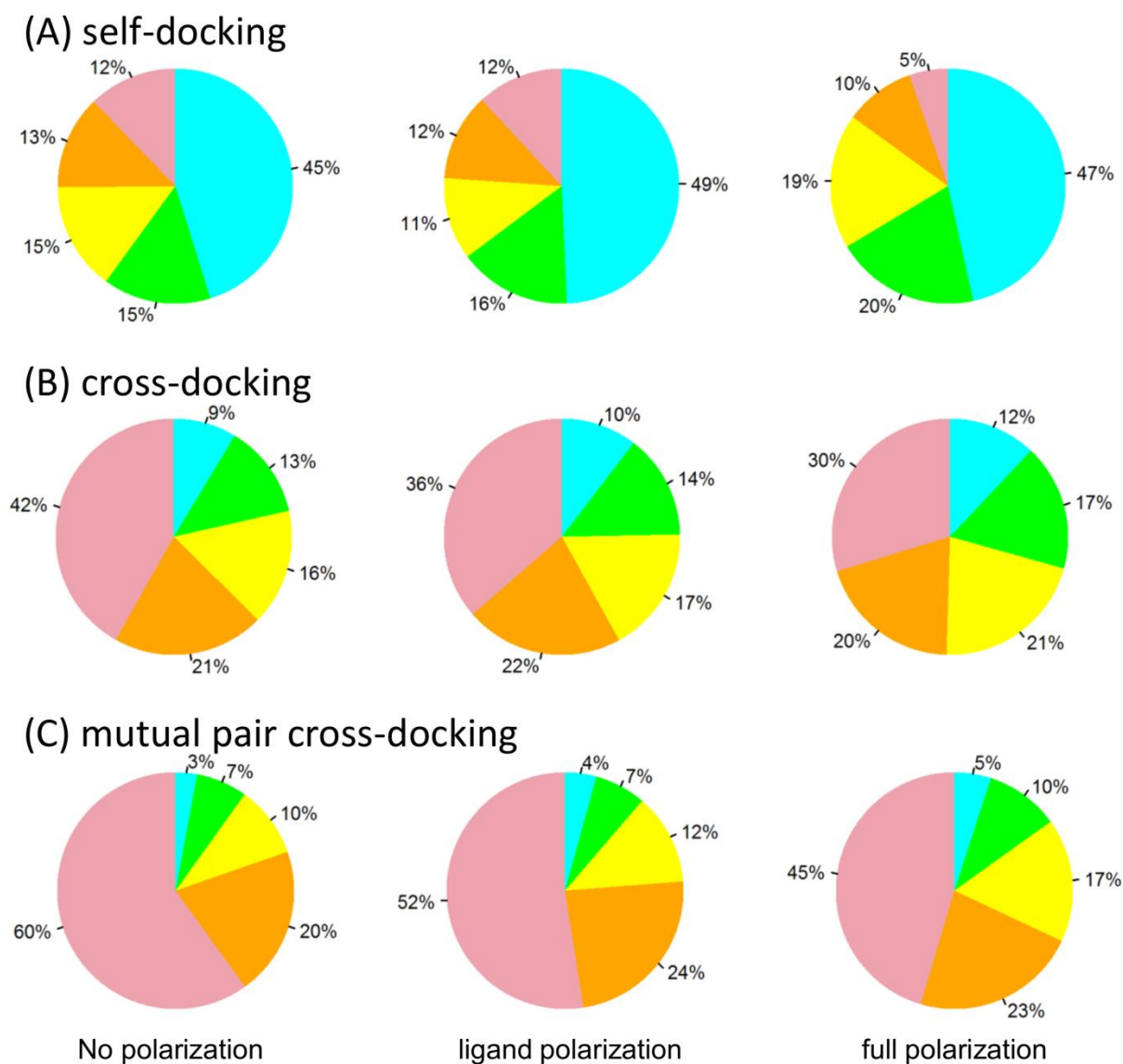
Figure 6.5. (A) The self-docking results, showing the percentage of ligands docked to within 0.5 Å (cyan), 1.0 Å (green), 2.0 Å (yellow), 4.0 Å (orange) and > 4.0 Å (pink). (B) The cross-docking results, colour-coded as for (A). (C). The mutual (pair-based) cross-docking results, colour-coded as for (A). Here we define mutual pairs as a stringent (pair-based) cross docking requirement, that if ligand A from protein A docks into protein B, then ligand B from protein B must also dock to protein A – both to within the given threshold.

RMSD analysis: cross-docking. Cross-docking, however, remains a difficult problem, as shown by the dramatic reduction in the number of poses docked to within 1.0 Å from 60% for

self-docking (Figure 6.5A) to 22% (no polarization), as shown in Figure 6.5B. Under cross-

docking, full polarization gives an improvement of 3%, 7%, 12% and 11% respectively at the

4 thresholds. However, this mean improvement of ~8% hides a number of interesting

features, as can be seen in Table 6,2, which shows that while six families show modest

improvement less than 10%, four families show greater than 15% improvement in cross-

docking. By way of contrast, Allen et al. obtained a very impressive success rate of 51% at

2.0 Å for cross-docking, but the enzyme – ligand complexes were pre-primed to accept the

ligand by minimization so the results (Allen et al., 2015) are not comparable to those

presented here as we did not minimize the complexes with the ligand prior to docking.

Table 6.2. Cross-docking performance per group for basic Glide (i.e. no polarization), ligand polarization and full polarization. The percentage of poses docked with an RMSD better than 2 Å is given, along with error bars determined by bootstrapping. The percentage increase in performance on including polarization is shown in **bold** if it is larger than the sum of the two error bars.

Name	Pdb code	Glide	Polar ligand	Full Polarizn	% increase for polar ligand	% increase for full polarizn
Herpes simplex virus type 1	1e2i	27.8±7.7	22.2±6.8	33.3±7.8	-5.6	5.5
Aldose reductase	1pwm	23.1±2.6	26.1±2.8	26.5±2.8	3.0	3.4
Progesterone receptor	1sqn	50.0±17.9	50.0±18.5	62.5±17.7	0.0	12.5
Isocitrate dehydrogenase	1t0l	7.1±5.0	14.3±6.4	25.0±8.1	7.2	**17.9**
Farnesyl diphosphate synthase	1yv5	17.8±5.7	20.0±5.8	75.6±6.3	2.2	**57.8**

methionine aminopeptidase	2gg7	11.1±4.7	24.4±6.3	22.2±6.3	**13.3**	**11.1**
microurokinase	1fv9	20.0±3.8	29.5±4.3	32.4±4.6	**9.5**	**12.4**
Oestrogen receptor	1gwq	47.4±2.3	50.1±2.3	52.6±2.3	2.7	**5.2**
Dipeptidyl peptidase IV	1n1m	7.7±2.8	13.2±3.5	23.1±4.5	5.5	**15.4**
nitric oxide synthase	1qwc	4.4±1.8	5.1±1.9	12.5±2.8	0.7	**8.1**
Cyclin-dependent kinase 2	1wcc	3.7±0.8	5.7±1.0	6.8±1.1	**2.0**	**3.1**
tRNA-Guanine transglycosylase	1s39	48.6±4.8	61.9±4.8	67.6±4.6	**13.3**	**19.0**
Thrombin	2c0-	10.7±5.8	7.1±4.7	7.1±4.9	-3.6	-3.6
HSP90	2jjc	13.6±1.6	20.3±1.9	23.8±2.0	**6.7**	**10.2**
β-secretase 1	2ohk	2.1±1.0	5.3±1.6	8.9±2.0	**3.2**	**6.8**

RMSD analysis: mutual cross-docking pairs. In Figure 6.5C we examine the more stringent requirement that if ligand A from protein A docks into protein B then ligand B from protein B must also dock to protein A – both to within the given threshold. Here the cross-docking success rates dramatically decrease from 49% in Fig. 6.5A for self-docking to 9% for cross-docking (Fig. 6.5B) and further down to 3% (Fig. 6.5C) for mutually successful cross-docking at the 0.5 Å threshold. At the more relevant 2.0 Å threshold, without polarization, the success rates decrease from 78% for self-docking to 41% for cross-docking and further down to 17% for mutual cross-docking. When full polarization is included, the corresponding success rates are 83%, 52% and 32% respectively. Thus, the improvement with full polarization included is 5% for self-docking (as reported above), 11% for cross-docking and 15% for mutual cross-docking. The error bars computed using bootstrapping (Table 6.2) indicate that in the majority of cases these improvements are significant (bootstrapping $\hat{\theta}$ statistics for this Table shown in appendix figures D.4 to D.6).

We investigated to see if there was a correlation between improved performance with polarization and various physical properties. These properties included: ligand charge, Number of ligand acceptor groups, Number of ligand donor groups, percentage of receptors with steric clashes with ligand (R(C-C) < 3 Å), Molshacs and Tanimoto ligand similarity, ligand polar surface area, ligand volume and pocket volume. The only notable observation was a weak correlation with formal ligand charge: the improvement is greatest for drug-like ligands for farnesyl diphosphate synthase that carry formal charges of -4 and for ligands for other targets that carry formal charges of -3. Allen et al. noted that cross-docking was more difficult in the presence of high electrostatic fields, and so these may be cases where polarization would be most beneficial (Allen et al., 2015). However, overall there were no convincing correlations (results not shown).

The overall success rate is 11%, but given that different enzyme targets behave differently for docking (Warren et al., 2006), it is better to average over each family so that results from large families do not swamp the results from small families. With this method the overall improvement in cross-docking in the presence of full polarization is 13.4%.

RMSD analysis: mean (μ) RMSD

The mean RMSDs for both self-docking and cross-docking are given in Figure 6.6. Again, it is clear that the RMSDs obtained for self-docking are (a) in line with those of other studies (the mean of 1.7 Å compares favourably with those of ~1.6 Å (Sandor et al., 2010) and (b) considerably lower than those for cross-docking, where the mean is 3.7 Å.
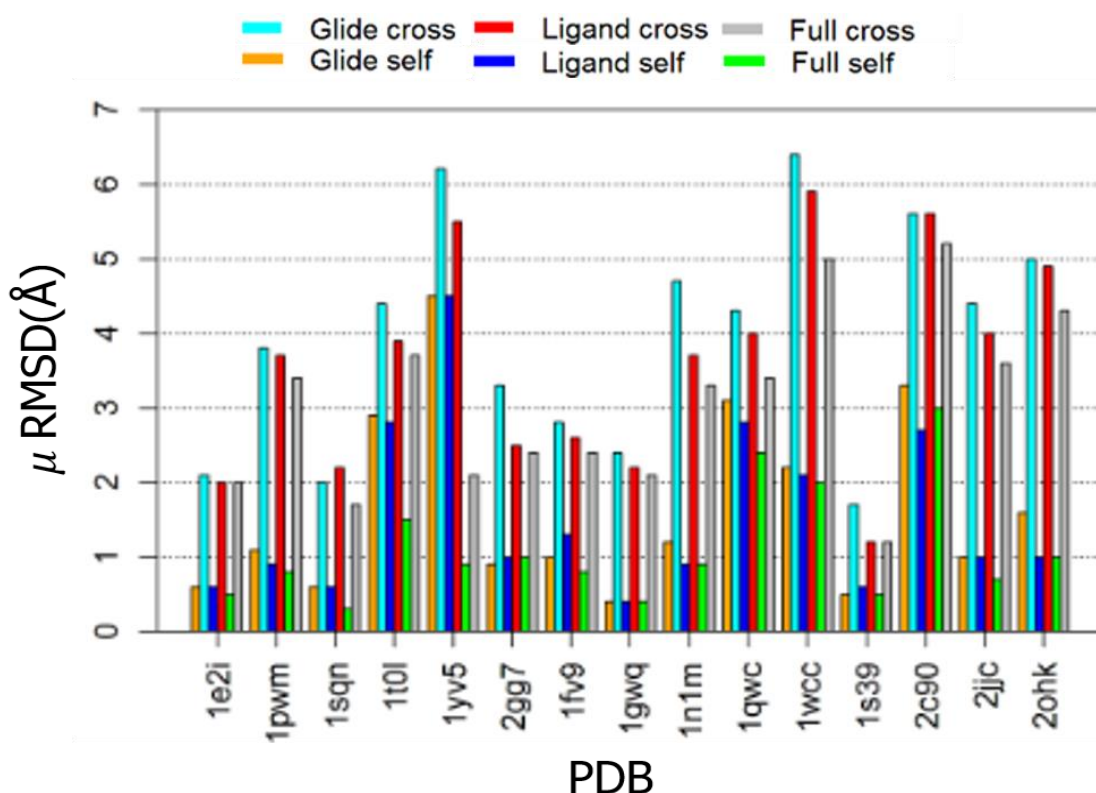
Figure 6.6 The mean (μ) RMSD in angstroms (Å) of the docked poses for both self-docking and cross-docking, for all 15 groups. SELF refers to native docks originating from co-crystallized structures, CROSS refers to cross docks originating from structures that were not co-crystallized or primed for *(in silico)*, OPLS refers to the Glide SP cVDW using IMPACT and OPLS 2005 partial charges. Ligand refers to the polarized ligand only partial charges, and Full refers to the fully polarized complex partial charges.

An alternative way to determine the best approach is to monitor which method generates or has equal to the lowest mean RMSD; these results are given in Fig. 6.7, which show that in the vast majority of cases, the lowest RMSD is given by full polarization. Graphical docking examples of the situation where full polarization performs well and Glide SP does not are shown in Fig. 6.8. The number of times this happens, and the number of refractory cases when Glide SP performs well and full polarization does not are covered by two criteria in appendix Figures D.2 and D.3.
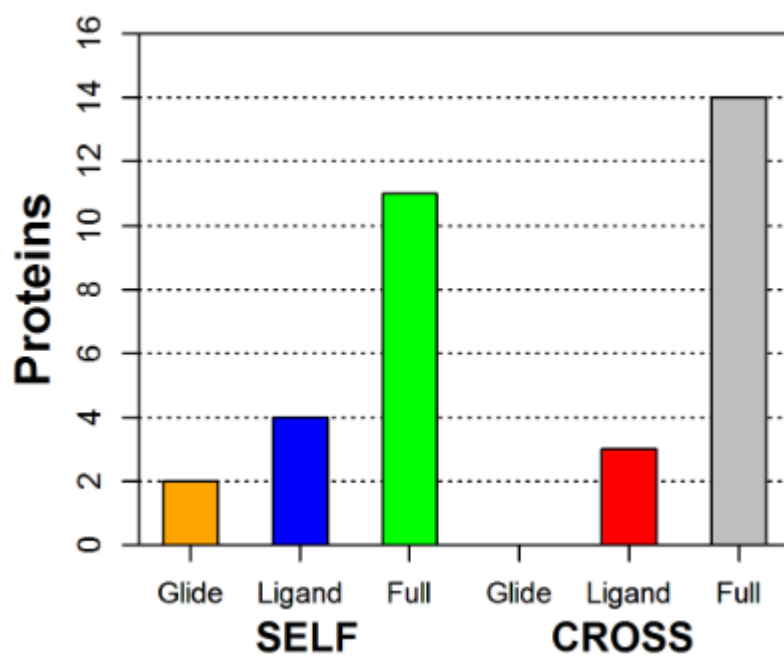
Figure 6.7. The association between method and the number of best μ RMSD poses (Å) generated, for both self-docking (left) and cross-docking (right).
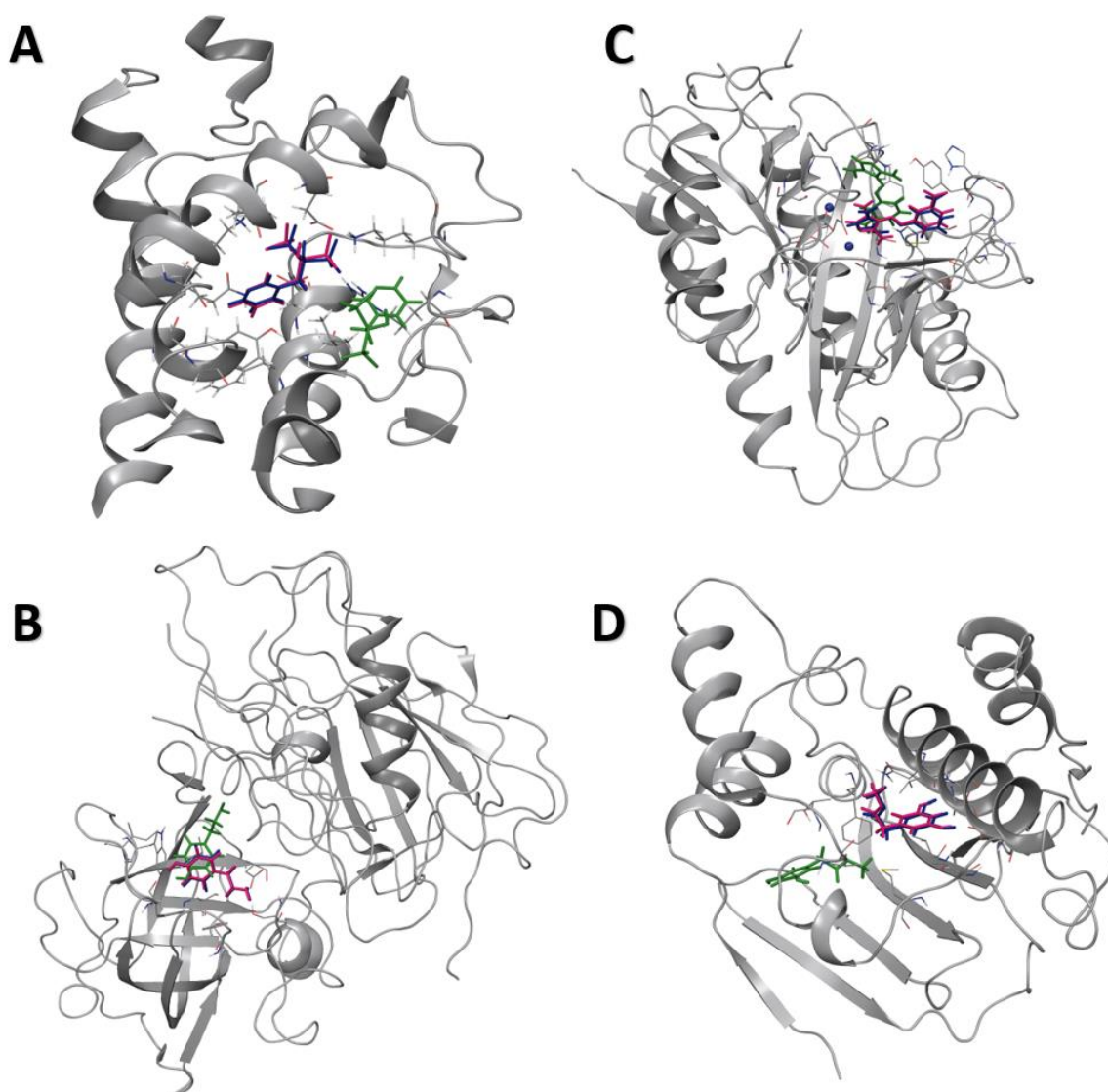
Figure 6.8. Clear improvements from including full polarization. In the four figures A to D, pink indicates a ligands x-ray experimental position in its co-crystallised receptor (silver). Blue and green indicates the same ligand, crossed docked in a non-native receptor (named below), and shown in the above receptor at the pre-docking global co-ordinates from the Maestro protein structural alignment tool. Blue is the top ranked pose from a full polarization cross-docking experiment, and Green is the same top ranked pose from the Glide SP (cVDW) cross-docking experiment. The four receptors, and their relevant pdb codes of their ligands are (A) Farnesyl diphosphate synthase,1yv5 ligand RIS cross docked in 4n9u. (B) β-secretase 1, 3bra ligand AEF cross docked in 2ohk (C) methionine aminopeptidase, 2gg7 ligand U14 cross docked in 2gg0 (D) Heat shock protein, HSP90, 2ye2 ligand

XQI cross docked in 4NH7. The relative RMSD by colour from the pink ligand are (A) blue 0.5 Å, green 10.5 Å, (B) blue 0.6 Å, green 4.2 Å, (C) blue 0.9 Å, green 6.8 Å, (A) blue 0.5 Å, green 10.4 Å. The relative occurances of these clear improvements and the refractory cases where Glide SP (cVDW) is better without full polarization are shown in appendix Tables D.2 and D.3.

6.4 Discussion

Our results have shown that cross-docking is a difficult problem, even when addressed with a well-validated docking program such as GLIDE. Cross-docking, however, rather than the easier self-docking problem is most closely related to the real-world virtual screening problem of trying to find a novel ligand for a drug target crystallized in the presence of a well-known tool compound. This is not surprising given the known limitations of current docking programs. For example, the benchmarks for accuracy in computational chemistry set by compute-intensive high level quantum chemical calculations and by long molecular dynamics simulations provide a backdrop against which docking calculations can be evaluated. Such calculations illustrate the importance of a good description of the electronic configuration, of the dynamics of the protein-ligand complex and of the solvent. Against this background, most theory/force-field based docking programs are of necessity somewhat deficient because of the need to handle large databases of compounds (e.g. ~ 1M) in a reasonably short amount of time. Indeed, knowledge-based and machine-learning based approaches implicitly recognize the limitations in the physics behind docking programs, and offer elegant alternative approaches. However, numerous successes for virtual screening, including studies using theory or force field-based methods have been reported and this has inspired much research for improving docking methods. Flexible docking can go some way towards approximating the dynamics of the protein-ligand complex, but typically this involves a small number of the more important flexible residues rather than a fully flexible protein. It is this issue of flexibility that creates problems in cross-docking and is also the reason why cross-docking over multiple targets is used as a surrogate for protein dynamics (Yuriev and Ramsland, 2013, Rueda et al., 2009, Moroy et al., 2015, Sandor et al., 2010).

Changes in protein conformation are however, not the only way in which proteins respond to the approach of a ligand. Polar ligands will generate an electric field at the protein and so the electronic distribution in the binding site should change as the ligand docks, and indeed the electrons can respond faster that the atoms. The resulting polarization energy is likely to be of the order of about 10-15% of the electrostatic energy (Ferenczy and Reynolds, 2001, Thompson et al., 1994, Illingworth et al., 2006, Gao, 1997), but more significantly polarization may help to alleviate electrostatic clashes (Illingworth et al., 2008, Gooding et al., 2000a) and so may alter the docked geometry; this appears to be the origin of the improved success rate for cross-docking with full polarization, as reported here.

A number of docking programs and scoring functions, e.g. GOLD (Jones et al., 1995) do not include a description of the electronic distribution and nevertheless obtain good results. However, with an increased realization that lipophilicity can contribute to drug attrition through promiscuous off-target interactions, hence the desire for a reasonably high ligand lipophilicity efficiency (Congreve et al., 2008a), there is probably more incentive for docking programs to identify high-scoring ligands that are not too lipophilic. This cause will be aided by docking programs that have a good treatment of electrostatics. Inclusion of polarization should benefit this cause.

6.5 References

2015. Small-Molecule Drug Discovery Suite 2015-4: Glide, version 6.9, Schrödinger, LLC. New York, NY: Schrödinger.

ALLEN, W. J., BALIUS, T. E., MUKHERJEE, S., BROZELL, S. R., MOUSTAKAS, D. T., LANG, P. T., CASE, D. A., KUNTZ, I. D. & RIZZO, R. C. 2015. DOCK 6: Impact of New Features and Current Docking Performance. *Journal of Computational Chemistry,* 36**,** 1132-1156.

BAJUSZ, D., RACZ, A. & HEBERGER, K. 2015. Why is Tanimoto index an appropriate choice for fingerprint-based similarity calculations? *Journal of Cheminformatics,* 7.

BANKS, J. L., BEARD, H. S., CAO, Y., CHO, A. E., DAMM, W., FARID, R., FELTS, A. K., HALGREN, T. A., MAINZ, D. T. & MAPLE, J. R. 2005a. Integrated modeling program, applied chemical theory (IMPACT). *Journal of computational chemistry,* 26**,** 1752-1780.

BANKS, J. L., BEARD, H. S., CAO, Y., CHO, A. E., DAMM, W., FARID, R., FELTS, A. K., HALGREN, T. A., MAINZ, D. T., MAPLE, J. R., MURPHY, R., PHILIPP, D. M., REPASKY, M. P., ZHANG, L. Y., BERNE,

B. J., FRIESNER, R. A., GALLICCHIO, E. & LEVY, R. M. 2005b. Integrated Modeling Program, Applied Chemical Theory (IMPACT). *J Comput Chem,* 26**,** 1752-80.

BECKE, A. D. 1993. Density-functional thermochemistry. III. The role of exact exchange. *The Journal of Chemical Physics,* 98**,** 5648-5652.

CHO, A. E., GUALLAR, V., BERNE, B. J. & FRIESNER, R. 2005. Importance of accurate charges in molecular docking: quantum mechanical/molecular mechanical (QM/MM) approach. *J Comput.Chem.,* 26**,** 915-931.

CONGREVE, M., CHESSARI, G., TISI, D. & WOODHEAD, A. J. 2008a. Recent developments in fragment-based drug discovery. *J.Med.Chem,* 51**,** 3661-3680.

CONGREVE, M., CHESSARI, G., TISI, D. & WOODHEAD, A. J. 2008b. Recent Developments in Fragment-Based Drug Discovery. *Journal of Medicinal Chemistry,* 51**,** 3661-3680.

DE LIMA, L. A. C. V. & NASCIMENTO, A. S. 2013. MolShaCS: A free and open source tool for ligand similarity identification based on Gaussian descriptors. *European Journal of Medicinal Chemistry,* 59**,** 296-303.

FAVIA, A. D., BOTTEGONI, G., NOBELI, I., BISIGNANO, P. & CAVALLI, A. 2011. SERAPhiC: A benchmark for in silico fragment-based drug design. *Journal of chemical information and modeling,* 51**,** 2882-2896.

FERENCZY, G. G. 1991. CHARGES DERIVED FROM DISTRIBUTED MULTIPOLE SERIES. *Journal of Computational Chemistry,* 12**,** 913-917.

FERENCZY, G. G. & REYNOLDS, C. A. 2001. Modelling polarization through induced atomic charges. *J.Phys.Chem.A,* 105**,** 11470-11479.

FERENCZY, G. G., WINN, P. J. & REYNOLDS, C. A. 1997. Toward improved force fields .2. Effective distributed multipoles. *Journal of Physical Chemistry A,* 101**,** 5446-5455.

FRIESNER, R. A., BANKS, J. L., MURPHY, R. B., HALGREN, T. A., KLICIC, J. J., DANIEL, T., REPASKY, M. P., KNOLL, E. H., SHELLEY, M. & PERRY, J. K. 2004a. Glide: a new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy. *Journal of medicinal chemistry,* 47**,** 1739-1749.

FRIESNER, R. A., BANKS, J. L., MURPHY, R. B., HALGREN, T. A., KLICIC, J. J., MAINZ, D. T., REPASKY, M. P., KNOLL, E. H., SHELLEY, M., PERRY, J. K., SHAW, D. E., FRANCIS, P. & SHENKIN, P. S. 2004b. Glide: A new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy. *J.Med.Chem.,* 47**,** 1739-1749.

FRISCH, M., TRUCKS, G., SCHLEGEL, H., SCUSERIA, G., ROBB, M., CHEESEMAN, J., MONTGOMERY JR, J., VREVEN, T., KUDIN, K. & BURANT, J. 2004. Gaussian 03, revision c. 02; Gaussian. *Inc., Wallingford, CT,* 4.

GAO, J. L. 1997. Energy components of aqueous solution: Insight from hybrid QM/MM simulations using a polarizable solvent model. *Journal of Computational Chemistry,* 18**,** 1061-1071.

GLEESON, M. P. & GLEESON, D. 2009. QM/MM as a tool in fragment based drug discovery. A cross-docking, rescoring study of kinase inhibitors. *J.Chem Inf.Model.,* 49**,** 1437-1448.

GOODING, S. R., WINN, P. J., MAURER, R. I., FERENCZY, G. G., MILLER, J. R., HARRIS, J. E., GRIFFITHS, D. V. & REYNOLDS, C. A. 2000a. Fully polarizable QM/MM calculations: An application to the nonbonded iodine-oxygen interaction in dimethyl-2- iodobenzoylphosphonate. *Journal of Computational Chemistry,* 21**,** 478-482.

GOODING, S. R., WINN, P. J., MAURER, R. I., FERENCZY, G. G., MILLER, J. R., HARRIS, J. E., GRIFFITHS, D. V. & REYNOLDS, C. A. 2000b. Fully polarizable QM/MM calculations: An application to the nonbonded iodine–oxygen interaction in dimethyl-2-iodobenzoylphosphonate. *Journal of Computational Chemistry,* 21**,** 478-482.

HEHRE, W. J., DITCHFIELD, R. & POPLE, J. A. 1972. Self—consistent molecular orbital methods. XII. Further extensions of gaussian—type basis sets for use in molecular orbital studies of organic molecules. *The Journal of Chemical Physics,* 56**,** 2257-2261.

ILLINGWORTH, C. J., GOODING, S. R., WINN, P. J., JONES, G. A., FERENCZY, G. G. & REYNOLDS, C. A. 2006. Classical polarization in hybrid QM/MM methods. *J Phys Chem A,* 110**,** 6487-97.

ILLINGWORTH, C. J., MORRIS, G. M., PARKES, K. E., SNELL, C. R. & REYNOLDS, C. A. 2008. Assessing the role of polarization in docking. *J Phys Chem A,* 112**,** 12157-63.

JONES, G., WILLETT, P. & GLEN, R. C. 1995. Molecular Recognition of Receptor-Sites Using a Genetic Algorithm with a Description of Desolvation. *Journal of Molecular Biology,* 245**,** 43-53.

KAMINSKI, G. A., FRIESNER, R. A., TIRADO-RIVES, J. & JORGENSEN, W. L. 2001. Evaluation and reparametrization of the OPLS-AA force field for proteins via comparison with accurate quantum chemical calculations on peptides. *J.Phys.Chem.B,* 105**,** 6474-6487.

KRAMER, B., RAREY, M. & LENGAUER, T. 1999. Evaluation of the FLEXX incremental construction algorithm for protein-ligand docking. *Proteins-Structure Function and Genetics,* 37**,** 228-241.

LEE, C., YANG, W. & PARR, R. G. 1988. Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. *Physical Review B,* 37**,** 785-789.

LIPINSKI, C. A., LOMBARDO, F., DOMINY, B. W. & FEENEY, P. J. 2001. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv Drug Deliv Rev,* 46**,** 3-26.

LIU, J. F., HE, X. & ZHANG, J. Z. H. 2013a. Improving the Scoring of Protein-Ligand Binding Affinity by Including the Effects of Structural Water and Electronic Polarization. *Journal of Chemical Information and Modeling,* 53**,** 1306-1314.

LIU, Y., ZHAO, L., LI, W. T., ZHAO, D. Y., SONG, M. & YANG, Y. L. 2013b. FIPSDock: A new molecular docking technique driven by fully informed swarm optimization algorithm. *Journal of Computational Chemistry,* 34**,** 67-75.

MANLY, B. F. 2006. *Randomization, bootstrap and Monte Carlo methods in biology*, CRC Press.

MILLER, K. J. 1990. Additivity methods in molecular polarizability. *Journal of the American Chemical Society,* 112**,** 8533-8542.

MOROY, G., SPERANDIO, O., RIELLAND, S., KHEMKA, S., DRUART, K., GOYAL, D., PERAHIA, D. & MITEVA, M. A. 2015. Sampling of conformational ensemble for virtual screening using molecular dynamics simulations and normal mode analysis. *Future Medicinal Chemistry,* 7**,** 2317-2331.

MORRIS, G. M., GOODSELL, D. S., HUEY, R. & OLSON, A. J. 1996. Distributed automated docking of flexible ligands to proteins: parallel applications of AutoDock 2.4. *J Comput.Aided Mol.Des,* 10**,** 293-304.

MORRIS, G. M., HUEY, R., LINDSTROM, W., SANNER, M. F., BELEW, R. K., GOODSELL, D. S. & OLSON, A. J. 2009. AutoDock4 and AutoDockTools4: Automated Docking with Selective Receptor Flexibility. *Journal of Computational Chemistry,* 30**,** 2785-2791.

OLSSON, M. H., SONDERGAARD, C. R., ROSTKOWSKI, M. & JENSEN, J. H. 2011. PROPKA3: Consistent Treatment of Internal and Surface Residues in Empirical pKa Predictions. *J Chem Theory Comput,* 7**,** 525-37.

RASSOLOV, V. A., RATNER, M. A., POPLE, J. A., REDFERN, P. C. & CURTISS, L. A. 2001. 6-31G* basis set for third-row atoms. *Journal of Computational Chemistry,* 22**,** 976-984.

RUEDA, M., BOTTEGONI, G. & ABAGYAN, R. 2009. Consistent Improvement of Cross-Docking Results Using Binding Site Ensembles Generated with Elastic Network Normal Modes. *Journal of Chemical Information and Modeling,* 49**,** 716-725.

SANDOR, M., KISS, R. & KESERU, G. M. 2010. Virtual Fragment Docking by Glide: a Validation Study on 190 Protein-Fragment Complexes. *Journal of Chemical Information and Modeling,* 50**,** 1165-1172.

SONDERGAARD, C. R., OLSSON, M. H., ROSTKOWSKI, M. & JENSEN, J. H. 2011. Improved Treatment of Ligands and Coupling Effects in Empirical Calculation and Rationalization of pKa Values. *J Chem Theory Comput,* 7**,** 2284-95.

SOUSA, S. F., RIBEIRO, A. J. M., COIMBRA, J. T. S., NEVES, R. P. P., MARTINS, S. A., MOORTHY, N. S. H. N., FERNANDES, P. A. & RAMOS, M. J. 2013. Protein-Ligand Docking in the New Millennium - A Retrospective of 10 Years in the Field. *Current Medicinal Chemistry,* 20**,** 2296-2314.

STONE, A. J. 1981. DISTRIBUTED MULTIPOLE ANALYSIS, OR HOW TO DESCRIBE A MOLECULAR CHARGE-DISTRIBUTION. *Chemical Physics Letters,* 83**,** 233-239.

STONE, A. J. 2005. Distributed multipole analysis: Stability for large basis sets. *Journal of Chemical Theory and Computation,* 1**,** 1128-1132.

TANIMOTO, T. 1957. IBM internal report 1957. November.

THOMPSON, M. A., GLENDENING, E. D. & FELLER, D. 1994. THE NATURE OF K+ CROWN-ETHER INTERACTIONS - A HYBRID QUANTUM MECHANICAL-MOLECULAR MECHANICAL STUDY. *Journal of Physical Chemistry,* 98**,** 10465-10476.

WARREN, G. L., ANDREWS, C. W., CAPELLI, A. M., CLARKE, B., LALONDE, J., LAMBERT, M. H., LINDVALL, M., NEVINS, N., SEMUS, S. F., SENGER, S., TEDESCO, G., WALL, I. D., WOOLVEN, J. M., PEISHOFF, C. E. & HEAD, M. S. 2006. A critical assessment of docking programs and scoring functions. *Journal of Medicinal Chemistry,* 49**,** 5912-5931.

WINN, P. J., FERENCZY, G. G. & REYNOLDS, C. A. 1997. Toward improved force fields .1. Multipole-derived atomic charges. *Journal of Physical Chemistry A,* 101**,** 5437-5445.

WINN, P. J., FERENCZY, G. G. & REYNOLDS, C. A. 1999. Towards improved force fields: III. Polarization through modified atomic charges. *Journal of Computational Chemistry,* 20**,** 704-712.

YURIEV, E., HOLIEN, J. & RAMSLAND, P. A. 2015. Improvements, trends, and new ideas in molecular docking: 2012-2013 in review. *Journal of Molecular Recognition,* 28**,** 581-604.

YURIEV, E. & RAMSLAND, P. A. 2013. Latest developments in molecular docking: 2010-2011 in review. *Journal of Molecular Recognition,* 26**,** 215-239.

7 Concluding remarks

In Chapter 3, we have shown that polarization of fragments can offer in the region of a 10-15% improvement in docking results, as judged by the percentage of poses within a rather tight threshold of 0.5 or 1.0 Å RMSD of the experimental structure, where accurate prediction of binding interactions are more likely. Clearly, such an improvement could make a significant difference to a fragment-based drug design program. These results are most apparent when the correct pose is known a priori, as under these circumstances the ligand and the protein can be polarized correctly, and polarizing the ligand and the protein gives better results than just polarizing the ligand. The improvement is more apparent for the CVDW results as these are based on a molecular mechanics force field and so the electrostatics (and hence polarization effects) are not scaled down as much and the improvement there continues up to the 4 Å threshold. Analysis of the results that are correct to within a given threshold suggests that polarization is more relevant to improving the good results (thresholds 0.5 – 1 Å, and possibly 2 Å) than it is for improving the less good results (thresholds 4 Å).

In Chapter 4 generally the DMA results were disappointing, despite the potential to include important factors such as anisotropic charge distributions and exponential repulsion through Orient. Of the methods studied here, the QM/MM results were generally the best, but on occasion DMA out-performed the QM/MM calculations in some respects. However, while QM/MM has many benefits over classical approaches, Glide was shown in Chapter 3 to generally work better (by virtue of its specific parameterization). There was no inclusion of a treatment of solvation in Chapter 4. This manifested itself in a rather subtle way, in that the results from optimization were not as good as the results for single point calculations – for both Orient (DMA) and Jaguar (QM) calculations. The most obvious explanation of this is that the 'gas phase' optimization moved the structures away from the solution structures, which can to some extent be modelled using Glide due to its treatment of solvent effects. Orient

was not designed as a drug design tool, meaning that it was not possible to treat the whole enzyme, but rather the calculations were restricted to a small shell of residues around the active site. Additionally, the exponential terms of the force field were not optimized to be included with the anisotropic electrostatics, hence the Lennard-Jones repulsion performed better than exponential repulsion. There are many factors that contribute towards accurate docking. The Glide program is designed as a drug design tool and is well optimized and so generally gives good results. This makes it difficult to improve these results. Nevertheless, there are several deficiencies in the Glide method that arise primarily from the requirement to be fast. One of these is the problem of the rigid enzyme/receptor. In Chapter 3, the electronic flexibility introduced by polarization was shown to alleviate this problem. Any improvement here through polarization, was probably prevented because the effects are being swamped by other errors. However, the results in this chapter are sufficiently good to suggest that these issues should be explored more fully in the future, but ideally a more protein-related piece of software should be used.

In chapter 5, the improvement in docking observed through inclusion of specific water molecules, is far more significant than that obtained through inclusion of polarization. The active site water molecules clearly have an important role in ligand binding and so their treatment is an important factor in drug design. Polarization generally only increases the magnitude of the interacting charges (and hence the resultant electrostatic energy) by about 10-15%, but the introduction of a polar water molecule has a far greater effect due to the large magnitude of the extra charges (q_H = 0.417, q_O = -0.834 for TIP3P water). For this reason, the polarization effect introduced in this chapter is swamped for a far greater effect and so the improvement due to polarization, though generally present, is not always apparent. Since these charges are large, they will enhance the steric effect – as long as the water is oriented appropriately. Glide does no fully take into account explicit water, but rather approximates them as isoelectronic 2.8 Å spheres. When polarization helps to improve the

docking in the presence of a water molecule, some general principles seem to emerge. If the water orientation is optimized by MM methods, then polarization of the ligand is generally more successful than polarization of the ligand and the protein. However, if the water orientation is optimized by QM/MM methods then polarization of the ligand and the protein is generally more successful. One possible conclusion from this is that if the water is not correctly orientated then the protein may not be correctly polarized and so omission of protein polarization is preferable. It would seem that QM/MM orientation of the water molecule is preferable. This raises an important issue with regards to treatment of hydration and polarization within binding sites. One of the issues is that polarization in protein– ligand – specific water systems may be ligand dependent. Water molecules may be absent in low resolution structures, but they might also be absent because of their dynamic nature. Moreover, the water pattern for one bound ligand may be different to that in other bound ligands. Thus, not all water molecules are the same. We have shown that the water molecule can be placed with some degree of reliability simply by docking when the ligand is present, and the initially good hydrogen bonding conditions are met. The presence of such a water molecule could then enhance virtual screening programs for the purpose of finding other good fragment hits. We have shown that docking a water molecule in the absence of a ligand is unlikely to be successful. There is much scope for studies on the conservation of water molecules in different X-ray crystal structures of the same enzyme to see whether the absence of conservation at particular positions in indicative of an 'unhappy' water molecule.

In Chapter 6, we investigated cross-docking by determining a set of 15 family members that had at least seven protein-ligand complexes. The self-docking success to these 257 proteins and the improvement arising from the inclusion of polarization was similar to that seen in Chapter 3. However, there was a more marked success in cross-docking and this was particularly evident in the cross-docking of mutual pairs, i.e. if ligand A from its complex with protein A can dock into protein B, then ligand B should also dock into protein A. The success in cross-docking was not uniform, but rather was more marked for highly charged ligands.

Future Work

We have largely concentrated on the effects of improvements to docking in reproducing the experimental pose since this is not a trivial problem. However, a similar but related and more difficult problem is that of cross-docking, since docking a ligand taken from one structure into another structural form of the same enzyme, typically crystallized with a different ligand, is not trivial, as shown in Chapter 6. The initial success in cross-docking is encouraging but further work is required, e.g. by including a more thorough treatment of hydration. We are hopeful that the method will show some improvement as the effect of polarization should be to reduce electrostatic repulsion, thus alleviating the need for quite so much receptor structural flexibility. It would also be interesting to test the methods for recovering the experimental pose or a set of known binders from a fragment database seeded with decoys.

Improved methods for placing water molecules are required. In keeping with the current work, a molecular dynamics approach which incorporates a polarizable water model, could be developed using a modified version of the tinker program, since this has been developed within the group. Such an approach could be consistent with methods for including distributed multipoles and softer repulsion since these are also included in Tinker.

APPENDIX A

Glide XP and alternate protocols with (200 pose requests)

A (GSCORE) Lowest Available also Lowest Scoring Pose

B (GSCORE) RMSD Success discrete cut-offs

C (CVDW) Lowest Available also Lowest Scoring Pose

D (CVDW) RMSD Success discrete cut-offs

E (EMODEL) Lowest Available also Lowest Scoring Pose

F (EMODEL) RMSD Success discrete cut-offs
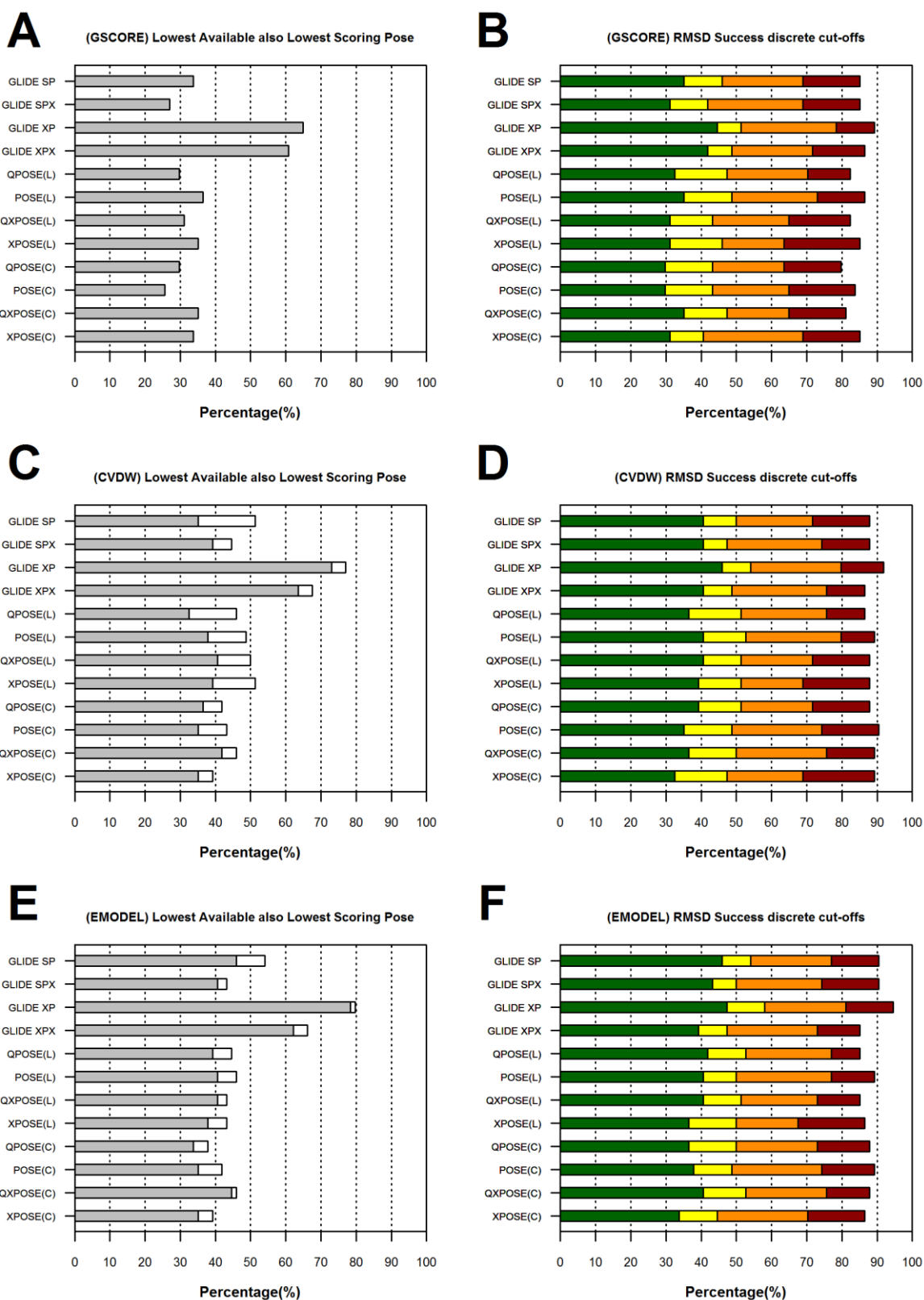
Figure A.1 Assessment of Initial docking settings and Polarization of the Ligand/Complex and the initial Polarized Geometry on Flexible re-docking success (200 Poses). Methods: Glide SP (Standard Precision), Glide XP (Extra Precision). Re-Docked with Polarized Ligand partial charges (L) or Complex partial charges (C); Then the Methods apply to the initial (L)/(C) Pose used for Flexible Re-Docking with Glide SP where: POSE is Polarized by the Highest Ranking Glide GSCORE (Also referred to as at @MM and if the pose is QSITE Jaguar QM/ Impact MM optimized before polarization given the suffix @QM); XPOSE, GLIDE SPX, (and GLIDE XPX) are re-runs with extended sampling option set. (CVDW), (EMODEL), (GSCORE) are the partitioned scoring function used by Glide. Plots A, C, E – Percentage over 74 molecules where highest ranked/lowest scoring Pose (Grey), Pose cluster (White*) is also lowest available pose. Plots B, D, F – Percentage over 74 molecules where highest ranked/lowest scoring pose is at discrete cut-offs <0.5 RMSD (Green);<1 RMSD (Yellow); <2 RMSD (Orange); <4 RMSD (Red). All Root Mean Square Deviation(s) (RMSD) are given in the unit Angstroms (Å). *Clustering parameters set at <1.1 kcal/mol and <1 RMSD. Initial Docking request was to keep 200 poses.
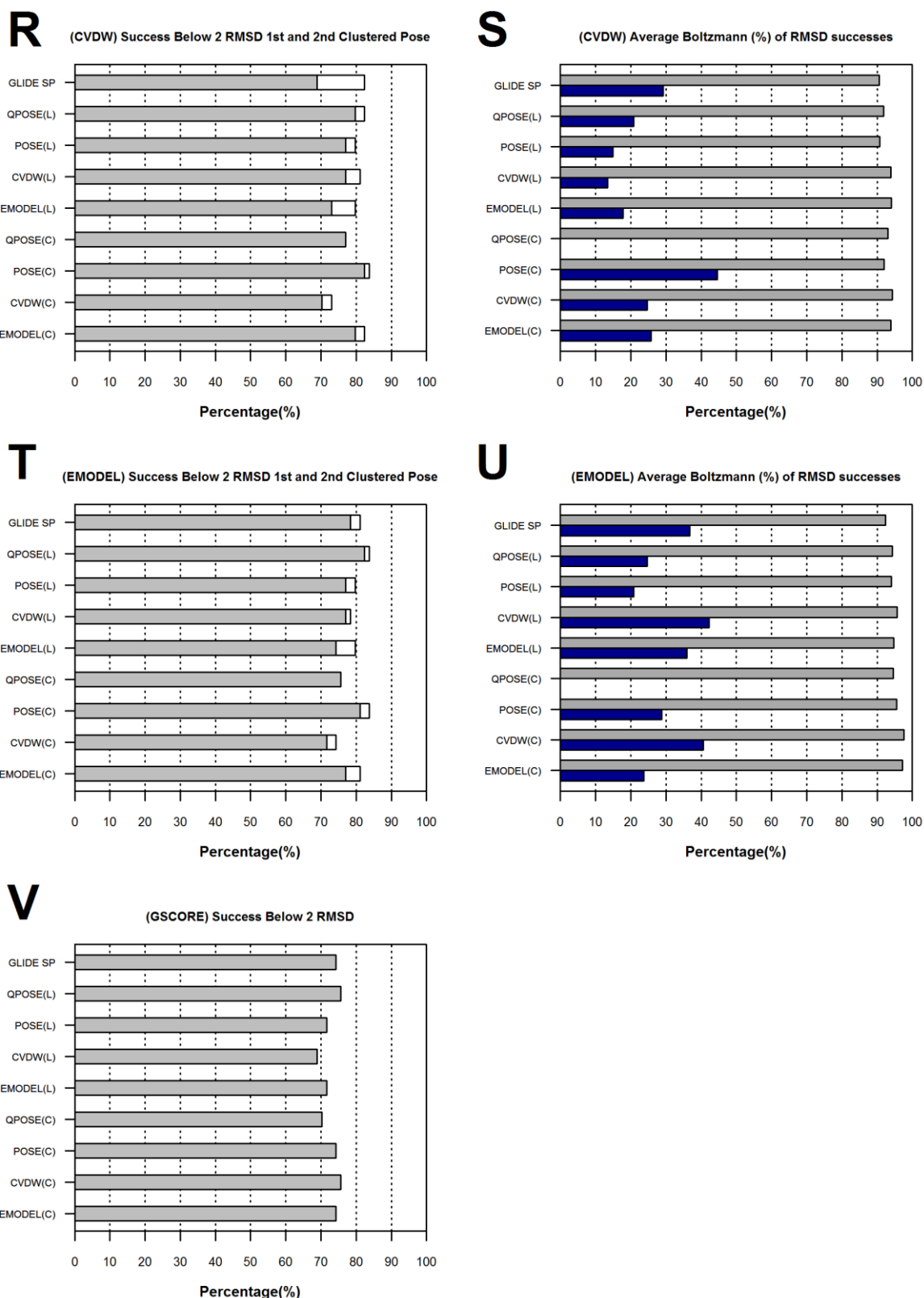
Figure A.2 Rough (<2 RMSD) Assessment of Additional Clustering on (200 pose) Docking Success (as Figure A.1) using Glide SP (Standard Precision), Re-scored with Polarized Ligand partial charges (L) or Complex partial charges (C); The Methods – Apply to the initial (L)/(C) Pose used for Re-Docking and Re-Scoring. Methods: Glide SP (Standard Precision), Glide XP (Extra Precision). Re-Docked with

Polarized Ligand partial charges (L) or Complex partial charges (C); Then the Methods apply to the initial (L)/(C) Pose used for Flexible* Re-Docking with Glide SP where: POSE is Polarized by the Highest Ranking Glide GSCORE (Also referred to as at @MM and if the pose is QSITE Jaguar QM/ Impact MM optimized before polarization given the suffix @QM); XPOSE, GLIDE SPX, (and GLIDE XPX) are re-runs with extended sampling option set. (CVDW), (EMODEL), (GSCORE) are the partitioned scoring function used by Glide. Plots G, I, K – The percentage over 74 molecules where the highest ranking/lowest scoring pose cluster (Grey) is a successful (below 2 RMSD). Then % increase when additionally accounting for 2nd Pose cluster (White*) <2 RMSD when 1st pose is not. Plots H, J – Given Boltzmann probabilty (%) averaged over 74 molecules of the Pose Population adopting the highest ranked/lowest scoring pose minima (Grey) and of adopting the 2nd Pose Cluster minima when the first isn't <2 RMSD (Dark Blue).

*Clustering parameters set at <1.1 kcal/mol and <1 RMSD. Initial Docking request was to keep 200 poses.

Glide SP Poses polarized at 1ˢᵗ Ranked Glide Scoring Function Geometries and QM/MM optimized Glide Pose
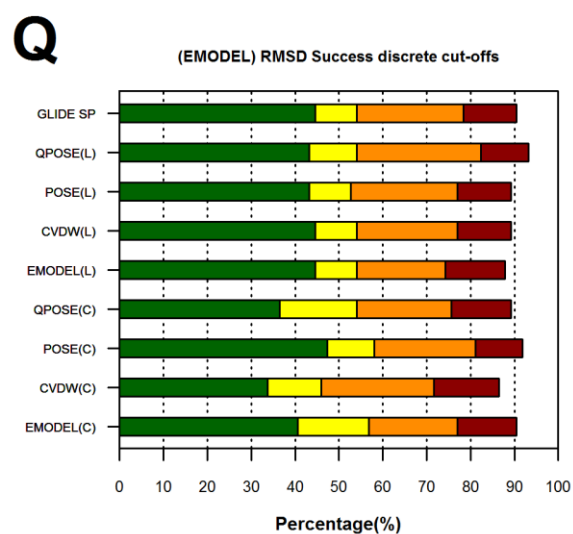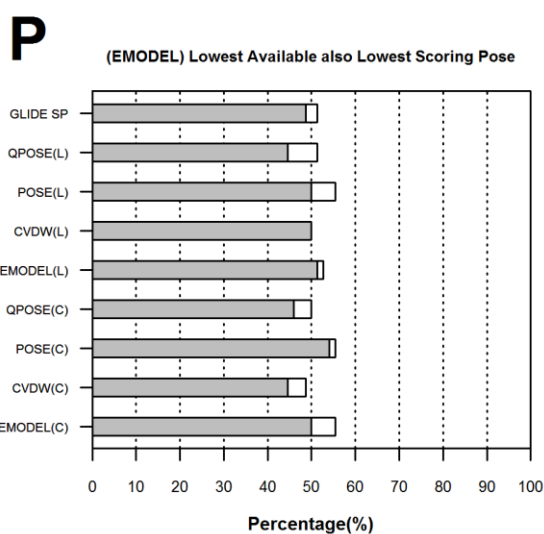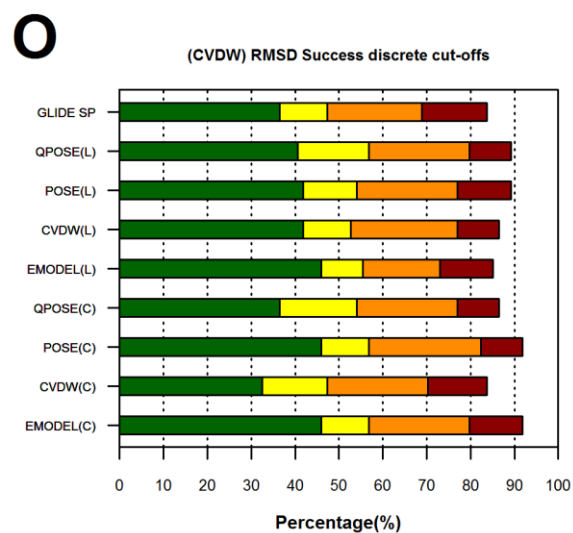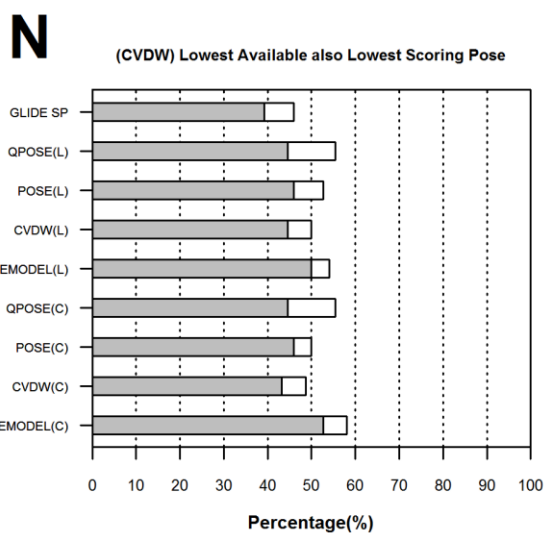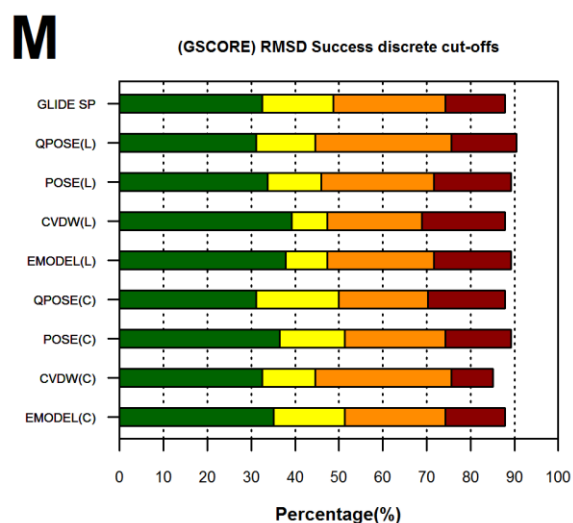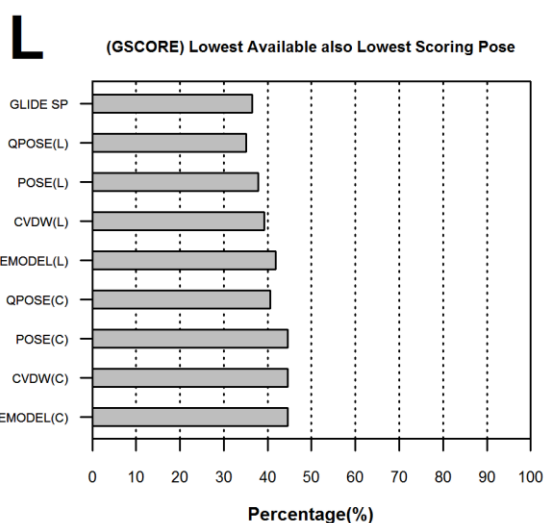
Figure A.3 Assessment of Polarization of the Ligand/Complex and the initial Polarized Geometry on Flexible* re-docking success (15 poses) using Glide SP (Standard Precision), Re-Docked with Polarized Ligand partial charges (L) or Complex partial charges (C); Then the Methods apply to the initial (L)/(C) Pose used for Flexible* Re-Docking where: POSE - Highest Ranking Glide GSCORE pose; CVDW - Highest Ranking Glide CVDW pose; POSE - Highest Ranking Glide EMODEL pose. (Also referred to as at @MM and if the pose is QSITE Jaguar QM/ Impact MM optimized before polarization given the suffix @QM e.g. JUST POSE here); (CVDW), (EMODEL), (GSCORE) in headings are the partitioned scoring function used by Glide. Plots A, C, E – Percentage over 74 molecules where highest ranked/lowest scoring Pose (Grey), Pose cluster (White**) is also lowest available pose. Plots B, D, F – Percentage over 74 molecules where highest ranked/lowest scoring pose is at discrete cut-offs <0.5 RMSD (Green);<1 RMSD (Yellow); <2 RMSD (Orange); <4 RMSD (Red). All Root Mean Square Deviation(s) (RMSD) are given in the unit Angstroms (Å). *INDIV - Re-scored using Glide Score in Place (Gives CVDW and GSCORE only). **Clustering parameters set at <1.1 kcal/mol and <1 RMSD. Initial Docking request was to keep 15 poses.
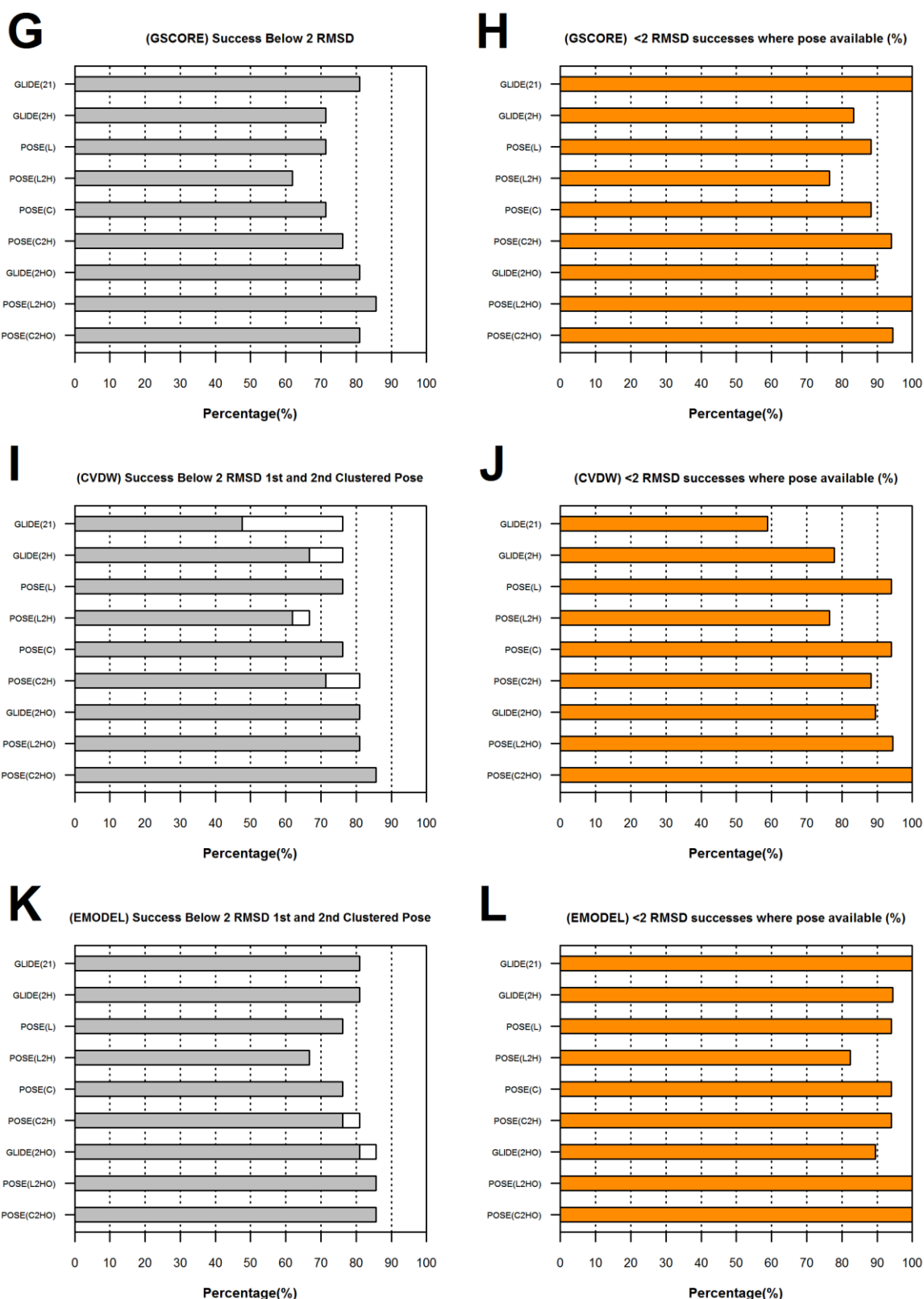
Figure A.4 Rough (<2 RMSD) Assessment of Additional Clustering on (15 pose) Docking Success (as Figure A.3) using Glide SP (Standard Precision), Re-Docked with Polarized Ligand partial charges (L) or Complex partial charges (C); Then the Methods apply to the initial (L)/(C) Pose used for Flexible Re-Docking where: POSE - Highest Ranking Glide GSCORE pose; CVDW - Highest Ranking Glide CVDW pose; POSE - Highest Ranking Glide EMODEL pose. (Also referred to as at @MM and if the pose is QSITE Jaguar QM/ Impact MM optimized before polarization given the suffix @QM e.g. JUST POSE here); (CVDW), (EMODEL), (GSCORE) in headings are the partitioned scoring function used by Glide.

Plots G, I, K – The percentage over 74 molecules where the highest ranking/lowest scoring pose cluster (Grey) is a successful (below 2 RMSD). Then % increase when additionally accounting for 2nd Pose cluster (White) <2 RMSD when 1st pose is not. Plots H, J – Given Boltzmann probabilty (%) averaged over 74 molecules of the Pose Population adopting the highest ranked/lowest scoring pose minima (Grey) and of adopting the 2nd Pose Cluster minima when the first isn't <2 RMSD (Dark Blue). **Clustering parameters set at <1.1 kcal/mol and <1 RMSD. Initial Docking request was to keep 15 poses.

Appendix B

'Lone Water' Ligand Docking results when only 2 non water-water Hydrogen bonds formed (1 Protein; 1 Ligand) – Dataset reduced to 21 molecules below.

Table B.1 Dataset structures containing a bridged interaction 'lone water' molecule. Here

(brackets) indicate ligand molecule used where more than one, also 3IME B refers to chain B.

PDB NAME	WATER ID	NO H-BONDS	PDB NAME	WATER ID	NO H-BONDS
1M2X	985	2	1MLW	506	2
1MLW	601	2	1TKU	803	2
1TKU	915	2	1YNH	2010	2
1F5F	307	2	1PWM	1552	2
1UWC	2335	2	2BKX	2364	2
2BRT	2085	2	2BRT	2128	2
1YV5	1010	2	2HDQ (501)	596	2
2ZVJ	309	2	3C0Z	1010	2
1N1M	1230	2	1WCC	2282	2
2C90	2188	2	2JJC	2292	2
3IME B	1032	2			
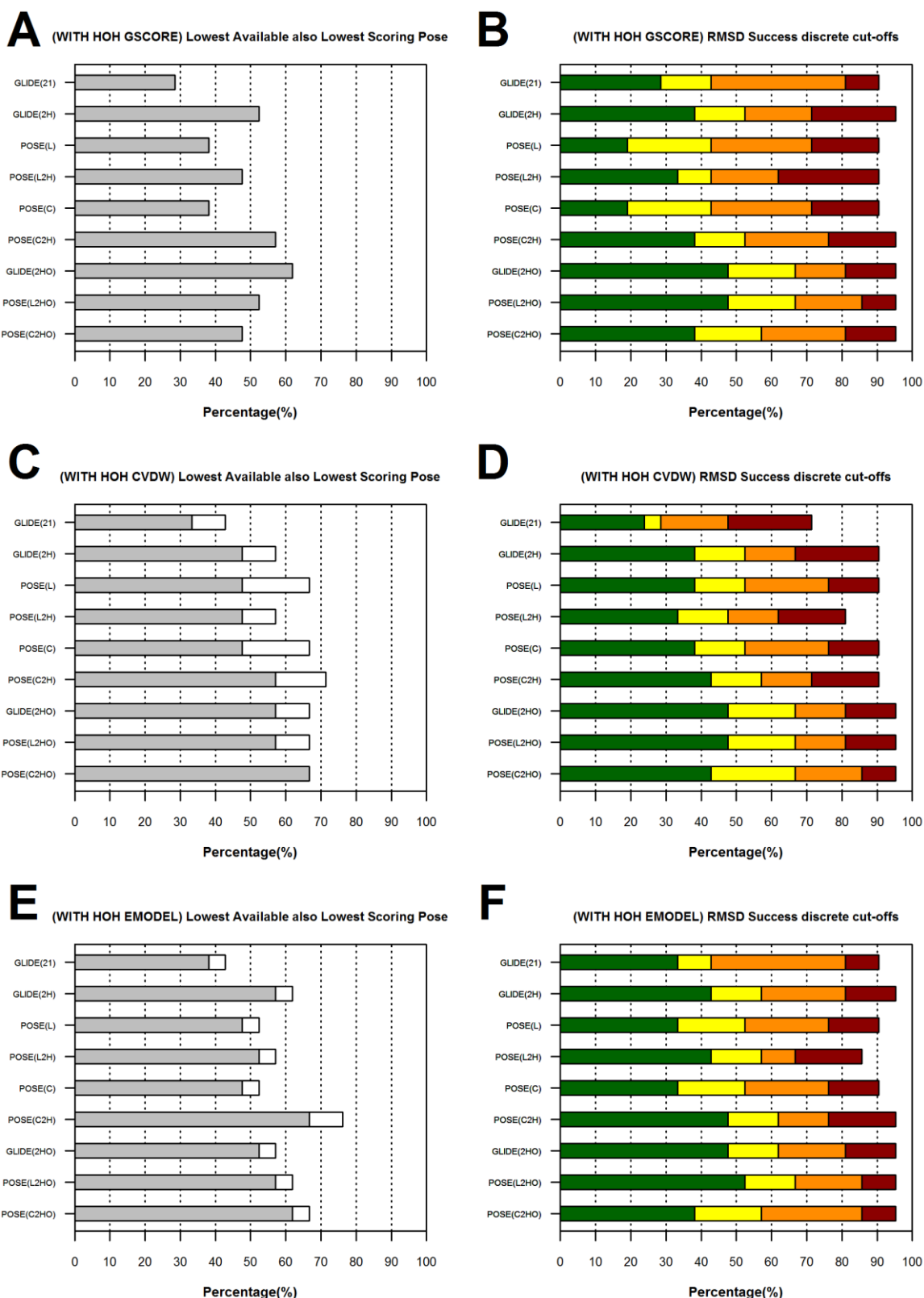
Figure B.1 Assessment of Polarization of the Ligand/Complex and the a Ligand attached water Flexible re-docking success (15 poses) using Glide where (21) refers to number of molecules from original dataset without single water molecule, (2H) refers to the number of protein/ligand hydrogen bonds formed namely 2, (L) is Re-Docked with Polarized Ligand partial charges or (C) is Complex partial charges (C); (2HO) is the same as (2H) only the single water molecule hydrogens have been orientated by QM/MM calculation, before commencing a docking run. Then POSE – is polarized at the

Highest Ranking Glide GSCORE pose geometry. (CVDW), (EMODEL), (GSCORE) are the partitioned scoring function used by Glide. Plots A, C, E – Percentage over 21 of the 74 molecules where highest ranked/lowest scoring Pose (Grey), Pose cluster (White*) is also lowest available pose. Plots B, D, F – Percentage over 21 of the 74 molecules where highest ranked/lowest scoring pose is at discrete cut-offs <0.5 RMSD (Green);<1 RMSD (Yellow); <2 RMSD (Orange); <4 RMSD (Red). All Root Mean Square Deviation(s) (RMSD) are given in the unit Angstroms (Å). *Clustering parameters set at <1.1 kcal/mol and <1 RMSD. Initial Docking request was to keep 15 poses.

Figure B.2 A Rough (<2 RMSD) Assessment of Additional Clustering on (15 pose) Docking Success (as Figure B.1) using Glide Standard Precision -GLIDE, where (21) refers to number of molecules from original dataset without single water molecule, (2H) refers to the number of protein/ligand hydrogen bonds formed namely 2, (L) is Re-Docked with Polarized Ligand partial charges or (C) is Complex partial charges (C); (2HO) is the same as (2H) only the single water molecule hydrogens have been orientated by QM/MM calculation, before commencing a docking run. Then POSE – is polarized at the

Highest Ranking Glide GSCORE pose geometry. (CVDW), (EMODEL), (GSCORE) are the partitioned scoring function used by Glide. Plots G, I, K – The percentage over 21 of 74 molecules where the highest ranking/lowest scoring pose cluster (Grey) is a successful (below 2 RMSD). Then % increase when additionally accounting for 2nd Pose cluster (White) <2 RMSD when 1st pose is not. Plots H, J, L – The percentage over 21 of 74 molecules where a pose exists below 2 RMSD and the highest ranking/lowest scoring pose is a successful (below 2 RMSD - orange). *Clustering parameters set at <1.1 kcal/mol and <1 RMSD. Initial Docking request was to keep 15 poses.

Appendix C

0.5 angstrom class binned probability plots
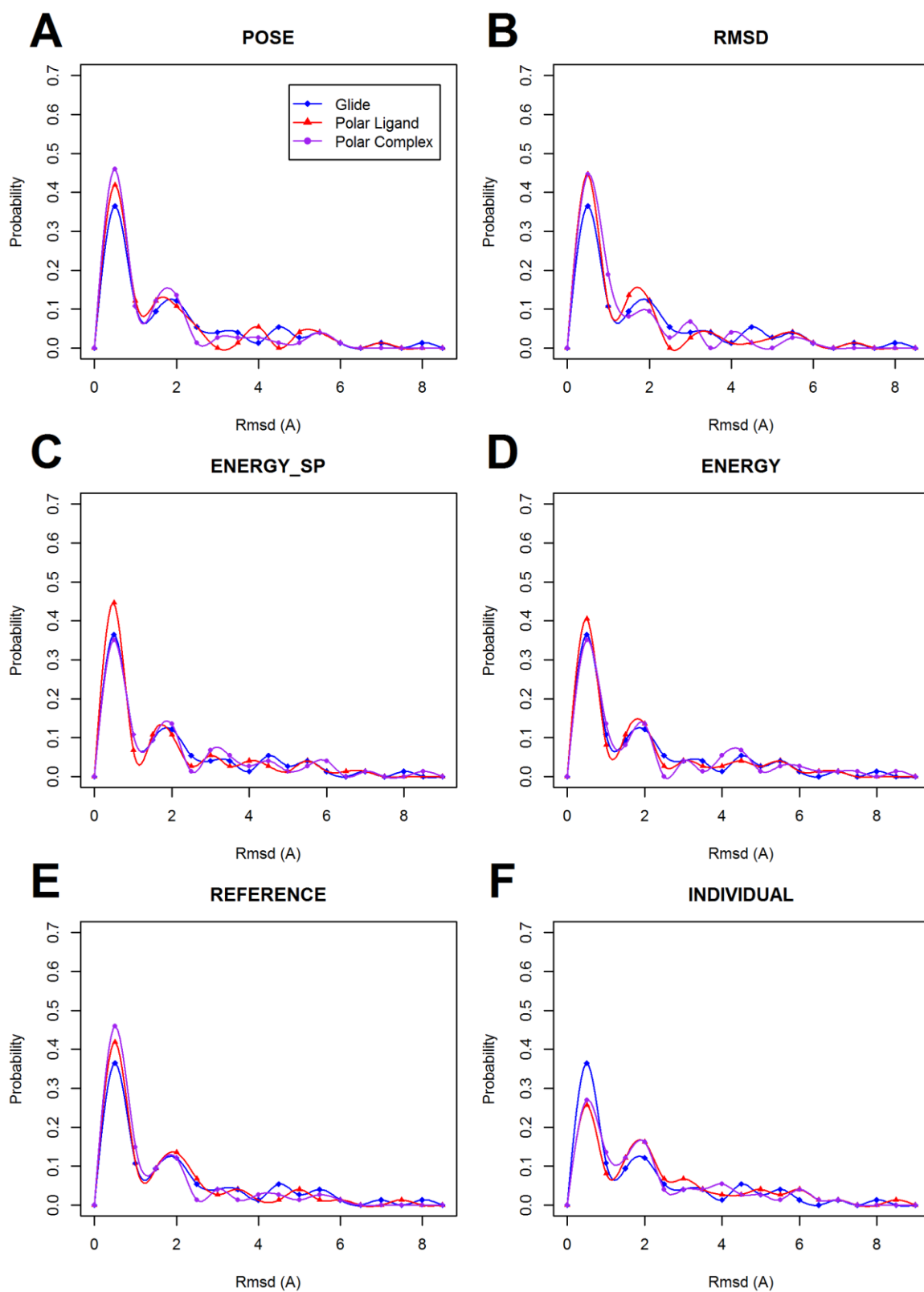
Figure C.1 Based on the probability that a pose is within a discrete 0.5 angstrom class bin, for a given value of x. The probability is calculated using variables for RMSD from the top-ranked CVDW poses.
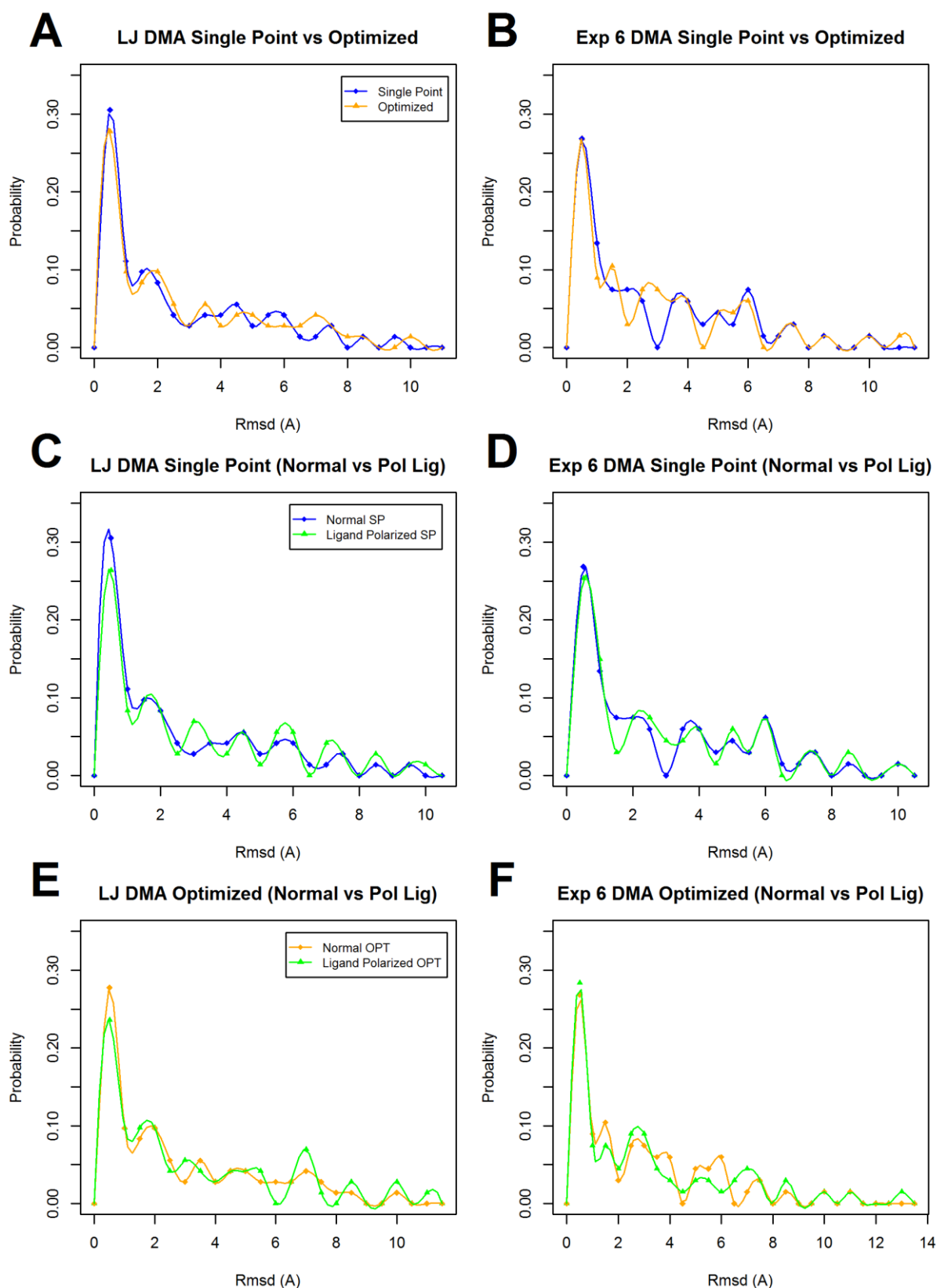
Figure C.2. RMSD across the 74 molecules top-ranked poses as a 0.5 angstrom binned probability plot (decimal) for DMA/Orient-based methods.
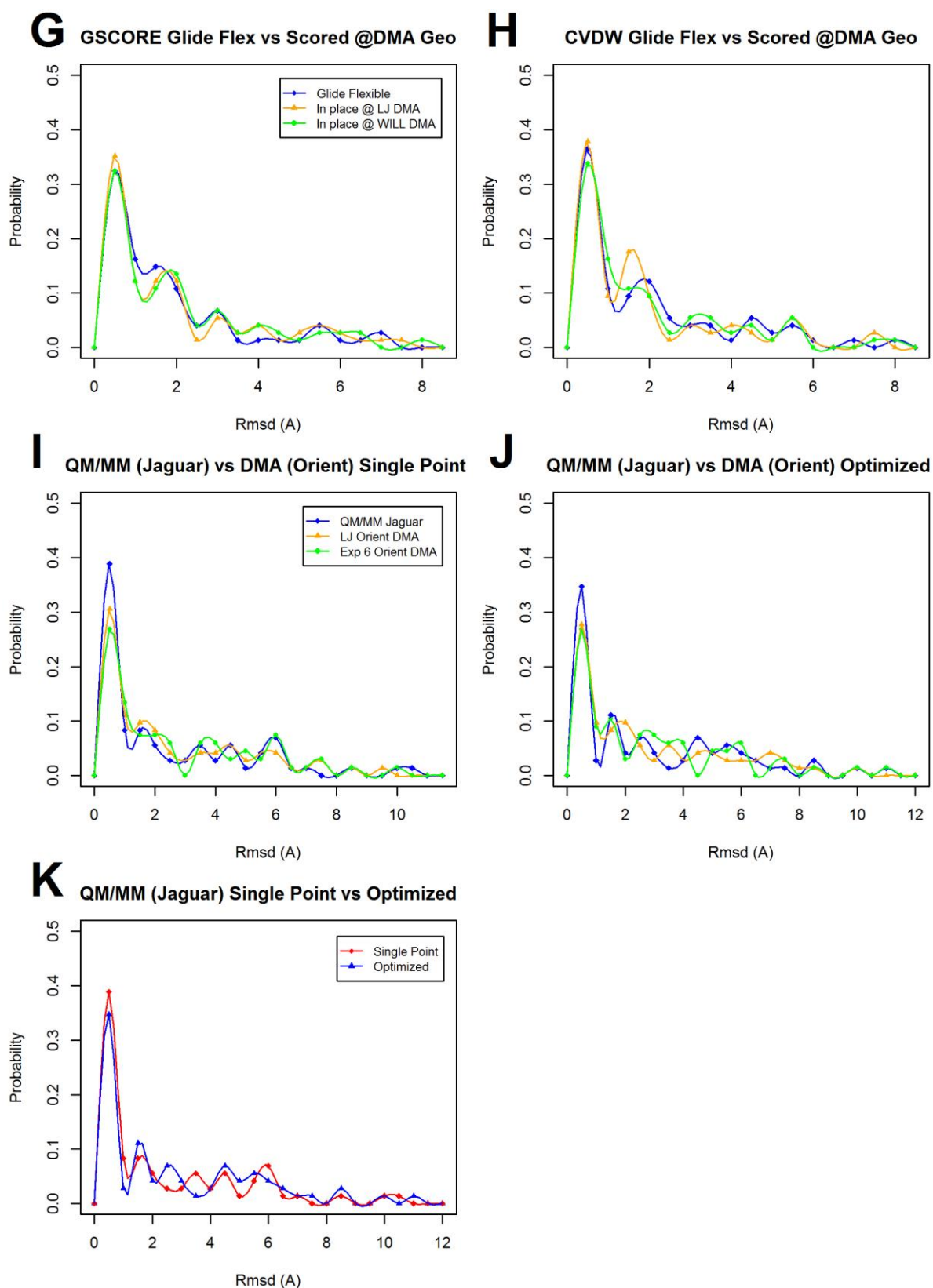
Figure C.3. RMSD across the 74 molecules top-ranked poses as a 0.5 angstrom binned probability plot(decimal) for Jaguar QM/MM and DMA/Orient-based methods and Glide score in place. WILL DMA includes exponential-6 repulsion.

Appendix D

Chapter 6 Supporting Info

Table D.1 Atomic Polarizabilities (α)

Atom type	α (Å^3)	Atom type	α (Å^3)	Atom type	α (Å^3)
C	1.433	**H**	0.387	**Zn**	0.387
CM	1.352	**HW**	0.387	**Mg**	0.387
CT	1.061	**HC**	0.387	S	3
NA	1.09	**HA**	0.387	P	1.538
O	0.569	**Fe**	0.387	Cl	2.315
OW	0.637	**Co**	0.387	Fl	2.315
OH	0.637	**Ni**	0.387	DM	0.001

Atom typing key: Carbons: C sp3 tetrahedral, CM sp2 aromatic, CT sp2 trigonal. Nitrogen: NA default. Oxygens: O non-hydroxyl, OW and OH hydroxyl. Hydrogens (Atom type **red** highlight). Metals Default (Atom type **purple** highlight).

Table D.2 Mean (μ) RMSD (Å) to X-ray experiment result half of other method, Full polarization vs No polarization.

pdb	Structures	Polarization	(%)	OPLS	(%)	BETTER Polar μ	Opls μ	WORSE Polar μ	Opls μ
1e2i	72	3	4.2	**4**	5.6	**1.5**	4	2.7	**0.8**
1pwm	528	**64**	12.1	24	4.5	**1.4**	5.6	5.7	**1.3**
1sqn	29	**2**	6.9	1	3.4	**1.1**	6.9	2.5	**0.7**
1t0l	56	**9**	16.1	1	1.8	**1.7**	4.8	**1.1**	**0.5**
1yv5	90	**42**	46.7	1	1.1	**1.1**	9.2	**0.9**	**0.4**
2gg7	90	**17**	18.9	3	3.3	**1.4**	5.2	5.9	2.2
1fv9	210	**32**	15.2	16	7.6	**1.6**	5.8	6.9	2.2
1gwq	1075	**99**	9.2	36	3.3	**1**	4.5	3.4	**1.2**
1n1m	182	**55**	30.2	4	2.2	**1.5**	6	7.8	2.5
1qwc	272	**44**	16.2	4	1.5	**1.5**	5.8	6.8	2.5
1w cc	1122	**240**	21.4	23	2	2.9	8.9	7.5	2.6
1s39	210	**39**	18.6	6	2.9	**0.7**	3.9	6	**1.6**
2c90	56	**7**	12.5	5	8.9	**1.4**	6.8	6	**1.9**
2jjc	866	**132**	15.2	29	3.3	**1.2**	5.4	4.4	**1.1**
2ohk	380	**55**	14.5	10	2.6	**1.8**	5.8	6.1	**1.9**
Mean Total	**5238**	**840**	17.2	**167**	3.6	1.5	5.9	4.9	1.6

Key: 3rd and 5th Column compare the number of structures from column 2 that half the rmsd of the other method, in these columns **black bold** indicates polarization is better, **red bold** indicates a refractory situation where polarization is worse. The last four columns report the mean rmsd (both methods) when polarization is better and worse RMSD (Å). **Black bold** and **red bold** indicate mean within 2 RMSD, for better or worse respectively. Last row: column means are in black, Totals in **purple**.

Table D.3 Evaluation of criteria RMSD to X-ray experiment result <2 Å for one method and >4 Å for the other method, Full polarization vs No polarization.

pdb source	Structures	Polarization	(%)	OPLS	(%)	BETTER Polar μ	OPLS μ	WORSE Polar μ	OPLS μ
1e2i	72	0	0	0	0	-	-	-	-
1pwm	528	**33**	6.2	8	1.5	1.2	6.7	8.9	1.2
1sqn	29	**2**	6.9	0	0	1.1	6.9	0	0
1t0l	56	**7**	12.5	0	0	1.6	4.8	0	0
1yv5	90	**35**	38.9	0	0	0.9	9.7	0	0
2gg7	90	**5**	5.6	1	1.1	1.3	6.4	5.7	1.3
1fv9	210	**15**	7.1	5	2.4	1.4	6.4	6	1.2
1gwq	1075	**45**	4.2	4	0.4	1	6	5.7	1.4
1n1m	182	**27**	14.8	0	0	1	6.4	0	0
1qwc	272	**21**	7.7	0	0	1.3	6.2	0	0
1w cc	1122	**68**	6.1	3	0.3	1.4	6.5	8.3	1.3
1s39	210	**16**	7.6	3	1.4	0.6	4.5	7.2	1.7
2c90	56	**3**	5.4	2	3.6	0.7	7.2	6.4	1.7
2jjc	866	**62**	7.2	10	1.2	1	6.8	6.8	1.4
2ohk	380	**20**	5.3	3	0.8	1.3	5.8	7.4	1.2
Mean Total	**5238**	**359**	9.0	39	0.8	1.1	6.0	4.2	0.8

Key: 3rd and 5th Column compare the number of structures from column 2 that have RMSD <2 Å for one method and RMSD >4 Å for the other method in relation to X-ray experiment result, in these columns **black bold** indicates polarization is better**, red bold** indicates a refractory situation where polarization is worse e.g., no overall refractory cases. The last four columns report the mean (μ) RMSD (both methods) when polarization is better and worse RMSD (Å). **Black bold** and **red bold** indicate mean within 2 RMSD, for better or worse respectively. Last row: column means are in black, Totals in **purple**.

Table D.4 Bootstrapping population <u>parameter ($\hat{\theta}$) statistics </u>for Table 6.2 Glide cross docking, where n=1000, for the mutual pairs cross docking <2 Å RMSD.

Name	pdb	Parent (%)	$\mu\hat{\theta}$	σ bias	σ	SE
Herpes simplex virus type 1	1e2i	27.778	27.547	0.231	7.231	0.448
Aldose reductase	1pwm	23.106	23.167	0.061	2.575	0.16
Progesterone receptor	1sqn	50	50.6	0.6	17.94	1.112
Isocitrate dehydrogenase	1t0l	7.143	6.989	0.154	4.976	0.308
Farnesyl diphosphate synthase	1yv5	17.778	17.351	0.427	5.705	0.354
methionine aminopeptidase	2gg7	11.111	11.138	0.027	4.655	0.289
microurokinase	1fv9	20	19.81	0.19	3.816	0.237
Oestrogen receptor	1gwq	47.416	47.39	0.026	2.259	0.14
Dipeptidyl peptidase IV	1n1m	7.692	7.804	0.112	2.798	0.173
nitric oxide synthase	1qwc	4.412	4.451	0.039	1.757	0.109
Cyclin-dependent kinase 2	1w cc	3.743	3.725	0.018	0.8	0.05
tRNA-Guanine transglycosylase	1s39	48.571	48.281	0.29	4.795	0.297
Thrombin	2c90	10.714	10.461	0.254	5.786	0.359
HSP90	2jjc	13.626	13.732	0.106	1.636	0.101
β−secretase 1	2ohk	2.105	2.128	0.023	1.045	0.065

Key: Here, σ bias is the difference between the parent group success and the population parameter mean $\mu\hat{\theta}$, σ is the estimated standard deviation from $\mu\hat{\theta}$, SE is the standard error in the population parameter $\hat{\theta}$ at n=1000.

Table D.5 Bootstrapping population parameter p$_i$ statistics for Table 6.2 Ligand cross docking, where n=1000, for the mutual pairs cross docking <2 A RMSD.

Name	pdb	Parent (%)	$\mu\hat{\theta}$	σ bias	σ	SE
Herpes simplex virus type 1	1e2i	22.222	21.897	0.325	6.824	0.423
Aldose reductase	1pwm	26.136	26.176	0.04	2.789	0.173
Progesterone receptor	1sqn	50	50.288	0.288	18.501	1.147
Isocitrate dehydrogenase	1t0l	14.286	14.286	0	6.364	0.394
Farnesyl diphosphate synthase	1yv5	20	19.651	0.349	5.827	0.361
methionine aminopeptidase	2gg7	24.444	24.573	0.129	6.287	0.39
microurokinase	1fv9	29.524	29.32	0.204	4.315	0.267
Oestrogen receptor	1gwq	50.112	50.079	0.033	2.311	0.143
Dipeptidyl peptidase IV	1n1m	13.187	13.286	0.099	3.527	0.219
nitric oxide synthase	1qwc	5.147	5.188	0.04	1.88	0.117
Cyclin-dependent kinase 2	1w cc	5.704	5.704	0	0.978	0.061
tRNA-Guanine transglycosylase	1s39	61.905	61.868	0.037	4.788	0.297
Thrombin	2c90	7.143	6.875	0.268	4.726	0.293
HSP90	2jjc	20.323	20.47	0.147	1.933	0.12
β−secretase 1	2ohk	5.263	5.373	0.11	1.619	0.1

Key: Here, σ bias is the difference between the parent group success and the population parameter mean $\mu\hat{\theta}$, σ is the estimated standard deviation from $\mu\hat{\theta}$, SE is the standard error in the population parameter $\hat{\theta}$ at n=1000.

Table D.6 Bootstrapping population parameter p$_i$ statistics for Table 6.2 Full cross docking, where n=1000, for the mutual pairs cross docking <2 A RMSD.

Name	pdb	Parent (%)	$\mu\hat{\theta}$	σ bias	σ	SE
Herpes simplex virus type 1	1e2i	33.333	33.117	0.217	7.755	0.481
Aldose reductase	1pwm	26.515	26.524	0.009	2.798	0.173
Progesterone receptor	1sqn	62.5	63.1	0.6	17.729	1.099
Isocitrate dehydrogenase	1t0l	25	24.986	0.014	8.057	0.499
Farnesyl diphosphate synthase	1yv5	75.556	75.456	0.1	6.314	0.391
methionine aminopeptidase	2gg7	22.222	22.302	0.08	6.254	0.388
microurokinase	1fv9	32.381	32.318	0.063	4.551	0.282
Oestrogen receptor	1gwq	52.584	52.599	0.015	2.324	0.144
Dipeptidyl peptidase IV	1n1m	23.077	23.195	0.118	4.537	0.281
nitric oxide synthase	1qwc	12.5	12.576	0.076	2.84	0.176
Cyclin-dependent kinase 2	1w cc	6.774	6.793	0.019	1.061	0.066
tRNA-Guanine transglycosylase	1s39	67.619	67.489	0.13	4.642	0.288
Thrombin	2c90	7.143	7.075	0.068	4.926	0.305
HSP90	2jjc	23.788	23.895	0.108	2.048	0.127
β−secretase 1	2ohk	8.947	8.933	0.015	2.005	0.124
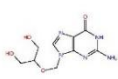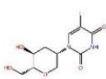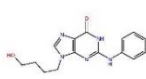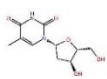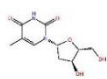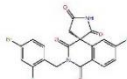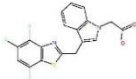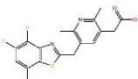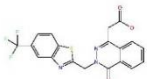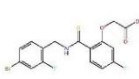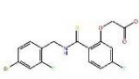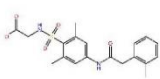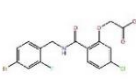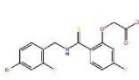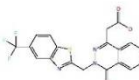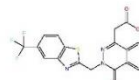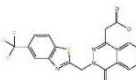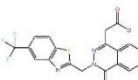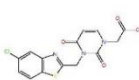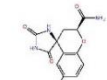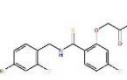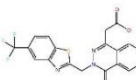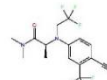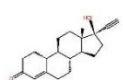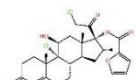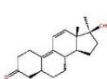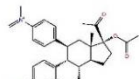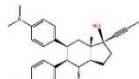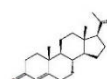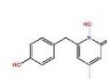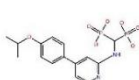
Key: Here, σ bias is the difference between the parent group success and the population parameter mean $\mu\hat{\theta}$, σ is the estimated standard deviation from $\mu\hat{\theta}$, SE is the standard error in the population parameter $\hat{\theta}$ at n=1000.

Table D.7 Initial proteins from fragment validation set that met protein selection criteria

PDB	Recepter	Ligand	Name or Synonym	www.drugbank.ca
1e2i	Herpes simplex virus type 1	**APS**	9-Hydroxypropyladenine,S-Isomer	
1pwm	Aldose reductase	**FID**	Fidarestat	
1sqn	Progesterone receptor	**NDR**	Norethisterone	Approved
1t0l	Isocitrate dehydrogenase	**ICT**	Isocitric Acid	
1yv5	Farnesyl diphosphate synthase	**RIS**	Risedronate	Approved
2gg7	methionine aminopeptidase	**U14**	3-(5-amino-3-imino-3H-pyrazol-4-ylazo)-benzoic acid	Experimental
1fv9	microurokinase	**172**	2-amino-5-hydroxy-benzidazole	
1gwq	Oestrogen receptor	**ZTW**	Raloxifene Core	Experimental
1n1m	Dipeptidyl peptidase IV	**A3M**	2-Amino-3-Methyl-1-Pyrrolidin-1-Yl-Butan-1-One	Experimental
1qwc	nitric oxide synthase	**14W**	N-(3-(Aminomethyl)Benzyl) Acetamidine	Experimental
1wcc	Cyclin-dependent kinase 2	**CIG**	2-Amino-6-Chloropyrazine	Experimental
1s39	tRNA-Guanine transglycosylase	**AQO**	2-Aminoquinazolin-4(3h)-One	Experimental
2c90	Thrombin	**C1M**	1-(4-chlorophenyl)-1H-tetrazole	
2jjc	HSP90	**LGA**	Pyramidine-2-Amine	
2ohk	β−secretase 1	**1SQ**	1-Amino-Isoquinoline	

Unique |Initial proteins from Fragment validation set, 1e2i, 1pwm, 1sqn, 1t0l, 1yv5, 2gg7 originate from SERAPhiC (Favia et al., 2009). The rest are representative fragment complexes largely from the generations of leads or retrospective decomposition of leads (Congreve et al., 2008).

Table D.8 The 257 ligands used in the cross docking experiments.

title: 1E2I_APS	title: 3F0T_NCV	title: 1KI2_GA2	title: 1KI6_AHU	title: 1QHI_BPG	title: 1OF1_SCT
title: 1KIM_THM	title: 1E2L_TMC	title: 1E2J_THM	title: 1PWM_FID	title: 1PWL_BFI	title: 1Z3N_3NA
title: 4GCA_2X9	title: 2PD5_ZST	title: 3LD5_LDT	title: 3LZ5_LDT	title: 1EL3_I84	title: 3LQG_388
title: 2PDK_SBI	title: 3LBO_LDT	title: 3RX2_SLO	title: 2PDM_ZST	title: 2PDB_ZST	title: 2PDI_ZST
title: 2PDF_ZST	title: 2AGT_FID	title: 2PDQ_47D	title: 1EF3_FID	title: 3ONB_LDT	title: 2PDX_ZST
title: 2ACU_CIT	title: 2IPW_2CL	title: 2IS7_2CL	title: 3G8O_30X	title: 1SQN_NDR	title: 1SR7_MOF
title: 1E3K_R18	title: 4OAR_2S0	title: 2W8Y_486	title: 1A28_STR	title: 1T0L_ICT	title: 4L03_AKG
title: 4L04_AKG	title: 3MAS_ICT	title: 3INM_AKG	title: 3MAP_ICT	title: 4I3K_1BX	title: 4L06_AKG
title: 1YV5_RIS	title: 1YQ7_RIS	title: 4DEM_YS4	title: 2QIS_RIS	title: 4KPJ_210	title: 4N9U_RIS

title: 4KFA_ZOL title: 2RAH_11P title: 4NUA_RIS title: 3CP6_RSX title: 2GG7_U14 title: 3MAT_AHH

title: 2GG0_U11 title: 1XNZ_FCD title: 4A6V_IKY title: 1C27_NLP title: 2GU4_NLP title: 2P98_YE7

title: 2GTX_NLP title: 2P99_YE6 title: 1FV9_172 title: 3MWI_B25 title: 1EJN_AGB title: 1OWD_497

title: 2VNT_QGG title: 1SC8_2IN title: 1W0Z_SI1 title: 2O8U_BEN title: 4FU7_1UP title: 4H42_11E

title: 4JNI_PZH title: 1GJ7_132 title: 1GI8_BMZ title: 1OWE_675 title: 1O5A_696 title: 1GWQ_ZTW

title: 2IOK_IOK title: 3DT3_369 title: 1ERE_EST title: 3ERD_DES title: 1QKU_EST title: 1A52_EST

title: 4PXM_EST title: 4Q50_OHT title: 2P15_EZT title: 3HLV_J2Z title: 1L2I_ETC title: 3OS8_KN0

title: 1ERR_RAL title: 2QXS_RAL title: 3UUC_0D1 title: 1G50_EST title: 4DMA_0L8 title: 4IU7_1GM

title: 3Q95_ESL title: 2QA8_GEN title: 3L03_4OH title: 1GWR_EST title: 1UOM_PTI title: 4MG7_27H

title: 3UUA_0CZ

title: 2QZO_KN1

title: 2OCF_EST

title: 4PP6_STL

title: 1X7E_244

title: 1SJ0_E4D

title: 4MG5_27E

title: 4MG8_27J

title: 3UU7_2OH

title: 1QKT_EST

title: 1PCG_EST

title: 1N1M_A3M

title: 2QT9_524

title: 2RGU_356

title: 2JID_GVB

title: 2QJR_PZF

title: 3NOX_6A5

title: 2RIP_34Q

title: 4PV7_CJP

title: 1X70_715

title: 3CCB_B2Y

title: 4A5S_N7F

title: 2ONC_SY1

title: 3Q8W_AZV

title: 3QBJ_NXZ

title: 1QWC_14W

title: 1QW6_3AR

title: 1OM5_INE

title: 1P6K_DP1

title: 1LZX_HAR

title: 2G6H_HAR

title: 3JX0_JI4

title: 3FC5_HAR

title: 3N65_XFJ

title: 3JWS_JI4

title: 1ZZQ_DP9

title: 3DQR_JI2

title: 4JSG_Q10

title: 4CX5_HW8

title: 1ZZR_DP9

title: 2G6J_HAR

title: 3JX6_JI5

title: 1WCC_CIG

title: 1FIN_ATP

title: 2W17_I19

title: 3EZR_EZR

title: 3PJ8_404

title: 1OIT_HDT

title: 1QMZ_ATP

title: 1VYW_292

title: 3BHT_MFR

title: 1JST_ATP

title: 4BCQ_TJF

title: 1PF8_SU9

title: 1OIR_HDY

title: 1H1P_CMG

title: 1OGU_ST8

title: 4CFU_2WC

title: 1GZ8_MBP

title: 4EOQ_ATP

title: 4ERW_STU

title: 4EOS_1RO

title: 2IW6_QQ2

title: 1E9H_INR

title: 4EOI_1RO

title: 4EOO_ATP

title: 2CJM_ATP

title: 4EOP_1RO

title: 1GII_1PU

title: 4I3Z_ADP

title: 3QHR_ADP

title: 4EON_1RO

title: 4EOJ_ATP

title: 2IW8_4SP

title: 4EOM_ATP

title: 4EOK_4SP

title: 1S39_AQO

title: 1R5Y_DQU

title: 1ENU_APZ

title: 4Q8V_2YX

title: 4PUM_2WU

title: 2POT_GUN

title: 4GHR_ITE

title: 1EFZ_PRF

title: 1Y5V_NE8

title: 4Q8P_CKR

title: 4HSH_QEI

title: 4GCX_PRF

title: 1PXG_PRF

title: 4HQV_QEI

title: 2NQZ_PRF

title: 2C90_C1M

title: 1JWT_BLI

title: 2BDY_UNB

title: 1D9I_00P

title: 1GJ5_130

title: 1NM6_L86

title: 1D6W_00R

title: 1WBG_L03

title: 2JJC_LGA

title: 2XDU_LGA

title: 4EGH_0OY

title: 3FT5_MO8

title: 2XJG_XJG

title: 4NH7_PIE

title: 2XJJ_L81

title: 2YE7_2GA

title: 2YE2_XQI

title: 3B24_B2J   title: 3BM9_BXZ   title: 3K98_1RC   title: 2YEE_2EC   title: 3EKO_PYU   title: 3D0B_SNX

title: 1UY6_PU3   title: 3HHU_819   title: 2BSM_BSM   title: 3QDD_94M   title: 1UYI_PUZ   title: 2YI0_YI0

title: 3R4M_WOE   title: 4AWO_99B   title: 4EEH_HH6   title: 2CCS_4BH   title: 3K97_4CD   title: 4B7P_9UN

title: 1BYQ_ADP   title: 3INW_JZB   title: 1YC1_4BC   title: 2OHK_1SQ   title: 2VA5_C8C   title: 3CKP_012

title: 2OF0_CMZ   title: 3TPR_5HA   title: 3HVG_EV0   title: 3L58_CS5   title: 2ZJN_F1N   title: 4YBI_4B2

title: 3VV6_B00   title: 2QU2_251   title: 2ZJH_F1H   title: 2ZJI_F1I   title: 3BRA_AEF   title: 1TQF_32P

title: 3QI1_C6A   title: 2ZJJ_F1J   title: 3KMX_G00   title: 2QK5_CS5   title: 4L7G_1W0

Titles show pdb code and ligand id separated by underscore.