Practical Considerations in Virtual Screening and Molecular Docking

27

Michael Berry[1], Burtram Fielding[2], and Junaid Gamieldien[1]

South African National Bioinformatics Institute, University of the Western Cape, Cape Town, South Africa[1]

Department of Medical Biosciences, University of the Western Cape, Cape Town, South Africa[2]

1 INTRODUCTION

In silico virtual screening, or high-throughput virtual screening (HTVS), has yielded an excellent complement to the time-consuming and expensive experimental techniques of high-throughput screening. The ability to virtually screen compound libraries to improve enrichment of ligands progressed to experimental validation has provided countless lead compounds. HTVS computationally screens large databases of virtual compounds that either possess similarity toward a known inhibitor (ligand-based) or complementarity toward the solved receptor structure (structure-based; Shoichet, 2004). This allows researchers to screen large databases or compound libraries in order to identify a highly focused subset from which actives can be confirmed experimentally (Ripphausen *et al.*, 2011) and, in the case of molecular docking, can predict the binding pose, thereby simplifying future lead optimization (Joseph-McCarthy *et al.*, 2007). HTVS is especially attractive to academic facilities, occasionally in parallel with HTS, as fiscal expenses of pure HTS screens are often too large for academic budgets (Zhu *et al.*, 2013). Improvements in computer hardware and the availability of relatively inexpensive clusters have also increased the speed of HTVS, contributing to its gain in popularity (Anderson, 2003). Because it mostly eliminates cost-of-ownership associated with computing infrastructure, it is likely that cloud computing will further contribute to this uptake.

Molecular docking was first described in 1982 (Kuntz *et al.*, 1982) and has since become the central idea in structure-based virtual screening. It comprises two major tasks for which separate algorithms are used. The sampling algorithm predicts the many confirmations, referred to as *poses,* which the ligand can assume within the binding or active pocket. A scoring function then predicts the binding energies between the ligand and receptor for each predicted pose. The generated binding poses are then ranked based on their binding energies, where the top-ranked pose should correspond to the correct confirmation of the ligand. Scoring functions

are, therefore, also capable of filtering through, and ranking, large databases of compounds in virtual screening, where the highest-ranked binding energies should correspond to a potential lead (Phatak *et al.*, 2009).

Used on its own, molecular docking is, however, plagued with weaknesses. The static nature of the receptor is a primary fault where the dynamic nature of the biological structures is not considered. Limitations in sampling algorithms and imperfections in scoring functions also lead to the generation of both false positives and false negatives (Lill, 2011; Wang *et al.*, 2003; Brooijmans and Kuntz, 2003; Alvarez, 2004) and the requirement for training sets in various algorithms often leads to accuracy being highly target dependent (Warren *et al.*, 2006). These inherent flaws are further exacerbated by user oversights and errors. This chapter will detail several practical aspects to consider prior to commencing a virtual screening study, while simultaneously providing a theoretical explanation of docking and scoring. This review will provide guidelines, but there is no "one-rule-fits-all" in molecular docking. Most docking programs have varying methods to deal with each topic discussed and describing details of each program is outside the scope of this review. It must also be taken into account that every receptor is different and the ability to replicate experimental and physiological findings is highly system dependent.

2 RECEPTOR STRUCTURE PREPARATION

Although most receptor preparation tools accurately complete processes that were not undertaken during X-ray crystal structure refinement, it is important to understand these processes and make adjustments where necessary. The most common receptor preparation procedures include adding hydrogens and atom-type charges, but it is also important to ensure that missing side-chains are added, missing bonds and molecule chain breaks are detected and fixed, bond orders are assigned, and where alternate locations are present, the atoms with highest frequencies must be selected. Other, more complicated, procedures in receptor preparation include accurate prediction of protonation states and identifying which water molecules (if any) should remain in the receptor structure. All of these procedures maximize the biological realism in the modeled system, which leads to the identification of a higher proportion of true bioactives.

2.1 PROTONATION STATES

The resolution of most crystal structures does not provide information on the location of hydrogens, commonly referred to as the *protonation state* (ten Brink and Exner, 2009). The accurate prediction of the correct protonation state, especially within the binding interface, is crucial to accurately predict the correct binding mode and, to a greater extent, binding affinity (Kalliokoski *et al.*, 2009; Fornabaio *et al.*, 2003; Onufriev and Alexov, 2013). This incorrect prediction of binding mode and affinity will inevitably lead to the identification of false positives, while true bioactives are

missed (Onufriev and Alexov, 2013). It is notable to point out that force field–based scoring functions are more susceptible to incorrect protonation states in comparison to knowledge-based scoring functions (Onufriev and Alexov, 2013). Assigning the incorrect protonation states further alters the state of hydrogen bond donors and acceptors, which substantially limits the accurate prediction of protein-ligand interactions (Polgár and Keserü, 2005).

Side-chains of ionizable amino acids can further vary their protonation states within a receptor depending on the local environment and pH. Ligand binding can also be accompanied by proton gain or release (Petukh *et al.*, 2013) but this is almost never incorporated into a molecular docking study (Onufriev and Alexov, 2013). One study pointed out that a residue's protonation state cannot be accurately replicated, as protons are not static and are readily transferred between molecules (Fornabaio *et al.*, 2003). The quantum mechanical simulations necessary to replicate proton movements are far beyond the scope and capabilities of molecular docking and at best, the protonation state, or an ensemble of protonation states, that is most suitable to ligand binding must be identified.

Histidine (His) provides a unique problem in terms of residue protonation, as it can be protonated in three different conformations. The imidazole ring of the His side-chain can be protonated in a neutral confirmation at the ε-nitrogen or the δ-nitrogen or in a charged (+1) conformation where both the ε- and δ-nitrogens are protonated (Kim *et al.*, 2013). To further complicate the correct conformation of the imidazole side-chain ring, ambiguities in crystal structures often switch the carbon and nitrogen, creating an additional three rotameric conformations, termed "flipped" (Glusker *et al.*, 1994). His also represents a weaker base, and for this reason, determining the protonation state is more complicated than for other ionizable residues and must be determined individually (Waszkowycz *et al.*, 2011). In the case of His, analysis of hydrogen bonding networks is likely to yield the most detail about the correct side-chain protonation.

The dynamic nature of a receptor means the protonation states of ionizable residues are constantly changing. In order to accurately predict the conformation of a ligand binding to a receptor, the protonation state of the receptor must be relevant to the bound conformation and in correspondence with crystal data (*i.e.,* absence of steric clashes and hydrogen bonds occurring at expected locations) and in accordance with the pH of the experimental conditions. Assigning protonation states to Asp, Glu, Arg, and Lys during receptor preparation is generally straightforward, with deprotonated acids (Asp and Glu) and protonated bases (Arg and Lys) (Kim *et al.*, 2013; Waszkowycz *et al.*, 2011). This is, however, a generalization and not a rule, and the microenvironment of the residue and physiological pH of the receptor must be taken into careful consideration. Calculating the theoretical pK_a of these residues at the physiological pH is possibly the most straightforward mechanism to determine or estimate their protonation state (Polgár and Keserü, 2005).

As scoring functions are highly dependent on the correct receptor protonation state, it can be assumed that a scoring function will favor the correct protonation state by scoring it above the incorrect state (Onufriev and Alexov, 2013). This provides a

mechanism to accurately predict the correct protonation state within an ensemble of pregenerated receptor states. The correct replication of hydrogen bond positions between ligand and receptor, as seen in the crystal structure or detailed in the literature, will further suggest the accurate placement of residue protons (Krieger et al., 2012; Hooft et al., 1996). Observable steric clashes between a ligand and receptor, after protonation, will further suggest incorrect proton placement (Word et al., 1999; Krieger et al., 2012). This approach will only account for ionizable groups within the binding interface and will not be able to account for the entire receptor, but this remains a far more attractive strategy than ignoring the issue entirely.

In summary, in order to accurately approximate a receptor's protonation state, the identification of its physiological pH is key. Second, calculated pK_a values for ionizable residues enables determination of the protonation state according to the given pK_a at the specified pH. Third, crystal structures and known, experimentally identified bioactives can yield a wealth of knowledge on the protonation state of a receptor by scoring function analysis and inspection of steric clashes and hydrogen bonding networks between ligand and receptor. Given these guidelines, the techniques used to accurately predict the correct protonation state of a receptor are largely dependent on the class of receptor being studied. For this reason, the techniques applied must be accurately verified for the receptor under investigation before virtual library screening.

2.2 SELECTING IMPORTANT ACTIVE SITE WATER MOLECULES

Active site water molecules are key determinants in ligand-receptor binding (Thilagavathi and Mancera, 2010; Barillari et al., 2007). Not only can they mediate hydrogen bonding between ligand and receptor, but their contribution to entropic and enthalpic changes are significant (Lie et al., 2011; Cheng et al., 2012; Kroemer, 2007). In a virtual screening context, the addition of water (an explicit solvent) is frequently neglected, as the intensive computational simulations required does not permit the rapid screening required for large libraries, often seen in high-throughput virtual screens, and accounting for water molecules in docking remains a significant challenge (Cheng et al., 2012; Huang and Shoichet, 2008; Schneider and Fechner, 2005).

The position of water molecules within an active site are also highly variable (Santos et al., 2009), and to account for them as static in nature would be biased toward ligands that complement the specific orientation and prejudice those that would physiologically replace the water molecules, leading to a drastic increase in false negatives (Kroemer, 2007). Several reports claim to more accurately predict the binding mode of crystal structure inhibitors by incorporating water molecules within the active site (Lemmon and Meiler, 2013). While these studies do possess a high degree of merit, the inclusion of waters within the active site greatly decreases the volume of the pocket and thereby the possible conformations that the ligand may assume, which is further biased toward the correct conformation (Lie et al., 2011; Hartshorn et al., 2007). As there is a constant compromise between speed and

accuracy in a high-throughput virtual screen, however, the presence of active site waters can greatly increase ligand enrichment. It is, therefore, important to determine which waters, if any, must be kept during a virtual screen and exclude those that are nonessential.

An initial step to assess the importance of active site waters would be to attempt to replicate the binding mode of experimental structures in the absence of explicit waters. If the accuracy is diminished by the absence of waters in the binding site, it is important to select which waters are pivotal to binding. Waters that are not hydrogen bonded to the receptor, and those that are located outside the binding pocket (more than 5Å), will obviously have little effect on ligand binding and can therefore be removed (Huang and Shoichet, 2008). Waters that possess three hydrogen bonds with the receptor, or those with low B-factors, are likely to be highly stable within the pocket and should be included in docking studies, as these waters may prove difficult to displace by ligand binding and likely function to stabilize the protein binding site (Yang *et al*., 2006; Hornak *et al*., 2006). Waters that form hydrogen bond bridges between the ligand and receptor are also likely to be important in ligand binding. This may, however, be highly ligand-specific and its importance in virtual screening, where a diverse set of ligand classes are under study, must be properly assessed and validated. Where essential water molecules are included in a virtual screen they should, ideally, be treated as flexible (Huang and Shoichet, 2008). It is also important to bear in mind that the accuracy of a docking algorithm may be highly dependent on the parameterization of the algorithm and suitability toward the class of receptor and inhibitor, which will be discussed later in this chapter.

3 ACCURATELY PREDICTING THE POSE OF SOLVED CRYSTAL STRUCTURES AND DIFFERENTIATING DECOYS FROM ACTIVES

It is commonly accepted that there is no "first-in-class" algorithm or molecular docking software for the prediction of correct ligand-binding pose or relative free energy of binding. Molecular docking algorithms are often calibrated on a training set of experimental ligand-protein complexes and accuracy of these docking programs is often highly dependent on the training set used (Ballester and Mitchell, 2010). This highlights the importance of confirming that the docking software used for virtual screening is capable of replicating the binding mode of known, experimental inhibitors for the class of receptor studied (Lim *et al*., 2011; Kroemer, 2007). To improve ligand enrichment in a virtual screening context, the docking algorithm selected must be properly validated for the class of receptor under investigation. Of course, in a virtual screen, where hundreds of thousands to millions of compounds are potentially being screened, validating for each class of potential inhibitor would be impossible, but accurate validation must be undertaken with the largest obtainable data set of true experimental leads, where the binding pose is known. A root-mean-square deviation (RMSD) below 2Å for heavy atoms (excluding hydrogens) between the experimental structure and predicted pose of docking is a well-defined benchmark to

assess the accuracy of molecular docking sampling algorithms (Houston and Walkinshaw, 2013).

A highly useful benchmarking strategy and metric to gauge the success of a molecular docking program is the ability to differentiate true actives from decoys. The Database of Useful Decoys–Enhanced (DUD-E; http://dude.docking.org/generate) can generate decoys for an active compound (Mysinger *et al.*, 2012). DUD generates decoys based on cheminformatic properties, including molecular weight, logP, number of rotatable bonds, and number of hydrogen bond donors and acceptors. As these decoys are not intended to bind to the target receptor, they are topologically distinct from the active inhibitors, thereby serving as suitable negative controls. The enrichment of the docking program can be assessed by its ability to rank true actives above decoy ligands (Mysinger *et al.*, 2012).

4 SIDE-CHAIN FLEXIBILITY AND ENSEMBLE DOCKING

Virtual screening simulations are typically performed on static structures, and it has previously been demonstrated that the use of a *holo* (ligand-bound) conformation provides better enrichment when compared to *apo* or homology modeled receptors (McGovern and Shoichet, 2003). Given this, addressing protein flexibility can substantially improve enrichment but remains one of the most challenging aspects of molecular docking. There are currently two approaches to incorporate the dynamic nature of protein structures; flexible receptor docking and ensemble docking (Lill, 2013). These approaches have shown to improve enrichment in docking studies (Craig *et al.*, 2010), but the compromise between speed and accuracy must be heavily weighted in high-throughput virtual screens.

Flexible docking most often only incorporates side-chains of residues within the active site and therefore does not cover the dynamic range of protein conformations (Meng *et al.*, 2011). It has been demonstrated that only a small number of side-chains within a binding pocket undergo structural changes upon ligand binding. This study suggested that, within 85% of studied receptors, only three or fewer side-chains exhibited movements upon ligand binding and further developed a scale of side-chain flexibility (Lys > Arg, Gln, Met > Glu, Ile, Leu > Asn, Thr, Val, Tyr, Ser, His, Asp > Cys, Trp, Phe; Najmanovich *et al.*, 2000). Utilizing this scale, it may be possible to identify which side-chains within a pocket must be made flexible and which may be left static, although the ability to accurately enrich active ligands must be displayed.

In ensemble docking, the ensemble of rigid structures can be generated by a molecular dynamic simulation where snapshots are isolated from the trajectory or when several structures are available from crystallography or nuclear magnetic resonance (NMR) experimental studies. There are two distinct classes of ensemble docking. In the first method, several protein conformations are generated prior to a docking screen and each ligand is docked into each receptor independently (Carlson, 2002; Carlson and McCammon, 2000; Barril and Morley, 2005), thereby

introducing receptor flexibility by multiple docking runs (Henzler and Rarey, 2010). This is, of course, computationally inefficient and the time required to conduct a screen increases with every protein structure included in the ensemble. The conformational diversity is also limited to the conformational representations included in the ensemble (B-Rao et al., 2009). The second method assesses an ensemble of protein structures in a single docking screen (B-Rao et al., 2009). This method either unites ensemble structures or uses a receptor grid averaged over all protein structures, and therefore reduces computational cost considerably (Totrov and Abagyan, 2008; Knegtel et al., 1997; Henzler and Rarey, 2010). To identify a suitable ensemble of structures to incorporate in a docking run, an enrichment docking screen of known actives can be performed. Both ensemble and flexible receptor docking is described in greater detail in several reviews (Cavasotto and Abagyan, 2004; Carlson, 2002; Therrien et al., 2014; Henzler and Rarey, 2010).

Molecular dynamic (MD) simulations is considered to be the most accurate method to determine the stability of a ligand within a binding pocket, while accounting for full side-chain and backbone flexibility and incorporating solvent effects (Marco and Gago, 2007; Alonso et al., 2006). Several docking studies have utilized MD simulations to confirm results obtained from docking studies. However, the intense computational costs make it practical for only a small set of ligands (Österberg and Åqvist, 2005; Han, 2012; Mukherjee et al., 2011; Segura-Cabrera et al., 2013).

5 CONSENSUS DOCKING

Scoring functions have been highlighted as the major weakness of molecular docking (Yang et al., 2005; Warren et al., 2006; Wang et al., 2003). As these functions are solely responsible for selecting and ranking the correct ligand pose within the binding site from the many possible conformations generated by the sampling algorithm, it can potentially lead to identification of an incorrect pose. The integration of a consensus approach to sampling and scoring, incorporating several algorithms to each task, has shown to greatly improve ligand enrichment in virtual screening and identifying the correct pose of experimental structures (Teramoto and Fukunishi, 2007; Houston and Walkinshaw, 2013; Kukol, 2011; Yang et al., 2005; Charifson et al., 1999; Plewczynski et al., 2011). Consensus scoring compensates for deficiencies in individual scoring functions and thereby improves the overall performance (Teramoto and Fukunishi, 2007), with the inclusion of a single extra scoring function being sufficient to improve binding affinity predictions (Chang et al., 2010). A similar technique to consensus scoring is the approach of consensus sampling, which is less well characterized. A recent study by Houston and Walkinshaw, 2013 utilized three sampling algorithms from Dock (Ewing et al., 2001), Autodock (Morris et al., 2009), and Autodock Vina (Trott and Olson, 2010) to identify the experimental pose of a diverse set of ligands. The study achieved an accuracy of 82%, compared to the 55%–64% accuracy of using a single algorithm (Houston and Walkinshaw, 2013). In

this study, a consensus result was confirmed when independently predicted poses were within an RMSD cutoff of 2Å, the same distance defined as correct sampling in comparison to experimental structures (Houston and Walkinshaw, 2013).

The approach of employing several algorithms to identify the correct pose with subsequent consensus scoring to identify top-ranked ligands can greatly improve the enrichment rate in a virtual screening context. The major cost of this approach is the increase in false negatives, which are therefore missed and do not progress to experimental testing. In an academic setting, or a lab where resources are limited, this is an acceptable consequence, as the quality of the results is more vital in a virtual screening context. The improvement in the identified hit list, with a decrease in false positives and subsequent decrease in resource waste, would largely compensate for the increase in false negatives (Houston and Walkinshaw, 2013).

6 MM-GBSA

Various elements of binding free energy, including long-range electrostatics, desolvation upon binding, and entropic contributions, are poorly defined in conventional scoring functions utilized in molecular docking (Rastelli *et al*., 2010a). These terms are better defined by more rigorous and computationally intensive calculations included in techniques such as free energy perturbation (FEP; Kollman, 1993), thermodynamic integration (TI; Lybrand *et al*., 1986), linear response (LR; Åqvist *et al*., 1994), molecular mechanics Poisson-Boltzamann/surface area (MM-PBSA; Kuhn and Kollman, 2000) and molecular mechanics generalized-Born/surface area (MM-GBSA; Kollman *et al*., 2000). Of these, MM-PBSA and MM-GBSA are faster by several orders of magnitude, making them favorable techniques for the rescoring and reranking of hit lists identified by virtual screening. As these techniques are computationally efficient and yield high correlations with experimental binding energies, the general opinion that docking results should be further analyzed by more advanced approaches is increasing (Rastelli *et al*., 2010a, Sgobba *et al*., 2012).

MM-PBSA and MM-GBSA previously required an ensemble of snapshots, generated by an MD simulation of the protein-ligand complex in water. This has been replaced by the use of a continuum implicit solvent model with a single minimized protein-ligand structure. This technique has given excellent correlations with experimental data (Guimarães and Cardozo, 2008; Greenidge *et al*., 2013) and is comparable with the more time-consuming and computer-intensive approach of averaging MD simulations in water (Rastelli *et al*., 2010a). The use of a single energy-minimized structure with a continuum implicit solvent has further improved the enrichment of virtual screens and can successfully discriminate between true binders and decoys (Rastelli *et al*., 2010a). Explicit solvent models have further shown to decrease this correlation (Greenidge *et al*., 2013).

MM-PBSA and MM-GBSA are force field–based methods that use a combination of molecular mechanics (MM) energies, polar and nonpolar solvation terms, and

an entropy term to calculate the free energy of binding (ΔG_{bind}; Massova and Kollman, 2000; Kollman *et al*., 2000) from the change between the bound complex (ΔG_{com}) and unbound receptor (ΔG_{rec}) and ligand (ΔG_{lig}) in solution [Eq. (27.1); Rastelli *et al*., 2010a; Guimarães and Cardozo, 2008]:

$$\Delta G_{bind} = \Delta G_{com} - \Delta G_{rec} - \Delta G_{lig}. \tag{27.1}$$

Each of these terms are decomposed into gas-phase MM energy (ΔE_{MM}), polar and nonpolar solvation terms (ΔS_{solv}), and an entropy term (ΔS) at a predefined temperature (T) [Eq. (27.2)].

$$\Delta G_{(com/rec/lig)} = \Delta E_{MM} + \Delta G_{solv} - T.\Delta S. \tag{27.2}$$

E_{MM} is calculated by the sum of chemical bonds, angles, and torsion terms (E_{bat}) predefined by the force field and van der Waals (E_{vdW}) and Coulombic terms (E_{coul}). The G_{solv} terms are further decomposed into a polar ($G_{solv,p}$) contribution and nonpolar ($G_{solv,np}$) contribution. The polar contributions are calculated by generalized-Born (GB) approximations in MM-GBSA (Kollman *et al*., 2000; Greenidge *et al*., 2013) and a Poisson-Boltzamn (PB) distribution in MM-PBSA (Kuhn and Kollman, 2000), where the nonpolar contribution is usually calculated as a linear function of the solvent accessible surface area (Hou *et al*., 2011a; Greenidge *et al*., 2013). With these functions, the binding free energy (ΔG_{bind}) is calculated:

$$\Delta G_{(com/rec/lig)} = \Delta E_{bat} + \Delta E_{vdW} + \Delta E_{coul} + \Delta G_{solv,p} + \Delta G_{solv,np} - T.\Delta S. \tag{27.3}$$

In most studies, the entropy (T. ΔS) term is neglected, as its calculation can be a major source of error (Rastelli *et al*., 2010b) and does not always improve the prediction accuracy (Hou *et al*., 2011a; Guimaraes, 2012); however, some researchers do still advocate its use (e.g., Lafont *et al*., 2007).

When comparing the PB and GB methods in calculation of solvation terms, the PB model is theoretically more rigorous and computationally intensive than GB but does not always give a stronger correlation with experimental binding free energy. The GB model is also more efficient and faster at ranking binding affinities of ligands, making it more suitable in a virtual screening context (Hou *et al*., 2011a, 2011b; Huang *et al*., 2010; Li *et al*., 2010). MM-GBSA has further been shown to be a more attractive option than the computationally heavy FEP and TI methodologies, as it can be as accurate and computationally more efficient, and handle structurally more diverse ligands because it requires no training set (Guimarães and Cardozo, 2008). In conclusion, MM-GBSA provides excellent correlation with experimental binding energy, improved enrichment in virtual screening of compound databases, is computationally suitable for medium-throughput screening or reranking a defined hit list and provides more accurate docking poses (Hou *et al*., 2011b). With this, the rescoring of docking complexes using MM-GBSA has emerged as a computationally important approach in structure-based drug design (Guimaraes, 2012).

7 INCORPORATING PHARMACOPHORIC CONSTRAINTS WITHIN THE VIRTUAL SCREEN

A *pharmacophore* is defined as an ensemble of structural features that are necessary for molecular recognition (Guner, 2000). These features predominantly include hydrophobic moieties and hydrogen bond donors and acceptors, but may also include aromatic rings, cations, and anions. A pharmacophore model can be used prior to a docking study to reduce the size of a ligand library, or it can be used to filter hits following a virtual screen. These pharmacophoric features can be defined by an ensemble of known, active inhibitors where features that are frequently repeated are included in the pharmacophore model (Yang, 2010), or they can be defined by the natural substrate. Identifying ligands that are able to replicate the interactions made between the natural substrate and a receptor can greatly improve the success and enrichment of a virtual screen. An example of substrate derived pharmacophoric constraints is the three chymotrypsin-like protease ($3CL^{pro}$) of coronaviruses. The S_1 pocket in this family of proteases has an absolute specificity for glutamine, which is mediated by a hydrogen bond between the substrate and His163, deep in the pocket. The S_2 pocket forms a deep hydrophobic region that displays preference for a hydrophobic moiety and the Glu166 residue increases substrate specificity via an additional hydrogen bond (Zheng *et al*., 2007; Chuck *et al*., 2011; Shoichet, 2004; Schapira *et al*., 2003). These pharmacophoric features have been extensively used to identify novel inhibitors of the $3CL^{pro}$ (Jacobs *et al*., 2013).

8 CONCLUSION

Despite its limitations, molecular docking has yielded the discovery of novel leads (Shoichet and Kobilka, 2012; Wang and Ekins, 2006) and, if used correctly, the speed and cost effectiveness at which molecular docking screens can be conducted can provide an excellent starting point in a project with few to no compelling leads (Alvarez, 2004). Possibly the most important consideration to make when commencing a structure-based drug design study is a question of project design, especially if the user is a beginner in the field. The more prior knowledge and availability of published data, the greater the chance of success in the project where proper scrutiny of available literature is essential. The availability of high-resolution crystals or NMR structures of the receptor are paramount prior to a virtual screen, as homology models have been proven to yield low enrichment when compared to *holo* or *apo* experimental structures. *Holo* structures have further proven to improve enrichment, and the state of the experimental structure should therefore be taken into account (McGovern and Shoichet, 2003). Efficient characterization of the active or allosteric binding site is essential. Detailed understanding of the location and flexibility of side-chains within the pocket, the presence or absence of active site waters, and protonation states of ionisable residues will contribute greatly to enrichment in a virtual

screen. The availability of known actives will also allow for essential benchmarking, validation, and potential generation of an effective pharmacophore model. It is important to characterize what class of inhibitors these actives belong to. Molecular docking is not capable of replicating covalent interactions between ligand and receptor, and therefore covalent inhibitors should be excluded. Large peptidomimetics are also difficult to dock with conventional docking methodologies. This is directly related to the inaccuracy of docking algorithms to predict the correct conformation of compounds with increased number of rotatable bonds. With each rotatable bond, the conformational space that must be sampled increases dramatically and thereby reduces the chance of successfully predicting the correct pose. Ligands in a molecular docking screen, therefore, should be limited to eight rotatable bonds (Houston and Walkinshaw, 2013). A final consideration covered in this chapter is the use of consensus scoring and sampling. This has been shown to greatly improve enrichment with MM-GBSA rescoring, yielding high correlations with experimental evidence and should be considered in a virtual screening context (Teramoto and Fukunishi, 2007; Chang *et al.*, 2010; Houston and Walkinshaw, 2013; Hou *et al.*, 2011b).

REFERENCES

Alonso, H., Bliznyuk, A.A., Gready, J.E., 2006. Combining docking and molecular dynamic simulations in drug design. Med. Res. Rev. 26, 531–568.

Alvarez, J.C., 2004. High-throughput docking as a source of novel drug leads. Curr. Opin. Chem. Biol. 8, 365–370.

Anderson, A.C., 2003. The Process of Structure-Based Drug Design. Chem. Biol. 10, 787–797.

Åqvist, J., Medina, C., Samuelsson, J.-E., 1994. A new method for predicting binding affinity in computer-aided drug design. Protein Eng. 7, 385–391.

Ballester, P.J., Mitchell, J.B., 2010. A machine learning approach to predicting protein–ligand binding affinity with applications to molecular docking. Bioinformatics 26, 1169–1175.

Barillari, C., Taylor, J., Viner, R., Essex, J.W., 2007. Classification of water molecules in protein binding sites. J. Am. Chem. Soc. 129, 2577–2587.

Barril, X., Morley, S.D., 2005. Unveiling the full potential of flexible receptor docking using multiple crystallographic structures. J. Med. Chem. 48, 4432–4443.

B-Rao, C., Subramanian, J., Sharma, S.D., 2009. Managing protein flexibility in docking and its applications. Drug Discov. Today 14, 394–400.

Brooijmans, N., Kuntz, I.D., 2003. Molecular recognition and docking algorithms. Annu. Rev. Biophys. Biomol. Struct. 32, 335–373.

Carlson, H.A., 2002. Protein flexibility and drug design: how to hit a moving target. Curr. Opin. Chem. Biol. 6, 447–452.

Carlson, H.A., Mccammon, J.A., 2000. Accommodating protein flexibility in computational drug design. Mol. Pharmacol. 57, 213–218.

Cavasotto, C.N., Abagyan, R.A., 2004. Protein flexibility in ligand docking and virtual screening to protein kinases. J. Mol. Biol. 337, 209–225.

Chang, M.W., Ayeni, C., Breuer, S., Torbett, B.E., 2010. Virtual screening for HIV protease inhibitors: a comparison of AutoDock 4 and Vina. PLoS One 5, e11955.

Charifson, P.S., Corkery, J.J., Murcko, M.A., Walters, W.P., 1999. Consensus scoring: a method for obtaining improved hit rates from docking databases of three-dimensional structures into proteins. J. Med. Chem. 42, 5100–5109.

Cheng, T., Li, Q., Zhou, Z., Wang, Y., Bryant, S.H., 2012. Structure-based virtual screening for drug discovery: a problem-centric review. AAPS J. 14, 133–141.

Chuck, C.-P., Chow, H.-F., Wan, D.C.-C., Wong, K.-B., 2011. Profiling of substrate specificities of 3C-like proteases from group 1, 2a, 2b, and 3 coronaviruses. PLoS One 6, e27228.

Craig, I.R., Essex, J.W., Spiegel, K., 2010. Ensemble docking into multiple crystallographically derived protein structures: an evaluation based on the statistical analysis of enrichments. J. Chem. Inf. Model. 50, 511–524.

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. J. Comput. Aided Mol. Des. 15, 411–428.

Fornabaio, M., Cozzini, P., Mozzarelli, A., Abraham, D.J., Kellogg, G.E., 2003. Simple, intuitive calculations of free energy of binding for protein-ligand complexes. 2. Computational titration and pH effects in molecular models of neuraminidase-inhibitor complexes. J. Med. Chem. 46, 4487–4500.

Glusker, J.P., Lewis, M., Rossi, M., 1994. Crystal Structure Analysis for Chemists and Biologists. John Wiley & Sons.

Greenidge, P.A., Kramer, C., Mozziconacci, J.C., Wolf, R.M., 2013. MM/GBSA binding energy prediction on the PDBbind data set: successes, failures, and directions for further improvement. J. Chem. Inf. Model. 53, 201–209.

Guimaraes, C.R., 2012. MM-GB/SA Rescoring of Docking Poses. Computational Drug Discovery and Design. Springer.

Guimarães, C.R., Cardozo, M., 2008. MM-GB/SA rescoring of docking poses in structure-based lead optimization. J. Chem. Inf. Model. 48, 958–970.

Guner, O.F., 2000. Pharmacophore Perception, Development, and Use in Drug Design. International University Line, La Jolla, CA, 29.

Han, S.-H., 2012. Docking and molecular dynamics simulations of celastrol binding to p23. Bulletin of the Korean Chemical Society 33, 322–324.

Hartshorn, M.J., Verdonk, M.L., Chessari, G., Brewerton, S.C., Mooij, W.T., Mortenson, P.N., Murray, C.W., 2007. Diverse, high-quality test set for the validation of protein-ligand docking performance. J. Med. Chem. 50, 726–741.

Henzler, A.M., Rarey, M., 2010. Pursuit of Fully Flexible Protein-Ligand Docking: Modeling the Bilateral Mechanism of Binding. In: Molecular Informatics, 29, pp. 164–173.

Hooft, R.W., Sander, C., Vriend, G., 1996. Positioning hydrogen atoms by optimizing hydrogen-bond networks in protein structures. Proteins: Structure, Function, and Bioinformatics 26, 363–376.

Hornak, V., Okur, A., Rizzo, R.C., Simmerling, C., 2006. HIV-1 protease flaps spontaneously close to the correct structure in simulations following manual placement of an inhibitor into the open state. J. Am. Chem. Soc. 128, 2812–2813.

Hou, T., Wang, J., Li, Y., Wang, W., 2011a. Assessing the performance of the molecular mechanics/Poisson Boltzmann surface area and molecular mechanics/generalized Born surface area methods. II. The accuracy of ranking poses generated from docking. J. Comput. Chem. 32, 866–877.

Hou, T., Wang, J., Li, Y., Wang, W., 2011b. Assessing the performance of the MM/PBSA and MM/GBSA methods. 1. The accuracy of binding free energy calculations based on molecular dynamics simulations. J. Chem. Inf. Model. 51, 69–82.

Houston, D.R., Walkinshaw, M.D., 2013. Consensus docking: improving the reliability of docking in a virtual screening context. J. Chem. Inf. Model. 53, 384–390.

Huang, N., Shoichet, B.K., 2008. Exploiting ordered waters in molecular docking. J. Med. Chem. 51, 4862–4865.

Huang, S.Y., Grinter, S.Z., Zou, X., 2010. Scoring functions and their evaluation methods for protein-ligand docking: recent advances and future directions. Phys. Chem. Chem. Phys. 12, 12899–12908.

Jacobs, J., Grum-Tokars, V., Zhou, Y., Turlington, M., Saldanha, S.A., Chase, P., Eggler, A., Dawson, E.S., Baez-Santos, Y.M., Tomar, S., Mielech, A.M., Baker, S.C., Lindsley, C.W., Hodder, P., Mesecar, A., Stauffer, S.R., 2013. Discovery, synthesis, and structure-based optimization of a series of N-(tert-butyl)-2-(N-arylamido)-2-(pyridin-3-yl) acetamides (ML188) as potent noncovalent small molecule inhibitors of the severe acute respiratory syndrome coronavirus (SARS-CoV) 3CL protease. J. Med. Chem. 56, 534–546.

Joseph-Mccarthy, D., Baber, J.C., Feyfant, E., Thompson, D.C., Humblet, C., 2007. Lead optimization via high-throughput molecular docking. Curr. Opin. Drug Discov. Devel. 10, 264–274.

Kalliokoski, T., Salo, H.S., Lahtela-Kakkonen, M., Poso, A., 2009. The effect of ligand-based tautomer and protomer prediction on structure-based virtual screening. J. Chem. Inf. Model. 49, 2742–2748.

Kim, M.O., Nichols, S.E., Wang, Y., Mccammon, J.A., 2013. Effects of histidine protonation and rotameric states on virtual screening of M. tuberculosis RmlC. J. Comput. Aided Mol. Des. 27, 235–246.

Knegtel, R., Kuntz, I.D., Oshiro, C., 1997. Molecular docking to ensembles of protein structures. J. Mol. Biol. 266, 424–440.

Kollman, P., 1993. Free energy calculations: applications to chemical and biochemical phenomena. Chem. Rev. 93, 2395–2417.

Kollman, P.A., Massova, I., Reyes, C., Kuhn, B., Huo, S., Chong, L., Lee, M., Lee, T., Duan, Y., Wang, W., 2000. Calculating structures and free energies of complex molecules: combining molecular mechanics and continuum models. Acc. Chem. Res. 33, 889–897.

Krieger, E., Dunbrack Jr., R.L., Hooft, R.W., Krieger, B., 2012. Assignment of Protonation States in Proteins and Ligands: Combining Pka Prediction with Hydrogen Bonding Network Optimization. Computational Drug Discovery and Design. Springer.

Kroemer, R.T., 2007. Structure-based drug design: docking and scoring. Current Protein and Peptide Science 8, 312–328.

Kuhn, B., Kollman, P.A., 2000. Binding of a diverse set of ligands to avidin and streptavidin: an accurate quantitative prediction of their relative affinities by a combination of molecular mechanics and continuum solvent models. J. Med. Chem. 43, 3786–3791.

Kukol, A., 2011. Consensus virtual screening approaches to predict protein ligands. Eur. J. Med. Chem. 46, 4661–4664.

Kuntz, I.D., Blaney, J.M., Oatley, S.J., Langridge, R., Ferrin, T.E., 1982. A geometric approach to macromolecule-ligand interactions. J. Mol. Biol. 161, 269–288.

Lafont, V., Armstrong, A.A., Ohtaka, H., Kiso, Y., Mario Amzel, L., Freire, E., 2007. Compensating enthalpic and entropic changes hinder binding affinity optimization. Chem. Biol. Drug Des. 69, 413–422.

Lemmon, G., Meiler, J., 2013. Towards ligand docking including explicit interface water molecules. PLoS One 8, e67536.

Li, Y., Liu, Z., Wang, R., 2010. Test MM-PB/SA on true conformational ensembles of protein − ligand complexes. J. Chem. Inf. Model. 50, 1682–1692.

Lie, M.A., Thomsen, R., Pedersen, C.N., Schiøtt, B., Christensen, M.H., 2011. Molecular docking with ligand attached water molecules. J. Chem. Inf. Model. 51, 909–917.

Lill, M.A., 2011. Efficient incorporation of protein flexibility and dynamics into molecular docking simulations. Biochemistry 50, 6157–6169.

Lill, M., 2013. Virtual screening in drug design. Methods Mol. Biol. 993, 1–12.

Lim, S.V., Rahman, M.B.A., Tejo, B.A., 2011. Structure-based and ligand-based virtual screening of novel methyltransferase inhibitors of the dengue virus. BMC bioinformatics 12, S24.

Lybrand, T.P., Mccammon, J.A., Wipff, G., 1986. Theoretical Calculation of Relative Binding Affinity in Host-Guest Systems. In: Proceedings of the National Academy of Sciences, 83, pp. 833–835.

Marco, E., Gago, F., 2007. Overcoming the inadequacies or limitations of experimental structures as drug targets by using computational modeling tools and molecular dynamics simulations. ChemMedChem 2, 1388–1401.

Massova, I., Kollman, P.A., 2000. Combined molecular mechanical and continuum solvent approach (MM-PBSA/GBSA) to predict ligand binding. Perspect Drug Discov Des 18, 113–135.

Mcgovern, S.L., Shoichet, B.K., 2003. Information decay in molecular docking screens against holo, apo, and modeled conformations of enzymes. J. Med. Chem. 46, 2895–2907.

Meng, X.-Y., Zhang, H.-X., Mezei, M., Cui, M., 2011. Molecular docking: a powerful approach for structure-based drug discovery. Curr. Comput. Aided Drug Des. 7, 146.

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. J. Comput. Chem. 30, 2785–2791.

Mukherjee, P., Shah, F., Desai, P., Avery, M., 2011. Inhibitors of SARS-3CLpro: virtual screening, biological evaluation, and molecular dynamics simulation studies. J. Chem. Inf. Model. 51, 1376–1392.

Mysinger, M.M., Carchia, M., Irwin, J.J., Shoichet, B.K., 2012. Directory of useful decoys, enhanced (DUD-E): better ligands and decoys for better benchmarking. J. Med. Chem. 55, 6582–6594.

Najmanovich, R., Kuttner, J., Sobolev, V., Edelman, M., 2000. Side-chain flexibility in proteins upon ligand binding. Proteins: Structure, Function, and Bioinformatics 39, 261–268.

Onufriev, A.V., Alexov, E., 2013. Protonation and pK changes in protein–ligand binding. Q. Rev. Biophys. 46, 181–209.

Österberg, F., Åqvist, J., 2005. Exploring blocker binding to a homology model of the open hERG K^+ channel using docking and molecular dynamics methods. FEBS Lett. 579, 2939–2944.

Petukh, M., Stefl, S., Alexov, E., 2013. The role of protonation states in ligand-receptor recognition and binding. Curr. Pharm. Des. 19, 4182.

Phatak, S.S., Stephan, C.C., Cavasotto, C.N., 2009. High-Throughput and in Silicon Screenings in Drug Discovery.

Plewczynski, D., Lazniewski, M., Von Grotthuss, M., Rychlewski, L., Ginalski, K., 2011. VoteDock: consensus docking method for prediction of protein-ligand interactions. J. Comput. Chem. 32, 568–581.

Polgár, T., Keserü, G.M., 2005. Virtual screening for β-secretase (BACE1) inhibitors reveals the importance of protonation states at Asp32 and Asp228. J. Med. Chem. 48, 3749–3755.

Rastelli, G., Del Rio, A., Degliesposti, G., Sgobba, M., 2010a. Fast and accurate predictions of binding free energies using MM-PBSA and MM-GBSA. J. Comput. Chem. 31, 797–810.

Rastelli, G., Rio, A.D., Degliesposti, G., Sgobba, M., 2010b. Fast and accurate predictions of binding free energies using MM-PBSA and MM-GBSA. J. Comput. Chem. 31, 797–810.

Ripphausen, P., Nisius, B., Bajorath, J., 2011. State-of-the-art in ligand-based virtual screening. Drug Discov. Today 16, 372–376.

Santos, R., Hritz, J., Oostenbrink, C., 2009. Role of water in molecular docking simulations of cytochrome P450 2D6. J. Chem. Inf. Model. 50, 146–154.

Schapira, M., Raaka, B.M., Das, S., Fan, L., Totrov, M., Zhou, Z., Wilson, S.R., Abagyan, R., Samuels, H.H., 2003. Discovery of Diverse Thyroid Hormone Receptor Antagonists by High-Throughput Docking. In: Proceedings of the National Academy of Sciences, 100, pp. 7354–7359.

Schneider, G., Fechner, U., 2005. Computer-based de novo design of drug-like molecules. Nat. Rev. Drug Discov. 4, 649–663.

Segura-Cabrera, A., García-Pérez, C.A., Guo, X., Rodríguez-Pérez, M.A., 2013. Repurposing of FDA-approved drugs for the discovery of inhibitors of dengue virus NS2B-NS3 protease by docking, consensus scoring, and molecular dynamics simulations. Biophys. J. 104, 404a.

Sgobba, M., Caporuscio, F., Anighoro, A., Portioli, C., Rastelli, G., 2012. Application of a post-docking procedure based on MM-PBSA and MM-GBSA on single and multiple protein conformations. Eur. J. Med. Chem. 58, 431–440.

Shoichet, B.K., 2004. Virtual screening of chemical libraries. Nature 432, 862–865.

Shoichet, B.K., Kobilka, B.K., 2012. Structure-based drug screening for G-protein-coupled receptors. Trends Pharmacol. Sci. 33, 268–272.

Ten Brink, T., Exner, T.E., 2009. Influence of protonation, tautomeric, and stereoisomeric states on protein − ligand docking results. J. Chem. Inf. Model. 49, 1535–1546.

Teramoto, R., Fukunishi, H., 2007. Supervised consensus scoring for docking and virtual screening. J. Chem. Inf. Model. 47, 526–534.

Therrien, E., Weill, N., Tomberg, A., Corbeil, C.R., Lee, D., Moitessier, N., 2014. Docking Ligands into Flexible and Solvated Macromolecules. 7. Impact of Protein Flexibility and Water Molecules on Docking-based Virtual Screening Accuracy.

Thilagavathi, R., Mancera, R.L., 2010. Ligand − Protein cross-docking with water molecules. J. Chem. Inf. Model. 50, 415–421.

Totrov, M., Abagyan, R., 2008. Flexible ligand docking to multiple receptor conformations: a practical alternative. Curr. Opin. Struct. Biol. 18, 178–184.

Trott, O., Olson, A.J., 2010. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. J. Comput. Chem. 31, 455–461.

Wang, B., Ekins, S., 2006. Computer Applications in Pharmaceutical Research and Development. John Wiley & Sons.

Wang, R., Lu, Y., Wang, S., 2003. Comparative evaluation of 11 scoring functions for molecular docking. J. Med. Chem. 46, 2287–2303.

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. J. Med. Chem. 49, 5912–5931.

Waszkowycz, B., Clark, D.E., Gancia, E., 2011. Outstanding challenges in protein–ligand docking and structure-based virtual screening. Wiley Interdisciplinary Reviews: Computational Molecular Science 1, 229–259.

Word, J.M., Lovell, S.C., Richardson, J.S., Richardson, D.C., 1999. Asparagine and glutamine: using hydrogen atom contacts in the choice of side-chain amide orientation. J. Mol. Biol. 285, 1735–1747.

Yang, S.-Y., 2010. Pharmacophore modeling and applications in drug discovery: challenges and recent advances. Drug Discov. Today 15, 444–450.

Yang, J.-M., Chen, Y.-F., Shen, T.-W., Kristal, B.S., Hsu, D.F., 2005. Consensus scoring criteria for improving enrichment in virtual screening. J. Chem. Inf. Model. 45, 1134–1146.

Yang, H., Bartlam, M., Rao, Z., 2006. Drug Design targeting the main protease, the Achilles heel of Coronaviruses. Curr. Pharm. Des. 12, 4573–4590.

Zheng, K., Ma, G., Zhou, J., Zen, M., Zhao, W., Jiang, Y., Yu, Q., Feng, J., 2007. Insight into the activity of SARS main protease: molecular dynamics study of dimeric and monomeric form of enzyme. Proteins 66, 467–479.

Zhu, T., Cao, S., Su, P.C., Patel, R., Shah, D., Chokshi, H.B., Szukala, R., Johnson, M.E., Hevener, K.E., 2013. Hit identification and optimization in virtual screening: practical recommendations based on a critical literature analysis. J. Med. Chem. 56, 6560–6572.