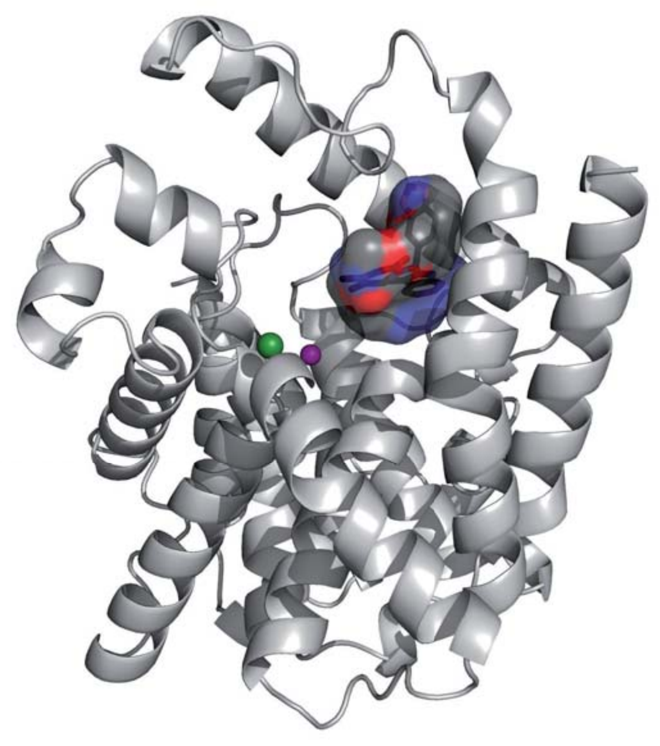
JYVÄSKYLÄ STUDIES IN BIOLOGICAL AND ENVIRONMENTAL SCIENCE

264

Salla Virtanen

Virtual Screening

Development of a Novel Structure-Based Method



JYVÄSKYLÄN YLIOPISTO

JYVÄSKYLÄ STUDIES IN BIOLOGICAL AND ENVIRONMENTAL SCIENCE 264

Salla Virtanen

Virtual Screening

Development of a Novel Structure-Based Method

Esitetään Jyväskylän yliopiston matemaattis-luonnontieteellisen tiedekunnan suostumuksella
julkisesti tarkastettavaksi yliopiston Ambiotica-rakennuksen salissa YAA303
elokuun 24. päivänä 2013 kello 12.

Academic dissertation to be publicly discussed, by permission of
the Faculty of Mathematics and Science of the University of Jyväskylä,
in building Ambiotica, hall YAA303, on August 24, 2013 at 12 o'clock noon.

JYVÄSKYLÄ 2013

Virtual Screening

Development of a Novel Structure-Based Method

JYVÄSKYLÄ STUDIES IN BIOLOGICAL AND ENVIRONMENTAL SCIENCE 264

Salla Virtanen

Virtual Screening

Development of a Novel Structure-Based Method

JYVÄSKYLÄ 2013

Editors
Varpu Marjomäki
Department of Biological and Environmental Science, University of Jyväskylä
Pekka Olsbo
Publishing Unit, University Library of Jyväskylä

Jyväskylä Studies in Biological and Environmental Science
Editorial Board

Jari Haimi, Anssi Lensu, Timo Marjomäki, Varpu Marjomäki
Department of Biological and Environmental Science, University of Jyväskylä

Cover picture by Salla Virtanen

URN:ISBN:978-951-39-5341-6
ISBN 978-951-39-5341-6 (PDF)

ISBN 978-951-39-5340-9 (nid.)
ISSN 1456-9701

Copyright © 2013, by University of Jyväskylä

Jyväskylä University Printing House, Jyväskylä 2013

ABSTRACT

Virtanen, Salla
Virtual screening: Development of a novel structure-based method
Jyväskylä: University of Jyväskylä, 2013, 51 p.
(Jyväskylä Studies in Biological and Environmental Science
ISSN 1456-9701; 264)
ISBN 978-951-39-5340-9 (nid.)
ISBN 978-951-39-5341-6 (PDF)
Yhteenveto: Virtuaaliseulonta; uuden rakennepohjaisen menetelmän kehitys
Diss.

Computational methods have become an integral part of today's drug discovery and drug design efforts. The computational methods commonly used in virtual screening (VS) can be classified into two categories: ligand-based methods, which use information of known ligand molecules, and protein structure-based methods, which require information of the target protein structure. However, the currently used VS methods have some disadvantages: ligand-based methods tend to discover molecules that are very similar to the existing ligands, and protein structure-based methods often employ approximated scoring functions, which can lead to unreliable results. In this thesis a novel method for VS was developed, in which the shape of the ligand binding site is depicted as a ligand-like negative image. This negative image can then be used in place of a real active molecule in VS. When compared to traditional ligand- and structure-based VS, the results with negative image-based (NIB) screening were often better or comparable. Since the chemistry of the ligand binding site is also important for efficient binding, the chemical information was added to the negative images in the form of partial charges. The results show, that the addition of chemical information can improve the VS results, particularly in cases, where the binding site contains polar amino acids. Methods that employ implicit solvent models combined with molecular mechanics force fields, such as Molecular Mechanics Generalized Born Surface Area (MM-GBSA), have also become important tools in predicting binding free energies of ligands. We found that NIB screening can greatly benefit from post-processing with MM-GBSA. The performance of MM-GBSA, along with Molecular Mechanics Poisson-Boltzmann Surface Area (MM-PBSA) and a similar method called Solvated Interaction Energy (SIE), was studied further, and our results show a highly case-specific performance both in VS and in predicting known binding affinities.

Keywords: Computational drug discovery; binding free energy; negative image-based screening; molecular dynamics; virtual screening.

Salla Virtanen, University of Jyväskylä, Department of Biological and Environmental Science, P.O. Box 35, FI-40014 University of Jyväskylä, Finland

Author's address Salla Virtanen
Department of Biological and Environmental Science
P.O. Box 35
40014 University of Jyväskylä
Finland
salla.i.virtanen@jyu.fi

Supervisor Adjunct professor Olli Pentikäinen, Ph.D.
Department of Biological and Environmental Science
P.O. Box 35
40014 University of Jyväskylä
Finland

Reviewers Professor Stefano Moro, Ph.D.
Department of Pharmaceutical and Pharmacological Sciences
University of Padova
Via Marzolo 5
35131 Padova
Italy

Associate professor Patrik Rydberg, Ph.D.
Department of Drug Design and Pharmacology
University of Copenhagen
Jagtvej 162
2100 Copenhagen
Denmark

Opponent Professor Jari Yli-Kauhaluoma, Ph.D.
Division of Pharmaceutical Chemistry
P.O. Box 56
FI-00014 University of Helsinki
Finland

CONTENTS

LIST OF ORIGINAL PUBLICATIONS
RESPONSIBILITIES OF SALLA VIRTANEN IN THE THESIS ARTICLES
ABBREVIATIONS

1 INTRODUCTION 11

2 REVIEW OF LITERATURE 13

2.1 Virtual screening 13
2.1.1 Ligand-based methods 14
2.1.2 Structure-based methods 17
2.1.3 Pharmacophore modeling..20
2.1.4 3D quantitative structure-activity relationship...............20
2.1.5 Negative image-based screening 21
2.2 Binding free energy calculations 22
2.2.1 MM-GBSA and MM-PBSA 22
2.2.2 SIE..23

3 AIMS OF THE STUDY 24

4 METHODS AND MATERIALS 25
4.1 Molecular databases 25
4.2 Negative images..25
4.3 3D similarity searching software..26
4.4 Docking 26
4.4.1 GOLD (I and III) 26
4.4.2 GLIDE (I and II)...26
4.5 Molecular dynamics simulations 27
4.5.1 MD simulations with NAMD (I).......................................27
4.5.2 MD simulations with SANDER (II and III) 27
4.6 MM-GB(PB)SA 28
4.7 SIE 28
4.8 Metrics in the virtual screening studies....................................28

5 RESULTS 30
5.1 Negative image based (NIB) screening 30
5.1.1 Optimization of VOIDOO/FLOOD parameters 30
5.1.2 Virtual screening efficiency with negative image 30
5.1.3 Use of multiple protein conformations 31
5.1.4 Negative images generated with PASS............................31
5.2 Virtual screening of PDE-5 inhibitors (II).................................32
5.2.1 Ligand shape-based screening 32

5.2.2 Protein structure-based screening 32
5.2.3 The effect of added electrostatic information.................................33
5.2.4 Post-processing with MD/MM-GBSA...34
5.3 Comparison of free energy calculation methods in virtual screening (III)..35
5.3.1 Virtual screening efficiency 35
5.3.2 Binding free energy calculations..36
5.3.3 Effect of the MD simulation length..37

6 DISCUSSION 39
6.1 Negative image-based screening increases the efficiency of virtual screening 39
6.2 The use of multiple protein structures for NIB screening can improve the efficiency...40
6.3 The addition of electrostatic information improves the negative-image-based screening for some targets 40
6.4 MM-GBSA is a viable method for enriching NIB screening results.....41
6.5 MM-GB(PB)SA and SIE in virtual screening 41
6.6 MM-GB(PB)SA in predicting experimental binding data 42

7 CONCLUSIONS...43

YHTEENVETO (RÉSUMÉ IN FINNISH)...45
REFERENCES 47

LIST OF ORIGINAL PUBLICATIONS

The thesis is based on the following original articles, which will be referred to in the text by their Roman numerals.

I Virtanen S.I. & Pentikäinen O.T. 2010. Efficient virtual screening using multiple protein structures described as negative images of the ligand-binding site. *Journal of Chemical Information and Modeling* 50: 1005–1011.

II Niinivehmas S.P.*, Virtanen S.I.*, Lehtonen J.V., Postila P.A. & Pentikäinen O.T. 2011. Comparison of virtual high-throughput screening methods for the identification of phosphodiesterase-5 inhibitors. *Journal of Chemical Information and Modeling* 51: 1353–1363.

III Virtanen S.I. & Pentikäinen O.T. Case-specific performance of MM-PBSA, MM-GBSA, and SIE in virtual screening. Submitted manuscript.

RESPONSIBILITIES OF SALLA VIRTANEN IN THE THESIS ARTICLES

Article I — I constructed the negative images and did all the virtual screening studies, in addition to the molecular dynamics simulations. I made the figures for the article. The article was written together with Olli Pentikäinen.

Article II — I did the virtual screening studies with docking and the MM-GBSA calculations for the docking results. I did the negative image- and ligand-based screening studies for the nuclear hormone receptors. The article was written together with the other authors.

Article III — I planned the computational experiments together with Olli Pentikäinen and did all the simulations. I wrote the article and made the figures.

ABBREVIATIONS

1D	1 dimensional
2D	2 dimensional
3D	3 dimensional
ALR-2	aldose reductase 2
AUC	area under curve
DUD	Directory of Useful Decoys
HTS	high-throughput screening
MC	Monte Carlo
MD	molecular dynamics
MM-GBSA	molecular mechanics generalized Born surface area
MM-PBSA	molecular mechanics Poisson-Boltzmann surface area
NIB	negative image-based
NMA	normal mode analysis
PDB	protein data bank
PDE-5	phosphodiesterase 5
PR	progesterone receptor
QSAR	quantitative structure-activity relationship
ROC	receiver operating characteristic
SIE	solvated interaction energy
vdW	van der Waals
VHTS	virtual high-throughput screening
VS	virtual screening

1 INTRODUCTION

A drug is a substance that is used to treat or prevent a disease. The actions of a drug molecule are usually mediated by a key protein that is involved in a metabolic or signaling pathway affected by the disease. Proteins form an essential group of macromolecules that participate in many important processes inside the cells. Examples of proteins include enzymes that catalyze chemical reactions, receptors that mediate signals, or structural proteins that provide structural support. Because of the versatility of proteins, the drug molecules can also produce their effects in different ways. A drug can for example work as an antagonist and inhibit enzymes or receptors by binding to the active site and thereby preventing the action of the natural substrate. Conversely, a drug can be an agonist by promoting the action of the target protein.

To bring a new drug to the market is, however, a long and costly process. Traditionally drug discovery has been made by trial and error of chemicals for their effects on cells or animals to treat a particular disease. Today these steps are automated and high throughput screening (HTS) facilities are routinely employed to screen hundreds of thousands of molecules for their biological activity. Despite the large number of tested molecules, only a few active compounds enter the subsequent drug development phase. This combined with the fact that only a fraction (10-20 %) of the drug candidates that enter clinical trials become accepted, motivates the search for more efficient methods to speed up the drug development process.

Virtual screening (VS) of molecules offers a way to enhance the drug discovery phase. In virtual screening large databases of molecules can be computationally evaluated for their likelihood of binding to the target protein. This enables the laborious and expensive experimental tests to focus on the top fraction of the screened library where the active molecules are expected to be enriched, thus saving time and resources. For virtual screening to be successful, structural information of the drug target is required. If an experimentally determined structure of the protein target is available, it can be used to find molecules that complement the shape and chemistry of the binding site. However, if there is no experimental data available for the target protein,

molecules that are known to bind to the protein can be used as templates to discover similar structures from the screening database. In this thesis a novel method for virtual screening was developed, which employs structural and chemical information of the binding site of the protein. In addition, several different methods were evaluated for their efficiency in virtual screening and in predicting experimentally determined binding data.

2 REVIEW OF LITERATURE

2.1 Virtual screening

Today virtual screening is becoming a standard procedure in drug development. There are already drugs on the market for which the initial discovery can be traced to a successful virtual screening campaign, such as the anticoagulant Aggrastat (tirofiban), which was found with a pharmacophore model (Hartman et al. 1992), and the HIV-integrase inhibitor Isentress (raltegravir) (Schames et al. 2004). Despite of the promising results, improvements to the accuracy and computational efficiency of the virtual screening methods are still needed, as has been shown in several retrospective virtual screening studies (McGaughey et al. 2007, Cross et al. 2009, von Korff et al. 2009).

The virtual screening workflow starts by preparation of the database of molecules used for screening (Fig. 1). Before the actual VS, the database is usually filtered in order to enrich it with molecules that have desirable properties, or, on the other hand, to eliminate compounds with undesirable properties. Often properties that are considered to be drug-like are used for this. For bioavailability the drugs should be relatively small in size, sufficiently water soluble for delivery in the blood or intracellular fluid, and sufficiently hydrophobic for crossing the cell membranes. A common set of drug-like filters is the Lipinski's Rule of Five (Lipinski et al. 2001), which consists of criterions that are shared by many drugs: 1) molecular weight less than 500 Daltons, 2) logP (octanol-water partition coefficient) less than 5, 3) maximum of 5 hydrogen bond donors, and 4) maximum of 10 hydrogen bond acceptors. Although the Lipinski's Rule has also been criticized for being a poor measure of drug-likeness, it is still widely used in drug discovery.

The choice of an appropriate VS method depends on many issues, e.g. on the availability of structural information of the target protein or known ligands, the size of the database used for screening, and computational resources. Different approaches can be used in combination in a virtual screen with the first steps made with coarse and computationally efficient methods, and only the top-

scoring hits from the primary screens passing forward to stages where more accurate and computationally demanding methods are employed. VS methods can be divided into two categories depending on the structural information they use: ligand-based and protein structure-based.

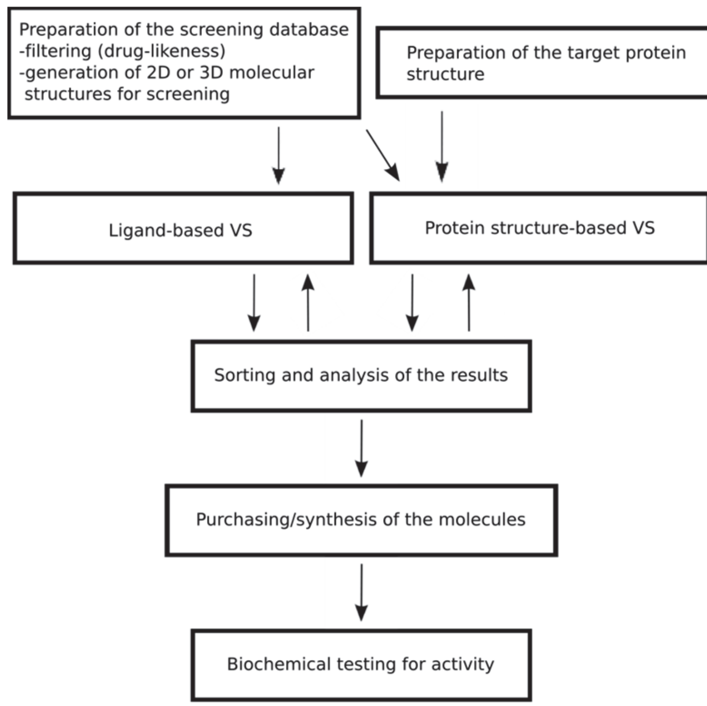


FIGURE 1 A general overview of virtual screening.

2.1.1 Ligand-based methods

Ligand-based virtual screening methods employ the idea of 'similar property principle' which states that similar compounds should have similar biological activity (Johnson & Maggiora 1990). The similarity between molecules is often evaluated with different kinds of molecular descriptors, which define the properties of the molecules. These descriptors can be divided into categories based on their dimensionality, which refers to the molecular representation where the descriptors are derived from (Fig. 2). 1 dimensional (1D) descriptors can be calculated from the chemical formula of the molecule (Fig. 2). These descriptors include such bulk properties as molecular weight and the number of specific atoms in the molecules. Many properties can be calculated from the 2 dimensional (2D) chemical structures of the molecules, such as substructure and connectivity information (Fig. 2). 3 dimensional (3D) molecular descriptors, such as solvent-accessible surface area or 3D pharmacophore properties, require additionally conformational information of the molecules (Fig. 2). In VS applications 2D and 3D descriptors are commonly used in similarity searches. 1D

descriptors, however, are usually not used alone in VS, but rather as a filter or in combination with 2D and 3D descriptors (Hong et al. 2008).

Inhibitor	1D structure	2D structure	3D structure
Sildenafil	$C_{22}H_{30}N_6O_4S$		
Tadalafil	$C_{22}H_{19}N_3O_4$		

FIGURE 2 Different representations for phosphodiesterase-5 inhibitors sildenafil and tadalafil. Similarity comparisons can be made with 1D, 2D, or 3D information of the molecules.

A very common ligand-based approach in virtual screening is the use of 2D fingerprints, where the 2D molecular descriptors have been expressed in a binary form. A fingerprint can be for example the presence or absence of a particular molecular fragment in a molecule, such as in MACCS fingerprints (Durant et al. 2002), where a predefined set of molecular substructures is used in the generation of the descriptors. The advantage in the use of 2D fingerpints is that the database searches can be made very efficiently, which enables screening of very large molecular databases. Table 1 summarizes some commonly used software for similarity searches with 2D fingerprints.

Another popular approach in ligand-based VS is the 3D shape-based screening, where the 3D shapes of molecules are compared (Fig. 2). The simplest comparisons can be made by using only shape similarity; however, also the chemical similarity of the molecules can be used for scoring the compounds. Examples of 3D similarity methods are ROCS (Rapid Overlay of Chemical Structures; Rush et al. 2005) and SHAEP (Vainio et al. 2009). With ROCS it is possible to compare the shape similarity of molecules in addition to their pharmacophoric features, whereas SHAEP can be used to compare the shape and electrostatic potential similarities of the molecules. Some commonly used 3D similarity searching software are listed in Table 1.

TABLE 1 List of commonly used software in ligand-based VS.

Software	Similarity searching method	References
CANVAS	2D fingerprints	Duan et al. 2010, Sastry et al. 2010
CHEMAXON SCREEN	2D fingerprints, 3D shape and pharmacophore	ChemAxon Kft., Budapest, Hungary
DISCOVERY STUDIO	2D fingerprints, 3D shape and pharmacophore	Accelrys Inc., San Diego, CA
ESSHAPE 3D	3D shape fingerprints	Cannon et al. 2008
MOE	2D fingerprints, 3D shape and pharmacophore	Chemical Computing Group Inc., Montreal, Canada
OPENBABEL	2D fingerprints	O'Boyle et al. 2011
PHASE SHAPE	3D shape and pharmacophore	Sastry et al. 2011
ROCS	3D shape and pharmacophore	Rush et al. 2005
SHAEP	3D shape and electrostatic potential	Vainio et al. 2009
UNITY	2D fingerprints, 3D shape and pharmacophore	Tripos Inc., St. Louis, MO
USR	3D shape	Ballester et al. 2010

Because of the increased complexity of the 3D similarity searching methods compared to 2D fingerprints, the 3D shape-based methods are computationally more demanding. Additionally, because the bioactive conformations are usually not known, multiple low-energy conformations have to be generated for the molecules in the screening database, increasing both computation time and the storage space required. Despite the increased complexity of the searches compared to 2D methods, the computation times per molecule are still reasonable for the screening of very large databases (Vainio et al. 2009).

Many retrospective virtual screening studies have shown that ligand-based methods can often identify the known active molecules with high precision (McGaughey et al. 2007, von Korff et al. 2009). However, the results depend heavily on the available template molecule(s) (Kirchmair et al. 2009). The 2D fingerprint methods by default tend to select molecules that have similar topology as the molecules used as targets (McGaughey et al. 2007). As a consequence, 2D fingerprints might not be the best choice for screening if scaffold hopping is the primary goal. 3D similarity methods, however, are expected to perform better in recognizing active molecules with diverse scaffolds, because they consider the volume of the molecules, not the topology (Vainio et al. 2009).

2.1.2 Structure-based methods

The most widely used method for structure-based VS is molecular docking (Fig. 3A, 3B). In docking the ligands are fitted into the binding site of the target protein, and the complementarity of steric and chemical properties is evaluated. The ligands can be treated either rigidly or flexibly, whereas the protein structure is usually static. The prediction of the best binding mode and the binding affinity to the protein is accomplished with scoring functions. There are three types of scoring functions: force field based, empirical, and knowledge based. Force field based scoring functions employ molecular mechanics force fields. In these scoring functions the van der Waals (vdW) and electrostatic interactions are calculated between the ligand and the protein, and additionally intramolecular energies can be included in the scores. Empirical scoring functions are parameterized to reproduce experimental data, such as binding affinity or conformations. These functions frequently include individual terms for hydrogen bonds, ionic interactions, hydrophobic effects, and entropy. Knowledge based scoring functions have been developed by statistical analysis of protein-ligand structural data. The assumption is that atoms or functional groups found frequently in close proximity to each other are energetically favorable, and therefore contribute favorably to the binding affinity. Some scoring functions use a combination of the scoring function types described above (Plewczynski et al. 2011).

Docking programs use different search algorithms for flexible docking of small molecules. The approaches can be systematic, stochastic or simulation based. Systematic methods often use an incremental construction algorithm where the rigid parts of the ligand are first docked into the binding site followed by docking of the flexible fragments with systematic sampling of the torsion angles. Stochastic algorithms generally use Monte Carlo (MC) based simulations or genetic algorithms to sample the conformational space. The simulation based approaches employ either molecular dynamics (MD) simulations or energy minimization (Kitchen et al. 2004). Because the docking methods are more complex compared to ligand-based VS, they also require more computation time. Table 2 summarizes the commonly used docking software.

Although docking is widely used in structure-based VS, there are certain limitations and challenges associated with it. One of the problems is protein flexibility. Proteins can change conformation when different ligands bind to them, and docking may fail if the protein conformation is not favorable for a particular ligand. One solution is to use multiple protein structures for docking and combine the results. These protein structures can be either other crystal structures or derived with computational methods, such as MD simulations. Some docking software allow the user to define key amino acids within the binding site to be treated flexibly (Jones et al. 1997, Morris et al. 1998), and also an approach called induced fit has been developed (Sherman et al. 2006a, b). The caveat in the use of multiple protein structures or flexible protein docking is naturally the increased computation time.

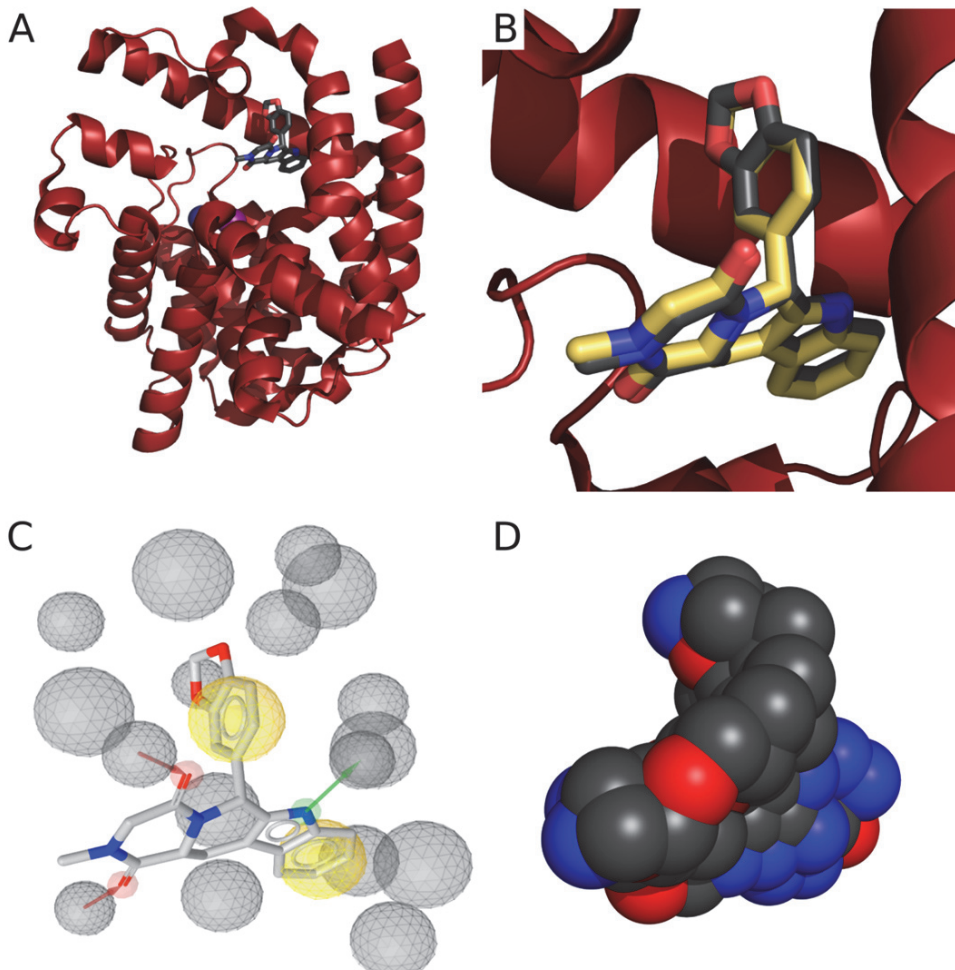


FIGURE 3 Different structure-based virtual screening methods. A) The crystal structure of phosphodiesterase 5 (PDE-5; pdb: 1xoz) with the inhibitor tadalafil (dark grey carbon atoms). Also visible are the Mg^{2+} and Zn^{2+} atoms within the binding site in blue and magenta space-filling models. B) Tadalafil docked with Glide into the binding site (yellow carbon atoms). The conformation of the ligand in the crystal structure is shown for reference (dark grey carbon atoms). C) A pharmacophore model of PDE-5 based on the 1xoz crystal structure. Red arrows indicate hydrogen bond acceptors, green arrow indicates a hydrogen bond donor, yellow spheres depict aromatic rings, and grey spheres indicate excluded volumes. The model was derived with LIGANDSCOUT (Wolber & Langer 2005). D) A negative image of the binding site of PDE-5 generated with VOIDOO/FLOOD (Kleywegt & Jones 1994). The colors in NIB model represent the electrostatic properties extracted from the protein structure: dark grey = neutral; red = positive; blue = negative.

TABLE 2 List of commonly used docking software.

Software	Conformational searching method	Scoring function	References
AUTODOCK	Genetic algorithm	Empirical	Morris et al. 1998
DOCK	Incremental construction	Force field	Ewing et al. 2001
EHITS	Exhaustive systematic	Knowledge based/ empirical	Zsoldos et al. 2007
FLEXX	Incremental construction	Empirical	Rarey et al. 1996
FRED	Shape matching of multiple conformers	Gaussian/ empirical	McGann et al. 2003
GLIDE	Incremental construction, Monte Carlo optimization	Empirical	Friesner et al. 2004
GOLD	Genetic algorithm	Force field	Jones et al. 1997
HAMMERHEAD	Incremental construction	Empirical	Welch et al. 1996
ICM	Monte Carlo simulation	Force field/ empirical	Abagyan et al. 1994
LIGANDFIT	Shape matching, Monte Carlo simulation	Empirical	Venkatachamal et al. 2003
SURFLEX	Surface-based molecular similarity, incremental construction	Empirical	Jain 2003
QXP	Monte Carlo minimization	Force field	McMartin & Bohacek 1997

Water molecules can mediate important interactions between the protein and the ligand. For docking this is challenging, because for some ligands the water molecule(s) can be crucial for binding while in other cases they can be replaced by ligand atoms. Also the optimal position of the water molecules can change depending on the ligand. The docking programs can usually accept waters as part of the protein structure. In some of the docking approaches user can also specify certain water molecules to be 'on' or 'off' (Jones et al. 1997), while in others explicit water molecules are docked to appropriate places within the binding site (Rarey et al. 1996, Friesner et al. 2004).

The most challenging issue in docking, however, is the 'scoring problem'. It has been shown that none of the currently available scoring functions can provide good results for a variety of different protein targets (McGaughey et al. 2007, Cross et al. 2009, von Korff et al. 2009). Although in many cases the binding mode can be predicted with high accuracy (Fig. 3B), the prediction of binding affinities is often case-specific. This inaccuracy in the scoring functions often leads to many false positive hits in the results. Thus, the applicability of a particular docking method and scoring function for a specific protein target should be carefully evaluated prior to VS to ensure reliable results.

2.1.3 Pharmacophore modeling

A pharmacophore model includes the features required for molecular recognition between the small molecule and the macromolecular target (Fig. 3C). These features include steric properties and chemical properties, such as hydrogen bond donors or acceptors, hydrophobic interactions and ring structures. The pharmacophore model can be derived either from known ligands, where several different known active molecules are used to identify the common important features, or from the target protein structure. However, the creation of a reliable pharmacophore model can be challenging and time-consuming (Yang 2010). To overcome the problem of complex pharmacophore models, PHASE (Dixon et al. 2006a, b) offers a ligand-based pharmacophore modeling method in which the biological activities of the molecules can be incorporated into the generation of the model. A threshold activity value can be set to affect the complexity of the resulting pharmacophore model. This approach can also create a 3D quantitative structure-activity relationship (3D-QSAR) model for the protein system under study, which can be used to find correlation between structural differences in the molecules and their biological activities. 3D-QSAR modeling is discussed further below.

2.1.4 3D quantitative structure-activity relationship

3D-QSAR is a statistical method to find correlations between experimental binding data and structural properties of molecules. Common uses for 3D-QSAR models include the prediction of binding affinities for new structures, the detection of regions that may have significant effects on activity, and providing information of the important interactions between the small molecule and the protein. For the generation of a 3D-QSAR model, active molecules with known affinities are needed. Some methods require that the molecules are aligned with each other. The alignment can be made for example by superimposing the molecules in order to find the best common 3D alignment, or the alignment can be made by using information from protein structures, such as by using multiple protein crystal structures or by docking the molecules into the binding site. The validity of the model can be tested by predicting the activities of a set of ligands not present in the training set. An example of a widely used 3D-QSAR method is CoMFA (Comparative Molecular Field Analysis; Cramer et al. 1988), which calculates steric and interaction fields for the aligned molecules by using an atom probe. Correlation between these field energy terms and activity is then calculated with partial least squares (PLS) method. The quality and applicability of the 3D-QSAR model depends heavily on the training dataset, and in the case of alignment-dependent methods, also on the alignment. Thus, it cannot predict the activities for structures that are not present in the dataset that was used to create the model or predict activities that are outside the range of the training dataset.

2.1.5 Negative image-based screening

In molecular recognition, the shape-complementarity between the small molecule and the macromolecular target is an important requirement. In negative image-based (NIB) screening the shape of the binding site is described as a ligand-like structure (Fig. 3D), which is then used with ligand shape-based comparison methods in place of known active molecule(s). Compared to ligand-based approaches, NIB screening has the advantage that it is not limited to known ligand structures, but instead searches for molecules that resemble the shape of the binding site. This also means that NIB screening can be used for targets that do not have any known small molecular modulators, or for targets for which totally different kinds of ligands to existing ones are searched for. The docking methods also take into account the shape complementarity of the protein and the ligand; however, since NIB screening utilizes ligand-based similarity comparison methods, it is generally more efficient computationally than docking. As the chemical complementarity is also of importance, the chemical similarity can be included in the searches in the form of pharmacophore points (Fig. 3C) or electrostatics (Fig. 3D) extracted from the binding site.

The first study reporting the use of negative image-based VS was done by Ebalunode et al. (2008). They generated the negative image of the binding site, and extracted the relevant pharmacophore features from the target protein structure. This ligand-like structure was then used with ROCS for similarity comparison. The method showed better performance compared to docking and ligand-based VS. They also showed that the addition of chemical information into the negative images improved the efficiency of VS.

In this thesis another kind of approach to NIB screening was developed, in which the chemical properties of the binding site can direct the shape of the NIB model (I). This means that in places were hydrogen bonding may occur the model is allowed to go closer to the binding site surface and mimic hydrogen bonding distances, while in places where van der Waals interactions dominate, the NIB model is farther from the surface. The results showed that shape-based NIB screening produced better results compared to docking and ligand shape-based VS. Additionally, it was discovered that incorporating the protein flexibility into NIB screening by using multiple protein structures to generate the NIB models, improves the VS efficiency. The method was developed further to account for the chemical interactions between the ligand and the protein (II). The chemical information was added to the NIB models in the form of electrostatics. The results showed that the addition of electrostatic information to the NIB models can improve the efficiency of VS, however, this is case-specific depending on the nature of the target protein binding site.

2.2 Binding free energy calculations

The ability to reliably predict the binding free energies of small molecules is of particular importance in drug development. The prediction of binding affinity can be rapidly made with scoring functions; however, as mentioned earlier, the approximated scoring functions do not always give reliable results. More rigorous free energy calculation methods include free energy perturbation (FEP) (Kollman 1993) and thermodynamic integration (TI) (Lybrand et al. 1986) but these are unsuitable for VS applications because of their computational intensity. Methods that combine molecular mechanics force fields with continuum solvent models have become popular in binding free energy calculations because of their accuracy and computational efficiency.

2.2.1 MM-GBSA and MM-PBSA

Massova and Kollman introduced a method for free energy calculations which takes into account the desolvation energy in addition to molecular mechanics force fields (Massova & Kollman 2000). The water molecules are replaced with a continuum solvent model, such as Poisson Boltzmann (PB) or Generalized Born (GB), which is used to calculate the polar contribution to solvation. Nonpolar contribution is proportional to the solvent accessible surface area (SA). The van der Waals and electrostatic energies are calculated with molecular mechanics (MM) force fields. The binding free energies are usually calculated for an ensemble of conformations derived with MD simulations. The binding free energies (ΔG_{bind}) can be estimated from the free energies of the three reactants, the receptor, the ligand, and the receptor-ligand complex, with the following equation:

$$\Delta G_{bind} = \langle G_{comp} \rangle - \langle G_{rec} \rangle - \langle G_{lig} \rangle$$

$\langle \ \rangle$ denotes an average over a set of snapshots along the MD trajectory, and comp, rec, and lig stand for complex, receptor, and ligand, respectively. The free energy for each of the reactants is estimated with the equation:

$$G = E_{MM} + G_{solv} - TS_{solute}$$

where E_{MM} is the molecular mechanics contribution in vacuo consisting of the sum of internal, electrostatics, and van der Waals energies; G_{solv} is the contribution of solvation free energies expressed as the sum of polar (G_{GB}) and nonpolar ($G_{nonpolar}$) solvation free energies; T is the temperature; and S_{solute} is the solute entropy. Nonpolar solvation free energy can be estimated with $G_{nonpolar} = \gamma$ SAS, where γ is the surface tension constant and SAS is the solvent accessible surface. The entropy term in the binding free energy is usually estimated with normal mode analysis (NMA).

MM-GB(PB)SA methods have been used previously in many studies, where their efficiency in predicting binding affinities has been evaluated (Wang et al. 2001, Ferrari et al. 2007, Guimarães & Cardozo 2008, Hou et al. 2011). The results have shown good correlation between the predicted and experimental binding data (Wang et al. 2001, Ferrari et al. 2007, Guimarães & Cardozo 2008). However, the results also indicate, that these methods are sensitive to the protocol used for generating the geometries for the protein-ligand complex, the solvent model used in free energy calculations, and also the length of the MD simulation (Ferrari et al. 2007, Hou et al. 2011). In some studies also the VS efficiency of these methods has been evaluated, however, the research has been limited to only small and limited databases/targets.

2.2.2 SIE

A similar method compared to MM-GB(PB)SA, called Solvated Interaction Energy (SIE), was developed by Naïm et al. (2007). The motivation for the development of this approach was that the implicit solvation models contain parameters that are fitted to experimental solvation free energies of small molecules, and thus are not optimal for protein-ligand binding free energy calculations. The SIE method contains five parameters that have been fitted to experimental binding free energies of 99 protein-ligand complexes. The SIE calculations can be made with the software SIETRAJ, which employs the equation:

$$\Delta G_{bind}(\rho, D_{in}, \alpha, \gamma, C) = \alpha * [E_C(D_{in}) + \Delta G_{bind}^R(\rho, D_{in}) + E_{vdw} + \gamma * \Delta SA(\rho)] + C$$

where E_C and E_{vdw} are the intermolecular Coulomb and van der Waals interaction energies in the bound state, respectively. ΔG_{bind}^R is the change in the reaction free energy between the bound and free states respectively, calculated by the Poisson-Boltzmann equation. The ΔSA term is the change in molecular surface area upon binding. The optimized values for the coefficients are: ρ = 1.1 (AMBER van der Waals radii linear scaling coefficient), α = 0.1048 (global proportionality coefficient relating to the loss off configurational entropy upon binding), γ = 0.0129 kcal mol^{-1} Å^{-2} (molecular surface area coefficient), D_{in} = 2.25 (solute internal dielectric constant), and C = -2.89 kcal mol^{-1}.

The SIE method has been used for example by Yang and colleagues to predict the binding affinities of phosphodiesterase 2 (PDE-2) inhibitors, showing predicted binding free energies that were in excellent correlation with experimental binding data (Yang et al. 2010). However, to our knowledge, the SIE method has not been studied in VS applications.

3 AIMS OF THE STUDY

The aim of this study was to develop a new efficient method for virtual screening. This method is based on the shape of the ligand binding site in addition to the chemical properties in the form of electrostatics. Additionally, different methods for free energy calculations were studied to determine their efficiency in virtual screening applications as well as their potential to accurately predict experimental binding affinity data.

4 METHODS AND MATERIALS

4.1 Molecular databases

Molecular databases available at the Directory of Useful Decoys (DUD) (Huang et al. 2006) were employed in the virtual screening studies. The DUD database contains a set of active ligand molecules and decoy molecules for 40 different protein targets. The decoy molecules have been chosen from the ZINC database (Irwin & Shoichet 2005) based on their similarity in their physicochemical properties compared to the ligands, but dissimilarity in their topology. The decoys are assumed to be inactive, although their activities have not necessarily been tested. Molecules with known affinities to target proteins were downloaded from the CHEMBL database (Bellis et al. 2011).

4.2 Negative images

VOIDOO/FLOOD (Kleywegt & Jones 1994) was used to generate the negative images for the ligand binding sites of the proteins. VOIDOO is a grid-based cavity detection software, which uses a probe to detect the solvent accessible surface area. The parameters of the program were modified to make the resulting negative images reflect more closely the chemical properties of the binding site. This means that the negative image was allowed to go closer to the polar groups of the amino acids, while in nonpolar areas the distance to the binding site was bigger. In order to accomplish this, the vdW radii of electronegative atoms, oxygen and nitrogen were reduced from their original values, whereas the radius for carbon was increased. The vdW radius for oxygen was reduced from 1.6 Å to 1.2 Å, and for nitrogen from 1.75 Å to 1.2 Å. The radii for carbon was increased from 1.85 Å to 2.25 Å. FLOOD can be used to fill the detected cavity with solvent molecules (or other atoms/molecules).

Another cavity detection tool, PASS (Putative Active Sites with Spheres; Brady & Stouten 2000), was also tested for negative image creation. Default

settings were used. PASS creates one file of the found cavities, and the correct one was manually selected with BODIL (Lehtonen et al. 2004).

In order to include protein flexibility into the NIB screening, several crystal structures for the target proteins were downloaded from the PDB (I, Table 1). The structures were superimposed, and the water molecules and ligands were removed with BODIL.

To add atom-centered partial charges to the data points in the negative image (II), MMFF94 charges (Halgren 1996) were assigned to the amino acids within the binding site. Then charges within 2.7 Å from each data point were averaged, and the opposite charge was assigned to each data point.

4.3 3D similarity searching software

In the study two similarity searching software, ROCS and SHAEP, were used. In article I, where only the shape of the ligand-binding site was compared, SHAEP was used with 'onlyshape' as well as ROCS was used with 'shapeonly' option. In II and III the full algorithm of SHAEP, which compares both the shape similarity combined with electrostatic potential similarity, was used. The default setting for the weighting of the shape and electrostatics were employed.

4.4 Docking

4.4.1 GOLD (I and III)

GOLD 4.1 was used in the docking studies. Hydrogens were added to the protein structures with TLEAP available in ANTECHAMBER (Wang et al. 2006). Default settings in GOLD were used for the genetic algorithm, which allows 10 docking trials for each molecule. The water-fill radius of 15 Å was used. GoldScore fitness function was used for scoring. The fitness function is in the form of:

$$F = S_{hb_ext} + 1.375 * S_{vdW_ext} + S_{hb_int} + S_{vdW_int}$$

where S_{hb_ext} is protein-ligand hydrogen bond energy, S_{vdW_ext} is protein-ligand vdW energy, S_{hb_int} is ligand internal hydrogen bond energy, and S_{vdW_int} is ligand internal vdW energy. The internal hydrogen bonding term is optional, and was not taken into account in the scoring.

4.4.2 GLIDE (I and II)

GLIDE 5.5 was used. The protein crystal structures were prepared for docking with the Protein Preparation Wizard in Maestro (Schrödinger, LLC, New York, NY). The receptor grid for docking was generated with GLIDE. Default settings

in the HTVS (I) and SP (II) were used. The scoring function *GScore* has the following form:

$$GScore = 0.065 * vdW + 0.130 * Coul + Lipo + Hbond + Metal + BuryP + RotB + Site$$

where *vdW* is vdW energy, *Coul* is coulombic energy, *Lipo* is the lipophilic term, *Hbond* is the hydrogen bonding term, *Metal* is the metal binding term, *BuryP* is penalty for buried polar groups, *RotB* is penalty for freezing rotatable bonds, and *Site* is polar interactions in the active site.

4.5 Molecular dynamics simulations

4.5.1 MD simulations with NAMD (I)

Molecular dynamics simulations with NAMD (Phillips et al. 2005) were performed to the proteins for conformational searching (I). The apo form of the protein was used. To prevent the closure of the binding site during the MD simulation, VOIDOO/FLOOD was used to add water molecules to the binding pocket. Solvent radius of 2.1 Å was used in FLOOD. TLEAP was used to (1) add hydrogens to the protein structures, (2) solvate the system with transferable intermolecular potential three-point (TIP3P) (Åqvist 1990) water molecules 13 Å in all directions, and (3) derive force field parameters for the protein (ff03) (Duan et al. 2003). MD simulations were performed as described previously (Postila et al. 2010). In short, the ligand-free protein structures were first energy-minimized in two steps: the water molecules and amino acid side chains were first minimized with a conjugate gradient algorithm (15,000 steps) with C^{α}-atoms restrained with a harmonic force of 5 kcal mol^{-1} $Å^{-2}$, after which the protein was minimized without constraints (15,000 steps). The MD simulations were first run in constant volume and temperature for 360 ps (time-step of 2 fs) with C^{α}-atoms restrained as in the minimization step. This was followed by the production simulation of 2.4 ns (time-step of 2 fs) in constant pressure and temperature.

4.5.2 MD simulations with SANDER (II and III)

In II and III MD simulations were used to derive conformations for protein-ligand complexes for free energy calculations. The SANDER module distributed in AMBER 10 (University of California, San Francisco, CA) package was used for MD simulations. The protein-ligand complex generated with either GOLD or SHAEP was used as a starting conformation for MD simulations. Charges for the ligands and for NADP were derived with AM1-BCC (Jakalian et al. 2000) available in ANTECHAMBER. TLEAP was used to create force field parameters for the protein (ff03) and the ligand/cofactor (gaff) (Wang et al. 2004), add hydrogens, and solvate the ligand-protein complex with a rectangular box of transferable intermolecular potential three-point water molecules (TIP3P) 4 Å in

all directions. The MD simulations were run as previously described. Briefly, the system was first minimized with conjugate-gradient method for 1,000 steps without restraints. This was followed by an equilibration step at constant volume by allowing the system to heat from 100 K to 300 K for 1,000 steps with NMR restraints. The production simulation without restraints was run for 20,000 steps (simulation time 40 ps) in (II) and 256,000 steps (simulation time 512 ps) in (III) at constant pressure. Temperature was maintained with the Berendsen thermostat (Berendsen et al. 1984). Electrostatics were treated with Particle-Mesh Ewald (PME) method (Darden et al. 1993, Petersen 1995) and cutoff value of 12 Å for nonbonded interactions was employed. The equilibration step and the production simulation were run under periodic boundary conditions. The SHAKE algorithm (Ryckaert et al. 1977) was used for bonds involving hydrogens, which allowed the use of 2 fs time step.

4.6 MM-GB(PB)SA

For the binding free energy analyses snapshots at 400 fs intervals (II) and 4 ps intervals (III) were extracted from the MD trajectory files and the free energies were averaged over the ensemble of conformers produced, leading to 100 snapshots (II) and 128 snapshots (III). The atomic cavity radii and charges were taken from the corresponding topology files. Dielectric constants of 1 and 80 were used for the solute and the solvent, respectively. Amber package offers 3 GB models for MM-GBSA calculations, IGB1, IGB2, and IGB5, respectively. In (II) only the model IGB1, developed by Hawkins et al. (1996) with the parameters of Tsui and Case (2001), was used. In (III) the MM-PBSA method along with all 3 GB models for MM-GBSA calculations were used. The hydrophobic contribution to the solvation free energy was estimated by calculating the solvent accessible surface area with Molsurf (Connolly 1983) and using a probe radius of 1.4 Å. The surface tension constant γ was set to 0.0072 kcal mol^{-1} $Å^{-2}$.

4.7 SIE

In III also SIE method was tested by using the software SIETRAJ. Default settings were used in SIETRAJ, and the same 128 MD snapshot structures extracted in 4 ps intervals were studied in free energy calculations.

4.8 Metrics in the virtual screening studies

The efficiency by which the method is able to enrich the active ligands from the database was depicted by calculating receiver-operating characteristics (ROC)

curves. The area under the curve (AUC) equals the probability of ranking randomly selected active molecule higher than randomly selected inactive molecule (Hanley & McNeil 1982). For a perfect screening method that recognizes all the actives before decoys the AUC value equals 1, whereas a diagonal line with AUC 0.5 describes a situation where random picking of the molecules is evenly good as the tested method.

Early enrichment was studied by calculating the enrichment factors at 1%, 5%, and 10% of the screened databases. The enrichment factors ($EF_{n\%}$) were calculated with the following equation:

$$EF_{n\%} = (Ligs_{n\%} * Mols_{all}) / (Mols_{n\%} * Ligs_{all})$$

Where $Ligs_{n\%}$ and $Mols_{n\%}$ are the number of ligands and molecules in the top n% of the ranked compounds. $Ligs_{all}$ and $Mols_{all}$ are the number of ligands and molecules, respectively, in the entire database.

5 RESULTS

5.1 Negative image based (NIB) screening

Nuclear hormone receptors were used as models for the NIB shape-based VS. These were optimal targets for NIB screening because the binding sites are within the protein and clearly definable. Negative images were created for the receptors, and the VS efficiency was evaluated by using the databases available in DUD. The results were compared to results from ligand shape-based VS and docking with GLIDE and GOLD.

5.1.1 Optimization of VOIDOO/FLOOD parameters

The solvent radius used in FLOOD determines the 'density' of the negative image. A low solvent radius leads to very dense NIB model, which defines well the shape of the binding site (I, Fig. 1A). Higher radius on the other hand leads to NIB model where the data points are located very sparsely, and thus the shape of the binding site might not be described optimally (I, Fig. 1B). The data point density naturally has an effect on the computation time and also on VS efficiency. With low solvent radius (< 0.5 Å) the computation time increases significantly (I, Fig. 2). The AUC values, however, gradually drop when the solvent radius of more than 1 Å is used. A compromise between the computational and VS efficiency was made by choosing the radius of 0.7 Å. This also leads to NIB models where the distance between data points is 1.4 Å, which is an average of single and double bonds between two carbon atoms, and thus represents adequately bond lengths in real organic molecules (I, Fig. 1C, Fig. 1D). In grid-based methods also the orientation can affect the volume of the detected cavity. However, the effects on the VS efficiency were only minor (I, Fig. 2).

5.1.2 Virtual screening efficiency with negative image

With the optimized protocol described above NIB models were generated for selected protein targets (I, Table 1). In 7 out of 10 studied cases NIB screening

produced equal or higher enrichment compared to the ligand-based screening with SHAEP (I, Table 2). This implies that the negative image is better in identifying the active molecules. ROCS did not perform as well as SHAEP for the negative image-based screening, where the negative image was better in only 5 out of 10 studied targets (I, Table 2).

The AUC values for progesterone receptor (PR) were near or below random with every studied method (I, Table 2). The highest enrichment was achieved with GLIDE (AUC 0.54 ± 0.06). When inspecting the set of active ligands, it is evident that the molecules are of various different shapes. Thus, the one ligand, or negative image that is based on one crystal structure does not represent ideally the wide variety of the ligands. For that reason another protein crystal structure (PDB code: 1a28), containing the natural ligand of the receptor (progesterone) was also tested. The AUC values increased from 0.50 ± 0.06 to 0.67 ± 0.06 with both the negative image and the ligand when SHAEP was used in shape comparison (I, Table 2). While the result is still quite poor, the improvement is clear and shows well the importance of choosing the right query in shape-based screening.

The docking results show that GLIDE gave clearly better enrichment than GOLD. However, ligand-based and NIB virtual screening were better compared to docking (I, Table 2).

5.1.3 Use of multiple protein conformations

To include protein flexibility into the search multiple protein structures were used in the generation of the negative images. The different protein structures were either other crystal structures or generated by MD simulations. In many cases the use of multiple NIB-models in VS produced higher enrichment than a single NIB model (I, Table 2). Only for retinoid X receptor alpha (RXR-alpha) the use of a single NIB model produced better result than the multi-NIB models, however, this might be due to the fact that the ligands are very similar to the ligand crystallized with protein structure 1mvc. In general, the NIB-models created from X-ray crystal structures showed better performance. This is understandable, because these protein structures had already been in complex with a ligand, and thus in one energy minimum, whereas the snapshots from MD simulations could contain some structures that are improbable for binding. In contrast with the other targets, for PR the multi-NIB model generated from MD structures produced higher AUC value than the multi-NIB model from X-ray structures with AUC values 0.87 ± 0.04 and 0.79 ± 0.05, respectively. The binding site of PR contains amino acids with highly flexible side chains, which allows the binding site to assume many different conformations (I, Fig. 3).

5.1.4 Negative images generated with PASS

Another software, PASS, was also tested for NIB-model generation (I, Table 3). The downside in PASS is that the parameters are not adjustable, and thus the generated model cannot take into account the hydrogen bonding or vdW

distances from the protein. Despite the limitations, also NIB models generated with PASS could produce reasonable VS results (I, Table 3).

5.2 Virtual screening of PDE-5 inhibitors (II)

The NIB-based virtual screening was developed further by adding electrostatics to the search. Phosphodiesterase-5 (PDE-5) was chosen as the target for studying the effects of NIB screening with added electrostatic information, because the binding site contains important polar interactions as well as hydrophobic/pi-pi-stacking interactions. The charges were added to the NIB models by calculating the average charge of the nearby amino acids within the binding pocket (II, Fig. 3). In the studies two different protein structures were used, one with sildenafil bound and the other with tadalafil. The studies were made by using either one of the protein structures or combining the results. In addition to the PDE-5 specific decoy set, the other ligands in the DUD set were used as an additional decoy set to make the enrichment studies more reliable. The PDE-5 specific decoy set is referred to as PDE-5 decoys and the additional decoy set is referred to as DUD decoys. The combined decoy set of PDE-5 and DUD decoys is therefore referred to as PDE-5/DUD decoys.

5.2.1 Ligand shape-based screening

Neither of the PDE-5 inhibitors, sildenafil or tadalafil, could produce high AUC values in ligand shape-based screening (II, Table 1). Sildenafil produced AUC values below 0.5 with both decoy sets (AUC = 0.48 ± 0.04, PDE-5/DUD; AUC = 0.42 ± 0.04, PDE-5). Tadalafil succeeded somewhat better with both decoy sets (AUC = 0.58 ± 0.04, PDE-5/DUD; AUC = 0.67 ± 0.04, PDE-5). The very poor results obtained with sildenafil stem from the fact that the ligand is much bulkier than most of the molecules in the PDE-5 ligand set. Combination of the results from the two inhibitors did not improve the results, but weakened them marginally with both decoy sets (AUC = 0.57 ± 0.04, PDE-5/DUD; AUC = 0.64 ± 0.04, PDE-5).

5.2.2 Protein structure-based screening

Docking. The enrichment with docking was relatively weak when the tadalafil-bound crystal structure was used (II, Table 2; AUC = 0.59 ± 0.04, PDE-5/DUD; AUC = 0.68 ± 0.04, PDE-5), but improved considerably with the sildenafil-crystal structure (AUC = 0.73 ± 0.04, PDE-5/DUD; AUC = 0.78 ± 0.04, PDE-5). One reason for the big difference between the two protein structures could be the different solvation states of the proteins. The combination of the docking results for the two protein structures produced the highest AUC value of 0.80 ± 0.03 (PDE-5 decoys). The docking was also the most efficient compared to other

methods when the very early enrichment at 0.5 % of the screened database was studied (II, Tables 3 and 4).

NIB screening. The shape-based NIB screening produced higher AUC value than the ligand shape-based screening (II, Table 1). The NIB model based on the tadalafil-bound structure (NIB solvation model 1, NIB-SM1) produced higher AUC values than the sildenafil-bound structure (NIB solvation model 2, NIB-SM2). Similarly to the inhibitor sildenafil, NIB-SM2 is also bulkier than tadalafil or NIB-SM1, and this explains why the smaller NIB-SM1 works better in the virtual screening studies. The combination of the NIB model results did not weaken the results in contrast to the ligand shape-based screening. The early enrichment when the top 1 % and 5 % were considered was good for NIB shape-based screening compared to ligand shape-based screening (II, Tables 3 and 4).

5.2.3 The effect of added electrostatic information

The shape-based screening produced weaker results than docking with both the ligand-based and NIB-based screening. The reason for this is probably the polar nature of the binding site of PDE-5, which is taken into account in docking but not in shape-based screening. The next step was to include electrostatics also to ligand-based and NIB-based screening.

Ligand-based screening. The addition of electrostatic information to the screening improved the results considerably (II, Table 1 vs. Table 2, Table 3, and Table 4) for the tadalafil-based search and combined results against PDE-5/DUD decoys. The sildenafil-based screening, however, remained at very low efficiency. This suggests that sildenafil structure does not contain information that could improve the efficiency of the ligand-based virtual screening at least with the used PDE-5 ligand set. The situation could be different, however, with a more diverse molecular set.

NIB-based screening. The AUC values improved by the addition of electrostatic information for the tadalafil-based NIB-model and for the combined results with PDE-5/DUD decoy set. Also the early enrichment improved; however, the docking was still more efficient when the top 0.5 - 1.0 % of the screened database was considered. The early enrichment at the top 0.5 - 1.0 % of the database for the NIB-SM2 was better than for the NIB-SM1 with the PDE-5 specific decoy set.

Effect of electrostatic information to other targets. The effect of added electrostatic information was studied also for targets studied in article I. These targets have binding sites that are largely hydrophobic and enclosed within the protein. The ligand-based virtual screening benefitted markedly from the addition of electrostatic information (II, Table 6). In contrast to PDE-5, docking produced weaker AUC values for the additional test cases compared to other methods, although the results were similar. For NIB screening the addition of electrostatic information improved the results slightly compared to shape-based screening, however, the early enrichment weakened considerably in some cases (II, Table 6). The weighting of shape and electrostatic potential in the scoring of the compounds produced similar results above and below equal weighting

(50:50). One exception is progesterone receptor, for which the highest AUC values were obtained with higher weighting for electrostatics (30:70). The reason for this is probably that the shape of the used NIB model is not optimal for all the ligands in the database, and thus overweighting the electrostatics helps to identify the true actives. This also implies that the usefulness of added electrostatic information to the NIB screening is more case-specific than with the ligand-based screening. The nature of the binding site, i.e. polarity or hydrophobicity are probably the key factors determining whether the electrostatics will improve the virtual screening results.

5.2.4 Post-processing with MD/MM-GBSA

More rigorous free energy estimation was employed for the top 5 % of the screened libraries to determine whether the early enrichment could be improved further. The concern was that the ligand conformations were not optimal in relation to the protein structure, and for this reason short MD simulations were run to create conformations for the protein ligand complexes. The free energies of binding were then calculated by using the MM-GBSA method and the compounds were re-ranked according to the results. Short (40 ps) MD simulations were used because it has been shown that longer simulations do not necessarily produce better results (Ferrari et al. 2007, Rastelli et al. 2010).

Rescoring of the docking results. The early enrichment at 0.5 - 1 % was very good for docking (II, Table 4), as was the AUC values with the PDE-5 specific decoy set (II, Table 2). The post-processing with MD/MM-GBSA did not improve the AUC values (II, Table 2) or the early enrichment (II, Table 4). In fact, the enrichment got considerably weaker. This might be due to poor alignment of the ligands in their docked conformation compared to the complexes generated with ligand-based or NIB methods.

Rescoring of the ligand-based screening results. According to the AUC value, the post-processing did not improve the virtual screening results (II, Table 2). However, the improvement in the results can be better seen when only the top 5 % of the compounds are used to calculate the ROC curves with both decoy sets (II, Fig. 3). The early enrichment improved considerably with the post-processing (II, Table 3). Especially the enrichment improved with the PDE-5/DUD decoy set with a 20-fold enrichment when the top 0.5 % of the rescored compounds was considered. The enrichment did not improve similarly with the PDE-5 specific decoys.

Rescoring of the NIB screening results. As was the case with rescoring the ligand-based screening results, the overall AUC values for NIB screening did not improve with the post-processing step (II, Table 2). However, the improvement is clearly visible when the ROC curves for the top 5 % of the compounds were studied (II, Figure 3). The number of found active hits (II, Table 3) and the early enrichment (II, Table 4) improved significantly with the rescoring with MD/MM-GBSA. The highest enrichment of all the methods was obtained with the PDE-5/DUD decoys: 39 for the top 0.5 % and 24 for the top 1 % (II, Table 4). The post-

processing of the NIB screening results thus enables the enrichment of the results to the degree that is required for efficient VS.

5.3 Comparison of free energy calculation methods in virtual screening (III)

The applicability of MM-GB(PB)SA and SIE for virtual screening was studied using the DUD set. Three different types of protein targets were chosen for the studies: aldose reductase 2 (ALR-2), PDE-5, and PR. The binding sites of the three targets differ from each other, so the applicability of the methods could be compared for different target systems. In ALR-2 the binding site is close to the surface of the protein and contains important polar and nonpolar interactions between the protein and the ligands (III, Fig. 2A and 2B). In PDE-5 the binding site is also close to the surface, which includes both important hydrophobic and polar interactions. Within the binding site there are also two metal ions, Zn^{2+} and Mg^{2+} (III, Fig. 2C and 2D). The binding site of PR is inside the protein and contains mostly hydrophobic residues (III, Fig. 2E and 2F). Since the generation of the protein-ligand complex can also affect the results of binding free energy calculations, two different approaches were tested: superimposition of molecules with SHAEP onto the ligand from the protein crystal structure according to similarity and molecular docking with GOLD. Even though GLIDE was better in VS than GOLD in the previous study (I), GOLD was used here because it was able to dock every ligand and decoy molecule into the protein targets, which is essential when comparing the VS efficiencies of the free energy calculation methods. The MM-GBSA calculations were made with all 3 GB models available in Amber, and these are later referred to as IGB1, IGB2, and IGB5, respectively.

5.3.1 Virtual screening efficiency

ALR-2. For ALR-2, the AUC values were low for every tested method, with the best obtained with MM-GBSA and the IGB5 solvation model for the structures generated with GOLD (AUC = 0.65 ± 0.07; III, Table 2, Fig. 2A and 2B). Overall, the GOLD-generated structures produced better results compared to structures generated with SHAEP. Despite of the poor overall enrichment as described by AUC values, the very early enrichment factors at 1 % and 5 % of the screened database were rather good for GOLD, MM-PBSA, and SIE. Although the enrichment with SHAEP benefitted from the post-processing with free energy calculations, the docking results did not significantly improve despite of the increased computational effort.

PDE-5. For PDE-5 the overall enrichments were comparable between GOLD and the free energy calculations with MM-PBSA, SIE, and MM-GBSA with IGB1 and IGB2 solvation models when GOLD was used in the structure-generation. The highest AUC value of 0.75 ± 0.03 was obtained with IGB1, as well as the highest early enrichments (III, Table 3, Fig. 2C and 2D). SHAEP was not as

successful in VS, producing an AUC value comparable to random picking of the molecules. However, the overall enrichments for the SHAEP-generated structures with IGB1 and IGB2 were good. In fact, the early enrichments were clearly better for IGB1 with the structures generated with SHAEP compared to structures generated with GOLD.

PR. For PR, docking did not succeed in VS, producing the lowest AUC values and enrichment factors compared to all the other tested methods (III, Table 4, Fig. 2E and 2F). However, overall the enrichments stayed at fairly low level for MM-GBSA with IGB1 and IGB2 models, and SIE, giving AUC values comparable to random picking of molecules. With MM-PBSA the AUC values obtained were slightly better (0.69 ± 0.05 at best) and MM-GBSA with IGB5 model was the best by producing clearly the highest AUC values (0.78 ± 0.05 at best). The early enrichments were, however, the highest for SIE and MM-PBSA.

5.3.2 Binding free energy calculations

ALR-2. The set of inhibitors for ALR-2 contains 6 molecules with diverse structures and wide range of activities (III, Table 1, Fig. S1). The MM-GBSA method with IGB1 and IGB2 solvation models produced the highest correlations between the predicted free energies and the experimental binding affinities, giving the correlation of 0.94 at best (III, Table 5). This is in sharp contrast with the third GB model, IGB5, for which the correlations were very low, or even negative. For MM-PBSA the highest correlation was 0.76, roughly the same as the correlations obtained for SHAEP-generated structures calculated with the SIE method. However, the correlations obtained with SIE for the GOLD generated structures were very low.

The results were analysed further by determining how many of the molecules with pIC50 values ≥ 6 would be detected within 1 %, 5 %, and 10 % of the screened DUD database (III, Table 9). For ALR-2 there was only 1 molecule with pIC50 ≥ 6 (tolrestat; III, Table 1), and it was found in the top 5 % only with SIE and the GOLD-generated structures. Otherwise the molecule was generally found in the top 10 %. The relatively poor identification of tolrestat among the top-scoring molecules is surprising since the protein crystal structure used in the studies had been crystallized with bound tolrestat, and further, the predicted binding conformations with both GOLD and SHAEP were very close to the crystal structure (RMSD values 1.90 and 0.48, respectively).

PDE-5. The inhibitor set for PDE-5 contains tadalafil and 8 analogs (III, Table 1, Fig. S1). In fact, the authors have measured the inhibitory activities for several different stereoisomers for the analogs. It is thus not surprising that in general the correlations are very low for all of the free energy calculation methods (III, Table 6). The best correlation of 0.35, was obtained with MM-GBSA and IGB5 model for the structures generated with SHAEP. The results show, that in spite of the overall correlation is very poor for all the cases, a closer look at the results reveals that in some cases the individual stereoisomers obtained rather good correlations (III, Table 8, Fig. 4). For example for the stereoisomers 4a-d the best correlation was 0.76, and for the 8a-d the best correlation was 0.94. However,

for the stereoisomers 7a-d the correlations were very poor or even negative, although the difference to the structures 8a-d is only a methyl group replaced by ethyl in 8a-d (III, Fig. S1).

For PDE-5 there are 11 molecules with pIC50 ≥ 6 (III, Table 1). MM-GBSA could not identify any of these molecules within the top 10 % of the screened database with the different GB models. Further, MM-PBSA could identify only 1 molecule within the top 5 % with the structures generated with SHAEP and 3 molecules within the top 10 % with the structures generated with GOLD, which also included the most potent inhibitor tadalafil. SIE could identify 2 molecules within the top 10 % with the GOLD-generated structures. The binding modes for tadalafil predicted with GOLD and SHAEP were not entirely optimal with RMSD values 3.02 and 2.43, respectively (III, Fig. 5), however, the results did not improve even when the crystal structure conformation for the ligand was used. This implies that the poor performance of the free energy calculation methods in ranking tadalafil analogs is probably caused by inaccurate parameters rather than incorrect binding modes.

PR. The set of inhibitors for PR contains mifepristone, used as reference, and 11 chromene-based analogs (III, Table 1, Fig. S1). The structures of the molecules are thus similar to each other, but also the affinities are in a rather narrow range. Considering this, the best correlation of 0.83 obtained with SIE and the GOLD-generated protein-ligand-complexes is rather good (III, Table 7). The other tested methods did not succeed as well, giving 0.71 at best for the IGB2 and the GOLD generated structures. In general, MM-GBSA with IGB1 and IGB2 models performed better compared to IGB5 and MM-PBSA. In fact, the correlations for IGB5 were close to zero or negative.

The molecule set for PR contains 8 molecules with pIC50 ≥ 6 (III, Table 1). According to the results, MM-GBSA with IGB5 model could identify all these molecules within the top 5 % with both complex generation methods (III, Table 9). SIE could also identify all these molecules within the top 5 % with the GOLD-generated structures. In addition, the most potent inhibitor, mifepristone, was identified within the top 1 % in all the studied cases. It is noteworthy that even though IGB5 could not predict the correct order of binding affinities for the molecules, it was the most successful in identifying the highly active molecules among the screened DUD database.

5.3.3 Effect of the MD simulation length

The effect of the length of the MD simulation was studied by calculating the AUC values and enrichment factors for structures obtained at 4 ps, 32 ps, 64 ps, 128 ps, 256 ps, 384 ps, and 512 ps of the MD simulation. Fig. 3 (III) shows how the AUC values change during the 512 ps simulation. The results show that in general the length of the simulation does not improve the VS efficiency. In fact, often the best results are obtained with a simulation of less than 128 ps. In addition, usually a single snapshot from the beginning of the simulation gave similar, or sometimes even better results than longer simulations.

The effect of the length of the simulation was also studied for the binding free energy calculations, and the correlation coefficients for the predicted binding free energies against experimental inhibition activity were calculated at the same time points. In contrast to the VS studies, there was more variation in the binding free energy calculations when evaluated in different time points. For ALR-2 and the structures generated with GOLD the correlations dropped significantly during the 512 ps simulation, whereas for the structures created with SHAEP the correlations stayed at a fairly constant level. For PR the correlations increased in general for structures generated with GOLD, whereas for the SHAEP-generated structures the correlations got worse for the majority of the cases. The reason for the different behavior caused by the different method of coordinate generation is that the starting conformation is different.

6 DISCUSSION

6.1 Negative image-based screening increases the efficiency of virtual screening

In this thesis a novel method for virtual screening was developed, where the binding pocket is described as a ligand-like entity to which molecular databases are compared. The results show that generally the NIB-based screening works better compared to ligand shape-based VS and docking. A related method, where the binding site was described as a negative image was used by Ebalunode and colleagues, and their results showed similar trend in NIB shape-based screening outperforming docking results (Ebalunode et al. 2008). The main difference in the method developed by Ebalunode et al. compared to ours is the way how the shape of the binding site is detected and defined. In our method the NIB model illustrates the chemical environment of the binding site by allowing the model go closer to the protein where hydrogen bonding is possible, while in hydrophobic areas the distance to the protein corresponds vdW distances. Additionally, in our method the distance between the data points corresponds to distances found in organic molecules. The reason for the better performance of NIB shape-based screening compared to ligand shape-based screening is presumably the fact that the screening with the shape of the binding site rather than with the ligand allows more ligands to be identified correctly, in contrast to the ligand-based search, which is limited to known active molecular structures. When comparing the NIB shape-based results to docking results, the difference can be explained by the inaccuracies in the scoring functions, which tend to yield many falsely identified positive hits. However, it must be taken into consideration that the DUD database used in this research for validation was originally developed to evaluate specifically docking algorithms, and thus it might be biased towards ligand-based methods (McGaughey et al. 2007).

6.2 The use of multiple protein structures for NIB screening can improve the efficiency

Protein flexibility can be a challenge in structure-based VS, because the binding of structurally different ligands can promote conformational changes also in the protein. In docking the protein flexibility can be taken into consideration by using different experimentally determined 3D structures, by creating rotamers for the amino acids, or by generating alternative conformations by computational methods, such as MD simulations, and dock all the molecules to each of the protein structures. Additionally, some docking algorithms allow the binding site to be treated partially flexibly. The downside is naturally the increased computational load. Similarly, in NIB screening the negative images can be created for multiple crystal structures or snapshot structures from MD simulations. Our results show that some targets can benefit enormously from the multiple binding site models. The most striking case is the PR, where the use of a different receptor structure with bound progesterone improved the VS efficiency of NIB shape-based screening from AUC 0.50 to 0.67. Further, the use of multiple crystal structures and conformations from MD simulations improved the results yet further to 0.79 and 0.87, respectively. The use of multiple protein structures in NIB screening is probably case-specific; some protein structures are known to be more flexible than others, and for these targets the use of multiple protein conformations may prove to be more beneficial, whereas in cases where the binding site is more rigid, a good "average" structure could probably suffice.

6.3 The addition of electrostatic information improves the negative-image-based screening for some targets

Even though the shape-based NIB screening often produced satisfactory results in terms of AUC values compared to both ligand shape-based screening and docking, it is evident that chemical properties important for binding could provide additional improvement to the NIB screening. For this reason the method was further developed by adding electrostatic information extracted from the binding site of the protein to the NIB model. The added electrostatic information did improve the results for PDE-5, which was used as a model target protein. However, the extent to which the added chemical information is useful probably depends on the nature of the binding site. More hydrophobic binding sites, such as those found in nuclear hormone receptors, might not benefit from the added electrostatic information as much as the protein targets with more polar binding sites. Ebalunode et al. (2008) also included chemical information in the form of pharmacophore properties to their negative image models. Accordingly, their results showed that the VS results benefitted from the added chemical information.

6.4 MM-GBSA is a viable method for enriching NIB screening results

In II we tested whether the screened database could be enriched further with MM-GBSA, and for this the top 5 % of the initial VS results from docking and NIB screening were rescored with MM-GBSA. The results show that the NIB screening benefitted significantly from the post-processing. Although for docking the early enrichment was better compared to ligand-based VS methods and NIB screening, the post-processing did not further improve the enrichment for the docking VS results, in fact the enrichment got worse. This is presumably caused by the different selection of molecules in the top 5 % of molecules in the VS results, and by the different initial conformations of the molecules.

6.5 MM-GB(PB)SA and SIE in virtual screening

The methods routinely used for VS include ligand-based similarity searching and molecular docking. However, the methods suffer from inherent limitations: ligand-based methods are limited by the structures available for similarity searches in addition to the lack of the protein structure information, whereas docking methods are limited by the inaccurate scoring functions. MM-GB(PB)SA and SIE methods could provide an alternative approach to VS. MM-GB(PB)SA methods have been tested earlier in their ability to distinguish between active and inactive molecules and the results have been promising, however, the studies have been limited to small datasets and the free energy calculations have been made to only single structures and not for an ensemble of snapshots (Thompson et al. 2008, Rastelli et al. 2010). In addition, to our knowledge, SIE has not been used in VS studies thus far.

In this research the VS efficiencies of MM-GB(PB)SA and SIE were studied by using three different protein targets: ALR-2, PDE-5, and PR. MM-PBSA was among the best of the tested methods for all the targets, however, the fact that it is computationally the most demanding of all the tested methods makes it the most unlikely choice for VS of large databases. Nevertheless, earlier studies as well as the results in this research show that also a single snapshot can provide equally good results as longer MD simulations (Thompson et al. 2008, Rastelli et al. 2010), and this makes also the MM-PBSA more applicable to VS. Of the MM-GBSA approaches, the IGB5 solvation model worked well for PR and ALR-2, whereas for PDE-5 the IGB1 and IGB2 produced the highest AUC values and enrichment factors. Two of the targets, namely ALR-2 and PR, are known to undergo ligand-dependent conformational changes upon binding, and this also allowed us to examine whether the conformational searching by minimization and subsequent MD simulation could help improve the VS results for flexible protein targets. For ALR-2 the VS results in terms of AUC values were very low

for every studied method, and the post-processing did not improve the outcome over docking. The VS efficiency for PR, however, benefitted from the post-processing, giving clearly better AUC values and early enrichments over docking. Two different methods were tested in the generation of the protein-ligand complex, ligand-based superimposition and molecular docking. In general, the structures obtained by docking produced slightly better results compared to ligand-based method, however, for PDE-5 the SHAEP-generated structures produced better early enrichment and overall enrichment with IGB1 and IGB2 solvation models in MM-GBSA calculations.

6.6 MM-GB(PB)SA in predicting experimental binding data

MM-GB(PB)SA and SIE have been previously studied to evaluate their ability to correctly predict the binding affinities of small molecules to their target proteins. The results have been promising, however, the results have also been reported to be highly dependent on many issues, such as the molecular targets, parameters used in the MD simulation and in the free energy calculations, the length of the simulation, and the method by which the starting coordinates are derived (Ferrari et al. 2007, Rastelli et al. 2010, Hou et al. 2011). In our studies we found that the free energy calculation methods could more successfully predict the affinities for molecules that were structurally diverse, and with wide range of affinities. When the molecules were more similar to each other, such as for PDE-5 and PR, the prediction of the binding affinities became more difficult. Despite of this, the methods were able to produce rather good correlations for some of the different stereoisomers of PDE-5 inhibitors. To our knowledge this is the first study where the ability of MM-GB(PB)SA or SIE to predict binding affinities of different stereoisomers has been evaluated.

In general, the MM-GBSA method with GB models IGB1 and IGB2 performed more consistently compared to the other methods, however, for PR the SIE method produced the highest correlation and MM-GBSA with IGB5 solvation model was the most efficient in identifying the highly active molecules among the screened DUD database. Even though MM-PBSA performed rather well in VS, the results in binding affinity predictions were not as good. This finding is supported by an earlier study by Hou et al. where they showed that despite of the increased computational effort, MM-PBSA does not necessarily give better results compared to MM-GBSA (Hou et al. 2011).

In contrast to the VS studies, in the binding affinity predictions the starting structures generated with SHAEP for MD simulations worked fairly well. In VS studies a long MD simulation (512 ps) was not essential for the accuracy of the results. However, in the binding affinity predictions, a longer simulation improved some of the results. This depended on the target and the method by which the protein-ligand complex was generated.

7 CONCLUSIONS

As a conclusion, in this research we have developed a novel method for virtual screening that employs the negative image of the binding site. We have shown that the VS results by this method are often better compared to docking and ligand-based VS. The protein flexibility can also be taken into account by using multiple protein structures in the creation of the NIB model, and the VS results can greatly benefit from this. As the complementary shape between the interacting molecules is not the only required property for efficient binding, the chemical information of the binding site was added to the NIB models in the form of electrostatics, which was found to improve the results for some targets. We also found that MM-GBSA is a viable method for enriching the NIB screening results to a level that is needed for successful VS.

We tested the free energy calculation methods MM-GB(PB)SA and SIE in VS and in their ability to predict experimentally determined binding affinities. We found that there are differences in the performance of the free energy calculation methods for different protein targets. In addition, we showed that despite of the high computational effort of these methods, they can be used for VS also for larger databases, as even a single snapshot is enough for the accuracy of the results. The results show, that the free energy calculations can be made with structures obtained from either docking or from ligand-based similarity comparisons, which makes it a possible post-processing tool to both structure-based and ligand-based VS. In predicting experimentally determined binding affinities, better results were obtained when the ligands were structurally different and with wide range of affinities, although in some cases good correlations were obtained even for different stereoisomers.

Acknowledgements

This research was carried out during 2008–2013 at the University of Jyväskylä, at the Department of Biological and Environmental Science. The thesis was funded by the Emil Aaltonen Foundation, the Sigrid Juselius Foundation, and the Drug Discovery Graduate School (DDGS). The computational resources were provided by CSC – IT Center for Science.

I would like to express my sincere gratitude to my thesis supervisor Dr. Olli Pentikäinen for all the advice and guidance during these past years. I am also grateful for the opportunity for a long research visit to Harvard University.

I would like to thank the reviewers of this thesis, Prof. Stefano Moro and Prof. Patrik Rydberg, for their good advice and constructive criticism. I am also very pleased that Prof. Jari Yli-Kauhaluoma accepted the invitation to be the opponent of this dissertation. I would like to thank my thesis committee members Dr. Tommi Nyrönen and Dr. Juhani Huuskonen. I am also grateful to Prof. Eugene Shakhnovich for allowing me to work in his group during 2010-2013.

Many thanks go to the whole CBL community: Pekka Postila, Elisa Nurminen, Jarkko Koivunen, Mikko Ylilauri, Sanna Niinivehmas, Jason Ahokas, and Antti Yli-Tainio, as well as to the members of the Shakhnovich group, especially Peter Kutchukian, Adrian Serohijos, Lee Yang, and Nicolas Cheron.

I am grateful to my friends and relatives for their support during these years. Especially I would like to thank my better half, Tommi, for being there the past 15 years.

YHTEENVETO (RÉSUMÉ IN FINNISH)

Virtuaaliseulonta; uuden rakennepohjaisen menetelmän kehitys

Virtuaaliseulonta on nykyään tärkeässä roolissa lääkeainekehityksessä. Virtuaaliseulonnalla on mahdollista tietokoneavusteisesti valita suurista, jopa miljoonia molekyylejä sisältävistä molekyylikirjastoista ne molekyylit, jotka suurimmalla todennäköisyydellä sitoutuvat lääkkeen kohdeproteiiniin. Tämän ansiosta aikaavievät ja kalliit biokemialliset testit voidaan kohdentaa vain näihin molekyyleihin ja siten tehostaa lääkemolekyylien etsintää.

Virtuaaliseulonnassa käytettävät menetelmät voidaan jakaa kahteen kategoriaan: ligandipohjaiset menetelmät ja proteiinin rakennepohjaiset menetelmät. Ligandipohjaisissa menetelmissä käytetään hyväksi tietoa tunnetuista sitoutuvista molekyyleistä ja etsitään molekyylikirjastoista samankaltaisia molekyylejä. Ligandipohjaisten menetelmien etuna on niiden nopeus, mikä mahdollistaa suurienkin tietokantojen seulonnan. Toisaalta näillä menetelmillä löydetään usein hyvin samankaltaisia molekyylejä kuin jo tunnetut sitoutuvat molekyylit, mikä ei ole toivottavaa jos tavoitteena on löytää täysin uudenlaisia lääkekandidaatteja. Proteiinin rakennepohjaisissa menetelmissä sen sijaan vaaditaan tietoa kohdeproteiinin kolmiulotteisesta rakenteesta, jonka perusteella etsitään muodoltaan ja kemiallisilta ominaisuuksiltaan proteiinin sitoutumistaskua vastaavia molekyylejä. Yleisin rakennepohjaisista menetelmistä on telakointi, jossa molekyylin sitoutumista proteiiniin arvioidaan erilaisten pisteytysfunktioiden avulla. Telakoinnin etuna ligandipohjaiseen virtuaaliseulontaan on sen riippumattomuus tunnetuista sitoutuvista molekyyleistä, mutta huonona puolena sen sijaan ovat epätarkat pisteytysfunktiot, joiden vuoksi saadaan usein vääriä positiivisia tuloksia.

Tässä tutkimuksessa on kehitetty uudenlainen rakennepohjainen virtuaaliseulontamenetelmä, jossa käytetään hyväksi proteiinin sitoutumistaskun muotoa ja kemiallisia ominaisuuksia. Menetelmässä sitoutumistaskusta muodostetaan ligandinkaltainen negatiivinen kuva, jonka perusteella voidaan hakea samankaltaisia molekyylejä molekyylitietokannoista käyttäen hyväksi perinteisiä ligandipohjaisia menetelmiä. Kemialliset ominaisuudet voidaan lisätä negatiiviseen kuvaan osittaisvarausten muodossa. Tämän menetelmän etuna perinteiseen ligandipohjaiseen virtuaaliseulontaan verrattuna on se, että tunnettujen ligandien asemasta käytetään sitoutumistaskun muotoa, mikä mahdollistaa täysin uudentyyppisten ligandien löytämisen. Tällä menetelmällä on myös mahdollista etsiä ligandeja proteiinikohteille, joille ei vielä edes tunneta ainuttakaan ligandia. Vastaavasti verrattuna telakointiin menetelmän etuina ovat sen nopeus ja riippumattomuus epätarkoista pisteytysfunktioista. Proteiinin joustavuus voidaan ottaa huomioon käyttämällä useita erilaisia proteiinirakenteita negatiivisen kuvan rakentamisessa. Tulosten mukaan tutkimuksessa kehitetty menetelmä tarjoaa nopean tavan seuloa suuriakin molekyylitietokantoja tuottaen parempia tai vastaavanlaisia virtuaaliseulontatuloksia verrattuna ligandipohjaiseen seulontaan ja telakointiin.

Molekyylien sitoutumisaffiniteetin ennustaminen luotettavasti on tärkeää lääkekehityksen eri vaiheissa. Laskennalliset menetelmät, joissa molekyylien sitoutumisaffiniteettia ennustetaan molekyylimekaniikan voimakenttien ja solvaatioenergian laskemisen avulla ovat herättäneet kiinnostusta myös virtuaaliseulonnassa, koska menetelmät ovat verrattain nopeita ja tulokset ovat tähän mennessä olleet lupaavia. Näihin menetelmiin kuuluvat MM-GBSA, MM-PBSA ja SIE. Tässä tutkimuksessa testattiin kyseisten menetelmien soveltuvuutta virtuaaliseulontaan sekä myös kokeellisesti määritettyjen sitoutumisaffiniteettien ennustamiseen kolmelle eri proteiinikohteelle. Tulosten perusteella MM-GBSA, MM-PBSA ja SIE voivat tarjota vaihtoehdon telakoinnille ja ligandipohjaiselle virtuaaliseulonnalle. Tätä tulosta tukee erityisesti se, että virtuaaliseulonnassa tuloksiin ei vaikuttanut molekyylidynamiikkasimulaation pituus, mikä on usein laskennallisesti vaativin osuus näissä menetelmissä. Sitoutumisaffiniteetin ennustamisessa parhaat tulokset puolestaan saatiin molekyyleille, joiden rakenteet poikkesivat toisistaan, ja joiden kokeellisesti mitattujen sitoutumisaffiniteettien erot olivat suuret. Tulosten perusteella kuitenkin MM-GBSA:n, MM-PBSA:n ja SIE:n suorituskyky niin virtuaaliseulonnassa kuin sitoutumisaffiniteettien arvioimisessa voi vaihdella huomattavasti eri proteiinikohteiden välillä ja on vaikea määritellä sääntöjä, joiden perusteella voisi ennustaa menetelmien sopivuutta tietylle proteiinille.

REFERENCES

Abagyan R.A., Totrov M.M. & Kuznetsov D.N. 1994. Icm - a new method for protein modeling and design. applications to docking and structure prediction from the distorted native conformation. J. Comp. Chem. 15: 488–506.

Åqvist J. 1990. Ion-water interaction potentials derived from free energy perturbations. J. Phys. Chem. 94: 8021–8024.

Ballester P.J., Westwood I., Laurieri N., Sim E. & Richards W.G. 2010. Prospective virtual screening with ultrafast shape recognition: the identification of novel inhibitors of arylamine n-acetyltransferases. *J. R. Soc. Interface*. 7: 335–342.

Bellis L.J., Akhtar R., Al-Lazikani B., Atkinson F., Bento A.P., Chambers J., Davies M., Gaulton A., Hersey A., Ikeda K., Krüger F.A., Light Y., McGlinchey S., Santos R., Stauch B. & Overington J.P. 2011. Collation and data-mining of literature bioactivity data for drug discovery. *Biochem. Soc. Trans.* 39: 1365–1370.

Berendsen H.J.C., Postma J.P.M., Van Gunsteren W.F., Dinola A. & Haak J.R. 1984. Molecular-dynamics with coupling to an external bath. *J. Chem. Phys*. 81: 3684–3690.

Brady G.P.J. & Stouten P.F. 2000. Fast prediction and visualization of protein binding pockets with pass. *J. Comput. Aided Mol. Des.* 14: 383–401.

Cannon E.O., Nigsch F. & Mitchell J.B.O. 2008. A novel hybrid ultrafast shape descriptor method for use in virtual screening. *Chem. Cent. J.* 2: 3.

Connolly M.L. 1983. Analytical molecular surface calculation. *J. Appl. Crystallogr*. 16: 548–558.

Cramer R.D., Patterson D.E. & Bunce J.D. 1988. Comparative molecular field analysis (comfa). 1. effect of shape on binding of steroids to carrier proteins. *J. Am. Chem. Soc.* 110: 5959–5967.

Cross J.B., Thompson D.C., Rai B.K., Baber J.C., Fan K.Y., Hu Y. & Humblet C. 2009. Comparison of several molecular docking programs: pose prediction and virtual screening accuracy. *J. Chem. Inf. Model*. 49: 1455–1474.

Darden T., York D. & Pedersen L. 1993. Particle mesh ewald: an w log(n) method for ewald sums in large systems. *J. Chem. Phys*. 98: 10089–10092.

Dixon S.L., Smondyrev A.M. & Rao S.N. 2006a. Phase: a novel approach to pharmacophore modeling and 3d database searching. *Chem. Biol. Drug Des.* 67: 370–372.

Dixon S.L., Smondyrev A.M., Knoll E.H., Rao S.N., Shaw D.E. & Friesner R.A. 2006b. Phase: a new engine for pharmacophore perception, 3d qsar model development, and 3d database screening. 1. methodology and preliminary results. *J. Comput. Aided Mol. Des.* 20: 647–671.

Duan J., Dixon S.L., Lowrie J.F. & Sherman W. 2010. Analysis and comparison of 2D fingerprints: insights into database screening performance using eight fingerprint methods. *J. Molec. Graph. Model.* 29: 157–170.

Duan Y., Wu C., Chowdhury S., Lee M.C., Xiong G., Zhang W., Yang R., Cieplak P., Luo R., Lee T., Caldwell J., Wang J. & Kollman P. 2003. A point-charge force field for molecular mechanics simulations of proteins based on condensed-phase quantum mechanical calculations. *J. Comput. Chem.* 24: 1999–2012.

Durant J.L., Leland B.A., Henry D.R. & Nourse J.G. 2002. Reoptimization of mdl keys for use in drug discovery. *J. Chem. Inf. Comput. Sci*. 42: 1273–1280.

Ebalunode J.O., Ouyang Z., Liang J. & Zheng W. 2008. Novel approach to structure-based pharmacophore search using computational geometry and shape matching techniques. *J. Chem. Inf. Model*. 48: 889–901.

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.

Ferrari A.M., Degliesposti G., Sgobba M. & Rastelli G. 2007. Validation of an automated procedure for the prediction of relative free energies of binding on a set of aldose reductase inhibitors. *Bioorg. Med. Chem.* 15: 7865–7877.

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. 2004. Glide: a new approach for rapid, accurate docking and scoring. 1. method and assessment of docking accuracy. *J. Med. Chem.* 47: 1739–1749.

Guimarães C.R.W. & Cardozo M. 2008. Mm-gb/sa rescoring of docking poses in structure-based lead optimization. *J. Chem. Inf. Model*. 48: 958–970.

Halgren T. 1996. Merck molecular force field. i. basis, form, scope, parameterization, and performance of mmff94. *J. Comp. Chem*. 17: 490–519.

Hanley J.A. & McNeil B.J. 1982. The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology* 143: 29–36.

Hartman G.D., Egbertson M.S., Halczenko W., Laswell W.L., Duggan M.E., Smith R.L., Naylor A.M., Manno P.D., Lynch R.J., Zhang G., Chang C.T.C. & Gould, R.J. 1992. Non-peptide fibrinogen receptor antagonists. 1. discovery and design of exosite inhibitors. *J. Med. Chem.* 35: 4640–4642.

Hawkins G.D., Cramer C.J. & Truhlar D.G. 1996. Parametrized models of aqueous free energies of solvation based on pairwise descreening of solute atomic charges from a dielectric medium. *J. Phys. Chem*. 100: 19824–19839.

Hong H., Xie Q., Ge W., Qian F., Fang H., Shi L., Su Z., Perkins R. & Tong W. 2008. Mold(2), molecular descriptors from 2d structures for chemoinformatics and toxicoinformatics. *J. Chem. Inf. Model*. 48: 1337–1344.

Hou T., Wang J., Li Y. & Wang W. 2011. 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.

Huang N., Shoichet B.K. & Irwin J.J. 2006. Benchmarking sets for molecular docking. *J. Med. Chem.* 49: 6789–6801.

Irwin J.J. & Shoichet B.K. 2005. Zinc--a free database of commercially available compounds for virtual screening. *J. Chem. Inf. Model*. 45: 177–182.

Jain A.N. 2003. Surflex: fully automatic flexible molecular docking using a molecular similarity-based search engine. *J. Med. Chem.* 46: 499–511.

Jakalian A., Bush B.L., Jack D.B. & Bayly C.I. 2000. Fast, efficient generation of high-quality atomic charges. am1-bcc model: 1. method. *J. Comp. Chem*. 21: 132–141.

Johnson A.M. & Maggiora G.M. 1990. *Concepts and applications of molecular similarity.* Wiley, New York.

Jones G., Willett P., Glen R.C., Leach A.R. & Taylor R. 1997. Development and validation of a genetic algorithm for flexible docking. *J. Mol. Biol.* 267: 727–748.

Kirchmair J., Distinto S., Markt P., Schuster D., Spitzer G.M., Liedl K.R. & Wolber G. 2009. How to optimize shape-based virtual screening: choosing the right query and including chemical information. *J. Chem. Inf. Model*. 49: 678–692.

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. & Jones T. 1994. Detection, delineation, measurement and display of cavities in macromolecular structures. *Acta Crystallogr. D Biol. Crystallogr.* 50: 178–185.

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

Lehtonen J.V., Still D., Rantanen V., Ekholm J., Björklund D., Iftikhar Z., Huhtala M., Repo S., Jussila A., Jaakkola J., Pentikäinen O., Nyrönen T., Salminen T., Gyllenberg M. & Johnson M.S. 2004. Bodil: a molecular modeling environment for structure-function analysis and drug design. *J. Comput. Aided Mol. Des.* 18: 401–419.

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.

Lybrand T.P., McCammon J.A. & Wipff G. 1986. Theoretical calculation of relative binding affinity in host-guest systems. *Proc. Natl. Acad. Sci. U.S.A.* 83: 833–835.

Massova I. & Kollman P.A. 2000. Combined molecular mechanical and continuum solvent approach (mm-pbsa/gbsa) to predict ligand binding. *Perspect. Drug Discovery Des.* 18: 113–135.

McGann M.R., Almond H.R., Nicholls A., Grant J.A. & Brown F.K. 2003. Gaussian docking functions. *Biopolymers* 68: 76–90.

McGaughey G.B., Sheridan R.P., Bayly C.I., Culberson J.C., Kreatsoulas C., Lindsley S., Maiorov V., Truchon J. & Cornell W.D. 2007. Comparison of topological, shape, and docking methods in virtual screening. *J. Chem. Inf. Model*. 47: 1504–1519.

McMartin C. & Bohacek R.S. 1997. Qxp: powerful, rapid computer algorithms for structure-based drug design. *J. Comput. Aided Mol. Des.* 11: 333–344.

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 an empirical binding free energy function. *J. Comput. Chem*. 19: 1639–1662.

Naïm M., Bhat S., Rankin K.N., Dennis S., Chowdhury S.F., Siddiqi I., Drabik P., Sulea T., Bayly C.I., Jakalian A. & Purisima E.O. 2007. Solvated interaction energy (sie) for scoring protein-ligand binding affinities. 1. exploring the parameter space. *J. Chem. Inf. Model*. 47: 122–133.

O'Boyle N.M., Banck M., James C.A., Morley C., Vandermeersch T. & Hutchison G.R. 2011. Open babel: an open chemical toolbox. *J. Cheminform*. 3: 33.

Petersen A. 1995. Accuracy and efficiency of the particle mesh ewald method. *J. Chem. Phys*. 103: 3668–3679.

Phillips J.C., Braun R., Wang W., Gumbart J., Tajkhorshid E., Villa E., Chipot C., Skeel R.D., Kalé L. & Schulten K. 2005. Scalable molecular dynamics with namd. *J. Comput. Chem*. 26: 1781–1802.

Plewczynski D., Łaźniewski M., Augustyniak R. & Ginalski K. 2011. Can we trust docking results? evaluation of seven commonly used programs on pdbbind database. *J. Comput. Chem*. 32: 742–755.

Postila P.A., Swanson G.T. & Pentikäinen O.T. 2010. Exploring kainate receptor pharmacology using molecular dynamics simulations. *Neuropharmacol*. 58: 515–527.

Rarey M., Kramer B., Lengauer T. & Klebe G. 1996. A fast flexible docking method using an incremental construction algorithm. *J. Mol. Biol.* 261: 470–489.

Rastelli G., Del Rio A., Degliesposti G. & Sgobba M. 2010. Fast and accurate predictions of binding free energies using mm-pbsa and mm-gbsa. *J. Comput. Chem*. 31: 797–810.

Rush T.S.3., Grant J.A., Mosyak L. & Nicholls A. 2005. A shape-based 3-d scaffold hopping method and its application to a bacterial protein-protein interaction. *J. Med. Chem.* 48: 1489–1495.

Ryckaert J., Ciccotti G. & Berendsen H. 1977. Numerical integration of the cartesian equations of protein and nucleic acids. *J. Comp. Phys*. 23: 327–341.

Sastry G.M., Dixon S.L. & Sherman W. 2011. Rapid shape-based ligand alignment and virtual screening method based on atom/feature-pair similarities and volume overlap scoring. *J. Chem. Inf. Model*. 51: 2455–2466.

Sastry G.M., Lowrie J.F., Dixon S.L. & Sherman W. 2010. Large-scale systematic analysis of 2D fingerprint methods and parameters to improve virtual screening enrichments. *J. Chem. Inf. Model.* 50: 771–784.

Schames J.R., Henchman R.H., Siegel J.S., Sotriffer C.A., Ni H. & McCammon J.A. 2004. Discovery of a novel binding trench in hiv integrase. *J. Med. Chem.* 47: 1879–1881.

Sherman W., Beard H.S. & Farid R. 2006a. Use of an induced fit receptor structure in virtual screening. *Chem. Biol. Drug Des*. 67: 83–84.

Sherman W., Day T., Jacobson M.P., Friesner R.A. & Farid R. 2006b. Novel procedure for modeling ligand/receptor induced fit effects. *J. Med. Chem.* 49: 534–553.

Thompson D.C., Humblet C. & Joseph-McCarthy D. 2008. Investigation of mm-pbsa rescoring of docking poses. *J. Chem. Inf. Model*. 48: 1081–1091.

Tsui V. & Case D.A. 2001. Theory and applications of the generalized born solvation model in macromolecular simulations. *Biopolymers* 56: 275–291.

Vainio M.J., Puranen J.S. & Johnson M.S. 2009. Shaep: molecular overlay based on shape and electrostatic potential. *J. Chem. Inf. Model*. 49: 492–502.

Venkatachalam C.M., Jiang X., Oldfield T. & Waldman M. 2003. Ligandfit: a novel method for the shape-directed rapid docking of ligands to protein active sites. *J. Mol. Graph. Model.* 21: 289–307.

von Korff M., Freyss J. & Sander T. 2009. Comparison of ligand- and structure-based virtual screening on the dud data set. *J. Chem. Inf. Model*. 49: 209–231.

Wang J., Morin P., Wang W. & Kollman P.A. 2001. Use of mm-pbsa in reproducing the binding free energies to hiv-1 rt of tibo derivatives and predicting the binding mode to hiv-1 rt of efavirenz by docking and mm-pbsa. *J. Am. Chem. Soc.* 123: 5221–5230.

Wang J., Wang W., Kollman P.A. & Case D.A. 2006. Automatic atom type and bond type perception in molecular mechanical calculations. *J. Mol. Graph. Model.* 25: 247–260.

Wang J., Wolf R.M., Caldwell J.W., Kollman P.A. & Case D.A. 2004. Development and testing of a general amber force field. *J. Comput. Chem*. 25: 1157–1174.

Welch W., Ruppert J. & Jain A.N. 1996. Hammerhead: fast, fully automated docking of flexible ligands to protein binding sites. *Chem. Biol.* 3: 449–462.

Wolber G. & Langer T. 2005. Ligandscout: 3-d pharmacophores derived from protein-bound ligands and their use as virtual screening filters. *J. Chem. Inf. Model*. 45: 160–169.

Yang B., Hamza A., Chen G., Wang Y. & Zhan C. 2010. Computational determination of binding structures and free energies of phosphodiesterase-2 with benzo[1,4]diazepin-2-one derivatives. *J. Phys. Chem. B*. 114: 16020–16028.

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

Zsoldos Z., Reid D., Simon A., Sadjad S.B. & Johnson A.P. 2007. Ehits: a new fast, exhaustive flexible ligand docking system. *J. Mol. Graph. Model.* 26: 198–212.