DISSERTATION / DOCTORAL THESIS

Titel der Dissertation /Title of the Doctoral Thesis

„ Ligand-based and Structure-based studies to understand the molecular basis of inhibition of ABC transporter expressed in the liver "

verfasst von / submitted by

Sankalp Jain

angestrebter akademischer Grad / in partial fulfilment of the requirements for the degree of

Doktor der Naturwissenschaften (Dr. rer. nat.)

Wien, 2018 / Vienna 2018

Studienkennzahl lt. Studienblatt /
degree programme code as it appears on the
student record sheet: A 796 610 449

Dissertationsgebiet lt. Studienblatt /   Pharmazie
field of study as it appears on the student
record sheet:

Betreut von / Supervisor: Univ.-Prof. Mag. Dr. Gerhard F. Ecker

Acknowledgements

discussions we had. Riccardo, you are simply great. Be it any technical glitch, you will fix it. Thank you for company and Italian pasta. Jenny, Florentina and Alzbeta thank you for the nice talks and wish you all the very best for your PhD. I am sure you will do great. Thank you Andi for the fun talks and the beer evenings. Thanks Amir for being the funny and warm person with an extensive chemistry background. Floriane and Daria, you are a great example of dynamic women in science. It was great to have you around. Thanks for your constant tips and motivation. Thank you Kevin for taking care of the administrative issues. I really enjoyed your positive attitude and fun talks we had. Anika and Kathi, thank you for doing what most of us (including me) hate the most: the administrative work, managing paperwork, contracts and keeping up with ever-changing bureaucratic rules. Having you around was really an asset. I would like to thank other colleagues of the group- Jana, Bernhard, Benjamin, Anna, Katrin, Eleni, Priska, Christine, Jovana, Ines, Jakob, Gulia, Vani, Claire and Ece for their cheerful company and all the good times. Thank you all for introducing me to Austrian culture, food, festivities, sports, in short for all the great time I had and for always being there for me. I wish everyone the best of luck in their career.

I was very fortunate to have some amazing friends who provided a great deal of ethical support even from thousands of kilometres away. Vishal, thank you for your crucial advice, support and motivation. You and Menorca have stood by me always and forever when needed. Thank you, Menorca for being my fantastic travel buddy who planned each of our travels (mini vacations) to get my motivation back for work. Karan, you have been always there for me, listening to all my stories whether it be work or be it anything. Thank for being there for me. I am so glad to have met you all. You all are irreplaceable parts of my life. Thank you all.

I am very grateful for my family's and parents' support, who always have my back and were so excited and proud of every small achievement along the way. Thank you "ma" and "papa" for always and forever supporting me. Special thanks to my grandmother for her constant well wishes. Bhai, you have been my strength always and the strongest pillar in my life. This would not have been possible without you. Bhabhi, I feel so lucky to have you in our family. Thank you for your support and motivation.

III

Preface

The work presented in this dissertation was performed between February 2015 and March 2018 at the Pharmacoinformatics Research Group of the University of Vienna, under the supervision of Prof. Gerhard F. Ecker.

Part I describes the motivation behind the work, provides the biological background of ABC-transporters and introduces the structure-based methods used in this thesis. It includes two book chapters containing major contributions of the thesis author. While a certain information overlap is inevitable, the individual scopes of the parent volumes are different, which has been stated in the introductory part of each section. Finally, the contributions of this thesis are listed.

Part II, Chapter 3 focuses on ligand-based approaches to address the prominent problem of imbalanced datasets in the field of drug discovery. It reports the results obtained after evaluating the performance of seven distinct meta-classifiers in predicting transporter-related hepatotoxicity endpoints.

Part II, Chapter 4 reports the structure-based work undertaken during this period. The results presented in the BCRP study are so far unpublished, but a synthesis-oriented manuscript is in preparation in collaboration with Dr. Vittorio Pace (University of Vienna).

Finally, part III contains the concluding discussion of the Thesis. The major contributions of each chapter are discussed as well as the main outcomes and take-home-messages of these studies.

The compounds used in the BCRP study (Chapter 4.2) were synthesised by Dr. Vittorio Pace (University of Vienna). *In vitro* assays for BCRP inhibition for those compounds were performed by Anna Cseke and Dr. Katrin Wlcek at the University of Vienna under the supervision of Prof. Gerhard F. Ecker.

Index

I. Background

1. Introduction

1.1 Motivation and aim of the thesis

The ATP-binding cassette transporters (ABC transporters) are a superfamily of active transmembrane proteins that selectively aid the movement of molecules in the cell by binding to them and undergoing a conformational change [1]. These transporters participate in active transport, i.e. they hydrolyze ATP and use the energy to transport their substrates. Some of these transporters transfer a large number of structurally and functionally diverse cytotoxic compounds including toxins of natural origin. The overexpression of such transporters has been implicated in multidrug resistance (MDR), a phenomenon in which a cell (cancerous or bacterial) becomes resistant to multiple drugs [2](Figure 1). Thus, besides protecting the cells and tissues against toxic agents, an increase in the efflux activity leads to resistance of tumor cells to a variety of drugs commonly used in chemotherapy [3–5]. Two primary members of the ABC family involved in cancer multidrug resistance are P-glycoprotein (P-gp, gene ABCB1) and the breast cancer resistance protein (BCRP, gene ABCG2).
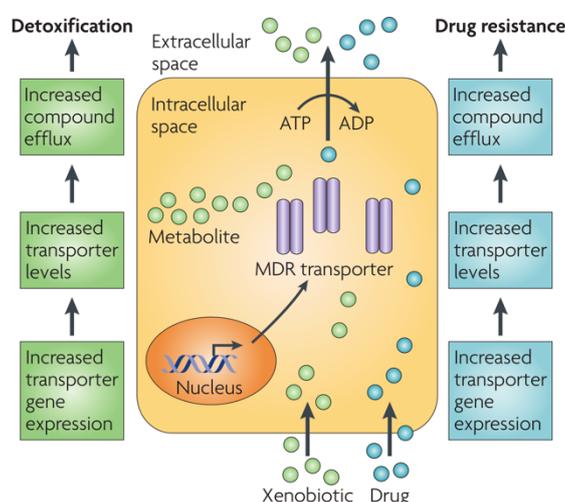
Figure 1: MDR as a result of the overexpression and/or increased efflux activity of ABC transporters [3–5]. Reprinted and edited by permission from Springer Nature: Nature Reviews Cancer, Fletcher et al.[6], copyright 2010.

Failure of several anticancer drug therapies has marked the MDR-related ABC-transporters as one of the widely studied transporters [7–11]. With an aim to overcome MDR, inhibitors of these transporters have been extensively studied [12, 13]. Due to toxicity concerns, none has reached the market yet [8, 14–16]. After several years of research, it can be understood that inhibiting ABC-transporters may not be the best solution to overcome MDR [9, 17, 18]. However, this had little impact on the increasing interest in studying these transporters.

Most ABC-transporters are expressed under normal physiological conditions in important tissues and membranes such as intestine, liver, kidney, placenta, testis and the capillary endothelial cells of the brain [19, 20]. They influence the absorption, distribution, metabolism, excretion and toxicity (ADMET) of pharmacological agents [21, 22]. Genetic variations in their related genes are known to cause a large number of disorders in humans, such as cystic fibrosis, cholesterol and bile transport defects, neurological disease, to name a few [23]. ABC transporters expressed in liver canaliculi in particular (P-gp, BSEP, BCRP, MRP2 and MRP4), are responsible for efflux of many drugs and other xenobiotics [24, 25]. Dysfunction of any of these transporters or their inhibition by small molecules is known to lead to drug-drug interactions and drug-induced liver injuries [26–30]. In this context, regulatory authorities and organizations such as the United States Food and Drugs Administration (US FDA) and the International Transporter Consortium (ITC) recommends screening of candidate drugs for inhibition of P-gp, BCRP and BSEP [31–33]. Therefore, understanding the molecular basis of inhibition of these relevant ABC-transporters by small molecules is highly essential to be able to develop comprehensive *in silico* models that can predict these interactions.

Lack of substantial structural information at higher resolutions, limits the structure-based drug design studies for predicting inhibitors of the ABC transporters [34–36]. Thus far, *in silico* studies to predict inhibitors primarily focused on ligand-based approaches such as quantitative structure-activity relationship (QSAR) modeling and machine learning [37]. While these models have proved to be efficient, they do not consider the properties of the protein and thus a lot of information necessary for understanding the inhibition process is ignored. Another problem associated with ligand-based studies is the increasing amount of data generated in drug discovery. In this context, problems data imbalance is being frequently reported in the literature [38–40] The field of toxicity is no exception and considering the

number of liver transporters implicated in serious adverse events, it is essential to deal with this issue and provide recommendations to handle such datasets.

The general purpose of this thesis is to provide the community with useful *in silico* models to evaluate the probability of a new compound to be a canalicular liver ABC-transporter inhibitor by employing structure-based modeling approaches. We hope to gain a better understanding of the mechanism of inhibition itself and also evaluate data transferability across species in development of predictive *in vivo* and *in vitro* models. Furthermore, a comprehensive comparison of different machine learning methods is expected to resolve the limitations associated with data imbalance and provide guidelines for handling highly imbalanced datasets.

In the light of this, we performed structure-based modeling of three liver canalicular transporters BSEP, BCRP and P-glycoprotein. The release of an experimentally determined crystal structure of BCRP facilitated us to propose a binding hypothesis that could explain the activity trends within an inhibitor class. Further, a comprehensive comparison of the binding sites of human, rat and mouse P-gp transporters helped us to evaluate the transferability of *in vitro* human P-gp data for development of models to predict *in vitro* and *in vivo* outcomes in rat and mouse. We also addressed the issue of learning on imbalanced datasets by evaluating seven distinct meta-classifiers on different datasets in the toxicity domain that are known to possess a varying degree of class imbalance.

We believe that the outcomes of this work would improve the understanding of the transporter mechanism at the molecular level and help us filter out unwanted compounds or prioritize interesting candidates in the early stages of drug discovery in an effort to save time and money.

1.2 Biological background of liver ABC transporters

ABC transporters can be further classified into exporters and importers. Depending on their architecture and mechanism, the importers can be further grouped into two classes (I and II) [41–44]. Humans possess a total of 49 ABC-transporters, which can be divided into seven subfamilies [45], ABCA to ABCG. These groups include transmembrane drug transporters, ion transporters, peptide transporters and others. The ubiquitous ABC transporters are characterized by two nucleotide-binding domains (NBD) and two transmembrane domains (TMD). Conformational changes in TMD, driven by the ATP hydrolysis on the NBD, result in an alternating access from inside and outside of the cell, facilitating a unidirectional transport across the lipid bilayer [44] (Figure 2). Few ABC-transporters are referred to as "half-transporters". Their genes encode only for one transmembrane and one nucleotide binding domain, which necessitates the dimerization of these transporters in order to be functional. The first structural insights on the tertiary structure of ABC transporters were based on the nucleotide-binding domain (NBD) of histidine permease, determined at atomic resolution [46]. By 2009, eight crystal structures of complete ABC transport proteins were solved by X-ray crystallography [47]. Figure 3 represents the ribbon representations of different ABC proteins and their localization within the bilayer membrane. Since then, several other full-length structures of ABC export proteins were solved at the atomic level, providing detailed insights about their conformational variability [48].
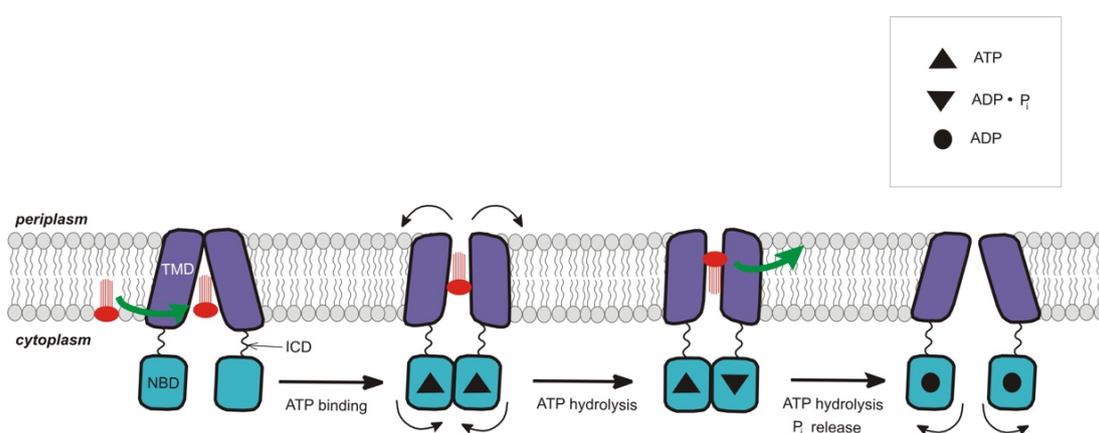
Figure 2: Schematic representation of the Transport cycle for ABC exporters. Reprinted and edited by permission from The American Association for the Advancement of Science: Science, Dong et al. [49], copyright 2005.
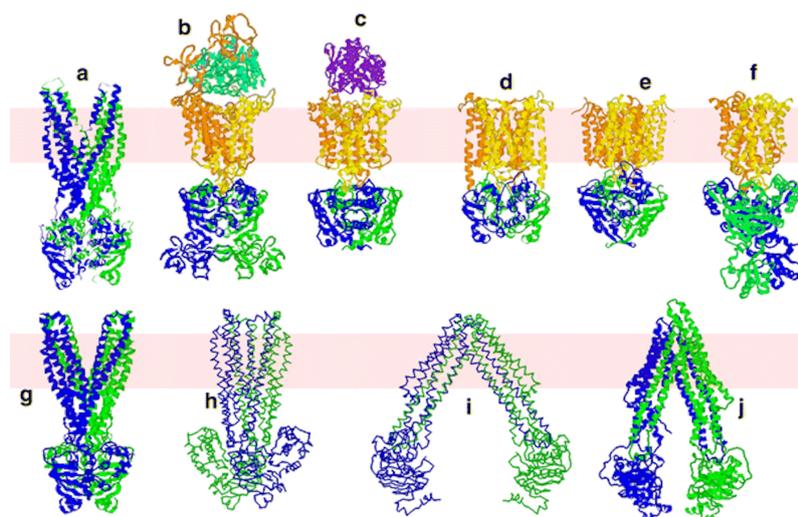
Figure 3: Ribbon representations of different ABC proteins. **(a)** Sav1866 (Dawson and Locher *et al.* [50]). **(b)** MalFGK2 in complex with MBP (Oldham *et al.*[51]). **(c)** ModBC in complex with ModA (Hollenstein *et al.*[52]). **(d)** BtuCD (Locher *et al.*[53]). **(e)** Putative metal chelate transporter H10796 (Pinkett *et al.*[54]). **(f)** Methionine transporter MetNI (Kadaba *et al.*[55]). **(g-i)** Lipid flippase MsbA from *Salmonella typhimurium*, *Vibrio cholera*, and *Escherichia coli*, respectively (Ward *et al.*[56]). **(j)** Mouse Pgp (Aller *et al.*[57]). Reprinted by permission from Springer Nature: Cellular and Molecular Life Sciences, Kos and Ford et al. [47], copyright 2009.

Till date, P-glycoprotein (Pgp, MDR1, ABCB1), an efflux transporter, is the most extensively studied ABC protein. Together with breast cancer resistance protein (BCRP, ABCG2) and multidrug resistance-associated protein 1 (MRP1, ABCC1), Pgp is well known for its role in MDR in tumor cells [58]. These transporters share a low sequence similarity when their transmembrane domains (TMDs) are compared, which could explain the differences in their substrate and inhibitor specificities [47].

In the following sections, we detail the structure, function, and small molecule interactions of the three liver ABC-transporters of significant relevance to this thesis work.

1.2.1 P-glycoprotein (P-gp)

P-glycoprotein (gene ABCB1) was the first membrane protein identified to be able to confer multidrug resistance to cancer cells [23]. In 1976, Juliano and Ling linked MDR to the expression of a membrane protein, P- glycoprotein in Chinese hamster ovary cell line [59].

Later, the structure was determined in 2009 by Aller and coworkers [57], which was further improved by Li et al. [60] in 2014. P-gp is a "full transporter," i.e. the ABCB1 gene encodes for two transmembrane domains (TMDs) and two nucleotide binding domains (NBD's) that constitute the transporter. The two TMD and NBD regions of the transporter differ in their amino acid sequence.

In humans, P-gp is expressed in the blood-brain barrier, placenta, testis, hepatocytes, exocrine cells of the pancreas, gastrointestinal tract, kidney, bladder, spleen and lungs among other tissues [61, 62]. In case of cancer, P-gp expression increases in colon, kidney, adrenal gland, pancreas and other tumour cells [63, 64].

P-gp influences the ADMET (absorption, distribution, metabolism, excretion, and toxicity) properties of many compounds. If a drug is a substrate of P-gp, it could face the risk of increased metabolism in intestinal cells. Besides, if the co-administered drugs are substrates or inhibitors of P-gp, their pharmacokinetic profiles can be altered by P-gp modulating compounds, due to drug-drug interaction, leading to severe side effects [65–67]. Digoxin, an inhibitor of the cardiac Na+/K+-ATPase used for treating heart failures or arrhythmia, is a classic example of drug-drug interactions in the context of P-gp. Digoxin is a substrate of P-gp and is excreted by the kidneys. Inhibition of this transporter by quinidine or ritonavir has caused decreased clearance of digoxin [68, 69], which could potentially lead to cardiotoxicity. Thus, early identification of P-gp inhibitors is highly important in drug safety considerations.

1.2.2 Bile salt export pump (BSEP)

BSEP (gene ABCB11) is an ABC transporter of the B subfamily and is primarily expressed in the cholesterol-rich canalicular membrane of hepatocytes [70]. It facilitates secretion of bile salts from the liver into the bile canaliculi [70–72]. Bile salts are conjugated bile acids which are negatively charged at physiological pH. Bile acids are products of the catabolism of cholesterol in the liver [73–75]. They conjugate with phospholipids to form micelles, which increase their excretability into bile and thus promote digestion and absorption of dietary fat [76]. Bile salts, through an enterohepatic cycle, are transported from the liver to bile to duodenum and again back into the enterohepatic blood circulation. They are then

picked up by the transporter Na^+-taurocholate cotransporting polypeptide (NTCP) at the basolateral membrane of hepatocytes [25].

Genetic variations in ABCB11 result in different forms of progressive familial intrahepatic cholestasis (PFIC) [77, 78]. PFIC is characterized by an early onset of cholestasis and eventually leads to liver cirrhosis and failure [79–81].

BSEP is inhibited by many drugs and drug metabolites [27, 82, 83]. This is a potential mechanism leading to drug-induced cholestasis. Thus BSEP is a crucial transporter protein that is often studied in the recent research on drug safety. Drugs such as bosentan, rifampicin, troglitazone [84] cause intracellular accumulation of bile salts which is an unwanted effect directly related to the inhibition of BSEP. In few cases, it could result in liver injury and thereby liver transplantation. Dysfunction of individual bile salt transporters such as BSEP, due to genetic mutation, suppression of gene expression, disturbed signaling, or steric inhibition, are other factors leading to cholestatic liver disease. Therefore, it is highly essential to screen for BSEP inhibition in the drug discovery pipeline to limit the post-marketing drug withdrawals associated with drug-induced liver toxicity.

1.2.3 Breast cancer resistance protein (BCRP)

Breast cancer research protein (BCRP) was first identified in 1998 [4, 85]. Thereafter, a large number of BCRP inhibitors and substrates were reported, which not only include therapeutic agents but also physiological substances such as estrone-3-sulfate and uric acid. Taylor et al. have recently reported a crystal structure of the transporter, determined by cryo-electron microscopy [86], that provides the first high-resolution insight into a human multidrug transporter. Two cholesterol molecules were observed to be bound in the multidrug binding pocket, which is located in a central, hydrophobic, inward-facing translocation pathway. Today, BCRP (~655 amino acids) is considered among the three major transporters responsible for drug resistance in mammalian cells [87]. It is a half ABC transporter, with one nucleotide-binding domain (NBD) and one membrane-spanning domain (MSD) [88, 89]. Topologically, the N-terminal of BCRP contains the cytoplasmic NBD while the C-terminal contains the TMD, which is a characteristic of the G-subfamily of the ABC transporters. While multimerized forms of BCRP have been reported [90–92], but be functional, it is

8

supposed to be in a homodimer state [93, 94]. Readers can refer to the work by Ni et al. [88] for greater details on the structural and functional aspects of BCRP.

BCRP is highly expressed in the intestinal epithelium, the liver hepatocytes, the renal proximal tubular cells, the endothelial cells of brain microvessels, and the apical membranes of the placental syncytiotrophoblasts [95]. Thus it plays an important role in the absorption, distribution, elimination of drugs and endogenous compounds, as well as tissue protection against xenobiotic exposure. Consequently, the FDA perceived it among the key drug transporters for clinical drug disposition [71, 95]. Although a large number of substrates and inhibitors are already known, the structure-activity relationship (SAR) trend is not clearly known for this elusive transporter [96]. Furthermore, several single nucleotide polymorphisms (SNPs) were already reported [97–100] for this transporter, including a few that may alter pharmacokinetics and lead to drug toxicity. For example, SNP Q141K, frequently found among the Asian population (35%) [101], leads to decrease in membrane expression and ATPase activity [102]. Variation of BCRP function by small molecule inhibitors could also lead to drug-drug interactions. For instance, when the chemotherapeutic agent topotecan (a substrate of BCRP) was administered orally along with elacridar (inhibitor of BCRP with an IC50 below 1 μM) [103, 104], it doubled the bioavailability and tripled the peak plasma concentration of topotecan [105]. Therefore, it is highly essential to prevent such drug-drug interactions that can lead to toxicity.

1.3 Structure-based Methods in Computational Drug Design

Development of faster computers has led to their increasing use in studying biomolecular processes. With a large number of protein structures yet to be resolved and increasing availability of tertiary structure prediction tools and servers, protein structure prediction serves as an appropriate alternative in cases where it is not feasible to determine the structure of interest using experimental techniques [106, 107]. Furthermore, computational modeling of 3D protein structures is among the most common starting points for drug design in both academic and industrial pharmaceutical research.

Below we describe the computational methods which were widely used in this thesis.

1.3.1 Homology modeling

Experimental techniques such as nuclear magnetic resonance (NMR) and X-ray diffraction can resolve the protein's three-dimensional structure [108, 109], but are challenged by several limitations which include the size of the protein, costs involved and difficulty in purification or crystallization [110, 111]. Homology modeling, also known as comparative modeling, refers to construction of an atomic resolution model of a protein from its amino acid sequence and the experimentally determined 3D structure of a template protein with which it shares reasonable sequence identity [106, 112–116]. It is based on the assumption that proteins sharing similar sequences form similar structure. Due to the challenges associated with other methods, homology modeling has become one of the most used and reliable method. Figure 4 lists different steps in homology modeling.
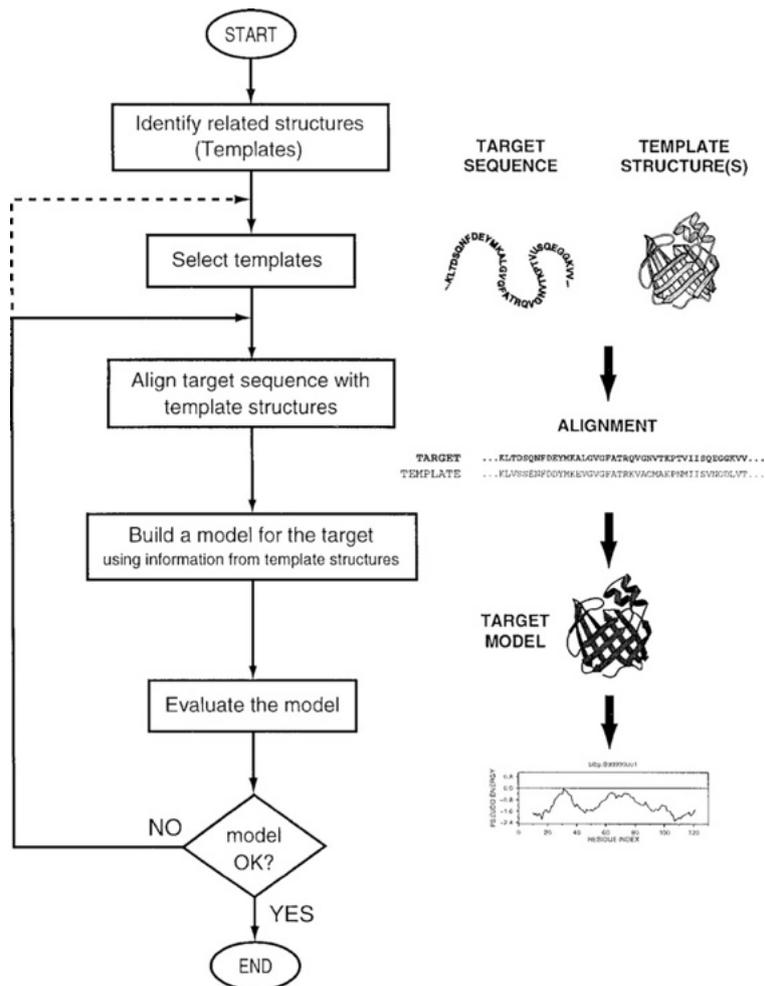
Figure 4: Steps of homology modeling. Reprinted by permission from Annual Reviews, Inc: Annual review of biophysics, Martí-Renom et al.[117], copyright 2008.

The first step is to construct a multiple sequence alignment between the sequences of the target protein and the identified template protein. For this, sequence similarity is performed by using BLAST searches [118–120] against sequences of known structures and other sequences of proteins from the same family as the query. This helps to get an estimate of consensus sequence motifs, the degree of conservation and general features of the family. Then the obtained alignment is corrected for positions of insertions and deletions, accurate alignment of active site residues and also conserved residues. Next step is to construct backbone and model loops, generate side chains and optimize conformations using the software like MODELER [121, 122]. This tool requires three input files: an alignment of the template and the target sequence, the template PDB structure and a script containing the commands and file paths.

In case of complex models, use of multiple softwares/servers and a comparison of the results is recommended. Finally, the generated structure model is optimized and validated. For this, the final model is selected from a pool of generated models, based on energy functions, Ramachandran plots and agreement with mutational data. The Discrete Optimized Protein Energy (DOPE) [123] function assesses the model quality using a statistical potential function. Third-party tools such as PROCHECK [124] also facilitate calculation of Ramachandran plots [125] while the QMEAN server can be used to obtain a Z-score [126, 127], which indicates the deviation from experimental structures. At times, errors in backbone could lead to an incorrect prediction of rotamers. These can be corrected by performing refinements and by applying energy minimization using different types of force fields. The models could still consist errors and require further validation if the bond angles, lengths, torsion angles, etc. are within the desired ranges.

Homology modeling has been successful in drug design, providing more insights into the architecture and function of the protein[128]. In our case, based on sequence identity and resolution, the corrected mouse P-gp (4M1M) structure was selected as the template for homology modeling of BSEP and P-gp.

1.3.2 Molecular docking

Molecular docking is a prominent computational technique in structural biology and computer-aided drug design, useful in predicting the binding modes of a ligand within the binding pockets in the three-dimensional structure of a protein. It is widely employed for tasks such as virtual screening, generation of hypotheses for target inhibition by a ligand and lead optimization. The technique was first reported in the 1980s [129]. Over the time, the algorithms have evolved and many standalone docking tools such as AutoDock [130], GLIDE [131, 132], GOLD [133, 134] and several online services (ZDOCK [135, 136], SwissDock [137]) are now available.

AutoDock is the most commonly used open-source docking software that is freely available for academic research [130]. It supports flexible side chains, checks the syntactic correctness of the input, verifies invariance of the covalent bond lengths and avoids imposing artificial restrictions. Another popular package Schrodinger [138, 139], though not freely available, is

a comprehensive software suite with packages for lead discovery, lead optimization, target preparation, docking and various modeling tools with options for automation. Schrodinger's Glide [131, 132] package enables docking of flexible ligands by grid construction in rigid protein models and rapid sampling of the conformational, orientational, and positional degrees of freedom of the ligand. Another method available in Schrodinger is induced-fit docking (IFD) [140, 141] wherein protein flexibility can also be accounted. It employs Prime package along with Glide to explore all possible binding modes and possible conformational changes in the receptor's active sites. In IFD, the ligand is docked using Glide which generates different ligand poses, followed by the structure prediction using the Prime module to accommodate the ligand by reorienting nearby side chains. Then the residues and ligand are minimized and all the ligands are re-docked into their corresponding low energy protein structures. Prime's advance refinement process further enhances the accuracy of Glide.

Genetic Optimization of Ligand Docking (GOLD) [133, 134] is an another widely used commercial docking software. GOLD facilitates users to define the protein binding pocket with a radius along the given coordinates or by a reference ligand from a co-crystal structure. On the protein and the ligand surfaces, hydrophobic and hydrogen bond (HB) fitting points are then created. Protein flexibility can be accounted by a brute force exploration of all possible angles or as defined in a rotamer database [142]. Additionally, the location and orientation of water molecules can also be predicted. Furthermore, a harmonic potential can be used to fix the distance between two atoms (within the protein and/or between protein and ligand atoms). This is especially useful to reduce the number of docking poses in concurrence with the experimental data (when available).

GOLD uses a genetic algorithm to generate a docking pose. For this,
1) Bit strings (chromosome) are generated from the ligand torsional angles.
2) A scoring function is then applied and two random poses are selected and weighted by their score.
3) Of those two chromosomes, genetic operations of mutation, crossover and migration are applied.
Steps 1-3 are repeated, until the desired number of poses are obtained.

The strength of molecular docking is its capability to provide insights into different binding possibilities, that can be used for screening of large compound libraries and also as a tool for

defining a starting complex for molecular dynamics (MD) simulations. Depending on the focus of the study, docking poses can be generated by either using constraints into a specific binding pocket or without any constraints, which provide a probability distribution of the binding mode. The docking poses can be further clustered on the basis of their placement into specific binding pockets. Further, it is also recommended to energy minimize the ligand and residues within a certain radius for re-scoring purposes. An elaborated overview of different software available for docking is provided by Pagadala et al. [143].

1.3.2.1 Scoring functions

Scoring functions are mathematical methods used to predict the binding affinity between the ligand and active site of the protein structure after they are docked. They can be divided into three classes:

Force-field based scoring functions - These scoring functions are estimated on the basis of intermolecular van der Waals and electrostatic interactions between all atoms of the protein and the ligand using a force field used in molecular dynamics (MD) simulations [144, 145].

Knowledge-based scoring functions - Also known as statistical potential functions. These are based on the probability of finding protein and ligand atoms within a certain distance estimated by observing intermolecular close contacts in 3D databases like Protein Data Bank (PDB) [146] and Cambridge Structural Database (CSD) [147]. This method is based on the assumption that close intermolecular interactions between certain functional groups which occur more frequently in comparison to others are likely to be energetically favourable [148–151].

Empirical scoring functions - It is the most commonly employed category of scoring functions, the reason being that these are faster than force-field based scoring functions and more reliable than the knowledge-based scoring functions. They are based on different types of interaction between the ligand and the receptor protein [152]. They consist of energy terms with coefficients determined by multiple linear regression (MLR) trained on experimental protein-ligand complexes.

Scoring functions can also be applied after docking, i.e. re-scoring. This would also allow obtaining a consensus scoring from multiple scoring functions, which could be used to prioritize a binding hypothesis [153]. The GOLD software provides several scoring functions, such as GoldScore [133, 154], ChemScore [152, 154], ChemPLP [155] and the Astex Statistical Potential (ASP) [156]. GoldScore is a force-field based scoring function. This performs well but is relatively slow compared to ChemPLP. ChemPLP is a piecewise linear potential function that uses the Ants algorithm[155]. ChemScore is mostly used in the case of metal complexes since it contains terms for that [152]. Glide [131, 132] scoring function from Schrodinger [138, 139] is reported to outperformed GOLD against the same target in a similar virtual screening experiment [143].

The external scoring function X-Score[157], which is validated against a set of 800 protein-ligand complexes is also known to perform well in most cases. It uses the van der Waals and electrostatic terms and also approximates the ligand entropic contribution by taking its number of rotatable bonds into account[158].

Though the scoring functions are very useful in screening large compound libraries, they are limited by low predictability. This is due to the fact that these scoring functions estimate score based on a single snapshot of the protein-ligand complex whereas binding affinity is related to a Boltzmann weighted average of different states of a complex. Furthermore, many scoring functions do not account for desolvation, the entropy of the binding pocket and interactions with water. Further, studies reported that machine-learning methods outperform Multiple Linear Regression (MLR) trained scoring functions[159]. Nonetheless, the choice of the scoring function strongly depends on the research question and a combination of several scoring functions, referred to as consensus scoring, is recommended [160].

1.3.3 Hierarchical Clustering

Hierarchical clustering is a useful method to cluster the poses of different ligands with same scaffold [161]. It is believed that ligands that share a common chemical scaffold fit in the same fashion in a protein binding pocket [161–163]. The docking poses can be clustered based on the root mean square distance (RMSD) matrix of the heavy atoms in the Euclidean space. The large number of poses in a typical range of 2 Å indicates that the binding modes

have higher probabilities of being active. For this, the scaffolds of ligands are first extracted from docking poses and saved as SMILES strings with coordinates in the database. Next, different clustering methods such as complete-linkage method can be used to cluster the poses with least distance together and construct a dendrogram. Further, a cut-off on the RMSD can be used to remove the outliers from the cluster. Visual inspection of the dendrogram facilitates the selection of a cut-off value on the RMSD.

In our study, clustering of arylmethyloxyphenyl derivatives docked into the BCRP binding pocket helped us to propose a binding hypothesis for the series of analogues.

1.3.4 Molecular Dynamics Simulations

Molecular dynamics (MD) is a computer-aided simulation method to study the dynamic movement of atoms and molecules[164]. It is an important tool in drug discovery [165], which facilitates simulation of both individual membrane proteins and more complex systems[166]. MD simulations provide a detailed description of particles in motion as a function of time by iteratively solving Newton's classical equation of motion for each molecule [167]. MD simulations are particularly useful when the system cannot be studied by the experimental methods such as mass spectroscopy or crystallization methods such as NMR or X-ray crystallography [168]. Therefore, they hold great significance in understanding the physical basis of the structure and function of proteins and other biological macromolecules [113].

In an MD simulation, interactions between the atoms can be defined by different potential energy functions of a given force field (OPLS [169], CHARMM [170, 171] or GROMOS [172]). The bonded interactions within the system such as stretching, bending and dihedral terms are modeled by employing harmonic potentials while the non-bonded interactions are described by the Lennard-Jones potential for van der Waals interactions and by the Coulomb's law for electrostatic interactions. Calculation of the non-bonded terms is computationally expensive, which necessitates the employment of algorithms such as SHAKE [173] or LINCS [174] that correct for the interatomic distance in every step.

Protein systems are typically simulated in a box using periodic boundary conditions in order to emulate crystal structure conditions and to prevent undesirable boundary effects [175–177]. Membrane proteins should be placed in a lipid bilayer which approximates their indigenous biological conditions. Further, the system is energy minimized to get rid of any overlapping van der Waals cores. The book "Molecular Modeling of Proteins" [178] provides an excellent review on various aspects of these issues.

A general protocol to setup an MD simulation, as described by Jurik et. al.[113], can be found below:

1) select the forcefield taking into account the parameters for the protein and the ligand
2) place the protein into the membrane
3) solvate the system and add ions to neutralize excess charges and adjust the final ion concentration
4) perform energy minimization on the system
5) run MD for ~5-10ns with restraints on all protein heavy atoms
6) equilibrate without restraints
7) run production MD
8) perform analysis

The main advantage of using MD is that it strives to mimic the structure of interest and can be effective in comprehending the structure-to-function relationships of macromolecular structures [164]. Multiple conformations are generated, that could describe protein-ligand interactions in the dimension of time. Furthermore, it is possible to achieve precise interaction energy values that facilitate the interpretation of ligand binding and unbinding events. A classical MD simulation, unlike Monte Carlo or Markov Chain methods, does not efficiently sample the conformational space. Based on the starting structure, the trajectories could be confined within a multidimensional energy minimum. This limitation can be handled by increasing the simulation time or by employing enhanced sampling techniques such as steered molecular dynamics or essential dynamics[164]. Further, recent advancement in computational hardware, especially the use of graphical processing units (GPUs) and high performance computing (HPC) clusters facilitate simulation of much larger systems in shorter times, allowing greater conformational changes to be sampled [179].

1.4 Contribution of this thesis

We hope that this thesis work sheds light on the thus far unexplored protein-ligand interactions to reveal the molecular basis of inhibition of ABC-transporters. This would facilitate the development of *in silico* prediction models and assist lead optimization. Due to the constraints on data availability and duration of the Ph.D. thesis, we focused on the ABC-transporters BSEP, BCRP and P-gp. Since multiple ligand-based studies have already reported prediction models for inhibition of these transporters, we primarily focused on structure-based approaches.

More precisely, the key contributions of this thesis are:

- A benchmarking study to evaluate the performance of seven different meta-classifiers in handling imbalanced drug discovery datasets: Bagging, under-sampled stratified bagging, cost-sensitive classifier, MetaCost, threshold selection, SMOTE and ClassBalancer.

- Comparative structural modeling of human BSEP and structure-based classification of BSEP/ABCB11 inhibitors.

- Protein-ligand interaction fingerprint (PLIF) based method and analysis for identification of functional group-binding site residue interactions that reveal the molecular basis of inhibition of the transporter protein by a wide range of ligands.

- A hypothesis for the molecular basis of the inhibition of BCRP by arylmethyloxyphenyl analogues using the BCRP crystal structure.

- Structure-based approaches to compare the binding site interaction profiles of human, rat and mouse P-gp to evaluate the transferability of *in vitro* human activity data in the development of *in vivo* prediction models for rat and mouse.

2. Status quo in field

2.1 Transporter in Hepatotoxicity

Eleni Kotsampasakou, **Sankalp Jain**, Daniela Digles Gerhard F. Ecker**, Transporter in Hepatotoxicity**, Computational Toxicology: Risk Assessment for Pharmaceutical and Environmental Chemicals, 2nd edition, Sean Ekins, ISBN: 978-1-119-28256-3

In the following chapter, we summarize the role of different transporters in hepatotoxicity. We then briefly describe the data sources available and the difficulties in obtaining the data for the related transporters proteins. Further, we present different ligand-based and structure-based studies performed to predict *in silico* whether a small molecule is an inhibitor or a substrate of a given transporter. We also mention the *in vitro* models available to predict liver toxicity. The majority of these models focus on predicting BSEP inhibition, which is directly related to liver toxicity.

E. Kotsampasakou wrote the introduction, basolateral transporter and canalicular transporter, D. Digles wrote Data Sources for Transporters in Hepatotoxicity, **S. Jain** performed the literature search and wrote In Silico Transporters Models, ligand-based approaches and structure-based approaches, G.F. Ecker supervised the work and revised the chapter.

6

Transporters in Hepatotoxicity

Eleni Kotsampasakou, Sankalp Jain, Daniela Digles, and Gerhard F. Ecker

Department of Pharmaceutical Chemistry, University of Vienna, Wien, Austria

---

CHAPTER MENU

6.1   Introduction

Transmembrane transporters are essential for regulation of the uptake and efflux of endobiotics and xenobiotics at the cellular level as well as in barrier tissues (e.g., blood–brain barrier, kidney, liver, enterocytes). Among them,
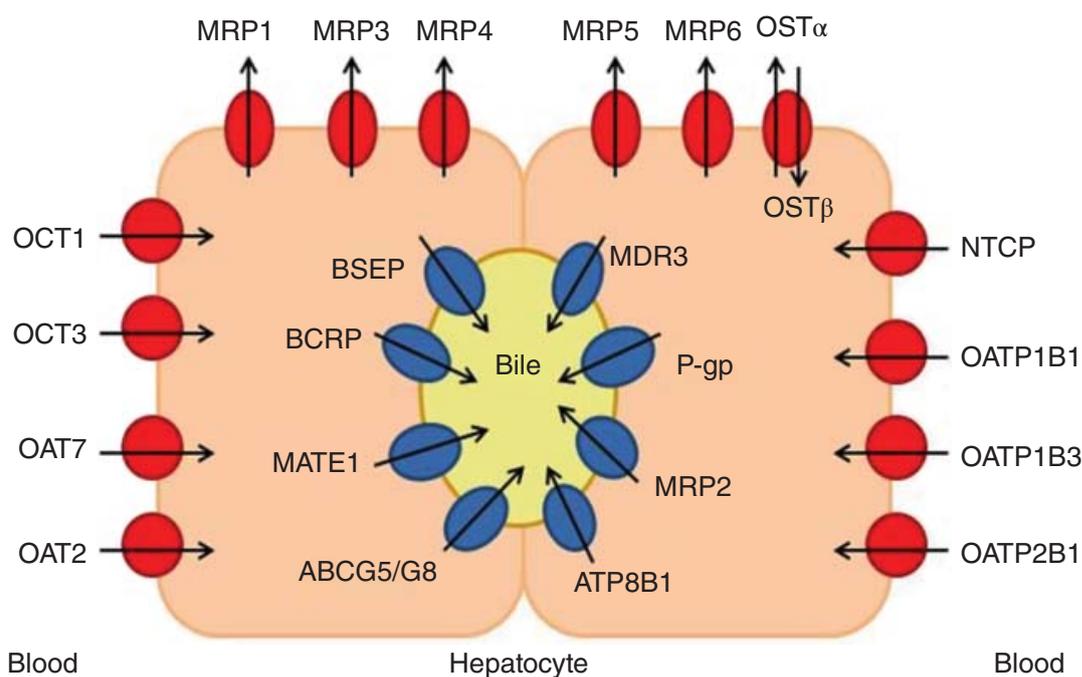
Figure 6.1 Transporters located in the hepatocyte. The medium grey symbols represent the canalicular transporters and dark grey ones the basolateral transporters. Cycles represent uptake transporters and ellipses refer to efflux transporters. The arrows define the direction of transport.

hepatic transporters possess a vital role, as the liver is the main organ of metabolism and detoxification [1, 2]. Figure 6.1 depicts the main hepatic transporters and their respective location in the hepatocyte. In the following section, we will briefly introduce their significance in selected liver toxicity manifestations.

6.2 Basolateral Transporters

Regarding the basolateral uptake transporters, the sodium (Na^+) taurocholate co-transporting polypeptide (NTCP) is quite important in the enterohepatic circulation of bile salts, thus contributing to liver homeostasis [3, 4]. It has been proposed that the mechanistic basis of some hepatotoxic – and, in particular, cholestatic - drugs includes the inhibition of NTCP [5]. In addition, the potential association of organic anion transporting polypeptides 1B1 and 1B3 (OATP1B1 and OATP1B3) inhibition with hyperbilirubinemia, a pathological accumulation of conjugated or unconjugated bilirubin in sinusoidal blood [6, 7], is worth mentioning. Hyperbilirubinemia can be drug-induced [6, 7] or genetically induced, such as in the case of the Rotor syndrome [7–13]. Figure 6.2 shows the cycle of bilirubin and how transporters might be involved in the development of this condition.
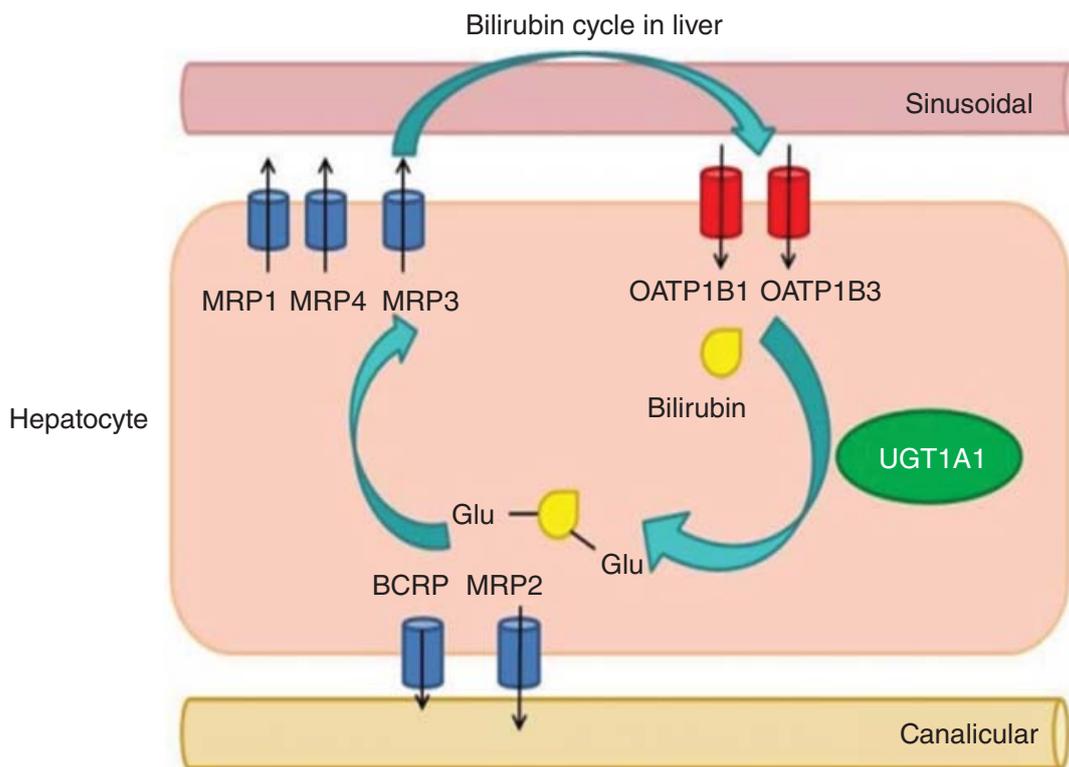
Figure 6.2 The cycle of bilirubin in the liver. Bilirubin is taken up from sinusoidal blood by OATP1B1 and OATP1B3. It is metabolized by UGT1A1 into mono- and bi-glucuronidated products that are exported into bile primarily by MRP2 and in smaller extent (smaller arrow) by BCRP. A portion of the glucuronidated or unglucuronidated bilirubin is effluxed into sinusoidal blood by MRP4 and the cycle is repeated. Source: Adapted from Sticova and Jirsa 2013 [11].

For the other major basolateral uptake transporters, such as the organic anion transporters (OATs) and the organic cation transporters (OCTs), there is low incidence for a potential role in toxicity phenotypes in the liver. However, there is one exception, namely, some polymorphisms and mutations in human OCT1 that lead to decreased transport activity of OCT1 in the liver, which can obstruct the biliary excretion of hydrophobic cationic drugs [14].

Regarding the basolateral efflux transporters, the organic solute transporter alpha-beta (OSTα–OSTβ) dimer is upregulated as a protective mechanism against the accumulation of toxic bile salts in the hepatocyte [15]. The same accounts for most of the multidrug resistance-associated proteins (MRPs). Several reviews describe an increase in mRNA levels of MRP1, MRP3, MRP4, and MRP5 [4], as well as an increase in protein levels of MRP3 and MRP4 [16] in hepatobiliary pathological conditions. Moreover, MRP3 as well as MRP1 may act as a compensatory mechanism to alleviate the potential toxic effects of high bile acid concentrations in the liver, when the canalicular efflux transporters such as the bile salt export pump (BSEP) and multidrug resistance-associated protein 2 (MRP2) are blocked [1, 17].

6.3   Canalicular Transporters

For canalicular transporters, the most prominent example is the contribution of both genetically – [3, 18–21] and drug-induced [18, 20, 22–24] BSEP inhibition in the development of cholestatic conditions. MRP2, due to its important role in bilirubin and bile salts transport, is also suggested to be correlated with drug-induced hyperbilirubinemia [11, 25] and cholestasis [26–28]. Similarly, BCRP is also believed possibly contribute to the efflux of bilirubin conjugates into bile [11]. Deficiency of BCRP is also suspected to result in accumulation of toxic bile salts in the liver, which induce toxicity issues [29]. MDR3 maintains the integrity of the membrane and conducts the phospholipid flow across the canalicular membrane of the hepatocyte [30]. It has also been associated with genetically – [1, 16, 26, 29–33] and drug-induced [16, 26, 29, 30, 34, 35] cholestatic conditions.

Furthermore, MDR1 (P-glycoprotein, P-gp) is also expressed in the liver. MDR1 plays a prominent role in drug resistance during cancer therapy [36, 37] and has also been associated with drug-drug interactions. Nevertheless, in most of the cases of drug-induced hepatotoxicity or cholestasis, the implication of P-gp is attributed to its localization in several organ membranes and its great number of its substrates, rather than to direct effects in the liver [38, 39].

The ATP-binding cassette subfamily G members 5 and 8 (ABCG5 and ABCG8) heterodimer, the ATPase class I type 8B member 1, also known as ATPase-aminophospholipid transporter (ATP8B1 or FIC1), the multidrug and toxin extrusion transporter 1 (MATE1), the cystic fibrosis transmembrane conductance regulator (CFTR), the copper-transporting P-type ATP-ase (ATP7B), and the manganese transporter SLC30A10 are also liver transporters with an important physiological role. Despite the fact that they are associated with several diseases – including manifestations of liver toxicity, to our knowledge they are not associated with any pathological drug-induced liver condition.

With this list of transporters and their important role it becomes evident, that any distortion in the proper function of hepatic transporters might result in manifestation of hepatotoxic phenomena. Therefore, knowledge of the inhibitory profile of drugs currently in the market, as well as the ones under development, is vital in order to avoid potential side effects. One step in this direction is the collection of the available data and another step further is the development of robust predictive models for these transporters.

6.4   Data Sources for Transporters in Hepatotoxicity

Currently several large-scale initiatives collect and predict toxicity data for both drugs and environmental chemicals. These include, among others,

projects funded by the innovative medicines initiative (IMI) such as eTOX (http://www.etoxproject.eu/) and MIP-DILI (http://www.mip-dili.eu/), the Horizon 2020 EU-ToxRisk project (www.eu-toxrisk.eu) and the Toxicology in the 21st Century (Tox21) initiative [40] (http://tox21.org). EU-ToxRisk aims at advancing *in vitro* and *in silico* tools for toxicology, thereby focusing on mechanism-based approaches. Adverse outcome pathways (AOPs) introduced by the Organisation for Economic Co-operation and Development (OECD) play an important role here. One example for an AOP relevant to hepatotoxicity is "cholestatic liver injury induced by inhibition of the BSEP (ABCB11)" [41].

Searching for data on hepatotoxicity in bioactivity databases, such as ChEMBL [42, 43] or PubChem [44], is difficult owing to the way biological data are organized. While searches for bioactivity data for protein targets are straightforward, hepatotoxicity as a "target" is more difficult to define. For example, an assay search in ChEMBL version 22 [43] (accessed October 5, 2016) for "hepatotoxicity" returns 585 different assays mentioning hepatotoxicity in the assay description. Here, the target is for example the tissue Liver, the cell-line hepatocyte, or the general target ADMET. However, the phenotype "hepatotoxicity" is available as target directly (CHEMBL1697861) and is connected with 31 assays. These include, among others, datasets mined from literature [45, 46], the drug induced liver injury prediction system (DILIps) training set [47], and the food and drug administration (FDA) liver toxicity knowledge base benchmark dataset (LTKB-BD) [48]. Of note for hepatotoxicity, but not yet available in ChEMBL, is a recent work by Chen *et al.* [49], where a reference list for drug-induced liver injury (DILI) was presented.

While identifying activity values for a specific transporter is more straightforward, interpreting the data can be challenging. As an example, a search for BCRP easily identifies the human protein (CHEMBL6020), which shows a total of 1799 bioactivity values. While a large portion of the values are reported as IC_{50} values in nanomolar (nM) units (615), others are given as inhibition in percentage (357), activity in percentage or fold increase of control (278), or EC_{50} in nM (213). Several activities are reported as ratios (58) or other activity types (275), for example, fluorescence intensity, drug transport, intrinsic activity, or permeability. This makes a direct comparison of the values rather difficult. In addition, measurements of different assay setups cannot always be directly compared, as shown for P-gp inhibitors [50].

To retrieve bioactivity values for transporters (e.g., to build computational models), a list of relevant transporters is needed first. This can be achieved by reviewing the literature, but data collections such as the Gene Ontology (geneontology.org) [51] can be helpful as well. For example, the molecular function of "canalicular bile acid transmembrane transporter activity" (GO:0015126) can be used to retrieve a list of BSEP proteins from different organisms.

6.5 *In Silico* Transporters Models

Table 6.1 summarizes some of the available computational models of hepatic transporters implicated in hepatotoxicity, namely, BSEP, MRP2, MDR1, BCRP, MATE1, OCT1, OCT2, OATP1B1, OATP1B3, MRP3, MRP4, NTCP, ASBT, and OATPs. Owing to the heterogeneity of experimental reports in terms of assay types, test concentrations, and experimental conditions, most computational studies focus on classification models of varying prediction performances. These models are built to distinguish inhibitors from non-inhibitors [79]. Only a few models for prediction of binding affinity or inhibition at a quantitative level are available. Their predictivity is usually limited to small sets of compounds with measurements from assays with similar experimental conditions [79].

6.6 Ligand-Based Approaches

Considerable progress has been made in the development of *in silico* prediction models for canalicular transporters such as BSEP, MRP2, MDR1, and BCRP. In addition, there were also recent advances for *in silico* models for basolateral transporters.

6.7 OATP1B1 and OATP1B3

Karlgren *et al.* proposed a computational model for OATP1B1 [52] based on 146 compounds (2/3 training set; 1/3 test set) using orthogonal partial least-squares discriminant analysis (OPLS-DA). The model used a set of molecular descriptors and achieved a performance of 80% sensitivity and 91% specificity for a test set. Subsequently, they reported classification models for OATP1B1, OATP1B3, and OATP2B1 inhibitors at a 20 µM potency threshold, with accuracies between 75% and 93% [53]. Following a proteochemometric modeling approach, De Bruyn *et al.* [80] combined protein-based and ligand-based molecular descriptors using random forest (RF) as classifier. They used 2,000 compounds for training and 54 compounds as an external test set. An additional OATP1B1 classification model was published by van de Steeg *et al.* [81] Their Bayesian model was based on a training set of 437 compounds (37 inhibitors and 400 non-inhibitors) and an internal set of 155 compounds for validation (12 inhibitors and 143 non-inhibitors), resulting from the screening of a commercial library of 640 FDA-approved drugs. The overall model performance was greater than 80%, both for leave-one-out cross-validation and external validation. Kotsampasakou *et al.* [54] developed

Table 6.1 Summary of the best-performing models for transporters.

Transporter	Model summary (best model)	Performance (TS, EV)	Dataset size (training set/TS, EV)	References
OATP1B1	OPLS-DA (Cl. inhib)	Correctly predicted 81% inhibitors and 93% non-inhibitors	98/48	Karlgren *et al.* [52]
	PLS (Cl. inhib)	Acc. = 79%	150/75	Karlgren *et al.* [53]
	RF (Cl. inhib)	Acc. = 85%	1708/201	Kotsampasakou *et al.* [54]
OATP1B3	PLS (Cl. inhib)	Acc. = 92%	150/75	Karlgren *et al.* [53]
	SVM (Cl. inhib)	Acc. = 87%	1725/209	Kotsampasakou *et al.* [54]
OATP2B1	PLS (Cl. inhib)	Acc. = 75%	118/60	Karlgren *et al.* [53]
	RF (Cl. inhib); *k*NN (Cl. inhib)	Acc. = 80% (RF, *k*NN)	109/27	Sedykh *et al.* [55]
	RF (Cl. subst.); *k*NN (Cl. subst.)	Acc. = 75% (RF, *k*NN)	42/11	Sedykh *et al.* [55]
NTCP	Pharmacophore	Acc. = 60%	5/10	Greupink *et al.* [56]
OCT1	OPLS-DA (Cl. inhib)	Acc. = 85%	95/96	Ahlin *et al.* [57]
OCT2	2D-QSAR	$R^2 = 0.81$	28/6	Suhre *et al.* [58]
	3D-QSAR (CoMFA)	$R^2 = 0.97$	28/6	Suhre *et al.* [58]
	Pharmacophore (combinations)	Acc. = 70%	162/299	Xu *et al.* [59]
MRP1	Stepwise multiple regression	$R^2 = 0.77$	29/0	Van Zanden *et al.* [60]
	Pharmacophore	$R^2 = 0.80$	60/20	Tawari *et al.* [61]
	CoMFA	$Q^2 = 0.71$	107/0	Pajeva *et al.* [62]
	CoMSIA	$Q^2 = 0.73$	107/0	Pajeva *et al.* [62]

(Continued)

Table 6.1 (Continued)

Transporter	Model summary (best model)	Performance (TS, EV)	Dataset size (training set/TS, EV)	References
MRP3	SVM (Cl. subst.)	Acc. = 98%	50/12	Sedykh et al. [55]
MRP4	SVM (Cl. subst.)	Acc. = 89%	74/18	Sedykh et al. [55]
	SVM (Cl. inhib)	Acc. = 67%	51/13	Sedykh et al. [55]
BSEP	SVM (Cl. inhib)	Acc. = 87%	437/187	Warner et al. [145]
	Multiple linear regression	R^2 = 0.95	37/0	Saito et al. [63], Pedersen et al. [64]
	OPLS-DA	Acc. = 89%	163/86	Pedersen et al. [64]
	RF (Cl. inhib)	Acc. = 80% (TS), 89% (EV)	670/168 (TS), 156 (EV)	Montanari et al. [65]
MRP2	SA-PLS (binding affinity, Ki)	R^2 = 0.82	20/5	Ng et al. [66]
	OPLS-DA (Cl. inhib)	Acc. = 72%	79/39	Pedersen et al. [64]
	SVM (Cl. inhib)	Acc. = 77%	257/61	Zheng et al. [67]
	RF (Cl. inhib)	Acc. = 75%	964/240	Pinto et al. [68]
	SVM (Cl. subst.)	Acc. = 87%	150/38	Sedykh et al. [55]
	SVM (Cl. inhib)	Acc. = 89%	77/19	Sedykh et al. [55]
MDR1	LDA (Cl. inhib)	Acc. = 85% (TS), 86% (EV)	772/85 (TS), 418 (EV)	Broccatelli et al. [69]
	Naive Bayes (Cl. inhib)	81%	973/300	Chen et al. [70]
	SVM (Cl. inhib)	75%	1201/407	Klepsch et al. [71]

	Algorithm	Performance	Data size	Reference
BCRP	ANN, SVM (Cl. inhib)	Acc. = 87% (TS), 67% (EV)	96/32 (TS), 147 (EV)	Erić et al. [72]
	OPLS-DA (Cl. inhib)	79%	80/43	Matsson et al. [73]
	Pharmacophore (Cl. inhib)	66%	30/79	Pan et al. [74]
	Logistic regression (Cl. inhib)	64% (leave-sources-out CV); 83% (10-fold CV)	978	Montanari et al. [75]
ASBT	Linear regression (binding affinity, Ki)	$R^2 = 0.73$	29/1	González et al. [76]
	Linear regression (binding affinity, Ki)	$R^2 = 0.68$	23/4	Zheng et al. [77]
	Linear regression (binding affinity, Ki)	$R^2 = 0.89$	31/1	Rais et al. [78]
	kNN (Cl. subst.)	Acc. = 94%	80/20	Sedykh et al. [55]
	RF (Cl. inhib)	Acc. = 88%	120/30	Sedykh et al. [55]

The type of transporter and the summary for the best model (algorithm, performance, data size, and original publication) are provided.
Cl. inhib., classification of inhibitors; Cl. subst., classification of substrates; Acc., accuracy; TS, test set; EV, external validation set; RF, random forest; SVM, support vector machine; PLS, partial least squares regression; OPLS-DA, orthogonal partial least-squares projection to latent structures discriminant analysis; CoMFA, comparative molecular field analysis; CoMSIA, comparative molecular similarity index analysis; LDA, linear discriminant analysis; SMO, Kohonen self-organizing maps; BPNN, back-propagation neural network; QSAR, quantitative structure–activity relationship; ANN, artificial neural network; kNN, k-nearest neighbor; SA-PLS, simulated annealing–partial least squares.

a set of classification models for OATP1B1 and OATP1B3 inhibition based on 1,700 curated compounds from the literature. Virtual screening of DrugBank drugs followed by biological testing of 10 top-ranked hits confirmed the validity of the models, yielding in an accuracy of 90% for OATP1B1 and 80% for OATP1B3, respectively.

6.8   NTCP

A study by Greupink *et al.* [56] proposed a ligand-based common feature pharmacophore model consisting of two hydrogen bond acceptors and three hydrophobic features. This model, based on five NTCP substrates, was then applied to screen large chemical libraries. In the virtual screening procedure, 10 compounds were selected out of which 6 notably inhibited taurocholate uptake in NTCP overexpressing cells.

6.9   OCT1

Three pharmacophore models have been reported for OCT1 so far [82–84]. Ahlin *et al.* [57] investigated the inhibition patterns of OCT1 using registered oral drugs to develop predictive computational models. Increased lipophilicity and positive net charge were found to be key physicochemical properties that positively correlated with OCT1 inhibitory activity. Moreover, dipole moment and multiple hydrogen bonds were found to be negatively correlated. The data were used to generate orthogonal partial least-squares projection to latent structures discriminant analysis (OPLS-DA) models for OCT1 inhibitors so as to discriminate the inhibitors from the non-inhibitors. The final model correctly predicted 82% of the inhibitors and 88% of the non-inhibitors from the test set.

6.10   OCT2

A 2D-QSAR model based on 34 OCT2 inhibitors that inhibit tetraethylammonium (TEA) transport was reported by Suhre *et al.* [58]. Another study by Zolk *et al.* [85] analyzed 26 commonly used drugs for inhibition of MPP^+ uptake. A significant correlation was found between the topological polar surface area (TPSA) and activity on MPP^+ uptake inhibition. Kido *et al.* [86] experimentally screened 910 compounds, of which 244 compounds inhibited OCT2-mediated transport of 4-(4-(dimethylamino)styryl)-*N*-methylpyridinium(ASP^+). Using computational analysis, molecular charge was identified as one of the key

properties for differentiating inhibitors from non-inhibitors. The 10 most potent OCT2 inhibitors were used to generate a two-point pharmacophore, showing a pattern of an ion-pair interaction site and a hydrophobic aromatic site separated by 5.0 Å.

Xu *et al.* [59] designed a scheme for screening combinations of pharmacophores based on hypotheses established using 162 OCT2 inhibitors. The final model comprises four individual pharmacophores. The combinatorial model provided an overall accuracy of about 70% on a test set containing 81 OCT2 inhibitors and 218 non-inhibitors.

6.11   MRP1, MRP3, and MRP4

van Zanden *et al.* [60] studied the effect of flavonoids on MRP1 and MRP2 transfected MDCKII cells. A QSAR model for the inhibition of MRP1 was obtained [60]. Pharmacophore-based models are reported for MRP1 inhibition by Chang *et al.* [87], Tawari *et al.* [61], and Pajeva *et al.* [62].

Owing to lack of experimental measurements, very few computational studies exist for the basolateral bile acid efflux transporters MRP3 and MRP4 (Table 6.1). Sedykh *et al.* [55] reported classification models of MRP4 inhibitors at a 10 µM threshold with accuracy of 70% on external dataset. The modeling was based on a rather small set of 64 molecules. In a recent study, Akanuma *et al.* [88] attempted structural analysis of MRP4 transport for several groups of β-lactam antibiotics.

6.12   BSEP

For the human BSEP, Warner *et al.* [20] used a recently described *in vitro* membrane vesicle BSEP inhibition assay to quantify transporter inhibition for a set of 624 compounds. A support vector machine (SVM) learning model, employing in-house descriptor sets comprising 2D, 3D, and fingerprint-like features, led to prediction accuracy of 87%. Relating a set of physicochemical properties of the compounds to BSEP inhibition, they demonstrated that lipophilicity and molecular size are significantly correlated with BSEP inhibition. The model could be further used to minimize the propensity of drug candidates to inhibit BSEP. Saito *et al.* [63] reported a BSEP inhibition model based on multiple linear regression using 37 diverse druglike compounds and their chemical fragment descriptors. However, the model was not validated further to evaluate its applicability. The model proposed by Hirano *et al*. [89], based on as few as 37 compounds, does not allow *in silico* profiling of chemically diverse compound libraries. Later, Pedersen *et al.* [90] built two OPLS-DA models on 163 compounds. They report an accuracy of 89% on a test set of randomly selected 86

compounds. Nevertheless, none of the aforementioned models were applied in prospective studies to mark BSEP inhibitors in real-life settings.

In a more recent study, Montanari *et al.* [65] developed a classification model based on a set of physicochemical descriptors. The model revealed the importance of hydrophobicity, aromaticity, and H-bond donor characteristics in distinguishing inhibitors from non-inhibitors. One major finding of these studies was bromocriptine - a known drug - being identified as BSEP inhibitor. The accuracies of the BSEP models on external datasets ranged from 70% to 90%.

6.13 MRP2

Several publications have proposed prediction models for MRP2 inhibition (Table 6.1) using linear and nonlinear modeling methods. For linear models, mainly partial least squares (PLS) regression and discriminant analysis were used, while nonlinear modeling methods include SVM, *k*-nearest neighbors (*k*NN), and RF [55, 64, 91]. Ng *et al.* developed a QSAR model of binding affinity to rat MRP2 for 25 methotrexate analogs as well as a pharmacophore for their binding mode [66]. Zhang *et al.* [91] have constructed a pharmacophore for MRP2 inhibitors, which performed slightly worse than their SVM-based model. Pinto *et al.* [68] applied different machine learning methods for the development of models for putative substrate/non-substrate classification for MRP2. Although the prediction performance is not excellent, the study can be marked as the first of its kind for classification of a huge set of putative MRP2 substrates and non-substrates.

6.14 MDR1/P-gp

P-gp is a thoroughly studied ABC transporter protein. A number of ligand-based approaches have been proposed already, including conventional methods such as Hansch analysis, linear and nonlinear classification algorithms, pharmacophore modeling, and even more advanced methods such as supervised and unsupervised artificial neural networks [92–97]. One of the groundbreaking contributions is the work of Broccatelli *et al.* [69], who used a combination of molecular field analysis, pharmacophore-based representation of the compounds, as well as physicochemical descriptors to develop both global and local models for P-gp inhibitors. The final model indicated that flexibility, hydrophobic surface area, and log *P* are the discriminating physicochemical properties for inhibitors and non-inhibitors. The model, which was

based on 1275 compounds extracted from 61 studies, also points toward shape, a 3D descriptor/feature, as a crucial discriminative property. With a reported accuracy of 86%, the model demonstrated a sensitivity of 0.9, a specificity of 0.8, and Cohen's kappa of 0.7 when tested on an external set. In addition to binary classifiers, a number of other 2D-QSAR models [98–107] and machine learning methods were successfully applied for prediction of P-gp substrates and inhibitors [108, 109].

Wang *et al.* [109] used unsupervised machine learning methods such as Kohonen self-organizing maps, which were also employed to predict P-gp substrates and inhibitors. The best model, based on a dataset of 206 compounds, correctly predicted 83% of substrates and 81% of inhibitors. Models based on recursive partitioning and Naïve Bayes methods were developed by Chen *et al.* [70] on a dataset containing 1273 compounds. The best model accurately predicted 81% of the compounds in the test dataset. Klepsch *et al.* [71] used BestFirst as a feature selection method using a dataset of 1608 P-gp inhibitors and non-inhibitors. Random forest and SVM models were reported as the best classifiers, accurately predicting a total of 86% and 83% of the training set compounds and 73% and 75% of the test set compounds, respectively.

Different studies, employing a range of simple to complex methods, showed satisfactory prediction performance and have contributed to identification of molecular features that are involved in P-gp mediated MDR reversal. However, the applicability of the models is questionable, taking into account the still relatively small number of molecules investigated in each of these studies [110].

6.15   MDR3

Multidrug resistance protein 3 (MDR3) is the closest homologe to P-gp sharing a sequence identity of 75%. Only five substrates could be identified in previous studies [111]. Regarding inhibitors, a study by He *et al.* [34] led to the discovery of nine drugs that inhibit MDR3, while a more recent study by Mahdi *et al.* showed inhibition of MDR3 by antifungal azoles. In addition, their data indicated a potential increased cholestatic effect in case of simultaneous inhibition of BSEP and MDR3 [35]. However, this information is not sufficient to establish *in silico* prediction models.

6.16   BCRP

Several global machine learning-based classification models have been proposed to predict BCRP inhibition. Erić *et al.* [72] extracted and merged

literature data on BCRP inhibition to build neural network and SVM models based on 96 compounds. The models provided test set accuracies over 82%, sensitivities over 83%, and specificities over 80%. Matsson and colleagues [73] developed models that could distinguish BCRP inhibitors from non-inhibitors using a diverse training set of 80 compounds and the descriptors log D and polarizability. The best model had a sensitivity of 83% and a specificity of 76% on a test set of 43 compounds. Pan *et al.* [74] developed a Bayesian classification model and a set of pharmacophores on 203 compounds. Screening the collaborative drug discovery (CDD) database [112] with these models led to selection and testing of 33 compounds. Among them, two compounds, flunarizine and pimozide, showed significant BCRP inhibition at $10\,\mu M$. All these models were built on rather small datasets, without using all the data available at the respective times of their studies.

Montanari *et al.* [113] compiled the largest set of 978 BCRP compounds available up to now by extracting information from 47 different studies. The authors reported an accuracy of 0.92 and an area under the ROC curve (AUC) of 0.85 in cross validation based on a naïve Bayes model. Later on, this dataset was used [75] to build a global binary classification model for prediction of BCRP inhibition. The final model was used to screen all the approved drugs in DrugBank to identify potential BCRP inhibitors. Ten drugs were selected and tested in BCRP-expressing PLB985 cells. Among them, two drugs, cisapride ($IC_{50} = 0.4\,\mu M$) and roflumilast ($IC_{50} = 0.9\,\mu M$), showed inhibition in the sub micromolar range.

6.17 MATE1

Protein-ligand interactions for organic cation transporters and the multidrug and toxin extrusion (MATE) transporter have been investigated using pharmacophores and quantitative structure-activity relationships [58, 82, 85, 86, 114]. In a recent study, Astorga *et al.* [114], characterized the relative selectivity of MATE1 and MATE2-K for some clinically important organic cations (OCs). Novel inhibitors for these transporters were identified and predictive models of MATE1 selectivity were developed. Using the IC_{50} values, a common-feature pharmacophore could be developed along with quantitative pharmacophores for hMATE1. Furthermore, a Bayesian model suggesting molecular features favoring and not favoring the interaction of ligands with hMATE1 was introduced [114].

In another study, Wittwer *et al.* [115] proposed an RF classification model to identify MATE1 inhibitors and non-inhibitors. The average AUC for 10 tests was 0.70 ± 0.05 (permutation test; p-value < 0.0001), indicating that models of good quality had been obtained.

6.18 ASBT

Efforts from Zheng *et al.* [67, 77], Rais *et al.* [78, 116], and González *et al.* [76] provided several QSAR models and pharmacophore models for ASBT binding affinity, with R^2 values between 0.68 and 0.89. All were trained on small congeneric series of conjugated bile acid derivatives. Classification QSARs of ASBT inhibitors based on 10 and 100 µM potency thresholds were reported by Sedykh *et al.* [55] and Zheng *et al.* [67], respectively.

To summarize this part, based on the data presented in Table 6.1, confined size of datasets has been a major limitation in developing highly accurate *in silico* prediction models to identify the drug interaction potential of hepatic transporters. The conformational flexibility of membrane transporters, the diverse chemical space covered by their substrates, and the inconsistency in data availability from experimental assays limit the predictive power of computational models even further.

6.19 Structure-Based Approaches

As stated earlier, the nonnavailability of resolved 3D structures of a number of membrane transporters is the reason for limited progress in structure-based approaches for transporter interaction prediction. However, in recent years, a number of 3D structures of ABC transporters have been resolved [117, 118]. Thus, improved performance of experimental approaches [119] has led to the development of structure-based models with decent performance.

Bikadi *et al.* [120] used SVM prediction and molecular docking approaches to predict P-gp substrate binding modes. Dolghih *et al.* [121] separated P-gp binders from non-binders via induced fit docking into the crystal structure of mouse P-gp (PDB ID: 3G60) [117] and using the docking score for subsequent classification. Further, Chen *et al.* [93] performed docking studies using 245 P-gp substrates and non-substrates, but could not clearly separate them on the basis of the Glide docking scores [122]. Klepsch *et al.* [123] docked a set of propafenones into a homology model of human P-gp. The study revealed that the binding poses are consistent with QSAR data, indicating that the observations can be exploited in identification of new P-gp inhibitors [124]. This study was further extended to structure-based classification of nearly 2000 compounds, which showed a prediction accuracy of 61% for the external test set compounds [71].

Although ligand-based approaches, owing to their high speed and accuracies, remain the method of choice for classification of transporter ligands, structure-assisted docking models show reasonable prediction accuracies in addition to providing valuable information on putative protein-ligand interactions at the molecular level.

6.20 Complex Models Incorporating Transporter Information

As described in the introduction, there is ample of evidence for the association between hepatic transporters and toxicity manifestations in the liver. This knowledge generated the idea that transporter information (inhibition, expression, or upregulation) could be incorporated within *in vitro* or *in silico* models, together with other assay data and physicochemical and/or biological descriptors. This is also in line with the FDA recommendations for transporters to be tested during drug development [125, 126]. Curiously, despite the fact that information on drug-transporter interactions is quite important and there are several *in vitro* and *in silico* models available for transporters *per se*, as outlined in the next section, there are only few studies combining the transporters information with other data.

6.21 *In Vitro* Models

There have been some well-established assays for hepatic transporters inhibition to predict liver toxicity. Especially in the case of BSEP, whose inhibition is linked with cholestasis, the respective screening is considered essential at the early stages of drug development. However, although there are several methods to measure BSEP inhibition, not all of them are equally suitable. In their review, Kis *et al.* [22] describe several appropriate *in vitro* methods that can predict BSEP-drug interactions. Furthermore, Szakács *et al.* present several *in vitro* methods and models for elucidating the ADMET profile of ABC transporters [127].

Thomson *et al.* have proposed a combination of assays for cytotoxicity [128]. Their suggestion is the use of a hazard matrix based on covalent binding, in conjunction with an array of five *in vitro* assays, addressing cytotoxicity in different cell lines and inhibition of the canalicular transporters BSEP and MRP2, with individual cutoff values for each assay. Aleo *et al.* have shown that the severity of human DILI is highly associated with the dual inhibition of mitochondrial function and BSEP, flagging them as two very important liability factors that should be checked during pharmaceutical screening [129]. Another study by Schadt *et al.* [130] proposed a methodology based on a compilation of assays to predict DILI for drug candidates. Among these assays are BSEP inhibition, glutathione adduct assay, CYP3A time-dependent inhibition, cytotoxicity in human hepatocytes, mitochondrial toxicity, and cytotoxicity in NIH 3T3 mouse fibroblasts. As a training set, 81 marketed or withdrawn compounds with differing DILI classes (according to FDA) were used. The resulted modeling approach yielded a performance of 79% overall accuracy, 76% sensitivity, and 82% specificity for the external test set composed of 39 compounds [130].

On a slightly different level, Dawson *et al*.s' [18] testing of 85 drugs for human BSEP inhibition, as well as its rat ortholog Bsep, followed by statistical analysis showed that inhibition of BSEP/Bsep correlates with the drug potential to cause DILI with an $r^2 = 0.94$. Moreover, all drugs with human BSEP IC$_{50}$ < 300 μM had molecular weight > 250, Clog$P$ > 1.5, and nonpolar surface area > 180 Å [18].

Similarly, in the work of Köck *et al.* [131], 88 drugs (100 μM) were investigated regarding their inhibitory effect on MRP3- and MRP4-mediated substrate transport. 50 BSEP non-inhibitors (24 non-cholestatic; 26 cholestatic) and 38 BSEP inhibitors (16 non-cholestatic; 22 cholestatic) were examined. MRP4 inhibition was associated with an increased cholestatic risk among BSEP non-inhibitors. In this group, for each 1% increase in MRP4 inhibition, the odds of the drug being cholestatic increased by 3.1%. By implementing a cutoff value of 21% for inhibition, which predicted a 50% chance of cholestasis, 62% of the cholestatic drugs inhibited MRP4 ($P < 0.05$). Nevertheless, merely 17% of non-cholestatic drugs were MRP4 inhibitors. Among BSEP inhibitors, MRP4 inhibition did not provide additional predictive value for cholestatic potential, as almost all BSEP inhibitors were also MRP4 inhibitors. The study failed to prove statistically significant association of MRP3 inhibition and cholestasis, regardless of the drug's capability to inhibit BSEP.

6.22   Multiscale Models

During the last decades, there has been a vast development in biomedical research, which allows the investigation of biological systems with higher level of detail and accuracy [132]. Multiscale models, that is, complex models that couple high- and low-resolution models thus allowing the study of biological systems from atomic to macroscopic levels [133], have made considerable contribution in this direction. The virtual liver network (VLN) is a characteristic example where several multiscale models are combined to simulate the function of a single organ [132]. Similar initiatives have also taken place previously for heart, such as the Virtual Heart (http://thevirtual heart.org/) [134] and the Living Heart Project (http://www.3ds.com/products-services/simulia/solutions/life-sciences/the-living-heart-project/) [135]. They combine information from the level of molecular targets, move toward molecular pathways/processes, then cellular/tissue processes, and end up at a tissue or whole-organ endpoint. This approach, apart from modeling the physiological function of an organ, can further be implemented for modeling whole-organ toxicity [136]. These multiscale models might facilitate the discovery of potentially hazardous drugs/chemicals at the early stages of drug discovery in a more efficient way than the single models, as more parameters that contribute to toxicity are taken into account.

In this direction, Diaz Ochoa *et al.* [137] developed a multiscale modeling framework for spatiotemporal prediction of substances' distribution that may result in hepatotoxicity. This framework consists of cellular models, a 2D liver model, and a whole-body model. Several mechanistic, genome-based *in silico* cells composite the 2D liver model and the whole-body model, including also the function of MRP2, MRP3, and MRP4. In principle, they use cellular systems for kinetic modeling and their aim was not only to calculate the drug concentration in the organ, but also the cell viability [137].

Another systems biology approach based on the analysis of dynamic adaptations in parameter trajectories (ADAPT) pointed out the important role of liver X receptor (LXR) activation for the development of steatosis [138]. Hijmans *et al.* showed that both input and output fluxes to hepatic triglyceride content can be induced by LXR activation, and during the early stages of LXR activation, steatosis can be induced by just a small imbalance between input/output fluxes of triglycerides. For the modeling analysis, mRNA levels of several mice genes were used, including Abcg1, which is known for its major role in cholesterol efflux from macrophage foam cells [139], and Abcg5, which forms a heterodimer with Abcg8 to translocate cholesterol and other plant sterols from the canalicular membrane into bile [16, 19, 39].

In addition, recent modeling approaches in our lab concerning prediction of hepatotoxicity endpoints by incorporating transporter interaction profiles follow the multiscale model concept. Apart from the prediction of hepatotoxicity endpoints, these models also aim to investigate the putative link of transporters inhibition with the respective toxicity endpoints. Initially, we used physicochemical descriptors of chemical compounds together with predictions of OATP1B1 and OATP1B3 inhibition [54] to predict hyperbilirubinemia [140]. In total a dataset of 836 compounds (86 positives and 749 negatives) for hyperbilirubinemia was used for training. Combination of MetaCost [141] and SMO (the SVM implementation in the WEKA [142] software package) using 93 interpretable 2D MOE [143] descriptors gave a performance of 68% accuracy and AUC. However, with respect to hyperbilirubinemia-transporter association, we only saw a weak relationship. For sure, more studies are expected in this field, which will allow targeting complex *in vivo* endpoints on a more sophisticated level than conventional machine learning methods currently allow.

6.23   Outlook

Transmembrane transport proteins represent a considerable fraction of the human genome. Their substrates cover a broad chemical space and range from

neurotransmitters via hormones up to a large panel of xenobiotics. Furthermore, they are also strongly involved in ADME and toxicity. One of the organs where a proper transporter homeostasis plays an important role is the liver. Imbalance in the function of the numerous transport proteins expressed in the liver has a big impact in its physiological function and subsequently in human health.

In the past decade, the community has faced a tremendous increase in knowledge on transmembrane transporters, their function, and their ligands. Several high-resolution structures were deposited in the Protein Data Bank, and specialized databases composed of inhibitors and substrates for transport proteins became available. These served in the development of *in silico* models for predicting transporter ligands. However, coverage is still quite limited and there is a strong need for high-quality data for particular transporters (NTCP, MRPs, MDR3) in order to develop more robust models for transporter inhibition. Furthermore, as generally observed for all target classes, the data available suffer from a "positive data bias," that is, they are heavily biased toward biologically active compounds. In addition, in most cases, the respective assay conditions are not available in a standardized form, which renders it difficult to compare data retrieved from different assays. Thus, it would be of major importance to have public available data depositories, which allow the deposition of both positive and negative data. These transporter data hubs should also follow the findable accessible, integratable reuse (FAIR) principles of data access [144] and allow data upload in a standardized format, especially with respect to assay conditions.

With respect to *in silico* toxicity prediction tools, multiscale models and virtual organs might be the near future of toxicity prediction. They are able to capture the necessary information from the molecular interaction with individual targets to the cellular response up to the whole tissue or organ. Of course, this is a complex challenge, but the first success stories for the heart demonstrate the advantage of a more holistic view on organ function and dysfunction. In addition, in this case, high-quality data are the key. They need to be provided on different levels, ranging from molecular interactions up to time/concentration series of solutes. In our opinion, all the tools necessary to pursue such a task for the liver are there already, and it just needs a concerted effort to make it happen.

Finally, following the increasing automation in life sciences, genotyping of patients will become routine soon. This opens up the whole field of single nucleotide polymorphisms (SNPs) and their consequences on response rates to medication. In addition, in the field of transporters, numerous SNPs are known which influence function and ligand recognition. This will add another layer of complexity to holistic prediction tools, but finally will link transporter informatics to precision medicine.

Acknowledgments

References

1 Faber, K.N., Muller, M., and Jansen, P.L. (2003) Drug transport proteins in the liver. *Adv. Drug Deliv. Rev.*, **55**, 107–124.

2 Jamei, M., Bajot, F., Neuhoff, S. *et al.* (2013) A mechanistic framework for in vitro–in vivo extrapolation of liver membrane transporters: prediction of drug-drug interaction between rosuvastatin and cyclosporine. *Clin. Pharmacokinet.*, **53**, 73–87.

3 Alrefai, W.A. and Gill, R.K. (2007) Bile acid transporters: structure, function, regulation and pathophysiological implications. *Pharm. Res.*, **24**, 1803–1823.

4 Roma, M.G., Crocenzi, F.A., and Sanchez Pozzi, E.A. (2008) Hepatocellular transport in acquired cholestasis: new insights into functional, regulatory and therapeutic aspects. *Clin. Sci. (London)*, **114**, 567–588.

5 Mita, S., Suzuki, H., Akita, H. *et al.* (2006) Inhibition of bile acid transport across Na^+/taurocholate cotransporting polypeptide (SLC10A1) and bile salt export pump (ABCB 11)-coexpressing LLC-PK1 cells by cholestasis-inducing drugs. *Drug Metab. Dispos.*, **34**, 1575–1581.

6 Chang, J.H., Plise, E., Cheong, J. *et al.* (2013) Evaluating the in vitro inhibition of UGT1A1, OATP1B1, OATP1B3, MRP2, and BSEP in predicting drug-induced hyperbilirubinemia. *Mol. Pharm.*, **10**, 3067–3075.

7 Campbell, S.D., de Morais, S.M., and Xu, J.J. (2004) Inhibition of human organic anion transporting polypeptide OATP 1B1 as a mechanism of drug-induced hyperbilirubinemia. *Chem. Biol. Interact.*, **150**, 179–187.

8 Hagenbuch, B. and Stieger, B. (2013) The SLCO (former SLC21) superfamily of transporters. *Mol. Aspects Med.*, **34**, 396–412.

9 Dhumeaux, D. and Erlinger, S. (2012) Hereditary conjugated hyperbilirubinaemia: 37 years later. *J. Hepatol.*, **58**, 388–390.

10 Keppler, D. (2014) The roles of MRP2, MRP3, OATP1B1, and OATP1B3 in conjugated hyperbilirubinemia. *Drug Metab. Dispos.*, **42**, 561–565.

11 Sticova, E. and Jirsa, M. (2013) New insights in bilirubin metabolism and their clinical implications. *World J. Gastroenterol.*, **19**, 6398–6407.

12 van de Steeg, E., Stranecky, V., Hartmannova, H. *et al.* (2012) Complete OATP1B1 and OATP1B3 deficiency causes human Rotor syndrome by interrupting conjugated bilirubin reuptake into the liver. *J. Clin. Invest.*, **122**, 519–528.

13 van de Steeg, E., Wagenaar, E., van der Kruijssen, C.M. *et al.* (2010) Organic anion transporting polypeptide 1a/1b-knockout mice provide insights into hepatic handling of bilirubin, bile acids, and drugs. *J. Clin. Invest.*, **120**, 2942–2952.

14 Koepsell, H., Lips, K., and Volk, C. (2007) Polyspecific organic cation transporters: structure, function, physiological roles, and biopharmaceutical implications. *Pharm. Res.*, **24**, 1227–1251.

15 Soroka, C.J., Ballatori, N., and Boyer, J.L. (2010) Organic solute transporter, OSTalpha-OSTbeta: its role in bile acid transport and cholestasis. *Semin. Liver Dis.*, **30**, 178–185.

16 Klaassen, C.D. and Aleksunes, L.M. (2010) Xenobiotic, bile acid, and cholesterol transporters: function and regulation. *Pharmacol. Rev.*, **62**, 1–96.

17 Ros, J.E., Libbrecht, L., Geuken, M. *et al.* (2003) High expression of MDR1, MRP1, and MRP3 in the hepatic progenitor cell compartment and hepatocytes in severe human liver disease. *J. Pathol.*, **200**, 553–560.

18 Dawson, S., Stahl, S., Paul, N. *et al.* (2011) In vitro inhibition of the bile salt export pump correlates with risk of cholestatic drug-induced liver injury in humans. *Drug Metab. Dispos.*, **40**, 130–138.

19 Chan, J. and Vandeberg, J.L. (2012) Hepatobiliary transport in health and disease. *Clin. Lipidol.*, **7**, 189–202.

20 Warner, D.J., Chen, H., Cantin, L.D. *et al.* (2012) Mitigating the inhibition of human bile salt export pump by drugs: opportunities provided by physicochemical property modulation, in silico modeling, and structural modification. *Drug Metab. Dispos.*, **40**, 2332–2341.

21 Garzel, B., Yang, H., Zhang, L. *et al.* (2013) The role of bile salt export pump gene repression in drug-induced cholestatic liver toxicity. *Drug Metab. Dispos.*, **42**, 318–322.

22 Kis, E., Ioja, E., Rajnai, Z. *et al.* (2012) BSEP inhibition: in vitro screens to assess cholestatic potential of drugs. *Toxicol. In Vitro*, **26**, 1294–1299.

23 Ogimura, E., Sekine, S., and Horie, T. (2011) Bile salt export pump inhibitors are associated with bile acid-dependent drug-induced toxicity in sandwich-cultured hepatocytes. *Biochem. Biophys. Res. Commun.*, **416**, 313–317.

24 Kock, K., Ferslew, B.C., Netterberg, I. *et al.* (2014) Risk factors for development of cholestatic drug-induced liver injury: inhibition of hepatic basolateral bile acid transporters multidrug resistance-associated proteins 3 and 4. *Drug Metab. Dispos.*, **42**, 665–674.

25 Templeton, I., Eichenbaum, G., Sane, R., and Zhou, J. (2014) Case study 5. Deconvoluting hyperbilirubinemia: differentiating between hepatotoxicity and reversible inhibition of UGT1A1, MRP2, or OATP1B1 in drug development. *Methods Mol. Biol.*, **1113**, 471–483.

26 Padda, M.S., Sanchez, M., Akhtar, A.J., and Boyer, J.L. (2011) Drug-induced cholestasis. *Hepatology*, **53**, 1377–1387.

27 Huang, L., Smit, J.W., Meijer, D.K., and Vore, M. (2000) Mrp2 is essential for estradiol-17beta(beta-D-glucuronide)-induced cholestasis in rats. *Hepatology*, **32**, 66–72.

28 Payen, L., Sparfel, L., Courtois, A. *et al.* (2002) The drug efflux pump MRP2: regulation of expression in physiopathological situations and by endogenous and exogenous compounds. *Cell Biol. Toxicol.*, **18**, 221–233.

29 Yang, K., Woodhead, J.L., Watkins, P.B. *et al.* (2014) Systems pharmacology modeling predicts delayed presentation and species differences in bile acid-mediated troglitazone hepatotoxicity. *Clin. Pharmacol. Ther.*, **96**, 589–598.

30 Rodrigues, A.D., Lai, Y., Cvijic, M.E. *et al.* (2013) Drug-induced perturbations of the bile acid pool, cholestasis, and hepatotoxicity: mechanistic considerations beyond the direct inhibition of the bile salt export pump. *Drug Metab. Dispos.*, **42**, 566–574.

31 Meier, Y., Pauli-Magnus, C., Zanger, U.M. *et al.* (2006) Interindividual variability of canalicular ATP-binding-cassette (ABC)-transporter expression in human liver. *Hepatology*, **44**, 62–74.

32 Park, H.J., Kim, T.H., Kim, S.W. *et al.* (2016) Functional characterization of ABCB4 mutations found in progressive familial intrahepatic cholestasis type 3. *Sci. Rep.*, **6**, 1–926872.

33 Sundaram, S.S. and Sokol, R.J. (2007) The multiple facets of ABCB4 (MDR3) deficiency. *Curr. Treat Options Gastroenterol.*, **10**, 495–503.

34 He, K., Cai, L., Shi, Q. *et al.* (2015) Inhibition of MDR3 activity in human hepatocytes by drugs associated with liver injury. *Chem. Res. Toxicol.*, **28**, 1987–1990.

35 Mahdi, Z.M., Synal-Hermanns, U., Yoker, A. *et al.* (2016) Role of multidrug resistance protein 3 in antifungal-induced cholestasis. *Mol. Pharmacol.*, **90**, 23–34.

36 Bodo, A., Bakos, E., Szeri, F. *et al.* (2003) The role of multidrug transporters in drug availability, metabolism and toxicity. *Toxicol. Lett.*, **140–141**, 133–143.

37 DeGorter, M.K., Xia, C.Q., Yang, J.J., and Kim, R.B. (2012) Drug transporters in drug efficacy and toxicity. *Annu. Rev. Pharmacol. Toxicol.*, **52**, 249–273.

38 Pauli-Magnus, C. and Meier, P.J. (2006) Hepatobiliary transporters and drug-induced cholestasis. *Hepatology*, **44**, 778–787.

39 Stieger, B., Kullak-Ublick, G.A., and DeLeve, L.D. (2013) Chapter 7 – Role of Membrane Transport in Hepatotoxicity and Pathogenesis of Drug-Induced Cholestasis A2, in *Drug-Induced Liver Disease*, 3rd edn (ed. N. Kaplowitz), Academic Press, Boston, pp. 123–133.

40 Tox21, http://tox21.org (accessed Aug 11, 2017).

41 Vinken, M., Landesmann, B., Goumenou, M. *et al.* (2016) AOP27: cholestatic liver injury induced by inhibition of the bile salt export pump (ABCB11).

42 Bento, A.P., Gaulton, A., Hersey, A. *et al.* (2013) The ChEMBL bioactivity database: an update. *Nucleic Acids Res.*, **42**, D1083–D1090.

43 ChEMBL v.22, DOI: 10.6019/CHEMBL.database.22.

44 Wang, Y., Suzek, T., Zhang, J. *et al.* (2014) PubChem BioAssay: 2014 update. *Nucleic Acids Res.*, **42**, D1075–D1082.

45 Greene, N., Fisk, L., Naven, R.T. *et al.* (2010) Developing structure–activity relationships for the prediction of hepatotoxicity. *Chem. Res. Toxicol.*, **23**, 1215–1222.

46 Fourches, D., Barnes, J.C., Day, N.C. *et al.* (2010) Cheminformatics analysis of assertions mined from literature that describe drug-induced liver injury in different species. *Chem. Res. Toxicol.*, **23**, 171–183.

47 Liu, Z., Shi, Q., Ding, D. *et al.* (2011) Translating clinical findings into knowledge in drug safety evaluation--drug induced liver injury prediction system (DILIps). *PLoS Comput. Biol.*, **7**, e1002310.

48 Chen, M., Vijay, V., Shi, Q. *et al.* (2011) FDA-approved drug labeling for the study of drug-induced liver injury. *Drug Discov. Today*, **16**, 697–703.

49 Chen, M., Suzuki, A., Thakkar, S. *et al.* (2016) DILIrank: the largest reference drug list ranked by the risk for developing drug-induced liver injury in humans. *Drug Discov. Today*, **21**, 648–653.

50 Zdrazil, B., Pinto, M., Vasanthanathan, P. *et al.* (2012) Annotating Human P-Glycoprotein Bioassay Data. *Mol. Inform.*, **31**, 599–609.

51 The Gene Ontology Consortium (2014) Gene Ontology Consortium: going forward. *Nucleic Acids Res.*, **43**, D1049–D1056.

52 Karlgren, M., Ahlin, G., Bergström, C.A. *et al.* (2012) In vitro and in silico strategies to identify OATP1B1 inhibitors and predict clinical drug–drug interactions. *Pharm. Res.*, **29**, 411–426.

53 Karlgren, M., Vildhede, A., Norinder, U. *et al.* (2012) Classification of inhibitors of hepatic organic anion transporting polypeptides (OATPs): influence of protein expression on drug-drug interactions. *J. Med. Chem.*, **55**, 4740–4763.

54 Kotsampasakou, E., Brenner, S., Jäger, W., and Ecker, G.F. (2015) Identification of novel inhibitors of organic anion transporting polypeptides 1B1 and 1B3 (OATP1B1 and OATP1B3) using a consensus vote of six classification models. *Mol. Pharm.*, **12**, 4395–4404.

55 Sedykh, A., Fourches, D., Duan, J. *et al.* (2013) Human intestinal transporter database: QSAR modeling and virtual profiling of drug uptake, efflux and interactions. *Pharm. Res.*, **30**, 996–1007.

56 Greupink, R., Nabuurs, S.B., Zarzycka, B. *et al.* (2012) In silico identification of potential cholestasis-inducing agents via modeling of Na(+)-dependent taurocholate cotransporting polypeptide substrate specificity. *Toxicol. Sci.*, **129**, 35–48.

57 Ahlin, G., Karlsson, J., Pedersen, J.M. *et al.* (2008) Structural requirements for drug inhibition of the liver specific human organic cation transport protein 1. *J. Med. Chem.*, **51**, 5932–5942.

58 Suhre, W.M., Ekins, S., Chang, C. *et al.* (2005) Molecular determinants of substrate/inhibitor binding to the human and rabbit renal organic cation transporters hOCT2 and rbOCT2. *Mol. Pharmacol.*, **67**, 1067–1077.

59 Xu, Y., Liu, X., Li, S. *et al.* (2013) Combinatorial pharmacophore modeling of organic cation transporter 2 (OCT2) inhibitors: insights into multiple inhibitory mechanisms. *Mol. Pharm.*, **10**, 4611–4619.

60 van Zanden, J.J., Wortelboer, H.M., Bijlsma, S. *et al.* (2005) Quantitative structure activity relationship studies on the flavonoid mediated inhibition of multidrug resistance proteins 1 and 2. *Biochem. Pharmacol.*, **69**, 699–708.

61 Tawari, N.R., Bag, S., and Degani, M.S. (2008) Pharmacophore mapping of a series of pyrrolopyrimidines, indolopyrimidines and their congeners as multidrug-resistance-associated protein (MRP1) modulators. *J. Mol. Model.*, **14**, 911–921.

62 Pajeva, I.K., Globisch, C., and Wiese, M. (2009) Combined pharmacophore modeling, docking, and 3D QSAR studies of ABCB1 and ABCC1 transporter inhibitors. *ChemMedChem*, **4**, 1883–1896.

63 Saito, H., Osumi, M., Hirano, H. *et al.* (2009) Technical pitfalls and improvements for high-speed screening and QSAR analysis to predict inhibitors of the human bile salt export pump (ABCB11/BSEP). *AAPS J.*, **11**, 581–589.

64 Pedersen, J.M., Matsson, P., Bergstrom, C.A. *et al.* (2008) Prediction and identification of drug interactions with the human ATP-binding cassette transporter multidrug-resistance associated protein 2 (MRP2; ABCC2). *J. Med. Chem.*, **51**, 3275–3287.

65 Montanari, F., Pinto, M., Khunweeraphong, N. *et al.* (2016) Flagging Drugs That Inhibit the Bile Salt Export Pump. *Mol. Pharm.*, **13**, 163–171.

66 Ng, C., Xiao, Y.D., Lum, B.L., and Han, Y.H. (2005) Quantitative structure–activity relationships of methotrexate and methotrexate analogues transported by the rat multispecific resistance-associated protein 2 (rMrp2). *Eur. J. Pharm. Sci.*, **26**, 405–413.

67 Zheng, X., Ekins, S., Raufman, J.P., and Polli, J.E. (2009) Computational models for drug inhibition of the human apical sodium-dependent bile acid transporter. *Mol. Pharm.*, **6**, 1591–1603.

68 Pinto, M., Trauner, M., and Ecker, G.F. (2012) An In Silico Classification Model for Putative ABCC2 Substrates. *Mol. Inform.*, **31**, 547–553.

69 Broccatelli, F., Carosati, E., Neri, A. *et al.* (2011) A novel approach for predicting P-glycoprotein (ABCB1) inhibition using molecular interaction fields. *J. Med. Chem.*, **54**, 1740–1751.

70 Chen, L., Li, Y., Zhao, Q. *et al.* (2011) ADME evaluation in drug discovery. 10. Predictions of P-glycoprotein inhibitors using recursive partitioning and naive Bayesian classification techniques. *Mol. Pharm.*, **8**, 889–900.

71 Klepsch, F., Vasanthanathan, P., and Ecker, G.F. (2014) Ligand and structure-based classification models for prediction of P-glycoprotein inhibitors. *J. Chem. Inf. Model.*, **54**, 218–229.

72 Erić, S., Kalinić, M., Ilić, K., and Zloh, M. (2014) Computational classification models for predicting the interaction of drugs with P-glycoprotein and breast cancer resistance protein. *SAR QSAR Environ. Res.*, **25**, 939–966.

73 Matsson, P., Englund, G., Ahlin, G. *et al.* (2007) A global drug inhibition pattern for the human ATP-binding cassette transporter breast cancer resistance protein (ABCG2). *J. Pharmacol. Exp. Ther.*, **323**, 19–30.

74 Pan, Y., Chothe, P.P., and Swaan, P.W. (2013) Identification of novel breast cancer resistance protein (BCRP) inhibitors by virtual screening. *Mol. Pharm.*, **10**, 1236–1248.

75 Montanari, F., Cseke, A., Wlcek, K., and Ecker, G.F. (2017) Virtual screening of drugbank reveals two drugs as new BCRP inhibitors. *J. Biomol. Screen.*, **22**, 86–93.

76 González, P.M., Acharya, C., Mackerell, A.D. Jr.,, and Polli, J.E. (2009) Inhibition requirements of the human apical sodium-dependent bile acid transporter (hASBT) using aminopiperidine conjugates of glutamyl-bile acids. *Pharm. Res.*, **26**, 1665–1678.

77 Zheng, X., Pan, Y., Acharya, C. *et al.* (2010) Structural requirements of the ASBT by 3D-QSAR analysis using aminopyridine conjugates of chenodeoxycholic acid. *Bioconjug. Chem.*, **21**, 2038–2048.

78 Rais, R., Acharya, C., Mackerell, A.D., and Polli, J.E. (2010) Structural determinants for transport across the intestinal bile acid transporter using C-24 bile acid conjugates. *Mol. Pharm.*, **7**, 2240–2254.

79 Yang, K., Kock, K., Sedykh, A. *et al.* (2013) An updated review on drug-induced cholestasis: mechanisms and investigation of physico-chemical properties and pharmacokinetic parameters. *J. Pharm. Sci.*, **102**, 3037–3057.

80 De Bruyn, T., van Westen, G.J., Ijzerman, A.P. *et al.* (2013) Structure-based identification of OATP1B1/3 inhibitors. *Mol. Pharmacol.*, **83**, 1257–1267.

81 van de Steeg, E., Venhorst, J., Jansen, H.T. *et al.* (2015) Generation of Bayesian prediction models for OATP-mediated drug–drug interactions based on inhibition screen of OATP1B1, OATP1B1 *15 and OATP1B3. *Eur. J. Pharm. Sci.*, **70**, 29–36.

82 Bednarczyk, D., Ekins, S., Wikel, J.H., and Wright, S.H. (2003) Influence of molecular structure on substrate binding to the human organic cation transporter, hOCT1. *Mol. Pharmacol.*, **63**, 489–498.

83 Moaddel, R., Patel, S., Jozwiak, K. *et al.* (2005) Enantioselective binding to the human organic cation transporter-1 (hOCT1) determined using an immobilized hOCT1 liquid chromatographic stationary phase. *Chirality*, **17**, 501–506.

84 Moaddel, R., Ravichandran, S., Bighi, F. *et al.* (2007) Pharmacophore modelling of stereoselective binding to the human organic cation transporter (hOCT1). *Br. J. Pharmacol.*, **151**, 1305–1314.

85 Zolk, O., Solbach, T.F., König, J., and Fromm, M.F. (2009) Structural determinants of inhibitor interaction with the human organic cation transporter OCT2 (SLC22A2). *Naunyn Schmiedebergs Arch. Pharmacol.*, **379**, 337–348.

86 Kido, Y., Matsson, P., and Giacomini, K.M. (2011) Profiling of a prescription drug library for potential renal drug-drug interactions mediated by the organic cation transporter 2. *J. Med. Chem.*, **54**, 4548–4558.

87 Chang, C., Ekins, S., Bahadduri, P., and Swaan, P.W. (2006) Pharmacophore-based discovery of ligands for drug transporters. *Adv. Drug Deliv. Rev.*, **58**, 1431–1450.

88 Akanuma, S., Uchida, Y., Ohtsuki, S. *et al.* (2011) Molecular-weight-dependent, anionic-substrate-preferential transport of beta-lactam antibiotics via multidrug resistance-associated protein 4. *Drug Metab. Pharmacokinet.*, **26**, 602–611.

89 Hirano, H., Kurata, A., Onishi, Y. *et al.* (2006) High-speed screening and QSAR analysis of human ATP-binding cassette transporter ABCB11 (bile salt export pump) to predict drug-induced intrahepatic cholestasis. *Mol. Pharm.*, **3**, 252–265.

90 Pedersen, J.M., Matsson, P., Bergström, C.A. *et al.* (2013) Early identification of clinically relevant drug interactions with the human bile salt export pump (BSEP/ABCB11). *Toxicol. Sci.*, **136**, 328–343.

91 Zhang, H., Xiang, M.L., Zhao, Y.L. *et al.* (2009) Support vector machine and pharmacophore-based prediction models of multidrug-resistance protein 2 (MRP2) inhibitors. *Eur. J. Pharm. Sci.*, **36**, 451–457.

92 Chang, C. and Swaan, P.W. (2006) Computational approaches to modeling drug transporters. *Eur. J. Pharm. Sci.*, **27**, 411–424.

93 Chen, L., Li, Y., Yu, H. *et al.* (2012) Computational models for predicting substrates or inhibitors of P-glycoprotein. *Drug Discov. Today*, **17**, 343–351.

94 Demel, M.A., Kramer, O., Ettmayer, P. *et al.* (2009) Predicting ligand inter-actions with ABC transporters in ADME. *Chem. Biodivers.*, **6**, 1960–1969.

95 Ecker, G.F., Stockner, T., and Chiba, P. (2008) Computational models for prediction of interactions with ABC-transporters. *Drug Discov. Today*, **13**, 311–317.

96 Gandhi, Y.A. and Morris, M.E. (2009) Structure–activity relationships and quantitative structure–activity relationships for breast cancer resistance protein (ABCG2). *AAPS J.*, **11**, 541–552.

97 Montanari, F. and Ecker, G.F. (2015) Prediction of drug–ABC-transporter interaction – recent advances and future challenges. *Adv. Drug Deliv. Rev.*, **86**, 17–26.

98 Dearden, J.C., Al-Noobi, A., Scott, A.C., and Thomson, S.A. (2003) QSAR studies on P-glycoprotein-regulated multidrug resistance and on its rever-sal by phenothiazines. *SAR QSAR Environ. Res.*, **14**, 447–454.

99 Kupsáková, I., Rybár, A., Docolomanský, P. *et al.* (2004) Reversal of P-glycoprotein mediated vincristine resistance of L1210/VCR cells by ana-logues of pentoxifylline: a QSAR study. *Eur. J. Pharm. Sci.*, **21**, 283–293.

100 Osterberg, T. and Norinder, U. (2000) Theoretical calculation and predic-tion of P-glycoprotein-interacting drugs using MolSurf parametrization and PLS statistics. *Eur. J. Pharm. Sci.*, **10**, 295–303.

101 Wang, R.B., Kuo, C.L., Lien, L.L., and Lien, E.J. (2003) Structure–activity relationship: analyses of p-glycoprotein substrates and inhibitors. *J. Clin. Pharm. Ther.*, **28**, 203–228.

102 Wang, Y.H., Li, Y., Yang, S.L., and Yang, L. (2005) An in silico approach for screening flavonoids as P-glycoprotein inhibitors based on a Bayesian-regularized neural network. *J. Comput. Aided Mol. Des.*, **19**, 137–147.

103 Jabeen, I., Wetwitayaklung, P., Chiba, P. *et al.* (2013) 2D- and 3D-QSAR studies of a series of benzopyranes and benzopyrano[3,4*b*][1,4]-oxazines as inhibitors of the multidrug transporter P-glycoprotein. *J. Comput. Aided Mol. Des.*, **27**, 161–171.

104 Shen, J., Cui, Y., Gu, J. *et al.* (2014) A genetic algorithm- back propagation artificial neural network model to quantify the affinity of flavonoids toward P-glycoprotein. *Comb. Chem. High Throughput Screen.*, **17**, 162–172.

105 Sousa, I.J., Ferreira, M.J., Molnar, J., and Fernandes, M.X. (2013) QSAR studies of macrocyclic diterpenes with P-glycoprotein inhibitory activity. *Eur. J. Pharm. Sci.*, **48**, 542–553.

106 Vázquez, R.N., Camargo, A.B., Marchevsky, E.J., and Luco, J.M. (2014) Molecular factors influencing the affinity of flavonoid compounds on P-glycoprotein efflux transporter. *Curr. Comput. Aided Drug Des.*, **10**, 250–258.

107 Wu, J., Li, X., Cheng, W. *et al.* (2009) Quantitative structure activity relationship (QSAR) approach to multiple drug resistance (MDR) modulators based on combined hybrid system. *QSAR & Comb. Sci.*, **28**, 969–978.

108 Sakiyama, Y. (2009) The use of machine learning and nonlinear statistical tools for ADME prediction. *Expert Opin. Drug Metab. Toxicol.*, **5**, 149–169.

109 Wang, Y.H., Li, Y., Yang, S.L., and Yang, L. (2005) Classification of substrates and inhibitors of P-glycoprotein using unsupervised machine learning approach. *J. Chem. Inf. Model.*, **45**, 750–757.

110 Ngo, T.D., Tran, T.D., Le, M.T., and Thai, K.M. (2016) Computational predictive models for P-glycoprotein inhibition of in-house chalcone derivatives and drug-bank compounds. *Mol. Divers.*, **20**, 945–961.

111 Smith, A.J., van Helvoort, A., van Meer, G. *et al.* (2000) MDR3 P-glycoprotein, a phosphatidylcholine translocase, transports several cytotoxic drugs and directly interacts with drugs as judged by interference with nucleotide trapping. *J. Biol. Chem.*, **275**, 23530–23539.

112 Ekins, S. and Bunin, B.A. (2013) The Collaborative Drug Discovery (CDD) database. *Methods Mol. Biol.*, **993**, 139–154.

113 Montanari, F. and Ecker, G.F. (2014) BCRP inhibition: from data collection to ligand-based modeling. *Mol. Inform.*, **33**, 322–331.

114 Astorga, B., Ekins, S., Morales, M., and Wright, S.H. (2012) Molecular determinants of ligand selectivity for the human multidrug and toxin extruder proteins MATE1 and MATE2-K. *J. Pharmacol. Exp. Ther.*, **341**, 743–755.

115 Wittwer, M.B., Zur, A.A., Khuri, N. *et al.* (2013) Discovery of potent, selective multidrug and toxin extrusion transporter 1 (MATE1, SLC47A1) inhibitors through prescription drug profiling and computational modeling. *J. Med. Chem.*, **56**, 781–795.

116 Rais, R., Acharya, C., Tririya, G. *et al.* (2010) Molecular switch controlling the binding of anionic bile acid conjugates to human apical sodium-dependent bile acid transporter. *J. Med. Chem.*, **53**, 4749–4760.

117 Aller, S.G., Yu, J., Ward, A. *et al.* (2009) Structure of P-glycoprotein reveals a molecular basis for poly-specific drug binding. *Science*, **323**, 1718–1722.

118 Klepsch, F. and Ecker, G.F. (2010) Impact of the recent mouse P-glycoprotein structure for structure-based ligand design. *Mol. Inform.*, **29**, 276–286.

119 Winter, S.S., Lovato, D.M., Khawaja, H.M. *et al.* (2008) High-throughput screening for daunorubicin-mediated drug resistance identifies mometasone furoate as a novel ABCB1-reversal agent. *J. Biomol. Screen.*, **13**, 185–193.

120 Bikadi, Z., Hazai, I., Malik, D. *et al.* (2011) Predicting P-glycoprotein-mediated drug transport based on support vector machine

and three-dimensional crystal structure of P-glycoprotein. *PLoS One*, **6**, e25815.

121 Dolghih, E., Bryant, C., Renslo, A.R., and Jacobson, M.P. (2011) Predicting binding to p-glycoprotein by flexible receptor docking. *PLoS Comput. Biol.*, **7**, e1002083.

122 Friesner, R.A., Murphy, R.B., Repasky, M.P. *et al.* (2006) Extra precision glide: docking and scoring incorporating a model of hydrophobic enclosure for protein–ligand complexes. *J. Med. Chem.*, **49**, 6177–6196.

123 Klepsch, F., Chiba, P., and Ecker, G.F. (2011) Exhaustive sampling of docking poses reveals binding hypotheses for propafenone type inhibitors of P-glycoprotein. *PLoS Comput. Biol.*, **7**, e1002036.

124 Prokes K. (2012) Development of "in silico" models for identification of new ligands acting as pharmacochaperones for P-glycoprotein. University of Vienna.

125 Zhang, L., Strong, J.M., Qiu, W. *et al.* (2006) Scientific perspectives on drug transporters and their role in drug interactions. *Mol. Pharm.*, **3**, 62–69.

126 FDA (2012) Guidance for industry: drug interaction studies – study design, data analysis, implications for dosing, and labeling recommendations, available at http://www.fda.gov/downloads/Drugs/ GuidanceComplianceRegulatoryInformation/Guidances/ucm292362.pdf (accessed September 27, 2017).

127 Szakács, G., Varadi, A., Ozvegy-Laczka, C., and Sarkadi, B. (2008) The role of ABC transporters in drug absorption, distribution, metabolism, excretion and toxicity (ADME-Tox). *Drug Discov. Today*, **13**, 379–393.

128 Thompson, R.A., Isin, E.M., Li, Y. *et al.* (2012) In vitro approach to assess the potential for risk of idiosyncratic adverse reactions caused by candidate drugs. *Chem. Res. Toxicol.*, **25**, 1616–1632.

129 Aleo, M.D., Luo, Y., Swiss, R. *et al.* (2014) Human drug-induced liver injury severity is highly associated with dual inhibition of liver mitochondrial function and bile salt export pump. *Hepatology*, **60**, 1015–1022.

130 Schadt, S., Simon, S., Kustermann, S. *et al.* (2015) Minimizing DILI risk in drug discovery – a screening tool for drug candidates. *Toxicol. In Vitro*, **30**, 429–437.

131 Köck, K., Ferslew, B.C., Netterberg, I. *et al.* (2014) Risk factors for development of cholestatic drug-induced liver injury: inhibition of hepatic basolateral bile acid transporters multidrug resistance-associated proteins 3 and 4. *Drug Metab. Dispos.*, **42**, 665–674.

132 Holzhütter, H.G., Drasdo, D., Preusser, T. *et al.* (2012) The virtual liver: a multidisciplinary, multilevel challenge for systems biology. *Wiley Interdiscip. Rev. Syst. Biol. Med.*, **4**, 221–235.

133 Ayton, G.S., Noid, W.G., and Voth, G.A. (2007) Multiscale modeling of biomolecular systems: in serial and in parallel. *Curr. Opin. Struct. Biol.*, **17**, 192–198.

134 Cherry, E. and Fenton, F. The Virtual Heart http://thevirtualheart.org/ (accessed September 28, 2017).

135 The Living Heart Project. http://www.3ds.com/products-services/simulia/ solutions/life-sciences/the-living-heart-project/ (accessed September 28, 2017).

136 Judson, R., Elloumi, F., Setzer, R.W. *et al.* (2008) A comparison of machine learning algorithms for chemical toxicity classification using a simulated multi-scale data model. *BMC Bioinformatics*, **9**, 1–16241.

137 Diaz Ochoa, J.G., Bucher, J., Pery, A.R. *et al.* (2013) A multi-scale modeling framework for individualized, spatiotemporal prediction of drug effects and toxicological risk. *Front. Pharmacol.*, **3**, 1–11204.

138 Hijmans, B.S., Tiemann, C.A., Grefhorst, A. *et al.* (2016) A systems biology approach reveals the physiological origin of hepatic steatosis induced by liver X receptor activation. *FASEB J.*, **29**, 1153–1164.

139 Yvan-Charvet, L., Wang, N., and Tall, A.R. (2010) Role of HDL, ABCA1, and ABCG1 transporters in cholesterol efflux and immune responses. *Arterioscler. Thromb. Vasc. Biol.*, **30**, 139–143.

140 Kotsampasakou, E., Escher, S.E., and Ecker, G.F. (2017) Linking transporter interaction profiles to toxicity – the hyperbilirubinemia use case. *Eur. J. Pharm. Sci.*, **100**, 9–16.

141 Pedro D. (1999) *MetaCost: A General Method for Making Classifiers Cost-Sensitive.* Proceedings of the fifth ACM SIGKDD international conference on Knowledge discovery and data mining, San Diego, California, USA, ACM.

142 Hall, M., Frank, E., Holmes, G. *et al.* (2009) The WEKA data mining software: an update. *SIGKDD Explor. NewsL.*, **11**, 10–18.

143 Molecular Operating Environment (MOE) (2015) 1010 Sherbooke St. West, Suite #910, Montreal, QC, Canada, H3A 2R7, Chemical Computing Group Inc.

144 Wilkinson, M.D., Dumontier, M., Aalbersberg, I.J. *et al.* (2016) The FAIR Guiding Principles for scientific data management and stewardship. *Sci. Data*, **3**, 1–9160018.

145 Warner, D.J., Chen, H., Cantin, L.D. *et al.* (2012) Mitigating the inhibition of human bile salt export pump by drugs: opportunities provided by physicochemical property modulation, in silico modelling, and structural modification. *Drug Metab. Dispos.*, **40**, 2332–2341.

2.2 Multi-target prediction

2.2.1 Linked open data: ligand-transporter interaction profiling and beyond

The shift in paradigm from 'one-drug-one-target' to 'one-drug-many-targets' is marked by several current drugs that specifically interact with multiple biological targets. Similarities and synergies and even more importantly side effects and adverse reactions demand a most thorough development process covering as many factors as possible. Computational methods often provide highly useful platforms to tackle these issues due to their immense flexibility and ability to deal with big data. In this chapter, we present a workflow for systematic extraction and curation of data for multiple drug targets from the public domain and provide insights into how such data can be employed in the development of ligand and structure-based approaches while discussing the bottlenecks to be considered with respect to data analysis.

S. Kickinger, E. Hellsberg, **S. Jain** performed the literature search and wrote the chapter. G.F. Ecker supervised the work and revised the chapter.

Linked Open Data: Ligand-Transporter Interaction Profiling and Beyond

Stefanie Kickinger, Eva Hellsberg, Sankalp Jain, and Gerhard F. Ecker

Abstract

Multi-target drug design is an innovative new paradigm in the drug development process. With the help of growing open data sources, in silico modeling approaches have become successful tools to discover and investigate multi-target drugs. In this chapter, we describe a workflow for retrieving and curating information for multiple drug targets from the open domain, provide insights into how the retrieved data can be employed in ligand and structure-based approaches, and discuss the hurdles to consider with respect to data analysis.

Keywords KNIME workflow, Ligand-based design, Molecular docking, Molecular dynamics simulation, Multi-target drug design, Open data, Protein homology modeling, Structure-based design

1 Introduction

Multi-target drug design is an emerging new paradigm to treat complex diseases by regulating multiple targets at the same time to achieve the desired physiological responses [1–4]. Traditionally, drugs have been designed to selectively modulate a so-called on-target in order to avoid side effects by modulating "off-targets." However, several approved drugs retrospectively have been shown to hit more than one target, which turned out to contribute to the therapeutic efficacy [5, 6]. Furthermore, in recent years many drugs failed in phase II clinical trials because of a lack of therapeutic efficacy [7]. Therefore, multi-target drug design represents an innovative principle to overcome lack of efficacy. Different approaches to discover and investigate multi-target drugs have been reviewed by Zhang et al. [8] addressing data-driven, ligand-based, or structure-based methods [4, 9–14]. Most of these methods focus on drug repurposing (i.e., to find new targets for known drugs) such as the ligand-based methods SPiDER [15] and SEA [16], which are based on 2D fingerprint or 3D shape similarity. Furthermore, structure-based methods such as TarFisDOCK [17], INVDOCK [14, 18], or

Stefanie Kickinger, Eva Hellsberg, and Sankalp Jain contributed equally to this work.

VinaMPI [18] could be used to dock potential ligands into many 36
target structures at the same time [19, 20]. With the help of growing 37
open data sources such as Open PHACTS [21], ChEMBl [22], and 38
freely available medicinal chemistry literature, data-driven in silico 39
modeling approaches have also proven to be capable of effectively 40
identifying protein-ligand interactions at an early stage in the drug 41
discovery pipeline [23]. However, increase in complexity and size 42
and diversity of public data sources necessitate judicious curation of 43
the data before using them. With the availability of workflow tools 44
like KNIME [24] or pipeline pilot [25], complex querying for 45
multiple drug targets became a feasible task without the need of 46
comprehensive programming skills [26]. In this chapter, we present 47
a protocol which starts with mining the Open PHACTS Discovery 48
Platform to collect a data-set of suitable size and quality for 49
subsequent structure-based selectivity profiling studies. As concrete 50
case study, we chose the human serotonin (hSERT) and dopamine 51
transporter (hDAT). Both proteins belong to the neurotransmitter 52
sodium symporter family which represents the largest group of 53
transporters in the human genome. hSERT and hDAT are responsi- 54
ble for the reuptake of serotonin and dopamine, respectively, from 55
the presynaptic cleft after signaling [27, 28]. Numerous drugs have 56
been developed which interact with these transporters and are used 57
as therapeutic agents to treat neurological disorders such as depres- 58
sion. In addition, there is a wealth of compounds which are abused as 59
illicit drugs [28–30]. Even though hSERT and hDAT share high 60
sequence and structural similarity, they fulfill different physiological 61
roles. Substances increasing dopamine levels in the mesolimbic path- 62
way of the brain can influence the reward system, whereas increased 63
levels of serotonin are involved in several other neurotransmitter 64
systems, most importantly influencing mood [31]. A profound 65
understanding of the structural basis for hSERT and hDAT ligand 66
selectivity is therefore of major interest for designing ligands that 67
either hit one of these transporters or both. This chapter will tackle 68
this research question by reviewing the data mining and curating 69
process for hSERT and hDAT bioactivities present in the linked open 70
data domain. This is followed by a comprehensive scaffold analysis in 71
order to analyze the chemical space, which allowed to identify a 72
congeneric series of compounds suitable for structure-activity rela- 73
tionship studies and experimental data guided ligand docking. The 74
power of this protocol is based on the combination of mining the 75
available knowledge in the open data domain and its breakdown to 76
concrete molecular interactions. This chapter thus gives an overview 77
of the overall workflow, points out the potential of retrieving data for 78
multiple drug targets from the open domain, provides insights into 79
structure-based approaches, and discusses the hurdles to be consid- 80
ered in data analysis. 81

2 Materials

82

Data retrieval and scaffold analysis

83

- Knime [24]: Knime is an open-source platform that provides an integrated solution for the data mining process across the drug discovery pipeline. It can be downloaded from https://www.knime.com/software. It also provides a visual assembly of data workflows drawn from an extensive repository of tools. Additionally, it also offers nodes for machine learning (classification and regression analysis).

84
85
86
87
88
89
90

Homology modeling

91

- MODELLER [32]: Modeller is a widely used open-source software for comparative modeling of protein three-dimensional structures. The program also incorporates limited functions for ab initio structure prediction of loop regions of proteins, which are often highly variable even among homologous proteins and thus difficult to predict by homology modeling. It can be downloaded from https://salilab.org/modeller.

92
93
94
95
96
97
98

Molecular docking and visualization

99

- Schrödinger [33]: Schrödinger is one of the leading commercial software packages in the field of drug design. It includes small molecule modeling and simulations, macromolecular modeling and simulations, lead discovery, and lead optimization, visualization, and automation (https://www.schrodinger.com/maestro). Glide [34] is the molecular docking module in Schrödinger that places the ligand in the protein binding pocket and ranks the generated poses with an empirical scoring function.

100
101
102
103
104
105
106
107

- Molecular Operating Environment (MOE) [35]: MOE is a commercial drug discovery software platform that integrates visualization, modeling, and simulations, as well as methodology development, in one package (http://www.chemcomp.com/).

108
109
110
111

Molecular dynamics

112

- Desmond [36]: Desmond is a freely available software package developed at D. E. Shaw Research to perform high-speed molecular dynamics simulations of biological systems (http://www.deshawresearch.com/index.html). Schrödinger provides an easy-to-use graphical user interface for performing molecular dynamics simulations with Desmond [37].

113
114
115
116
117
118

3   Methods

Sophisticated approaches are necessary to tackle multi-target drug design. The great variety of methodological possibilities demands well-informed decisions on which individual path to embark. In this section, we describe the methods in detail which we used to retrieve and curate information on two drug targets. Note that this example was driven by the solid basis of available experimental data and previous findings on these drug targets. All technical parameters described in the methods section are either the default options recommended by the software developers or adapted due to specific biological evidence relevant for the focus of the study.

3.1   Data Collection and Data Mining

Open data sources such as ChEMBL [22], DrugBank [38], KEGG [39], or Open PHACTS [21] provide a large collection of linked information on compounds including their structures, biological targets, pathways, bioactivities, and experimental details on biological assays. ChEMBL and other resources extract their information from the literature in an automated or semiautomated fashion. The collected data therefore originate from a variety of different resources resulting in a collection of bioactivity data of different activity endpoints (K_i, IC_{50}, % inhibition, etc.) that was measured in different assay types and under varying assay conditions (*see* **Note 1**). However, using such diverse data for modeling or virtual screening was reported to show inconsistent performance, and hence recommendations were proposed to deal with the experimental uncertainty associated with such data [40, 41]. For our case study, bioactivity data for hSERT and hDAT were extracted from the Open PHACTS Discovery Platform via a KNIME workflow. The application programming interface (API) call was used to retrieve pharmacology data from ChEMBL20 for both proteins. In the present case study, we decided to include the bioactivity endpoints IC_{50} and K_i, because these bioactivities have been demonstrated to be most reliably in large data analysis [41, 42] and because they can be correlated with each other. In order to investigate the uncertainty of the data that was introduced by combining these different activity endpoints from different assays, the correlation between pIC_{50} and pK_i (p = negative log) values from duplicate measurements for hSERT and hDAT was calculated. This showed that the observed correlations are within the same range as the calculated correlations for duplicate measurements within only one of the activity endpoints [43]. As a next step, classification of the data into active and inactive compounds has to be performed in order to extract the actives. Setting reasonable activity thresholds is a challenging task, and it requires considering the focus of the study. In the present case, the thresholds were tailored according to the lowest known activity endpoints (IC_{50} and K_i) that still showed pharmacological activity.

If a dataset is used for calculating structure-activity relationships 164
(SAR), the compounds must be measured for the same activity 165
endpoint (i.e., either IC_{50} alone or K_i alone). However, if a dataset 166
is designed for the construction of machine learning models, also the 167
use of activity annotations is possible (i.e., active, 1, inactive, 0). In 168
this scenario, the data from different endpoints can be merged 169
(as described above). To increase the accuracy of the classification 170
of the dataset, data points close to the activity thresholds might be 171
omitted. Inconsistent data points with conflicting activity data 172
should in general be omitted from the dataset. In order to visualize 173
the diversity of the dataset and to see if there are scaffolds showing 174
pronounced selectivity for one or both targets, Bemis-Murcko scaf- 175
fold analysis [44] was performed. Out of the 53 most populated 176
scaffolds, four scaffolds were identified as hDAT selective, 10 as 177
hSERT selective, and 24 as promiscuous. In order to perform quan- 178
titative structure-activity relationship (SAR) calculations, scaffolds 179
that contained congeneric series of compounds, which showed selec- 180
tivity for one of the targets and were measured in the same assay, 181
were prioritized. A congeneric series of 56 compounds sharing a 182
cathinone substructure was identified that showed pronounced 183
selectivity for hDAT over hSERT. A detailed description of the 184
KNIME workflow for data retrieval, filtering, preprocessing, and 185
analyses can be found in [43]. The whole workflow can be down- 186
loaded from myExperiment [45]. Out of the whole set of derivatives, 187
six compounds were further selected for subsequent structure-based 188
studies in order to link the observed selectivity profile to specific 189
molecular interactions. 190

191

3.2 Ligand-Based Methods

In general, ligand-based methods can be used to find trends in the 192
data (as discussed above) or to classify compounds with machine 193
learning methods. However, their application depends strongly on 194
the data quality. In our case study, we analyzed the SAR of the 195
56 cathinones to get first insights which molecular features trigger 196
their selectivity profiles. Since the compounds show selectivity for 197
hDAT (over hSERT), we performed multiple linear regression 198
(classical Hansch analysis) with hDAT pIC_{50} values and selectivity 199
($= \log(\text{hSERT } IC_{50}/\text{hDAT } IC_{50})$) as dependent variables using a 200
limited set of descriptors characterizing the molecules (Van der 201
Waals volume (overall, Cα- and N-substituents), partition coeffi- 202
cient ($\log P$ (o/w)), molar refractivity, constants for the substitu- 203
ents on the aromatic ring, and indicator variables for meta- and 204
para-substitutions). Briefly, both calculated equations showed a 205
first trend that the substituent on the Cα-atom to the carbonyl 206
group of the compounds influences hDAT activity and selectivity. 207
Details on the approach can be found in [43]. This information is 208
subsequently used to guide the prioritization of docking poses. 209

210

3.3 Structure-Based Methods

Structure-based methods require 3D coordinates from available high-resolution crystal structures, NMR experiments, or homologous template structures. A plethora of crystal structures is deposited in the Protein Data Bank [46] (PDB, www.rcsb.org) and can be downloaded free of charge. All selected crystal structures should be checked thoroughly whether the resolution and B-factors are appropriate, if certain amino acids are annotated with multiple possible rotamers, and if there are relevant amino acids missing (*see* **Note 2**). This procedure can be performed with commercial protein visualization software (MOE [35] or Schrödinger Suite [33]) or free software (VMD [47] or pymol [48]). A lot of information can be already taken from the downloaded pdb files themselves, as they are written in text format and include the experimental data and setup. A visual inspection of PDB structures is also possible in a web browser using the LiteMol viewer [49, 50] in PDBe (https://www.ebi.ac.uk/pdbe/) [51]. Since many crystal structures are models retrieved by X-ray crystallography based on experimentally measured diffraction patterns, it is furthermore advisable to check the placement of the protein and its ligands in the experimentally measured electron density map [31, 52]. Electron density maps can be visualized with commercial software (Schrödinger [33], MOE [35]) and free software (Coot [53]) or in the web browser (LiteMol [49, 50], PDBe [51]). By considering the abovementioned procedures, one can identify the areas of the crystal structure where the structure can be trusted or should be taken with caution. In the case of our study, no crystal structures of hSERT and hDAT were resolved back then. Consequently, homology modeling needs to be performed to obtain decent models based on suitable template crystal structures.

3.3.1 Homology Modeling

Homology modeling or comparative modeling refers to the technique of using a resolved crystal structure to model an unknown homologous protein structure. It is believed that overall fold is far more conserved among different proteins than sequence identity [54]. There are four crucial steps in homology modeling. First, a suitable crystal structure is chosen as a template. At the time this analysis was performed, the PDB provided two different types of homologous template structures for modeling hSERT and hDAT: crystal structures of the bacterial leucine transporter (LeuT, sequence 20%) [55] and the drosophila dopamine transporter (dDAT, sequence identity 70%) [56]. In the present case study, the dDAT PDB structure 4M48 was chosen as the most suitable template due to higher sequence identity and the fact that it shows the desired outward-open conformation (*see* **Note 3**). Second, the desired protein sequence needs to be aligned with the template structure. This task was performed with the tool ClustalX [57]. All 12 transmembrane helices (TMs) of hSERT and hDAT are highly

conserved and can be easily aligned with the template structure. 258
Third, models are generated and refined, e.g., with the program 259
Modeller, which was also the program of choice in this study 260
[32]. Within Modeller it is possible to also implement experimental 261
data in the model generation process by setting restraints for sec- 262
ondary structure elements, disulfides or salt bridges. Fourth, the 263
models' quality needs to be assessed with help of, e.g., the DOPE 264
score (*see* **Note 4**) [58]. Additional quality assessment can be 265
performed with ProCHECK (https://www.ebi.ac.uk/thornton- 266
srv/software/PROCHECK/index.html) [58, 59] and ProQM 267
(http://bioinfo.ifm.liu.se/ProQM/index.php) [60]. Procheck 268
additionally provides Ramachandran plots and information on resi- 269
due properties. ProQM was specifically optimized for membrane 270
proteins. Nevertheless, the quality of the homology model depends 271
highly on the quality of the available crystal structures and the 272
amount of available structural information. A more detailed 273
description of homology modeling was recently provided by Lush- 274
ington [61]. The generated hSERT and hDAT homology models 275
were then further used for molecular docking experiments. 276

3.3.2 Docking

Molecular Docking is a common method in structure-based drug 278
design to calculate the possible positions of a ligand in the binding 279
site of its target protein. A great variety of software packages is 280
available that provide different algorithms and all kinds of settings 281
[62]. In the present example, six selected compounds of a congeneric 282
series sharing a cathinone scaffold were docked into the central 283
binding site of both the homology models of hDAT and hSERT 284
with Glide 6.8 [34] from the Schrödinger release 2015-2 [33]. In 285
Glide, the protein is kept rigid during the docking process, and the 286
ligands are placed into the space between defined binding site resi- 287
dues. This setting was sufficient for our task, as we were docking 288
small compounds with respect to the outward-open binding site of 289
the transporters and we wanted to keep the side chain rotamers of 290
the homology models as close as possible to the dDAT template 291
crystal structure 4M48 at this stage. Furthermore, we restrained the 292
cationic amine function of the compounds to be placed within 2–4 Å 293
to the carbonyl oxygen of F76 in hDAT and Y95 of hSERT, because 294
several X-ray structures of related proteins with co-crystallized 295
ligands are available in the PDB showing a similar distance (for 296
further details, *see* [43]). The decision on how much flexibility 297
should be allowed during the docking process is strongly depending 298
on the availability of experimental data—which is very rich in this 299
case. Consequently, the introduced bias caused by applying docking 300
constraints was justified by the available experimental data. The 301
models and ligands were prepared in the Schrödinger suite using 302
default options (*see* **Note 5**). Once the docking output is generated, 303
which usually results in about 100 poses per ligand, a reasonable pose 304

analysis and interpretation approach are needed. The poses are 305
ranked by a specific docking score, which gives an orientation how 306
well the program was able to place the ligand into the defined 307
binding site. The docking score includes relevant energetic and steric 308
terms to achieve a most accurate placement and ranking. The Glide- 309
Score (used in this study) consists of such components (van der 310
Waals energy, Coulomb energy, lipophilic term, hydrogen-bonding 311
term, metal-binding term, as well as several rewards and penalties for 312
relevant features) [33] to predict the binding mode of the ligand 313
most accurately. However, these algorithms cannot include individ- 314
ual information such as the details known from biological experi- 315
ments about proposed binding modes for a certain target. In this 316
case a common scaffold clustering approach of all gained poses is 317
recommended [63]. In this approach, the common scaffold shared 318
by all docked ligands is extracted, and an RMSD matrix of all poses is 319
generated from these atoms. Subsequently, the clusters are calculated 320
at a defined similarity level which corresponds to the maximal dis- 321
tance within a cluster in Ångström. This helps to bundle the large 322
amount of poses into assessable bins which can be analyzed for 323
common characteristics and compared with the knowledge from 324
biological experiments in a more quantitative way. The analysis of 325
the docking study revealed certain trends explaining the observed 326
ligand selectivity of hSERT over hDAT showing slightly more nega- 327
tive overall glide scores, less steric clashes, and hydrogen bonding 328
exclusively in hDAT. 329

330

3.3.3 Molecular Dynamics Simulations

In general, molecular dynamics (MD) simulations are used to study 331
the motions of molecules over time and are therefore the method of 332
choice to characterize dynamic interactions within and between 333
biomolecules. Using such methods requires a lot of considerations 334
regarding the force field, ligand parameters, membrane and solvent 335
type, ion concentration, system size, and many more. Experimental 336
data about the respective systems and facts from profound literature 337
ideally guide these decisions. The book Molecular Modeling of 338
Proteins [64] provides an excellent review on various aspects of 339
these issues. This case study focuses on the protein-ligand interac- 340
tions between cathinone compounds and hSERT and hDAT. Inves- 341
tigating the structure-activity relationships of these compounds and 342
a subsequent docking study showed trends in the ligand selectivity 343
and provided possible binding modes. To further evaluate these 344
hypotheses, MD simulations of one compound representing the 345
previous findings (*see* [43]) were conducted. In this context, the 346
primary aim is to verify the stability of the complexes gained from 347
docking and to review the motions of the ligand inside the binding 348
site over time. MD simulations are computationally expensive and 349
need comprehensive analysis, so it is crucial to take the actual 350
research question into consideration before choosing the simulation 351

settings. For example, the simulation time to check the ligand stabil- 352
ity can be short (20 ns) if the binding mode is well defined, whereas 353
free simulation of unbinding might take up to micro or even milli- 354
seconds [65–67]. For this study, a system instability or an unfavor- 355
able starting pose of the ligand would already be observed within the 356
first nanoseconds of the simulation, because the biological data 357
provide a solid basis for our current understanding. The major 358
criteria to prove stability is a convergence of the root-mean-square 359
deviation (RMSD) of the protein and the ligand in unrestrained 360
simulations. For the protein, it is important to solely consider the 361
RMSD of the backbone atoms as the higher side chain movement 362
could hide major conformational changes in the backbone. The 363
stability of the protein-ligand interactions can be observed by inves- 364
tigating all interactions of the ligand with the protein residues over 365
the whole simulation time. This identifies the involved residues and 366
shows the duration of each interaction. Key interactions should be 367
present over the whole simulation time. The structure-based part of 368
this work was all done in the Schrödinger software suite [33]. The 369
MD simulations were prepared in Maestro 10.2 [68] and conducted 370
in 20 ns simulations with Desmond 4.2 [69]. The MD studies 371
showed that the selected poses were stable and could also confirm 372
the observed trends in the ligand selectivity profiles for the two target 373
proteins. 374

375

3.4 Summary   Designing ligands which target multiple targets with a defined 376
affinity pattern represents a powerful approach to overcome lack 377
of efficacy. With this case study, we present a holistic workflow 378
starting from data mining across public data sources and ending 379
with molecular dynamics simulations of a concrete ligand- 380
transporter complex, which revealed the stability of the ligand- 381
binding mode suggested by experimental data guided docking. As 382
parts of the protocols described are implemented in KNIME work- 383
flows, they can be easily adapted to other targets of interest. 384

385

4 Notes
386

1. In ChEMBL, more than 5000 measurement types are consid- 387
   ered including, e.g., "%max," "Activity," "Efficacy," "EC50," 388
   "Kd," and "Residual Activity" [41]. Depending on the focus of 389
   the study, these filters can be modified. 390

2. If there are several rotamers possible fitting in the observed 391
   electron density, the "right" rotamer is not necessarily the one 392
   selected by the crystallographer! High *B*-values are also an indi- 393
   cator for high flexibility. Make sure to check which rotamer is 394
   relevant for the specific research question. 395

3. When dealing with flexible proteins such as transporters, choosing the right conformation of your template structure is essential. We believe that classical inhibitors most probably bind and stabilize the outward-open conformation of the transporter and therefore hinder the transporter from adopting other conformations in the transport cycle [55]. Substrates most likely bind to the occluded transporter state as the translocation process requires among others the adaptation of an outward-occluded transporter conformation [70].

4. The DOPE score is the most widely used quality assessment parameter even though it is only optimized for soluble proteins [58]. It has been successfully used for scoring homology models of different membrane proteins [71, 72], nevertheless, it is advisable to not only rely on this parameter when modeling membrane proteins. Scores specifically optimized for membrane proteins such as the ProQM score should be taken into consideration as well for selecting the best model.

5. The Schrödinger Suite [33] offers preparation modules for both proteins and ligands. It is strongly recommended to conduct both preparation and docking procedure in the same software package as the used algorithms are compatible.

Acknowledgment

References

1. Morphy R, Rankovic Z (2005) Designed multiple ligands. An emerging drug discovery paradigm. J Med Chem 48:6523–6543

2. Hopkins AL, Mason JS, Overington JP (2006) Can we rationally design promiscuous drugs? Curr Opin Struct Biol 16:127–136

3. Peters J-U (2013) Polypharmacology – foe or friend? J Med Chem 56:8955–8971

4. Anighoro A, Bajorath J, Rastelli G (2014) Polypharmacology: challenges and opportunities in drug discovery. J Med Chem 57:7874–7887

5. Bolognesi ML, Cavalli A (2016) Multitarget drug discovery and polypharmacology. ChemMedChem 11:1190–1192

6. Lu J-J, Pan W, Hu Y-J, Wang Y-T (2012) Multi-target drugs: the trend of drug research and development. PLoS One 7:e40262

7. Harrison RK (2016) Phase II and phase III failures: 2013–2015. Nat Rev Drug Discov 15:817–818

8. Zhang W, Pei J, Lai L (2017) Computational multitarget drug design. J Chem Inf Model 57:403–412

9. Ma XH, Shi Z, Tan C, Jiang Y, Go ML, Low BC, Chen YZ (2010) In-silico approaches to multi-target drug discovery. Pharm Res 27:739–749

10. Koutsoukas A, Simms B, Kirchmair J et al (2011) From in silico target prediction to multi-target drug design: current databases, methods and applications. J Proteome 74:2554–2574

11. Lavecchia A, Cerchia C (2016) In silico methods to address polypharmacology: current

459 status, applications and future perspectives.
460 Drug Discov Today 21:288–298

461 12. Taboureau O, Baell JB, Fernández-Recio J,
462 Villoutreix BO (2012) Established and
463 emerging trends in computational drug discov-
464 ery in the structural genomics era. Chem Biol
465 19:29–41

466 13. Kuyoc-Carrillo VF, Medina-Franco JL (2014)
467 Progress in the analysis of multiple activity pro-
468 file of screening data using computational
469 approaches. Drug Dev Res 75:313–323

470 14. Ellingson SR, Smith JC, Baudry J (2014) Poly-
471 pharmacology and supercomputer-based dock-
472 ing: opportunities and challenges. Mol Simul
473 40:848–854

474 15. Reker D, Rodrigues T, Schneider P, Schneider
475 G (2014) Identifying the macromolecular tar-
476 gets of de novo-designed chemical entities
477 through self-organizing map consensus. Proc
478 Natl Acad Sci U S A 111:4067–4072

479 16. Keiser MJ, Roth BL, Armbruster BN,
480 Ernsberger P, Irwin JJ, Shoichet BK (2007)
481 Relating protein pharmacology by ligand
482 chemistry. Nat Biotechnol 25:197–206

483 17. Li H, Gao Z, Kang L et al (2006) TarFisDock:
484 a web server for identifying drug targets with
485 docking approach. Nucleic Acids Res 34:
486 W219–W224

487 18. Chen YZ, Zhi DG (2001) Ligand-protein
488 inverse docking and its potential use in the
489 computer search of protein targets of a small
490 molecule. Proteins 43:217–226

491 19. Meslamani J, Rognan D, Kellenberger E
492 (2011) sc-PDB: a database for identifying var-
493 iations and multiplicity of "druggable" binding
494 sites in proteins. Bioinformatics 27:1324–1326

495 20. Meslamani J, Li J, Sutter J, Stevens A, Bertrand
496 H-O, Rognan D (2012) Protein-ligand-based
497 pharmacophores: generation and utility assess-
498 ment in computational ligand profiling. J
499 Chem Inf Model 52:943–955

500 21. Williams AJ, Harland L, Groth P et al (2012)
501 Open PHACTS: semantic interoperability for
502 drug discovery. Drug Discov Today
503 17:1188–1198

504 22. Bento AP, Gaulton A, Hersey A et al (2014)
505 The ChEMBL bioactivity database: an update.
506 Nucleic Acids Res 42:D1083–D1090

507 23. Montanari F, Zdrazil B, Digles D, Ecker GF
508 (2016) Selectivity profiling of BCRP versus
509 P-gp inhibition: from automated collection of
510 polypharmacology data to multi-label learning.
511 J Cheminform. https://doi.org/10.1186/
512 s13321-016-0121-y

513 24. Berthold MR, Cebron N, Dill F, Di Fatta G,
514 Gabriel TR, Georg F, Meinl T, Ohl P, Sieb C,

515 Wiswedel B (2008) KNIME: The Konstanz
516 Information Miner

517 25. Pipeline pilot. http://accelrys.com/products/
518 collaborative-science/biovia-pipeline-pilot/

519 26. Montanari F, Zdrazil B (2017) How open data
520 shapes in silico transporter modeling. Mole-
521 cules. https://doi.org/10.3390/
522 molecules22030422

523 27. César-Razquin A, Snijder B, Frappier-Brinton
524 T et al (2015) A call for systematic research on
525 solute carriers. Cell 162:478–487

526 28. Kristensen AS, Andersen J, Jørgensen TN,
527 Sørensen L, Eriksen J, Loland CJ,
528 Strømgaard K, Gether U (2011) SLC6 neuro-
529 transmitter transporters: structure, function,
530 and regulation. Pharmacol Rev 63:585–640

531 29. Koldsø H, Christiansen AB, Sinning S, Schiøtt
532 B (2013) Comparative modeling of the human
533 monoamine transporters: similarities in sub-
534 strate binding. ACS Chem Neurosci
535 4:295–309

536 30. Sitte HH, Freissmuth M (2015) Ampheta-
537 mines, new psychoactive drugs and the mono-
538 amine transporter cycle. Trends Pharmacol Sci
539 36:41–50

540 31. Schultz W (2010) Dopamine signals for reward
541 value and risk: basic and recent data. Behav
542 Brain Funct 6:24

543 32. Webb B, Sali A (2016) Comparative protein
544 structure modeling using MODELLER. Curr
545 Protoc Protein Sci:2.9.1–2.9.37

546 33. (2015) Schrödinger Release 2015-2. Schrödin-
547 ger, LLC, New York, NY

548 34. Halgren TA, Murphy RB, Friesner RA, Beard
549 HS, Frye LL, Pollard WT, Banks JL (2004)
550 Glide: a new approach for rapid, accurate dock-
551 ing and scoring. 2. Enrichment factors in data-
552 base screening. J Med Chem 47:1750–1759

553 35. Molecular Operating Environment (MOE),
554 2013.08. Chemical Computing Group Inc.,
555 Montreal, Canada

556 36. Bowers K, Chow E, Xu H, et al (2006) Scalable
557 algorithms for molecular dynamics simulations
558 on commodity clusters. In: ACM/IEEE SC
559 2006 conference (SC'06). https://doi.org/
560 10.1109/sc.2006.54

561 37. Desmond Molecular Dynamics System, D. E.
562 Shaw Research, New York, NY, 2017.
563 Maestro-Desmond interoperability tools,
564 Schrödinger, New York, NY, 2017

565 38. Wishart DS, Feunang YD, Guo AC et al (2018)
566 DrugBank 5.0: a major update to the Drug-
567 Bank database for 2018. Nucleic Acids Res 46:
568 D1074–D1082

569 39. Aoki-Kinoshita KF, Kanehisa M (2007) KEGG
570 primer: an introduction to pathway analysis

using KEGG. NCI Nature Pathway Interaction Database. https://doi.org/10.1038/pid.2007.2

40. Kramer C, Fuchs JE, Whitebread S, Gedeck P, Liedl KR (2014) Matched molecular pair analysis: significance and the impact of experimental uncertainty. J Med Chem 57:3786–3802

41. Hu Y, Bajorath J (2014) Influence of search parameters and criteria on compound selection, promiscuity, and pan assay interference characteristics. J Chem Inf Model 54:3056–3066

42. Hu Y, Bajorath J (2015) Structural and activity profile relationships between drug scaffolds. AAPS J 17:609–619

43. Zdrazil B, Hellsberg E, Viereck M, Ecker GF (2016) From linked open data to molecular interaction: studying selectivity trends for ligands of the human serotonin and dopamine transporter. Medchemcomm 7:1819–1831

44. Bemis GW, Murcko MA (1996) The properties of known drugs. 1. Molecular frameworks. J Med Chem 39:2887–2893

45. myExperiment – Workflows – KNIME workflow without hERG labels included from Zdrazil et al., MedChemComm, 2016: "From linked open data to molecular interaction: studying selectivity trends for ligands of the human serotonin and dopamine transporter" (Barbara Zdrazil) [KNIME Workflow]. https://www.myexperiment.org/workflows/4911.html. Accessed 5 Feb 2018

46. Berman HM (2000) The Protein Data Bank. Nucleic Acids Res 28:235–242

47. Humphrey W, Dalke A, Schulten K (1996) VMD: visual molecular dynamics. J Mol Graph 14:33–38

48. PyMOL, The PyMOL Molecular Graphics System, version 2.0. Schrödinger, LLC

49. Conroy MJ, Sehnal D, Deshpande M, Svobodova R, Mir S, Berka K, Midlik A, Velankar S, Koca J (2017) LiteMol: web-based three-dimensional visualization of macromolecular structure data. Acta Crystallogr A 73:C669

50. Sehnal D, Deshpande M, Vařeková RS, Mir S, Berka K, Midlik A, Pravda L, Velankar S, Koča J (2017) LiteMol suite: interactive web-based visualization of large-scale macromolecular structure data. Nat Methods 14:1121–1122

51. Mir S, Alhroub Y, Anyango S et al (2018) PDBe: towards reusable data delivery infrastructure at protein data bank in Europe. Nucleic Acids Res 46:D486–D492

52. Pozharski E, Weichenberger CX, Rupp B (2013) Techniques, tools and best practices for ligand electron-density analysis and results from their application to deposited crystal structures. Acta Crystallogr D Biol Crystallogr 69:150–167

53. Emsley P, Lohkamp B, Scott WG, Cowtan K (2010) Features and development of Coot. Acta Crystallogr D Biol Crystallogr 66:486–501

54. Qu X, Swanson R, Day R, Tsai J (2009) A guide to template based structure prediction. Curr Protein Pept Sci 10:270–285

55. Singh SK, Piscitelli CL, Yamashita A, Gouaux E (2008) A competitive inhibitor traps LeuT in an open-to-out conformation. Science 322:1655–1661

56. Penmatsa A, Wang KH, Gouaux E (2013) X-ray structure of dopamine transporter elucidates antidepressant mechanism. Nature 503:85–90

57. Thompson JD, Gibson TJ, Higgins DG (2002) Multiple sequence alignment using ClustalW and ClustalX. Curr Protoc Bioinformatics. Chapter 2:Unit 2.3

58. Shen M-Y, Sali A (2006) Statistical potential for assessment and prediction of protein structures. Protein Sci 15:2507–2524

59. Laskowski RA, MacArthur MW, Moss DS, Thornton JM (1993) PROCHECK: a program to check the stereochemical quality of protein structures. J Appl Crystallogr 26:283–291

60. Ray A, Lindahl E, Wallner B (2010) Model quality assessment for membrane proteins. Bioinformatics 26:3067–3074

61. Lushington GH (2015) Comparative modeling of proteins. Methods Mol Biol 1215:309–330

62. Chen Y-C (2015) Beware of docking! Trends Pharmacol Sci 36:78–95

63. Richter L, de Graaf C, Sieghart W, Varagic Z, Mörzinger M, de Esch IJP, Ecker GF, Ernst M (2012) Diazepam-bound GABAA receptor models identify new benzodiazepine binding-site ligands. Nat Chem Biol 8:455–464

64. Kukol A (2017) Molecular modeling of proteins. Humana Press, New York

65. Schuetz DA, de Witte WEA, Wong YC et al (2017) Kinetics for drug discovery: an industry-driven effort to target drug residence time. Drug Discov Today 22:896–911

66. De Vivo M, Masetti M, Bottegoni G, Cavalli A (2016) Role of molecular dynamics and related methods in drug discovery. J Med Chem 59:4035–4061

67. Huang D, Caflisch A (2011) The free energy landscape of small molecule unbinding. PLoS Comput Biol 7:e1002002

683 68. Schrödinger Release 2015-2: Maestro, version
684   10.2. Schrödinger, LLC, New York, NY
685 69. Schrödinger Release 2015-2: Desmond Molec-
686   ular Dynamics System, version 4.2. Schrödin-
687   ger, LLC, New York, NY
688 70. Wang H, Gouaux E (2012) Substrate binds in
689   the S1 site of the F253A mutant of LeuT, a
690   neurotransmitter sodium symporter homo-
691   logue. EMBO Rep 13:861–866

71. Jurik A, Zdrazil B, Holy M, Stockner T, Sitte   692
    HH, Ecker GF (2015) A binding mode   693
    hypothesis of tiagabine confirms liothyronine   694
    effect on γ-aminobutyric acid transporter   695
    1 (GAT1). J Med Chem 58:2149–2158   696
72. Saha K (2015) "Second generation" mephe-   697
    drone analogs, 4-MEC and 4-MePPP, differ-   698
    entially affect monoamine transporter   699
    function. Intrinsic Activity 3:A2.18   700

II. Result and Discussion

3. Ligand-based studies

Quantitative structure-activity relationship (QSAR) methods have been highly successful in modeling physicochemical and biological properties of small molecules. They facilitate screening of millions of compounds with a goal to accurately distinguish active compounds from inactive compounds. These methods are also beneficial in understanding the change in activity of a molecule due to changes in its structure. Besides being a low-cost approach, modeling of large chemical libraries has become highly productive with QSAR modeling. Additionally, it is possible to predict properties of non-existing and non-synthesized compounds [180, 181]. These factors significantly affect the success of drug discovery and development.

On the other hand, unfavourable safety, efficacy and pharmacokinetic profiles have been the major reasons contributing to the failure of the majority of candidate drugs, thereby hampering the success of drug discovery projects to incur huge burden on pharmaceutical companies [182]. Therefore, early identification of lead compounds with unacceptable ADMET profile is highly essential. In this respect, data mining techniques employing machine- learning methods (e.g., support vector machines and decision trees) are highly essential to construct models using these large datasets and establish a relationship between compounds and observed activity. However, the non-balanced and diversified nature of chemical datasets present a challenging problem in the successful application of these techniques and need to be dealt with.

3.1 Comparing the performance of meta-classifiers – A case study on a set of imbalanced data sets relevant for prediction of liver toxicity

Sankalp Jain, Eleni Kotsampasakou, and Gerhard F. Ecker

In the following study, we evaluated the performance of seven distinct meta-classifiers namely 1) Bagging, 2) Under-sampled stratified bagging, 3) Cost-sensitive classifier, 4) MetaCost, 5) Threshold Selection, 6) SMOTE and 7) ClassBalancer on four datasets that are directly (cholestasis) or indirectly (*via* inhibition of organic anion transporting polypeptide 1B1 and 1B3) related to hepatotoxicity with varying degree of class imbalance. We used three different sets of molecular descriptors for model development. From the investigated meta-classifiers, Stratified Bagging provided the highest balanced accuracies while MetaCost and CostSensitiveClassifier achieved better sensitivity. The findings are expected to improve the understanding and selection of an optimal strategy to handle imbalanced datasets.

E. Kotsampasakou compiled the datasets, generated the models developed in WEKA (for Random Forest, Cost-sensitive classifier, MetaCost, Threshold Selection, SMOTE and ClassBalancer), did the statistical testing. **S. Jain** performed the modeling on OCHEM for Bagging and Stratified Bagging, wrote the R code to generate the plots and wrote the manuscript. G.F. Ecker supervised the work and revised the manuscript. All three authors participated in the original design of the study.

CrossMark

Comparing the performance of meta-classifiers—a case study on selected imbalanced data sets relevant for prediction of liver toxicity

Sankalp Jain[1] · Eleni Kotsampasakou[1,2] · Gerhard F. Ecker[1]

Abstract

Cheminformatics datasets used in classification problems, especially those related to biological or physicochemical properties, are often imbalanced. This presents a major challenge in development of in silico prediction models, as the traditional machine learning algorithms are known to work best on balanced datasets. The class imbalance introduces a bias in the performance of these algorithms due to their preference towards the majority class. Here, we present a comparison of the performance of seven different meta-classifiers for their ability to handle imbalanced datasets, whereby Random Forest is used as base-classifier. Four different datasets that are directly (cholestasis) or indirectly (via inhibition of organic anion transporting polypeptide 1B1 and 1B3) related to liver toxicity were chosen for this purpose. The imbalance ratio in these datasets ranges between 4:1 and 20:1 for negative and positive classes, respectively. Three different sets of molecular descriptors for model development were used, and their performance was assessed in 10-fold cross-validation and on an independent validation set. Stratified bagging, MetaCost and CostSensitiveClassifier were found to be the best performing among all the methods. While MetaCost and CostSensitiveClassifier provided better sensitivity values, Stratified Bagging resulted in high balanced accuracies.

Graphical Abstract
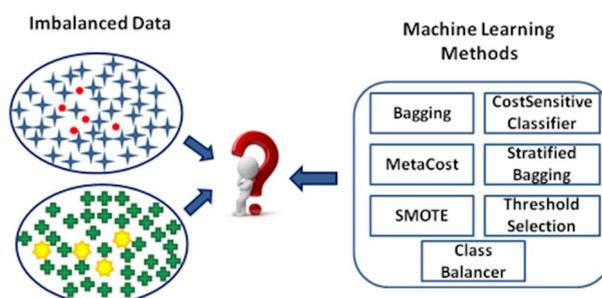
Abbreviations

AUC	Area under the ROC curve
HTS	High throughput screening
MCC	Matthews correlation coefficient
OATP1B1	Organic anion transporting polypeptide 1B1
OATP1B3	Anion transporting polypeptide 1B3
RF	Random Forest
sd	Standard deviation

Sankalp Jain and Eleni Kotsampasakou have contributed equally to this manuscript.

Extended author information available on the last page of the article

Springer

| SMOTE | Synthetic minority over-sampling technique |
| SVM | Support vector machines |

Introduction

A wide range of classification and regression methods have been applied in QSAR studies. However, many classification methods assume that datasets are balanced in terms of the number of instances of each class and thus give equal importance to all classes, often resulting in classification models of poor accuracy [1, 2]. A major problem that arises in this context is class imbalance, i.e. the number of instances of one class substantially differ from those of the other classes. Especially in the field of drug discovery, imbalanced datasets [2–4] need to be frequently dealt with [2]. Characteristically, a classifier developed on an imbalanced data set shows a low error rate for the majority class and a high error rate for the minority class [5, 6]. Nevertheless, a few studies pointed out that the class imbalance is not a main obstacle in learning [7, 8], and several methods have been developed to address this issue. These methods can be broadly divided into (1) data-oriented/re-sampling techniques; (2) algorithm-oriented methods; and (3) combinatorial/ensemble/hybrid techniques [2, 3, 7, 9, 10].

Several studies compared classifiers that handle imbalanced datasets. Schierz et al. [11] compared four WEKA classifiers (Naïve Bayes, SVM, Random Forest and J48 tree) and reported SVM and J48 to be the best performing for bioassay datasets. Lin and Chen in 2013 found SVM threshold adjustment as the best performing classifier (among linear discriminant analysis, Random Forest, SVM and SVM-threshold adjustment) to deal with imbalanced HTS datasets [9]. Later, Zakarov et al. used under-sampling and threshold selection techniques on several imbalanced PubChem HTS assays to test and develop robust QSAR models in the program GUSAR [12]. In a recent study, Razzaghi et al. reported multilevel SVM-based algorithms to outperform conventional SVM, weighted SVM, neural networks, linear regression, Naïve Bayes and C4.5 tree using public benchmark datasets having imbalanced classes and missing values and real data in health applications [13].

A comprehensive comparison of the performance of different meta-classifiers on datasets with different levels of class imbalance, which would provide guidance for choosing the appropriate method for an imbalanced dataset, has not been attempted so far. Herein, we evaluated the performance of seven distinct meta-classifiers from the three aforementioned categories on four datasets from the toxicology domain. The imbalance ratio of the datasets ranges from 1:4 to 1:20 for the positive and the negative class, respectively. The meta-classifiers were applied to build classification models based on three different sets of descriptors.

Considering its wide applicability in modeling imbalanced datasets, Random Forest was used as the common base-classifier for all models [14–18]. Further, we discuss the reasons behind the superior performance of certain meta-classifiers in comparison to the others while explaining their intrinsic limitations.

Methods

Training datasets

Four different datasets from the biomedical sciences domain were used in this study. Two of these are the OATP1B1 and OATP1B3 inhibition datasets consisting of 1708 and 1725 compounds, respectively. Both were compiled and used in our previous study that reported classification models for OATP1B1 and 1B3 inhibition [19]. The other two datasets come from the toxicology domain and are related to drug-induced cholestasis for human data and animal data which comprise 1766 and 1578 compounds, respectively. Both datasets were published in a previous study that reported computational models for hepatotoxicity and other liver toxicity endpoints [20].

External test datasets

The external test sets for OATP1B1 and 1B3 inhibition from our previous study served as test datasets in this study [19]. The test set for human cholestasis was compiled in two stages from two previous studies [21]. The positives for human cholestasis were compiled from literature [22–25] and from the SIDER v2 database [26, 27]. As cholestasis is one of the three types of drug induced liver injury (DILI), and the compounds that are negative for DILI will also be negative for cholestasis, the negatives for drug-induced liver injury compiled in a previous study [21] were used as negatives for cholestasis. Overall, the external human cholestasis dataset consisted of 231 compounds. No data were available for animal cholestasis to be used as an external test dataset. The composition and degree of class imbalance of each training and test dataset is presented in Table 1.

The chemotypes in the datasets were curated using the following protocol:

– Removed all inorganic compounds according to chemical formula in MOE 2014.09 [28].
– Removed salts and compounds containing metals and/or rare or special atoms.
– Standardized chemical structures using Francis Atkinson Standardiser tool [29].
– Removed duplicates and permanently charged compounds using MOE 2014.09 [28].

Table 1 An overview of the training and test datasets used in this study

Dataset name	Total number of compounds	Number of positives	Number of negatives	Imbalance ratio (negatives: positives)	Source
OATP1B1 inhibition training	1708	190	1518	8:1	Kotsampasakou et al. [19]
OATP1B1 inhibition testing	201	64	137	2:1	Kotsampasakou et al. [19]
OATP1B3 inhibition training	1725	124	1601	13:1	Kotsampasakou et al. [19]
OATP1B3 inhibition testing	209	40	169	4:1	Kotsampasakou et al. [19]
Cholestasis human training	1766	347	1419	4:1	Mulliner et al. [20]
Cholestasis human testing	231	53	178	3:1	Kotsampasakou et al. [21]
Cholestasis animal training	1578	75	1503	20:1	Mulliner et al. [20]

– 3D structures were then generated using CORINA (version 3.4) [30], and energy minimized with MOE 2014.09 [28], using default settings (Forcefield MMF94x, gradient 0.05 RMS kcal/mol/A^2, preserving chirality).

Molecular descriptors

Three different sets of descriptors were calculated for each of the datasets:

1. All 2D MOE [28] descriptors (192 descriptors in total).
2. ECFP6 fingerprints (1024 bits) calculated with RDKit [31].
3. MACCS fingerprints (166 bits), calculated with PaDEL software [32].

Machine learning methods

Random Forest [33] implemented in the WEKA software suite [34, 35] was used as a base-classifier along with all the meta-learning methods evaluated in this study. The number of trees was arbitrarily set to 100 (default), since it has been shown that the optimal number of trees is usually 64–128, while further increasing the number of trees does not necessarily improve the model's performance [36]. The following meta-classifiers were investigated: (1) Bagging, (2) Under-sampled stratified bagging, (3) Cost-sensitive classifier, (4) MetaCost, (5) Threshold Selection, (6) SMOTE and (7) ClassBalancer.

1. *Bagging* (*Bootstrap AGGregatING*) [37] is a machine learning technique that is based on an ensemble of models developed using multiple training datasets sampled from the original training set. It calculates several models and averages them to produce a final ensemble model [37]. A traditional bagging method generates multiple copies of the training set by selecting the molecules with replacement from training set in a random fashion.

Because of random sampling, about 37% of the molecules are not selected and left out in each run. These samples create the "out-of-the-bag" sets, which are used for testing the performance of the final model. A total of 64 models were used for our analysis, since it was shown in an earlier study by Tetko et al. [38] that larger numbers of models per ensemble (e.g. 128, 256, 512 and 1024) did not significantly increase the balanced accuracy of models.

2. *Under-sampled stratified bagging* [2, 8, 38] In this method, the total bagging training set size is double the number of the minority class molecules. Although a small set of samples was selected each time, the majority of molecules contributed to the overall bagging procedure, since the datasets were generated randomly. The performance of the developed models is tested with molecules from the "out-of-the-bag" set [38]. Since only one way of stratified learning, i.e., under-sampling stratified bagging, was used in the study, we refer to it as "Stratified Bagging".

Bagging and Stratified Bagging were used as implemented in the Online Chemical Modeling Environment (OCHEM) [39, 40]. For other meta-classifiers, WEKA(v. 3-7-12) [34, 35] was used.

3. *Cost sensitive classifier* [2–4, 10, 11] is a meta-classifier that renders the base classifier cost-sensitive. Two methods can be used to introduce cost-sensitivity: (i) reweighting training instances according to the total cost assigned to each class, i.e. the weights are applied during learning, or; (ii) predicting the class with minimum expected misclassification cost (rather than the most likely class), i.e. the "cost-sensitive" is introduced in the test phase. In our case, the cost sensitivity was introduced according to method (i) using the CostSensitiveClassifier from the set of meta-classifiers of the WEKA software [34, 35].

4. *MetaCost* [41] is another application that provides the methodology to perform cost-sensitive training of a classifier in a generalized meta-learning manner independent of the underlying classifier. It is a combination of Cost-

sensitive meta-classifier and Bagging [37]. The algorithm uses class-relabeling, i.e. it modifies the original training set by changing the class labels to the so-called "optimal classes". The classifier is then trained on this modified training set, which results in having the error rate minimized according to the cost matrix provided to the MetaCost algorithm. This implementation uses all bagging iterations when reclassifying training data. MetaCost is advantageous as, unlike CostSensitiveClassifier, a single cost-sensitive classifier of the base learner is generated, thus giving the benefits of fast classification and interpretable output (if the base learner itself is interpretable). MetaCost further differs from traditional bagging by the fact that the number of examples in each resample may be smaller than the training set size. This variation improves the efficiency of the algorithm. More details about the method can be found in [41].

For both CostSensitiveClassifier and MetaCost, several trials of different cost matrices were applied, until a satisfactory outcome was retrieved.

5. *ThresholdSelector* [42] is a meta-classifier implemented in WEKA [34, 35] that sets a threshold on the probability output of a base-classifier. Threshold adjustment for the classifier's decision is one of the methods used for dealing with imbalanced datasets [2, 43]. By default, the WEKA probability threshold to assign a class is 0.5, i.e. if an instance is attributed with a probability of equal or less than 0.5, it is classified as negative for the respective class, while if it is greater than 0.5, the instance is classified as positive. For our study, the optimal threshold was selected automatically by the meta-classifier by applying internal fivefold cross validation to optimize the threshold according to FMeasure (Eq. 7), a measure of a model's accuracy which considers both precision and sensitivity [44].

6. *SMOTE* [45] (*Synthetic minority over-sampling technique*) increases the minority class by generating new "synthetic" instances based on its number of nearest neighbours. SMOTE, as implemented in WEKA, was used to generate synthetic examples. For our study, five nearest neighbours of a real existing instance (minority class) were used to compute a new synthetic one. For different datasets, different percentages of SMOTE instances were created, which can be found in the supplementary information (Table S1). The complete algorithm is explained in [45].

7. *ClassBalancer* [34, 35, 46] reweights the instances so that the sum of weights for all classes of instances in the data is the same, i.e. the total sum of weights across all instances is maintained. This is an additional way to treat class imbalance, unlike CostSensitiveClassifier or MetaCost, which try to minimize the total misclassification cost.

With respect to parameters, not for all classifiers a parameter optimization was performed. For instance, no parameters were adjusted for ClassBalancer since it automatically reassigns weights to the instances in the dataset such that each class has the same total weight [46]. For Bagging and Stratified Bagging, the only parameter to optimize would be the number of bags. In our case, the number of bags was adjusted to 64 as a previous study [38] suggests that generation of 64 models provides satisfactory results without exponentially increasing the computational cost. In case of ThresholdSelector, an optimal threshold was selected automatically via fivefold cross-validation before selecting the final model on the basis of FMeasure. For both CostSensitiveClassifier and MetaCost, the cost for misclassification was initially applied in accordance with the imbalance ratio, which, in case it did not provide a sensitivity of at least 0.5, was further increased to arrive at the final model. In case of SMOTE, similar principles were applied: initially, the number of the synthetic instances created was set to a number that balances the two classes. If insufficient, it was further increased until no further improvement in sensitivity (with no reduction in specificity) was observed. The detailed parameter settings of the best performing models for each method are provided in the supplementary material (Table S1).

Validation

All models were evaluated in a 10-fold cross-validation followed by an external validation performed on independent test sets, except for Bagging and Stratified Bagging. For Bagging and Stratified Bagging, since multiple training datasets were generated by selecting the molecules with replacement from training set in a random fashion, this leaves out about 37% of the instances in each run. Therefore, these molecules that constitute the 'out-of-the-bag' sets are later used for testing the performance of the final model.

Model performance assessment: selection of the optimal method

Prior to identifying the best performing method, an optimal model for each meta-classifier was selected. The best parameters for the model were selected using linear search (as explained in the "Methods" section). For all models, different performance measures including sensitivity (Eq. 1), specificity (Eq. 2), accuracy (Eq. 3), balanced accuracy (Eq. 4), Matthews correlation coefficient (MCC, Eq. 5), area under the curve (AUC) and precision (Eq. 6) were calculated. A model was considered eligible for selection if the 10-fold cross-validation provided a sensitivity value of at least 0.5 and a specificity value not less than 0.5. As the datasets are relevant to different toxicological endpoints, sensitivity was

considered more important. For a highly imbalanced dataset, accuracy may be misleading. Therefore we considered balanced accuracy (which considers both sensitivity and specificity) as a more appropriate performance measure to compare different classifiers for their ability to handle imbalanced datasets. If two models provided the same sensitivity, the model that demonstrated higher balanced accuracy was prioritized for selection. Furthermore, 20 iterations were performed by varying the seed for cross validation [by assigning values from 1 (default) to 20]. For Bagging and Stratified Bagging, the 20 iterations were performed by changing the random seed for the Random Forest generation by assigning values from 1 (default) to 20. After cross-validation, average values for different performance measures were calculated and compared. The best method was then evaluated by performing a statistical t-test in R [47], as well as on the basis of the performance on external test sets. The individual settings used in selecting the best model for each meta-classifier can be found in the supplementary information (Table S1).

$$Sensitivity = \frac{TP}{(TP + FN)} \quad (1)$$

$$Specificity = \frac{TN}{(TN + FP)} \quad (2)$$

$$Accuracy = \frac{(TP + TN)}{(TP + FP + TN + FN)} \quad (3)$$

$$Balanced\ Accuracy = \frac{1}{2}\left(\frac{(TP)}{(TP + NP)} + \frac{(TN)}{(TN + FP)}\right) \quad (4)$$

$$MCC = \frac{\{(TP \times TN) - (FP \times FN)\}}{\{(TP + FP) \times (TP + FN) \times (TN + FP) \times (TN + FN)\}^{1/2}} \quad (5)$$

$$Precision = \frac{(TP)}{(TP + FP)} \quad (6)$$

$$FMeasure = \frac{2TP}{(2TP + FP + FN)} \quad (7)$$

TP: true positives; TN: true negatives; FP: false positives; FN: false negatives.

Results and discussion

Tables S2–S5 in the supplementary material report the performance measures for predictions on all datasets used in this study. The performance values of the base-classifier (Random Forest) are also reported to facilitate a comparison with the investigated methods. For each dataset, the mean and the standard deviation values of performance of the best performing models (based on 20 iterations) were calculated and are reported in Tables S6–S9 (supplementary material). Figure 1a–c, Figure S1(a–d) in the supplementary material provide a comparison of performances of different meta-classifiers on the three test datasets (no test set available for animal cholestasis) and four training sets respectively.

Irrespective of the dataset and the descriptor set used, Random Forest was found to be the weakest performing classifier as anticipated. Except on the test dataset for human cholestasis, Random Forest alone did not yield a sensitivity greater than 0.5, which indicates that assistance of a meta-classifier indeed consistently improves performance when handling imbalanced datasets. Among the Meta-Classifier based methods, bagging provided the lowest performance. A simple reason behind the failure of Bagging is that it only
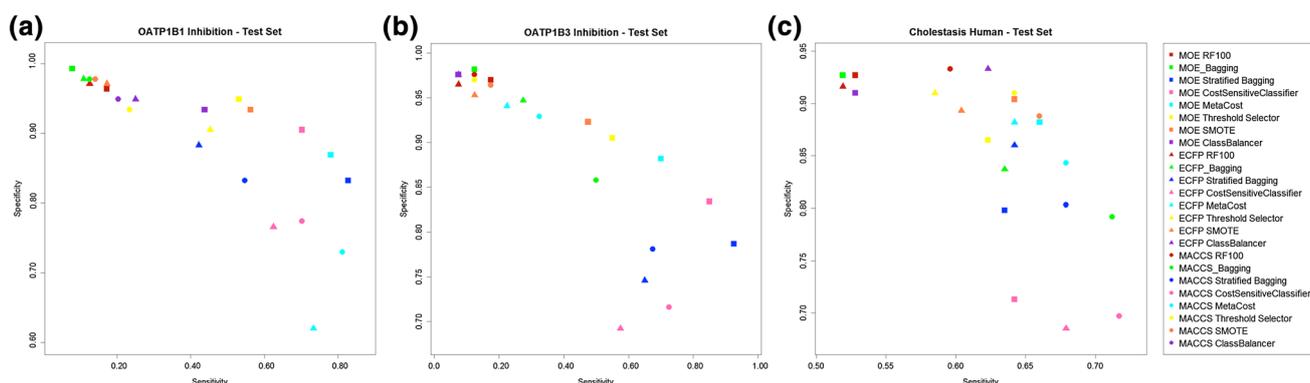
Fig. 1 Comparison of performances of different meta-classifiers on test sets **a** OATP1B1 inhibition **b** OATP1B3 inhibition **c** human cholestasis. *x-axis* corresponds to the sensitivity and on the *y-axis* is the specificity. The squares correspond to MOE descriptors, the triangles correspond to ECFP6 fingerprints and the circles correspond to MACCS fingerprints. Each classifier is depicted in a different color: red for RF standalone, green for Bagging, blue for Stratified Bagging, dark pink for CostSensitiveClassifier, cyan for MetaCost, yellow for ThresholdSelector, orange for SMOTE and dark violet for ClassBalancer. Please note that the scaling for the two axes are different

does resampling without any effort to balance or weight the two classes.

Threshold Selection was frequently found to be among the good performing methods. In many cases, this classifier could handle imbalance very well. However, the sensitivity measures were poor in comparison to other classifiers. This could be due to the fact that the thresholds were selected on the basis of FMeasure, as accuracy and specificity are not suitable due to the high impact of the majority class. If the selection of best models is done purely on the basis of sensitivity, this classifier yields very good sensitivity values (0.8–1.0), however with a radical decrease in specificity (0.2–0). Notably, Threshold Selection provided better results in combination with a second meta-classifier. But since the aim of the study was to compare the classifiers individually, this trend was not investigated further.

Stratified Bagging, CostSensitiveClassifier and Meta-Cost were consistently the best performing classifiers in both cross-validation and test set validation for all the datasets (see Fig. 1, Figure S1 in the supplementary material). Further, the t-test on the basis of 95% confidence interval (exact p-values not shown here) indicated a statistically significant difference in performance between the selected methods (meta-classifiers). The statistical test was performed pair-wise for all the obtained performance measures, with more stress on sensitivity and balanced accuracy. Both MetaCost and CostSensitiveClassifier tended to yield higher sensitivities while Stratified Bagging, on the other hand, was found to be superior in terms of MCC, balanced accuracy and AUC. An advantage of Stratified Bagging is that it is a straightforward method with only one parameter to optimize, i.e. the number of bags. On the other hand, cost-sensitive approaches tend to give more weight to sensitivity when needed, which is an advantage for toxicity prediction. Although both methods provided comparable performances, the cost that had to be applied was greater in case of Cost-SensitiveClassifier in comparison to MetaCost. This is due to the fact that the latter is a hybrid classifier which combines Bagging with the application of a cost, thus equilibrating the dataset more easily. It should further be noted that the computational cost for MetaCost is higher than that for CostSensitiveClassifier. On the other hand, Stratified Bagging is not computationally demanding (for the optimal parameter of 64 bags). Since each bag is double the size of the minority class, the calculation of models using Stratified Bagging requires less computational time, compared to the models built using Bagging (the bags are of the same size as the training set) and MetaCost (includes both bagging and weighting).

SMOTE and ClassBalancer were only in a few cases able to provide a sensitivity of at least 0.5 in both cross-validation and test set evaluation. Considering its reputation in handling such problems, the poor performance of SMOTE was quite surprising. We assume that the small size of the datasets could be the primary reason behind SMOTE's poor performance. The datasets used in this study are much smaller in size compared to the HTS datasets in which the minority class has enough instances for SMOTE to generate synthetic instances, although the overall imbalance ratio is typically in the range of 100:1 [12, 45, 48].

With respect to the different sets of descriptors used, the performance of the classifiers on different datasets remained almost the same. Of all the descriptors, 2D MOE descriptors and MACCS fingerprints provided the best performance across many of the datasets, while ECFP6 fingerprints consistently performed lower. Considering the amount of information encoded in ECFP6 (1024 bits) in comparison to MACCS fingerprints (166 bits) and the MOE descriptors, it might be assumed that the poor performance of ECFP6 is subject to the individual datasets in this study. This also highlights the fact that sometimes simple set of descriptors could provide better results than complex and highly populated descriptors. Moreover, in other recent studies [49–51] different descriptor and fingerprint combinations did not demonstrate significant differences in performance.

Overall, the best classifiers performed well regardless of the type of data (toxicity endpoint or a general or specific in vitro endpoint), the type and number of descriptor sets used, or the degree of class imbalance. However, there were instances where a dataset related to in vivo toxicity (animal cholestasis) could not be successfully handled by the best classifiers. Finally, highly sophisticated meta-classifiers such as Stratified Bagging and MetaCost, that combine re-sampling and a way to weight the two classes, performed in principle better than Bagging and ClassBalancer.

Conclusions

In this study, we compared the performance of seven different meta-classifiers for their ability to handle imbalanced datasets. We demonstrated that, for all datasets used in the study, Stratified Bagging performed at least as good as cost-sensitive approaches such as MetaCost and CostSensitive-Classifier and in most cases outperformed them. Random Forest (as a standalone classifier) and Bagging were unable to address the imbalance issue. Interestingly, the choice of descriptors did not play a substantial role in ranking the performance of different classifiers. Thus, considering that Stratified Bagging can be directly used in combination with any machine-learning method without parameter optimization, a general recommendation for handling imbalanced datasets is to wrap the modeling process in the stratified bagging loop. However, one should also consider the computational cost, as extensive re-sampling can be computationally expensive. Therefore, a method that balances between the

complexity of the algorithm and computational cost would be an ideal choice to obtain optimal results.

References

1. Kotsiantis SB (2008) Handling imbalanced data sets with a modification of Decorate algorithm. Int J Comput Appl Technol 33:91–98. https://doi.org/10.1504/IJCAT.2008.021931
2. Kotsiantis S, Kanellopoulos D, Pintelas P (2006) Handling imbalanced datasets: a review. GESTS Int Trans Comput Sci Eng 30(1):25–36
3. Ali A, Shamsuddin SM, Ralescu AL (2015) Classification with class imbalance problem: a review. Int J Adv Soft Comput Appl 7:176–204
4. López V, Fernández A, Moreno-Torres JG, Herrera F (2012) Analysis of preprocessing vs. cost-sensitive learning for imbalanced classification: open problems on intrinsic data characteristics. Expert Syst Appl 39:6585–6608. https://doi.org/10.1016/j.eswa.2011.12.043
5. Qiao X, Liu Y (2009) Adaptive weighted learning for unbalanced multicategory classification. Biometrics 65:159–168. https://doi.org/10.1111/j.1541-0420.2008.01017.x
6. Fernández A, Jesus MJ, del Herrera F (2010) Multi-class imbalanced data-sets with Linguistic fuzzy rule based classification systems based on pairwise learning. In: Hüllermeier E, Kruse R, Hoffmann F (eds) Computational intelligence for knowledge-based systems design. Springer, Berlin, pp 89–98
7. Galar M, Fernández A, Barrenechea E et al (2012) A review on ensembles for the class imbalance problem: bagging-, boosting-, and hybrid-based approaches. IEEE Trans Syst Man Cybern Part C 42:463–484. https://doi.org/10.1109/TSMCC.2011.2161285
8. He H, Garcia EA (2009) Learning from imbalanced data. IEEE Trans Knowl Data Eng 21:1263–1284. https://doi.org/10.1109/TKDE.2008.239
9. Lin W-J, Chen JJ (2013) Class-imbalanced classifiers for high-dimensional data. Brief Bioinform 14:13–26. https://doi.org/10.1093/bib/bbs006
10. Garcia V, Sanchez JS, Mollineda RA, Alejo R, Sotoca JM (2007) The class imbalance problem in pattern classification and learning. In: II Congreso Español de Informática, Tamida, Saragossa, Spain, pp 283–291
11. Schierz AC (2009) Virtual screening of bioassay data. J Cheminform 1:21. https://doi.org/10.1186/1758-2946-1-21
12. Zakharov AV, Peach ML, Sitzmann M, Nicklaus MC (2014) QSAR modeling of imbalanced high-throughput screening data in PubChem. J Chem Inf Model 54:705–712. https://doi.org/10.1021/ci400737s
13. Razzaghi T, Roderick O, Safro I, Marko N (2016) Multilevel weighted support vector machine for classification on healthcare data with missing values. PLoS ONE 11:e0155119. https://doi.org/10.1371/journal.pone.0155119
14. Schlieker L, Telaar A, Lueking A et al (2017) Multivariate binary classification of imbalanced datasets-A case study based on high-dimensional multiplex autoimmune assay data. Biom J Biom Z 59:948–966. https://doi.org/10.1002/bimj.201600207
15. Chen J, Tang YY, Fang B, Guo C (2012) In silico prediction of toxic action mechanisms of phenols for imbalanced data with Random Forest learner. J Mol Graph Model 35:21–27. https://doi.org/10.1016/j.jmgm.2012.01.002
16. Khalilia M, Chakraborty S, Popescu M (2011) Predicting disease risks from highly imbalanced data using Random Forest. BMC Med Inform Decis Mak 11:51. https://doi.org/10.1186/1472-6947-11-51
17. Barta G (2016) Identifying biological pathway interrupting toxins using multi-tree ensembles. Front Environ Sci. https://doi.org/10.3389/fenvs.2016.00052
18. Koutsoukas A, St. Amand J, Mishra M, Huan J (2016) Predictive toxicology: modeling chemical induced toxicological response combining circular fingerprints with Random Forest and support vector machine. Front Environ Sci. https://doi.org/10.3389/fenvs.2016.00011
19. Kotsampasakou E, Brenner S, Jäger W, Ecker GF (2015) Identification of novel inhibitors of organic anion transporting polypeptides 1B1 and 1B3 (OATP1B1 and OATP1B3) using a consensus vote of six classification models. Mol Pharm 12:4395–4404. https://doi.org/10.1021/acs.molpharmaceut.5b00583
20. Mulliner D, Schmidt F, Stolte M et al (2016) Computational models for human and animal hepatotoxicity with a global application scope. Chem Res Toxicol 29:757–767. https://doi.org/10.1021/acs.chemrestox.5b00465
21. Kotsampasakou E, Ecker GF (2017) Predicting drug-induced cholestasis with the help of hepatic transporters—an in silico modeling approach. J Chem Inf Model 57:608–615. https://doi.org/10.1021/acs.jcim.6b00518
22. Kullak-Ublick G (2003) Drug-induced cholestatic liver disease. In: Trauner M, Jansen P, (eds) Mol Pathog Cholestasis. Springer, New York, pp 271–280
23. Mita S, Suzuki H, Akita H et al (2006) Inhibition of bile acid transport across Na+/taurocholate co transporting polypeptide (SLC10A1) and bile salt export pump (ABCB 11)-coexpressing LLC-PK1 cells by cholestasis-inducing drugs. Drug Metab Dispos Biol Fate Chem 34:1575–1581. https://doi.org/10.1124/dmd.105.008748
24. Padda MS, Sanchez M, Akhtar AJ, Boyer JL (2011) Drug induced cholestasis. Hepatol Baltim Md 53:1377–1387. https://doi.org/10.1002/hep.24229
25. Van den Hof WFPM., Coonen MLJ, van Herwijnen M et al (2014) Classification of hepatotoxicants using HepG2 cells: a proof of principle study. Chem Res Toxicol 27:433–442. https://doi.org/10.1021/tx4004165
26. Kuhn M, Campillos M, Letunic I et al (2010) A side effect resource to capture phenotypic effects of drugs. Mol Syst Biol 6:343. https://doi.org/10.1038/msb.2009.98
27. Kuhn M, Letunic I, Jensen LJ, Bork P (2016) The SIDER database of drugs and side effects. Nucleic Acids Res 44:D1075-1079. https://doi.org/10.1093/nar/gkv1075
28. Molecular Operating Environment (MOE), 2013.08. Chemical Computing Group Inc., 1010 Sherbooke St. West, Suite #910. Montreal, QC
29. Atkinson F (2014) Standardiser
30. Sadowski J, Gasteiger J, Klebe G (1994) Comparison of automatic three-dimensional model builders using 639 X-ray structures. J Chem Inf Comput Sci 34:1000–1008. https://doi.org/10.1021/ci00020a039

31. Landrum G (2006) RDKit: Open-source cheminformatics
32. Yap CW (2011) PaDEL-descriptor: an open source software to calculate molecular descriptors and fingerprints. J Comput Chem 32:1466–1474. https://doi.org/10.1002/jcc.21707
33. Breiman L (2001) Random Forests. Mach Learn 45:5–32
34. Hall M, Frank E, Holmes G et al (2009) The WEKA data mining software: an update. SIGKDD Explor Newsl 11:10–18. https://doi.org/10.1145/1656274.1656278
35. University of Waikato: Waikato, New Zeland Weka: Weka Environment for Knowledge Analysis. http://www.cs.waikato.ac.nz/ml/weka/. Accessed 2 Nov 2010
36. Oshiro TM, Perez PS, Baranauskas JA (2012) How many trees in a Random Forest? In: Machine learning and data mining in pattern recognition. Springer, Berlin, pp 154–168
37. Breiman L (1996) Bagging predictors. Mach Learn 24:123–140. https://doi.org/10.1023/A:1018054314350
38. Tetko IV, Novotarskyi S, Sushko I et al (2013) Development of dimethyl sulfoxide solubility models using 163,000 molecules: using a domain applicability metric to select more reliable predictions. J Chem Inf Model. https://doi.org/10.1021/ci400213d
39. Sushko I, Novotarskyi S, Körner R et al (2011) Online chemical modeling environment (OCHEM): web platform for data storage, model development and publishing of chemical information. J Comput Aided Mol Des 25:533–554. https://doi.org/10.1007/s10822-011-9440-2
40. On-line CHEmical database and Modelling environment (OCHEM). https://www.ochem.eu. Accessed 7 Apr 2013
41. Domingos P (1999) MetaCost: a general method for making classifiers cost-sensitive. In: Proceedings of the Fifth International Conference on Knowledge Discovery and Data Mining. ACM Press, pp 155–164
42. ThresholdSelector. http://weka.sourceforge.net/doc.packages/thresholdSelector/weka/classifiers/meta/ThresholdSelector.html. Accessed 16 Jul 2017
43. Chawla NV, Japkowicz N, Kotcz A (2004) Editorial: special issue on learning from imbalanced data sets. SIGKDD Explor Newsl 6:1–6. https://doi.org/10.1145/1007730.1007733
44. Powers D (2011) Evaluation: from precision, recall and f-measure to roc., informedness, markedness & correlation. J Mach Learn Technol 2:37–63
45. Chawla NV, Bowyer KW, Hall LO, Kegelmeyer WP (2002) SMOTE: synthetic minority over-sampling technique. J Artif Int Res 16:321–357
46. ClassBalancer. http://weka.sourceforge.net/doc.dev/weka/filters/supervised/instance/ClassBalancer.html. Accessed 16 Jul 2017
47. R Core Team (2013). R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna
48. Su B-H, Tu Y-S, Lin OA et al (2015) Rule-based classification models of molecular autofluorescence. J Chem Inf Model 55:434–445. https://doi.org/10.1021/ci5007432
49. Duan J, Dixon SL, Lowrie JF, Sherman W (2010) Analysis and comparison of 2D fingerprints: insights into database screening performance using eight fingerprint methods. J Mol Graph Model 29:157–170. https://doi.org/10.1016/j.jmgm.2010.05.008
50. Drwal MN, Siramshetty VB, Banerjee P et al (2015) Molecular similarity-based predictions of the Tox21 screening outcome. Front Environ Sci. https://doi.org/10.3389/fenvs.2015.00054
51. Drwal MN, Banerjee P, Dunkel M et al (2014) ProTox: a web server for the in silico prediction of rodent oral toxicity. Nucleic Acids Res 42:W53–W58. https://doi.org/10.1093/nar/gku401

Affiliations

Sankalp Jain[1] · Eleni Kotsampasakou[1,2] · Gerhard F. Ecker[1]

✉ Gerhard F. Ecker
gerhard.f.ecker@univie.ac.at

1 Department of Pharmaceutical Chemistry, University of Vienna, Althanstrasse 14, 1090 Vienna, Austria

2 Present Address: Computational Toxicology Group, CMS, R&D Platform Technology & Science, GSK, Park Road, Ware, Hertfordshire SG12 0DP, UK

4. Structure-based studies

The modern pharmaceutical research aims to develop novel molecules with a desired bioactivity profile against one or more drug targets and, at the same time, avoid unwanted side effects. In this regard, it is very important to elucidate drug-target interactions as this information could provide insights into the mode of action for a particular bioactive molecule[112]. Increasing availability of protein 3D structures in the Protein Data Bank (PDB) [183] and advancements in the computational techniques has motivated researchers all over the world for a structure-based elucidation of protein targets.

4.1 Structure-based modeling studies on BSEP

4.1.1 Structure based classification for bile salt export pump (BSEP) inhibitors using comparative structural modeling of human BSEP

Sankalp Jain, Melanie Grandits, Lars Richter, Gerhard F. Ecker

In this study, we present a homology model of BSEP developed using the corrected mouse P-glycoprotein structure (PDB ID: 4M1M) that was used for molecular docking, in order to predict BSEP inhibitors and non-inhibitors. Among the several docking protocols employed, the best performing one correctly predicted 88% of the compounds in the training set and 77% of the compounds in an external test set. Further, we analyzed the protein-ligand interaction fingerprints, which revealed certain functional group-binding site residue interactions that could play a key role in ligand binding. Finally, combining the structure-based model with our previously published ligand-based classification model in a sequential order (sequential modeling) improved the precision and reduced the calculation time.

S. Jain performed the study and wrote the manuscript. M. Grandits assisted with molecular dynamics simulation and revised the manuscript. **S. Jain** and L. Richter performed the protein-ligand interaction fingerprint (PLIF) analysis. G.F. Ecker supervised the work and revised the manuscript.

CrossMark

Structure based classification for bile salt export pump (BSEP) inhibitors using comparative structural modeling of human BSEP

Sankalp Jain[1] · Melanie Grandits[1] · Lars Richter[1] · Gerhard F. Ecker[1]

Abstract The bile salt export pump (BSEP) actively transports conjugated monovalent bile acids from the hepatocytes into the bile. This facilitates the formation of micelles and promotes digestion and absorption of dietary fat. Inhibition of BSEP leads to decreased bile flow and accumulation of cytotoxic bile salts in the liver. A number of compounds have been identified to interact with BSEP, which results in drug-induced cholestasis or liver injury. Therefore, *in silico* approaches for flagging compounds as potential BSEP inhibitors would be of high value in the early stage of the drug discovery pipeline. Up to now, due to the lack of a high-resolution X-ray structure of BSEP, *in silico* based identification of BSEP inhibitors focused on ligand-based approaches. In this study, we provide a homology model for BSEP, developed using the corrected mouse P-glycoprotein structure (PDB ID: 4M1M). Subsequently, the model was used for docking-based classification of a set of 1212 compounds (405 BSEP inhibitors, 807 non-inhibitors). Using the scoring function ChemScore, a prediction accuracy of 81% on the training set and 73% on two external test sets could be obtained. In addition, the applicability domain of the models was assessed based on Euclidean distance. Further, analysis of the protein–ligand interaction fingerprints revealed certain functional group-amino acid residue interactions that could play a key role for ligand binding. Though ligand-based models, due to

their high speed and accuracy, remain the method of choice for classification of BSEP inhibitors, structure-assisted docking models demonstrate reasonably good prediction accuracies while additionally providing information about putative protein–ligand interactions.

Keywords BSEP · Structure-based classification · Drug-induced cholestasis · Inhibiton · Transporters · Classification model

✉ Gerhard F. Ecker
   gerhard.f.ecker@univie.ac.at

1 Department of Pharmaceutical Chemistry, University of Vienna, Althanstrasse 14, 1090 Vienna, Austria

Introduction

Transmembrane transport proteins selectively aid in the translocation of molecules across biological membranes by binding the substrate molecules followed by a conformational change [1]. Members of the ATP-binding cassette (ABC) superfamily facilitate the transport of their solutes by using the energy from hydrolysis of ATP. While some ABC-transporters allow specific passage of inorganic ions, others facilitate ATP-dependent transport of organic compounds including xenotoxins, short peptides, lipids, bile acids, glutathione, and glucuronide conjugates. Therefore, ABC-transporters affect the absorption, distribution, metabolism, excretion and toxicity of numerous pharmacological agents. Genetic variations in the genes that encode these transporters lead to disorders such as cystic fibrosis, cholesterol and bile transport defects, as well as neurological diseases [2].

The bile salt export pump (BSEP, gene ABCB11) is a canalicular-specific exporter predominantly expressed in the cholesterol-rich apical membrane of hepatocytes [3]. BSEP facilitates secretion of bile salts from the liver into the bile canaliculi [4–6]. The main function of bile acids is to promote digestion and absorption of dietary fat via

Springer

formation of micelles [7]. Apart from this, they are increasingly being shown to have hormonal actions throughout the body [8, 9]. Variations in the ABCB11 gene result in different forms of progressive familial intrahepatic cholestasis (PFIC) [10, 11]. PFIC is characterized by an early onset of cholestasis and eventually leads to liver cirrhosis and failure [12–14].

Inhibition of BSEP can result in accumulation of bile salts in the liver, which is considered to be a primary mechanism leading to drug-induced cholestasis—one of the reasons for drug-induced liver injury (DILI) [15–17]. By inhibiting BSEP, drugs such as bosentan, rifampicin and troglitazone cause intracellular accumulation of bile salts and decreased bile flow [18]. Dysfunction due to suppression of gene expression, disturbed signaling or steric inhibition are other important factors leading to DILI [19]. In its Guideline on the Investigation of Drug Interactions (effective: January 2013), the European Medicines Agency (EMA) indicated that BSEP inhibition assessment should be "preferably investigated". Additionally, EMA states: "If in vitro studies indicate BSEP inhibition, adequate biochemical monitoring including serum bile salts is recommended during drug development" [20]. Furthermore, studies indicate that a majority of drugs that showed in vitro inhibition of BSEP have led to DILI, suggesting that decreased BSEP inhibition is likely to be associated with reduced risk for DILI [17, 21, 22].

With the increasing knowledge of the importance of ABC-transporter for ADMET, also *in silico* models for predicting ligand-transporter interaction became available [23]. With respect to BSEP, QSAR modeling was applied by Warner et al. [24] in which a support vector machine (SVM) model provided the highest accuracy of 87% in the classification of BSEP inhibitors and non-inhibitors on a dataset of 624 compounds [24]. Our group recently published a classification model based on a set of 670 compounds, which allowed the identification of bromocriptine as a BSEP inhibitor [25]. With first X-ray structures of ABC-transporters being published, also structure-based models became available. Bikadi et al. used SVM to predict P-gp substrate binding modes [26, 27]. Dolghih et al. separated P-gp binders from non-binders by applying induced fit docking into the crystal structure of mouse P-gp using the docking score for classification [28]. High area under the curve (AUC) scores of 0.93 and 0.90, respectively were observed for two independent datasets (126 and 64 compounds, respectively). Also Chan et al. [29] evaluated the prediction capability of docking by using 245 P-gp substrates and non-substrates, but the classes were not clearly separated based on the Glide docking scores.

Klepsch et al. [30] showed that docking of a set of propafenones into a homology model of human P-gp reveals poses consistent with QSAR data, and that this can

be exploited for the identification of new P-gp inhibitors [31]. Recently, this was enhanced towards a structure-based classification of almost 2000 compounds [32]. Although the docking-based classification showed significantly lower performance than ligand-based models derived from machine learning, it offers information on the molecular basis of protein ligand interaction.

Up to now, due to the lack of a high-resolution X-ray structure of BSEP, no structure-based studies have been performed for this protein. In the present study, we use comparative modeling [33] to create a protein homology model for BSEP by using the corrected mouse P-glycoprotein structure (PDB ID: 4M1M) as template. Subsequently, we developed structure-based classification models using a dataset comprising 408 compounds (113 inhibitors and 295 non-inhibitors) as training set and two external test sets containing 166 compounds (44 inhibitors and 122 non-inhibitors) and 638 compounds (248 inhibitors and 390 non-inhibitors), respectively.

Materials and methods

Dataset

A set of 408 compounds (113 inhibitors and 295 non-inhibitors) from the work of Warner et al. [24] was used as the training set and another set containing 166 compounds (44 inhibitors and 122 non-inhibitors) from Pedersen et al. [34] was used as external test set. Both studies provide in vitro inhibition data on human BSEP. While Warner et al. classified compounds with a mean $IC_{50} \leq 300$ µM as BSEP inhibitors, in our study we decided to use a much lower threshold (mean $IC_{50} \leq 10$ µM) in order to retain only strong inhibitors. Compounds with mean $IC_{50} > 300$ µM were considered non-inhibitors, and the remaining compounds were excluded from the dataset. Finally, we have a total of 113 strong inhibitors and 295 non-inhibitors. The Pedersen et al. data set is based on inhibition of bile salt export pump (BSEP)-mediated taurocholate (TA) transport in inverted membrane vesicles. After removal of compounds that overlapped with those in our training set, we had a total of 166 compounds (44 strong inhibitors and 122 non-inhibitors) to be used as external test set. In addition, a dataset provided by AstraZeneca within the framework of the IMI project eTOX (http://www.etoxproject.eu) was used as a second external test set to further evaluate our models. The data was measured in a [3H]-taurocholate transport assay performed in Sf21 membrane vesicles using the protocol as described by Dawson et al. [17] and contains the BSEP inhibitory potencies of 1092 compounds as IC_{50} values. Removing the overlapping compounds from the first two datasets resulted in 638 compounds (248 inhibitors and 390

non-inhibitors). All datasets were standardized using the protocol previously described in Montanari et al. [25] and Pinto et al. [35].

Homology modeling

For human BSEP (UNIPROT ID: O95342), based on sequence identity and atomic resolution, the corrected mouse P-glycoprotein structure (PDB ID: 4M1M) was selected as the most structurally related template protein. Multiple homology models were constructed using MODELLER 9.13 [36] and the Prime module in Maestro [37, 38]. Energy minimized models were then evaluated using DOPE score [39], and GA341 score [40, 41]. The quality of the stereochemical parameters and the normality of the structures were checked using the PROCHECK program included in the PDBsum analysis [42]. Ramachandran plot [43] and G-factor [44], and finally the Q-score [45, 46] values were evaluated to identify the top ranked homology model.

Molecular dynamics simulation

Molecular dynamics (MD) simulation was carried out in Gromacs 5.0.4 [47–50] using the GROMOS 54a7 force-field [51]. The protein was placed inside a rectangular box of size $16 \times 16 \times 16$ nm^3 including approximately 34,000 simple point charge (SPC) water molecules [52]. Sodium and chloride ions were added to gain a neutral system. Energy minimization was carried out with a maximum force of 1000 kJ/mol/nm using the steepest descent algorithm. After the minimization, a NVT equilibration was performed at a constant temperature of 300 K for 100 ps. Followed by a NPT equilibration step for 1 ns, with the pressure set constant at 1 atm and a constant temperature of 300 K. The production simulation was performed at 300 K for 20 ns. The LINCS algorithm [53] was used to constrain the covalent bonds and PME [54] was used to calculate the electrostatic interactions during the simulation. The stability of the protein structure was evaluated by calculating the secondary structure over the simulation time according to the Kabsch and Sander rules [55] and the root-mean-square fluctuation (rmsf) of active site residues (Fig. S1 in the supplementary material). All graphs were created using the XMGrace tool [56].

Molecular docking and scoring

In order to avoid any bias in the docking studies, the binding site was defined as the complete TM region, taking 20 Å around the coordinate of the center point to allow subsequent flexible docking studies of a series of BSEP inhibitors. The protein was prepared using Protein Preparation Wizard of the Schrödinger Suite (2015) [57, 58]. During this process, hydrogen atoms were added, and optimal protonation states and ASN/GLN/HIS flips were determined. To assess their correct protonation states, ligands were prepared using the LigPrep module of Schrödinger Suite [58, 59] which produces low-energy 3D structures that can be further used for docking studies. The OPLS_2005 force field was used for the minimization of the structures. Different ionization states were generated by adding or removing protons from the ligand at a target pH of 7.0 ± 2.0 using Epik version 3.1 [60, 61]. Tautomers were generated for each ligand. To generate stereoisomers, the information on chirality from the input file for each ligand was retained as is for the entire calculation. This gave a dataset of 1865 structures (318 inhibitors and 1547 non-inhibitors) for the training set, 2009 structures (858 inhibitors and 1151 non-inhibitors) for the external test set from Pedersen et al. and 1560 structures (668 inhibitors and 892 non-inhibitors) for the external test set from AstraZeneca, which were used for docking with the genetic algorithm-based GOLD suit (version 5.2.0) [62, 63].

All the docking runs were performed in high-throughput mode with GOLD. The fitness functions GoldScore (GS) and ChemScore (CS) were used. GlideXP [64, 65] docking from Maestro was also used in order to compare different scoring functions. Finally, all the poses were rescored using an external scoring function, XScore [66]. To gain deeper insights on the binding modes of BSEP inhibitors and non-inhibitors, the protein–ligand interaction fingerprints (PLIF) of the resultant complexes were retrospectively analyzed.

Machine learning-based model building

The open source software WEKA (version 3.7.10) [67] was used for building binary classification models. The machine learning classifiers: J48, Random Forest, REP-Tree, LibSVM and Naive Bayes were used with the default parameters along with tenfold internal cross-validation.

Network-based representation of the dataset

Tanimoto (Tc) similarities between the inhibitors and non-inhibitors of the training set were calculated using MACCS fingerprints [68]. A chemical space network (CSN) [69, 70] was constructed and analyzed in order to assess the structural similarity shared by the compounds of both groups. To show connections between the compounds, a threshold value of 0.7 was set based on the average of Tanimoto max-similarity in the dataset.

Functional group analysis

Functional group analysis was performed in two stages. First, the substructure patterns of 100 functional groups in SMARTS notation were extracted from the Daylight website (http://www.daylight.com/dayhtml_tutorials/languages/smarts/smarts_examples.html#GROUP). Next, the pattern matching was performed using the SMARTSQueryTool implemented in the Chemistry Development Kit (CDK) [71]. For each functional group, the occurrences of the fragments in a given set of molecules were calculated.

Protein ligand interaction fingerprints (PLIF)

A PLIF summarizes the interactions between a ligand and a protein using a fingerprint scheme. Here we generated three types of PLIFs that differ in the information encoded. In the first approach, the PLIF encodes the residues involved in an interaction with the ligand in each bit. The second one encodes not only the residue but also the nature of the interaction (e.g. hydrogen bond donor) with the ligand. The third category encodes the functional group of the ligand that interacts with the residue. All the PLIF bits were calculated with the MOE [72] built-in function CalculateRawInteractions using a 1% threshold for molecular interactions and a 20% threshold for surface contacts. The function was embedded in an SVL in-house script and was post processed to enable to calculate functional group PLIFs.

Applicability domain assessment

An applicability domain (AD) analysis was performed to evaluate if the chemical space covered by the training set used for developing the model is applicable to predict the outcomes of the test sets used to evaluate the model performance. Therefore, AD could provide a first hint if a new chemical structure is covered within the chemical structures or descriptor space of the training set. Many approaches were proposed to estimate AD, for instance based on descriptor ranges, Euclidean distance or probability density, each having their pros and cons. In this study, we implemented the Euclidean distance approach using the KNIME [73] node APD [74, 75] to evaluate if the test sets are within the AD of the training set.

Performance evaluation

In order to evaluate the quality of our classification models based on the docking studies, we used standard parameters such as count of true positives (TP), false positives (FP), true negatives (TN) and false negatives (FN). Sensitivity (Eq. 1), specificity (Eq. 2) and accuracy (Eq. 3) values were calculated for each model based on the aforementioned

parameters to estimate its performance in classifying inhibitors and non-inhibitors. To measure the overall quality of the model, the G-mean (Eq. 4), which takes into account both sensitivity and specificity, and the Matthews's correlation coefficient (MCC, Eq. 5) were also calculated.

$$Sensitivity = \frac{TP}{(TP + FN)} \tag{1}$$

$$Specificity = \frac{TN}{(TN + FP)} \tag{2}$$

$$Accuracy = \frac{(TP + TN)}{(TP + FP + TN + FN)} \tag{3}$$

$$G-mean = \sqrt{Sensitivity \times Specificity} \tag{4}$$

$$MCC = \frac{\{(TP \times TN) - (FP \times FN)\}}{\{(TP + FP) \times (TP + FN) \times (TN + FP) \times (TN + FN)\}^{1/2}} \tag{5}$$

Calculating the probability of prediction

We examined the distribution of docking scores [Chemscore, Goldscore, GlideXP, Xscore (Chemscore) and Xscore (Goldscore)] for the training set molecules. Based on the minimum and maximum score values, the scores were binned in different intervals. Each bin is characterized by the corresponding number of inhibitors and non-inhibitors. Based on these values, we calculated the probability for a molecule to be an inhibitor or a non-inhibitor. A p value (Chi square test) is calculated for each bin to identify the best scoring range that can be used to separate inhibitors from non-inhibitors.

Results and discussion

Chemical space network of the dataset

Figure 1 shows the CSN with well-resolved community structures for a set of inhibitors and non-inhibitors from the training set. The representative compounds of some communities are shown in Fig. S2 in the supplementary material. Major community structures [69] (communities with at least five representative members) were algorithmically detected and are color-coded. For our CSN designs, the Fruchterman–Reingold algorithm [76] was applied. The node size is proportional to the activity value (pIC_{50}) i.e. the more active the compound, the bigger the node size and vice versa.

A majority of the nodes do not have a connection indicating a high structural diversity in the training dataset. The test dataset from Pedersen et al., showed only three clusters

Fig. 1 CNS representation of the training set compounds based on MACCS Tc similarity threshold of 0.70. Communities with at least five representative members are color coded
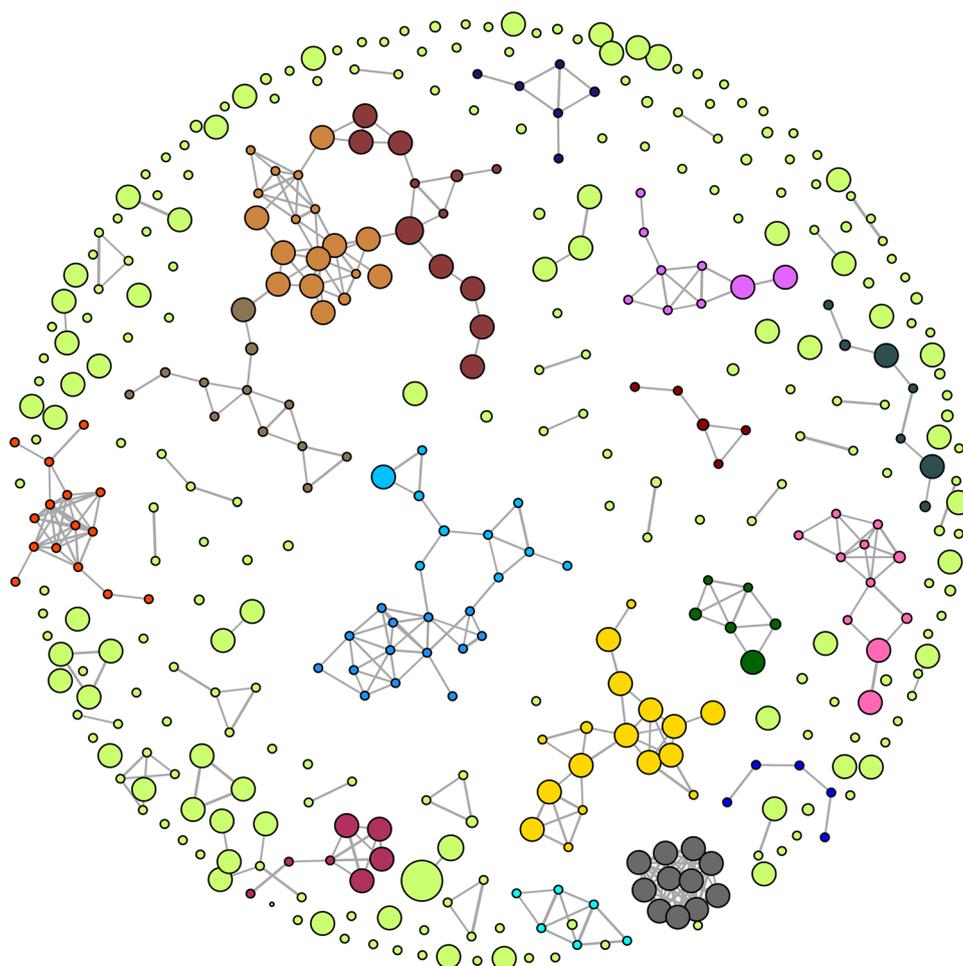
in the CSN with at least five representative members (Fig. S3 in the supplementary material).

Homology modeling

Applying the Prime module from Maestro (Schrödinger, Inc. V-10.1.013), a set of homology models of BSEP were created and refined, using the refined mouse P-gp structure as template (PDB ID: 4M1M). The sequence alignment was done using Prime's alignment program STAin maestro [37, 38] (Fig. S4 in the supplementary material). Analyzing the models with the structure assessment program PROCHECK [42], the best model had a normalized Dope score of −0.625, G-factor −0.12, and Qmean score of 0.597. Furthermore, the Ramachandran plot (Fig. S5 in the supplementary material) showed excellent results, with only 1.9% of residues in generously allowed or disallowed regions. These were all located in the nucleotide binding domains (NBD) or extracellular loops (ECL), and are therefore not involved in drug binding (Fig. S6 in the supplementary material). Based on the study by Mochizuki et al., Asn109, Asn116, Asn122, and Asn125 are residues

predicted to be potential glycosylation sites in the extracellular loop (No.1) (EL No.1) of human BSEP [77]. In our final BSEP homology model (Fig. 2), these residues were also found in EL No.1, thus occurring in the correct region of the transmembrane domain (TMD, Fig. S7 in the supplementary material). For further validation, the best model based on normalized Dope score and Qmean score was subject to molecular dynamics simulations for 20 ns. Both the secondary structure of the protein (Fig. 3) as well as the root mean square fluctuation (RMSF < 0.25 nm) of active site residues showed the stability of the structure.

Docking (structure-based classification)

We recently could demonstrate that a validated homology model of P-glycoprotein allowed docking-based classification of inhibitors and non-inhibitors with reasonable performance [32]. Thus, in this study we extended this approach also to BSEP, using a set of 408 compounds (113 inhibitors and 295 non-inhibitors) published by Warner et al. [24] as training set and two data sets as external test set (see "Materials and methods" section). The scores obtained
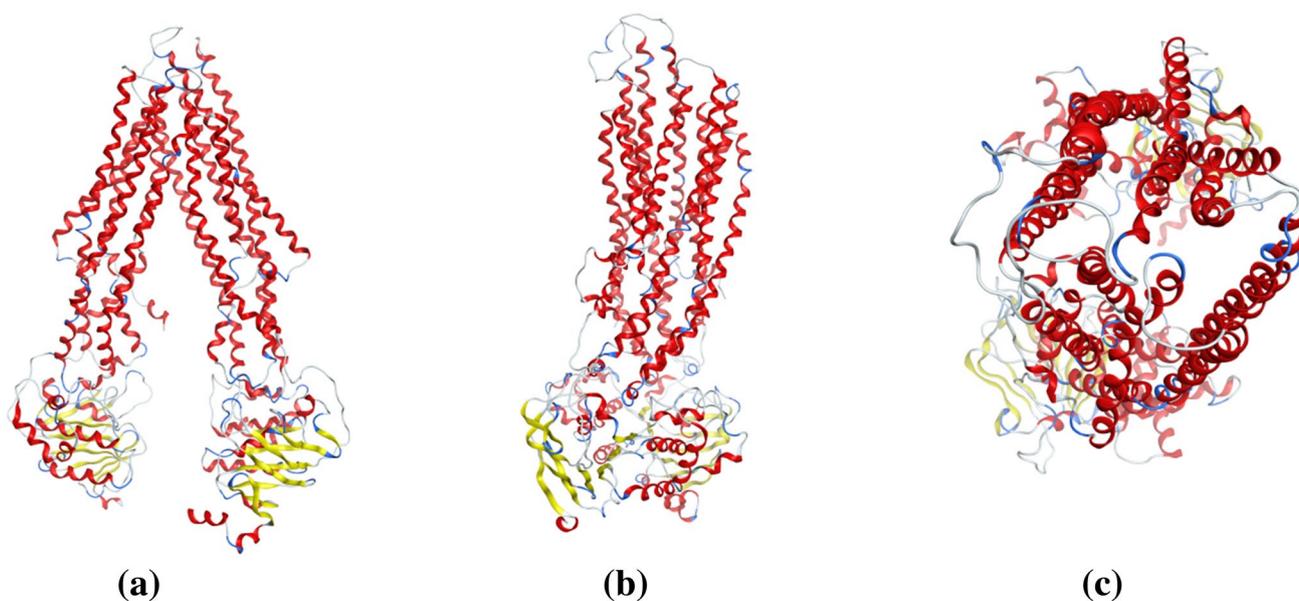
Fig. 2 Homology model structure of human BSEP in the inward-facing state. **a** Front view of the transporter. **b** Side view after a 90° rotation. **c** Top view from the extracellular space
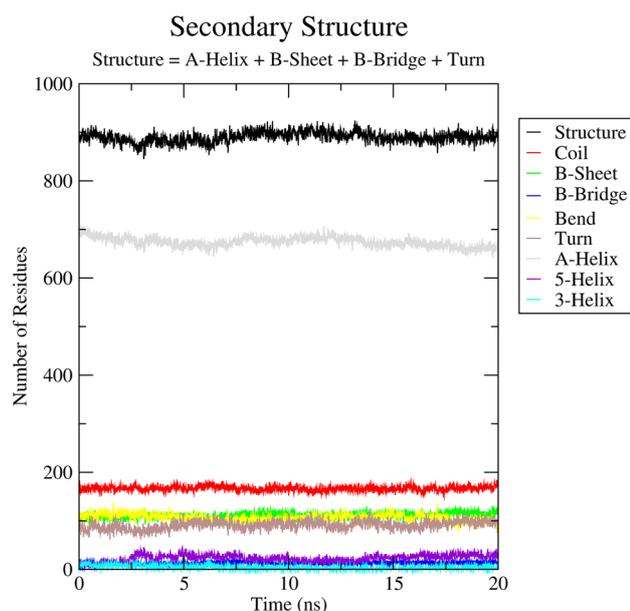
Fig. 3 Secondary structure of the protein over the simulation time

from different fitness functions were binned and the intersection point of the curves for inhibitors and non-inhibitors in the training set served as classification criterion (Fig. 4). Respective confusion matrix parameters and other performance measures are summarized in Table 1. The ChemScore docking run using Xscore as rescoring function retrieved the best performing model with AUC (0.918) and MCC (0.689) measures comparable to the models

developed by Warner et al. [24] and Montanari et al. [25]. This model accurately predicted 88% of the training set compounds and 72% of the external test set compounds derived from Pedersen et al. [34] as well as 77% of a set of AstraZeneca internal compounds. The area under the ROC curve (AUC) measure, being independent from class distribution [78, 79], is a good metric for evaluating performance of virtual screening approaches. High AUC values (above 0.8) were observed, indicating a high capacity of the model in ranking compounds by their probability of being inhibitors of BSEP (Figs. S8–S12 in the supplementary material). The results from the AD assessment also show that all compounds from both test sets were found to be within the chemical domain of the training compounds (Table S1 in the supplementary material). Interestingly, the accuracy of predictions did not improve when a consensus of different scoring functions was used.

Probability of prediction

For the training set using ChemScore scoring, bin 35–40 gave the maximum number of inhibitors. 88% of inhibitors and 12% of non-inhibitors had the docking score in this range with a p value of 5.9×10^{-8}. For both test sets, at least 75% of the inhibitors were found to be in this range. Results for different scoring functions can be found in the Table S2 in the supplementary material. Also with the rescoring of ChemScore using Xscore, a particular range could be defined which significantly distinguishes between inhibitors and non-inhibitors. However, this is not the case
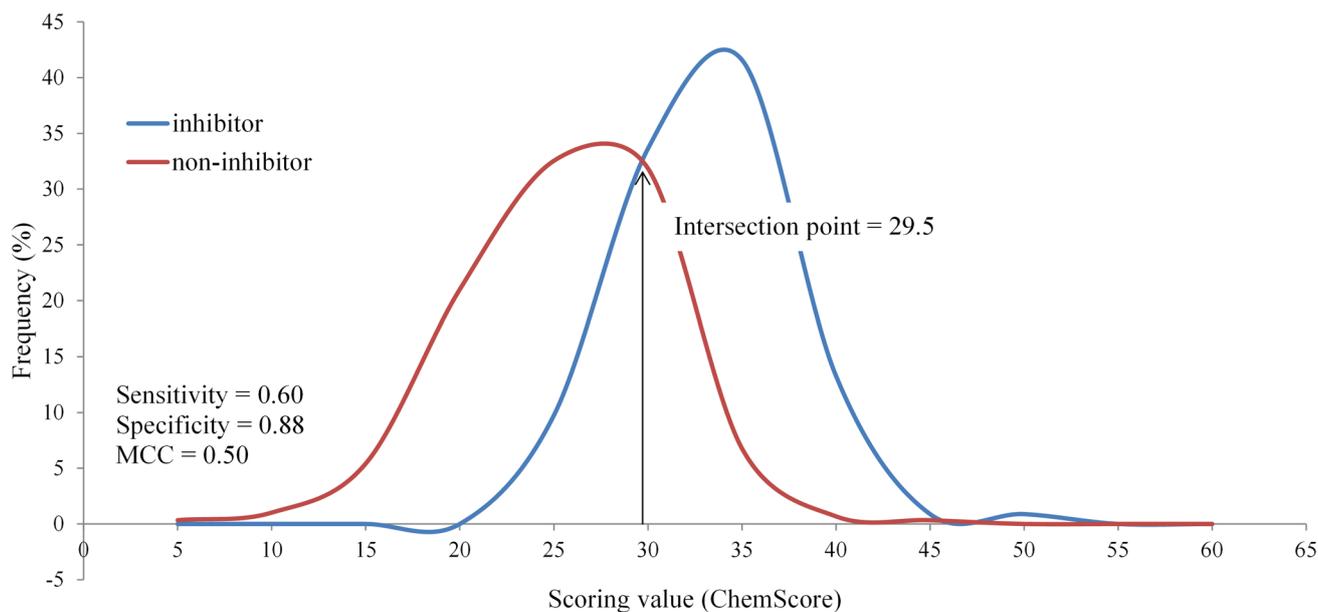
Fig. 4 Distribution of BSEP inhibitors and non-inhibitors (training set) based on ChemScore scoring. Sensitivity, specificity, precision and MCC were calculated from the confusion matrix based on the intersection point of both curves

Table 1 Models obtained from different scoring functions based on the training set

Scoring function	Intersection point	AUC	Sensitivity	Specificity	Accuracy	G-mean	MCC
ChemScore	29.50	0.87	0.60	0.88	0.81	0.73	0.50
GoldScore	53.50	0.82	0.74	0.75	0.75	0.74	0.45
GlideXP	−6.80	0.77	0.80	0.65	0.69	0.72	0.39
Xscore (ChemScore)	6.15	0.92	0.71	0.95	0.88	0.82	0.69
Xscore (GoldScore)	6.10	0.93	0.68	0.95	0.88	0.80	0.68

The scoring function in brackets were used to generate the docking poses

for GoldScore scoring. With this scoring function no particular docking score range could be identified for the three sets (training set, both test sets) to differentiate between the two classes of compounds with a significant p value. Similar results were obtained using the GlideXP scoring function.

Analysis of protein ligand interactions

The Maestro tool allows the computation of different molecular interactions between binding site residues and the corresponding ligand conformation. In this study, the receptor–ligand interaction fingerprint analysis was performed both for the true positives (TPs) and for the true negatives (TNs) on the basis of the docking poses generated. For the training set (Fig. 5) and the two external test sets (Figs. S13, S14 in the supplementary material), the inhibitors showed significantly more hydrophobic interactions with Phe334, Leu364, Tyr772, Phe776 and Leu1026 than non-inhibitors. More than 75% of the inhibitors in the

training set and the external test sets showed hydrophobic interactions with Phe334 and Tyr772 (Fig. 5a). In contrast, non-inhibitors showed a higher number of hydrogen bond interactions than inhibitors (Fig. 5b), which points towards the fact that non-inhibitors are more hydrophilic.

The significant contribution of hydrophobic interactions prompted us to assess the importance of simple molecular descriptors such as logP and molecular weight. Figure 6 represents the distribution of molecular weight and logP(o/w), respectively, for the training set compounds. Similar distributions, represented in Fig. S15 in the supplementary material, were observed with the external test sets from Pedersen et al. [34] and from AstraZeneca (Fig. S16 in the supplementary material). As proposed by Warner et al. [24], molecular properties such as molecular weight (MW) and logP(o/w) could separate the groups quite well (Table 2). At the intersection of MW = 390 and logP(o/w) = 3.6, 79 and 77% of the compounds were classified correctly. Accordingly, compounds with a molecular weight of 390 or higher or a logP of 3.6 or higher were

Fig. 5 **a** Hydrophobic interaction, **b** hydrogen bond interaction fingerprints of true positives (TP) and true negatives (TN) of the training set. The classification of the compounds is based on the ChemScore scoring function

considered as inhibitors while others were considered as non-inhibitors.

The models based on docking scores (ChemScore and XScore) in combination with molecular weight and logP(o/w) (each normalized) outperformed the other models in terms of MCC and precision. ChemScore and XScore based models, when combined with the physicochemical properties [molecular weight and logP(o/w)] correctly predicted 87 and 88% of training set compounds, giving a MCC value of 0.673 and 0.701 respectively. These models also showed high accuracies as compared to other models for the two external test sets. Detailed accuracy measures are presented in Table S3 in the supplementary material.

Also when poses, generated with GoldScore scoring function and rescored with XScore, were combined

with the normalized molecular weight and logP(o/w), it provided accuracies comparable to the former models (Table S3 in the supplementary material). This indicates that considering physicochemical properties of molecules that influence their activity significantly improves the performance of structure-based prediction models.

Distribution of BSEP inhibitors and non-inhibitors using different scoring functions and in combination with physicochemical properties (molecular weight, logP) are presented in Figs. S17–S32 in the supplementary material. A single intersection point could not be obtained, when the rescoring using Xscore (pose generated with GoldScore) was combined with logP(o/w) and thus was not used for the classification of inhibitors and non-inhibitors (Fig. S31 in the supplementary material).
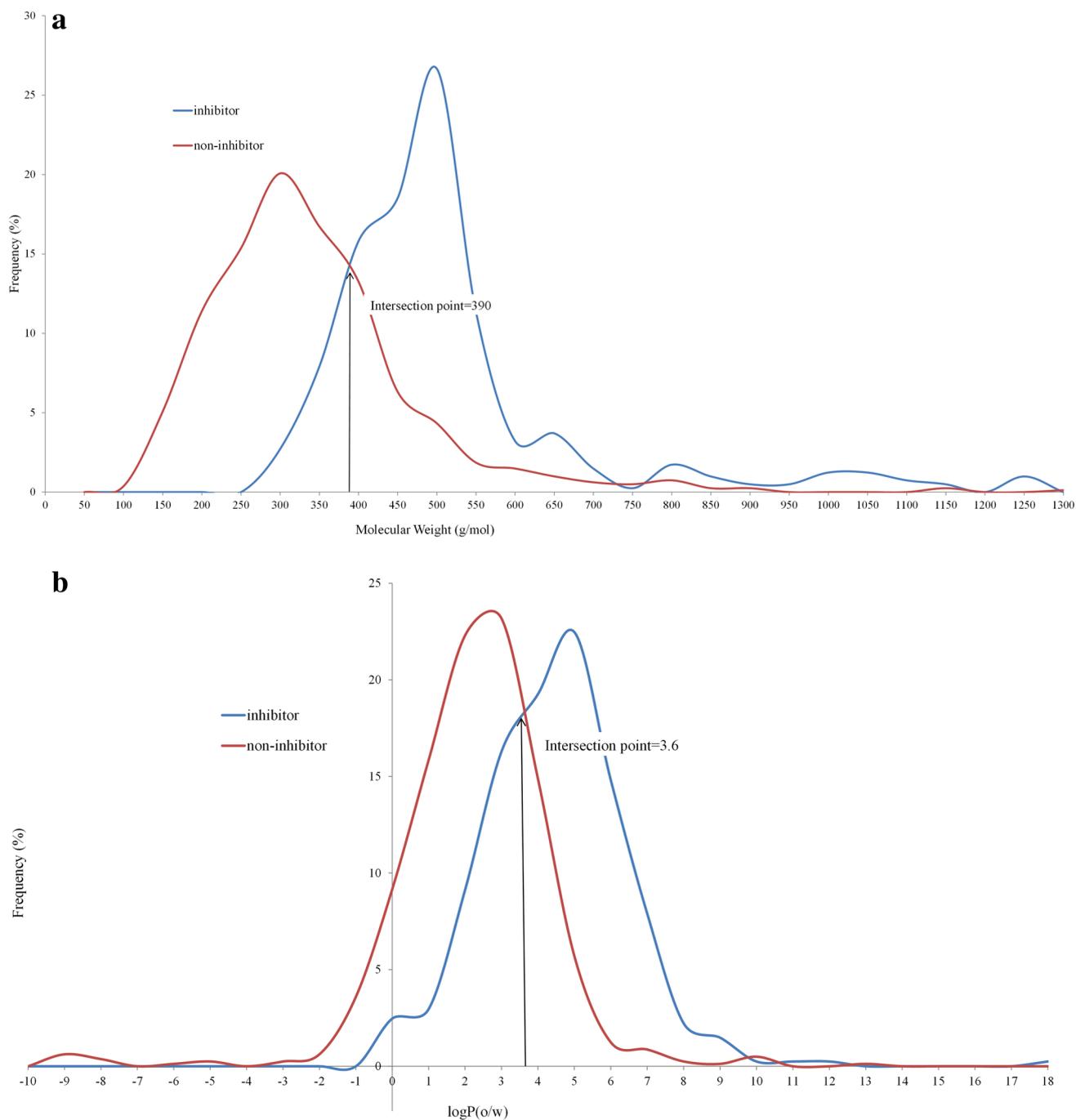
Fig. 6 Distribution of BSEP inhibitors and non-inhibitors based on the **a** molecular weight, **b** logP(o/w) of the training set

Table 2 Models based on physicochemical properties

Molecular property	Intersection point	Sensitivity	Specificity	Accuracy	G-mean	MCC
Molecular weight	390	0.76	0.80	0.79	0.78	0.54
logP	3.6	0.57	0.87	0.77	0.71	0.47

Using the best performing docking scores (Chem-Score, XScore) and the descriptors (molecular weight and logP(o/w)) as parameters, we additionally developed machine-learning based binary classification models using J48, Random Forest, REPTree, LibSVMand Naive Bayes in WEKA [67]. These models performed well with accuracies and MCC values (Table S4 in the supplementary material) comparable to those from machine-learning based classification models of Warner et al. [24] and our models previously developed [25].

Analysis of functional groups and protein–ligand interactions

Next, we investigated the distribution of functional groups between inhibitors and non-inhibitors to identify structural features that are responsible for differences in the activity (inhibitor vs. non-inhibitor). About 70 SMARTS patterns representing the most common functional groups were extracted from the Daylight website (http://www.daylight.com/dayhtml_tutorials/languages/smarts/smarts_examples.html). Basically, groups such as halide/halogen, ether, carbonyl, vinyl carbons (sp2 hybridized) and amide were more frequently found in the inhibitors compared to the non-inhibitors (Fig. 7, S33 in the supplementary material). This further points towards more hydrophobic-driven interactions for inhibitors.

In addition, we also identified the most frequently occurring interactions between residues and functional groups for the training set compounds. A heat map (Fig. 8a) was generated to illustrate the outcomes of PLIF analysis by displaying the contact residues against the functional groups of the interacting ligands. The color scale represents the amount of ligands which are involved in interactions. Therefore, the most significant interactions between a specific residue and a specific functional group could be visually detected.

We found that the interactions of arene and carbonyl functional groups with tyrosine and leucine are more prominently found among the inhibitors in comparison to the non-inhibitors. We furthered with retrospective assessment of the docking results to check the presence of the aforementioned interactions and evaluated the chances to prioritize a compound as a BSEP inhibitor. Figure 8b represents the docking pose of Glimepiride (yellow) in which its carbonyl groups interact with the residues Tyr337, Tyr772 and Asn996. The residue Leu364 shows a hydrophobic interaction with the arene moiety of the ligand. Similarly, the functional group-residue interactions were confirmed to be present in the docking results of both external test datasets (Figs. S34–S36 in the supplementary material).

Although the functional groups analysis suggests that halide/halogen, carbonyl, ether, vinyl and amide groups were significantly over represented in the inhibitors, only carbonyl group, amide were found to frequently interact with the protein. According to the heat map (Fig. 8a), halide/halogen and vinyl groups do not appear to have a significant number of contacts with the residues. At the same time, arene was found at a similar rate in inhibitors (nearly 95%) and non-inhibitors (nearly 85%), but the PLIF analysis revealed that the arene moiety participates in a significant number of interactions with residues such as Leu364 and Leu1026. This indicates that significant differences in the functional group composition between inhibitors and non-inhibitors (Fig. 7) does not necessarily indicate or provide an outlook on the nature of interactions. This would rather depend on the position of these functional groups in the molecular structure, nature of the binding site residues as well as the size of the binding pocket.
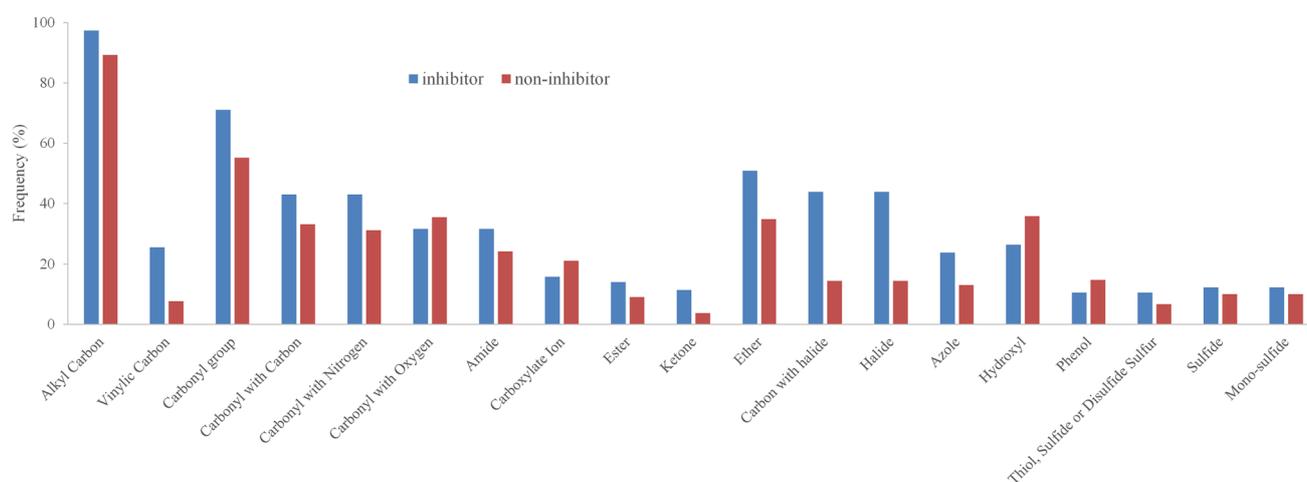
Fig. 7 Distribution of functional groups in the training dataset

a

b
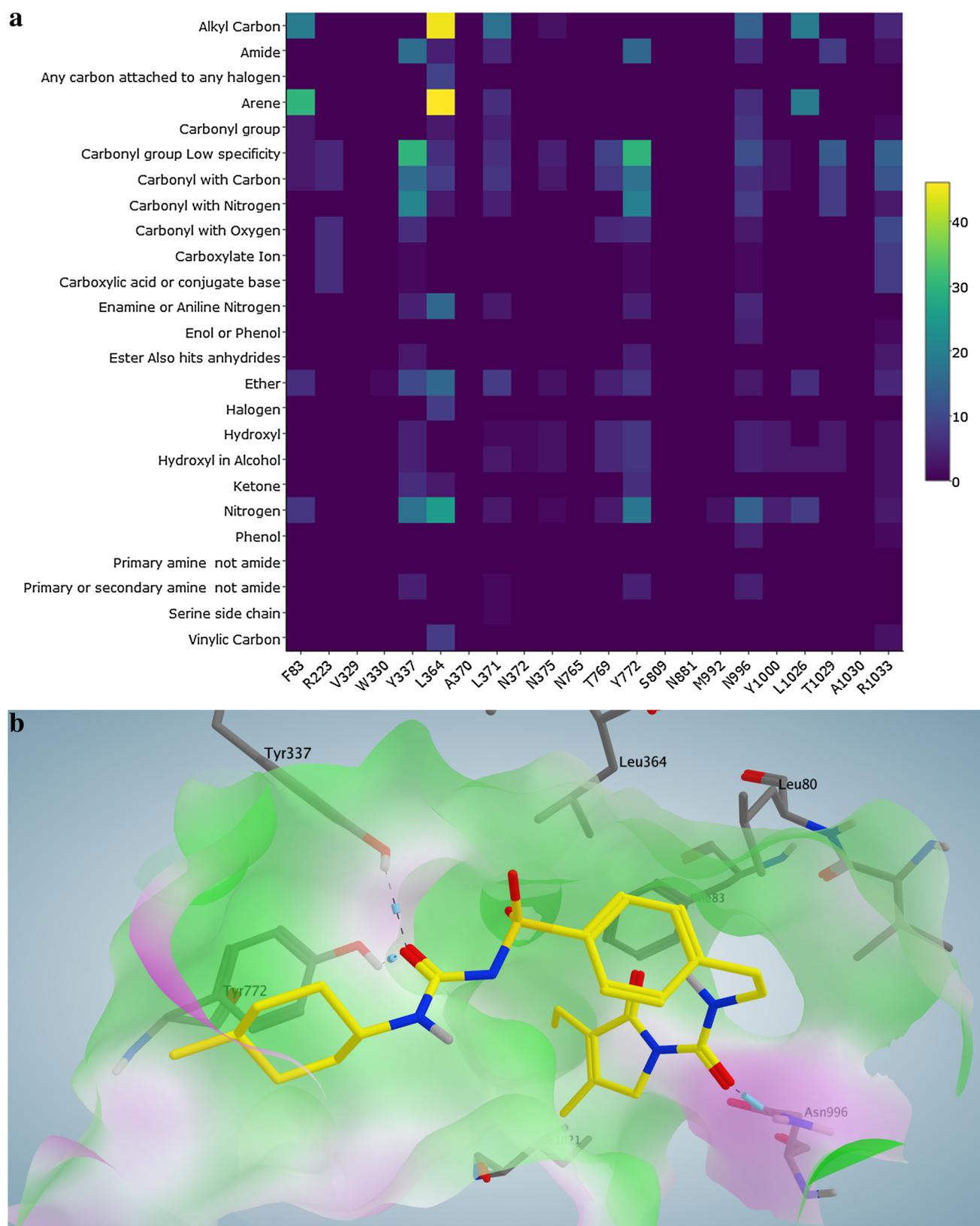
Fig. 8 **a** Heat map illustrating the PLIF analysis of the training set inhibitors (*x-axis* contact residues; *y-axis* functional groups of the ligand showing an interaction with the residue; *color scale* number of interacting ligands). **b** Docking pose of Glimepiride (*yellow*) in which its carbonyl groups interact with the residues Tyr337, Tyr772 and Asn996

Table 3 Ligand-based and structure-based classification

Model type	TP	TN	FP	FN	Sensitivity	Specificity	Accuracy	MCC	Precision
LBC	**30**	**104**	**9**	**9**	**0.77**	**0.92**	**0.88**	**0.69**	**0.77**
SBC_C	27	91	22	12	0.69	0.81	0.78	0.47	0.55
SBC_G	26	79	34	13	0.67	0.70	0.69	0.33	0.43
SBC_C_X	27	96	17	12	0.69	0.85	0.81	0.52	0.61
LBC+SBC_C	24	107	6	15	0.62	0.95	0.86	0.62	0.80
LBC+SBC_C_X	**25**	**108**	**5**	**14**	**0.64**	**0.96**	**0.88**	**0.66**	**0.83**
Consensus	27	106	7	12	0.69	0.94	0.88	0.66	0.79

The best model of the combined approach is highlighted in bold as well as the ligand-based classification

TP true positives, *TN* true negatives, *FP* false positives, *FN* false negatives, *LBC* Ligand-based classification (Montanari et al. [25]), *SBC_C* Structure-based classification using ChemScore scoring function, *SBC_G* Structure-based classification using GoldScore scoring function, *SBC_C_X* Structure-based classification using ChemScore scoring function (rescoring using Xscore). Consensus Combination of LBC, SBC_C and SBC_C_X

Finally, preliminary results show that the PLIF can also be used as predictor for inhibitor/non inhibitor properties by calculating the Tanimoto distance to known inhibitors. A more detailed description of this approach can be found in the supplementary material.

Analysis of misclassified compounds

Nearly 90 compounds, altogether from different datasets, were incorrectly classified by all the four scoring functions used in the study. More than 59% of the training set compounds and 48% of the test set compounds were correctly classified by all the scoring functions. Of the 19 misclassified compounds from the training set, nine were predicted as inhibitors and ten were predicted as non-inhibitors.

The training set compound Ebselen was wrongly predicted as non-inhibitor by all scoring functions. Examining its molecular properties revealed that both molecular weight (274) and logP(2.74) fall in the range of non-inhibitors (Table 2). Moreover, the structure of Ebselen was found to be structurally more similar to a set of non-inhibitors compared to the set of inhibitors. Benzylpenicillin (Penicillin G) also belongs to the property space of non-inhibitors (molecular weight = 333.38 and logP = 1.74). Interestingly, both Ebselen and Benzylpenicillin are strong inhibitors (IC$_{50}$ < 10 µM) [24]. On the other hand, Phytomenadione (molecular weight = 450.70, logP = 9.05), despite being a non-inhibitor (IC$_{50}$ Y > 1000), was always misclassified as inhibitor. Similar trend was noticed in both external test sets. In total, six inhibitors and 13 non-inhibitors were misclassified from the Pedersen et al. [34] dataset. Interestingly, all six inhibitors were found to be strongly hydrophobic and the molecular properties of about 80% of the non-inhibitors fall in the range of inhibitors. This strengthens the inclusion of this physicochemical properties into the classification model.

Combining ligand- and structure-based classification (sequential modeling)

Although the structure-based models performed reasonably well, ligand-based methods are considerably faster and perform equally well. Thus, we evaluated if a sequential approach that starts with a ligand-based method and proceeds with screening the positives using structure-based models would improve the precision and reduce the false positives. Therefore, we used an external test set containing 39 inhibitors and 113 non-inhibitors as a starting point. After applying ligand-based classification using the workflow from Montanari et al. [25], 30 inhibitors were correctly predicted (TPs) and there were nine FPs, which leads to a precision of 0.77. After application of our structure-based model based on ChemScore and rescoring using XScore, the precision improved to 0.83, reducing the number of FPs to 5. Further performance measures on the sequential approach are provided in Table 3. Thus, combining ligand- and structure-based models in a sequential setting increased the precision and reduced the calculation time. This might be a versatile approach to reduce the number of FPs when performing large scale *in silico* screening.

Conclusion

Development of structure-based methods for transmembrane transporters of the ABC-family has been less pronounced due to limited availability of experimentally determined 3D structures. However, recent efforts that used homology models of P-glycoprotein provide promising evidences that structure-based classification methods can be applied to these highly flexible and promiscuous proteins. In this study, we used comparative modeling to generate a homology model for the ABC-transporter BSEP and developed structure-based models to classify

inhibitors and non-inhibitors. Including logP and molecular weight as an additional layer of information besides the scoring function further increased the performance of the models. PLIF analysis revealed certain functional group-residue interactions that could help to understand the molecular basis of inhibition of the transporter protein by a wide range of ligands. Applicability domain of the models was assessed using Euclidean distance. Furthermore, we estimated the probability of prediction by employing a binning scheme and identified a docking score range that can distinguish a majority of inhibitors from non-inhibitors with high confidence. Finally, combining the structure-based model with our previously published ligand-based classification model in a sequential order provided additional improvement.

Combining ligand- and structure-based models to enhance the performance of virtual screening is of course not a new approach. For receptors and enzymes identification of new ligands quite often starts with a pharmacophore-based screening followed by docking of the top-ranked hits to further refine the shopping list [80]. However, in case of ABC-transporters such as P-glycoprotein, which shows a pronounced polyspecificity in its ligand profile, there is a broad variety of pharmacophore models available. This would render a sequential approach quite challenging. Furthermore, due to the eminent role of ABC-transporters like P-gp, BSEP, and the breast cancer protein (BCRP) in ADME and toxicity, the focus for *in silico* screening lays more on flagging potentially toxic compounds rather than on the identification of new inhibitors for further development as drug candidates. In this setting, machine learning-based classification models might be a better tool for a first computational pre-screening. Therefore, a workflow comprising of prescreening with simple descriptors, classification by machine learning techniques and post processing by structure-based methods might be the workflow of choice to provide accurate prediction combined with additional information on the molecular basis of compound-transporter interaction.

References

1. Lodish H, Berk A, Zipursky SL, Matsudaira P, Baltimore D, Darnell J (2000) Overview of membrane transport proteins. In: Lodish H (ed) Molecular cell biology, 4th edn. W. H. Freeman, New York
2. Dean M, Rzhetsky A, Allikmets R (2001) The human ATP-binding cassette (ABC) transporter superfamily. Genome Res 11:1156–1166
3. Kim S-R, Saito Y, Itoda M, Maekawa K, Kawamoto M, Kamatani N, Ozawa S, Sawada J (2009) Genetic variations of the ABC transporter gene ABCB11 encoding the human bile salt export pump (BSEP) in a Japanese population. Drug Metab Pharmacokinet 24:277–281
4. Glavinas H, Krajcsi P, Cserepes J, Sarkadi B (2004) The role of ABC transporters in drug resistance, metabolism and toxicity. Curr Drug Deliv 1:27–42
5. Giacomini KM, Huang S-M, Tweedie DJ et al (2010) Membrane transporters in drug development. Nat Rev Drug Discov 9:215–236
6. Cheng X, Buckley D, Klaassen CD (2007) Regulation of hepatic bile acid transporters Ntcp and Bsep expression. Biochem Pharmacol 74:1665–1676
7. Hofmann AF, Borgström B (1964) The intraluminal phase of fat digestion in man: the lipid content of the micellar and oil phases of intestinal content obtained during fat digestion and absorption*. J Clin Invest 43:247–257
8. Fiorucci S, Mencarelli A, Palladino G, Cipriani S (2009) Bile-acid-activated receptors: targeting TGR5 and farnesoid-X-receptor in lipid and glucose disorders. Trends Pharmacol Sci 30:570–580
9. Kuipers F, Groen AK (2008) Chipping away at gallstones. Nat Med 14:715–716
10. Strautnieks SS, Byrne JA, Pawlikowska L et al (2008) Severe bile salt export pump deficiency: 82 different ABCB11 mutations in 109 families. Gastroenterology 134:1203–1214
11. Perez M-J, Briz O (2009) Bile-acid-induced cell injury and protection. World J Gastroenterol 15:1677–1689
12. Amer S, Hajira A (2014) A comprehensive review of progressive familial intrahepatic cholestasis (PFIC): genetic disorders of hepatocanalicular transporters. Gastroenterol Res 7:39–43
13. Alonso EM, Snover DC, Montag A, Freese DK, Whitington PF (1994) Histologic pathology of the liver in progressive familial intrahepatic cholestasis. J Pediatr Gastroenterol Nutr 18:128–133
14. JANSEN P, MULLER M (2000) The molecular genetics of familial intrahepatic cholestasis. Gut 47:1–5
15. Drug Transport. In: Sigma–Aldrich. http://www.sigmaaldrich.com/technical-documents/articles/biofiles/drug-transport.html. Accessed 17 March 2015
16. Kosters A, Karpen SJ (2008) Bile acid transporters in health and disease. Xenobiotica Fate Foreign Compd Biol Syst 38:1043–1071
17. Dawson S, Stahl S, Paul N, Barber J, Kenna JG (2012) In vitro inhibition of the bile salt export pump correlates with risk of cholestatic drug-induced liver injury in humans. Drug Metab Dispos Biol Fate Chem 40:130–138
18. Sahi J, Sinz MW, Campbell S et al (2006) Metabolism and transporter-mediated drug-drug interactions of the endothelin-A receptor antagonist CI-1034. Chem Biol Interact 159:156–168
19. Kullak-Ublick GA, Stieger B, Meier PJ (2004) Enterohepatic bile salt transporters in normal physiology and liver disease. Gastroenterology 126:322–342
20. Guideline on the investigation of drug interactions. http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2012/07/WC500129606.pdf

21. Morgan RE, Trauner M, van Staden CJ, Lee PH, Ramachandran B, Eschenberg M, Afshari CA, Qualls CW, Lightfoot-Dunn R, Hamadeh HK (2010) Interference with bile salt export pump function is a susceptibility factor for human liver injury in drug development. Toxicol Sci Off J Soc Toxicol 118:485–500

22. Kis E, Ioja E, Rajnai Z, Jani M, Méhn D, Herédi-Szabó K, Krajcsi P (2012) BSEP inhibition: in vitro screens to assess cholestatic potential of drugs. Toxicol Vitro Int J Publ Assoc BIBRA 26:1294–1299

23. Montanari F, Ecker GF (2015) Prediction of drug–ABC-transporter interaction—recent advances and future challenges. Adv Drug Deliv Rev 86:17–26

24. Warner DJ, Chen H, Cantin L-D, Kenna JG, Stahl S, Walker CL, Noeske T (2012) Mitigating the inhibition of human bile salt export pump by drugs: opportunities provided by physicochemical property modulation, *in silico* modeling, and structural modification. Drug Metab Dispos Biol Fate Chem 40:2332–2341

25. Montanari F, Pinto M, Khunweeraphong N et al (2016) Flagging drugs that inhibit the bile salt export pump. Mol Pharm 13:163–171

26. Bikadi Z, Hazai I, Malik D et al (2011) Predicting P-glycoprotein-mediated drug transport based on support vector machine and three-dimensional crystal structure of P-glycoprotein. PLoS ONE 6:e25815

27. Blower PE, Yang C, Fligner MA, Verducci JS, Yu L, Richman S, Weinstein JN (2002) Pharmacogenomic analysis: correlating molecular substructure classes with microarray gene expression data. Pharmacogenomics J 2:259–271

28. Dolghih E, Bryant C, Renslo AR, Jacobson MP (2011) Predicting binding to P-glycoprotein by flexible receptor docking. PLoS Comput Biol 7:e1002083

29. Chen L, Li Y, Yu H, Zhang L, Hou T (2012) Computational models for predicting substrates or inhibitors of P-glycoprotein. Drug Discov Today 17:343–351

30. Klepsch F, Chiba P, Ecker GF (2011) Exhaustive sampling of docking poses reveals binding hypotheses for propafenone type inhibitors of P-glycoprotein. PLoS Comput Biol 7:e1002036

31. Prokes K (2012) Development of "in silico" models for identification of new ligands acting as pharmacochaperones for P-glycoprotein. Diploma Thesis, University of Vienna, Austria

32. Klepsch F, Vasanthanathan P, Ecker GF (2014) Ligand and structure-based classification models for prediction of P-glycoprotein inhibitors. J Chem Inf Model 54:218–229

33. Xiang Z (2006) Advances in homology protein structure modeling. Curr Protein Pept Sci 7:217–227

34. Pedersen JM, Matsson P, Bergström CAS, Hoogstraate J, Norén A, LeCluyse EL, Artursson P (2013) Early identification of clinically relevant drug interactions with the human bile salt export pump (BSEP/ABCB11). Toxicol Sci 136:328–343

35. Pinto M, Trauner M, Ecker GF (2012) An *in silico* classification model for putative ABCC2 substrates. Mol Inform 31:547–553

36. Eswar N, Webb B, Marti-Renom MA, Madhusudhan MS, Eramian D, Shen M-Y, Pieper U, Sali A (2007) Comparative protein structure modeling using MODELLER. Curr Protoc Protein Sci Editor Board John E Coligan Al Chap 2:2.9

37. Jacobson MP, Pincus DL, Rapp CS, Day TJF, Honig B, Shaw DE, Friesner RA (2004) A hierarchical approach to all-atom protein loop prediction. Proteins 55:351–367

38. Jacobson MP, Friesner RA, Xiang Z, Honig B (2002) On the role of the crystal environment in determining protein side-chain conformations. J Mol Biol 320:597–608

39. Shen M, Sali A (2006) Statistical potential for assessment and prediction of protein structures. Protein Sci Publ Protein Soc 15:2507–2524

40. Melo F, Sánchez R, Sali A (2002) Statistical potentials for fold assessment. Protein Sci Publ Protein Soc 11:430–448

41. John B, Sali A (2003) Comparative protein structure modeling by iterative alignment, model building and model assessment. Nucleic Acids Res 31:3982–3992

42. Laskowski R, Macarthur M, Moss D, Thornton J (1993) PROCHECK: a program to check the stereochemical quality of protein structures. J Appl Cryst 26:283–291

43. Zhou AQ, O'Hern C, Regan L (2011) Revisiting the Ramachandran plot from a new angle. Protein Sci Publ Protein Soc 20:1166–1171

44. Engh R, Huber R (1991) Accurate bond and angle parameters for X-ray protein structure refinement. Acta Cryst A 47:392–400

45. Benkert P, Künzli M, Schwede T (2009) QMEAN server for protein model quality estimation. Nucleic Acids Res 37:W510–W514

46. Benkert P, Tosatto SCE, Schomburg D (2008) QMEAN: a comprehensive scoring function for model quality assessment. Proteins 71:261–277

47. Van Der Spoel D, Lindahl E, Hess B, Groenhof G, Mark AE, Berendsen HJC (2005) GROMACS: fast, flexible, and free. J Comput Chem 26:1701–1718

48. Hess B, Kutzner C, van der Spoel D, Lindahl E (2008) GROMACS 4: algorithms for highly efficient, load-balanced, and scalable molecular simulation. J Chem Theory Comput 4:435–447

49. Lindahl E, Hess B, Spoel D van der (2001) GROMACS 3.0: a package for molecular simulation and trajectory analysis. Mol Model Annu 7:306–317

50. Berendsen HJC, van der Spoel D, van Drunen R (1995) GROMACS: a message-passing parallel molecular dynamics implementation. Comput Phys Commun 91:43–56

51. Schmid N, Eichenberger AP, Choutko A, Riniker S, Winger M, Mark AE, van Gunsteren WF (2011) Definition and testing of the GROMOS force-field versions 54A7 and 54B7. Eur Biophys J 40:843–856

52. Berendsen HJC, Postma JPM, Gunsteren WF van, Hermans J (1981) Interaction models for water in relation to protein hydration. In: Pullman B (ed) Intermolecular forces. Springer, Dordrecht, pp 331–342

53. Hess B, Bekker H, Berendsen HJC, Fraaije JGEM (1997) LINCS: a linear constraint solver for molecular simulations. J Comput Chem 18:18–1463

54. Essmann U, Perera L, Berkowitz ML, Darden T, Lee H, Pedersen LG (1995) A smooth particle mesh Ewald method. J Chem Phys 103:8577–8593

55. Kabsch W, Sander C (1983) Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. Biopolymers 22:2577–2637

56. Turner PJ (2005) XMGRACE. Center for Coastal and Land-Margin Research. Oregon Graduate Institute of Science and Technology, Beaverton

57. Sastry GM, Adzhigirey M, Day T, Annabhimoju R, Sherman W (2013) Protein and ligand preparation: parameters, protocols, and influence on virtual screening enrichments. J Comput Aided Mol Des 27:221–234

58. Schrödinger Release 2015-1 (2015) Maestro, version 10.1, Schrödinger. LLC, New York

59. Schrödinger Release 2015-1 (2015) LigPrep, version 3.3, Schrödinger. LLC, New York

60. Shelley JC, Cholleti A, Frye LL, Greenwood JR, Timlin MR, Uchimaya M (2007) Epik: a software program for pK(a) prediction and protonation state generation for drug-like molecules. J Comput Aided Mol Des 21:681–691

61. Greenwood JR, Calkins D, Sullivan AP, Shelley JC (2010) Towards the comprehensive, rapid, and accurate prediction of the favorable tautomeric states of drug-like molecules in aqueous solution. J Comput Aided Mol Des 24:591–604

62. Verdonk ML, Cole JC, Hartshorn MJ, Murray CW, Taylor RD (2003) Improved protein–ligand docking using GOLD. Proteins Struct Funct Bioinform 52:609–623

63. Jones G, Willett P, Glen RC, Leach AR, Taylor R (1997) Development and validation of a genetic algorithm for flexible docking. J Mol Biol 267:727–748

64. Halgren TA, Murphy RB, Friesner RA, Beard HS, Frye LL, Pollard WT, Banks JL (2004) Glide: a new approach for rapid, accurate docking and scoring. 2. Enrichment factors in database screening. J Med Chem 47:1750–1759

65. Friesner RA, Murphy RB, Repasky MP, Frye LL, Greenwood JR, Halgren TA, Sanschagrin PC, Mainz DT (2006) Extra precision glide: docking and scoring incorporating a model of hydrophobic enclosure for protein–ligand complexes. J Med Chem 49:6177–6196

66. Wang R, Lai L, Wang S (2002) Further development and validation of empirical scoring functions for structure-based binding affinity prediction. J Comput Aided Mol Des 16:11–26

67. Hall M, Frank E, Holmes G, Pfahringer B, Reutemann P, Witten IH (2009) The WEKA data mining software: an update. SIGKDD Explor Newsl 11:10–18

68. MACCS Structural keys 2011, Accelrys, San Diego

69. Vogt M, Stumpfe D, Maggiora GM, Bajorath J (2016) Lessons learned from the design of chemical space networks and opportunities for new applications. J Comput Aided Mol Des 30:191–208

70. Zwierzyna M, Vogt M, Maggiora GM, Bajorath J (2015) Design and characterization of chemical space networks for different compound data sets. J Comput Aided Mol Des 29:113–125

71. Steinbeck C, Han Y, Kuhn S, Horlacher O, Luttmann E, Willighagen E (2003) The Chemistry Development Kit (CDK): an open-source Java library for chemo- and bioinformatics. J Chem Inf Comput Sci 43:493–500

72. Molecular Operating Environment (MOE), 2013.08. Chemical Computing Group Inc., Montreal, Canada

73. Berthold MR, Cebron N, Dill F, Gabriel TR, Kötter T, Meinl T, Ohl P, Thiel K, Wiswedel B (2009) KNIME—the Konstanz information miner: version 2.0 and beyond. SIGKDD Explor Newsl 11:26–31

74. Melagraki G, Afantitis A, Sarimveis H, Igglessi-Markopoulou O, Koutentis PA, Kollias G (2010) In silico exploration for identifying structure-activity relationship of MEK inhibition and oral bioavailability for isothiazole derivatives. Chem Biol Drug Des 76:397–406

75. Afantitis A, Melagraki G, Koutentis PA, Sarimveis H, Kollias G (2011) Ligand-based virtual screening procedure for the prediction and the identification of novel β-amyloid aggregation inhibitors using Kohonen maps and Counterpropagation Artificial Neural Networks. Eur J Med Chem 46:497–508

76. Fruchterman TMJ, Reingold EM (1991) Graph drawing by force-directed placement. Softw Pract Exp 21:1129–1164

77. Mochizuki K, Kagawa T, Numari A, Harris MJ, Itoh J, Watanabe N, Mine T, Arias IM (2007) Two N-linked glycans are required to maintain the transport activity of the bile salt export pump (ABCB11) in MDCK II cells. Am J Physiol Gastrointest Liver Physiol 292:G818–G828

78. Fawcett T (2006) An introduction to ROC analysis. Pattern Recogn Lett 27:861–874

79. Bradley AP (1997) The use of the area under the ROC curve in the evaluation of machine learning algorithms. Pattern Recogn 30:1145–1159

80. Küblbeck J, Jyrkkärinne J, Poso A, Turpeinen M, Sippl W, Honkakoski P, Windshügel B (2008) Discovery of substituted sulfonamides and thiazolidin-4-one derivatives as agonists of human constitutive androstane receptor. Biochem Pharmacol 76:1288–1297

4.2 Structure-based modeling studies on BCRP

BCRP and P-gp have multiple common substrates and inhibitors [184, 185]. Several drugs including anti-cancer agents, statins, antibiotics and environmental toxins are BCRP substrates [186]. As stated earlier, BCRP also plays a major role in cancer resistance and tumor progression/development [186–188]. Recently, a cryo-electron microscopy structure of BCRP was published by Taylor et al. [86] that provided the first high-resolution insight into this human multidrug transporter. This motivated us to perform structure-based studies on BCRP.

4.2.1 A hypothesis of the molecular basis for inhibition of BCRP by arylmethyloxyphenyl analogues using the BCRP crystal structure

In the following chapter, we employed a structure-based modeling approach to elucidate molecular hypothesis for the binding of arylmethyloxyphenyl derivatives to BCRP. The structure-activity relationship knowledge from ligand-based investigations guided us through the quest for a flexible depiction of the protein side. Our binding hypothesis suggests that the activity of arylmethyloxyphenyl derivatives is driven by strong hydrophobic interactions and provides a rationale for the development of highly potent derivatives.

This work was performed in collaboration with Dr. Vittorio Pace (University of Vienna) and a synthesis-oriented manuscript is in preparation and is planned to be submitted soon.

A hypothesis of the molecular basis for inhibition of BCRP by arylmethyloxyphenyl analogues using the BCRP crystal structure

Introduction

In 2006, Colabufo et.al. [1] published a medicinal chemistry study in which arylmethyloxyphenyl derivatives and their potential use as P-glycoprotein inhibitors was explored. A set of eight derivatives was synthesized and tested for their ability to revert P-gp-mediated vinblastine transport in human epithelial colorectal adenocarcinoma cells (Caco-2). Two compounds obtained EC50 values below 30 μM and showed no ATPase activation. Further exploration of the arylmethyloxyphenyl scaffold followed [2, 3].

The most potent P-gp modulators from these studies [2, 3] were tested for their ability to inhibit the bile salt export pump (BCRP) in [3H]-mitoxantrone displacement assays. The tests revealed high inhibitory activity of amino derivatives (EC_{50} <2 μM) towards the transporter. Based on these findings, Dr. Vittorio Pace group at University of Vienna further synthesized arylmethyloxyphenyl analogues. Six of these analogues were tested in an intracellular mitoxantrone accumulation assay in PLB985 cells overexpressing BCRP [4, 5] and a spread of 0.12 – 18 μM in IC50 was observed (Figure 1). From these findings, some structure-activity relationships (SAR) could be inferred. It was shown that a lipophilic linker (SM-562, SM-565), connecting ring B and C, was more favourable than a hydrophilic linker (GP199-1, GP196-2). Furthermore, it was noted that a carbonyl moiety was less favourable than a hydroxyl moiety in the linker region. Additional increase in activity was observed by the introduction of methoxy moieties at ring A (GP199-1 versus GP196-2, GS4 versus GS3).

The inferred SAR in conjunction with the release of the BCRP crystal structure in May 2017 [6] motivated us to conduct structure based studies with the aim to propose a binding mode that could explain the spread in activity within the arylmethyloxyphenyl series. BCRP is a half transporter containing one nucleotide binding domain and one transmembrane domain. Thus, in order to be functional, the transporter has to undergo dimerization [7, 8], which renders the transporter symmetric.

93
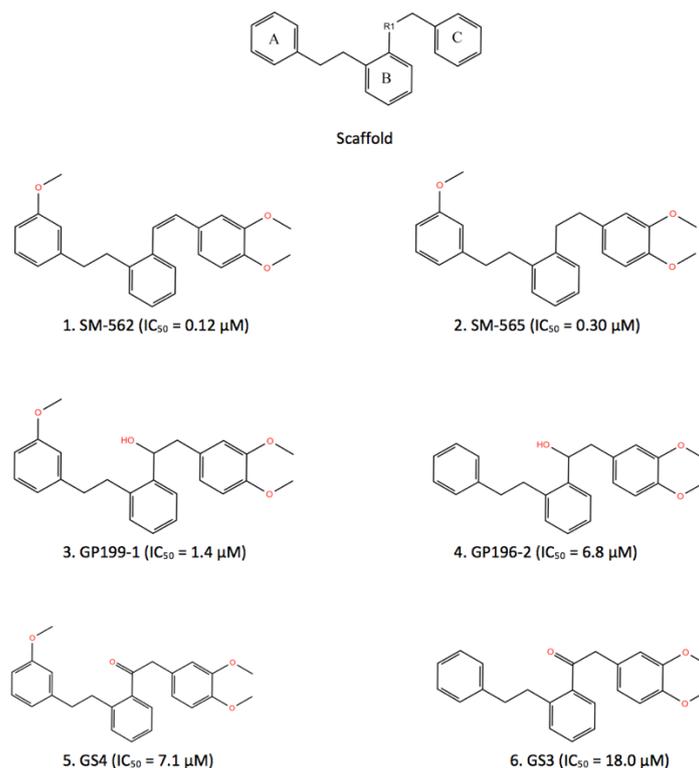
Figure 1: Ligand structures and codes along with their activity values (IC$_{50}$) used in the study.

Method

Molecular Docking

For the docking studies, 6 arylmethyloxyphenyl derivatives (Figure 1) were selected with their known SAR. LigPrep module of Schrödinger Suite [9, 10] was then used to generate their correct protonation states for these derivatives. The OPLS_2005 force field was applied for the minimization of the structures. Different ionization states were generated by adding or removing protons from the ligand at a target pH of 7.0 ± 2.0 using Epik version 3.1 [11, 12] . Tautomers were also generated for each ligand. To generate stereoisomers, the information on chirality from the input file for each ligand was retained as is for the entire calculation. Further ConfGen module from Schrödinger Suite [9, 13] was used to generate maximum possible conformations of the input ligand, which were then used for the docking studies. This gave us a dataset of 1588 ligands. PDB structure (PDB: 5NJ3) retrieved from Protein Data Bank database was prepared for docking procedure using Protein Preparation Wizard of the Schrödinger Suite (2015) [9, 14]. During the protein preparation, hydrogen atoms were

added, water molecules were removed, and correct bond types were set. As the active site is not known, complete transmembrane domain was defined as the binding site (Figure 2).
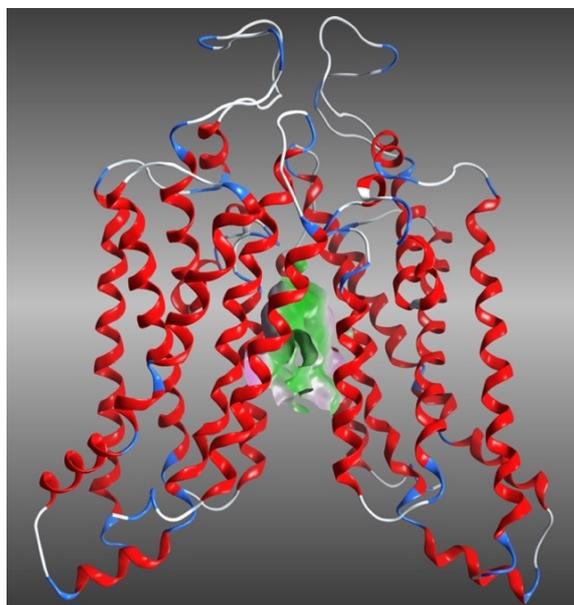
Figure 2: Structure of human ABCG2 and its potential ligand-binding site.

All the docking runs were performed in high-throughput mode with GOLD [15, 16]. The implemented Gold scoring function "GoldScore" was used for evaluation of the complexes. A total of 5 poses per conformation were generated, which led us to 7940 poses. This would help us avoid any bias introduced by scoring functions, as large amount of docking poses was generated.

Clustering of docking poses

A RMSD matrix of all 7940 poses was generated on basis of the common scaffold of the 6 arylmethyloxyphenyl derivatives. The matrix was used for cluster analysis applying complete linkage algorithm in R[17]. A clustering height of 2 Å was used.

Result and Discussion

Cluster Analysis

Although docking simulations have their limitations depending on the validity of the target structure, the results of docking of the 6 arylmethyloxyphenyl derivatives into BCRP crystal

structure (PDB: 5NJ3) are very consistent.

A total of 109 clusters were obtained. Highly populated clusters that contained poses of all docked compounds are considered the most promising. Cluster 2 was the only cluster that contained greater than 50 poses per ligand and was selected for further analysis (Figure 3). In cluster 2, the top docking poses of the 6 compounds are largely overlapping (Figure 4). While ring B consistently shows pi-pi and hydrophobic interactions with Phe439 (chain A), ring A and C are accommodated in identical hydrophobic sub pockets of the homodimer shaped by residues Phe431, Phe432, Asn436, Val546. Interestingly, the main scaffold of the ligands shows itself features of symmetry (Figure 1). All derivatives show strong hydrophobic interaction of ring A with Phe432 (chain A), ring B with Phe439 (chain A), ring C with Phe439 (chain B) (Figure 5).

Our binding hypothesis suggest that the activity of arylmethyloxyphenyl derivatives is driven by strong hydrophobic interactions.
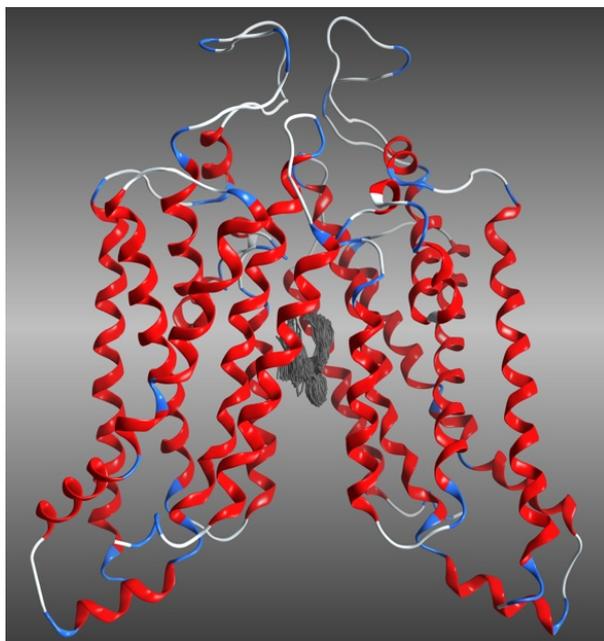
Figure 3: Distribution of poses of cluster 2 in the human ABCG2
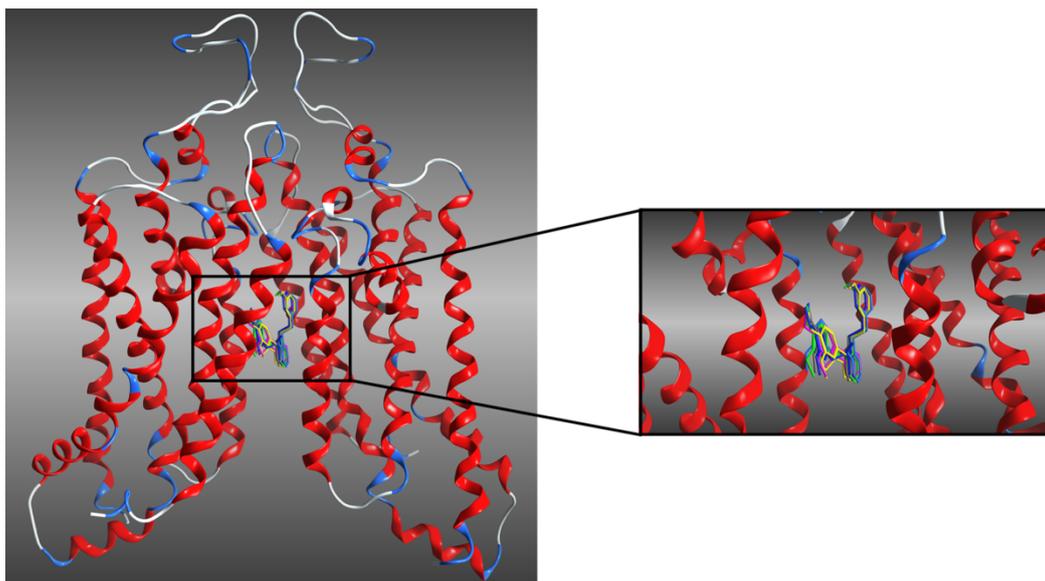
Figure 4: Top scored poses of the 6 arylmethyloxyphenyl derivatives. SM-562 (Green), SM-565 (Blue) GP199-1 (Yellow), GP196-2 (Pink), GS4 (Grey), GS3 (Dark blue)
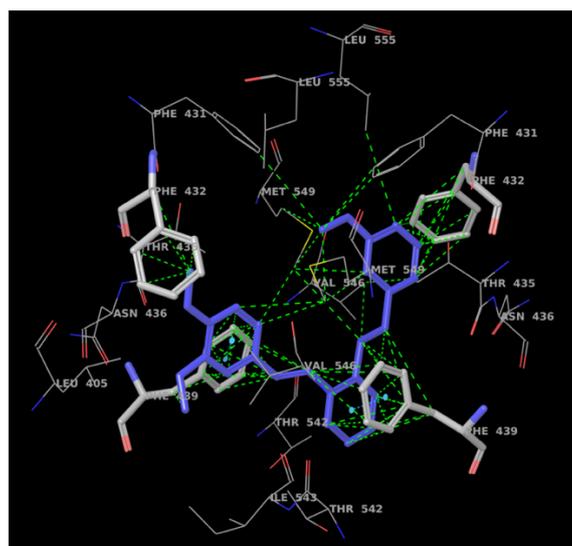
Figure 5: Hydrophobic interactions for SM-562. Ring A interacts with Phe432 (chain A), ring B with Phe439 (chain A), ring C with Phe439 (chain B).

The retrieved binding mode (cluster 2) was compared with the SAR found in the six analogues. Our binding hypothesis provides a rationale for the highest activity of SM-562 (IC50 = 0.12 μM) and SM-565 (IC50 = 0.30 μM) in the dataset. Poses of these ligands can adopt a conformation that allows additional pi-pi interaction of ring C with Phe439 (chain B) (Figure 6). The SAR shows that additional -OCH3 groups at ring A or C lead to increase in activity, as reflected in compounds SM-562 and SM-565. These facts align with our binding mode as the additional -OCH3 groups would occupy the hydrophobic sub pockets

surrounding ring A and C more efficiently (Figure 7). In detail, the methoxy moiety (-OCH3) at ring A interacts with Phe431 (chain A, chain B) and Met 549 (chain A) through hydrophobic interactions. The methoxy moiety at ring C at meta and para position show strong hydrophobic interaction with Phe 439 (chain B) and Phe 432 (chain B) respectively (Figure 5). Further evidence is found in the compound pair GP199-1 and GP196-2. Here the loss of methoxy group in ring A (GP196-2) leads to 4 times lower activity. This could be due to the loss of hydrophobic interactions between the methoxy moiety and Phe 431 (chain A, chain B), and Met 549 (chain A). Similar observation was obtained for GS4 (I50= 7.1 µM) and GS3 (I50= 18 µM).

(A)   **(B)**
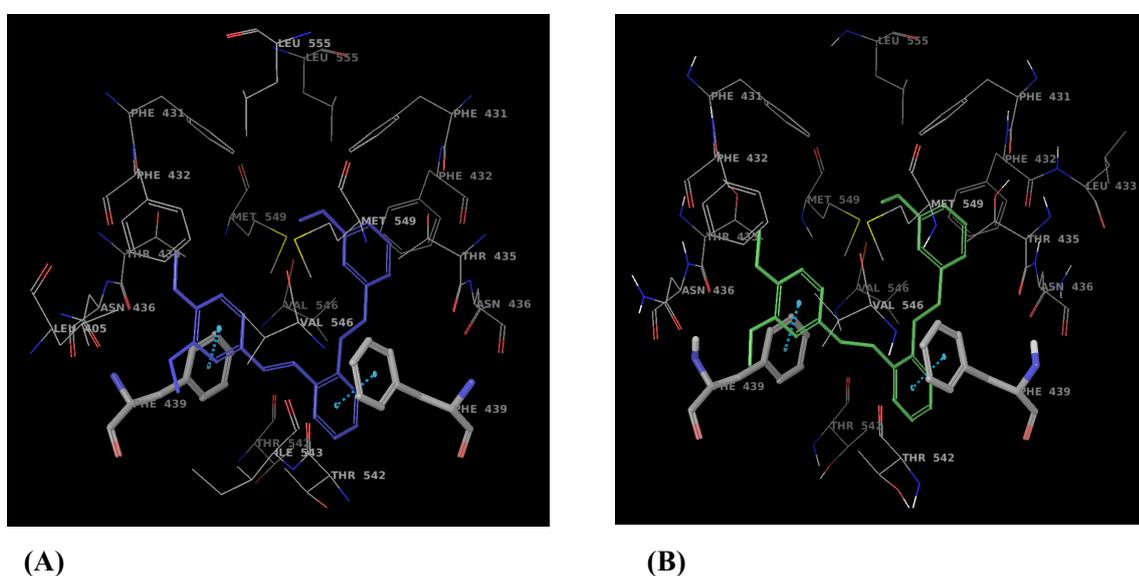
Figure 6: (A) SM-562, (B) SM-565 showing Pi-Pi interactions of ring B with Phe439 (chain A) and ring C with Phe439 (chain B).
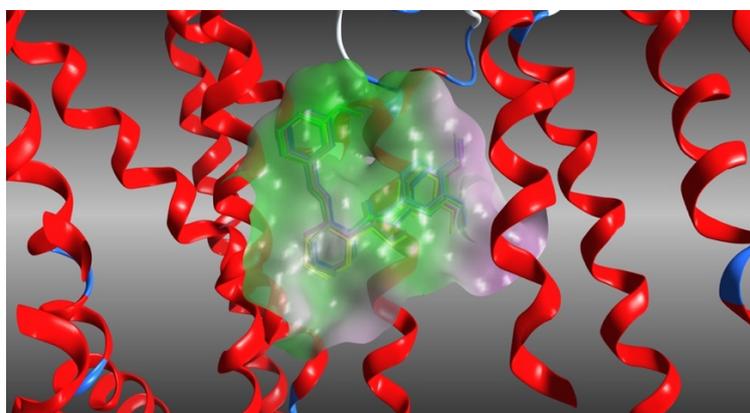
Figure 7: Top poses of the 6 arylmethyloxyphenyl derivatives aligned in the hydrophobic binding pocket. SM-562 (Green), SM-565 (Blue) GP199-1 (Yellow), GP196-2 (Pink), GS4 (Grey), GS3 (Dark blue)

GP199-1 differs from SM-565 by an additional hydroxyl group in the linker between ring B and C. The addition of the polar substituent led to a 4-fold decrease in activity, which can also be reflected in our binding hypothesis. The introduced polar hydroxyl-group is partly solvent exposed but it also placed in unfavourable hydrophobic environment shaped by phenylalanine. The activity of arylmethyloxyphenyl is further diminished by exchanging the hydroxyl group in the linker by a carbonyl moiety, exemplified by the GP199-1(1.4) and GS4 (7.1) pair. While GP199-1 is flexible enough to sustain partial solvation in our binding hypothesis, the introduction of a carbonyl-moiety leads to a twist in ligand conformation, forcing the carbonyl-moiety to be deeply buried in the hydrophobic pocket (Figure 8). Additionally we performed docking pose analysis using SeeSAR[18] to check for the desolvation penalty for the carbonyl analogues GS4 and GS3 versus its hydroxy analogues GP199-1 and GP196-2, respectively. For the carbonyl (=O) of GS4, we observed Hyde score[19, 20] of +5.5 KJ/mol (ligand desolvation energy of +6.4 KJ/mol and receptor desolvation energy of -1.1KJ/mol). When this carbonyl moiety was replaced by hydroxy (GP199-1), the Hyde score was only -0.1 KJ/mol (ligand desolvation energy of +1.1 KJ/mol and receptor desolvation energy of -1.2 KJ/mol). Similar outcome was observed for GS3(ligand desolvation energy of +6.3 KJ/mol and receptor desolvation energy of -3.6KJ/mol) and GP196-2 (ligand desolvation energy of +0.7 KJ/mol and receptor desolvation energy of -1.2 KJ/mol)). Thus the drop in GS4 and GS3 activity to GP199-1 and GP196-2, respectively could be due to the desolvation penalty of carbonyl moiety at the linker.
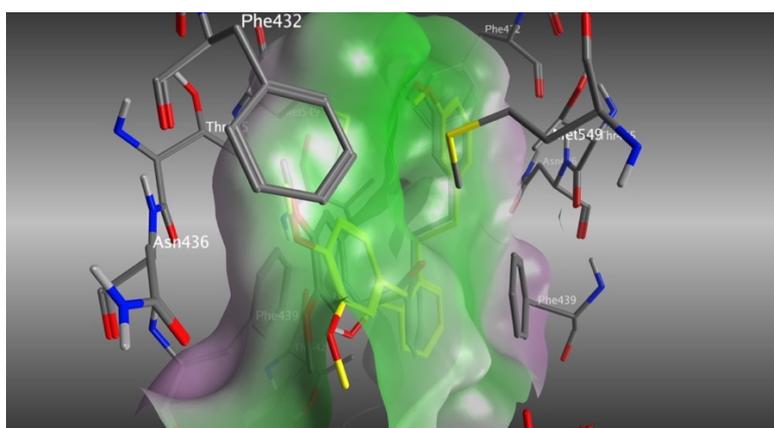
Figure 8: Pose orientation of GP199-1 (Yellow), GS4 (Grey).

Conclusion

In this study, we identified the binding mode of arylmethyloxyphenyl analogues in BCRP by means of molecular docking. Our binding hypothesis suggests that the activity of arylmethyloxyphenyl derivatives is driven by strong hydrophobic interactions. In order to overcome the difficulties of docking scoring functions in pose ranking, we applied an unconventional protocol that prioritized poses which show a high degree of SAR congruency. The pose evaluation leads to one sound binding mode, which after additional experimental validation can guide rational optimization of this compound class towards high potency. Furthermore, the uncovered ligand orientation may also be helpful to improve the mechanistic understanding of BCRP inhibition and could invoke the design of novel experiments. While further validations remain to be performed, we report here for the first time a binding hypothesis for arylmethyloxyphenyl inhibitors of BCRP that fit with the experimental data.

References

1. Colabufo NA, Berardi F, Perrone R, et al (2006) Arylmethyloxyphenyl derivatives: small molecules displaying P-glycoprotein inhibition. J Med Chem 49:6607–6613 . doi: 10.1021/jm060639z

2. Colabufo NA, Berardi F, Perrone R, et al (2008) Synthesis and biological evaluation of (hetero)arylmethyloxy- and arylmethylamine-phenyl derivatives as potent P-glycoprotein modulating agents. J Med Chem 51:1415–1422 . doi: 10.1021/jm701267q

3. Colabufo NA, Berardi F, Perrone R, et al (2008) 2-[(3-Methoxyphenylethyl)phenoxy]-based ABCB1 inhibitors: effect of different basic side-chains on their biological properties. J Med Chem 51:7602–7613 . doi: 10.1021/jm800928j

4. Montanari F, Cseke A, Wlcek K, Ecker GF (2017) Virtual Screening of DrugBank Reveals Two Drugs as New BCRP Inhibitors. SLAS Discov Adv Life Sci RD 22:86–93 . doi: 10.1177/1087057116657513

5. Schwarz T, Montanari F, Cseke A, et al (2016) Subtle Structural Differences Trigger Inhibitory Activity of Propafenone Analogues at the Two Polyspecific ABC Transporters: P-Glycoprotein (P-gp) and Breast Cancer Resistance Protein (BCRP). ChemMedChem 11:1380–1394 . doi: 10.1002/cmdc.201500592

6. Taylor NMI, Manolaridis I, Jackson SM, et al (2017) Structure of the human multidrug transporter ABCG2. Nature 546:504–509 . doi: 10.1038/nature22345

7. Kage K, Tsukahara S, Sugiyama T, et al (2002) Dominant-negative inhibition of breast cancer resistance protein as drug efflux pump through the inhibition of S-S dependent homodimerization. Int J Cancer 97:626–630

8. Litman T, Jensen U, Hansen A, et al (2002) Use of peptide antibodies to probe for the mitoxantrone resistance-associated protein MXR/BCRP/ABCP/ABCG2. Biochim Biophys Acta 1565:6–16

9. Schrödinger Release 2015-1: Maestro, version 10.1, Schrödinger, LLC, New York, NY, 2015.

10. Schrödinger Release 2015-1: LigPrep, version 3.3, Schrödinger, LLC, New York, NY, 2015.

11. Shelley JC, Cholleti A, Frye LL, et al (2007) Epik: a software program for pK(a) prediction and protonation state generation for drug-like molecules. J Comput Aided Mol Des 21:681–691 . doi: 10.1007/s10822-007-9133-z

12. Greenwood JR, Calkins D, Sullivan AP, Shelley JC (2010) Towards the comprehensive, rapid, and accurate prediction of the favorable tautomeric states of drug-like molecules in aqueous solution. J Comput Aided Mol Des 24:591–604 . doi: 10.1007/s10822-010-9349-1

13. Watts KS, Dalal P, Murphy RB, et al (2010) ConfGen: a conformational search method for efficient generation of bioactive conformers. J Chem Inf Model 50:534–546 . doi: 10.1021/ci100015j

14. Sastry GM, Adzhigirey M, Day T, et al (2013) Protein and ligand preparation: parameters, protocols, and influence on virtual screening enrichments. J Comput Aided Mol Des 27:221–234 . doi: 10.1007/s10822-013-9644-8

15. Verdonk ML, Cole JC, Hartshorn MJ, et al (2003) Improved protein–ligand docking using GOLD. Proteins Struct Funct Bioinforma 52:609–623 . doi: 10.1002/prot.10465

16. Jones G, Willett P, Glen RC, et al (1997) Development and validation of a genetic algorithm for flexible docking. J Mol Biol 267:727–748 . doi: 10.1006/jmbi.1996.0897

17. R Core Team (2013). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria

18. BioSolveIT GmbH - SeeSAR. https://www.biosolveit.de/SeeSAR/. Accessed 23 Jan 2018

19. Reulecke I, Lange G, Albrecht J, et al (2008) Towards an Integrated Description of Hydrogen Bonding and Dehydration: Decreasing False Positives in Virtual Screening with the HYDE Scoring Function. ChemMedChem 3:885–897 . doi: 10.1002/cmdc.200700319

20. Schneider N, Lange G, Hindle S, et al (2013) A consistent description of HYdrogen bond and DEhydration energies in protein–ligand complexes: methods behind the HYDE scoring function. J Comput Aided Mol Des 27:15–29 . doi: 10.1007/s10822-012-9626-2

4.3 Data transferability for Predictive *in silico* Modeling

The efflux transporter P-glycoprotein (P-gp) is a protein of high interest in drug discovery among other major anti-targets. In early stages of drug development, the pharmacokinetic and toxicity profiles of a drug candidate are determined in animal models (usually rodents) before being tested in humans. European Union initiatives such as the Horizon 2020 EU-ToxRisk project (www.eu-toxrisk.eu) drive the required paradigm shift in toxicological testing from 'black box' animal testing towards a toxicological assessment based on human cell responses [189–191]. Similar initiatives across the world are progressing towards the 3R goals - refinement, reduction and replacement of animal trials [192–194].

In the light of this, besides developing predictive *in silico* models for the identification of inhibitors of human P-gp, it is beneficial to establish predictive models for mouse and rat to reduce the number of compounds to be tested in later stages. Though a substantial amount of experimental data against human P-gp is already available and has been utilized for the development of *in silico* models [195, 196], sufficient data is not available to build predictive models for rat and mouse P-gp. Further, lack of availability of an experimentally determined three-dimensional (3D) structure for human P-gp also limits the development of reliable structure-based models. Thus, employing the human P-gp data in the structure-based modeling of resolved 3D structures, for instance the mouse P-gp structure, would reveal potential ligand-target interactions with high certainty.

4.3.1 Interspecies comparison of putative ligand binding sites of human, rat and mouse P-glycoprotein

Sankalp Jain, Melanie Grandits and Gerhard F. Ecker

In the following manuscript, we used a structure-based approach to compare the binding site interaction profiles of human, rat and mouse P-gp to assess if *in vitro* human activity data could be successfully employed for development of *in vivo* prediction models for rodents. A comparison of the per-residue interaction energies of the docking poses and analysis of the protein-ligand interaction fingerprints indicate a significant overlap between the binding site interacting residues across the three species. This would help to improve our understanding of protein-ligand interactions at the molecular level, stimulating scientists to conduct new experiments and thus aid to extrapolation of molecular hypotheses from rodents to humans and *vice-versa*.

S. Jain performed the study and wrote the manuscript. M. Grandits and G.F. Ecker supervised the work and revised the manuscript.

Manuscript Number:

Title: Interspecies comparison of putative ligand binding sites of human, rat and mouse P-glycoprotein

Article Type: Research Paper

Keywords: Species differences; P-glycoprotein; binding site comparison; transmembrane domain; protein-ligand interaction fingerprint

Corresponding Author: Professor Gerhard F. Ecker, Dr.

Corresponding Author's Institution: University of Vienna

First Author: Sankalp Jain, MSc

Order of Authors: Sankalp Jain, MSc; Melanie Grandits, Dr.; Gerhard F Ecker, Univ.-Prof. Dr.

Manuscript Region of Origin: AUSTRIA

Abstract: Prior to the clinical phases of testing, safety, efficacy and pharmacokinetic profiles of lead compounds are evaluated in animal studies. These tests are primarily performed in rodents, such as mice and rats. In order to reduce the number of animal experiments, computational models that predict the outcome of these studies and thus aid in prioritization of preclinical candidates are heavily needed. However, although computational models for human off-target interactions with decent quality are available, they cannot easily be transferred to rodents due to lack of respective data. In this study, we assess the transferability of human P-glycoprotein activity data for development of in silico models to predict in vivo effects in rats and mouse using a structure-based approach. P-glycoprotein (P-gp) is an ATP-dependent efflux transporter that transports xenobiotic compounds such as toxins and drugs out of cells and has a broad substrate and inhibitor specificity. It influences the bioavailability and toxicity of drugs and plays a major role in multidrug resistance. Comparing the binding site interaction profiles of human, rat and mouse P-gp derived from docking studies with a set of common inhibitors suggests that the inhibitors share potentially similar binding modes. These findings encourage the use of in vitro human P-gp data for predicting in vivo effects in rodents and thus contributes to the 3Rs of animal experiments.

Interspecies comparison of putative ligand binding sites of human, rat and mouse P-glycoprotein

Sankalp Jain[a], Melanie Grandits[a] and Gerhard F. Ecker[a*]

[a]University of Vienna, Department of Pharmaceutical Chemistry, Althanstrasse 14, 1090 Vienna, Austria

*Corresponding author: E-Mail: gerhard.f.ecker@univie.ac.at; Phone: +43-1-4277-55110; eFax: +43-1-4277-855110

Graphical Abstract

Abstract

Prior to the clinical phases of testing, safety, efficacy and pharmacokinetic profiles of lead compounds are evaluated in animal studies. These tests are primarily performed in rodents, such as mice and rats. In order to reduce the number of animal experiments, computational models that predict the outcome of these studies and thus aid in prioritization of preclinical candidates are heavily needed. However, although computational models for human off-target interactions with decent quality are available, they cannot easily be transferred to

rodents due to lack of respective data. In this study, we assess the transferability of human P-glycoprotein activity data for development of *in silico* models to predict *in vivo* effects in rats and mouse using a structure-based approach. P-glycoprotein (P-gp) is an ATP-dependent efflux transporter that transports xenobiotic compounds such as toxins and drugs out of cells and has a broad substrate and inhibitor specificity. It influences the bioavailability and toxicity of drugs and plays a major role in multidrug resistance. Comparing the binding site interaction profiles of human, rat and mouse P-gp derived from docking studies with a set of common inhibitors suggests that the inhibitors share potentially similar binding modes. These findings encourage the use of *in vitro* human P-gp data for predicting *in vivo* effects in rodents and thus contributes to the 3Rs of animal experiments.

Keywords:

Species differences, P-glycoprotein, binding site comparison, transmembrane domain, protein-ligand interaction fingerprint.

1. Introduction

The efflux transporter P-glycoprotein (P-gp) is a protein of high interest among other major anti-targets (Cramer et al., 2007). It is expressed in tissues such as intestine, liver, kidney, placenta, testis, and in the capillary endothelial cells of the brain (Seelig, 1998; Thiebaut et al., 1987), and plays an important role in the absorption, distribution and excretion of many drugs. Overexpression of P-gp has been implicated in resistance to multiple chemotherapeutic drugs and is a widely accepted mechanism underlying multidrug resistance (Aller et al., 2009; Fojo et al., 1987; Widmer et al., 2003). Co-administration of a P-gp inhibitor with a drug can lead to altered disposition of the latter, resulting in elevated plasma levels of the drug which could lead to adverse effects (Bussey, 1982; Tsuji, 2002; Verschraagen et al., 1999). In this respect, the United States Food and Drug Administration (US FDA) guidance requires new drug candidates to be routinely screened against P-gp as part of the clinical drug interaction studies ("Clinical Drug Interaction Studies — Study Design, Data Analysis, and Clinical Implications Guidance for Industry," 2017; Klepsch et al., 2011). Therefore, computational methods that characterize P-gp interactions and thus guide the prioritization of compounds in the early phase of the drug discovery process are of considerable interest (Schneider, 2010).

In early stages of drug development, pharmacokinetic and toxicity profiles of a candidate drug are evaluated in animal models (typically rats or mouse) prior to the clinical phases of testing in humans. A substantial amount of experimental data against human P-gp is already available and has been utilized for the development of *in silico* models (see e.g. livertox.univie.ac.at). However, besides developing *in silico* models for the prediction of ligands for human P-gp, it would be beneficial to also establish models for rat and mouse P-gp in order to predict the outcomes of preclinical animal studies. Unfortunately, limited availability of experimental data for rat and mouse P-gp restricts the development of such models. In this context the question arises, whether predicted interaction profiles of ligands with human P-gp could be transferred to rodent P-gp. This would require a comprehensive comparison of the putative binding sites of the P-gp structures across species. Literature sheds little light on this, suggesting the need for exploration of species-related differences in P-gp mediated drug transport activity (Martignoni et al., 2006; Schwab et al., 2003; Suzuyama et al., 2007).

Inhibition of P-gp activity as a result of drug interactions has been reported in both animals and humans (Bussey, 1982; Choo et al., 2000; Pedersen, 1985), but only a few studies discussed species-related differences in the inhibitory effects on the P-gp function (Chu et al., 2013; Suzuyama et al., 2007; Zolnerciks et al., 2011). A few studies proposed moderate species differences, human *vs*. rat (Molden et al., 2000), human *vs*. mouse (Adachi et al., 2001; Lin and Yamazaki, 2003) and also among the three species (human *vs*. rat *vs*. mouse) (Katoh et al., 2006), while a few other studies reported no significant differences between human, rat and mouse P-gp (Chu et al., 2013; Feng et al., 2008; Hsiao and Unadkat, 2012). However, it must be noted that only a small number of compounds were tested in these studies. It might thus well be that the inhibitory effects on P-gp-mediated drug transport are subjective to both the chemical structure of substrates/inhibitors and to the species. Moreover, it is not yet clear if the possible species differences in the inhibitory effects of P-gp activity are due to differences in binding site residues of P-gp, which is therefore worth investigating.

To the best of our knowledge, no computational study compared the binding site interaction profiles of P-gp across different species (human, rat and mouse) so far. In this study, we used a structure-based approach to compare their binding sites in order to derive information concerning potential species differences in P-gp-mediated drug transport. Since an X-ray crystal structure is available for mouse P-gp alone, homology modeling was performed to

construct the models for human P-gp and for rat P-gp. Subsequently, docking of common inhibitors of rat, mouse and human P-gp was performed. Next, known inhibitors of human P-gp were docked into the models of the three species followed by an analysis of the interactions between the inhibitors and binding site residues. The interaction profiles of the P-gp binding sites of the three species were then compared to evaluate the transferability of *in vitro* human P-gp data for development of models to predict effects in rat and mouse.

2. Methodology

2.1. Dataset

A substantial amount of human P-gp data is made publicly available through previous literature reports (Broccatelli et al., 2011; Chen et al., 2011; Klepsch et al., 2014). However, due to the limited availability of rat P-gp data in public domain bioactivity databases such as ChEMBL(Gaulton et al., 2012; Willighagen et al., 2013) and BindingDB(Liu et al., 2007), an exhaustive literature search was performed. A total of 18 rat P-gp inhibitors could be identified that are known to also inhibit both human P-gp and mouse P-gp. Due to the inconsistencies in the assay conditions, these compounds unfortunately could not be utilized to compare inhibitory profiles across the species. Suzuyama *et al*. (Suzuyama et al., 2007) studied the species differences (human, monkey, canine, rat and mouse) in the inhibitory effects of the prototype P-gp inhibitors quinidine and verapamil. These two drugs served as the starting point for in silico comparison of binding site interaction profiles across the species. Further, we also extracted the human P-gp data from Broccatelli *et al*. (Broccatelli et al., 2011) in order to perform protein-ligand interaction fingerprint (PLIF) analysis and to identify the common functional group residue interactions among the three species. The dataset was standardized according to the procedure described in Pinto *et al*., 2012. (Pinto et al., 2012) The final dataset contained a total of 1161 compounds (612 inhibitors and 549 non-inhibitors).

2.2. Homology modeling

For human P-gp (UNIPROT ID: P08183), rat P-gp (MDR1a-UNIPROT ID: Q9JK64; MDR1b-UNIPROT ID: P43245) and mouse P-gp (mdr1b-UNIPROT ID: P06795), the corrected mouse P-gp structure (mdr1a-UNIPROT ID: P21447; PDB ID: 4M1M) was selected as the most structurally related template protein. Rat and mouse P-gp proteins are encoded by two paralogous genes namely MDR1a and MDR1b that show a sequence identity

of 83% (Chu et al., 2013; Devault and Gros, 1990). Therefore, we constructed in total four homology models to consider the paralogs too. Homology models were constructed using MODELLER 9.13 (Eswar et al., 2007) and the Prime module in Maestro (Schrödinger, Inc. V-10.1.013)(Jacobson et al., 2004, 2002). The energy minimized models were further evaluated using DOPE score (Shen and Sali, 2006) and GA341 score (John and Sali, 2003; Melo et al., 2002). Quality of the stereochemical parameters and the normality of the structures were checked using the PROCHECK program, included in the PDBsum analysis (Laskowski et al., 1993). Ramachandran plot (Zhou et al., 2011) and G-factor (Engh and Huber, 1991), and finally the Q-score (Benkert et al., 2008, 2009) values were evaluated to identify the best homology models. The electrostatic potential surface (EPS) of each of the three best models for the three species was also calculated and compared using MOE 2013 (*Molecular Operating Environment (MOE), 2013.08*, n.d.).

2.3. Sequence alignment

Sequence alignment was performed using ClustalX (Larkin et al., 2007) and verified by including secondary structure predictions. Subsequently, the alignment was analyzed using Jalview (Supplementary Fig. S1-S4) (Clamp et al., 2004; Waterhouse et al., 2009).

2.4. Binding site identification and molecular docking

In order to avoid any bias, the binding site for all five structures was defined as the complete transmembrane region, taking 20 Å around the coordinate of the center point to allow subsequent flexible docking of a series of P-gp inhibitors. The protein was prepared using the Protein Preparation Wizard of the Schrödinger Suite (2015) (Sastry et al., 2013; *Schrödinger Release 2015-1: Maestro, version 10.1, Schrödinger, LLC, New York, NY, 2015.*, n.d.). Hydrogen atoms were added, and optimal protonation states and ASN/GLN/HIS flips were determined. To assess their correct protonation states, ligands were prepared using the LigPrep module of the Schrödinger Suite, (*Schrödinger Release 2015-1: LigPrep, version 3.3, Schrödinger, LLC, New York, NY, 2015.*, n.d., *Schrödinger Release 2015-1: Maestro, version 10.1, Schrödinger, LLC, New York, NY, 2015.*, n.d.) which produces low-energy 3D structures that can be used for docking. The OPLS_2005 force field was used for minimization of the structures. Different ionization states were generated by adding or removing protons from the ligand at a target pH of 7.0 ± 2.0 using Epik version 3.1., (Greenwood et al., 2010; Shelley et al., 2007) and tautomers were generated for each ligand. To generate stereoisomers, the information on chirality from the input file for each ligand

was retained as is for the entire calculation. All docking runs were performed in high-throughput mode with GlideXP(Friesner et al., 2006; Halgren et al., 2004) docking in Maestro. We also used the genetic algorithm-based GOLD suit (version 5.2.0) (Jones et al., 1997; Verdonk et al., 2003) for docking.

2.5. Protein ligand interaction fingerprint (PLIF)

A PLIF summarizes the interactions between a ligand and a protein using a molecular fingerprint scheme. We generated two types of PLIFs that differ in the information encoded. The first PLIF encodes residues involved in an interaction with the ligand at each bit position. The second type encodes the functional group of the ligand that interacts with the residue. For this, the substructure patterns of 100 functional groups (in SMARTS notation) were extracted from the Daylight website (http://www.daylight.com/dayhtml_tutorials/languages/smarts/smarts_examples.html#GROU P). All PLIF bits were calculated with the MOE 2013 (*Molecular Operating Environment (MOE), 2013.08*, n.d.) built-in function CalculateRawInteractions using a 1% threshold for molecular interactions and a 20% threshold for surface contacts. The function was embedded in an in-house SVL script and was post-processed to enable calculation of functional group PLIFs.

3. Results and Discussion

Predicting interactions of small molecules with membrane protein structures has always been challenging. Nevertheless, visualization of the 3D models contributes to the comprehension of the physical and chemical properties of these biomolecules, and of their intermolecular interactions with endogenous and exogenous compounds. Due to the lack of crystal structures for human and rat P-gp, homology modeling and computational ligand docking were used to generate structure-based hypotheses for protein-ligand-interactions.

3.1. Homology modeling

ABC transporters are transmembrane proteins that are in general difficult to be resolved *via* crystallization (Klepsch et al., 2010). In such cases, homology modelling is the method of choice for structure-based studies. The homology models generated in this study resemble the

open-inward (or apo/ground) state of P-gp. This state was considered because it resembles the first step of the basic catalytic cycle for drug-binding in P-gp (Wilkens, 2015).

Since January 2014, a refined X-ray structure of a eukaryotic ABC efflux pump, ABCB1 (mouse) is available (Li et al., 2014) (PDB code: 4M1M, resolution: 3.8 Å). High sequence identities with human MDR1 (86%), rat MDR1a (94%), rat MDR1b (82%), mouse mdr1b (83%) and a moderate resolution of 3.8 Å renders 4M1M a reasonable template for homology modeling (Pajeva et al., 2009). Moreover, the secondary structure elements (NBDs and TMDs) are also conserved among the species. When only the TMD was analysed, the sequence identity is greater than 85% for all structures (Supplementary Fig. S5). The best models had a normalized Dope score of less than −0.6, G-factors less than −0.12, and Qmean scores of greater than 0.60 (see Table 1). For all modelled structures, the Ramachandran plot (Supplementary Fig. S6-S9) showed excellent results with less than 1.9% of residues in generously allowed or disallowed regions. All of these residues are located in the nucleotide binding domains (NBD) or extracellular loops (ECL) and are therefore not involved in drug binding (Supplementary Fig. S10-S13). Table 1 summarizes the model assessment details for the best structure. The X-ray crystal structure and site directed mutagenesis studies on ABCB1 serve as validity tests for both helix orientation in the template (Ward et al., 2007), and the alignment used for ABC transporter modelling (Supplementary Fig. S1-S4). The homology models as well as the crystal structure displayed a V-shaped structure with analogos domain orientations.

Table 1: Results from the stereochemical validation of the homology models.

Model	Dope score	G-factor	Qmean score	Residues in generously allowed or disallowed regions (%)
Human MDR1	-0.633	-0.13	0.68	1.7
Rat MDR1a	-0.795	-0.03	0.70	1.7
Rat MDR1b	-0.703	-0.16	0.65	1.8
Mouse mdr1b	-0.808	-0.06	0.69	2.0

3.2. Sequence alignment and binding site analysis

The amino acid sequence is highly conserved among the three species (Supplementary Fig. S5), suggesting a high structural similarity (see Table 2).

Table 2: Sequence identity/similarity [%] between human, rat and mouse P-gp.

	Human-MDR1	Rat MDR1a	Rat MDR1b	Mouse mdr1a	Mouse mdr1b
Human-MDR1	100	87/93	80/90	86/92	80/90
Rat MDR1a	87/93	100	84/91	94/97	84/92
Rat MDR1b	80/90	84/91	100	82/91	93/97
Mouse mdr1a	86/92	94/97	82/91	100	83/91
Mouse mdr1b	80/90	84/92	93/97	83/91	100

Experimental techniques such as cysteine and arginine scanning and photoaffinity labeling were previously employed to determine the drug binding sites of P-gp (Loo and Clarke, 2008; Pleban et al., 2005; Seeger and van Veen, 2009; Shilling et al., 2006). Multiple binding sites were identified and binding at different sites could lead to different inhibitory effects. Well characterized binding sites are the ones of Hoechst 33342 and Rhodamine, the so called H-sites and R-site (Loo and Clarke, 2002; Qu and Sharom, 2002). Studies also suggest the presence of an allosteric regulatory site as well as a progesterone and prazosin binding region (Martin et al., 2000; Shapiro et al., 1999). The H-site and R-site residues, (characterized by Ferreira et.al (Ferreira et al., 2013)) of the three species were compared and showed a high sequence identity. This would indicate the similar arrangement of the binding sites residues and thus further pointing to the presence of a similar binding/interaction profile of the inhibitors. Mostly identical or similar residues were present in the five structures. The H-site and R-site had 77% and 65% residues identical within the three species. Those residues which show a difference, have mostly similar properties. For example, Glu180 in mouse mdr1a is replaced with Aspartic Acid in mouse mdr1b and in ratMDR1b. Both residues are

charged and have acidic properties. In a few instances, charged (basic) amino acids are replaced by polar (neutral) or hydrophobic (aliphatic) amino acids in the other species but most of these residues did not participate in interactions with docked ligands. In general, the H-site has a higher percentage of charged residues (lysine, histidine, and glutamic acid residues), while the R-site has a high number of glycine, glutamine, and proline residues. Interestingly, threonine and tyrosine were not found in the H-site and R-sites, respectively. A detailed comparison of the H-site and R-site residues of the five structures is shown in Supplementary Table S1. These observations signify the harmony of electrostatic properties and molecular features of the drug recognition site (central binding cavity) in the three species. Supplementary Fig. S14-S18 represents the electrostatic potential surface (EPS) of the substrate recognition area of each of the ABC-transporter models. The EPS of the substrate recognition area in the TMDs of the human model is neutral with negative and weakly positive areas, similar to the EPS of rat and mouse models.

3.3. Molecular Docking

In order to analyze the putative binding pocket of the transport protein in the three species, we proceeded with docking of a set of inhibitors. Ligand docking is a commonly used approach to identify ligand-protein interactions. However, in case of P-gp, this appears to be challenging due to various reasons. Firstly, P-gp possesses a high degree of flexibility with a large binding cavity consisting of multiple binding sites. Secondly, it can harbor more than one ligand simultaneously (Loo et al., 2003a; Lugo and Sharom, 2005). And finally, lack of a high resolution crystal structure of human P-gp necessitates the use of homology models, which add additional layers of uncertainty. A large binding pocket could also be seen in a recent structure (PDB id : 4M1M) wherein large cyclopeptides bind at different sites with partially overlapping residues (Li et al., 2014). Some of these residues are identical to those involved in rhodamine or verapamil binding (Loo et al., 2006; Loo and Clarke, 1997). Other studies reported different prazosin binding sites in hamster (Isenberg et al., 2001) and human P-gp (Ambudkar et al., 2003). Overall, it is understood that P-gp possesses a huge binding pocket with at least four distinct binding sites, with TM 6 as the helix primarily involved in binding (Klepsch et al., 2010). Therefore, we considered the complete TMD as drug binding site (DBS) and generated a large number of docking poses to prevent any bias introduced by scoring functions.

We started with docking of verapamil and quinidine into the binding pocket (complete TMD) of all models. These two compounds were chosen since IC_{50} values measured under the same assay conditions were available for all three species. Our study revealed that the top ranked docking poses of verapamil were found in the R-site of P-gp in all three species, which is in agreement with previous reports (Ferreira et al., 2013). The top scored docking pose for each of the five models was found in the same region of the binding pocket (R-site) are shown in Fig. 1.
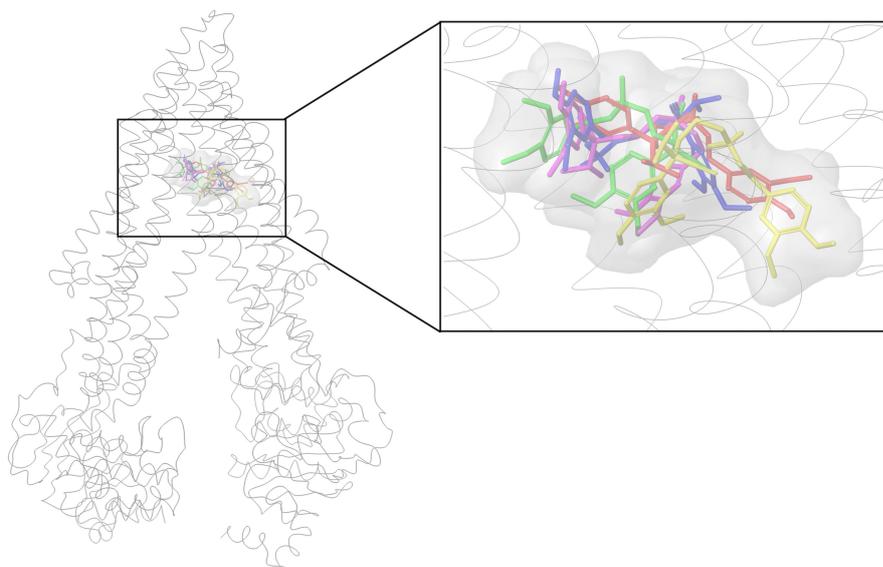
Fig. 1: Best scored docking pose of verapamil: green (human MDR1), yellow (rat MDR1a), pink (rat MDR1b), red (mouse mdr1a), blue (mouse mdr1b), secondary structure of human P-gp.

We used the GlideXP scoring function from Maestro (Friesner et al., 2006; Halgren et al., 2004) to evaluate the binding poses. GlideXP docking also provides the per residue interaction energies for a particular docking pose. For each model, the residue interaction energy (RIE) for the top scored docking poses was calculated. Phe303, Tyr307, Tyr310, Phe336, Phe343, Phe728, Phe983, Met986 and Gln990 (numbering according to human-MDR1) are residues that showed more negative interaction energy values in all three species, indicating their higher contribution to binding. Residue interaction energies for all residues which are involved in interactions with verapamil can be seen in Supplementary Table S2. For example, the residues corresponding to Tyr307 in human MDR1 are Tyr306 (RIE:-3.229 kcal/mol), Tyr306 (RIE:-3.714 kcal/mol), Try303 (RIE:-6.96 kcal/mol) and Tyr299 (RIE:-5.1

kcal/mol) in rat MDR1a, rat MDR1b, mouse mdr1a and mouse mdr1b, respectively. Each of these residues contributes to the binding with more negative interaction energy. We observed less negative RIE values with residues which are different within the species (e.g. human MDR1: Met68, rat MDR1a: Leu67, rat MDR1b: Leu66, mouse mdr1a and mdr1b: Met67), suggesting their small influence on - and involvement in - interactions with verapamil. Thus, the comparison of the per residue interaction energies of the best docking pose of the three species revealed that similar binding site residues (as per the alignment) are involved in strong interactions with the ligand (see Fig. 2).
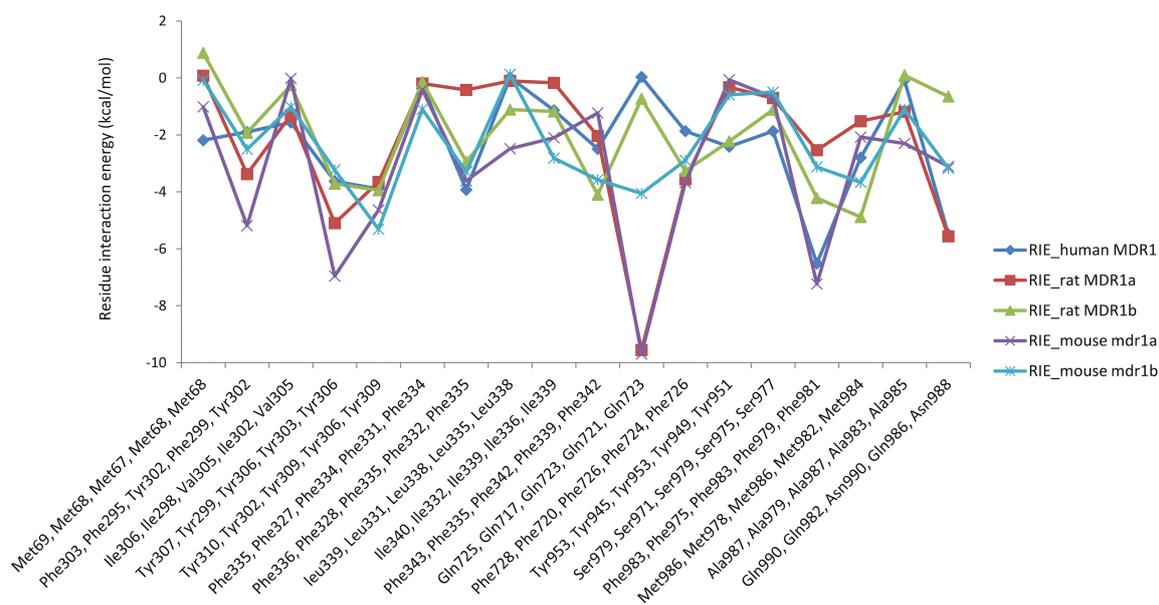
Fig. 2: Residue interaction energy for common interaction residues in human MDR1, rat MDR1a, rat MDR1b, mouse mdr1a and mouse mdr1b. x-axis denotes residue number in the order human MDR1, rat MDR1a, rat MDR1b, mouse mdr1a and mouse mdr1b, y-axis denotes the corresponding residue interaction energy (kcal/mol).

In case of quinidine, the human MDR1 residues Phe336 (RIE:-2.41 kcal/mol), Gln725 (RIE:-11.577 kcal/mol), Phe728 (RIE:-2.479 kcal/mol), Ser979 (RIE:-1.535 kcal/mol), Phe983 (RIE:-8.114 kcal/mol) and Met986 (RIE:-2.162 kcal/mol) interacted with greater negative interaction energies. The corresponding residues in rat MDR1a, rat MDR1b, mouse mdr1a and mouse mdr1b that demonstrated more negative interaction energies can be found in Supplementary Table S3. Supplementary Fig. S19 shows the RIE for common residues involved in interaction in human MDR1, rat MDR1a, rat MDR1b, mouse mdr1a and mouse mdr1b with the top scored docking pose of quinidine. Replacing a phenyl alanine in human

116

(Phe303) and mouse (Phe299 in mdr1a and mdr1b) with another hydrophobic residue, for instance Tyr302 in MDR1a and MDR1b, still showed negative RIE values. Supplementary Fig. S20 shows the top scored docking poses of quinidine in the five models.

Thus, similar amino acids, as observed with verapamil and quinidine, also confirm the homogeneous nature of the binding site residues in different species. This would further support the hypothesis of similarity in their binding sites. Site directed mutagenesis studies on human ABCB1 also indicated that Ile306 (TMH5) (Loo et al., 2006; Loo and Clarke, 2005), Ile340 (TMH6) (Loo and Clarke, 2002), Phe343 (TMH6) (Loo et al., 2003b, 2006), Phe728 (TMH7) (Loo et al., 2006), and Val982 (TMH12) (Loo and Clarke, 2002, 2005) may participate in ligand binding. As shown in Fig. 3, these residues may form a substrate recognition site in the human ABCB1 model. The involvement of these residues in ligand binding was also confirmed by Li et al (Aller et al., 2009; Li et al., 2014).
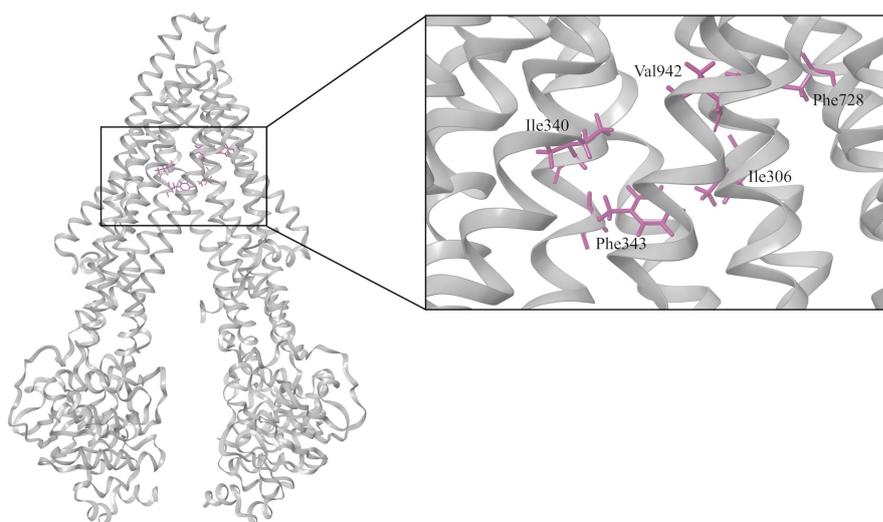
Fig. 3: Key residues of the substrate recognition site in the human ABCB1 model from literature (Loo et al., 2003b, 2006, Loo and Clarke, 2002, 2005).

In our previous work, we demonstrated that the Chemscore scoring function from the GOLD docking suit facilitated docking-based classification of inhibitors and non-inhibitors for P-gp(Klepsch et al., 2014) and the bile salt export pump (BSEP) (Jain et al., 2017) with reasonable accuracies. Therefore, we used the Chemscore scoring function to perform docking of all human P-gp inhibitors into human MDR1, rat MDR1a, rat MDR1b, mouse

mdr1a, and mouse mdr1b structures in order to compare the interaction profiles of the binding site residues in the three species *via* PLIF analysis.

3.4. Protein ligand interaction fingerprint (PLIF)

Maestro allows computation of different molecular interactions between binding site residues and a ligand in a specific pose. A PLIF summarizes the interactions between a ligand and a protein using a fingerprint scheme. It provides a detailed picture of the binding modes of different inhibitors. We retrospectively analyzed the PLIFs of complexes of verapamil and quinidine with structures of all three species (human MDR1, rat MDR1a, rat MDR1b, mouse mdr1a and mouse mdr1b) derived from docking, in order to compare their interaction profiles. In case of verapamil, in human MDR1, around 70% of the poses showed hydrophobic interactions with Phe336, Ile340, Phe343, Phe728 and Met986. Also, over 85% of the poses displayed interaction with Met69, Tyr310, Tyr 953, Phe983 (Fig. 4). In the rat structure, more than 73% of the residues showed interaction with Phe328, Phe335, Phe720, Met978 and over 85% residues showed interaction with Phe295, Ile298, Tyr299, Tyr302, Phe975 (Fig. 4). The percentage of binding poses in which specific residues are involved in interactions with verapamil in rat MDR1b, mouse mdr1a, and mouse mdr1b models can also be seen in Fig. 4. Supplementary Fig. S21 provides the same information for quinidine.
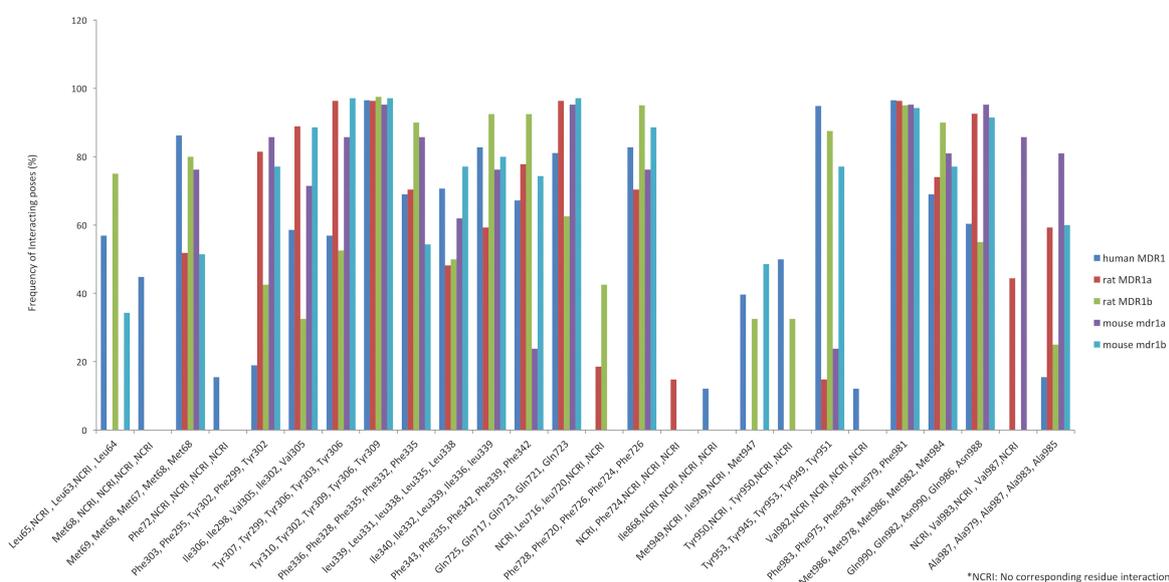
Fig. 4: Hydrophobic interactions common in human MDR1, rat MDR1a, rat MDR1b, mouse mdr1a and mouse mdr1bfor verapamil. X-axis denotes residue number in the order human MDR1, rat MDR1a, rat MDR1b, mouse mdr1a and mouse mdr1b, Y-axis denotes frequency of interacting residues (%).

For both verapamil and quinidine, the interacting residues that contributed to a significant number of binding poses were at similar positions in the 3D structures of the five transporters (Fig. 4, Supplementary Fig. S21). Specifically, for verapamil, Tyr310 (human MDR1), Tyr302 (rat MDR1a), Tyr310 (rat MDR1b), Tyr309 (mouse mdr1a and mdr1b) showed an interaction in more than 95% of the poses in all five docked structures. Interestingly, when an amino acid in one species was replaced with another amino acid in another species, a similar percentage of docking poses interacted with this residue. For instance, the exchange of Ile340 in human MDR1 with Leucine in rat MDR1b and mouse mdr1b showed hydrophobic interactions in almost 80% of the poses for verapamil, indicating that the interaction pattern did not change when two hydrophobic residues were interchanged. Similar PLIF-based observations could be inferenced after evaluation of the docking poses of quinidine for the three species. However, due to the lower degree of freedom (flexibility) of quinidine, relatively fewer docking poses could be obtained.

We also identified the interacting residues for a set of 612 human P-gp inhibitors that were docked into these five structures. For human P-gp, more than 70% of the inhibitors interacted with Ile306, Tyr310, Phe336, Phe343, Tyr953, Phe983 and Met986. Fig. 5 shows interacting residues common to human, rat and mouse structures. Supplementary Table S4 lists the occurrence of commonly interacting residues in the three species. PLIFs obtained from docking of a diverse set of human P-gp inhibitors into the five models revealed that similar residues were involved in the interactions, thereby further strengthening the existence of analogous binding site residues and interaction profiles in the three species.

Fig. 5: Hydrophobic interactions common in human MDR1, rat MDR1a, rat MDR1b, mouse mdr1a and mouse mdr1b for human P-gp inhibitors. X-axis denotes residue number in the order human MDR1, rat MDR1a, rat MDR1b, mouse mdr1a and mouse mdr1b, y-axis denotes frequency of interacting residues (%).

3.5. Analysis of the interactions of functional groups with protein residues

Investigation of the functional group-residue interactions for the set of 612 P-gp inhibitors docked revealed similar interaction patterns with all three species. Functional groups ether, carbonyl, alkyl carbon, nitrogen and arene showed more prominent interactions with Tyr310, Phe343, Phe983, and Met986 (numbering as per human MDR1). Corresponding residues in other species that participated in interactions are shown in Supplementary Table S5. To illustrate the outcomes, a heat map (Fig. 6, Supplementary Fig. S22-S25) plotting the interacting residues against the functional groups of ligands was generated. The color scale of the heat map represents the number of inhibitors involved in a particular interaction between a specific residue and a specific functional group. Thus, the most significant interactions could be visually identified.

120

Fig. 6: Heat map illustrating the PLIF analysis of the human P-gp inhibitors for human MDR1. X-axis denotes contact residues. Y-axis denotes functional groups of the ligand which are showing an interaction with the residue. Color scale signifies the number of interacting ligands.

In the three species, the H-site and R-site showed high sequence identity, suggesting the presence of similar residues at specific positions in the 3D structure. Additionally, similar functional group-residue interaction patterns were observed for human (MDR1), rat (rat MDR1a, rat MDR1b) and mouse (mouse mdr1a, and mouse mdr1b) (Supplementary Fig. S22-S25). This further strengthens the idea of utilizing human P-gp activity data, collated from *in vitro* studies, for structure-based modeling of rodent ligand-target interactions.

A study by Schwab et al. (Schwab et al., 2003) reported comparable IC$_{50}$ values in a calcein-AM assay for human MDR1, mouse mdr1a and mouse mdr1b for 28 reference compounds. Zolnerciks et al.(Zolnerciks et al., 2011) also observed comparable IC$_{50}$ values for a set of compounds against human and rat P-gp transporter and also suggested that multiple P-gp substrates would be needed to accurately predict clinically significant P-gp drug interactions, in both *in vitro* and *in vivo* (including human) drug-drug interaction studies. As mentioned earlier, Suzuyama et al. (Suzuyama et al., 2007) evaluated the inhibitory effects of quinidine and verapamil on P-gp-mediated drug transport using MDR1 transfected cell lines of different species. As a common observation, although the IC$_{50}$ values differ between the species, it was less than 10-fold. This along with our molecular docking and PLIF analysis

121

results signify the possibility of similar interaction profiles in the three species (human, rat and mouse), suggesting the usability and transferability of *in vitro* human data for development of prediction models for rat and mouse.

4. Conclusion

P-glycoprotein is a transmembrane efflux transporter that plays an important role in drug absorption, disposition, metabolism, and toxicity. It is essential to investigate the interactions of P-gp with candidate drugs not only to understand the contribution of P-gp to the pharmacological properties of candidate drugs, but also to evaluate their drug-drug interaction (DDI) profiles and thereby their clinical implications. In this regard, it is important to understand the binding site interaction profiles of P-gp in rodents which is poorly addressed so far due to the limited availability of experimental data. In this communication, we compared the P-gp binding sites across human, rat and mouse using molecular docking and protein-ligand interaction fingerprint analysis. To the best of our knowledge, this is the first *in silico* study of its kind that compares the binding sites across three different species with emphasis on their inhibitory interaction profile. Our results show a significant overlap between the binding site interacting residues across the three species. This strengthens the likelihood of similar binding mode of human, rat and mouse P-gp inhibitors, thus supporting the transferability of *in vitro* human P-gp data for development of computational models to predict effects in rat and mouse. As shown recently, the incorporation of predicted ligand transporter interaction profiles increases the performance of selected in vivo toxicity prediction models. The transferability of human P-gp data to rodent in silico models might thus increase the predictivity of rodent *in vivo* toxicological outcomes, which was a major aim of the eTOX project (www.etoxproject.eu) (Briggs et al., 2012; Hartmann and Pognan, 2017). This will subsequently improve the quality of drug candidates while lowering the attrition rate during subsequent phases of drug development, and, most remarkably, reduce the number of animal experiments in preclinical studies.

Conflicts of interest

The authors have no conflicts of interest to declare.

Acknowledgment

References

Adachi, Y., Suzuki, H., Sugiyama, Y., 2001. Comparative studies on in vitro methods for evaluating in vivo function of MDR1 P-glycoprotein. Pharm. Res. 18, 1660–1668.

Aller, S.G., Yu, J., Ward, A., Weng, Y., Chittaboina, S., Zhuo, R., Harrell, P.M., Trinh, Y.T., Zhang, Q., Urbatsch, I.L., Chang, G., 2009. Structure of P-glycoprotein reveals a molecular basis for poly-specific drug binding. Science 323, 1718–1722. https://doi.org/10.1126/science.1168750

Ambudkar, S.V., Kimchi-Sarfaty, C., Sauna, Z.E., Gottesman, M.M., 2003. P-glycoprotein: from genomics to mechanism. Oncogene 22, 7468–7485. https://doi.org/10.1038/sj.onc.1206948

Benkert, P., Künzli, M., Schwede, T., 2009. QMEAN server for protein model quality estimation. Nucleic Acids Res. 37, W510-514. https://doi.org/10.1093/nar/gkp322

Benkert, P., Tosatto, S.C.E., Schomburg, D., 2008. QMEAN: A comprehensive scoring function for model quality assessment. Proteins 71, 261–277. https://doi.org/10.1002/prot.21715

Briggs, K., Cases, M., Heard, D.J., Pastor, M., Pognan, F., Sanz, F., Schwab, C.H., Steger-Hartmann, T., Sutter, A., Watson, D.K., Wichard, J.D., 2012. Inroads to Predict in Vivo Toxicology—An Introduction to the eTOX Project. Int. J. Mol. Sci. 13, 3820–3846. https://doi.org/10.3390/ijms13033820

Broccatelli, F., Carosati, E., Neri, A., Frosini, M., Goracci, L., Oprea, T.I., Cruciani, G., 2011. A novel approach for predicting P-glycoprotein (ABCB1) inhibition using molecular interaction fields. J. Med. Chem. 54, 1740–1751. https://doi.org/10.1021/jm101421d

Bussey, H.I., 1982. The influence of quinidine and other agents on digitalis glycosides. Am. Heart J. 104, 289–302.

Chen, L., Li, Y., Zhao, Q., Peng, H., Hou, T., 2011. ADME Evaluation in Drug Discovery. 10. Predictions of P-Glycoprotein Inhibitors Using Recursive Partitioning and Naive Bayesian Classification Techniques. Mol. Pharm. 8, 889–900. https://doi.org/10.1021/mp100465q

Choo, E.F., Leake, B., Wandel, C., Imamura, H., Wood, A.J., Wilkinson, G.R., Kim, R.B., 2000. Pharmacological inhibition of P-glycoprotein transport enhances the distribution of HIV-1 protease inhibitors into brain and testes. Drug Metab. Dispos. Biol. Fate Chem. 28, 655–660.

Chu, X., Bleasby, K., Evers, R., 2013. Species differences in drug transporters and implications for translating preclinical findings to humans. Expert Opin. Drug Metab. Toxicol. 9, 237–252. https://doi.org/10.1517/17425255.2013.741589

Clamp, M., Cuff, J., Searle, S.M., Barton, G.J., 2004. The Jalview Java alignment editor. Bioinforma. Oxf. Engl. 20, 426–427. https://doi.org/10.1093/bioinformatics/btg430

Clinical Drug Interaction Studies — Study Design, Data Analysis, and Clinical Implications Guidance for Industry, 2017.

Cramer, J., Kopp, S., Bates, S.E., Chiba, P., Ecker, G.F., 2007. Multispecificity of drug transporters: probing inhibitor selectivity for the human drug efflux transporters ABCB1 and ABCG2. ChemMedChem 2, 1783–1788. https://doi.org/10.1002/cmdc.200700160

Devault, A., Gros, P., 1990. Two members of the mouse mdr gene family confer multidrug resistance with overlapping but distinct drug specificities. Mol. Cell. Biol. 10, 1652–1663.

Engh, R., Huber, R., 1991. Accurate bond and angle parameters for {X}-ray protein structure refinement. Acta Cryst A47, 392–400.

Eswar, N., Webb, B., Marti-Renom, M.A., Madhusudhan, M.S., Eramian, D., Shen, M.-Y., Pieper, U., Sali, A., 2007. Comparative protein structure modeling using MODELLER. Curr. Protoc. Protein Sci. Editor. Board John E Coligan Al Chapter 2, Unit 2.9. https://doi.org/10.1002/0471140864.ps0209s50

Feng, B., Mills, J.B., Davidson, R.E., Mireles, R.J., Janiszewski, J.S., Troutman, M.D., de Morais, S.M., 2008. In vitro P-glycoprotein assays to predict the in vivo interactions of P-glycoprotein with drugs in the central nervous system. Drug Metab. Dispos. Biol. Fate Chem. 36, 268–275. https://doi.org/10.1124/dmd.107.017434

Ferreira, R.J., Ferreira, M.-J.U., dos Santos, D.J.V.A., 2013. Molecular Docking Characterizes Substrate-Binding Sites and Efflux Modulation Mechanisms within P-Glycoprotein. J. Chem. Inf. Model. 53, 1747–1760. https://doi.org/10.1021/ci400195v

Fojo, A.T., Ueda, K., Slamon, D.J., Poplack, D.G., Gottesman, M.M., Pastan, I., 1987. Expression of a multidrug-resistance gene in human tumors and tissues. Proc. Natl. Acad. Sci. U. S. A. 84, 265–269.

Friesner, R.A., Murphy, R.B., Repasky, M.P., Frye, L.L., Greenwood, J.R., Halgren, T.A., Sanschagrin, P.C., Mainz, D.T., 2006. Extra Precision Glide: Docking and Scoring Incorporating a Model of Hydrophobic Enclosure for Protein−Ligand Complexes. J. Med. Chem. 49, 6177–6196. https://doi.org/10.1021/jm051256o

Gaulton, A., Bellis, L.J., Bento, A.P., Chambers, J., Davies, M., Hersey, A., Light, Y., McGlinchey, S., Michalovich, D., Al-Lazikani, B., Overington, J.P., 2012. ChEMBL: a large-scale bioactivity database for drug discovery. Nucleic Acids Res. 40, D1100–D1107. https://doi.org/10.1093/nar/gkr777

Greenwood, J.R., Calkins, D., Sullivan, A.P., Shelley, J.C., 2010. Towards the comprehensive, rapid, and accurate prediction of the favorable tautomeric states of drug-

like molecules in aqueous solution. J. Comput. Aided Mol. Des. 24, 591–604. https://doi.org/10.1007/s10822-010-9349-1

Halgren, T.A., Murphy, R.B., Friesner, R.A., Beard, H.S., Frye, L.L., Pollard, W.T., Banks, J.L., 2004. Glide: A New Approach for Rapid, Accurate Docking and Scoring. 2. Enrichment Factors in Database Screening. J. Med. Chem. 47, 1750–1759. https://doi.org/10.1021/jm030644s

Hartmann, T.S., Pognan, F., 2017. The eTOX Consortium: To Improve the Safety Assessment of New Drug Candidates. Pharm. Med. 2017.

Hsiao, P., Unadkat, J.D., 2012. P-glycoprotein-based loperamide-cyclosporine drug interaction at the rat blood-brain barrier: prediction from in vitro studies and extrapolation to humans. Mol. Pharm. 9, 629–633. https://doi.org/10.1021/mp200563a

Isenberg, B., Thole, H., Tümmler, B., Demmer, A., 2001. Identification and localization of three photobinding sites of iodoarylazidoprazosin in hamster P-glycoprotein. Eur. J. Biochem. 268, 2629–2634. https://doi.org/10.1046/j.1432-1327.2001.02155.x

Jacobson, M.P., Friesner, R.A., Xiang, Z., Honig, B., 2002. On the role of the crystal environment in determining protein side-chain conformations. J. Mol. Biol. 320, 597–608.

Jacobson, M.P., Pincus, D.L., Rapp, C.S., Day, T.J.F., Honig, B., Shaw, D.E., Friesner, R.A., 2004. A hierarchical approach to all-atom protein loop prediction. Proteins 55, 351–367. https://doi.org/10.1002/prot.10613

Jain, S., Grandits, M., Richter, L., Ecker, G.F., 2017. Structure based classification for bile salt export pump (BSEP) inhibitors using comparative structural modeling of human BSEP. J. Comput. Aided Mol. Des. 31, 507–521. https://doi.org/10.1007/s10822-017-0021-x

John, B., Sali, A., 2003. Comparative protein structure modeling by iterative alignment, model building and model assessment. Nucleic Acids Res. 31, 3982–3992.

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. https://doi.org/10.1006/jmbi.1996.0897

Katoh, M., Suzuyama, N., Takeuchi, T., Yoshitomi, S., Asahi, S., Yokoi, T., 2006. Kinetic analyses for species differences in P-glycoprotein-mediated drug transport. J. Pharm. Sci. 95, 2673–2683. https://doi.org/10.1002/jps.20686

Klepsch, F., Chiba, P., Ecker, G.F., 2011. Exhaustive Sampling of Docking Poses Reveals Binding Hypotheses for Propafenone Type Inhibitors of P-Glycoprotein. PLoS Comput Biol 7, e1002036. https://doi.org/10.1371/journal.pcbi.1002036

Klepsch, F., Jabeen, I., Chiba, P., Ecker, G.F., 2010. Pharmacoinformatic approaches to design natural product type ligands of ABC-transporters. Curr. Pharm. Des. 16, 1742–1752.

Klepsch, F., Vasanthanathan, P., Ecker, G.F., 2014. Ligand and Structure-Based Classification Models for Prediction of P-Glycoprotein Inhibitors. J. Chem. Inf. Model. 54, 218–229. https://doi.org/10.1021/ci400289j

Larkin, M.A., Blackshields, G., Brown, N.P., Chenna, R., McGettigan, P.A., McWilliam, H., Valentin, F., Wallace, I.M., Wilm, A., Lopez, R., Thompson, J.D., Gibson, T.J., Higgins, D.G., 2007. Clustal W and Clustal X Version 2.0. Bioinformatics 23, 2947–2948. https://doi.org/10.1093/bioinformatics/btm404

Laskowski, R., Macarthur, M., Moss, D., Thornton, J., 1993. {PROCHECK}: a program to check the stereochemical quality of protein structures. J Appl Cryst 26, 283–291.

Li, J., Jaimes, K.F., Aller, S.G., 2014. Refined structures of mouse P-glycoprotein. Protein Sci. Publ. Protein Soc. 23, 34–46. https://doi.org/10.1002/pro.2387

Lin, J.H., Yamazaki, M., 2003. Role of P-glycoprotein in pharmacokinetics: clinical implications. Clin. Pharmacokinet. 42, 59–98. https://doi.org/10.2165/00003088-200342010-00003

Liu, T., Lin, Y., Wen, X., Jorissen, R.N., Gilson, M.K., 2007. BindingDB: a web-accessible database of experimentally determined protein–ligand binding affinities. Nucleic Acids Res. 35, D198–D201. https://doi.org/10.1093/nar/gkl999

Loo, T.W., Bartlett, M.C., Clarke, D.M., 2006. Transmembrane segment 7 of human P-glycoprotein forms part of the drug-binding pocket. Biochem. J. 399, 351–359. https://doi.org/10.1042/BJ20060715

Loo, T.W., Bartlett, M.C., Clarke, D.M., 2003a. Simultaneous binding of two different drugs in the binding pocket of the human multidrug resistance P-glycoprotein. J. Biol. Chem. 278, 39706–39710. https://doi.org/10.1074/jbc.M308559200

Loo, T.W., Bartlett, M.C., Clarke, D.M., 2003b. Methanethiosulfonate Derivatives of Rhodamine and Verapamil Activate Human P-glycoprotein at Different Sites. J. Biol. Chem. 278, 50136–50141. https://doi.org/10.1074/jbc.M310448200

Loo, T.W., Clarke, D.M., 2008. Mutational analysis of ABC proteins. Arch. Biochem. Biophys. 476, 51–64. https://doi.org/10.1016/j.abb.2008.02.025

Loo, T.W., Clarke, D.M., 2005. Recent progress in understanding the mechanism of P-glycoprotein-mediated drug efflux. J. Membr. Biol. 206, 173–185. https://doi.org/10.1007/s00232-005-0792-1

Loo, T.W., Clarke, D.M., 2002. Location of the rhodamine-binding site in the human multidrug resistance P-glycoprotein. J. Biol. Chem. 277, 44332–44338. https://doi.org/10.1074/jbc.M208433200

Loo, T.W., Clarke, D.M., 1997. Identification of residues in the drug-binding site of human P-glycoprotein using a thiol-reactive substrate. J. Biol. Chem. 272, 31945–31948.

Lugo, M.R., Sharom, F.J., 2005. Interaction of LDS-751 and rhodamine 123 with P-glycoprotein: evidence for simultaneous binding of both drugs. Biochemistry (Mosc.) 44, 14020–14029. https://doi.org/10.1021/bi0511179

Martignoni, M., Groothuis, G.M.M., de Kanter, R., 2006. Species differences between mouse, rat, dog, monkey and human CYP-mediated drug metabolism, inhibition and induction. Expert Opin. Drug Metab. Toxicol. 2, 875–894. https://doi.org/10.1517/17425255.2.6.875

Martin, C., Berridge, G., Higgins, C.F., Mistry, P., Charlton, P., Callaghan, R., 2000. Communication between multiple drug binding sites on P-glycoprotein. Mol. Pharmacol. 58, 624–632.

Melo, F., Sánchez, R., Sali, A., 2002. Statistical potentials for fold assessment. Protein Sci. Publ. Protein Soc. 11, 430–448. https://doi.org/10.1002/pro.110430

Molden, E., Christensen, H., Sund, R.B., 2000. Extensive metabolism of diltiazem and P-glycoprotein-mediated efflux of desacetyl-diltiazem (M1) by rat jejunum in vitro. Drug Metab. Dispos. Biol. Fate Chem. 28, 107–109.

Molecular Operating Environment (MOE), 2013.08, n.d. . Chemical Computing Group Inc., 1010 Sherbooke St. West, Suite #910, Montreal, QC, Canada.

Pajeva, I.K., Globisch, C., Wiese, M., 2009. Combined Pharmacophore Modeling, Docking, and 3D QSAR Studies of ABCB1 and ABCC1 Transporter Inhibitors. ChemMedChem 4, 1883–1896. https://doi.org/10.1002/cmdc.200900282

Pedersen, K.E., 1985. Digoxin interactions. The influence of quinidine and verapamil on the pharmacokinetics and receptor binding of digitalis glycosides. Acta Med. Scand. Suppl. 697, 1–40.

Pinto, M., Trauner, M., Ecker, G.F., 2012. An In Silico Classification Model for Putative ABCC2 Substrates. Mol. Inform. 31, 547–553. https://doi.org/10.1002/minf.201200049

Pleban, K., Kopp, S., Csaszar, E., Peer, M., Hrebicek, T., Rizzi, A., Ecker, G.F., Chiba, P., 2005. P-glycoprotein substrate binding domains are located at the transmembrane domain/transmembrane domain interfaces: a combined photoaffinity labeling-protein homology modeling approach. Mol. Pharmacol. 67, 365–374. https://doi.org/10.1124/mol.104.006973

Qu, Q., Sharom, F.J., 2002. Proximity of bound Hoechst 33342 to the ATPase catalytic sites places the drug binding site of P-glycoprotein within the cytoplasmic membrane leaflet. Biochemistry (Mosc.) 41, 4744–4752.

Sastry, G.M., Adzhigirey, M., Day, T., Annabhimoju, R., Sherman, W., 2013. Protein and ligand preparation: parameters, protocols, and influence on virtual screening enrichments. J. Comput. Aided Mol. Des. 27, 221–234. https://doi.org/10.1007/s10822-013-9644-8

Schneider, G., 2010. Virtual screening: an endless staircase? Nat. Rev. Drug Discov. 9, 273–276. https://doi.org/10.1038/nrd3139

Schrödinger Release 2015-1: LigPrep, version 3.3, Schrödinger, LLC, New York, NY, 2015., n.d.

Schrödinger Release 2015-1: Maestro, version 10.1, Schrödinger, LLC, New York, NY, 2015., n.d.

Schwab, D., Fischer, H., Tabatabaei, A., Poli, S., Huwyler, J., 2003. Comparison of in vitro P-glycoprotein screening assays: recommendations for their use in drug discovery. J. Med. Chem. 46, 1716–1725. https://doi.org/10.1021/jm021012t

Seeger, M.A., van Veen, H.W., 2009. Molecular basis of multidrug transport by ABC transporters. Biochim. Biophys. Acta 1794, 725–737. https://doi.org/10.1016/j.bbapap.2008.12.004

Seelig, A., 1998. A general pattern for substrate recognition by P-glycoprotein. Eur. J. Biochem. 251, 252–261.

Shapiro, A.B., Fox, K., Lam, P., Ling, V., 1999. Stimulation of P-glycoprotein-mediated drug transport by prazosin and progesterone. Evidence for a third drug-binding site. Eur. J. Biochem. 259, 841–850.

Shelley, J.C., Cholleti, A., Frye, L.L., Greenwood, J.R., Timlin, M.R., Uchimaya, M., 2007. Epik: a software program for pK(a) prediction and protonation state generation for drug-like molecules. J. Comput. Aided Mol. Des. 21, 681–691. https://doi.org/10.1007/s10822-007-9133-z

Shen, M., Sali, A., 2006. Statistical potential for assessment and prediction of protein structures. Protein Sci. Publ. Protein Soc. 15, 2507–2524. https://doi.org/10.1110/ps.062416606

Shilling, R.A., Venter, H., Velamakanni, S., Bapna, A., Woebking, B., Shahi, S., van Veen, H.W., 2006. New light on multidrug binding by an ATP-binding-cassette transporter. Trends Pharmacol. Sci. 27, 195–203. https://doi.org/10.1016/j.tips.2006.02.008

Suzuyama, N., Katoh, M., Takeuchi, T., Yoshitomi, S., Higuchi, T., Asashi, S., Yokoi, T., 2007. Species differences of inhibitory effects on P-glycoprotein-mediated drug transport. J. Pharm. Sci. 96, 1609–1618. https://doi.org/10.1002/jps.20787

Thiebaut, F., Tsuruo, T., Hamada, H., Gottesman, M.M., Pastan, I., Willingham, M.C., 1987. Cellular localization of the multidrug-resistance gene product P-glycoprotein in normal human tissues. Proc. Natl. Acad. Sci. U. S. A. 84, 7735–7738.

Tsuji, A., 2002. Transporter-mediated Drug Interactions. Drug Metab. Pharmacokinet. 17, 253–274.

Verdonk, M.L., Cole, J.C., Hartshorn, M.J., Murray, C.W., Taylor, R.D., 2003. Improved protein–ligand docking using GOLD. Proteins Struct. Funct. Bioinforma. 52, 609–623. https://doi.org/10.1002/prot.10465

Verschraagen, M., Koks, C.H., Schellens, J.H., Beijnen, J.H., 1999. P-glycoprotein system as a determinant of drug interactions: the case of digoxin-verapamil. Pharmacol. Res. 40, 301–306. https://doi.org/10.1006/phrs.1999.0535

Ward, A., Reyes, C.L., Yu, J., Roth, C.B., Chang, G., 2007. Flexibility in the ABC transporter MsbA: Alternating access with a twist. Proc. Natl. Acad. Sci. U. S. A. 104, 19005–19010. https://doi.org/10.1073/pnas.0709388104

Waterhouse, A.M., Procter, J.B., Martin, D.M.A., Clamp, M., Barton, G.J., 2009. Jalview Version 2--a multiple sequence alignment editor and analysis workbench. Bioinforma. Oxf. Engl. 25, 1189–1191. https://doi.org/10.1093/bioinformatics/btp033

Widmer, N., Colombo, S., Buclin, T., Decosterd, L.A., 2003. Functional consequence of MDR1 expression on imatinib intracellular concentrations. Blood 102, 1142. https://doi.org/10.1182/blood-2003-03-0993

Wilkens, S., 2015. Structure and mechanism of ABC transporters. F1000prime Rep. 7, 14. https://doi.org/10.12703/P7-14

Willighagen, E.L., Waagmeester, A., Spjuth, O., Ansell, P., Williams, A.J., Tkachenko, V., Hastings, J., Chen, B., Wild, D.J., 2013. The ChEMBL database as linked open data. J. Cheminformatics 5, 23. https://doi.org/10.1186/1758-2946-5-23

Zhou, A.Q., O'Hern, C., Regan, L., 2011. Revisiting the Ramachandran plot from a new angle. Protein Sci. Publ. Protein Soc. 20, 1166–1171. https://doi.org/10.1002/pro.644

Zolnerciks, J.K., Booth-Genthe, C.L., Gupta, A., Harris, J., Unadkat, J.D., 2011. Substrate- and Species-dependent Inhibition of P-glycoprotein-mediated Transport: Implications for Predicting in vivo Drug Interactions. J. Pharm. Sci. 100, 3055–3061. https://doi.org/10.1002/jps.22566

III. Concluding Discussion

This thesis aims to investigate the potential of structural-based modeling methods to provide detailed insights into the mechanism of inhibition of membrane-associated liver transporters (BSEP, BCRP and P-gp), which might assist in the development of *in silico* prediction models and lead optimization. The transporters studied are implicated in multidrug resistance and hepatotoxicity. Exploring the mechanisms of inhibition of these transporters is highly essential not only to understand the pharmacological behavior of candidate drugs, but also to evaluate the potential drug-drug interaction liabilities and their clinical implications. Part I, section 1.2 provides the biological background of these transporters and emphasizes on their role in liver toxicity.

A majority of the *in silico* studies related to these transporters focused on ligand-based approaches that include QSAR modeling, pharmacophore modeling and machine learning methods, among others [37]. However, ligand-based models do not consider the structural aspects of the protein that are valuable in understanding the inhibition process. The lack of high-resolution structural information has been a primary reason behind the limited focus on structure-based approaches. Section 2.1 provides a detailed overview of the currently available ligand-based and structure-based models to predict inhibitors of different liver transporters. Experimentally resolved protein structures deposited in the Protein Data Bank and the inhibitors and substrates available from other dedicated resources serve in the development of *in silico* models for predicting transporter ligands. However, the coverage is still limited as high-quality data is still not available for certain transporters (e.g. NTCP, MRPs and MDR3).

Currently, a vast amount of open data is being generated in the drug discovery domain. In the light of this, issues with imbalanced datasets are frequently reported [38, 39, 197–199]. Chapter 3, section 3.1 in part II emphasizes on the problems with learning from imbalanced data and details various approaches to address them. Seven distinct meta-classifiers were evaluated on four highly imbalanced datasets to identify that while MetaCost and CostSensitiveClassifier achieve better sensitivities, Stratified Bagging provides the best balanced accuracies. An additional advantage of Stratified Bagging is that it is computationally less expensive and can be directly combined with any machine-learning method without any parameter optimization. In general, a method that balances between the complexity of the algorithm and the computational cost should be considered an ideal choice to obtain optimal results. On this basis, we provided a general recommendation to wrap the

modeling process in the stratified bagging loop when handling imbalanced data sets. Nevertheless, the performance of an *in silico* model depends on both quantity and quality of the underlying data. With few exceptions, such as P-gp, BCRP and BSEP, the limited availability of activity data in the public domain has been a major limiting factor in developing reliable models for ABC transporters. This highlights the need for publicly available data repositories that facilitate the deposition of high-confidence activity data comprising both positive and negative results.

Chapter 4 in part II presents the results of the structure-based studies on the three liver transporters (BSEP, BCRP and P-gp). Recent studies that employed homology models of P-glycoprotein provide promising evidences that structure-based classification methods could be valuable in studying these highly flexible and promiscuous transporters [196]. Section 4.1.1 reports a homology model for BSEP and the structure-based models to classify inhibitors and non-inhibitors. The significance of hydrophobic interactions of the inhibitors guided us to use molecular weight and logP(o/w) as additional descriptors, which further improved the prediction performance. Molecular docking enables the exploration of protein-ligand interactions, which facilitates understanding the biology at the molecular level and provides the rationale for the discovery, design, and development of safer and effective drugs. In our study, PLIF analysis revealed that certain functional group-amino acid residue interactions play a key role in ligand binding. While the functional groups halide, carbonyl, ether, vinyl and amide are overrepresented among the inhibitors, specific groups such as carbonyl and amide frequently participated in the interactions with the protein. The interactions of arene and carbonyl groups with tyrosine and leucine residues were more prominently noticed among inhibitors as compared to the non-inhibitors. These insights could further guide lead optimization. Thus, a sequential modeling approach, i.e. combining the structure-based model with ligand-based classification model would be a valuable approach to reduce the number of false positives in large-scale virtual screening efforts.

Structure-based methods can only be as good as the information they are provided with. The recent release of the BCRP crystal structure [86] motivated us to conduct structure-based studies with the aim to propose a binding mode that could explain the spread in activity within the arylmethyloxyphenyl series (Section 4.2.1). Our binding hypothesis, based on the results from docking studies, suggests that the activity of arylmethyloxyphenyl analogues is driven by strong hydrophobic interactions with residues Phe431, Phe432 and Phe439 that are

consistently involved in aromatic (pi-pi) and hydrophobic interactions. Thus, structure-based exploration of protein-ligand interactions is valuable in understanding the SAR of ligands which would be further useful in the development of potent and selective inhibitors for BCRP.

Development of *in silico* models that can predict *in vitro* and *in vivo* outcomes in animals is a valuable approach to reduce the number of animal experiments in preclinical development. However, limited availability of experimental data on rat and mouse P-gp activity restricts the development of such models. Section 4.3.1 presents the results from our structure-based assessment of the transferability of *in vitro* human P-gp data for development of *in silico* models to predict outcomes in rodents. We identified that similar binding site residues are involved in interactions across the three species, which strengthens the likelihood of similar binding modes for their inhibitors. To the best of our knowledge, this is the first *in silico* study of its kind that compares the binding sites of a protein across three different species with an emphasis on the interaction profiles of their inhibitors. However, only a small number of compounds were employed to validate the docking studies due to the limited availability of high-confidence experimental data in the public domain. Data from proprietary sources such as the pharmaceutical industry should be valuable for a more comprehensive validation.

Taken together, availability of high-resolution structures is a prerequisite, especially when studying membrane proteins. The lack of resolution therefore generates a blurry layer of uncertainty on top of the investigated problem and presents a challenging scenario to reveal useful structural insights. Performing docking or applying scoring functions on low-resolution structures for pose selection adds another layer of uncertainty and may as well lead to artefacts that do not represent the correct binding modes [200]. Therefore, these approaches have to be cautiously employed and must be completed by evidences gathered about the protein and the ligands of interest [201]. Nevertheless, since understanding the mechanism of inhibition of transporters is crucial, structure-based methods are essentially an ideal choice.

In a nutshell, this thesis work provides structural insights into the inhibition of three liver transporters (BSEP, BCRP and P-gp). The comparative modeling approach was successful in facilitating a better understanding of the mechanisms of inhibition while also emphasizing that structural information from the protein structure is essential for complete understanding

of the ligand SAR. Further, the protein-ligand interaction fingerprint (PLIF) analysis identified the most frequently occurring interactions between binding site residues and specific functional groups that provide detailed insights to understand the molecular basis of inhibition of the transporter proteins by a wide range of ligands.

IV. Appendix

5. Supplements to Section 3.1

Supplementary material

Comparing the performance of meta-classifiers − A case study on selected imbalanced data sets relevant for prediction of liver toxicity

Sankalp Jain[a,§], Eleni Kotsampasakou[a,§,#] and Gerhard F. Ecker[a,*]

[a]University of Vienna, Department of Pharmaceutical Chemistry, Althanstrasse 14, 1090 Vienna, Austria

[#]present address: Computational Toxicology Group, CMS, R&D Platform Technology & Science, GSK, Park Road, Ware, Hertfordshire, SG12 0DP, United Kingdom

[§]both authors contributed equally to this manuscript

*Corresponding author: E-Mail: gerhard.f.ecker@univie.ac.at; Phone: +43-1-4277-55110; eFax: +43-1-4277-855110

Tables

Table S1. Tuned settings of the best performing models for each meta-classifier/method

a. OATP1B1 dataset

Method	2D MOE descriptors	ECFP6 fingerprints	MACCS fingerprints
Stratified Bagging	-	-	-
CostSensitiveClassifier	cost 30:1 matrix: [0.0, 1.0; 30.0, 0.0]	cost 100:1 matrix: [0.0, 1.0; 100.0, 0.0]	cost 100:1 matrix: [0.0, 1.0; 100.0, 0.0]
MetaCost	cost 10:1 matrix: [0.0, 1.0; 10.0, 0.0]	cost 30:1 matrix: [0.0, 1.0; 30.0, 0.0]	cost 25:1 matrix: [0.0, 1.0; 25.0, 0.0]
SMOTE	1500% synthetic instances	2000% synthetic instances	1500% synthetic instances

b. OATP1B3 dataset

Method	2D MOE descriptors	ECFP6 fingerprints	MACCS fingerprints
Stratified Bagging	-	-	-
CostSensitiveClassifier	cost 70:1	cost 280:1	cost 200:1

	matrix: [0.0, 1.0; 70.0, 0.0]	matrix: [0.0, 1.0; 280.0, 0.0]	matrix: [0.0, 1.0; 200.0, 0.0]
MetaCost	cost 13:1 matrix: [0.0, 1.0; 13.0, 0.0]	cost 50:1 matrix: [0.0, 1.0; 50.0, 0.0]	cost 40:1 matrix: [0.0, 1.0; 40.0, 0.0]
SMOTE	1500% synthetic instances	2000% synthetic instances	1300% synthetic instances

c. Cholestasis human dataset

Method	2D MOE descriptors	ECFP6 fingerprints	MACCS fingerprints
Stratified Bagging	cost 2:1	cost 2:1	cost 2:1
CostSensitiveClassifier	cost 14:1 matrix: [0.0, 1.0; 14.0, 0.0]	cost 12:1 matrix: [0.0, 1.0; 12.0, 0.0]	cost 12:1 matrix: [0.0, 1.0; 12.0, 0.0]
MetaCost	cost 8:1 matrix: [0.0, 1.0; 8.0, 0.0]	cost 8:1 matrix: [0.0, 1.0; 8.0, 0.0]	cost 8:1 matrix: [0.0, 1.0; 8.0, 0.0]
SMOTE	1300% synthetic instances	3000% synthetic instances	1300% synthetic instances

cost 2:1 for Stratified Classifier: Stratified bagging used in combination with MetaCost with matrix: [0.0, 1.0; 2.0, 0.0]. For the case of human cholestasis dataset, Stratified Bagging on its own was not able to handle the dataset in such satisfactory way. Thus Stratified Bagging was combined with the application of a slight cost of 2:1 in favor of the minority class

d. Cholestasis animal dataset

Method	2D MOE descriptors	ECFP6 fingerprints	MACCS fingerprints
Stratified Bagging	cost 2:1	cost 2:1	cost 2:1
CostSensitiveClassifier	cost 450:1 matrix: [0.0, 1.0; 450.0, 0.0]	cost 500:1 matrix: [0.0, 1.0; 500.0, 0.0]	cost 500:1 matrix: [0.0, 1.0; 500.0, 0.0]
MetaCost	cost 45:1 matrix: [0.0, 1.0; 45.0, 0.0]	cost 45:1 matrix: [0.0, 1.0; 45.0, 0.0]	cost 50:1 matrix: [0.0, 1.0; 50.0, 0.0]
SMOTE	3000% synthetic instances	3000% synthetic instances	3000% synthetic instances

cost 2:1 for Stratified Classifier: Stratified bagging used in combination with MetaCost with matrix: [0.0, 1.0; 2.0, 0.0]. For the case of animal cholestasis dataset, Stratified Bagging on its own was not able to handle the dataset in such satisfactory way. Thus Stratified Bagging was combined with the application of a slight cost of 2:1 in favor of the minority class.

Table S2. Results on OATP1B1 inhibition dataset for all calculated statistics metrics : Accuracy, Balanced Accuracy, Sensitivity, Specificity, MCC, AUC, Precision. The performance is given for both 10-fold cross-validation and on the external test set. With bold font are depicted those models that gave a satisfactory result of sensitivity > 0.5 and they were further investigated by performing 20 iterations.

Model Settings	Descriptors	Validation	Accuracy	Balanced Accuracy	Sensitivity	Specificity	MCC	AUC	Precision
Random Forest	MOE	10 CV	0.893	0.614	0.253	0.974	0.322	0.809	0.545
		Test set	0.711	0.568	0.172	0.964	0.233	0.837	0.688
	ECFP6	10 CV	0.892	0.573	0.163	0.983	0.256	0.798	0.544
		Test set	0.701	0.548	0.125	0.971	0.188	0.804	0.667
	MACCS	10 CV	0.899	0.623	0.268	0.978	0.356	0.778	0.6
		Test set	0.711	0.560	0.141	0.978	0.233	0.768	0.75
Bagging	MOE	Training set	0.897	0.616	0.254	0.978	0.340	0.74	0.585
		Test set	0.701	0.535	0.078	0.993	0.194	0.724	0.833
	ECFP6	Training set	0.892	0.585	0.190	0.979	0.272	0.694	0.529
		Test set	0.701	0.544	0.109	0.978	0.211	0.572	0.700
	MACCS	Training set	0.904	0.636	0.291	0.981	0.394	0.701	0.655
		Test set	0.706	0.552	0.125	0.978	0.187	0.572	0.727
Stratified Bagging	**MOE**	**Training set**	**0.809**	**0.768**	**0.714**	**0.821**	**0.395**	**0.819**	**0.333**
		Test set	**0.831**	**0.830**	**0.828**	**0.832**	**0.634**	**0.887**	**0.697**
	ECFP6	**Training set**	**0.807**	**0.736**	**0.646**	**0.827**	**0.354**	**0.790**	**0.317**
		Test set	0.736	0.653	0.422	0.883	0.347	0.774	0.628
	MACCS	**Training set**	**0.783**	**0.757**	**0.725**	**0.790**	**0.365**	**0.798**	**0.300**
		Test set	0.741	0.689	0.547	0.832	0.390	0.809	0.603
CostSensitive Classifier	**MOE**	**10 CV**	**0.843**	**0.719**	**0.621**	**0.817**	**0.399**	**0.822**	**0.376**
		Test set	**0.841**	**0.804**	**0.703**	**0.905**	**0.625**	**0.856**	**0.776**
	ECFP6	**10 CV**	**0.653**	**0.711**	**0.784**	**0.637**	**0.269**	**0.791**	**0.213**
		Test set	**0.721**	**0.670**	**0.625**	**0.766**	**0.38**	**0.789**	**0.556**
	MACCS	**10 CV**	**0.645**	**0.701**	**0.774**	**0.628**	**0.257**	**0.79**	**0.207**
		Test set	**0.751**	**0.739**	**0.703**	**0.774**	**0.458**	**0.779**	**0.592**
MetaCost	**MOE**	**10 CV**	**0.819**	**0.746**	**0.653**	**0.839**	**0.376**	**0.826**	**0.337**

Category	Fingerprint								
		Test set	**0.841**	**0.825**	**0.781**	**0.869**	**0.64**	**0.87**	**0.735**
	ECFP6	**10 CV**	**0.622**	**0.693**	**0.784**	**0.602**	**0.245**	**0.769**	**0.198**
		Test set	**0.657**	**0.677**	**0.734**	**0.62**	**0.331**	**0.758**	**0.475**
	MACCS	**10 CV**	**0.673**	**0.703**	**0.742**	**0.664**	**0.263**	**0.767**	**0.217**
		Test set	**0.756**	**0.772**	**0.813**	**0.73**	**0.509**	**0.769**	**0.584**
Threshold Selector	MOE	**10 CV**	**0.881**	**0.721**	**0.516**	**0.926**	**0.423**	**0.806**	**0.467**
		Test set	**0.816**	**0.740**	**0.531**	**0.949**	**0.555**	**0.837**	**0.829**
	ECFP6	**10 CV**	**0.868**	**0.712**	**0.511**	**0.912**	**0.390**	**0.797**	**0.422**
		Test set	0.761	0.679	0.453	0.905	0.41	0.804	0.69
	MACCS	10 CV	0.880	0.656	0.368	0.944	0.342	0.775	0.452
		Test set	0.711	0.584	0.234	0.934	0.242	0.768	0.625
SMOTE	MOE	**10 CV**	**0.869**	**0.710**	**0.505**	**0.914**	**0.389**	**0.807**	**0.425**
		Test set	**0.816**	**0.749**	**0.563**	**0.934**	**0.555**	**0.823**	**0.800**
	ECFP6	10 CV	0.896	0.620	0.263	0.976	0.341	0.791	0.575
		Test set	0.716	0.572	0.172	0.971	0.253	0.767	0.733
	MACCS	10 CV	0.898	0.657	0.347	0.966	0.391	0.777	0.564
		Test set	0.711	0.560	0.141	0.978	0.233	0.787	0.75
ClassBalance r	MOE	**10 CV**	**0.694**	**0.748**	**0.816**	**0.679**	**0.321**	**0.823**	**0.241**
		Test set	0.776	0.686	0.438	0.934	0.447	0.839	0.757
	ECFP6	10 CV	0.893	0.657	0.353	0.961	0.377	0.799	0.532
		Test set	0.726	0.6	0.25	0.949	0.291	0.805	0.696
	MACCS	10 CV	0.875	0.665	0.395	0.935	0.344	0.774	0.434
		Test set	0.711	0.56	0.203	0.949	0.237	0.786	0.65

Table S3. Results on OATP1B3 inhibition dataset for all calculated statistics metrics : Accuracy, Balanced Accuracy, Sensitivity, Specificity, MCC, AUC, Precision. The performance is given for both 10-fold cross-validation and on the external test set. With bold font are depicted those models that gave a satisfactory result of sensitivity > 0.5 and they were further investigated by performing 20 iterations.

Model Settings	Descriptors	Validation	Accuracy	Balanced Accuracy	Sensitivity	Specificity	MCC	AUC	Precision
Random Forest	MOE	10 CV	0.926	0.593	0.202	0.983	0.276	0.868	0.472
		Test set	0.818	0.573	0.175	0.97	0.246	0.912	0.583
	ECFP6	10 CV	0.926	0.540	0.089	0.991	0.168	0.841	0.423
		Test set	0.804	0.526	0.075	0.976	0.112	0.795	0.429
	MACCS	10 CV	0.926	0.596	0.210	0.981	0.278	0.813	0.464
		Test set	0.804	0.526	0.075	0.976	0.112	0.821	0.429
Bagging	MOE	Training set	0.879	0.531	0.075	0.986	0.137	0.708	0.421
		Test set	0.818	0.554	0.125	0.982	0.220	0.645	0.625
	ECFP6	Training set	0.930	0.568	0.145	0.991	0.261	0.632	0.563
		Test set	0.797	0.520	0.075	0.965	0.078	0.609	0.333
	MACCS	Training set	0.929	0.571	0.153	0.989	0.253	0.646	0.514
		Test set	0.813	0.551	0.125	0.976	0.196	0.547	0.556
Stratified Bagging	**MOE**	**Training set**	**0.842**	**0.800**	**0.750**	**0.849**	**0.392**	**0.814**	**0.278**
		Test set	**0.813**	**0.856**	**0.925**	**0.787**	**0.588**	**0.915**	**0.507**
	ECFP6	**Training set**	**0.882**	**0.747**	**0.589**	**0.905**	**0.379**	**0.789**	**0.324**
		Test set	0.818	0.611	0.275	0.947	0.297	0.772	0.550
	MACCS	**Training set**	**0.798**	**0.724**	**0.637**	**0.811**	**0.278**	**0.800**	**0.207**
		Test set	0.789	0.679	0.500	0.858	0.345	0.817	0.455
CostSensitive Classifier	**MOE**	**10 CV**	**0.874**	**0.802**	**0.718**	**0.886**	**0.428**	**0.873**	**0.327**
		Test set	**0.852**	**0.842**	**0.825**	**0.858**	**0.603**	**0.9**	**0.579**
	ECFP6	**10 CV**	**0.647**	**0.725**	**0.815**	**0.634**	**0.237**	**0.814**	**0.147**
		Test set	**0.727**	**0.698**	**0.650**	**0.746**	**0.331**	**0.766**	**0.377**
	MACCS	**10 CV**	**0.733**	**0.737**	**0.742**	**0.732**	**0.267**	**0.819**	**0.177**
		Test set	**0.761**	**0.728**	**0.675**	**0.781**	**0.389**	**0.818**	**0.422**
MetaCost	**MOE**	**10 CV**	**0.863**	**0.796**	**0.718**	**0.874**	**0.409**	**0.872**	**0.307**

Group	Descriptor	Set							
	ECFP6	**Test set**	**0.837**	**0.842**	**0.850**	**0.834**	**0.589**	**0.894**	**0.548**
		10 CV	**0.683**	**0.736**	**0.798**	**0.674**	**0.254**	**0.796**	**0.159**
	MACCS	**Test set**	**0.670**	**0.634**	**0.575**	**0.692**	**0.219**	**0.742**	**0.307**
		10 CV	**0.717**	**0.751**	**0.790**	**0.711**	**0.277**	**0.816**	**0.175**
Threshold Selector	**MOE**	**Test set**	**0.718**	**0.721**	**0.725**	**0.716**	**0.36**	**0.767**	**0.377**
		10 CV	**0.908**	**0.754**	**0.573**	**0.934**	**0.433**	**0.868**	**0.403**
	ECFP6	**Test set**	**0.847**	**0.791**	**0.700**	**0.882**	**0.544**	**0.912**	**0.583**
		10 CV	**0.912**	**0.722**	**0.500**	**0.944**	**0.406**	**0.838**	**0.411**
	MACCS	Test set	0.804	0.583	0.225	0.941	0.227	0.795	0.474
		10 CV	0.915	0.676	0.395	0.956	0.356	0.814	0.408
	MOE	Test set	0.813	0.627	0.325	0.929	0.308	0.821	0.52
SMOTE		10 CV	0.886	0.686	0.452	0.92	0.311	0.742	0.304
		Test set	**0.837**	**0.728**	**0.55**	**0.905**	**0.464**	**0.886**	**0.579**
	ECFP6	10 CV	0.926	0.585	0.185	0.984	0.263	0.829	0.469
		Test set	0.804	0.526	0.075	0.976	0.112	0.823	0.429
	MACCS	10 CV	0.922	0.638	0.306	0.97	0.328	0.831	0.442
		Test set	0.809	0.548	0.125	0.97	0.176	0.852	0.5
ClassBalancer	**MOE**	10 CV	0.918	0.695	0.435	0.955	0.388	0.873	0.429
	ECFP6	Test set	0.837	0.699	0.475	0.923	0.435	0.892	0.594
		10 CV	0.923	0.643	0.315	0.971	0.338	0.827	0.453
	MACCS	Test set	0.794	0.539	0.125	0.953	0.126	0.815	0.385
		10 CV	0.922	0.638	0.306	0.97	0.328	0.831	0.442
		Test set	0.813	0.57	0.175	0.964	0.227	0.827	0.538

Table S4. Results on Cholestasis human dataset for all calculated statistics metrics : Accuracy, Balanced Accuracy, Sensitivity, Specificity, MCC, AUC, Precision. The performance is given for both 10-fold cross-validation and on the external test set. With bold font are depicted those models that gave a satisfactory result of sensitivity > 0.5 and they were further investigated by performing 20 iterations.

Model Settings	Descriptors	Validation	Accuracy	Balanced Accuracy	Sensitivity	Specificity	MCC	AUC	Precision
Random Forest	MOE	10 CV	0.839	0.622	0.265	0.979	0.382	0.772	0.754
		Test set	**0.835**	**0.728**	**0.528**	**0.927**	**0.501**	**0.81**	**0.683**
	ECFP6	10 CV	0.833	0.606	0.231	0.98	0.35	0.773	0.741
		Test set	**0.823**	**0.719**	**0.528**	**0.91**	**0.469**	**0.835**	**0.635**
	MACCS	10 CV	0.831	0.635	0.311	0.958	0.364	0.774	0.643
		Test set	**0.861**	**0.778**	**0.623**	**0.933**	**0.589**	**0.844**	**0.733**
Bagging	MOE	Training set	0.837	0.617	0.254	0.980	0.375	0.691	0.759
		Test set	0.835	0.723	0.519	0.927	0.492	0.73	0.675
	ECFP6	Training set	0.835	0.613	0.248	0.979	0.364	0.701	0.741
		Test set	0.826	0.717	0.519	0.916	0.734	0.471	0.643
	MACCS	Training set	0.838	0.634	0.297	0.970	0.387	0.685	0.710
		Test set	0.857	0.764	0.596	0.933	0.567	0.763	0.721
Stratified Bagging +cost2:1	MOE	**Training set**	**0.781**	**0.719**	**0.617**	**0.821**	**0.394**	**0.768**	**0.457**
		Test set	**0.761**	**0.716**	**0.635**	**0.798**	**0.395**	**0.747**	**0.478**
	ECFP6	**Training set**	**0.804**	**0.717**	**0.573**	**0.860**	**0.413**	**0.773**	**0.501**
		Test set	**0.791**	**0.736**	**0.635**	**0.837**	**0.445**	**0.761**	**0.532**
	MACCS	**Training set**	**0.785**	**0.728**	**0.634**	**0.822**	**0.410**	**0.775**	**0.466**
		Test set	0.774	0.752	0.712	0.792	0.451	0.807	0.500
CostSensitive Classifier	MOE	10 CV	0.724	0.701	0.663	0.739	0.337	0.78	0.383
		Test set	0.797	0.769	0.717	0.82	0.492	0.795	0.543
	ECFP6	10 CV	0.773	0.714	0.614	0.813	0.381	0.789	0.445
		Test set	0.810	0.751	0.642	0.860	0.483	0.825	0.576
	MACCS	10 CV	0.751	0.710	0.643	0.777	0.362	0.78	0.414
		Test set	0.775	0.741	0.679	0.803	0.44	0.823	0.507
MetaCost	MOE	10 CV	0.669	0.670	0.671	0.668	0.276	0.741	0.331

Technique	Descriptor	Split							
	ECFP6	Test set	0.697	0.678	0.642	0.713	0.310	0.724	0.400
	ECFP6	10 CV	0.750	0.697	0.608	0.785	0.343	0.762	0.409
	MACCS	Test set	0.684	0.682	0.679	0.685	0.313	0.746	0.391
	MACCS	10 CV	0.694	0.696	0.700	0.692	0.32	0.771	0.357
Threshold Selector	MOE	Test set	0.701	0.707	0.717	0.697	0.355	0.773	0.413
	MOE	10 CV	0.798	0.670	0.536	0.863	0.385	0.771	0.488
	ECFP6	Test set	0.831	0.771	0.660	0.882	0.532	0.81	0.625
	ECFP6	10 CV	0.816	0.683	0.464	0.902	0.387	0.77	0.537
	MACCS	Test set	0.827	0.762	0.642	0.882	0.517	0.835	0.618
	MACCS	10 CV	0.775	0.702	0.582	0.822	0.368	0.774	0.445
SMOTE	MOE	Test set	0.805	0.761	0.679	0.843	0.490	0.844	0.563
	MOE	10 CV	0.780	0.697	0.559	0.834	0.364	0.785	0.451
	ECFP6	Test set	0.810	0.744	0.623	0.865	0.476	0.825	0.579
	ECFP6	10 CV	0.835	0.748	0.308	0.965	0.381	0.777	0.682
	MACCS	Test set	0.836	0.637	0.585	0.910	0.517	0.849	0.660
	MACCS	10 CV	0.818	0.651	0.375	0.927	0.353	0.774	0.556
ClassBalancer	MOE	Test set	0.848	0.776	0.642	0.91	0.563	0.849	0.68
	MOE	10 CV	0.824	0.697	0.438	0.919	0.396	0.776	0.569
	ECFP6	Test set	0.844	0.773	0.642	0.904	0.554	0.788	0.667
	ECFP6	10 CV	0.840	0.687	0.435	0.939	0.437	0.78	0.637
	MACCS	Test set	0.827	0.749	0.604	0.893	0.504	0.833	0.627
	MACCS	10 CV	0.809	0.678	0.464	0.893	0.371	0.776	0.514
	MACCS	Test set	0.835	0.774	0.660	0.888	0.541	0.835	0.636

Table S5. Results on Cholestasis animal dataset for all calculated statistics metrics : Accuracy, Balanced Accuracy, Sensitivity, Specificity, MCC, AUC, Precision. The performance is given for both 10-fold cross-validation and on the external test set. With bold font are depicted those models that gave a satisfactory result of sensitivity > 0.5 and they were further investigated by performing 20 iterations.

Model Settings	Descriptors	Validation	Accuracy	Balanced Accuracy	Sensitivity	Specificity	MCC	AUC	Precision
Random Forest	MOE	10 CV	0.953	0.500	0.000	1.000	0.000	0.703	0.000
	ECFP6	10 CV	0.953	0.500	0.000	1.000	0.000	0.629	0.000
	MACCS	10 CV	0.951	0.511	0.027	0.997	0.083	0.700	0.333
Bagging	MOE	Training set	0.952	0.500	0.000	0.999	-0.006	0.503	0.000
	ECFP6	Training set	0.953	0.500	0.000	1.000	0.000	0.498	0.000
	MACCS	Training set	0.952	0.512	0.027	0.998	0.093	0.521	0.400
Stratified Bagging+cost2:1	**MOE**	**Training set**	**0.636**	**0.594**	**0.547**	**0.641**	**0.083**	**0.715**	**0.070**
	ECFP6	**Training set**	**0.722**	**0.639**	**0.547**	**0.731**	**0.131**	**0.686**	**0.092**
	MACCS	**Training set**	**0.623**	**0.637**	**0.653**	**0.621**	**0.119**	**0.732**	**0.079**
CostSensitive Classifier	**MOE**	**10 CV**	**0.632**	**0.623**	**0.613**	**0.633**	**0.108**	**0.665**	**0.077**
	ECFP6	**10 CV**	**0.532**	**0.527**	**0.520**	**0.533**	**0.023**	**0.531**	**0.052**
	MACCS	**10 CV**	**0.579**	**0.633**	**0.693**	**0.573**	**0.114**	**0.690**	**0.075**
MetaCost	**MOE**	**10 CV**	**0.582**	**0.597**	**0.613**	**0.580**	**0.083**	**0.644**	**0.068**
	ECFP6	**10 CV**	**0.599**	**0.587**	**0.573**	**0.600**	**0.075**	**0.600**	**0.066**
	MACCS	**10 CV**	**0.588**	**0.645**	**0.707**	**0.582**	**0.124**	**0.674**	**0.077**
Threshold Selector	MOE	10 CV	0.875	0.580	0.253	0.906	0.112	0.686	0.118
	ECFP6	10 CV	0.874	0.567	0.227	0.906	0.094	0.624	0.107
	MACCS	10 CV	0.848	0.635	0.4	0.87	0.163	0.687	0.132
SMOTE	MOE	10 CV	0.943	0.533	0.080	0.985	0.105	0.728	0.214
	ECFP6	10 CV	0.953	0.500	0.000	1.000	0.000	0.638	0.000
	MACCS	10 CV	0.949	0.511	0.027	0.995	0.057	0.708	0.2
ClassBalancer	MOE	10 CV	0.948	0.697	0.4	0.993	0.079	0.681	0.231
	ECFP6	10 CV	0.951	0.505	0.013	0.997	0.04	0.636	0.2
	MACCS	10 CV	0.945	0.509	0.027	0.991	0.037	0.693	0.125

Table S6. Results on OATP1B1 inhibition dataset only for the best performing methods on the appropriate set of descriptors (Sensitivity $\geq$ 0.5) for all calculated statistics metrics : Accuracy, Balanced Accuracy, Sensitivity, Specificity, MCC, AUC, Precision. The mean performance out of 20 iterations and the standard deviation values are provided.

Model Settings	Descriptors	Statistical Value	Accuracy	Balanced Accuracy	Sensitivity	Specificity	MCC	AUC	Precision
Stratified Bagging	MOE	mean	0.769	0.817	0.823	0.810	0.334	0.715	0.334
		sd	0.002	0.005	0.010	0.002	0.007	0.005	0.004
	ECFP6	mean	0.805	0.734	0.642	0.826	0.351	0.795	0.315
		sd	0.002	0.005	0.009	0.003	0.007	0.006	0.004
	MACCS	mean	0.721	0.724	0.728	0.721	0.299	0.803	0.245
		sd	0.003	0.004	0.007	0.003	0.006	0.004	0.003
CostSensitive Classifier	MOE	mean	0.847	0.754	0.634	0.873	0.413	0.804	0.385
		sd	0.003	0.011	0.020	0.003	0.016	0.067	0.008
	ECFP6	mean	0.641	0.701	0.778	0.624	0.256	0.785	0.206
		sd	0.009	0.014	0.018	0.010	0.013	0.007	0.006
	MACCS	mean	0.646	0.707	0.784	0.629	0.264	0.798	0.212
		sd	0.005	0.011	0.017	0.005	0.011	0.006	0.013
MetaCost	MOE	mean	0.817	0.747	0.656	0.837	0.376	0.822	0.335
		sd	0.006	0.009	0.011	0.006	0.011	0.005	0.009
	ECFP6	mean	0.625	0.694	0.782	0.605	0.245	0.770	0.201
		sd	0.008	0.013	0.017	0.008	0.012	0.005	0.012
	MACCS	mean	0.666	0.705	0.755	0.655	0.264	0.772	0.215
		sd	0.007	0.011	0.014	0.008	0.011	0.005	0.005
Threshold Selector	MOE	mean	0.879	0.721	0.519	0.924	0.420	0.813	0.460
		sd	0.005	0.017	0.027	0.006	0.018	0.007	0.018
	ECFP6	mean	0.875	0.703	0.483	0.924	0.391	0.794	0.442
		sd	0.005	0.014	0.021	0.007	0.015	0.007	0.017
SMOTE	MOE	mean	0.870	0.715	0.517	0.914	0.398	0.811	0.430
		sd	0.005	0.011	0.017	0.005	0.019	0.003	0.018

Table S7. Results on OATP1B3 inhibition dataset only for the best performing methods on the appropriate set of descriptors (Sensitivity ≥ 0.5) for all calculated statistics metrics : Accuracy, Balanced Accuracy, Sensitivity, Specificity, MCC, AUC, Precision. The mean performance out of 20 iterations and the standard deviation values are provided.

Model Settings	Descriptors	Statistical Value	Accuracy	Balanced Accuracy	Sensitivity	Specificity	MCC	AUC	Precision
Stratified Bagging	MOE	mean	0.841	0.804	0.761	0.847	0.395	0.819	0.278
		sd	0.002	0.005	0.044	0.004	0.060	0.031	0.004
	ECFP6	mean	0.882	0.755	0.606	0.904	0.388	0.789	0.328
		sd	0.002	0.005	0.010	0.002	0.008	0.009	0.006
	MACCS	mean	0.799	0.729	0.647	0.811	0.285	0.800	0.210
		sd	0.003	0.010	0.019	0.003	0.012	0.008	0.006
CostSensitive Classifier	MOE	mean	0.871	0.792	0.695	0.890	0.420	0.874	0.328
		sd	0.024	0.009	0.015	0.003	0.010	0.0037	0.007
	ECFP6	mean	0.651	0.725	0.811	0.639	0.238	0.809	0.148
		sd	0.009	0.014	0.018	0.010	0.010	0.008	0.004
	MACCS	mean	0.737	0.729	0.720	0.739	0.260	0.820	0.176
		sd	0.006	0.013	0.021	0.006	0.013	0.007	0.005
MetaCost	MOE	mean	0.864	0.7964	0.7174	0.8754	0.410	0.870	0.308
		sd	0.003	0.012	0.021	0.003	0.013	0.004	0.008
	ECFP6	mean	0.688	0.732	0.783	0.681	0.251	0.797	0.160
		sd	0.007	0.012	0.019	0.006	0.012	0.005	0.004
	MACCS	mean	0.711	0.738	0.769	0.706	0.262	0.802	0.169
		sd	0.007	0.016	0.024	0.007	0.014	0.009	0.006
Threshold Selector	MOE	mean	0.907	0.748	0.562	0.934	0.423	0.872	0.397
		sd	0.004	0.024	0.041	0.006	0.022	0.006	0.017
	ECFP6	mean	0.903	0.725	0.518	0.937	0.388	0.820	0.374
		sd	0.006	0.018	0.030	0.007	0.022	0.012	0.021

Table S8. Results on human cholestasis dataset only for the best performing methods on the appropriate set of descriptors (Sensitivity ≥ 0.5) for all calculated statistics metrics : Accuracy, Balanced Accuracy, Sensitivity, Specificity, MCC, AUC, Precision. The mean performance out of 20 iterations and the standard deviation values are provided.

Model Settings	Descriptors	Statistical Value	Accuracy	Balanced Accuracy	Sensitivity	Specificity	MCC	AUC	Precision
Stratified Bagging +cost2:1	MOE	mean	0.777	0.713	0.607	0.819	0.384	0.768	0.450
		sd	0.002	0.005	0.011	0.003	0.008	0.004	0.005
	ECFP6	mean	0.806	0.722	0.583	0.860	0.421	0.773	0.505
		sd	0.003	0.003	0.006	0.004	0.006	0.004	0.006
	MACCS	mean	0.782	0.723	0.625	0.820	0.400	0.772	0.460
		sd	0.005	0.004	0.007	0.005	0.009	0.005	0.008
CostSensitive Classifier	MOE	mean	0.731	0.707	0.667	0.74685	0.346	0.786	0.392
		sd	0.005	0.009	0.013	0.005	0.011	0.005	0.006
	ECFP6	mean	0.771	0.704	0.596	0.812	0.369	0.782	0.440
		sd	0.006	0.014	0.017	0.012	0.015	0.006	0.011
	MACCS	mean	0.753	0.705	0.629	0.782	0.343	0.776	0.415
		sd	0.006	0.010	0.011	0.010	0.073	0.005	0.008
MetaCost	MOE	mean	0.671	0.681	0.698	0.664	0.293	0.755	0.337
		sd	0.008	0.014	0.019	0.010	0.015	0.007	0.007
	ECFP6	mean	0.750	0.699	0.614	0.783	0.346	0.768	0.409
		sd	0.006	0.011	0.015	0.007	0.013	0.006	0.009
	MACCS	mean	0.690	0.692	0.695	0.689	0.313	0.764	0.353
		sd	0.005	0.006	0.013	0.007	0.010	0.004	0.006
Threshold Selector	MOE	mean	0.801	0.694	0.518	0.870	0.382	0.775	0.495
		sd	0.008	0.013	0.015	0.011	0.015	0.006	0.018
	MACCS	mean	0.782	0.694	0.551	0.838	0.363	0.771	0.455
		sd	0.008	0.021	0.027	0.015	0.011	0.005	0.014

147

Table S9. Results animal cholestasis dataset only for the best performing methods on the appropriate set of descriptors (Sensitivity ≥ 0.5) for all calculated statistics metrics : Accuracy, Balanced Accuracy, Sensitivity, Specificity, MCC, AUC, Precision. The mean performance out of 20 iterations and the standard deviation values are provided.

Model Settings	Descriptors	Statistical Value	Accuracy	Balanced Accuracy	Sensitivity	Specificity	MCC	AUC	Precision
Stratified Bagging +cost2:1	MOE	mean	0.648	0.608	0.564	0.653	0.096	0.710	0.075
		sd	0.015	0.011	0.021	0.016	0.010	0.008	0.003
	ECFP6	mean	0.713	0.633	0.545	0.721	0.124	0.678	0.088
		sd	0.009	0.008	0.018	0.010	0.008	0.009	0.003
	MACCS	mean	0.624	0.636	0.649	0.623	0.118	0.729	0.079
		sd	0.007	0.009	0.022	0.008	0.008	0.008	0.002
CostSensitive Classifier	MOE	mean	0.6304	0.6122	0.592	0.632	0.098	0.659	0.074
		sd	0.009	0.017	0.030	0.009	0.015	0.015	0.005
	ECFP6	mean	0.530	0.533	0.536	0.523	0.026	0.541	0.053
		sd	0.008	0.023	0.048	0.008	0.023	0.014	0.004
	MACCS	mean	0.588	0.645	0.708	0.582	0.125	0.683	0.078
		sd	0.008	0.022	0.044	0.008	0.019	0.018	0.005
MetaCost	MOE	mean	0.586	0.610	0.637	0.5829	0.095	0.666	0.070
		sd	0.009	0.014	0.018	0.009	0.009	0.011	0.003
	ECFP6	mean	0.587	0.599	0.6126	0.586	0.085	0.610	0.098
		sd	0.014	0.032	0.048	0.016	0.019	0.025	0.005
	MACCS	mean	0.5894	0.6453	0.708	0.5826	0.1245	0.6752	0.0776
		sd	0.008	0.024	0.039	0.010	0.016	0.012	0.004

Figures

Figure S1 (a-d). Comparison of performances of different meta-classifiers on the four training datasets (after one round of 10-fold cross validation). *x-axis* corresponds to the sensitivity and on the *y-axis* is the specificity. The squares correspond to MOE descriptors, the triangles correspond to ECFP6 fingerprints and the circles correspond to MACCS fingerprints. Each classifier is depicted in a different color: red for RF standalone, green for Bagging, blue for Stratified Bagging, dark pink for CostSensitiveClassifier, cyan for MetaCost, yellow for ThresholdSelector, orange for SMOTE and dark violet for ClassBalancer.

(a)

(b)

149

(c)

Cholestasis Human - Training Set

(d)

Cholestasis Animal - Training Set

150
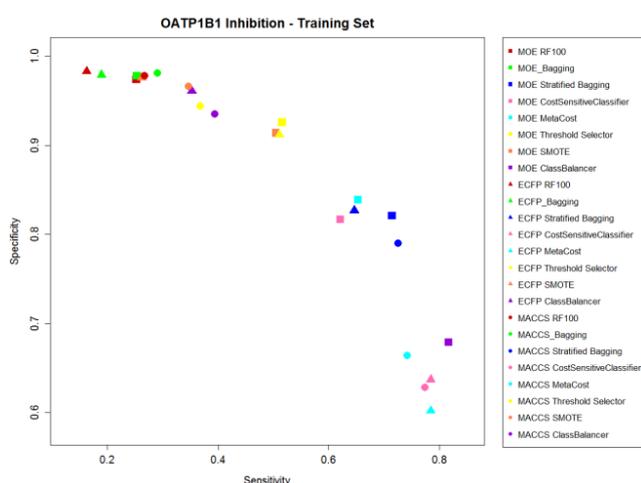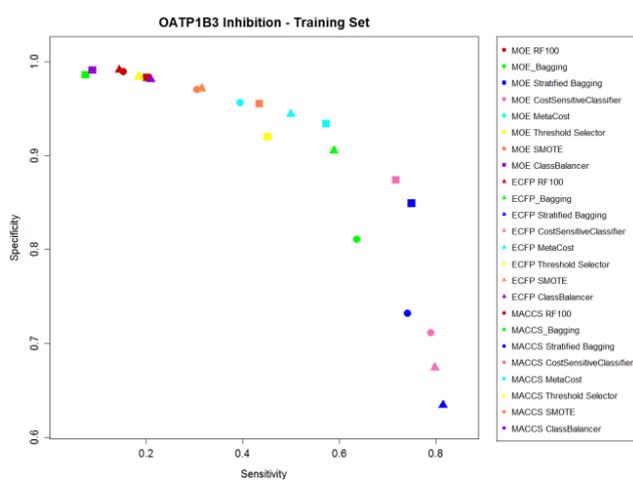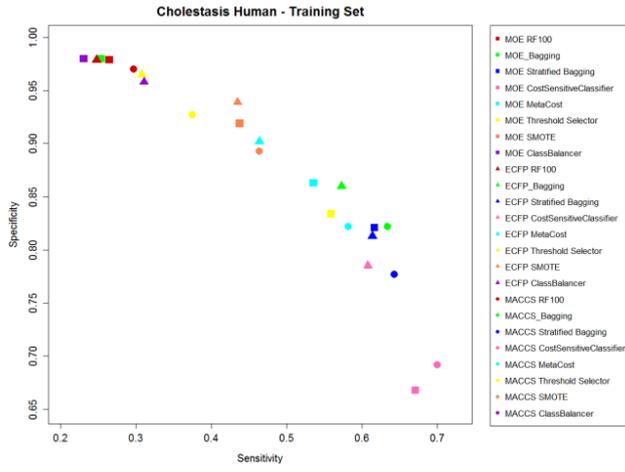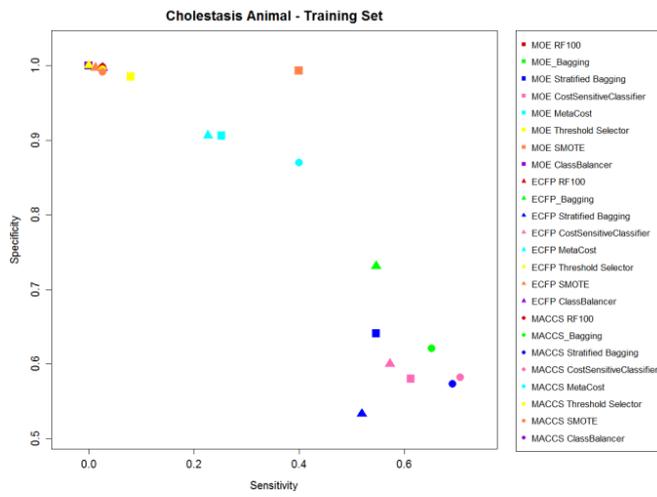
6. Supplements to Section 4.1.1

<u>Supplementary Material</u>

Structure Based Classification for Bile Salt Export Pump (BSEP) Inhibitors using Comparative Structural Modeling of Human BSEP

Sankalp Jain[a], Melanie Grandits[a], Lars Richter[a], Gerhard F. Ecker[a]

[a]University of Vienna, Department of Pharmaceutical Chemistry, Althanstrasse 14, 1090 Vienna, Austria

E-Mail: gerhard.f.ecker@univie.ac.at; Phone: +43-1-4277-55110; eFax: +43-1-4277-855110

Tables

Table S1. Summary view of the applicability domain (AD) analysis with information about the percentage of the reliable and the unreliable predictions.

Test set (Pedersen et al.)			
Scoring function	**Reliable percentage (%)**	**Unreliable percentage (%)**	**AD limit (Threshold)**
Chemscore	100 (166/166)	0 (0/166)	4.302
Xscore_chemscore	100 (166/166)	0 (0/166)	0.564
Goldscore	100 (166/166)	0 (0/166)	8.809
Xscore_goldscore	100 (166/166)	0 (0/166)	0.552
Glidescore	99.4 (165/166)	0.60 (1/166)	1.393
Test set (AstraZeneva-unpublished et al.)			
Chemscore	100 (638/638)	0 (0/638)	4.302
Xscore_chemscore	99.7 (636/638)	0.30 (2/638)	0.564
Goldscore	99.8 (637/638)	0.20 (1/638)	8.809
Xscore_goldscore	99.8 (637/638)	0.20 (1/638)	0.552

Table S2. Probability of prediction for different scoring function.

ChemScore scoring function

Bin	Probability of Inhibitor			Probability of non-inhibitor			p-value		
	Training Set	Test set (Pedersen et al.)	Test set (AstraZeneca-unpublished)	Training Set	Test set (Pedersen et al.)	Test set (AstraZeneca-unpublished)	Training Set	Test set (Pedersen et al.)	Test set (AstraZeneca-unpublished)
0-5	0	0	0	1	0	1	1.00E+000	NA	1.00E+000
5-10	0	0	0	1	0	1	6.68E-001	NA	6.87E-001
10-15	0	0	0	1	1	1	2.51E-002	5.21E-001	5.68E-002
15-20	0	0	0.16	1	1	0.84	2.78E-07	5.38E-002	2.08E-003
20-25	0.10	0.09	0.13	0.90	0.91	0.87	5.09E-06	2.64E-002	2.12E-010
25-30	0.29	0.14	0.32	0.71	0.86	0.68	8.24E-001	2.74E-002	6.27E-003
30-35	0.70	0.40	0.58	0.30	0.60	0.42	<2.20E-016	1.84E-002	4.79E-009
35-40	0.88	0.75	0.78	0.12	0.25	0.22	5.93E-08	1.52E-005	1.82E-008
40-45	0.5	0.75	0.94	0.5	0.25	0.06	1.00E+00	0.09873	3.14E-006
45-50	1	0	1	0	0	0	0.6178	NA	0.8193
50-55	0	0	0	0	0	0	NA	NA	NA

GoldScore scoring function

Bin	Probability of Inhibitor			Probability of non-inhibitor			p-value		
	Training Set	Test set (Pedersen et al.)	Test set (AstraZeneca-unpublished)	Training Set	Test set (Pedersen et al.)	Test set (AstraZeneca-unpublished)	Training Set	Test set (Pedersen et al.)	Test set (AstraZeneca-unpublished)
0-5	0	0	0	0	0	0	NA	NA	NA
5-10	0	0	0	0	0	0	NA	NA	NA
10-15	0	0	0	0	0	0	NA	NA	NA
15-20	0	0	1	0	0	0	NA	NA	0.8193
20-25	0	0	0	1	1	0	1.00E+000	1.00E+000	NA
25-30	0	0	0	1	1	1	1.71E-001	9.61E-001	4.29E-001
30-35	0	0	0	1	1	1	6.15E-003	9.61E-001	1.23E-001
35-40	0	0	0.03	1	1	0.97	8.65E-06	4.22E-002	3.27E-005
40-45	0.11	0	0.02	0.89	1	0.98	7.12E-003	2.58E-002	5.95E-008
45-50	0.11	0.23	0.07	0.89	0.77	0.93	5.73E-004	2.32E-002	4.49E-013
50-55	0.36	0.06	0.35	0.64	0.94	0.65	1.40E-001	1.02E-001	4.35E-001
55-60	0	0	0	0	0	0	NA	NA	NA

Bin	Probability of Inhibitor			Probability of non-inhibitor			p-value		
	Training Set	Test set (Pedersen et al.)	Test set (AstraZeneca-unpublished)	Training Set	Test set (Pedersen et al.)	Test set (AstraZeneca-unpublished)	Training Set	Test set (Pedersen et al.)	Test set (AstraZeneca-unpublished)
55-60	0.4	0.21	0.47	0.6	0.79	0.53	5.65E-002	6.67E-001	4.39E-002
60-65	0.71	0.5	0.68	0.29	0.5	0.32	1.73E-08	1.55E-002	1.30E-009
65-70	0.72	0.4	0.64	0.28	0.6	0.36	2.00E-08	3.50E-001	1.29E-005
70-75	0.44	0.42	0.62	0.56	0.58	0.38	4.48E-001	3.70E-001	8.45E-003
75-80	0.6	1	0.75	0.4	0	0.25	2.62E-001	1.09E-003	2.19E-002
80-85	0.33	1	0.86	0.67	0	0.14	1.00E+000	1.18E-001	3.03E-002
85-90	0	0.5	0.67	1	0.5	0.33	1.00E+000	6.14E-001	6.92E-001

Xscore(Chemscore) scoring function

Bin	Probability of Inhibitor			Probability of non-inhibitor			p-value		
	Training Set	Test set (Pedersen et al.)	Test set (AstraZeneca-unpublished)	Training Set	Test set (Pedersen et al.)	Test set (AstraZeneca-unpublished)	Training Set	Test set (Pedersen et al.)	Test set (AstraZeneca-unpublished)
0-0.5	0	0	0	0	0	0	NA	NA	NA
0.5-1	0	0	0	0	0	0	NA	NA	NA
1-1.5	0	0	0	0	0	0	NA	NA	NA
1.5-2	0	0	0	0	0	0	NA	NA	NA
2-2.5	0	0	0	0	0	0	NA	NA	NA

Xscore(Goldscore) scoring function

Bin	Probability of Inhibitor			Probability of non-inhibitor			p-value		
	Training Set	Test set (Pedersen	Test set (AstraZeneca-	Training Set	Test set (Pedersen	Test set (AstraZeneca-	Training Set	Test set (Pedersen	Test set (AstraZeneca-
2.5-3	0	0	0	0	0	0	NA	NA	NA
3-3.5	0	0	0	0	0	0	NA	NA	NA
3.5-4	0	0	0	1	1	1	1.71E-001	1.00E+000	1.00E+000
4-4.5	0	0	0	1	1	1	1.18E-003	1.83E-001	3.39E-004
4.5-5	0.03	0	0.02	0.97	1	0.98	1.24E-06	2.36E-001	1.85E-007
5-5.5	0.05	0.04	0.08	0.95	0.96	0.92	1.23E-07	1.93E-002	4.31E-014
5.5-6	0.19	0.16	0.25	0.81	0.84	0.75	2.97E-002	2.20E-001	1.49E-004
6-6.5	0.71	0.22	0.59	0.29	0.78	0.41	<2.20E-016	5.72E-001	8.29E-010
6.5-7	0.91	0.63	0.74	0.08	0.37	0.26	1.22E-11	3.57E-004	3.50E-011
7-7.5	0.83	0.64	0.79	0.17	0.36	0.21	5.25E-05	1.13E-002	1.56E-006
7.5-8	1	0.89	0.5	0	0.11	0.5	6.18E-001	1.42E-004	8.88E-001
8-8.5	0.8	0	0.83	0.2	1	0.17	3.34E-002	1.00E+000	6.81E-002
8.5-9	1	1	1	0	0	0	6.18E-001	5.93E-001	4.53E-002

	et al.)	unpublished)		et al.)	unpublished)		et al.)	unpublished)
0-0.5	0	0	0	0	0	NA	NA	NA
0.5-1	0	0	0	0	0	NA	NA	NA
1-1.5	0	0	0	0	0	NA	NA	NA
1.5-2	0	0	0	0	0	NA	NA	NA
2-2.5	0	0	0	0	0	NA	NA	NA
2.5-3	0	0	0	0	0	NA	NA	NA
3-3.5	0	0	0	0	0	NA	NA	NA
3.5-4	0	0	1	0	0	0.9363	NA	NA
4-4.5	0	0	1	1	1	5.79E-004	3.96E-001	2.36E-004
4.5-5	0.01	0	0.99	1	1	1.55E-08	5.69E-003	4.72E-010
5-5.5	0.04	0.03	0.96	0.95	0.97	5.04E-08	3.15E-002	7.31E-016
5.5-6	0.22	0.32	0.77	0.91	0.68	2.25E-001	2.08E-002	5.30E-002
6-6.5	0.76	0.56	0.24	0.66	0.44	<2.20E-016	2.35E-001	2.50E-007
6.5-7	0.88	0.72	0.12	0.35	0.28	4.23E-12	1.01E-004	3.91E-010
7-7.5	0.92	0.81	0.08	0.42	0.19	6.50E-07	2.42E-002	2.57E-006
7.5-8	0.75	0.93	0.25	0.2	0.07	1.18E-001	2.53E-002	9.12E-005

Bin	Probability of Inhibitor			Probability of non-inhibitor			p-value		
	Training Set	Test set (Pedersen et al.)	Test set (AstraZeneca-unpublished)	Training Set	Test set (Pedersen et al.)	Test set (AstraZeneca-unpublished)	Training Set	Test set (Pedersen et al.)	Test set (AstraZeneca-unpublished)
8-8.5	1	0	0	0.86	1	0.14	1.34E-001	1.00E+000	3.03E-002
8.5-9	0	0	1	0	0	0	NA	NA	2.94E-001

GlideXP scoring function

Bin	Probability of Inhibitor			Probability of non-inhibitor			p-value		
	Training Set	Test set (Pedersen et al.)	Test set (AstraZeneca-unpublished)	Training Set	Test set (Pedersen et al.)	Test set (AstraZeneca-unpublished)	Training Set	Test set (Pedersen et al.)	Test set (AstraZeneca-unpublished)
-16 - -15.5	0	0	0	0	0	0	NA	NA	NA
-15.5 - -15	0	0	0	0	0	0	NA	NA	NA
-15 - -14.5	0	0	0	1	0	0	1.00E+000	NA	NA
-14.5 - -14	0	1	0	0	0	0	NA	5.93E-001	NA
-14 - -13.5	0	0	0	0	0	0	NA	NA	NA
-13.5 - -13	0	0	0	1	0	0	1.00E+000	NA	NA
-13 - -12.5	0	0	1	0	0	0	NA	NA	8.19E-001
-12.5 - -12	1	0	1	0	1	0	6.12E-001	9.61E-001	8.19E-001
-12 - -11.5	0	0	0	0	1	1	NA	9.61E-001	1.00E+000
-11.5 - -11	0.33	1	0	0.67	0	1	1.00E+000	2.30E-004	1.00E+000

-11 - -10.5	1	0.25	0.33	0	0.75	0.67	6.12E-001	1.00E+000	1.00E+000
-10.5 - -10	0	0.33	0.2	1	0.67	0.8	3.80E-001	1.00E+000	6.83E-001
-10 - -9.5	0.54	0.6	0.58	0.46	0.4	0.42	6.31E-002	2.27E-001	1.37E-001
-9.5 - -9	0.71	0.47	0.69	0.29	0.53	0.31	2.72E-002	1.22E-001	4.17E-004
-9 - -8.5	0.68	0.27	0.55	0.32	0.73	0.45	3.10E-05	1.00E+000	9.93E-002
-8.5 - -8	0.56	0.33	0.48	0.44	0.67	0.52	1.51E-003	9.29E-001	1.88E-001
-8 - -7.5	0.55	0.43	0.58	0.45	0.57	0.42	1.09E-003	2.58E-001	1.88E-001
-7.5 - -7	0.41	0.24	0.6	0.59	0.76	0.4	6.88E-002	9.97E-001	5.75E-005
-7 - -6.5	0.33	0.18	0.41	0.67	0.82	0.59	6.21E-001	5.60E-001	7.79E-001
-6.5 - -6	0.20	0.06	0.33	0.80	0.94	0.67	3.67E-001	8.12E-002	3.44E-001
-6 - -5.5	0.12	0.17	0.14	0.88	0.83	0.86	5.33E-002	6.44E-001	5.20E-005
-5.5 - -5	0.09	0	0.2	0.91	1	0.8	5.57E-003	1.43E-001	6.93E-003
-5 - -4.5	0.11	0.14	0.19	0.89	0.86	0.81	4.39E-002	7.56E-001	2.23E-002
-4.5 - -4	0.11	0.17	0.15	0.89	0.83	0.85	6.73E-002	9.32E-001	4.64E-002
-4 - -3.5	0.11	0	0.11	0.89	1	0.89	1.54E-001	1.00E+000	2.74E-002
-3.5 - -3	0	0	0.13	1	1	0.87	1.76E-001	1.00E+000	2.40E-001
-3 - -2.5	0	0	0	1	1	1	6.72E-002	6.97E-001	5.68E-002

-2.5 - -2	0	0	1	0	1	9.40E-001	NA	4.29E-001
-2 - -1.5	0.5	0	1	0.5	1	1.00E+000	1.00E+000	1.00E+000
-1.5 - -1	0	0	1	0	0	5.02E-001	NA	NA
-1 - -0.5	0	0.5	0	1	0.5	NA	1.00E+000	1.00E+000
-0.5 - 0	0	0	1	0	0	1.00E+000	NA	NA
0 - 0.5	0	0	0	0	0	NA	NA	NA

The scoring function in brackets was used to generate the docking poses.

Table S3 (a). Summary of models obtained using different scoring functions for the training set

Scoring Function	Intersection Point	TP	TN	FP	FN	Sensitivity	Specificity	Accuracy	G-mean	MCC
ChemScore	29.50	68	261	34	45	0.60	0.88	0.81	0.73	0.50
GoldScore	53.50	83	221	74	30	0.74	0.75	0.75	0.74	0.45
Maestro_docking score(XP)	-6.80	89	192	105	23	0.80	0.65	0.69	0.72	0.39
Consensus		88	236	59	25	0.78	0.80	0.79	0.79	0.54
Xscore (ChemScore)	6.15	80	279	16	33	0.71	0.95	0.88	0.82	0.69
Xscore (GoldScore)	6.10	77	280	15	36	0.68	0.95	0.86	0.80	0.68
ChemScore+Molecular Weight (MW)+logP (Normalized)										

159

ChemScore+MW	0.80	91	266	29	22	0.81	0.90	0.88	0.85	0.69
ChemScore+logP	1.10	77	269	26	36	0.68	0.91	0.85	0.79	0.61
ChemScore+MW+logP	1.20	85	270	25	28	0.75	0.92	0.87	0.83	0.67
Xscore(ChemScore)+Molecular Weight (MW)+logP (Normalized)										
Xscore(ChemScore) +MW	0.70	96	263	32	17	0.85	0.89	0.88	0.87	0.72
Xscore(ChemScore) +logP	1.2	78	282	13	35	0.69	0.96	0.89	0.81	0.86
Xscore(ChemScore) +MW+logP	0.90	92	266	29	21	0.81	0.90	0.88	0.86	0.70
GoldScore+Molecular Weight (MW)+logP(Normalized)										
GoldScore+MW	0.40	91	230	65	22	0.81	0.78	0.79	0.79	0.58
GoldScore+logP	0.70	91	259	36	22	0.81	0.88	0.86	0.84	0.66
GoldScore+MW+logP	0.70	99	250	45	14	0.88	0.85	0.86	0.86	0.68
Xscore(GoldScore)+Molecular Weight (MW)+logP(Normalized)										
Xscore(GoldScore) +MW	0.5	97	251	44	16	0.86	0.85	0.85	0.86	0.67
Xscore(GoldScore)+logP	----									
Xscore(GoldScore)+MW+logP	0.80	97	267	28	16	0.86	0.91	0.89	0.88	0.74

The scoring function in brackets was used to generate the docking poses.

Table S3 (b). Summary of models obtained using different scoring functions for the test set (Pedersen et al.)

Scoring Function	Intersection Point	TP	TN	FP	FN	Sensitivity	Specificity	Accuracy	G-mean	MCC
ChemScore	29.50	35	86	36	9	0.80	0.71	0.73	0.75	0.45
GoldScore	53.50	37	69	53	7	0.84	0.57	0.64	0.69	0.36
Maestro_docking score(XP)	-6.80	38	49	73	6	0.86	0.40	0.52	0.59	0.25
Consensus		38	67	55	6	0.86	0.55	0.63	0.69	0.37
Xscore										
Xscore (ChemScore)	6.15	38	82	40	6	0.86	0.67	0.72	0.76	0.47
Xscore (GoldScore)	6.10	35	83	39	9	0.80	0.68	0.71	0.74	0.42
ChemScore+Molecular Weight (MW)+ logP(Normalized)										
ChemScore+MW	0.80	34	109	13	10	0.77	0.89	0.86	0.83	0.65
ChemScore+logP	1.10	31	122	0	13	0.71	1.00	0.92	0.84	0.80
ChemScore+MW+logP	1.20	36	120	2	8	0.82	0.98	0.94	0.90	0.84
Xscore(ChemScore)+Molecular Weight (MW)+ logP(Normalized)										
Xscore(ChemScore)+MW	0.70	38	100	22	6	0.86	0.82	0.83	0.84	0.63
Xscore(ChemScore)+logP	1.2	24	113	9	20	0.55	0.93	0.83	0.71	0.52

161

Scoring Function	Intersection Point	TP	TN	FP	FN	Sensitivity	Specificity	Accuracy	G-mean	MCC
Xscore(ChemScore)+MW+logP	0.90	37	106	16	7	0.84	0.87	0.86	0.86	0.67
GoldScore+MolecularWeight (MW)+ logP(Normalized)										
GoldScore+MW	0.40	35	94	28	9	0.80	0.77	0.78	0.78	0.52
Goldsscore+logP	0.70	34	106	16	10	0.77	0.87	0.84	0.82	0.62
GoldScore+MW+logP	0.70	35	110	12	9	0.80	0.90	0.87	0.85	0.68
Xscore(GoldScore)+Molecular Weight (MW)+ logP(Normalized)										
Xscore(GoldScore)+MW	0.50	38	95	27	6	0.86	0.78	0.80	0.82	0.58
Xscore(GoldScore)+logP	----									
Xscore(GoldScore)+MW+logP	0.80	40	104	18	4	0.91	0.85	0.87	0.88	0.71

The scoring function in brackets was used to generate the docking poses.

Table S3 (c). Summary of models obtained using different scoring function for the the test set (AstraZeneca-unpublished)

Scoring Function	Intersection Point	TP	TN	FP	FN	Sensitivity	Specificity	Accuracy	G-mean	MCC
ChemScore	29.50	162	297	93	86	0.65	0.76	0.72	0.71	0.41
GoldScore	53.50	213	240	150	35	0.86	0.62	0.71	0.73	0.47
Maestro_docking score(XP)	-6.80	199	205	185	49	0.80	0.53	0.63	0.65	0.33
Consensus		211	257	133	37	0.85	0.66	0.73	0.75	0.50

Xscore										
Xscore (ChemScore)	6.15	179	310	80	69	0.72	0.80	0.77	0.76	0.51
Xscore (GoldScore)	6.10	175	306	84	73	0.71	0.79	0.75	0.74	0.49
ChemScore+Molecular Weight (MW)+ logP(Normalized)										
ChemScore+MW	0.80	121	347	43	127	0.49	0.89	0.73	0.66	0.42
ChemScore+logP	1.10	121	352	38	127	0.49	0.90	0.74	0.66	0.44
ChemScore+MW+logP	1.20	131	356	34	117	0.53	0.91	0.76	0.69	0.49
Xscore(ChemScore)+Molecular Weight (MW)+ logP(Normalized)										
Xscore(ChemScore)+MW	0.70	134	344	46	114	0.54	0.88	0.75	0.69	0.46
Xscore(ChemScore)+logP	1.2	104	363	27	144	0.42	0.93	0.73	0.63	0.42
Xscore(ChemScore)+MW+logP	0.90	143	346	44	105	0.58	0.89	0.77	0.72	0.50
GoldScore+Molecular Weight (MW)+ logP(Normalized)										
GoldScore+MW	0.40	165	308	82	83	0.67	0.79	0.74	0.73	0.46
Goldsscore+logP	0.70	142	342	48	106	0.57	0.88	0.76	0.71	0.48
GoldScore+MW+logP	0.70	153	332	58	95	0.62	0.85	0.76	0.73	0.49
Xscore(GoldScore)+Molecular Weight (MW)+ logP(Normalized)										
Xscore(GoldScore)+MW	0.50	156	328	62	92	0.63	0.84	0.76	0.73	0.48

Descriptor		TP	TN	FP	FN	Sensitivity	Specificity	Accuracy	G-mean	MCC
Xscore(GoldScore)+logP	---									
Xscore(GoldScore)+MW+logP	0.80	156	345	45	92	0.63	0.89	0.79	0.75	0.54

The scoring function in brackets was used to generate the docking poses.

Table S4 (a). Machine learning models obtained using ChemScore scoring function combined with physicochemical properties as descriptors for the training set

Descriptor	Machine Learning method	TP	TN	FP	FN	Sensitivity	Specificity	Accuracy	G-mean	MCC
MolecularWeight (MW)+logP										
MW+logP	Naive Bayes	63	284	11	50	0.56	0.96	0.85	0.73	0.60
MW+logP	LibSVM	37	287	8	76	0.33	0.97	0.79	0.56	0.43
MW+logP	J48	88	265	30	25	0.78	0.90	0.87	0.84	0.67
MW+logP	RF	95	266	29	18	0.84	0.90	0.88	0.87	0.72
MW+logP	REPTree	85	265	30	28	0.75	0.90	0.86	0.82	0.65
ChemScore+ MolecularWeight (MW)										
ChemScore+MW	Naive Bayes	75	272	23	38	0.66	0.92	0.85	0.78	0.61
ChemScore+MW	LibSVM	53	284	11	60	0.47	0.96	0.83	0.67	0.53
ChemScore+MW	J48	77	280	15	36	0.68	0.95	0.88	0.80	0.68

164

ChemScore+MW	RF	85	267	28	28	0.75	0.91	0.86	0.83	0.66
ChemScore+MW	REPTree	78	273	22	35	0.69	0.93	0.86	0.80	0.64
ChemScore+logP										
ChemScore+logP	Naive Bayes	78	263	32	35	0.69	0.89	0.84	0.78	0.59
ChemScore+logP	LibSVM	34	291	4	79	0.30	0.99	0.80	0.54	0.44
ChemScore+logP	J48	69	268	27	44	0.61	0.91	0.83	0.74	0.55
ChemScore+logP	RF	75	256	39	38	0.66	0.87	0.81	0.76	0.53
ChemScore+logP	REPTree	69	267	28	44	0.61	0.91	0.82	0.74	0.54
ChemScore+ Molecular Weight (MW)+logP										
ChemScore+MW+logP	Naive Bayes	83	271	24	30	0.74	0.92	0.87	0.82	0.66
ChemScore+MW+logP	LibSVM	53	288	7	60	0.47	0.98	0.84	0.68	0.56
ChemScore+MW+logP	J48	93	273	22	20	0.82	0.93	0.90	0.87	0.74
ChemScore+MW+logP	RF	91	267	28	22	0.81	0.91	0.88	0.85	0.70
ChemScore+MW+logP	REPTree	83	269	26	30	0.73	0.91	0.86	0.82	0.65
Xscore(ChemScore)										
Xscore_C+Molecular Weight (MW)										
Xscore_C+MW	Naive Bayes	82	276	19	31	0.73	0.94	0.88	0.82	0.69

Xscore_C+MW	LibSVM	72	286	9	41	0.64	0.97	0.88	0.79	0.68
Xscore_C+MW	J48	85	276	19	28	0.75	0.94	0.88	0.84	0.71
Xscore_C+MW	RF	85	267	28	28	0.75	0.91	0.86	0.83	0.66
Xscore_C+MW	REPTree	82	274	21	31	0.73	0.93	0.87	0.82	0.67
Xscore_C+logP										
Xscore_C+logP	Naive Bayes	80	276	19	33	0.71	0.94	0.87	0.81	0.67
Xscore_C+logP	LibSVM	67	285	10	46	0.59	0.97	0.86	0.76	0.64
Xscore_C+logP	J48	83	270	25	30	0.73	0.92	0.87	0.82	0.66
Xscore_C+logP	RF	79	273	22	34	0.70	0.93	0.86	0.80	0.65
Xscore_C+logP	REPTree	79	273	22	34	0.70	0.93	0.86	0.80	0.65
Xscore_C+MW+logP										
Xscore_C+MW+logP	Naive Bayes	80	282	13	30	0.73	0.96	0.89	0.83	0.72
Xscore_C+MW+logP	LibSVM	68	288	7	45	0.60	0.98	0.87	0.77	0.67
Xscore_C+MW+logP	J48	88	272	23	25	0.78	0.92	0.88	0.85	0.70
Xscore_C+MW+logP	RF	90	271	24	23	0.80	0.92	0.88	0.86	0.71
Xscore_C+MW+logP	REPTree	89	269	26	24	0.79	0.91	0.88	0.85	0.70

The scoring function in brackets was used to generate the docking poses. Xscore_C means Xscore(ChemScore).

Table S4 (b). Machine learning models obtained using ChemScore scoring function combined with physicochemical properties as descriptors for the test set (Pedersen et al.)

Descriptor	Machine Learning method	TP	TN	FP	FN	Sensitivity	Specificity	Accuracy	G-mean	MCC
MolecularWeight (MW)+logP										
MW+logP	Naive Bayes	31	110	12	13	0.70	0.90	0.85	0.80	0.61
MW+logP	LibSVM	26	112	10	18	0.59	0.92	0.83	0.74	0.55
MW+logP	J48	33	107	15	11	0.75	0.88	0.84	0.81	0.61
MW+logP	RF	35	104	18	9	0.80	0.85	0.84	0.82	0.61
MW+logP	REPTree	39	101	21	5	0.89	0.83	0.84	0.86	0.66
ChemScore+ MolecularWeight (MW)										
ChemScore+MW	Naive Bayes	33	103	19	11	0.75	0.84	0.82	0.80	0.57
ChemScore+MW	LibSVM	29	111	11	15	0.66	0.91	0.84	0.77	0.59
ChemScore+MW	J48	32	108	14	12	0.73	0.89	0.84	0.80	0.60
ChemScore+MW	RF	32	100	22	12	0.73	0.82	0.80	0.77	0.52
ChemScore+MW	REPTree	32	106	16	12	0.73	0.87	0.83	0.79	0.58
ChemScore+logP										

167

ChemScore+logP	Naive Bayes	37	100	22	7	0.84	0.82	0.83	0.83	0.61
ChemScore+logP	LibSVM	24	114	8	20	0.55	0.93	0.83	0.71	0.54
ChemScore+logP	J48	38	93	29	6	0.86	0.76	0.79	0.81	0.56
ChemScore+logP	RF	30	96	26	14	0.68	0.79	0.76	0.73	0.44
ChemScore+logP	REPTree	31	102	20	13	0.70	0.84	0.80	0.77	0.52
ChemScore+ Molecular Weight (MW) +logP										
ChemScore+MW+logP	Naive Bayes	35	104	18	9	0.80	0.85	0.84	0.82	0.61
ChemScore+MW+logP	LibSVM	29	111	11	15	0.66	0.91	0.84	0.77	0.59
ChemScore+MW+logP	J48	33	107	15	11	0.75	0.88	0.84	0.81	0.61
ChemScore+MW+logP	RF	36	103	19	8	0.82	0.84	0.84	0.83	0.62
ChemScore+MW+logP	REPTree	40	100	22	4	0.91	0.82	0.84	0.86	0.66
Xscore(ChemScore)										
Xscore_C+ Molecular Weight (MW)										
Xscore_C+MW	Naive Bayes	37	97	25	7	0.84	0.80	0.81	0.82	0.58
Xscore_C+MW	LibSVM	34	100	22	10	0.77	0.82	0.81	0.80	0.55
Xscore_C+MW	J48	37	97	25	7	0.84	0.80	0.81	0.82	0.58
Xscore_C+MW	RF	39	86	36	5	0.89	0.70	0.75	0.79	0.52

Xscore_C+MW	REPTree	38	86	36	6	0.86	0.70	0.75	0.78	0.50
Xscore_C+logP										
Xscore_C+logP	Naive Bayes	39	99	23	5	0.89	0.81	0.83	0.85	0.64
Xscore_C+logP	LibSVM	34	104	18	10	0.77	0.85	0.83	0.81	0.59
Xscore_C+logP	J48	38	80	42	6	0.86	0.66	0.71	0.75	0.46
Xscore_C+logP	RF	35	93	29	9	0.80	0.76	0.77	0.78	0.51
Xscore_C+logP	REPTree	38	79	43	6	0.86	0.65	0.70	0.75	0.45
Xscore_C+ Molecular Weight (MW)+logP										
Xscore_C+MW+logP	Naive Bayes	38	99	23	6	0.86	0.81	0.83	0.84	0.62
Xscore_C+MW+logP	LibSVM	36	108	14	8	0.82	0.89	0.87	0.85	0.68
Xscore_C+MW+logP	J48	37	93	29	7	0.84	0.76	0.78	0.80	0.54
Xscore_C+MW+logP	RF	34	100	22	10	0.77	0.82	0.81	0.80	0.55
Xscore_C+MW+logP	REPTree	36	90	32	8	0.82	0.74	0.76	0.78	0.50

The scoring function in brackets was used to generate the docking poses. Xscore_C means Xscore(ChemScore).

Table S4 (c). Machine learning models obtained using ChemScore scoring function combined with physicochemical properties as descriptors for the test set (AstraZeneca-unpublished)

Descriptor	Machine Learning method	TP	TN	FP	FN	Sensitivity	Specificity	Accuracy	G-mean	MCC
MolecularWeight (MW)+logP										
MW+logP	Naive Bayes	241	136	254	7	0.97	0.35	0.59	0.58	0.38
MW+logP	LibSVM	222	266	124	26	0.90	0.68	0.77	0.78	0.57
MW+logP	J48	136	247	143	112	0.55	0.63	0.60	0.59	0.18
MW+logP	RF	152	234	156	96	0.61	0.60	0.61	0.61	0.21
MW+logP	REPTree	240	204	186	8	0.97	0.52	0.70	0.71	0.51
ChemScore+ MolecularWeight (MW)										
ChemScore+MW	Naive Bayes	201	302	88	47	0.81	0.77	0.79	0.79	0.57
ChemScore+MW	LibSVM	155	332	58	93	0.81	0.77	0.79	0.79	0.57
ChemScore+MW	J48	176	328	62	72	0.71	0.84	0.79	0.77	0.56
ChemScore+MW	RF	190	298	92	58	0.77	0.76	0.77	0.77	0.52
ChemScore+MW	REPTree	172	333	57	76	0.69	0.85	0.79	0.77	0.56
ChemScore+logP										

ChemScore+logP	Naive Bayes	238	118	272	10	0.96	0.30	0.56	0.54	0.32
ChemScore+logP	LibSVM	204	230	160	44	0.82	0.59	0.68	0.70	0.41
ChemScore+logP	J48	1818	93	297	67	0.96	0.24	0.84	0.48	0.30
ChemScore+logP	RF	170	122	268	78	0.69	0.31	0.46	0.46	0.00
ChemScore+logP	REPTree	212	214	176	36	0.86	0.55	0.67	0.69	0.40
ChemScore+ Molecular Weight (MW)+logP										
ChemScore+MW+logP	Naive Bayes	238	144	246	10	0.96	0.37	0.60	0.60	0.38
ChemScore+MW+logP	LibSVM	211	284	106	37	0.85	0.73	0.78	0.79	0.56
ChemScore+MW+logP	J48	136	247	143	112	0.55	0.63	0.60	0.59	0.18
ChemScore+MW+logP	RF	204	228	162	44	0.82	0.59	0.68	0.69	0.40
ChemScore+MW+logP	REPTree	239	212	178	9	0.96	0.54	0.71	0.72	0.52
Xscore(ChemScore)										
Xscore_C+ Molecular Weight (MW)										
Xscore_C+MW	Naive Bayes	129	345	45	119	0.52	0.89	0.74	0.68	0.44
Xscore_C+MW	LibSVM	106	358	32	142	0.43	0.92	0.73	0.63	0.41
Xscore_C+MW	J48	129	348	42	120	0.52	0.89	0.75	0.68	0.45
Xscore_C+MW	RF	114	327	63	104	0.52	0.84	0.73	0.66	0.38

Xscore_C+MW	REPTree	175	325	65	73	0.71	0.83	0.78	0.77	0.54
Xscore_C+logP										
Xscore_C+logP	Naive Bayes	223	203	187	25	0.90	0.52	0.67	0.68	0.43
Xscore_C+logP	LibSVM	168	326	64	80	0.68	0.84	0.77	0.75	0.52
Xscore_C+logP	J48	141	346	44	107	0.57	0.89	0.76	0.71	0.49
Xscore_C+logP	RF	132	325	65	116	0.53	0.83	0.72	0.67	0.39
Xscore_C+logP	REPTree	178	310	80	70	0.72	0.80	0.77	0.76	0.51
Xscore_C+ Molecular Weight (MW)+logP										
Xscore_C+MW+logP	Naive Bayes	226	242	148	22	0.91	0.62	0.73	0.75	0.53
Xscore_C+MW+logP	LibSVM	170	335	55	78	0.69	0.86	0.79	0.77	0.56
Xscore_C+MW+logP	J48	167	331	59	81	0.67	0.85	0.78	0.76	0.53
Xscore_C+MW+logP	RF	183	292	98	65	0.74	0.75	0.75	0.74	0.48
Xscore_C+MW+logP	REPTree	175	325	65	73	0.71	0.83	0.78	0.77	0.54

The scoring function in brackets was used to generate the docking poses. Xscore_C means Xscore(ChemScore).

Table S5 (a). Summary of models obtained using PLIF results for the training set

Scoring Function	Critical value	TP	TN	FP	FN	Sensitivity	Specificity	Accuracy	G-mean	MCC	Precision
Residues	0.105	86	220	75	27	0.76	0.75	0.75	0.75	0.46	0.53
Resudues+Interaction Type	0.0739	83	219	76	30	0.734	0.74	0.74	0.74	0.44	0.52
Residue+functional groups	0.0316	91	235	60	22	0.82	0.80	0.80	0.80	0.56	0.60
Sequential approach—TP+FP--ChemScore											
Residues	0.105	57	281	14	56	0.50	0.95	0.83	0.69	0.54	0.80
Resudues+Interaction Type	0.0739	56	281	14	57	0.50	0.95	0.83	0.69	0.53	0.80
Residue+functional groups	0.0316	58	286	9	55	0.51	0.97	0.84	0.71	0.58	0.87

173

Table S5 (b). Summary of models obtained using PLIF results for the test set (Pedersen et. al)

Scoring Function	Critical value	TP	TN	FP	FN	Sensitivity	Specificity	Accuracy	G-mean	MCC	Precision
Residues	0.105	41	80	32	3	0.93	0.71	0.78	0.82	0.58	0.56
Resudues+Interaction Type	0.0739	37	82	40	7	0.84	0.67	0.72	0.75	0.45	0.48
Residue+functional groups	0.0316	33	95	27	11	0.75	0.78	0.77	0.76	0.49	0.55

Sequential approach—TP+FP--ChemScore

	Critical value	TP	TN	FP	FN	Sensitivity	Specificity	Accuracy	G-mean	MCC	Precision
Residues	0.105	33	109	13	11	0.75	0.89	0.86	0.82	0.64	0.72
Resudues+Interaction Type	0.0739	29	110	12	15	0.66	0.90	0.84	0.77	0.57	0.71
Residue+functional groups	0.0316	28	111	11	16	0.64	0.91	0.84	0.76	0.57	0.72

Table S5 (c). Summary of models obtained using PLIF results for the test set (AstraZeneca-unpublished)

Scoring Function	Critical value	TP	TN	FP	FN	Sensitivity	Specificity	Accuracy	G-mean	MCC	Precision
Residues	0.105	205	287	103	43	0.83	0.74	0.77	0.78	0.55	0.67
Resudues+Interaction Type	0.0739	207	290	100	41	0.84	0.74	0.78	0.79	0.56	0.67
Residue+functional groups	0.0316	191	306	84	57	0.77	0.79	0.78	0.78	0.55	0.70
Sequential approach—TP+FP--ChemScore											
Residues	0.105	135	350	40	113	0.54	0.90	0.76	0.70	0.48	0.77
Resudues+Interaction Type	0.0739	138	355	35	110	0.56	0.91	0.77	0.71	0.51	0.80
Residue+functional groups	0.0316	131	356	34	117	0.53	0.91	0.76	0.69	0.49	0.79

Figures

Figure S1. Residues which show hydrophobic interactions with a high interaction rate and a low root mean square fluctuation.
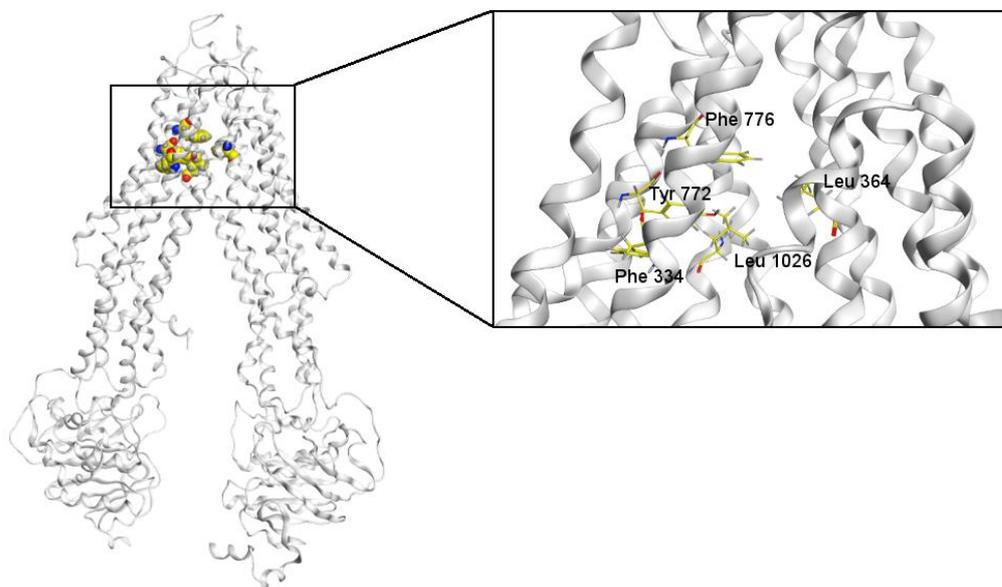
Figure S2. CNS representation of the training set compounds based on MACCS Tc similarity threshold of 0.70. Communities with at least five representative members are color coded. Also shown below are the few exemplary compound (with their IC50 value in µM) of highlighted communities.
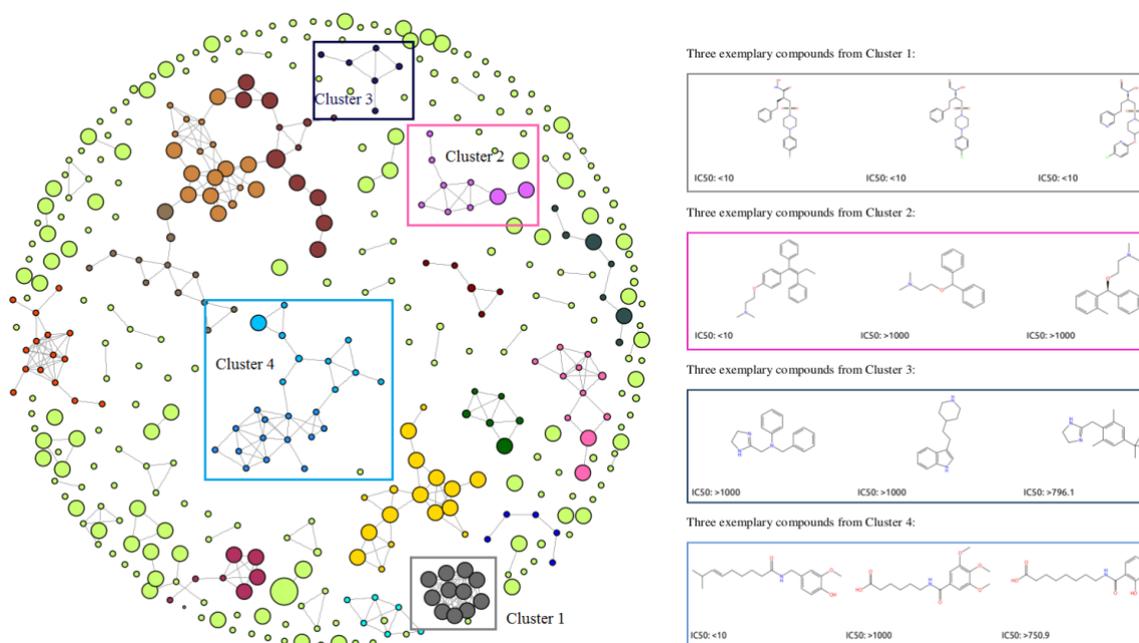
Figure S3. CNS representation of the test set compounds (Pedersen et al.) based on MACCS Tc similarity threshold of 0.70. Communities with at least five representative members are color coded. Also shown are the few exemplary compound of these communities.
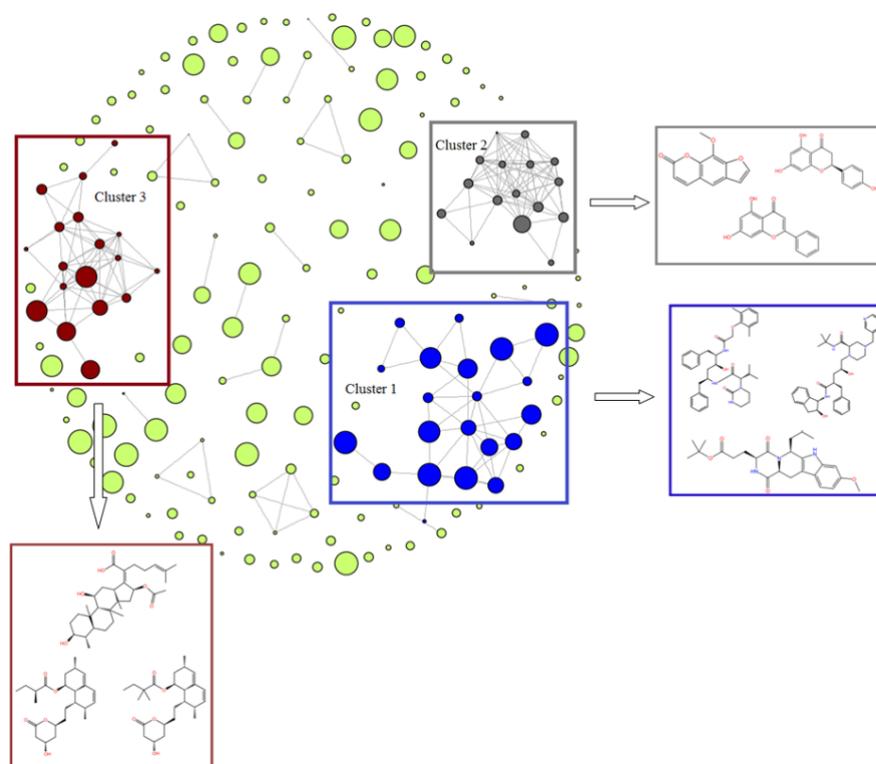
Figure S4. Sequence alignment of human BSEP with corrected mouse P-glycoprotein structure (PDB ID: 4M1M). The residues are colored according to the ClustalX scheme using Jalview.

Figure S5. Ramachandran plot for the final homology model of human BSEP taken from PDBsum.
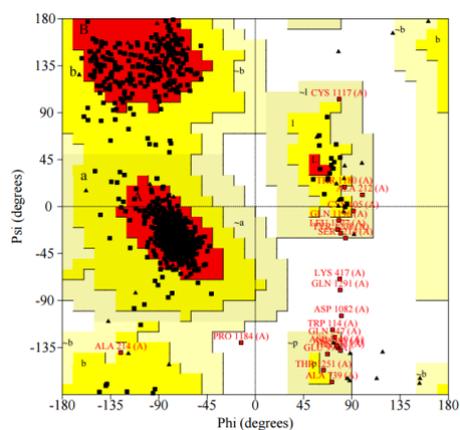
Figure S6. Residues that are present in the disallowed region in the final BSEP homology model.
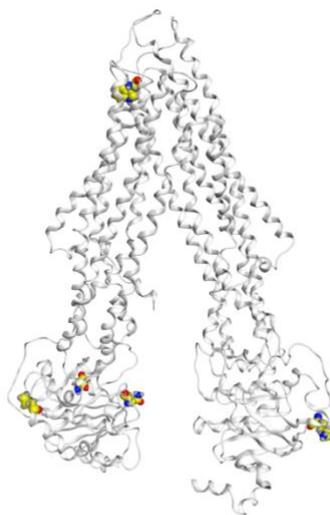
Figure S7. The location of Asn109, Asn116, Asn122 and Asn125 in EL1 of the BSEP homology model. The carbon atoms of the amino acids are colored in yellow for a better visibility.

177
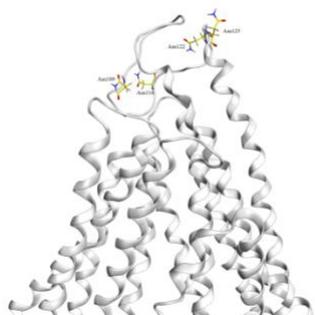
Figure S8. The ROC curve of ChemScore scores of training set compounds. The area under the ROC curve is 0.87.
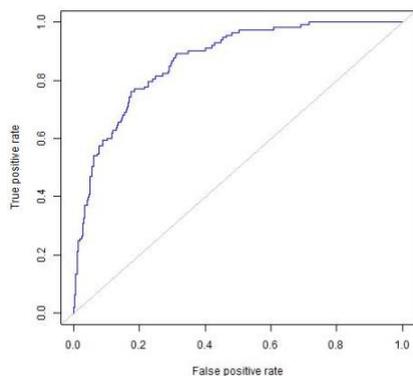
Figure S9. The ROC curve of GoldScore scores of training set compounds. The area under the ROC curve is 0.82.
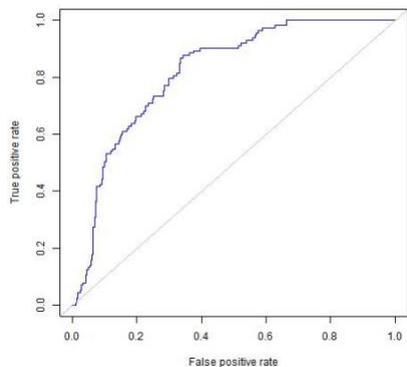
Figure S10. The ROC curve of GlideXP scores of training set compounds. The area under the ROC curve is 0.77.
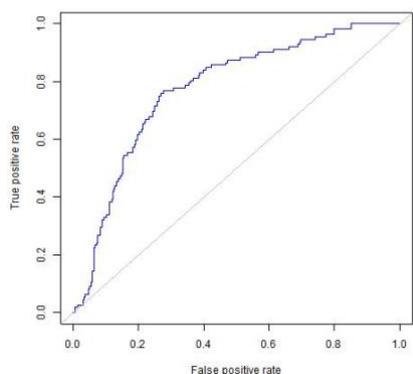
Figure S11. The ROC curve of Xscore(ChemScore) scores of training set compounds. The area under the ROC curve is 0.92.
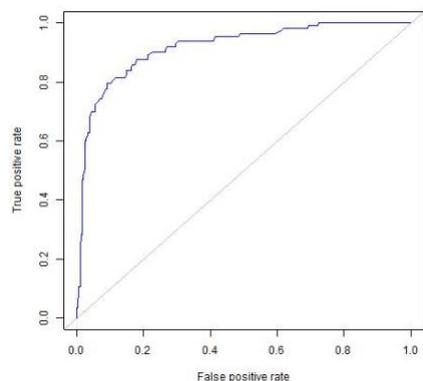
Figure S12. The ROC curve of Xscore(GoldScore) scores of training set compounds. The area under the ROC curve is 0.93.
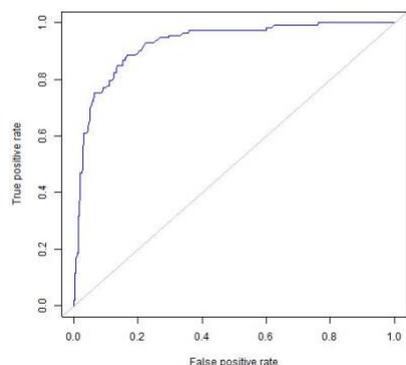
Figure S13. (a) Hydrophobic interaction - (b) hydrogen bond interaction fingerprints of true positives (TPs) and true negatives (TNs) of the test set (Pedersen et al.). The classification of the compounds is based on the ChemScore scoring function.

a
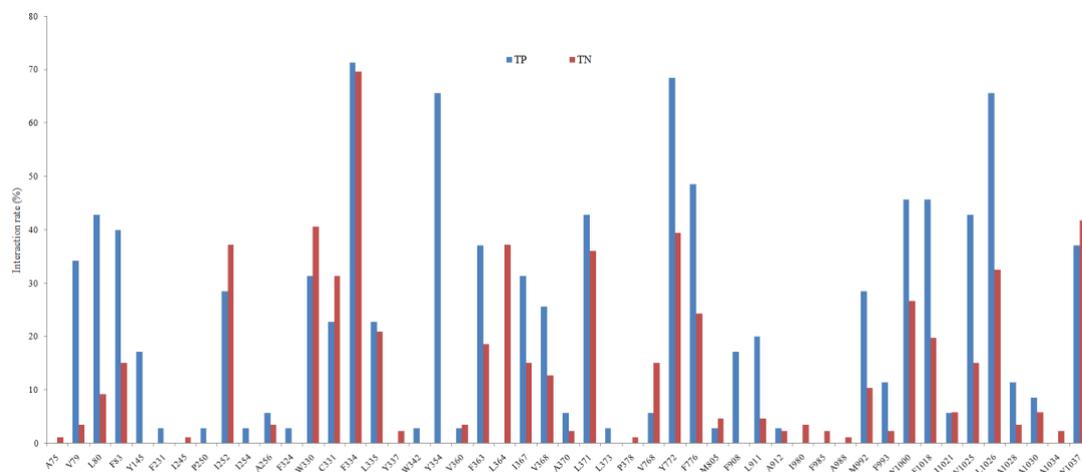
b
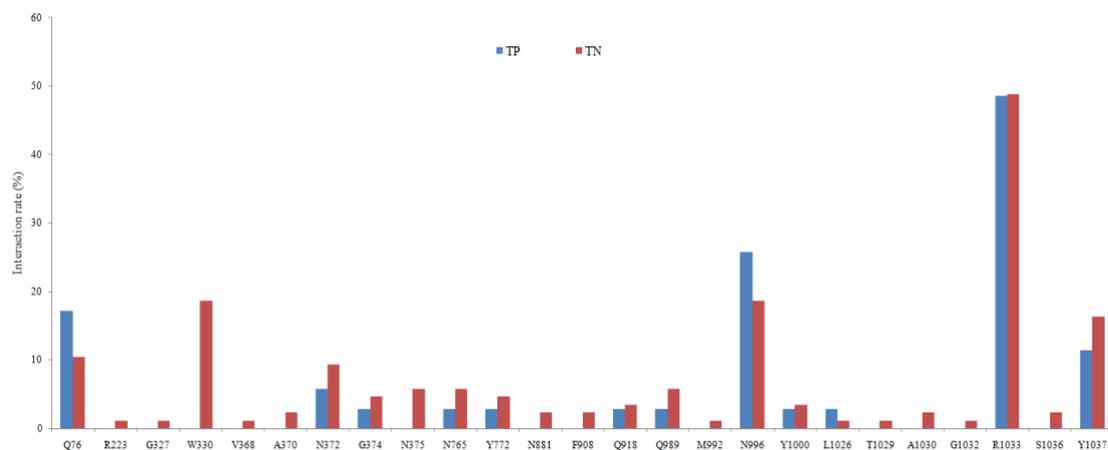
Figure S14. (a) Hydrophobic interaction - (b) hydrogen bond interaction fingerprints of true positives (TPs) and true negatives (TNs) of the test set (AstraZeneca-unpublished). The classification of the compounds is based on the ChemScore scoring function.

a

b

180
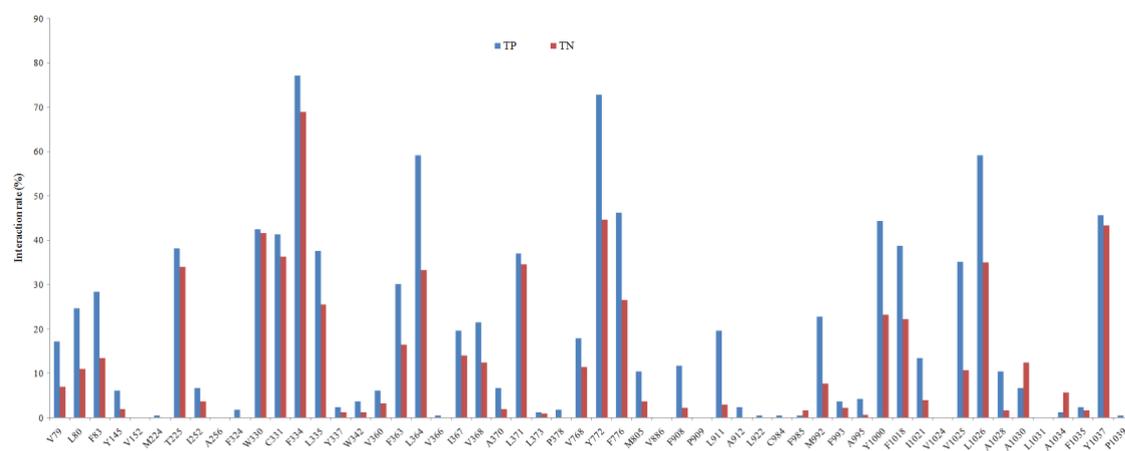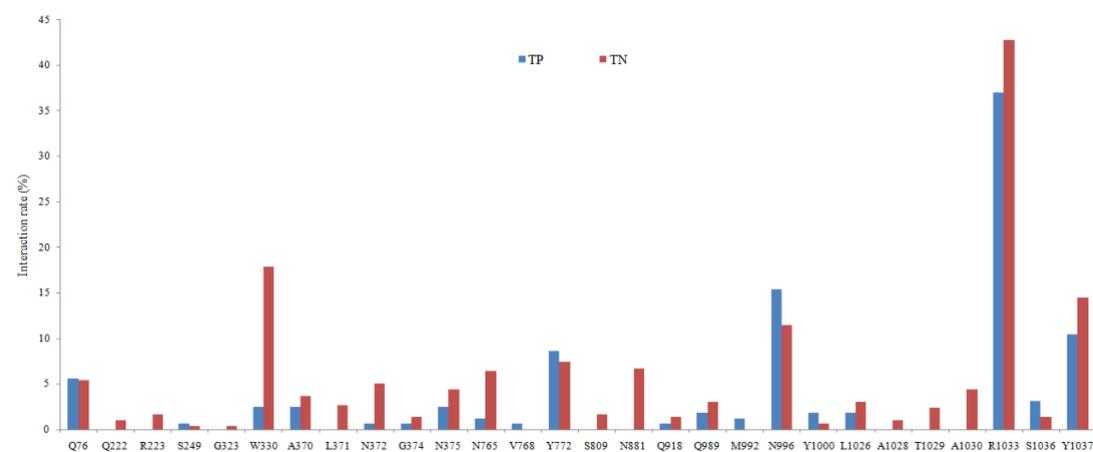
Figure S15. Distribution of BSEP inhibitors and non-inhibitors based on the (a) Molecular Weight (b) logP(o/w) of the test set (Pedersen et al.)

a
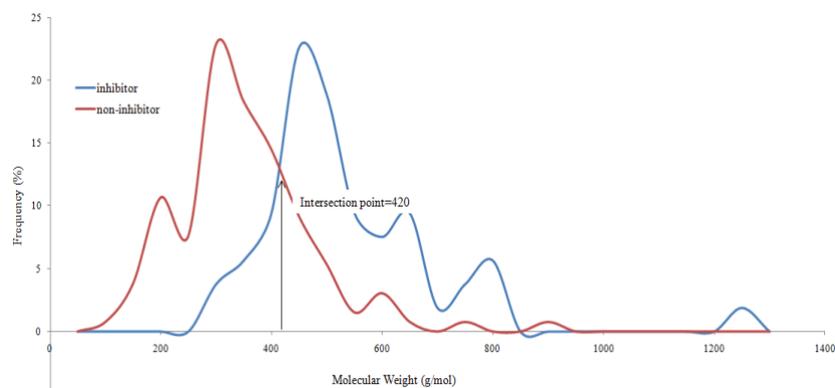
b
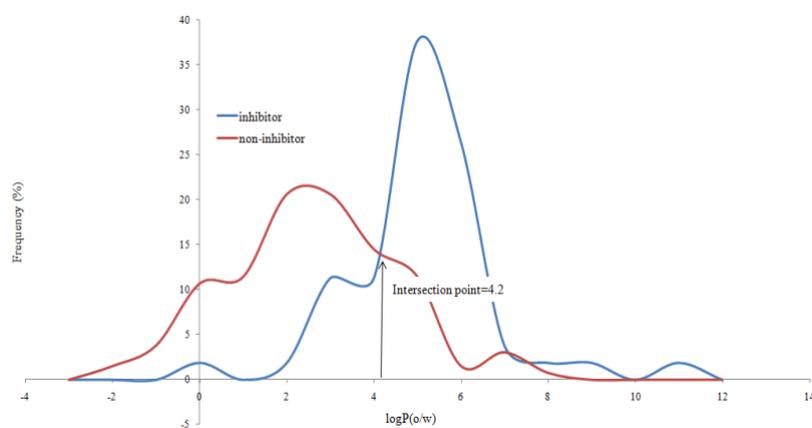
Figure S16. Distribution of BSEP inhibitors and non-inhibitors based on the (a) Molecular Weight (b) logP(o/w) of the test set (AstraZeneca-unpublished)

a
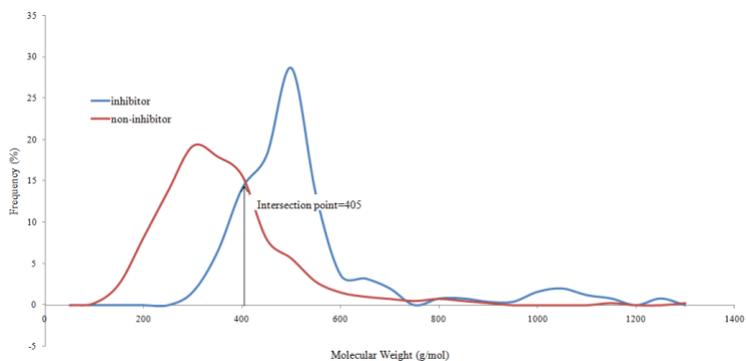
b
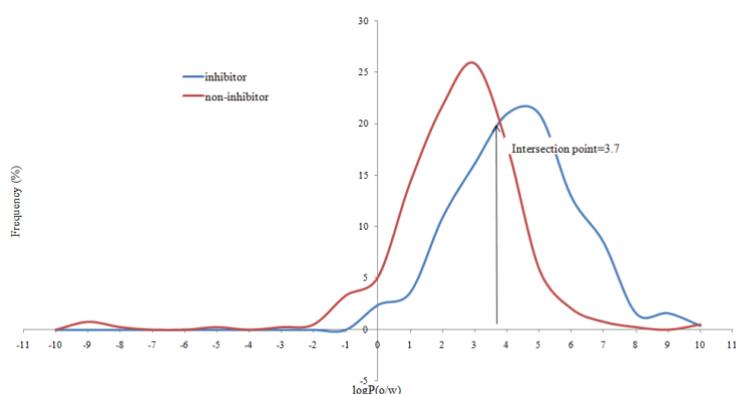
Figure S17. Distribution of BSEP inhibitors and non-inhibitors (training set) based on GoldScore scoring. Sensitivity, specificity, precision and MCC were calculated from the confusion matrix based on the intersection point of both curves.
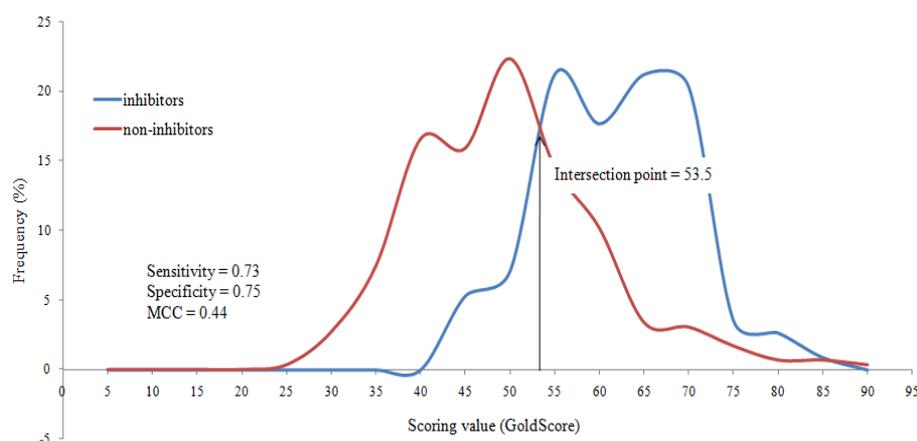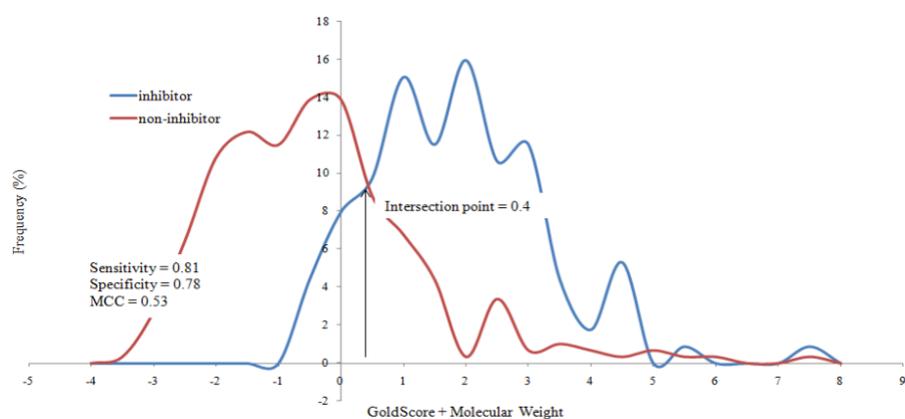

Figure S18. Distribution of BSEP inhibitors and non-inhibitors (training set) based on GlideXP scoring. Sensitivity, specificity, precision and MCC were calculated from the confusion matrix based on the intersection point of both curves.
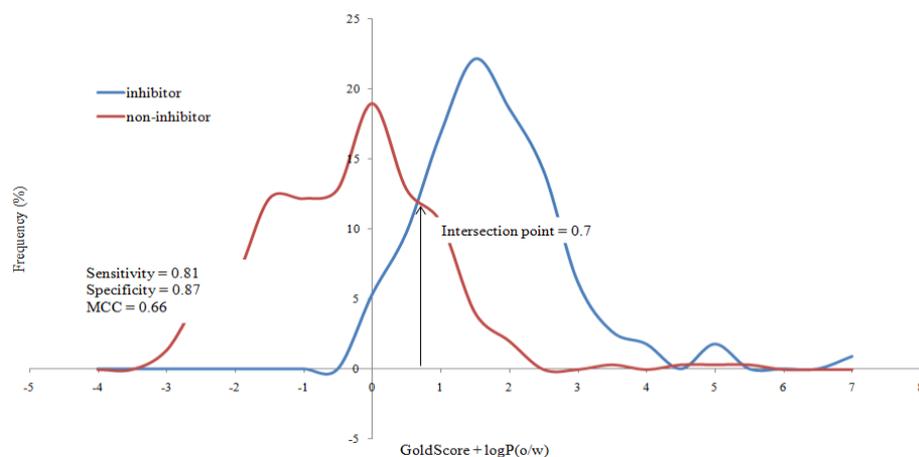

182
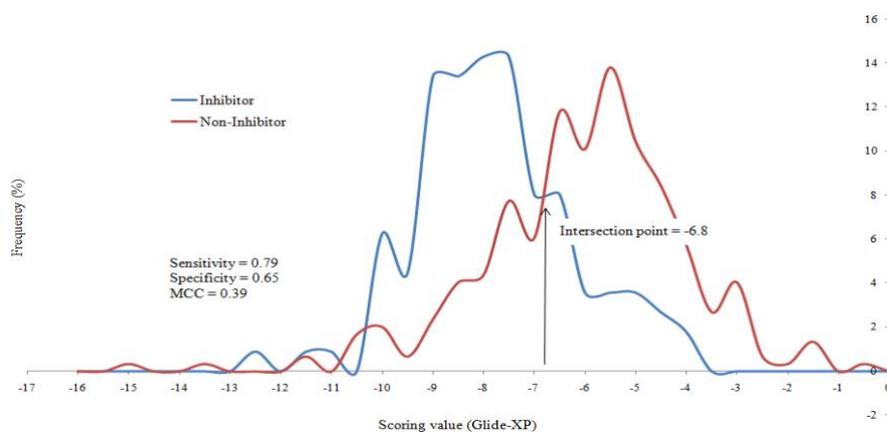
Figure S19. Distribution of BSEP inhibitors and non-inhibitors (training set) based on rescoring using Xscore score (poses generated using ChemScore). Sensitivity, specificity, precision and MCC were calculated from the confusion matrix based on the intersection point of both curves.
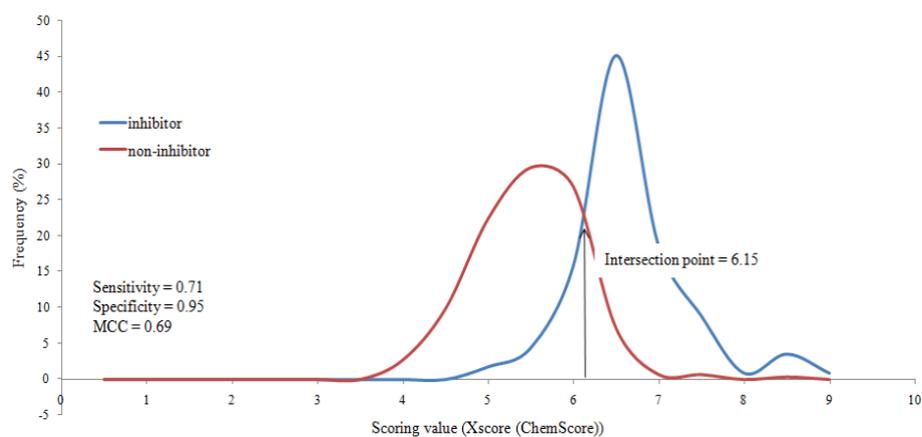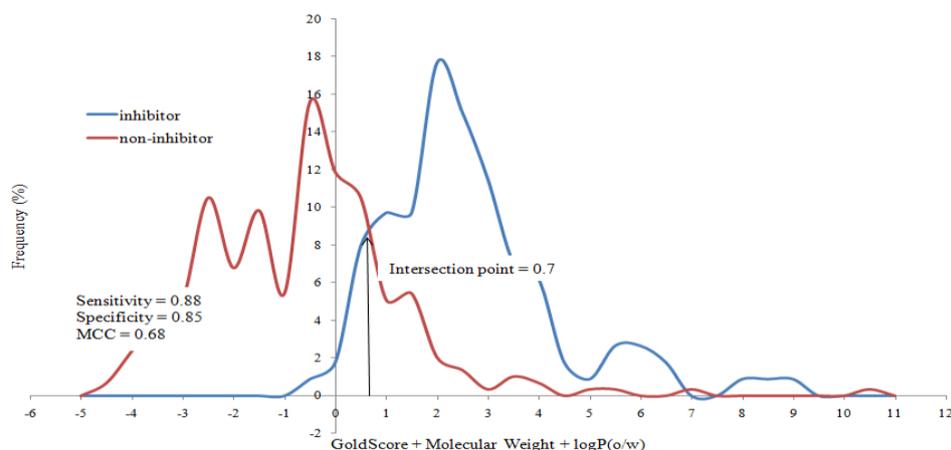

Figure S20. Distribution of BSEP inhibitors and non-inhibitors (training set) based on rescoring using Xscore score (poses generated using GoldScore). Sensitivity, specificity, precision and MCC were calculated from the confusion matrix based on the intersection point of both curves.
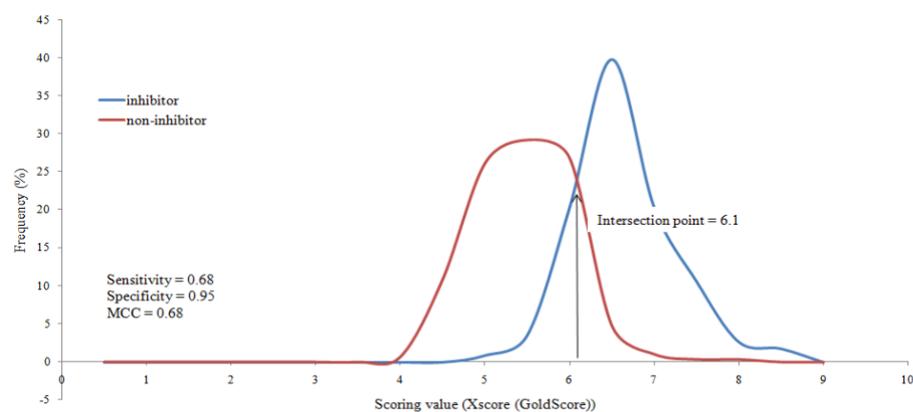
Figure S21. Distribution of BSEP inhibitors and non-inhibitors (training set) based on ChemScore scoring and molecular weight. Sensitivity, specificity, precision and MCC were calculated from the confusion matrix based on the intersection point of both curves.
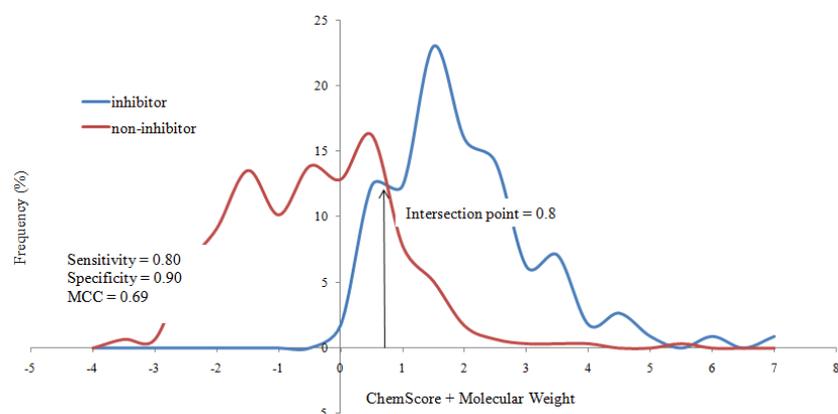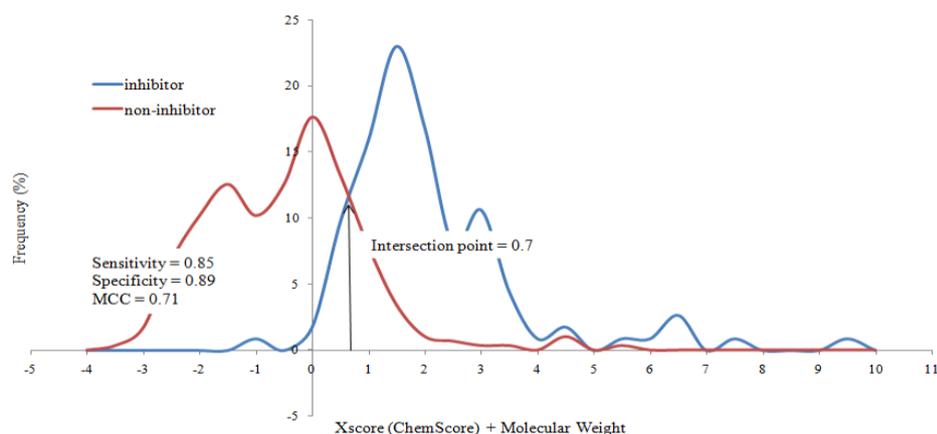
Figure S22. Distribution of BSEP inhibitors and non-inhibitors (training set) based on ChemScore scoring and logP. Sensitivity, specificity, precision and MCC were calculated from the confusion matrix based on the intersection point of both curves.
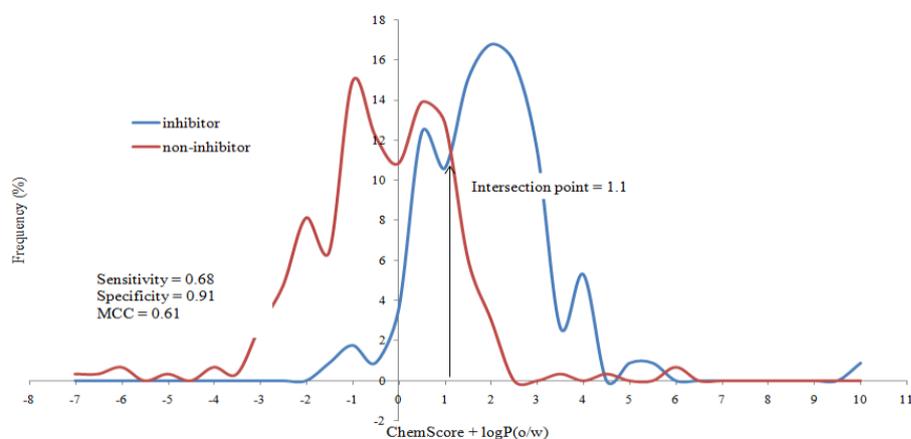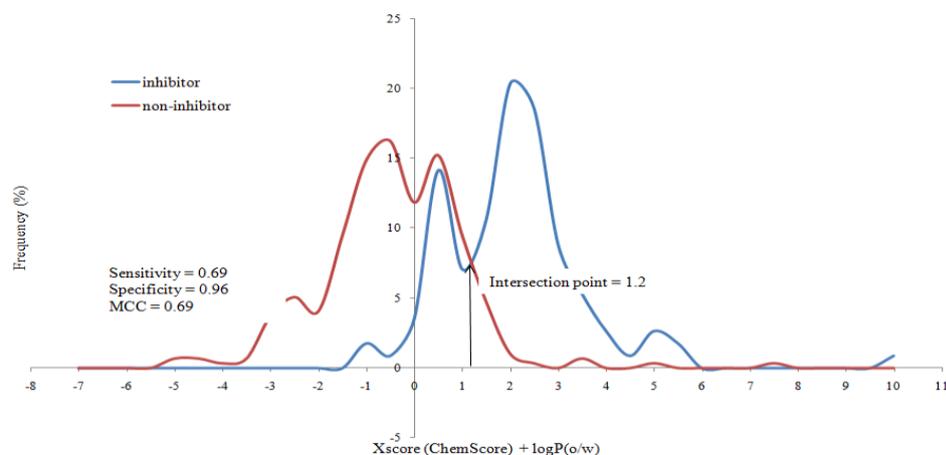

Figure S23. Distribution of BSEP inhibitors and non-inhibitors (training set) based on ChemScore scoring and Molecular Weight and logP. Sensitivity, specificity, precision and MCC were calculated from the confusion matrix based on the intersection point of both curves.
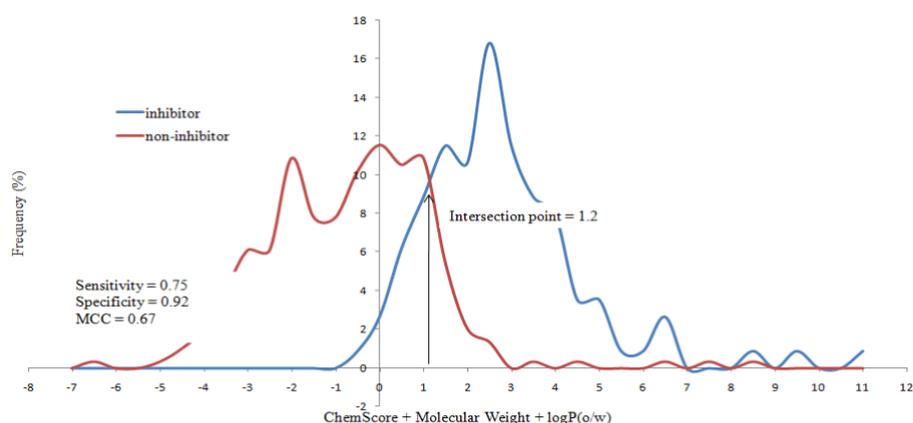
Figure S24. Distribution of BSEP inhibitors and non-inhibitors (training set) based on GoldScore rescoring and Molecular Weight. Sensitivity, specificity, precision and MCC were calculated from the confusion matrix based on the intersection point of both curves.
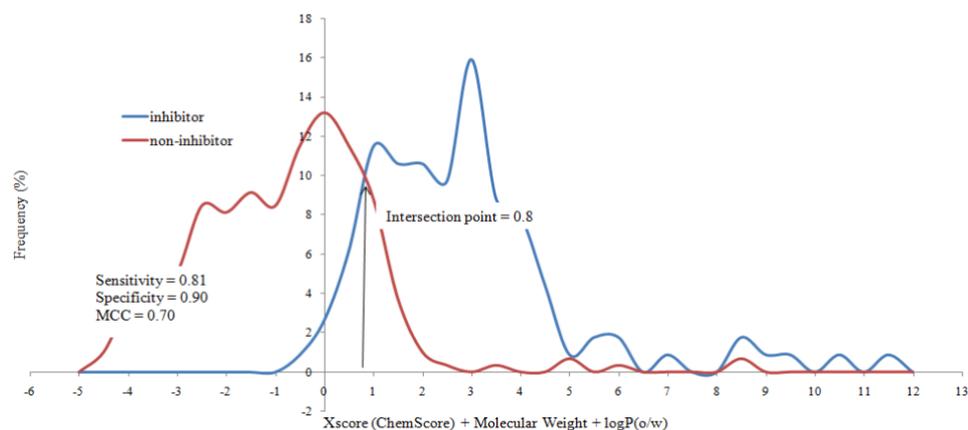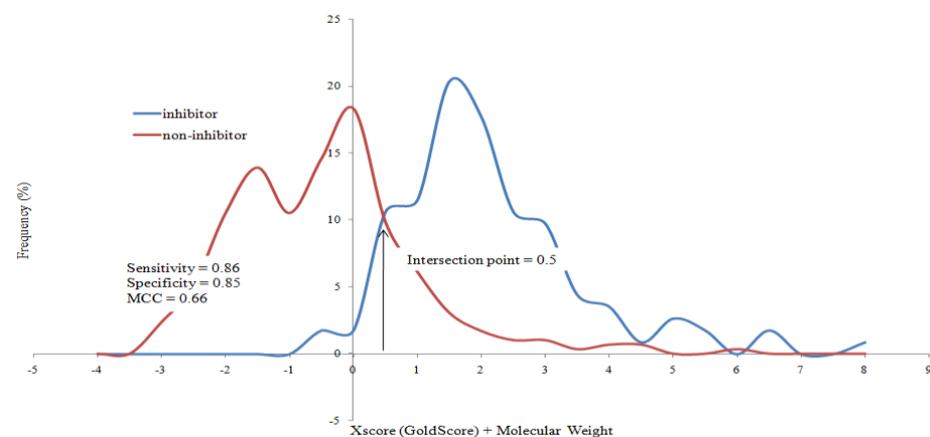
Figure S25. Distribution of BSEP inhibitors and non-inhibitors (training set) based on GoldScore scoring and logP. Sensitivity, specificity, precision and MCC were calculated from the confusion matrix based on the intersection point of both curves.

Figure S26. Distribution of BSEP inhibitors and non-inhibitors (training set) based on GoldScore scoring and Molecular Weight and logP. Sensitivity, specificity, precision and MCC were calculated from the confusion matrix based on the intersection point of both curves.
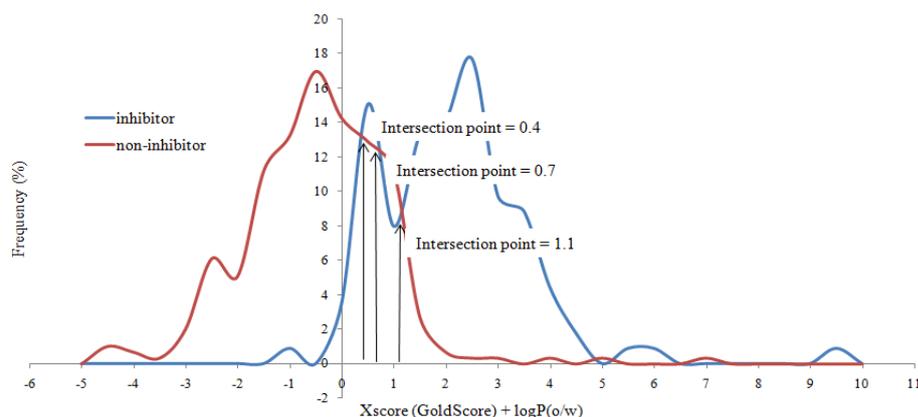

Figure S27. Distribution of BSEP inhibitors and non-inhibitors (training set) based on Xscore (ChemScore) and Molecular Weight. Sensitivity, specificity, precision and MCC were calculated from the confusion matrix based on the intersection point of both curves.

Figure S28. Distribution of BSEP inhibitors and non-inhibitors (training set) based on Xscore (ChemScore) and logP. Sensitivity, specificity, precision and MCC were calculated from the confusion matrix based on the intersection point of both curves.

**Figure S29.** Distribution of BSEP inhibitors and non-inhibitors (training set) based on Xscore (ChemScore) and Molecular Weight and logP. Sensitivity, specificity, precision and MCC were calculated from the confusion matrix based on the intersection point of both curves.


Figure S30. Distribution of BSEP inhibitors and non-inhibitors (training set) based on Xscore (GoldScore) and Molecular Weight. Sensitivity, specificity, precision and MCC were calculated from the confusion matrix based on the intersection point of both curves.

Figure S31. Distribution of BSEP inhibitors and non-inhibitors (training set) based on Xscore (GoldScore) and logP. Sensitivity, specificity, precision and MCC were calculated from the confusion matrix based on the intersection point of both curves.

Figure S32. Distribution of BSEP inhibitors and non-inhibitors (training set) based on Xscore (GoldScore) and Molecular Weight and logP. Sensitivity, specificity, precision and MCC were calculated from the confusion matrix based on the intersection point of both curves.
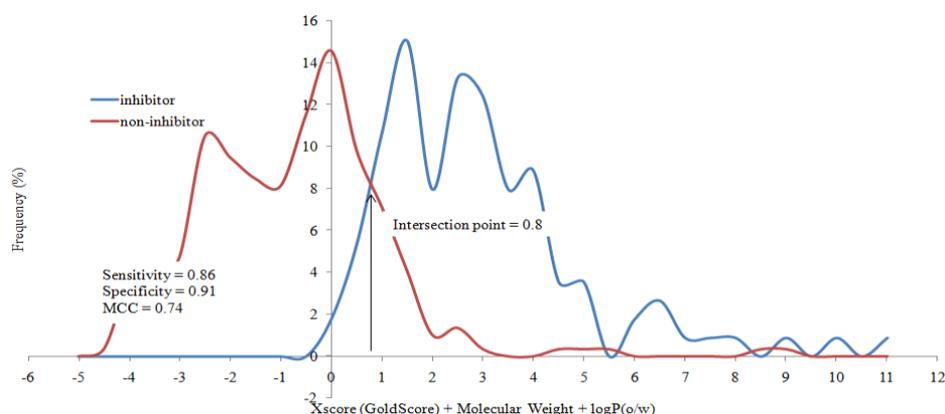
Figure S33. Distribution of functional groups in the test set (a) Pedersen et al. (b) AstraZeneca (unpublished) dataset classified using ChemScore rescoring function.

a
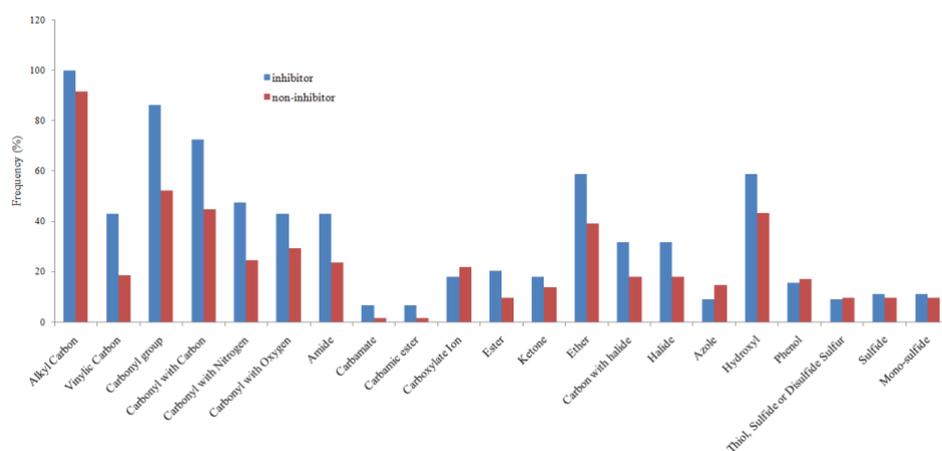
b
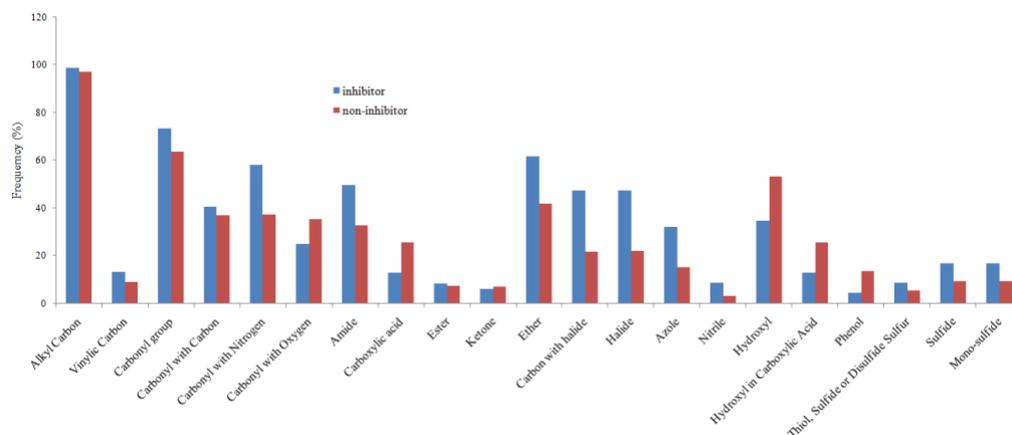
Figure S34. Heat map from PLIF analysis for training set non-inhibitors (x-axis: contact residues; y-axis: functional groups in the ligand showing interaction with the residue; color scale: number of interacting ligands).
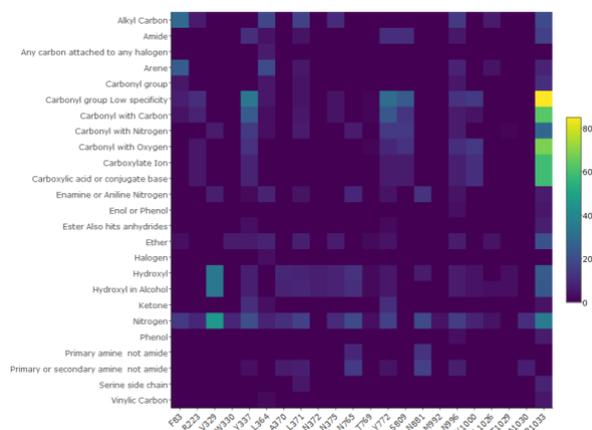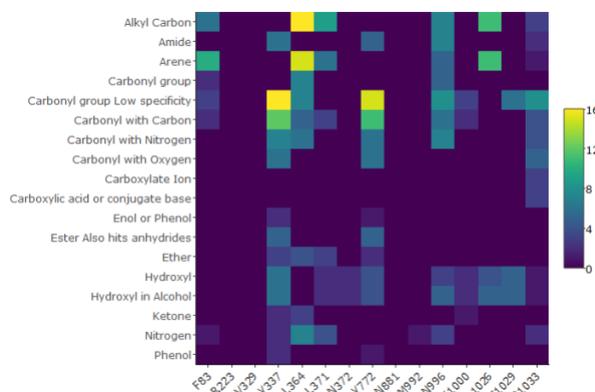

Figure S35. Heat map from PLIF analysis for test set (a) inhibitors (b) non-inhibitors (Pedersen et al.) (x-axis: contact residues; y-axis: functional groups in the ligand showing interaction with the residue; color scale: number of interacting ligands).
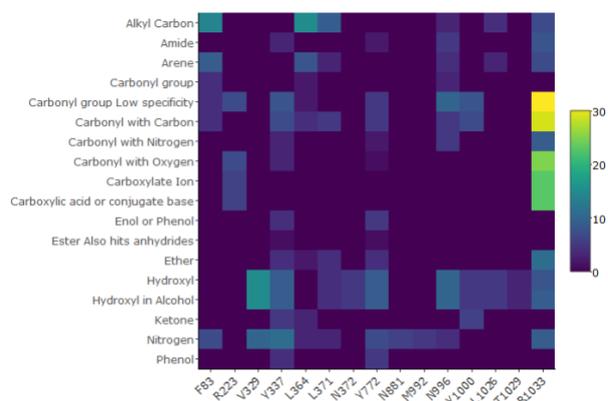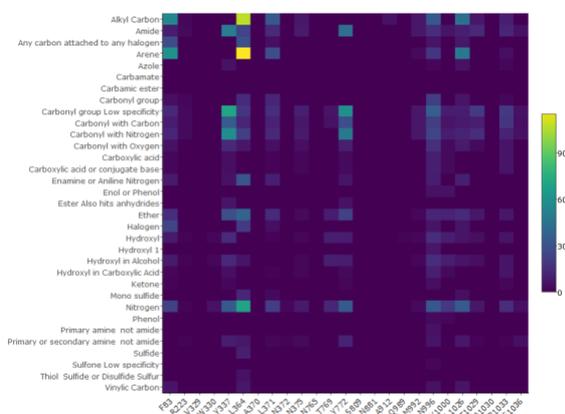
a

b

Figure S36. Heat map from PLIF analysis for test set (a) inhibitors (b) non-inhibitors (AstraZeneca-unpublished) (x-axis: contact residues; y-axis: functional groups in the ligand showing interaction with the residue; color scale: number of interacting ligands).
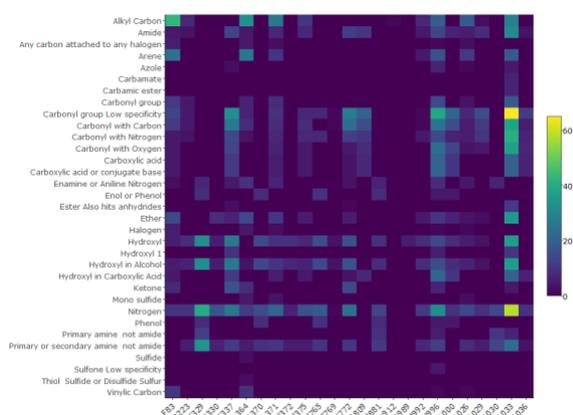
a

b

Classification using Protein Ligand Interaction Fingerprints (PLIF)

Finally, we wanted to assess a measure of PLIF homogeneity within the inhibitors of the training set. Therefore, we calculated the Tanimoto coefficients for each inhibitor versus the remaining inhibitors on basis of their PLIFs, and finally averaged the resulting coefficients. The averaged PLIF Tanimoto coefficient describes an inhibitor's PLIF similarity in relation to all inhibitors. The same procedure was undertaken for all inhibitors in the training set. From the distribution of averaged coefficients we calculated the mean and the standard deviation. Finally, a critical value was defined by subtraction of the standard deviation from the mean. This critical value was used as a threshold to classify compounds as inhibitors or non-inhibitors from the test dataset. To classify a test compound with this approach, the PLIF vector of the compound is used to calculate Tanimoto similarites against all compounds of the inhibitors in the training set. After averaging the calculated coefficients of the test compound, the resulting mean is compared against the critical value. If the averaged Tanimoto coefficients of the test compound is greater than the critical value, it is classified as an inhibitor, otherwise as a non-inhibitor. The PLIF-based classification provided accuracy measures comparable to those obtained from the docking score based classification (Table S5 in the supplementary material).

Moreover, information obtained using PLIF analysis in a sequential fashion i.e. reassessment of true positives and false positives obtained via the docking score based classification using PLIF-based similarity, improved the classification precision for both the training and the external test datasets (Table S5 in the supplementary material). The highest precision was obtained using the third PLIF approach that encoded residues along with the functional groups of the interacting ligand. Using this method, we achieved a precision of 0.87 (accuracy = 84%) for the training set. The same model showed a precision of 0.72 for the test dataset from Pedersen et al. [34] (accuracy = 84%) and 0.79 for the test dataset from AstraZeneca (accuracy = 76%). Overall, the number of false positives could be significantly reduced using the PLIF based classification.

7. Supplements to Section 4.3.1

<u>Supplementary Material</u>

Interspecies comparison of putative ligand binding sites of human, rat and mouse P-glycoprotein

Sankalp Jain[a], Melanie Grandits[a] and Gerhard F. Ecker[a*]

[a]*University of Vienna, Department of Pharmaceutical Chemistry, Althanstrasse 14, 1090 Vienna, Austria*

*Corresponding author: E-Mail: gerhard.f.ecker@univie.ac.at; Phone: +43-1-4277-55110; eFax: +43-1-4277-855110

Appendix

Supplementary Tables

Supplementary Table S1. Mapped residues in the H-site and R-site in the five models. For example, residue His61 in human corresponds to His60 (rat MDR1a), His59 (rat MDR1b), His60 (mouse mdr1a) and His60 (mouse mdr1b).

Human MDR1	Rat MDR1a	Rat MDR1b	Mouse mdr1a	Mouse mdr1b
H-site				
His61	His60	His59	His60	His60
Val125	Val117	Val124	Val121	Val124
Leu126	Leu118	Leu125	Leu122	Leu125
Ala129	Ala121	Ala128	Ala125	Ala128
Gln132	Gln124	Gln131	Gln128	Gln131
Val133	Val125	Val132	Val129	Val132
Trp136	Trp128	Trp135	Trp132	Trp135
Cys137	Cys129	Cys136	Cys133	Cys136
Asn183	Asn175	Asn182	Asn179	Asn182
Glu184	Glu176	Asp183	Glu180	Asp183
Gly185	Gly177	Gly184	Gly181	Gly184
Ile186	Ile178	Ile185	Ile182	Ile185
Gly187	Gly179	Gly186	Gly183	Gly186
Asp188	Asp180	Asp187	Asp184	Asp187
Ile190	Ile182	Leu189	Ile186	Ile189
Gly191	Gly183	Gly190	Gly187	Gly190
Met192	Met184	Met191	Met188	Met191
Phe194	Phe186	Phe193	Phe190	Phe193
Gln195	Gln187	Gln194	Gln191	Gln194
Leu245	Gln237	Gln244	His241	Gln244
Ser344	Ser336	Ser343	Ser340	Ser343
Val345	Val337	Ile344	Val341	Ile344
Gln347	Gln339	His346	Gln343	His346

Ser349	Ser341	Ala348	Ser345	Ala348
Pro350	Pro342	Pro349	Pro346	Pro349
Ser351	Asn343	Asn350	Asn347	Asn350
Glu353	Glu345	Glu352	Glu349	Glu352
Ala354	Ala346	Ala353	Ala350	Ala353
Ala355	Ala347	Ala354	Ala351	Ala354
Arg680	Gly642	Arg678	Arg676	Arg678
Lys681	Glu673	Arg679	Lys677	Arg679
Leu682	Leu674	Leu680	Leu678	Leu680
Leu879	Leu871	Leu879	Leu875	Leu877
Ser880	Ser872	Ser880	Ser876	Ser878
Leu884	Leu876	Leu884	Leu880	Leu882
Ala901	Ala893	Ala901	Ala897	Ala899
Lys934	Lys926	Lys934	Lys930	Lys932
Phe938	Phe930	Phe938	Phe934	Phe936
Phe942	Phe934	Phe942	Phe938	Phe940
Ser943	Ser935	Ala943	Ser939	Ser941
Gln946	Gln938	Gln946	Gln942	Gln944
Ala947	Ala939	Ala947	Ala943	Ala945
Tyr950	Tyr942	Tyr950	Tyr946	Tyr948
Asp997	Asp989	Asp997	Asp993	Asp995
Lys1000	Lys992	Lys1000	Lys996	Lys998
R-site				
Ala233	Ala225	Ala232	Ala229	Ala232
Thr240	Thr232	Thr239	Thr236	Thr239
Asp241	Asp233	Asn240	Asp237	Asn240
Leu244	Leu236	Leu243	Leu240	Leu243
Leu245	Gln237	Gln244	His241	Gln244
Ile293	Ile285	Ile292	Ile289	Ile292
Asn296	Asn288	Asn295	Asn292	Ser295
Ile299	Met291	Ile298	Met295	Ile298
Gly300	Gly292	Gly299	Gly296	Gly299
Phe303	Phe295	Tyr302	Phe299	Tyr302
Leu304	Leu296	Leu303	Leu300	Leu303
Ile340	Ile332	Leu339	Ile336	Leu339
Phe343	Phe335	Phe342	Phe339	Phe342
Ser344	Ser336	Ser343	Ser340	Ser343
Val345	Val337	Ile344	Val341	Ile344
Gly346	Gly338	Gly345	Gly342	Gly345
Gln347	Gln339	His346	Gln343	His346
Ala348	Ala340	Leu347	Ala344	Leu347
Ser349	Ser341	Ala348	Ser345	Ala348
Pro350	Pro342	Pro349	Pro346	Pro349
Glyu353	Glu345	Glu352	Glu349	Glu352
Gln678	Gln670	Gln676	Gln674	Gln676
Asp679	Asp671	Glu677	Asn675	Glu677
Arg680	Gly672	Arg678	Arg676	Arg678
Asn721	Asn713	Asn719	Asn717	Asn719
Leu724	Leu716	Ile722	Leu720	Ile722
Gln725	Gln717	Gln723	Gln721	Gln723
Phe728	Phe720	Phe726	Phe724	Phe726
Ser766	Ser758	Ser764	Ser762	Ser764
Thr769	Thr761	Thr767	Thr765	Thr767
Phe770	Phe762	Tyr768	Phe766	Tyr768

Gln773	Gln765	Gln771	Gln769	Gln771
Gly774	Gly766	Gly772	Gly770	Gly772
Phe777	Phe769	Phe775	Phe773	Phe775
Gly778	Gly770	Gly776	Gly774	Gly776
Gly782	Glu774	Glu780	Glu778	Glu780
Ala823	Ala815	Ser821	Ala819	Ser821
Gln824	Gln816	Asn822	Gln820	Ser822
Val825	Val817	Val823	Val821	Val823
Lys826	Lys818	Lys824	Lys822	Lys824
Gly827	Gly819	Gly825	Gly823	Gly825
Gly989	Gly981	Gly989	Gly985	Gly987
Gln990	Gln982	Asn990	Gln986	Asn988
Ser992	Ser984	Ser992	Ser988	Ser990
Ser993	Ser985	Ser993	Ser989	Ser991
Phe994	Phe986	Phe994	Phe990	Phe992
Ala995	Ala987	Ala995	Ala991	Ala993
Pro996	Pro988	Pro996	Pro992	Pro994
Asp997	Asp989	Asp997	Asp993	Asp995

Supplementary Table S2. Residue interaction energy for all residues which are involved in interactions with the top scored docking pose of verapamil. For example, residue Leu65 in human corresponds to Leu64 (rat MDR1a), Leu63 (rat MDR1b), Leu64 (mouse mdr1a) and Leu64 (mouse mdr1b).

Residue_human MDR1	RIE_human MDR1 (kcal/mol)	Residue_rat MDR1a	RIE_Rat MDR1a (kcal/mol)	Residue_Rat MDR1b	RIE_rat MDR1b (kcal/mol)	Residue_Mouse mdr1a	RIE_Mouse mdr1a (kcal/mol)	Residue_Mouse mdr1b	RIE_Mouse mdr1b (kcal/mol)
Leu65	-0.887	Leu64	-0.005	Leu63	-0.428	Leu64	-0.075	Leu64	-0.038
		Pro65	-0.082	Pro64	-0.078			Pro65	0.023
Met68	0.09	Met67	0.172	Leu66	-0.118	Met67	0.15	Leu67	0.102
Met69	-2.19	Met68	0.077	Met67	0.878	Met68	-1.015	Met68	-0.074
Phe72	0.091			Phe70	-0.305	Phe71	0.032		
		Thr191	-0.208	Thr198	-0.134	Thr195	0.028	Thr198	0.086
Phe303	-1.897	Phe295	-3.378	Tyr302	-1.925	Phe299	-5.19	Tyr302	-2.503
Ile306	-1.565	Ile298	-1.304	Val305	-0.254	Ile302	-0.016	Val305	-1.015
Tyr307	-3.627	Tyr299	-5.1	Tyr306	-3.714	Tyr303	-6.96	Tyr306	-3.229
Ser309	-0.081	Ser301	0.975			Ser305	1.059		
Tyr310	-3.898	Tyr302	-3.654	Tyr309	-3.944	Tyr306	-4.638	Tyr309	-5.312
Phe314	0.263	Phe306	0.484	Phe313	0.455	Phe310	0.496	Phe313	0.272
Leu332	-0.821	Leu324	-0.365	Leu331	-0.381	Leu328	-0.344		
Phe335	-0.242	Phe327	-0.196	Phe334	-0.134	Phe331	-0.415	Phe334	-1.099
Phe336	-3.931	Phe328	-0.421	Phe335	-2.944	Phe332	-3.616	Phe335	-3.306
Ser337	-0.151	Ser329	-0.011	Ser336	-0.041	Ser333	-0.056	Ser336	-0.444
Val338	0.22	Val330	0.084	Ile337	-0.071	Val334	-0.056	Ile337	-0.22
leu339	0.014	Leu331	-0.104	Leu338	-1.113	Leu335	-2.484	Leu338	0.124
Ile340	-1.133	Ile332	-0.173	Ile339	-1.177	Ile336	-2.098	Ile339	-2.817
Gly341	0.229	Gly333	0.091	Gly340	-0.049	Gly337	0.006	Gly340	-0.214
Ala342	0.696	Ala334	0.42	Thr341	-0.241	Ala338	0.419	Thr341	-0.029
Phe343	-2.493	Phe335	-2.039	Phe342	-4.102	Phe339	-1.228	Phe342	-3.577
Ser344	0.593	Ser336	0.294	Ser343	0.313	Ser340	0.314	Ser343	0.288
Gln347	-0.753	Gln339	0.044	His346	0.255	Gln343	-0.144	His346	0.375
Gln725	0.033	Gln717	-9.559	Gln723	-0.732	Gln721	-9.697	Gln723	-4.055
Phe728	-1.868	Phe720	-3.56	Phe726	-3.263	Phe724	-3.676	Phe726	-2.895

Residue_human MDR1	RIE_human MDR1	Residue_rat MDR1a	RIE_Rat MDR1a	Residue_Rat MDR1b	RIE_rat MDR1b	Residue_Mouse mdr1a	RIE_Mouse mdr1a	Residue_Mouse mdr1b	RIE_Mouse mdr1b
Ala729	-0.325	Ser721	-0.384	Ala727	-0.169	Ser725	-0.376	Ala727	-0.194
Phe732	-0.27	Phe724	0.523	Phe730	0.288	Phe728	0.247	Phe730	0.059
Asn842	-0.574			Asn840	0.183	Asn838	-0.817	Asn840	-0.439
Val865	0.073			Ile865	0.064	Val861	0.063	Ile863	-0.061
Ile868	0.109			Ile868	0.112	Ile864	0.099	Ile866	0
Gln946	-0.193			Gln946	-0.265				
Met949	-0.389	Met941	-0.218	Ile949	-0.171	Met945	-0.297	Met947	-0.495
Tyr950	-0.118					Tyr946	-0.268		
						Ser948	-0.076	Ser950	0.185
Tyr953	-2.413	Tyr945	-0.326	Tyr953	-2.227	Tyr949	-0.067	Tyr951	-0.593
Phe957	0.03			Phe957	-0.024	Phe953	0.05	Phe955	0.011
Leu975	-0.645							Met973	-0.338
Phe978	-0.481	Phe970	-0.195	Phe978	-0.666	Phe974	-0.302	Phe976	-0.493
Ser979	-1.872	Ser971	-0.71	Ser979	-1.116	Ser975	-0.676	Ser977	-0.497
Ala980	-0.325			Ala980	-0.546	Ala976	-0.592	Ala978	-0.372
Val981	0.05			Val981	-0.245	Ile977	-0.356	Val979	-0.267
Val982	0.162	Val974	-0.031	Val982	-0.414	Val978	-0.462	Val980	-0.676
Phe983	-6.517	Phe975	-2.537	Phe983	-4.217	Phe979	-7.235	Phe981	-3.126
Gly984	-0.251	Gly976	-0.385	Gly984	-0.539	Gly980	-0.896	Gly982	-0.449
Ala985	0.398	Ala977	0.016	Ala985	0.084	Ala981	-0.102	Ala983	-0.044
Met986	-2.801	Met978	-1.52	Met986	-4.887	Met982	-2.077	Met984	-3.665
Ala987	-0.081	Ala979	-1.188	Ala987	0.096	Ala983	-2.297	Ala985	-1.099
Val988	0.194	Val980	-0.175	Ala988	0.131	Val984	-0.333	Ala986	-0.093
Gly989	0.425	Gly981	0.155	Gly989	0.619	Gly985	0.521	Gly987	0.413
Gln990	-5.492	Gln982	-5.564	Asn990	-0.651	Gln986	-3.109	Asn988	-3.175
Val991	0.125			Thr991	0.193	Val987	-0.588	Thr989	0.102
				Ser993	0.216	Ser989	0.503	Ser991	0.284

Supplementary Table S3. Residue interaction energy for all residues which are involved in interactions with the top scored docking pose of quinidine. For example, residue Leu65 in human corresponds to Leu64 (rat MDR1a), Leu63 (rat MDR1b), Leu64 (mouse mdr1a) and Leu64 (mouse mdr1b).

195

	(kcal/mol)		(kcal/mol)		(kcal/mol)		(kcal/mol)		(kcal/mol)
Leu65	-0.003	Leu64	-0.015	Leu63	0.05	Leu64	-0.041	Leu64	-0.013
		Pro65	-0.048	Pro64	0.002			Pro65	-0.048
Met68	0.109	Met67	0.175	Leu66	0.376	Met67	0.124	Leu67	0.112
Met69	-0.266	Met68	-0.104	Met67	0.286	Met68	-0.264	Met68	-0.039
Phe72	0.014	Thr191	-0.183	Phe70	0.117	Phe71	0.018	Thr198	0.019
		Gly195	0.106	Thr198	-0.231	Thr195	0.023		
		Ile210	-0.14						
		Leu217	0.096						
		Ale294	-0.497						
Phe303	-0.716	Phe295	-1.46	Tyr302	-0.972	Phe299	-1.993	Tyr302	-0.773
Ile306	-0.403	Ile298	-1.148	Val305	-0.915	Ile302	-0.254	Val305	-1.548
Tyr307	3.557	Tyr299	-4.056	Tyr306	-1.176	Tyr303	-6.213	Tyr306	-3.815
Ser309	0.553	Ser301	0.806			Ser305	0.528		
Tyr310	2.25	Tyr302	-5.209	Tyr309	-5.08	Tyr306	-5.899	Tyr309	-4.699
Phe314	0.139	Phe306	0.328	Phe313	0.158	Phe310	0.292	Phe313	0.418
Leu332	-0.332	Leu324	-0.33	Leu331	-0.591	Leu328	-0.35		
		Thr325	-0.264						
		Val326	-0.2						
Phe335	-0.537	Phe327	-0.555	Phe334	-1.092	Phe331	-0.499	Phe334	-0.58
Phe336	-2.41	Phe328	-2.059	Phe335	-3.339	Phe332	-2.368	Phe335	-1.494
Ser337	-0.065	Ser329	-0.111	Ser336	-0.368	Ser333	-0.012	Ser336	-0.058
Val338	-0.005	Val330	-0.09	Ile337	-0.171	Val334	0.031	Ile337	0.001
leu339	-0.595	Leu331	-2.698	Leu338	-1.305	Leu335	-0.275	Leu338	-1.143
Ile340	-0.686	Ile332	-1.436	Ile339	-1.8	Ile336	-0.472	Ile339	-1.282
Gly341	0.06	Gly333	-0.039	Gly340	-0.101	Gly337	0.038	Gly340	0.009
Ala342	0.296	Ala334	0.251	Thr341	0.007	Ala338	0.325	Thr341	-0.133
Phe343	-0.289	Phe335	-3.234	Phe342	-3.302	Phe339	-0.187	Phe342	-1.806
Ser344	0.303	Ser336	0.153	Ser343	0.63	Ser340	0.277	Ser343	0.116
		Val337	0.239						
		Gly338	0.484						
Gln347	-0.485	Gln339	0.201	His346	0.438	Gln343	-0.089	His346	0.005
				Ile722	-0.224				

Gln725	-11.577	Gln717	-0.976	Gln723	-4.79	Gln721	-11.095	Gln723	-4.616
Phe728	-2.479	Phe720	-2.062	Phe726	-2.064	Phe724	-4.597	Phe726	-4.058
Ala729	0.111	Ser721	-0.373	Ala727	-0.098	Ser725	-0.497	Ala727	-0.125
Phe732	-0.721	Phe724	0.282	Phe730	-0.091	Phe728	-1.042	Phe730	0.576
		Phe751	-0.588						
Asn842	-1.129			Asn840	-0.016	Asn838	-0.167	Asn840	-0.249
						Leu839	0.102		
Val865	0.077			Ile865	-0.004	Val861	0.095	Ile863	0.073
Ile868	0.092			Ile868	0.13	Ile864	0.114	Ile866	0.095
								val867	0.048
Gln946	-0.308			Gln946	-0.346				
Met949	-0.126	Met941	-0.105	Ile949	0.037	Met945	-0.259	Met947	-0.153
Tyr950	-0.086					Tyr946	-0.253		
						Ser948	-0.104	Ser950	0.094
Tyr953	-0.045	Tyr945	-0.219	Tyr953	0.034	Tyr949	-0.124	Tyr951	-0.25
Phe957	0.031			Phe957	0.112	Phe953	0.032	Phe955	0.028
Leu975	-0.485							Met973	-0.306
				Met975	-0.308				
Phe978	-0.505	Phe970	-0.383	Phe978	-0.527	Phe974	-0.416	Phe976	-0.351
Ser979	-1.535	Ser971	-1.056	Ser979	-0.553	Ser975	-1.18	Ser977	-0.633
Ala980	-0.595			Ala980	-0.4	Ala976	-0.828	Ala978	-0.692
Val981	-0.381			Val981	-0.166	Ile977	-0.478	Val979	-0.355
Val982	-0.596	Val974	-0.37	Val982	-0.315	Val978	-0.595	Val980	-0.169
Phe983	-8.114	Phe975	-4.85	Phe983	-5.064	Phe979	-7.466	Phe981	-4.837
Gly984	-0.952	Gly976	-0.804	Gly984	-0.405	Gly980	-0.987	Gly982	-0.534
Ala985	-0.013	Ala977	-0.173	Ala985	0.246	Ala981	-0.089	Ala983	-0.051
Met986	-2.162	Met978	-0.778	Met986	-4.61	Met982	-1.593	Met984	-1.336
Ala987	-0.046	Ala979	0.073	Ala987	-0.505	Ala983	0.184	Ala985	-0.119
Val988	-0.059	Val980	-0.216	Ala988	0.204	Val984	-0.083	Ala986	-0.057
Gly989	0.896	Gly981	0.533	Gly989	0.872	Gly985	0.712	Gly987	0.495
Gln990	1.336	Gln982	-3.983	Asn990	2.143	Gln986	-2.831	Asn988	2.71
Val991	0.424			Thr991	-0.049	Val987	0.448	Thr989	0.259
				Ser993	0.19	Ser989	0.668	Ser991	0.623

Supplementary Table S4. Percentage (%) of interacting ligands/inhibitors to commonly interacting residues in the three species. Residue number in the order human MDR1, rat MDR1a, rat MDR1b, mouse mdr1a and mouse mdr1b. For example, residue Leu65 in human corresponds to Leu64 (rat MDR1a), Leu63 (rat MDR1b), Leu64 (mouse mdr1a) and Leu64 (mouse mdr1b).

Residue number_human MDR1, Residue number_Rat MDR1a, Residue number_rat MDR1b, Residue number_mouse mdr1a, Residue number_mouse mdr1b	% ligand Interaction_human MDR1	% ligand Interaction_rat MDR1a	% ligand Interaction_rat MDR1b	% ligand Interaction_mouse mdr1a	% ligand Interaction_mouse mdr1b
Leu65, Leu64, Leu63, Leu64, Leu64	28.115	33.706	27.955	43.291	24.760
Met68, Met67, Leu66, NCRI, Leu67	7.188	3.834	3.195	0.000	7.188
Met69, Met68, Met67, Met68, Met68	50.799	77.955	61.022	56.230	67.732
Phe72, Phe71, Phe70, Phe71, Phe71	19.649	42.013	9.904	8.626	35.144
Leu225, Leu217, Ile224, Leu221, Ile224	6.390	4.153	10.383	2.077	7.508
Ala229, Ala221, Ile298,NCRI, NCRI	2.077	1.757	20.767	0.000	0.000
Ile299, Met291,NCRI,NCRI, Ile298	22.524	3.514	0.000	0.000	7.188
Ala302, Ala294, Ala301,NCRI, Ala301	19.649	12.780	33.706	0.000	16.613
Phe303, Phe295, Tyr302,Phe299, Tyr302	52.236	38.658	63.099	24.760	37.859
Ile306, Ile298, Val305, Ile302, Val305	77.955	71.565	80.192	67.252	63.578
Tyr307, Tyr299, Tyr306, Tyr303, Tyr306	65.016	63.099	72.204	76.038	61.821
Tyr310, Tyr302, Tyr309, Tyr306, Tyr309	96.006	89.297	97.604	96.486	94.728
Phe336, Phe328, Phe335, Phe332, Phe335	78.594	83.546	68.530	71.725	72.364
Leu339, Leu331, Leu338, Leu335, Leu338	64.058	74.760	43.131	51.278	62.939
Ile340, Ile332, Leu339, Ile336, Leu339	74.760	71.406	68.211	84.665	86.422
Phe343, Phe335, Phe342, Phe339, Phe342	68.530	66.134	86.102	69.649	66.134
Phe728, Phe720, Phe726, Phe724, Phe726	71.406	66.454	79.393	76.358	72.364
Phe732, Phe724, Phe730, Phe728, Phe730	4.313	7.188	2.236	3.195	4.633
Met949, Met941, Ile949, Met945, Met947	12.460	11.661	12.780	14.377	13.898
Tyr950, Tyr942, Tyr950, Tyr946, Tyr948	19.329	3.674	5.911	15.974	4.313
Tyr953, Tyr945, Tyr953,NCRI, Tyr951	58.147	65.655	33.706	0.000	67.732
Phe957, Phe949, Phe953, Phe957, Phe955	2.716	26.358	6.869	4.473	13.898
NCRI, Leu967, Met975,NCRI, Met973	0.000	16.773	5.272	0.000	23.802
Phe978, Phe970, Phe974, Phe978, Phe976	14.537	42.332	8.946	7.987	21.246
Phe983, Phe975, Phe979, Phe983, Phe981	88.978	90.256	69.010	86.102	89.297
Met986, Met978, Met982, Met986, Met984	41.214	33.706	44.249	39.297	36.741

| Ala987, Ala979, Ala987, Ala983, Ala985 | 2.077 | 5.112 | 2.236 | 10.224 | 6.070 |
| Phe994,NCRI, Phe994,NCRI ,NCRI | 15.176 | 0.000 | 2.396 | 0.000 | 0.000 |

Supplementary Table S5. Key residues obtained by the heat map analysis showing prominent interactions with functional groups (ether, carbonyl group, alkyl carbon, nitrogen and arene) in the five transporter models. For example, residue Tyr310 in human corresponds to Tyr302 (rat MDR1a), Tyr309 (rat MDR1b), Tyr306 (mouse mdr1a) and Tyr309 (mouse mdr1b).

Human MDR1	Rat MDR1a	Rat MDR1b	Mouse mdr1a	Mouse mdr1b
Tyr310	Tyr302	Tyr309	Tyr306	Tyr309
Phe343	Phe335	Phe342	Phe339	Phe342
Phe983	Phe975	Phe983	Phe979	Phe981
Met986	Met978	Met986	Met982	Met984

Supplementary Figures

Supplementary Fig. S1. Sequence alignment of human MDR1 with corrected mouse P-glycoprotein structure (PDB ID: 4M1M). The residues are colored according to the ClustalX scheme using Jalview.
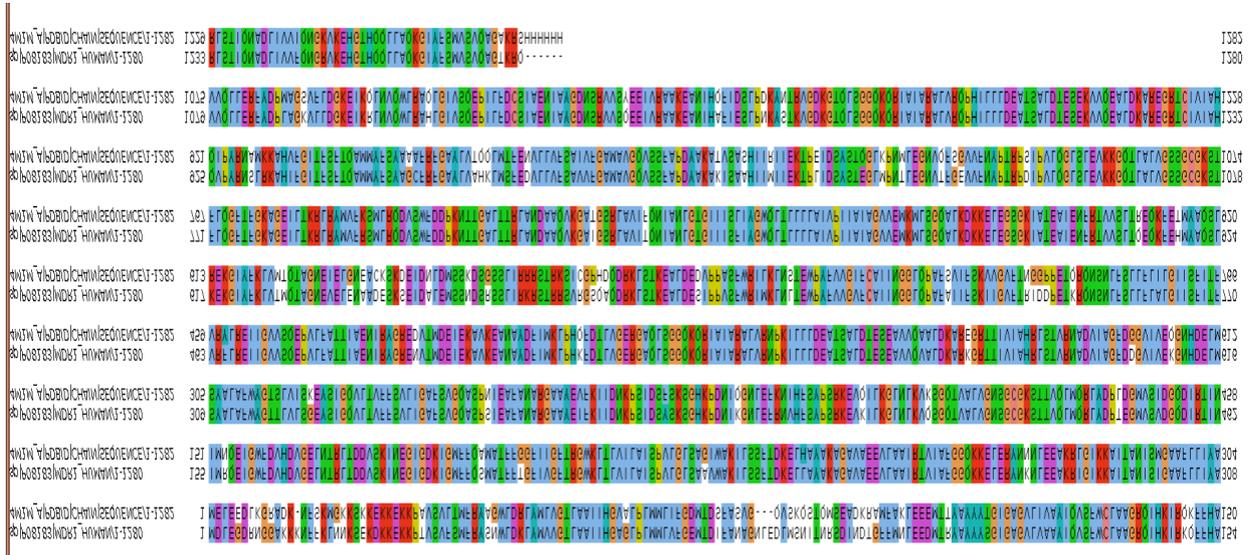
Supplementary Fig. S2. Sequence alignment of rat MDR1a with corrected mouse P-glycoprotein structure (PDB ID: 4M1M). The residues are colored according to the ClustalX scheme using Jalview.
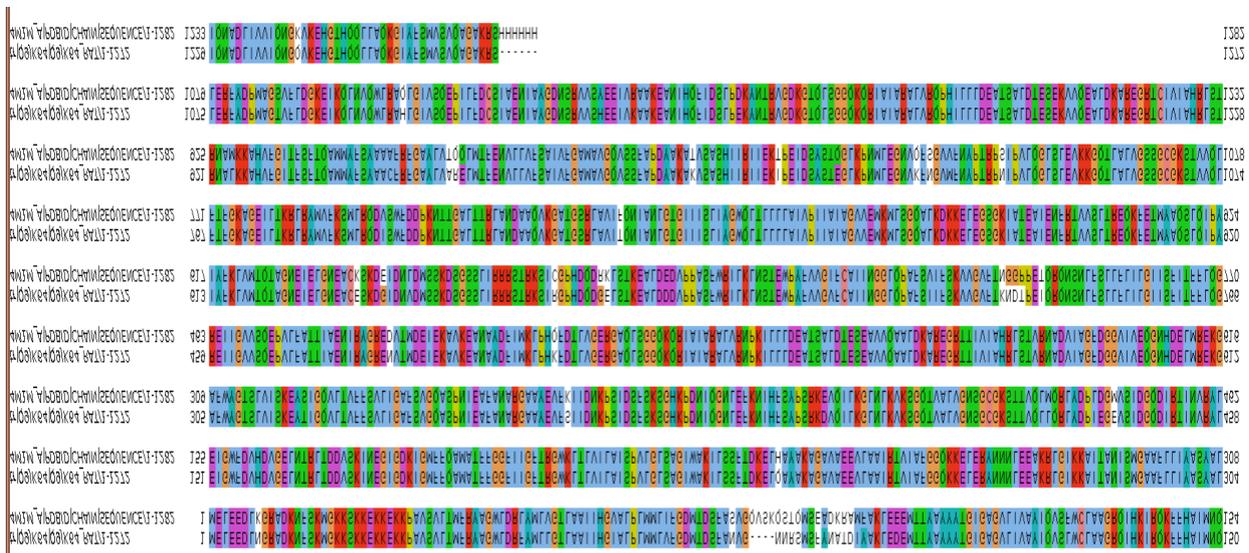
Supplementary Fig. S3. Sequence alignment of rat MDR1b with corrected mouse P-glycoprotein structure (PDB ID: 4M1M). The residues are colored according to the ClustalX scheme using Jalview.
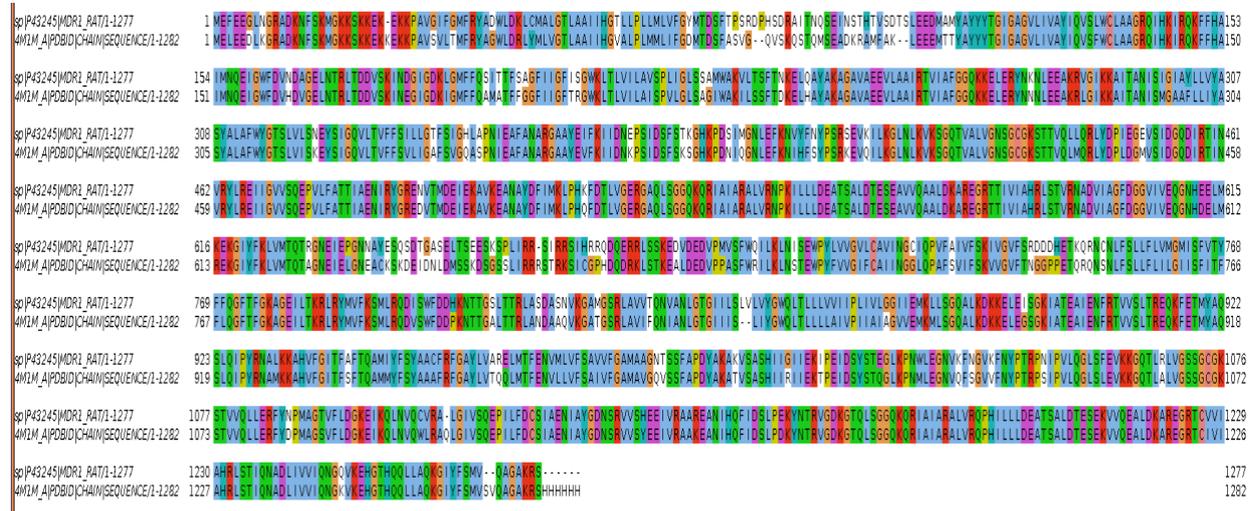
Supplementary Fig. S4. Sequence alignment of mouse mdr1b with corrected mouse P-glycoprotein structure (PDB ID: 4M1M). The residues are colored according to the ClustalX scheme using Jalview.
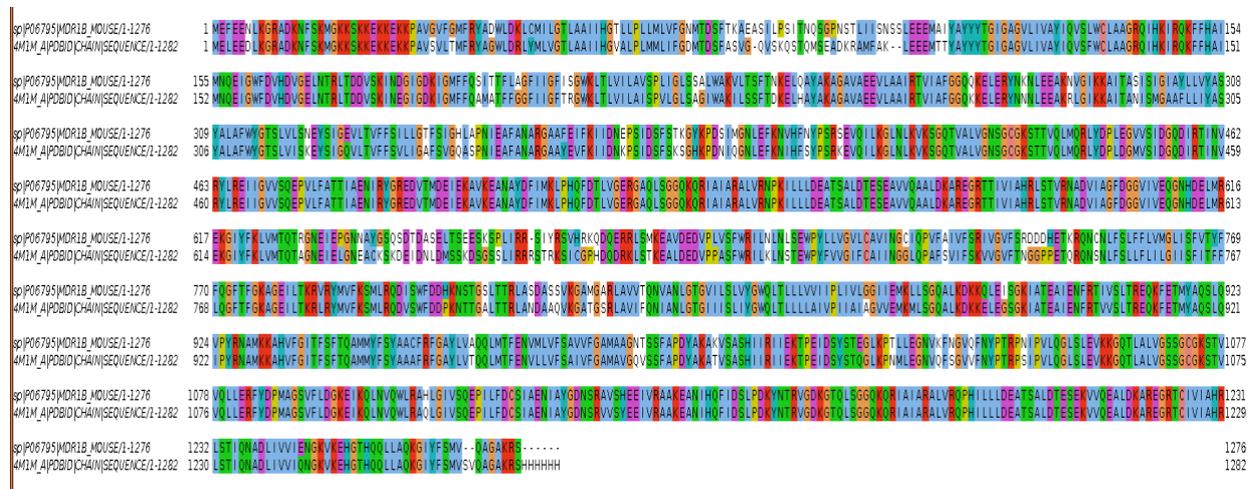
Supplementary Fig. S5. Sequence alignment of human MDR1, rat MDR1a, rat MDR1b, mouse mdr1a and mouse mdr1b. TMD's are indicated in boxes.
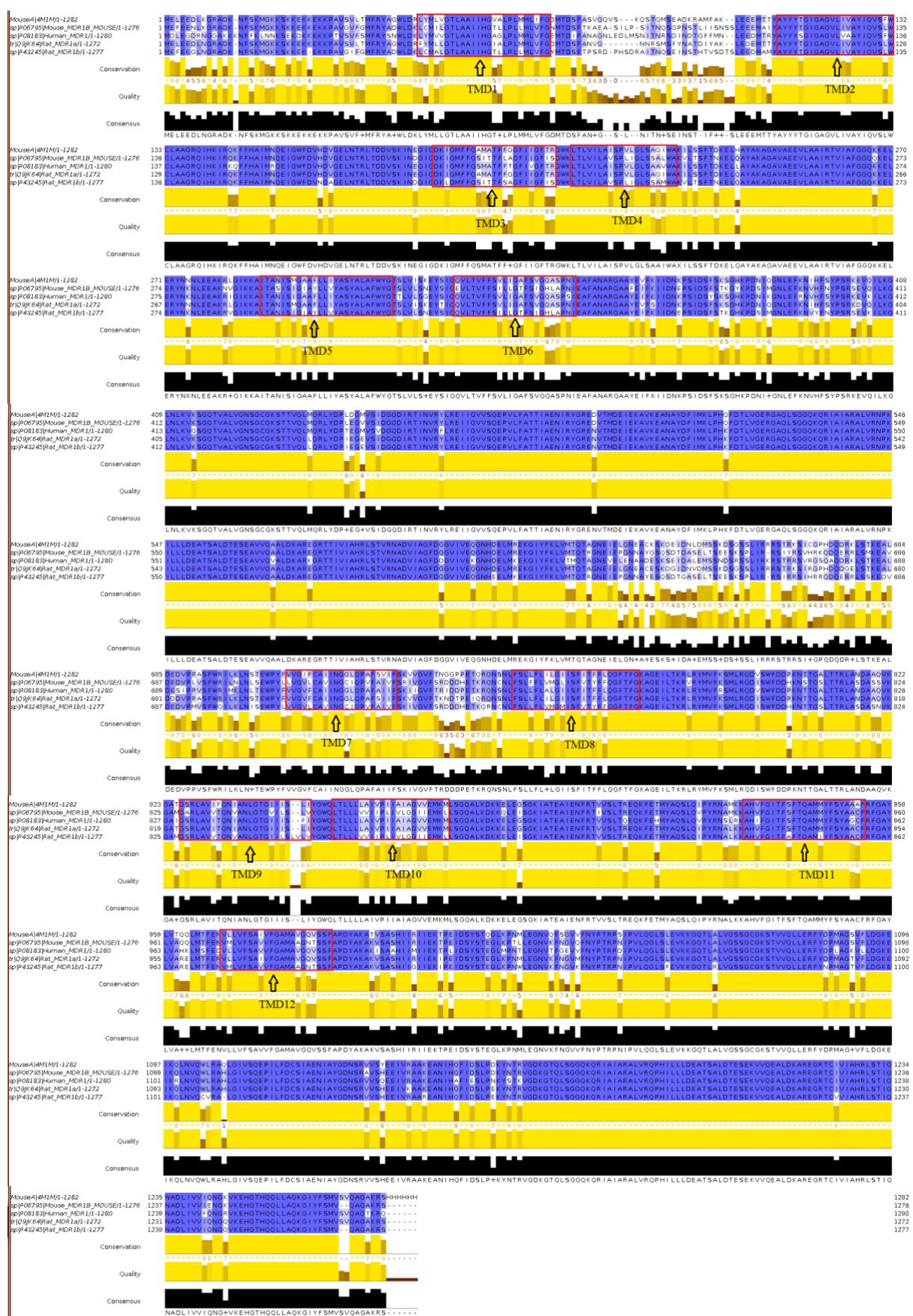
Supplementary Fig. S6. Ramachandran plot for the final homology model of human MDR1 taken from PDBsum.
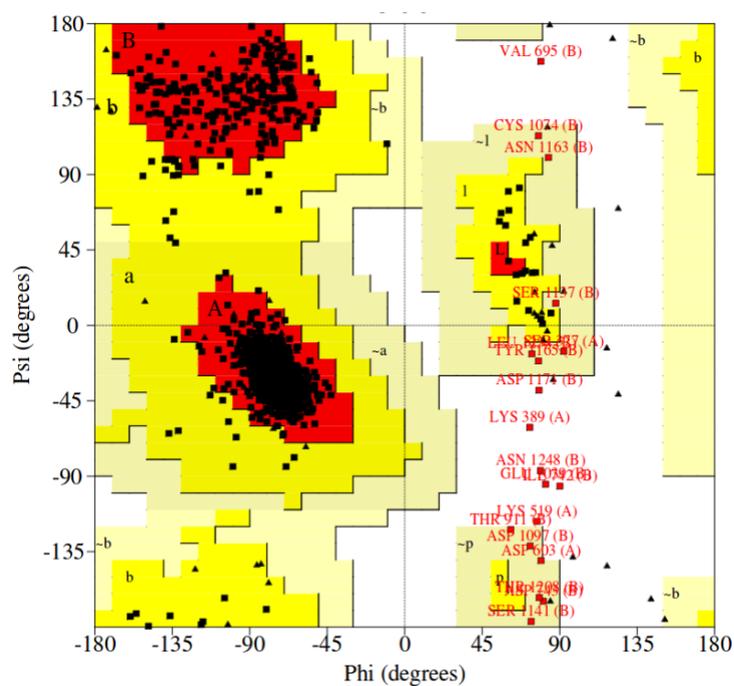
Supplementary Fig. S7. Ramachandran plot for the final homology model of rat MDR1a taken from PDBsum.
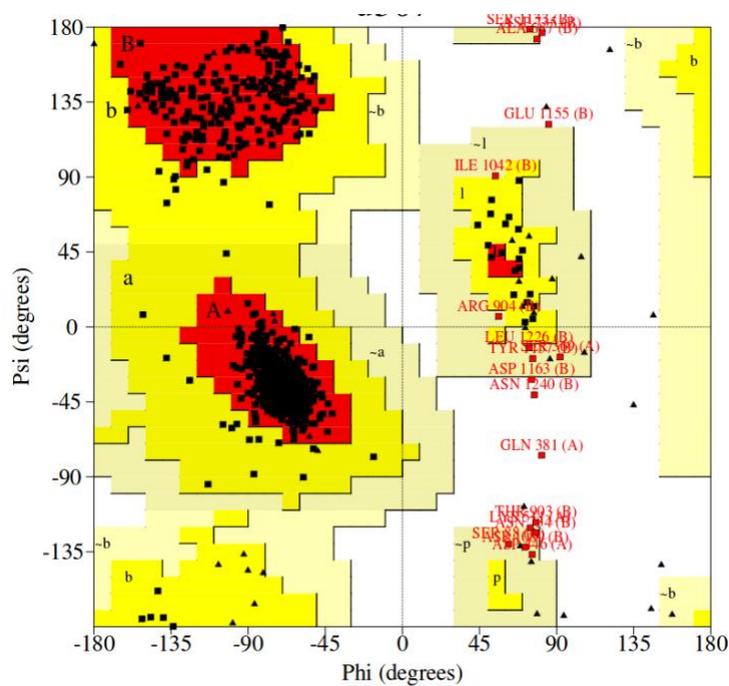
Supplementary Fig. S8. Ramachandran plot for the final homology model of rat MDR1b taken from PDBsum.
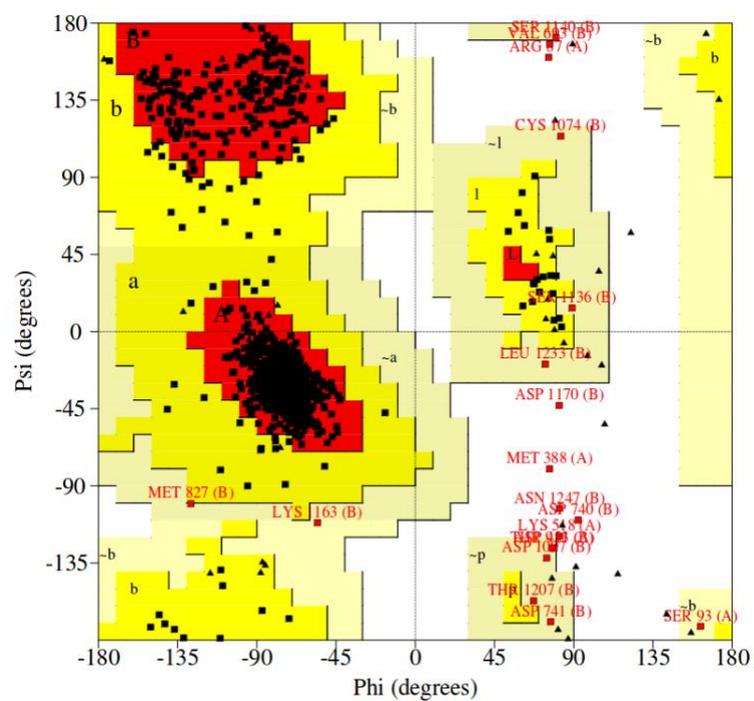
Supplementary Fig. S9. Ramachandran plot for the final homology model of mouse mdr1b taken from PDBsum.
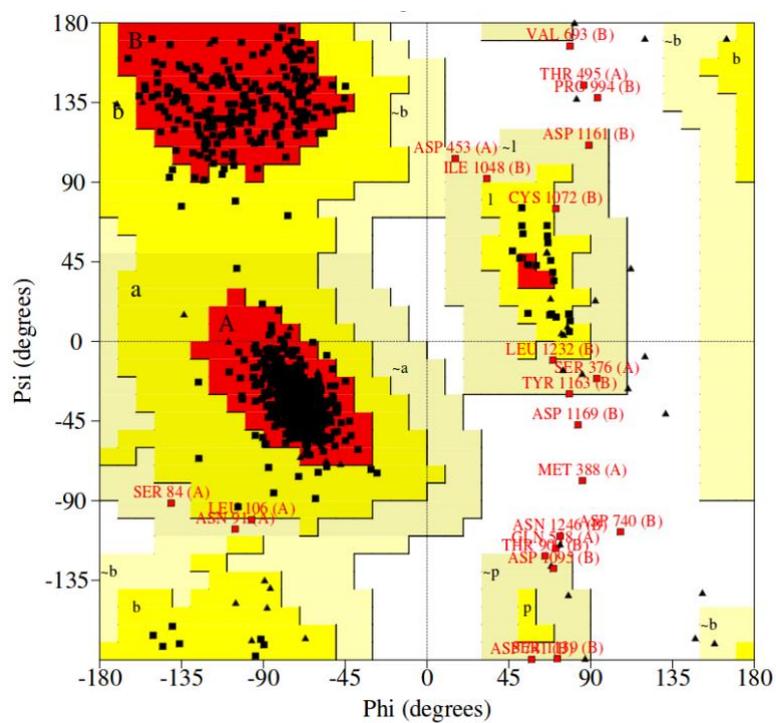
Supplementary Fig. S10. Residues that are present in the disallowed region in the final human MDR1 homology model
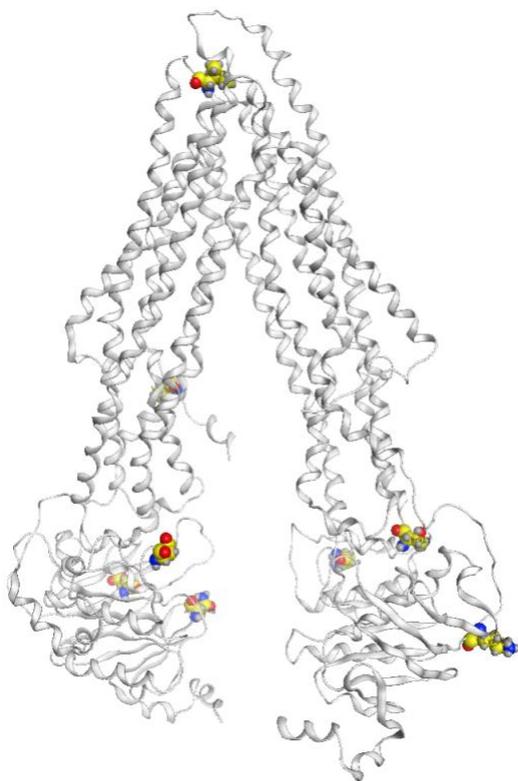
Supplementary Fig. S11. Residues that are present in the disallowed region in the final rat MDR1a homology model
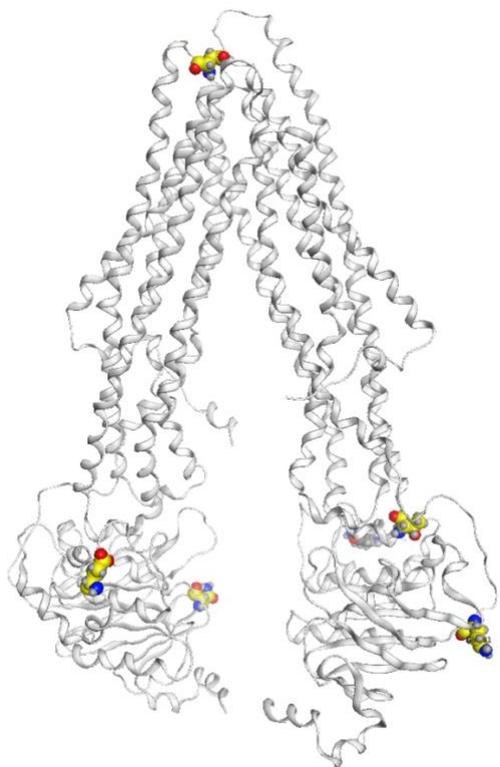
Supplementary Fig. S12. Residues that are present in the disallowed region in the final rat MDR1b homology model
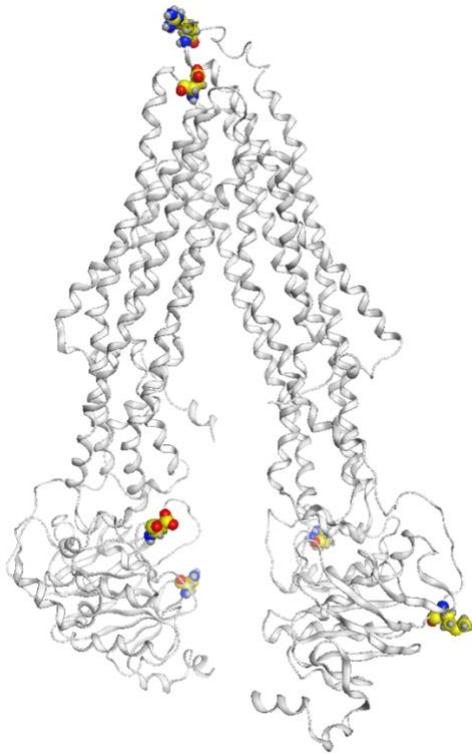
Supplementary Fig. S13. Residues that are present in the disallowed region in the final mouse mdr1b homology model
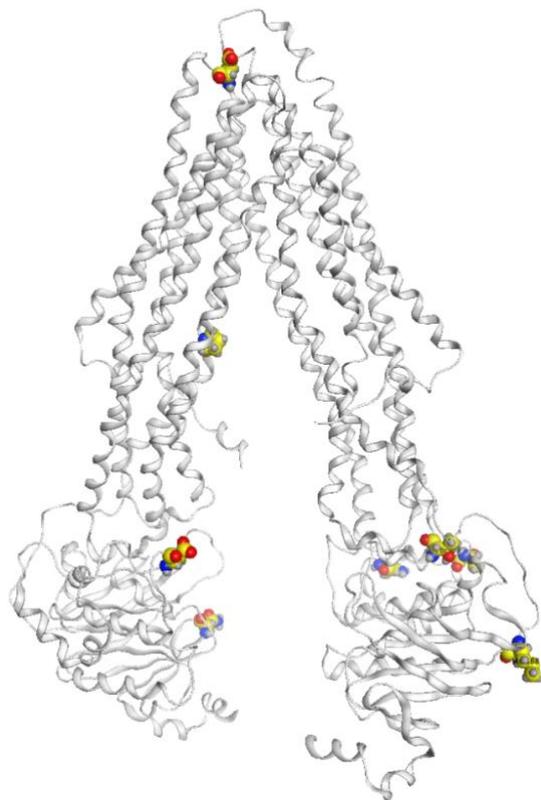
Supplementary Fig. S14. Electrostatic potential surface (EPS) of the central binding cavity of human MDR1. Binding site surface detected using Site Finder from MOE 2013. EPS generated at the dummy atoms in the central binding cavity.
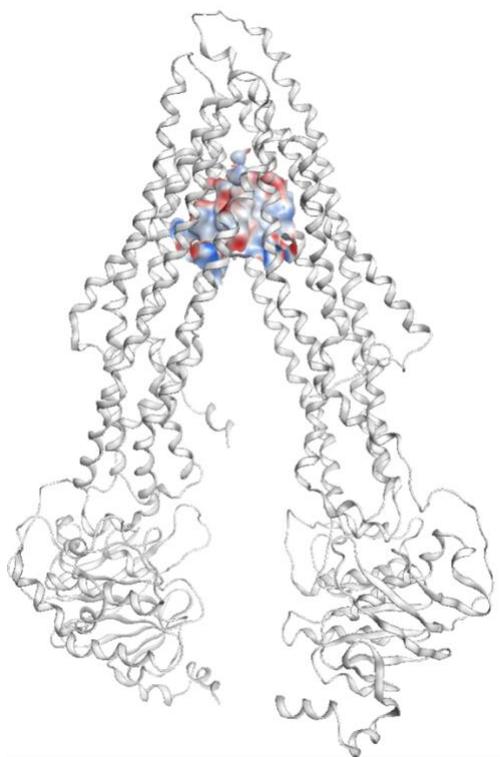
Supplementary Fig. S15. Electrostatic potential surface (EPS) of the central binding cavity of rat MDR1a. Binding site surface detected using Site Finder from MOE 2013. EPS generated at the dummy atoms in the central binding cavity.
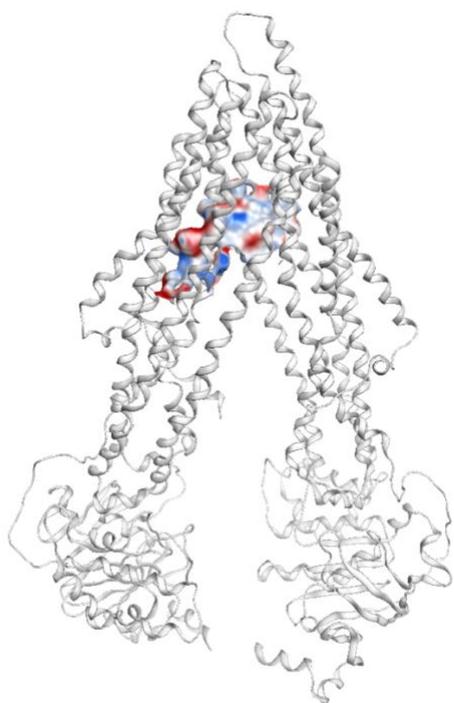
Supplementary Fig. S16. Electrostatic potential surface (EPS) of the central binding cavity of rat MDR1b. Binding site surface detected using Site Finder from MOE 2013. EPS generated at the dummy atoms in the central binding cavity.
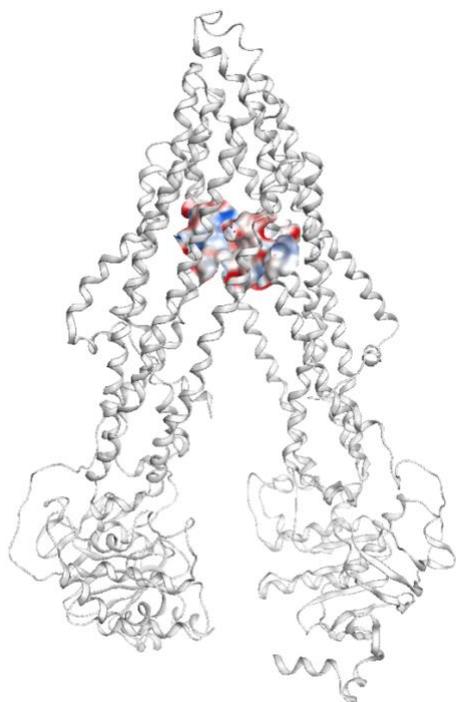
Supplementary Fig. S17. Electrostatic potential surface (EPS) of the central binding cavity of mouse mdr1a. Binding site surface detected using Site Finder from MOE 2013. EPS generated at the dummy atoms in the central binding cavity.
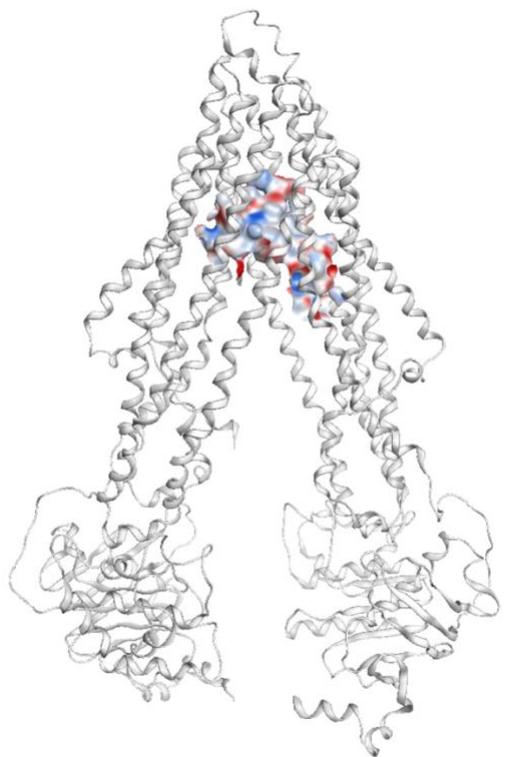
Supplementary Fig. S18. Electrostatic potential surface (EPS) of the central binding cavity of mouse mdr1b. Binding site surface detected using Site Finder from MOE 2013. EPS generated at the dummy atoms in the central binding cavity.
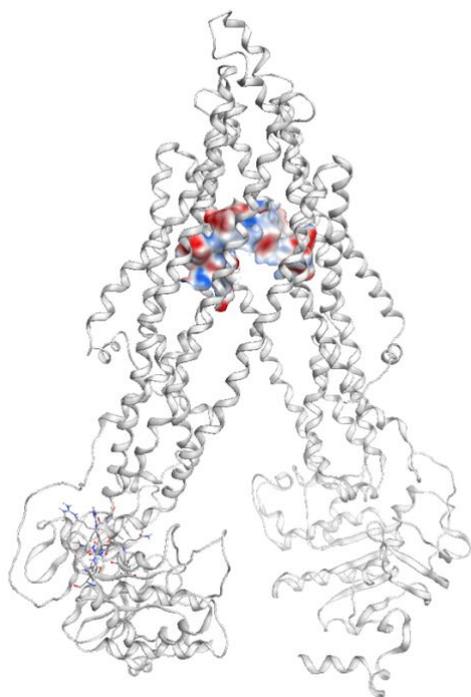
Supplementary Fig. S19 . Residue interaction energy for common interaction residues in human MDR1, rat MDR1a, rat MDR1b, mouse mdr1a and mouse mdr1b. x-axis denotes residue number in the order human MDR1, rat MDR1a, rat MDR1b, mouse mdr1a and mouse mdr1b, y-axis denotes the corresponding residue interaction energy (kcal/mol).
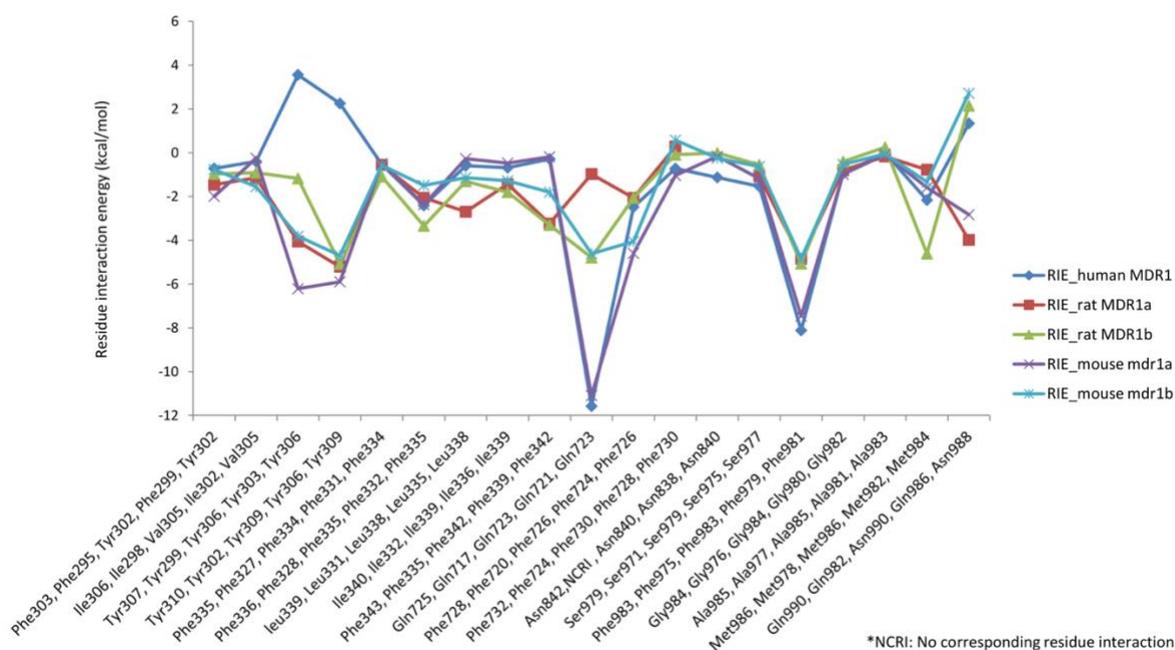
Supplementary Fig. S20. Binding poses for Quinidine.
Residues: Human MDR1 (grey), Rat MDR1a (orange), Rat MDR1b (blue purple), Mouse mdr1a (maroon), Mouse mdr1b (turquoise).
Quinidine: Green (Human MDR1), Yellow (Rat MDR1a), Pink (Rat MDR1b), Red (Mouse mdr1a), Blue (Mouse mdr1b)
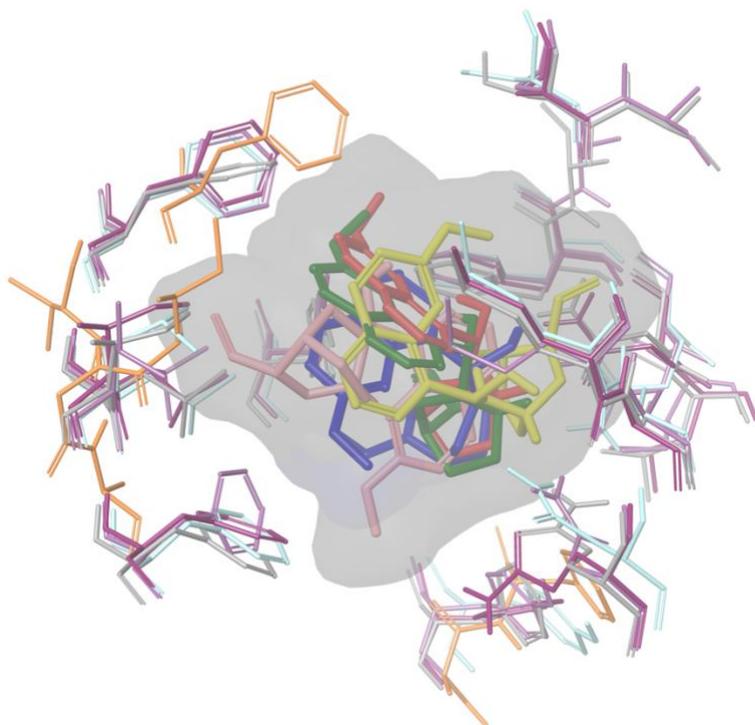
Supplementary Fig. S21. Hydrophobic interactions common in human MDR1, rat MDR1a, rat MDR1b, mouse mdr1a and mouse mdr1b for quinidine. x-axis denotes residue number in the order human MDR1, rat MDR1a, rat MDR1b, mouse mdr1a and mouse mdr1b, y-axis denotes frequency of interacting residues (%).
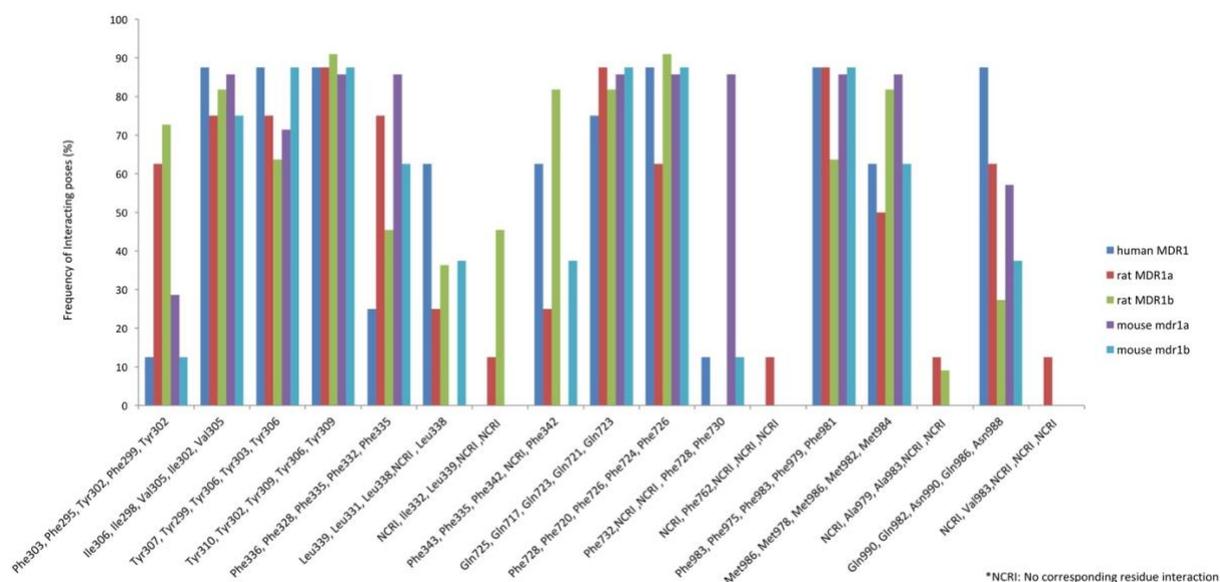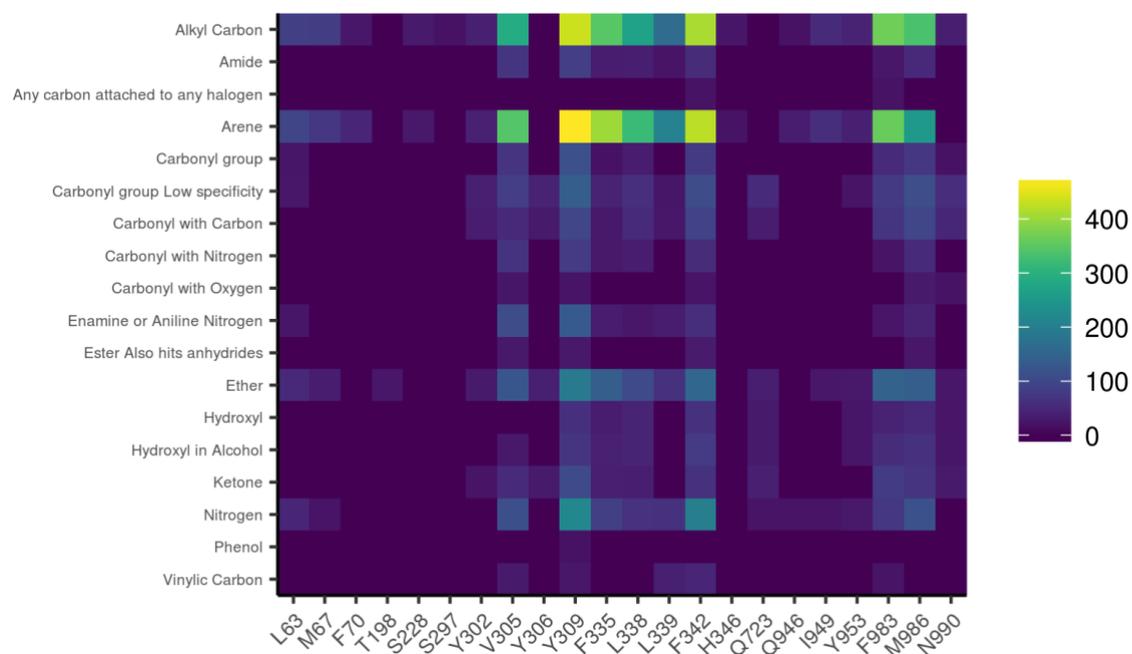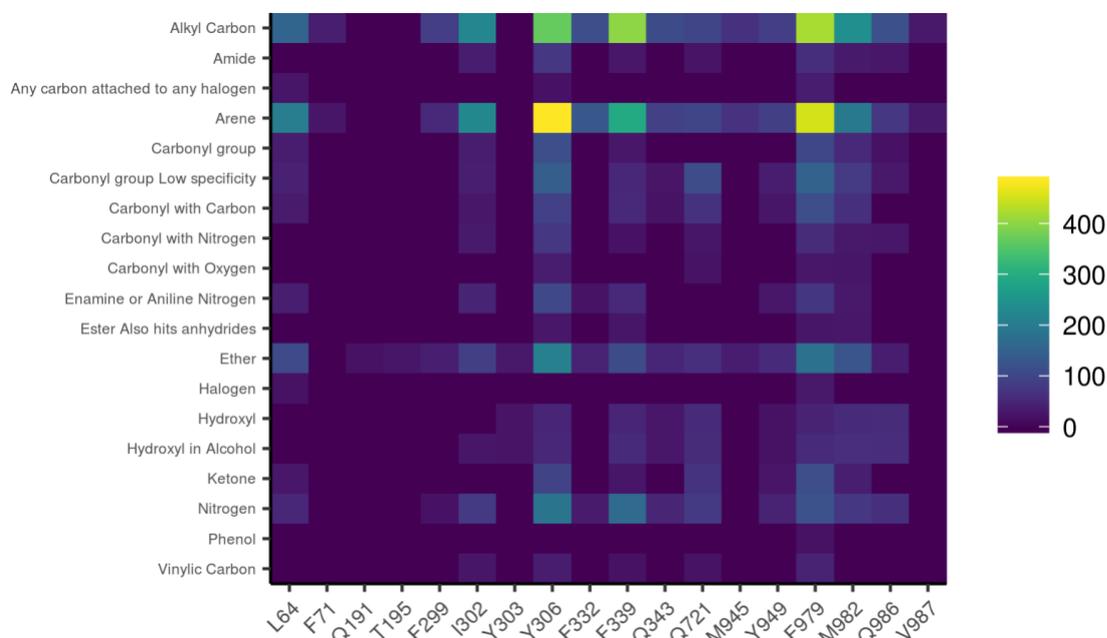
Supplementary Fig. S22. Heat map illustrating the PLIF analysis of the human P-gp inhibitors for rat MDR1a. x-axis denotes contact residues. y-axis denotes functional groups of the ligand which are showing an interaction with the residue. Color scale denotes number of interacting ligands.
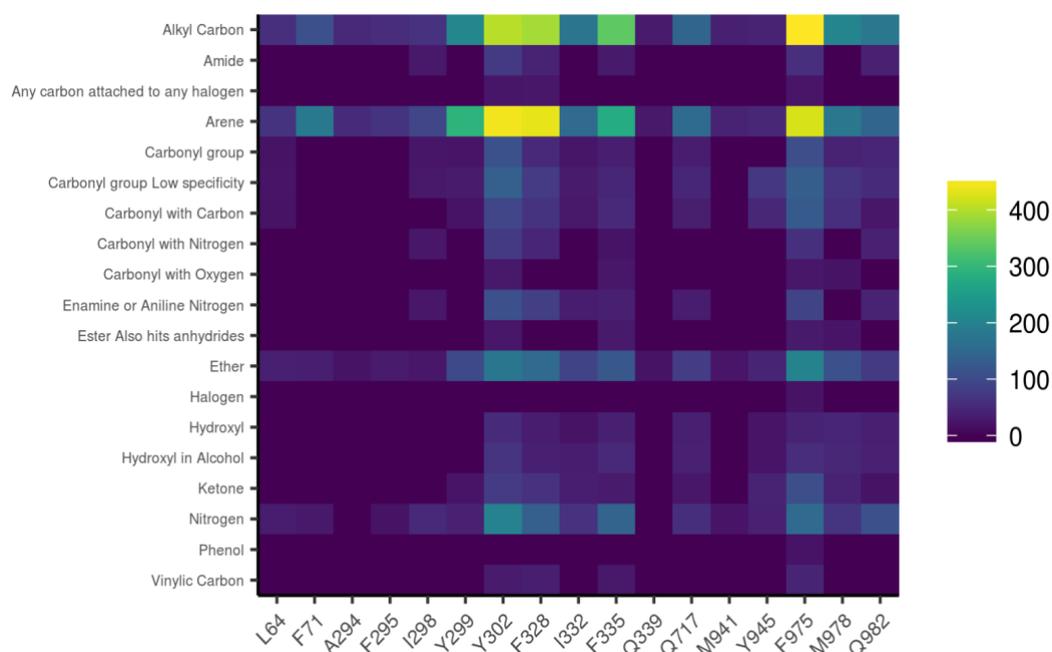
Supplementary Fig. S23. Heat map illustrating the PLIF analysis of the human P-gp inhibitors for rat MDR1b. x-axis denotes contact residues. y-axis denotes functional groups of the ligand which are showing an interaction with the residue. Color scale denotes number of interacting ligands.

Supplementary Fig. S24. Heat map illustrating the PLIF analysis of the human P-gp inhibitors for mouse mdr1a. x-axis denotes contact residues. y-axis denotes functional groups of the ligand which are showing an interaction with the residue. Color scale denotes number of interacting ligands.

Supplementary Fig. S25. Heat map illustrating the PLIF analysis of the human P-gp inhibitors for mouse mdr1b. x-axis denotes contact residues. y-axis denotes functional groups of the ligand which are showing an interaction with the residue. Color scale denotes number of interacting ligands.
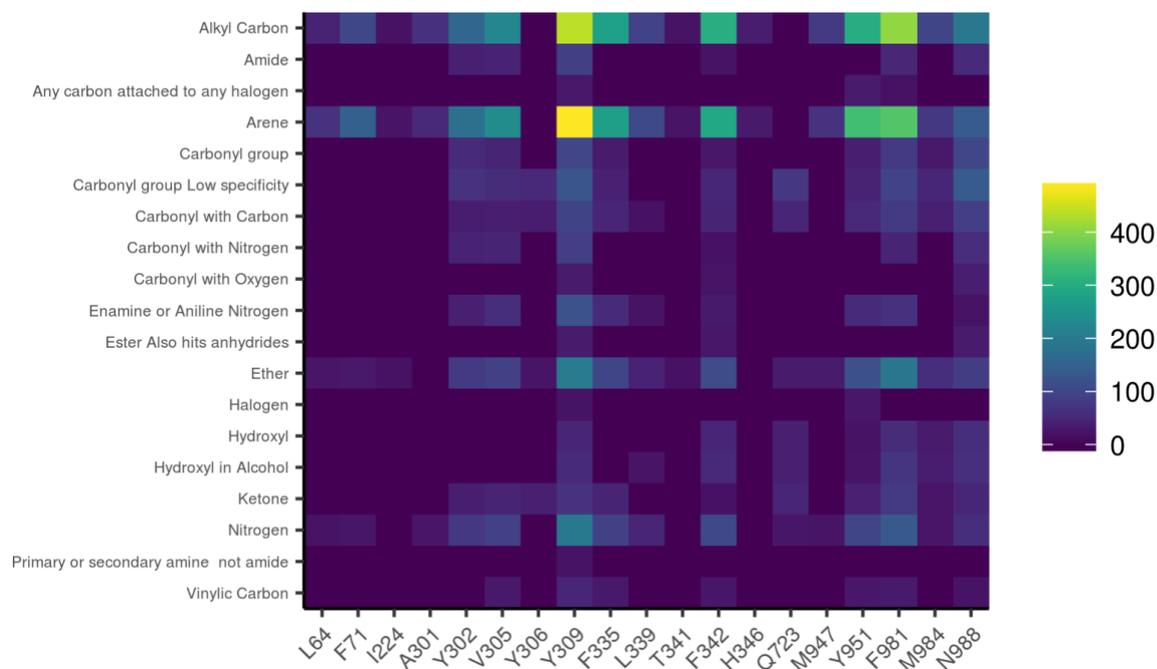
8. Publications and posters

Book Chapters

- Eleni Kotsampasakou, <u>**Sankalp Jain**</u>, Daniela Digles and Gerhard F. Ecker**, Transporter in Hepatotoxity, Computational Toxicology: Risk Assessment for Pharmaceutical and Environmental Chemicals**, 2nd edition, Sean Ekins, ISBN: 978-1-119-28256-3
- Stefanie Kickinger, Eva Hellsberg, <u>**Sankalp Jain**</u> and Gerhard F. Ecker, **Linked open data: ligand-transporter interaction profiling and beyond, Multi-Target Drug Design Using Chem-Bioinformatic Approaches**, *(Submitted on 15th March 2018)*

Publications

- <u>**Sankalp Jain**</u>, Melanie Grandits, Gerhard F. Ecker, **Interspecies comparison of ligand binding sites of the human, rat and mouse P-glycoprotein,** *European Journal of Pharmaceutical Sciences (Submitted on 21st March 2018; under peer review)*
- <u>**Sankalp Jain**</u>, Eleni Kotsampasakou, Gerhard F. Ecker**, Comparing the performance of meta-classifiers—a case study on selected imbalanced data sets relevant for prediction of liver toxicity,** *J Comput Aided Mol Des 1-8. doi: 10.1007/s10822-018-0116-z.*
- <u>**Sankalp Jain,**</u> Melanie Grandits, Lars Richter, Gerhard F. Ecker, **Structure based classification for bile salt export pump (BSEP) inhibitors using comparative structural modeling of human BSEP**, *J Comput Aided Mol Des 31:507–521. doi: 10.1007/s10822-017-0021-x.*
- Pradeep Kumar Naik, <u>**Sankalp Jain**</u>, Piyush Ranjan and Dipankar Sengupta, **TpPred: A tool for hierarchical prediction of transport proteins using cluster of neural networks and sequence derived features**, *International Journal of Computational Biology, 0003:44-58, 2012.*
- Pradeep Kumar Naik, Piyush Ranjan, Pooja Kesari and <u>**Sankalp Jain**</u>, **MetalloPred: A tool for hierarchical prediction of metal ion binding proteins using cluster of neural**

networks and sequence derived features, *Journal of Biophysical Chemistry, vol. 2 no. 2, 2011.*

- Pradeep Kumar Naik, Mani Srivastava, Prasad Bajaj, <u>**Sankalp Jain**</u>, Abhishek Dubey, Piyush Ranjan, Rishay Kumar and Harvinder Singh, **The binding modes and binding affinities of artemisinin derivatives with Plasmodium falciparum Ca2+-ATPase (PfATP6)**, *Journal of Molecular Modeling, vol. 16, no. 6, 2010.*

Selected Posters

- **Sankalp Jain**, Melanie Grandits, Gerhard F. Ecker; "Comparison of P-glycoprotein binding sites reveals a conservation of ligand binding modes in human, mouse and rat"; German Conference on Chemoinformatics , Mainz, Germany

- **Sankalp Jain**, Melanie Grandits, Gerhard F. Ecker; "Structure Based Classification for Bile Salt Export Pump (BSEP) Inhibitors by Comparative Structural Modeling of Human BSEP"; Gordon Research Conference, Multi-Drug Efflux Systems, Galveston, TX/USA

- **Sankalp Jain**, Melanie Grandits, Gerhard F. Ecker; "Interspecies comparison of ligand binding sites of the human, mouse and rat P-glycoprotein transporters"; MolTag Science Day, Vienna, AUS

- **Sankalp Jain**, Melanie Grandits, Gerhard F. Ecker; "Ligand binding site comparison of human, mouse and rat P-glycoprotein transpoeter"; 9th SFB 35 Symposium Vienna

- **Sankalp Jain**, Melanie Grandits, Gerhard F. Ecker; "Interspecies comparison of ligand binding sites of the human, mouse and rat P-glycoprotein transporters"; 21st EuroQSAR - 21st European Symposium on Quantitative Structure-Activity Relationship, Verona, Italy

- **Sankalp Jain**, Melanie Grandits, Gerhard F. Ecker; "Comparative structural modeling of human BSEP and Structure based classification for BSEP/ABCB11 Inhibitors" 8th SFB 35 Symposium, Vienna

- **Sankalp Jain**, Melanie Grandits, Gerhard F. Ecker; "Structure based classification for BSEP/ABCB11 Inhibitors using comparative structural modeling of human BSEP"; 24th Scientific Congress of the Austrian Pharmaceutical Society, Vienna

- **Sankalp Jain**, Gerhard F. Ecker; "Comparative structural modeling of human BSEP and Structure based classification for BSEP/ABCB11 Inhibitors"; 10th European Workshop in Drug Design, Certosa di Pontignano, Siena, Italy

9. List of Abbreviations

ABC-transporter: ATP-binding cassette transporter

MDR: multi-drug resistance

P-gp: P-glycoprotein

BCRP: breast cancer resistance protein

ADMET: absorption, distribution, metabolism, excretion and toxicity

BSEP: bile salt export pump

MRP: multidrug resistance-related protein

DILI: drug-induced liver injury

FDA: Food and Drugs Administration

ITC: International Transporter Consortium

QSAR: quantitative structure-activity relationship

NBD: nucleotide-binding domains

TMD: transmembrane domains

NTCP: Na^+-taurocholate cotransporting polypeptide

PFIC: progressive familial intrahepatic cholestasis

MSD: membrane-spanning domain

SAR: structure-activity relationship

SNPs: single nucleotide polymorphisms

NMR: nuclear magnetic resonance

BLAST: Basic Local Alignment Search Tool

DOPE: Discrete Optimized Protein Energy

GOLD: Genetic Optimization of Ligand Docking

HB: hydrogen bond

MD: molecular dynamics

PDB: Protein Data Bank

CSD: Cambridge Structural Database

MLR: multiple linear regression

RMSD: root mean square distance

GPU: graphical processing units

HPC: high performance computing

Cryo-EM: cryo-electron microscopy

AUC: area under the curve

CV: cross-validation

MCC: Matthews correlation coefficient

RF: random forest

ROC: receiver operating characteristics

Bibliography

1. Lodish H, Berk A, Zipursky SL, et al (2000) Overview of Membrane Transport Proteins. In: Molecular Cell Biology, 4th edition. W.H. Freeman and Company, New York

2. Chang G (2003) Multidrug resistance ABC transporters. FEBS Lett 555:102–105

3. Robey RW, Polgar O, Deeken J, et al (2007) ABCG2: determining its relevance in clinical drug resistance. Cancer Metastasis Rev 26:39–57 . doi: 10.1007/s10555-007-9042-6

4. Doyle LA, Yang W, Abruzzo LV, et al (1998) A multidrug resistance transporter from human MCF-7 breast cancer cells. Proc Natl Acad Sci U S A 95:15665–15670

5. Bunting KD (2002) ABC transporters as phenotypic markers and functional regulators of stem cells. Stem Cells Dayt Ohio 20:11–20 . doi: 10.1634/stemcells.20-3-274

6. Fletcher JI, Haber M, Henderson MJ, Norris MD (2010) ABC transporters in cancer: more than just drug efflux pumps. Nat Rev Cancer 10:147–156 . doi: 10.1038/nrc2789

7. Wandel C, Kim RB, Kajiji S, et al (1999) P-glycoprotein and cytochrome P-450 3A inhibition: dissociation of inhibitory potencies. Cancer Res 59:3944–3948

8. Ozben T (2006) Mechanisms and strategies to overcome multiple drug resistance in cancer. FEBS Lett 580:2903–2909 . doi: 10.1016/j.febslet.2006.02.020

9. Wu C-P, Calcagno AM, Ambudkar SV (2008) Reversal of ABC drug transporter-mediated multidrug resistance in cancer cells: evaluation of current strategies. Curr Mol Pharmacol 1:93–105

10. Sparreboom A, Danesi R, Ando Y, et al (2003) Pharmacogenomics of ABC transporters and its role in cancer chemotherapy. Drug Resist Updat Rev Comment Antimicrob Anticancer Chemother 6:71–84

11. Sun Y-L, Patel A, Kumar P, Chen Z-S (2012) Role of ABC transporters in cancer chemotherapy. Chin J Cancer 31:51–57 . doi: 10.5732/cjc.011.10466

12. Ford JM (1996) Experimental reversal of P-glycoprotein-mediated multidrug resistance by pharmacological chemosensitisers. Eur J Cancer Oxf Engl 1990 32A:991–1001

13. Gottesman MM, Fojo T, Bates SE (2002) Multidrug resistance in cancer: role of ATP-dependent transporters. Nat Rev Cancer 2:48–58 . doi: 10.1038/nrc706

14. Dinarvand R, Varshochian R, Kamalinia G, et al (2013) Recent approaches to overcoming multiple drug resistance in breast cancer using modified liposomes. Clin Lipidol 8:391–394 . doi: 10.2217/clp.13.33

15. Thomas H, Coley HM (2003) Overcoming multidrug resistance in cancer: an update on the clinical strategy of inhibiting p-glycoprotein. Cancer Control J Moffitt Cancer Cent 10:159–165 . doi: 10.1177/107327480301000207

16. Jabeen I, Wetwitayaklung P, Chiba P, et al (2013) 2D- and 3D-QSAR studies of a series of benzopyranes and benzopyrano[3,4b][1,4]-oxazines as inhibitors of the multidrug transporter P-glycoprotein. J Comput Aided Mol Des 27:161–171 . doi: 10.1007/s10822-013-9635-9

17. Szakács G, Hall MD, Gottesman MM, et al (2014) Targeting the Achilles Heel of Multidrug-Resistant Cancer by Exploiting the Fitness Cost of Resistance. Chem Rev 114:5753–5774 . doi: 10.1021/cr4006236

18. Moitra K (2015) Overcoming Multidrug Resistance in Cancer Stem Cells. In: BioMed Res. Int. https://www.hindawi.com/journals/bmri/2015/635745/. Accessed 24 Mar 2018

19. Fojo AT, Ueda K, Slamon DJ, et al (1987) Expression of a multidrug-resistance gene in human tumors and tissues. Proc Natl Acad Sci U S A 84:265–269

20. Thiebaut F, Tsuruo T, Hamada H, et al (1987) Cellular localization of the multidrug-resistance gene product P-glycoprotein in normal human tissues. Proc Natl Acad Sci U S A 84:7735–7738

21. Deng J, Shao J, Markowitz JS, An G (2014) ABC transporters in multi-drug resistance and ADME-Tox of small molecule tyrosine kinase inhibitors. Pharm Res 31:2237–2255 . doi: 10.1007/s11095-014-1389-0

22. Montanari F, Ecker GF (2015) Prediction of drug–ABC-transporter interaction — Recent advances and future challenges. Adv Drug Deliv Rev 86:17–26 . doi: 10.1016/j.addr.2015.03.001

23. Dean M, Rzhetsky A, Allikmets R (2001) The human ATP-binding cassette (ABC) transporter superfamily. Genome Res 11:1156–1166 . doi: 10.1101/gr.184901

24. Klaassen CD, Aleksunes LM (2010) Xenobiotic, bile acid, and cholesterol transporters: function and regulation. Pharmacol Rev 62:1–96 . doi: 10.1124/pr.109.002014

25. Meier PJ, Stieger B (2002) Bile salt transporters. Annu Rev Physiol 64:635–661 . doi: 10.1146/annurev.physiol.64.082201.100300

26. Aleo MD, Shah F, He K, et al (2017) Evaluating the Role of Multidrug Resistance Protein 3 (MDR3) Inhibition in Predicting Drug-Induced Liver Injury Using 125 Pharmaceuticals. Chem Res Toxicol 30:1219–1229 . doi: 10.1021/acs.chemrestox.7b00048

27. Morgan RE, Trauner M, van Staden CJ, et al (2010) Interference with bile salt export pump function is a susceptibility factor for human liver injury in drug development. Toxicol Sci Off J Soc Toxicol 118:485–500 . doi: 10.1093/toxsci/kfq269

28. Rodrigues AD, Lai Y, Cvijic ME, et al (2014) Drug-induced perturbations of the bile acid pool, cholestasis, and hepatotoxicity: mechanistic considerations beyond the direct inhibition of the bile salt export pump. Drug Metab Dispos Biol Fate Chem 42:566–574 . doi: 10.1124/dmd.113.054205

29. Yucha RW, He K, Shi Q, et al (2017) In Vitro Drug-Induced Liver Injury Prediction: Criteria Optimization of Efflux Transporter IC50 and Physicochemical Properties. Toxicol Sci Off J Soc Toxicol 157:487–499 . doi: 10.1093/toxsci/kfx060

30. Kotsampasakou E, Ecker GF (2017) Predicting Drug-Induced Cholestasis with the Help of Hepatic Transporters—An in Silico Modeling Approach. J Chem Inf Model 57:608–615 . doi: 10.1021/acs.jcim.6b00518

31. Clinical Drug Interaction Studies — Study Design, Data Analysis, and Clinical Implications. Data Anal 32

32. (2017) Clinical Drug Interaction Studies — Study Design, Data Analysis, and Clinical Implications Guidance for Industry

33. Brouwer KLR, Keppler D, Hoffmaster KA, et al (2013) In vitro methods to support transporter evaluation in drug discovery and development. Clin Pharmacol Ther 94:95–112 . doi: 10.1038/clpt.2013.81

34. Stockner T, de Vries SJ, Bonvin AMJJ, et al (2009) Data-driven homology modelling of P-glycoprotein in the ATP-bound state indicates flexibility of the transmembrane domains. FEBS J 276:964–972 . doi: 10.1111/j.1742-4658.2008.06832.x

35. Jain S, Grandits M, Richter L, Ecker GF (2017) Structure based classification for bile salt export pump (BSEP) inhibitors using comparative structural modeling of human BSEP. J Comput Aided Mol Des 31:507–521 . doi: 10.1007/s10822-017-0021-x

36. Ahuja S, Rougé L, Swem DL, et al (2015) Structural Analysis of Bacterial ABC Transporter Inhibition by an Antibody Fragment. Structure 23:713–723 . doi: 10.1016/j.str.2015.01.020

37. Kotsampasakou E, Jain S, Digles D, F. Ecker G (2018) Transporters in Hepatotoxicity. pp 145–174

38. Kotsiantis S, Kanellopoulos D, Pintelas P Handling imbalanced datasets: A review

39. Ali A, Shamsuddin SM, Ralescu AL (2015) Classification with class imbalance problem: A Review. Int J Adv Soft Compu Appl 7:176–204

40. López V, Fernández A, Moreno-Torres JG, Herrera F (2012) Analysis of Preprocessing vs. Cost-sensitive Learning for Imbalanced Classification. Open Problems on Intrinsic Data Characteristics. Expert Syst Appl 39:6585–6608 . doi: 10.1016/j.eswa.2011.12.043

41. Holland IB, Blight MA (1999) ABC-ATPases, adaptable energy generators fuelling transmembrane movement of a variety of molecules in organisms from bacteria to humans. J Mol Biol 293:381–399 . doi: 10.1006/jmbi.1999.2993

42. ter Beek J, Guskov A, Slotboom DJ (2014) Structural diversity of ABC transporters. J Gen Physiol 143:419–435 . doi: 10.1085/jgp.201411164

43. Rees DC, Johnson E, Lewinson O (2009) ABC transporters: The power to change. Nat Rev Mol Cell Biol 10:218–227 . doi: 10.1038/nrm2646

44. Wilkens S (2015) Structure and mechanism of ABC transporters. F1000prime Rep 7:14 . doi: 10.12703/P7-14

45. Vasiliou V, Vasiliou K, Nebert DW (2009) Human ATP-binding cassette (ABC) transporter family. Hum Genomics 3:281–290 . doi: 10.1186/1479-7364-3-3-281

46. Hung LW, Wang IX, Nikaido K, et al (1998) Crystal structure of the ATP-binding subunit of an ABC transporter. Nature 396:703–707 . doi: 10.1038/25393

47. Kos V, Ford RC (2009) The ATP-binding cassette family: a structural perspective. Cell Mol Life Sci CMLS 66:3111–3126 . doi: 10.1007/s00018-009-0064-9

48. Moeller A, Lee SC, Tao H, et al (2015) Distinct Conformational Spectrum of Homologous Multidrug ABC Transporters. Structure 23:450–460 . doi: 10.1016/j.str.2014.12.013

49. Dong J, Yang G, McHaourab HS (2005) Structural basis of energy transduction in the transport cycle of MsbA. Science 308:1023–1028 . doi: 10.1126/science.1106592

50. Dawson RJP, Locher KP (2006) Structure of a bacterial multidrug ABC transporter. Nature 443:180–185 . doi: 10.1038/nature05155

51. Oldham ML, Khare D, Quiocho FA, et al (2007) Crystal structure of a catalytic intermediate of the maltose transporter. Nature 450:515–521 . doi: 10.1038/nature06264

52. Hollenstein K, Frei DC, Locher KP (2007) Structure of an ABC transporter in complex with its binding protein. Nature 446:213–216 . doi: 10.1038/nature05626

53. Locher KP, Lee AT, Rees DC (2002) The E. coli BtuCD structure: a framework for ABC transporter architecture and mechanism. Science 296:1091–1098 . doi: 10.1126/science.1071142

54. Pinkett HW, Lee AT, Lum P, et al (2007) An inward-facing conformation of a putative metal-chelate-type ABC transporter. Science 315:373–377 . doi: 10.1126/science.1133488

55. Kadaba NS, Kaiser JT, Johnson E, et al (2008) The high-affinity E. coli methionine ABC transporter: structure and allosteric regulation. Science 321:250–253 . doi: 10.1126/science.1157987

56. Ward A, Reyes CL, Yu J, et al (2007) Flexibility in the ABC transporter MsbA: Alternating access with a twist. Proc Natl Acad Sci U S A 104:19005–19010 . doi: 10.1073/pnas.0709388104

57. Aller SG, Yu J, Ward A, et al (2009) Structure of P-glycoprotein reveals a molecular basis for poly-specific drug binding. Science 323:1718–1722 . doi: 10.1126/science.1168750

58. Hoffmann U, Kroemer HK (2004) The ABC transporters MDR1 and MRP2: multiple functions in disposition of xenobiotics and drug resistance. Drug Metab Rev 36:669–701 . doi: 10.1081/DMR-200033473

59. Juliano RL, Ling V (1976) A surface glycoprotein modulating drug permeability in Chinese hamster ovary cell mutants. Biochim Biophys Acta 455:152–162

60. Li J, Jaimes KF, Aller SG (2014) Refined structures of mouse P-glycoprotein. Protein Sci Publ Protein Soc 23:34–46 . doi: 10.1002/pro.2387

61. Cordon-Cardo C, O'Brien JP, Boccia J, et al (1990) Expression of the multidrug resistance gene product (P-glycoprotein) in human normal and tumor tissues. J Histochem Cytochem Off J Histochem Soc 38:1277–1287 . doi: 10.1177/38.9.1974900

62. Uhlén M, Fagerberg L, Hallström BM, et al (2015) Proteomics. Tissue-based map of the human proteome. Science 347:1260419 . doi: 10.1126/science.1260419

63. Goldstein LJ, Galski H, Fojo A, et al (1989) Expression of a multidrug resistance gene in human cancers. J Natl Cancer Inst 81:116–124

64. Bentires-Alj M, Barbu V, Fillet M, et al (2003) NF-kappaB transcription factor induces drug resistance through MDR1 expression in cancer cells. Oncogene 22:90–97 . doi: 10.1038/sj.onc.1206056

65. Keogh JP (2012) Membrane transporters in drug development. Adv Pharmacol San Diego Calif 63:1–42 . doi: 10.1016/B978-0-12-398339-8.00001-X

66. Marchetti S, Mazzanti R, Beijnen JH, Schellens JHM (2007) Concise review: Clinical relevance of drug drug and herb drug interactions mediated by the ABC transporter ABCB1 (MDR1, P-glycoprotein). The Oncologist 12:927–941 . doi: 10.1634/theoncologist.12-8-927

67. Hochman JH, Yamazaki M, Ohe T, Lin JH (2002) Evaluation of drug interactions with P-glycoprotein in drug discovery: in vitro assessment of the potential for drug-drug interactions with P-glycoprotein. Curr Drug Metab 3:257–273

68. Ding R, Tayrouz Y, Riedel K-D, et al (2004) Substantial pharmacokinetic interaction between digoxin and ritonavir in healthy volunteers. Clin Pharmacol Ther 76:73–84 . doi: 10.1016/j.clpt.2004.02.008

69. Leahey EB, Bigger JT, Butler VP, et al (1981) Quinidine-digoxin interaction: time course and pharmacokinetics. Am J Cardiol 48:1141–1146

70. Cheng X, Buckley D, Klaassen CD (2007) Regulation of hepatic bile acid transporters Ntcp and Bsep expression. Biochem Pharmacol 74:1665–1676 . doi: 10.1016/j.bcp.2007.08.014

71. Giacomini KM, Huang S-M, Tweedie DJ, et al (2010) Membrane transporters in drug development. Nat Rev Drug Discov 9:215–236 . doi: 10.1038/nrd3028

72. Glavinas H, Krajcsi P, Cserepes J, Sarkadi B (2004) The role of ABC transporters in drug resistance, metabolism and toxicity. Curr Drug Deliv 1:27–42

73. Chiang JYL (2004) Regulation of bile acid synthesis: pathways, nuclear receptors, and mechanisms. J Hepatol 40:539–551 . doi: 10.1016/j.jhep.2003.11.006

74. Chiang JY (1998) Regulation of bile acid synthesis. Front Biosci J Virtual Libr 3:d176–193

75. Russell DW (2003) The enzymes, regulation, and genetics of bile acid synthesis. Annu Rev Biochem 72:137–174 . doi: 10.1146/annurev.biochem.72.121801.161712

76. Hofmann AF, Borgström B (1964) The Intraluminal Phase of Fat Digestion in Man: The Lipid Content of the Micellar and Oil Phases of Intestinal Content Obtained during Fat Digestion and Absorption*. J Clin Invest 43:247–257

77. Perez M-J, Briz O (2009) Bile-acid-induced cell injury and protection. World J Gastroenterol WJG 15:1677–1689

78. Strautnieks SS, Byrne JA, Pawlikowska L, et al (2008) Severe bile salt export pump deficiency: 82 different ABCB11 mutations in 109 families. Gastroenterology 134:1203–1214 . doi: 10.1053/j.gastro.2008.01.038

79. JANSEN P, MULLER M (2000) The molecular genetics of familial intrahepatic cholestasis. Gut 47:1–5 . doi: 10.1136/gut.47.1.1

80. Alonso EM, Snover DC, Montag A, et al (1994) Histologic pathology of the liver in progressive familial intrahepatic cholestasis. J Pediatr Gastroenterol Nutr 18:128–133

81. Amer S, Hajira A (2014) A Comprehensive Review of Progressive Familial Intrahepatic Cholestasis (PFIC): Genetic Disorders of Hepatocanalicular Transporters. Gastroenterol Res 7:39–43 . doi: 10.14740/gr.v7i2.612

82. Chen M, Suzuki A, Borlak J, et al (2015) Drug-induced liver injury: Interactions between drug properties and host factors. J Hepatol 63:503–514 . doi: 10.1016/j.jhep.2015.04.016

83. Dawson S, Stahl S, Paul N, et al (2012) In vitro inhibition of the bile salt export pump correlates with risk of cholestatic drug-induced liver injury in humans. Drug Metab Dispos Biol Fate Chem 40:130–138 . doi: 10.1124/dmd.111.040758

84. Sahi J, Sinz MW, Campbell S, et al (2006) Metabolism and transporter-mediated drug-drug interactions of the endothelin-A receptor antagonist CI-1034. Chem Biol Interact 159:156–168 . doi: 10.1016/j.cbi.2005.11.001

85. Allikmets R, Schriml LM, Hutchinson A, et al (1998) A human placenta-specific ATP-binding cassette gene (ABCP) on chromosome 4q22 that is involved in multidrug resistance. Cancer Res 58:5337–5339

86. Taylor NMI, Manolaridis I, Jackson SM, et al (2017) Structure of the human multidrug transporter ABCG2. Nature 546:504–509 . doi: 10.1038/nature22345

87. Staud F, Pavek P (2005) Breast cancer resistance protein (BCRP/ABCG2). Int J Biochem Cell Biol 37:720–725 . doi: 10.1016/j.biocel.2004.11.004

88. Ni Z, Bikadi Z, Rosenberg MF, Mao Q (2010) Structure and function of the human breast cancer resistance protein (BCRP/ABCG2). Curr Drug Metab 11:603–617

89. Natarajan K, Xie Y, Baer MR, Ross DD (2012) Role of breast cancer resistance protein (BCRP/ABCG2) in cancer drug resistance. Biochem Pharmacol 83:1084–1103 . doi: 10.1016/j.bcp.2012.01.002

90. Ni Z, Bikadi Z, Cai X, et al (2010) Transmembrane helices 1 and 6 of the human breast cancer resistance protein (BCRP/ABCG2): identification of polar residues important for drug transport. Am J Physiol Cell Physiol 299:C1100-1109 . doi: 10.1152/ajpcell.00160.2010

91. McDevitt CA, Collins RF, Conway M, et al (2006) Purification and 3D structural analysis of oligomeric human multidrug transporter ABCG2. Struct Lond Engl 1993 14:1623–1632 . doi: 10.1016/j.str.2006.08.014

92. Xu J, Liu Y, Yang Y, et al (2004) Characterization of oligomeric human half-ABC transporter ATP-binding cassette G2. J Biol Chem 279:19781–19789 . doi: 10.1074/jbc.M310785200

93. Litman T, Jensen U, Hansen A, et al (2002) Use of peptide antibodies to probe for the mitoxantrone resistance-associated protein MXR/BCRP/ABCP/ABCG2. Biochim Biophys Acta 1565:6–16

94. Kage K, Tsukahara S, Sugiyama T, et al (2002) Dominant-negative inhibition of breast cancer resistance protein as drug efflux pump through the inhibition of S-S dependent homodimerization. Int J Cancer 97:626–630

95. Mao Q, Unadkat JD (2014) Role of the Breast Cancer Resistance Protein (BCRP/ABCG2) in Drug Transport—an Update. AAPS J 17:65–82 . doi: 10.1208/s12248-014-9668-6

96. Robey RW, To KKK, Polgar O, et al (2009) ABCG2: a perspective. Adv Drug Deliv Rev 61:3–13 . doi: 10.1016/j.addr.2008.11.003

97. Bosch TM, Kjellberg LM, Bouwers A, et al (2005) Detection of single nucleotide polymorphisms in the ABCG2 gene in a Dutch population. Am J Pharmacogenomics Genomics-Relat Res Drug Dev Clin Pract 5:123–131

98. de Jong FA, Marsh S, Mathijssen RHJ, et al (2004) ABCG2 pharmacogenetics: ethnic differences in allele frequency and assessment of influence on irinotecan disposition. Clin Cancer Res Off J Am Assoc Cancer Res 10:5889–5894 . doi: 10.1158/1078-0432.CCR-04-0144

99. Bäckström G, Taipalensuu J, Melhus H, et al (2003) Genetic variation in the ATP-binding cassette transporter gene ABCG2 (BCRP) in a Swedish population. Eur J Pharm Sci Off J Eur Fed Pharm Sci 18:359–364

100. Honjo Y, Morisaki K, Huff LM, et al (2002) Single-nucleotide polymorphism (SNP) analysis in the ABC half-transporter ABCG2 (MXR/BCRP/ABCP1). Cancer Biol Ther 1:696–702

101. Lepper ER, Nooter K, Verweij J, et al (2005) Mechanisms of resistance to anticancer drugs: the role of the polymorphic ABC transporters ABCB1 and ABCG2. Pharmacogenomics 6:115–138 . doi: 10.1517/14622416.6.2.115

102. Imai Y, Nakane M, Kage K, et al (2002) C421A polymorphism in the human breast cancer resistance protein gene is associated with low expression of Q141K protein and low-level drug resistance. Mol Cancer Ther 1:611–616

103. Kühnle M, Egger M, Müller C, et al (2009) Potent and selective inhibitors of breast cancer resistance protein (ABCG2) derived from the p-glycoprotein (ABCB1) modulator tariquidar. J Med Chem 52:1190–1197 . doi: 10.1021/jm8013822

104. Boumendjel A, Macalou S, Ahmed-Belkacem A, et al (2007) Acridone derivatives: design, synthesis, and inhibition of breast cancer resistance protein ABCG2. Bioorg Med Chem 15:2892–2897 . doi: 10.1016/j.bmc.2007.02.017

105. Kruijtzer CMF, Beijnen JH, Rosing H, et al (2002) Increased oral bioavailability of topotecan in combination with the breast cancer resistance protein and P-glycoprotein inhibitor GF120918. J Clin Oncol Off J Am Soc Clin Oncol 20:2943–2950 . doi: 10.1200/JCO.2002.12.116

106. Schwede T (2013) Protein modeling: what happened to the "protein structure gap"? Struct Lond Engl 1993 21:1531–1540 . doi: 10.1016/j.str.2013.08.007

107. Schwede T, Sali A, Honig B, et al (2009) Outcome of a Workshop on Applications of Protein Models in Biomedical Research. Structure 17:151–159 . doi: 10.1016/j.str.2008.12.014

108. Shi Y (2014) A glimpse of structural biology through X-ray crystallography. Cell 159:995–1014 . doi: 10.1016/j.cell.2014.10.051

109. Berg JM, Tymoczko JL, Stryer L (2002) Three-Dimensional Protein Structure Can Be Determined by NMR Spectroscopy and X-Ray Crystallography

110. Carpenter EP, Beis K, Cameron AD, Iwata S (2008) Overcoming the challenges of membrane protein crystallography. Curr Opin Struct Biol 18:581–586 . doi: 10.1016/j.sbi.2008.07.001

111. Smith SM (2011) Strategies for the purification of membrane proteins. Methods Mol Biol Clifton NJ 681:485–496 . doi: 10.1007/978-1-60761-913-0_29

112. Schmidt T, Bergner A, Schwede T (2014) Modelling three-dimensional protein structures for applications in drug design. Drug Discov Today 19:890–897 . doi: 10.1016/j.drudis.2013.10.027

113. Jurik A, Klepsch F, Zdrazil B (2012) Molecular Modeling and Simulation of Membrane Transport Proteins. pp 373–406

114. Venselaar H, Joosten RP, Vroling B, et al (2010) Homology modelling and spectroscopy, a never-ending love story. Eur Biophys J EBJ 39:551–563 . doi: 10.1007/s00249-009-0531-0

115. Xiang Z (2006) Advances in Homology Protein Structure Modeling. Curr Protein Pept Sci 7:217–227

116. Hongmao S (2016) Chapter 4 - Homology Modeling and Ligand-Based Molecule Design. In: A Practical Guide to Rational Drug Design. Woodhead Publishing, pp 109–160

117. Martí-Renom MA, Stuart AC, Fiser A, et al (2000) Comparative protein structure modeling of genes and genomes. Annu Rev Biophys Biomol Struct 29:291–325 . doi: 10.1146/annurev.biophys.29.1.291

118. Altschul SF, Madden TL, Schäffer AA, et al (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res 25:3389–3402

119. Altschul SF, Gish W, Miller W, et al (1990) Basic local alignment search tool. J Mol Biol 215:403–410 . doi: 10.1016/S0022-2836(05)80360-2

120. Madden TL, Tatusov RL, Zhang J (1996) Applications of network BLAST server. Methods Enzymol 266:131–141

121. Eswar N, Webb B, Marti-Renom MA, et al (2007) Comparative protein structure modeling using MODELLER. Curr Protoc Protein Sci Editor Board John E Coligan Al Chapter 2:Unit 2.9 . doi: 10.1002/0471140864.ps0209s50

122. Webb B, Sali A (2016) Comparative Protein Structure Modeling Using MODELLER. Curr Protoc Bioinforma 54:5.6.1-5.6.37 . doi: 10.1002/cpbi.3

123. Shen M, Sali A (2006) Statistical potential for assessment and prediction of protein structures. Protein Sci Publ Protein Soc 15:2507–2524 . doi: 10.1110/ps.062416606

124. Laskowski R, Macarthur M, Moss D, Thornton J (1993) {PROCHECK}: a program to check the stereochemical quality of protein structures. J Appl Cryst 26:283–291

125. Zhou AQ, O'Hern C, Regan L (2011) Revisiting the Ramachandran plot from a new angle. Protein Sci Publ Protein Soc 20:1166–1171 . doi: 10.1002/pro.644

126. Benkert P, Künzli M, Schwede T (2009) QMEAN server for protein model quality estimation. Nucleic Acids Res 37:W510-514 . doi: 10.1093/nar/gkp322

127. Benkert P, Tosatto SCE, Schomburg D (2008) QMEAN: A comprehensive scoring function for model quality assessment. Proteins 71:261–277 . doi: 10.1002/prot.21715

128. Hillisch A, Pineda LF, Hilgenfeld R (2004) Utility of homology models in the drug discovery process. Drug Discov Today 9:659–669 . doi: 10.1016/S1359-6446(04)03196-4

129. Kuntz ID, Blaney JM, Oatley SJ, et al (1982) A geometric approach to macromolecule-ligand interactions. J Mol Biol 161:269–288

130. Goodsell DS, Morris GM, Olson AJ (1996) Automated docking of flexible ligands: applications of AutoDock. J Mol Recognit JMR 9:1–5 . doi: 10.1002/(SICI)1099-1352(199601)9:1<1::AID-JMR241>3.0.CO;2-6

131.   Friesner RA, Murphy RB, Repasky MP, et al (2006) Extra Precision Glide: Docking and Scoring Incorporating a Model of Hydrophobic Enclosure for Protein−Ligand Complexes. J Med Chem 49:6177–6196 . doi: 10.1021/jm051256o

132.   Halgren TA, Murphy RB, Friesner RA, et al (2004) Glide: A New Approach for Rapid, Accurate Docking and Scoring. 2. Enrichment Factors in Database Screening. J Med Chem 47:1750–1759 . doi: 10.1021/jm030644s

133.   Jones G, Willett P, Glen RC, et al (1997) Development and validation of a genetic algorithm for flexible docking. J Mol Biol 267:727–748 . doi: 10.1006/jmbi.1996.0897

134.   Verdonk ML, Cole JC, Hartshorn MJ, et al (2003) Improved protein–ligand docking using GOLD. Proteins Struct Funct Bioinforma 52:609–623 . doi: 10.1002/prot.10465

135.   Chen R, Li L, Weng Z (2003) ZDOCK: an initial-stage protein-docking algorithm. Proteins 52:80–87 . doi: 10.1002/prot.10389

136.   Pierce BG, Hourai Y, Weng Z (2011) Accelerating Protein Docking in ZDOCK Using an Advanced 3D Convolution Library. PLOS ONE 6:e24657 . doi: 10.1371/journal.pone.0024657

137.   Grosdidier A, Zoete V, Michielin O (2011) SwissDock, a protein-small molecule docking web service based on EADock DSS. Nucleic Acids Res 39:W270-277 . doi: 10.1093/nar/gkr366

138.   Schrödinger Release 2015-1: Maestro, version 10.1, Schrödinger, LLC, New York, NY, 2015.

139.   Schrödinger Release 2018-1: Maestro, version 10.1, Schrödinger, LLC, New York, NY, 2018.

140.   Sherman W, Beard HS, Farid R (2006) Use of an induced fit receptor structure in virtual screening. Chem Biol Drug Des 67:83–84 . doi: 10.1111/j.1747-0285.2005.00327.x

141.   Sherman W, Day T, Jacobson MP, et al (2006) Novel procedure for modeling ligand/receptor induced fit effects. J Med Chem 49:534–553 . doi: 10.1021/jm050540c

142.   Lovell SC, Word JM, Richardson JS, Richardson DC (2000) The penultimate rotamer library. Proteins 40:389–408

143.   Pagadala NS, Syed K, Tuszynski J (2017) Software for molecular docking: a review. Biophys Rev 9:91–102 . doi: 10.1007/s12551-016-0247-1

144.   Huang S-Y, Grinter SZ, Zou X (2010) Scoring functions and their evaluation methods for protein-ligand docking: recent advances and future directions. Phys Chem Chem Phys PCCP 12:12899–12908 . doi: 10.1039/c0cp00151a

145.   Liu J, Wang R (2015) Classification of Current Scoring Functions. J Chem Inf Model 55:475–482 . doi: 10.1021/ci500731a

146. Berman HM, Westbrook J, Feng Z, et al (2000) The Protein Data Bank. Nucleic Acids Res 28:235–242

147. Groom CR, Allen FH (2014) The Cambridge Structural Database in retrospect and prospect. Angew Chem Int Ed Engl 53:662–671 . doi: 10.1002/anie.201306438

148. Muegge I (2006) PMF Scoring Revisited. J Med Chem 49:5895–5902 . doi: 10.1021/jm050038s

149. Muegge I (2000) A knowledge-based scoring function for protein-ligand interactions: Probing the reference state. In: Virtual Screening: An Alternative or Complement to High Throughput Screening? Springer, Dordrecht, pp 99–114

150. Gohlke H, Hendlich M, Klebe G (2000) Knowledge-based scoring function to predict protein-ligand interactions. J Mol Biol 295:337–356 . doi: 10.1006/jmbi.1999.3371

151. Huang S-Y, Zou X (2006) An iterative knowledge-based scoring function to predict protein-ligand interactions: I. Derivation of interaction potentials. J Comput Chem 27:1866–1875 . doi: 10.1002/jcc.20504

152. Eldridge MD, Murray CW, Auton TR, et al (1997) Empirical scoring functions: I. The development of a fast empirical scoring function to estimate the binding affinity of ligands in receptor complexes. J Comput Aided Mol Des 11:425–445 . doi: 10.1023/A:1007996124545

153. Jabeen I, Wetwitayaklung P, Klepsch F, et al (2011) Probing the stereoselectivity of P-glycoprotein—synthesis, biological activity and ligand docking studies of a set of enantiopure benzopyrano[3,4-b][1,4]oxazines. Chem Commun 47:2586–2588 . doi: 10.1039/C0CC03075A

154. Jones G, Willett P, Glen RC (1995) Molecular recognition of receptor sites using a genetic algorithm with a description of desolvation. J Mol Biol 245:43–53 . doi: 10.1016/S0022-2836(95)80037-9

155. Korb O, Stützle T, Exner TE (2009) Empirical Scoring Functions for Advanced Protein−Ligand Docking with PLANTS. J Chem Inf Model 49:84–96 . doi: 10.1021/ci800298z

156. Mooij WTM, Verdonk ML (2005) General and targeted statistical potentials for protein-ligand interactions. Proteins 61:272–287 . doi: 10.1002/prot.20588

157. Wang R, Lai L, Wang S (2002) Further development and validation of empirical scoring functions for structure-based binding affinity prediction. J Comput Aided Mol Des 16:11–26

158. Cheng T, Li X, Li Y, et al (2009) Comparative assessment of scoring functions on a diverse test set. J Chem Inf Model 49:1079–1093 . doi: 10.1021/ci9000053

159. Li H, Leung K-S, Wong M-H, Ballester PJ (2014) Substituting random forest for multiple linear regression improves binding affinity prediction of scoring functions: Cyscore as a case study. BMC Bioinformatics 15:291 . doi: 10.1186/1471-2105-15-291

160.   Charifson PS, Corkery JJ, Murcko MA, Walters WP (1999) Consensus scoring: A method for obtaining improved hit rates from docking databases of three-dimensional structures into proteins. J Med Chem 42:5100–5109

161.   Richter L, de Graaf C, Sieghart W, et al (2012) Diazepam-bound GABAA receptor models identify new benzodiazepine binding-site ligands. Nat Chem Biol 8:455–464 . doi: 10.1038/nchembio.917

162.   Klepsch F, Chiba P, Ecker GF (2011) Exhaustive Sampling of Docking Poses Reveals Binding Hypotheses for Propafenone Type Inhibitors of P-Glycoprotein. PLoS Comput Biol 7:e1002036 . doi: 10.1371/journal.pcbi.1002036

163.   Wang H, Goehring A, Wang KH, et al (2013) Structural basis for action by diverse antidepressants on biogenic amine transporters. Nature 503:141–145 . doi: 10.1038/nature12648

164.   Hospital A, Goñi JR, Orozco M, Gelpí JL (2015) Molecular dynamics simulations: advances and applications. Adv Appl Bioinforma Chem AABC 8:37–47 . doi: 10.2147/AABC.S70333

165.   De Vivo M, Masetti M, Bottegoni G, Cavalli A (2016) Role of Molecular Dynamics and Related Methods in Drug Discovery. J Med Chem 59:4035–4061 . doi: 10.1021/acs.jmedchem.5b01684

166.   Stansfeld PJ, Sansom MSP (2011) Molecular simulation approaches to membrane proteins. Struct Lond Engl 1993 19:1562–1572 . doi: 10.1016/j.str.2011.10.002

167.   Adcock SA, McCammon JA (2006) Molecular Dynamics: Survey of Methods for Simulating the Activity of Proteins. Chem Rev 106:1589–1615 . doi: 10.1021/cr040426m

168.   Chavent M, Duncan AL, Sansom MS (2016) Molecular dynamics simulations of membrane proteins and their interactions: from nanoscale to mesoscale. Curr Opin Struct Biol 40:8–16 . doi: 10.1016/j.sbi.2016.06.007

169.   Jorgensen WL, Tirado-Rives J (1988) The OPLS [optimized potentials for liquid simulations] potential functions for proteins, energy minimizations for crystals of cyclic peptides and crambin. J Am Chem Soc 110:1657–1666 . doi: 10.1021/ja00214a001

170.   Brooks Bernard R., Bruccoleri Robert E., Olafson Barry D., et al (2004) CHARMM: A program for macromolecular energy, minimization, and dynamics calculations. J Comput Chem 4:187–217 . doi: 10.1002/jcc.540040211

171.   Brooks BR, Bruccoleri RE, Olafson BD, et al (1983) CHARMM: A program for macromolecular energy, minimization, and dynamics calculations. J Comput Chem 4:187–217 . doi: 10.1002/jcc.540040211

172.   Scott WRP, Hünenberger PH, Tironi IG, et al (1999) The GROMOS Biomolecular Simulation Program Package. J Phys Chem A 103:3596–3607 . doi: 10.1021/jp984217f

173.   Ryckaert J-P, Ciccotti G, Berendsen HJC (1977) Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. J Comput Phys 23:327–341 . doi: 10.1016/0021-9991(77)90098-5

174.   Hess B, Bekker H, Berendsen HJC, Fraaije JGEM (1997) LINCS: A Linear Constraint Solver for Molecular Simulations. J Comput Chem 18:18–1463

175.   Nam K (2014) Acceleration of Ab Initio QM/MM Calculations under Periodic Boundary Conditions by Multiscale and Multiple Time Step Approaches. J Chem Theory Comput 10:4175–4183 . doi: 10.1021/ct5005643

176.   Kasahara K, Sakuraba S, Fukuda I (2018) Enhanced Sampling of Molecular Dynamics Simulations of a Polyalanine Octapeptide: Effects of the Periodic Boundary Conditions on Peptide Conformation. J Phys Chem B 122:2495–2503 . doi: 10.1021/acs.jpcb.7b10830

177.   Frenkel D (2013) Simulations: The dark side. Eur Phys J Plus 128:10 . doi: 10.1140/epjp/i2013-13010-8

178.   Kukol A (2017) Molecular Modeling of Proteins. Springer New York

179.   Karplus M, McCammon JA (2002) Molecular dynamics simulations of biomolecules. Nat Struct Biol 9:646–652 . doi: 10.1038/nsb0902-646

180.   Papa E, Kovarich S, Gramatica P (2010) QSAR Modeling and Prediction of the Endocrine-Disrupting Potencies of Brominated Flame Retardants. Chem Res Toxicol 23:946–954 . doi: 10.1021/tx1000392

181.   Schneider G (2013) Modeling Structure-Activity Relationships. Landes Bioscience

182.   Khanna I (2012) Drug discovery in pharmaceutical industry: productivity challenges and trends. Drug Discov Today 17:1088–1102 . doi: 10.1016/j.drudis.2012.05.007

183.   Berman HM, Kleywegt GJ, Nakamura H, Markley JL (2013) The Future of the Protein Data Bank. Biopolymers 99:218–222 . doi: 10.1002/bip.22132

184.   Schwarz T, Montanari F, Cseke A, et al (2016) Subtle Structural Differences Trigger Inhibitory Activity of Propafenone Analogues at the Two Polyspecific ABC Transporters: P-Glycoprotein (P-gp) and Breast Cancer Resistance Protein (BCRP). ChemMedChem 11:1380–1394 . doi: 10.1002/cmdc.201500592

185.   Cramer J, Kopp S, Bates SE, et al (2007) Multispecificity of drug transporters: probing inhibitor selectivity for the human drug efflux transporters ABCB1 and ABCG2. ChemMedChem 2:1783–1788 . doi: 10.1002/cmdc.200700160

186.   DeGorter MK, Xia CQ, Yang JJ, Kim RB (2012) Drug transporters in drug efficacy and toxicity. Annu Rev Pharmacol Toxicol 52:249–273 . doi: 10.1146/annurev-pharmtox-010611-134529

187.   Bodó A, Bakos E, Szeri F, et al (2003) The role of multidrug transporters in drug availability, metabolism and toxicity. Toxicol Lett 140–141:133–143

188.   Köck K, Brouwer KLR (2012) A perspective on efflux transport proteins in the liver. Clin Pharmacol Ther 92:599–612 . doi: 10.1038/clpt.2012.79

189.   Daneshian M, Kamp H, Hengstler J, et al (2016) Highlight report: Launch of a large integrated European in vitro toxicology project: EU-ToxRisk. Arch Toxicol 90:1021–1024 . doi: 10.1007/s00204-016-1698-7

190.   Van De Water B (2016) The EU-ToxRisk project: A European flagship program for mechanism-based safety sciences and risk assessment. Toxicol Lett 258:S20 . doi: 10.1016/j.toxlet.2016.06.1185

191.   EU-ToxRisk - EU-ToxRisk – An Integrated European 'Flagship' Programme Driving Mechanism-based Toxicity Testing and Risk Assessment for the 21st century. http://www.eu-toxrisk.eu/. Accessed 18 Aug 2017

192.   Flecknell P (2002) Replacement, reduction and refinement. ALTEX 19:73–78

193.   Liebsch M, Grune B, Seiler A, et al (2011) Alternatives to animal testing: current status and future perspectives. Arch Toxicol 85:841–858 . doi: 10.1007/s00204-011-0718-x

194.   Graham ML, Prescott MJ (2015) The multifactorial role of the 3Rs in shifting the harm-benefit analysis in animal models of disease. Eur J Pharmacol 759:19–29 . doi: 10.1016/j.ejphar.2015.03.040

195.   Chen L, Li Y, Zhao Q, et al (2011) ADME Evaluation in Drug Discovery. 10. Predictions of P-Glycoprotein Inhibitors Using Recursive Partitioning and Naive Bayesian Classification Techniques. Mol Pharm 8:889–900 . doi: 10.1021/mp100465q

196.   Klepsch F, Vasanthanathan P, Ecker GF (2014) Ligand and Structure-Based Classification Models for Prediction of P-Glycoprotein Inhibitors. J Chem Inf Model 54:218–229 . doi: 10.1021/ci400289j

197.   Zakharov AV, Peach ML, Sitzmann M, Nicklaus MC (2014) QSAR Modeling of Imbalanced High-Throughput Screening Data in PubChem. J Chem Inf Model 54:705–712 . doi: 10.1021/ci400737s

198.   Chen J, Tang YY, Fang B, Guo C (2012) In silico prediction of toxic action mechanisms of phenols for imbalanced data with Random Forest learner. J Mol Graph Model 35:21–27 . doi: 10.1016/j.jmgm.2012.01.002

199.   Khalilia M, Chakraborty S, Popescu M (2011) Predicting disease risks from highly imbalanced data using random forest. BMC Med Inform Decis Mak 11:51 . doi: 10.1186/1472-6947-11-51

200.   Ferrara P, Jacoby E (2007) Evaluation of the utility of homology models in high throughput docking. J Mol Model 13:897–905 . doi: 10.1007/s00894-007-0207-6

201.   Klepsch F, Ecker GF (2010) Impact of the Recent Mouse P-Glycoprotein Structure for Structure-Based Ligand Design. Mol Inform 29:276–286 . doi: 10.1002/minf.201000017

Abstract

ABC-transporters such as the bile salt export pump (BSEP), the breast cancer resistance protein (BCRP) and P-glycoprotein (P-gp) play an important role in the pharmacokinetics of several drugs and small molecules. Predicting inhibition of these transporters by small molecules facilitates identification of potential drug-drug interactions and adverse effects such as drug-induced liver injuries. Thus far, *in silico* identification of inhibitors is dominated by ligand-based approaches that most often employed Quantitative structure–activity relationship (QSAR) and machine learning methods. Although the models based on these methods are reported to be efficient, they do not consider the properties of the protein and thus fail to provide insights into the mechanism of inhibition. While structure-based studies could investigate these details, the lack of high-resolution structural information and the polyspecific binding behaviour of these transporters pose a serious obstacle.

This thesis outlines three independent studies that explore structure-based methods to investigate the molecular basis of inhibition of transporter proteins relevant to liver toxicity and another study that employs ligand-based methods to deal with the imbalanced datasets. The structure-based studies presented here describe the use of homology modeling and molecular docking to uncover the protein-ligand interactions involved in the mechanism of inhibition.

In our first study, a homology model was constructed for BSEP, followed by the development of structure-assisted, docking-based classification models for prediction of BSEP inhibitors. Further, we analyzed the protein-ligand interaction fingerprints which revealed specific functional group-amino acid residue interactions that could play a key role in ligand binding. In the BCRP study, a structure-based modeling approach facilitated elucidation of binding hypothesis for arylmethyloxyphenyl derivatives, which after experimental validations could guide rational optimization of this compound class to improve potency. In the third study, we compared the binding site interaction profiles of human, rat and mouse P-gp structures to reveal a significant overlap between the binding site interacting residues which suggests the transferability of *in vitro* human P-gp activity data in the development of *in silico* models to predict *in vitro* and *in vivo* effects in rodents. In our ligand based study, we dealt with the problem of learning on imbalanced datasets relevant to toxicity by evaluating the

performance of seven distinct meta-classifiers and provided recommendations in choosing an appropriate classifier depending on the dataset in hand.

The results of this thesis work further improve our understanding of protein-ligand interactions at the molecular level, stimulating scientists to conduct new experiments and thus also aid in the extrapolation of molecular hypotheses from rodents to humans and *vice-versa*. Furthermore, combining ligand-based and structure-based approaches would significantly enhance the performance of virtual screening experiments in drug discovery and provide detailed insights on the molecular features involved in crucial interactions, thereby assisting lead optimization.

Zusammenfassung

ABC-Transporter wie z.B. die Gallensalzexportpumpe BSEP (bile salt export pump), der Effluxtransporter BCRP (breast cancer resistance protein) oder das P-Glykoprotein (P-gp) spielen eine wichtige Rolle in der Pharmakokinetik zahlreicher Wirkstoffe und kleiner organischer Moleküle. Die Vorhersage der Transporterhemmung durch chemische Verbindungen ermöglicht die Identifizierung von potenziellen Arzneistoffwechselwirkungen und unerwünschten Wirkungen wie z.B. der arzneistoffinduzierten Leberschädigung. Heutzutage wird die Identifizierung von Hemmern mittels computergestützter Methoden von ligandenbasierten Studien (z.B. QSAR (Quantitative Struktur Wirkungs Beziehung), Machine Learning Methoden) dominiert. Obwohl die resultierenden Modelle als effizient gelten, können sie die Proteineigenschaften nicht miteinbeziehen und daher keine Informationen über den Mechanismus der Hemmung liefern. Diese Details können anhand strukturbasierter Studien untersucht werden, jedoch ist sowohl der Mangel an hochaufgelösten 3D-Strukturen als auch die Polypharmakologie dieser Transporter problematisch.

Diese Dissertation umfasst drei unabhängige Studien, die strukturbasierte Methoden zur Untersuchung der Transporterhemmung im Bereich der Lebertoxizität auf molekularer Ebene vorstellen sowie eine weitere ligandenbasierte Studie über Machine Learning für imbalancierte Datensätze. Die hier präsentierten strukturbasierten Studien beschreiben die Verwendung von Homologiemodellen und molekularem Docking zur Untersuchung der Protein-Liganden-Wechselwirkungen, die dem Mechanismus der Hemmung zugrunde liegen.

In unserer ersten Studie wurde ein Homologiemodell von BSEP erstellt und anschließend strukturunterstützte, dockingbasierte Klassifikationsmodelle zur Vorhersage von BSEP-Inhibitoren entwickelt. Weiters haben wir die Protein Ligand Interaction Fingerprints analysiert, welche spezifische Interaktionen zwischen funktionellen Gruppen der Liganden und den Aminosäuren des Proteins aufzeigen und damit eine Schlüsselrolle in der Ligandenbindung spielen könnten. In der BCRP Studie ermöglichte strukturbasiertes Modeling die Aufklärung der Bindungshypothese von Arylmethyloxyphenylderivaten welche nach experimenteller Validierung zur rationalen Optimierung mit Potenzsteigerung dieser Substanzklasse verwendet werden kann. In der dritten Studie verglichen wir die Interaktionsprofile in der Bindungstasche der P-gp Strukturen von Mensch, Ratte und Maus. Die Resultate zeigen signifikante Überschneidungen bei den interagierenden Aminosäuren der

Bindungstaschen, welche die Übertragbarkeit humaner in vitro P-gp Aktivitätsdaten für die Entwicklung von in silico Modellen zur Vorhersage von Effekten in vitro als auch in vivo bei Nagetieren nahelegen. In unserer ligandenbasierten Studie stellten wir uns der Herausforderung durch unausgewogene Datensätze mit Toxizitätsrelevanz mittels Evaluierung der Performance von sieben unterschiedlichen Meta-Klassifizierern und konnten Empfehlungen zur Auswahl angemessener Klassifizierer in Abhängigkeit desvorliegenden Datensatzes abgeben.

Die Ergebnisse dieser Dissertation verbessern unser Verständnis von Protein-Liganden-Interaktionen auf der molekularen Ebene, inspirieren damit neue Experimente und unterstützen die Extrapolierung molekularer Hypothesen vom Tierversuch zum Menschen und wieder zurück. Darüber hinaus erhöht die Kombination von liganden- und strukturbasierten Methoden die Qualität virtueller Screenings in der Medikamentenentwicklung und verschafft uns detailierte Einblicke in die relevanten molekularen Eigenschaften wichtiger Wechselwirkungen, welche zur Unterstützung der Leitstruktur-Optimierung beitragen.