

DMD # 85167

**Computational prediction of the site(s) of metabolism (SOM) and binding modes of
protein kinase inhibitors metabolised by CYP3A4**

Pramod C. Nair, Ross A. McKinnon and John O. Miners

Department of Clinical Pharmacology (PCN, JOM) and Flinders Centre for Innovation in
Cancer (PCN, RAM, JOM), College of Medicine and Public Health, Flinders University,
Adelaide, Australia.

DMD # 85167

Running title: CYP3A4 mediated SOM prediction of protein kinase inhibitors

Corresponding author: Dr Pramod C. Nair,

Department of Clinical Pharmacology,

College of Medicine and Public Health,

Flinders University,

GPO Box 2100, Adelaide, SA 5001,

Australia.

Email pramod.nair@flinders.edu.au;

telephone 61-8-82043155; fax 61-8-82045114.

Text pages: 33

Number of tables: 6

Number of figures: 4

Number of supplemental table(s): 1

Number of supplemental figure(s): 1

Number of references: 117

Abstract: 250 words

Introduction: 775 words

Discussion: 1419 words

Abbreviations: KI, kinase inhibitor; SOM, site(s) of metabolism; BEC, bromoergocryptine

DMD # 85167

Abstract

Protein kinase inhibitors (KIs), which are mainly biotransformed by CYP3A4-catalyzed oxidation, represent a rapidly expanding class of drugs used primarily for the treatment of cancer. Ligand- and structure- based methods were applied here to investigate whether computational approaches may be employed to predict the site(s) of metabolism (SOM) of KIs, and to identify amino acids within the CYP3A4 active site involved in KI binding. A dataset of the experimentally determined SOMs of 31 KIs known to undergo biotransformation by CYP3A4 was collated. The structure based (molecular docking) approach employed three CYP3A4 X-ray crystal structures to account for structural plasticity of this enzyme. Docking pose and SOM predictivity were influenced by the X-ray crystal template used for docking and the scoring function used for ranking binding poses. The best prediction of SOM (77%) was achieved using the substrate (bromoergocryptine) bound X-ray crystal template together with the PMF-score. Binding interactions of KIs with CYP3A4 active site residues were generally similar to those observed for other substrates of this enzyme. The ligand-based molecular superposition approach, using bromoergocryptine from the X-ray co-crystal structure as a template, poorly predicted (42%) the SOM of KIs, although predictivity improved to 71% when the docked conformation of sorafenib was used as the template. Among the web-based approaches examined, all web servers provided excellent predictivity, with Xenosite predicting the SOM of 87% of the dataset molecules. Computational approaches may be used to predict the SOM of KIs, and presumably other classes of CYP3A4 substrates, but predictivity varies between methods.

DMD # 85167

Introduction

Protein kinases catalyze the phosphorylation of proteins, an essential post-translational signalling mechanism that influences numerous cellular processes. Abnormal regulation of kinase signalling cascades is known to cause a myriad of diseases, including cancer. For instance, mutation of kinase receptors (e.g. the EGFR and VEGFR families) and downstream signalling proteins (e.g. RAS, RAF, MEK, MAPK, etc.) are associated with constitutive pathway activation, uncontrolled cell proliferation, and malignancy (Rowland et al. 2017). Protein kinase inhibitors (KIs) compete with ATP at the ATP-binding site of the kinase, thereby preventing phosphorylation of the target protein (Duckett and Cameron, 2010). Consequently, strategies to target both the active and inactive conformations of protein kinases have been used to develop inhibitors. However, specific targeting to a single kinase has proved challenging. Nonetheless, KIs have led to significantly improved outcomes for several difficult to treat malignancies such as renal cell and hepatocellular carcinomas (e.g. sorafenib, axitinib), metastatic melanoma (e.g. vemurafenib, dabrafenib), HER-2 positive breast cancer (e.g. lapatinib) and pancreatic cancer (e.g. sunitinib) (Hochhaus et al., 2017; Ferraro and Zalcberg, 2014; Druker et al., 2006; Zhu et al., 2011; Chapman et al., 2011; Burstein et al., 2008; Faivre et al. 2017).

As of November 2016, the FDA had approved 31 KIs for clinical use (Zhang et al. 2017; Rowland et al., 2017), although several others have been approved since that time (e.g. brigatinib, neratinib) and more are in clinical development (Wu et al., 2015 and 2016). It is thus anticipated that the number of KIs available for the treatment of cancer and certain other diseases (e.g. rheumatoid arthritis, idiopathic pulmonary fibrosis) will increase in the future.

The metabolism of KIs occurs primarily in the liver and gastrointestinal tract. Although multiple enzymes have been implicated in KI biotransformation (Rowland et al., 2017),

DMD # 85167

CYP3A4 is the main enzyme involved in the metabolism of the majority of drugs in this class (**Table 1**). Consequently, CYP3A4 activity may influence KI clearance, and hence dosage and response. Moreover, clinically significant drug-drug interactions have been reported for several KIs when simultaneously dosed with CYP3A4 inhibitors or inducers (Duckett and Cameron, 2010; Keller et al., 2018). In addition, the CYP3A4-catalyzed biotransformation of several KIs (e.g. erlotinib, dasatinib, gefitinib, and lapatinib) is known to result in the formation of reactive metabolites capable of covalently binding to cellular macromolecules (Duckett and Cameron, 2010). Despite the importance of CYP3A4 in KI metabolism, the binding of these drugs within the CYP3A4 active site, and how this determines the site of metabolism (SOM), has received little attention. Numerous computational approaches have been developed for SOM prediction. The *in silico* methods provide a useful starting point for the characterization of drug metabolism pathways and complement experimental data (Tarcsay and Keseru, 2011; Kirchmair et al., 2012; Ford et al., 2015).

CYP3A4 is a major drug metabolizing enzyme, which is responsible for the metabolism of approximately 30% of clinically used drugs (Rendic and Guengerich 2015; Zanger and Schwab 2013). This enzyme is abundantly expressed in liver, and in addition to KIs, metabolizes drugs from numerous therapeutic classes, including benzodiazepines (e.g. alprazolam, midazolam), macrolide antibiotics (e.g. clarithromycin, erythromycin), calcium channel blockers (e.g. diltiazem, verapamil), carbamazepine, and HIV-1 protease inhibitors (e.g. indinavir, saquinavir). A feature of CYP3A4 substrates is the diverse range of sizes (molecular mass) and shapes. CYP3A4 has a large active site, ranging in volume from 950 Å³ to 2000 Å³ in the absence and presence of bound ligand (Yano et al., 2004; Ekroos & Sjögren, 2006). Not surprisingly, X-ray crystal structures demonstrate that CYP3A4 displays significant structural plasticity in order to accommodate structurally diverse ligands (Sevrioukova and Poulos 2013 and 2017). Almost 40 X-ray crystal structures of human CYP3A4 have been published to date.

DMD # 85167

Three of the structures were crystallized without a bound ligand (1W0E, 1TQN, 4I3Q), while five structures have been co-crystallized with a substrate (viz., progesterone, bromoergocryptine (2 independent structures), midazolam, and erythromycin A) and 31 with an inhibitor, mostly ritonavir analogues. Despite advances in CYP3A4 structure-function, there is limited structural information on how KIs are recognized by this enzyme, and which domains within the active site are associated with substrate recognition and binding.

The objectives of the present study were to: (i) collate SOM data for kinase inhibitors metabolized by human CYP3A4 from published in vivo and in vitro studies; (ii) investigate whether molecular docking of KIs in selected CYP3A4 X-ray crystal structure(s) and ligand-based alignment approaches using the structure co-crystallized with bromoergocryptine may be used as templates to predict the SOM of KIs; (iii) evaluate existing Web-based methods for SOM prediction of KIs metabolized by CYP3A4; and (iv) identify amino acids important for KI binding within the CYP3A4 active site from the docking analyses.

Materials and Methods

KI Dataset

A dataset of 31 marketed KIs was collated (**Figure 1**). Information on the SOMs of the KIs metabolized by human CYP3A4 was obtained from the published literature and from Product Information (Alecensa Product Information, 2017; Bosulif Product Information, 2014; Cabometyx Product Information, 2018; Caprelsa Product Information, 2013; Cotellic Product Information, 2016; Glivec Product Information, 2001; Ibrance Product Information, 2017; Iclusig Product Information, 2014; Imbruvica Product Information, 2015; Inlyta Product Information, 2017; Iressa Product Information, 2003; Jakavi Product Information, 2013; Lenvima Product Information, 2016; Mekinist Product Information, 2014; Nexavar Product Information, 2006; Sprycel Product Information, 2007; Stivarga Product Information, 2013;

DMD # 85167

Sutent Product Information, 2006; Tafinlar Product Information, 2013; Tagrisso Product Information, 2016; Tarceva Product Information, 2006; Tasigna Product Information, 2008; Tykerb Product Information, 2007; Votrient Product Information, 2010; Xalkori Product Information, 2013; Xeljanz Product Information, 2015; Zelboraf Product Information, 2012; Zydelig Product Information, 2015; Zykadia Product Information, 2016; Morcos et al., 2017; Abbas and Hsyu 2016; Smith et al., 2014; Lacy et al., 2015; Takahashi et al., 2016; Johnson et al., 2015; Bershas et al., 2013; Christopher et al., 2008a; Ling et al., 2006; McKillop et al., 2004; Scheers et al., 2015; Gschwind et al., 2005; Castellino et al., 2012; Dubbelman et al., 2016; Dickinson et al., 2016; Deng et al., 2013; Ye et al., 2017; Gerisch et al., 2018; Shilling et al., 2010; Minami et al. 2008; Speed et al., 2012; Goldinger et al., 2015; Dowty et al., 2014; Ho et al., 2014; Ding J et al. 2014). Additionally, data from in vitro studies with human liver microsomes, hepatocytes and/or recombinant CYP3A4, where available, were used for confirmation of metabolite formation and the role of CYP3A4 in metabolite formation (Stopfer et al., 2012; Nakagawa et al., 2017; Zientek et al., 2016; Lawrence et al., 2014; Christopher et al., 2008b; Wang et al., 2008; Jin et al., 2015; Inoue et al., 2014; Ghassabian et al., 2012). The major SOM was identified as the predominant primary metabolite(s), determined from in vivo data (faecal/urine excretion and/or plasma concentrations). In cases where multiple major sites of metabolism were reported, all sites were considered. Precise sites of metabolism were not identified experimentally for some of the dataset molecules (alectinib, ceritinib, cobimetinib, ponatinib); in these cases, the SOM(s) was taken as a substructure within the molecule (e.g. morpholine, phenyl rings, etc.) where biotransformation was reported (**Figure 1**).

CYP3A4 appears not to be the main enzyme involved in afatinib and nintedanib metabolism. The major metabolites of afatinib are apparently formed non-enzymatically (by Michael addition), although in vitro studies indicate some involvement of FMO3 and CYP3A4 in the metabolism of this drug (Giotrif Product Information, 2013). Nintedanib is primarily

DMD # 85167

metabolized by esterases and UGT1A1, with a lesser contribution of CYP3A4 (Roth et al. 2015; Ofev Product Information, 2015).

CYP3A4 Structure and Molecular Modelling

Three X-ray crystal structures were selected for docking studies to take into account the plasticity of the CYP3A4 protein. These included: the 5VCC structure, which was obtained from a codon-optimized cDNA and resolved to 1.7 Å (the highest resolution achieved thus far for a CYP3A4 protein); the 3UA1 structure, co-crystallized with bromoergocryptine (a substrate); and the ritonavir (an inhibitor) bound structure, 5VCO. The 5VCC structure was co-crystallized with glycerol and ethylene glycol (from crystallization medium), but not with a drug substrate or inhibitor. Thus, it is subsequently referred as the ‘unliganded’ structure. The CYP3A4 structures were prepared for docking studies by adding unresolved residues/atoms, H atoms and assigning Kollman all-atom charges to the pdb structures using the Biopolymer module of SYBYL X-2.1

The three-dimensional coordinates (sdf format) of dataset molecules were obtained from the Pubchem server (<https://pubchem.ncbi.nlm.nih.gov/>). The molecules were imported into SYBYL (version X-2.1, CERTARA, Princeton, NJ) and assigned a protonation state at pH 7.4, calculated using the Calculator Plugins implemented in ChemAxon (Marvin 16.6.20). The structures were additionally assigned Gasteiger–Huckel partial atomic charges (Gasteiger and Marsili 1980) and energy minimized by Powell's conjugate gradient method using the Tripos force field (Powell 1977). A minimum energy difference of 0.001 kcal/mol was set as the convergence criterion. All molecular modelling was performed using SYBYL installed on a Linux workstation running the Red Hat Linux 6.9 operating system.

Molecular Docking Calculations

DMD # 85167

Docking experiments were conducted using the Surflex-Dock docking suite (Jain 2003), which combines Hammerhead's empirical scoring function with a molecular similarity method (morphological similarity) to generate putative poses of ligand fragments. First, an idealized binding site is generated from the protein structure by placing three different types of molecular fragments (CH₄, C=O, and N-H) into putative sites in multiple positions, which are then optimized for binding to the protein. High-scoring nonredundant fragments, collectively referred to as the 'protomol', serve as a target for alignment of ligands (or ligand fragments) based on molecular similarity. The ligand to be docked is fragmented, and each fragment is conformationally searched and aligned to the protomol to yield poses that maximize molecular similarity to the protomol. The aligned fragments are scored and pruned on the basis of the scoring function and the degree of protein interpenetration (Jain 2003).

The putative binding-site protomol for ligand docking was calculated for the three CYP3A4 structures described previously. Ligands and water molecules were removed from the structures before protomol creation and docking calculations. The protomol for the 5VCC structure was defined automatically, whereas the respective co-crystallized ligands were used as a reference for generating the protomol with the 5VCO and 3UA1 structures. The active site, which includes the heme porphyrin moiety, was identified as the principal site for docking by the protomol. The docking module Surflex-Dock GeomX (SFXC) was used for all docking calculations.

The aim of the docking studies was to assess whether the first ranked pose identified by each of the scoring functions can accurately predict the experimentally determined major SOMs for KIs biotransformed by CYP3A4 (i.e. only the top ranked binding pose obtained from different scoring functions was evaluated for prediction of the experimentally identified SOM, which essentially tests scoring accuracy). Surflex-Dock predicts the ligand binding poses based on

DMD # 85167

the total score (as the default). However, binding poses based on other scoring functions (D-score, PMF-score, G-Score, ChemScore, and CSCORE) can also be evaluated using this program (Kuntz et al. 1982; Muegge and Martin, 1999; Jones et al. 1997; Eldridge et al. 1997; Clark et al. 2002; SYBYL X-2.1, CERTARA, Princeton, NJ).

The D-score is derived from the molecular docking algorithm DOCK, which is based on the charge and van der Waals interactions between the protein and the ligand (Kuntz et al. 1982). Here, the van der Waals energy is calculated with a Lennard-Jones 12-6 potential and the electrostatic energy is calculated with the Coulombic equation. The potential of mean force (PMF)-score is drawn from the work of Muegge and Martin (1999), who analyzed a large set (697) of protein–ligand complexes drawn from the Protein Data Bank and developed a set of Helmholtz free energies of interactions for protein-ligand atom pairs. The ChemScore is an empirical scoring function that includes terms for hydrogen bonding, metal-ligand interactions, lipophilic contacts, and rotational entropy, along with an intercept term (Wang et al. 2008). This scoring function was originally calibrated by reproducing the measured dissociation constants of 82 protein–ligand complexes (Eldridge et al. 1997). The G-Score is based on the scoring function implemented in the molecular docking program GOLD and is the sum of three terms, namely protein–ligand complexation, hydrogen bonding, and internal energy (Jones et al. 1997). In this scoring function the complexation term is calculated with a reparametrized Lennard-Jones 8-4 potential, while the hydrogen bonding term is the sum of the individual energies from all the donor–acceptor pairs. The internal energy of the ligand includes a dispersion–repulsion energy and a torsional energy calculated according to the Tripos force field. The G-score was validated by testing the scoring function of a set of 100 protein–ligand complexes. Total Score is the Surflex-Dock scoring function, which is an empirically derived scoring function based on the binding affinities of protein-ligand complexes from a set of X-ray crystal structures. The scoring function primarily comprises hydrophobic, polar, entropic,

DMD # 85167

and solvation terms (Jain 1996). The consensus score (CScore) allows integration of different scoring functions for ranking the affinity of ligands bound to the active site of a receptor. The strengths of individual scoring functions combine to produce a consensus that is more robust and accurate than any single function for evaluating ligand-receptor interactions (Clark et al. 2002). In this study, the CScore was derived based on the consensus of D-Score, PMF-Score, ChemScore and G-Score.

The binding poses of the docked substrates were considered positive if the experimentally known SOM occurred within 6 Å from the Fe atom of the heme moiety, as reported by others (de Graff et al. 2007; Vasanathanathan et al. 2010). Ligands docked away from the active site or with a SOM distance > 6 Å from the Fe atom were considered to be in an unproductive pose and the docking was classified as unsuccessful.

Molecular superposition by morphological similarity

Besides the molecular docking method, we also examined the molecular superposition approach to predict the SOM of the dataset molecules. The structural superposition (or overlay) of molecules was undertaken using the Surflex-Sim program (Jain, 2000, 2004), which utilizes the morphological similarity approach to generate alignments of molecules. Similarity is defined as a Gaussian function of the differences in the molecular surface distances of two molecules at weighted observation points on a uniform grid. The computed surface represents distances to the nearest atomic surface and distances to donor and acceptor surfaces. Bromoergocryptine (BEC) from the 3UA1 CYP3A4 X-ray structure was used as a template for the molecular superpositioning of the dataset molecules, as other substrate bound CYP3A4 structures were considered suboptimal for the comparison of the aligned dataset molecules with respect to the CYP3A4 structure. In particular, in one structure the substrate (progesterone) bound at a site distant from heme, and binding of the other structure (midazolam) caused

DMD # 85167

substantial alteration in the active site. Neither of these substrates can be used as templates as they cannot be aligned efficiently for SOM prediction. In addition to BEC, we also employed the docked conformation of three KIs with differing chemical structures, molecular masses and shapes, namely tofacitinib, sorafenib and lapatinib (see Results and Discussion) as templates for the molecular superpositioning of dataset molecules.

Web-based SOM prediction

In addition to molecular docking and molecular superposition, three web-servers were assessed for SOM prediction: MetaPrint2D (www-metaprint2d.ch.cam.ac.uk/metaprint2d/), Regioselectivity (RS)-WebPredictor 1.0 (<http://reccr.chem.rpi.edu/Software/RS-WebPredictor/>) and Xenosite (<http://swami.wustl.edu/xenosite/>). MetaPrint2D utilizes a database of atom environments based on the biotransformation of xenobiotics from a curated database extracted from the literature (Carlsson et al., 2010). SOM prediction is achieved by calculating the atom environments within a molecule, which are then searched and compared for similar environments in the molecular database. An occurrence ratio is calculated for each atom in the molecule by measuring how often a similar environment has been found at a reaction centre relative to how many times it has been observed in total. Ratios generated in this manner represent the relative likelihood of metabolism occurring at each atom, which is scaled and normalized to an occurrence ratio of one. In MetaPrint2D, the predicted metabolic sites/atoms are represented by a colour code, indicating the probability of biotransformation. The most probable SOM is shown in red and the least probable in grey, with probability values ranging from 0-1 [red (0.66-1), orange (0.33-0.66), green (0.15-0.33), white (0.00-0.15), and grey (little/no data)]. Thus, multiple probable major SOMs may be identified for some molecules. Where five-colour codes are not noted for a molecule, the major predicted SOM is based on the probability scale of red to grey. For example, if the site(s) of metabolism were predicted by green and grey codes only (e.g. dabrafenib, ibrutinib), then the positions coded green would be considered as the major predicted SOM based on the probability scale.

DMD # 85167

RS-WebPredictor is a QSAR based method for predicting the SOM by specific CYP enzymes that takes into account the regioselectivity of substrate metabolism (Zaretski et al., 2012). The models used for predictions are trained using a combination of topological, quantum chemical (atom-pair based) and ‘SMARTCyp reactivities’ applied to substrate sets (i.e. substrate databases) of individual CYP enzymes (1A2, 2A6, 2B6, 2C8, 2C9, 2C19, 2D6, 2E1, and 3A4). Models are generated using MIRan, a support vector machine (SVM)-like ranking and multiple instance learning method specifically designed to correctly rank metabolophores associated with oxidized SOM(s) over metabolophores of nonoxidized SOMs on the same substrate.

Xenosite uses a machine learning approach based on neural networks for predicting the SOM of substrates of CYP enzymes (Zaretski et al., 2013). Each atom in a molecule is associated with a vector of numbers, with each number encoding a chemical property of the SOM. This approach is somewhat similar to RS-WebPredictor, which also employs a combination of several descriptors (topological, quantum chemical (QC), SMARTCyp reactivity, refined QC, molecule-level (MOL) and fingerprint similarity), which are analyzed by machine-learning algorithms. The server predicts a score for each atom, which ranges from 0 to 1, that can be interpreted as a probability or statistical likelihood of metabolism occurring at a particular atom. The major SOM of the dataset molecules is identified based on a colour coded probability scale.

Results

Kinase inhibitors and their sites of metabolism

In general, KIs are metabolized by CYP3A4 to produce numerous products, several of which are classified as ‘major’ metabolites. Hence, in many cases multiple major SOMs were considered. A precise site of primary biotransformation was unavailable for some drugs; for example, the morpholine ring in alectinib and the piperidine ring in crizotinib, which have

DMD # 85167

multiple potential oxidation sites (**Figure 1**). These regions were considered for SOM predictions. Approximately 50 primary biotransformation pathways for the 31 KI substrates were collated (**Figure 1**).

Molecular Docking

Given the structural flexibility of CYP3A4, molecular docking was performed using three CYP3A4 X-ray crystal structures as templates. The binding poses of the 31 KIs based on the distance criterion (viz. within 6 Å of the heme iron) for the prediction of SOM were evaluated using six scoring functions. In total, 558 binding poses for the dataset molecules were assessed, which comprised of 18 individual poses of each KI obtained from docking in the three different templates. The number of dataset molecules predicted correctly by each scoring function in the X-ray crystal structures is shown in **Tables 2a-c** and **Figure 2**.

The PMF-score and CScore predicted the experimentally known major SOM of 23 and 21 of the 31 dataset molecules, respectively, docked in the unliganded X-ray crystal structure 5VCC (**Figure 2, Table 2a**). By contrast, the total score and G-score predicted the major SOM of 18 dataset molecules. The docked poses ranked by D-score were poorer still, predicting the SOM of 16 KIs. Docking in the 5VCC (unliganded) structure failed to generate binding poses (using four or more scoring functions) of 9 KIs (viz. afatinib, alectinib, crizotinib dabrafenib, imatinib, nintedanib, palbociclib, ruxolitinib and sunitinib) that predicted the experimentally determined SOM.

SOM prediction using the 3UA1 structure (CYP3A4 co-crystallized with bromoergocryptine) provided better outcomes compared to docking in the 5VCC structure. The best prediction was obtained with the PMF-score and total score, with each scoring function predicting the SOM of 24 (77%) dataset molecules (**Figure 2, Table 2b**). As with docking in the 5VCC structure, the D-score performed poorly in ranking binding poses in the 3UA1 structure with the

DMD # 85167

experimentally known SOM. The majority of the failed predictions from ranking with the D-score either docked the KI distant from the active site or showed a binding orientation where the major SOM was > 6 Å from the heme. Molecular docking in the 3UA1 structure resulted in fewer failed predictions. The major SOM of only 6 KIs (afatinib, crizotinib dabrafenib, imatinib, palbociclib, and sunitinib) was predicted inaccurately by four or more scoring functions.

Docking in the CYP3A4 structure (5VCO) co-crystallized with ritonavir provided slightly better predictivity compared to the 5VCC structure, but not SOM predictions obtained from docking in the 3UA1 structure (**Figure 2, Table 2c**). Use of 5VCO as the template failed to predict the SOM of 9 KIs (bosutinib, crizotinib, dabrafenib, imatinib, nintedanib, osimertinib, palbociclib, tofacitinib and vandetanib). The PMF-score predicted the SOM of 21 dataset molecules, whereas the G-score and CScore each correctly predicted the SOM of 20 KIs. In contrast to the 3UA1 and 5VCC structures, the D-score performed relatively better and correctly predicted the SOM of 19 dataset molecules. The docked conformations of the dataset molecules in the 5VCC structure were more widely distributed due to the larger active site compared to the substrate-bound structure. Of note, prediction of the SOM of 4 KIs (crizotinib, dabrafenib, imatinib, and palbociclib) consistently failed with all three X-ray crystal templates. Overall, the PMF-score provided best predictivity of the known SOMs of the dataset KIs with all three CYP3A4 structures.

The ratio (expressed as a percentage) of the number of docked poses with a specific SOM relative to the total number of successful docked poses obtained from the six different scoring functions was calculated to identify the most probable SOM for each KI (**Figure 1**). All binding orientations of KIs were taken into account for calculating the ratio, with the exception of those that were classified as unproductive (i.e. docked at a distant site or SOM > 6 Å from heme).

DMD # 85167

The ratios correctly predicted the experimentally known major SOM of 24 (77%) KIs (**Figure 1**). Furthermore, the docking based SOM prediction was useful for predicting the metabolic sites of those molecules whose experimental SOM was poorly defined (e.g. ceritinib, cabozantinib and ponatanib). Several binding poses with multiple SOMs were predicted for osimertinib and vandetanib, both of which are known to undergo extensive metabolism mediated by CYP3A4 (**Figure 1**).

To investigate the role of individual amino acids in KI binding, CYP3A4-ligand binding interactions were explored based on a consensus docked pose (maximum number of poses with a common SOM) from different scoring functions with the 3UA1 (substrate bound) structure. The docked complexes were analyzed by the protein–ligand interaction profiler (PLIP) server (<https://projects.biotec.tu-dresden.de/plip-web/plip/index>) to classify the types of interactions between the dataset molecules and individual amino acids within the CYP3A4 active site (Salentin et al. 2015). Representative docked poses (bostunib, dasatinib, ibrutinib, and sorafenib) are shown in **Figure 3**. For each of the four substrates, the SOM (represented as a sphere(s)) is located within 5 Å of the heme Fe atom. By way of example, key residues within the CYP3A4 active site in close proximity to bostunib are: Arg212 and Arg372 (H-bonding); Glu374 (salt-bridge interaction with the positively charged N atom of the piperazine ring); Phe57 (aromatic interaction); Phe108 (aromatic pi stacking interaction).

Most KIs are predicted to bind primarily via a hydrophobic interaction(s) within the CYP3A4 active site. **Table 3** shows CYP3A4 active site residues that are associated with KI binding. Phe57 and/or Phe215 are associated with the binding of approximately 80% of the dataset molecules via hydrophobic/aromatic contacts (e.g. **Figure 3**). Pi-cation interactions are observed most commonly with Arg106 (e.g. cabozantinib, ceritinib), but may also involve Arg105 (e.g. alectinib, dasatinib) and Arg212 (e.g. nilotinib and ruxolitinib). The dataset

DMD # 85167

molecules form H-bonds with a range of residues, most importantly Ser119, Arg212, Thr224, Arg372, and Glu374. Several kinase inhibitors contain non-aromatic rings, such as piperidine (e.g. ceritinib, cobimetinib) and piperazine (e.g. bosutinib, ponatinib), or an aliphatic tertiary amine group(s) (e.g. afatinib, osimertinib), which are predicted to exist in a protonated state at physiological pH. The catalytically favorable docked poses of several of these protonated KIs show a charge (salt-bridge) interaction(s) with Asp76 and/or Glu374 (**Table 3**).

Ligand superposition by morphological similarity

Molecular overlay is a useful approach for predicting the structural similarity of ligands and has been used to predict the SOM of substrates of several CYP enzymes (Sykes et al., 2008; de Bruyn Kops et al., 2017). The CYP3A4 substrate bromoergocryptine (BEC), which was co-crystallized in the 3UA1 structure, was initially employed as the template for ligand superposition. Importantly, BEC is the only substrate bound within the active site of a CYP3A4 X-ray crystal structure in a catalytically favorable orientation and which has a molecular mass similar to many of the dataset KIs. However, the known SOM was predicted for only 13 of the dataset molecules using BEC as the template for structural overlay (**Table 4**).

As there was no other suitable co-crystal substrate that could be used for ligand superposition, we hypothesized that SOM prediction might be improved if a template from the dataset was selected based on its docked conformation. First, we evaluated the docked conformation of sorafenib as the template for the structural overlay. The docked conformation of sorafenib is positioned at a catalytically favorable distance such that the major metabolic sites, N-oxide and hydroxylation (**Figure 1**), are 3.3 and 5.4 Å from the Fe atom of the heme. Ghassabian et al. (2018) similarly reported that the distance of the site of N-oxidation from the heme (Fe atom) was 3.3 Å when sorafenib was docked in another unliganded CYP3A4 structure (ITQN). Use of sorafenib as the template for ligand superposition predicted the SOM of 22 of the 31 dataset

DMD # 85167

molecules (prediction accuracy of ~71%) (**Table 4**). Alignment of the representative substrates cabozantinib, dasatinib, lenvatinib, and regorafenib with sorafenib is shown in **Figure 4**. Interestingly, the SOMs of dabrafenib and imatinib, which was not predicted by docking in the X-ray crystal structures, were similarly not predicted using the ligand superposition method with sorafenib as the template. On the other hand, imatinib overlaid well with the respective SOM positions (separated only by 1.2 Å) using lapatinib as the template. Overall predictivity with lapatinib as the template was similar to that with sorafenib (**Table 4**). SOM prediction was also evaluated using the structurally dissimilar and lower molecular mass KI, tofacitinib. However, the dataset molecules showed a poor overlay (Supplementary **Figure S1**) with this template; the known major SOM was predicted correctly for only 10 KIs (**Table 4**).

Web-based SOM prediction

Three web-based methods were used to predict the SOM of dataset molecules. MetaPrint2D, RS-WebPredictor, and Xenosite correctly predicted the major SOM of 24, 26 and 27 of the 31 KIs, respectively (Supplementary **Table S1**). MetaPrint2D failed to predict the major SOM for crizotinib, lapatinib, lenvatinib, and sunitinib, whereas both RS-WebPredictor and Xenosite correctly predicted the major SOM of these compounds. Additionally, MetaPrint2D failed to predict the known SOM for tofacitinib. On the other hand, MetaPrint2D was the only webserver that correctly predicted the major SOM for dasatinib and pazopanib. The performance of RS-WebPredictor and Xenosite was similar, correctly predicting the SOM of 84% and 87% of the KIs, respectively. All the webserver failed to predict the SOM of dabrafenib, as also noted for the docking and ligand superposition methods.

Discussion

Metabolism provides a clearance mechanism for approximately 75% of clinically used drugs. Although biotransformation primarily functions as a deactivation and detoxification

DMD # 85167

mechanism, metabolism may result in the formation of compounds with enhanced pharmacological, physiological and/or toxicological properties. Thus, metabolism is a significant consideration in drug efficacy and safety. Accordingly, metabolite characterization is of fundamental importance in drug discovery and development. Site(s) of metabolism (SOM) prediction tools provide a complementary approach to in vitro and in vivo approaches, for example metabolite profiling of structural analogues prior to synthesis (Kirchmair et al. 2015). In this study, several in silico approaches were compared to assess their ability to predict the SOM of KIs, which are mainly, or in part, metabolized by CYP3A4. The studies further aimed to utilize computational approaches to identify binding modes and interactions of KIs with specific amino acids in the CYP3A4 active site.

The SOM prediction of substrates using docking approaches is not straightforward due to the known plasticity of CYP proteins that may adopt substantially different conformations upon ligand binding (Nair et al. 2016). Hence, three X-ray crystal structures were investigated as templates for molecular docking to account for the molecular flexibility. Studies of the docking of the dataset molecules within the CYP3A4 active site show that the choice of X-ray crystal template is an important consideration. Docked ligands overlapped well with the structure of bound BEC in the 3UA1 co-crystal structure. By contrast, the binding modes of ligands docked in the 5VCC and 5VCO structures only partially overlapped with bound BEC. Notably, of the three X-ray crystal structures used as templates, the 3UA1 structure has a relatively constricted active site that is narrower and lower in volume (779 Å³) compared to both the unliganded (5VCC, 823 Å³) and the inhibitor bound (5VCO, 954 Å³) structures (<http://sts.bioe.uic.edu/castp/index.html?2011>). The data indicate that multiple ligand-bound X-ray crystal templates may need to be assessed to identify the most suitable structure for docking studies.

DMD # 85167

As noted above, the docked KIs aligned well with BEC in the active site of the 3UA1 structure, with the SOM of ~77% of the dataset molecules located within ~3 Å of the SOM of BEC. Six scoring functions were evaluated to ascertain whether the top ranked binding poses could correctly predict the SOM of the dataset molecules. Predictions based on the scoring functions were generally dependent on the X-ray crystal template employed, although, the PMF-score was superior among the scoring functions evaluated in ranking the binding poses of KIs with the experimentally known SOM. By contrast, the D-Score was relatively poor for predicting the correct SOM in all X-ray crystal templates. The ranking of binding poses by different scoring functions for SOM prediction improved when the ligand bound crystal structures (e.g. 3UA1) was used as a template for docking studies, except with the D-score. Overall, however, the docking approach provided very good SOM predictivity for most of the dataset molecules.

Docking in the 3UA1 structure in combination with the various scoring functions failed to predict the SOM of only four KIs (afatinib, dabrafenib, imatinib, and sunitinib) (**Table 2a**). No obvious molecular fingerprint could be identified for these molecules. The inability of the docking approach to predict the SOMs could be due to multiple reasons. 1) The ligands might bind in a unique conformation that may involve reshaping of the CYP3A4 active site or by changes in the conformation of amino acid side chains, as noted in multiple CYP3A4 X-ray crystal structures (Sevrioukova and Poulos 2013 and 2017). In these cases, the scoring functions may fail to rank the pose(s). 2) Where KIs are modified non-enzymatically or metabolism by CYP3A4 is a minor metabolic pathway, binding may be weak. For instance, afatinib, a poor CYP3A4 substrate, was incorrectly ranked in the 3UA1 and 5VCC structures by more than 4 scoring functions in each structure. 3) As KIs are a new drug class, metabolic pathways for some drugs may not be completely characterized. Thus, currently unidentified SOM(s) may exist. For example, although methyl hydroxylation of dabrafenib (**Figure 1**) is a

DMD # 85167

known metabolic pathway, this SOM was not predicted by any of the *in silico* approaches investigated here, which suggests the possibility of an alternate site of metabolism.

An additional benefit of the docking approach is that it provides insights into the structure of the ligand-enzyme complex. Docking in the active site of the CYP3A4 structures identified a number of residues important for KI binding particularly, Phe57, Asp76, Arg105, Arg106, Ser119, Arg212, Phe215, Thr224, Arg372, and Glu374 (**Table 3**). The binding interactions of the docked dataset molecules were consistent with the BEC (substrate) bound CYP3A4 X-ray structure and, more generally, those reported to be important for CYP3A4 ligand recognition. For example, based on the 3UA1 X-ray crystal structure the lysergic acid moiety of BEC is sandwiched between the side chains of Arg106 and Phe215 due to hydrophobic contacts. Similarly, the Thr224 and Arg212 side chains are known to form H-bonds with the lysergic acid group of BEC, while Ser119 has been shown to be associated with H-bonding/polar interactions in multiple CYP3A4 X-ray crystal structures (Kaur et al., 2016; Sevrioukova and Poulos 2013 and 2017).

The ligand superposition approach evaluated in this study performed poorly (< 50% prediction of SOM) using BEC as the template. This contrast to the success of the alignment-based approach reported previously from this laboratory for CYP2C9 substrates and UGT2B10 inhibitors (Sykes et al., 2008; Pattanawongsa et al., 2016). It should be noted, however, that the CYP2C9 substrate flurbiprofen, used for molecular overlay, exhibited marked structural similarity with the CYP2C9 dataset molecules (Sykes et al., 2008), most of which contained an acidic functional group (e.g. the carboxylate of NSAIDs) that forms a salt-bridge with Arg108 (Wester et al., 2004). However, substrates metabolized by CYP3A4 are more diverse, both chemically and structurally (Rendic and Di Carlo 1997), and KIs are no exception in this regard. Thus, the selection of a suitable template for ligand superposition and SOM prediction

DMD # 85167

using this approach is less straightforward. The structure of BEC is comparatively more rigid with significant chemical differences compared to KIs (e.g. bosutinib, dasatinib, gefitinib), which influences alignment and overlay between the template and dataset molecules. We therefore additionally explored the novel approach of employing the docked conformation of selected KIs to predict the SOM of other compounds in the dataset. Use of sorafenib as the template correctly predicted the SOM of ~71% of the dataset molecules. It is acknowledged, however, that this approach may be dataset sensitive given the structural diversity of CYP3A4 ligands. Nevertheless, the data strongly suggest that SOM prediction can be improved by using an appropriate template molecule from the dataset under investigation when there are limited experimental data relating to substrate binding mode.

The web-based approaches evaluated in this study utilized three online servers. The servers predicted the major SOM for 24 to 27 KIs. RS-WebPredictor, which is a predecessor of Xenosite, predicted the SOM of ~84% of the dataset molecules. The performance of Xenosite, was marginally superior, predicting ~87% of the dataset molecules. These data suggest that the web-based methods are useful for SOM prediction, even for relatively large, complex molecules such as KIs.

In summary, data presented here demonstrate that the docking pose and SOM predictivity are influenced by both the CYP3A4 X-ray crystal template employed for docking and the scoring function used for ranking the poses, although the PMF-score consistently proved superior for ranking the binding poses of KIs against the experimentally determined SOM. The results underscore the importance of accounting for CYP protein plasticity in ligand docking studies and indicate that the performance of scoring functions may vary for datasets of structurally diverse ligands. Nevertheless, very good predictivity of the SOM of the KIs was achieved for several protein template – scoring function combinations (e.g. 3UA1/PMF-score). The study highlighted the challenges with the ligand superposition approach where the lack of availability

DMD # 85167

of a suitable substrate co-crystal template limits the effectiveness of this method. However, the use of a representative dataset substrate as the template for ligand superposition may provide an alternative approach. While the use of web-based methods predicted the SOM of the dataset KIs extremely well, unlike ligand docking this approach does not provide insights into substrate binding within the CYP3A4 active site and the contribution of individual amino acids to substrate binding. The data indicate that computational approaches may be used to predict the SOM of KIs, but predictivity is dependent on the approach adopted. Similar considerations would be expected to be important for other classes of CYP3A4 substrates.

DMD # 85167

Acknowledgments

PCN acknowledges Flinders Foundation and Flinders Centre for Innovation in Cancer (FCIC) for Early Career Research Funding. RAM is a recipient of a Beat Cancer Professorial Fellowship from Cancer Council SA.

DMD # 85167

AUTHORSHIP CONTRIBUTIONS

Participated in research design: Nair, Miners, McKinnon

Conducted experiments: Nair

Performed data analysis: Nair, Miners, McKinnon

Wrote or contributed to writing of the manuscript: Nair, Miners, McKinnon

References

- Abbas R and Hsyu P-H (2016) Clinical Pharmacokinetics and Pharmacodynamics of Bosutinib. *Clinical Pharmacokinetics* **55**:1191-1204.
- Alecensa (2017) Roche Products Pty Limited, Dee Why, NSW, Australia.
- Bershas DA, Ouellet D, Mamaril-Fishman DB, Nebot N, Carson SW, Blackman SC, Morrison RA, Adams JL, Jurusik KE, Knecht DM, Gorycki PD, and Richards-Peterson LE (2013) Metabolism and Disposition of Oral Dabrafenib in Cancer Patients: Proposed Participation of Aryl Nitrogen in Carbon-Carbon Bond Cleavage via Decarboxylation following Enzymatic Oxidation. *Drug Metabolism and Disposition* **41**:2215-2224.
- Bosulif (2014) Pfizer Australia Pty Ltd, Sydney, NSW, Australia.
- Burstein HJ, Storniolo AM, Franco S, Forster J, Stein S, Rubin S, Salazar VM, and Blackwell KL (2008) A phase II study of lapatinib monotherapy in chemotherapy-refractory HER2-positive and HER2-negative advanced or metastatic breast cancer. *Annals of Oncology* **19**:1068-1074.
- Cabometyx (2018) Ipsen Pty Ltd, Glen Waverley, Victoria, Australia.
- Caprelsa (2013) AstraZeneca Pty Ltd, North Ryde, NSW, Australia.
- Carlsson L, Spjuth O, Adams S, Glen RC, and Boyer S (2010) Use of historic metabolic biotransformation data as a means of anticipating metabolic sites using MetaPrint2D and Bioclipse. *BMC Bioinformatics* **11**:362.
- Castellino S, O'Mara M, Koch K, Borts DJ, Bowers GD, and MacLauchlin C (2012) Human Metabolism of Lapatinib, a Dual Kinase Inhibitor: Implications for Hepatotoxicity. *Drug Metabolism and Disposition* **40**:139-150.
- Chapman PB, Hauschild A, Robert C, Haanen JB, Ascierto P, Larkin J, Dummer R, Garbe C, Testori A, et al. (2011) Improved Survival with Vemurafenib in Melanoma with BRAF V600E Mutation. *New England Journal of Medicine* **364**:2507-2516.
- Christopher LJ, Cui D, Wu C, Luo R, Manning JA, Bonacorsi SJ, Lago M, Allentoff A, Lee FY, McCann B, Galbraith S, Reitberg DP, He K, Barros A Jr, Blackwood-Chirchir A, Humphreys WG, Iyer RA. (2008a) Metabolism and disposition of dasatinib after oral administration to humans. **36**:1357-64.
- Christopher LJ, Cui D, Li W, Barros A, Arora VK, Zhang H, Wang L, Zhang D, Manning JA, He K, Fletcher AM, Ogan M, Lago M, Bonacorsi SJ, Humphreys WG, and Iyer RA (2008b) Biotransformation of [¹⁴C]Dasatinib: In Vitro Studies in Rat, Monkey, and Human and Disposition after Administration to Rats and Monkeys. *Drug Metabolism and Disposition* **36**:1341-1356.
- Clark RD, Strizhev A, Leonard JM, Blake JF, Matthew JB. (2002) Consensus scoring for ligand/protein interactions. *J Mol Graph Model* **20**:281-95.
- Cotellic (2016) Roche Products Pty Limited, Sydney, NSW, Australia.

DMD # 85167

de Bruyn Kops C, Friedrich NO, Kirchmair J (2017) Alignment-Based Prediction of Sites of Metabolism. *Journal of Chemical Information and Modeling* **57**:1258-1264.

de Graaf C, Oostenbrink C, Keizers PH, van Vugt-Lussenburg BM, van Waterschoot RA, Tschirret-Guth RA, Commandeur JN, Vermeulen NP (2007) Molecular modeling-guided site-directed mutagenesis of cytochrome P450 2D6. *Current Drug Metabolism* **8**:59-77.

Deng Y, Sychterz C, Suttle AB, Dar MM, Bershas D, Negash K, Qian Y, Chen EP, Gorycki PD, and Ho MYK (2013) Bioavailability, metabolism and disposition of oral pazopanib in patients with advanced cancer. *Xenobiotica* **43**:443-453.

Dickinson PA, Cantarini MV, Collier J, Frewer P, Martin S, Pickup K, and Ballard P (2016) Metabolic Disposition of Osimertinib in Rats, Dogs, and Humans: Insights into a Drug Designed to Bind Covalently to a Cysteine Residue of Epidermal Growth Factor Receptor. *Drug Metabolism and Disposition* **44**:1201-1212.

Ding J, (2014) Handbook of Metabolic Pathways of Xenobiotics, Nilotinib part 2. Compound articles, DOI: 10.1002/9781118541203

Dowty ME, Lin J, Ryder TF, Wang W, Walker GS, Vaz A, Chan GL, Krishnaswami S, and Prakash C (2014) The Pharmacokinetics, Metabolism, and Clearance Mechanisms of Tofacitinib, a Janus Kinase Inhibitor, in Humans. *Drug Metabolism and Disposition*. **42**:759-73.

Druker BJ, Guilhot F, O'Brien SG, Gathmann I, Kantarjian H, Gattermann N, Deininger MWN, Silver RT, et al. (2006) Five-Year Follow-up of Patients Receiving Imatinib for Chronic Myeloid Leukemia. *New England Journal of Medicine* **355**:2408-2417.

Dubbelman A-C, Nijenhuis CM, Jansen RS, Rosing H, Mizuo H, Kawaguchi S, Critchley D, Shumaker R, Schellens JHM, and Beijnen JH (2016) Metabolite profiling of the multiple tyrosine kinase inhibitor lenvatinib: a cross-species comparison. *Investigational New Drugs* **34**:300-318.

Duckett DR and Cameron MD (2010) Metabolism considerations for kinase inhibitors in cancer treatment. *Expert Opinion on Drug Metabolism & Toxicology* **6**:1175-1193.

Ekroos M, Sjögren T. (2006). Structural basis for ligand promiscuity in cytochrome P450 3A4. *Proc Natl Acad Sci USA* **103**:13682–13687.

Eldridge MD, Murray CW, Auton TR, Paolini GV, Mee RP. (1997) Empirical scoring functions: I. The development of a fast empirical scoring function to estimate the binding affinity of ligands in receptor complexes. *Journal of Computer-Aided Molecular Design* **11**:425-45.

Faivre S, Niccoli P, Castellano D, Valle JW, Hammel P, Raoul JL, Vinik A, Van Cutsem E, Bang YJ, Lee SH, et al. (2017) Sunitinib in pancreatic neuroendocrine tumors: updated progression-free survival and final overall survival from a phase III randomized study. *Annals of Oncology* **28**:339-343.

Ferraro D and Zalberg J (2014) Regorafenib in gastrointestinal stromal tumors: clinical evidence and place in therapy. *Therapeutic Advances in Medical Oncology* **6**:222-228.

DMD # 85167

Ford KA, Ryslik G, Sodhi J, Halladay J, Diaz D, Dambach D, Masuda M. (2015) Computational predictions of the site of metabolism of cytochrome P450 2D6 substrates: comparative analysis, molecular docking, bioactivation and toxicological implications. *Drug Metabolism Reviews* **47**:291-319.

Gasteiger J and Marsili M (1980) Iterative partial equalization of orbital electronegativity—a rapid access to atomic charges. *Tetrahedron* **36**:3219-3228.

Gerisch M, Hafner F-T, Lang D, Radtke M, Diefenbach K, Cleton A, and Lettieri J (2018) Mass balance, metabolic disposition, and pharmacokinetics of a single oral dose of regorafenib in healthy human subjects. *Cancer Chemotherapy and Pharmacology* **81**:195-206.

Ghassabian S, Rawling T, Zhou F, Doddareddy MR, Tattam BN, Hibbs DE, Edwards RJ, Cui PH, and Murray M (2012) Role of human CYP3A4 in the biotransformation of sorafenib to its major oxidized metabolites. *Biochemical Pharmacology* **84**:215-223.

Ghassabian S, Gillani TB, Rawling T, Crettol S, Nair PC, Murray M. (2019) Inhibition of human hepatic CYP3A4 by sorafenib and its N-oxide metabolite. *The AAPS Journal*, **21**:1-11.

Giotrif (2013) Boehringer Ingelheim Pty Limited, North Ryde NSW, Australia.

Glivec (2001) Novartis Pharmaceuticals Australia Pty Ltd, Macquarie Park, NSW, Australia.

Goldinger SM, Rinderknecht J, Dummer R, Kuhn FP, Yang KH, Lee L, Ayala RC, Racha J, Geng W, Moore D, Liu M, Joe AK, Bazan SPG, and Grippo JF (2015) A single dose mass balance and metabolite profiling study of vemurafenib in patients with metastatic melanoma. *Pharmacology Research & Perspectives* **3**:e00113.

Gschwind H-P, Pfaar U, Waldmeier F, Zollinger M, Sayer C, Zbinden P, Hayes M, Pokorny R, Seiberling M, Ben-Am M, Peng B, and Gross G (2005) Metabolism and disposition of imatinib mesylate in healthy volunteers. *Drug Metabolism and Disposition* **33**:1503-1512.

Ho MYK, Morris MJ, Pirhalla JL, Bauman JW, Pendry CB, Orford KW, Morrison RA, and Cox DS (2014) Trametinib, a first-in-class oral MEK inhibitor mass balance study with limited enrollment of two male subjects with advanced cancers. *Xenobiotica* **44**:352-368.

Hochhaus A, Larson RA, Guilhot F, Radich JP, Branford S, Hughes TP, Baccarani M, Deininger MW, Cervantes F, Fujihara S, Ortmann C-E, Menssen HD, Kantarjian H, O'Brien SG, and Druker BJ (2017) Long-Term Outcomes of Imatinib Treatment for Chronic Myeloid Leukemia. *New England Journal of Medicine* **376**:917-927.

Ibrance (2017) Pfizer Australia Pty Ltd, West Ryde, NSW, Australia.

Inoue K, Mizuo H, Kawaguchi S, Fukuda K, Kusano K, and Yoshimura T (2014) Oxidative Metabolic Pathway of Lenvatinib Mediated by Aldehyde Oxidase. *Drug Metabolism and Disposition* **42**:1326-1333.

Iclusig (2014) Takeda Pharmaceuticals Australia Pty Ltd, Sydney, NSW, Australia.

Imbruvica (2015) JANSSEN-CILAG Pty Ltd, Macquarie Park, NSW, Australia.

DMD # 85167

Inlyta (2017) Pfizer Australia Pty Ltd, West Ryde, NSW, Australia.

Iressa (2003) AstraZeneca Pty Ltd, Macquarie Park, NSW, Australia.

Jakavi (2013) Novartis Pharmaceuticals Australia Pty Limited, North Ryde, NSW, Australia.

Jain AN (1996) Scoring noncovalent protein-ligand interactions: A continuous differentiable function tuned to compute binding affinities. *Journal of Computer-Aided Molecular Design* **10**:427-40.

Jain AN (2000) Morphological similarity: a 3D molecular similarity method correlated with protein-ligand recognition. *Journal of Computer-Aided Molecular Design* **14**:199–213.

Jain AN (2003) Surflex: fully automatic flexible molecular docking using a molecular similarity-based search engine. *Journal of Medicinal Chemistry* **46**:499-511.

Jain AN (2004) Ligand-based structural hypotheses for virtual screening. *Journal of Medicinal Chemistry* **47**:947–961.

Jin F, Robeson M, Zhou H, Hisoire G, and Ramanathan S (2015) The pharmacokinetics and safety of idelalisib in subjects with moderate or severe hepatic impairment. *The Journal of Clinical Pharmacology* **55**:944-952.

Johnson TR, Tan W, Goulet L, Smith EB, Yamazaki S, Walker GS, O’Gorman MT, Bedarida G, Zou HY, Christensen JG, Nguyen LN, Shen Z, Dalvie D, Bello A, and Smith BJ (2015) Metabolism, excretion and pharmacokinetics of [¹⁴C]crizotinib following oral administration to healthy subjects. *Xenobiotica* **45**:45-59.

Jones G, Willett P, Glen RC, Leach AR, Taylor R. (1997) Development and validation of a genetic algorithm for flexible docking. *Journal of Molecular Biology* **267**:727-48.

Kaur P, Chamberlin AR, Poulos TL, and Sevrioukova IF (2016) Structure-Based Inhibitor Design for Evaluation of a CYP3A4 Pharmacophore Model. *Journal of Medicinal Chemistry* **59**:4210-4220.

Keller KL, Franquiz MJ, Duffy AP, and Trovato JA (2018) Drug–drug interactions in patients receiving tyrosine kinase inhibitors. *Journal of Oncology Pharmacy Practice* **24**:110-115.

Kirchmair J, Williamson MJ, Tyzack JD, Tan L, Bond PJ, Bender A, Glen RC. (2012) Computational prediction of metabolism: sites, products, SAR, P450 enzyme dynamics, and mechanisms. *Journal of Chemical Information and Modeling* **52**:617-48.

Kuntz ID, Blaney JM, Oatley SJ, Langridge R, Ferrin TE. (1982) A geometric approach to macromolecule-ligand interactions. *Journal of Molecular Biology* **161**:269-88.

Lacy S, Hsu B, Miles D, Aftab D, Wang R, and Nguyen L (2015) Metabolism and Disposition of Cabozantinib in Healthy Male Volunteers and Pharmacologic Characterization of Its Major Metabolites. *Drug Metabolism and Disposition* **43**:1190-1207.

Lawrence SK, Nguyen D, Bowen C, Richards-Peterson L, and Skordos KW (2014) The Metabolic Drug-Drug Interaction Profile of Dabrafenib: In Vitro Investigations and

DMD # 85167

Quantitative Extrapolation of the P450-Mediated DDI Risk. *Drug Metabolism and Disposition* **42**:1180-1190.

Lenvima (2016), Eisai (Australia) Pty Ltd, Melbourne, VIC, Australia.

Ling J, Johnson KA, Miao Z, Rakhit A, Pantze MP, Hamilton M, Lum BL, and Prakash C (2006) Metabolism and excretion of erlotinib, a small molecule inhibitor of epidermal growth factor receptor tyrosine kinase, in healthy male volunteers. *Drug Metabolism and Disposition* **34**:420-426.

Loi C-M, Vaz A, Hoffman J, O’Gorman M, Kirkovsky L, Wang D, and Dalvie D. (2015) P340: A phase one open-label single-radiolabeled dose study to investigate the absorption, metabolism, and excretion of [14c]pd-0332991 in healthy male volunteers. (19th North American ISSX) *Drug Metabolism Reviews* **47**(S1): 41–297.

Marvin 16.6.20, ChemAxon (<https://www.chemaxon.com>)

McKillop D, Hutchison M, Partridge EA, Bushby N, Cooper CM, Clarkson-Jones JA, Herron W, Swaisland HC. (2004) Metabolic disposition of gefitinib, an epidermal growth factor receptor tyrosine kinase inhibitor, in rat, dog and man. *Xenobiotica* **34**:917-34.

Mekinist (2014) GlaxoSmithKline Australia Pty Ltd, Abbotsford, Victoria, Australia.

Minami H, Kawada K, Ebi H, Kitagawa K, Kim YI, Araki K, Mukai H, Tahara M, Nakajima H, Nakajima K. (2008) Phase I and pharmacokinetic study of sorafenib, an oral multikinase inhibitor, in Japanese patients with advanced refractory solid tumors. *Cancer Science* **99**:1492-8.

Morcos PN, Yu L, Bogman K, Sato M, Katsuki H, Kawashima K, Moore DJ, Whayman M, Nieforth K, Heinig K, Guerini E, Muri D, Martin-Facklam M, Phipps A. (2017) Absorption, distribution, metabolism and excretion (ADME) of the ALK inhibitor alectinib: results from an absolute bioavailability and mass balance study in healthy subjects. *Xenobiotica* **47**:217-229.

Muegge I, Martin YC. (1999) A general and fast scoring function for protein-ligand interactions: a simplified potential approach. *Journal of Medicinal Chemistry* **42**:791-804.

Nair PC, McKinnon RA, Miners JO. (2016) Cytochrome P450 structure-function: insights from molecular dynamics simulations, *Drug Metabolism Reviews*, **48**:434-452.

Nakagawa T, Fowler S, Takanashi K, Youdim K, Yamauchi T, Kawashima K, Sato-Nakai M, Yu L, and Ishigai M (2017) In vitro metabolism of alectinib, a novel potent ALK inhibitor, in human: contribution of CYP3A enzymes. *Xenobiotica* **48**:546-554.

Nexavar (2006) Bayer Australia Ltd, Pymble, NSW, Australia.

Ofev (2015) Boehringer Ingelheim Pty Limited, North Ryde, NSW, Australia.

Pattana Wongsa A, Nair PC, Rowland A, Miners JO. (2016) Human UDP-Glucuronosyltransferase (UGT) 2B10: Validation of Cotinine as a Selective Probe Substrate, Inhibition by UGT Enzyme-Selective Inhibitors and Antidepressant and Antipsychotic Drugs,

DMD # 85167

and Structural Determinants of Enzyme Inhibition. *Drug Metabolism and Disposition* **44**:378-88.

Powell MJD (1977) Restart procedures for the conjugate gradient method. *Mathematical Programming* **12**:241-254.

Rendic S, Di Carlo FJ, (1997) Human cytochrome P450 enzymes: a status report summarizing their reactions, substrates, inducers, and inhibitors. *Drug Metabolism and Disposition* **29**:413-580.

Rendic S, Guengerich FP. (2015) Survey of Human Oxidoreductases and Cytochrome P450 Enzymes Involved in the Metabolism of Xenobiotic and Natural Chemicals. *Chemical Research in Toxicology*, **28**:38-42.

Roth GJ, Binder R, Colbatzky F, Dallinger C, Schlenker-Herceg R, Hilberg F, Wollin SL, Kaiser R. (2015) Nintedanib: from discovery to the clinic. *Journal of Medicinal Chemistry*, **58**:1053-63.

Rowland A, van Dyk M, Mangoni AA, Miners JO, McKinnon RA, Wiese MD, Rowland A, Kichenadasse G, Gurney H, Sorich MJ (2017) Kinase inhibitor pharmacokinetics: comprehensive summary and roadmap for addressing inter-individual variability in exposure. *Expert Opin Drug Metab Toxicol*. **13**:31-49.

Salentin S, Schreiber S, Haupt VJ, Adasme MF, Schroeder M (2015) PLIP: fully automated protein-ligand interaction profiler. *Nucleic Acids Research*. **43**:W443-W447

Scheers E, Leclercq L, de Jong J, Bode N, Bockx M, Laenen A, Cuyckens F, Skee D, Murphy J, Sukbuntherng J, and Mannens G (2015) Absorption, Metabolism, and Excretion of Oral ¹⁴C Radiolabeled Ibrutinib: An Open-Label, Phase I, Single-Dose Study in Healthy Men. *Drug Metabolism and Disposition* **43**:289-297.

Sevrioukova IF and Poulos TL (2013) Understanding the mechanism of cytochrome P450 3A4: recent advances and remaining problems. *Dalton Transactions* **42**:3116-3126.

Sevrioukova IF and Poulos TL (2017) Structural basis for regiospecific midazolam oxidation by human cytochrome P450 3A4. *Proceedings of the National Academy of Sciences* **114**:486-491.

Shilling AD, Nedza FM, Emm T, Diamond S, McKeever E, Punwani N, Williams W, Arvanitis A, Galya LG, Li M, Shepard S, Rodgers J, Yue T-Y, and Yeleswaram S (2010) Metabolism, Excretion, and Pharmacokinetics of [¹⁴C]INCB018424, a Selective Janus Tyrosine Kinase 1/2 Inhibitor, in Humans. *Drug Metabolism and Disposition* **38**:2023-2031.

Smith BJ, Pithavala Y, Bu HZ, Kang P, Hee B, Deese AJ, Pool WF, Klamerus KJ, Wu EY, Dalvie DK. (2014) Pharmacokinetics, metabolism, and excretion of [¹⁴C]axitinib, a vascular endothelial growth factor receptor tyrosine kinase inhibitor, in humans. *Drug Metabolism and Disposition* **42**:918-31.

Speed B, Bu H-Z, Pool WF, Peng GW, Wu EY, Patyna S, Bello C, and Kang P (2012) Pharmacokinetics, Distribution, and Metabolism of [¹⁴C]Sunitinib in Rats, Monkeys, and Humans. *Drug Metabolism and Disposition* **40**:539-555.

DMD # 85167

Sprycel (2007) Bristol-Myers Squibb Australia Pty Ltd, Mulgrave, Victoria, Australia.

Sykes MJ, McKinnon RA, and Miners JO (2008) Prediction of Metabolism by Cytochrome P450 2C9: Alignment and Docking Studies of a Validated Database of Substrates. *Journal of Medicinal Chemistry* **51**:780-791.

Stivarga (2013) Bayer Australia Limited, Pymble, NSW, Australia.

Stopfer P, Marzin K, Narjes H, Gansser D, Shahidi M, Uttreuther-Fischer M, and Ebner T (2012) Afatinib pharmacokinetics and metabolism after oral administration to healthy male volunteers. *Cancer Chemotherapy and Pharmacology* **69**:1051-1061.

Sutent (2006) Pfizer Australia Pty Ltd, West Ryde, NSW, Australia.

Tafinlar (2013) Novartis Pharmaceuticals Australia Pty Limited, Macquarie Park, NSW, Australia.

Tagrisso (2016) AstraZeneca Pty Ltd, Macquarie Park, NSW, Australia.

Takahashi RH, Choo EF, Ma S, Wong S, Halladay J, Deng Y, Rooney I, Gates M, Hop CECA, Khojasteh SC, Dresser MJ, and Musib L (2016) Absorption, Metabolism, Excretion, and the Contribution of Intestinal Metabolism to the Oral Disposition of [14C]Cobimetinib, a MEK Inhibitor, in Humans. *Drug Metabolism and Disposition* **44**:28-39.

Tarceva (2006) Roche Products Pty Limited, Dee Why, NSW, Australia.

Tarcsay Á, Keseru GM. (2011) In silico site of metabolism prediction of cytochrome P450-mediated biotransformations. *Expert Opinion on Drug Metabolism & Toxicology* **7**:299-312.

Tasigna (2008) Novartis Pharmaceuticals Australia Pty Ltd, Macquarie Park, NSW, Australia.

Tykerb (2007) GlaxoSmithKline Australia Pty Ltd, Abbotsford, Victoria, Australia.

Vasanthanathan P, Olsen L, Jørgensen FS, Vermeulen NP, Oostenbrink C (2010) Computational prediction of binding affinity for CYP1A2-ligand complexes using empirical free energy calculations. *Drug Metabolism and Disposition* **38**:1347-54.

Votrient (2010) Novartis Pharmaceuticals Australia Pty Ltd, Macquarie Park, NSW, Australia.

Wang L, Christopher LJ, Cui D, Li W, Iyer R, Humphreys WG, and Zhang D (2008) Identification of the Human Enzymes Involved in the Oxidative Metabolism of Dasatinib: An Effective Approach for Determining Metabolite Formation Kinetics. *Drug Metabolism and Disposition* **36**:1828-1839.

Wang R, Lu Y, Wang S. (2003) Comparative evaluation of 11 scoring functions for molecular docking. *Journal of Medicinal Chemistry* **46**:2287-303.

Wu P, Nielsen TE, and Clausen MH (2015) FDA-approved small-molecule kinase inhibitors. *Trends in Pharmacological Sciences* **36**:422-439.

Wu P, Nielsen TE, and Clausen MH (2016) Small-molecule kinase inhibitors: an analysis of FDA-approved drugs. *Drug Discovery Today* **21**:5-10.

DMD # 85167

Xalkori (2013) Pfizer Australia Pty Ltd, West Ryde, NSW, Australia.

Xeljanx (2015) Pfizer Australia Pty Ltd, West Ryde, NSW, Australia.

Yano JK, Wester MR, Schoch GA, et al. (2004). The structure of human microsomal cytochrome P450 3A4 determined by X-ray crystallography to 2.05-Å resolution. *Journal of Biological Chemistry* **279**:38091–38094.

Ye YE, Woodward CN, and Narasimhan NI (2017) Absorption, metabolism, and excretion of [¹⁴C]ponatinib after a single oral dose in humans. *Cancer Chemotherapy and Pharmacology* **79**:507-518.

Zanger UM, Schwab M (2013) Cytochrome P450 enzymes in drug metabolism: regulation of gene expression, enzyme activities, and impact of genetic variation. *Pharmacology & Therapeutics* **138**:103-41.

Zaretski J, Matlock M, and Swamidass SJ (2013) XenoSite: Accurately Predicting CYP-Mediated Sites of Metabolism with Neural Networks. *Journal of Chemical Information and Modeling* **53**:3373-3383.

Zaretski J, Rydberg P, Bergeron C, Bennett KP, Olsen L, and Breneman CM (2012) RS-Predictor Models Augmented with SMARTCyp Reactivities: Robust Metabolic Regioselectivity Predictions for Nine CYP Isozymes. *Journal of Chemical Information and Modeling* **52**:1637-1659.

Zelboraf (2012) Roche Products Pty Limited, Sydney, NSW, Australia.

Zhang J, Salminen A, Yang X, Luo Y, Wu Q, White M, Greenhaw J, Ren L, Bryant M, Salminen W, Papoian T, Mattes W, Shi Q. (2017) Effects of 31 FDA approved small-molecule kinase inhibitors on isolated rat liver mitochondria. *Archives of Toxicology* **91**:2921-2938.

Zhu X-D, Zhang J-B, Fan P-L, Xiong Y-Q, Zhuang P-Y, Zhang W, Xu H-X, Gao D-M, Kong L-Q, Wang L, Wu W-Z, Tang Z-Y, Ding H, and Sun H-C (2011) Antiangiogenic effects of pazopanib in xenograft hepatocellular carcinoma models: evaluation by quantitative contrast-enhanced ultrasonography. *BMC Cancer* **11**:28.

Zientek MA, Goosen TC, Tseng E, Lin J, Bauman JN, Walker GS, Kang P, Jiang Y, Freiwald S, Neul D, and Smith BJ (2016) In Vitro Kinetic Characterization of Axitinib Metabolism. *Drug Metabolism and Disposition* **44**:102-114.

Zydelig (2015) Gilead Sciences Pty Ltd, Melbourne, Victoria, Australia.

Zykadia (2016) Novartis Pharmaceuticals Australia Pty Ltd, Macquarie Park, NSW, Australia.

DMD # 85167

Footnotes

This study was supported by a project grant [1120137] from the National Health and Medical Research Council of Australia.

DMD # 85167

Table 1 Contribution of CYP3A4 and other enzymes to the metabolism of kinase inhibitors based on biotransformation data from in vivo and in vitro studies.

Kinase inhibitor	Major enzyme(s) involved in metabolism	Minor enzyme(s) involved in metabolism
Afatinib	-	FMO3, CYP3A4
Alectinib	CYP3A4	
Axitinib	CYP3A4/5	CYP1A2, CYP2C19, UGT1A1
Bosutinib	CYP3A4	
Cabozantinib	CYP3A4	
Ceritinib	CYP3A, CYP2C9	
Cobimetinib	CYP3A, UGT2B7	
Crizotinib	CYP3A4/5	
Dabrafenib	CYP2C8, CYP3A4	
Dasatinib	CYP3A4	FMO3, UGT
Erlotinib	CYP3A4	CYP1A2
Gefitinib	CYP3A4	
Ibrutinib	CYP3A	CYP2D6
Idelalisib	Aldehyde oxidase, CYP3A	
Imatinib	CYP3A4	CYP1A2, CYP2D6, CYP2C9, CYP2C19, CYP2C8
Lapatinib	CYP3A4, CYP3A5	
Lenvatinib	CYP3A, aldehyde oxidase	
Nilotinib	CYP3A4	
Nintedanib	Esterases, UGT 1A1, UGT 1A7, UGT 1A8, UGT 1A10	CYP3A4
Osimertinib	CYP3A	
Palbociclib	CYP3A, SULT2A1	
Pazopanib	CYP3A4	CYP1A2, CYP2C8
Ponatinib	CYP3A4	CYP2C8, CYP2D6, CYP3A5, esterases and/or amidases
Regorafenib	CYP3A4, UGT1A9	
Ruxolitinib	CYP3A4	
Sorafenib	CYP3A4, UGT1A9	
Sunitinib	CYP3A4	
Tofacitinib	CYP3A4	CYP2C19
Trametinib ^a	Deacetylation/mono-oxygenation/and or glucuronidation	CYP3A4
Vandetanib	CYP3A4	FMO1, FMO3
Vemurafenib	CYP3A4	

^aDeacetylated product metabolized by CYP3A4.

DMD # 85167

Table 2a Prediction of the SOM of KIs by molecular docking in the 5VCC structure using different scoring functions.

Kinase inhibitor	D-score	PMF-score	G-Score	ChemScore	CSCORE	Total Score
Afatinib	x	✓	x	x	x	✓
Alectinib	x	x	x	✓	✓	x
Axitinib	x	✓	x	x	✓	✓
Bosutinib	✓	✓	✓	✓	✓	x
Cabozantinib	✓	✓	✓	✓	✓	✓
Ceritinib	✓	✓	✓	✓	✓	✓
Cobimetinib	✓	✓	✓	✓	✓	✓
Crizotinib	x	x	x	x	x	x
Dabrafenib	x	x	x	x	x	x
Dasatinib	✓	✓	✓	✓	✓	✓
Erlotinib	✓	✓	✓	✓	✓	✓
Gefitinib	✓	✓	✓	✓	✓	✓
Ibrutinib	✓	✓	✓	✓	✓	✓
Idelalisib	x	✓	x	✓	✓	✓
Imatinib	x	x	x	x	x	x
Lapatinib	x	✓	x	x	✓	✓
Lenvatinib	✓	✓	✓	✓	✓	✓
Nilotinib	✓	✓	✓	✓	✓	✓
Nintedanib	x	x	x	✓	x	x
Osimertinib	x	✓	✓	✓	x	x
Palbociclib	x	x	x	x	x	x
Pazopanib	✓	✓	✓	✓	x	✓
Ponatinib	✓	✓	✓	✓	✓	✓
Regorafenib	✓	✓	✓	x	✓	✓
Ruxolitinib	✓	x	x	x	✓	x
Sorafenib	✓	✓	✓	x	✓	x
Sunitinib	x	x	x	x	x	x
Tofacitinib	x	✓	✓	x	✓	x
Trametinib	x	✓	✓	✓	✓	✓
Vandetanib	x	✓	x	✓	x	x
Vemurafenib	✓	✓	✓	✓	✓	✓
Number of molecules predicted^a	16(52%)	23(74%)	18(58%)	19(61%)	21(68%)	18(58%)

^a Percentage of dataset molecules with SOM predicted correctly.

DMD # 85167

Table 2b Prediction of the SOM of KIs by molecular docking in the 3UA1 structure using different scoring functions.

Kinase inhibitor	D-score	PMF-score	G-Score	ChemScore	CSCORE	Total Score
Afatinib	x	x	x	x	x	x
Alectinib	✓	✓	✓	✓	✓	✓
Axitinib	x	✓	✓	x	x	✓
Bosutinib	✓	✓	✓	✓	✓	✓
Cabozantinib	✓	✓	✓	✓	✓	✓
Ceritinib	x	✓	x	✓	✓	✓
Cobimetinib	x	✓	✓	✓	✓	✓
Crizotinib	x	x	x	x	✓	x
Dabrafenib	x	x	x	x	x	x
Dasatinib	✓	✓	✓	✓	✓	✓
Erlotinib	x	✓	✓	✓	x	x
Gefitinib	x	✓	✓	✓	x	✓
Ibrutinib	✓	✓	✓	✓	✓	✓
Idelalisib	✓	✓	✓	✓	✓	✓
Imatinib	x	x	x	x	x	x
Lapatinib	x	✓	✓	x	x	✓
Lenvatinib	✓	✓	✓	✓	✓	✓
Nilotinib	✓	✓	✓	✓	✓	✓
Nintedanib	✓	✓	✓	✓	✓	✓
Osimertinib	x	✓	✓	✓	x	✓
Palbociclib	x	x	x	x	x	✓
Pazopanib	✓	✓	✓	✓	✓	✓
Ponatinib	x	✓	✓	✓	✓	✓
Regorafenib	✓	✓	✓	x	✓	✓
Ruxolitinib	✓	✓	x	x	✓	x
Sorafenib	x	✓	✓	x	✓	✓
Sunitinib	x	x	x	x	x	x
Tofacitinib	x	✓	x	x	x	x
Trametinib	x	✓	x	✓	✓	✓
Vandetanib	x	x	✓	✓	✓	✓
Vemurafenib	✓	✓	✓	✓	✓	✓
Number of molecules predicted	13(42%)	24(77%)	21(68%)	19(61%)	20(65%)	23(74%)

^a Percentage of dataset molecules with SOM predicted correctly.

DMD # 85167

Table 2c Prediction of the SOM of KIs by molecular docking in the 5VCO structure using different scoring functions.

Kinase inhibitor	D-score	PMF-score	G-Score	ChemScore	CSCORE	Total Score
Afatinib	✓	✓	✓	✓	X	✓
Alectinib	✓	✓	✓	✓	✓	✓
Axitinib	X	✓	✓	✓	X	X
Bosutinib	X	✓	X	X	✓	X
Cabozantinib	✓	✓	✓	✓	✓	✓
Ceritinib	✓	✓	✓	✓	✓	✓
Cobimetinib	✓	✓	✓	✓	✓	✓
Crizotinib	X	X	X	X	✓	X
Dabrafenib	X	X	X	X	X	X
Dasatinib	✓	X	✓	✓	✓	✓
Erlotinib	✓	✓	✓	✓	X	✓
Gefitinib	✓	✓	✓	✓	X	✓
Ibrutinib	✓	✓	✓	✓	✓	✓
Idelalisib	✓	✓	✓	✓	✓	✓
Imatinib	X	X	X	X	X	X
Lapatinib	✓	✓	✓	✓	X	✓
Lenvatinib	✓	✓	✓	✓	✓	✓
Nilotinib	✓	✓	✓	✓	✓	✓
Nintedanib	X	X	X	X	✓	X
Osimertinib	X	X	X	X	X	X
Palbociclib	X	X	X	X	X	X
Pazopanib	✓	✓	✓	✓	✓	✓
Ponatinib	✓	✓	✓	✓	✓	✓
Regorafenib	✓	✓	✓	X	✓	✓
Ruxolitinib	✓	✓	✓	✓	✓	✓
Sorafenib	X	✓	✓	✓	✓	X
Sunitinib	✓	✓	X	X	X	X
Tofacitinib	X	X	X	X	X	X
Trametinib	✓	✓	X	X	✓	✓
Vandetanib	X	X	X	X	✓	X
Vemurafenib	X	X	✓	X	✓	✓
Number of molecules predicted ^a	19(61%)	21(68%)	20(65%)	18(58%)	20(65%)	19(61%)

^a Percentage of dataset molecules with SOM predicted correctly.

DMD # 85167

Table 3 CYP3A4 amino acids involved in non-covalent interactions with KIs based on docking (PMF-score) in the 3UA1 structure.

Kinase inhibitor	Hydrophobic	Hydrogen Bonds	π-Stacking	π-Cation	Salt bridge	Halogen bond
Alectinib	Phe 57 Asp 76 Thr 224 Phe 304	Thr 224 Arg 372	Phe 215	Arg 106		
Axitinib	Phe 57 Phe 215 Thr 309 Ile 369 Ala 370					
Bosutinib		Arg 105 Glu 374 Arg 375			Asp 76 Glu 374	
Cabozantinib	Phe 57 Phe 108 Phe 215		Phe 215	Arg 105		
Ceritinib	Phe 57 Ile 120 Phe 215 Phe 241 Ile 301 Phe 304 Ala 305	Arg 106 Ser 119 Ala 305 Arg 372		Arg 105		
Cobimetinib	Ala 305	Arg 212 Arg 372 Glu 374				Ser 119
Dasatinib	Tyr 53 Phe 57 Asp 76 Thr 224		Phe 215	Arg 106		
Erlotinib	Phe 57	Arg 105 Ser 119 Arg 212 Glu 374				
Gefitinib	Phe 57 Phe 215	Arg 212				Thr 224
Ibrutinib	Phe 57 Phe 215 Ala 370	Arg 105 Arg 212 Glu 374				
Idelalisib	Ile 120 Phe 215 Phe 304					

DMD # 85167

	Ala 370			
Lapatinib	Phe 57 Gln 79 Phe 215 Thr 224 Pro 227	Arg 212		Arg 106
Lenvatinib	Phe 57 Phe 215	Arg 212 Ala 305 Arg 372		
Nilotinib	Phe 57 Phe 215 Ile 223 Thr 224 Pro 227 Phe 304	Arg 106 Arg 372		Arg 106 Arg 212
Nintedanib	Phe 57 Phe 215 Arg 372 Glu 374	Tyr 53 Ser 119 Thr 224	Phe 108 Phe 215	Arg 106
Osimertinib	Phe 57 Phe 215	Ala 305 Ala 370	Phe 215	Arg 106
Pazopanib	Phe 215 Thr 309 Ala 370	Arg 105 Ala 305 Arg 372 Glu 374		Arg 106
Ponatinib	Phe 215 Ala 305 Thr 309 Ala 370 Leu373	Ile 223 Glu 374		
Regorafenib	Phe 215	Glu 374		Thr 224
Ruxolitinib	Phe 215	Phe 213 Ala 305		Arg 212
Sorafenib	Phe 57	Arg 372 Glu 374		Arg 106
Tofacitinib	Thr 309			
Trametinib	Arg 106 Phe 108 Ile 120 Phe 215 Phe 241 Phe 304 Ala 305 Thr 309 Ile 369	Ser 119 Arg 212	Phe 108	

DMD # 85167

	Ala 370		
Vemurafenib	Phe 57	Phe 215	Asp 76
	Phe 215		
	Phe 304		

DMD # 85167

Table 4 Prediction of the SOM of KIs using different template structures for ligand superposition.

Kinase inhibitor	Template molecule for ligand superposition			
	BEC	Sorafenib	Tofacitinib	Lapatinib
Afatinib	x	✓	x	✓
Alectinib	x	✓	x	✓
Axitinib	x	✓	✓	✓
Bosutinib	x	✓	✓	✓
Cabozantinib	✓	✓	x	x
Ceritinib	✓	x	x	x
Cobimetinib	✓	✓	x	✓
Crizotinib	x	✓	x	✓
Dabrafenib	x	x	x	✓
Dasatinib	x	✓	✓	✓
Erlotinib	x	✓	✓	✓
Gefitinib	x	✓	x	✓
Ibrutinib	x	✓	✓	x
Idelalisib	x	x	x	x
Imatinib	x	x	x	✓
Lapatinib	✓	x	x	✓
Lenvatinib	✓	✓	x	✓
Nilotinib	x	x	✓	✓
Nintedanib	x	✓	✓	x
Osimertinib	x	x	x	✓
Palbociclib	x	x	x	✓
Pazopanib	✓	✓	✓	✓
Ponatinib	✓	✓	x	✓
Regorafenib	✓	✓	x	✓
Ruxolitinib	x	x	x	x
Sorafenib	✓	✓	x	x
Sunitinib	✓	✓	✓	x
Tofacitinib	✓	✓	x	x
Trametinib	✓	✓	x	x
Vandetanib	✓	✓	x	✓
Vemurafenib	x	✓	✓	✓
Number of molecules predicted^a	13(42%)	22(71%)	10(32%)	21(68%)

^a Percentage of dataset molecules with SOM predicted correctly.

DMD # 85167

Figure Legends

Figure 1 Chemical structures of kinase inhibitors (KIs) and bromoergocryptine. The experimentally known major site(s) of metabolism (SOM) of KIs are shown by black arrows, and the predicted SOM by magenta arrows. The number adjacent to the magenta arrow shows the ratio (expressed as a percentage) of the number of docked poses with the indicated SOM relative to the total number of successful docked poses (see Results: Molecular Docking). Black and magenta spheres represent multiple sites identified from experimental and docking predictions, respectively.

^aThe deacetylated product was investigated in this study.

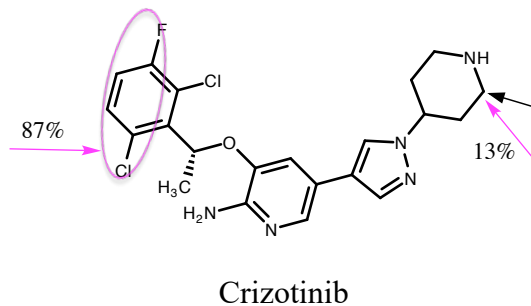
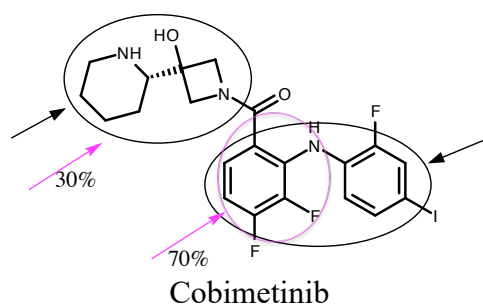
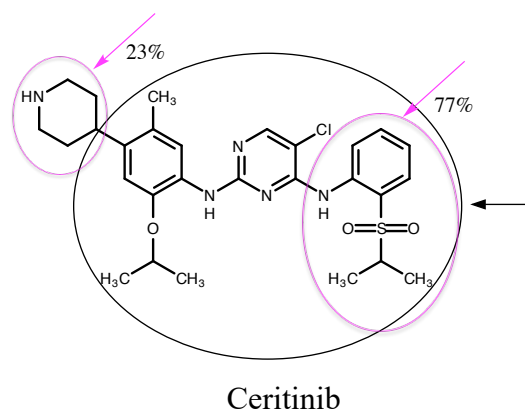
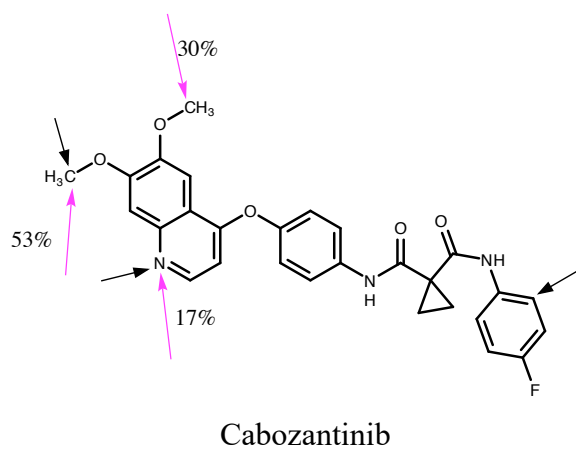
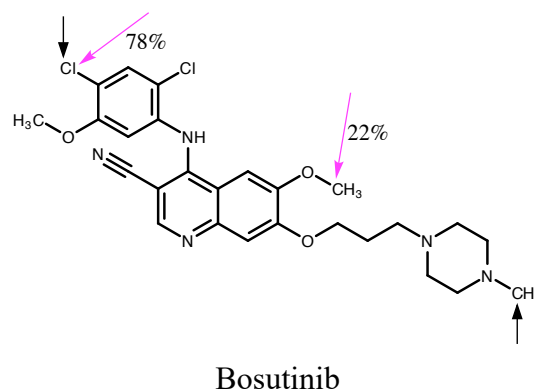
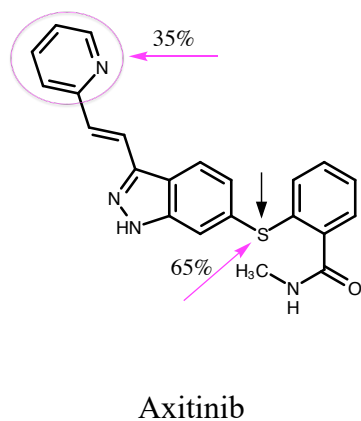
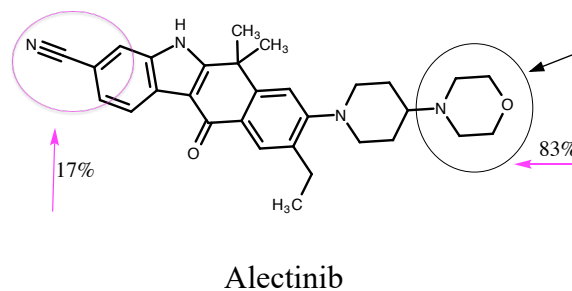
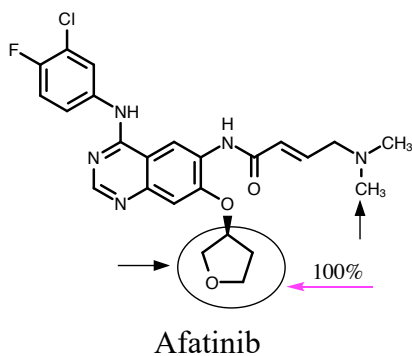
Figure 2 Number of KIs for which the major SOM was correctly predicted by molecular docking using the six scoring functions: A) 5VCC as the template structure for docking; B) 3UA1 as the template structure for docking; and C) 5VCO as the template structure for docking.

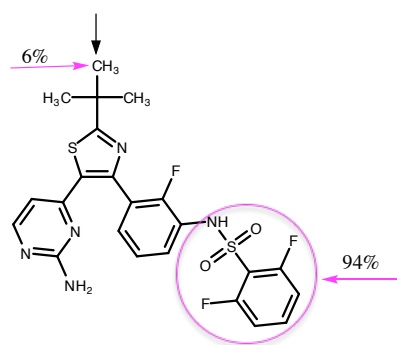
Figure 3 Consensus docked poses of kinase inhibitors in the CYP3A4 catalytic site: A) bostunib (C atoms shown as yellow); B) dasatinib (C atoms shown as cyan); C) ibrutinib (C atoms shown as purple); and D) sorafenib (C atoms shown as magenta). The SOM atom(s) are shown as spheres (with arrows to aid identification). Protein C atoms are in green. O, N, S and halogens (Cl/F) are shown in red, blue, yellow and green, respectively. The SOM distance(s) are shown in Å (dotted black line(s)).

Figure 4 Predicted SOM(s) of KIs by ligand superposition using sorafenib (C atoms in yellow) as the template: A) cabozantinib, B) dasatinib, C) lenvatinib, and D) regorafenib. The major SOM(s) is shown as a sphere(s) (with arrows to aid identification). C-atoms for the overlayed

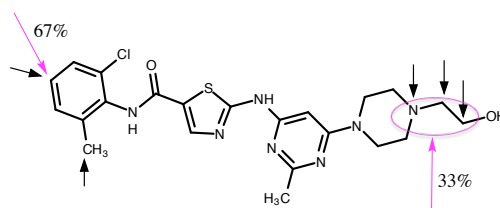
DMD # 85167

molecules (A-D) are shown in magenta. O, N and halogen (F/Cl) atoms are shown in red, blue and green, respectively.

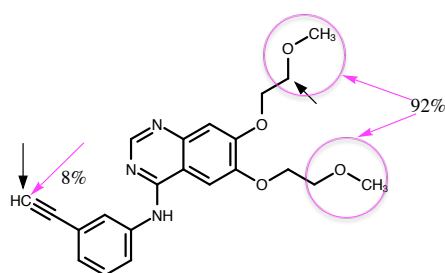




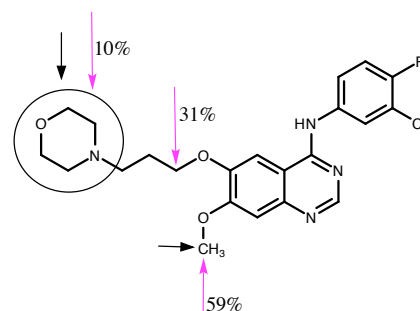
Dabrafenib



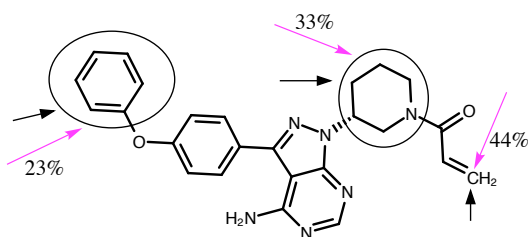
Dasatinib



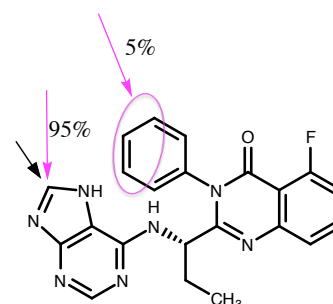
Erlotinib



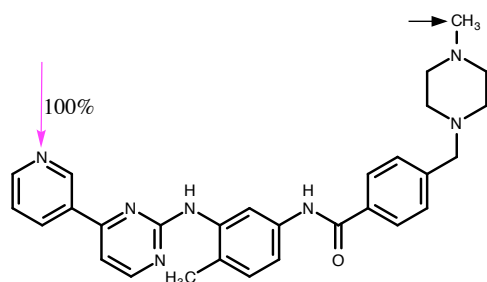
Gefitinib



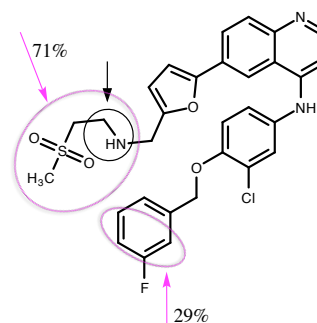
Ibrutinib



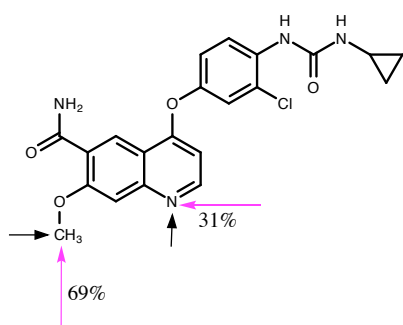
Idelalisib



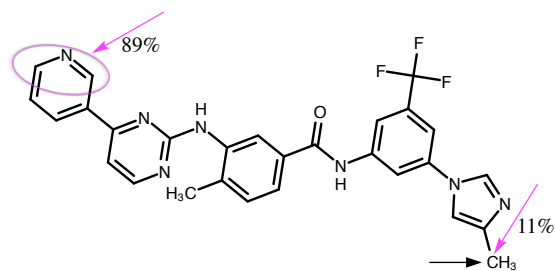
Imatinib



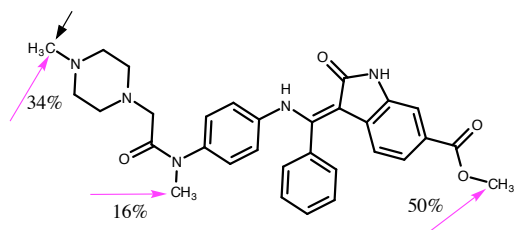
Lapatinib



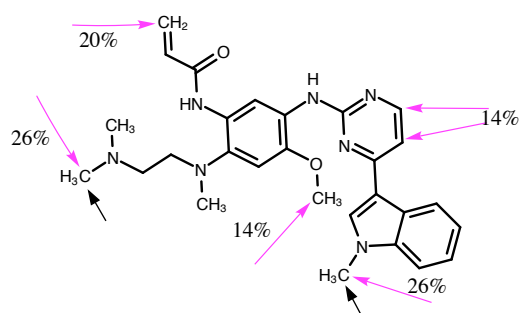
Lenvatinib



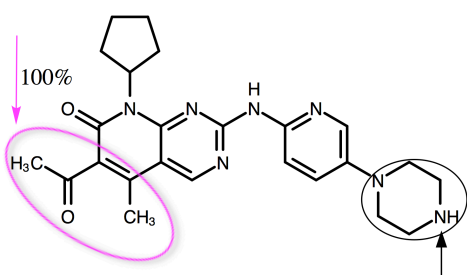
Nilotinib



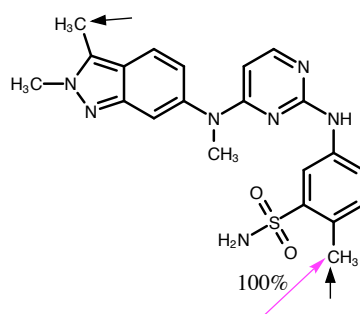
Nintedanib



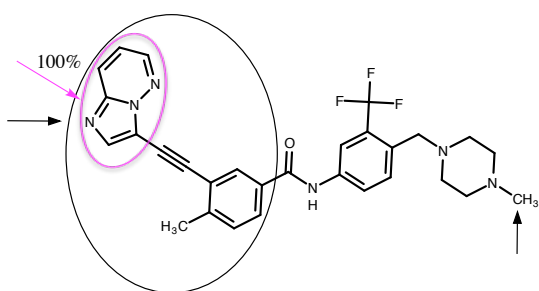
Osimertinib



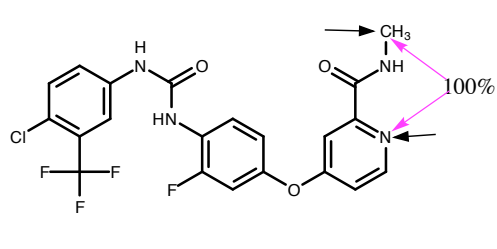
Palbociclib



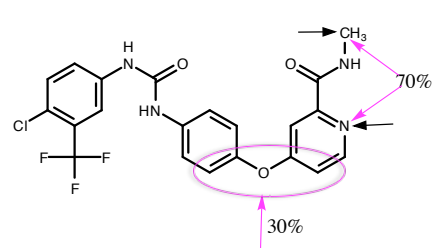
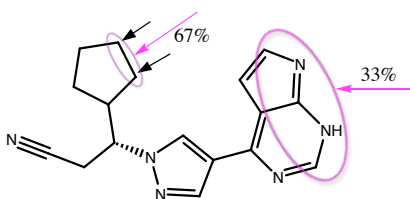
Pazopanib



Ponatinib



Regorafenib



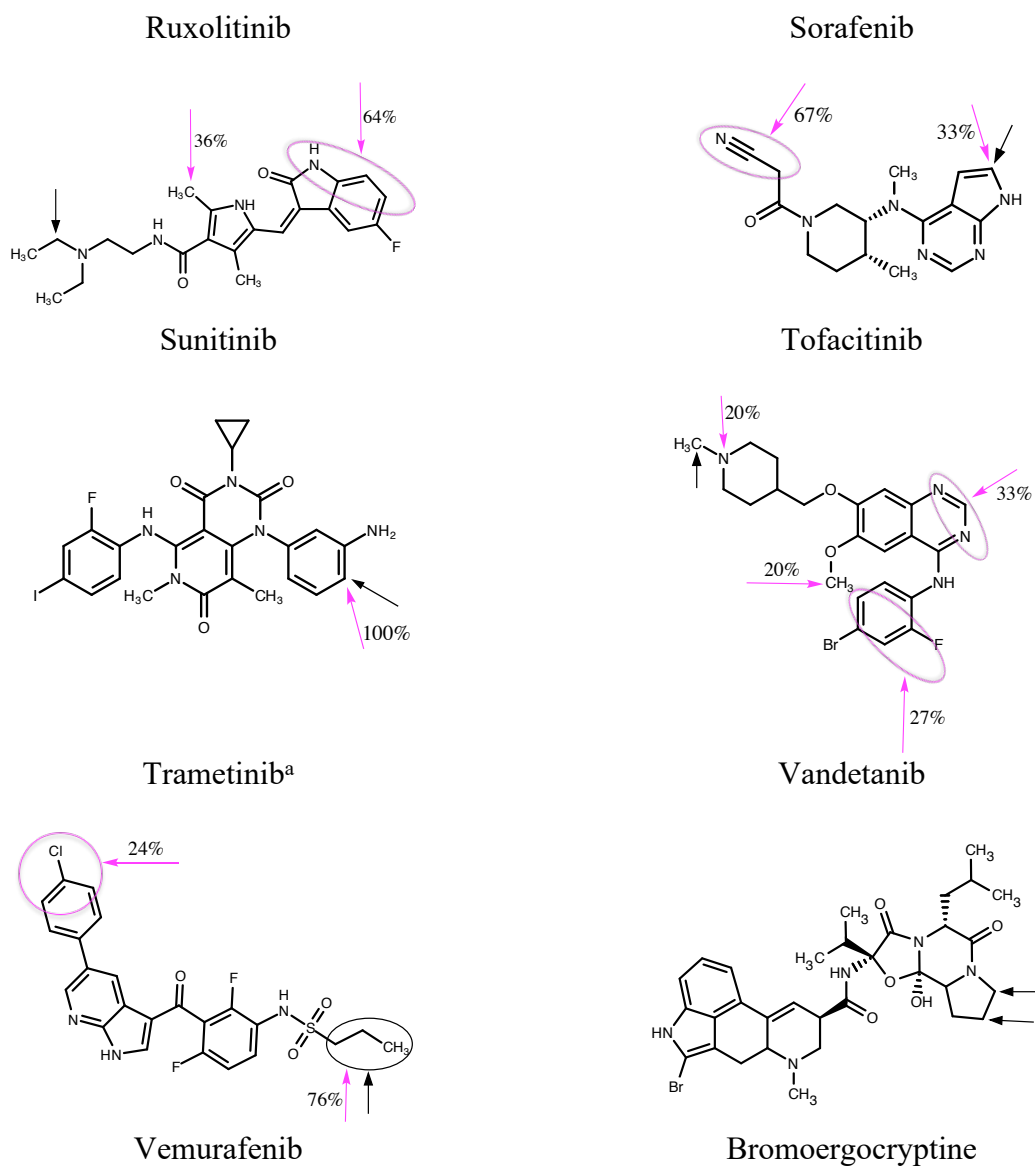


Figure 1

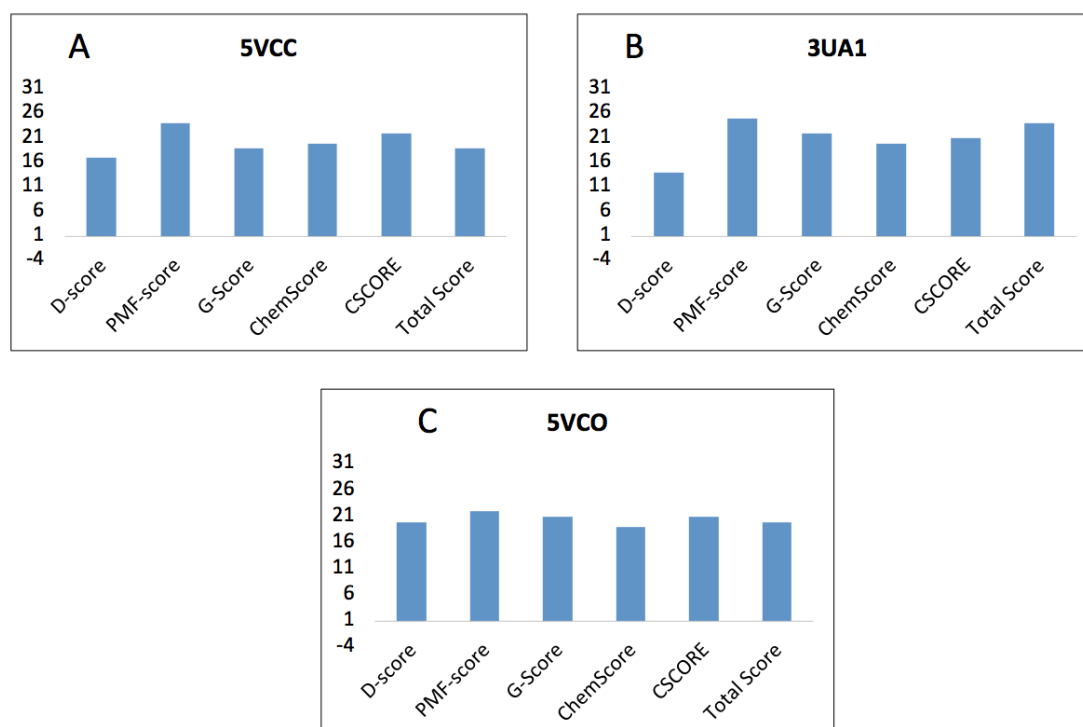


Figure 2

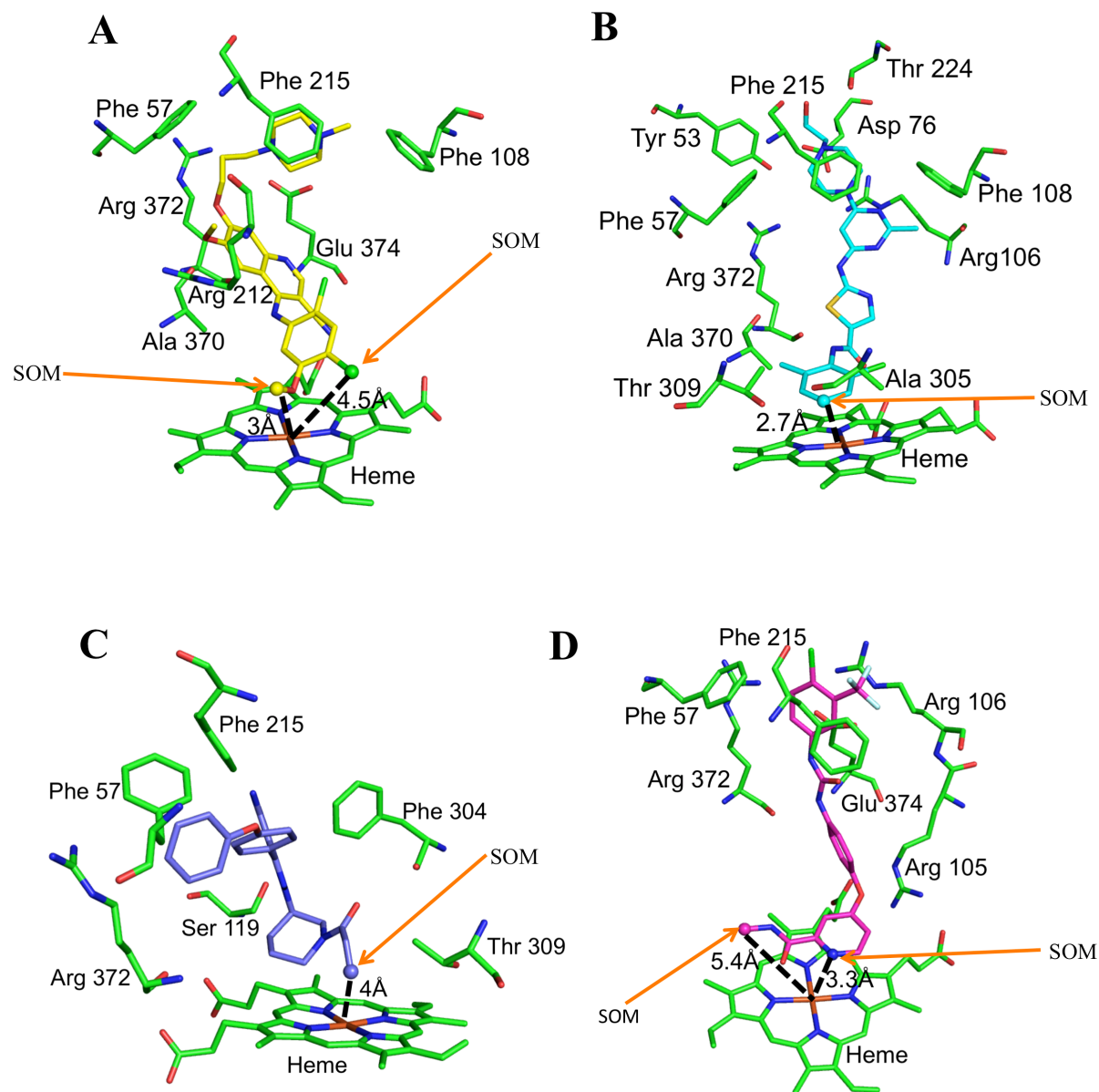


Figure 3

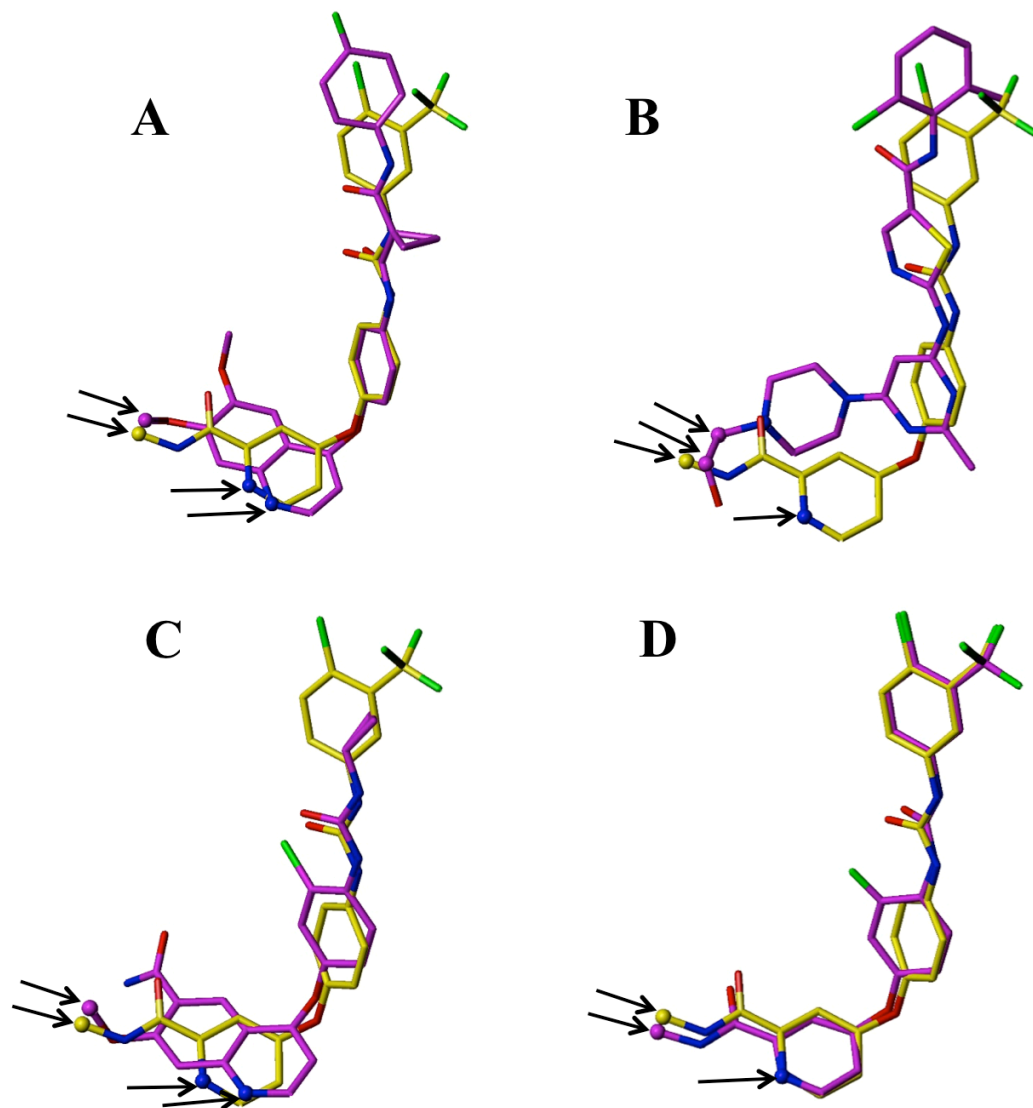
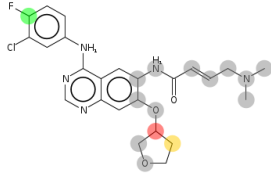
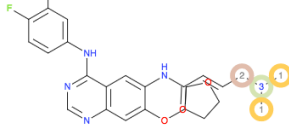
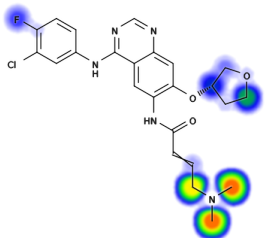
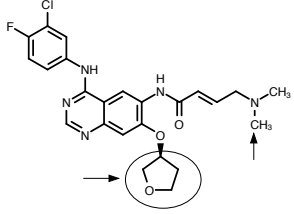
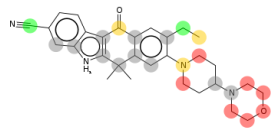
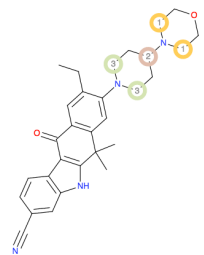
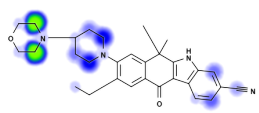
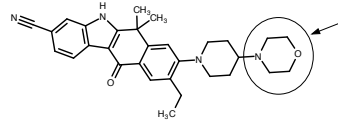
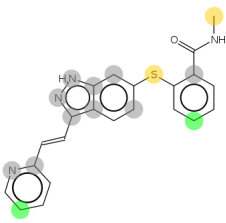
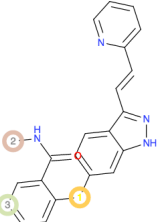
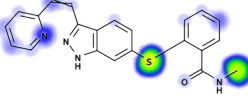
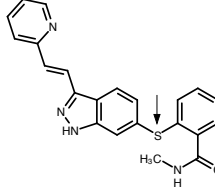
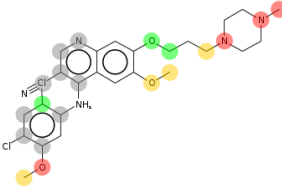
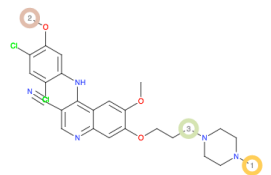
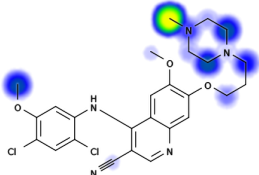
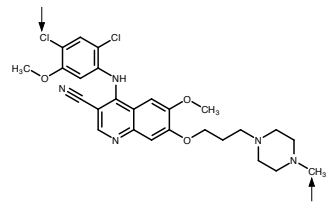
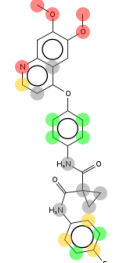
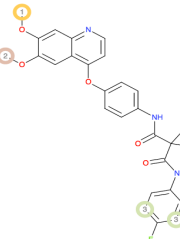
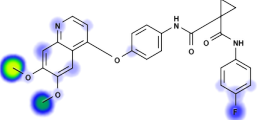
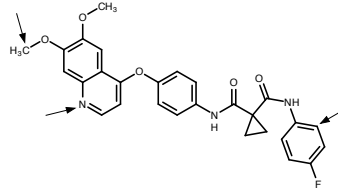


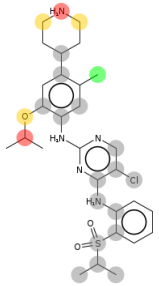
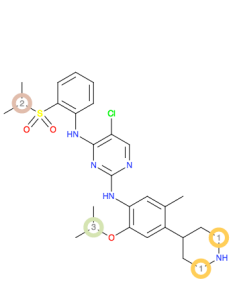
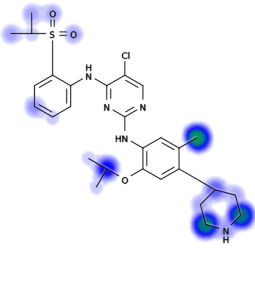
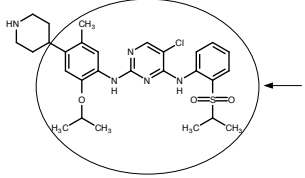
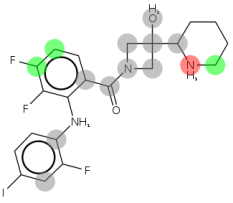
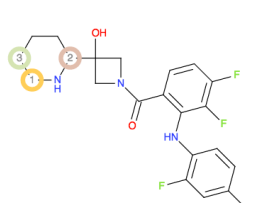
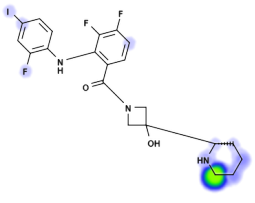
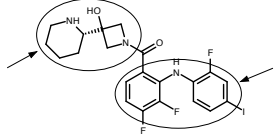
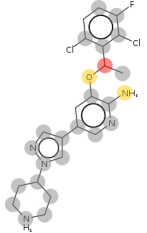
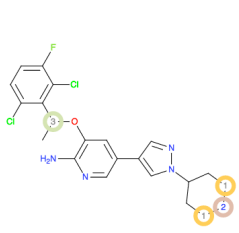
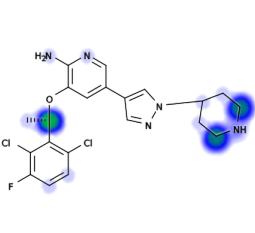
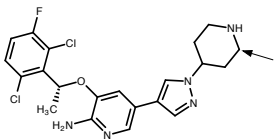
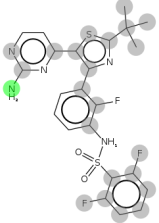
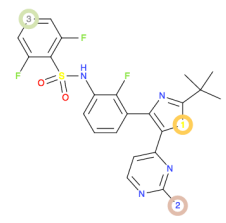
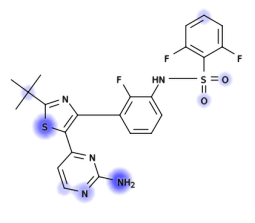
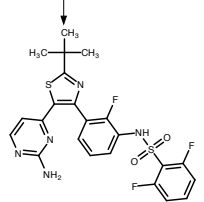
Figure 4

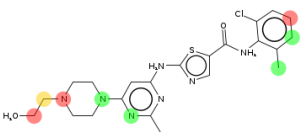
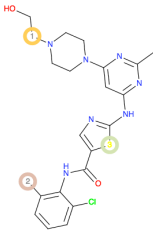
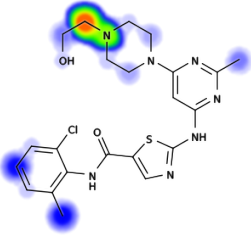
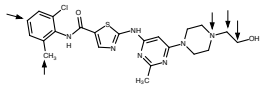
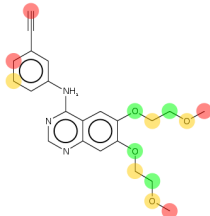
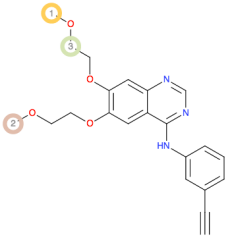
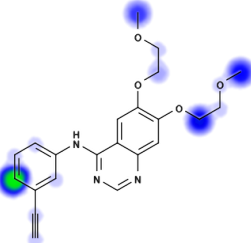
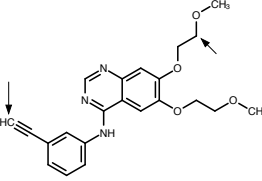
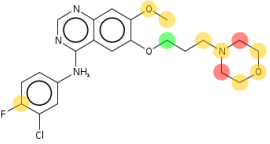
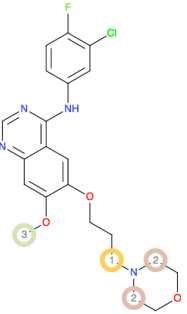
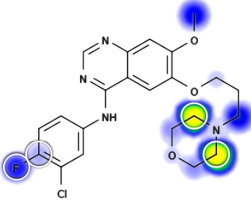
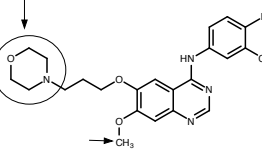
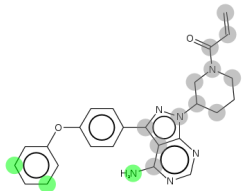
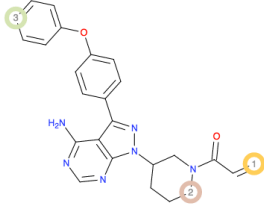
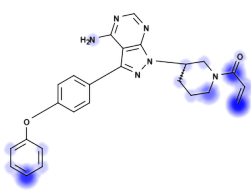
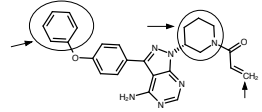
Supplemental Table S1 and Figure S1. Pramod C. Nair, Ross A. McKinnon and John O. Miners. Computational prediction of the site(s) of metabolism (SOM) and binding modes of protein kinase inhibitors metabolised by CYP3A4

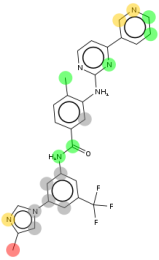
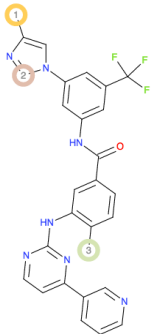
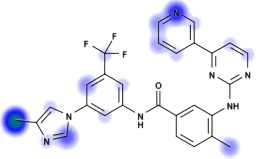
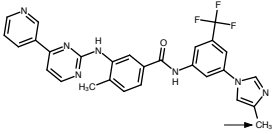
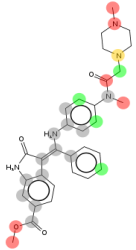
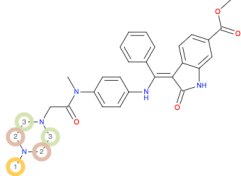
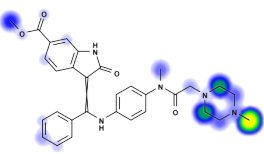
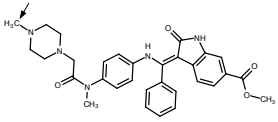
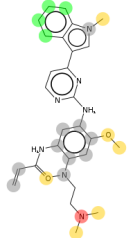
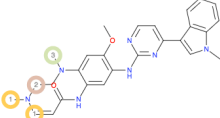
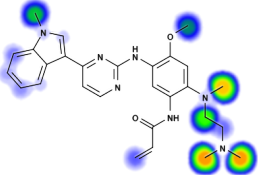
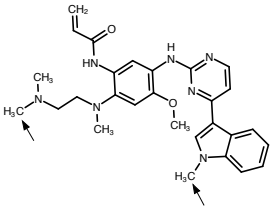
Table S1 Web based methods approaches for site of metabolism (SOM) prediction.

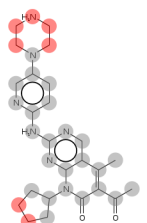
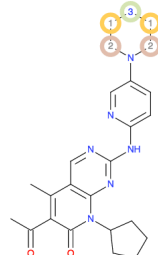
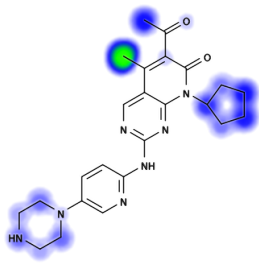
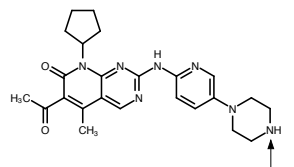
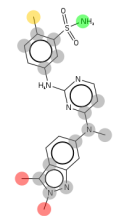
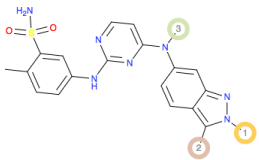
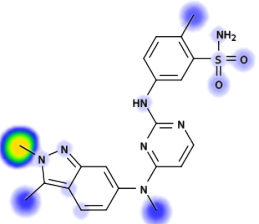
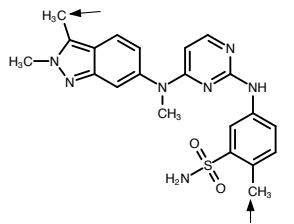
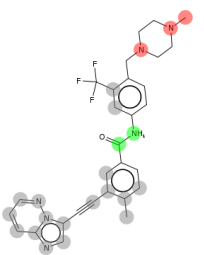
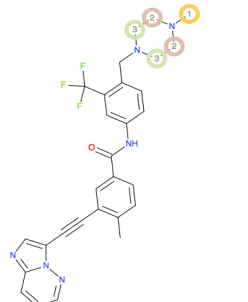
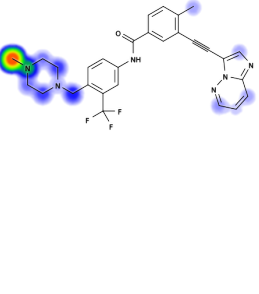
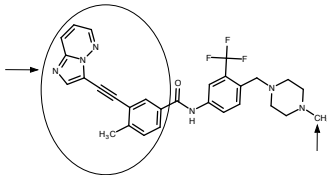
KI	MetaPrint2D	RS-WebPredictor 1.0	Xenosite	Experimental site(s) of metabolism
Afatinib				
Alectinib				

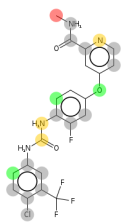
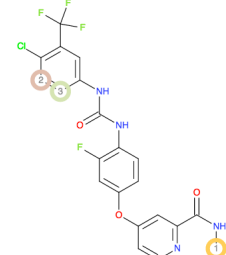
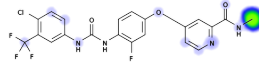
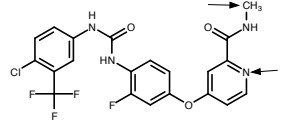
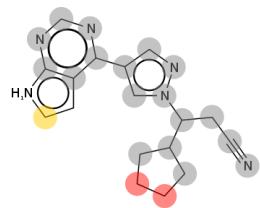
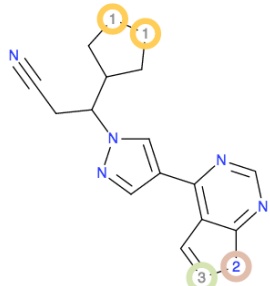
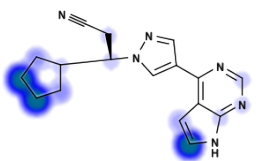
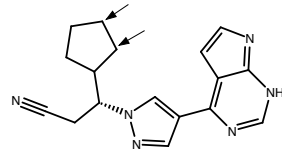
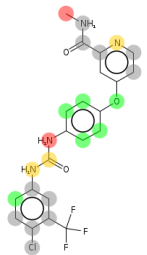
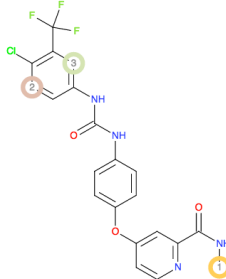
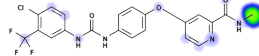
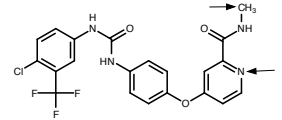
<p>Axitinib</p>				
<p>Bosutinib</p>				
<p>Cabozantinib</p>				

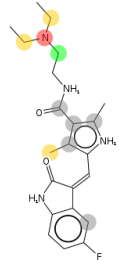
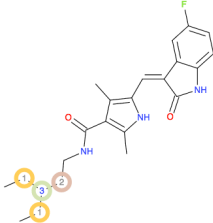
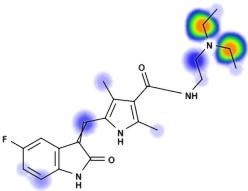
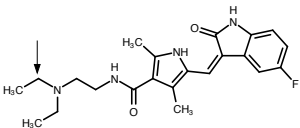
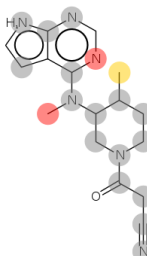
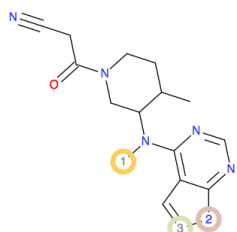
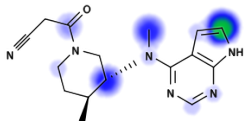
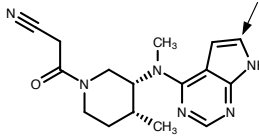
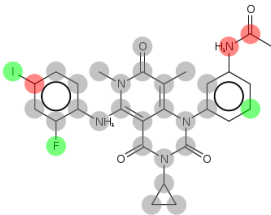
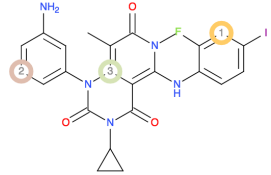
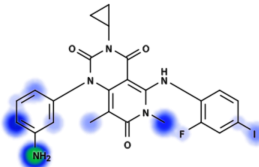
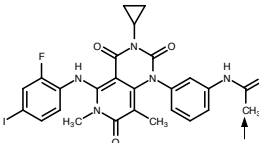
<p>ceritinib</p>				
<p>cobimetinib</p>				
<p>Crizotinib</p>				
<p>Dabrafenib</p>				

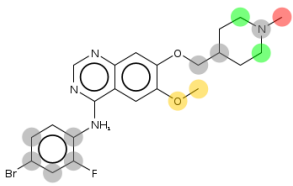
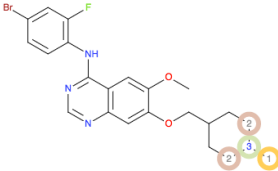
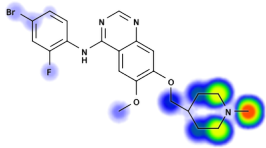
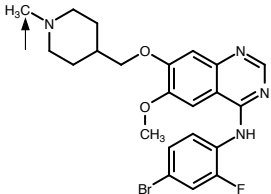
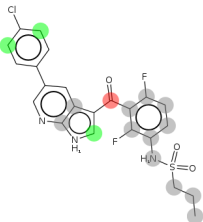
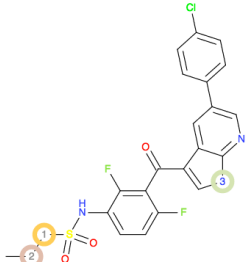
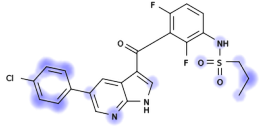
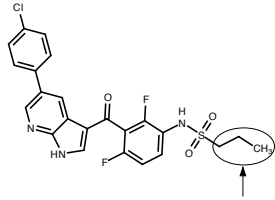
<p>Dasatinib</p>				
<p>erlotinib</p>				
<p>Gefitinib</p>				
<p>Ibrutinib</p>				

<p>Nilotinib</p>				
<p>Nintedanib</p>				
<p>Osimertinib</p>				

<p>Palbociclib</p>				
<p>pazopanib</p>				
<p>Ponatinib</p>				

<p>Regorafenib</p>				
<p>Ruxolitinib</p>				
<p>sorafenib</p>				

<p>Sunitinib</p>				
<p>Tofacitinib</p>				
<p>Trametinib</p>				

vandetanib				
Vemurafenib				

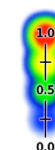
MetaPrint2D

The color highlighting an atom indicates its normalised occurrence ratio (NOR). A high NOR indicates a more frequently reported SOM in the metabolic database.

Red	0.66 <= NOR <= 1.00
Orange	0.33 <= NOR < 0.66
Green	0.15 <= NOR < 0.33
White	0.00 <= NOR < 0.15
Grey	Little/no data

Xenosite

A probability or statistical likelihood of metabolism occurring at a particular atom (ranging from 0 to 1) is denoted by a colour code.



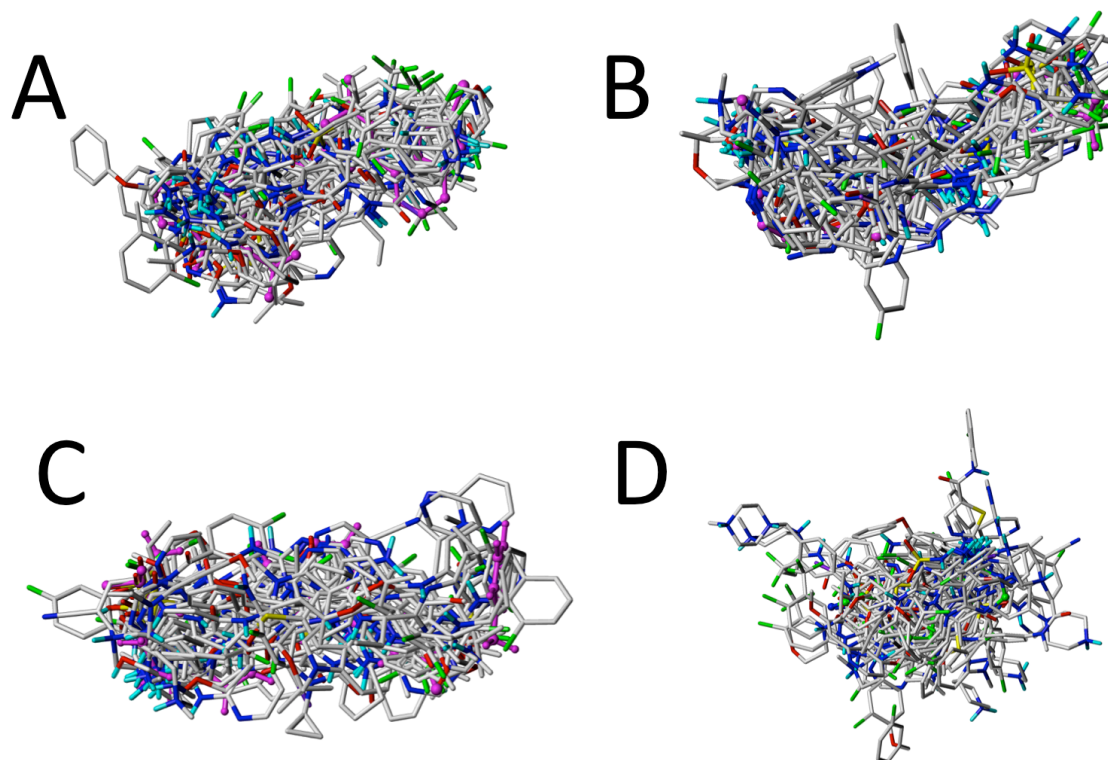


Figure S1 Ligand superposition by morphological similarity using different templates A) bromoergocryptine B) sorafenib C) laptatinib D) tofacitinib. Templates are shown in ball and stick green (bromoergocryptine, sorafenib, laptatinib) and magenta (tofacitinib). Dataset molecules (overlaid on template) are shown in sticks with C, O, N, Cl/F in white, red, blue, and green respectively.