



Review

Descriptors of Cytochrome Inhibitors and Useful Machine Learning Based Methods for the Design of Safer Drugs

Tyler C. Beck ^{1,2} , Kyle R. Beck ³, Jordan Morningstar ², Menny M. Benjamin ^{1,2} and Russell A. Norris ^{2,*}

- ¹ Drug Discovery & Biomedical Sciences, Medical University of South Carolina, 280 Calhoun Street, QF204, Charleston, SC 29424-2303, USA; beckt@muscd.edu (T.C.B.); benjamim@muscd.edu (M.M.B.)
- ² Regenerative Medicine & Cell Biology, Medical University of South Carolina, 171 Ashley Avenue, 604F CRI, Charleston, SC 29424-2303, USA; morningj@muscd.edu
- ³ College of Pharmacy, The Ohio State University, 217 Lloyd M Parks Hall, 500 West 12th Ave., Columbus, OH 43210, USA; beck.735@buckeyemail.osu.edu
- * Correspondence: norrisra@muscd.edu

Abstract: Roughly 2.8% of annual hospitalizations are a result of adverse drug interactions in the United States, representing more than 245,000 hospitalizations. Drug–drug interactions commonly arise from major cytochrome P450 (CYP) inhibition. Various approaches are routinely employed in order to reduce the incidence of adverse interactions, such as altering drug dosing schemes and/or minimizing the number of drugs prescribed; however, often, a reduction in the number of medications cannot be achieved without impacting therapeutic outcomes. Nearly 80% of drugs fail in development due to pharmacokinetic issues, outlining the importance of examining cytochrome interactions during preclinical drug design. In this review, we examined the physiochemical and structural properties of small molecule inhibitors of CYPs 3A4, 2D6, 2C19, 2C9, and 1A2. Although CYP inhibitors tend to have distinct physiochemical properties and structural features, these descriptors alone are insufficient to predict major cytochrome inhibition probability and affinity. Machine learning based in silico approaches may be employed as a more robust and accurate way of predicting CYP inhibition. These various approaches are highlighted in the review.

Keywords: CYP3A4; CYP2D6; CYP2C19; CYP2C9; CYP1A2; cheminformatics; cytochrome



Citation: Beck, T.C.; Beck, K.R.; Morningstar, J.; Benjamin, M.M.; Norris, R.A. Descriptors of Cytochrome Inhibitors and Useful Machine Learning Based Methods for the Design of Safer Drugs.

Pharmaceuticals **2021**, *14*, 472. <https://doi.org/10.3390/ph14050472>

Academic Editors: In-hwan Baek and Jung-woo Chae

Received: 10 April 2021

Accepted: 13 May 2021

Published: 17 May 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Roughly seventy-four percent (74%) of all physician office visits involve drug therapy, with 2.9 billion drugs prescribed annually [1]. These prescriptions accounted for USD 335 billion in patient costs in 2020 [2]. Half of Americans take at least one prescription drug and nearly a quarter of persons use three or more medications concomitantly [1]. An estimated 6.7% of hospitalized patients experience an adverse drug reaction with a fatality rate of 0.32%, responsible for 106,000 deaths annually [3]. These data suggest adverse drug reactions are the fourth leading cause of death in the United States [3–5]. Furthermore, these statistics exclude ambulatory adverse drug reactions and are therefore likely an underrepresentation of the morbidity and mortality burden of adverse drug reactions. Nearly 350,000 adverse drug reactions occur in nursing homes annually, where polypharmacy is commonplace [6]. Many of these drug–drug interactions are avoidable with proper patient medication management and by deprescribing. While efforts to reduce drug–drug interactions are important, a reduction in the number of medications cannot always be achieved without impacting therapeutic outcomes.

Adverse drug reactions commonly occur due to drug–drug interactions as a result of major cytochrome P450 (CYP) inhibition [7]. CYPs are heme-containing membrane-bound enzymes that predominantly reside in the smooth endoplasmic reticulum and mitochondria of hepatocytes and in the intestines [8,9]. In mammals, 57 CYP isoforms have been identified with the function of performing the oxidative metabolism of xenobiotics and

endogenous compounds [8]. Of the 57 CYP isoforms, 5 CYPs (CYPs 3A4, 2D6, 2C19, 2C9, and 1A2) are responsible for the metabolism of more than 80% of clinically used drugs [7] (Figure 1). Enzyme inhibition occurs when two co-administered drugs share an identical mechanism of biotransformation and compete for metabolism in the same enzyme receptor site [9]. As such, the more potent inhibitor will prevail, interrupting biotransformation, resulting in the diminished metabolism of the competing drug. This results in elevated serum drug levels of the unmetabolized drug, increasing the probability of adverse toxicological outcomes. This is especially important for drugs with a narrow therapeutic index, such as chemotherapeutics. It is important to note that inhibitors of major cytochromes can promote adverse drug reactions when administered as monotherapy. Inhibitors of major cytochromes may alter their own metabolism at high doses, leading to toxic plasma drug accumulation. These interactions pose a significant safety risk and should be examined closer during the initial drug design process. The CYP system and drug biotransformation has been reviewed extensively elsewhere [10].

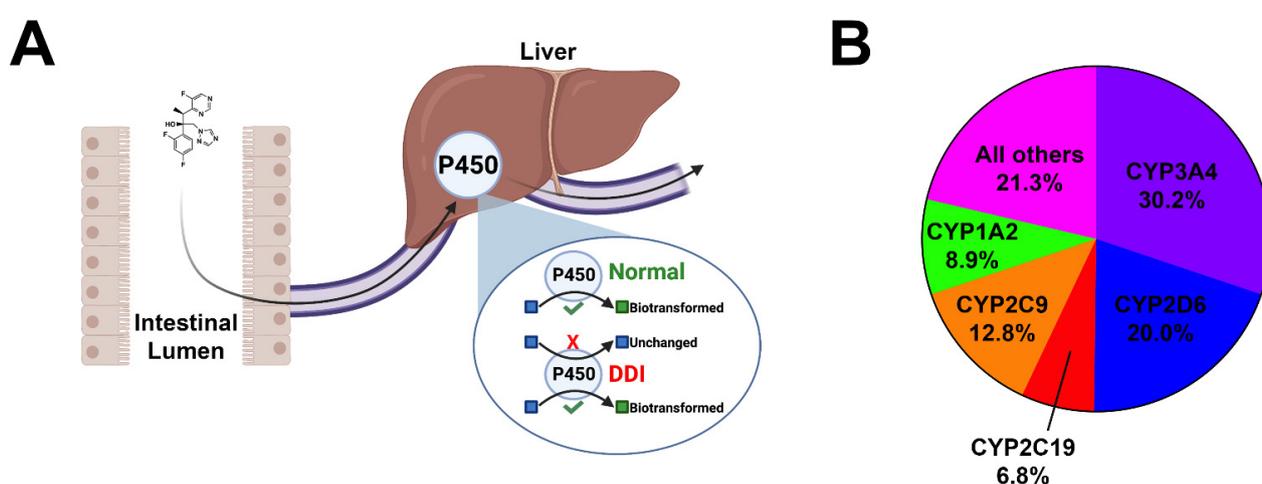


Figure 1. Cartoon depiction of cytochrome inhibition and pie chart. (A) Orally bioavailable drugs are absorbed from the intestinal lumen into the mesenteric capillaries and transported to the liver via the portal vein. Many drugs undergo phase I biotransformation by major cytochromes within hepatocytes. This process enzymatically converts lipid-soluble compounds to more water-soluble compounds to facilitate the excretion. In the event of polypharmacy and/or the co-administration of CYP inhibitors, serum accumulation of unmetabolized drugs may occur, leading to untoward toxicities. (B) Pie chart demonstrating the breakdown of the five major cytochromes and their contributions to the oxidative metabolism of xenobiotics and organic endogenous molecules.

Traditionally, pharmacokinetics has been neglected in preclinical drug design. Roughly four in five investigational new drugs fail as a result of absorption, distribution, metabolism, excretion and toxicity (ADMET) issues [11]. Tools that assist in ADMET prediction may be advantageous in early drug design by reducing the drug attrition rate in advanced preclinical- and early clinical-stage studies. In particular, the design of drugs devoid of major cytochrome inhibition would reduce the burden of drug–drug interactions in patients where polypharmacy is unavoidable. Several reviews exist that focus on platforms that utilize a diverse range of prediction methods [12–20]. Many of these reviews focus on outdated tools, as well as platforms that are exclusively available through a fee-based licensing model. More recently, numerous machine learning-based platforms have emerged that are user-friendly, are not as computationally expensive as pre-existing methods, and demonstrate improved prediction accuracy. In this review, we will explore the various physiochemical and structural descriptors of CYP inhibitors, as well as superior machine learning prediction platforms freely available to investigators.

2. Physicochemical Properties

Historically, physicochemical properties have been used as key descriptors to predict CYP inhibition liability. Cytochrome P450 enzymes are inherently lipophilic and thus have a predilection for lipophilic substrates [21–25]. It has been hypothesized that molecular weight and lipophilicity represent key physicochemical characteristics useful in predicting CYP binding affinity [21,26–28]. In further exploring this hypothesis, we examined CYP inhibitors deposited in the ChEMBL database to further illustrate the relationship between molecular weight, lipophilicity, and the IC₅₀ of various major cytochromes (Figure 2). ChEMBL hosts a database of over one million small molecules with reported bioactivities and calculated physicochemical properties [29]. Data on the inhibitors of each individual cytochrome were extracted for further analysis. Duplicates, as well as data points that were flagged as potentially unreliable by ChEMBL, were withheld from the analysis. The inhibitors of major cytochromes tend to be lipophilic in nature (Figure 2A). Inhibitors of CYP3A4, CYP2C19, and CYP2C9 were especially lipophilic (ALogP = 4.21, 4.18, and 4.28, respectively). Inhibitors of CYP3A4 and CYP2C9 tend to be larger structures, whereas inhibitors of CYP1A2 tend to be small. These data make sense considering the shape and size of the CYP 3A4, 2C9, and 1A2 enzyme receptor sites, which will be reviewed in Section 3. It should be noted, however, that although inhibitors of CYP3A4 and CYP2D6 tend to be larger in size, many of their substrates are small-sized molecules, such as furanocoumarin, diltiazem, mifepristone, bupropion, and fluoxetine.

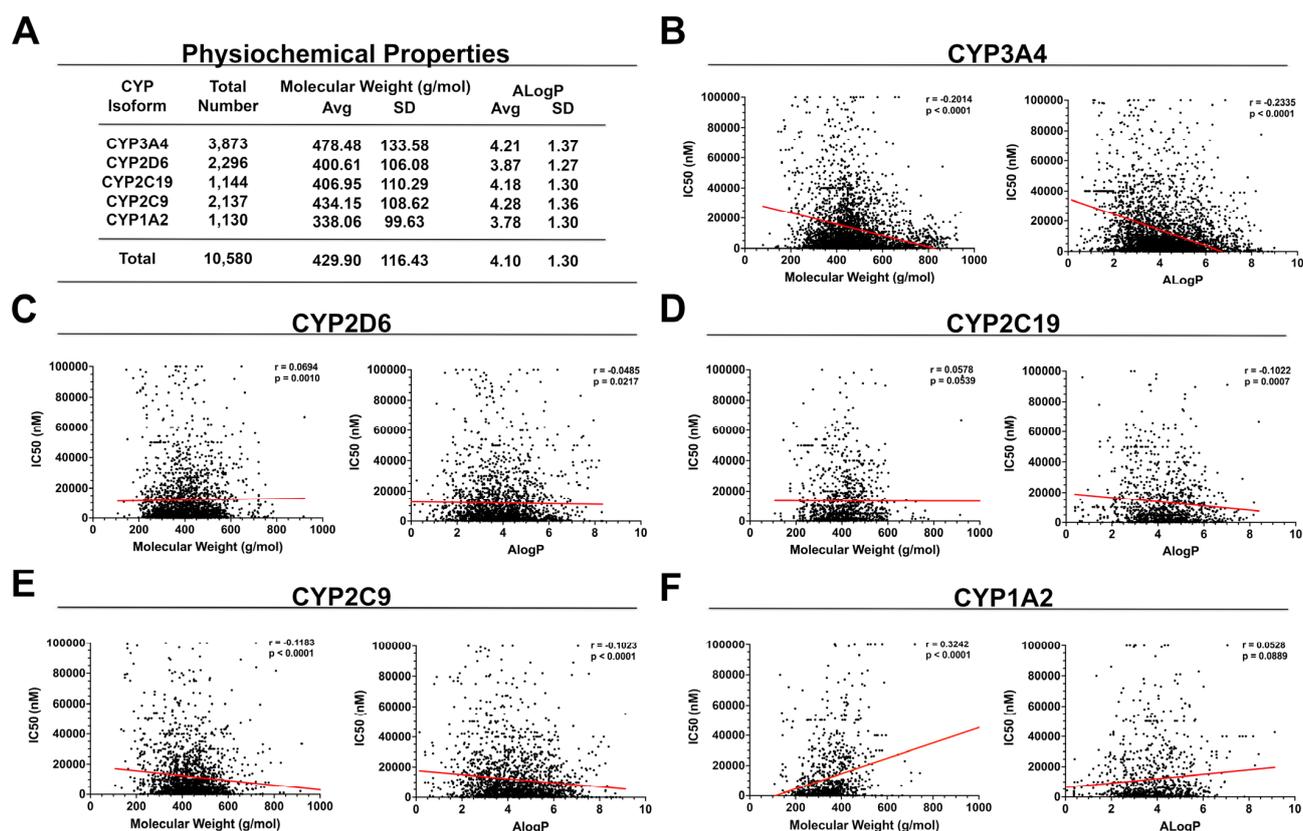


Figure 2. Summary of the physicochemical properties of major cytochrome inhibitors and their impact on half-maximal inhibitor concentration (IC₅₀). (A) Molecular weight and lipophilicity of validated inhibitors of each individual cytochrome. (B–F) Graphs representing IC₅₀ versus either molecular or ALogP for CYPs 3A4 (B), 2D6 (C), 2C19 (D), 2C9 (E), and 1A2 (F). CYP3A4 and CYP2C9 inhibitors demonstrated a weak statistically significant negative relationship between molecule size, lipophilicity, and IC₅₀ (i.e., larger, more lipophilic molecules, tend to have lower IC₅₀ values); however, the exact opposite trend was observed for CYP1A2. No relationship was identified for CYPs 2D6 and 2C19.

We then examined the relationship between molecular weight, lipophilicity (A_{LogP}) and IC₅₀ by performing simple linear regression and Pearson's correlation analyses. In doing so, we identified several statistically significant correlations between physiochemical properties and CYP half-maximal inhibitory concentration; however, these mild correlations do not adequately explain the majority of the data (Figure 2B–F). Next, we performed a multiple linear regression analysis for each major cytochrome to further assess how multiple explanatory variables, such as molecular weight and lipophilicity, impact the half-maximum inhibitory concentration (Supplemental Materials Tables S1–S5). A statistically significant relationship ($p < 0.05$) between molecular weight, lipophilicity, and IC₅₀ was identified for CYPs 3A4 ($R = 0.136$), 2C19 (0.125), 2C9 (0.132), and 1A2 (0.309). A statistically significant relationship was not identified for CYP2D6 ($R = 0.024$, $p = 0.510$). As observed with the simple linear regression data, exclusively weak relationships were observed that poorly describe the data. Although CYP inhibitors tend to fit a distinct physiochemical profile, these descriptors alone are insufficient to predict the probability of cytochrome inhibition and CYP half-maximal inhibitory concentration.

3. Structural Properties

The individual CYP isoforms come in different shapes, sizes, and compositions, contributing to unique substrate specificity profiles. The key structural features of a few important CYP inhibitors have been elucidated using X-ray crystallography, pharmacophore modeling, and various other *in silico* approaches [30–38]. In particular, the structural features of CYP3A4 inhibitors have been well studied [30,39–43]; however, data on the specific structural features of inhibitors for the other four major cytochromes are limited and warrant further investigation. In this section, we review the structural features of CYP enzymes and highlight the structural features of well-known, potent CYP inhibitors. In particular, emphasis is placed on molecular weight, surface area, lipophilicity (LogP), rotatable bonds, hydrogen bond acceptors, and hydrogen bond donors. The molecular weights and surface areas of small molecules are key driving factors that dictate the strength of binding interactions. In general, an enzyme or protein with a small binding cavity will be more compatible with a smaller drug, etc. Lipophilicity is an important descriptor that is used to predict CYP interactions. A majority of CYP inhibitors possess numerous aromatic rings and hydrophobic moieties capable of interacting with lipophilic CYP-active sites. Rotatable bonds determine the flexibility of a drug molecule and are important for inhibitor binding interactions with certain CYPs (i.e., CYP3A4), and less important for other CYPs (i.e., CYPs 2D6 and 1A2). Lastly, hydrogen bond acceptor and donor moieties form specific binding interactions with active site amino acids. These interactions were further explored using molecular docking, which was performed using previously described methods [44–46].

3.1. CYP3A4

CYP3A4 is the most widely expressed and clinically significant cytochrome responsible for the biotransformation of 30.2% of xenobiotics that undergo hepatic clearance [47]. CYP3A4 is capable of oxidizing a wide array of bulky substrates with many different structural features. The versatile functionality of CYP3A4 is due to its large substrate-binding cavity [47,48]. The active site cavity of CYP3A4 is open and flexible, capable of accommodating volumes ranging from 1173 to 2862 cubic angstroms (Å³) [39]. The CYP3A4–ligand complexes formed upon substrate binding are incredibly flexible, explaining the observed ligand promiscuity and diverse range of substrates capable of high-affinity binding interactions [49]. In further examining the 20 most frequently prescribed CYP3A4 inhibitors, we identified shared features in comparing the compounds (Figure 3). The most frequently prescribed high-affinity CYP3A4 inhibitors have a very high molecular weight (595.7 g/mol) and a large surface area (251.3 Å²) (Figure 4). Additionally, these structures are highly lipophilic (LogP = 4.0) and have many rotatable bonds and hydrogen bond acceptors (9.7 and 8.6, respectively). These hydrogen bond acceptors are available to interact with key amino acids within the CYP3A4 active site, including arginine 106 (R106), serine 119 (S119),

threonine 309 (T309), and arginine 372 (R372). These data are consistent with previous reports, suggesting that potent inhibitors of CYP3A4 are large, lipophilic molecules with a significant degree of flexibility. It is important to note that while CYP3A4 inhibitors are capable of diverse binding poses, there are a few high-energy binding interactions that occur among all high-affinity binders. Molecules with a tertiary amine or imidazole nitrogen frequently coordinate with heme-601 (HEM601) [12,50]. High-affinity CYP3A4 inhibitors are lipophilic, containing one or more aromatic rings capable of forming numerous pi-pi stacking interactions with active phenylalanine (F) residues, including F108, F213, F215, F304, and occasionally F57. A majority of high-affinity inhibitors make polar interactions with D76, R106, S119, and R372. Additionally, CYP3A4 inhibitors typically possess a nitrogen or allylic position as a site of oxidation [48]. Lastly, high-affinity CYP3A4 inhibitors have approximately two hydrogen bond acceptors. These moieties commonly accept hydrogens from aspartate 76 (D76) and E374. Due to the highly flexible nature of the enzyme-binding pocket and the diversity of ligand-binding interactions, predicting drug-drug interactions involving CYP3A4 is difficult.

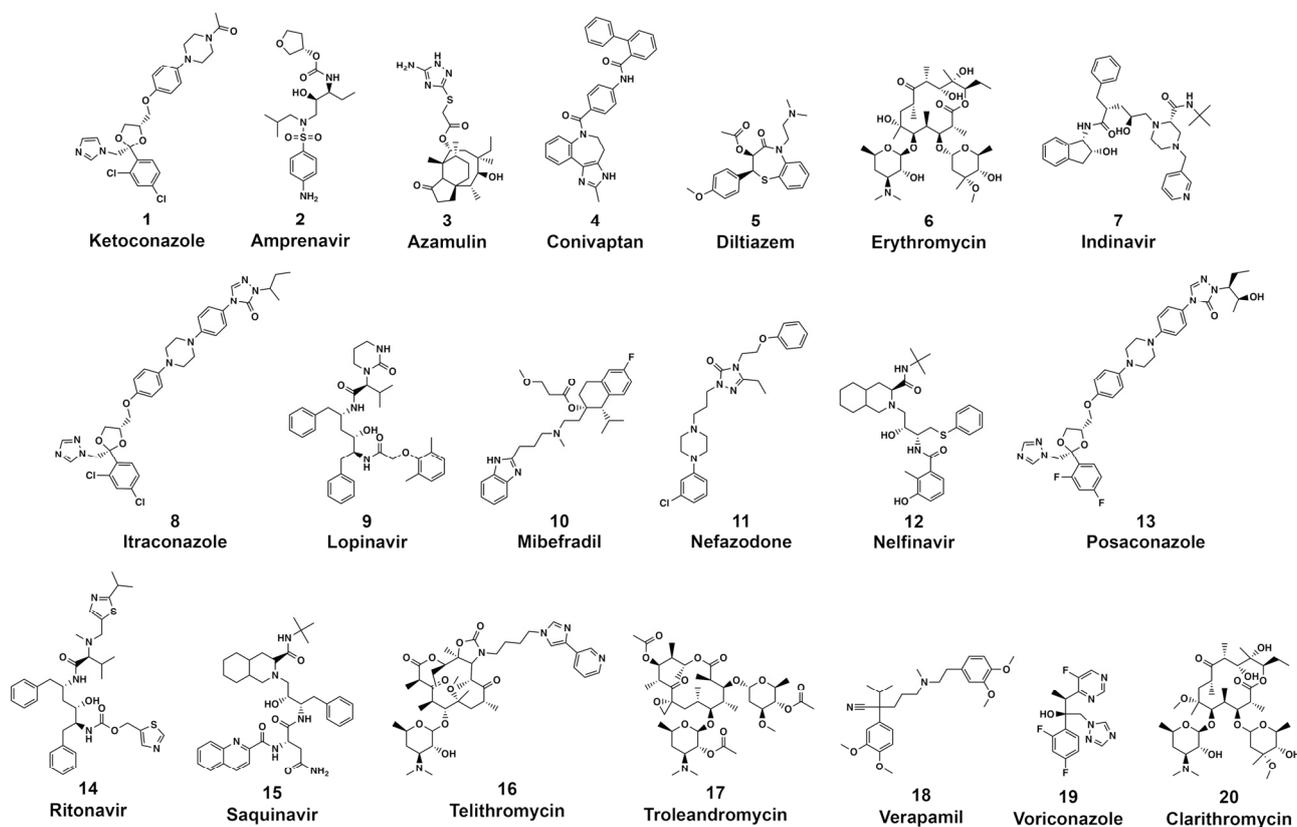


Figure 3. Structures of the 20 most commonly prescribed high-affinity CYP3A4 inhibitors.

In addition to hepatic metabolism, CYP3A4 drug metabolism also occurs in the gastrointestinal (GI) tract. This is true for the oral chemotherapeutic ibrutinib, used to treat chronic lymphocytic leukemia (CLL), which makes roughly equal contributions to hepatic and intestinal CYP3A4 metabolism [51]. Moreover, some agents exclusively inhibit CYP3A4 in the GI tract (minimally effective hepatic CYP3A4), resulting in increased drug concentrations. One specific example is the co-administration of grapefruit juice with drugs that are metabolized by CYP3A4. Grapefruit juice irreversibly inhibits intestinal, but not hepatic, CYP3A4, increasing systemic drug exposure [52]. These examples highlight the significance of CYP3A4 inhibition in tissues outside of the liver.

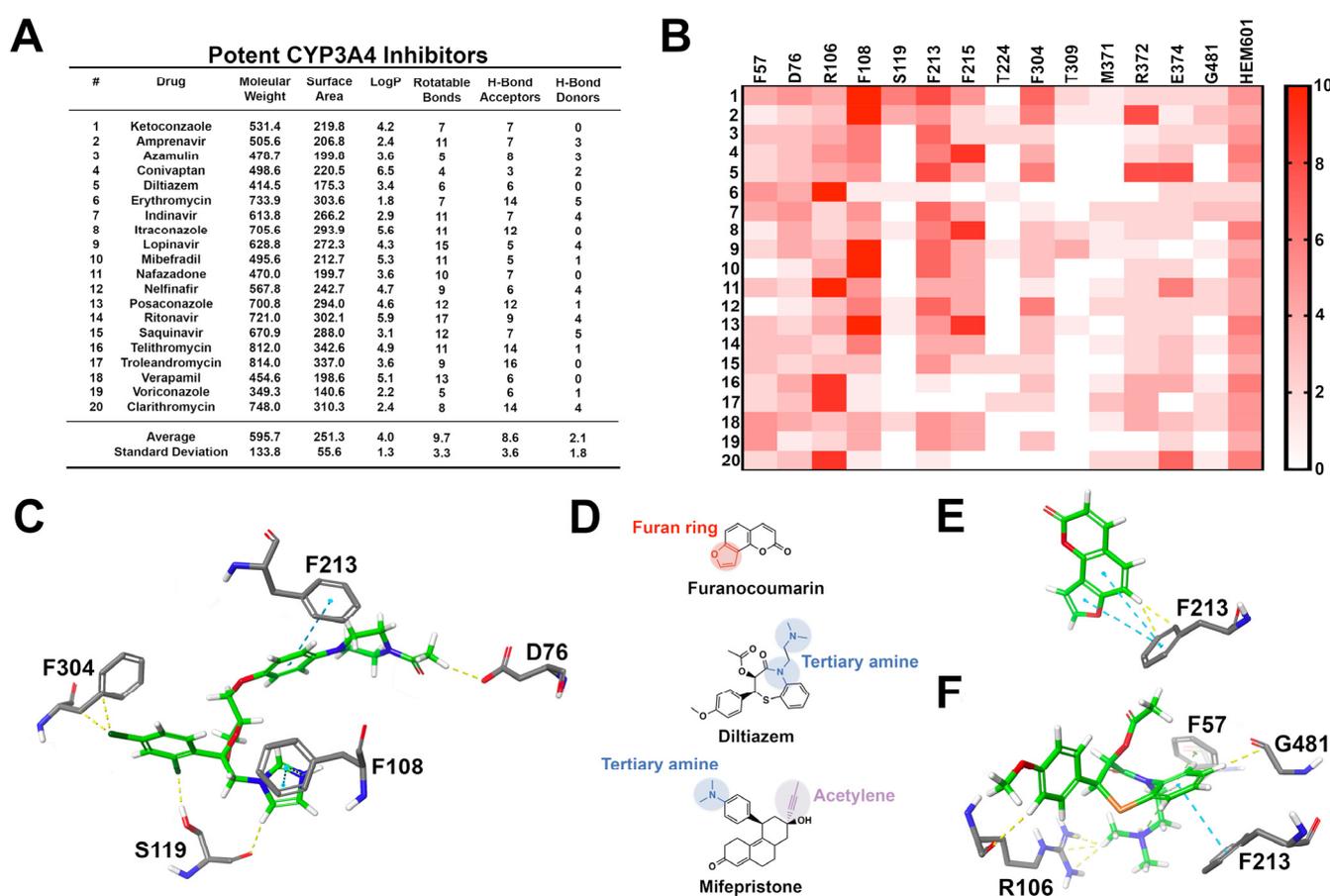


Figure 4. Structural investigation of the 20 most commonly prescribed CYP3A4 inhibitors. (A) CYP3A4 inhibitors are large, lipophilic, and flexibility molecules capable of accepting numerous hydrogen bonds. Molecular weight is represented as g/mol or Daltons (Da) and surface area is represented in angstroms (\AA). Values were calculated using pkCSM ADMET [38]. (B) Heat map representing the frequency of ligand-to-amino acid binding interactions in the top ten most energetically favorable binding poses in docking against CYP3A4 (PDB ID: 1W0E). (C) Docking image demonstrating ketoconazole, a potent CYP3A4 inhibitor, making numerous pi–pi stacking interactions (blue dotted lines) with F108, F213, and F304, as well as energetically favorable contacts (yellow dotted lines) with D76 and S119. (D) Moieties commonly associated with high-affinity CYP inhibition in the literature, including the presence of a furan ring, tertiary amine, or acetylene group [39]. (E) 3D docking image of furanocoumarin, a CYP3A4 inhibitor found in the skin of grapefruit. Furanocoumarin makes several pi–pi stacking interactions and energetically favorable contacts. The furan ring of furanocoumarin specifically lends to pi–pi stacking interactions within the CYP3A4 receptor site. (F) 3D docking image of diltiazem, an antihypertensive drug and calcium channel blocker known for inhibiting CYP3A4. Diltiazem forms many pi–pi stacking interactions and energetically favorable contacts.

3.2. CYP2D6

CYP2D6 comprises 2% of hepatocyte CYP content and is responsible for the biotransformation of 20% of drugs that undergo hepatic metabolism [53]. The active site cavity of CYP2D6 has a more restricted volume of 510 cubic angstroms (\AA^3) [39,54]. The active site of CYP2D6 is acidic in nature and contains several phenylalanine residues capable of pi–pi stacking with substrates [55]. As such, it has been reported that CYP2D6 has a predilection for basic substrates [56]. In examining six well-known high-affinity CYP2D6 inhibitors, the structural similarities include a primary or secondary amine, heterocycles, and numerous aromatic moieties (Figure 5). These CYP2D6 inhibitors are relatively small in composition and are moderately lipophilic. It should be noted that in examining 2296 CYP2D6 inhibitors, the average size was 400.61 Da, in comparison to 319.6 Da observed in the six most frequently prescribed drugs. Regardless of the discrepancy, the typical

CYP2D6 inhibitor is smaller than the average FDA-approved drug (480 Da) [57]. Additionally, CYP2D6 inhibitors are typically flat, planar structures capable of procuring a positive charge. The active site contains two carboxylic acid residues from Q216 and D301 that are critical for substrate recognition and binding selectivity [12]. These acidic residues frequently interact with the primary and/or secondary amine groups present on CYP2D6 inhibitors. The aromatic moieties make energetically favorable pi–pi stacking interactions with F120, F483, and less commonly F112. Potent inhibitors commonly contain heterocycles, primary amines, and carbonyl groups that coordinate with the iron of heme-601 (HEM601) within the CYP2D6-active site. Additionally, the average inhibitor contained 4.2 rotatable bonds and 3.3 hydrogen bond acceptors, most frequently interacting with glutamine 214 (Q214) and T309. CYP2D6 inhibitors contain approximately one hydrogen bond donor, which occasionally interacts with aspartate 301 (D301). Lastly, valine 374 (V374) forms van der Waals interactions with hydrophobic residues of inhibitors.

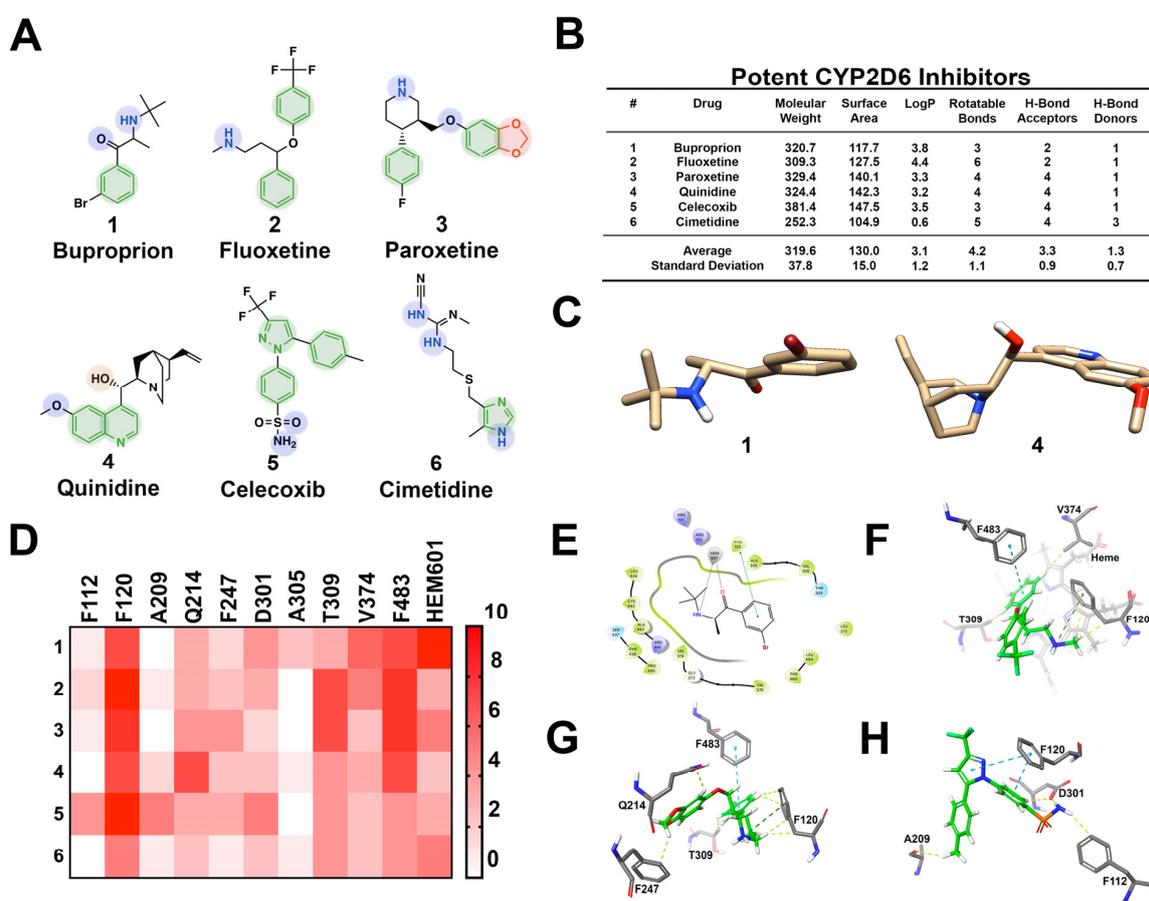


Figure 5. Structural investigation of 6 well-known high-affinity CYP2D6 inhibitors. **(A)** CYP2D6 inhibitors commonly have one or multiple aromatic moieties, heterocycles, and primary or secondary amines. Red and green moieties interact with F112, F120, F247, F483, and/or heme-601; blue moieties interact with D301 and/or heme-601, and orange interact with D301. **(B)** CYP2D6 inhibitors are rigid, moderately sized structures with varying physiochemical properties. **(C)** CYP2D6 inhibitors are flat, planar structures capable of becoming positively charged. **(D)** Heat map representing the frequency of ligand-to-amino acid binding interactions in the top ten most energetically favorable binding poses in docking against CYP2D6 (PDB ID: 3QM4). Key amino acids include F120, Q214, D301, T309, V374, and F483. Frequent interactions are observed with heme as well. **(E)** Two-dimensional ligand interaction map showing the secondary amine and carbonyl group of bupropion coordinating with Heme-601. **(F)** Three-dimensional docking image depicting fluoxetine interaction with F120, T309, V374, and F483. **(G)** Three-dimensional docking image showing paroxetine forming pi–pi stacking interactions with F120, F247, and F83, as well as energetically favorable contacts with Q214 and T309. **(H)** Three-dimensional docking image showing celecoxib interacting with F120, F112, A209, and D301.

As described with CYP3A4, it is difficult to predict half-maximal inhibition concentration using physiochemical and structural features alone given the diverse binding capabilities of CYP2D6. A study by Vandenbrink et al. demonstrated that CYP2D6's inhibition potency is substrate-dependent [58]. In examining the binding kinetics of 20 known inhibitors of CYP2D6 in the presence of four clinically relevant substrates, atypical inhibition kinetics were observed for each individual inhibitor–substrate pairing. The substrates with constrained amine moieties impaired inhibitor binding affinity by more than 5-fold when compared to other structurally unrelated substrates. These data add an additional layer of difficulty in making reliable predictions for cytochromes such as CYP2D6 that interact with a wide range of structurally diverse substrates. Lastly, CYP2D6 is highly polymorphic, with the largest phenotypic variation among the human CYPs [59]. Approximately 80 different allelic variants and 130 genetic variations have been characterized in patients, with phenotypic consequences ranging from loss of function to ultra-rapid metabolism. Therefore, it should be acknowledged that while CYP2D6 predictions may be accurate in a majority of cases, the data will be subject to significant variability due to the genetic heterogeneity present in the global population, which will be an unavoidable roadblock in preclinical drug design.

3.3. CYP2C19

CYP2C19 metabolizes 6.8% of drugs that undergo hepatic biotransformation. Despite being considered a minor hepatic cytochrome based on its relative contribution, CYP2C19 is responsible for metabolizing several clinically important drugs, including omeprazole and phenytoin [60]. As such, many pharmacologists view CYP2C19 as a major cytochrome. Inhibitors of CYP2C19 share frequently observed structural features (Figure 6). Structures contain numerous aromatic moieties, heterocycles, carbonyl groups, and aromatic nitrogen atoms. Several well-characterized high-affinity CYP2C19 inhibitors are small-sized molecules, but they tend to be less lipophilic than inhibitors of other cytochromes. In general, however, CYP2C19 inhibitors tend to be medium-sized molecules with variable lipophilicity. These results are consistent with previous reports, claiming that lipophilicity does not impact CYP2C19 inhibitory potency [12,61]. On average, these molecules are less flexible than typical CYP inhibitors (3.8 rotatable bonds) and have approximately three hydrogen bond acceptors. The hydrogen bond acceptor moieties commonly interact with asparagine 204 (N204). It is important to note that the location of the hydrogen bond acceptors helps dictate binding affinity. CYP2C19 inhibitors should have a minimum of two hydrogen bond acceptors within close proximity to a site of oxidation [12]. CYP2C19 inhibitors contain heterocycles and heterocyclic nitrogen atoms that frequently coordinate with Heme-501 (HEM501). Numerous pi–pi stacking interactions occur between small-molecule arylene groups and F114 and F476. Additionally, G296 forms frequent interactions with hydroxyl and carbonyl groups present on CYP2C19 inhibitors. CYP2C19 inhibitors make numerous hydrophobic contacts with isoleucine 205 (I205), alanine 292 (A292), glycine 296 (G296), A297, and I362.

Similar to CYP2D6, a variety of genetic polymorphisms exist within the CYP2C19 isoform. Roughly 5% of Caucasians and 20% of Asian populations possess genetic variations resulting in reduced or complete loss of enzyme function [61–64]. Although less common, genetic variations may also result in enhanced metabolic activity, which also impacts drug handling and excretion. These polymorphisms complicate predictions at the population level and should be taken into consideration.

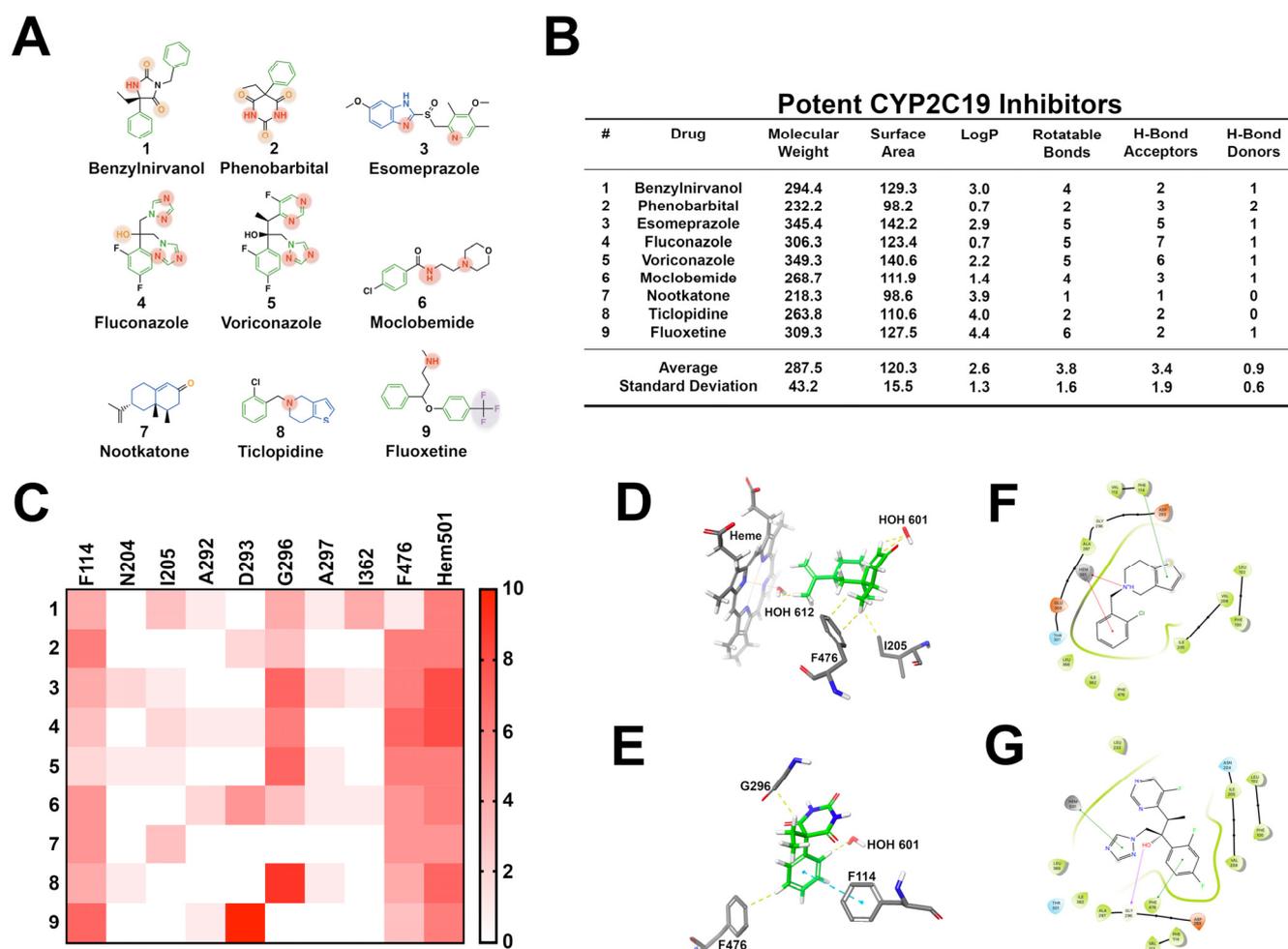


Figure 6. Structural investigation of 9 well-known high-affinity CYP2C19 inhibitors. (A) Common structural features seen in CYP2C19 inhibitors include aromatic moieties, heterocycles, carbonyl groups, and aromatic nitrogens. Green moieties interact with F114, F476, and/or heme; blue and red moieties interact with heme; orange interact with G296, and purple interact with water and I205. (B) The nine most frequently prescribed CYP2C19 inhibitors are small; however, in examining 1144 CYP219 inhibitors, the average molecule is moderately sized. (C) Heat map representing the frequency of ligand-to-amino acid binding interactions in the top ten most energetically favorable binding poses in docking against CYP2C19 (PDB ID: 4GQS). Key amino acids include F114, G296, F476. Frequent interactions are observed with heme-501 as well. (D) Three-dimensional ligand docking showing nootkatone interacting with I205, F476 and water molecules within the active site. (E) Three-dimensional docking image depicting phenobarbital interacting with F114, G296, F476, and water molecules within the active site. (F) Two-dimensional ligand interaction diagram showing the cyclic nitrogen and arylene group of ticlopidine coordinating with heme-501. (G) Two-dimensional interaction map showing voriconazole coordinating with heme, while simultaneously interacting with G296 and F476.

3.4. CYP2C9

CYP2C9 is responsible for the biotransformation of 12.8% of drugs that undergo hepatic metabolism. In particular, CYP2C9 plays a key role in the metabolism of commonly used medications, such as non-steroidal anti-inflammatory drugs (NSAIDs) and warfarin [12]. The active site of CYP2C9 is rather large and has a cavity volume ranging from 978 to 1271 cubic angstroms (\AA^3) [39,65,66]. Inhibitors of CYP2C9 are typically aromatic compounds with heterocycles, aromatic nitrogens, primary amines, and halogens (Figure 7). Heterocycles, aromatic nitrogens, and primary amines coordinate to the iron of heme-501 (HEC501). Aromatic moieties form extensive pi–pi stacking interactions with F114 and F476. Halogens commonly interact with T364. Additionally, carbonyl, hydroxyl, and carboxylate groups interact with R108. These particular interactions lock the substrate

into a position with close proximity to the iron allowing for oxidative metabolism to occur [67]. CYP2C9 inhibitors perform a large array of binding poses; however, binding with R108, F114, and R476 is most commonly observed in the most energetically favorable docking interactions. CYP2C9 inhibitors tend to be medium-sized semi-flexible molecules that are moderately lipophilic and acidic, with a minimum of two to four hydrogen bond acceptors. These hydrogen bond acceptors most commonly interact with R97, R108, N217, T364, and S365. As seen with the other major cytochromes, CYP2C9 is highly polymorphic, with more than 20 activity-altering genetic variants identified [67,68].

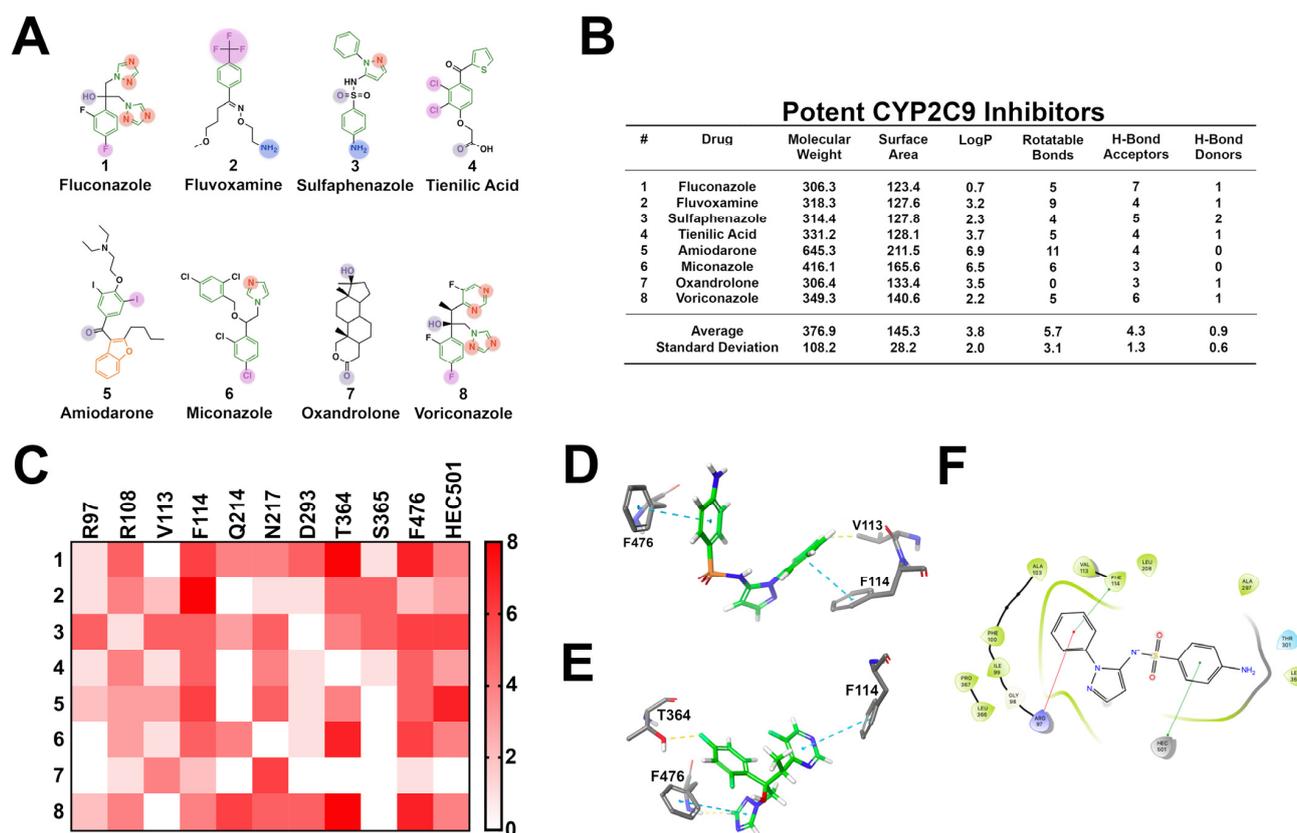


Figure 7. Structural investigation of 8 well-known high-affinity CYP2C9 inhibitors. **(A)** The common structural features seen in CYP2C9 inhibitors include aromatic rings, aromatic nitrogens, heterocycles, primary amines, and halogens. Purple moieties interact with N217; magenta interact with T364; green interact with F114, F476, and/or heme; blue interact with S365; red interact with Q214; and orange interact with heme. **(B)** Molecules tend to be moderately sized with varying lipophilicity. CYP2C9 inhibitors are moderately flexible and possess numerous H-bond acceptors. **(C)** Heat map representing the frequency of ligand-to-amino acid binding interactions in the top ten most energetically favorable binding poses in docking against CYP2C9 (PDB ID: 1OG5). Ligand interactions are highly diverse, with key interactions including R108, F114, F476, and heme-501. **(D)** Three-dimensional docking image depicting sulfaphenazole forming pi-pi stacking interactions between F114 and F476, as well as energetically favorable contacts with V113. **(E)** Three-dimensional docking image portraying voriconazole interacting with F114, T364, and F476. **(F)** Two-dimensional ligand interaction map depicting sulfaphenazole interacting with R97 and F114, while coordinating with the iron of heme-501.

3.5. CYP1A2

CYP1A2 is responsible for the biotransformation of 8.9% of drugs that undergo hepatic metabolism, including antipsychotics and antibiotics [69]. Additional CYP1A2 substrates include caffeine, paracetamol, and aflatoxin B1. Interestingly, CYP1A2 activity is commonly induced by dietary constituents, including various vegetables and polyaromatic hydrocarbons, whereas ingredients such as cumin and turmeric can inhibit enzyme function [70,71]. The active site of CYP1A2 is small and planar with a volume of 375 to 390 cubic angstroms

(Å³), and thus CYP1A2 substrates tend to be small, planar molecules [12,39]. In examining eight well-known high-affinity CYP1A2 inhibitors, the shared structural features include multiple aromatic moieties, heterocycles, secondary amines, and halogens (Figure 8). These high-affinity CYP1A2 inhibitors are small, moderately lipophilic, aromatic structures with 5.6 rotatable bonds and 4.5 hydrogen bond acceptors. Molecular docking studies indicate a diverse range of binding interactions. Key interactions include pi–pi stacking with F125, F226, and F260, the coordination of secondary amines and aromatic moieties with heme-900 (HEM900), and halogen interactions with T124 and T321. Amino acids such as T124, T223, and T321 can function as hydrogen donors to inhibitors. Lastly, I386 and L497 form Van der Waals interactions with hydrophobic moieties.

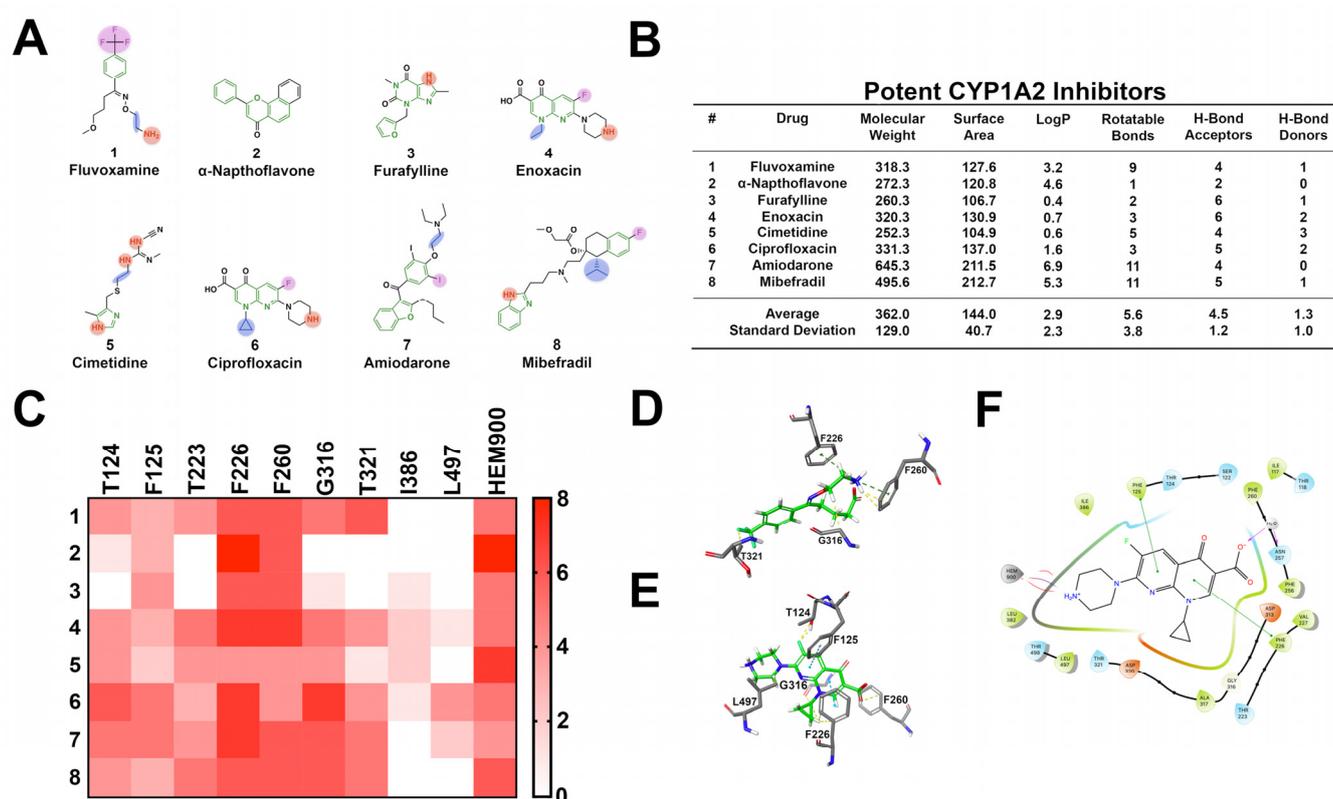


Figure 8. Structural investigation of the 8 most well-characterized high-affinity CYP1A2 inhibitors. (A) Common structural features seen in CYP1A2 inhibitors include aromatic rings, heterocycles, primary amines, and halogens. Magenta moieties interact with T124 or T321; green interact with F125, F226, F260, and/or heme; blue interact with G316; red interact with F226, F260 and/or heme. (B) CYP1A2 inhibitors are small, planar structures and are moderately flexible with numerous H-bond acceptors. (C) Heat map representing the frequency of ligand-to-amino acid binding interactions in the top ten most energetically favorable binding poses in docking against CYP1A2 (PDB ID: 2HI4). Ligand interactions are highly diverse, with key interactions including T124, T223, F226, F260, G316, T321, and heme-900. (D) Three-dimensional docking image depicting Fluvoxamine interacting with F226, F260, G316, and T321. (E) Three-dimensional docking image portraying ciprofloxacin interacting with T124, F125, F226, F260, G316, and L497. (F) Two-dimensional ligand interaction map depicting ciprofloxacin interacting with F125 and F226, while coordinating with heme-900.

3.6. Summary

Understanding the ADMET properties of compounds early in pre-clinical drug design is paramount to lead selection and clinical success. Although key physicochemical and structural descriptors are shared among CYP inhibitors (Table 1), these features on their own are insufficient for predicting CYP inhibition and binding affinity. The most effective means of predicting CYP inhibition is through the use of computational approaches that incorporate physicochemical and structural descriptors from large data sets, while using

advanced algorithms to take into consideration molecule flexibility (number of rotatable bonds), the number of hydrogen bond donors and acceptors, pharmacophores, molecule charge, prospective metabolites, and other key descriptors. Machine learning-based in silico approaches will be reviewed in Section 4.

Table 1. Key characteristics of major cytochrome p450 enzymes and their inhibitors.

CYP	Enzyme		Inhibitors	
	Cavity Size (Å ³)	Description	Description	Common Descriptors
3A4	1173–2862	Large, open, flexible, aromatic, diverse range of substrates, oxidation site is commonly a nitrogen or allylic position	Large, structurally diverse, lipophilic, aromatic, highly flexible, high hydrogen bond accepting capacity	One of more aromatic moieties, furan rings, tertiary amines, acetylene groups
2D6	510	Flat, restricted volume, acidic, aromatic, site of oxidation proximal to a primary or secondary amine	Flat, planar, aromatic structures capable of procuring a positive charge, with 2–3 hydrogen bond acceptors	One of more aromatic moieties, heterocycles, primary or secondary amines capable of carrying a positive charge
2C19	Not reported	Aromatic, moderately flexible, similar to CYP2C9	Medium sized molecules, variable lipophilicity, aromatic, oxidation site close to two hydrogen bond acceptors	Several aromatic moieties, heterocycles, carbonyl groups, and aromatic nitrogen atoms
2C9	978–1271	Larger cavity volume, moderately flexible	Aromatic, lipophilic, moderately flexible, several hydrogen bond acceptors	Aromatic, heterocycles, aromatic nitrogens, primary amines, and halogens
1A2	375–390	Small cavity volume, planar, rigid	Small, planar, aromatic, lipophilic, slightly acidic	Several aromatic moieties, heterocycles, secondary amines, and halogens

4. Machine Learning-Based Methods

Current in vitro and in vivo experimental approaches to investigate drug ADMET properties are highly effective in assessing pharmacokinetics; however, such approaches are extraordinarily expensive and are commonly performed late in drug development during IND-enabling studies [13]. In silico investigation of CYP interactions early in drug development allows investigators to select for compounds that are less likely to fail in late-stage preclinical or early clinical trials due to undesirable pharmacokinetics (Figure 9). Such tools exist and are effective at predicting CYP inhibitors, common sites of metabolism (SoMs), and prospective drug metabolites. In this section, we highlight computational approaches available for ADMET predictions, focusing primarily on CYP inhibition and metabolism. The aim of this discussion will be to highlight emerging platforms that utilize machine learning-based methods, with a special emphasis placed on applications that are free to the user and available through a public webserver. A complete overview of platforms available through a fee-based licensing model, as well as older free prediction methods and methods for analyzing CYP molecular dynamics, are reviewed elsewhere [13–20].

Various platforms capable of making enzyme-inhibitor predictions exist. These techniques use an array of methods, ranging from molecular docking with flexible multi-dimensional QSAR to machine learning based approaches. It should be noted that flexible docking-based methods, such as VirtualToxLab, are computationally expensive, take longer to receive results, and commonly require a software license [13,72]. Machine learning-based tools are capable of producing accurate results in a timely fashion and can be hosted via a publicly available web server. These platforms tend to be more user-friendly, and are commonly able to screen a large number of structures. These platforms each use unique descriptors and utilize their own methods of data generation. As such, the accuracy of

said platforms is subject to variability, contingent on the descriptors and molecular fingerprints used, the number of input structures used to build and train the model, and how the data are interpreted. In this section, we review the best publicly available machine learning-based platforms free to the scientific community.

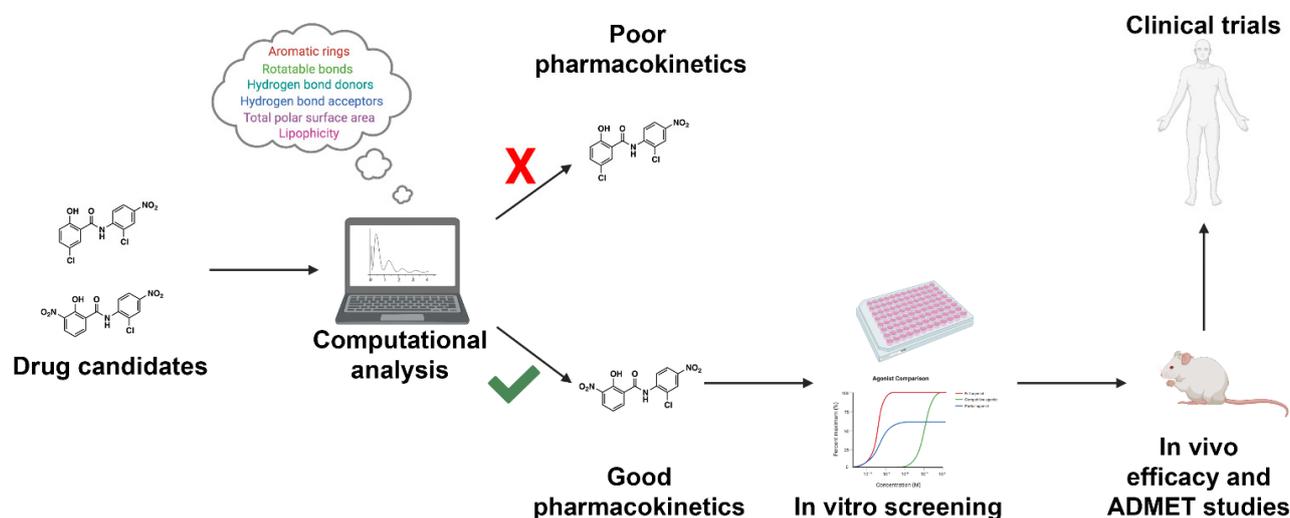


Figure 9. Workflow for incorporating ADMET predictions into early-stage drug design and development. Drug candidates can be readily screened using high-throughput computational approaches. Drugs with acceptable pharmacokinetic parameters as defined by the investigator will proceed to in vitro screening. Drugs with unacceptable pharmacokinetic properties will be triaged early in drug development. The following approach will result in a reduction in ADMET-related drug failures, reducing the cost of drug discovery efforts.

4.1. pkCSM

pkCSM is a free, publicly available web service (<http://structure.bioc.cam.ac.uk/pkcsml>; accessed on 8 February 2021) that offers a wide range of ADMET predictions [50,73]. pkCSM retains no chemical information submitted to the server and is capable of producing results in five minutes or less. Structural information is input to the server as a SMILES code and submitted for analysis. The platform is able to predict the results of key ADMET experiments commonly performed in preclinical drug discovery. Absorption predictions include caco-2 permeability, water solubility, human intestinal absorption, p-glycoprotein substrate or inhibitor, and skin permeability. Distribution predictions include the volume of distribution, the fraction unbound, and blood–brain barrier permeability. Metabolic predictions include cytochrome P450 inhibitors (CYPs 3A4, 2D6, 2C19, 2C9, and 1A2), as well as CYP2D6/3A4 substrates. The excretion predictions made include renal OCT2 substrate and total clearance. Lastly, toxicity predictions include rat LD50, AMES toxicity, T. Pyriformis toxicity, minnow toxicity, maximum tolerated dose, oral rat chronic toxicity, hepatotoxicity, skin sensitization, and hERG I and II inhibitors. pkCSM predictions are made using graph-based signatures to train, test, and predict central ADMET properties for drug development. Exact details on the number of structures used to train each individual parameter are available online (http://biosig.unimelb.edu.au/pkcsml/static/help/pkcsml_theory.pdf; accessed on 8 February 2021). Input molecules are evaluated by the platform based on key descriptors: (1) toxicophore fingerprint, (2) atomic pharmacophore frequency count (hydrophobic, aromatic, hydrogen acceptor, hydrogen donor, positive ionizable, and negative ionizable), and (3) physicochemical and structural properties, including lipophilicity (logP), molecular weight, surface area, number of rotatable bonds, and many more. pkCSM is highly accurate, with Q/AUC values meeting or exceeding those of industry standard ADMET prediction methods (CYP3A4: 0.780/0.847; CYP2D6: 0.853/0.843; CYP2C19: 0.808/0.879; CYP2C9: 0.807/0.868; CYP1A2: 0.802/0.876). Although platforms with comparable or slightly superior CYP inhibitor predictive values exist, pkCSM is capable of making

highly accurate predictions on a wide range of relevant ADMET parameters, designating it one of the most well-rounded applications available.

4.2. DeepCyp

DeepCYP is a free, publicly available web service (<http://repharma.pku.edu.cn/deepcyp/home.php>; accessed on 8 February 2021) that utilizes a multitask model for concurrent inhibition prediction of the five most clinically relevant CYPs: 3A4, 2D6, 2C19, 2C9, and 1A2 [74]. Test item information is submitted as a mol.2 or comparable file type with results achieved in minutes. Results appear as inhibition probabilities for each individual cytochrome. DeepCYP was built and trained using a multitask auto-encoder deep neural network (DNN) with chemical input from the PubChem and BioAssay databases (>13,000 compounds). Descriptors and chemical fingerprints of inhibitors and non-inhibitors were calculated using PaDEL-Descriptor software [74,75]. Totals of 1253 and 688 descriptors for PaDEL-1D&2D and PubChem datasets were used in model training, respectively. Examples of dataset descriptors and fingerprints include atom type electrotopological state (Estate) descriptors, ClogP, molecular linear free energy, element counts, ring counts, simple atom pairs, atom nearest neighbors, and SMARTS patterns. Exact details on model development and training are available online [74]. The multi-task DNN DeepCYP platform outperformed single-task models and traditional machine learning methods, with average prediction accuracies of 0.864 for 10-fold cross-validation and 0.887 for the external test datasets. Additionally, the method utilizes linear regression models to determine how each task contributes to predictions, allowing the investigators to identify conditions under which DeepCYP outperforms single task methods. These data may be used to further optimize the model with time. The predictive power of DeepCYP makes it one of the most accurate freely accessible methods available for concurrent CYP inhibition predictions.

4.3. SuperCYPsPred

SuperCYPsPred is a free, user-friendly, publicly available web service (<http://insilico-cyp.charite.de/SuperCYPsPred/>; accessed on 8 February 2021) that uses well-established machine learning methods to predict inhibitor specificity for CYPs 3A4, 2D6, 2C19, 2C9, and 1A2 [7]. The webserver allows for the input of molecules in three different ways: (1) input of a SMILES code, (2) PubChem structure search, or (3) drawing of custom structures in ChemDoodle. The model was developed based on traditional machine learning random forests methods, a commonly employed ensemble learning algorithm that is less prone to class bias and overfitting [7,76]. Two chemical-based fingerprints were used in model development: MACCS molecular fingerprints and Morgan circular fingerprints. The server allows the user to select the preferred fingerprint, or both fingerprints, to be used in prediction generation. Results are shown as active or inactive for each individual CYP with assigned probabilities for the prediction. SuperCYPsPred is remarkably accurate; cross-validation accuracy averaged 0.930 (CYP3A4: 0.92; CYP2D6: 0.84; CYP2C19: 0.97; CYP2C9: 0.97; CYP1A2: 0.95) and external validation averaged 0.882 (CYP3A4: 0.86; CYP2D6: 0.80; CYP2C19: 0.95; CYP2C9: 0.90; CYP1A2: 0.90). Similar to DeepCYP, SuperCYPsPred is one of the most accurate free resources for CYP specificity, making it an excellent screening tool for pre-clinical drug design and development.

4.4. vNN-ADMET

vNN-ADMET is a free, publicly available web service (<https://vnnadmet.bhsai.org/>; accessed on 8 February 2021) that offers a wide range of ADMET predictions, including drug cytotoxicity, mutagenicity, cardiotoxicity, microsomal stability, hepatotoxicity, and cytochrome specificity [77]. Additionally, the webserver allows for public use of the versatile variable nearest neighbor (vNN) method for the custom design of novel models. The vNN method is similar to the k-nearest neighbor (kNN) method: a non-parametric method that uses pattern recognition of chemical features for classification and regression. One caveat

of the kNN method is that it always gives a prediction for a test compound regardless of the degree of structural disagreement from the dataset. The vNN method is built using a predetermined similarity criterion, which only considers the nearest neighbors that share a defined degree of structural similarity (based on Tanimoto distance) to the input structure for prediction generation. A prediction is only made when valid nearest neighbors are identified. The Accelrys extended-connectivity fingerprints with a diameter of four chemical bonds (ECFP4) are used to identify structurally related compounds. Similar to the aforementioned machine learning models, a 10-fold cross-validation procedure was used to validate the model. The CYP inhibitor specificity results are portrayed as a categorical response (yes or no). vNN-ADMET is capable of making reliable CYP specificity predictions, with cross-validation accuracy averaging 0.890 (CYP3A4: 0.88; CYP2D6: 0.89; CYP2C19: 0.87; CYP2C9: 0.91; CYP1A2: 0.90), making the webserver an excellent resource for medicinal chemists.

4.5. AdmetSAR 2.0

AdmetSAR is a free resource useful in the prediction of ADMET properties of novel chemical entities (<http://lmm.d.ecust.edu.cn/admetSar2/>; accessed on 8 February 2021) [78,79]. AdmetSAR 2.0 is an optimized version of AdmetSAR (1.0), now with 47 prediction models (formerly contained 27 models). Training data were procured from DrugBank, ChEMBL, CPDB, CYP450, and Tox21. Three chemical-based fingerprints were used in model development: MACCS molecular fingerprints, Morgan circular fingerprints, and AtomParis; all of which were implemented with RDKit. The scikit-learn package with python scripts was used to implement various machine learning algorithms, including random forests, support vector machines, and kNN. Supplementary details regarding method design and development are available online [78]. Six descriptors were used to define the applicability domain, including molecular weight, AlogP, number of atoms, number of rings, hydrogen bond acceptors, and hydrogen bond donors. The performance of binary models in CYP specificity prediction via AdmetSAR averaged 0.784 (CYP3A4: 0.645; CYP2D6: 0.855; CYP2C19: 0.806; CYP2C9: 0.802; CYP1A2: 0.815). Notably, CYP3A4 accuracy is lower than that of the previously mentioned platforms (0.645). One exciting novel feature of AdmetSAR 2.0 is ADMETopt—a module that allows users to perform molecule ADMET optimization by scaffold hopping. Scaffold hopping is a procedure that replaces the query molecule with a structurally similar scaffold that possess improved ADMET properties. AdmetSAR, and more specifically ADMETopt, is an excellent free resource that can be used in lead optimization.

4.6. SwissADME

SwissADME is a free, publicly available tool (<http://www.swissadme.ch>; accessed on 8 February 2021) capable of predicting physicochemical properties, pharmacokinetics, drug-likeness and medicinal chemistry friendliness [80]. Additionally, SwissADME provides graphical representations of biological barrier permeation (BOILED-Egg) and a Bioavailability Radar. Structures may be drawn directly into the webserver or submitted as a SMILES code, with categorical cytochrome inhibition predictions made on CYPs 3A4, 2D6, 2C19, 2C9, and 1A2. SwissADME was built using the PubChem Bioassay dataset, which contains 16,561 compounds with 50 chemical descriptors. SwissADME makes fast and reliable predictions, with an average 10-fold cross-validation accuracy of 0.794 (CYP3A4: 0.77; CYP2D6: 0.79; CYP2C19: 0.80; CYP2C9: 0.78; CYP1A2: 0.72); a 10-fold cross-validation AUROC (area under the receiver operating characteristic curve) average of 0.862 (CYP3A4: 0.85; CYP2D6: 0.85; CYP2C19: 0.86; CYP2C9: 0.85; CYP1A2: 0.90); an external validation accuracy average of 0.788 (CYP3A4: 0.78; CYP2D6: 0.81; CYP2C19: 0.80; CYP2C9: 0.71; CYP1A2: 0.84), and an external AUROC average of 0.864 (CYP3A4: 0.86; CYP2D6: 0.87; CYP2C19: 0.87; CYP2C9: 0.81; CYP1A2: 0.91). SwissADME affords unique advantages, allowing for the input of multiple molecules and for the user to export data in various file formats. SwissADME analyzes small molecule composition to determine whether it

violates or is in compliance with Lipinski's Rule of Five and PAINS, two drug triage tools used in industry to predict drug likeness.

4.7. CypRules

CypRules is a freely available webserver (<http://cyprules.cmdm.tw/>; accessed on 8 February 2021) that predicts inhibitors and non-inhibitors of the five major CYPs 3A4, 2D6, 2C19, 2C9, and 1A2 [81]. CypRules utilizes the same dataset as SwissADME for predictions; however, CypRules utilizes decision trees in combination with the theory of information entropy [13,81]. CypRules uses three descriptor tools: PubChem 2D Fingerprint, PaDEL-Descriptor, and Mold2 [75,82]. Classification models utilize a C5.0 algorithm, which chooses an attribute from the data that most effectively splits the initial dataset into enriched subsets [81]. CypRules makes accurate predictions, with an average accuracy of 0.812 (CYP3A4: 0.73; CYP2D6: 0.90; CYP2C19: 0.86; CYP2C9: 0.77; CYP1A2: 0.80). An added benefit of CypRules is the ability to perform virtual high-throughput screening of a large set of testing compounds.

4.8. CypReact

CypReact is a freely available tool (https://bitbucket.org/Leon_Ti/cypreact; accessed on 8 February 2021) capable of predicting reactants of the nine most important CYP enzymes: CYPs 3A4, 2E1, 2D6, 2C19, 2C9, 2C8, 2B6, 2A6, and 1A2 [83]. In addition to predicting whether a test item will serve as a reactant of a given cytochrome, CypReact is also capable of predicting sites of metabolism (SoMs), as well as the end result of the interaction ("endpoint prediction"). This allows for the user to generate predictions on the query molecule, as well as prospective metabolites. CypReact generates binary data on all parameters (1 = true; 0 = false). The CypReact dataset was built using 1632 compounds extracted from Zaretski et al. [84]. An additional 1053 non-reactant compounds were gathered in order to improve the quality and predictive capacity of the dataset. Three chemical-based fingerprints were used in model development: MACCS molecular fingerprints, PubChem fingerprints, and ClassyFire27 fingerprint. CypReact employs an LBM learning algorithm that considers a set of 2279 features of dataset compounds, including 36 physiochemical properties. Additional information on model training and dataset selection is available online [83]. CypReact makes reliable predictions, with AUROC values ranging from 0.83 to 0.92. Furthermore, CypReact outperforms other reactant prediction tools such as ADMET Predictor (AUROC = 0.75) and SMARTCyp (AUROC = 0.53). ADMET Predictor and SMARTCyp are reviewed elsewhere [13,85,86]. The resulting metabolites may be screened for CYP inhibition specificity and toxicity risk using the previously described platforms. One criticism of CypReact is the user interface. The platform utilizes bitbucket, requiring the use of rudimentary coding skills, making the model less user-friendly than the previously described methods. However, the inclusion of SoM prediction(s) makes CypReact a gold standard prediction model.

5. Conclusions

More than half of drug failures are a result of poor ADMET. More specifically, drug–drug interactions as a result of major cytochrome inhibition are commonplace and can lead to life-threatening adverse drug reactions. A majority of pre-clinical-stage drug discovery programs focus solely on binding affinity and selectivity [50]. Although binding affinity is of utmost importance, ADMET properties ensure that a drug is properly absorbed, distributed, and cleared; features that equally contribute to the success of a drug in the clinic. As such, *in silico* approaches have been developed to be introduced during early drug screening, with the goal of reducing the rate of drug attrition (Table 2).

Table 2. Summary of publicly available machine learning-based methods for cytochrome inhibition and metabolism.

Name	CYPs	Prediction Type	ML Method	No. Structures	Avg. Accuracy ¹	Additional Features
pkCSM	3A4, 2D6, 2C19, 2C9, 1A2	Inhibition	Graph-based signatures	18,000	0.810 (0.780–0.853)	Comprehensive ADMET predictions (23 total)
DeepCYP	3A4, 2D6, 2C19, 2C9, 1A2	Inhibition	Multitask autoencoder deep neural network	13,000	0.864 (0.809–0.968)	Assigns probabilities for CYP inhibition
SuperCYPs-Pred	3A4, 2D6, 2C19, 2C9, 1A2	Inhibition	Random forests	41,963 ²	0.930 (0.840–0.970)	Assigns probabilities for CYP inhibition
vNN-ADMET	3A4, 2D6, 2C19, 2C9, 1A2	Inhibition	Variable nearest neighbors	6261	0.890 (0.870–0.910)	
AdmetSAR 2.0	3A4, 2D6, 2C19, 2C9, 1A2	Inhibition	Random forests, support vector machines, k-nearest neighbors	96,000 ³	0.784 (0.645–0.855)	Comprehensive ADMET predictions (47 total); ADMETopt for lead optimization
SwissADME	3A4, 2D6, 2C19, 2C9, 1A2	Inhibition	Support vector machines	16,561 ⁴	0.794 (0.720–0.800)	Predictions of physicochemical properties, pharmacokinetics, and drug likeness; high throughput
CypRules	3A4, 2D6, 2C19, 2C9, 1A2	Inhibition	Decision trees	16,561	0.812 (0.730–0.900)	High throughput
CypReact	3A4, 2E1, 2D6, 2C19, 2C9, 2C8, 2B6, 2A6, 1A2	Sites of metabolism	LBM learning algorithm	2685	Unavailable	Metabolite predictions; Additional CYPs

¹ Cross-validation accuracy. ² Sum of the number of compounds used for all five major cytochromes. The dataset includes inhibitors and non-inhibitors and may include duplicates. ³ Includes compounds used to build ADMET models in addition to CYP inhibition models. ⁴ Includes compounds used to establish physicochemical properties, pharmacokinetics, drug likeness and medicinal chemistry friendliness models.

In silico approaches utilize a wide range of molecular descriptors to build and train the models. Physicochemical descriptors are considered to be highly predictive of CYP specificity and binding affinity. Thus, we explored the ChEMBL database, which contains over 10,000 CYP inhibitors with individual items of data on molecular weight, lipophilicity (ALogP), and half-maximal inhibitory concentration (IC₅₀). Although CYP inhibitors share common physicochemical features, molecular weight and lipophilicity alone are unable to predict CYP half-maximal inhibitory concentration. CYP3A4 and CYP2C9 inhibitors demonstrated a weak statistically significant negative relationship between molecule size, lipophilicity, and IC₅₀ (i.e., larger, more lipophilic molecules tend to have lower IC₅₀ values); however, the exact opposite trend was observed for CYP1A2. No relationship was identified for CYPs 2D6 and 2C19. We then explored the common structural features of major cytochrome inhibitors, which are summarized above (Table 1). These features are critically important in making predictions; however, these data alone have limited predictive value without the help of machine learning-based methods. In this review, we highlight eight machine learning-based approaches for the prediction of cytochrome inhibition. These eight platforms are publicly available and free to the user. Based on the average accuracies provided in the literature, SuperCYPsPred (0.930) appears to be the most accurate tool for predicting CYP inhibition, followed by vNN-ADMET (0.890) and DeepCYP (0.864). It should be noted, however, that each of the mentioned platforms are capable of making reliable predictions on CYP inhibitor specificity and possess their own unique features. pkCSM and AdmetSAR 2.0 are capable of predicting a suite of ADMET properties. More specifically, pkCSM is capable of predicting the outcomes of

commonly used in vitro ADMET assays, as well as human gastrointestinal absorption and toxicity. Additionally, pkCSM provides the data on human volume of distribution and clearance, which can be used to estimate drug half-life. Although pkCSM may not be the most accurate machine learning-based platform discussed in this review, it is capable of making rapid predictions on a broad range of parameters and is user-friendly. In addition to predicting ADMET parameters, AdmetSar 2.0 hosts ADMETopt, a platform that assists the user with lead optimization. All of the platforms described assign binary “yes” or “no” predictions for CYP inhibition, with the exception of DeepCyp and SuperCypsPred; these tools assign probabilities to their predictions. Probability assignment is useful, as it allows the user to gauge the relative accuracy and utility of the prediction. SwissADME makes predictions on physicochemical properties, pharmacokinetics, and drug likeness. Physicochemical properties such as solubility are important for drug formulation and in vivo exposure. Additionally, SwissADME assesses whether a test molecule violates industry standard rules such as Lipinski’s Rule of Five. These predictions are important for predicting the future success of drug candidates and should be considered during the early stages of drug development. SwissADME and CypRules are capable of performing high-throughput screening of large drug libraries. Lastly, CypReact features site of metabolism and metabolite predictions. These predictions can be used to assess CYP reactivity and prospective metabolite-driven toxicities. The goal of this article is to promote the use of ADMET prediction software early in the drug design process in order to improve drug discovery outcomes.

Supplementary Materials: The following are available online at <https://www.mdpi.com/article/10.3390/ph14050472/s1>, Table S1: CYP3A4 Multiple Linear Regression; Table S2: CYP2D6 Multiple Linear Regression; Table S3: CYP2C19 Multiple Linear Regression; Table S4: CYP2C9 Multiple Linear Regression; Table S5: CYP3A4 Multiple Linear Regression.

Author Contributions: Conceptualization, T.C.B. and R.A.N.; methodology, T.C.B. and R.A.N.; software, T.C.B.; validation, T.C.B., K.R.B., M.M.B., J.M. and R.A.N.; formal analysis, T.C.B.; data curation, T.C.B.; writing—original draft preparation, T.C.B.; writing—review and editing, T.C.B., K.R.B., M.M.B., J.M. and R.A.N.; visualization, T.C.B.; supervision, R.A.N.; funding acquisition, T.C.B. and R.A.N. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Heart, Lung, and Blood Institute, HL007260. The APC was funded by National Heart, Lung, and Blood Institute, HL123974.

Data Availability Statement: Individual data are available via FigShare [87].

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

References

1. Therapeutic Drug Use. Available online: <https://www.cdc.gov/nchs/fastats/drug-use-therapeutic.htm> (accessed on 28 February 2021).
2. Mikulic, M. Prescription Drug Expenditure in the United States 1960–2020. Available online: <https://www.statista.com/statistics/184914/prescription-drug-expenditures-in-the-us-since-1960/> (accessed on 28 January 2021).
3. Donaldson, M.S.; Corrigan, J.M.; Kohn, L.T. (Eds.) *In to Err Is Human: Building a Safer Health System*; National Academies Press: Washington, DC, USA, 2000.
4. Lazarou, J.; Pomeranz, B.H.; Corey, P.N. Incidence of Adverse Drug Reactions in Hospitalized Patients. *JAMA* **1998**, *279*, 1200–1205. [[CrossRef](#)] [[PubMed](#)]
5. Preventable Adverse Drug Reactions: A Focus on Drug Interactions. Available online: <https://www.fda.gov/drugs/drug-interactions-labeling/preventable-adverse-drug-reactions-focus-drug-interactions> (accessed on 2 February 2021).
6. Gurwitz, J.H.; Field, T.S.; Avorn, J.; McCormick, D.; Jain, S.; Eckler, M.; Benser, M.; Edmondson, A.C.; Bates, D.W. Incidence and preventability of adverse drug events in nursing homes. *Am. J. Med.* **2000**, *109*, 87–94. [[CrossRef](#)]
7. Banerjee, P.; Dunkel, M.; Kemmler, E.; Preissner, R. SuperCYPsPred—A web server for the prediction of cytochrome activity. *Nucleic Acids Res.* **2020**, *48*, W580–W585. [[CrossRef](#)] [[PubMed](#)]

8. Redlich, G.; Zanger, U.M.; Riedmaier, S.; Bache, N.; Giessing, A.B.M.; Eisenacher, M.; Stephan, C.; Meyer, H.E.; Jensen, O.N.; Marcus, K. Distinction between Human Cytochrome P450 (CYP) Isoforms and Identification of New Phosphorylation Sites by Mass Spectrometry. *J. Proteome Res.* **2008**, *7*, 4678–4688. [[CrossRef](#)]
9. Ogu, C.C.; Maxa, J.L. Drug Interactions Due to Cytochrome P450. In *Baylor University Medical Center Proceedings*; Informa UK Limited: London, UK, 2000; Volume 13, pp. 421–423.
10. Vrbanac, J.; Slauter, R. ADME in Drug Discovery. In *A Comprehensive Guide to Toxicology in Preclinical Drug Development*; Faqi, A.S., Ed.; Academic Press: Cambridge, MA, USA, 2013.
11. McDonnell, A.M.; Dang, C.H. Basic Review of the Cytochrome P450 System. *J. Adv. Pract. Oncol.* **2013**, *4*, 263–268. [[CrossRef](#)]
12. Raunio, H.; Kuusisto, M.; Juvonen, R.O.; Pentikainen, O.T. Modeling of interactions between xenobiotics and cytochrome P450 (CYP) enzymes. *Front. Pharmacol.* **2015**, *6*, 123. [[CrossRef](#)]
13. Tyzack, J.D.; Kirchmair, J. Computational methods and tools to predict cytochrome P450 metabolism for drug discovery. *Chem. Biol. Drug Des.* **2019**, *93*, 377–386. [[CrossRef](#)]
14. Martiny, V.Y.; Carbonell, P.; Chevillard, F.; Moroy, G.; Nicot, A.B.; Vayer, P.; Villoutreix, B.O.; Miteva, M.A. Integrated structure- and ligand-based in silico approach to predict inhibition of cytochrome P450 2D6. *Bioinformatics* **2015**, *31*, 3930–3937. [[CrossRef](#)]
15. Refsgaard, H.H.F.; Jensen, B.F.; Christensen, I.T.; Hagen, N.; Brockhoff, P.B. In silico prediction of cytochrome P450 inhibitors. *Drug Dev. Res.* **2006**, *67*, 417–429. [[CrossRef](#)]
16. Rudik, A.; Dmitriev, A.; Lagunin, A.; Filimonov, D.; Poroikov, V. MetaPASS: A Web Application for Analyzing the Biological Activity Spectrum of Organic Compounds Taking into Account their Biotransformation. *Mol. Inform.* **2021**, *40*, 2000231. [[CrossRef](#)]
17. Kar, S. Recent Advances of Computational Modeling for Predicting Drug Metabolism: A Perspective. *Curr. Drug Metab.* **2018**, *18*, 1106–1122. [[CrossRef](#)]
18. Kirchmair, J.; Göller, A.H.; Lang, D.; Kunze, J.; Testa, B.; Wilson, I.D.; Glen, R.C.; Schneider, G. Predicting drug metabolism: Experiment and/or computation? *Nat. Rev. Drug Discov.* **2015**, *14*, 387–404. [[CrossRef](#)] [[PubMed](#)]
19. Shaik, S.; Chen, H.; Usharani, D.; Thiel, W. QM/MM Studies of Structure and Reactivity of Cytochrome P450 Enzymes: Methodology and Selected Applications. *Early Drug Dev.* **2014**, *311*, 133–178. [[CrossRef](#)]
20. Manikandan, P.; Nagini, S. Cytochrome P450 Structure, Function and Clinical Significance: A Review. *Curr. Drug Targets* **2018**, *19*, 38–54. [[CrossRef](#)]
21. Kiani, Y.S.; Jabeen, I. Exploring the Chemical Space of Cytochrome P450 Inhibitors Using Integrated Physicochemical Parameters, Drug Efficiency Metrics and Decision Tree Models. *Computation* **2019**, *7*, 26. [[CrossRef](#)]
22. Nassar, A.-E.F.; Kamel, A.M.; Clarimont, C. Improving the decision-making process in structural modification of drug candidates: Reducing toxicity. *Drug Discov. Today* **2004**, *9*, 1055–1064. [[CrossRef](#)]
23. Lewis, D.F.V.; Dickins, M. Substrate SARs in human P450s. *Drug Discov. Today* **2002**, *7*, 918–925. [[CrossRef](#)]
24. Lewis, D.F.V.; Dickins, M. Baseline Lipophilicity Relationships in Human Cytochromes P450 Associated with Drug Metabolism. *Drug Metab. Rev.* **2003**, *35*, 1–18. [[CrossRef](#)]
25. Gleeson, M.P. Generation of a Set of Simple, Interpretable ADMET Rules of Thumb. *J. Med. Chem.* **2008**, *51*, 817–834. [[CrossRef](#)]
26. Arnott, J.; Planey, S.L. The influence of lipophilicity in drug discovery and design. *Expert Opin. Drug Discov.* **2012**, *7*, 863–875. [[CrossRef](#)]
27. Lewis, D.F.; Lake, B.G.; Ito, Y.; Dickins, M. Lipophilicity Relationships in Inhibitors of CYP2C9 and CYP2C19 Enzymes. *J. Enzym. Inhib. Med. Chem.* **2006**, *21*, 385–389. [[CrossRef](#)] [[PubMed](#)]
28. Lewis, D.F.V.; Lake, B.G.; Dickins, M. Quantitative structure-activity relationships (QSARs) in inhibitors of various cytochromes P450: The importance of compound lipophilicity. *J. Enzym. Inhib. Med. Chem.* **2007**, *22*, 1–6. [[CrossRef](#)]
29. Gaulton, A.; Bellis, L.J.; Bento, A.P.; Chambers, J.; Davies, M.; Hersey, A.; Light, Y.; McGlinchey, S.; Michalovich, D.; Al-Lazikani, B.; et al. ChEMBL: A large-scale bioactivity database for drug discovery. *Nucleic Acids Res.* **2011**, *40*, D1100–D1107. [[CrossRef](#)] [[PubMed](#)]
30. Kaur, P.; Chamberlin, A.R.; Poulos, T.L.; Sevrioukova, I.F. Structure-Based Inhibitor Design for Evaluation of a CYP3A4 Pharmacophore Model. *J. Med. Chem.* **2016**, *59*, 4210–4220. [[CrossRef](#)]
31. Eagling, V.A.; Back, D.J.; Barry, M.G. Differential inhibition of cytochrome P450 isoforms by the protease inhibitors, ritonavir, saquinavir and indinavir. *Br. J. Clin. Pharmacol.* **1997**, *44*, 190–194. [[CrossRef](#)]
32. De Groot, M.J.; Vermeulen, N.P.E.; Kramer, J.D.; Van Acker, F.A.A.; Kelder, G.M.D.-O.D. A Three-Dimensional Protein Model for Human Cytochrome P450 2D6 Based on the Crystal Structures of P450 101, P450 102, and P450 108. *Chem. Res. Toxicol.* **1996**, *9*, 1079–1091. [[CrossRef](#)] [[PubMed](#)]
33. Kemp, C.A.; Flanagan, J.U.; Van Eldik, A.J.; Maréchal, J.-D.; Wolf, C.R.; Roberts, G.C.K.; Paine, M.J.I.; Sutcliffe, M.J. Validation of Model of Cytochrome P450 2D6: An in Silico Tool for Predicting Metabolism and Inhibition. *J. Med. Chem.* **2004**, *47*, 5340–5346. [[CrossRef](#)]
34. Egnell, A.-C.; Eriksson, C.; Albertson, N.; Houston, B.; Boyer, S. Generation and Evaluation of a CYP2C9 Heteroactivation Pharmacophore. *J. Pharmacol. Exp. Ther.* **2003**, *307*, 878–887. [[CrossRef](#)]
35. Ekins, S.; Bravi, G.; Binkley, S.; Gillespie, J.S.; Ring, B.J.; Wikel, J.H.; A Wrighton, S. Three- and four-dimensional-quantitative structure activity relationship (3D/4D-QSAR) analyses of CYP2C9 inhibitors. *Drug Metab. Dispos.* **2000**, *28*, 994–1002.
36. Foti, R.S.; Rock, D.A.; Han, X.; Flowers, R.A.; Wienkers, L.C.; Wahlstrom, J.L. Ligand-Based Design of a Potent and Selective Inhibitor of Cytochrome P450 2C19. *J. Med. Chem.* **2012**, *55*, 1205–1214. [[CrossRef](#)] [[PubMed](#)]

37. Zhang, T.; Wei, D.-Q.; Chou, K.-C. A pharmacophore model specific to active site of CYP1A2 with a novel molecular modeling explorer and CoMFA. *Med. Chem.* **2012**, *8*, 198–207. [PubMed]
38. Korhonen, L.E.; Rahnasto, M.; Mähönen, N.J.; Wittekindt, C.; Poso, A.; Juvonen, R.O.; Raunio, H. Predictive Three-Dimensional Quantitative Structure–Activity Relationship of Cytochrome P450 1A2 Inhibitors. *J. Med. Chem.* **2005**, *48*, 3808–3815. [CrossRef]
39. Gay, S.C.; Roberts, A.G.; Halpert, J.R. Structural features of cytochromes P450 and ligands that affect drug metabolism as revealed by X-ray crystallography and NMR. *Future Med. Chem.* **2010**, *2*, 1451–1468. [CrossRef] [PubMed]
40. Sevrioukova, I.F.; Poulos, T.L. Structure and mechanism of the complex between cytochrome P4503A4 and ritonavir. *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 18422–18427. [CrossRef]
41. Rock, B.M.; Hengel, S.M.; Rock, D.A.; Wienkers, L.C.; Kunze, K.L. Characterization of Ritonavir-Mediated Inactivation of Cytochrome P450 3A4. *Mol. Pharmacol.* **2014**, *86*, 665–674. [CrossRef] [PubMed]
42. He, K.; Woolf, T.F.; Hollenberg, P.F. Mechanism-based inactivation of cytochrome P-450-3A4 by mifepristone (RU486). *J. Pharmacol. Exp. Ther.* **1999**, *288*, 791–797.
43. Zhou, S.; Chan, E.; Lim, L.; Boelsterli, U.; Li, S.; Wang, J.; Zhang, Q.; Huang, M.; Xu, A. Therapeutic Drugs that Behave as Mechanism-Based Inhibitors of Cytochrome P450 3A4. *Curr. Drug Metab.* **2004**, *5*, 415–442. [CrossRef] [PubMed]
44. Friesner, R.A.; Banks, J.L.; Murphy, R.B.; Halgren, T.A.; Klicic, J.J.; Mainz, D.T.; Repasky, M.P.; Knoll, E.H.; Shelley, M.; Perry, J.K.; et al. Glide: A New Approach for Rapid, Accurate Docking and Scoring. 1. Method and Assessment of Docking Accuracy. *J. Med. Chem.* **2004**, *47*, 1739–1749. [CrossRef]
45. Friesner, R.A.; Murphy, R.B.; Repasky, M.P.; Frye, L.L.; Greenwood, J.R.; Halgren, T.A.; Sanschagrin, P.C.; Mainz, D.T. Extra Precision Glide: Docking and Scoring Incorporating a Model of Hydrophobic Enclosure for Protein–Ligand Complexes. *J. Med. Chem.* **2006**, *49*, 6177–6196. [CrossRef]
46. Halgren, T.A.; Murphy, R.B.; Friesner, R.A.; Beard, H.S.; Frye, L.L.; Pollard, W.T.; Banks, J.L. Glide: A New Approach for Rapid, Accurate Docking and Scoring. 2. Enrichment Factors in Database Screening. *J. Med. Chem.* **2004**, *47*, 1750–1759. [CrossRef]
47. Šarić Mustapić, D.; Debeljak, Ž.; Maleš, Ž.; Bojić, M. The Inhibitory Effect of Flavonoid Aglycones on the Metabolic Activity of CYP3A4 Enzyme. *Molecules* **2018**, *23*, 2553. [CrossRef]
48. Dai, D.; Tang, J.; Rose, R.; Hodgson, E.; Bienstock, R.J.; Mohrenweiser, H.W.; A Goldstein, J. Identification of variants of CYP3A4 and characterization of their abilities to metabolize testosterone and chlorpyrifos. *J. Pharmacol. Exp. Ther.* **2001**, *299*, 825–831.
49. Ekroos, M.; Sjögren, T. Structural basis for ligand promiscuity in cytochrome P450 3A4. *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 13682–13687. [CrossRef]
50. Pires, D.E.V.; Blundell, T.L.; Ascher, D.B. pkCSM: Predicting Small-Molecule Pharmacokinetic and Toxicity Properties Using Graph-Based Signatures. *J. Med. Chem.* **2015**, *58*, 4066–4072. [CrossRef] [PubMed]
51. Van Hoppe, S.; Rood, J.J.M.; Buil, L.; Wagenaar, E.; Sparidans, R.W.; Beijnen, J.H.; Schinkel, A.H. P-Glycoprotein (MDR1/ABC B1) Restricts Brain Penetration of the Bruton’s Tyrosine Kinase Inhibitor Ibrutinib, While Cytochrome P450-3A (CYP3A) Limits Its Oral Bioavailability. *Mol. Pharm.* **2018**, *15*, 5124–5134. [CrossRef]
52. Greenblatt, D.J.; Patki, K.C.; Von Moltke, L.L.; Shader, R.I. Drug Interactions with Grapefruit Juice: An Update. *J. Clin. Psychopharmacol.* **2001**, *21*, 357–359. [CrossRef]
53. Vuppalanchi, R. *Metabolism of Drugs and Xenobiotics*; Saxena, R., Ed.; Elsevier: Cambridge, MA, USA, 2011.
54. Rowland, P.; Blaney, F.E.; Smyth, M.G.; Jones, J.J.; Leydon, V.R.; Oxbrow, A.K.; Lewis, C.J.; Tennant, M.G.; Modi, S.; Eggleston, D.S.; et al. Crystal Structure of Human Cytochrome P450 2D6. *J. Biol. Chem.* **2006**, *281*, 7614–7622. [CrossRef]
55. Cytochrome 2D6 Inhibition. Available online: <https://www.cambridgemedchemconsulting.com/resources/ADME/cyp2d6inhibition.html> (accessed on 9 February 2021).
56. Roskoski, R. Properties of FDA-approved small molecule protein kinase inhibitors. *Pharmacol. Res.* **2019**, *144*, 19–50. [CrossRef] [PubMed]
57. VandenBrink, B.M.; Foti, R.S.; Rock, D.A.; Wienkers, L.C.; Wahlstrom, J.L. Prediction of CYP2D6 Drug Interactions from In Vitro Data: Evidence for Substrate-Dependent Inhibition. *Drug Metab. Dispos.* **2011**, *40*, 47–53. [CrossRef] [PubMed]
58. Demkow, U. *Next Generation Sequencing in Pharmacogenomics*; Płoski, U.D.R., Ed.; Academic Press: Cambridge, MA, USA, 2016.
59. Derayea, S.M.; Tsujino, H.; Oyama, Y.; Ishikawa, Y.; Yamashita, T.; Uno, T. Investigation on drug-binding in heme pocket of CYP2C19 with UV–visible and resonance Raman spectroscopies. *Spectrochim. Acta Part A Mol. Biomol. Spectrosc.* **2019**, *209*, 209–216. [CrossRef]
60. Cytochrome 2C19 Inhibition. Available online: <https://www.cambridgemedchemconsulting.com/resources/ADME/cyp2c19inhibition.html> (accessed on 2 February 2021).
61. Ibeanu, G.C.; Goldstein, J.; Meyer, U.; Benhamou, S.; Bouchardy, C.; Dayer, P.; Ghanayem, B.; Blaisdell, J. Identification of new human CYP2C19 alleles (CYP2C19*6 and CYP2C19*2B) in a Caucasian poor metabolizer of mephenytoin. *J. Pharmacol. Exp. Ther.* **1998**, *286*, 1490–1495. [PubMed]
62. Fukushima-Uesaka, H.; Saito, Y.; Maekawa, K.; Ozawa, S.; Hasegawa, R.; Kajio, H.; Kuzuya, N.; Yasuda, K.; Kawamoto, M.; Kamatani, N.; et al. Genetic Variations and Haplotypes of CYP2C19 in a Japanese Population. *Drug Metab. Pharmacokinet.* **2005**, *20*, 300–307. [CrossRef] [PubMed]
63. Sim, S.C.; Risinger, C.; Dahl, M.-L.; Aklillu, E.; Christensen, M.; Bertilsson, L.; Ingelman-Sundberg, M. A common novel CYP2C19 gene variant causes ultrarapid drug metabolism relevant for the drug response to proton pump inhibitors and antidepressants. *Clin. Pharmacol. Ther.* **2006**, *79*, 103–113. [CrossRef] [PubMed]

64. Cytochrome 2C9 Inhibition. Available online: https://www.cambridgemedchemconsulting.com/resources/ADME/cyp2c9_inhibition.html (accessed on 4 February 2021).
65. Williams, P.A.; Cosme, J.; Ward, A.; Angove, H.C.; Vinković, D.M.; Jhoti, H. Crystal structure of human cytochrome P450 2C9 with bound warfarin. *Nature* **2003**, *424*, 464–468. [[CrossRef](#)]
66. Wester, M.R.; Yano, J.K.; Schoch, G.A.; Yang, C.; Griffin, K.J.; Stout, C.D.; Johnson, E.F. The Structure of Human Cytochrome P450 2C9 Complexed with Flurbiprofen at 2.0-Å Resolution. *J. Biol. Chem.* **2004**, *279*, 35630–35637. [[CrossRef](#)] [[PubMed](#)]
67. Pratt, V.M.; Cavallari, L.H.; Del Tredici, A.L.; Hachad, H.; Ji, Y.; Moyer, A.M.; Scott, S.A.; Whirl-Carrillo, M.; Weck, K.E. Recommendations for Clinical CYP2C9 Genotyping Allele Selection: A Joint Recommendation of the Association for Molecular Pathology and College of American Pathologists. *J. Mol. Diagn.* **2019**, *21*, 746–755. [[CrossRef](#)] [[PubMed](#)]
68. García-Martín, E.; Martínez, C.; Ladero, J.M.; Agúndez, J.A.G. Interethnic and Intraethnic Variability of CYP2C8 and CYP2C9 Polymorphisms in Healthy Individuals. *Mol. Diagn. Ther.* **2006**, *10*, 29–40. [[CrossRef](#)]
69. Ghodke-Puranik, Y.; Lamba, J.K. *Pharmacogenomics*; Academic Press: Cambridge, MA, USA, 2017.
70. Fontana, R.J.; Lown, K.S.; Paine, M.F.; Fortlage, L.; Santella, R.M.; Felton, J.S.; Knize, M.G.; Greenberg, A.; Watkins, P.B. Effects of a char-grilled meat diet on expression of CYP3A, CYP1A, and P-glycoprotein levels in healthy volunteers. *Gastroenterology* **1999**, *117*, 89–98. [[CrossRef](#)]
71. Sanday, K. South Asians and Europeans React Differently to Common Drugs. Available online: <https://medicalxpress.com/news/2011-10-south-asians-europeans-react-differently.html> (accessed on 8 February 2021).
72. Vedani, A.; Dobler, M.; Smieško, M. VirtualToxLab—A platform for estimating the toxic potential of drugs, chemicals and natural products. *Toxicol. Appl. Pharmacol.* **2012**, *261*, 142–153. [[CrossRef](#)]
73. Pires, D.E.V.; Blundell, T.L.; Ascher, D.B. pkCSM Pharmacokinetics. Available online: <http://biosig.unimelb.edu.au/pkcsm/> (accessed on 8 February 2021).
74. Li, X.; Xu, Y.; Lai, L.; Pei, J. Prediction of Human Cytochrome P450 Inhibition Using a Multitask Deep Autoencoder Neural Network. *Mol. Pharm.* **2018**, *15*, 4336–4345. [[CrossRef](#)]
75. Yap, C.W. PaDEL-descriptor: An open source software to calculate molecular descriptors and fingerprints. *J. Comput. Chem.* **2010**, *32*, 1466–1474. [[CrossRef](#)] [[PubMed](#)]
76. Banerjee, P.; Siramshetty, V.B.; Drwal, M.N.; Preissner, R. Computational methods for prediction of in vitro effects of new chemical structures. *J. Cheminform.* **2016**, *8*, 51. [[CrossRef](#)] [[PubMed](#)]
77. Schyman, P.; Liu, R.; Desai, V.; Wallqvist, A. vNN Web Server for ADMET Predictions. *Front. Pharmacol.* **2017**, *8*, 889. [[CrossRef](#)] [[PubMed](#)]
78. Yang, H.; Lou, C.; Sun, L.; Li, J.; Cai, Y.; Wang, Z.; Li, W.; Liu, G.; Tang, Y. admetSAR 2.0: Web-service for prediction and optimization of chemical ADMET properties. *Bioinformatics* **2018**, *35*, 1067–1069. [[CrossRef](#)] [[PubMed](#)]
79. Rostkowski, M.; Spjuth, O.; Rydberg, P. WhichCyp: Prediction of cytochromes P450 inhibition. *Bioinformatics* **2013**, *29*, 2051–2052. [[CrossRef](#)]
80. Daina, A.; Michielin, O.; Zoete, V. SwissADME: A free web tool to evaluate pharmacokinetics, drug-likeness and medicinal chemistry friendliness of small molecules. *Sci. Rep.* **2017**, *7*, 42717. [[CrossRef](#)]
81. Shao, C.-Y.; Su, B.-H.; Tu, Y.-S.; Lin, C.; Lin, O.A.; Tseng, Y.J. CypRules: A rule-based P450 inhibition prediction server. *Bioinformatics* **2015**, *31*, 1869–1871. [[CrossRef](#)]
82. Hong, H.; Xie, Q.; Ge, W.; Qian, F.; Fang, H.; Shi, L.; Su, Z.; Perkins, R.; Tong, W. Mold2, Molecular Descriptors from 2D Structures for Chemoinformatics and Toxicoinformatics. *J. Chem. Inf. Model.* **2008**, *48*, 1337–1344. [[CrossRef](#)]
83. Tian, S.; Djoumbou-Feunang, Y.; Greiner, R.; Wishart, D.S.; Djoumbou, Y.; Greiner, R. CypReact: A Software Tool for in Silico Reactant Prediction for Human Cytochrome P450 Enzymes. *J. Chem. Inf. Model.* **2018**, *58*, 1282–1291. [[CrossRef](#)]
84. Zaretski, J.; Matlock, M.; Swamidass, S.J. XenoSite: Accurately Predicting CYP-Mediated Sites of Metabolism with Neural Networks. *J. Chem. Inf. Model.* **2013**, *53*, 3373–3383. [[CrossRef](#)]
85. Olsen, L.; Montefiori, M.; Tran, K.P.; Jørgensen, F.S. SMARTCyp 3.0: Enhanced cytochrome P450 site-of-metabolism prediction server. *Bioinformatics* **2019**, *35*, 3174–3175. [[CrossRef](#)]
86. Ghosh, J.; Lawless, M.S.; Waldman, M.; Gombar, V.; Fraczkiewicz, R.; Benfenati, E. Modeling ADMET. *Adv. Struct. Saf. Stud.* **2016**, *1425*, 63–83. [[CrossRef](#)]
87. Cleaned ChEMBL Data on Major Cytochrome Inhibitors. 2021. Available online: https://figshare.com/articles/dataset/Cleaned_ChEMBL_Data_on_Major_Cytochrome_Inhibitors/14272661 (accessed on 8 February 2021).