



Structure-based virtual screening of CYP1A1 inhibitors: towards rapid tier-one assessment of potential developmental toxicants

Janice Jia Ni Goh¹ · Julian Behn¹ · Cheng-Shoong Chong^{1,2} · Guorui Zhong¹ · Sebastian Maurer-Stroh^{1,2,3} · Hao Fan^{1,4,5} · Lit-Hsin Loo^{1,6}

Received: 25 May 2021 / Accepted: 17 June 2021 / Published online: 28 June 2021
© The Author(s) 2021

Abstract

Cytochrome P450 1A1 (CYP1A1) metabolizes estrogens, melatonin, and other key endogenous signaling molecules critical for embryonic/fetal development. The enzyme has increasing expression during pregnancy, and its inhibition or knockout increases embryonic/fetal lethality and/or developmental problems. Here, we present a virtual screening model for CYP1A1 inhibitors based on the orthosteric and predicted allosteric sites of the enzyme. Using 1001 reference compounds with CYP1A1 activity data, we optimized the decision thresholds of our model and classified the training compounds with 68.3% balanced accuracy (91.0% sensitivity and 45.7% specificity). We applied our final model to 11 known CYP1A1 orthosteric binders and related compounds, and found that our ranking of the known orthosteric binders generally agrees with the relative activity of CYP1A1 in metabolizing these compounds. We also applied the model to 22 new test compounds with unknown/unclear CYP1A1 inhibitory activity, and predicted 16 of them are CYP1A1 inhibitors. The CYP1A1 potency and modes of inhibition of these 22 compounds were experimentally determined. We confirmed that most predicted inhibitors, including drugs contraindicated during pregnancy (amiodarone, bicalutamide, cyproterone acetate, ketoconazole, and tamoxifen) and environmental agents suspected to be endocrine disruptors (bisphenol A, diethyl and dibutyl phthalates, and zearalenone), are indeed potent inhibitors of CYP1A1. Our results suggest that virtual screening may be used as a rapid tier-one method to screen for potential CYP1A1 inhibitors, and flag them out for further experimental evaluations.

Keywords CYP1A1 inhibition · Developmental toxicity · Endocrine disruptors · Docking · Allosteric · Orthosteric

Introduction

Cytochrome P450 1A1 (CYP1A1) is a highly conserved enzyme that metabolizes many xenobiotics and endogenous signaling molecules (Santes-Palacios et al. 2016). The enzyme is structurally very similar to CYP1A2, another

member from the same CYP family; but the tissue distributions of these two enzymes are quite different. CYP1A1 is highly expressed in the adult lungs, liver, gastrointestinal tract, skin, and other tissues (Nishimura et al. 2003; Choudhary et al. 2005); but CYP1A2 is mostly expressed in the liver (Nishimura et al. 2003; Choudhary et al. 2005). Furthermore, CYP1A1, but not CYP1A2, can be detected in the placenta, embryo, and fetus (Omiecinski et al. 1990;

Janice Jia Ni Goh and Julian Behn have contributed equally to this work.

✉ Hao Fan
fanh@bii.a-star.edu.sg

✉ Lit-Hsin Loo
loolh@bii.a-star.edu.sg

¹ Innovations in Food and Chemical Safety Programme and Bioinformatics Institute, Agency for Science, Technology, and Research, Singapore, Singapore

² Integrative Sciences and Engineering Programme, NUS Graduate School, National University of Singapore, Singapore, Singapore

³ Department of Biological Sciences, National University of Singapore, Singapore, Singapore

⁴ Synthetic Biology Translational Research Program, School of Medicine, National University of Singapore, Singapore, Singapore

⁵ Duke-NUS Medical School, Singapore, Singapore

⁶ Department of Pharmacology, Yong Loo Lin School of Medicine, National University of Singapore, Singapore, Singapore

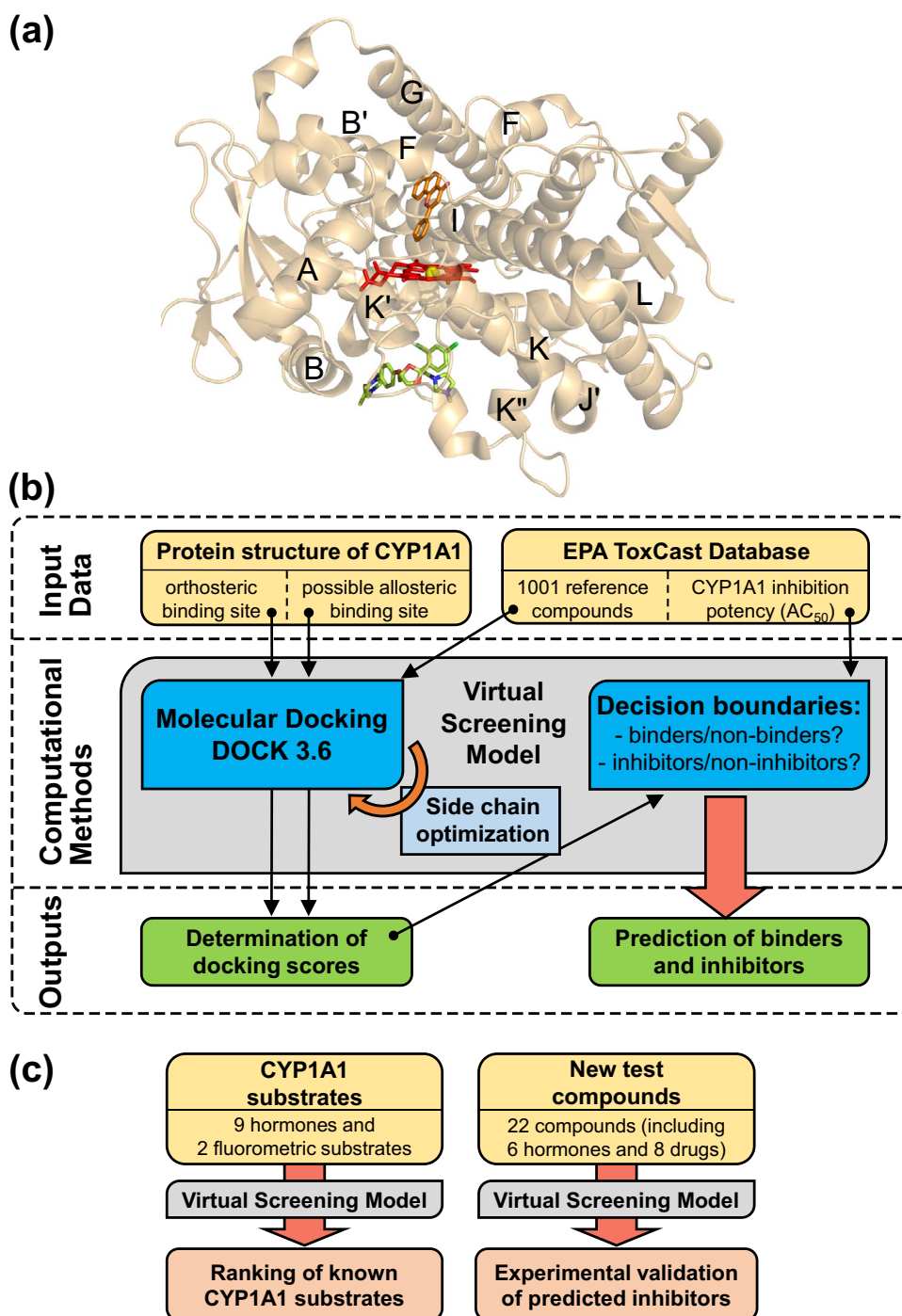
Hakkola et al. 1996; Nishimura et al. 2003; Choudhary et al. 2003, 2005). The enzyme is constitutively active in certain embryonic tissues within specific time windows during the early development of mouse (Choudhary et al. 2003; Campbell et al. 2005) and zebra fish (Otte et al. 2010), as early as the gastrulation stage (Otte et al. 2010). It has high catalytic activity for the oxidative metabolism (including 2-, 4-, 15 α - and 16 α -hydroxylations) of female sex hormones, estradiol (E₂) and estrone (E₁) (Lee et al. 2003). Hydroxylated estrogen metabolites, such as 15 α -hydroxyestriol (also called estetrol or E₄) and 16 α -hydroxyestradiol (also called estriol or E₃), are produced by human fetus (Gurpide et al. 1966) and increase substantially during pregnancy (Adlercreutz and Martin 1976; Holinka et al. 2008). Besides estrogens, CYP1A1 also metabolizes melatonin (Ma et al. 2005), and arachidonic acid (AA) and eicosapentaenoic acid (Schwarz et al. 2004), which are essential for normal embryonic and/or fetal development (Tamura et al. 2008; Yanes et al. 2010). Therefore, it is not surprising that previous animal studies have found that exogenous CYP1A1 inhibitors disrupt embryonic development (Wassenberg and Di Giulio 2004; Yin et al. 2014); and CYP1 gene knockouts increase the rates of embryonic lethality and birth defects (Dragin et al. 2008), and decrease the weights of key organs and the body of fetuses that manage to develop into adults (Dragin et al. 2008; Agbor et al. 2012). All of these results show that functioning CYP1A1 is required for normal embryonic and/or fetal development. Therefore, exogenous inhibitors of CYP1A1, including drugs or environmental agents, may disrupt important endocrine signaling processes, and lead to unintended developmental toxicity effects.

CYP1A1 catalyzes the oxidative metabolisms of specific substrates. The heme enzyme has a narrow and planar active site that restricts substrate orientation (Walsh et al. 2013). The heme forms the “floor” of the pocket (Fig. 1a). Several residues, such as Phe 224, that were previously found to be important for the recognition and binding of substrates provide constraints to the pocket (Walsh et al. 2013). Therefore, xenobiotics that bind to the active site may compete with other endogenous substrates, such as estrogens or melatonin, and inhibit the metabolism of these substrates. X-ray crystallography of CYP1A1 and several potent inhibitors, including α -naphthoflavone, bergamottin, and erlotinib, have found that these inhibitors bind to the active site (Walsh et al. 2013; Bart and Scott 2018). These crystal structures enable the use of computational methods, such as docking analysis, to rapidly screen for molecules that may bind to the active site. Several previous studies have built virtual screening models based on the active site to identify CYP1A1 inhibitors as candidates for anti-cancer drugs (Sridhar et al. 2012; Joshi et al. 2017).

However, the activity of a CYP may also be modulated through allosteric sites away from the active site (Hardy and Wells 2004). There are very few previous reports of allosteric sites that can inhibit the activity of a CYP. It has been suggested that a CYP interacts with cytochrome P450 oxidoreductase (POR) at its proximal side, where the heme group of the CYP is closest to its surface (Bridges et al. 1998), to facilitate electron transfer from NADPH to the CYP active site and oxidization of the bound substrate (Denisov et al. 2005). For efficient electron transfer, stable complex formation and correct positioning of these proteins at the contact regions are crucial (Hlavica et al. 2003). Electrostatic interactions between CYP and POR were postulated to be responsible for their binding (Mayuzumi et al. 1993). In particular, basic residues on the proximal CYP surface were found to be important for direct interaction or correct geometric positioning with acidic residues on POR (Im and Waskell 2011). More recently, it has been proposed that surface exposed hydrophobic residues may also play a vital role in the interaction (Kanaan et al. 2011). Notably, the structure of a self-sufficient bacterial CYP enzyme (Sevrioukova et al. 1999) containing both heme and flavine mononucleotide (FMN) domains revealed that the latter is positioned at the proximal surface of the heme domain. Since the FMN domain can also be found in human POR, similar interactions may occur in the human CYP–POR complex. Although there are more than 50 CYPs present in the human genome, there is only one gene encoding POR (Kandel and Lampe 2014). Thus, CYP1A1 is likely to interact with POR via similar mechanisms. Previous studies have found that POR mutations can cause disordered steroidogenesis in human (Huang et al. 2005), and affect the activities of many CYP enzymes, including CYP1A1 (Hayashi et al. 2003) and 1A2 (Agrawal et al. 2008). Furthermore, similar to CYP1A1, POR knockout in mice results in embryonic lethality and multiple developmental defects (Shen et al. 2002). Thus, the inclusion of additional allosteric sites for CYP1A1 located at or near potential CYP–POR binding sites may help us to more accurately predict CYP1A1 inhibitors.

Here, we present a virtual screening model for CYP1A1 inhibitors based on the orthosteric and a predicted allosteric site of CYP1A1 that overlaps with residues that may be critical for the interaction between CYP and POR. Our strategy was to use docking analysis (Fan et al. 2009; Jaladanki et al. 2021) to estimate the binding affinities via “docking scores” of a large set of reference compounds to the two sites, systematically identify possible relationships between the docking scores and the changes in CYP1A1 activity induced by these compounds, and derive docking-score thresholds that may optimally separate between binders vs non-binders, in particular between

Fig. 1 Building a virtual screening model for CYP1A1 inhibitors based on the orthosteric and allosteric sites of the enzyme. **a** Overview of the previously reported orthosteric binding site and the predicted allosteric binding site on human CYP1A1 (black letters = helices names (Walsh et al. 2013), red sticks = heme, orange sticks = co-crystallized ligand α -naphthoflavone, green sticks = docking pose of ketoconazole in the allosteric site; protein structure is based on PDB ID: 4I8V). **b** Flowchart showing the overall computational procedure to train and use the virtual screening model. **c** Three sets of test compounds unused during the training process were used to validate the virtual screening model (color figure online)



inhibitors vs non-inhibitors (Fig. 1b). We hypothesize that our virtual screening model can be used to computationally screen for CYP1A1 inhibitors, which may cause embryonic and/or fetal developmental toxicities. To test this hypothesis, we applied our final model to 11 known CYP1A1 orthosteric binders and related compounds, and compared the docking scores of these compounds to the known activity levels of CYP1A1 in metabolizing these compounds. We further tested the model by applying it to 22 new compounds with unknown/

unclear CYP1A1 activity, many of which are drugs known to induce fetal development toxicities in humans (Fig. 1c). We then experimentally validated the potency and modes of inhibition of those compounds predicted to be CYP1A1 inhibitors.

Results

Identification of potential allosteric CYP1A1 binding sites

Our virtual screening model is based on a previously published structure (PDB ID: 4I8V) (Walsh et al. 2013) of CYP1A1 and α -naphthoflavone, which binds to the orthosteric site (Supplementary Material 1—Table S1) surrounded by helices *F*, *G* and *I* and the loops around the *B'* region (Fig. 1a). To identify potential allosteric sites based on the structure, we used a computational program—“metaPocket 2.0” that predicts potential binding sites on the protein surface based on a consensus of eight geometry- and energy-based computational methods (Zhang et al. 2011); and another program—“CryptoSite” that predicts cryptic binding sites that may become apparent only after a conformational change of the protein (Cimermancic et al. 2016). CryptoSite computes a per-residue-score that indicates the chance of a residue being part of a cryptic site, and residues with scores larger than 10 are more likely to be part of such sites.

Four clusters of residues were predicted and ranked by metaPocket as potential binding sites. Two of them are located at the known orthosteric site, and thus not further considered. The third cluster is located on the protein surface on the “proximal” side of the heme (opposite to the active site) (Fig. 2a). In this cluster, 68.8% of the residues have CryptoSite scores above the threshold. The fourth cluster, which has a lower rank than the third cluster, is located on the protein surface above the active site. However, only 22.8% of the residues in this cluster have CryptoSite scores above the threshold (Fig. 2a). Therefore, we selected the third cluster as a candidate of an allosteric site (Supplementary Material 1—Table S2). The site is surrounded by helices *B*, *J'*, *K* and *L* and the loop region between the *K'* and *L* helices (Fig. 1a).

Binders to the allosteric site may disrupt CYP–POR interaction

We found that the potential allosteric site is located at the proximal surface of CYP1A1 with most of the predicted binding residues concentrated at the meander coil of the enzyme (cyan regions in Fig. 2b). To examine the possibility

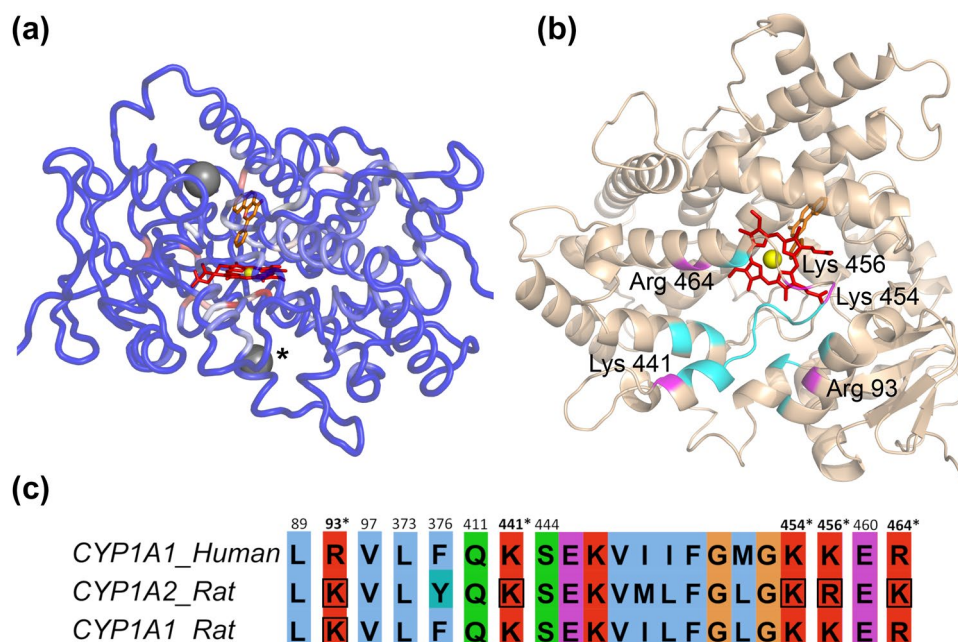


Fig. 2 Predicting a novel CYP1A1 allosteric site at the CYP–POR interface. **a** Visualization of both “metaPocket 2.0” and “CryptoSite” outputs on human CYP1A1 (grey spheres=predicted metaPocket clusters not located at the orthosteric site; *=the final selected allosteric site; red sticks=heme; orange sticks=the co-crystallized ligand, α -naphthoflavone). The backbone is colored according to the CryptoSite scores with a cutoff at 10 (blue=minimum, white=middle, red=maximum score values). **b** Structural model of the proximal surface of human CYP1A1. The predicted allosteric site is cen-

tered at the meander region of the enzyme (in cyan) and overlaps (in magenta) with five labelled basic POR-binding residues. **c** Multiple sequence alignment of mammalian CYP1 enzymes with the 21 residues making up the predicted allosteric binding site (numbers above the alignment=residue positions with respect to human CYP1A1, bolded and asterisked positions=overlaps of the residues with possible POR-binding residues, boxes=POR binding residues that have been experimentally verified in the other mammalian CYP1 enzymes) (color figure online)

that there may be an overlap between the allosteric and POR binding sites, we performed a literature search and identified 17 reported POR-binding residues in the CYP1A family proteins in rats and rabbits (Supplementary Material 1—Table S3). By performing a multi-species sequence alignment (Methods), we mapped the identified residues to the surface of human CYP1A1 and then compared them to the allosteric site. Of the 21 residues defining the allosteric site (Supplementary Material 1—Table S2), we found five basic residues that overlapped with the mapped POR-binding residues on CYP1A1 (Fig. 2c). Three of them (Lys 441, Lys 454 and Lys 456) are located at the meander coil region between helices K' and L, while the other two residues (Arg 93 and Arg 464) are, respectively, positioned in helices B and L (Fig. 2b). Other mapped POR-binding residues in CYP1A1 are far (more than 10 Å) away from the allosteric site, except for Arg 98 that is completely solvent-exposed and partially blocked by Arg 93 from the allosteric site.

Site-directed mutagenesis and chemical modifications of homologous basic residues in mammalian CYP1A1 and CYP1A2 have shown that these five overlapping residues (Fig. 2c and Supplementary Material 1—Table S3) are essential to the CYP–POR interaction. For example, covalent modification or swapping of Lys 453 (mapped to Lys 454 here) in mammalian CYP1A2 with glutamic acid was demonstrated to increase the apparent dissociation constant (K_d) of the interaction and/or diminish POR-dependent catalytic activity (Furuya et al. 1989; Shimizu et al. 1991; Mayuzumi et al. 1993). Similarly, ablation of positive charge at the other four overlapping residues can also perturb CYP–POR interactions (Shimizu et al. 1991; Shen and Strobel 1992; Mayuzumi et al. 1993), underscoring the importance of these residues as contact regions for an electrostatic interaction with POR. Therefore, the overlap between POR-binding residues and the allosteric site suggests that an allosteric binder may disrupt CYP1A1–POR binding in a direct manner. The high affinity (~5–110 nM) (Kandel and Lampe 2014) and small contact area (~967 Å²) (Sevrioukova et al. 1999) of a CYP–POR interaction may make it more amenable to disruptions by small molecule inhibitors (Thompson et al. 2012). Therefore, we hypothesize that the inclusion of this site into our virtual screening model may improve the accuracy in predicting CYP1A1 inhibitors.

Reference compounds and activity data from ToxCast

To develop a virtual screening model for CYP1A1 inhibitors, we used established docking protocols (Fan et al. 2009; Jaladanki et al. 2021) (Methods) to estimate the binding affinities via “docking scores” of a large set of reference compounds to the two sites. We used the US EPA's ToxCast

database (InVitroDB v2.0) (Kavlock et al. 2012), which provides bioactivity data for > 2000 chemical compounds, including a high-throughput CYP1A1 activity assay based on a fluorometric substrate, resorufin benzyl ether (BzRes). After performing quality control and removing compounds that overlap with our validation compounds or are not suitable for docking analysis (Methods), we ended up with 1001 reference compounds (78 are CYP1A1 inhibitors, and 923 are non-CYP1A1 inhibitors). Among the inhibitors, 17 of them have “activity concentration at 50% of maximal activity” (AC_{50}) < 1 μM, and these compounds are called “strong” inhibitors. The other 61 inhibitors with $1 \leq AC_{50} < 100$ μM are called “weak” inhibitors. For some of the compounds, the docking algorithm could not determine any valid docking geometry and all poses got “bumped” by the orthosteric site. For these cases, we applied an unfavorable docking score of $+\infty$ to the respective compounds. Furthermore, if there was more than one valid docking pose for a compound, we only considered the pose with the best docking score. The results were a set of CYP1A1 orthosteric and allosteric docking scores for the 1001 reference compounds (Fig. 3a and Supplementary Material 2).

The large set of reference compounds allow us to systematically identify possible relationships between the CYP1A1 docking scores and inhibition potency, and derive decision thresholds that may optimally separate inhibitors from non-inhibitors. We preferred to use simple linear thresholds, so that the classification of the compounds can be easily interpreted. The intended application of our virtual screening model is to perform tier-one in silico screens for drug-candidate or environmental-agent safety assessments. Further experimental tests will always be performed to confirm the “hits” and detect false positives misidentified during the in silico screens. We expect that the percentage of inhibitors in the tested compounds is usually low. For example, only 7.79% of our reference compounds are inhibitors. Therefore, we prefer a virtual screening model with higher sensitivity than specificity. For this study, we always searched for decision thresholds that provide the maximum sensitivity while maintaining a balanced accuracy at or higher than the 95th percentile level among all possible thresholds.

Threshold for allosteric binders

When studying the distribution of the allosteric docking scores, we found that none of the 1001 tested compounds is being bumped from the allosteric site (Fig. 3b). The result suggests that the site is promiscuous and most compounds may bind to the site but to different extents. Since the site is closed to the potential CYP–POR interacting site, these binders may inhibit the activity of CYP1A1. However, there are other mechanisms to inhibit CYP1A1

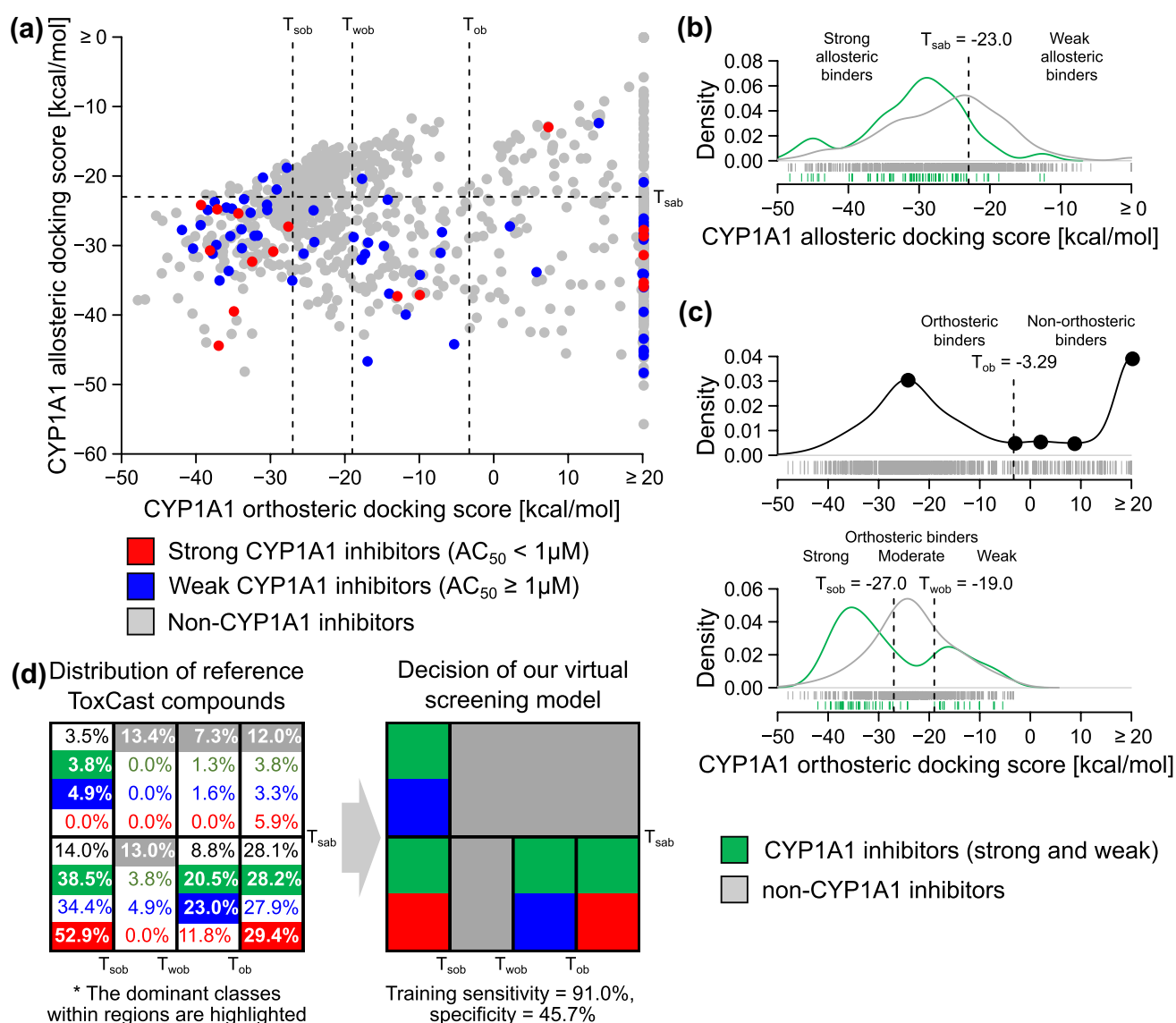


Fig. 3 Determining decision thresholds for inhibitors based on ToxCast compounds. **a** Scatter plot showing the allosteric and orthosteric docking scores for the 1001 reference compounds from ToxCast (blue or red dots=weak or strong CYP1A1 inhibitors based on ToxCast activity data, respectively; gray dots=non-CYP1A1 inhibitors; T_{sob} or T_{wob} =strong or weak orthosteric binder threshold, respectively; T_{ob} =orthosteric binder threshold; T_{sab} =strong allosteric binder threshold). **b** Density plot showing the probability distribution functions of the allosteric docking scores for CYP1A1 inhibitors (green) or non-inhibitors (gray). **c** Density plots showing the probability

distribution functions of the orthosteric docking scores for all compounds (black), CYP1A1 inhibitors (green), or non-inhibitors (gray). **d** Distributions of the reference compounds at the eight regions (left grid map) defined by the four docking-score thresholds (T_{sob} , T_{wob} , T_{ob} , and T_{sab}). The dominant classes within the regions are highlighted (green=CYP1A1 inhibitors, blue=weak CYP1A1 inhibitors, red=strong CYP1A1 inhibitors, gray=non-CYP1A1 inhibitors), and used to derive the final decision rule of our virtual screening model (right grid map) (color figure online)

(such as binding to the orthosteric site), thus an inhibitor may not necessarily be an allosteric binder. Interestingly, we found that inhibitors in the reference compounds generally have more favorable (negative) allosteric docking scores than non-inhibitors (Fig. 3b). The results suggest that predicted stronger allosteric-site binders are more likely to inhibit CYP1A1 catalytic activity than weaker

allosteric-site binders. By performing a grid search across all the allosteric docking scores (Methods), we found that the 95th percentile of all achievable balanced accuracy values was 63.4%. Among all the thresholds with balanced accuracy at or higher than this level, we further found an optimum threshold for strong allosteric binders ($T_{sab} = -23.0$ kcal/mol, Fig. 3b) that maximizes the

sensitivity (91.0%). The balanced accuracy and specificity of this threshold are 63.6% and 36.2%, respectively.

Thresholds for orthosteric binders

However, we observed a very different trend for the distribution of the orthosteric docking scores. We found that 282 (28.2%) of the reference compounds have docking scores larger than 20 kcal/mol, 127 of them are bumped (Fig. 3c). The results suggest that the orthosteric site is more selective than the allosteric site, likely due to the narrow and planar structure of the orthosteric site (Introduction). We note that an orthosteric binder may be an inhibitor and/or a substrate of CYP1A1, but a substrate may not necessarily have a strong potency in inhibiting the activity of CYP1A1. Therefore, unlike the allosteric docking scores, we cannot use the CYP1A1 activity level to separate binders vs non-binders. Instead, we determined the inflection points along the docking score distribution (Fig. 3c, upper panel), and used the first local minima after the first local maxima as a threshold for orthosteric binders ($T_{ob} = -3.29$ kcal/mol). We assume that most of the compounds with docking scores lower than T_{ob} are binders to the orthosteric site.

Surprisingly, when studying the docking-score distributions of inhibitors and non-inhibitors predicted to be orthosteric binders, we found a “gap” in the distribution of inhibitors (Fig. 3c, lower panel). Most of the inhibitors have either highly or lowly negative docking scores, but very few inhibitors have moderately negative docking scores. Importantly, there are also many non-inhibitors with docking scores within this gap, which is around the peak of the orthosteric score distribution for all orthosteric binders (Fig. 3c, upper panel). Therefore, two thresholds are required to optimally separate the inhibitors and non-inhibitors. As before, by performing a grid search across the orthosteric docking scores (Methods), we found that the 95th-percentile balanced accuracy is 66.4%. We further identified two optimum thresholds for strong orthosteric binders ($T_{sob} = -27.0$ kcal/mol) and weak orthosteric binders ($T_{wob} = -19.0$ kcal/mol) that maximize the sensitivity (94.3%) in classifying the orthosteric binders into either inhibitors or non-inhibitors. The balanced accuracy and specificity of these thresholds are 69.2% and 44.1%, respectively. When we used all the three orthosteric thresholds together and considered all the reference compounds, the performance became 64.1% sensitivity and 66.5% specificity (65.3% balanced accuracy). Therefore, a virtual screening model based on orthosteric docking scores alone would have higher specificity but lower sensitivity than a model based on allosteric docking scores alone. The lower sensitivity mostly stems from the large number of inhibitors with unfavorable orthosteric docking scores (orthosteric scores $> T_{ob}$ and allosteric scores $< T_{sab}$, Fig. 3a). Overall, we observed associations between the

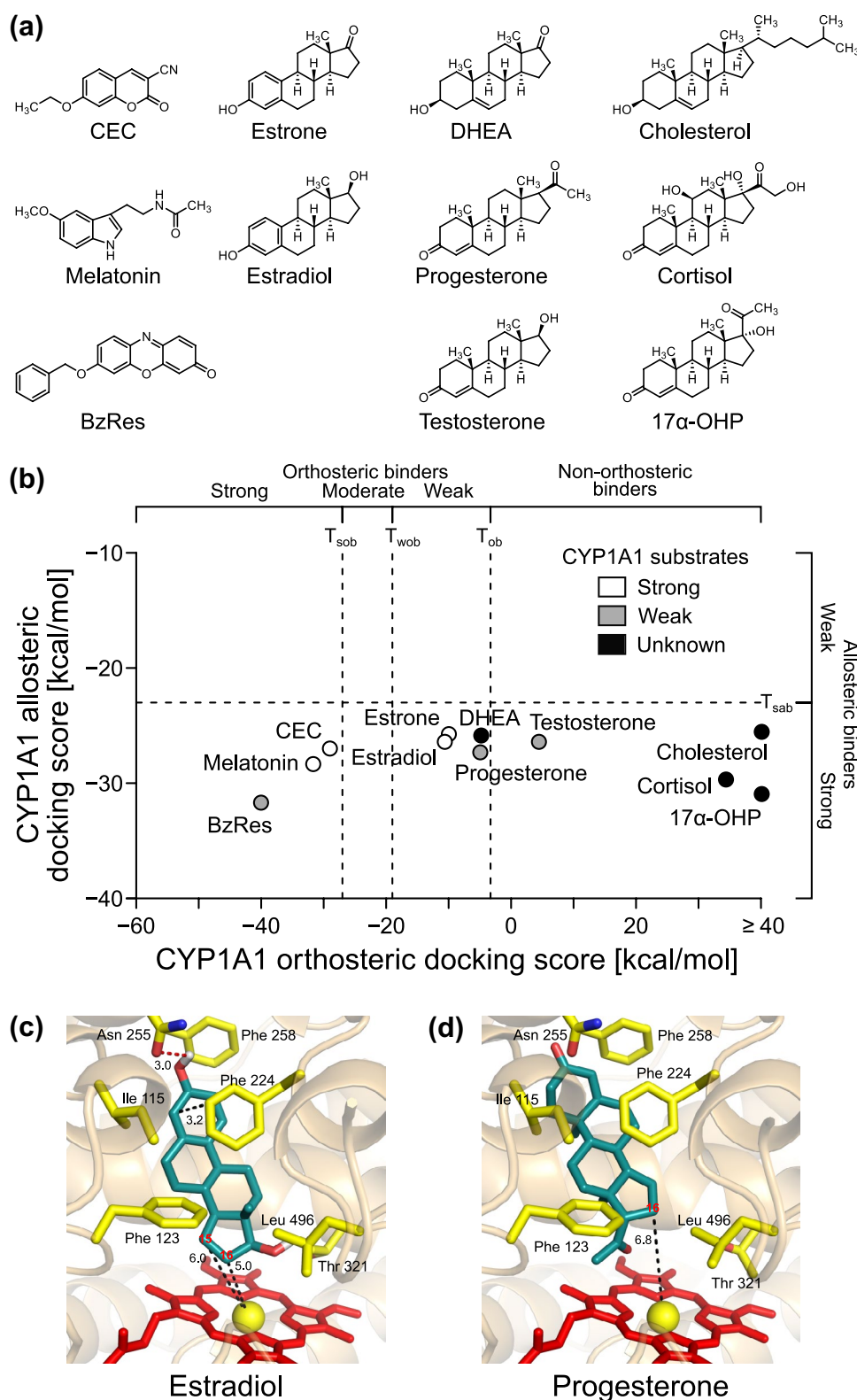
orthosteric and/or allosteric docking scores and CYP1A1 inhibition potency, which agree with our hypothesis that binders to one site or both sites may inhibit CYP1A1.

Final virtual screening model

Using all the four thresholds, we divided all the reference compounds into eight classes of compounds with different predicted CYP1A1 orthosteric and allosteric binding profiles (Fig. 3a). Then, we estimated the distributions of all non-inhibitors, all inhibitors, only strong inhibitors, and only weak inhibitors in these classes of compounds based on the ToxCast CYP1A1 activity data (Fig. 3d, left panel). Because the total numbers of different categories of inhibitors were not the same within a class of compounds, the percentage of all inhibitors was usually not the sum of the percentages of strong and weak inhibitors. We found that the probability of finding a non-inhibitor is higher than finding an inhibitor in predicted weak allosteric binders (allosteric scores $> T_{sab}$), except when the compound is also predicted to be a strong orthosteric binder (orthosteric scores $< T_{sob}$). Conversely, the probability of finding an inhibitor is higher than finding a non-inhibitor in predicted strong allosteric binders (allosteric scores $< T_{sab}$), except when the compound is also predicted to be a moderate orthosteric binder ($T_{sob} < \text{orthosteric scores} < T_{wob}$). Interestingly, the strong inhibitors have the highest probabilities to be found in either the strong allosteric-only binders (allosteric scores $< T_{sab}$ and orthosteric scores $> T_{ob}$) or the strong allosteric and orthosteric binders (allosteric scores $< T_{sab}$ and orthosteric scores $< T_{sob}$). Based on these observations, we derived a final rule to classify a compound as CYP1A1 inhibitors or non-inhibitors according to its allosteric and orthosteric docking scores (Fig. 3d, right panel). The performance of the final rule on the 1001 reference compounds was 91.0% sensitivity and 45.7% specificity (68.4% balanced accuracy).

Application to known binders and other related compounds

To verify that our virtual screening model can properly recognize orthosteric binders of CYP1A1, we applied the final model without any further modification to 11 known CYP1A1 substrates and related compounds, which include nine hormones and two fluorometric substrates (Fig. 4a). Based on experimental measurements from other previous studies, we categorized the compounds into three categories: “strong” CYP1A1 substrates (known to be hydroxylated in at least one reaction with metabolic activity > 1 nmol/min/nmol CYP1A1), “weak” CYP1A1 substrates (known to be hydroxylated in at least one reaction between 0.1 and 1 nmol/min/nmol CYP1A1), and “unknown” compounds (not known to be hydroxylated, or no hydroxylation in the



checked reactions beyond the sensitivity limits of the assays used).

The nine hormones were melatonin, estrone (E_1), estradiol (E_2), dehydroepiandrosterone (DHEA), testosterone,

progesterone, cortisol, 17 α -hydroxyprogesterone (17 α -OHP), and cholesterol (Fig. 4a). A previous study of melatonin metabolism by human CYPs found that this hormone is a strong substrate of CYP1A1 (Ma et al. 2005). A recent

Fig. 4 Validating the model with known CYP1A1 substrates and other related compounds. **a** Structures of the 11 known CYP1A1 substrates and other related compounds (CEC=3-cyano-7-ethoxycoumarin, BzRes=resorufin benzyl ether, DHEA=dehydroepiandrosterone, 17 α -OHP=17 α -hydroxyprogesterone). **b** Scatter plot showing the allosteric and orthosteric docking scores for the compounds (white dots=strong CYP1A1 substrates, gray dots=weak CYP1A1 substrates, black dots=unknown compounds, please refer to the [Results](#) section for the definitions of these categories; T_{sob} or T_{wob} =strong or weak orthosteric binder threshold, respectively; T_{ob} =orthosteric binder threshold; T_{sab} =strong allosteric binder threshold). Docking poses of **c** estradiol or **d** progesterone at the CYP1A1 orthosteric site (red sticks=heme; yellow sticks=selected sidechains; turquoise sticks=estradiol or progesterone; black dashed lines=distances in Å; red numbers=key sites of metabolism). For progesterone, the refined pose is shown (color figure online)

review of steroid hormones hydroxylated by human CYPs (Niwa et al. 2015) found that E_1 and E_2 are strong substrates of CYP1A1; testosterone and progesterone are weak substrates of CYP1A1; and DHEA, cortisol, 17 α -OHP, and cholesterol are known substrates of other CYPs. The review did not provide usable information about CYP1A1's activity on the last group of hormones, suggesting that either no CYP1A1 study has been performed on these hormones or they may not be the main substrates of CYP1A1. DHEA is known to inhibit CYP1A1 (Klinger et al. 2002), and thus may be a CYP1A1 substrate. Despite of the lack of information, this group of hormones were still included in our analysis, because they are structurally related to E_2 (Fig. 4a).

We docked these nine hormones to the orthosteric and allosteric sites of CYP1A1 (Fig. 4b and Supplementary Material 3). All of them were not part of the reference compounds, and thus they were not used to derive the thresholds. We found that all the three strong CYP1A1 substrates (melatonin, E_2 , and E_1) have the best orthosteric dockings scores (−31.79, −10.73, and −10.08 kcal/mol, respectively), which are all below the orthosteric binder threshold (T_{ob} = −3.29 kcal/mol). Our model correctly predicted that melatonin, E_2 , and E_1 are orthosteric binders. The docking pose of E_2 is characterized by an aromatic π – π -interaction between Phe 224 and the steroidal ring A moiety at ~3.2 Å distance (Fig. 4c). Furthermore, the key C15 and C16 positions are the closest to the center of the heme group, with C–Fe distances of 6.0 and 5.0 Å, respectively. E_1 also has a very similar docking pose (data not shown). The results are in strong agreement with the expected hydroxylation at these key positions ([Introduction](#)). Progesterone and DHEA were the other two hormones predicted to be weak orthosteric binders, but their orthosteric docking scores (−5.05 and −4.90 kcal/mol, respectively) are very close to T_{ob} (Fig. 4b). The docking pose of progesterone (Fig. 4d) agrees with a previous finding that CYP1A1 catalyzes the 16 α -hydroxylation of the hormone (Schwarz et al. 2000).

Among all the steroid hormones, E_2 , E_1 , progesterone, and DHEA have the lowest orthosteric docking scores, and these four hormones are predicted to be weak orthosteric binders of CYP1A1 (Fig. 4b). Previous studies have found that these four hormones may inhibit CYP1A1 (Eugster et al. 1993; Klinger et al. 2002), but their relative potency levels are unclear. Thus, we decided to include these hormones as part of our test compounds, whose ability to inhibit CYP1A1 were later experimentally determined and compared. All the other steroid hormones have docking scores > T_{ob} , including another weak CYP1A1 substrate, testosterone. Most of these hormones have very similar allosteric docking scores (Fig. 4b); therefore, the difference in their activity is likely due to the difference in their binding affinity to the orthosteric site.

We also docked two fluorometric substrates of CYP1A1, namely, 3-cyano-7-ethoxycoumarin (CEC) and resorufin benzyl ether (BzRes). The ToxCast program (Kavlock et al. 2012) used BzRes to measure CYP1A1 activity, and we used the resulting data to optimize the thresholds for our virtual screening model (Fig. 3). However, a previous comparison of the activities of 29 rat and human CYPs on nine fluorometric substrates found that CEC is a much stronger substrate for human CYP1A1 than BzRes (40.2 and 0.457 pmol/min/pmol CYP1A1, respectively) (Stresser et al. 2002). Using our model, we found that the docking scores of these two substrates are −29.1 and −40.1 kcal/mol (for orthosteric site), and −26.9 and −31.6 kcal/mol (for allosteric site), respectively (Fig. 4b). Therefore, our model correctly predicted that both compounds are strong orthosteric binders and inhibitors of CYP1A1 (< T_{sob} = −27.0 kcal/mol). However, CEC has an allosteric docking score much closer to the allosteric binder threshold (T_{sab} = −23.0 kcal/mol), and thus is less likely to bind to the allosteric site compared to BzRes. To avoid potential interference with the allosteric site and also due to the stronger reported metabolic activity of the substrate (Stresser et al. 2002), we decided to use CEC, but not BzRes, as the fluorometric CYP1A1 substrate for our validation experiments.

Application to new test compounds

We applied our model to a new set of 18 test compounds completely unused during the threshold optimization process. They include two steroid hormones (E_3 and E_4), eight known or potential endocrine disruptors (anastrozole, bisphenol A or BPA, cyproterone acetate, diethyl phthalate, dibutyl phthalate, propyl paraben, tamoxifen, and zearalenone), and eight known lung toxicants from a previous study of pulmonary toxicity (Lee et al. 2018) (amiodarone, bicalutamide, cyclophosphamide, 2,4'-DDT, diacetyl, ketoconazole, β -myrcene, and *p*-phenylenediamine). The lung

Table 1 Potency and mode of inhibition of the test compounds

Name	CASN	FDA pregnancy category	IC ₅₀	(95%CI)	K _i	(95%CI)	α	(95% CI)	MOI
Ketoconazole	65277-42-1	C	0.099	(0.066–0.13)	0.11	(0.063–0.24)	0.68	(0.27–1.4)	Noncompetitive
Zearalenone	17924-92-4	–	0.7	(0.49–0.91)	0.66	(0.43–1.1)	3.5	(1.5–8.3)	Mixed
Bisphenol A	80-05-7	–	2.4	(2–2.8)	0.53	(0.36–0.85)	9.1	(3.8–40)	Competitive
Amiodarone	19774-82-4	D	7.8	(4.3–11)	1.4	(0.85–2.7)	4.7	(1.7–15)	Competitive
Cyproterone acetate	427-51-0	D	9.9	(1.5–18)	8.8	(5.7–15)	7	(2.8–28)	Competitive
Dibutyl phthalate	84-74-2	–	10	(5.6–15)	1.6	(1.2–2.4)	13	(5.5–210)	Competitive
Tamoxifen	10540-29-1	D	12	(7.2–16)	3.6	(2–8.2)	7.2	(2.2–44)	Competitive
Bicalutamide	90357-06-5	X	14	(11–18)	6.6	(4.4–11)	2.7	(1.2–5.9)	Mixed
Diethyl phthalate	84-66-2	–	38	(21–56)	10	(6.4–18)	8.4	(3.2–72)	Competitive
Propyl paraben	94-13-3	–	39	(29–48)	17	(11–27)	6.2	(2.5–21)	Competitive
Progesterone	57-83-0	B	41	(19–63)	20	(12–41)	4.3	(1.5–18)	Competitive
Estradiol	50-28-2	X	43	(17–70)	13	(8.2–22)	4.3	(1.7–15)	Competitive
β-Myrcene	123-35-3	–	52	(42–62)	11	(7.3–17)	3.5	(1.6–7.6)	Mixed
DHEA	53-43-0	–	130	(67–200)	–	–	–	–	–
Anastrozole	120511-73-1	D	190	(19–370)	–	–	–	–	–
Estetrol	15183-37-6	–	> 200	–	–	–	–	–	–
2,4'-DDT	789-02-6	–	> 200	–	–	–	–	–	–
Estrone	53-16-7	–	> 200	–	–	–	–	–	–
p-Phenylenediamine	106-50-3	–	> 200	–	–	–	–	–	–
Cyclophosphamide	6055-19-2	D	∞	–	–	–	–	–	–
Diacetyl	431-03-8	–	∞	–	–	–	–	–	–
Estriol	50-27-1	–	∞	–	–	–	–	–	–

DHEA dehydroepiandrosterone; MOI mode of inhibition; 95% CI 95% confidence interval; unit for IC₅₀ and K_i=μM; ∞=the best fitted model was the constant model; > 200=the estimated IC₅₀ value was larger than the maximum tested concentration, 200 μM; – not tested, computed, or available; see [Methods](#) for the rules used to assign MOI based on α)

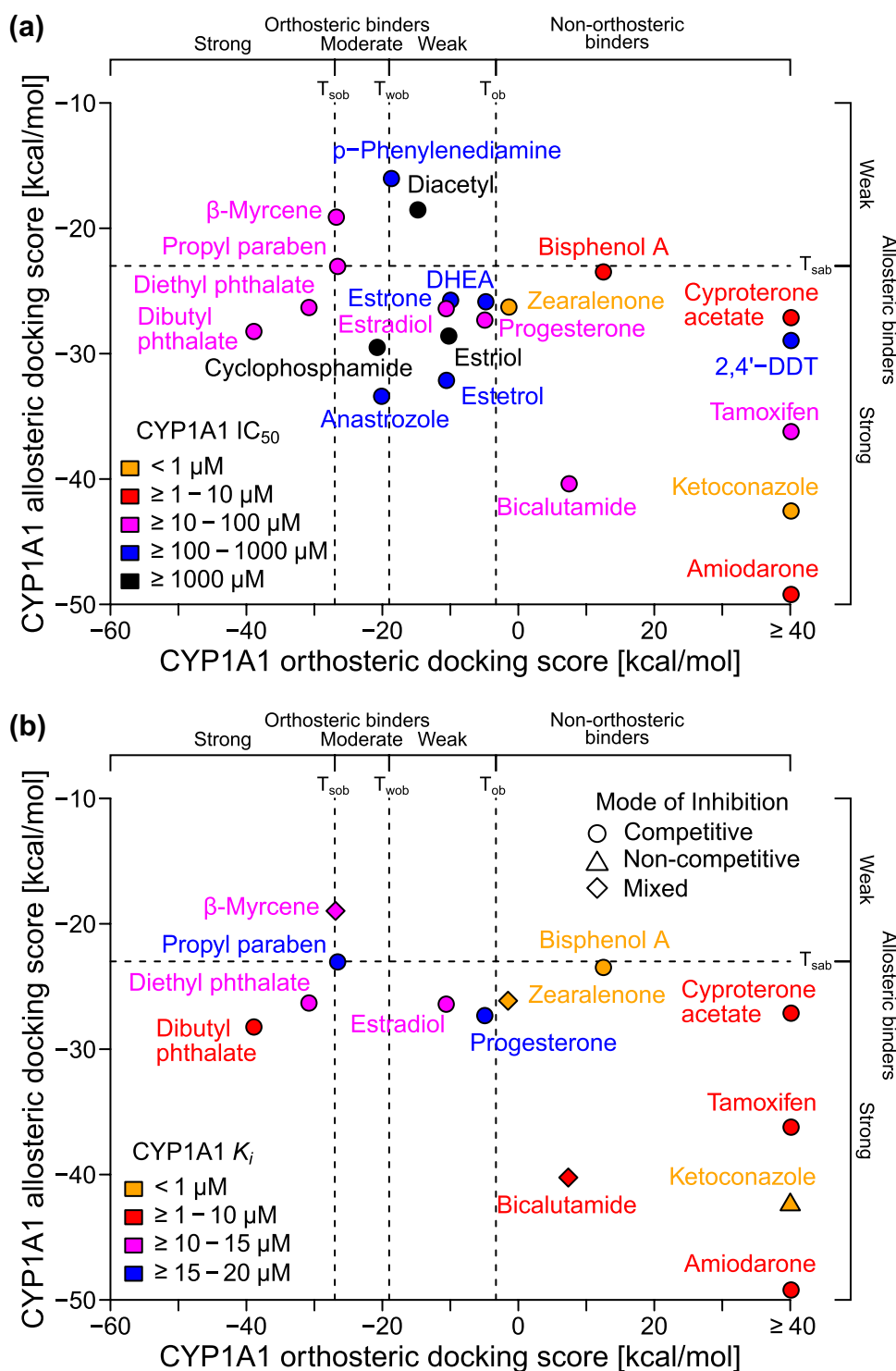
toxicants were included, because CYP1A1 is also highly expressed in the adult lung tissues (Nishimura et al. 2003), and thus these compounds may interact with the enzyme. Together with the four steroid hormone substrates predicted to be weak CYP1A1 inhibitors earlier, these 22 test compounds provide a rich and diverse set of molecules to validate our virtual screening model.

Nine of the test compounds are FDA-approved drugs, namely, amiodarone, anastrozole, bicalutamide, cyclophosphamide, cyproterone acetate, E₂, ketoconazole, progesterone, and tamoxifen. Thus, the human fetal developmental effects of most of these drugs have been clinically evaluated, and we obtained their assigned FDA pregnancy categories from their packaging labels (Table 1 and [Methods](#)). Two of them (bicalutamide and E₂) are of Category X, which have studies in animals or humans showing fetal abnormalities and the risk of using these drugs in a pregnant woman clearly outweighs any possible benefit. Five of them (amiodarone, anastrozole, cyclophosphamide, cyproterone acetate, and tamoxifen) are of Category D, which have positive evidence of human fetal risk based on adverse reaction data from investigational or marketing experience or studies

in humans, but potential benefits may warrant the use of these drugs in pregnant women despite of the risks. Finally, ketoconazole is of Category C, which cause some adverse fetal effects on animal reproduction studies, but there is no adequate or well-controlled human study to confirm these effects; and progesterone is of Category B, which have low risks to fetus based on animal studies. Thus, most of these nine drugs are known development toxicants to humans, except for ketoconazole and progesterone, whose human developmental effects are unclear/unstudied.

Our models predicted that 16 of the test compounds are CYP1A1 inhibitors and six are non-CYP1A1 inhibitors (Fig. 5a). To validate the predictions, we experimentally determined the half maximal inhibitory concentration (IC₅₀) of these compounds in inhibiting the activity of CYP1A1 in metabolizing CEC (Table 1). We found 13 compounds with IC₅₀ < 100 μM (“measured inhibitors”) and 9 compounds with IC₅₀ ≥ 100 μM (“measured very weak or non-inhibitors”). Our model correctly identified 11 of the 13 measured inhibitors (sensitivity = 84.6%) and 4 of the 9 measured very weak or non-inhibitors (specificity = 44.4%). The validation performance is very similar to the training performance

Fig. 5 Validating the model with new test compounds using a CYP1A1 activity inhibition assay. Scatter plot showing the experimentally determined **a** IC_{50} values and **b** K_i values and predicted modes of inhibition for 22 test compounds (DHEA = dehydroepiandrosterone; T_{sob} or T_{wob} = strong or weak orthosteric binder threshold, respectively; T_{ob} = orthosteric binder threshold; T_{sab} = strong allosteric binder threshold)



that we obtained during the earlier model building process (Fig. 3d). The only two mis-detected inhibitors (β -myrcene and propyl paraben) are very close to the T_{sob} decision boundary. Furthermore, the two most potent measured inhibitors (zearalenone and ketoconazole, both with $IC_{50} < 1 \mu M$) were predicted to be allosteric-only binders (Fig. 5a). The results agree with our earlier observation from the reference

compounds that 29.4% of strong CYP1A1 inhibitors are predicted to be allosteric-only binders (Fig. 3d).

Among all the tested steroid hormones, we confirmed that both E_2 and progesterone have the highest potency, $IC_{50} = 43 \mu M$ (17–70 μM , 95%CI) and 41 μM (19–63 μM , 95%CI), respectively. Four other steroid hormones (E_1 , E_3 , E_4 , and DHEA) were false positives. In earlier analysis

(Fig. 4b), we found that the orthosteric docking score of DHEA is very close to the orthosteric binder threshold (T_{ob}). We confirmed that DHEA is a very weak CYP1A1 inhibitor with $IC_{50} = 130 \mu M$ (67–200 μM , 95%CI). Furthermore, among the two tested phthalates, dibutyl phthalate has both more favorable orthosteric and allosteric docking scores than diethyl phthalate (Fig. 5a), despite of the fact that dibutyl phthalate has longer chain length than diethyl phthalate. We predicted that both phthalates are strong CYP1A1 inhibitors, but dibutyl phthalate may have a higher potency than diethyl phthalate. Again, we confirmed that the IC_{50} for dibutyl and diethyl phthalates are 10 and 38 μM , respectively.

Dissociation constants and modes of inhibition

IC_{50} is a potency measurement determined at a single substrate concentration, and thus the measurement depends on the concentration of the substrate used. To determine the potency and mode of inhibition of the compounds independent of the substrate concentration used, we performed additional experiments at multiple substrate concentrations (Methods). The resulted data allowed us to model the enzymatic kinetics and estimate the equilibrium dissociation constant for the enzyme–inhibitor complex (“ K_i ”), and also the degree to which the binding of inhibitor changes the affinity of the enzyme for substrate (“ α ”) (Methods). When α is much larger than one, the binding of an inhibitor prevents the binding of a substrate, thus the inhibition is “competitive”; when α is close to one, the binding of an inhibitor does not change the binding of a substrate, thus the inhibition is “noncompetitive”; otherwise, the inhibition is “mixed”. We only determined K_i and α for the 13 CYP1A1 inhibitors with $IC_{50} < 100 \mu M$.

Again, we found that the K_i values generally agree with our model predictions (Table 1). The highly potent compounds ($K_i < 10 \mu M$) were predicted to be either strong allosteric-only binders or strong orthosteric and allosteric binders (Fig. 5b). Interestingly, ketoconazole, zearalenone, and BPA are the three compounds with the highest potency ($K_i < 1 \mu M$). We also found that most of the inhibitors induced competitive inhibition, except for ketoconazole, which is non-competitive, and β -myrcene, which is mixed (Fig. 5b). Together, our results suggest that our virtual screening model is working as expected, and correctly predicted most of the CYP1A1 inhibitors while maintaining a reasonable specificity.

Discussion

We have developed a virtual screening model for CYP1A1 inhibitors based on the orthosteric site and a predicted allosteric site of the enzyme. Binders to the allosteric site

of CYP1A1 may directly disrupt the interaction and electron transfers between POR and CYP1A1, and thus inhibit CYP1A1’s catalytic activity. However, binders to the allosteric site may also induce other long-range effects that can result in conformational changes at the enzyme’s active site (Grover 2013) or indirect interferences of the interaction between CYP1A1 and POR or other proteins (Modell et al. 2016). This may explain why we found that many compounds predicted to be allosteric-only binders induce competitive inhibition of CYP1A1 (Fig. 5b and Table 1). Regardless of the underlying mechanisms, our results suggest that a more sensitive predictor for CYP1A1 inhibition can be obtained when both the orthosteric and allosteric sites are considered (91% sensitivity) compared to when only the orthosteric site is considered (64.1% sensitivity) while maintaining similar or slightly higher levels of balanced accuracy (68.4% and 65.3%, respectively). Due to the known roles and knockout/inhibition effects of CYP1A1 to embryonic and/or fetal development (Introduction), it is very likely that CYP1A1 inhibitors may cause developmental toxicity.

In our study, we found that several drugs, including those from FDA Pregnancy Category C, D or X, are predicted to be CYP1A1 inhibitors (Fig. 5). Among them, we have experimentally verified that amiodarone, bicalutamide, ketoconazole, tamoxifen, and cyproterone acetate are potent CYP1A1 inhibitors (all with $K_i < 10 \mu M$). Amiodarone is an antiarrhythmic drug known to inhibit several human CYPs (Ohyama et al. 2000). We found that the drug has $K_i = 1.4 \mu M$ (0.85–2.7 μM , 95%CI) and $\alpha = 4.7$ (1.7–15, 95%CI) for CYP1A1 inhibition, suggesting that it has a mixed mode of inhibition. Our measurements were very close to the values determined by another previous study for desethylamiodarone (DEA), an active metabolite of amiodarone ($K_i = 1.5 \mu M$, $\alpha = 5.7$) (Ohyama et al. 2000). Interestingly, that study compared eight CYPs, and found that DEA has the highest potency in inhibiting CYP1A1. Therefore, both amiodarone and DEA are potent CYP1A1 inhibitors. Previous studies have found that gestational exposure to amiodarone may lead to intrauterine growth restriction in rats (Hill and Reasor 1991) and humans (Widerhorn et al. 1991). Therefore, our results suggest that CYP1A1 inhibition may be one of the mechanisms underlying the observed amiodarone effects.

Another drug that we tested was ketoconazole, an antifungal drug known to inhibit CYP3A4’s orthosteric site by ligating its azole nitrogen to the enzyme’s heme group (Ekroos and Sjogren 2006). However, our model predicted that ketoconazole is a strong CYP1A1 inhibitor that only binds to CYP1A1’s allosteric site. We experimentally verified that the drug has $K_i = 110 \text{ nM}$ (63–240 nM , 95%CI) and $\alpha = 0.68$ (0.27–1.4, 95%CI). Again, our measurements were very close to the values determined from a previous study of the same compound for CYP1A1 inhibition ($K_i = 36.6 \text{ nM}$)

(Paine et al. 1999). The measured α value suggests that ketoconazole is a non-competitive inhibitor of CYP1A1, which agrees with our prediction that the drug does not bind to the orthosteric site of CYP1A1. These results provide a strong support to our hypothesis that binding to the allosteric site alone is sufficient to inhibit CYP1A1 activity. Interestingly, another related antifungal drug, fluconazole, was one of our reference chemicals and predicted to be a CYP1A1 inhibitor by our model (Supplementary Material 2). A previous study found that fluconazole can inhibit CYP1A1's metabolism of AA, and the formation of biologically active AA metabolites, such as hydroxyeicosatetraenoic acids (HETE) (El-Sherbeni and El-Kadi 2016). Other previous animal studies have found that both ketoconazole and fluconazole induce embryo lethality and teratogenicity at high doses, but similar effects have so far been observed only for fluconazole in humans (Pilmis et al. 2015). Therefore, FDA classified ketoconazole as Category C, but fluconazole as Category D. Due to their more severe developmental toxicity effects than CYP1A1 knockouts, the two azole antifungals are likely to inhibit additional targets beyond CYP1A1.

We also found that several suspected endocrine disruptors, namely, BPA, zearalenone, dibutyl phthalate (DBP), diethyl phthalate (DEP), and propyl paraben are CYP1A1 inhibitors. Phthalates and BPA are used in plastics, parabens are preservatives used in personal care products, and zearalenone is a naturally occurring mycotoxin that can be found in high concentrations in dairy products and cereals. All of them are suspected endocrine disruptors that may activate estrogen receptors under in vitro conditions (Blair et al. 2000). In previous studies, these compounds were also found to inhibit ethoxyresorufin-*O*-deethylation (EROD), a marker for CYP1A1 activity, in rat liver microsomes (Ozaki et al. 2016), human MCF-7 cell line (Yu et al. 2004), or mouse Hepa-1c1c7 cell line (Jeong et al. 2000). However, due to the existence of other CYPs and/or proteins that may regulate CYP1A1's expression (such as the aryl hydrocarbon receptor, AhR) in the assays used by these studies, the observed effects may not be specifically attributed to direct CYP1A1 inhibition (Burke et al. 1994). In our study, we experimentally confirmed that human CYP1A1 can be inhibited by these compounds (K_i = 0.53, 0.66, 1.6, 10, and 17 μ M, respectively). The high potency of BPA, zearalenone, and DBP (K_i < 2 μ M) suggests that these suspected endocrine disruptors may interfere with the endocrine systems not only by directly modulating the activity of key signaling receptors (such as estrogen receptors), but also by indirectly inhibiting the activity of enzymes that generate or convert key signaling molecules (such as estrogens) in these systems.

Besides CYP1A1, several other CYPs are also highly expressed in the fetus or placenta. They include CYP1B1, CYP2C8, CYP2D6, CYP2E1, CYP3A4, CYP3A5 and CYP3A7 (Hakkola et al. 1998), which may also play

important roles during embryonic or fetal development. Our approach of building a virtual screening model based on orthosteric and allosteric sites is general, and thus should be applicable to these other CYPs as well. By combining models for several CYPs and other key enzymes required for embryonic/fetal development, we may be able to more accurately predict compounds that can cause developmental toxicity, and filter out potential developmental toxicants for further experimental tests during drug-candidate or environmental-agent safety screening.

Methods

Topological modelling of the spatial distribution of potential POR-binding residues

To identify CYP1A1 residues important for the binding to POR, a literature search was conducted to obtain collective evidence on possible binding residues based on past studies in enzymes from the CYP1 family. A multi-species sequence alignment was then generated by MAFFT (<https://mafft.cbrc.jp/alignment/server/>) using parameters for the L-INS-i iterative refinement method (Kato et al. 2002). The illustration for the multi-sequence alignment was made using Jalview 2.10.3b13 and colored according to physicochemical properties using the Clustal X scheme at <http://www.jalview.org/help/html/colourSchemes/clustal.html>. Protein sequences used for the alignment were obtained from the UniProt database (The UniProt Consortium 2021) and their respective accession IDs, related enzyme and POR binding residue information as well as the literature references they were derived from are summarized in Supplementary Material 1—Table S3. Based on the results of the alignment, these binding residues were then mapped to their corresponding residues on the human CYP1A1 structure used. Finally, a comparison was made between the POR-binding residues and the predicted allosteric binding site residues to identify overlapping residues.

Reference compounds and activity data

The October 2015 release version of the “ToxCast & Tox21 Summary Files” data set was retrieved from US EPA server (ftp://newftp.epa.gov/COMPTOX/High_Throughput_Screening_Data/Previous_Data/ToxCast_Data_Release_Oct_2015/). For this study, the assay with endpoint name NVS_ADME_hCYP1A1 was used. The assay was based on the fluorescence intensity signal changes due to the metabolism of BzRes catalyzed by human CYP1A1. The reference compound set was derived from the original ToxCast data set through a multi-step scheme. First, all compounds that

overlap with the test compounds used in the validation studies were removed. Then, we used the assay variables and model fitting parameters from the ToxCast data to decide if a compound is an inhibitor or not. Compounds that were found to be inactive at the single-concentration tests and thus not selected for multi-concentration tests were assigned as “non-inhibitors”. In addition, compounds tested in multiple concentrations but with a “constant” fitted concentration–response model were also assigned as “non-inhibitors”. Compounds with a non-constant fitted model, but with either AC_{50} value larger than 10 μ M or maximum response values less than 10% were discarded from the reference compound set, because the activity data for these compounds was not sufficient to clearly distinguish them as either “inhibitors” or “non-inhibitors”. All the remaining compounds were assigned as “inhibitors”. Finally, compounds with missing or non-valid SMILES strings or compounds with less than 5 atoms had to be removed from the reference set due to their unsuitability for the docking approach. The final CYP1A1 reference data set consisted of 1001 compounds, of which 78 were CYP1A1 inhibitors (Supplementary Material 2).

Protein structure

The protein structure for CYP1A1 (PDB ID: 4I8V, UniProt accession: P04798) (Walsh et al. 2013) was downloaded from the RCSB PDB (<https://www.rcsb.org/>). Only chain A was considered. The co-crystallized ligand α -naphthoflavone was removed from the protein structure before docking, while the cofactor heme was kept. The procedures used to identify the allosteric site are described in the Results section. Graphics for all structural models were created using Pymol 1.7.2.1.

Docking

We used the DOCK 3.6 (Mysinger and Shoichet 2010) for virtual screening of the ToxCast reference data set against the CYP1A1 protein structure. Docking was performed at the known orthosteric binding pocket and at our predicted allosteric binding site. For each site, we performed a cycle of docking runs to improve the docking poses of the ToxCast reference compounds. Between two consecutive docking runs, we always performed a local sidechain optimization using the “Protein Local Optimization Program” (PLOP) program (Jacobson et al. 2002). For progesterone, we found some steric clashes with the rigid orthosteric pocket from the CYP1A1 protein structure. Therefore, we did small sidechain refinement to the crystal structure and re-dock the hormone. The docked compounds were ranked by the docking energy function that is the sum of van der Waals, Poisson–Boltzmann electrostatic, and ligand desolvation

penalty terms. For some compounds the algorithm could not determine a valid docking geometry (in case all poses got “bumped”). In this case, we applied an unfavorable docking score of $+\infty$ to the bumped compound. Furthermore, if there was more than one valid docking pose for a certain compound, we only considered the best scoring pose.

Allosteric and orthosteric thresholds

For the allosteric docking scores, we first capped the maximum scores to 0 kcal/mol, i.e., assigning 0 kcal/mol to compounds with docking scores larger than the value. Then, a grid search was used to determine the sensitivity and specificity in classifying inhibitors and non-inhibitors based on a threshold: $T_{\text{sab}} \in \{-60, -59, -58, \dots, -2, -1, 0\}$. After that, all the thresholds with balanced accuracy levels equal or larger than the 95th percentile value were identified, and the final threshold was chosen to have the maximum sensitivity among the identified thresholds.

For the orthosteric docking scores, we first capped the maximum scores to 20 kcal/mol, i.e., assigning 20 kcal/mol to compounds with docking scores larger than the value. Then, we estimated the probability density function (PDF) of the scores using the density() function with smoothing bandwidth = 3 under the R environment (v.3.5.2; The R Foundation, Vienna, Austria). Inflection points were determined from the smoothed curve, and the threshold for orthosteric binders (T_{ob}) was determined to be the first local minima after the first local maxima. Then, for all predicted orthosteric binders (orthosteric score $< T_{\text{ob}}$), a grid search was used to determine the sensitivity and specificity in classifying inhibitors and non-inhibitors based on two thresholds: $T_{\text{sob}} \in \{-60, -59, -58, \dots, -5, -4\}$ and $T_{\text{wob}} \in \{-60, -59, -58, \dots, -5, -4\}$, where $T_{\text{sob}} < T_{\text{wob}}$. After that, all the threshold pairs with balanced accuracy levels equal or larger than the 95th percentile value were identified, and the final threshold pair was chosen to have the maximum sensitivity among the identified pairs.

Materials

High-performance liquid chromatography (HPLC)-grade acetonitrile (ACN) was used (VWR, Pennsylvania, USA). All test compounds are of analytical-grade ($\geq 99.0\%$ purity; from Sigma-Aldrich, St. Louis, USA or Cayman Chemicals, Michigan, USA). Water was obtained using a water purification system (Milli-Q; Millipore, Billerica, MA, USA).

CYP1A1 activity assays

The compounds were evaluated for their potency in inhibiting CYP1A1’s metabolism of CEC. The assay was based

on bactosomes containing human recombinant CYP1A1 (rCYP1A1) and cytochrome P450 oxidoreductase (#EZ014; Cypex, Scotland, UK), and NADPH regenerating system containing NADPH A (NADP⁺ and glucose-6-phosphate) and B (glucose-6-phosphate dehydrogenase) solutions (#451220 and #451200, respectively; BD Gentest, Woburn, MA, USA). First, the test compounds were dissolved and titrated in ACN, and preincubated with a 30 μ L incubation mixture consisting of 5 μ M CEC, 5 pmol/mL rCYP1A1, 1:100 NADPH B, and 100 mM potassium phosphate buffer (pH 7.4) for 15 min at 37 °C. The final concentrations of the test compounds were 0.00, 0.091, 0.274, 0.823, 2.47, 7.41, 22.2, 66.7, and 200 μ M; except for ketoconazole were 0.00, 0.009, 0.027, 0.082, 0.247, 0.741, 2.22, 6.67, and 20 μ M. The test solutions were prepared in 384-well plates (Corning, New York, USA), and in quadruple technical replicates. Then, the reactions were initiated by adding 10 μ L NADPH A to the test solutions, yielding 50 μ L final test mixtures with 2% ACN (v/v). After incubation for 2.5 min at 37 °C, the reactions were quenched with 30 μ L ice-cold ACN. Finally, the fluorescence intensity of the generated metabolites was immediately quantified using a multimode plate reader (Synergy Mx; BioTek, VT, USA) at $\lambda_{\text{ex}} = 405$ nm and $\lambda_{\text{em}} = 460$ nm. In every batch of experiments, ketoconazole was always included as a positive control. At least two independent experimental replicates were performed for each compound. To determine the K_i and α values of a compound, the same CYP1A1 activity assay was repeated six times in different CEC concentrations (1.56, 3.125, 6.25, 12.5, 25.0, and 50.0 μ M). The final concentrations of the compounds were kept the same.

IC₅₀ computation

For each experimental replicate, the median fluorescence intensity value across the four technical replicates was first determined. Then, an average activity value was obtained by taking the mean of the median values from at least two experimental replicates. For each compound, the ratios of the average activity levels at all the tested concentrations with respect to the 0 μ M control were determined, and used to fit two concentration–response–curve models. The first model is a two-parameter log-logistic model, $\Delta_{\text{model}}([I]) = 1/(1 + (b(\log([I]) - \log(\text{IC}_{50}))))$, where $[I]$ is the inhibitor concentration, b is the steepness of the fitted curve, and IC_{50} is the half maximal inhibitory concentration. The second model is a constant model, $\Delta_{\text{model}}([I]) = 1$. The Akaike's Information Criterion (AIC) (Akaike 1974) was then used to identify the best fitted model. Compounds optimally fitted by the constant model were deemed to be “non-inhibitors”, and their IC_{50} values were set to ∞ . The DRC library (v3.0–1) under the R environment (v3.5.2; The R Foundation, Vienna, Austria) was used to fit the models,

including the estimation of the 95% confidence intervals of the estimated IC_{50} values.

K_i and α computation

First, we estimated the maximum rate of the reaction (V_{max}) and the Michaelis–Menten constant of the enzyme (K_m) using a non-linear regression model, $v = (V_{\text{max}} \cdot [S]) / (K_m + [S])$, where v is the velocity of the reaction and $[S]$ is the concentration of the substrate. The model was fitted only with measurements obtained without the inhibitor (i.e., the concentration of the inhibitor, $[I] = 0$). Then, we estimated the inhibition constant of the inhibitor values of (K_i) and the degree to which the binding of inhibitor changes the affinity of the enzyme for substrate (α) using another non-linear regression model, $v = (V_{\text{max}} \cdot [S]) / \left\{ K_m \cdot \left(1 + \frac{[I]}{K_i} \right) + [S] \cdot \left(1 + \frac{[I]}{\alpha \cdot K_i} \right) \right\}$, with all the measurements and estimated V_{max} and K_m values. If the lower and upper limits of the 95% confidence interval (CI) of α were larger than 1 and 10, respectively, the mode of inhibition (MOI) was predicted to be “competitive”; if the lower and upper limits of the 95% CI of α were smaller than 1 and between 1 and 10, respectively, the MOI was predicted to be “Noncompetitive”; and the MOI for all other conditions was predicted to be “Mixed”. The `nls()` function under the R environment (v.3.5.2; The R Foundation, Vienna, Austria) was used to perform the non-linear regression model fitting, including the estimations of the 95% confidence intervals of all the estimated parameters.

FDA pregnancy category

We searched for the labels of all the clinically approved drugs from the U.S. National Library of Medicine (NLM) DailyMed website (<https://dailymed.nlm.nih.gov/dailymed/>).

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s00204-021-03111-2>.

Acknowledgements We thank for our colleagues from the Toxicity Mode of Action Discovery (ToxMAD) Platform and the Innovations in Food and Chemical Safety (IFCS) Program for their support and discussions. The work was funded by a grant from the Biomedical Research Council (BMRC)—Economic Development Board (EDB) Industry Alignment Fund (IAF-PP6, H18/01/a0/B14), Agency for Science, Technology and Research (A*STAR), Singapore. JJNG was supported by National Science Scholarship (NSS), A*STAR. JB was supported by Singapore International Pre-Graduate Award (SIPGA), A*STAR.

Author contributions LHL and HF conceived and designed the study. JJNG performed the compound selection and enzymatic reaction experiments. JB and HF performed the allosteric site prediction,

molecular docking, and structural analyses. JJNG and LHL performed the enzymatic reaction modeling and other computational analyses. GZ performed CYP validation experiments. CSC and SMS performed the protein sequence homology analysis. All authors participated in the writing of the manuscript.

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Adlercreutz H, Martin F (1976) Oestrogen in human pregnancy faeces. *Eur J Endocrinol* 83:410–419. <https://doi.org/10.1530/acta.0.0830410>
- Agbor LN, Walsh MT, Boberg JR, Walker MK (2012) Elevated blood pressure in cytochrome P4501A1 knockout mice is associated with reduced vasodilation to omega-3 polyunsaturated fatty acids. *Toxicol Appl Pharmacol* 264:351–360. <https://doi.org/10.1016/j.taap.2012.09.007>
- Agrawal V, Huang N, Miller WL (2008) Pharmacogenetics of P450 oxidoreductase: effect of sequence variants on activities of CYP1A2 and CYP2C19. *Pharmacogenet Genomics* 18:569. <https://doi.org/10.1097/FPC.0b013e32830054ac>
- Akaike H (1974) A new look at the statistical model identification. *IEEE Trans Autom Control* 19:716–723. <https://doi.org/10.1109/TAC.1974.1100705>
- Bart AG, Scott EE (2018) Structures of human cytochrome P450 1A1 with bergamottin and erlotinib reveal active-site modifications for binding of diverse ligands. *J Biol Chem* 293:19201–19210. <https://doi.org/10.1074/jbc.RA118.005588>
- Blair RM, Fang H, Branham WS et al (2000) The estrogen receptor relative binding affinities of 188 natural and xenochemicals: structural diversity of ligands. *Toxicol Sci* 54:138–153. <https://doi.org/10.1093/toxsci/54.1.138>
- Bridges A, Gruenke L, Chang Y-T et al (1998) Identification of the binding site on cytochrome P450 2B4 for cytochrome b 5 and cytochrome P450 reductase. *J Biol Chem* 273:17036–17049. <https://doi.org/10.1074/jbc.273.27.17036>
- Burke MD, Thompson S, Weaver RJ et al (1994) Cytochrome P450 specificities of alkoxyresorufin O-dealkylation in human and rat liver. *Biochem Pharmacol* 48:923–936. [https://doi.org/10.1016/0006-2952\(94\)90363-8](https://doi.org/10.1016/0006-2952(94)90363-8)
- Campbell SJ, Henderson CJ, Anthony DC et al (2005) The murine Cyp1a1 gene is expressed in a restricted spatial and temporal pattern during embryonic development. *J Biol Chem* 280:5828–5835. <https://doi.org/10.1074/jbc.M412899200>
- Choudhary D, Jansson I, Schenkman JB et al (2003) Comparative expression profiling of 40 mouse cytochrome P450 genes in embryonic and adult tissues. *Arch Biochem Biophys* 414:91–100. [https://doi.org/10.1016/S0003-9861\(03\)00174-7](https://doi.org/10.1016/S0003-9861(03)00174-7)
- Choudhary D, Jansson I, Stoilov I et al (2005) Expression patterns of mouse and human CYP orthologs (families 1–4) during development and in different adult tissues. *Arch Biochem Biophys* 436:50–61. <https://doi.org/10.1016/j.abb.2005.02.001>
- Cimercancic P, Weinkam P, Rettenmaier TJ et al (2016) CryptoSite: expanding the druggable proteome by characterization and prediction of cryptic binding sites. *J Mol Biol* 428:709–719. <https://doi.org/10.1016/j.jmb.2016.01.029>
- Denisov IG, Makris TM, Sligar SG, Schlichting I (2005) Structure and chemistry of cytochrome P450. *Chem Rev* 105:2253–2278. <https://doi.org/10.1021/cr0307143>
- Dragin N, Shi Z, Madan R et al (2008) Phenotype of the Cyp1a1/1a2/1b1(-/-) triple-knockout mouse. *Mol Pharmacol* 73:1844–1856. <https://doi.org/10.1124/mol.108.045658>
- Ekroos M, Sjogren T (2006) Structural basis for ligand promiscuity in cytochrome P450 3A4. *Proc Natl Acad Sci* 103:13682–13687. <https://doi.org/10.1073/pnas.0603236103>
- El-Sherbeni AA, El-Kadi AOS (2016) Repurposing resveratrol and fluconazole to modulate human cytochrome P450-mediated arachidonic acid metabolism. *Mol Pharm* 13:1278–1288. <https://doi.org/10.1021/acs.molpharmaceut.5b00873>
- Eugster HP, Probst M, Würzler FE, Sengstag C (1993) Caffeine, estradiol, and progesterone interact with human CYP1A1 and CYP1A2. Evidence from cDNA-directed expression in *Saccharomyces cerevisiae*. *Drug Metab Dispos* 21:43–49
- Fan H, Irwin JJ, Webb BM et al (2009) Molecular Docking screens using comparative models of proteins. *J Chem Inf Model* 49:2512–2527. <https://doi.org/10.1021/ci9003706>
- Furuya H, Shimizu T, Hirano K et al (1989) Site-directed mutageneses of rat liver cytochrome P-450d: catalytic activities toward benzphetamine and 7-ethoxycoumarin. *Biochemistry* 28:6848–6857. <https://doi.org/10.1021/bi00443a011>
- Grover AK (2013) Use of allosteric targets in the discovery of safer drugs. *Med Princ Pract* 22:418–426. <https://doi.org/10.1159/000350417>
- Gurpide E, Schwers J, Welch MT et al (1966) Fetal and maternal metabolism of estradiol during pregnancy. *J Clin Endocrinol Metab* 26:1355–1365. <https://doi.org/10.1210/jcem-26-12-1355>
- Hakkola J, Pasanen M, Hukkanen J et al (1996) Expression of xenobiotic-metabolizing cytochrome P450 Forms in human full-term placenta. *Biochem Pharmacol* 51:403–411. [https://doi.org/10.1016/0006-2952\(95\)02184-1](https://doi.org/10.1016/0006-2952(95)02184-1)
- Hakkola J, Pelkonen O, Pasanen M, Raunio H (1998) Xenobiotic-metabolizing cytochrome P450 enzymes in the human fetoplacental unit: role in intrauterine toxicity. *Crit Rev Toxicol* 28:35–72. <https://doi.org/10.1080/10408449891344173>
- Hardy JA, Wells JA (2004) Searching for new allosteric sites in enzymes. *Curr Opin Struct Biol* 14:706–715. <https://doi.org/10.1016/j.sbi.2004.10.009>
- Hayashi S, Omata Y, Sakamoto H et al (2003) Purification and characterization of a soluble form of rat liver NADPH-cytochrome P-450 reductase highly expressed in *Escherichia coli*. *Protein Expr Purif* 29:1–7. [https://doi.org/10.1016/S1046-5928\(03\)00023-8](https://doi.org/10.1016/S1046-5928(03)00023-8)
- Hill D, Reasor M (1991) Effects of amiodarone administration during pregnancy in Fischer 344 rats. *Toxicology* 65:259–269. [https://doi.org/10.1016/0300-483X\(91\)90085-F](https://doi.org/10.1016/0300-483X(91)90085-F)
- Hlavica P, Schulze J, Lewis DFV (2003) Functional interaction of cytochrome P450 with its redox partners: a critical assessment and update of the topology of predicted contact regions. *J Inorg Biochem* 96:279–297. [https://doi.org/10.1016/S0162-0134\(03\)00152-1](https://doi.org/10.1016/S0162-0134(03)00152-1)

- Holinka CF, Diczfalussy E, Coelingh Bennink HJT (2008) Estetrol: a unique steroid in human pregnancy. *J Steroid Biochem Mol Biol* 110:138–143. <https://doi.org/10.1016/j.jsbmb.2008.03.027>
- Huang N, Pandey AV, Agrawal V et al (2005) Diversity and function of mutations in P450 oxidoreductase in patients with Antley-Bixler syndrome and disordered steroidogenesis. *Am J Hum Genet* 76:729–749. <https://doi.org/10.1086/429417>
- Im S-C, Waskell L (2011) The interaction of microsomal cytochrome P450 2B4 with its redox partners, cytochrome P450 reductase and cytochrome b5. *Arch Biochem Biophys* 507:144–153. <https://doi.org/10.1016/j.abb.2010.10.023>
- Jacobson MP, Kaminski GA, Friesner RA, Rapp CS (2002) Force field validation using protein side chain prediction. *J Phys Chem B* 106:11673–11680. <https://doi.org/10.1021/jp021564n>
- Jaladanki CK, He Y, Zhao LN et al (2021) Virtual screening of potentially endocrine-disrupting chemicals against nuclear receptors and its application to identify PPAR γ -bound fatty acids. *Arch Toxicol* 95:355–374. <https://doi.org/10.1007/s00204-020-02897-x>
- Jeong HG, Kimand JY, Choi CY (2000) Down-Regulation of murine Cyp1a-1 in mouse hepatoma hepa-1c1c7 cells by bisphenol A. *Biochem Biophys Res Commun* 277:594–598. <https://doi.org/10.1006/bbrc.2000.3717>
- Joshi P, McCann GJP, Sonawane VR et al (2017) Identification of potent and selective CYP1A1 inhibitors via combined ligand and structure-based virtual screening and their in vitro validation in saccharosomes and live human cells. *J Chem Inf Model* 57:1309–1320. <https://doi.org/10.1021/acs.jcim.7b00095>
- Kandel SE, Lampe JN (2014) Role of protein-protein interactions in cytochrome P450-mediated drug metabolism and toxicity. *Chem Res Toxicol* 27:1474–1486. <https://doi.org/10.1021/tx500203s>
- Katoh K, Misawa K, Kuma K, Miyata T (2002) MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res* 30:3059–3066. <https://doi.org/10.1093/nar/gkf436>
- Kavlock R, Chandler K, Houck K et al (2012) Update on EPA's ToxCast Program: providing high throughput decision support tools for chemical risk management. *Chem Res Toxicol* 25:1287–1302. <https://doi.org/10.1021/tx3000939>
- Kenaan C, Zhang H, Shea EV, Hollenberg PF (2011) Uncovering the role of hydrophobic residues in cytochrome P450–cytochrome P450 reductase interactions. *Biochemistry* 50:3957–3967. <https://doi.org/10.1021/bi1020748>
- Klinger W, Lupp A, Karge E et al (2002) Estradiol, testosterone, dehydroepiandrosterone and androstenedione: novel derivatives and enantiomers. Interactions with rat liver microsomal cytochrome P450 and antioxidant/radical scavenger activities in vitro. *Toxicol Lett* 128:129–144. [https://doi.org/10.1016/S0378-4274\(01\)00540-9](https://doi.org/10.1016/S0378-4274(01)00540-9)
- Lee AJ, Cai MX, Thomas PE et al (2003) Characterization of the oxidative metabolites of 17 β -estradiol and estrone formed by 15 selectively expressed human cytochrome P450 isoforms. *Endocrinology* 144:3382–3398. <https://doi.org/10.1210/en.2003-0192>
- Lee J-YJ, Miller JA, Basu S et al (2018) Building predictive in vitro pulmonary toxicity assays using high-throughput imaging and artificial intelligence. *Arch Toxicol* 92:2055–2075. <https://doi.org/10.1007/s00204-018-2213-0>
- Ma X, Idle JR, Krausz KW, Gonzalez FJ (2005) Metabolism of melatonin by human cytochromes P450. *Drug Metab Dispos* 33:489–494. <https://doi.org/10.1124/dmd.104.002410>
- Mayuzumi H, Sambongi C, Hiroya K et al (1993) Effect of mutations of ionic amino acids of cytochrome P450 1A2 on catalytic activities toward 7-ethoxycoumarin and methanol. *Biochemistry* 32:5622–5628. <https://doi.org/10.1021/bi00072a018>
- Modell AE, Blosser SL, Arora PS (2016) Systematic targeting of protein-protein interactions. *Trends Pharmacol Sci* 37:702–713. <https://doi.org/10.1016/j.tips.2016.05.008>
- Mysinger MM, Shoichet BK (2010) Rapid context-dependent ligand desolvation in molecular docking. *J Chem Inf Model* 50:1561–1573. <https://doi.org/10.1021/ci100214a>
- Nishimura M, Yaguti H, Yoshitsugu H et al (2003) Tissue distribution of mRNA expression of human cytochrome P450 isoforms assessed by high-sensitivity real-time reverse transcription PCR. *Yakugaku Zasshi* 123:369–375. <https://doi.org/10.1248/yakushi.123.369>
- Niwa T, Murayama N, Imagawa Y, Yamazaki H (2015) Regioselective hydroxylation of steroid hormones by human cytochromes P450. *Drug Metab Rev* 47:89–110. <https://doi.org/10.3109/03602532.2015.1011658>
- Ohyama K, Nakajima M, Suzuki M et al (2000) Inhibitory effects of amiodarone and its N-deethylated metabolite on human cytochrome P450 activities: prediction of in vivo drug interactions. *Br J Clin Pharmacol* 49:244–253. <https://doi.org/10.1046/j.1365-2125.2000.00134.x>
- Omicinski CJ, Redlich CA, Costa P (1990) Induction and developmental expression of cytochrome P450IA1 messenger RNA in rat and human tissues: detection by the polymerase chain reaction. *Cancer Res* 50:4315–4321
- Otte JC, Schmidt AD, Hollert H, Braunbeck T (2010) Spatio-temporal development of CYP1 activity in early life-stages of zebrafish (*Danio rerio*). *Aquat Toxicol* 100:38–50. <https://doi.org/10.1016/j.aquatox.2010.07.006>
- Ozaki H, Sugihara K, Watanabe Y et al (2016) Cytochrome P450-inhibitory activity of parabens and phthalates used in consumer products. *J Toxicol Sci* 41:551–560. <https://doi.org/10.2131/jts.41.551>
- Paine MF, Schmiedlin-Ren P, Watkins PB (1999) Cytochrome P-450 1A1 expression in human small bowel: interindividual variation and inhibition by ketoconazole. *Drug Metab Dispos* 27:360–364
- Pilmis B, Jullien V, Sobel J et al (2015) Antifungal drugs during pregnancy: an updated review. *J Antimicrob Chemother* 70:14–22. <https://doi.org/10.1093/jac/dku355>
- Santes-Palacios R, Ornelas-Ayala D, Cabañas N et al (2016) Regulation of human cytochrome P4501A1 (hCYP1A1): a plausible target for chemoprevention? *BioMed Res Int* 2016:e5341081. <https://doi.org/10.1155/2016/5341081>
- Schwarz D, Kisselev P, Schunck W-H et al (2000) Allelic variants of human cytochrome P450 1A1 (CYP1A1): effect of T461N and I462V substitutions on steroid hydroxylase specificity. *Pharmacogenet Genomics* 10:519
- Schwarz D, Kisselev P, Ericksen SS et al (2004) Arachidonic and eicosapentaenoic acid metabolism by human CYP1A1: highly stereoselective formation of 17(R),18(S)-epoxyeicosatetraenoic acid. *Biochem Pharmacol* 67:1445–1457. <https://doi.org/10.1016/j.bcp.2003.12.023>
- Sevrioukova IF, Li H, Zhang H et al (1999) Structure of a cytochrome P450–redox partner electron-transfer complex. *Proc Natl Acad Sci* 96:1863–1868. <https://doi.org/10.1073/pnas.96.5.1863>
- Shen S, Strobel HW (1992) The role of cytochrome P450 lysine residues in the interaction between cytochrome P450IA1 and NADPH-cytochrome P450 reductase. *Arch Biochem Biophys* 294:83–90. [https://doi.org/10.1016/0003-9861\(92\)90140-R](https://doi.org/10.1016/0003-9861(92)90140-R)
- Shen AL, O'Leary KA, Kasper CB (2002) Association of multiple developmental defects and embryonic lethality with loss of microsomal NADPH-cytochrome P450 oxidoreductase. *J Biol Chem* 277:6536–6541. <https://doi.org/10.1074/jbc.M111408200>
- Shimizu T, Tateishi T, Hatano M, Fujii-Kuriyama Y (1991) Probing the role of lysines and arginines in the catalytic function of cytochrome P450d by site-directed mutagenesis. Interaction with NADPH-cytochrome P450 reductase. *J Biol Chem* 266:3372–3375
- Sridhar J, Liu J, Foroozesh M, Klein Stevens CL (2012) Inhibition of cytochrome P450 enzymes by quinones and anthraquinones.

- Chem Res Toxicol 25:357–365. <https://doi.org/10.1021/tx2004163>
- Stresser DM, Turner SD, Blanchard AP et al (2002) Cytochrome P450 fluorometric substrates: identification of isoform-selective probes for rat CYP2D2 and human CYP3A4. Drug Metab Dispos 30:845–852. <https://doi.org/10.1124/dmd.30.7.845>
- Tamura H, Nakamura Y, Terron MP et al (2008) Melatonin and pregnancy in the human. Reprod Toxicol 25:291–303. <https://doi.org/10.1016/j.reprotox.2008.03.005>
- The UniProt Consortium (2021) UniProt: the universal protein knowledgebase in 2021. Nucleic Acids Res 49:D480–D489. <https://doi.org/10.1093/nar/gkaa1100>
- Thompson AD, Dugan A, Gestwicki JE, Mapp AK (2012) Fine-tuning multiprotein complexes using small molecules. ACS Chem Biol 7:1311–1320. <https://doi.org/10.1021/cb300255p>
- Walsh AA, Szklarz GD, Scott EE (2013) Human cytochrome P450 1A1 structure and utility in understanding drug and xenobiotic metabolism. J Biol Chem 288:12932–12943. <https://doi.org/10.1074/jbc.M113.452953>
- Wassenberg DM, Di Giulio RT (2004) Synergistic embryotoxicity of polycyclic aromatic hydrocarbon aryl hydrocarbon receptor agonists with cytochrome P4501A inhibitors in *Fundulus heteroclitus*. Environ Health Perspect 112:1658–1664. <https://doi.org/10.1289/ehp.7168>
- Widerhorn J, Bhandari AK, Bughi S et al (1991) Fetal and neonatal adverse effects profile of amiodarone treatment during pregnancy. Am Heart J 122:1162–1166. [https://doi.org/10.1016/0002-8703\(91\)90489-5](https://doi.org/10.1016/0002-8703(91)90489-5)
- Yanes O, Clark J, Wong DM et al (2010) Metabolic oxidation regulates embryonic stem cell differentiation. Nat Chem Biol 6:411–417. <https://doi.org/10.1038/nchembio.364>
- Yin J, Yang J, Zhang F et al (2014) Individual and joint toxic effects of cadmium sulfate and α -naphthoflavone on the development of zebrafish embryo. J Zhejiang Univ Sci B 15:766–775. <https://doi.org/10.1631/jzus.B1400091>
- Yu Z, Hu D, Li Y (2004) Effects of zearalenone on mRNA expression and activity of cytochrome P450 1A1 and 1B1 in MCF-7 cells. Ecotoxicol Environ Saf 58:187–193. <https://doi.org/10.1016/j.ecoenv.2004.03.001>
- Zhang Z, Li Y, Lin B et al (2011) Identification of cavities on protein surface using multiple computational approaches for drug binding site prediction. Bioinformatics 27:2083–2088. <https://doi.org/10.1093/bioinformatics/btr331>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.