Electronic Supplementary Material (ESI) for Medicinal Chemistry Communications

Supplementary Information

Experimental Section

Materials

All chemicals and reagents were obtained from standard chemical providers. 7-ethoxyresorufin (ER) was synthesized by the method of Burke et al.[1] and compound was confirmed by high-performance liquid chromatography (HPLC), mass spectrometry and ^{1}H NMR. The plasmid BMX100/h1A2 with CYP NADPH reductase was kindly provided by Dr. Michel Kranendonk (Lisbon, Portugal). α-naphthoflavone, DMSO, buffer (potassium phosphates, pH 7.4) and resorufin (standard) were purchased from Sigma-Alrich. NADPH was purchased from AppliChem GmbH, Germany. Screening compounds (Chembridge compound ID: 6655806, 5135529, 5353195, 6041943,5355944, 5539903, 6838304, 8902536, 5317944, 5306674, 5142560, 5308300, 5160525, 5141055, 5210442, 5523686, 5872529, 5142736, 6145131, 5211920, 5186801, 6239040, 5746436, 7641952, 5186503, 5255053, 5791580, 5770740, 5146691, 5161057, 6087958, 5737694, 5137323, 5661457, 5577002, 5401337, 6564528, 7907769, 5161996, 5937099 and 5548700) were purchased from ChemBridge Corporation, San Diego, CA, USA.[2] The purity of the compounds was at least 90 %.

CYP1A2 Preparation and Inhibition Assay

A set of 1 mg of 41 screening compounds was purchased from ChemBridge chemical store. Stocks with a concentration of 5 mM of screening compounds and ER in 100% DMSO were prepared.

Recombinant cDNA of human CYP1A2 plasmid was initially transformed into *E. coli* strain JM109. Human CYP1A2 was expressed into the media (in a 3 liter flask) containing 300 mL terrific broth (TB) with 1mM δ-aminolevulinic acid, 0.5 mM thiamine, 400 µL/L trace elements,[3] 100 µg/mL ampicillin and 1 mM isopropyl-β-D-thiogalactopyranoside (IPTG). The media was inoculated with 3 mL overnight culture of bacteria containing CYP1A2 plasmid. The cell culture was incubated for about 40 hours at 28^{o}C and 125 rpm and CYP1A2 content was determined according to the Omura and Sato procedure.[4] The cell was pelleted by centrifugation and resuspended in 15 mL KPi-Glycerol buffer (100 mM KPi (pH 7.5), 20% Glycerol, 0.25 mM EDTA and 0.1 mM DTT). The cell was treated with 0.5 mg/mL lysozyme prior to disruption by French press. Next, the material was sonicated (2 times 2 min) and centrifuged for 15 minutes at 3000 x g at 4˚C to spin down remaining whole cells. From the supernatant, cell membranes containing human CYP1A2 were isolated by ultracentrifugation (160.000 x g, 1 hour at 4˚C) and resuspended in 4 ml of KPi-Glycerol buffer for use.

Initial experiments were performed in a *Perkin Elmer 3000 spectrofluorometer.* In 1 mL of the cuvette, 10 µL of 7-ethoxyresorufin (10000 nM to 19.53 nM), 20 µL of protein (15 nM) and 870 µL of buffer (100 mM) were added to record the fluorescence at an excitation wavelength of 530 nm and an emission wavelength of 582 nm. The reaction was initiated by adding 100 µL of NADPH (0.2 mM) and the enzymatic reaction was monitored for 3-5 minutes. V_{max} and K_M were calculated from the Michaelis-Menten equation by measuring the resorufin formation (V_{max} and K_M were 2.94 nmol/min/nmol and 1.65 µM respectively). The EROD assay has been used to check the catalytic activity of EROD and fluorescence was calibrated by using a resorufin standard (250 nM). Using the same procedure, with a fixed EROD concentration of 1.65 µM, and 30, 3 or 0.3 µM of the screening compounds, the percentage inhibition at these concentrations were measured. Initially, the relevant concentration range was estimated by measuring inhibition at different concentrations of compound **O** (randomly selected, see Figure S1). Compound **O** showed 40% inhibition at 20 µM and therefore the starting point for the experimental testings was set at 30 µM. This is in line with the highest concentration s in the CYP1A2 inhibitor screening in PubChem.

The IC_{50} values of compounds D, H, J, Q, AB, AD, AH, A, AM and α-naphthoflavone were determined as previously described.[5] In short, black coaster 96-well plates were prepared containing final concentrations of 1.5 µM EROD, 15 nM protein, 0.8 % DMSO and varying amount of screening compound (range 0.01 nM to 20 µM). At 37°, reactions were started by adding NADPH (final concentration 0.2 mM and monitoring the linear increase in fluorescence for 10 min on a Victor2 1420 multilabel counter (excitation at 530 nm, emission at 586 nm). Reaction mixtures had a total volume of 200 µL. All measurements were performed in triplo. 100 % activity was estimated from 9 control measurements in the absence of inhibitor.

For data analysis, Origin 8.0[6] and GraphPad Prism software (version 4.0)[7] were used for determination of IC_{50}, V_{max}, and K_M values.

Database Processing

A subset of the ZINC database[8] consisting of 19997 entries from 16338 compounds was obtained by using two filters. First, a less restrictive filtering than the Lipiniski's "Rule of Five"[9] (number of H-bond donors ≤ 5; number of H-bond acceptors ≤ 7; xLogP ≤ 6;

Mw: 190–700). Then, to reduce the number of compounds further, our previously developed decision tree model for CYP1A2 ligand-like features was used as well.[10]

Random forest calculations

The random forest method for prediction of CYP1A2 inhibition was applied as described by Vasanthanathan *et al*. [10] The model is based on 39 descriptors calculated using MOE and Volsurf.[11] Weka was used to apply the random forest model.[12] If more than one entry exists per compound only that with the best random forest probability was kept. This resulted in a total of 16338 unique compounds.

Docking calculations

All compounds were docked into the CYP1A2 active site using the ChemScore scoring function[13] in the GOLD docking software (version 3.2).[14] The same settings as in Vasanthanathan et al.[6] were applied, i.e. excluding all water molecules. All 19997 entries in the compound database were docked, but only 19906 entries were successfully docked. If more than one entry exists per compound, only that with the best score was kept. This resulted in a total of 16330 unique compounds.

GRID

The GRID program (version 22)[15,16] was used to determine favorable interactions with the C3, Dry and O probes.

Multivariate data analysis

A principal component analysis was performed using Simca (version 11).[17]

Substructure search

A substructure search was done with Canvas.[18]

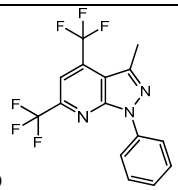
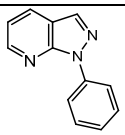
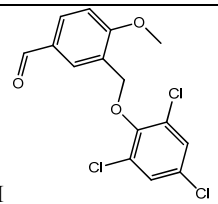
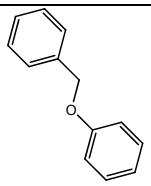
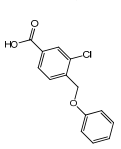
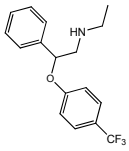
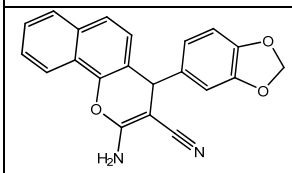
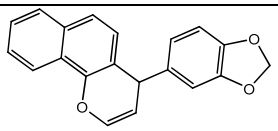
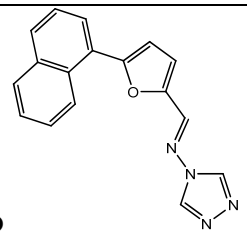
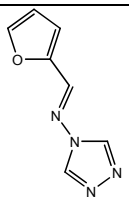
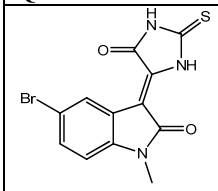
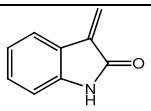
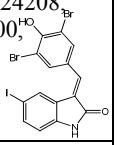
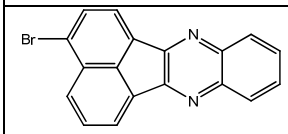
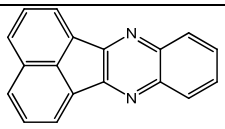
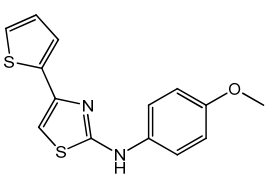
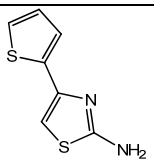
Table S1: ChemScore values, random forest probabilities, and % inhibition at different concentrations of the compounds tested experimentally.

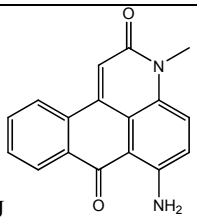
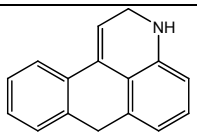
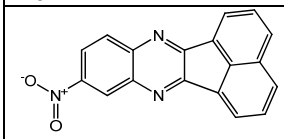
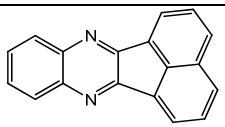
			% Inhibition		
Compound	ChemScore	Random forest	0.3uM	3uM	30uM
A	2.2	100	0.1	11.7	70.5
B	7.6	100	6.0	-16.1	54.3
C	29.6	100	-22.5	14.3	67.4
D	28.1	100	77.3	96.7	98.1
E	11.5	100	-4.9	32.8	77.3
F	25.5	100	18.4	89.8	97.9
G	21.4	100	-13.1	46.7	85.8
H	32.3	100	93.8	100.8	99.9
I	10.4	100	-2.9	49.5	87.9
J	14.9	100	75.4	97.0	98.3
K	29.7	100	24.2	45.9	77.3
L	27.7	100	38.4	62.8	91.0
M	20.0	100	32.5	61.1	79.8
N	13.2	100	44.1	79.9	80.6
O[a]	22.3	100	14.0	20.0	50.0
P	22.9	100	48.7	69.1	84.1
Q	34.1	100	83.2	101.5	100.3
R	28.2	100	59.0	90.5	97.4
S	31.9	100	52.2	71.4	86.6
T	32.1	60	71.5	90.3	98.5
U	14.4	10	38.5	75.2	84.7
V	16.5	60	16.6	68.2	89.0
W	21.0	40	18.2	58.6	90.9
X	30.0	60	-0.4	62.5	74.7
Y	25.5	40	-12.1	43.8	71.3

Z	41.2	70	4.6	64.4	98.1
AA	29.5	70	24.8	78.4	97.5
AB	29.1	40	90.7	92.2	97.3
AC	22.5	60	10.8	61.6	73.3
AD	46.0	60	99.5	100.2	105.0
AE	27.3	50	73.3	89.0	95.9
AF	12.6	30	26.6	0.4	79.3
AG	41.3	70	67.7	90.9	98.0
AH	29.6	60	94.9	99.1	99.6
AI	28.0	40	33.0	19.9	90.3
AJ	42.2	70	95.5	87.4	68.8
AK	19.6	30	35.0	27.4	93.5
AL	22.2	10	45.4	33.0	95.9
AM	40.8	80	92.1	90.3	100.7
AN	23.2	40	58.4	76.8	97.5
AO	37.5	30	67.5	76.7	92.2
αNF[b]	46.9	60	ND	ND	ND

[a] Interpolated from slightly different values. [b] *αNF* is *α-napthaflavone*

Table S2. Substructure search in PubChem (CYP1A2 bioassay database) for similar scaffolds as those identified in the most potent ligands (with IC_{50} < 250 nM, cf. Table 1).

Compound	**Scaffold**	**Search in CYP1A2 bioassay database in PubChem**
 **D**		No compounds identified with this scaffold
 **H**		18 compounds identified. PubChem ID: 7638, 3246678, 3246677, 3246675, 5459399, 54841, 4534086, 6603938, 1548968, 6093336, 419053, 65856, 6604909, 5145950,328505, 3246576, 308481, 419051, 308481.   **308481** Not CYP1A2 inhibitor **1548968(Fluoxetine)** CYP1A2 inhibitor
 **J**		No compounds identified with this scaffold
 **Q**		No compounds identified with this scaffold
 **AB**		6 compounds identified. PubChem ID: 3246667, 5924208, 5353593, 5312137, 5288600, 5800609 (see picture).  All are CYP1A2 inhibitors
 **AD**		No compounds identified with this scaffold
 **AH**		One compound identified (identical to scaffold) PubChem ID: 308481 Not a CYP1A2 inhibitor

 **AJ**		No compounds identified with this scaffold
 **AM**		No compounds identified with this scaffold

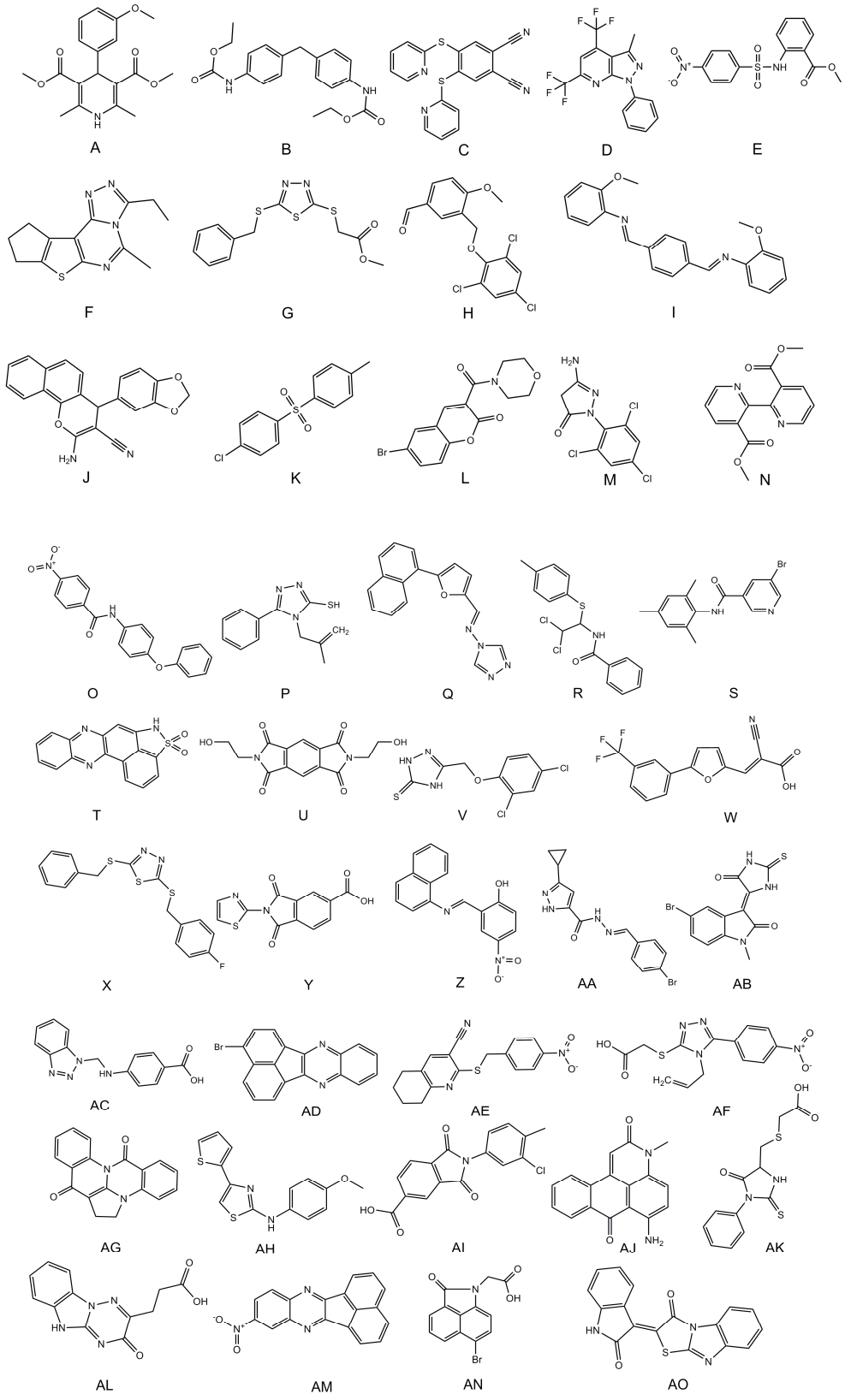


Figure S1. Compounds selected from virtual screening using the ligand-based random forest method (**A-S**) and the structure-based docking method (**T-AO**) that were tested experimentally for CYP1A2 inhibition.

Reference List

1. Burke, M. D.; Thompson, S.; Elcombe, C. R.; Halpert, J.; Haaparanta, T.; and Mayer, R. T. Ethoxy-, pentoxy- and benzyloxyphenoxazones and homologues: a series of substrates to distinguish between different induced cytochromes P-450. *Biochem. Pharmacol.* **1985,** *34*, 3337-3345.

2. www.hit2lead.com.

3. Bauer, S. and Shiloach, J. Maximal exponential growth rate and yield of E. coli obtainable in a bench-scale fermentor. *Biotechnol. Bioeng.* **1974,** *16*, 933-941.

4. Omura, T. and Sato, R. The Carbon Monoxide-binding Pigment of Liver Microsomes. *J. Biol. Chem.* **1964,** *239*, 2370-2378.

5. Appiah-Opong, R.; Commandeur, J. N. M.; van Vugt-Lussenburg, B.; and Vermeulen, N. P. E. Inhibition of human recombinant cytochrome P450s by curcumin and curcumin decomposition products. *Toxicology* **2007,** *235*, 83-91.

6. Vasanthanathan, P.; Hritz, J.; Taboureau, O.; Olsen, L.; Jørgensen, F. S.; Vermeulen, N. P.; and Oostenbrink, C. Virtual screening and prediction of site of metabolism for cytochrome P450 1A2 ligands. *J. Chem. Inf. Model.* **2009,** *49*, 43-52.

7. Swift, M. L. GraphPad prism, data analysis, and scientific graphing. *J. Chem. Inf. Comput. Sci.***1997,** *37*, 411-412.

8. Irwin, J. J. and Shoichet, B. K. ZINC - A free database of commercially available compounds for virtual screening. *J. Chem. Inf. Model.* **2005,** *45*, 177-182.

9. Lipinski, C. A.; Lombardo, F.; Dominy, B. W.; and Feeney, P. J. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv. Drug Deliv. Rev.* **2001,** *46*, 3-26.

10. Vasanthanathan, P.; Taboureau, O.; Oostenbrink, C.; Vermeulen, N. P.; Olsen, L.; and Jørgensen, F. S. Classification of cytochrome P450 1A2 inhibitors and noninhibitors by machine learning techniques. *Drug Metab. Dispos.* **2009,** *37*, 658-664.

11. Volsurf version 4.1.4.1 Molecular Discovery Ltd., Middlesex, United Kingdom.

12. Weka, version 3.2, Waikato Environment for Knowledge Analysis; University of Waikato, New Zealand.

13. Eldridge, M. D.; Murray, C. W.; Auton, T. R.; Paolini, G. V.; and Mee, R. P. Empirical scoring functions .1. The development of a fast empirical scoring function to estimate the binding affinity of ligands in receptor complexes. *J. Comput. Aided Mol. Des.* **1997,** *11*, 425-445.

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

15. GRID, version 22, Molecular Discovery Ltd., Middlesex, United Kingdom.

16. Goodford, P. J. A computational procedure for determining energetically favorable binding sites on biologically important macromolecules. *J. Med. Chem.* **1985,** *28*, 849-857.

17. SIMCA-P (v11.0.0.0) Umetrics, Box 7960, S-90719 Umeå, Sweden.

18. Canvas, version 1.3, Schrödinger, LLC, New York, NY, 2010