Novel QSAR Models for Prediction of Reversible and Time-Dependent Inhibition of Cytochrome P450 Enzymes

S. Faramarzi[1], X. Yang[1], A. Bassan[2], K. P. Kross[2], G.J. Myatt[2], D.A. Volpe[1], L. Stavitskaya[1]

[1]Office of Clinical Pharmacology, Center for Drug Evaluation and Research, Food and Drug Administration, Silver Spring, Maryland, USA; [2]Instem Inc., Columbus, Ohio, USA

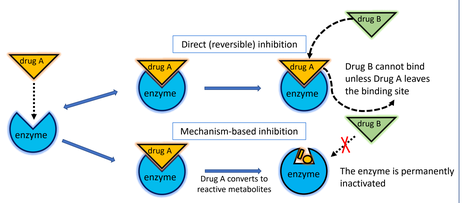
Abstract

FDA's *in vitro* drug interaction guidance[1] states that *in vivo* drug-drug interaction (DDIs) caused by metabolites may be possible even if *in vitro* studies suggest that the parent drug alone does not inhibit any major cytochrome P450 (CYP) enzymes. As a result, the guidance recommends that sponsors evaluate certain metabolites *in vitro* for their inhibitory effects on a panel of CYP enzymes. Considerations for testing metabolites include inhibitory potency of metabolite(s), *in vivo* metabolite-to-parent exposure , and the presence of a possible mechanism-based inhibition (MBI) alert, since such inhibition leads to a higher risk of DDI due to prolonged inhibition effect compared to reversible inhibition.

To facilitate identification of metabolites that are most likely to be inhibitors, five quantitative structure-activity relationship (QSAR) models for reversible inhibition of CYP3A4, 2C9, 2C19 and 2D6, and MBI of CYP3A4 were developed. The non-proprietary training sets for the models were harvested from FDA drug approval packages and published literature to give a total of 10,286 chemical structures. The cross-validation performance statistics for the models range from 79% to 83% in concordance and 77% to 83% negative predictivity. Additionally, the performance of the newly developed models was assessed using external validation sets. Overall performance statistics showed up to 85% in concordance and up to 97% in negative predictivity. The newly developed models will provide a faster and more effective evaluation of potential DDIs caused by metabolites.

Introduction

- CYP enzymes are a family of heme containing enzymes that catalyze the oxidative metabolism of endogenous and exogenous compounds. Drugs may inhibit CYP enzymes, and therefore alter the metabolism of co-administered drugs. This phenomenon makes up the majority of pharmacokinetic DDI.



- There are at least 57 CYP enzymes in human body, among which 12 are involved in drug metabolism[2]. Drugs are routinely evaluated as substrates and inhibitors of CYP1A2, 2B6, 2C8, 2C9, 2C19, 2D6, and 3A[1].
- 52% of small molecule drugs approved by the FDA from 2015-2020 are metabolized by CYP3A4, making it the major CYP subtype[3]. Other important CYP subtypes include CYP2C9, CYP2C19, and CYP2D6 which are involved in metabolism of drugs containing polar acidic groups, hydrogen bond donors/acceptors, and amines, respectively[4].
- QSAR models provide a rapid assessment of CYP inhibition of a molecule based solely on its chemical structure. These models can identify structural alerts and mitigating features in drugs.
- In a regulatory environment, high sensitivity and negative predictivity are important characteristics of QSAR models used to support drug safety decisions, thereby minimizing risk to patients.

Objectives

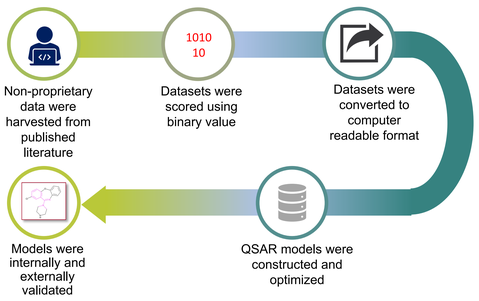
Develop inhibition datasets for major CYP enzymes

Construct and optimize QSAR models

Assess the performance of models

Materials and Methods

Construction of the Models



Data Collection

Experimental CYP inhibition data for 10,287 unique chemicals were collected from publicly available literature and databases (Table 1).

Experimental data consisted of studies conducted using human liver microsomes or recombinant CYP enzymes.

Table 1. CYP enzyme inhibition database

Enzyme type	Inhibition type	Number of compounds	Number of positives	Number of negatives
CYP2D6	direct	4163	1697	2466
CYP3A4	direct	7044	2668	4376
CYP2C9	direct	4065	1387	2678
CYP2C19	direct	2847	789	2058
CYP3A4	MBI	623	306	317

To determine if a compound is a direct or mechanism-based inhibitor of CYP enzymes, criteria listed in Table 2 was used[5].

The predictive performance of the CYP models was assessed using proprietary external validation set harvested from drugs reviews (Table 3).

Table 2. Thresholds for identification of direct and mechanism-based inhibitors

Inhibition type	Parameters and thresholds			
Direct		IC_{50} (μM)	K_i (μM)	R_1*
		10	5	1.02
Mechanism based	IC_{50} fold shift	Change in inhibition (%)	k_{obs} (min^{-1})	R_2*
	1.5	20	0.01	1.25

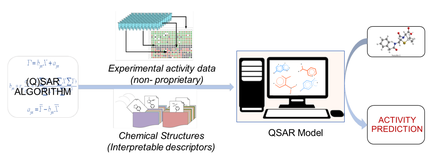
*From the FDA guidance[1].

Table 3. External validation sets composed of drugs approved in 2006-2021.

Enzyme type	Inhibition type	Number of compounds	Number of positives	Number of negatives
2D6	direct	221	24	197
3A4	direct	221	43	178
2C9	direct	224	28	196
2C19	direct	222	19	203
3A4	MBI	109	52	57

QSAR Modeling

- Leadscope (LS) Enterprise v4.9.3-2 (Instem Inc., USA) was used for model building and testing.



Results and Discussion

Evaluation of Molecular Properties

Table 4 shows the functional groups present in the training sets with corresponding activities. Degrees of red and blue shading indicates the extent of the positive and negative chemicals, respectively. Ethers and amines are among mechanism-based inhibition alerts. Carboxylic acids and carbonyls are among the groups with lowest activities for all enzymes.

Table 4. Functional groups analysis for each training set. Numbers are activities (z-scores)

									Carbonyls																
Compound Features	alcohol	aldehyde	alkene	alkyne	amidine	amines	azide	carbamate	carbonyl	e	carboxylate	carboxylic acid	ether	halide	hydrazine	hydroxylamine	imine	iminomethyl	ketone	nitrile	quinones	sulfide	sulfonamide	sulfone	sulfoxide	urea
CYP3A4	4.673	-1.657	1.259	-0.5551	-2.915	5.568	1.027	9.7	0.22	3.645	1.812	-8.266	-3.206	-0.72	4.486	-2.023	3.258	-1.178	-4.696	-1.171	0.2793	2.299	2.656	-2.057	-2.223	4.487
CYP3A4 MBI	0.6763		-2.592	-0.8933	-1.348	1.111		0.5651	-1.025	1.056	1.124	-3.735	-3.305	2.279	0.2916	-0.5531	1.76	-1.762	-1.122	0.0924	-1.604	-0.185	0.7485	-0.8378	-1.71	-2.675
CYP2C9	-1.714	-1.019	0.5386	-1.888	-0.4591	-4.956	0.6673	-1.543	-5.244	-1.749	-2.953	-2.057	1.877	2.274	-3.064	0.9572	1.215	-0.9967	-2.176	1.688	-1.019	-0.5739	5.121	5.445	-1.431	-1.756
CYP2C19	-2.273	-1.076	2.023	-1.003	0.6657	-0.6112		0.8963	-7.442	-2.295	-0.02369	-3.314	0.04706	-1.89	-2.443	0.2077	0.062	-1.207	-3.293	-0.1096	1.859	4.388	-0.701	-2.874	-2.715	-3.454
CYP2D6	5.566	-1.377	-3.444	-3.224	-2.803	10.12	-0.9736	-1.175	-16.24	-7.884	-2.954	-11.59	-0.9727	-2.91	-4.197	-1.804	0.6502	-3.126	-2.019	-0.9496	-1.377	3.191	-4.559	-3.354	-1.54	1.906

Figure 1 shows physicochemical properties present in direct CYP inhibition training sets

Panel A, shows that CYP3A4 is inhibited by large molecules when compared to other CYP enzymes.

Panel B shows that CYP2D6 inhibitors are on average less polar than the rest of the enzymes.

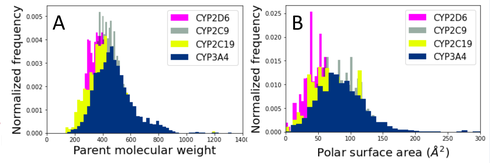


Figure 1. Physicochemical descriptor analysis for direct CYP inhibition training sets

Identification of Structural Alerts Responsible for CYP Inhibition

Structural features with highest direct inhibition activities were identified and evaluated for each average model (Table 5).

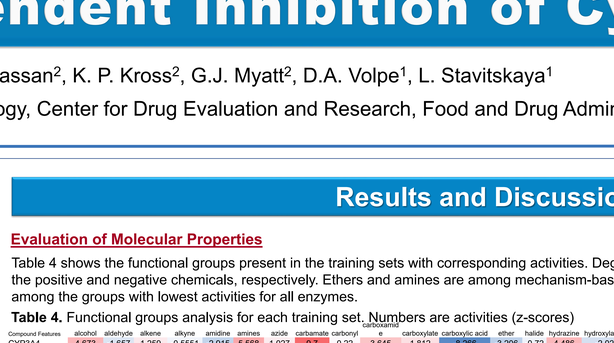
Overall, planar aromatic rings and secondary amines are among top direct inhibitory alerts for all CYP enzymes.

Figure 2 shows representative compounds containing alerts with high and low activity scores and their frequency in CYP3A4 MBI training set.

Features consist of newly identified alerts as well as chemical moieties which are known to be associated with CYP inhibition. For example, cyclopropylamines are reported to cause MBI of CYP enzymes through heme alkylation[6].

Carboxylic acids are already oxidized and therefore have the lowest MBI activity.

Table 5. Examples of structural features with highest direct inhibition activities for each model.



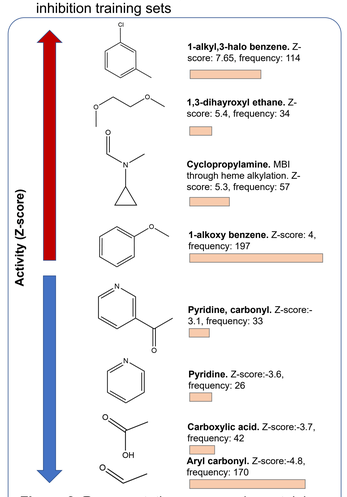


Figure 2. Representative compounds containing alerts with high and low activity scores and their frequency (beige bars) in CYP3A4 MBI training set.

Model Performances

Table 6 shows the predictive performance of newly constructed models using a 10 x 10% leave-many-out (LMO) method and external validation.

Sensitivity and negative predictivity (bolded in Table 6) are critical parameters for the safety assessment of drug products under review by FDA/CDER.

Unlike training sets, the external validation sets are almost exclusively made of approved drugs. The difference in chemical space between the two databases may explain relatively lower sensitivities for external validations.

Table 6. Cross-validation and external validation performance of newly constructed QSAR models with Leadscope.

Model	Cross-validation					External validation				
	CYP3A4	CYP3A4 (MBI)	CYP2C9	CYP2C19	CYP2D6	CYP3A4	CYP3A4 (MBI)	CYP2C9	CYP2C19	CYP2D6
Sensitivity (%)	**76**	**80**	**81**	**83**	**83**	**57**	**59**	**53**	**67**	**61**
Specificity (%)	81	82	79	84	82	72	81	88	76	87
Pos Pred* (%)	71 (80)	81 (81)	65 (78)	61 (80)	69 (82)	32 (66)	59 (83)	38 (81)	15 (65)	37 (83)
Neg Pred*(%)	**85 (77)**	**81 (80)**	**89 (81)**	**93 (83)**	**91 (83)**	**88 (63)**	**81 (57)**	**93 (66)**	**97 (77)**	**95 (70)**
Concordance (%)	79	81	80	81	83	69	69	84	75	85
Coverage (%)	80	84	84	86	87	72	59	70	69	76
Chi-squared	1775	197	1116	820	1705	9	11	20	8	26
MCC coefficient	0.56	0.61	0.57	0.58	0.63	0.24	0.41	0.36	0.23	0.39

*Numbers in parenthesis show normalized predictivities.

Conclusion

- Publicly available experimental data for 10,287 unique chemicals were collected and curated to serve as training sets for CYP inhibition models.
- Secondary amines are among top direct inhibitory alerts for all CYP enzymes.
- Ethers and heme alkylating groups show the highest mechanism-based inhibition.
- The overall cross-validation performance statistics for the models range from 76% to 83% in sensitivity and 77% to 83% in negative predictivity.
- Despite the small sample size, up to 85% accuracy was achieved in the external validation.
- These new models may facilitate the identification of structural alerts for direct and mechanism-based inhibition.

Acknowledgement

- Leadscope Enterprise was used by FDA / CDER under Research Collaboration Agreements with Instem Inc.
- This project was also supported in part by an appointment to the Oak Ridge Institute for Science and Education (ORISE) Research Participation Program at CDER administered through an agreement between the U.S. Department of Energy and CDER.

Disclaimer

This poster reflects the views of the authors and should not be construed to represent FDA's views or policies. The mention of commercial products, their sources, or their use in connection with material reported herein is not to be construed as either an actual or implied endorsement of such products by the Department of Health and Human Services.

References

1. Food & Drug Administration. 2020. In Vitro Drug Interaction Studies - Cytochrome P450 Enzyme- and Transporter-Mediated Drug Interactions Guidance for Industry. [https://www.fda.gov/media/134582/download]
2. Lewis DF, Ito Y. 2010. Expert Opin Drug Metab Toxicol. 6(6):661-674.
3. Guengerich FP. 2022. Biomol Ther. 30(1):1-18.
4. Sridhar J, *et. al.* 2012. Molecules. 17(8):9283-305.
5. Thresholds for inhibitions:

IC_{50}: concentration of the drug required to reduce the enzymatic activity to half its normal value

K_i: reversible inhibition constant

$R_1 = 1 + \frac{I_{max,u}}{K_{i,u}}$ where $I_{max,u}$ is the maximal unbound plasma concentration of the interacting drug at steady state and $K_{i,u}$ is the unbound inhibition constant determined in vitro

IC_{50} fold shift: shift in from IC_{50} from direct inhibition and after preincubation of the enzyme with nicotinamide adenine dinucleotide phosphate (NADPH)

k_{inact}: maximal inactivation rate constant

k_{obs}: first order rate constants for loss of CYP activity defined as $k_{obs} = k_{obs,[I]=0} + \frac{k_{inact}[I]}{K_I+[I]}$ where $k_{obs,[I]=0}$ is k_{obs} in the absence of the substrate, and *[I]* is the concentration of the inactivator. K_I: is the inactivator concentration that yields half of the maximum inactivation rate

$R_2 = \frac{k_{obs}+k_{deg}}{k_{obs}}$ where $k_{obs} = \frac{k_{inact} \times 50[I_{max}]}{K_I+50[I_{max}]}$ and k_{deg} is the apparent first-order degradation rate constant of the affected enzyme

6. Orr STM, *et al.* 2012. J Med Chem. 55(11):4896-933.