# Structure-Based Predictions of CYP selectivity, Reactivity, and Sites of Metabolism


**Michael Drummond[1], Olivier Rousseau[1], Chris Williams[1], and Rocky Goldsmith[2]**

**1. Chemical Computing Group 1010 Sherbrooke St., W., Suite 910, Montreal QC H3A 2R7, Canada**
**2. Bayer CropScience, 700 Chesterfield Parkway West, Chesterfield, MO, 63017, USA**



## Abstract

Cytochrome P450 oxidases (CYPs) are heme-containing enzymes responsible for clearing xenobiotics (including drug molecules) through oxidative metabolism. Thus, understanding the interactions between drug molecules and CYPs is critical for evaluating drug efficacy, clearance, toxicity, and drug-drug interactions. Although dozens of crystal structures of the five predominant CYP isoforms have been solved, most of the modeling tools that predict drug-CYP interactions completely neglect this structural information. In this work, both 2D and 3D methods are used to predict isoform selectivity, small molecule reactivity, and regioselectivity of CYPs. The 2D-based methods are parsimonious yet accurate, and can be used to quickly evaluate selectivity and reactivity. The 3D approach utilizes a pharmacophoric approach, providing a rapid and flexible way to predict CYP isoform selectivity and regioselectivity. Incorporating 3D CYP structural information into the models confers unique advantages over 2D-based approaches, such as the ability to distinguish reactivity differences among stereoisomers. Finally, predicted results can be readily visualized in a CYP pocket, and thus potential CYP liabilities are not merely flagged in a binary fashion, but can also be designed against in a structure-based design context – a clear improvement over the pass/fail filtering standards prevalent in CYP modeling efforts to date.


## Introduction

Cytochrome P450 oxidases (CYPs) are the primary macromolecular class of enzymes involved in oxidative metabolism and clearance of xenobiotics. CYPs exist in multiple isoforms. Five (5) of the 57 drug metabolizing human CYPs (hCYPs) are expressed at roughly 80% in human tissues (1A2, 2C9, 2C19, 2D6, 3A4). Such variety explains the multitude of oxidative reactions performed by CYPs (hydroxylation, epoxidation, deamination, dehalogenation, dealkylation, etc) as well as their selectivity towards different drugs/xenobiotics. Due to their diverse reactivity, CYPs are implicated in both Drug-Drug Interactions (DDIs) and adverse biotransformations. Complications can occur when two or more drugs are co-administered. Competitive inhibition can occur and create toxic buildup of a drug (Figure 1A). Biotransformations can be used to convert a pro-drug to an active drug; however adverse biotransformations can lead to toxic metabolite production (Figure 1B). A drug can act as substrate for one CYP and as an inhibitor for another (Figure 1C).

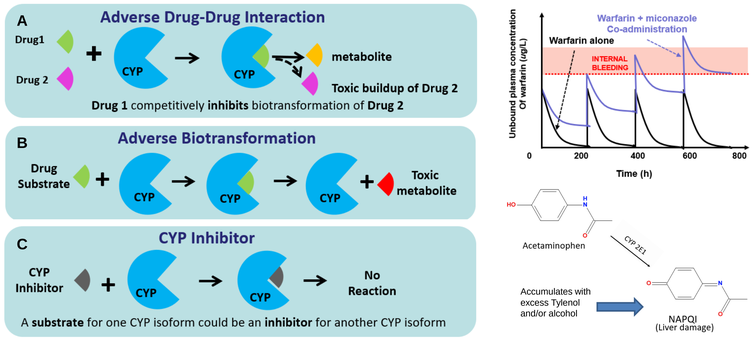





***Figure 1**. Illustration of the different undesired reactions occurring with CYPs such as toxic buildups with co-administration (A), toxic metabolites generation (B) or CYP inhibition (C).*

Understanding selectivity and reactivity can predict DDI liabilities. Understanding regioselectivity can lead to proactive modifications. The ideal CYP modeling method needs to be accurate, "fast enough", predictive for **Selectivity**, **Reactivity**, **Regioselectivity** ("one-stop shop") and suggest a path for rational improvement.

## Predictive CYP Modeling: 2D-based Models

**Binary-classification Trees (BCT) or Decision Trees**: Predict isoform class using a non-parametric model constructed by performing binary recursive partitioning
**BinaryQSAR**: Predict isoform class from Bayesian inference based on decorrelated descriptors distributions
**Fingerprint**: Predict isoform class by nearest-neighbor similarity searching

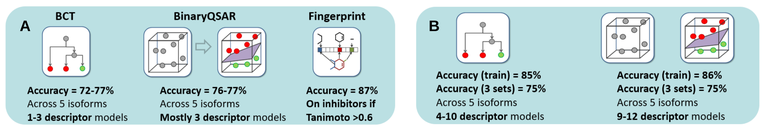


***Figure 2**. Performance of different reported 2D-based pharmacophore model for CYP isoform selectivity.*

- Simple selectivity models **(1-3 descriptors)** capture the spatial and physicochemical requirement of CYP isoform selectivity for ligands with **~75% accuracy** (Figure 2A).
- Simple reactivity models **(4-10 descriptors)** generally capture the reactive nature of substrates vs inhibitors with **> 75% accuracy** (Figure 2B).

## Predictive CYP Modeling 3D-based Models

**Why move to 3D?:** Stereoisomers are challenging in 2D, regioselectivity is inherently 3D and there is a clear signal in the CYP-ligand cocrystal structures.

**Starting point**: Use a CYP co-crystallized with a substrate in a productive binding mode as a template for predicting regioselectivity for this reaction (3UA1).

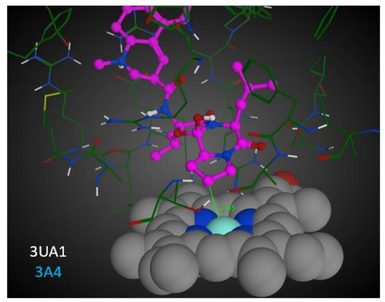


***Figure 3**. Starting point of the 3D-based pharmacophore model for CYP regioselectivity prediction using structure 3UA1.*

**Pharmacophore**: Using pharmacophore queries to describe the CYP binding has key advantages such as being able to incorporate features that encode reaction chemistry, and exclusion volume to encode protein shape and allowed binding pocket regions.
It uses three components:




Three chemistry specific features | Excluded volume for CYP receptor and heme | Excluded volume where the ligand can't go

Evaluation of the pharmacophore model was performed on a FDA Clinical test set (inherently drug-like) with experimentally-confirmed substrates for simplicity. 13 substrates were metabolized by 2D6 but not 3A4 while 22 substrates were metabolized by 3A4 but not 2D6.

- All substrates must have at least 1 secondary carbon in order to pass the pharmacophore.
- The pharmacophore should position a secondary carbon near the heme.

Matches **10/13** of the **2D6** Substrates (True Positives)
Also matches **3/22** of the **3A4** Substrates (False Positives)

| TP: 10/13, | TN: 19/22, | FP: 3/22, | FN: 3/13 |
|---|---|---|---|
| (77%) | (86%) | (14%) | (23%) |

- The missed 2D6 ligands are not simply the biggest, and the hit 3A4 ligands are not simply the smallest.
- **Clearly the pharmacophore can discriminate between 2D6 and 3A4 substrates.**

***Figure 4**. FDA clinical test set prediction using the 3D-based pharmacophore model.*

The 3D pharmacophore model can distinguish reactivity differences between stereoisomers. For example, a R conformer of the compound below passes the 2D6 CYP pharmacophore while the S conformer (manually superposed) clashes with the 2D6 CYP wall (Figure 5).

- **CYP pharmacophore model can distinguish R/S reactivity**

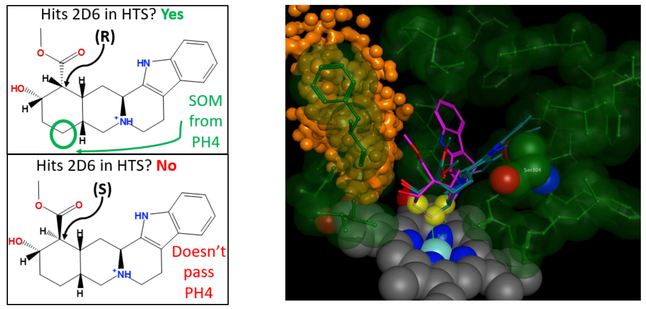


***Figure 5**. Stereoisomers reactivity prediction with the 3D-based pharmacophore model.*

## Predicting Regioselectivity

**Aliphatic Hydroxylation.** We used a dataset (Zaretzki *et al., JCIM* **2011**, 51, 1667) containing 21 true SOMs and 305 non-metabolized 2° carbons for aliphatic hydroxylation. Initial pharmacophore performed fairly well, but can be optimized. Adding a hydrogen bond donor feature on Glu216 halved the FP% (23% to 11%) and TP% only decreased by 5% (62% to 57%). Further optimization of feature radii and pocket volumes gives 76% TP and 7% FP.

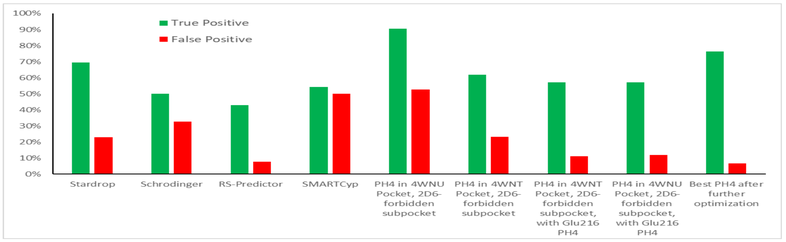


***Figure 6**. CYP aliphatic hydroxylation prediction with different pharmacophore models.*

**Aromatic Hydroxylation.** No crystal structures exist for aromatic hydroxylation (pre-reactive complex). Docking is performed with an initial, general pharmacophore feature: 5 Å radius "Aro" centroid near the $Fe^{IV}$=O. The pharmacophore was combined with the initial one and optimized with Extended Hückel Theory features.

**Careful:** Sigma + pi charge on C must be < -0.03q and union of two 2D6 pocket atoms drastically reduce FP (60% to 6%; TP at 51%)
**Balanced:** Conformational strain filter (dE < 3 kcal/mol) and intersection of two 2D6 pockets (more TPs than current state-of-the-art)
**Aggressive:** Set the dE filter to < 7 kcal/mol (even more TPs)

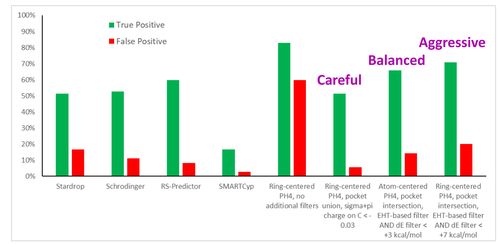


***Figure 7**. CYP aromatic hydroxylation prediction with different pharmacophore models including EHT.*

- **Stardrop:** Semi-empirical (AM1) calculated reaction barriers on actual system
- **Schrodinger:** Ensemble docking
- **RS-Predictor:** 540 2D ligand-based descriptors
- **SmartCYP**: Lookup of DFT-calculated reaction barriers on model systems.

- **Other advantages of 3D method are still present: distinguish stereoisomers, SBDD-like approach for "salvaging" a molecule, etc.**

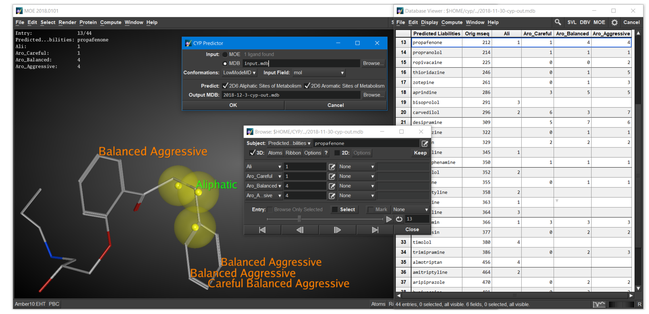


***Figure 8**. CYP Predictor panel in MOE.*

**CYP Predictor** is used to analyze molecules from the main MOE window or an MDB.
- Generate conformations on-the-fly or utilize pre-generated conformations
- Summary output MDB shows predicted liabilities across all pharmacophores
- Browse individual PH4 output MDBs to enable Structure-Based CYP avoidance
- Liability detection list will grow over time (inhibitors, N-dealkylation, different isoforms, 2D prefilters, etc.)

## Conclusion

The application of the 3D-based pharmacophore models for selectivity and regioselectivity prediction in CYPs has numerous advantages in drug discovery. It is fast, accurate, modular and can deal with stereoisomers. The precision of pharmacophore model can deeply influence its accuracy as demonstrated by the prediction of aliphatic and aromatic hydroxylation. For further improvement, 3d-based methods should be combined with 2D-based methods and as yet to be expanded to other isoforms, CYP reactions, and inhibitor vs. substrate reactivity.