Integrating Predictions from Complementary Cytochrome P450 (CYP) Models

Robert D. Clark, David Miller, Michael S. Lawless, and Marvin Waldman
Simulations Plus, Inc., Lancaster, CA 93534, USA (www.simulations-plus.com)

ABSTRACT

Many structure-activity classification models have been published for predicting whether a given compound is likely to inhibit and/or be subject to metabolism by a given cytochrome P450 (CYP) isoform, and several of them are commercially available. Some products predict the sites of metabolism (SoMs), liver microsomal stability, and which metabolites are most likely to be produced.

ADMET Predictor™ 8.0 includes substrate classification models for CYPs 1A2, 2A6, 2B6, 2C8, 2C9, 2C19, 2D6, 2E1, and 3A4, as well as models that predict the corresponding isoform-specific SoMs. In addition, models for inhibitor classification and site-specific kinetic parameters (maximal velocity (Vmax), Michaelis-Menten constant (Km), and intrinsic clearance (CLint)) are provided for five of the major hepatic CYPs: 1A2, 2C9, 2C19, 2D6 and 3A4.

ADMET Predictor integrates the substrate classification, SoM, and kinetic predictions and presents them in a readily interpretable way that identifies the most likely metabolites and predicts the contribution each will make to CYP metabolism *in vivo*.

EXAMPLES

Seven representative molecules (Figure 1) were selected from those in the ChEMBL database for which assay data on CYP oxidation rates in human liver microsomes (HLMs) were available. These compounds illustrate the kinds of metabolic predictions available in ADMET Predictor 8.0 and how the property predictions for the various CYP isoforms relate to one another.

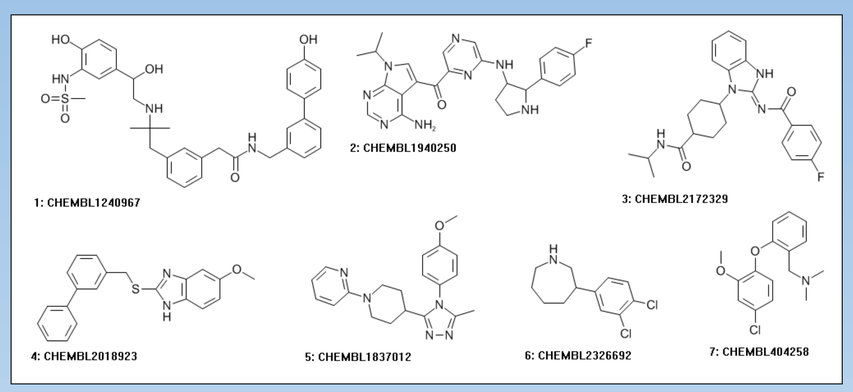

Figure 1. Compounds selected to illustrate how data is displayed in ADMET Predictor 8.0.

PROPERTIES PREDICTED

ADMET Predictor generates substrate classifications and sites of metabolism for CYPs 1A2, 2B6, 2C8, 2C9, 2C19, 2D6, 2E1 and 3A4. In addition, it provides estimates of the Michaelis constant (Km), maximal velocity (Vmax) and intrinsic clearance (CLint) for five major CYPs: 1A2, 2C9, 2C19, 2D6 and 3A4. The program provides inhibitor classification predictions for the same five CYPs as well as specific models for CYP3A4 inhibition of midazolam and testosterone metabolism.

Recombinant assays are "cleaner" systems for determining enzyme kinetics than HLMs, but the relationship between kinetics in the two systems can be complex. Therefore we also include Km, Vmax, and CLint models for total CYP CLint in HLMs and for CYP3A4 assayed in intact HLMs. The Km values for both HLM assays have been corrected for the estimated fraction unbound in microsomes (an additional model called S+fumic).

Figure 2 illustrates how metabolic property estimates are displayed in ADMET Predictor's main spreadsheet. The alphanumeric data columns visible relate to CYPs 1A2, 2A6, and 2B6. Predictions for other metabolic properties lie further to the right.

ANNOTATION

Classification confidence values calculated by applying beta binomial error analysis [2] are shown in parentheses. Units pop up as tooltips when the mouse hovers over a column header. All rCYP CLints are adjusted for microsomal abundance.

Star plots are used to summarize groups of relevant property predictions, which are also displayed in individual columns. Those for compound **3** have been enlarged and annotated in Figure 3. The various plots show:

- Key **physicochemical properties** (molecular weight, number of rotatable bonds, S+logP (lipophilicity), S+Peff (effective jejunal permeability) and log S+Sw (water solubility)
- CYP substrate classifications. A prediction that the compound is not a substrate yields a small colored sector (e.g., 2C9 in Figure 3).
- Predicted **CLint** values on a log scale that runs from 0 to 3.5. An empty sector indicates (e.g., for CYP2C9) indicates that no prediction was made because the compound is classified as a nonsubstrate .
- Potential drug development risks: limited intestinal absorption, CYP metabolism, mutagenicity, and toxicity. A sector indicating overall ADMET Risk™ incorporates includes all risk components. Companion spreadsheet columns indicate which Risk rules are violated by the compound [1].

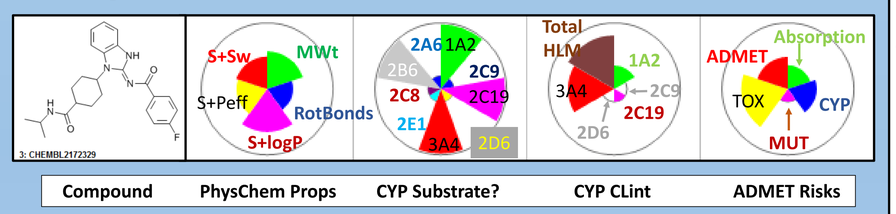

Figure 3. Star chart sector keys. The corresponding annotations pop up in the spreadsheet display whenever the mouse pointer hovers over a sector.

The spreadsheet displays molecular properties. Atomic properties can be displayed in the Atom Properties Viewer (Figures 4 and 5) by double-clicking on any highlighted cell in the spreadsheet.

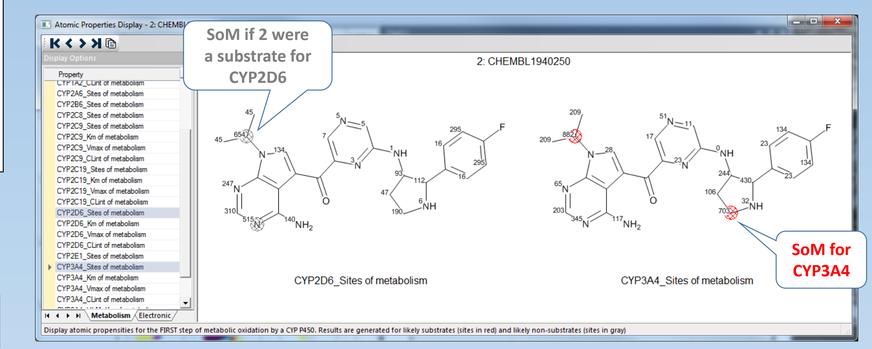

Figure 4. Predicted sites of metabolism for compound **2.**

Each potential SoM is labeled with its relative susceptibility to attack, with those most likely to be attacked highlighted by a hashed circle (Figure 4). **Gray highlighting** indicates that the compound is a not a substrate for the CYP in question (CYP2C9). **Red highlighting** is used when the compound is predicted to be a substrate for the corresponding CYP (here, for CYP3A4). Atomic Michaelis-Menten kinetic parameters can also be displayed for CYPs for which the compound is predicted to be a substrate (Figure 5).

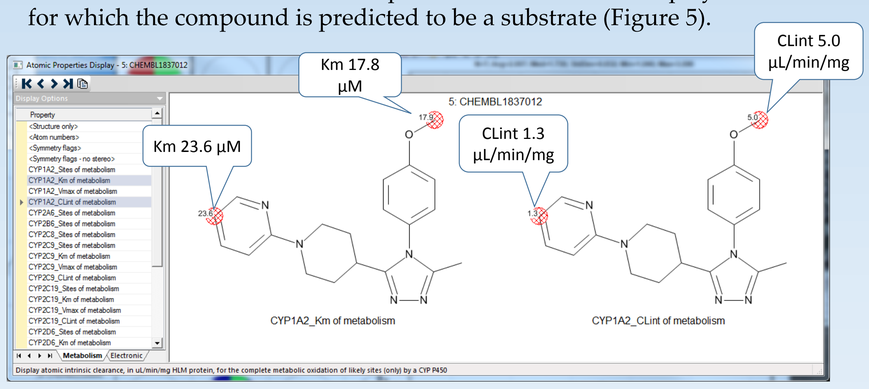

Figure 5. Estimated atomic CYP 1A2 Km and CLint values for compound **5.**

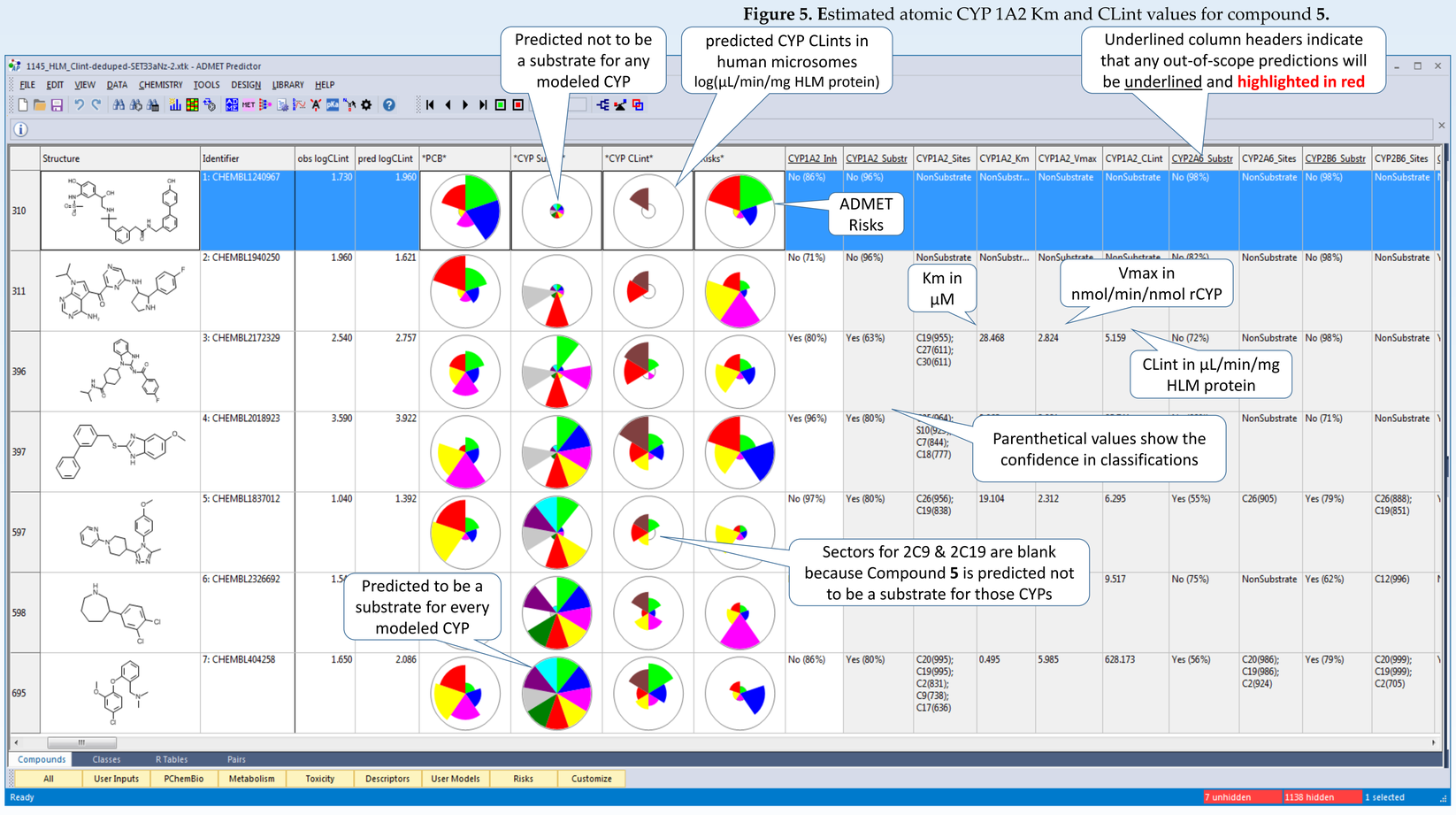

Figure 2. Spreadsheet showing selected star plots (left-of-center) and tabular (right hand columns) presentations of the data. All but seven rows have been hidden from view.

METABOLITE PREDICTION

The SoM predictions shown in Figure 4 are qualitative and cannot be compared meaningfully across CYPs. Quantitative predictions from the corresponding kinetic models need to be taken into account to discriminate between potential sites and those which are likely to dominate under *in vitro* assay conditions and *in vivo*.

Compound **1** is predicted not to be a CYP substrate for any of the modeled isoforms. It is, in fact, a substrate for CYP3A4 [3]. Hence it is a false negative, but – given that the associated confidence is only 58% - it is a marginal one.

Figure 5 shows the metabolic pathways predicted for four other examples.

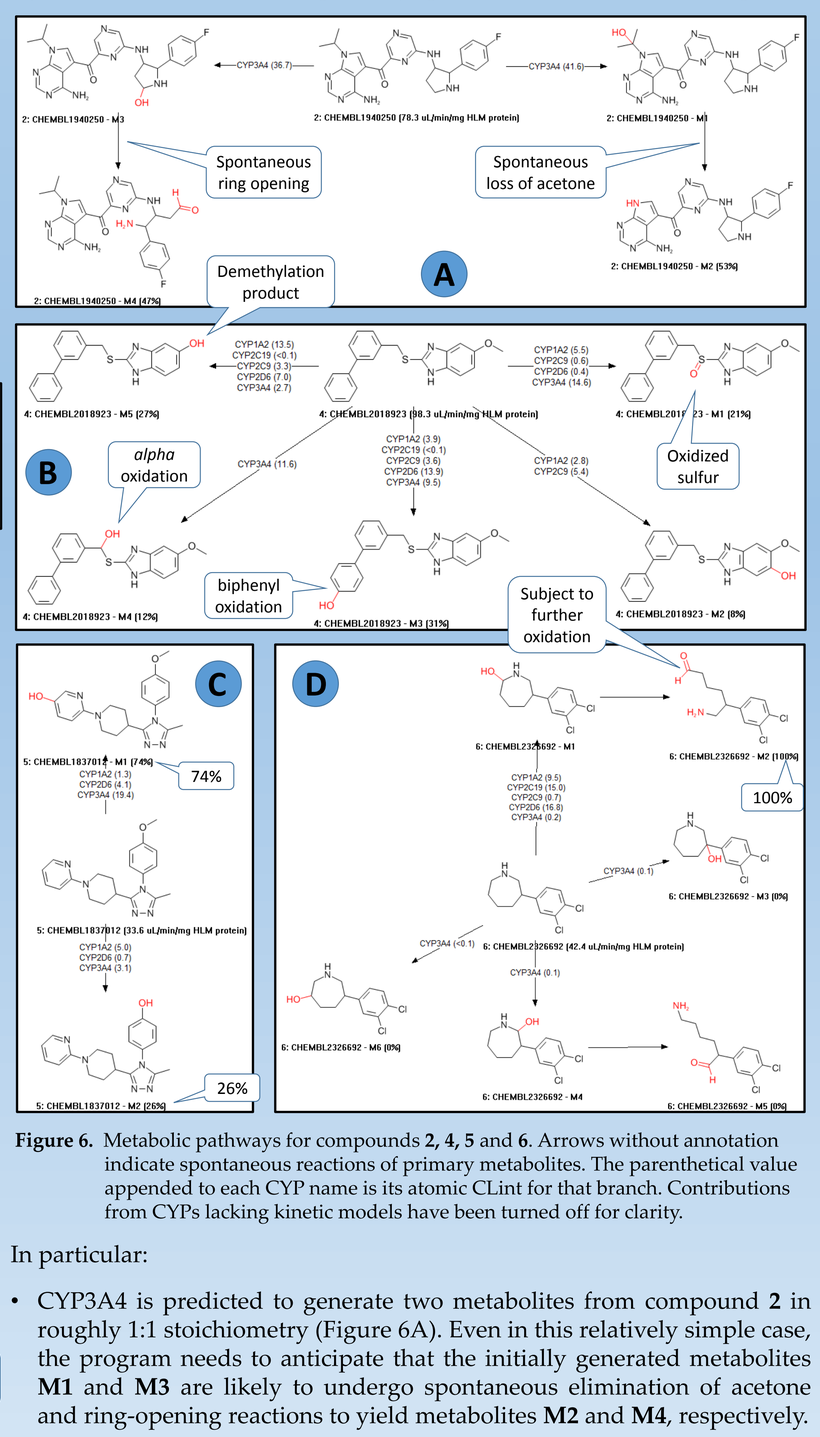

Figure 6. Metabolic pathways for compounds **2, 4, 5** and **6**. Arrows without annotation indicate spontaneous reactions of primary metabolites. The parenthetical value appended to each CYP name is its atomic CLint for that branch. Contributions from CYPs lacking kinetic models have been turned off for clarity.

In particular:

- CYP3A4 is predicted to generate two metabolites from compound **2** in roughly 1:1 stoichiometry (Figure 6A). Even in this relatively simple case, the program needs to anticipate that the initially generated metabolites **M1** and **M3** are likely to undergo spontaneous elimination of acetone and ring-opening reactions to yield metabolites **M2** and **M4**, respectively.
- Several CYPs are projected to oxidize the sulfur in compound **4** to give **M1**, hydroxylate the biphenyl group to yield **M3**, and demethylate it to form **M5** (Figure 6B). Hydroxylation *alpha* to the sulfur to form **M4** and *ortho* to the methoxy group to form **M2**, in contrast, are predicted to be catalyzed exclusively by CYP3A4 in the former case and by CYPs 1A2 and 2C9 in the latter.
- Oxidation of compound **5** to **M1** by CYP3A4 (Figure 6C) can be expected to dominate its metabolism. The predictive RMSE for logCLint of our models, however, is roughly 0.5 log units, which corresponds to a 3-fold error range. Hence the contribution of **M2** should not be discounted.
- Compound **7** presents a qualitatively different scenario (Figure 6D). Even allowing for predictive uncertainty, one of its metabolites (**M2**) is expected to completely dominate its metabolite profile. In fact, generating the metabolite profile indicates for this aldehyde indicates that it will be oxidized to the corresponding carboxylic acid by CYPs 1A2 and 2D6 as fast as it forms (details not shown).

CONCLUSIONS

ADMET Predictor 8.0 provides a broad suite of property predictions relevant to absorption, distribution, metabolism, excretion and toxicology (ADMET) in general, but its complement of CYP metabolism models in particular is remarkably complete. More importantly, the predictions from those models are well-integrated across isoforms via star plots; across atoms within molecules by visualization of metabolic "hot spots"; and most broadly, in terms of annotated metabolic pathways that include spontaneous secondary reactions. Taken together, these features make the program a powerful tool that is extremely well-suited to guide design, lead optimization, and DMPK studies.

REFERENCES

1. M.S. Lawless, M. Waldman, R. Fraczkiewicz & R.D. Clark. Using Cheminformatics in Drug Discovery. In: *Handbook of Experimental Pharmacology;* U. Nielsch, U. Fuhrmann & S. Jaroch, Eds.; Springer International, Berlin, **2015** (doi: 10.1007/ 164_2015_23).
2. R.D. Clark, W. Liang, A.C. Lee, M.S. Lawless, R. Fraczkiewicz & M. Waldman. *J. Cheminfo.* **2014**, *6*, 34.
3. Glossop *et al. J. Med. Chem.* **2010**, *53*, 6640–6652.

SimulationsPlus
SCIENCE + SOFTWARE = SUCCESS