

KOHONEN MAPS FOR PREDICTION OF BINDING TO HUMAN CYTOCHROME P450 3A4

Konstantin V. Balakin, Sean Ekins, Andrey Bugrim, Yan A. Ivanenkov, Dmitry Korolev, Yuri V. Nikolsky, Andrey V. Skorenko, Andrey A. Ivashchenko, Nikolay P. Savchuk, and Tatiana Nikolskaya

Chemical Diversity Labs, Inc., San Diego, California (K.V.B., Y.A.I., D.K., A.V.S., A.A.I, N.P.S.); and GeneGo, Inc., St. Joseph, Michigan (S.E., A.B., Y.V.N., T.N.)

Received April 20, 2004; accepted July 1, 2004

ABSTRACT:

The drug development process utilizes the parallel assessment of activity at a therapeutic target as well as absorption, distribution, metabolism, excretion, and toxicity properties of molecules. The development of novel, reliable, and inexpensive computational methods for the early assessment of metabolism and toxicity is becoming increasingly an important part of this process. We have used a computational approach for the assessment of drugs and drug-like compounds which bind to the cytochromes P450 (P450s) with experimentally determined K_m values. The physicochemical properties of these compounds were calculated using molecular descriptor software and then analyzed using Kohonen self-organizing maps.

This approach was applied to generate a P450-specific classification of nearly 500 drug compounds. We observed statistically significant differences in the molecular properties of low K_m molecules for various P450s and suggest a relationship between 33 of these compounds and their CYP3A4-inhibitory activity. A test set of additional CYP3A4 inhibitors was used, and 13 of 15 of these molecules were collocated in the regions of low K_m values. This computational approach represents a novel method for use in the generation of metabolism models, enabling the scoring of libraries of compounds for their K_m values to numerous P450s.

A number of pharmaceutical compounds have been withdrawn from the market due to toxicity, metabolism (including drug-drug interactions), and pharmacokinetic issues (Estabrook, 1996; Ioannides, 1996). It is assumed that a larger number of proprietary molecules fail at earlier stages in pharmaceutical companies for similar reasons. Hence companies are trying to improve late stage success by removing problematic molecular series earlier in the drug discovery pipeline. The metabolic transformations of pharmaceuticals and other xenobiotics in the human body profoundly affect bioavailability, efficacy, chronic toxicity, and excretion rate and route. Both the parent molecule and the products of its metabolic transformations may also interfere with endogenous metabolism or other coadministered compounds. The inhibition of metabolizing enzymes can be associated with drug-drug interactions, which can have potentially fatal consequences for the patient. This behavior is traditionally studied in vitro, but due to the rapid accumulation of this empirical data, they can be predicted computationally (Ekins et al., 2001, 2003a). Such computational models can be based solely on the molecular descriptors derived from the structure of compounds and may allow the removal of potentially undesirable compounds from the early drug discovery process.

The majority of xenobiotics undergo phase I metabolism via the

cytochrome P450 (P450) enzymes, predominantly in the liver (Ioannides, 1996). P450s are mixed-function monooxygenases capable of either inactivating or activating xeno- and endobiotic molecules alike. Of over 50 human P450 genes cloned and classified according to sequence homology, three P450 families and fewer than a dozen unique enzymes have been shown to play a substantial role in human hepatic metabolism of drugs. P450s display high sequence homology yet often have highly distinct roles in xenobiotic metabolism with active sites that enable broad and overlapping substrate specificity. This ligand binding promiscuity of many P450s complicates the prediction of therapeutic or toxic effects of xenobiotic metabolism. It has been shown that substrate selectivity of human P450s is related to the substrate structure and the key features of the active sites, namely, the disposition of certain amino acid residues within the heme environment (Lewis, 1996). From in vitro kinetic studies it became apparent that many of these P450s displayed autoactivation or heteroactivation kinetics that resemble allosteric kinetics (Ekins et al., 1998; Korzekwa et al., 1998). The recently published crystal structure of CYP2C9 also revealed an additional possible site for substrate binding (Williams et al., 2003). These factors all complicate building accurate predictive models for P450 binding.

To date, specific enzyme-substrate/inhibitor recognition interactions have been studied extensively, and several quantitative structure-activity relationships and pharmacophore models have been built for a limited number of P450s (Smith et al., 1997a,b; Lewis et al., 1998; Ekins et al., 2001; de Groot and Ekins, 2002; Szklarz and Paulsen, 2002). These have generally shown the importance of hydrophobic,

This work is supported by National Institutes of Health Grant 1-R43-GM069124-01 "In Silico Assessment of Drug Metabolism and Toxicity."

Article, publication date, and citation information can be found at <http://jpet.aspetjournals.org>.

doi:10.1124/dmd.104.000356.

ABBREVIATIONS: P450, cytochrome P450; PC, principal component; SOM, self-organizing map; HBA, hydrogen bond acceptor; HBD, hydrogen bond donor; PNSA-1, partial negative surface area 1; LY213829, tazofelone; LY303870, (*R*)-*N*-[2-[acetyl(³H₃)][(2-methoxyphenyl)-methyl]amino]-1-(1*H*-indol-3-ylmethyl)ethyl][1,4'-bipiperidine]-1'-acetamide.

hydrogen-bonding, and ionizable features for substrates based on K_m data as well as inhibitors using K_i , IC_{50} , and percentage inhibition data (Ekins et al., 1999a,b, 2000, 2001). In the current study, a computational algorithm is described for the classification of drug-like molecules based on their K_m values for human P450s. This approach complements the assessment of product-substrate specificity for the same enzymes and provides further novel insight into methods for predicting P450 involvement in metabolism (Korolev et al., 2003).

Materials and Methods

Databases. A data set of over 500 literature compounds with experimental apparent K_m values for 12 human P450s was obtained from the commercially available MetaDrug database (GeneGo, Inc., St Joseph, MI). Each compound was assigned to at least one enzyme-specific group within which the compounds were conditionally divided into three nonoverlapping categories: low K_m ($K_m < 10 \mu\text{M}$), moderate K_m ($K_m = 10\text{--}100 \mu\text{M}$), and high K_m ($K_m > 100 \mu\text{M}$). Before modeling experiments, the molecules were also filtered to ensure that they were "drug-like," based on molecular weight (range 150–700) and atom type content (only C, N, O, H, S, P, F, Cl, Br, and I were permitted) (Walters and Murcko, 2002). Some specific compound classes typically not related to drug-like agents, such as polyaromatic compounds or long-chain linear molecules (e.g., leukotrienes, fatty acids), were excluded from this reference set. Ultimately, 491 compounds remained from the initial database. The main descriptive statistics for these data set are described in Table 1.

Molecular Descriptors. Initially, 60 molecular descriptors representing lipophilicity, charge distribution, topological features, steric and surface parameters, and other physicochemical parameters were calculated for the entire data set using ChemoSoft (Chemical Diversity Labs, Inc., San Diego, CA) and Cerius²/Descriptor software (Accelrys, San Diego, CA). This initial descriptor space was reduced via a principal component (PC) analysis using ChemoSoft. About 90% of the variance was explained by the first seven PCs, the first six of which were found to be significant using standard Kaiser-Guttman and scree tests (Guttman, 1954; Catell, 1966; Jolliffe, 1986). Finally, six descriptors (Table 2) maximally contributing to the first six PCs were selected, based on these results, as the most relevant and were used as input parameters in all further neural network experiments.

Kohonen Self-Organizing Maps. The generation of the Kohonen self-organizing maps (SOMs) (Kohonen, 1989) was conducted using the ChemoSoft software. The training parameters for the SOM were as follows: the number of interactions for the training runs was 2000, the starting adjustment radius for the training runs was 0.1, and the decay factor was 0.001. Only one SOM was generated for the entire training set (491 compounds). After the SOM was generated, we studied the distribution of various compound groups (such as strong or poor binders, strong binders to particular isozymes, etc.) as separate maps.

Two external test sets were used for assessment of a relationship between substrates and inhibitors of human P450s. One set comprised 33 compounds which were classified in the MetaDrug database as reversible competitive CYP3A4 inhibitors. In addition, another 15 CYP3A4 competitive inhibitors

were compiled from the literature and selected as an independent test set (He et al., 1998; Iribarne et al., 1998; Ekins et al., 1999a; Gibbs et al., 1999; Katoh et al., 2000; Zhang et al., 2002).

Results

Molecular Features Important for P450 K_m . Upon binding to P450s, a molecule can interact either with the heme prosthetic group or with the other regions of the active site. The heme prosthetic group is the oxidation center for P450-catalyzed reactions; thus, compounds with lone electron pairs tend to form stronger complexes (Yan and Caldwell, 2001). For example, many compounds with nitrogen-containing heterocycles (such as imidazole, quinoline, pyridine, etc.) bind tightly to the heme iron of P450s. The intermolecular interactions involving polypeptide chains, such as hydrophobic and electrostatic interactions, Van der Waals forces, and H-bond formation, are also important for binding. The specific local microenvironment of the active site of a particular P450 determines the molecular features that a molecule should possess to bind to the site (Lewis, 2000; Lewis et al., 2002). Based on these known experimental observations and on the results of principal component analysis, we selected six descriptors (Table 2), which adequately describe the P450 K_m values. In terms of relative importance for P450 K_m , the properties of molecules in descending order are as follows: topologic complexity (Zagreb), H-binding capacity (HBA, HBD), flexibility (B_{rot}), surface charges (PNSA-1), and lipophilicity (log P).

Differences between Low and High K_m Molecules for P450s Based on Individual Descriptors. To uncover the differences between the low K_m ($<10 \mu\text{M}$) and high K_m ($>100 \mu\text{M}$) binders, we analyzed the corresponding distribution histograms based on individual molecular descriptors. The two-tailed t test was used for evaluation of statistical significance. Few statistically significant differences were revealed for most of the P450s studied, due to the low number of compounds in these isozyme-specific data sets shown in Table 1. However, for the most extensively studied enzyme, CYP3A4, the two groups of compounds were statistically significantly separated based on log P, the number of rotatable bonds, Zagreb index, and PNSA-1 (Fig. 1). For the normally distributed CYP3A4 set studied in this work, t values higher than 5 indicate statistically significant differences in mean values. As is evident from the histograms, the binding affinity to the active site of CYP3A4 increases with higher molecular lipophilicity, more rotatable bonds, larger topological complexity, and partial negative surface area. The presence of a large overlapping area between the groups indicates that effective differentiation between high and low K_m CYP3A4 substrates based on individual molecular

TABLE 1

Descriptive statistics for the P450 database derived from the commercially available MetaDrug database (GeneGo, Inc., St. Joseph, MI)

Enzyme	Number of Compounds	Number of Compounds per K_m		
		$K_m < 10 \mu\text{M}$	$K_m = 10\text{--}100 \mu\text{M}$	$K_m > 100 \mu\text{M}$
CYP1A1	12	7	4	1
CYP1A2	43	17	16	10
CYP2A6	15	1	3	11
CYP2B6	51	15	19	17
CYP2C8	13	6	5	2
CYP2C9	41	12	21	8
CYP2C19	48	18	21	9
CYP2D6	75	45	23	7
CYP2E1	19	2	8	9
CYP3A4	126	38	56	32
CYP3A5	12	5	6	1
CYP19	18	18	0	0
Total	491	180	208	103

TABLE 2

The six most significant PCs for the initially calculated descriptors (ChemoSoft or Cerius²) subsequently used in the SOM

Coefficients larger than 0.30 are shown in boldface.

Descriptor	Definition	PC1	PC2	PC3	PC4	PC5	PC6
Log P	Log of 1-octanol/water partition coefficient	0.386	0.047	0.581	0.019	0.054	0.131
B_rot	Number of rotatable bonds	0.786	0.035	0.084	0.141	0.334	0.093
HBA	Number of H-bond acceptors	0.800	0.357	0.169	0.376	0.033	0.107
HBD	Number of H-bond donors	0.459	0.007	0.481	0.023	0.270	0.327
PNSA-1	Partial negative surface area	0.580	0.743	0.273	0.007	0.009	0.038
Zagreb	Sum of the squares of vertex valencies	0.964	0.002	0.049	0.314	0.153	0.123

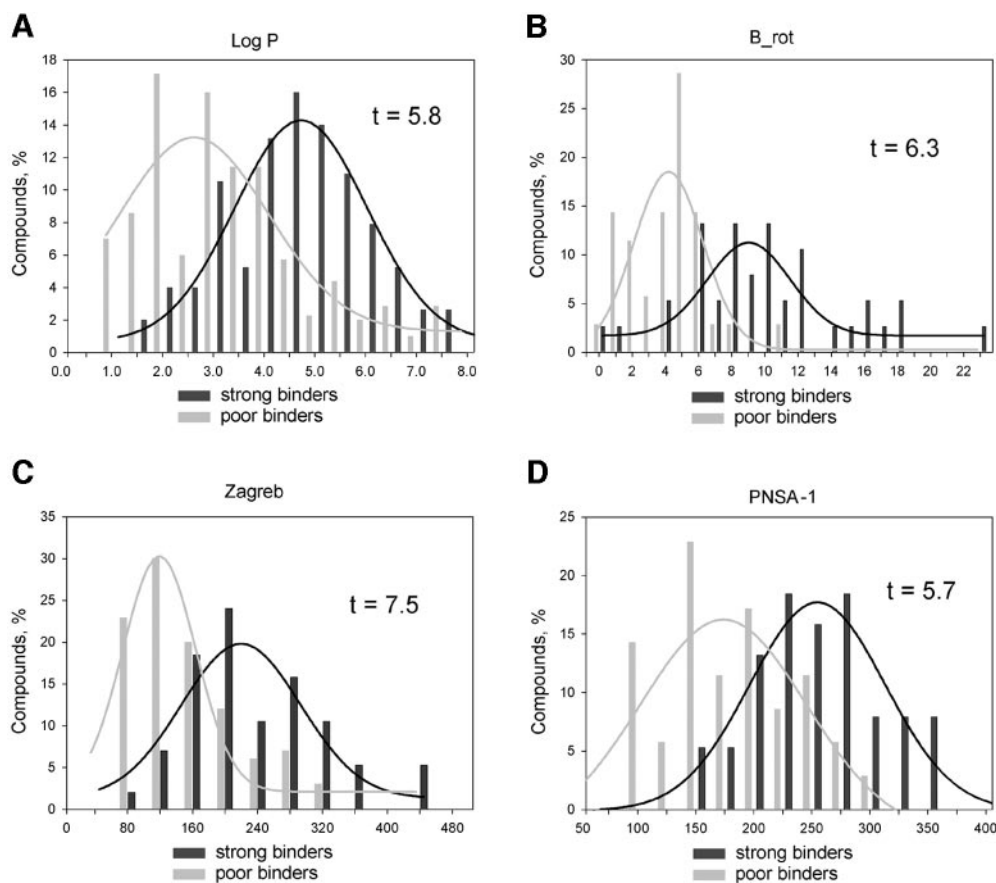


FIG. 1. Differences between CYP3A4 low and high K_m molecules based on individual molecular descriptors: log P, B_rot, Zagreb, and PNSA-1. The plots show calculated t -indices, which are standard measures of significance of difference between the mean values. The height of bars on the histograms is equal to the percentage of compounds falling into the defined range of the descriptor value.

descriptors may be problematic and not ideal, indicating the need for more complex approaches.

Nonlinear Classification Modeling. Recently, we developed an effective SOM classification scheme for substrate/nonsubstrate segregation of human P450s (Korolev et al., 2003). In the current study, a comprehensive set of 491 drug-like compounds with experimentally determined apparent K_m values versus 12 human P450s were used. Each P450 was placed in one of three categories depending on its K_m values. Based on this categorization, we generated the computational models differentiating between the high and low K_m compounds for each P450. After calculating the previously described molecular descriptors, we generated an SOM for the entire reference P450 K_m data set using the unsupervised learning procedure. The high K_m ($>100 \mu\text{M}$) and low K_m ($<10 \mu\text{M}$) groups occupied distinctly different sites on the map, and the moderate affinity binders typically occupied the intermediate positions. For illustration, we have shown the positions of low K_m and high K_m molecules for both CYP3A4 and CYP2D6

enzymes, which represent the two largest data sets studied (Fig. 2). Compounds with low K_m for both CYP3A4 and CYP2D6 occupied somewhat different sites on the map, although there was some substantial overlap between these enzymes (Fig. 2).

The distance between nodes on the SOM is a dimensionless parameter; it represents an abstract, discrete distance between the points in a multidimensional property space. For each isozyme-specific group, the areas of strong/poor binders can be identified as the nodes on the map, in which the percentage of strong/poor binders (with respect to their total number equal to 100%) is higher than the percentage of compounds belonging to the opposite category. In the case of CYP3A4, the model correctly classified 91% high K_m and 97% low K_m molecules as defined by their localization in the corresponding areas of the SOM (Table 3). The quality of this discrimination is statistically significant only in the case of CYP3A4, for which a relatively large number of low K_m and high K_m molecules are available. Although the study suggests the method may be able to

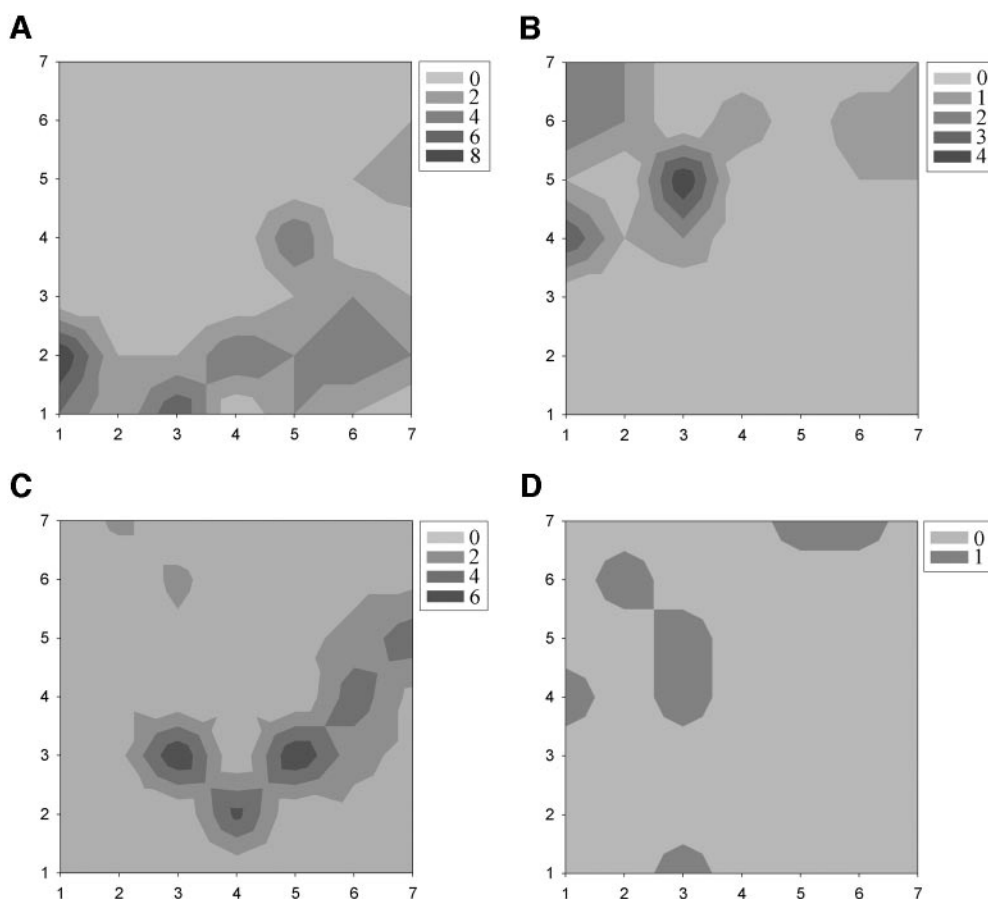


FIG. 2. Distribution of low K_m ($K_m < 10 \mu\text{M}$) and high K_m ($K_m > 100 \mu\text{M}$) molecules for two major P450 enzymes on the SOM map: a, CYP3A4 low K_m (38 compounds); b, CYP3A4 high K_m (32 compounds); c, CYP2D6 low K_m (45 compounds); and d, CYP2D6 high K_m (7 compounds). The data have been smoothed for presentation purposes.

TABLE 3

Classification quality for the developed classification Kohonen model

	Classified as Low K_m^*	Classified as High K_m^{**}
CYP3A4		
High K_m	3 (9.3%)	29 (90.7%)
Low K_m	37 (97.4%)	1 (2.6%)
CYP2C9		
High K_m	1 (12.5%)	7 (87.5%)
Low K_m	11 (91.7%)	1 (8.3%)
CYP2B6		
High K_m	2 (13.3%)	13 (86.7%)
Low K_m	16 (94.1%)	1 (5.9%)
CYP2D6		
High K_m	0 (0%)	7 (100%)
Low K_m	45 (100%)	0 (0%)

* $K_m < 10 \mu\text{M}$.

** $K_m > 100 \mu\text{M}$.

discriminate between the other P450s, more data are required for a statistically valid result.

We also applied the SOM to discriminate between low K_m and high K_m molecules across the whole panel of P450 enzymes. It should be taken into account that a molecule may have a low K_m with one P450 and a high K_m for another P450. Accordingly, the same compound can be considered either a low K_m or a high K_m compound, depending on the specific P450 being considered. Such compounds were assigned to the low K_m category, because this is likely the most important. In Fig. 3, the distribution of low K_m molecules with $K_m < 10$ with respect to at least one P450 isozyme is shown as the green area, and the high K_m molecules with $K_m > 100$ for at least one P450 isozyme (and no low

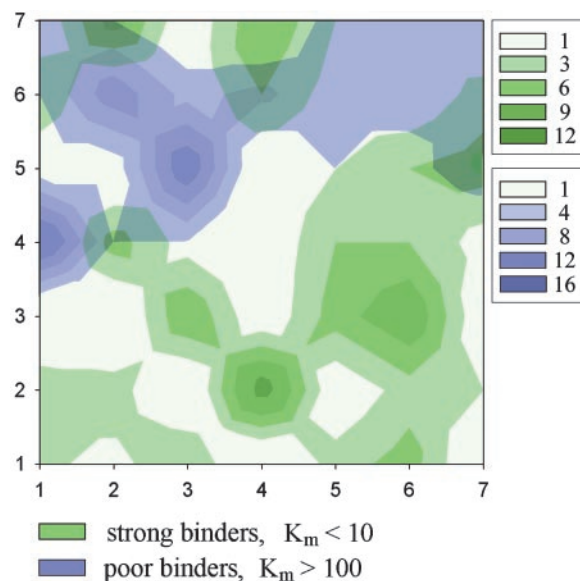


FIG. 3. An SOM map of low K_m ($K_m < 10 \mu\text{M}$) and high K_m ($K_m > 100 \mu\text{M}$) molecules for the whole panel of 12 P450 isozymes. The data have been smoothed for presentation purposes.

K_m values for any other isozyme) is shown as the blue area. These two compound categories occupy distinctly different sites on the map. The classification quality was 65% for low K_m and 82% for high K_m molecules, which suggests some utility of this model for predicting

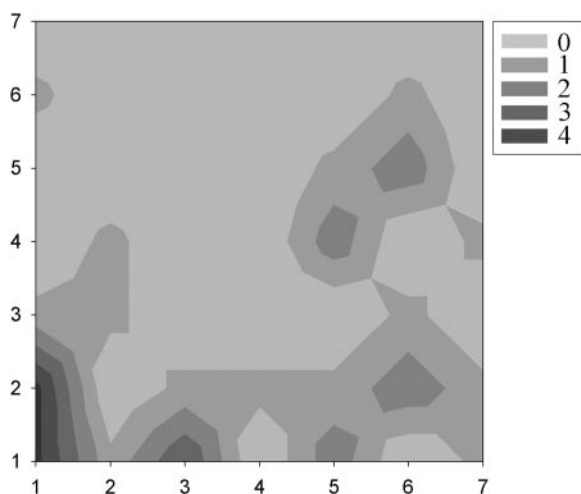


FIG. 4. Distribution of 33 competitive CYP3A4 inhibitors on the SOM map. Of these, 31 compounds (94% of all inhibitors) fall into the area of low K_m ($K_m < 10 \mu\text{M}$) CYP3A4 substrates for this isozyme (also see Fig. 2a and Fig. 5 for comparison). The data have been smoothed for presentation purposes.

global binding to human P450s, but which could clearly be improved upon.

Colocalization of P450 Inhibitors and Substrates. A significant issue in drug metabolism is whether there is a relationship between substrates and inhibitors of human P450s. We have addressed this problem in experiments with independent groups of CYP3A4 inhibitors due to the major role of this enzyme in metabolizing at least 50% of marketed drugs (Yan and Caldwell, 2001), and the availability of a large set of K_m and inhibition data. Overall, 33 compounds were classified in the MetaDrug database as reversible competitive CYP3A4 inhibitors and were processed on an SOM (Fig. 4); of these, 94% were located in the area of low K_m CYP3A4 molecules.

In addition, another 15 CYP3A4 competitive inhibitors were compiled from the literature and selected as an independent test set (Table 4) (He et al., 1998; Iribarne et al., 1998; Ekins et al., 1999a; Gibbs et al., 1999; Katoh et al., 2000; Zhang et al., 2002). The molecular descriptors were calculated for these 15 molecules and then they were positioned on the same SOM (Fig. 5) as described for the previous set of inhibitors. In this case, 87% of these molecules were located in the areas of low K_m CYP3A4 molecules. Only two compounds of the 15, namely **4** and **14** (LY213829 and LY303870) were misclassified yet are still located close to low K_m CYP3A4 molecules on the SOM.

Discussion

Previous observations indicate that compounds that bind to human P450s as substrates may also act as inhibitors in some cases. We have shown that the molecular properties of low K_m molecules are somewhat unique for different P450s (Fig. 2, a and c) because they occupy unique sites on the SOM, despite a significant overlapping region. Therefore, the developed computational models may be used to identify selective or nonselective P450 inhibitors in early drug discovery. The number of selective reversible inhibitors of P450s has increased over the past decade, providing researchers with probes to gain important insights into the molecular mechanisms of these enzymes (Branch et al., 2000; Hollenberg, 2002). These inhibitors can be used as potential therapeutic agents because P450 enzymes are responsible for the metabolic activation and detoxification of various chemical carcinogens and other toxins. Molecules capable of selective inhibition of these enzymes may shift the balance between the various metabolic pathways, so that metabolic activation is minimized,

whereas detoxification is enhanced. The coadministration of specific P450 inhibitors can reduce first-pass drug metabolism, improve drug dosage, and ultimately enhance drug bioavailability. However, the concurrent administration of two or more drugs is a common therapeutic practice and occurs more frequently in the aging population. The administered drugs can compete for the same or other sites (Ekins et al., 1998; Korzekwa et al., 1998) in the P450 enzyme, which may inhibit the elimination of drugs and result in undesirable toxic effects (Ito et al., 1998; Thummel and Wilkinson, 1998). Drug interactions are therefore a leading cause of death of hospitalized patients in the United States (Yuan et al., 1999; Marroum et al., 2000). The reliable prediction of drug-drug interactions is a significant issue for absorption, distribution, metabolism, excretion, and toxicity research in general. In vitro P450 inhibition assays have proven to be valuable in predicting interactions, although they are relatively expensive and time-consuming, and obviously require the synthesis of the molecules. Although inhibition of a P450 enzyme in vitro is not necessarily associated with drug-drug interactions in clinical studies, lead compounds with weak P450 inhibition are apparently favored for drug development (Yan and Caldwell, 2001). Therefore, reliable computational methods for the assessment of P450-inhibitory activity may be a viable complementary approach (not requiring molecule synthesis) with higher throughput and cost effectiveness, enabling use much earlier in the virtual stages of drug discovery.

In this study, we have demonstrated a relationship between the predicted K_m and the reversible competitive CYP3A4 inhibition of drugs. Based on these results, it would appear that generally low K_m CYP3A4 compounds correspond to competitive CYP3A4 inhibitors. These results are in agreement with literature data, where reversible competitive P450 inhibition is associated with a high affinity (low K_m) for the P450 active sites (Yuan et al., 1999; Hollenberg, 2002). Because this type of P450 inhibition is thought to be the most common cause of drug-drug interactions (Ito et al., 1998; Thummel and Wilkinson, 1998), the developed computational models are likely applicable for predicting P450 interactions.

We have classified a comprehensive set of drug compounds according to their K_m values for the active sites of several major P450 enzymes, applying a nonlinear SOM to interpretable molecular descriptors. Two compound categories, low K_m and high K_m ligands for P450s, were effectively separated based on the preselected set of physicochemical molecular descriptors. The groups of low K_m molecules for different P450s clustered in distinctly unique locations, forming P450-specific groups. Since low K_m values for P450s also correlate with the potential for competitive P450 inhibition, these models may be useful for the development of selective P450-specific inhibitors for potential therapeutic and mechanistic applications. The limited accuracy of the general model can be naturally explained by different and only partially overlapping substrate/inhibitor specificity for various members of the cytochrome P450 family. The model can be used as a filter at the stage of presynthetic library design, which should reduce the number of potential P450 substrates/inhibitors in a high-throughput fashion. This work therefore represents the continuation of our research on structure-activity relationships for drug candidates and human P450s involved in metabolism (Ekins, 2003; Ekins et al., 2003a,b; Korolev et al., 2003). Our results combined with the prior literature data for numerous applications demonstrate how a small number of simple molecular descriptors can be used with SOMs to provide an efficient clustering, classification, and visualization tool for P450s. SOMs also have the added advantage that the molecules do not need alignment. Combinations of computational models for P450 metabolism are applicable to aiding the selection of molecules during early drug discovery and therefore represent an approach to filtering

TABLE 4

Structures of competitive CYP3A4 inhibitors used for the independent external validation of the SOM

Number	Name	Structure
1	Clotrimazole ²⁶	
2	Ketoconazole ²⁷	
3	Tioconazole ²⁷	
4	LY213829 ²⁸	
5	Sulconazole ²⁷	
6	Miconazole ²⁷	
7	Barnidipine ²⁹	
8	Nicadipine ²⁹	
9	Benidipine ²⁹	
10	LY335979 ²⁸	
11	LY350965 ²⁸	
12	Indinavir ⁵⁰	
13	Bergamottin ³¹	
14	LY303870 ²⁸	
15	Terfenadine ²⁸	

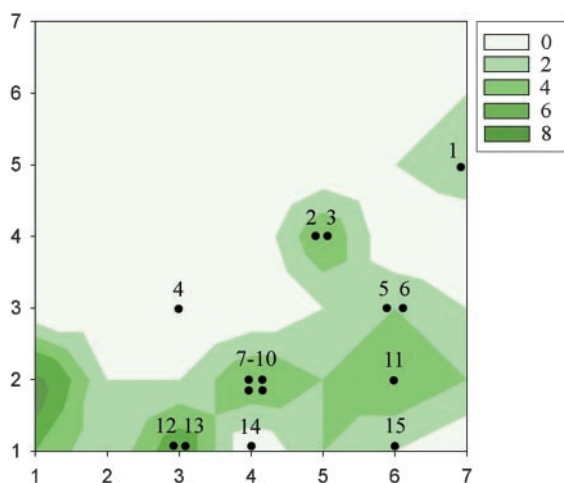


Fig. 5. Distribution of 15 competitive CYP3A4 inhibitors from the internal test set within the SOM map of low K_m ($K_m < 10 \mu\text{M}$) CYP3A4 substrates (identical to those shown in Fig. 2a). Compound structures are shown in Table 4. Of these, 13 inhibitors (87%) fall into the area of low K_m molecules. The data have been smoothed for presentation purposes.

large libraries alongside other predicted absorption, distribution, metabolism, excretion, and toxicity properties (Ekins et al., 2002; Shimada et al., 2002). With the addition of further in vitro data, it is likely that computational models can be generated for the currently less well studied P450s as well as other metabolic enzymes.

References

- Branch RA, Adedoyin A, Frye RF, Wilson JW, and Romkes M (2000) In vivo modulation of CYP enzymes by quinidine and rifampin. *Clin Pharmacol Ther* **68**:401–411.
- Catell RB (1966) The scree test for the number of factors. *Multi Behav Res* **1**:245–276.
- de Groot MJ and Ekins S (2002) Pharmacophore modeling of cytochromes P450. *Adv Drug Deliv Rev* **54**:367–383.
- Ekins S (2003) In silico approaches to predicting drug metabolism, toxicology and beyond. *Biochem Soc Trans* **31**:611–614.
- Ekins S, Berbaum J, and Harrison RK (2003a) Generation and validation of rapid computational filters for CYP2D6 and CYP3A4. *Drug Metab Dispos* **31**:1077–1080.
- Ekins S, Boulanger B, Swaan PW, and Hupcey MA (2002) Towards a new age of virtual ADME/TOX and multidimensional drug discovery. *J Comput-Aided Mol Des* **16**:381–401.
- Ekins S, Bravi G, Binkley S, Gillespie JS, Ring BJ, Wikel JH, and Wrighton SA (1999a) Three- and four-dimensional quantitative structure activity relationship analyses of cytochrome P-450 3A4 inhibitors. *J Pharmacol Exp Ther* **290**:429–438.
- Ekins S, Bravi G, Binkley S, Gillespie JS, Ring BJ, Wikel JH, and Wrighton SA (2000) Three- and four-dimensional-quantitative structure activity relationship (3D/4D-QSAR) analyses of CYP2C9 inhibitors. *Drug Metab Dispos* **28**:994–1002.
- Ekins S, Bravi G, Wikel JH, and Wrighton SA (1999b) Three-dimensional-quantitative structure activity relationship analysis of cytochrome P-450 3A4 substrates. *J Pharmacol Exp Ther* **291**:424–433.
- Ekins S, de Groot MJ, and Jones JP (2001) Pharmacophore and three-dimensional quantitative structure activity relationship methods for modeling cytochrome P450 active sites. *Drug Metab Dispos* **29**:936–944.
- Ekins S, Ring BJ, Binkley SN, Hall SD, and Wrighton SA (1998) Autoactivation and activation of the cytochrome P450s. *Int J Clin Pharmacol Ther* **36**:642–651.
- Ekins S, Stresser DM, and Williams JA (2003b) In vitro and pharmacophore insights into CYP3A enzymes. *Trends Pharmacol Sci* **24**:161–166.
- Estabrook RW (1996) Cytochrome P450: from a single protein to a family of proteins—with some personal reflections, in *Cytochromes P450: Metabolic and Toxicological Aspects* (Ioannides C and Parke DV eds) pp 4–28, CRC Press LLC, Boca Raton, FL.
- Gibbs MA, Kunze KL, Howald WN, and Thummel KE (1999) Effect of inhibitor depletion on inhibitory potency: tight binding inhibition of CYP3A4 by clotrimazole. *Drug Metab Dispos* **27**:596–599.
- Guttman L (1954) Some necessary conditions for common factor analysis. *Psychometrika* **19**:149–162.
- He K, Iyer K, Hayes R, Sinz M, Woolf T, and Hollenberg P (1998) Inactivation by cytochrome P450 3A4 by bergamottin, a component of grapefruit juice. *Chem Res Toxicol* **11**:252–259.
- Hollenberg PF (2002) Characteristics and common properties of inhibitors, inducers and activators of CYP enzymes. *Drug Metab Rev* **34**:17–35.
- Ioannides C, editor (1996) *Cytochromes P450: Metabolic and Toxicological Aspects*. CRC Press LLC, Boca Raton, FL.
- Iribarne C, Berthou F, Carlhant D, Dreano Y, Picard D, Lohezic F, and Riche C (1998) Inhibition of methadone and buprenorphine N-dealkylations by three HIV-1 protease inhibitors. *Drug Metab Dispos* **26**:257–260.
- Ito K, Iwatsubo T, Kanamitsu S, Ueda K, Suzuki H, and Sugiyama Y (1998) Prediction of pharmacokinetic alterations caused by drug-drug interactions: metabolic interaction in the liver. *Pharmacol Rev* **50**:387–411.
- Jolliffe IT (1986) *Principal Component Analysis*. Springer-Verlag, New York.
- Katoh M, Nakajima M, Shimada H, Yamazaki H, and Yokoi T (2000) Inhibition of human cytochrome P450 enzymes by 1,4-dihydropyridine calcium antagonists: prediction of in vivo drug-drug interactions. *Eur J Clin Pharmacol* **55**:843–852.
- Kohonen T (1989) *Self-Organization and Associative Memory*, 3rd ed. Springer-Verlag, Berlin.
- Korolev D, Balakin KV, Nikolsky Y, Kirillov E, Ivanenkov YA, Savchuk NP, Ivashchenko AA, and Nikolskaya T (2003) Modeling of human cytochrome P450-mediated drug metabolism using unsupervised machine learning approach. *J Med Chem* **46**:3631–3643.
- Korzekwa KR, Krishnamachary N, Shou M, Ogai A, Parise RA, Rettie AE, Gonzalez FJ, and Tracy TS (1998) Evaluation of atypical cytochrome P450 kinetics with two-substrate-models: evidence that multiple substrates can simultaneously bind to cytochrome P450 active sites. *Biochemistry* **37**:4137–4147.
- Lewis DF, Modi S, and Dickens M (2002) Structure-activity relationship for human cytochrome P450 substrates and inhibitors. *Drug Metab Rev* **34**:69–82.
- Lewis DFV (1996) *Cytochromes P450*. Taylor & Francis, London.
- Lewis DFV (2000) On the recognition of mammalian microsomal cytochrome P450 substrates and their characteristics. *Biochem Pharmacol* **60**:293–306.
- Lewis DFV, Eddershaw PJ, Dickens M, Tarbit MH, and Goldfarb PS (1998) Structural determinants of cytochrome P450 substrate specificity, binding affinity and catalytic rate. *Chem-Biol Interact* **115**:175–199.
- Marroum PJ, Uppoor RS, Parmelee T, Ajayi F, Burnett A, Yuan R, Svadjian R, Lesko LJ, and Balian JD (2000) In vivo drug-drug interaction studies—a survey of all new molecular entities approved from 1987–1997. *Clin Pharmacol Ther* **68**:280–285.
- Shimada J, Ekins S, Elkin C, Shaknovich EI, and Wery J-P (2002) Integrating computer-based de novo drug design and multidimensional filtering for desirable drugs. *Targets* **1**:196–205.
- Smith DA, Ackland MJ, and Jones BC (1997a) Properties of cytochrome P450 isoenzymes and their substrates. Part 1: Active site characteristics. *Drug Discovery Today* **2**:406–414.
- Smith DA, Ackland MJ, and Jones BC (1997b) Properties of cytochrome P450 isoenzymes and their substrates. Part 2: Properties of cytochrome P450 substrates. *Drug Discovery Today* **2**:479–486.
- Szklarz GD and Paulsen MD (2002) Molecular modeling of cytochrome P450 1A1: enzyme-substrate interactions and substrate binding affinities. *J Biomol Struct Dyn* **20**:155–162.
- Thummel KE and Wilkinson GR (1998) In vitro and in vivo drug interactions involving CYP3A. *Annu Rev Pharmacol Toxicol* **38**:389–430.
- Walters WP and Murcko MA (2002) Prediction of 'drug-likeness'. *Adv Drug Deliv Rev* **54**:255–271.
- Williams PA, Cosme J, Ward A, Angove HC, Matak Vinkovic D, and Jhoti H (2003) Crystal structure of human cytochrome P450 2C9 with bound warfarin. *Nature (Lond)* **424**:464–468.
- Yan Z and Caldwell GW (2001) Metabolism profiling and cytochrome P450 inhibition and induction in drug discovery. *Curr Top Med Chem* **1**:403–425.
- Yuan R, Parmelee T, Balian JD, Uppoor RS, Ajayi F, Burnett A, Lesko LJ, and Marroum PJ (1999) In vitro interaction studies: experience of the food and drug administration. *Clin Pharmacol Ther* **66**:9–15.
- Zhang W, Ramamoorthy Y, Kilicarslan T, Nolte H, Tyndale RF, and Sellers EM (2002) Inhibition of cytochromes P450 by antifungal imidazole derivatives. *Drug Metab Dispos* **30**:314–318.

Address correspondence to: Dr. Sean Ekins, Vice President, Computational Biology, GeneGo, 500 Renaissance Drive, Suite 106, St. Joseph, MI 49085. E-mail: sean@genego.com