Three- and Four-Dimensional Quantitative Structure Activity Relationship Analyses of Cytochrome P-450 3A4 Inhibitors

SEAN EKINS,1 GIANPAOLO BRAVI,2 SHELLY BINKLEY, JENNIFER S. GILLESPIE, BARBARA J. RING, JAMES H. WIKEL, and STEVEN A WRIGHTON


Accepted for publication February 17, 1999 This paper is available online at http://www.jpet.org

ABSTRACT

The program Catalyst was used to build three-dimensional quantitative structure activity relationship (3D-QSAR) pharmacophore models of the structural features common to competitive-type inhibitors of cytochrome P-450 (CYP) 3A4. These were compared with 3D- and four-dimensional (4D)-QSAR partial least-squares (PLS) models built using molecular surface-weighted holistic invariant molecular (MS-WHIM) descriptors for size and shape of the inhibitor. The Catalyst pharmacophore model generated from multiple conformers of competitive inhibitors of CYP3A4-mediated midazolam 1'-hydroxylation (n = 14) yielded a high correlation of observed and predicted Ki values of r = 0.91. Similarly, PLS MS-WHIM was used to produce 3D- and 4D-QSARs for this data set and produced models that were statistically predictable after cross-validation. Two additional Catalyst pharmacophores were constructed from literature Ki values (n = 32) derived from the inhibition of CYP3A-mediated cyclosporin A metabolism and IC50 data (n = 22) from the inhibition of CYP3A4-mediated quinine 3-hydroxylation. These Catalyst pharmacophores illustrated correlations of observed and predicted inhibition for CYP3A4 of r = 0.77 and 0.92, respectively. The corresponding 4D-QSARs generated by PLS MS-WHIM for these data sets were of comparable quality as judged by cross-validation. Both Ki pharmacophores generated with Catalyst were also validated by predicting the Ki(apparent) values of a test set of eight CYP3A4 inhibitors not included in either model. In seven of eight cases, the residuals of the predicted K_i(apparent) values were within 1 log unit of the observed values. The 3D- and 4D-QSAR models produced in this study suggest the utility of future in silico prediction of CYP3A4-mediated drug-drug interactions.

Techniques applicable for screening new chemical entities in terms of metabolism and toxicology (Wrighton et al., 1993) cover a wide range in terms of cellular integrity and cost. Refining the pharmaceutical screening process and accelerating drug discovery requires the use of these techniques in an earlier stage of drug development to identify the potential interaction of the new chemical entity with particular drug-metabolizing enzymes, namely cytochrome P-450s (CYPs). It is widely recognized that CYP is the major class of oxidative enzymes involved in drug metabolism (Wrighton and Stevens, 1992). To date, our understanding of the structural requirements for the binding of substrates and inhibitors, as well as metabolism by CYPs in vitro and in vivo, is limited (Smith et al., 1997a,b). Once these characteristics of CYPs are known, it will be possible to predict potential drug-drug interactions of a molecule before it reaches the later stages of drug development. These interactions are particularly important when multiple drugs are coadministered because drug-drug interactions are responsible for 1 to 2% of clinically relevant adverse drug reactions (Fuhr et al., 1996). Therefore, being able to screen for drugs likely to be potent or specific CYP inhibitors would be advantageous to the pharmaceutical industry.

CYP3A4 is particularly important in the metabolism of many classes of drugs (von Moltke et al., 1995), and from this point of view, it can be classified as the major human CYP involved in drug metabolism (Lewis et al., 1996). CYP3A4 is also expressed at the highest level relative to the other hepatic CYPs (Shimada et al., 1994). Because CYP3A4 is considered the major oxidative enzyme involved in drug metabolism, numerous drug-drug interactions have been reported, where the inhibition of this enzyme by a drug results in the decreased clearance of other drugs. Besides drugs, CYP3A4 can be inhibited by other common xenobiotics, such as flavonoids, which are consumed in large quantities in the hu-
man diet (Fukuda et al., 1997). A further consideration is the extensive intestinal expression of CYP3A4 and its inhibition by drugs and xenobiotics, including ketoconazole and troleandomycin (Raessi et al., 1997). Because humans are exposed to many potential CYP3A4 inhibitors as well as therapeutic agents that are metabolized by this enzyme, the likelihood of clinically relevant interactions, whether in the liver or intestine, is therefore substantial.

Due to the membrane-bound nature of mammalian CYPs, information is lacking on their crystal structures. Thus, computational methods are required to understand the structure and particularly the active site or sites of these enzymes. A number of differing three-dimensional (3D) quantitative structure activity relationships (3D-QSAR) techniques exist that have been widely used to infer active and/or binding site requirements for substrates and inhibitors of enzymes other than CYPs (Koehler et al., 1996). It is generally advantageous if the molecules used in such 3D-QSAR studies both cover a wide range in molecular structure and bioactivity (Grigov et al., 1997) to maximize conformational space and gain as much information as possible regarding the binding site. Using a combination of the relevant ligand conformers, the 3D distribution of structural features in relationship to the activities of the ligand defines a 3D pharmacophore. In the case of pharmacophore modeling drug-drug interactions for a particular enzyme, classically a $K_i$ value is determined (Bertz and Granneman, 1997) using enzymes where a large number of competitive inhibitors and an enzyme-selective assay are available. Because the $K_i$ value is a direct measure of the binding site properties of the enzyme (Nelsestuen and Martinez, 1997), this 3D-QSAR methodology has been successfully used to produce pharmacophore models and deduce requirements of the active sites of CYP2D6 (Strobl et al., 1993) and CYP2C9 (Jones et al., 1996b), but as yet, this technique has not been applied to CYP3A4.

The aim of the present study was to evaluate two computational approaches, Catalyst and partial least-squares (PLS) molecular surface-weighted holistic invariant molecular (MS-WHIM), for modeling CYP3A4 in vitro competitive inhibitor data from our drug interaction studies and those previously published. In addition, although the 3D-QSAR PLS MS-WHIM technique originally used the alignment of single conformations and distribution of atom types in the training set (Bravi et al., 1997), it has been modified to use multiple conformers and alignments of molecules (Bravi and Wikel, 1998a,b) and therefore may be used and classified as
a four-dimensional (4D)-QSAR (Klein and Hopfinger, 1998). The results of the 3D- and 4D-QSAR methods were compared to determine the most appropriate models for future in silico prediction of drug-drug interactions mediated by CYP3A4 inhibition.

**Experimental Procedures**

**Materials.** Flunitrazepam and NADPH were obtained from Sigma Chemical Co. (St. Louis, MO). Methanol and acetonitrile were obtained from Burdick and Jackson (Muskegan, MI). Midazolam and 1-hydroxy midazolam were gifts from Hoffman La Roche (Nutley, NJ).

**Liver Specimens.** Human liver specimens were obtained from the Liver Transplant Unit at the Medical College of Wisconsin and the Pathology Department of the Indiana School of Medicine, under protocols approved by the appropriate committee for the conduct of human research. Microsomes were prepared from these specimens using differential centrifugation (van der Hoeven and Coon, 1974).

**Midazolam 1-Hydroxylation.** Incubations containing human liver microsomes were carried out with 1 mM NADPH in 100 mM sodium phosphate, pH 7.4, and underwent a 3-min preincubation at 37°C. Using a modification of a method previously described (Kronbach et al., 1989), five different concentrations of midazolam (5–100 μM) were incubated with or without four different concentrations of inhibitor for 1 or 5 min and terminated by the addition of 200 μl of methanol and 10 μl of flunitrazepam (0.01 mg/ml). The vials were then placed on ice for approximately 10 min. After centrifugation, a 200-μl aliquot of supernatant was removed, and 50 μl was injected and analyzed by HPLC.

The isocratic HPLC system (Shimadzu, Columbia, MD) used a mobile phase consisting of 10 mM potassium phosphate/acetonitrile/methanol (45:21:34) and used a YMC Basic column (4.6 × 150 mm; YMC, Inc., Wilmington, NC) with a YMC Basic guard column. The flow rate was 1 ml/min, and the UV detector was set at a wavelength of 220 nm.

**Kinetic Analysis.** The Ki values for the inhibition of CYP3A4 were determined for our data set using nonlinear regression analysis as described in detail elsewhere (Ring et al., 1995).

**Molecular Modeling.** The computational molecular modeling studies were carried out using Silicon Graphics Indigo and Onyx workstations.

**Catalyst CYP3A4 Pharmacophore Models.** The 3D structures of inhibitors that competitively bind to the active site of CYP3A4 were built interactively using Catalyst Version 2.3 or 3.1 (Molecular Simulations, San Diego, CA). The number of conformers generated for each inhibitor was limited to a maximum of 255 with an energy range of 20 kcal/mol. Ten hypotheses were generated using these conformers for each of the molecules and the Ki values (Fig. 1, Table 1) after selection of the following features for the inhibitors: hydrogen bond donor, hydrogen bond acceptor, hydrophobic and negative ionizable. After assessment of all 10 hypotheses generated, the lowest energy cost hypothesis was considered the best because this possessed features representative of all the hypotheses. This procedure was repeated with data sets (Ki and IC50 values) from the literature.

The goodness of the structure activity correlation was estimated by means of r values. Statistical significance of the retrieved hypothesis was verified by permuting (randomizing) the response variable (i.e., the activities of the training set compounds were mixed a number of times so that each value was no longer assigned to the original molecule, and the Catalyst hypothesis generation procedure was repeated).

**Catalyst CYP3A4 Pharmacophore Validation Using a Test Set of K_{\text{apparent}} Values.** In an attempt to further validate the Catalyst pharmacophores, the models generated with our data set or that of Pichard et al. (1990, 1996) were used to predict the K_{\text{apparent}} values of a test set of eight molecules not included in the initial training sets. These were fit by aligning the pharmacophore features of the structures with the hypothesis deduced for each model. To predict a K_{\text{apparent}} value, two algorithms were used for this alignment: fast fit and best fit. “Fast fit” refers to the method of finding the optimum fit of the substrate to the hypothesis among all the conformers of the molecule without performing an energy minimization on the conformers of the molecule. The “best fit” procedure starts with fast fit and allows individual conformers to flex over an energy threshold of 20 kcal/mol. This allows examination of more conformational space and minimizes the distance between the hypothesis features and the atoms to which they map onto (Catalyst Tutorials Release 3.0; MSI, San Diego, CA.). The predictions using best and fast fit were compared by means of a residual that was calculated from the difference (in log units) between predicted and observed K_{\text{apparent}} values. A predicted K_{\text{apparent}} value within 1 log unit of the observed K_{\text{apparent}} value was considered to be a valid prediction of fit.

**3D- and 4D-QSAR Modeling with MS-WHIM and PLS.** Each molecule was coded into a SMILES (simplified molecular input line entry system) string format (Weininger, 1988). Atomic 3D coordinates and Gasteiger Huckel charges were generated by CONCORD 3.2.1 (CONCORD User's Manual; Tripos Inc., St. Louis, MO). MS-WHIM (Bravi et al., 1997; Bravi and Wikel, 1998a,b) descriptors are a set of statistical parameters that contain information about the structure of the molecules in terms of size, shape, symmetry, and distribution of molecular surface point coordinates after “weighted” centering and principal component analysis. The following weights are applied: 1) unweighted value, 2) positive and 3) negative electrostatic potential, 4) hydrogen bonding acceptor and 5) donor capacity, and 6) hydrophobicity, which yield a total of 72 descriptors (17 for each weight; see Bravi et al., 1997, for details on MS-WHIM calculation).

Table 1

<table>
<thead>
<tr>
<th>Inhibitor</th>
<th>K_{\text{apparent}} μM</th>
</tr>
</thead>
<tbody>
<tr>
<td>LY303870</td>
<td>1.8 ± 0.2</td>
</tr>
<tr>
<td>LY350965</td>
<td>2.2 ± 0.3</td>
</tr>
<tr>
<td>LY335979</td>
<td>3.8 ± 0.8</td>
</tr>
<tr>
<td>Terfenadine</td>
<td>4.7 ± 0.7</td>
</tr>
<tr>
<td>LY213628</td>
<td>6.6 ± 0.9</td>
</tr>
<tr>
<td>Nifedipine a</td>
<td>22.1 ± 2.4</td>
</tr>
<tr>
<td>LY334370</td>
<td>43.05 ± 3.7</td>
</tr>
<tr>
<td>Omeprazole</td>
<td>43.6 ± 5.7</td>
</tr>
<tr>
<td>LY307640 b</td>
<td>59.4 ± 6.0</td>
</tr>
<tr>
<td>Erythromycin a</td>
<td>194.1 ± 29.9</td>
</tr>
<tr>
<td>Cimetidine a</td>
<td>267.7 ± 20.4</td>
</tr>
<tr>
<td>LY277359 b</td>
<td>297.7 ± 55.6</td>
</tr>
<tr>
<td>LY237216</td>
<td>492.7 ± 72.3</td>
</tr>
<tr>
<td>LY024410</td>
<td>701.3 ± 106.3</td>
</tr>
</tbody>
</table>

a Wrighton and Ring (1994).

b VandenBraden et al. (1996).

c Ring et al. (1994).
formed as \( \log(1/K_i) \) or \( \log(1/IC_{50}) \). The optimum number of components in each PLS model generated was determined through two cross-validation procedures: 1) leave-one-out (LOO) and 2) five random groups repeated up to 100 times (5RG×100) (Baroni et al., 1993). Within the latter protocol, training set molecules are randomly assigned to five groups. Because results are strongly influenced by initial random assignment, the entire cross-validation procedure is repeated many times to achieve a meaningful statistical result. The predictive power of the PLS model was evaluated by means of \( q^2 \) and S.E.M. values of prediction (Baroni et al., 1993) computed as follows:

\[
q^2 = 1 - \frac{\sum(Y_{\text{observed}} - Y_{\text{predicted}})^2}{\sum(Y_{\text{observed}} - Y_{\text{mean}})^2}
\]

\[
\text{S.D. of error of prediction} = \left(\frac{\sum(Y_{\text{observed}} - Y_{\text{predicted}})^2}{n}\right)^{1/2}
\]

where \( n \) is the number of compounds. When using the 5RG×100 protocol, the reported \( q^2 \) and S.E.M. of prediction values represent mean values over 100 cross-validation cycles. The reliability of the PLS model was further verified by permutation of the response variable several times (as described for Catalyst) and repeat of cross-validation. MS-WHIM weight selection (Bravi and Wikel, 1998a,b) was carried out in attempt to improve the PLS models and to determine the most relevant contributions to structure activity correlation.

**Results**

**Catalyst CYP3A4 Pharmacophore Models.** Catalyst uses a collection of molecules with inhibitory activity spanning several orders of magnitude for the enzyme to construct a model of the chemical features necessary for competitive inhibition. The resultant hypothesis then explains the variability of the potency of inhibition with respect to the geometric localization of these pharmacophore features present in the molecules used to build them. In our first data set, the \( K_i \) values for the inhibition of midazolam 1'-hydroxylase cover 2 orders of magnitude (Table 1, Fig. 1). The lowest energy pharmacophore produced four features necessary for the inhibition of CYP3A4, namely, three hydrophobes at distances of 5.2 to 8.8 Å from a hydrogen bond acceptor (Fig. 2 and inset). The most potent inhibitor in this data set LY303870 was fit to all four features of this pharmacophore (Fig. 2). The pharmacophore demonstrated an excellent correlation of observed versus estimated \( K_i(\text{apparent}) \) values (\( r = 0.91 \), Table 2).

Previously published data derived in vitro represent a useful resource for CYP interaction data that may be modeled in silico. In the current study, we used literature \( K_i \) values obtained from studies on the inhibition of CYP3A4-mediated cyclosporin A metabolism that covers 3 orders of magnitude (Pichard et al., 1990, 1996) (Table 2). The lowest energy pharmacophore derived from the data from Pichard et al. (1990, 1996) produced five features for the inhibition of CYP3A4, namely, three hydrophobes at distances of 4.2 to 7.1 Å from a hydrogen bond acceptor and an additional 5.2 Å from another hydrogen bond acceptor (Fig. 3 and inset). The most potent inhibitor in this data set, ketoconazole, was fit to all five features (Fig. 3). The Catalyst pharmacophore demonstrated a reasonable correlation of observed versus estimated \( K_i(\text{apparent}) \) (\( r = 0.77 \), Table 2).

A second literature-derived data set was used in which IC\(_{50}\) values obtained from the CYP3A4-mediated quinine hydroxylation covered 4 orders of magnitude (Zhao and Ishizaki, 1997) (Table 2). The lowest energy pharmacophore produced four features necessary for inhibition of CYP3A4, namely, one hydrophobe at distances from 8.1 to 16.3 Å from the two farthest of three hydrogen bond acceptors (Fig. 4 and inset). The most potent inhibitor in this data set, ketocon-
azole, was fit to all five features (Fig. 4). The Catalyst pharmacophore demonstrated an excellent correlation of observed versus estimated IC50 ($r = 0.92$, Table 2).

The total energy cost of the generated pharmacophores can be calculated from the deviation between the estimated activity and the observed activity, combined with the complexity of the hypothesis (i.e., the number of pharmacophore features). A null hypothesis can also be calculated that presumes there is no relationship in the data and that experimental activities are normally distributed about their mean. Hence, the greater the difference between the energy cost of the generated hypothesis and the energy cost of the null hypothesis, the less likely it is that the hypothesis reflects a chance correlation. After permuting the initial activity values and structures for the CYP3A4 $K_i$ pharmacophore (derived from midazolam 19-hydroxylase data) twice, there was little difference between the energy cost of the hypothesis and the null hypothesis (Table 3). In addition, for the model, the pharmacophore features had changed, and the difference between the energy cost of the hypothesis and the null hypothesis had decreased (0.5 and 1.3 units, Table 3) compared with the original data (4.4 units, Table 2). Both the literature-derived CYP3A4 $K_i$ and IC50 pharmacophores possessed a larger difference versus the midazolam $K_i$ data, between total cost and null hypothesis, 9.7 and 14 units, respectively. These differences between total cost and the null hypothesis decreased on permutation, suggesting an alteration of the model (Tables 2 and 3) that was confirmed by a change in the number of pharmacophore features. However, the fit ($r$) did not change significantly in both of these cases (Table 3).

**Catalyst CYP3A4 Pharmacophore Validation Using a Test Set of $K_i$(apparent) Values.** After construction of the Catalyst 3D-QSAR models for our $K_i$(apparent) data set and that obtained from the literature (Pichard et al., 1990, 1996), we used the models to predict $K_i$(apparent) values of a test set of eight molecules excluded from the training set of $K_i$(apparent) (Table 4). The predicted $K_i$(apparent) values of 4-[[6-hydroxy-7-[[1-[(1-hydroxy-1-methyl)ethyl]-4-methyl-6-(7-oxo-7H-furo[3,2-g][1]benzopyran-4-yl]-4-hexenyl]oxy]-3,7-dimethyl-2-octenyl]oxy]-7H-furo[3,2-g][1]benzopyran-7-one (GF-I-4), 4-[[6-hydroxy-7-[[4-methyl-1-(1-methylethenyl)ethyl]-6-(7-oxo-7H-furo[3,2-g][1]benzopyran-4-yl]-4-hexenyl]oxy]-3,7-dimethyl-2-octenyl]oxy]-7H-furo[3,2-g][1]benzopyran-7-one (GF-I-1), N-desmethyl diltiazem, N,N-didesmethyl diltiazem, quinine, ritonavir, saquinavir, and indinavir were then compared with their observed literature values generated in...
human liver microsomes. The $K_i$ values of all eight molecules were better predicted using the best fit function for the Catalyst pharmacophore of our data than with the fast fit function. Only the $N,N$-didesmethyl diltiazem $K_i$ value was poorly predicted as the residual of predicted and observed $K_i$ values were greater than the stated cutoff of 1 log unit necessary for acceptability. For the model from the published $K_i$ data set, all eight molecules were also better predicted using the best fit method, although the $N,N$-didesmethyl diltiazem $K_i$ value was outside the 1 log residual cutoff. When ritonavir, saquinavir, and indinavir are excluded from the analysis, the fast fit function for the literature $K_i$ model and the best fit function for our data correctly rank ordered the predictions for the five remaining inhibitors.

### Table 3

Summary of permuted Catalyst CYP3A4 hypotheses

<table>
<thead>
<tr>
<th>CYP</th>
<th>Permutation Hypothesis 1</th>
<th>Permutation Hypothesis 2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pharmacophore features</td>
<td>Hypothesis-null energy cost</td>
</tr>
<tr>
<td>3A4*</td>
<td>1 HBA, 1 HBD, 1 H</td>
<td>69.9–70.4</td>
</tr>
<tr>
<td>3A4*</td>
<td>1 HBA, 2 HBD, 1 H</td>
<td>145.5–149.7</td>
</tr>
<tr>
<td>3A4*</td>
<td>2 HBA, 1 HBD, 1 H</td>
<td>104.6–110.1</td>
</tr>
</tbody>
</table>

*H, hydrophobe; HBA, hydrogen bond acceptor; HBD, hydrogen bond donor.

*Present study.


*Zhao et al. (1997).

### Table 4

Observed and predicted $K_{i\text{apparent}}$ for inhibitors of CYP3A4 fit to different Catalyst pharmacophores

<table>
<thead>
<tr>
<th>Inhibitor</th>
<th>Reference</th>
<th>Observed Human Liver Microsomal $K_{i\text{apparent}}$</th>
<th>CYP3A4 Catalyst Prediction Fast Fit $K_{i\text{apparent}}$</th>
<th>CYP3A4 Catalyst Prediction Best Fit $K_{i\text{apparent}}$</th>
<th>CYP3A4 Catalyst Prediction Fast Fit $K_{i\text{apparent}}$</th>
<th>CYP3A4 Catalyst Prediction Best Fit $K_{i\text{apparent}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Quinine</td>
<td>Zhang et al. (1997)</td>
<td>11</td>
<td>280 (1.8)</td>
<td>12 (0.03)</td>
<td>72 (0.8)</td>
<td>2.7 (0.6)</td>
</tr>
<tr>
<td>GF-I-4</td>
<td>Fukuda et al. (1997)</td>
<td>0.029</td>
<td>5.5 (2.3)</td>
<td>0.031 (0.03)</td>
<td>3.8 (2.1)</td>
<td>0.037 (0.1)</td>
</tr>
<tr>
<td>GF-I-1</td>
<td>Fukuda et al. (1997)</td>
<td>0.05</td>
<td>5.8 (2.1)</td>
<td>0.034 (0.2)</td>
<td>8.4 (2.2)</td>
<td>0.26 (0.7)</td>
</tr>
<tr>
<td>N-Desmethyl diltiazem</td>
<td>Sutton et al. (1997)</td>
<td>2</td>
<td>930 (2.7)</td>
<td>6.3 (0.5)</td>
<td>44 (1.3)</td>
<td>2.6 (0.1)</td>
</tr>
<tr>
<td>N,N-Desmethyl diltiazem</td>
<td>Sutton et al. (1997)</td>
<td>0.1</td>
<td>930 (4.0)</td>
<td>5.8 (1.8)</td>
<td>35 (2.8)</td>
<td>2.8 (1.4)</td>
</tr>
<tr>
<td>Ritonavir</td>
<td>Eagling et al. (1997)</td>
<td>0.019</td>
<td>7.8 (2.6)</td>
<td>0.036 (0.3)</td>
<td>9.4 (2.7)</td>
<td>0.059 (0.5)</td>
</tr>
<tr>
<td>Saquinavir</td>
<td>Eagling et al. (1997)</td>
<td>2.99</td>
<td>29 (1.0)</td>
<td>1.3 (0.4)</td>
<td>68 (1.4)</td>
<td>0.91 (0.5)</td>
</tr>
<tr>
<td>Indinavir</td>
<td>Eagling et al. (1997)</td>
<td>0.17</td>
<td>7 (1.6)</td>
<td>0.16 (0.02)</td>
<td>18 (2.2)</td>
<td>0.038 (0.6)</td>
</tr>
</tbody>
</table>

*Present study.


Values in parentheses represent residual (log units) of predicted − observed $K_{i\text{apparent}}$ values.
tors. Also, these three protease inhibitors are correctly rank ordered by the same fitting functions for each $K_i$ model ($K_i$ for saquinavir > indinavir > ritonavir).

**PLS MS-WHIM 3D- and 4D-QSAR Models of CYP3A4**

**Data.** MS-WHIM is a QSAR modeling approach that captures the molecules physicochemical properties at the molecular surface level and then condenses this information into a numerical vector. This is a useful technique because it is suggested that enzymes and ligands will recognize each other at their molecular surface, so the binding of inhibitors with CYPs appears to be dependent on this type of interaction. MS-WHIM has previously been used successfully to study structure property and structure activity problems (Bravi et al., 1997; Bravi and Wikel, 1998a,b). One of the test sets used for one of these previous studies (Bravi and Wikel, 1998a) was inhibitor data for CYP2A5 derived from a previous report (Poso et al., 1995).

MS-WHIM autoscaled descriptors were calculated for one conformer of all inhibitors (3D-QSAR) in each CYP3A4 inhibitor data set (Bravi et al., 1997); alternatively, multiple conformers (4D-QSAR) were generated and then correlated by means of PLS with $K_i$ data expressed as $\log(1/K_i)$ or $IC_{50}$ data expressed as $\log(1/IC_{50})$. In some cases, inhibitors were excluded from analysis if the molecular size or flexibility exceeded the parameters imposed by PLS MS-WHIM. The models were then cross-validated using the LOO and 5RG×100 protocols to yield their respective $q^2$ values (Table 5) as described in *Experimental Procedures*. Cramer et al. (1993) suggested that a $q^2$ value of more than 0.3 indicates that the model is meaningful (predictive). Due to the nature of this 3D-QSAR approach, there is no graphic output, and the models are therefore solely assessed and validated by the $q^2$ value.

The 3D-QSAR using PLS MS-WHIM of the CYP3A4-mediated midazolam 1'-hydroxylation was shown to be predictive, yielding a LOO $q^2$ value of 0.32 (Table 5) for the molecular surface weights: unweighted and positive potential. 4D-QSAR using multiple conformers of inhibitors (Constatf) with MS-WHIM PL5s demonstrated an improvement on the 3D-QSAR models as the LOO $q^2$ value increased to 0.44 (Table 5). 5RG cross-validation did not supply a significant model for this data. Modeling previously published data sets with PLS MS-WHIM allows us to test these QSAR techniques and make comparisons with the previously described Catalyst models, as well as with the models for CYP3A4 generated with our $K_i$ data for inhibition of midazolam 1'-hydroxylation. No significant 3D-QSAR model could be generated for the published CYP3A4 $K_i$ data (Pichard et al., 1990, 1996). Although 4D-QSAR of these inhibitors resulted in a LOO $q^2$ value of 0.41 with two components and the molecular surface weights: unweighted, positive electrostatic potential, hydrogen bond donor, and hydrophobicity. The more credible cross-validation procedure, 5RG, resulted in a predictive $q^2$ value of 0.38 (Table 5). Similarly, no significant 3D-QSAR model could be generated for the CYP3A4 $IC_{50}$ data (Zhao and Ishizaki 1997), although a 4D-QSAR of these inhibitors generated a predictive model with a LOO $q^2$ value of 0.56 with six components and the molecular surface property of positive electrostatic potential (Table 5). 5RG cross-validation did not provide a significant prediction for this $IC_{50}$ data set. In all three cases modeled with PLS MS-WHIM less than the full complement of structures in the data set initially used for the Catalyst modeling was used for building the PLS MS-WHIM models. Random permutation of the response variable a number of times [so that the $K_i$(apparent) or $IC_{50}$ values no longer corresponded to the same molecules] was adopted to test the validity of the obtained models because this procedure should not result in a predictive model. In each case, a $q^2$ of zero or lower was obtained, indicating that permutation of the response variable does not result in a predictive PLS MS-WHIM model. This approach confirms that the cross-validated PLS MS-WHIM 3D-QSAR models obtained for CYP3A4 are statistically valid.

**Discussion**

Recent reviews have described potential active site characteristics and physicochemical properties of substrates for CYP1A2, CYP2C9, CYP2C19, CYP2D6, CYP2E1, and CYP3A4 obtained from 3D-QSAR, protein homology modeling, NMR, and site-directed mutagenesis techniques (Smith et al., 1997a,b). It is important to note that the use of computational 3D-QSAR techniques as applied to CYPs is still in the very early stages of development. The aim of the present study was to compare different techniques for generating 3D- and 4D-QSARs of CYP3A4 inhibitors in silico.

Hepatic and intestinal CYP3A4 catalyze a wide range of reactions, making this an important enzyme involved in the metabolism of xenobiotic agents and most pharmaceutical agents (Lewis and Lake, 1996). To date, the active site of CYP3A4 has only been modeled from the point of view of docking of numerous substrates (Lewis et al., 1996). CYP3A4 inhibitors have not been extensively studied from a computational modeling approach. However, the inhibitors ketoconazole and gestodene have been docked into the CYP3A4 active site homology modeled on the basis of crystallized bacterial CYPBM-3 (Lewis et al., 1996). This enabled identification of hydrogen bonding of the Asn74 residue of CYP3A4 with these two inhibitors as an important factor in binding (Lewis et al., 1996), which was also found to occur with CYP3A4 substrates (Lewis and Lake, 1998). Structural requirements of CYP3A4 substrates have been suggested to include a hydrogen bond acceptor atom 5.5 to 7.8 Å from the site of metabolism and 3Å from the oxygen molecule associated with the heme (Lewis et al., 1996).

<table>
<thead>
<tr>
<th>TABLE 5</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>MS-WHIM PLS results for CYP3A4 data sets</strong></td>
</tr>
<tr>
<td>Model</td>
</tr>
<tr>
<td>No. of molecules in model&lt;sup&gt;d&lt;/sup&gt;</td>
</tr>
<tr>
<td>Molecular surface weights</td>
</tr>
<tr>
<td>LOO ($q^2$) (3D-QSAR)</td>
</tr>
<tr>
<td>LOO SDEP&lt;sup&gt;f&lt;/sup&gt;</td>
</tr>
<tr>
<td>LOO ($q^2$) Constat&lt;sup&gt;e&lt;/sup&gt; (4D-QSAR)</td>
</tr>
<tr>
<td>LOO Constat SDEP&lt;sup&gt;s&lt;/sup&gt;</td>
</tr>
<tr>
<td>5RG&lt;sup&gt;g&lt;/sup&gt; ($q^2$) Constat (SD)</td>
</tr>
<tr>
<td>Components</td>
</tr>
</tbody>
</table>

<sup>a</sup> Present study.
<sup>c</sup> Zhao et al. (1997).
<sup>d</sup> Number of molecules used for this model of the total number used in the original model.
<sup>e</sup> Molecular surface weights: 1, unitary; 2, positive potential; 3, negative potential;
<sup>f</sup> SDEP, standard deviation estimate predictions.
<sup>g</sup> Constat, multiple conformers.
<sup>s</sup> 5RG, five random groups repeated up to 100 times cross-validation.
<sup>t</sup> Not a significant value according to the published criteria (Cramer et al., 1993).
A CYP3A4 Catalyst model was constructed from a data set generated within our laboratory that contained 14 competitive inhibitor $K_{i(apparent)}$ values for midazolam 1'-hydroxylase activity. To test our limited understanding of requirements for CYP3A4 inhibitors, we have also used previously published inhibition $K_i$ and IC$_{50}$ values to construct 3D- and 4D-QSAR models. In the first instance, by combination of the $K_i$ data for two studies using cyclosporin A as the CYP3A4-selective catalytic probe (Pichard et al., 1990, 1996), we were able to generate a Catalyst model. One approach to test the usefulness of this technique is to examine the difference between the hypothesis cost and the null hypothesis cost before and after permutation of the activity values with the molecular structures. The use of this procedure with our data resulted in a low hypothesis-null cost difference (Table 2). After permutation of these data, the hypothesis-null cost difference decreased further along with the fit value (Table 3), which is indicative of a less significant model. For the literature $K_i$ data, this cost and the null hypothesis difference was very small and did not change remarkably after permutation (Table 3). This observation is similar to that observed with our CYP2B6 substrate pharmacophore (Ekins et al., 1999) and suggests too small an activity range. Although the CYP3A4 Catalyst model constructed from IC$_{50}$ data for the 3-hydroxylation of quinine (Zhao and Ishizaki, 1997) possessed a hypothesis-null cost difference that was slightly larger than that observed with the other $K_{i(apparent)}$ CYP3A4 models, it, too, was unchanged by permutation. Because Catalyst is still able to construct pharmacophores with high fit values ($r$) after permutation of these data sets, we believe this question the validity of either of these Catalyst two pharmacophores as assessed using this technique.

A further test of the validity of 3D-QSAR models is to predict the activity of molecules excluded from the training set and then to compare these values with those observed by means of a log residual. In the current study, we selected eight $K_{i(apparent)}$ values obtained from the literature and generated using well documented CYP3A4 assays. Both of the CYP3A4 $K_{i(apparent)}$ Catalyst models predicted the $K_i$ values similarly. The lowest log residuals for predictions of $K_{i(apparent)}$ were obtained using the best fit algorithm compared with the fast fit algorithm. Seven of eight of these best fit predictions were within 1 log unit residual for both models. The largest residual of all of the predictions was demonstrated for N,N-didesmethyl diltiazem, which was predicted to have a $K_{i(apparent)}$ value similar to that of N-desmethyl diltiazem. The failure to predict the $K_{i(apparent)}$ value of N,N-didesmethyl diltiazem appears to be related to the fact that, generally, best fit examines the ability of a higher energy conformer to interact with CYP3A4. This conformer may not represent the physiologically relevant conformation due to unfavorable van der Waals interactions within the enzyme active site. In this study, we believe it is more important to generate predictions closer to the observed value to validate the predictive ability of Catalyst models, although we have also observed the correct rank ordering of subsets of molecules (i.e., protease inhibitors) within this test set. It is interesting to note that Catalyst is able to successfully extrapolate from both data sets in our studies (Table 1) to predict potent inhibitors in the test set like GF-I-4 (Table 4). The variance of the test set of approximately 3 orders of magnitude ($-1.72$ to $1.30$ log units) was not concentrated around the mean value ($-0.24$ log unit). Instead, this mean value was considerably lower than the training set range ($0.25$-$2.84$ log units) and also lower than the training set mean value of $1.55$. We are therefore satisfied that the test set is sufficiently different from the training set to provide a useful validation exercise for the CYP3A4 Catalyst inhibitor pharmacophores. Even though the literature $K_i$ model is not valid according to the hypothesis-null energy cost criteria, it still can be used to make useful predictions for the test set.

The pharmacophore of our data (Fig. 2) was spatially in agreement with the IC$_{50}$ model (Fig. 4). In fact, all the CYP3A4 pharmacophores contained hydrophobic features $4.2$ to $8.8$ Å from at least one hydrogen bond acceptor feature (see pharmacophore figure insets), which is in agreement with the findings of Lewis et al. (1996) for CYP3A4 substrates. All of the Catalyst CYP3A4 inhibitor models contain pharmacophore features that cover a wide area, a likely characteristic of the larger molecules incorporated in these data sets compared with inhibitors of other CYPs (Smith et al., 1997a). There was some overlap in the molecules included in each inhibitor data set. Most notably, nifedipine, omeprazole, and erythromycin were present in all three models. By comparing both $K_{i(apparent)}$ pharmacophores, more detail regarding the features necessary for inhibition should be obtained. In this case, the two hydrophobes $5.2$ and $7$ Å from the hydrogen bond acceptor in our $K_i$ pharmacophore (Fig. 2, inset) would overlap with the three hydrophobes indicated in the model derived from the data of Pichard et al. (1990, 1996) (Fig. 3, inset). This results in two hydrophobic regions separated by hydrogen bond acceptors, suggesting similar domains and sites of hydrogen bond donation in the CYP3A4 active site.

Eleven of the 14 inhibitors of midazolam 1'-hydroxylase activity (after excluding erythromycin, LY237216, and LY024410 due to excessive molecular size or flexibility) were used for PLS MS-WHIM model construction. The 3D-QSAR for a single low energy conformer of each inhibitor (generated from Catalyst) illustrated a low LOO $q^2$ value of $0.32$ (Table 5), according to published criteria for the $q^2$ value (Cramer et al., 1993). The $q^2$ value improved to $0.44$ with the use of multiple conformers to produce a more predictive 4D-QSAR. The two literature data sets described previously used $25$ of $32$ (excluding FK506, cyclosporin G, midecamycin, triacetyleandomycin, josamycin, erythromycin, and virginiamycin) and $19$ of $22$ (excluding oleandomycin, triacetyleandomycin, and erythromycin) inhibitors for cyclosporin A metabolism and quinine hydroxylation, respectively. The suboptimal use of molecules from these data sets is a disadvantage of this technique because the molecule size or flexibility exceeds the parameters imposed by PLS MS-WHIM. In both of the literature data sets, LOO $q^2$ values generated by 4D-QSAR were superior to those generated for 3D-QSAR, in that in some cases, the $q^2$ for the 3D models were not more than $0.3$. This illustrates the advantage of assessing more than one conformer of a molecule to cover more conformational space in the model.

In the present study, the Catalyst 3D-QSAR $K_{i(apparent)}$ pharmacophore constructed with the midazolam 1'-hydroxylase data illustrated the predictive ability of the $K_{i(apparent)}$ database and appeared to be comparable with inhibitor CYP3A4 QSAR models built in this study using previously...
published data sets. As shown in this study, pharmacophores generated in silico with Catalyst appear to be useful in enabling analysis of whether a molecule may be a competitive inhibitor of CYP3A4, comparable to conventional in vitro methods. Minor differences between each CYP3A4 model could be explained by structurally diverse data sets and the low orders of magnitude for the $K_i^{(apparent)}$ in our data set compared with the literature. In the meantime, the three separate pharmacophores generated with Catalyst agreed the described CYP3A4 substrate template (Lewis et al., 1996; Lewis and Lake, 1998). Our results also suggest two pharmacophores based on the compact nature of the model derived from the 32 molecules used by Pichard et al. (Fig. 4), which fits within the area occupied by the features in our model produced from 14 inhibitors (Fig. 5).

Currently, the use of in vitro data to predict in vivo drug-drug interactions is receiving a great deal of attention as various groups test their predictions obtained with new chemical entities (Lin and Lu, 1997). It has been suggested as unnecessary to provide an exact quantitative prediction of drug interaction because classification into a category of inhibition is more useful (Fuhr et al., 1996). Our group (Wrighton et al., 1995) and others (Lin, 1997) have suggested that to make good predictions of drug interaction potential, we should understand the underlying mechanisms of inhibition by a drug, the metabolic fate of the drug, the enzymes involved in each metabolic pathway, and the concentration of the drug at the enzyme. We have described 3D- and 4D-QSAR model building approaches for CYPs that may allow qualitative or quantitative prediction of kinetic parameters, substrate affinity, or inhibitory potential long before the drug is placed in an in vitro model. For example, a recent comparative molecular field analysis 3D-QSAR model was used to predict Clint of volatile organic compounds in the rat (Waller et al., 1996), and models like this may be similarly applied to predicting the human Clint of drugs. We suggest that by using substrate- and inhibitor-related hypotheses with a tool such as Catalyst that is capable of searching 3D databases (Kaminski et al., 1997), it will be possible to identify potent substrates and inhibitors of each CYP. After the types of validation described in the present study, it will eventually become viable to screen molecular databases in parallel using 3D- or 4D-QSAR models for both inhibitory and affinity potential for CYPs or the other enzymes and transporters involved in human hepatic drug metabolism.

It is intriguing to speculate that eventually, using the strategy described above, in silico analysis of drug metabolism and drug interactions will preempt the presently accepted discovery process to economically guide molecular design. This could be pertinent if the potential for interaction with drug-metabolizing enzymes is important for a particular class of compounds. The ultimate goal of this research will be to minimize potential drug-drug interactions by using this high-throughput method for rational selection of molecules with a low enzyme interaction profile. The use of 3D-QSAR in combination with other allied modeling and conventional techniques will be invaluable for defining the active site of the enzyme as we await the crystallization and structural elucidation of a membrane-bound mammalian CYP. Because CYP 3D- and 4D-QSAR models indirectly provide information regarding active site characteristics that can be ratified with the homology modeled proteins, they may also be the ultimate reductionist tool for examining the details of CYP function that remain controversial, such as activation and autoactivation (Ekins et al., 1998; Korzekwa et al., 1998).

Acknowledgments

We gratefully acknowledge Dr. David Cummins for his help in the implementation of the F test in the PLS algorithm, Robert Coner for assistance with the permutation algorithm, and Dr. Patrick J. Murphy for his initial encouragement in pursuing this direction.

References


Send reprint requests to: Steven A. Wrighton, Ph.D., Department of Drug Disposition, Lilly Research Laboratories, Eli Lilly and Co., Lilly Corporate Center, Drop Code 0825, Indianapolis, IN 46285-0001.