In Silico Modeling of Non-specific Binding to Human Liver Microsomes

Hua Gao, Lili Yao, Heather W. Mathieu, Ying Zhang, Tristan S. Maurer, Matthew D. Troutman, Dennis O. Scott, Roger B. Ruggeri and Jing Lin*

Department of Chemistry (HG and RBR)
Department of Pharmacokinetics, Dynamics and Metabolism (LY, HWM, YZ, TSM, MDT, DOS and JL)
Pfizer Global Research and Development, Groton, Connecticut 06340, USA.
RUNNING TITLE: In Silico Microsomal Binding Model

**Corresponding author:** Jing Lin, Department of Pharmacokinetics, Dynamics and Metabolism, MS8220-4018, Groton, CT 06340, USA. Tel: 860-715-5180  Fax: 860-686-1060  E-mail: jing.lin@pfizer.com

The number of text pages: 12

Number of Tables: 1

Number of figures: 6

Number of references: 27

Number of words in Abstract: 208

Number of words in Introduction: 621

Number of words in Discussion: 691

**Abbreviations:**

ADMET: Absorption, Disposition, Metabolism, Excretion and Toxicity

QSAR: Quantitative Structure Activity Relationship

RMSE: Root Mean Square Error

CV: Coefficient Variation

NN: Nearest Neighbors
ABSTRACT

Estimation of unbound fraction of substrate in microsomal incubation media is important in accurately predicting hepatic intrinsic clearance and drug-drug interactions. In this study, the unbound fraction of 1223 drug-like molecules in human liver microsomal incubation media has been determined using equilibrium dialysis. These compounds including 27 marketed drug molecules cover a much broader range of physiochemical properties such as hydrophobicity, molecular weight, ionization state, and degree of binding than examined in previous work. In developing the in silico model, we have used two-dimensional molecular descriptors including cLogP, Kier connectivity, shape, and E-state indices, a subset of MOE descriptors, and a set of ADME structural keys used for our in-house ADMET modeling. Hydrophobicity is the most important molecular property contributing to the non-specific binding of substrate to microsomes. The prediction accuracy of the model is validated using a subset of 100 compounds, and 92% of the variance is accounted for by the model with a RMSE of 0.10. For the training set of compounds, 99% of variance is accounted for by the model with a RMSE of 0.02. The performance of the developed model has been further tested using the 27 marketed drug molecules with a RMSE of 0.10 between the observed and the predicted unbound fraction values.
INTRODUCTION

In drug discovery, the ultimate goal of a medicinal chemist is to rationally design compounds that are safe and efficacious at a marketable dose regimen. Such chemical design efforts frequently rely upon intrinsic clearance (CL_{int}) and CYP P450 inhibitory potency (IC_{50}) estimates derived from human liver microsomal incubations. With relatively few assumptions, these estimates can be expected to have proportional implications for oral dose requirements and therapeutic windows against pharmacokinetic drug-drug interactions. For example, the classical well stirred model of hepatic clearance can be used to illustrate the directly proportional relationship between CL_{int} and oral dose by eq. 1 (Wagner et al., 1968). Other oral dose determinants include the unbound steady-state average concentration necessary for the desired pharmacologic response (C_{ss,u}), dose interval (\tau) and fraction of dose absorbed (fa).

\[
\text{Oral dose} = \frac{C_{ss,u} \cdot CL_{int} \cdot \tau}{fa} \quad (1)
\]

Additionally, assuming inhibition of CYP P450 occurs via a competitive mechanism, equation 2 is a valid and commonly used approach to calculate the therapeutic window against pharmacokinetic drug-drug interactions.

\[
\text{Tx window} = \frac{C_{ss,u}}{IC_{50}} \quad (2)
\]

It is now well established in the scientific literature that nonspecific binding of compounds into microsomal phospholipids will confound the experimental estimation of CL_{int} and IC_{50} by decreasing the fraction of unbound drug (fu_{mic}) to metabolizing enzymes within the incubation (Kalvass et al., 2001; Margolis and Obach, 2003; Obach, 1997 and 1999; Tran et al., 2002). In such cases, estimates of CL_{int} and IC_{50} derived from microsomes are best deemed “apparent”
(i.e. CL_{int,app} and IC_{50,app}) since they represent a mixture of kinetic determinants with both relevant (CL_{int} and IC_{50}) and irrelevant (fu_{mic}) implications for drug design by eq. 3 and 4 (Obach, 1996).

\[
CL_{int,app} = CL_{int} \cdot fu_{mic}
\]  \hspace{1cm} (3)

\[
IC_{50,app} = \frac{IC_{50}}{fu_{mic}}
\]  \hspace{1cm} (4)

From these relationships, it follows that design efforts based simply upon CL_{int,app} and IC_{50,app} values will be misguided in the presence of significant nonspecific binding (i.e. fu_{mic} \ll 1). This can be particularly damaging considering the general influence of lipophilicity over CL_{int} (increases), IC_{50} (decreases) and fu_{mic} (decreases). As such, it is important to note that microsomal binding may mask worsening values of CL_{int} and IC_{50} and entice design efforts into increasingly lipophilic space. It is also important to note that extensive nonspecific microsomal binding is common among both marketed drugs and chemical libraries of the pharmaceutical industry, particularly among lipophilic chemical entities. As such, it is essential for these kinetic parameters to be corrected for fu_{mic} in order to ensure effective design in drug discovery (Obach, 1996; Bjornsson et al., 2003; Grime and Riley, 2006; Tucker et al., 2001). Unfortunately, experimental methods for the determination of fu_{mic} (e.g. equilibrium dialysis and ultracentrifugation) are far less amenable to high throughput screening than that microsomal stability incubations, thereby limiting the flow of information available to support design efforts. The need for this information has led to the development of several in silico methods for predicting fu_{mic} from physicochemical properties (most notably, LogD and LogP) (Austin et al., 2002; Austin et al., 2005; Hallifax and Houston, 2006; Sykes et al., 2006). Although the number of compounds used to develop these published models is limited (e.g., 56) and the data
are gathered from a variety of sources (e.g. different laboratories, conditions, species), the results indicate that $f_{\text{mic}}$ is highly amenable to prediction via physiochemical properties. This is perhaps not surprising, given that $f_{\text{mic}}$ likely represents a passive partitioning into microsomal phosphopholipds. In this work, we developed a novel \textit{in silico} model using Cubist with human $f_{\text{mic}}$ data generated in our laboratory for 1223 compounds including 27 commercial available chemicals via a uniform 96-well equilibrium dialysis methodology. In addition to a relatively larger number of compounds, this data set encompasses a much broader range of physiochemical properties such as lipophilicity (LogP and LogD), molecular weight (MW), ionization state (pKa), and degree of binding ($\log \left[ f_{\text{mic}}/(1-f_{\text{mic}}) \right]$) than examined in previous work.

\textbf{Methods}

\textbf{Materials.} Human liver microsome mix (HL-Mix-101) was generated from a pool of 59 individual livers and was obtained from BD Biosciences Gentest (Woburn, MA). The protein content of the microsome mix was 20.4 mg/ml and the Cytochrome P450 concentration was determined to be 0.33 nmol/mg. The same batch of microsomes was used for all the studies and was stored -80°C. Albendazole, amitriptyline, caramazepine, chloroimpramine, chlorpromazine, clozapine, colchicine, desipramine, diazepam, erythromycin, imipramine, indomethacin, keoprofen, ketoconazole, midazolam, pimozone, piroxicam, prednisone, promethazine, propranolol, quinidine, sulindac, tenidap, tolbutamide, trimeprazine, verapamil and warfarin were obtained from the Pfizer compound collection (Table 1). All the Pfizer proprietary compounds were obtained from Pfizer internal sample bank (Groton, CT). A 96-well dialysis apparatus and cellulose dialysis membrane strips with a molecular weight cut off at 12-
14 kDa were purchased from HTDialysis (Gales Ferry, CT). Breathe Easy gas permeable sealing membranes were obtained from Diversified Biotech (Boston, MA).

**Experimental Methods:**

**Microsomal Binding Incubations:** Microsomal binding in human liver microsomes was performed in triplicate using a 96 well equilibrium dialysis and adapted from a previous described method (Banker et al., 2003).

Each dialysis membrane strip was conditioned sequentially for 15 minutes in deionized water, 30% ethanol and then 0.1 M potassium phosphate buffer, pH 7.4. Conditioned membrane strip was then loaded into a 96-well dialysis apparatus. Experimental compound was diluted in DMSO to a concentration of 100 µM. Human liver microsomes were thawed and diluted with 0.1 M potassium phosphate buffer (pH 7.4) to a final protein concentration of 0.76 mg/ml. Experimental compound was then further diluted 100 fold to yield 1 µM final concentration with the diluted microsomal solution. The compound and microsomes were thoroughly mixed, and 150 µL of the resulting solution was added to one side of the dialysis well as a donor side and a 20 µL of this solution was retained for analysis. An equivalent 150 µL of 0.1 M potassium phosphate buffer (pH 7.4) was added to the other side of the dialysis well as a receiving side. The 96 well HTDialysis apparatus was sealed with a gas permeable membrane and incubated at 37 ºC, 10% CO₂ for 5 hours on an orbital shaker rotating at 200 rpm. After the 5 hour incubation period, the final microsomal solution and buffer samples (20 and 80 µL, respectively) were removed from each side of dialysis unit. The 20 µl of the microsomal solution from the donor side was mixed with 80 µl of phosphate buffer and the 80 µl of buffer sample from the receiving side was mixed with 20 µl of control diluted microsomes
to generate an equivalent matrix prior to sample extraction. Samples were frozen at -20°C until LC/MS/MS analysis.

**LC/MS/MS Analysis:** LC/MS/MS analysis was performed by using Shimadzu LC-20 AD gradient HPLC pump with Shimadzu DGU-20A3 degasser (Shimadzu, Columbia, MD), a CTC PAL autoinjector (Leap Technologies, Carrboro, NC), and a PE-Sciex API 3000 triple quadrupole mass spectrometer (Sciex, Thornhill, Ontario, Canada) with a turbo ionspray interface. Separation of the analyte was achieved by using an ACE C18 2.1 x 30 mm with 5-µm particle size operated under room temperature (ACE, Chadds Ford, PA). Mobile phase system A consisted of 95% of 10 mM ammonium formate (adjusted to pH 3.2 with formic acid) with 5% of acetonitrile. Mobile phase system B consisted of 95% of acetonitrile with 5% of 10 mM ammonium formate (adjusted to pH 3.2 with formic acid). The proportion of mobile phase B was held at 0% for 0.45 minute, then linear increase to 95% over next 0.75 minute. Then, mobile phase B will be held at 95% for 0.2 minute then return to 0% over 0.1 minute. After each injection the column was allowed to re-equilibrate for over 0.2 of a minute at 0% of mobile phase B. The system was operated with a flow rate of 0.75 ml/min.

The frozen samples were thawed at room temperature and quenched by 200 µL of acetonitrile containing an appropriate internal standard compound (IS). The plate of quenched samples was then centrifuged at 3000 rpm at 4 °C for 20 minutes in order to precipitate proteins, and then 10 µl of supernant was injected to LC/MS. Approximately 10% of the eluent was introduced into mass spectrometer source. The source temperature of the mass spectrometer was maintained at 450 °C. The most abundant product ion of each molecular ion (M+H⁺) was chosen in multiple reaction monitoring (MRM) mood based on collision-induced dissociations of precursor occurring at collision energy of 25, 35 or 45 eV, respectively, for accommodating the
needs of large quantity of compounds and the detection sensitivity of the mass spectrometer. Quantitation of each compound was achieved by comparison of the analyte/internal standard peak area ratios of the donor side (microsomal solution) and the receiving side (buffer). The unbound fraction of each compound in human liver microsomes (\(f_{u_{mic}}\)) was calculated according to eq. 5.

\[
f_{u_{mic}} = \frac{C_b}{C_m} \quad (5)
\]

Where \(C_b\) and \(C_m\) denote peak area ratio of a compound in dialysis apparatus from buffer and microsomal solution chambers, respectively.

**Model development**

For compounds with measurement done at different microsomal protein concentrations (internal historical data), the unbound fraction was normalized to a value at a microsomal protein concentration of 0.76 mg/ml using a modified version of an equation previously described (Kalvass and Maurer, 2002) (eq. 6).

\[
f_{u_{mic2}} = \frac{1}{\left(\frac{0.76}{C_1} \frac{1 - f_{u_{mic1}}}{f_{u_{mic1}}} \right) + 1} \quad (6)
\]

Where \(f_{u_{mic2}}\) is the unbound fraction at a microsomal concentration of 0.76 mg/ml, \(f_{u_{mic1}}\) is the unbound fraction at a microsomal concentration of \(C_1\).

A logit transformation of all the data was performed according to equation 7, which will enable the predicted \(f_{u_{mic}}\) within a boundary of 0 to 1.

\[
A = \log \left( \frac{f_{u_{mic}}}{1 - f_{u_{mic}}} \right) \quad (7)
\]
The whole set of 1223 compounds was randomly divided into two subsets. A subset of 1123 compounds was selected as a training set to develop a model, and a subset of 100 compounds was used as a validation set to validate the developed model. In addition, a set of 27 drug molecules (Table 1) was used as an external validation set to test the model performance.

In this study, we have used two-dimensional (2D) molecular descriptors including cLogP (BioByte Corp.), Kier connectivity, shape, and E-state indices (Kier, 1987, 1989 and 1999), a subset of MOE descriptors (Chemical Computing group, 2004), and a set of ADME keys which are structural features used for our in-house ADMET modeling (Lee et al., 2007; Tu and Li, 2004). Some of the descriptors such as Kier’s shape indices contain implicit three-dimensional (3D) information. Explicit 3D molecular descriptors were not used to avoid bias of the analysis due to predicted conformational effects and speed of calculation for fast prediction as stated in previous publication (Gao et al., 1999). cLogD and pKa calculations in our analysis were performed using ACD software version 9.03 (ACD/Labs).

The statistical model was derived using Cubist committee plus composite models (Rulequest Research, 2006). Cubist is a tool for generating rule-based predictive models and resembles a piecewise linear regression model, except that the rules can overlap. It does this by building a model containing one or more rules, where each rule is a conjunction of conditions associated with a linear regression. If a new case satisfies all the conditions, then the associated linear regression model is used to predict the target value. If two or more rules apply to a case, then the values are averaged to arrive at a final prediction. The predictive accuracy of a rule-based model can be improved by combining it with an instance-based or nearest-neighbor model. The latter predicts the target value of a new case by finding the n most similar cases in the training data, and averaging their target values. Then Cubist combines the rule-based
prediction with instance-based prediction to give a final predicted value. Prediction reliability was estimated using similarity to the training set of compounds. The similarity matrix used in the study is atom pair similarity (Carhart et al., 1985; Sheridan et al., 2004).

**Results**

A set of 1223 Pfizer proprietary compounds was selected that covers a range of lipophilicity (cLogP from -1.45 to 9.20 and cLogD from -4.6 to 9.0), molecular weight (from 200 to 650), ionization class (from an acidic pKa of 2.0 to a basic pKa of 11.2), and the degree of binding (Log[fumic/(1-fumic)] from 6 to -8). The selected 27 commercial compounds cover a range of lipophilicity (cLogP from 1.2 to 5.9 and cLogD from -0.25 to 4.8) and molecular weight (from 236 to 734), ionization class (from an acidic pKa of 3.96 to basic pKa of 10.4) and degree of binding (Log[fumic/(1-fumic)] from -4.6 to 5.0) (Table 1). An equilibrium dialysis in 96-well format was used to determine the extent of binding in triplicates for all the compounds in human liver microsomes. The 27 commercial molecules were used to validate the accuracy and precision of the experimental design and were used as quality control during the 4 month study period (standard error < 10% from quadruplicates). Data from studies with recovery >80% and CV (coefficient of variation) <20% from triplicates were included in the data set.

A predictive model was developed using Cubist. The observed and predicted microsomal unbound fractions (fumic) using the developed model are plotted in Figure 1. Since the model was built using committee plus composite models, a molecular descriptor can be used in multiple regression equations. It is not possible to describe the importance of a particular descriptor using correlation coefficients. The 10 most important molecular descriptors are cLogP, ADME Key 179 (smarts: [O,o]!:*:*), ADME Key 171 (smarts: [NH]C=[O,S,N]), kS_dO (Kier E-stare for =O), vsa_don (VDW donor surface area), ADME Key 10 (smarts:
*:=[c.C:H0,H1]O[C;H1,H2,H3]), ADME Key 302 (smarts: N~[!H]~[N,O]), number of carboxylic groups, and $^3\chi_3$ (Kier connectivity index). Hydrophobicity (cLogP) is the most important molecular property in this model. For the test set of 100 compounds, 92% of the variance is accounted for by this model with a RMSE (root mean square error) of 0.10, and for the training set of compounds, 99% of variance is accounted for by this model with a RMSE of 0.02. To further test the performance of the developed model, we determined the $f_{u\text{mic}}$ for 27 drug molecules and predicted with the model. The observed and predicted $f_{u\text{mic}}$ values for these 27 drugs are listed in Table 1 and graphed in Figure 1, and 52% of them are predicted within an error of 0.10.

Discussion

The distribution of measured microsomal unbound fraction ($f_{u\text{mic}}$) is shown in Figure 2, and $f_{u\text{mic}}$ from this study is evenly distributed across the whole range for neutral, basic, and acidic compounds.

The applicability of QSAR models are limited by the chemistry space of data from which the models are derived. If a model is trained on a narrow chemistry space, we intuitively expect the model only apply to a narrow range of chemical compounds. On another hand, if the training set is very diverse covering broad chemistry spaces, we would expect the model to be applicable to a wider chemical scope and therefore more powerful in molecular design. Generally it is believed that if a new chemical structure is similar to the training set of compounds, it is more likely to be well-predicted; otherwise there is a significant extrapolation and the estimated value contains high uncertainty. In this study, we have used atom-pair similarity as the similarity matrix to indicate whether a new molecule is similar to the training set of compounds. The
number of nearest neighbors (NN) were calculated using a similarity threshold of 0.7. This similarity index has been shown to perform well in many in-house ADMET models as well as in literature report (Sheridan et al., 2004). For the test set of 100 compounds, 84% of them with 1 or 2 NNs are predicted within an error of ±0.1, while only 55% of the compounds with 0 NN are predicted within an error of ±0.1. Similarly, 52% of the 27 commercial molecules are predicted within an error of ±0.10. The relationship of prediction error against number of NN is plotted in Figure 3, and it is apparent that the precision of the prediction increases with the increase of NN.

The non-specific binding of compounds to in vitro microsomal incubation matrix is believed to be mainly due to the partitioning of compounds to the hydrophobic phospholipids compartment. It is expected that the most important molecular property underlying the non-specific binding is molecular hydrophobicity. The correlations of unbound fraction with calculated hydrophobicity parameters cLogD and cLogP are shown in Figure 4. A similar observation has also been found in rat microsomal binding (Austin et al., 2002). Within similar cLogD range, basic compounds tend to have lower unbound fraction than neutral and acidic compounds, and this is mainly due to the fact that positively charged molecules have higher affinity to phospholipids. It is also noted that there is no significant difference among the three classes of compounds when plotted against cLogP, which is different from the phenomena of plasma protein binding ($f_{\text{uplasma}}$). In plasma protein, acidic compounds in general are more tightly bound than basic compounds. As Figure 5 shows, $f_{\text{umic}}$ does not track well with $f_{\text{uplasma}}$. This demonstrates that $f_{\text{umic}}$ and $f_{\text{uplasma}}$ are unlikely to cancel out each other in the prediction of human clearance using well-stirred or parallel-tube models (Obach, 1996). Both the unbound fraction in microsomal incubation media and the unbound fraction in human plasma are important in estimating in vivo clearance from in vitro metabolism data.
The free concentration of compound in microsomal incubation media significantly affects the observed *in vitro* apparent intrinsic metabolic clearance (CL\textsubscript{int,app}, eq. 3) and apparent CYP P450 inhibitory potency (IC\textsubscript{50}, eq. 4). The observed CL\textsubscript{int,app} vs. unbound CL\textsubscript{int} (corrected by fu\textsubscript{mic} using eq. 3) is plotted in Figure 6, and it has been noted that for compounds with low fu\textsubscript{mic} (<<1), the difference can be more than 1000 folds, where the observed CL\textsubscript{int,app} was adapted from a previous described method (Obach, 1999). Non-specific binding to incubation matrix renders fewer molecules available being metabolized by P450 enzymes thus masking the true metabolic liability of a compound. A retrospective Pfizer internal analysis has found that incorporation of human microsomal and plasma binding parameters could significantly increase the prediction accuracy of human PK parameters, including hepatic metabolic clearance and CYP P450 inhibitory potency (data not shown). This provides a strong driving force to develop a predictive *in silico* model to estimate non-specific binding to microsomal incubation media that will further assist molecule design with appropriate PK profiles to achieve desirable efficacious concentration and safety window.
References


Bjornsson TD; Callaghan JT; Einolf HJ; Fischer V; Gan L; Grimm S; Kao J; King SP; Miwa G; Ni L; Kumar G; McLeod J; Obach RS; Roberts S; Roe A; Shah A; Snikeris F; Sullivan JT; Tweedie D; Vega JM; Walsh J and Wrighton SA (2003) The conduct of in vitro and in vivo drug-drug interaction studies: A Pharmaceutical Research and Manufacturers of America (PhRMA) perspective. *Drug Metab Dispos* **31**: 815-832.


Legends For Figures

Figure 1. The observed vs. predicted microsomal unbound fraction (fu\textsubscript{mic}) (+ training set; ▲ test set; * marketed drugs)

Figure 2. Distribution of measured microsomal unbound fraction (fu\textsubscript{mic}) (grey: neutral compounds; blue: basic compounds; red: acidic compounds).

Figure 3. Plot of absolute prediction error vs. number of nearest neighbors (NN) to the training set (■ training set; ■ test set)

Figure 4. Relationship of microsomal unbound fraction with hydrophobicity (a. ACD LogD; b. clogP) (+ basic; ▲ acidic; ○ neutral)

Figure 5. Plot of unbound fraction in microsome (fu\textsubscript{mic}) and in plasma (fu\textsubscript{plasma})

Figure 6. The observed in vitro apparent human microsomal clearance (CL\textsubscript{int,app}) vs. calculated unbound clearance (CL\textsubscript{int})
Table 1. The physicochemical properties and observed $f_{u_{mic}}$ vs. predicted $f_{u_{mic}}$ for 27 commercial compounds

<table>
<thead>
<tr>
<th>Name</th>
<th>MW</th>
<th>cLogP</th>
<th>cLogD</th>
<th>Acidic pKa</th>
<th>Basic pKa</th>
<th>Measured $f_{u_{mic}}$ mean</th>
<th>STD error</th>
<th>Calculated $f_{u_{mic}}$</th>
<th>$\Delta^*$</th>
<th>NN**</th>
</tr>
</thead>
<tbody>
<tr>
<td>ALBENDAZOLE</td>
<td>250.22</td>
<td>3.46</td>
<td>3.06</td>
<td>11.2</td>
<td>5.62</td>
<td>0.69</td>
<td>0.02</td>
<td>0.65</td>
<td>0.04</td>
<td>1</td>
</tr>
<tr>
<td>AMITRIPTYLINE</td>
<td>254.23</td>
<td>4.85</td>
<td>3.15</td>
<td>9.18</td>
<td>0.27</td>
<td>0.02</td>
<td>0.14</td>
<td>0.13</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>CARBAMAZEPINE</td>
<td>224.18</td>
<td>2.38</td>
<td>2.67</td>
<td>9.46</td>
<td>0.97</td>
<td>0.04</td>
<td>0.82</td>
<td>0.15</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>CHLORIMIPRAMINE</td>
<td>291.68</td>
<td>5.92</td>
<td>3.54</td>
<td>9.46</td>
<td>0.09</td>
<td>0.01</td>
<td>0.08</td>
<td>0.01</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>CHLORPROMAZINE</td>
<td>299.72</td>
<td>5.3</td>
<td>2.87</td>
<td>9.41</td>
<td>0.15</td>
<td>0.01</td>
<td>0.03</td>
<td>0.12</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>CLOZAPINE</td>
<td>307.68</td>
<td>3.71</td>
<td>3.35</td>
<td>7.14</td>
<td>0.45</td>
<td>0.03</td>
<td>0.3</td>
<td>0.15</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>COLCHICINE</td>
<td>374.24</td>
<td>1.2</td>
<td>0.92</td>
<td>1.01</td>
<td>1.01</td>
<td>0.04</td>
<td>0.93</td>
<td>0.07</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>DESIPRAMINE</td>
<td>244.21</td>
<td>4.47</td>
<td>1.43</td>
<td>10.4</td>
<td>0.35</td>
<td>0.01</td>
<td>0.26</td>
<td>0.09</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>DIAZEPAM</td>
<td>271.64</td>
<td>2.96</td>
<td>2.96</td>
<td>3.4</td>
<td>0.77</td>
<td>0.03</td>
<td>0.55</td>
<td>0.22</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>ERYTHROMYCIN</td>
<td>666.40</td>
<td>1.61</td>
<td>2.08</td>
<td>8.16</td>
<td>0.84</td>
<td>0.02</td>
<td>0.85</td>
<td>0.01</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>IMIPRAMINE</td>
<td>256.22</td>
<td>5.04</td>
<td>2.79</td>
<td>9.49</td>
<td>0.35</td>
<td>0.02</td>
<td>0.57</td>
<td>0.22</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>INDOMETHACIN</td>
<td>341.66</td>
<td>4.18</td>
<td>-0.16</td>
<td>3.96</td>
<td>0.91</td>
<td>0.03</td>
<td>0.72</td>
<td>0.19</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>KETOCONAZOLE</td>
<td>503.22</td>
<td>3.38</td>
<td>2.71</td>
<td>6.33</td>
<td>0.27</td>
<td>0.08</td>
<td>0.27</td>
<td>0.02</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>KETOPROFEN</td>
<td>240.17</td>
<td>2.76</td>
<td>-0.25</td>
<td>4.23</td>
<td>0.92</td>
<td>0.06</td>
<td>0.93</td>
<td>0.01</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>MIDAZOLAM</td>
<td>312.67</td>
<td>3.42</td>
<td>3.92</td>
<td>5.59</td>
<td>0.65</td>
<td>0.01</td>
<td>0.64</td>
<td>0.01</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>PIMOZIDE</td>
<td>432.32</td>
<td>6.4</td>
<td>4.77</td>
<td>12</td>
<td>9.42</td>
<td>0.01</td>
<td>0.00</td>
<td>0.01</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>PIROXICAM</td>
<td>318.25</td>
<td>1.89</td>
<td>-1.15</td>
<td>4.5</td>
<td>3.8</td>
<td>1.02</td>
<td>0.02</td>
<td>0.98</td>
<td>0.02</td>
<td>1</td>
</tr>
</tbody>
</table>

*: Difference between measured and calculated $f_{u_{mic}}$

**NN: Number of nearest neighbors.
Table 1. The physicochemical properties and observed $f_{\text{mic}}$ vs. predicted $f_{\text{mic}}$ for 27 commercial compounds (continued)

<table>
<thead>
<tr>
<th>Name</th>
<th>MW</th>
<th>cLogP</th>
<th>cLogD</th>
<th>Acidic pKa</th>
<th>Basic pKa</th>
<th>Measured $f_{\text{mic}}$</th>
<th>Calculated $f_{\text{mic}}$</th>
<th>$\Delta^*$</th>
<th>NN**</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Measured $f_{\text{mic}}$</td>
<td>Measured $f_{\text{mic}}$</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>mean</td>
<td>STD error</td>
<td></td>
<td></td>
<td></td>
<td>mean</td>
<td>STD error</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PREDNISONE</td>
<td>332.23</td>
<td>1.66</td>
<td>1.57</td>
<td>12.4</td>
<td>0.98</td>
<td>0.06</td>
<td>0.84</td>
<td>0.14</td>
<td>1</td>
</tr>
<tr>
<td>PROMETHAZINE</td>
<td>264.27</td>
<td>4.4</td>
<td>2.29</td>
<td>8.98</td>
<td>0.29</td>
<td>0.01</td>
<td>0.2</td>
<td>0.09</td>
<td>0</td>
</tr>
<tr>
<td>PROPRANOLOL</td>
<td>238.18</td>
<td>2.75</td>
<td>1.27</td>
<td>9.14</td>
<td>0.62</td>
<td>0.02</td>
<td>0.46</td>
<td>0.16</td>
<td>1</td>
</tr>
<tr>
<td>QUINIDINE</td>
<td>302.25</td>
<td>2.79</td>
<td>1.6</td>
<td>12.8</td>
<td>9.28</td>
<td>0.72</td>
<td>0.06</td>
<td>0.65</td>
<td>0.07</td>
</tr>
<tr>
<td>SULINDAC</td>
<td>339.28</td>
<td>3.16</td>
<td>0.55</td>
<td>4.26</td>
<td>1.08</td>
<td>0.01</td>
<td>0.89</td>
<td>0.11</td>
<td>0</td>
</tr>
<tr>
<td>TENIDAP</td>
<td>311.68</td>
<td>2.3</td>
<td>-0.16</td>
<td>4.5</td>
<td>0.63</td>
<td>0.03</td>
<td>0.95</td>
<td>0.32</td>
<td>0</td>
</tr>
<tr>
<td>TOLBUTAMIDE</td>
<td>252.21</td>
<td>2.5</td>
<td>0.39</td>
<td>5.12</td>
<td>0.96</td>
<td>0.06</td>
<td>0.98</td>
<td>0.02</td>
<td>0</td>
</tr>
<tr>
<td>TRIMEPRAZINE</td>
<td>416.31</td>
<td>4.8</td>
<td>2.53</td>
<td>9.34</td>
<td>0.23</td>
<td>0.01</td>
<td>0.12</td>
<td>0.11</td>
<td>0</td>
</tr>
<tr>
<td>VERAPAMIL</td>
<td>292.20</td>
<td>4.47</td>
<td>2.33</td>
<td>8.97</td>
<td>0.59</td>
<td>0.03</td>
<td>0.43</td>
<td>0.16</td>
<td>1</td>
</tr>
<tr>
<td>WARFARIN</td>
<td>276.28</td>
<td>2.9</td>
<td>0.61</td>
<td>4.5</td>
<td>0.80</td>
<td>0.08</td>
<td>0.85</td>
<td>0.05</td>
<td>0</td>
</tr>
<tr>
<td>ALBENDAZOLE</td>
<td>250.22</td>
<td>3.46</td>
<td>3.06</td>
<td>11.2</td>
<td>5.62</td>
<td>0.62</td>
<td>0.02</td>
<td>0.46</td>
<td>0.16</td>
</tr>
</tbody>
</table>
Figure 1
No. of compounds

Binned figure

Figure 2
Figure 5