Short Communication

A Fragment-Based Approach for the Computational Prediction of the Nonspecific Binding of Drugs to Hepatic Microsomes

Received May 29, 2016; accepted August 18, 2016

ABSTRACT

Correction for the nonspecific binding (NSB) of drugs to liver microsomes is essential for the accurate measurement of the kinetic parameters $K_m$ and $K_i$, and hence in vitro–in vivo extrapolation to predict hepatic clearance and drug–drug interaction potential. Although a number of computational approaches for the estimation of drug microsomal NSB have been published, they generally rely on compound lipophilicity and charge state at the expense of other physicochemical and chemical properties. In this work, we report the development of a fragment-based hologram quantitative structure activity relationship (HQSAR) approach for the prediction of NSB using a database of 132 compounds. The model has excellent predictivity, with a noncross-validated $r^2$ of 0.966 and cross-validated $r^2$ of 0.680, with a predictive $r^2$ of 0.748 for an external test set comprising 34 drugs. The HQSAR method reliably predicted the fraction unbound in incubations of 95% of the training and test set drugs, excluding compounds with a steroid or morphinan 4,5-epoxide nucleus. Using the same data set of compounds, performance of the HQSAR method was superior to a model based on logP/D as the sole descriptor (predictive $r^2$ for the test set compounds, 0.534). Thus, the HQSAR method provides an alternative approach to laboratory-based procedures for the prediction of the NSB of drugs to liver microsomes, irrespective of the drug charge state (acid, base, or neutral).

Introduction

Human liver microsomes (HLM) are used widely as the enzyme source for the reaction phenotyping and kinetic characterization of drug metabolism and inhibitory drug–drug interaction potential of both established therapeutic agents and new chemical entities (Houston, 1994; Miners et al., 1994, 2010; Obach, 1999; Rostami-Hodjegan and Tucker, 2007). However, it is well established that numerous drugs and other xenobiotics may partition into the microsomal membrane during incubations with HLM and rat liver microsomes (RLM) (Herbette et al., 1983; Nussio et al., 2007; McLure et al., 2011; Nagar and Korzekwa, 2012). Nonspecific binding (NSB) to the microsomal membrane generally occurs to the greatest extent with lipophilic organic bases, with lesser binding of neutral and acidic compounds (Austin et al., 2002; Sykes et al., 2006; McLure et al., 2011).

As a result of NSB to HLM, and indeed other enzyme sources used for in vitro drug metabolism studies, the free concentration of the compound present in incubations is lower than the added concentration. Failure to account for NSB results in overestimation of key kinetic parameters (viz. $K_m$, $S_{50}$, and $K_i$) and hence underprediction of in vitro intrinsic clearance and drug–drug interaction potential (Obach, 1999; McLure et al., 2000; Margolis and Obach, 2003; Grime and Riley, 2006; Miners et al., 2006, 2010). NSB, measured as the fraction unbound in the incubation medium ($f_{unb}$), is determined experimentally by a range of techniques, including equilibrium dialysis (conventional and rapid equilibrium dialysis devices) and ultrafiltration. However, both techniques are time-consuming, requiring measurement of the drug concentration in each compartment of the dialysis apparatus or, in the case of ultrafiltration, in the reservoir and filtrate. Solubility is also a problem for highly lipophilic compounds (Burns et al., 2015). Given the broad dependence of NSB on lipophilicity and charge state, several algorithms based on these properties have been developed to predict $f_{unb}$, thereby circumventing the need for experimental methods (Austin et al., 2002; Hallifax and Houston, 2006; Sykes et al., 2006; SimCYP calculator).

It is clear, however, that physicochemical and chemical properties other than charge and lipophilicity determine NSB and hence $f_{unb}$. For example, we and others have reported that molecular mass, polar surface area, charge distribution, the number of hydrogen bond donors/acceptors, and the presence of halogen atoms (especially the trifluoromethyl group) may all influence NSB (Gao et al., 2008; Li et al., 2009; McLure et al., 2011; Burns et al., 2015). Moreover, we demonstrated recently that algorithms commonly used to calculate logP gave widely discrepant values of this parameter, differing by as much as a factor of 2 (i.e., 2 log units) (Burns et al., 2015). Perhaps not surprisingly, we observed that the published algorithms that use logP (or logD for acids) and charge state poorly predicted the $f_{unb}$ values of a series of protein kinase inhibitors (Burns et al., 2015).

We describe in this study a hologram-based approach for the computational prediction of the NSB of drugs to hepatic microsomes. Hologram quantitative structure activity relationship (HQSAR) is a technique that employs fragment fingerprints (molecular holograms) as predictive variables of biologic activity or structural data. Each molecule in the data set of interest is fragmented into structural components to generate a molecular hologram that provides a count of the number of times a specific component is identified. Partial least
squares analysis is then used to generate a statistical model that relates the molecular holograms (descriptor variables) to a measured property [viz. \( \log([1 - f_{\text{unib}}]/f_{\text{unim}}) \) in this study] (Wold et al., 1984). HQSAR has been applied to the prediction of several biologic variables, including inhibition/antagonism of different drug targets, toxicity prediction, drug oral bioavailability, and identification of novel chemical scaffolds (Tong et al., 1998; Moda et al., 2007; Nair et al., 2008; Nair and Sobhia, 2008a,b). In the current study, a database of 132 drugs with a broad range of physicochemical properties was collated to establish quantitative relationships between \( f_{\text{unib}} \) and the molecular fragments. Models generated were evaluated using an independent set of test drugs. The models show excellent discrimination between high, intermediate, and low microsomal binders, with a high predictive power.

**Materials and Methods**

**Data Set for Analysis.** The \( f_{\text{unib}} \) values of 98 drugs from studies with HLM were collated from published data (Obach, 1999; McLure et al., 2000, 2011; Nair et al., 2001; Gertz et al., 2008; Gao et al., 2008; Kilford et al., 2008; Raungrut et al., 2010; Uchaipichat et al., 2011; Wattanachai et al., 2011; Burns et al., 2015; Pattanawongsa et al., 2015) (Supplemental Tables 1 and 2). Because \( f_{\text{unib}} \) does not differ appreciably between HLM and RLM (Obach, 1997; Zhang et al., 2010), \( f_{\text{unib}} \) values for the binding of 34 drugs to RLM were additionally incorporated into the data set (Obach, 1997; Austin et al., 2002). Thus, the data set in total contained 132 compounds, comprising 28 acidic, 55 neutral, and 49 basic drugs. The classification of acids (pKa for the formation of an anion <7.4) and bases (pKa for the formation of a cation >7.4) was as defined previously (Austin et al., 2002; Sykes et al., 2006). The data set molecules were randomly segregated into 82 training and 34 test set molecules, with 16 compounds excluded as outliers (see Results and Discussion). Based on the UNITY fingerprint (generated using SYBYL-X 2.1, Tripos International, Certara, St Louis, MO; Supplemental Fig 1), training and test set compounds exhibited an acceptable balance of similarity/dissimilarity. Where \( f_{\text{unib}} \) values were determined at microsomal concentrations other than 0.25 and 1 mg/ml, data were normalized to 0.25 mg/ml and 1 mg/ml, using the expression (Austin et al., 2002):

\[
\frac{fu_2}{fu_1} = \frac{1}{\frac{fu_1}{fu_2} (1 - fu_1) + 1}
\]

where \( fu_1 \) is the fraction unbound obtained at particular microsomal concentration \( C_1 \), \( fu_2 \) is the normalized \( f_{\text{unib}} \) at concentration \( C_2 \) (0.25 mg/ml or 1 mg/ml). The \( f_{\text{unib}} \) values were transformed as \( \log([1 - f_{\text{unib}}]/f_{\text{unib}}) \), given the similarity of this expression to an equilibrium constant (Austin et al., 2002; Sykes et al., 2006). Values of \( f_{\text{unib}} \geq 1 \) were rounded to 0.99, for the ease of calculation and to prevent exclusion of data.

**Molecular Modeling and HQSAR Model Development.** Molecular modeling studies were performed using the molecular modeling package SYBYL-X 2.1 installed on a Macintosh Workstation running on OS X 10.9.5 operating system (SYBYL-X 2.1). The structural coordinates of drug molecules were obtained from the Pubchem server (https://pubchem.ncbi.nlm.nih.gov/) and imported into SYBYL.

Molecular fragments were generated with different atom counts (see Table 1) and were hashed into 85 bins (i.e., bins 1–85) of the fingerprint. The molecular hologram contains information about the number of fragments hashed into each bin. The optimal HQSAR model was derived from screening through the hologram lengths, set as 12 prime numbers ranging from 53 to 401 (SYBYL-X 2.1). Model development involved generation of a range of fragment sizes together with fragment features. Fragment size refers to the minimum and maximum size of the fragment in a hologram fingerprint. Fragment features include an atom feature (based on elemental atom types) and bonds and connections, which consider bond orders and hybridization states within fragments, respectively (SYBYL-X 2.1). The models shown in Table 1 were obtained based on atoms, bonds, and connection features (default parameters) of the molecule. Individual models were evaluated statistically using the \( r_{cv}^2, r_{ncv}^2 \) and S.E. (Table 1).

The best model (model C; Table 1) was further validated by predicting the \( \log([1 - f_{\text{unib}}]/f_{\text{unib}}) \) for an external set of test molecules and calculating the predictive \( r_{pred}^2 \), defined as:

\[
r_{\text{pred}}^2 = \frac{SD - PRESS}{SD}
\]

where SD is the sum of the squared deviations between the measured \( f_{\text{unib}} \) values [viz. \( \log([1 - f_{\text{unib}}]/f_{\text{unib}}) \)] of molecules in the test set and the mean \( \log([1 - f_{\text{unib}}]/f_{\text{unib}}) \) values of the training set molecules, and PRESS is the sum of the squared deviations between predicted and actual activity values for all molecules in the test set (Waller et al., 1993; Nair et al., 2008; Nair and Sobhia, 2008b).

**Results and Discussion**

Currently available models [including quantitative structure activity relationship (QSAR)] for the prediction of drug NSB generally rely on, or at least incorporate, logP/D, charge state, or structural similarity. As noted earlier, we have recently demonstrated that approaches based on logP and charge state may be unreliable due to variation in the logP values between algorithms and the contribution of other physicochemical and chemical features to NSB (McLure et al., 2011; Burns et al., 2015). Thus, a hologram fragment approach was used in this study to develop a QSAR model based on published \( f_{\text{unib}} \) values for 132 drugs. Published \( f_{\text{unib}} \) values employed for model generation ranged from 0.002 to 0.99.

Several models were generated using various HQSAR parameters. Evaluation of the training set molecules across a range of atom counts (fragment size) (Table 1) and atoms, bonds, and connection parameters demonstrated significant correlations with NSB. The quality of the model is assessed from statistical parameters that include noncross-validated \( r_{ncv}^2 \), cross-validated \( r_{cv}^2 \) by leave-one-out, predicted \( r_{\text{pred}}^2 \), and S.E. During model development, it was noted that the correlation improves as the atom count is increased from 1 to 3–6. However, beyond atom counts 4–7 the predictive power of the model decreased, as shown by a lower \( r_{cv}^2 \) and higher S.E. The best model (C; Table 1) was obtained with atom counts 3–6, giving a \( r_{ncv}^2 = 0.966 \) and \( r_{cv}^2 = 0.680 \) (\( F = 1545, P = 1.58 \times 10^{-25} \)). Incorporation of additional features, namely chirality, hydrogen atoms, and hydrogen bond donor or acceptor features, resulted in poor predictive and statistically less significant models (data not shown). These models were not considered further.

Earlier studies investigating the NSB of drugs were generally measured at a HLM (or RLM) concentration of 1 mg/ml. However, in vitro metabolism and inhibition studies are now often performed at lower HLM concentrations (≤0.25 mg/ml). To examine the predictivity of the hologram QSAR approach, \( f_{\text{unib}} \) values obtained at two (0.25 mg/ml and 1 mg/ml) microsomal concentrations were considered. If the generated fragments truly determine NSB (and \( f_{\text{unib}} \)), then similar prediction is expected irrespective of the microsome concentration.

**TABLE 1**

<table>
<thead>
<tr>
<th>Atom counts</th>
<th>Models</th>
<th>( f_{cv}^2 )</th>
<th>( f_{ncv}^2 )</th>
<th>( r_{cv} )</th>
<th>( r_{ncv} )</th>
<th>( r_{pred}^2 )</th>
<th>SD-PRESS</th>
</tr>
</thead>
<tbody>
<tr>
<td>1–3</td>
<td>Model A</td>
<td>0.513</td>
<td>0.680</td>
<td>0.470</td>
<td>0.426</td>
<td>0.389</td>
<td>0.002</td>
</tr>
<tr>
<td>4–6</td>
<td>Model B</td>
<td>0.561</td>
<td>0.470</td>
<td>0.426</td>
<td>0.389</td>
<td>0.389</td>
<td>0.002</td>
</tr>
<tr>
<td>7–9</td>
<td>Model C</td>
<td>0.898</td>
<td>0.909</td>
<td>0.931</td>
<td>0.839</td>
<td>0.839</td>
<td>0.002</td>
</tr>
<tr>
<td>10–12</td>
<td>Model D</td>
<td>0.944</td>
<td>0.966</td>
<td>0.909</td>
<td>0.921</td>
<td>0.921</td>
<td>0.002</td>
</tr>
<tr>
<td>13–15</td>
<td>Model E</td>
<td>0.267</td>
<td>0.198</td>
<td>0.377</td>
<td>0.456</td>
<td>0.456</td>
<td>0.002</td>
</tr>
<tr>
<td>16–18</td>
<td>Model F</td>
<td>0.362</td>
<td>0.377</td>
<td>0.456</td>
<td>0.489</td>
<td>0.489</td>
<td>0.002</td>
</tr>
</tbody>
</table>

BHL, best hologram length; NOC, number of optimum components; \( r_{cv}^2 \), cross-validated correlation coefficient; \( r_{ncv}^2 \), noncross-validated correlation coefficient.

[Downloaded from dmd.aspetjournals.org at ASIET Journals on November 28, 2017]
Importantly, the models generated in this work (A–F, Table 1) show identical correlations at both microsomal concentrations for the same training and test set compounds with near identical predictivity (Fig. 1, A and B). With respect to Fig. 1, it should be noted that \( \log[(1 - f_{\text{mic}})/f_{\text{mic}}] \) values in the range \(-0.95 \) to \(-2.00 \) represent compounds with an \( f_{\text{mic}} \) \(> 0.9 \), whereas \( \log[(1 - f_{\text{mic}})/f_{\text{mic}}] \) values \(> 2.00 \) represent compounds with a \( f_{\text{mic}} < 0.01 \) (for 1 mg/ml microsomal concentrations).

The training set is comprised of 82 drugs: 18 acids, 31 bases, and 33 neutral molecules. The \( f_{\text{mic}} \) values of drugs in the training set ranged from 0.002 (very high binders) to 0.99 (essentially unbound), all of which were predicted with high accuracy (Fig. 1; Supplemental Table 1). Notably, the hologram method is not biased toward a specific physicochemical class, and predicts the binding of acids, bases, and neutral compounds with high accuracy (Fig. 1). This contrasts to models based on logP and charge state where poorer correlations were noted for acidic and neutral drugs compared with bases, and hence for a combined database of acids, bases, and neutral compounds (Austin et al., 2002; Sykes et al., 2006). The QSAR model of Li et al. (2009) similarly predicted the \( f_{\text{mic}} \) values of acidic drugs poorly.

The robustness of the HQSAR model is apparent from the test set predictivity, reflected by the predictive \( r^2 \) value. Like the \( r^2 \), the \( r^2 \) can assume a negative value reflecting a complete lack of predictive ability of the training set (Cramer et al., 1988). The test set of 34 drugs is comprised of 10 acids, 11 bases, and 13 neutral compounds with diverse chemical structures and \( f_{\text{mic}} \) values ranging from 0.02 to 0.99. The \( f_{\text{mic}} \) values of the drugs in each class were predicted well (Fig. 1), with \( r^2 = 0.748 \). In particular, the model accurately predicted high binders (84–98% binding at a microsome concentration of 1 mg/ml) in the test set, including ezlopitant (\( f_{\text{mic}} = 0.02 \)), sorafenib (\( f_{\text{mic}} = 0.04 \)), nicardipine (\( f_{\text{mic}} = 0.04 \)), chlorpromazine (\( f_{\text{mic}} = 0.08 \)), lapatinib (\( f_{\text{mic}} = 0.08 \)), saquinavir (\( f_{\text{mic}} = 0.10 \)), promethazine (\( f_{\text{mic}} = 0.11 \)), and imipramine (\( f_{\text{mic}} = 0.16 \)) (Supplemental Table 2). Overall, the HQSAR method predicted the \( f_{\text{mic}} \) values of all data set compounds (excluding outliers), within 1 log units (at a HLM concentration of 0.25 mg/ml), and 88% within 0.50 log units of the transformed \( f_{\text{mic}} \) data.

The model of Austin et al. (2002) excluded compounds with a \( f_{\text{mic}} > 0.90 \), as these compounds may be difficult to discriminate using equilibrium dialysis. In the current study, compounds with a \( f_{\text{mic}} > 0.90 \) were included to examine the predictivity of the hologram approach, and the majority were predicted well. However, predicted transformed values [i.e., \( \log[(1 - f_{\text{mic}})/f_{\text{mic}}] \)] for a number of compounds deviated from the measured values. Overall, of the 132 data set molecules, 16 were predicted with a deviation \( > \pm 1 \) in the \( \log[(1 - f_{\text{mic}})/f_{\text{mic}}] \), and these were excluded from the training and test sets. Inclusion of the outliers in the training set reduced the \( r^2 \) and \( q^2 \) values to 0.833 and 0.401, respectively. Notably, 10 of these outliers possess either a steroid (aldosterone, androstenedione, \( \beta \)-estradiol cortisol, progesterone, and testosterone—all neutral molecules) or a morphinan nucleus incorporating a 4,5-epoxy ring system (buprenorphine, codeine, morphine, and naloxone—all bases), suggesting that the HQSAR approach is not consistently predictive for these structural classes. Dextromethorphan and levallorphan, which contain the morphinan nucleus without the 4,5-epoxide, and dexamethasone and prednisone, were, however, correctly predicted as low binders, with a \( f_{\text{mic}} > 0.90 \). Atenolol, milbefradil, simvastatin, tacrolimus, tamoxifen, and trametinib were the other molecules with a deviation \( > \pm 1 \) in \( \log[(1 - f_{\text{mic}})/f_{\text{mic}}] \). Although the deviation in logarithmic transformed data is \( > 1 \) for atenolol, the difference between measured and predicted \( f_{\text{mic}} \) is relatively small (0.99 versus 0.90, respectively; Supplemental Table 1). Experimental error in the determination of \( f_{\text{mic}} \) concentration-dependent binding.

![Fig. 1.](image-url) (A and B) Plots of measured versus HQSAR predicted \( \log[(1 - f_{\text{mic}})/f_{\text{mic}}] \) of the training and test set molecules for microsome concentrations of 0.25 mg/ml (A) and 1 mg/ml (B). (C) Plot of measured \( \log[(1 - f_{\text{mic}})/f_{\text{mic}}] \) versus logP/D. (D) Plot of measured \( \log[(1 - f_{\text{mic}})/f_{\text{mic}}] \) versus predicted \( \log[(1 - f_{\text{mic}})/f_{\text{mic}}] \) for the logP/D model. The training set and test set molecules are shown as blue and orange squares, respectively.
(McLure et al., 2011) may potentially contribute to the poor predictivity observed for mibefradil, simvastatin, tacrolimus, tamoxifen, and trametinib, whose inclusion in the training set significantly decreased overall model predictivity. As we have observed recently, the experimental determination of the NSB of highly lipophilic values can be problematic due to their limited aqueous solubility (Burns et al., 2015).

Our data set was additionally employed to investigate the predictivity of calculated logP/D (logP for bases, logD_{7.4} for acids and neutrals; see Supplemental Table 3) as the sole descriptor of fumic, as described by Austin et al. (2002). The training set was described by \( \log\left(\frac{1 - f_{\text{mic}}}{f_{\text{mic}}}\right) = 0.43\log P/D - 1.43 \) \((\text{r}^2 = 0.540, \text{F} = 113, P = 7.03 \times 10^{-18})\) (Fig. 1C). This expression was subsequently used to predict \( \log\left(\frac{1 - f_{\text{mic}}}{f_{\text{mic}}}\right) \) (Fig. 1D). Compared with the HQSAR, the logP/D model shows lower test set predictivity with a \( r_{\text{pred}}^2 = 0.534 \) (cf. 0.748 for the HQSAR model; see above). Exclusion of the outliers from the HQSAR method had minimal effect on the performance of the logP/D model. Consistent with previous data from this laboratory (Sykes et al., 2006), the correlation for acids and neutral molecules was worse than that for bases using the logP/D method. As noted earlier, calculated values of logP/D vary between algorithms, and it is possible that the use of alternative algorithms may result in better (or worse) correlations. Furthermore, logP/D can be combined with other descriptors such as polar surface area, van der Waal donor surface area, connectivity indices, etc., to improve the prediction of \( f_{\text{mic}} \) (Gao et al., 2008; Li et al., 2009).

The HQSAR approach further provides insights into the contribution of specific atoms and functional groups toward NSB by atomic color coding. Figures 2 and 3 show the color coding for representative drugs of different charge state (base and neutral) and which exhibit a wide a range of \( f_{\text{mic}} \) values. Generally, drugs that are high binders tend to show more atoms/groups with a positive contribution (noted as green/yellow) toward NSB (e.g., amiodarone, thiordazine, regorafenib), whereas poor binders (e.g., piroxicam, indapamide) show more atoms/groups with a negative contribution (red/orange end of the spectrum) to NSB. Although color coding is qualitative, atomic/molecular features associated with strong microsomal binding can be identified. Features that contribute to NSB include an amine group (e.g., amiodarone, thiordizaine, astemizole), a halogen atom (e.g., amiodarone, astemizole, regorafenib, sorafenib), the \(-\text{CF}_3\) group (e.g., regorafenib, sorafenib), a methylpiperidine ring (e.g., thiordizaine), and hydrocarbon features (e.g., amiodarone, imipramine, paclitaxel, thiordizaine, regorafenib). By contrast, no features that consistently contribute unfavorably to NSB were highlighted, including the carboxylic acid group, which is the common feature of low-binding organic acids.

In conclusion, the HQSAR method reported in this work reliably predicted the \( f_{\text{mic}} \) values of 114 of 120 (i.e., 95%) training and test set drugs when compounds with a steroid or morphinan 4,5-epoxide nucleus are excluded. Thus, the HQSAR method provides an alternative approach to laboratory-based procedures for the prediction of the NSB of drugs to liver microsomes, irrespective of the drug charge state (acid, base, or neutral). The method would appear to be particularly valuable for assessing the NSB of large data sets of new drug-like molecules.

\[ \text{Department of Clinical Pharmacology (P.C.N., J.O.M.) and Flinders Centre for Innovation in Cancer (P.C.N., R.A.M., J.O.M.}, \]
\[ \text{School of Medicine, Flinders University, Adelaide, Australia} \]

**Authorship Contributions**

Participated in research design: Nair, Miners.
Conducted experiments: Nair.
Performed data analysis: Nair, Miners.
Wrote or contributed to the writing of the manuscript: Nair, Miners, McKinnon.
References


Address correspondence to: Dr. Pramod C. Nair, Department of Clinical Pharmacology, Flinders University School of Medicine, GPO Box 2100, Adelaide, SA 5001, Australia. E-mail: pramod.nair@flinders.edu.au