QSAR Modeling of Human Liver Microsomal Stability

Alexey Zakharov

CADD Group
Chemical Biology Laboratory
Frederick National Laboratory for Cancer Research
National Cancer Institute, National Institutes of Health, DHHS
Causes of major drug failures

- Global Business Intelligence Research released a report related to the causes of major drug failures during 2005-2010, in which more than 20 drug failures were analyzed. The main reasons for failures were: 68% related to efficacy, 21% to safety.

- Efficacy and safety is strongly influenced by metabolic degradation and excretion.
Metabolic stability assessment

- The most important value that can be measured to quantify metabolic excretion and thus stability of compounds is their half-life time ($t_{1/2}$) determined in human liver microsomes.

- High-throughput in vitro metabolic stability assays are widely used for investigation of the stability of compounds.

- An alternative are computational approaches (QSAR methods), which can be applied to prioritize large numbers of compounds for in vivo measurements.
# Microsomal stability data sets

1) Evolvus database (commercial), Elvolvus Group, India

2) ChEMBL (public), Assay ID 1614674


4) Sanford Burnham Medical Research Institute (SBMRI), (see PubChem AID 1555, AID1940)

<table>
<thead>
<tr>
<th>Data Set</th>
<th>Number of Compounds</th>
<th>Unstable</th>
<th>Stable</th>
</tr>
</thead>
<tbody>
<tr>
<td>Evolvus data set</td>
<td>1242</td>
<td>345</td>
<td>897</td>
</tr>
<tr>
<td>ChEMBL human external set</td>
<td>669</td>
<td>5</td>
<td>664</td>
</tr>
<tr>
<td>Goodman &amp; Gilman human external set</td>
<td>246</td>
<td>5</td>
<td>241</td>
</tr>
<tr>
<td>SBMRI human external set</td>
<td>80</td>
<td>21</td>
<td>59</td>
</tr>
</tbody>
</table>

Unstable: $t_{1/2} \leq 15$ min; stable: $t_{1/2} > 15$ min.
Diversity analysis of Evolvus data set

- Fingerprints from KNIME
- Tanimoto distance
- Sammons embedding approach
Significant differences in $t_{1/2}$ measurements

Experimental $t_{1/2}$ data in the ChEMBL set vs. the G&G set

These two sets have 156 structures in common
Significant differences in $t_{1/2}$ measurements

The ChEMBL and G&G sets have the same end-point data for 156 structures, but obtained from different sources

<table>
<thead>
<tr>
<th>$t_{1/2}$, min.</th>
<th>0 -15</th>
<th>15 – 60</th>
<th>60 - 360</th>
<th>360-720</th>
<th>720-1440</th>
<th>1440-6000</th>
</tr>
</thead>
<tbody>
<tr>
<td>$t_{1/2}$, hours</td>
<td>0.25</td>
<td>0.25 – 1</td>
<td>1 - 6</td>
<td>6 - 12</td>
<td>12 - 24</td>
<td>24 - 100</td>
</tr>
<tr>
<td>Number of compounds</td>
<td>3</td>
<td>10</td>
<td>82</td>
<td>28</td>
<td>19</td>
<td>14</td>
</tr>
<tr>
<td>Average difference, min.</td>
<td>12</td>
<td>9</td>
<td>50</td>
<td>197</td>
<td>188</td>
<td>411</td>
</tr>
<tr>
<td>Average difference (%)</td>
<td>43</td>
<td>18</td>
<td>23</td>
<td>31</td>
<td>18</td>
<td>20</td>
</tr>
</tbody>
</table>
QSAR Methods

GUSAR (commercial), Ver. 2011

StarDrop (commercial), Ver. 5.0

KNIME (public), Ver. 2.4.2
QSAR Methods

GUSAR software

QNA (Quantitative Neighborhoods of Atoms) descriptors


PASS* Predictions as independent variables


Self-Consistent Regression (SCR)


*PASS: Prediction of Activity Spectra for Substances
QNA: Quantitative Neighborhoods of Atoms descriptors

\[ P_i = B_i \sum_k (\exp(-\frac{1}{2}C))_{ik} B_k \]
\[ Q_i = B_i \sum_k (\exp(-\frac{1}{2}C))_{ik} B_k A_k \]

\[ A = \frac{1}{2}(IP + EA), \]
\[ B = (IP - EA)^{-\frac{1}{2}}, \]

**IP** is the first ionization potential,

**EA** is the electron affinity.


QNA: Quantitative Neighborhoods of Atoms descriptors

\[
C = \begin{bmatrix}
0 & 1 & 1 & 1 & 0 \\
1 & 0 & 0 & 0 & 1 \\
1 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0
\end{bmatrix}
\]

\[\text{Exp}\left(-\frac{1}{2}C\right) = \begin{bmatrix}
1.40 & -0.59 & -0.57 & -0.57 & 0.14 \\
-0.59 & 1.27 & 0.14 & 0.14 & -0.54 \\
-0.57 & 0.14 & 1.13 & 0.13 & -0.02 \\
-0.57 & 0.14 & 0.13 & 1.13 & -0.02 \\
0.14 & -0.54 & -0.02 & -0.02 & 1.13
\end{bmatrix}\]

(a) structural formula;
(b) connectivity matrix;
(c) exponent of the connectivity matrix;
(d) electron affinities (EA), ionization potentials (IP), parameters A and B, P and Q values for each of the atoms of formic acid molecule.
StarDrop uses several different **QSAR techniques**:  

- Partial Least Squares 
- Radial Basis Function fitting (RBF SD) 
- Gaussian Processes (GP SD) 
- Decision Trees (DT SD)
QSAR Methods

KNIME software

Descriptors used:

MOLD2 descriptors*:
• Physicochemical properties, fragmental descriptors, structural features and functional groups
• Total number of descriptors: 777

KNIME uses several different QSAR techniques:

• K Nearest Neighbor (kNN)
• Multilayer Perceptron (MLP)
• Support Vector Machine (SVM)
• Bayes Network (BayesNet)
• Radial basis function network (RBF network)
• Logistic regression (Logistic)

QSAR Workflow

Training Set
998 compounds

Model building

GUSAR
StarDrop
KNIME

Model selection

Test Set
976 compounds

Prediction

External Test Set
80 compounds

External validation

SBMRI Dataset

Evolvus Test set
ChEMBL Test set
G&G Test set

Evolvus set
# Calculation of prediction accuracy

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Accuracy</strong></td>
<td>$\frac{TN + TP}{TN + TP + FP + FN}$</td>
<td>Accuracy: probability of correctly classifying compounds.</td>
</tr>
<tr>
<td><strong>Sensitivity</strong></td>
<td>$\frac{TP}{FN + TP}$</td>
<td>Sensitivity: probability of predicting positive (unstable) when true outcome is positive.</td>
</tr>
<tr>
<td><strong>Specificity</strong></td>
<td>$\frac{TN}{TN + FP}$</td>
<td>Specificity: probability of predicting negative (stable) when true outcome is negative.</td>
</tr>
<tr>
<td><strong>CCR</strong></td>
<td>$(\text{Sensitivity} + \text{Specificity}) / 2$</td>
<td>Correct Classification Rate: shows balance between Sensitivity and Specificity.</td>
</tr>
</tbody>
</table>

Prediction results for Test set

- GUSAR
- RBF StarDrop
- BayesNet KNIME
- MLP KNIME
- GP StarDrop
- Logistic KNIME
- kNN KNIME
- RBF network KNIME
- SVM KNIME
- DT StarDrop

Accuracy & CCR
Prediction results for SBMRI test set

The chart shows the prediction results for various models on the SBMRI test set. The models include GUSAR, RBF StarDrop, RBF network KNIME, KNN KNIME, SVM KNIME, GP StarDrop, MLP KNIME, BayesNet KNIME, Logistic KNIME, and DT StarDrop. The accuracy and CCR (Correlation Coefficient Ratio) are represented by blue and red bars, respectively.
The NCI database includes more than 250,000 compounds, which are the publicly available part of the half-million structures assembled by the U.S. National Cancer Institute (NCI) in the course of its 55 year long efforts in screening compounds against cancer and AIDS.

Each compound from the NCI database was classified as stable or unstable using the best GUSAR models.

The prediction output also included an assessment of the applicability domain as provided by GUSAR (196460 comp.).

Summary

• Information about chemical structures and their half-life data was collected from several public and commercial sources and used for the construction of categorical QSAR models.

• Predictive QSAR models were developed using both commercial (StarDrop, GUSAR) and open-source software (KNIME).

• For estimation of the predictivity of the models, several external sets were used.

• The obtained QSAR models showed generally high accuracy of prediction.

• The best obtained model was used to predict metabolic stability of about 196,460 structures from the NCI database. These data have been made available for free download.
Acknowledgments

CADD Group, Chemical Biology Laboratory, Center for Cancer Research
Megan L. Peach   Markus Sitzmann   Igor V. Filippov   Marc C. Nicklaus

Vanderbilt University
Heather J. McCartney

Sanford Burnham Medical Research Institute
Layton H. Smith

Computer-Aided Drug Design at Cancer Research UK, Beatson Laboratories
Angelo Pugliese