What is ADMET Predictor™?

ADMET Predictor™ is an advanced computer program that enables pharmaceutical researchers to rapidly estimate a number of ADMET properties of new chemical entities (NCE’s) from their molecular structure. It’s predictive models are grouped into optional modules listed below. ADMET Predictor models have been consistently ranked as the most accurate in independent published comparisons.*

Physico-chemical and Biopharmaceutical Module

- **NEW!** Blood-brain barrier permeation (two models: qualitative and quantitative)
- Enhanced pKa (ionization constants; a multiprotic model)
- Permeability through rabbit cornea
- Human effective permeability in jejunum (Peff)
- MDCK apparent permeability (Papp)
- Solubility
  - Biorelevant solubility in FaSSIF, FeSSIF & SGF
  - Native solubility (solubility in pure water @ 25°C)
  - Native pH at saturation in pure water @ 25°C
  - Intrinsic solubility in pure water @ 25°C
  - Salt solubility factor @ 25°C
  - Water solubility at user-specified pH
- Supersaturation ratio
- logP (two models: artificial neural network ensemble and Moriguchi)
- logD (estimation of octanol-water distribution coefficient at user-defined pH)
- Diffusion coefficient in water (Hayduk-Laudie formula)
- Molal volume (Schroeder formula)
- Human volume of distribution
- Human plasma protein binding as percent unbound
- Blood-to-plasma concentration ratio
- Activity models
  - Inhibition of HIV integrase mediated via strand transfer and 3’ processing

Metabolite Module

- **NEW!** Likely sites of metabolic attack by five CYP P450 enzymes: 1A2, 2C9, 2C19, 2D6, 3A4
- **NEW!** Classification of whether a molecule will be a substrate of one of the five CYP P450 enzymes: 1A2, 2C19, 2C9, 2D6, and 3A4

Simulation Module

- Human fraction absorbed (by simulation at 1 mg, 10 mg, 100 mg, and 1000 mg dose levels)
- SimDOSE - assess the likely dose needed to achieve a therapeutic concentration at steady-state
ADMET Modeler Module

ADMET Modeler is an integrated module of ADMET Predictor™ that automates the difficult and tedious process of making high quality predictive structure-property models from sets of experimental data. It works seamlessly with ADMET Predictor structural descriptors as its inputs, and appends the selected final model back to ADMET Predictor as an additional predicted property.

The following modeling methods are offered by ADMET Modeler:

- Kohonen Self-Organizing Maps
- Artificial Neural Network Ensembles for regression and binary classification
- Support Vector Machine Ensembles for regression and classification
- Kernel Partial Least Squares and Ordinary PLS for regression
- Multiple Linear Regression
- SALI analysis

Enstein Metabolism Module

- Michaelis-Menten kinetic $K_m$ and $V_{max}$ constants for hydroxylation reaction catalyzed by five CYP P450 enzymes: 1A2, 2C9, 2C19, 2D6, 3A4
- Intrinsic clearance, $Cl_{int}$, resulting from metabolic activity of five CYP P450 enzymes: 1A2, 2C9, 2C19, 2D6, 3A4
- NEW! General inhibitory properties against five CYP P450 enzymes: 1A2, 2C9, 2C19, 2D6, and 3A4
- Specific inhibitory properties against 3A4-mediated metabolism of midazolam and testosterone
- Specific inhibitory constant $K_i$ against 3A4-mediated metabolism of midazolam and testosterone
- Probability of metabolism by human UGT (1A1, 1A3, 1A4, 1A6, 1A9, 1A10, and 2B7)

Other Features

Customizable ADMET Risk Filters
- Risk of low absorption from an oral dose (three models: one derived from a focused subset of the World Drug Index, one trained on measured fraction absorbed in human, and one identical to the Lipinski’s Rule of 5)
- Risk of mutagenicity
- Risk of overall toxicity
- Risk of metabolic liability
- Global ADMET Risk summarizing all of the above in one

Descriptor Sensitivity Analysis
- Interpretation of model predictions in structural terms for guided design of molecule's derivatives with desired ADMET properties

Toxicity Module

- Allergenic skin sensitization
- Environmental bioconcentration factor
- Acute lethal toxicity in rat as $LD_{50}$
- Estrogen receptor toxicity
- NEW! Androgen receptor toxicity
- Maximum recommended therapeutic dose (MRTD)
- Fathead minnow lethality as $LC_{50}$
- NEW! Chromosomal aberrations
- NEW! Acute toxicity in Daphnia magna (water fleas) as $pLC_{50}$
- NEW! Phospholipidosis

- NEW! Allergenic respiratory sensitization
- NEW! Reproductive/developmental toxicity
- Carcinogenicity in rats as $TD_{50}$
- Carcinogenicity in mice as $TD_{50}$
- Qualitative filter of mutagenicity in 10 strains of *Salmonella* bacteria
- NEW! hERG-encoded $K^+$ channel affinity as $pIC_{50}$
- Acute toxicity in *Tetrahymena pyriformis* as $IGC_{50}$
- Human liver adverse effects

Associative modeling
- Extending the scope of ADMET predictions using your own data, without the need to retrain the underlying models

4D Data Mining
- New display in a rotatable 3D scatter plot. Individual points may be colored by a fourth variable.
- An integrated Miner3D graphics component

Run Modes
- Interactive
- Batch
- Command line
- Integrated with Pipeline Pilot™

Accepted Input Formats
- SMILES
- SDF and RDF (ISIS/Base™)
- MaC (Sybyl™)
- MOL (ISIS/Draw™, ChemDraw™)
Examples of Metabolic Site Predictions — from the NEW Metabolite Module!

Many authors, or associates thereof, of metabolic regioselectivity prediction software publish "comparisons" of their product's performance against competing software. Somehow, the author's own product always comes out on top. Usually, the vast majority of compounds used in these "studies" are standard CYP substrates, most of which turn out to be in the compared models' training sets.

We'll leave the side-by-side comparisons of different predictive regioselectivity models to our users and independent researchers. The purpose of this page is the validation of our regioselectivity models on a few concrete examples and the illustration of several factors influencing their predictivity. In the examples shown below we have deliberately chosen compounds predominantly present in the external test sets of our models. Also, we do not shy away from showing a worst case prediction scenario.

**Examples of model selectivity**

Our first example is the molecule of sertraline present in the external test sets of 4 out of 5 our CYP_._Sites models:

<table>
<thead>
<tr>
<th>Model</th>
<th>Membership</th>
</tr>
</thead>
<tbody>
<tr>
<td>1A2</td>
<td>Test Set</td>
</tr>
<tr>
<td>2C19</td>
<td>Training Set</td>
</tr>
<tr>
<td>2C9</td>
<td>Test Set</td>
</tr>
<tr>
<td>2D6</td>
<td>Test Set</td>
</tr>
<tr>
<td>3A4</td>
<td>Test Set</td>
</tr>
</tbody>
</table>

Experimental studies show two major kinds of metabolic attack on sertraline: 1) N-demethylation catalyzed by all 5 CYP isoforms, and 2) deamination leading to a ketone product catalyzed by 2C19 and 3A4 only (Obach et al; 2005):

Our models selectivity follows the observed trend: N-demethylation is predicted for all enzymes as primary target, while deamination is predicted for 2C19, 3A4, and 1A2:

Predicted and marked in red sites of sertraline metabolism

The 1A2-catalyzed deamination looks like a false positive, but a closer examination of Obach's data reveals that ketone formation is catalyzed by 1A2, as well, albeit in trace amounts. Apparently, the kinetics of this reaction is not fast enough to compete with other metabolic pathways.

Observed rates of sertraline ketone formation in recombinant CYP assays. (Obach et al; 2005)
Examples of Metabolic Site Predictions continued

The second example is Terbuthylazine also present in the external test sets of 4 models:

<table>
<thead>
<tr>
<th>Model</th>
<th>Membership</th>
</tr>
</thead>
<tbody>
<tr>
<td>1A2</td>
<td>Test Set</td>
</tr>
<tr>
<td>2C19</td>
<td>Test Set</td>
</tr>
<tr>
<td>2C9</td>
<td>Test Set</td>
</tr>
<tr>
<td>2D6</td>
<td>Test Set</td>
</tr>
<tr>
<td>3A4</td>
<td>Training Set</td>
</tr>
</tbody>
</table>

Lang observed that terbuthylazine was primarily N-dealkylated by all 5 CYPs. (Lang; 1997) Only 1A2 attacked the t-butyl group:

![Terbuthylazine shown with major sites of CYP metabolism.]

A 2004 study reports only the thioether S as a site of metabolic attack by all 5 isoforms of cytochrome (Usmani; 2004):

![Terbuthylazine shown with major sites of CYP metabolism.]

Although our models give this site the highest score, they also predict a large number of false positives. In particular, there is an observed tendency to pick other sulfurs as sites:

![Predicted sites of terbuthylazine metabolism.]

A case of the worst prediction

Sulprofos is a sulfur-rich insecticide present in test sets of all the CYP models:

<table>
<thead>
<tr>
<th>Model</th>
<th>Membership</th>
</tr>
</thead>
<tbody>
<tr>
<td>1A2</td>
<td>Test Set</td>
</tr>
<tr>
<td>2C9</td>
<td>Test Set</td>
</tr>
<tr>
<td>2C19</td>
<td>Test Set</td>
</tr>
<tr>
<td>2D6</td>
<td>Test Set</td>
</tr>
<tr>
<td>3A4</td>
<td>Test Set</td>
</tr>
</tbody>
</table>

![Predicted sites of sulprofos metabolism.]

Why we think our models are robust
In October 2010, early in our model development, we mined available electronic databases of observed metabolic reactions "as is". Our preliminary models produced in some cases "incorrect" predictions that later, after consulting with the original or more recent literature, turned out to be correct. We would like to share two such cases.

All of the examples shown below were in the training sets of our preliminary models. In each case, in spite of the incorrect training data submitted to ADMET Modeler, the ANNE training algorithm refused to accept the wrong sites. This innate resistance to overtraining is the reason why we believe in the robustness of our models.

The first example is metoprolol with only two atoms (O-methyl and benzyl-C) initially reported as sites of 2D6-mediated metabolism (Mautz; 1995):

<table>
<thead>
<tr>
<th>October 2010 Model</th>
<th>Membership</th>
<th>Training Set</th>
</tr>
</thead>
<tbody>
<tr>
<td>2D6</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The third site, N-isopropyl, was submitted to our preliminary CYP2D6 model as negative. Regardless, the model was stubbornly marking this atom as a center of 2D6 attack giving it the highest score! Puzzled by this egregious case of a "false positive", we conducted a literature search and found newer results (Hayhurst; 2001, Colbourne; 1998) These results not only confirmed N-isopropyl as true positive, but that the rate of N-deisopropylation competes with the rate of O-demethylation!

The 2D6 model was later retrained with the correct assignments for metoprolol.

An example at the opposite end, that of a "false negative" that turned out to be a true negative, is dextromethorphan. The setting is the same: October 2010, preliminary 2D6 model, sites of dextromethorphan taken "as is" from the commercial database. Reported three sites of 2D6 metabolism included O-methyl, N-methyl, and one of the ring carbons (Matsunaga; 2009):

<table>
<thead>
<tr>
<th>October 2010 Model</th>
<th>Membership</th>
<th>Training Set</th>
</tr>
</thead>
<tbody>
<tr>
<td>2D6</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The second site, N-methyl, was initially submitted to our preliminary CYP2D6 model as positive. Regardless, the model was stubbornly marking this atom as negative. Puzzled by this egregious case of a "false negative", we conducted a literature search and found newer results (Hayhurst; 2001, Colbourne; 1998) These results not only confirmed N-methyl as true negative, but that the rate of N-deisopropylation competes with the rate of O-demethylation!
Again, the model refused to accept the marked ring carbon as a site of 2D6 attack giving it a very low score, even though this atom was submitted in training as positive. A closer inspection of the Matsunaga et al paper revealed that the indicated carbon is not metabolized by the native CYP2D6, but by its unusual mutant, CYP2D6.49, occurring in some members of the Japanese population. Quoting the authors: "CYP2D6.49 formed a 7-hydroxydextromethorphan, with a roughly similar V(max)/K(m) value to that of O-demethylation." This crucial fact was omitted from the commercial database.

We would like to present the last example illustrating the issue of false positives predicted not only by ours, but in all silico models of regioselectivity. A good example is found in the case of chlorpromazine. Its chemical structure, predicted by our recent model and reported (Yoshii; 2000) sites of metabolism mediated by CYP1A2 are shown below:

<table>
<thead>
<tr>
<th>October’2010 Model</th>
<th>Membership</th>
</tr>
</thead>
<tbody>
<tr>
<td>1A2</td>
<td>Training Set</td>
</tr>
</tbody>
</table>

Predicted and observed (in 2009) sites of chlorpromazine metabolism by CYP1A2.

As of 2009, the only reported metabolite (of which we were aware) was hydroxylation of C-7, identified by the right arrow in the above diagram, mediated by 1A2 and 2D6 (Yoshii; 2000) However, the 1A2 dataset has two similar structures, perazine and promazine, shown below with arrows showing their reported CYP1A2 mediated sites of metabolism (Wójcikowski; 2004, Wójcikowski; 2003)

Observed sites of perazine and promazine metabolism by CYP1A2.

Based on these results, it is not unreasonable for the 1A2 model to predict sites of metabolism for the sulfur and N-methyl carbons of chlorpromazine. Such assignments, as of 2009, would be considered false positives! However, in late 2010, both N-demethylation and sulfoxidation of chlorpromazine were reported as metabolites formed by CYP1A2 (Wójcikowski; 2010). Hence, false positives in the model in 2009 became true positives in 2010! Out of four, only one site (with a low score) thus remains as a false positive prediction, perhaps pending future verification:

Predicted and observed (in 2010) sites of chlorpromazine metabolism by CYP1A2.
The scientific literature contains further discussion on the issue of false positives in metabolic site models (Hennemann; 2009). A false positive corresponds to a model prediction of a metabolite which has not been observed experimentally. This may not mean the metabolite is not produced; only that it has not been observed or reported. These two statements are not the same! In many cases, unknown metabolites are observed, but their chemical structures are not determined. So, a false positive may correspond to such an unidentified metabolite. In other cases, experimenters are interested in a specific metabolite due to its pharmacologic or toxic properties. They may determine which CYP P450 isoforms are responsible for its formation, but are not interested in other metabolites that may be produced by those CYPs. We assign sites of metabolism to the metabolites which are reported and treat the other sites as negatives in the dataset. However, based on the properties of similar molecules, the model may assign sites of metabolism which may actually occur, but have not (yet) been found experimentally.

References


An Illustrated Case Study

Part 1: Model Building

The following case study is typical of situations encountered in the modeling of real data, in which one is trying to build ANNE’s that predict a property of interest using descriptors generated from molecular structures. ADMET Modeler can build models from descriptors generated by ADMET Predictor or by any other source, so long as the data is available in tab-delimited, ASCII format.

The screenshot above shows the Molecular Record spreadsheet after opening a 2D SDF file containing 371 structures plus experimentally determined pharmacokinetic Volumes of Distribution in human (Vd). Such a complex biological property is quite difficult to model. Nevertheless, a useful predictive model for studying variability trends can be built with ADMET Modeler. After the data is loaded, a user chooses Vd as the dependent variable. The program already preselects all ADMET Predictor descriptors in the right panel. The Logify checkbox asks the program to calculate the decimal logarithm of each raw measurement prior to modeling. Each button in the left panel initiates a separate modeling method: Kohonen Mapping (for compound clustering in descriptor space only), regression by Artificial Neural Network Ensembles (ANNE), Support Vector Machine Ensembles (SVME), Kernel Partial Least Squares (KPLS, includes ordinary PLS), and Multiple Linear Regression (MLR). Specific parameters for each available modeling method are accessible through the Adv. Modeller Settings tab. In this run, Input Gradient is chosen as the descriptor selection method and the maximum number of descriptors is limited to 131.

The ANNE button is clicked. After internal filters remove invariant and highly correlated descriptors, 226 are left. This descriptor space becomes the playing field for automatic data division into the training plus verification (306 compounds), and external test (67 compounds) subsets by Kohonen map clustering, which is the default. ANN ensembles are then trained for different network architectures. With default settings the modeling run involves training 160 neural networks and subsequent automatic selection of the 32 best networks for each of the 30 tried architectures. The run takes about 7 minutes on a laptop PC and produces a grid of results where one cell represents one ensemble. A variety of statistics can be used to assess each data set (training (black), verification (red), and test (blue)) of each ensemble. The program also offers automatic selection of the best ensemble signified by the green cell background. Blank entries in the table correspond to ANNE architectures that were skipped because their complexity put them beneath a user-defined acceptable data-parameter ratio.
The ensemble model containing 60 descriptors and 5 neurons appears to have the best all-around statistics, and its performance is illustrated in the plot below. The blue circles denote compounds belonging to the training/verification pool, and the red circles indicate compounds belonging to the external test set. Placing the mouse cursor over any circle on this plot lets the user see its name or ID number. This feature is useful for the rapid identification of outliers.

By clicking the Export Current Ensemble as New Model, the chosen ANNE model is then easily appended to the array of ADMET Predictor models. The whole procedure of model creation from start to finish takes only 10 minutes on a 2.4 GHz laptop PC.
Part 2: Model Applications
The newly added ANNE model of Vd in part 1 of this study is ready for useful practical applications. For example, a synthetic chemist might wonder how to decrease volume of distribution of a new compound, since high Vd usually requires higher doses to keep the same level of free drug in blood plasma. Chloroquine is used here for illustration. The Descriptor Sensitivity Analysis tool immediately characterizes the ANNE output manifold at the chloroquine point in the chemical space. It turns out, Vd is highly sensitive to chloroquine’s ionization characteristics at pH=7.4 signified by ADMET Predictor’s descriptors FCation (fraction cationic) and FUnion (fraction un-ionized). It is clear from Vd model gradient bars that decreasing both fractions (by, e.g., introducing acidic groups) lowers Vd. In summary, increasing values of the most sensitive descriptors influence Vd in the following ways:

- Vd goes up with:
  1. FCation = Fraction cationic at pH=7.4
  2. FUnion = Fraction un-ionized at pH=7.4
  3. N_Fluor = Number of fluorine atoms
  4. TerAmine> = Number of tertiary amine groups

- Vd goes down with:
  1. Pi_MaxFPl = Maximum pi Fukui(+) index (a measure of nucleophilic susceptibility of the chloroquine’s ring system)
  2. MaxQ = Maximum PEOE partial atomic charge
  3. NPA_AQon = Sum of absolute NPA partial atomic charges on oxygens and nitrogens
  4. urea=NC(O)N = Number of urea groups
  5. M_NO = Count of oxygen and nitrogen atoms
  6. M_BLM = Number of beta lactam groups
  7. QAvgNeg = Population average of the aqueous ionized species with net formal negative charge at pH=7.4

The above list provides valuable clues to a medicinal chemist on how to lower the Vd for chloroquine. Because of drug activity concerns, this goal should be achieved with only minimal changes to chloroquine’s molecular structure. Hence, not all of the above descriptors will apply. The general strategy will be lowering the chloroquine’s basicity, preferably by introducing acidic groups (in consistency with FCation, FUnion, QAvgNeg), increasing its polarity, preferably by adding N and/or O atoms (as pointed by MaxQ, NPA_AQon, M_NO), and increasing nucleophilicity of it’s aromatic rings.

An argument supporting this strategy can be made from the experimentally measured Vd values for chloroquine (93 L/kg) and its derivative hydroxychloroquine (50 L/kg). The lower Vd in the latter case is perfectly correlated with lower basicity and higher values of other descriptors discussed above. Let us see if a reduction in Vd can be achieved with a change not as dramatic as the introduction of a hydroxyl group. Following our Descriptor Sensitivity Analysis guidelines, in a few minutes we drew and loaded into ADMET Predictor six proximal derivatives of chloroquine, shown below. The basicity and N,O-dependent polarity have changed appreciably in the desired direction, while other descriptors were influenced only weakly. Derivative D exhibits the largest reduction in Vd thanks to an ether linkage replacing a methylene group leading to both significant reduction in basicity and increase in polarity.
Naturally, the remaining concern is whether the Derivative D is going to be active. Chloroquine has long been used in the treatment or prevention of malaria (see its Wikipedia entry). The most critical step in its mechanism of action is an irreversible diffusion into the digestive vacuole of the *Plasmodium* parasite. This happens because the singly protonated chloroquine in plasma (at pH=7.4) becomes doubly protonated in the acidic environment of the digestive vacuole (pH=4.7). Hence, it is critical that Derivative D maintains these properties. That it is indeed the case as shown in a series of figures on the following page (pKₐ predicted from structure by ADMET Predictor). At pH=4.7 chloroquine exists primarily (99.8%) as doubly protonated species, marked by "2/+2", and at pH=7.4 it is split 55% : 45% between doubly- and singly protonated species, "1/+1" (see next page). The 99.8% - 55% difference accounts for its diffusion gradient. In spite of its lower pKa estimates, at pH=4.7 Derivative D still remains doubly protonated at a healthy 99.6%. Its pH=7.4 distribution is 37% : 63% for "2/+2" : "1/+1" species, respectively. Hence, its diffusion properties should be even better than those of chloroquine.

<table>
<thead>
<tr>
<th>Molecular Data</th>
<th>Basic Pred. pKₐ</th>
<th>%</th>
<th>TS (cal/mol)</th>
<th>FLogP</th>
<th>TPSA</th>
<th>MLogP</th>
<th>NPA</th>
<th>XLogP</th>
<th>NBO</th>
<th>3D Data Mining</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hydroxychloroquine mol</td>
<td>9.02 ± 7.39 ± 0.88</td>
<td>8.71</td>
<td>0.9961</td>
<td>0.0119</td>
<td>0.203</td>
<td>0.21</td>
<td>2.39</td>
<td>4</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Chloroquine mol</td>
<td>9.76 ± 7.47 ± 0.94</td>
<td>14.07</td>
<td>0.998</td>
<td>0.002</td>
<td>0.203</td>
<td>0.13</td>
<td>1.64</td>
<td>3</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Chloroquine derivative A mol</td>
<td>9.71 ± 5.08</td>
<td>12.09</td>
<td>0.9951</td>
<td>0.0046</td>
<td>0.198</td>
<td>0.09</td>
<td>1.50</td>
<td>3</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Chloroquine derivative B mol</td>
<td>9.74 ± 4.59 ± 0.85</td>
<td>14.19</td>
<td>0.9964</td>
<td>0.0046</td>
<td>0.207</td>
<td>0.13</td>
<td>1.67</td>
<td>3</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Chloroquine derivative C mol</td>
<td>9.70 ± 6.50 ± 1.60</td>
<td>13.53</td>
<td>0.9956</td>
<td>0.0044</td>
<td>0.201</td>
<td>0.13</td>
<td>1.64</td>
<td>3</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Chloroquine derivative D mol</td>
<td>9.31 ± 7.08 ± 1.78</td>
<td>12.19</td>
<td>0.9237</td>
<td>0.0763</td>
<td>0.203</td>
<td>0.13</td>
<td>2.3</td>
<td>4</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Chloroquine derivative E mol</td>
<td>9.85 ± 3.13</td>
<td>17.67</td>
<td>0.9965</td>
<td>0.0035</td>
<td>0.216</td>
<td>0.13</td>
<td>1.13</td>
<td>2</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Chloroquine derivative F mol</td>
<td>9.75 ± 6.34 ± 1.64</td>
<td>14.33</td>
<td>0.9959</td>
<td>0.0041</td>
<td>0.20</td>
<td>0.13</td>
<td>1.55</td>
<td>3</td>
<td>0</td>
<td></td>
</tr>
</tbody>
</table>
Graphical user interface
**Molecular Data**
Input data, predicted ADME properties and calculated molecular descriptors are conveniently displayed in a user-friendly Excel-like spreadsheet for inspection, one-click sorting, and editing. Molecular structures are depicted in the first column.

**pKa Prediction**
Because the pKa model predictions carry much more information than the regular ADMET models, its results are displayed in a separate window. Apparent pK_a, complete microspeciation, and microstate percentages are displayed for each macroscopic protonation state. In addition, proton dissociation probabilities and microconstants for each microstate can be presented on demand. No less important is the dependence of individual protonation micro- and macrostates on pH. This, plus the average number of protons and formal charge, are shown in the graphing window featuring the mouse-sensitive marker for instant numerical output.

**Interpretation of Predictive Model Results**
Nowadays, providing mere numerical estimates of ADMET properties is not enough. Much more valuable is answering the question of why a given molecule has a particular value of property X, what structural aspects influence X the most, and giving valuable synthetic hints on how to modify the molecule to improve its ADMET properties. The Descriptor Sensitivity Analysis tool is able to answer these questions through its intuitive interface to descriptor dependence gradients... and individual descriptor dependence graphs.
Detailed Structure Visualization
ADMET Predictor’s Structure Visualization tool is a real treat for chemists who like analyzing their favorite molecules atom by atom. Many important atomic properties can be mapped onto structure depiction.

Run Options
ADMET Predictor runs can be customized on-the-fly with an easy to use set of run options.

Input / Output
ADMET Predictor can read input files containing either 2D or 3D molecular records, interactively or in batch, in one of the following formats:

- SMILES strings
- SDF (ISIS/Base™)
- RDF (ISIS/Base)
- MOL (ISIS/Draw™, ChemDraw™, etc.) or MAC (SYBYL™)

The program preserves a list of most recently open files

ADMET Predictor results can be saved as:

- GastroPlus™ import tables
- Tab-delimited text tables (for MS Excel™ import)
- A copy of the original input file with inserted ADMET properties
- A 2D SDF- or RDF-formatted file with inserted ADMET properties (for, e.g., an easy import into ISIS/Base)
... or discrete spectrum:

Property/Descriptor Correlations
The "Property/Descriptor Correlations" tab displays correlation plots between any two numerical columns read or calculated by ADMET Predictor. Click a button to automatically calculate and display statistics of a linear fit between the two variables. Click on a graph point to display corresponding structure.

A special Plot $pK_a$ button does an automatic pairing of observed vs predicted $pK_a$ (multiple values per module)

allowing for easy and instantaneous evaluation of the $pK_a$ model predictivity.

ADMET Risk Editor
ADMET Risk rules (analog of Lipinski's Rule of Five) are 100% customizable - existing rules can be edited or removed, new rules can be added, and program default rules may be specified.

4D Data Mining
The distribution of input molecules in the chemical space can be visualized with the aid of the 4D Data Mining module. Any one of the X, Y, Z variables can be assigned to either a molecular descriptor column, a predicted property, or a local principal component pertaining to a given chemical dimension. The fourth variable, C, can be used to color data points.

Variables selected in the 4D Data Mining panel are displayed in a rotatable 3D chart. A normalized trend vector shows the magnitude and direction of the fastest increase of the coloring variable C.
Alternatively, the same 3D chart can be displayed with presentation quality in Miner3D™ graphics component built into ADMET Predictor.

On Line Help
The user manual, as well as on-line help files, contains a complete tutorial reducing the learning curve to a bare minimum and making expert users out of beginners in a very short time.

Model Editor
Users can add up to 500 in-house predictive models to the program by simply appending records to the model table in Model Editor window.