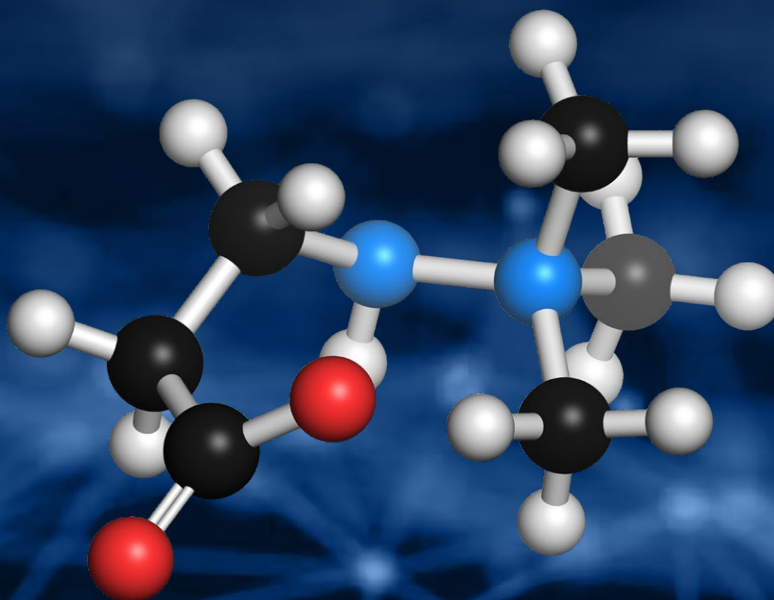


**S+** *SimulationsPlus*  
SCIENCE + SOFTWARE = SUCCESS

# ADMET Predictor®







## AI-Powered ADMET



[www.simulations-plus.com](http://www.simulations-plus.com)



+1-661-723-7723

Connect with us:    

ADMET Predictor is an advanced computer program that enables researchers to rapidly estimate a number of ADMET properties of new chemical entities from their molecular structure.

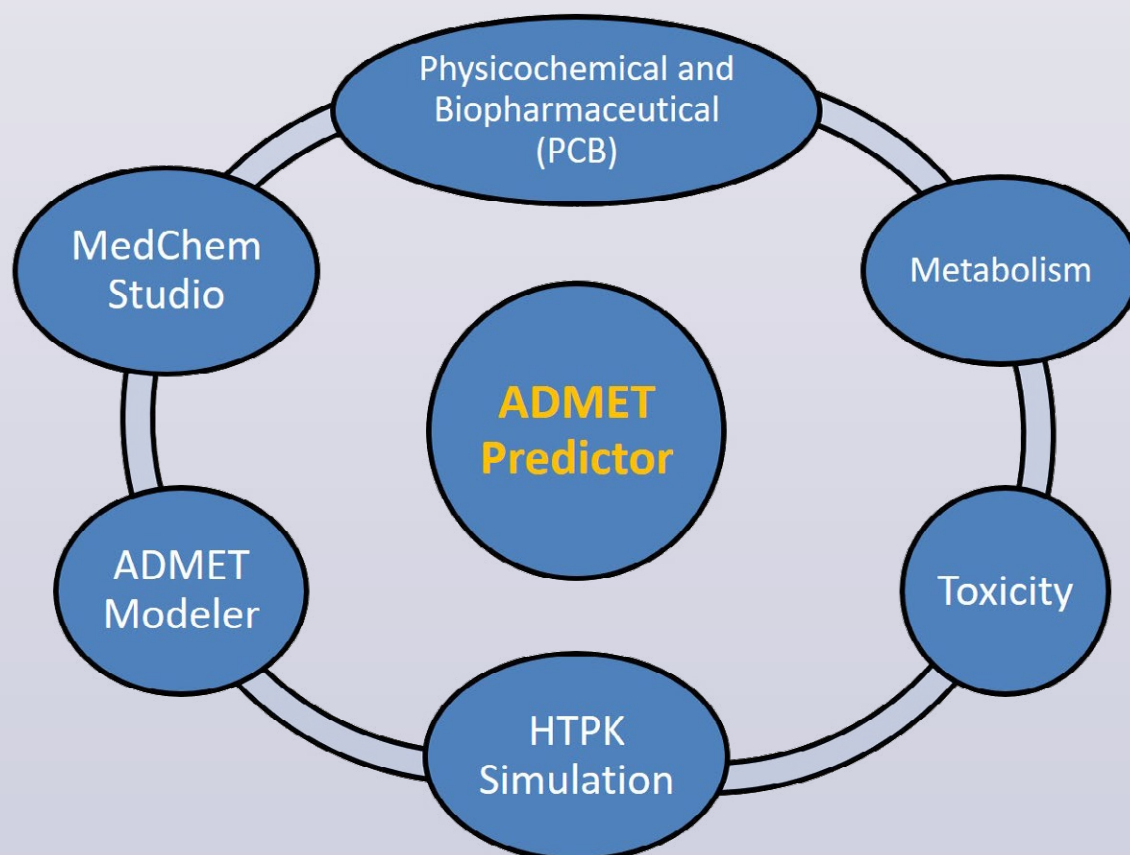
The six (6) modules in ADMET Predictor are shown in the diagram below. Its predictive models are grouped into Physicochemical and Biopharmaceutical (PCB), Metabolism, and Toxicity modules. ADMET Predictor models have been consistently ranked as the most accurate in independent published comparisons.<sup>1-4</sup> Liabilities for one or multiple compounds can be quickly assessed using our ADMET Risk™ scoring feature that combines 24 predicted properties associated with absorption, metabolism, and toxicity.

The HTPK Simulation module predicts fraction absorbed, oral bioavailability, and dose required to achieve a user-specified plasma concentration.

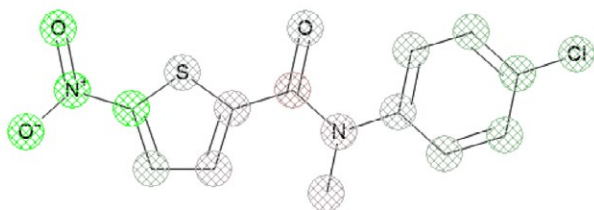
The ADMET Modeler™ module allows researchers to build high quality QSAR/QSPR models using state of the art machine learning algorithms and atomic and molecular descriptors.

The MedChem Studio™ module contains tools to cluster compounds into sets that contain the same scaffold, generate and analyze R Tables, create matched molecular pairs analysis (MMPA), *de novo* structure generation, and much more!

1. Dearden JC. "In silico prediction of aqueous solubility". *Expert Opin Drug Discovery* 2006; 1 (1): 31-52.
2. Tetko IV and Poda GI. in "Molecular Drug Properties: Measurement and Prediction." ed. Mannhold R., Weinheim, Germany, Wiley-VCH: 381-406. (2007).
3. Mannhold R, Poda GI, Ostermann C, Tetko IV. "Calculation of Molecular Lipophilicity: State of the Art and Comparison Methods." *J Pharm Sci.* 2008; 98 (3): 84.
4. Oyarzabal J, Pastor J, Howe TJ. "Optimizing the Performance of *In Silico* ADMET General Models According to Local Requirements"; *J. Chem. Model* 2009; 49 (11): 2572.



146795-38-2 -- Ames mutagenic



Aromatic nitro flagged for Ames mutagenicity

Ames [R/G - G=Toxic]

**General**

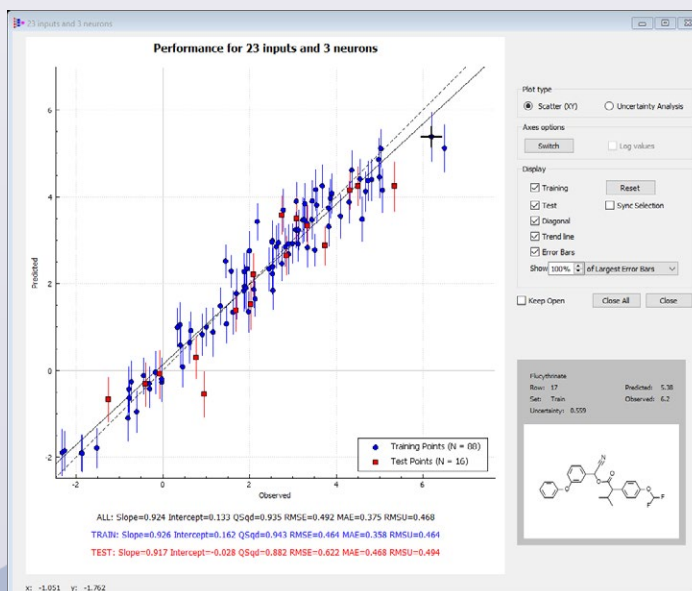
- Structure sensitivity analysis: Highlight fragments that contribute the most to the prediction
- Uncertainty analysis for regression models
- Enhanced metabolism predictions
  - Esterase, UGT, and aldehyde oxidase metabolites
- New models to predict Ames mutagenicity
- Bile salt export pump (BSEP) inhibition, organic cation (OCT2) transporter 2 inhibition, and breast cancer resistance protein (BCRP) substrate models
- Updated Pipeline Pilot and KNIME components
- Regression model prediction will include uncertainty estimates

**MedChem Designer Improvements**

- pKa microstate display
- Solubility and logD versus pH profiles
- HTPK module properties and plots
  - %Fa and %Fb
  - Cp-time profile
  - Parameter sensitivity analysis

**ADMET Modeler Module**

- Regression uncertainty analysis is automatically performed
- Error bars can be turned on and off
- Cumulative distribution of ensemble standard deviation and Q-Q plots are available



**Interface features and spreadsheet functions**

Customizable star plots

Structure	Identifier	Star(Risks)	Star(PCB)	Star(CYP Sub)	Star(CYP CLint)	Star(UGT Subst)	Star(Ames Mut)	Star(Tox)
	alprazolam							
	amlodipine							
	aprepitant							
	atomoxetine							
	bupropione							
	carbamazepine							
	chlorpheniramine							
	chlorpromazine							
	clonazepam							

Advanced text & structure query options

Distribution, scatter, & 3D plots



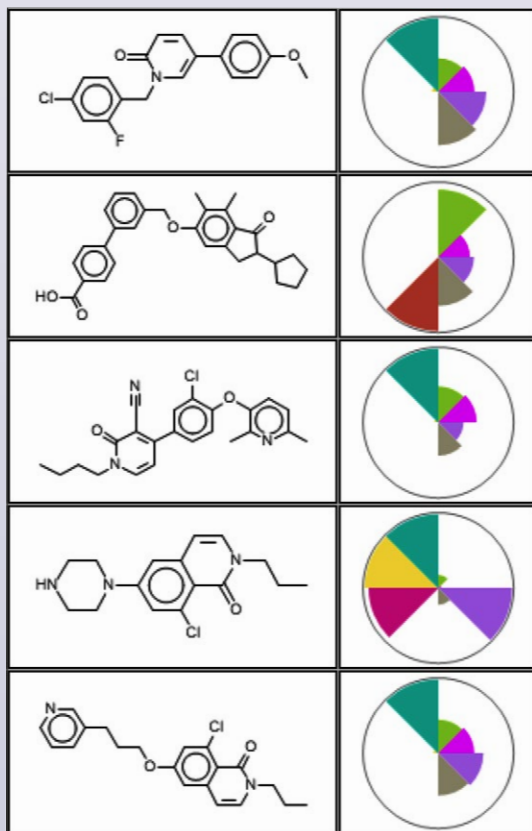
## Heat maps

Heat maps can be created for any numeric attribute. The color progression can be red-yellow-green or blue-white-red. One can also specify the direction of the progression, e.g., the largest values are colored blue.

Structure	Identifier	Solubility [ug/mL]	S+logP	N_Atoms	N_AromR
	10037895	1.200	3.420	19	3
	10467247	17.700	1.757	18	1
	11838365	11.700	3.315	23	2
	12143299	19.000	3.057	22	3
	12230894	0.200	3.237	24	3
	15793062	1.700	1.874	20	1
	16187949	20.000	4.014	22	2
	16189514	51.100	4.301	24	3
	16192552	2.000	3.498	23	3
	16194448	0.400	3.416	27	3
	16194448	27.100	3.860	21	3

Structure	Identifier	Solubility [ug/mL]	S+logP	N_Atoms	N_AromR
	10037895	1.300	3.420	19	3
	10467247	17.700	1.757	18	1
	11838365	11.700	3.315	23	2
	12143299	19.000	3.057	22	3
	12230894	0.200	3.237	24	3
	15793062	1.700	1.874	20	1
	16187949	20.000	4.014	22	2
	16189514	51.100	4.301	24	3
	16192552	2.000	3.498	23	3
	16194448	0.400	3.416	27	3
	16194448	27.100	3.860	21	3

## Automatic Star Plots



Star plots are automatically created after ADMET properties are computed. One can specify up to 16 properties per Star plot. The length of the wedge is proportional to the property range represented by the data. Optionally, one can specify the range minimum and maximum values for each wedge. Additionally, the wedge length can be proportional to the log of the property value.

Star plot

Selected attributes will be represented by colored wedges, which will be arranged in clockwise order starting at the top.

Available attributes

- % Potentiation
- ABSQ
- ABSQon
- AcidAtoms
- ADMET\_Code
- ADMET\_Risk
- AlCbxyL\_COOH
- AlHdxL\_OH
- AlphaAA
- AlphaAE
- Amide\_CI(=O)N<

Selected attributes

- S+logP
- S+logD
- S+Sw
- S+Peff
- FAnion
- FCation
- S+PrUnbnd
- S+BBB\_Filter

Attribute: S+logP

Color:  Change

Min: 1.527

Max: 8.013

Log Scale

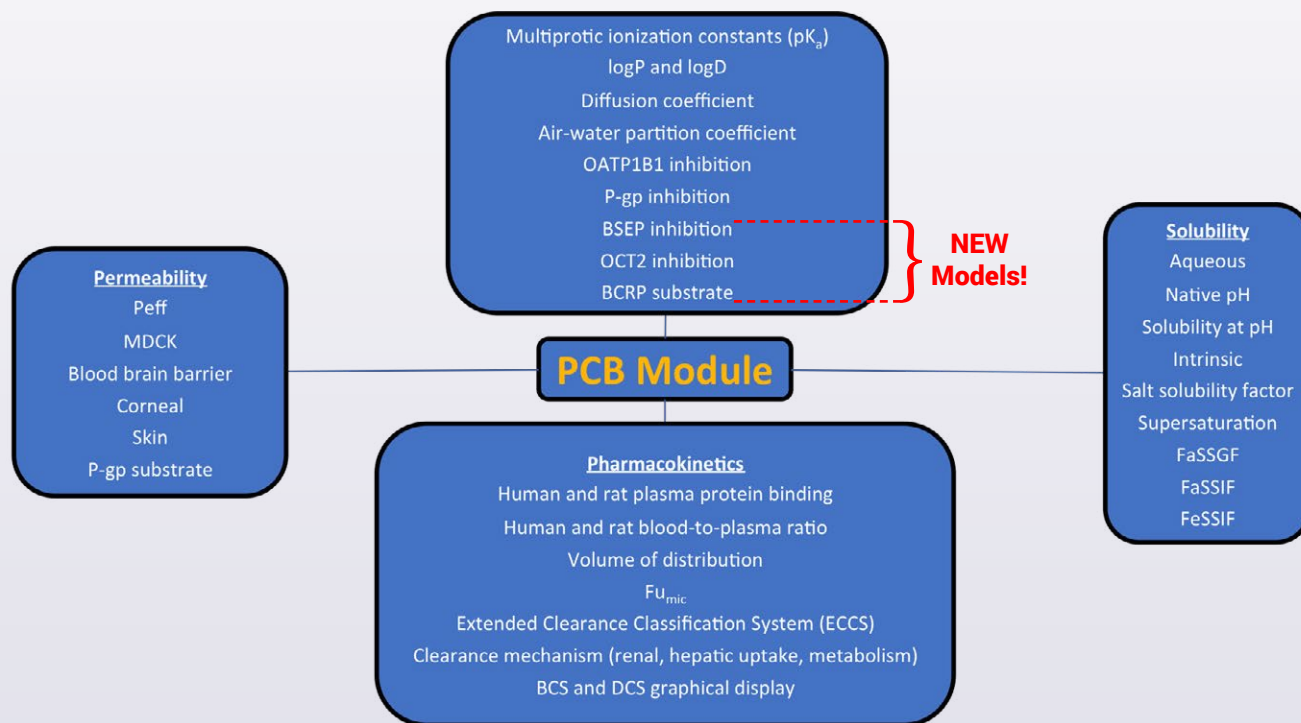
Column name (optional)

Star(PCB)

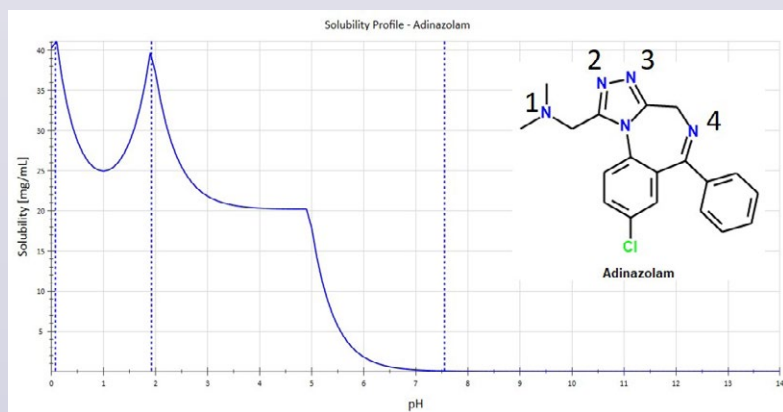
Log Scale For All Log Scale For None

OK Cancel

The models in the PCB module are shown in the figure below.

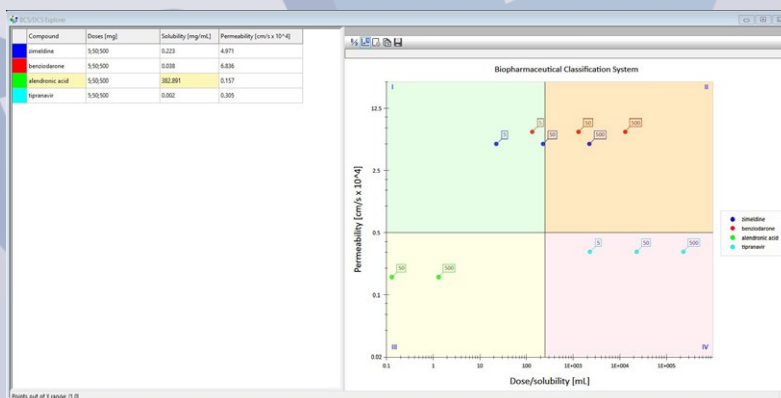


Graphics in ADMET Predictor 9.5 include sol. v. pH & BCS shown in the images below.

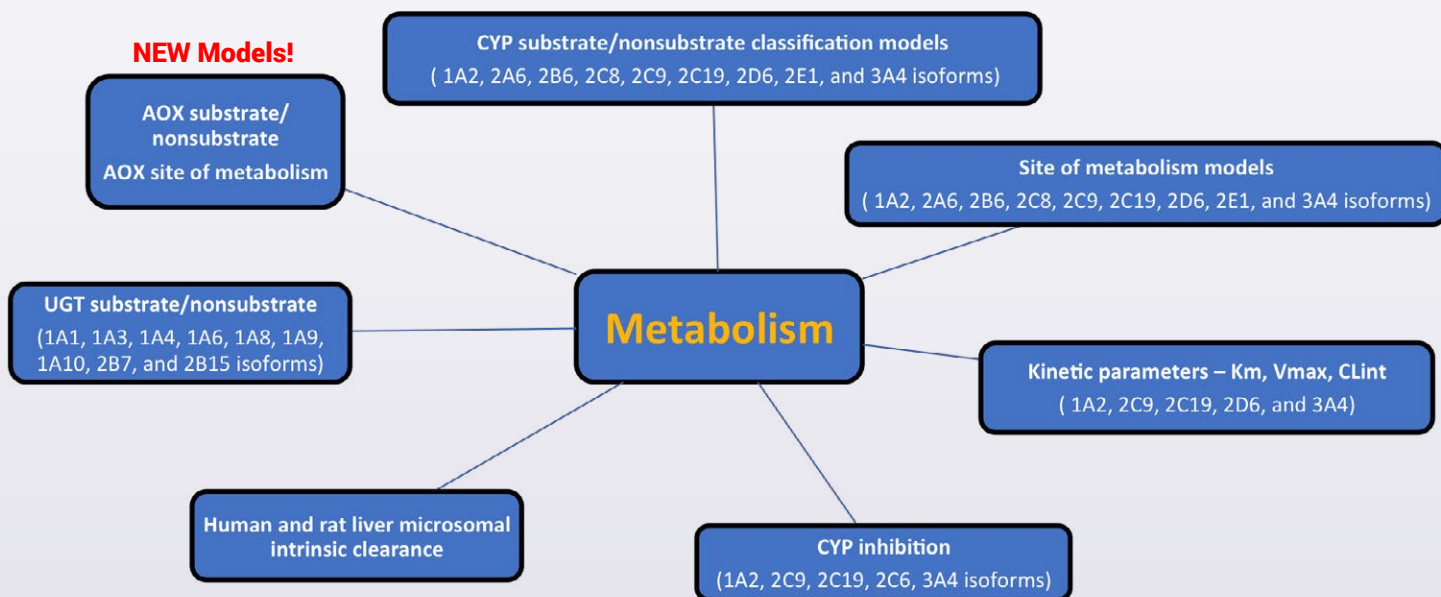


The predicted solubility versus pH profile for adinazolam. There are 4 basic amines, labeled 1 – 4. The vertical dashed lines represent  $pK_a$  of the macrostates. The highest  $pK_a$  is 7.56, where the solubility starts to increase with decreasing pH. Macrostates that include protonation of either nitrogens 3 or 4, in addition to nitrogen 1, begin to affect the solubility starting at just below pH 5.

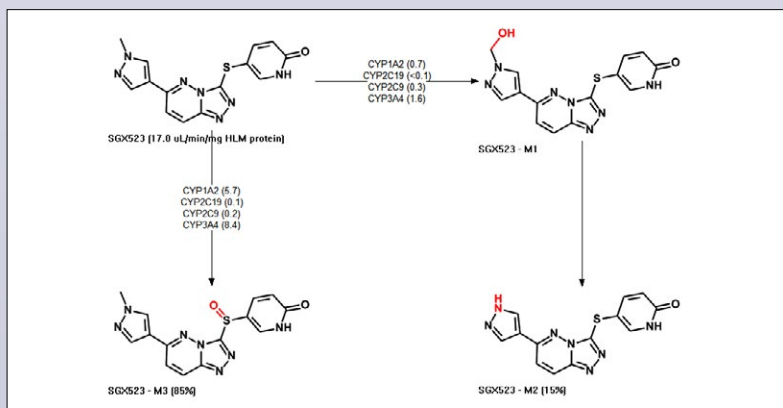
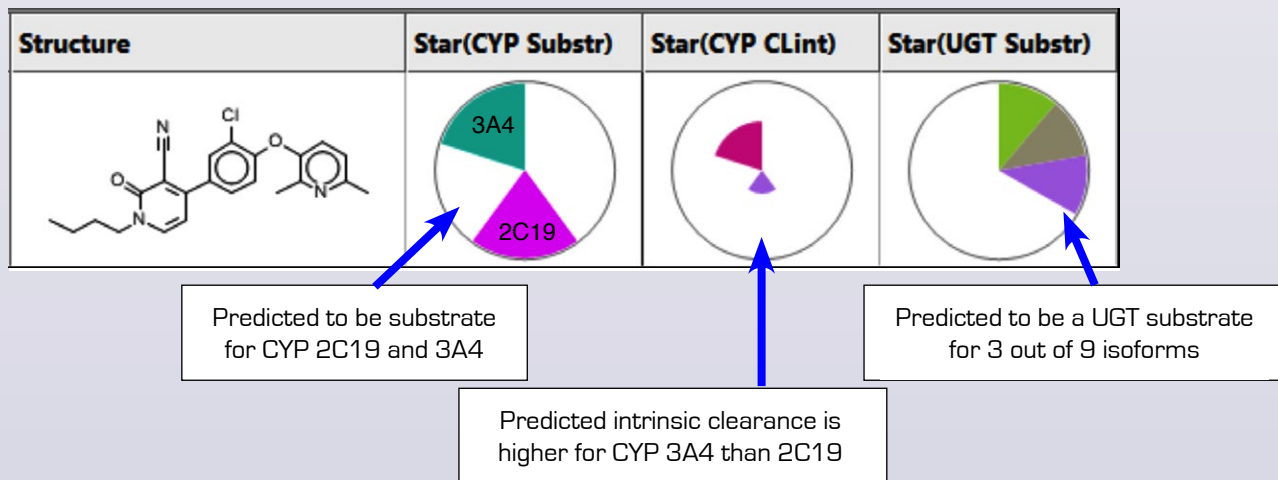
The Biopharmaceutical Classification System (BCS) viewer allows one to plot the BCS class of up to sixteen (16) compounds at user specified doses.



The models in the Metabolism module are shown in the image below.

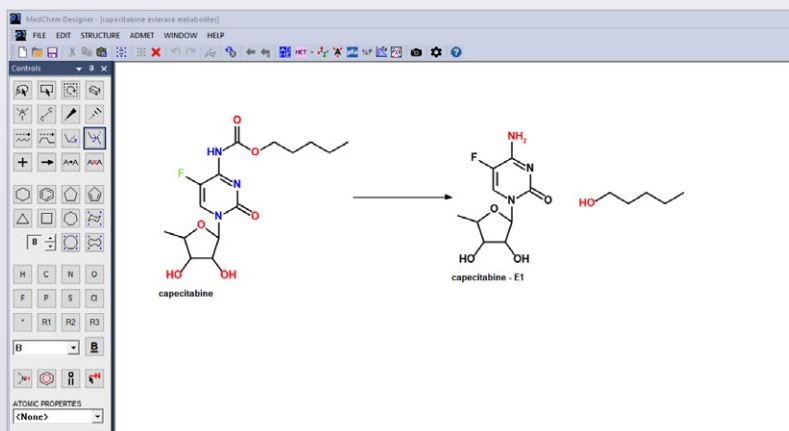
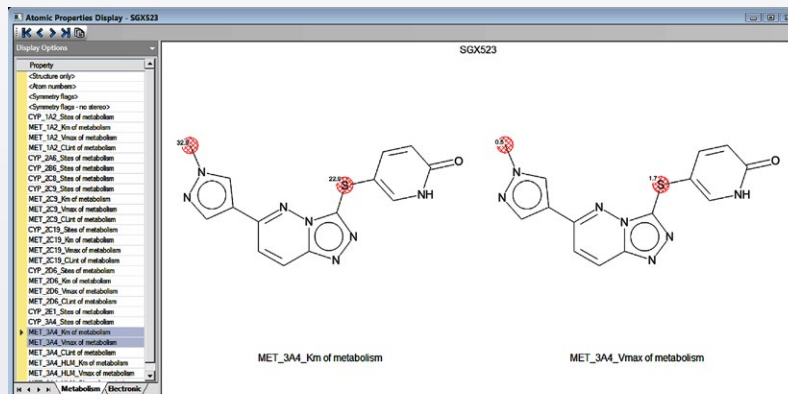


“Star” plots allow one to quickly assess model results.



Metabolite prediction: the  $CL_{int}$  of the parent molecule is displayed along with the percentage of each metabolite produced. Hydroxylation of the methyl group results in an unstable intermediate that decomposes to the desmethyl product and formaldehyde (not shown).

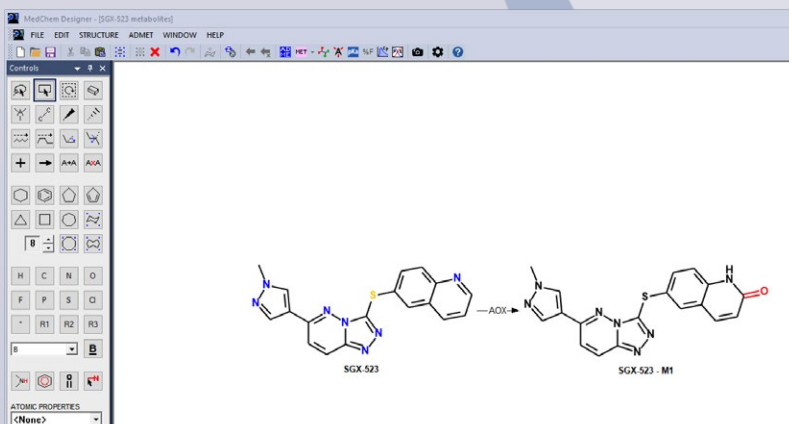
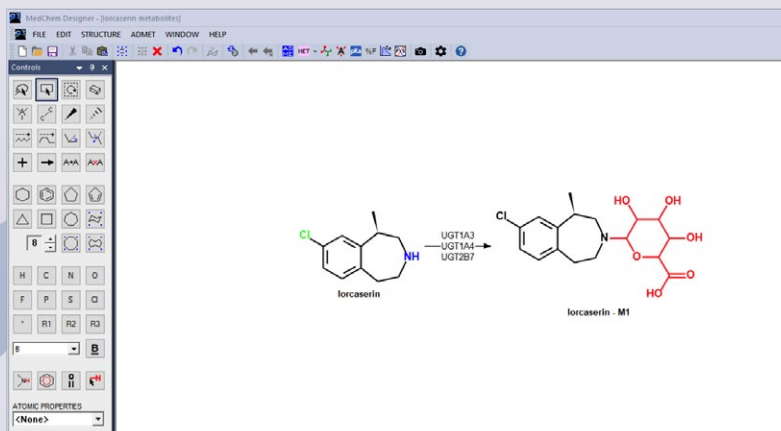
ADMET Predictor predicts Km and Vmax for each site of metabolism.



Cleavage of the carbamate group by an esterase is the first step in activation of the prodrug capecitabine.

- Expert rules help ensure that only legitimate substrates produce metabolites
- Users can easily customize rules
- Generate esterase metabolites

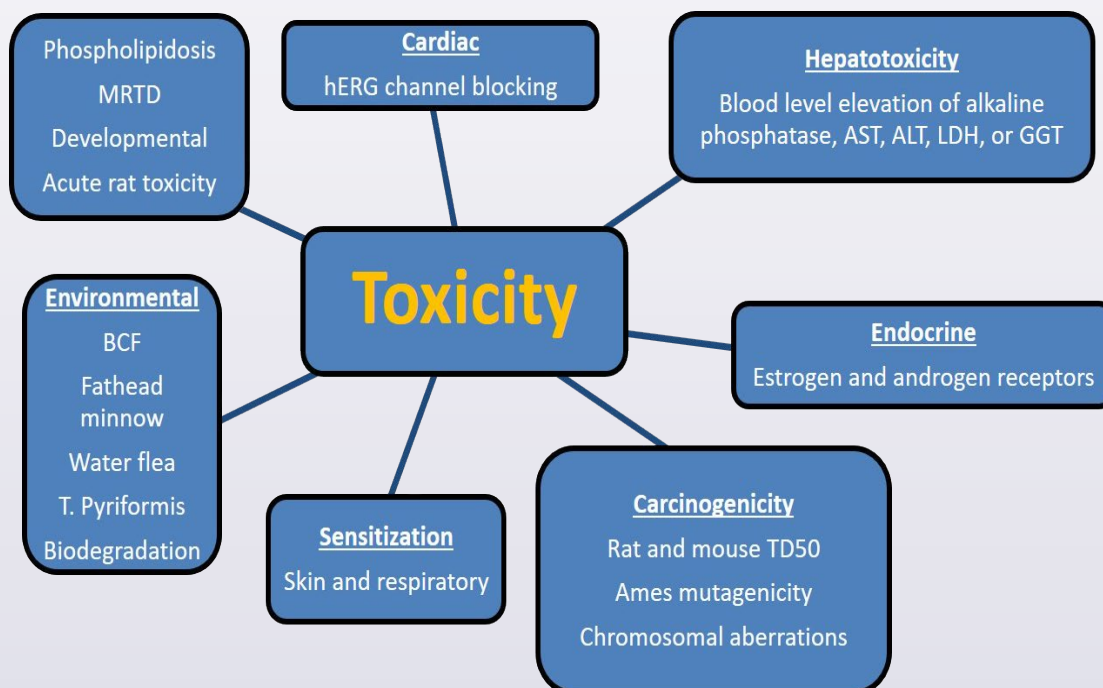
MedChem Designer will now display **glucuronidated metabolites**.



AOX\_Substr and AOX\_Sites models are applied to a molecule and **AOX metabolites** are displayed in MedChem Designer.

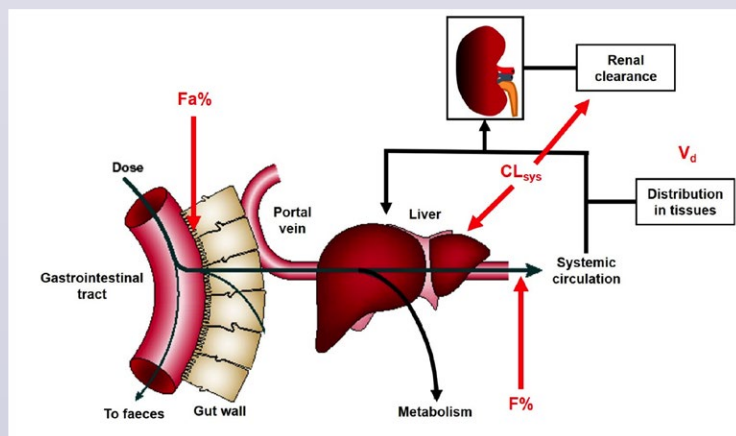
## Toxicity Module

No early compound candidate screening tool should neglect toxicity aspects. Living up to its name, ADMET Predictor features a rapidly growing array of toxicity prediction models. The module features models covering a large range of toxicities included cardiac, hepatotoxicity, endocrine, carcinogenicity, sensitivity and environmental.



## HTPK Simulation Module

The HTPK Simulation Module combines the **#1-ranked mechanistic absorption/PBPK models** with the **#1-ranked ADME property predictions** in order to predict the following *in vivo* properties in **rats** and **humans**:



- Fraction absorbed ( $F_a\%$ )
- Oral bioavailability ( $F_o\%$ )
- Mechanistic Volume of distribution ( $V_d$ )
- Dose ( $D$ ) needed to reach a user-specified plasma concentration ( $C_{eff}$ )
- $C_{max}$
- $T_{max}$
- AUC

The image on the left illustrates processes that occur *in vivo*. The gastrointestinal tract is represented by our ACAT™ (Advanced

Compartmental Absorption and Transit) model that consists of nine compartments. The drug molecule is represented in the unreleased, undissolved, and dissolved states. Molecules can be absorbed into the enterocyte via trans and para cellular diffusion. First pass liver clearance is computed using our rat or human intrinsic clearance models. Users can also supply *in vitro* liver microsomal or hepatocyte intrinsic clearance values. Sixteen compartments, e.g., adipose, brain, heart, are used in the PBPK model to compute volume of distribution ( $V_d$ ). Systemic clearance is computed from liver and kidney clearance or it can be provided by the user. These values are incorporated in the prediction of the dose required to reach a user specified plasma concentration.



ADMET Modeler automates the difficult and tedious process of developing high quality predictive structure-property models from experimental data. It works seamlessly with ADMET Predictor structural descriptors as inputs and appends the selected final models back to ADMET Predictor as an additional predicted property. Below are three keys to developing good predictive models:

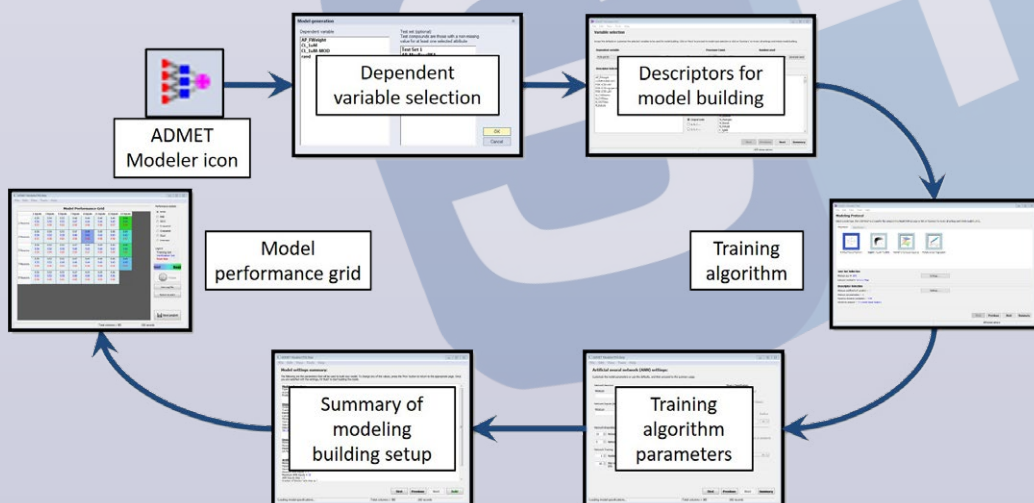
1. Clean and consistent data
2. Appropriate and accurate descriptors
3. Good training algorithms

ADMET Predictor provides many tools for assessing the quality of the data. For example, it is easy to find duplicate structures within a spreadsheet. If duplicates exist, do they have the same experimental value. Matched molecular pair analysis can be used to find pairs of molecules that have high molecular similarity and very different experiment end points. It is important to verify the experimental results in these cases. ADMET Predictor computes over 300 descriptors from the 2D structure of a molecule. This includes atomic and reactivity descriptors produced by our EEM [Electronegativity Equalization Method] Hückel model that is parameterized to reproduce high level ab initio calculations. Training algorithms include artificial neural network (ANN), support vector machine (SVM), kernel partial least squares (KPLS), and multiple linear regression (MLR).

The diagram below illustrates the intuitive model building workflow in ADMET Modeler. The first step is to select the dependent variable, i.e., the property you'd like to model. Next, one can select the descriptors to use during the model building process. Users can select models, e.g., S+logP as descriptors in the new model. Additionally, one can select their own descriptors, provided that they are present in the spreadsheet. The next step is to select the training algorithm and parameters for model building. A summary of the setup is presented and if it is acceptable then models are built using the specified algorithm.

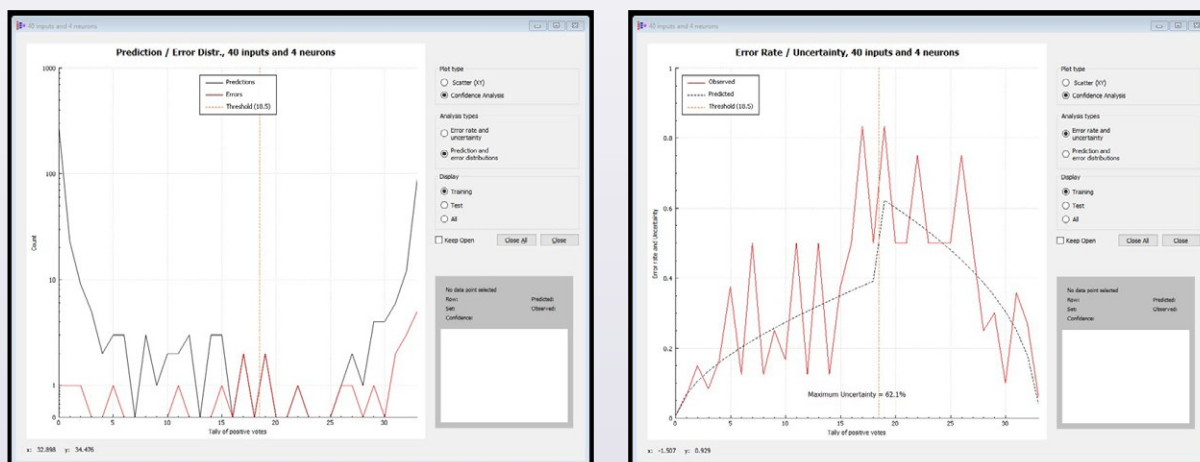
Additional features:

- Multi-class (greater than 2) classification models can now be built using Artificial Neural Network Ensembles (ANNs)
- Performance plots for binary classification models can now be colored coded by confidence value
- DELTA model approach allows users to incorporate in-house data into our global models



## Confidence Analysis

Confidence analysis is based on the degree of concordance among the individual neural networks in an artificial neural network ensemble (ANNE) model. An example of the analysis is shown below. The image on the left is a plot of the count of total predictions (black line) and number of incorrect predictions (red line) at each “count of positive votes” [x-axis]. There are over 100 predictions that have zero (0) positive votes, i.e., all of the 33 artificial neural networks predicted a negative. The red line shows that only one (1) of these predictions was incorrect. This data is converted to an error rate in the plot on the right. The dashed line represents two beta binomials, one for the negatives and one for the positives. Confidence estimates for arbitrary compounds are based on the number of positive votes and then derived from this type of plot.

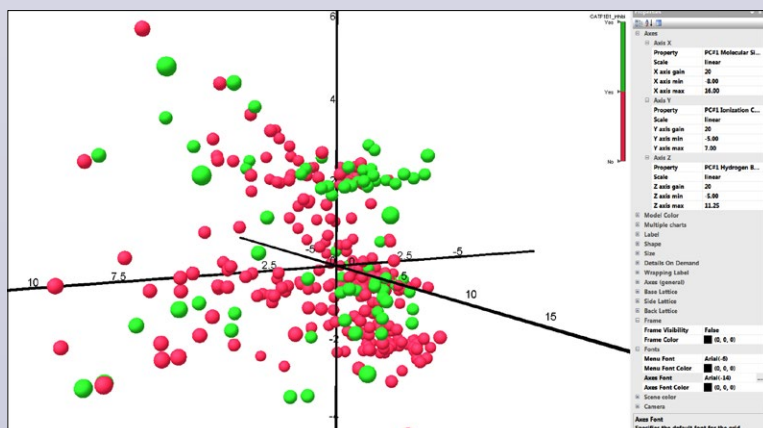


## Customizable ADMET Risk™ Filters

A drug molecule's potential liability can be quickly assessed using our ADMET Risk scoring function. Property predictions were calculated for a subset of greater than 2,000 compounds from the World Drug Index (WDI). Cutoff values for each property were determined by considering the overall distribution of the property and setting the threshold so that 10% of the WDI compounds would be removed. The score is increased by one for each prediction that falls outside of the acceptable boundaries.

- **Absorption Risk** - Risk of low absorption from an oral dose
  - Incorporates size, charge, number of rotatable bonds, h-bond acceptors and donors, lipophilicity, solubility and permeability
- **CYP Risk** - Risk of high clearance or possible CYP inhibition
- **TOX Risk** - Risk of overall toxicity
  - hERG inhibition, acute rat or mouse toxicity, hepatotoxicity, and mutagenicity
- Global **ADMET Risk** summarizing all of the above in one
- Scoring functions can be customized to incorporate user's preferred cutoffs or properties

## 4D Data Mining

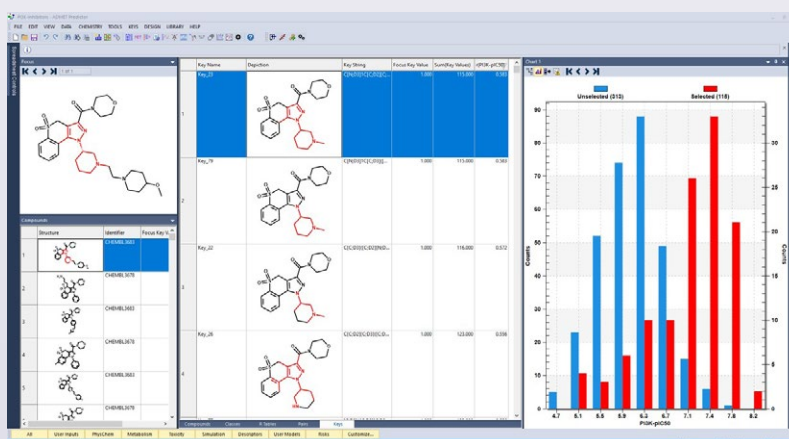


Molecular descriptors calculated by ADMET Predictor can be used to visualize chemical space. The graph above shows the principal components of molecular size and shape, ionization characteristics, and hydrogen bonding for a set of compounds that were tested for OATP1B1 inhibition. The green colored points are OATP1B1 inhibitors.

In ADMET Predictor 9.5, one can compute various types of fingerprints including Extended-Connectivity Fingerprints (ECFPs), internal (MACCS) fingerprints, unbranched fragments, and ring systems. Each fingerprint is represented by a row in the “Keys” tab. The portion of the molecule corresponding to the key is highlighted in red and the key in SMARTS format is displayed in the “Key String” column. The Pearson correlation coefficient between the fingerprint and a property, e.g. IC50, can be added as a column. The image below shows some of the fingerprints for a dataset of 428 PI3K inhibitors. The highlighted fragment (Key\_23) is present in 115 out of the 428 compounds in the data set. The red bars in the PI3K-pIC50 chart represent the pIC50 values for these compounds. Compounds that contain this fragment have higher PI3K IC50 values, e.g., they are more active.

Additional features:

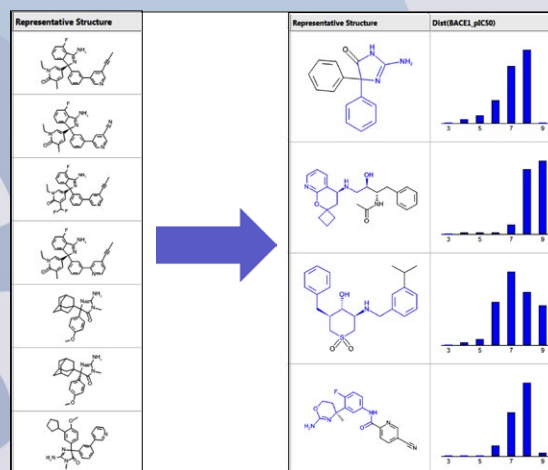
- Functionality to compute fingerprints, including Extended-Connectivity Fingerprints (ECFPs)
- Visualize fingerprints using new interface that highlights atoms on the molecule that correspond to the fingerprint
- Add column containing Pearson correlation coefficient between fingerprint and molecular property, e.g., IC<sub>50</sub> value



MedChem Studio’s class generation technology reproduces a chemist’s reasoning by automatically organizing molecules into chemotypes based on shared scaffolds:

- Uses maximum common substructures (MCSs) rather than fingerprints, so results are chemically intuitive
- Generates scaffolds from the data rather than reading them from a pre-defined list, so novel structural motifs can be discovered
- Provides an ideal starting point for local QSAR generation, molecule design, and other analysis tasks
- User-defined scaffolds can also be used to generate classes
- “Frameworks” class generation option similar to the Murcko assemblies method

Example: The MedChem Studio Module was used to group 718 BACE1 inhibitors into 27 families containing common scaffolds. The largest class contains 243 compounds.



Each row corresponds to a unique class. A representative molecule in the class is displayed with the scaffold highlighted in blue. A distribution chart of the pIC50 values for the compounds in the class is also shown. Property values for the class can be shown as the average, minimum, maximum, or a variety of other statistical functions.

Structure 1	Structure 2	BACE1_pIC50 1	BACE1_pIC50 2	Change(BACE1_pIC50)
		6.553	5.493	1.059
		6.585	7.658	1.073
		5.777	6.886	1.109
		4.086	5.215	1.128

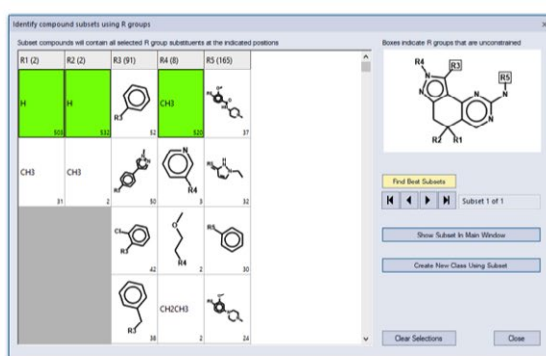
The pairs above result in less dramatic activity differences but they illustrate interesting SAR.

Matched molecular pairs (MMP) can be used to identify activity cliffs in a data set or to perform an analysis to identify and possibly exploit structural trends in property values:

- A MMP is a pair of molecules with a small, specific structural transformation
- MMP analysis (MMPA) is the automated identification of matched molecule pairs and analysis of the change in property values between the pairs
- The average change in property value is the expected size of the property value when the transformation is applied to another molecule
- MedChem Studio automatically creates a SMIRKS string that encodes each transformation rule
- These rules can be added to our “Combinatorial Transform...” feature in order to generate novel molecules potentially containing the desired change in property value

“Find Subsets using R-groups” finds the largest subset(s) of molecules with only two R-groups. In the example below, R1=R2=H and R4=CH3 represents 489 of the 534 compounds. This simplifies the data set and allows one to “drill down” into the data to understand SAR associated with R3 and R5.

Structure	Identifier	R1 (2)	R2 (2)	R3 (91)	R4 (8)	R5 (165)	R Pattern	R Count	Family	IGFR_pIC50	AP_FWeight
	CHEMBL3629	H	H		CH3		R1_4_5	3		6.682	427.509
	CHEMBL3629	H	H		CH3		R1_4_5	3		9.523	397.891
	CHEMBL3629	H	H		CH3		R1_4_5	3		7.721	415.301
	CHEMBL3629	H	H		CH3		R1_4_5	3		6.301	445.911
	CHEMBL3629	H	H		CH3		R1_4_5	3		7.194	517.676
	CHEMBL3629	H	H		CH3		R1_4_5	3		7.197	455.520
	CHEMBL3629	H	H		CH3		R1_4_5	3		7.602	475.938



## License management with FlexNet Manager™ from Flexera Software LLC

ADMET Predictor uses the FlexNet Manager software for license management. This offers several advantages over our previous method:

- Program files are not split between server and client machines
- Improved license administration and handling between applications
  - Centralizes license structure: multiple vendor license files can be hosted by the same server
  - Monitor license status and usage details

- Control access to the Flexera license server by configuring the communications port
- Efficiently perform license server administration (start, stop, restart license servers, etc.) in real time

- Ease of implementation for companies already using Flexera based applications
- Ease of obtaining a license
  - No need to call in order to exchange licensing codes
  - Licenses will be made available through download
- Reduces license denials by allowing server redundancies