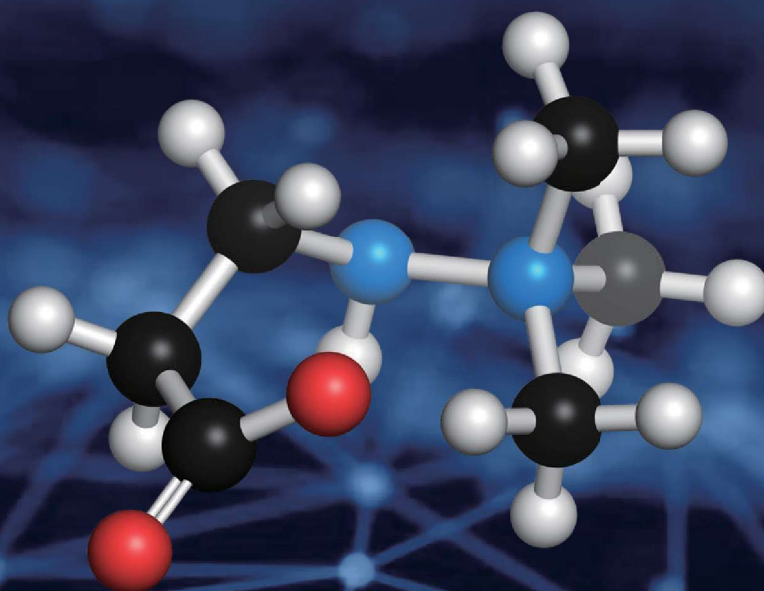


S+ *SimulationsPlus*
SCIENCE + SOFTWARE = SUCCESS

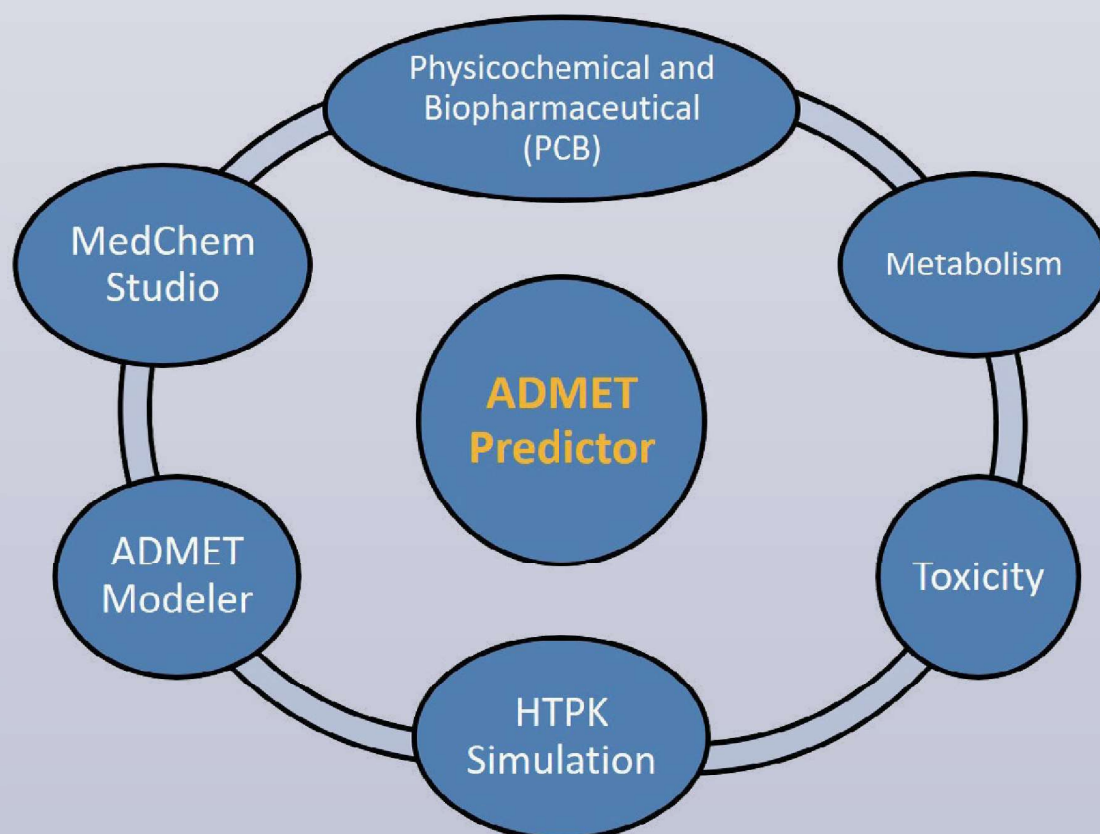
ADMET Predictor®



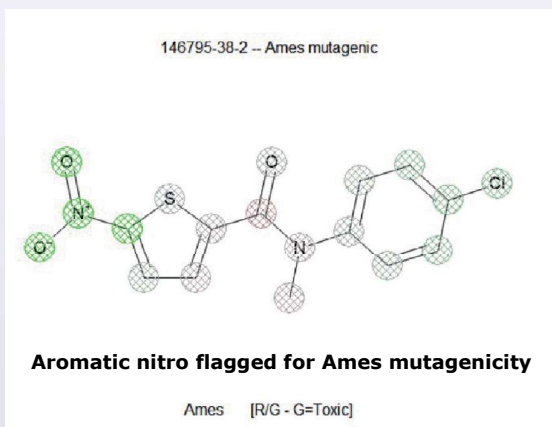
 *Northern Science Consulting Inc*

ADMET Predictor とは？

ADMET Predictor は、新規化学物質の数多くの **ADMET** プロパティをその分子構造から瞬時に予測する最新のソフトウェアです。 **ADMET Predictor** にある **6** モジュールを以下の図に示します。 **Physicochemical and Biopharmaceutical (PCB)**、**Metabolism**、**Toxicity** の各モジュールにその予測モデルが分類されています。 **ADMET Predictor** のモデルは、公開されている精度比較で最も正確であるソフトウェアとして一貫してランク付けされています。 吸収、代謝、毒性に関連する **24** もの予測プロパティを組み合わせた **ADMET Risk™** スコアリング機能を使用して、化合物の不利となる面を素早く評価することが可能です。 **HTPK Simulation** モジュールは、吸収率、経口バイオアベイラビリティ、ユーザー指定の血中濃度に達するために必要となる投与量を予測します。 **ADMET Modeler** モジュールでは、最先端の機械学習アルゴリズムと、原子・分子ディスクリプターを用いて、質の高い **QSAR/QAPR** モデルの構築を実現します。 **MedChem Studio** モジュールでは、同じ骨格構造もつ化合物セットのクラスタリング、**R** テーブルの発生と分析、**Matched Molecular Pairs Analysis (MMPA)** の生成、**de novo** 構造の生成などを行う数多くのツール群が提供されています。



ADMET Predictor 9.5 の新機能



全般：

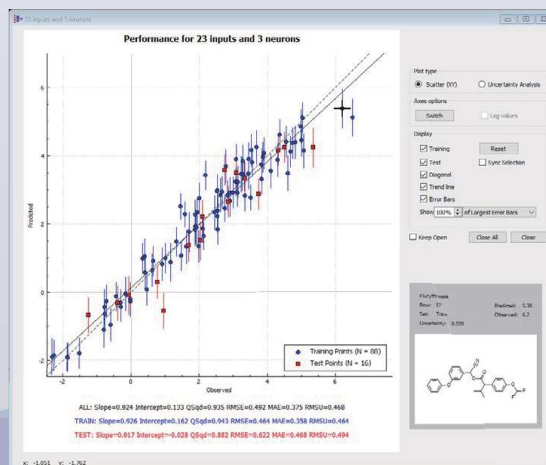
- 構造感度分析：予測に強く寄与しているフラグメントの視覚化
- 回帰モデルの不確かさ分析
- 強化された代謝産物予測 (**Esterase, UGT, AOX**)
- **Ames** 変異原性を予測する新規モデルの搭載
- **BSE** トランスポーター阻害、**OCT2** トランスポーター 2 阻害、**BCRP** 基質モデル
- 最新の **Pipeline Pilot** 及び **KNIME** コンポーネント
- 回帰モデル予測に不確かさ推定の追加
- **Ames** 変異原性のフラグが印された芳香族ニトロ化合物

MedChem Designer の改良：

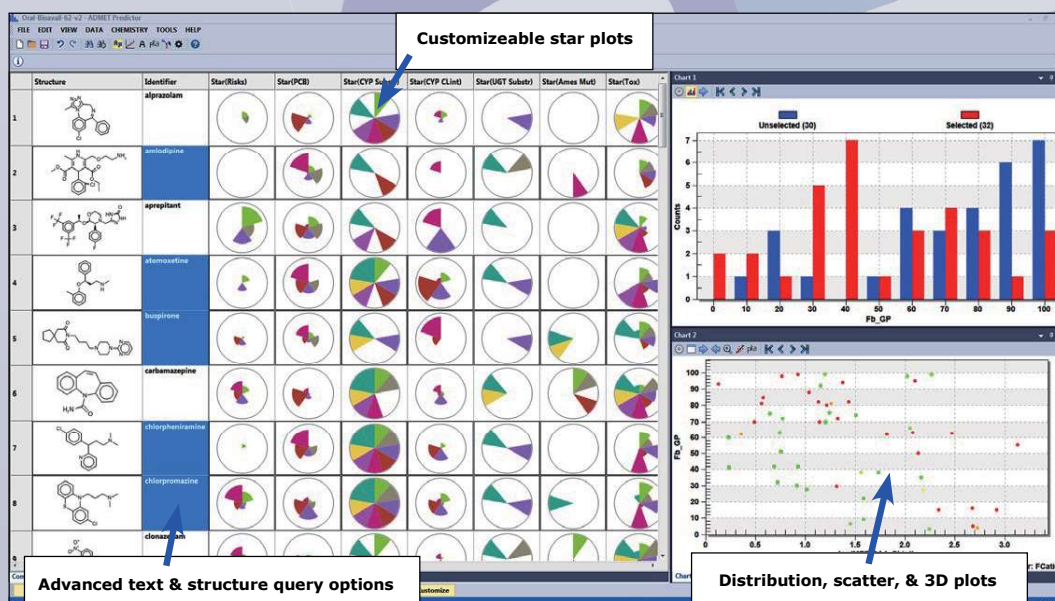
- **pKa** マイクロステート表示
- **Solubility vs. pH** プロファイル、**logD vs. pH** プロファイル
- **HTPK** モジュールプロパティとプロット
 - %Fa および %Fb
 - Cp-time プロファイル
 - パラメーター感度分析

ADMET Modeler モジュール

- 回帰不確かさ分析の自動実行
- エラーバー表示のオンオフ
- アンサンブル標準偏差と **Q-Q** プロットの累積分布が利用可能



インターフェイスとスプレッドシート



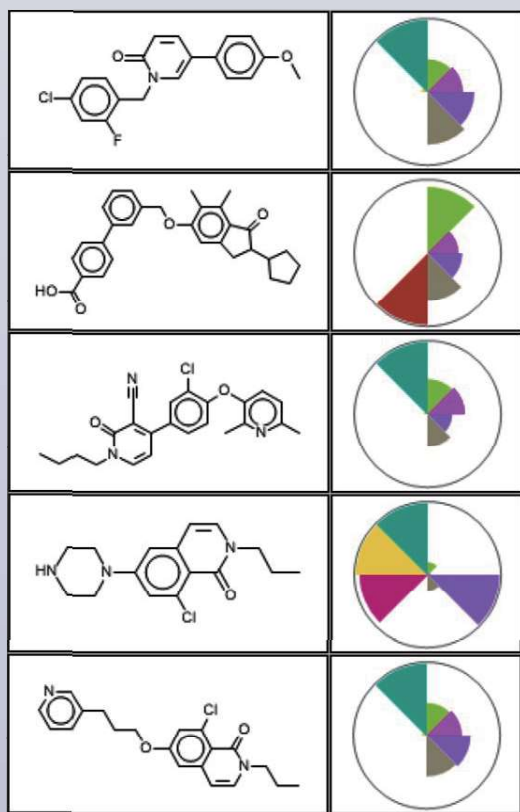
ヒートマップ

属性が数値であればヒートマップを作成することができます。色の連続変化は、赤黄緑または青白赤のいずれかであり、変化の方向を指定することもできます。この例では、最大値は青で表示されています。

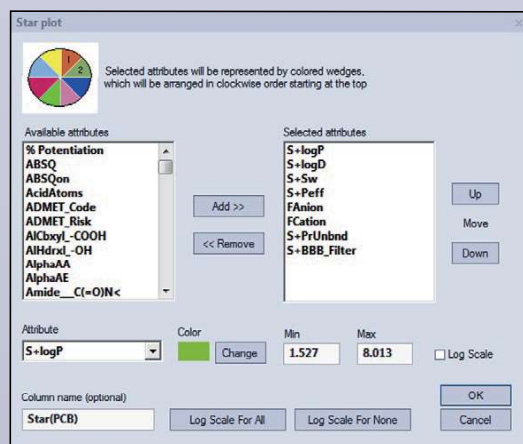
Structure	Identifier	Solubility (ug/ml)	S+logP	N_Atoms	N_AromR
	10037895	1.300	3.420	19	3
	10467247	17.700	1.757	18	1
	11838365	11.700	3.315	23	2
	12143299	19.000	3.057	22	3
	12230894	0.200	3.237	24	3
	15793062	1.700	1.874	20	1
	16187949	20.000	4.014	22	2
	16189514	51.100	4.301	24	3
	16192552	2.000	3.498	23	3
	16194448	0.400	3.416	27	3
	16194448	27.100	3.950	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	3.950	21	3

スタープロット

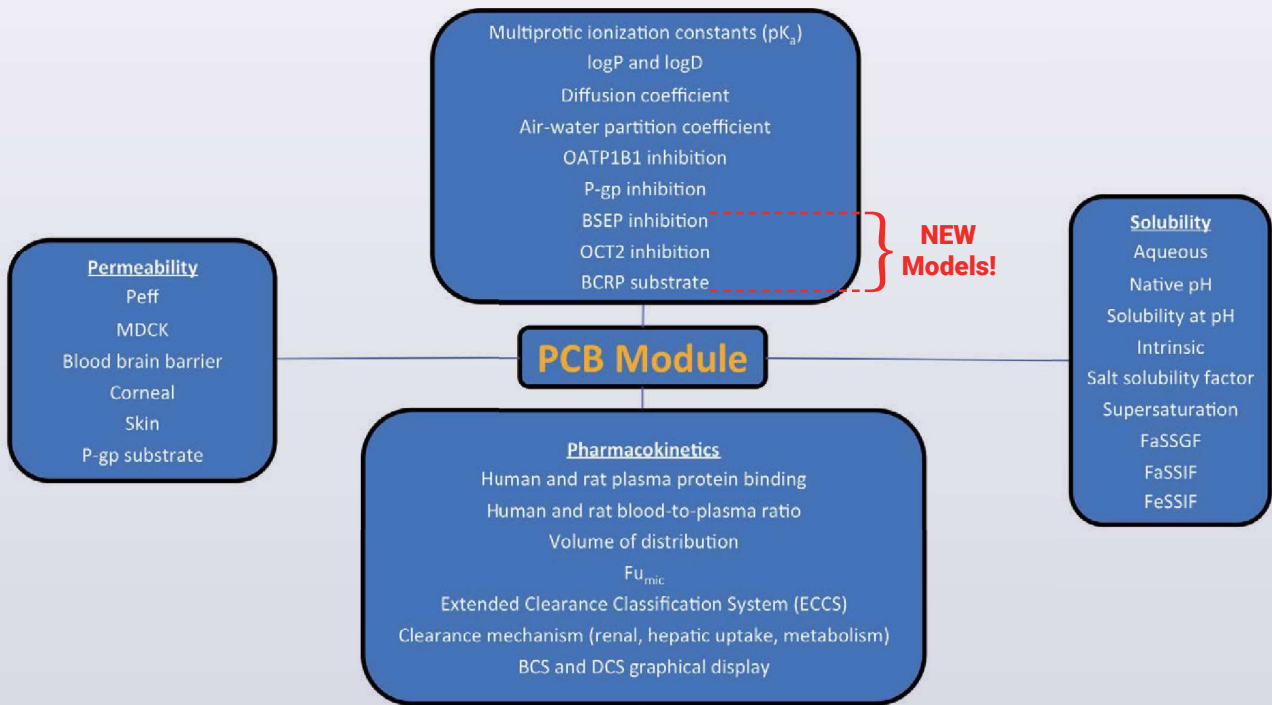


ADMET プロパティが計算されると、スタープロットが自動的に生成されます。ひとつのスタープロットにつき最大 **16** 種のプロパティを指定することができます。扇型の半径長はデータが表すプロパティ値の範囲に比例します。オプションで、各扇型の範囲の最小値と最大値の指定ができます。さらに、扇型の半径をプロパティ値の対数に比例されることも可能です。

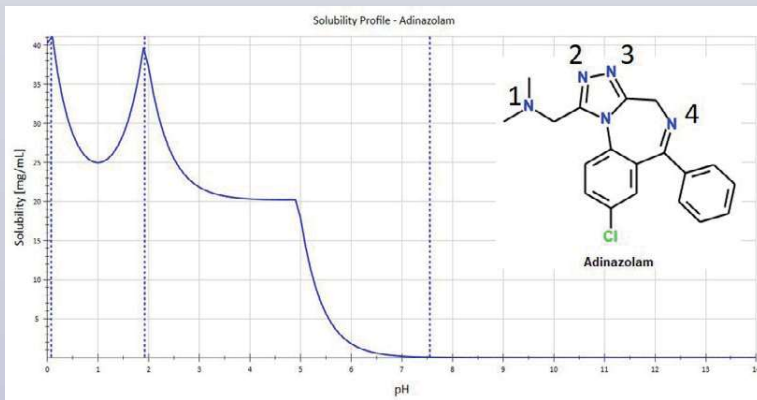


Physicochemical and Biopharmaceutical (PCB) モジュール

PCB モジュールのモデルは以下の図のとおりです。

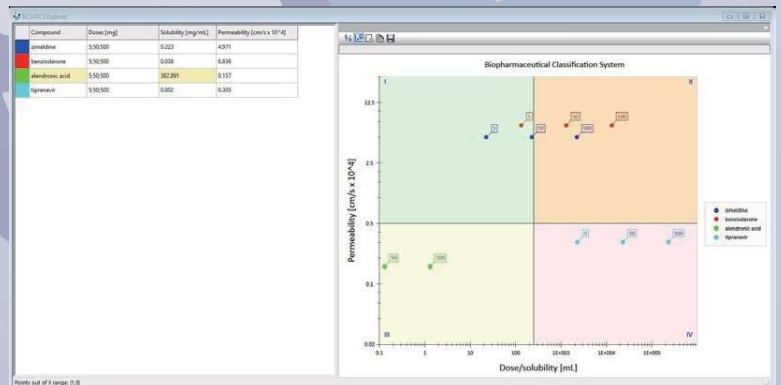


ADMET Predictor 9.5 のグラフには、下の図のような **solubility. vs. pH** と **BCS** があります。



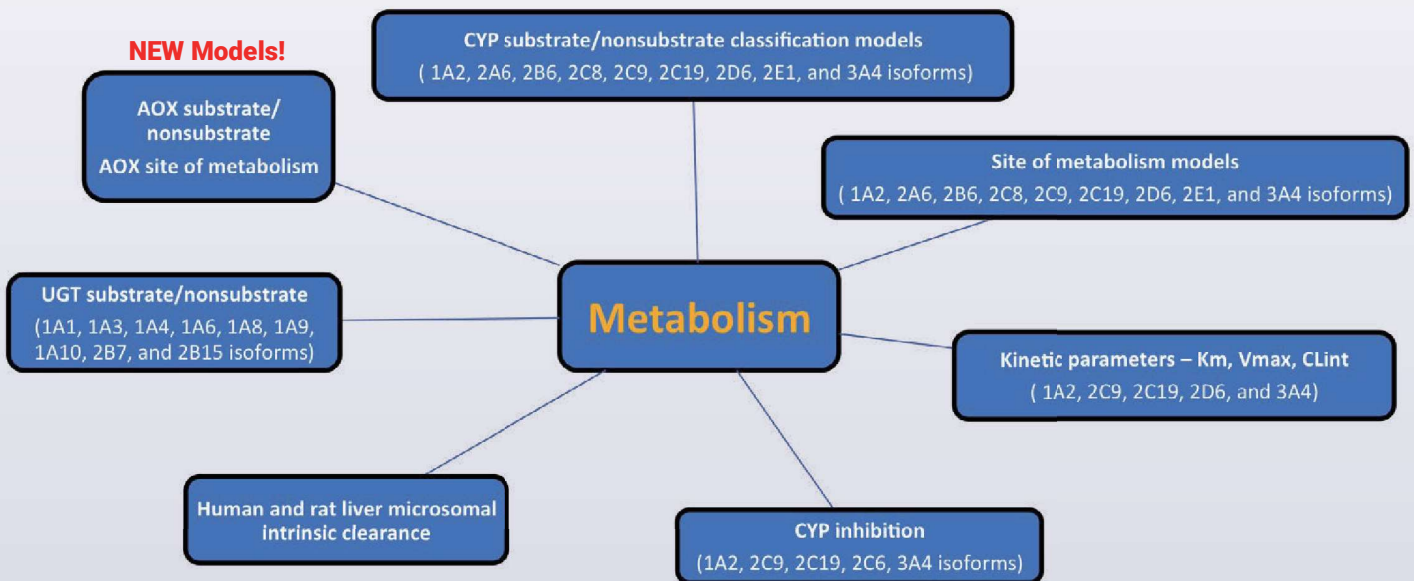
Adinazolam には 4 つの塩基性アミンがありそれらが 1-4 でラベルされた構造図とその予測溶解度対 pH のプロファイルが左図にあります。縦の点線はマクロステートでの **pKa** を表しています。最も高い **pKa** は **7.56** であり、**pH** の低下とともに溶解度が上がり始めます。窒素 **1** に加え、窒素 **3** または **4** のどちらかのプロトン化を含んでいるマクロステートは、ちょうど **pH5** 以下になると溶解度に影響を与え始めます。

Biopharmaceutical Classification System (BCS) 表示機能では、ユーザーが指定した投与量で最大 **16** の化合物まで **BCS** クラスをプロットすることができます。

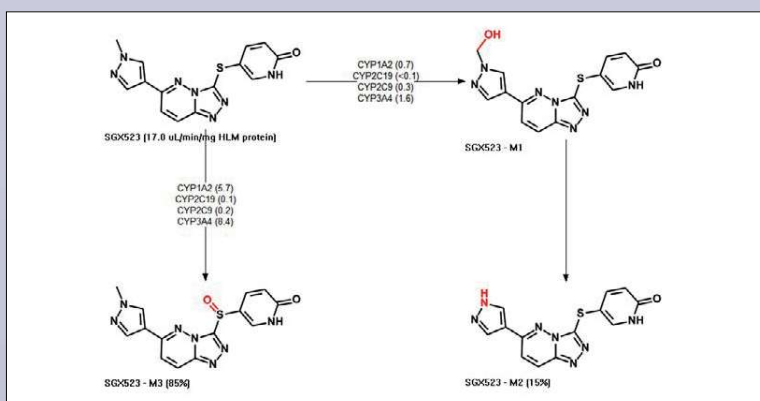
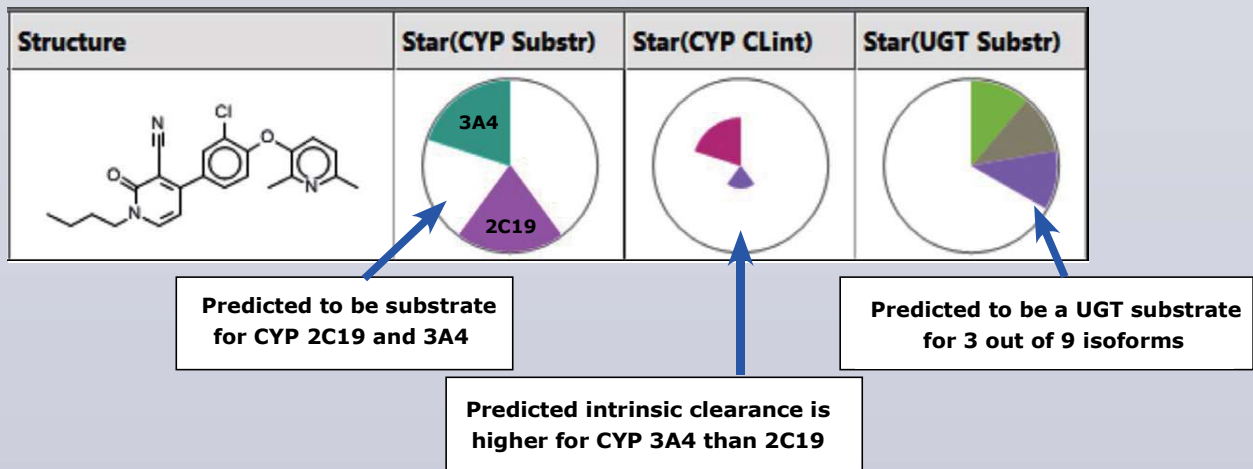


Metabolism モジュール

Metabolism モジュールにあるモデルが以下の図で示されています。

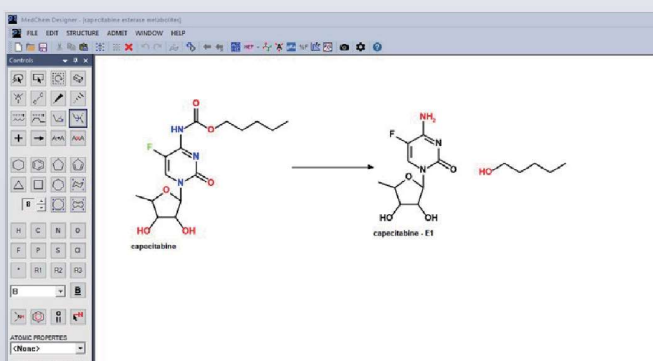
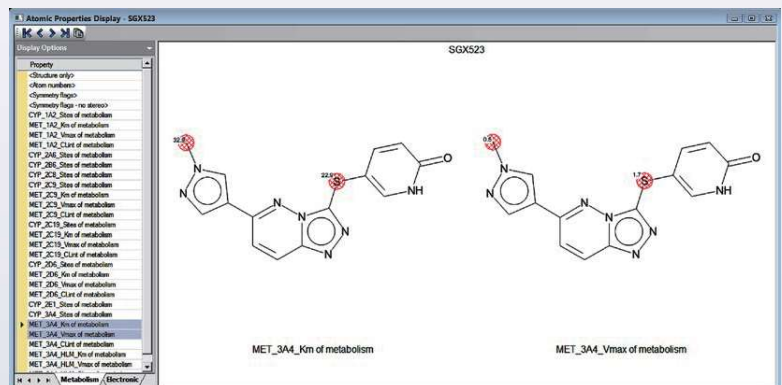


スタープロットでモデルによる結果を素早く評価することができます。



代謝物予測：親分子の **CLint** が、生成された各代謝物の割合とともに表示されます。メチル基のヒドロキシル化によって、デスメチル生成物とホルムアルデヒドに分解する不安定な中間体が生じます。

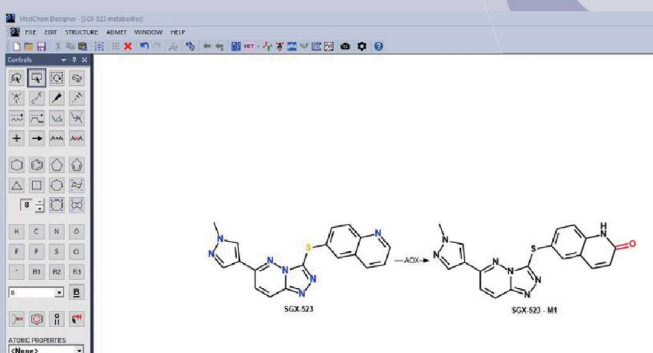
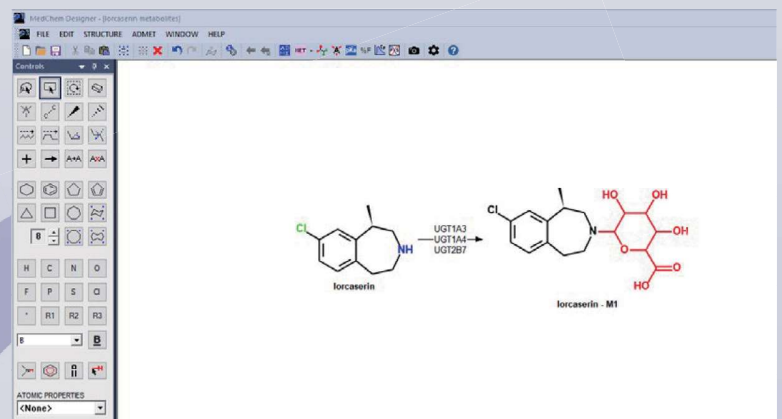
ADMET Predictor は、各代謝部位の K_m と V_{max} を予測します。



エステラーゼによるカルバメート基の開裂は、プロドラッグであるカペシタビン活性化の第一段階です。

- エキスパートルールに従う基質のみが代謝産物を生成
- ルールはユーザーがカスタマイズ可能
- エステラーゼ代謝産物を生成

MedChem Designer でグルクロン酸化代謝産物の表示が可能になりました。



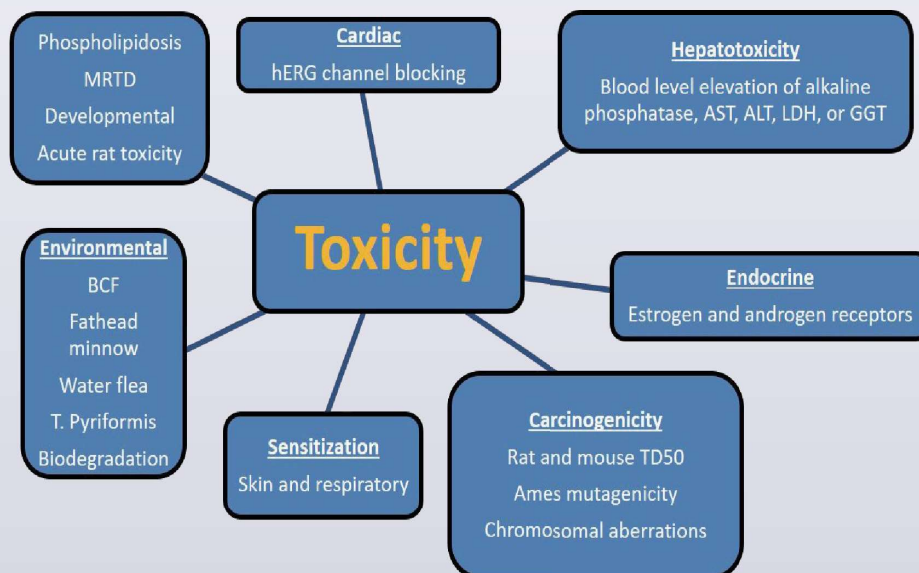
AOX_Substr と **AOX_Sites** モデルが分子に適用され、**AOX** 代謝産物が **MedChem Designer** に表示されます。

Toxicity モジュール

候補化合物スクリーニングの初期段階であっても毒性面について軽視すべきではありません。

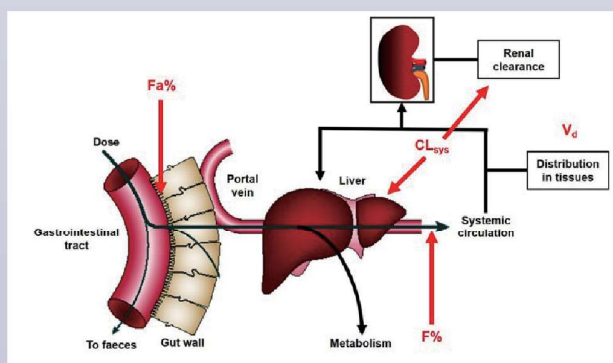
ADMET Predictor の毒性予測モデルはその数を急速に増やしておりそれを特徴としています。

Toxicity モジュールは、心毒性、肝毒性、内分泌、発がん性、過敏性、環境など、広範囲な毒性をカバーするモデルを備えています。



HTPK Simulation モジュール

HTPK Simulation モジュールは、**#1** ランクのメカニスティック吸収 / **PBPK** モデルと **#1** ランクの **ADME** プロパティ予測を組み合わせ、ラット、ヒトについて次の **in vivo** プロパティを予測します：



- 吸収率 (**Fa%**)
- **Cmax**
- **Tmax**
- **AUC**
- 経ロバイオアベイラビリティ (**F%**)
- メカニスティック分布容積 (**Vd**)
- ユーザー指定の血漿中濃度 (**Ceff**) 到達に必要な投与量 (**D**)

左の図は、生体内で発生するプロセスを示しています。消化管は **9** つのコンパートメントで構成される **ACAT™** (**Advanced Compartmental Absorption and Transit**)モデルによって表現され、薬物分子は、未放出、

未溶出、溶出状態で表されます。分子はトランスセルラー、パラセルラー経路の拡散を介して腸細胞に吸収されます。初回通過肝クリアランスは、ラット、ヒト固有クリアランスモデルを用いて計算されます。ユーザーが **in vitro** 肝臓ミクロソームや肝細胞固有クリアランスデータを与えることも可能です。**PBPK** モデルでは、**16** 個のコンパートメントを使用して分布容積 (**Vd**) を計算します。全身クリアランスは肝クリアランスと腎クリアランスから計算されますが、ユーザーが与えることも可能です。これらの値はユーザーが指定した血漿中濃度に達するために必要な投与量の予測に取り入れられます。

ADMET Modeler は、実験データから高品質な構造 - プロパティ予測モデルを開発する面倒で嫌になるプロセスを自動化します。 **ADMET Predictor** の構造ディスクリプターが入力として使用され、処理後、最終選択されたモデルが新たな予測プロパティとして **ADMET Predictor** に追加されます。これらはシームレスに行われます。以下は、優れた予測モデルを構築するための **3** つのキーです：

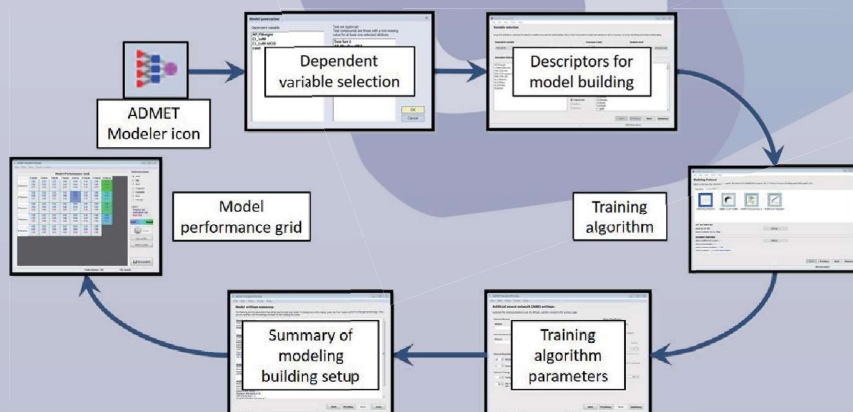
1. データがクリーンであり、一貫性があること
2. ディスクリプターが適切かつ正確なこと
3. 優れたトレーニングアルゴリズムを用いること

ADMET Predictor ではデータの品質を評価するための多くのツールが提供されています。例えば、スプレッドシート内で重複する構造を容易に見つけ出すことができます。重複があればそれらは同じ実験値を持っています。一致分子ペア分析では、分子類似性が高く実験エンドポイントが異なる分子のペアを見つけます。この場合、実験結果を検証することが重要になります。 **ADMET Predictor** は、分子の **2D** 構造から **300** 以上のディスクリプターを計算しますが、この中には、高水準 **ab initio** 計算を再現するためにパラメーター化された **EEM (Electronegativity Equalization Method) Hückel** モデルによって生成された原子、反応性ディスクリプターも含まれています。トレーニングアルゴリズムとして、ニューラルネット (**ANN**)、サポートベクターマシン (**SVM**)、カーネル部分最小平方 (**KPLS**)、重回帰 (**MLR**) が準備されています。

下の図は、 **ADMET Predictor** での直観的なモデル構築のワークフローを示しています。最初のステップでモデルを作りたいプロパティを選び、次にモデル構築で使用するディスクリプターを選択します。ユーザーは新規モデルのディスクリプターとして、例えば、 **S+lop** などのモデルを選ぶことも可能です。さらに、スプレッドシートに存在すれば、自身のディスクリプターを選ぶこともできます。次のステップは、モデル構築に用いるトレーニングアルゴリズムとパラメーターの選択になります。セットアップの要約が表示され、その内容でよければ指定されたアルゴリズムを用いてモデルが構築されます。

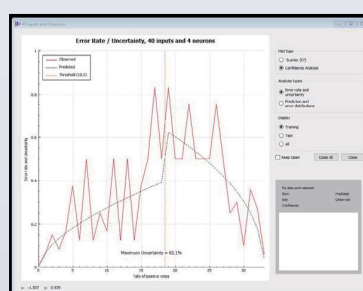
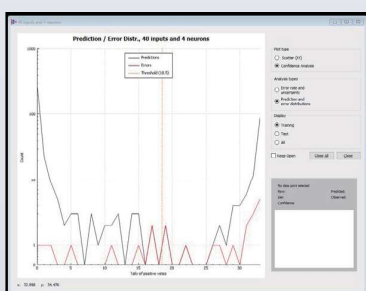
更なる機能として：

- ニューラルネットアンサンブルを用いた多クラス (**3** 以上) 分類モデルの構築
- 信頼値によって色分けされた二項分類モデルのパフォーマンスプロット
- **ADMET Predictor** グローバルモデルに社内データを組み込む **DELTA** モデルアプローチ



信頼性分析

信頼性分析は、ニューラルネットワークアンサンブルモデル内の個々のニューラルネット間の一致の度合いに基づいています。この分析の例が以下に示されています。左の図は各「せいの投票数」(x 軸)での、総予測数(黒線)と誤った予測の数(赤線)のプロットです。**100**以上の予測がゼロ正投票、すなわち、**33**のニューラルネットワークすべてが負を予測しました。赤い線はこれらの予測のひとつだけが正しくなかったことを示しています。このデータは右のプロット内で誤差率に変換されます。点線は、陰性と陽性の二つのベータ二項式を表しています。任意の化合物の信頼性の推定値は賛成票の数に基づいており、このタイプのプロットから導き出されます。



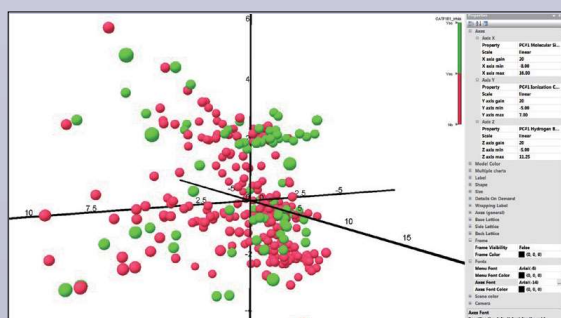
カスタマイズ可能な ADMET Risk Filter

ADMET Risk スコアリング機能を用いて、薬物分子の潜在的ライアビリティを素早く評価することが可能です。**World Drug Index (WDI)** からの **2,000** 化合物を超えるサブセットについて物性予測が計算されました。各プロパティの **Cutoff** の値はプロパティの全体的な分布を考慮し、**WDI** の化合物の **10%** が削除されるように閾値を設定して決定されました。

このスコアは許容範囲外のそれぞれの予測に対しては 1 ずつ増加します。

- **Absorption Risk** - 経口投与による低吸収リスク(サイズ、電荷、回転可能な結合の数、**h**- 結合受容体および供与体、親油性、溶解性および透過性を含む)
- **CYP Risk** - 高クリアランスまたは **CYP** 阻害のリスク
- **TOX Risk** - 一般的な毒性リスク (**hERG** 阻害、ラットマウスの急性毒性、肝毒性、変異原性)
- 上記すべてをひとつにまとめた **Global ADMET Risk**
- スコアリング機能はカスタマイズ可能でユーザーの好みのカットオフやプロパティを組み込むことができます

4D Data Mining

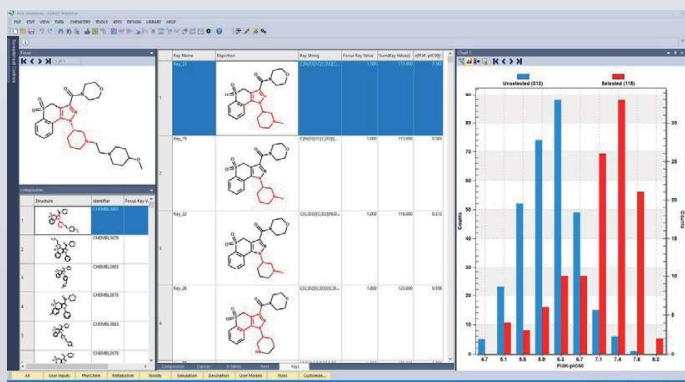


ADMET Predictor で計算される分子ディスクリプターは化学空間の視覚化に利用することができます。左のグラフは **OATP1B1** 阻害について試験された一連の化合物の分子サイズと形状、イオン化特性、水素結合の主成分を示しています。グリーン色のポイントが **OATP1B1** 阻害剤になります。

ADMET Predictor 9.5 では、**Extended-Connectivity Fingerprint (ECFP)**、**Internal fingerprints (MACCS)**、**Unbranched fragments**、**Ring systems** などの様々なタイプのフィンガープリントを計算することができます。それぞれのフィンガープリントは“**Keys**” タブ内で行によって表されます。その分子の、キーに対応する部分が赤で強調表示され、**SMART** フォーマットで **Key String** カラムに表示されます。フィンガープリントとプロパティ間のピアソン相関係数、例えば **IC50**、は、カラムとして追加することが可能です。下の画像は **428** ある **PI3K** 阻害剤のデータセットのフィンガープリントの一部を示しています。ハイライトされたフラグメント (**Key_23**) はデータセットの **428** 化合物のうちの **115** 個に存在しています。**PI3K-pIC50** チャートの赤いバーはこれら化合物の **pIC50** の値を表しています。このフラグメントを含んでいる化合物はより高い **PI3K IC50** 値を持ち、より活性があります。

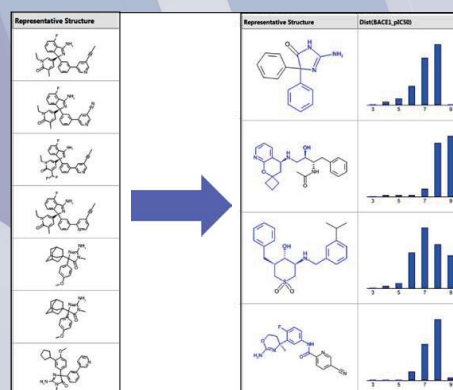
その他の機能：

- **Extended-Connectivity Fingerprint (ECFP)**などのフィンガープリントを計算する機能
- フィンガープリントに対する分子上の原子をハイライトするインターフェイスによるフィンガープリントの視覚化
- フィンガープリントと分子特性間のピアソン相関係数を含んだカラムの追加(例：**IC50** 値)



MedChem Studio のクラス生成テクノロジーは、共通骨格構造に基づいて分子を気もタイプに自動的に編成することにより、ケミストの論理的思考を再現します：

- フィンガープリントではなく最大共通部分構造(**MCS**)を利用することにより、結果が化学的に直観的なものになります。
- 骨格構造を事前定義されたリストから読み込むのではなくデータから生成させるため、新規構造モチーフを発見することが可能です。
- ローカルな **QSAR** 生成、分子設計、その他の分析業務などへ理想的な開始点を提供します。
- ユーザー定義の骨格構造をクラス生成に利用することも可能です。
- **Murcko** アセンブリ法に似た **Frameworks** クラス生成オプション。



例：**MedChem Studio** モジュールを使用して、**718** の **BACE1** 阻害剤を、共通骨格構造を持つ **27** ファミリーにグループ化しました。最大クラスには **243** 化合物が含まれています。

それぞれの行が固有のクラスに対応しています。そのクラスの代表分子が骨格構造部分を青でハイライトされ表示されます。また、そのクラスにある化合物の **pIC50** 値の分布図も示されます。

クラスのプロパティ値を、平均、最小、最大、他の様々な統計関数として表示することもできます。

MedChem Studio™ Module - Matched Molecular Pairs

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

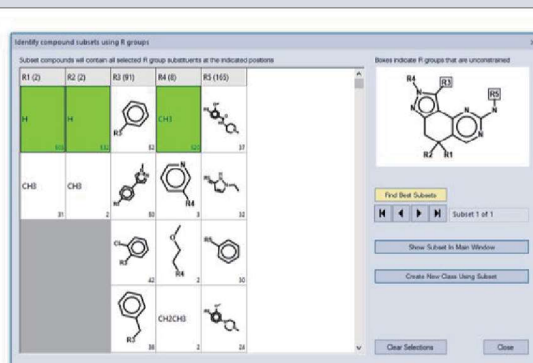
上記のペアはさほど劇的な活性の違いをもたらしません興味深い SAR を示しています。

Matched molecular pairs (MMP) はデータセットにあるアクティビティ・クリフの特定やプロパティ値にある構造的な傾向の同定、場合によっては活用のための分析に使用できます：

- **MMP** は、小さな特定の構造変換をもつ分子のペアです。
- **MMP 分析 (MMPA)** は一致した分子ペアの自動同定であり、ペア間のプロパティ値の変化の分析です。
- プロパティ値の平均変化は変換が別の分子に適用されたときのプロパティ値の期待サイズです。
- **MedChem Studio** は各変換則をコード化する **SMIRKS** 文字列を自動的に生成します。
- これらのルールを **“Combinatorial Transform...”** 機能に加えて、プロパティ値の希望する変化を潜在的に含む新規分子を生成することができます。

“Find Subsets using R-groups” は 2 つの R-基のみをもつ分子からなる最大のサブセットを見つけます。以下の例で、**R1=R2=H** と **R4=CH3** は 534 化合物の 489 を表しています。データセットがこれによって簡素化され、データを「掘り下げて」**R3** と **R5** に関連する **SAR** を理解することができます。

Structure	Identifier	R1 (2)	R2 (2)	R3 (R1)	R4 (R)	R5 (R5)	R-Count	Family	IGPR_pIC50	AP_Weight
	CH-EMBL3576	H	H	CH3	CH3	CH3	3		6.662	429.000
	CH-EMBL3616	H	H	CH3	CH3	CH3	3		6.523	367.887
	CH-EMBL3616	H	H	CH3	CH3	CH3	3		7.721	415.500
	CH-EMBL3616	H	H	CH3	CH3	CH3	3		6.361	446.911
	CH-EMBL3616	H	H	CH3	CH3	CH3	3		1.194	517.676
	CH-EMBL3616	H	H	CH3	CH3	CH3	3		1.131	416.500
	CH-EMBL3616	H	H	CH3	CH3	CH3	3		7.602	476.933



Aug,2019

■お問い合わせ先■

ノーザンサイエンスコンサルティング株式会社

〒060-0005 札幌市中央区北 5 条西 6 丁目 2-2 札幌センタービル

<http://www.northernsc.co.jp/>