

Reverse Fingerprint Modeling の紹介

Reverse Fingerprint Modeling は、MOE ユーザーコミュニティのための SVL コード交換サイト SVL Exchange で公開されている統合計算化学システム MOE のアドオンプログラムです。Reverse Fingerprint Modeling では、化合物のさまざまな分子フィンガープリントと活性値の情報を用いて、活性化化合物に共通の部分構造や活性 / 不活性に寄与する部分構造を特定したり、ファーマコフォアモデルを構築 / 評価したりできます。本稿では、Reverse Fingerprint Modeling の概要と 4 つの機能を紹介します。

■ Reverse Fingerprint Modeling の概要

分子フィンガープリント (FP) は、分子構造を表現する方法の一つで、主に官能基やファーマコフォアなどを基にした特徴の有無を 1 と 0 から構成されるビット列で表現したものです。FP には類似度の算出方法が定義されており、類似化合物の探索に利用できます。Reverse Fingerprint Modeling (RFM) では、さまざまな FP の各ビットとそれに割り当てられた特徴の元となった原子群と活性値の情報を用いて MIBALS 法¹⁾により以下の 3 種類のスコアを求めます。

■ ビットスコア

FP の各ビットについて算出されるスコアです。これが正の値のとき、そのビットに割り当てられた特徴が不活性化化合物よりも活性化化合物において高い頻度で存在することを示します。逆に、これが負の値のとき、そのビットに割り当てられた特徴が活性化化合物よりも不活性化化合物において高い頻度で存在することを示します。0 に近いときは、そのビットに割り当てられた特徴が活性化化合物と不活性化化合物に同程度の頻度で存在することを示します。

■ 原子スコア

各原子について、ビットスコアをもとに算出されるスコアです。ビットスコアと同様に、これが正の値をもつ原子は、活性化化合物に多く含まれるビットに関与しており、負の値をもつ原子は、不活性化化合物に多く含まれるビットに関与していることを示します。

■ 分子スコア

分子全体のスコアです。これはすべての一意なビットのビットスコアの総和であり、すべての原子の原子スコアの総和でもあります。正の値のとき、活性化化合物である確率が高く、負の値のとき、不活性化化合物である確率が高いので、バーチャルスクリーニングにおける化合物の順位付けに使用できます。

■ Motif Detection

Motif Detection は、活性化化合物群からビットスコア、原子スコア、分子スコアを計算して可視化する RFM の基本機能です。原子スコアは、その値のラベルと絶対値の大きさに比例した半径を持つ球体で表現されます。正の値のときはソリッドの球体で、負の値のときは網掛けの球体で描かれます。使用例として、殺虫剤と殺ダニ剤から構成される 80 個の化合物群を用いて、殺ダニ剤を

活性とした場合の 2 個の殺ダニ剤の原子スコアを可視化しました (図 1)。Fluvalinate のシアノ基に大きな原子スコアがあるのが分かります。これはタイプ II ピレスロイド系の典型的な部分構造です。この化合物は、ミツバチを襲うダニであるバロアダニの駆除のため養蜂家に使用されています。一方、Halfenprox にあるエーテル結合は、殺ダニ剤としては好ましくない部分構造であることが分かります。

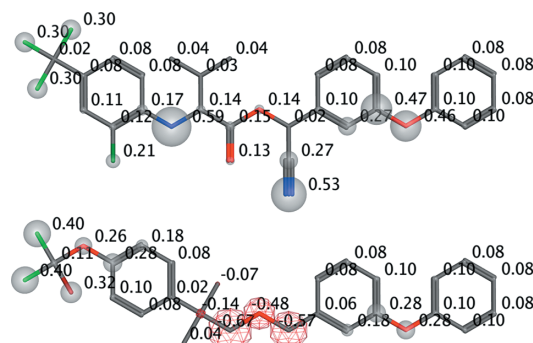


図 1. Motif Detection で求めた Fluvalinate (上) と Halfenprox (下) の原子スコア

■ Virtual Pockets

Virtual Pockets は、リガンドの配座データを使用して、結合ポケットの形状や特性を予測 / 可視化し、ファーマコフォア検索やドッキングシミュレーションに使用できるファーマコフォアモデルを作成します。

まず、リガンド構造と活性値から原子スコアを求めます。そして、原子スコアが高い原子を含む特性球を検出し、それらで構成される基本ファーマコフォアモデルを作成します。次に、それらのファーマコフォアモデルについて、多くのリガンドに当てはまるものだけに絞り込みます。その後、残ったファーマコフォアモデルをクラスタリングし、統合することでコンセンサス・ファーマコフォアモデルを作成します。さらに、コンセンサス・ファーマコフォアモデルにリガンド配座を重ね合わせた結果から、リガンドの形状や結合ポケットの表面やその特性を予測して、可視化します (図 2)。

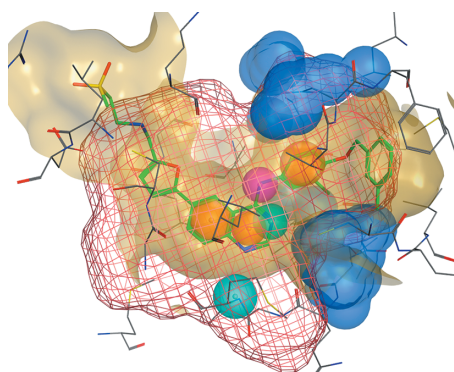


図2. EGFRの阻害剤を用いた Virtual Pockets の計算結果（黄色：実際の受容体表面（PDB ID: 1XKK）、赤色の網掛け：予測したリガンド形状、青色：予測した受容体原子の領域、橙色：芳香環の特性球、水色：アクセプターの特性球、紫色：ドナーの特性球）

Fragment Placement

Fragment Placement では、分子フラグメントの結合ポケットへの配置やファーマコフォアモデルの評価を行えます。分子フラグメントから Virtual Pockets と同様にファーマコフォアモデル（フラグメント・ファーマコフォアモデル）を作成し、ユーザーが提供した結合ポケットにあるファーマコフォアモデルと重ね合わせ、各特性球の一致度を評価/可視化します（図3）。また、重ね合わせにより、結合ポケットに配置されたフラグメント・ファーマコフォアモデルを使用して、分子フラグメントを結合ポケットに配置するドッキングシミュレーションも行えます。

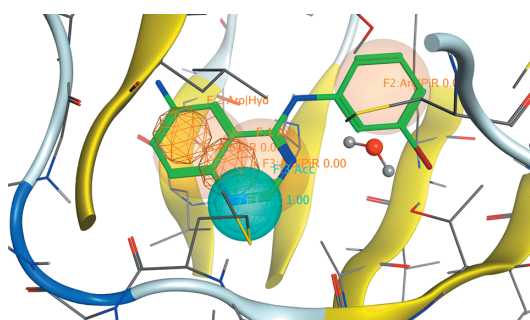


図3. EGFRの阻害剤の分子フラグメントを用いた Fragment Placement の計算結果。ユーザー指定のファーマコフォアモデル（ソリッド）とフラグメント・ファーマコフォアモデル（網掛け）が部分的に一致している。

Activity Scan

Activity Scan では、活性値の範囲に4つの活性/不活性の閾値を設定し、それぞれで Motif Detection が実行されます。各閾値に設定した色と原子スコアから、表示される球体の色と大きさが決定されます。活性値の範囲を分けた計算結果を組み合わせているため、活性値の大きさを反映した可視化ができます。使用例として、ジヒ

ドロ葉酸レダクターゼ（DHFR）の397個の阻害剤について、活性値の pIC_{50} を0、4、6、7の閾値を設定して計算しました（図4）。これらの閾値の色には、灰色、赤色、黄色、緑色をそれぞれ設定しています。カテコール基や環構造間のリンカーのメチル基は、高活性な阻害剤に頻出する部分構造であり、5員環の硫黄原子は、低活性な阻害剤に頻出する部分構造であることが分かります。これらの内容はビットスコアを可視化する機能でも確認できます（図5）。

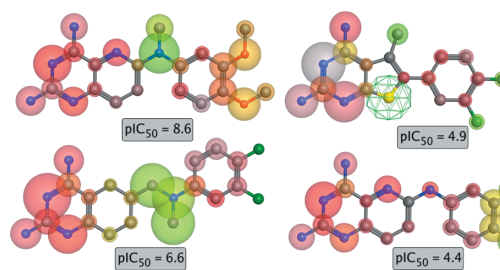


図4. DHFRの阻害剤を用いた Activity Scan の計算結果。灰色、赤色、黄色、緑色はそれぞれ pIC_{50} が0、4、6、7より大きい阻害剤に多い部分構造。球体が網掛けの場合は、逆の解釈になる。

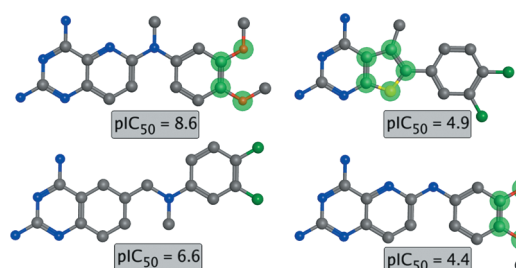


図5. DHFRの阻害剤を用いた Activity Scan によるビットスコアの可視化。ヘテロ原子を含む5員環（ビットスコア：-7.2）や芳香族原子に酸素原子が結合した構造（ビットスコア：7.9）を強調している。

まとめ

RFMは、FPのビットと部分構造の対応関係を応用した4つの機能を搭載しており、リガンドベースだけでなく、受容体構造ベースやフラグメントベースの創薬設計に使用できます。RFMはSVL Exchangeで公開されているMOEのアドオンプログラムで、MOEのユーザーなら無償で利用できます。

- Williams, C.; Schreyer, S. K. Reverse fingerprinting and mutual information-based activity labeling and scoring (MIBALS). *Comb. Chem. High Throughput Screening* **2009**, *12* (4), 424–439.