

(東大院工) ○荒川 正幹  
(豊橋技科大) 溝渕 創一郎  
(東大院工) 船津 公人\*

## 1. はじめに

近年の急速な情報技術の進歩にともない、分子設計の分野においてもコンピュータを用いた解析が必須のものとなりつつある。特に創薬の分野では、開発の初期段階においてケムインフォマティクス、バイオインフォマティクス技術を利用した解析を行うことによって、コストの削減や開発期間の短縮を図ることが重要である。コンピュータを用いてこれらの解析を実行するためには対応したソフトウェアが必要であり、多機能で使い易いソフトウェアの開発が望まれている。

そこで我々の研究室では、分子設計に関する解析を統合的に行うためのコンピュータシステムとして、分子設計トータルシステム ToMoCo の開発を行っている[1]。現在、本システムでは CoMFA (Comparative Molecular Field Analysis) による構造活性相関解析、HNN (Hopfield Neural Network) を用いた分子構造の重ね合わせ、GARGS 法による CoMFA 領域選択、LigConstructor による新規薬物候補構造の自動創出などを行うことが可能である。本システムは Microsoft Windows 2000/XP 以上で動作するシステムであり、ユーザは GUI を通してすべての解析を実行し結果を視覚的に確認することができる。本システムの開発は Microsoft の統合開発環境である Visual Studio .NET 2003 を用いて C++言語により行っている。また画面描画のためのグラフィックライブラリとして OpenGL を利用している。図 1 に ToMoCo の実行画面を示す。

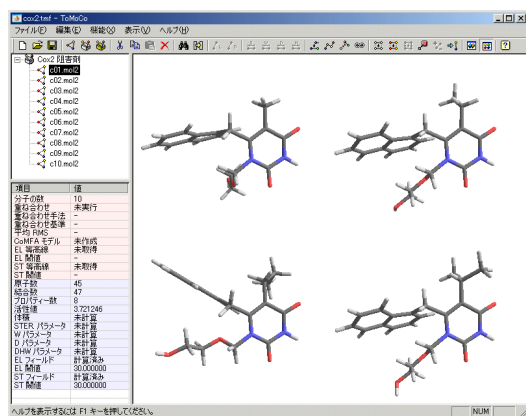


図 1 ToMoCo 実行画面

## 2. 手法

以下では本システムに実装されている各手法について説明する。

### CoMFA 法[2]

CoMFA 法は 3 次元構造活性相関解析の代表的な手法であり、さまざまな種類の化合物群に対して適用され研究成果が報告されている。CoMFA では、まず対象とする化合物全体を含む直方体領域を定義し、その内部に一定間隔の格子点を設定する。そして各格子点にプローブ原子と呼ばれる仮想的な原子を配置し、対象分子との間の立体的、静電的相互作用を求めて構造記述子とする。そしてこの記述子を説明変数とし、活性値との間で PLS (Partial Least Squares) 法による回帰分析を行い構造活性相関モデルを構築する。このモデルを用いることで、活性が未知の化合物の活性を予測することができる。また回帰係数の等高線図を描くことで、活性に大きな影響を与えらる領域を特定することが可能である。この結果に対して考察を加えることによって、より理想的な活性を示す化合物の設計を行うための指針となる情報を得ることができる。

本システムを用いることによって CoMFA 記述子の計算、PLS モデリング、等高線図の表示などの一連の解析を容易に行うことが可能である。

### 分子構造の重ね合わせ

CoMFA などの 3 次元構造活性相関解析手法を用いて優れたモデルを構築するためには、対象とする化合物について適切な重ね合わせを行うことが必要である。分子構造の重ね合わせを行うための手法については数多く提案されているが、重ね合わせの目的や対象とする分子の特徴によって最適な重ね合わせ手法は異なるため、全ての分子に適用可能な手法は存在しない。そこで本研究室では、3 次元構造活性相関解析に適した重ね合わせを行う手法として、HNN による分子構造重ね合わせ方法を提案している[3]。この手法では、まず重ね合わせを行う 2 つの分子上に化学プロパティを定義しそれらの対応付けを行う。そして対応付けられたプロパティ間距離の 2 乗和を最小化し、3 次元構造の重ね合わせを実現する。

なお化学プロパティとしては疎水性、水素結合供与性、水素結合受容性などを用いることができる。またプロパティの対応付けを行うために HNN を用いている。

本システムでは化学プロパティの自動設定および編集、プロパティを基準とした重ね合わせなどを実行することができる。また、対応させる原子を指定して重ね合わせを行うことも可能である。

### LigConstructor

進化的計算手法を用いて新規薬物候補構造の自動創出を行うシステムとして LigBuilder[4]が提案されている。LigBuilder は標的タンパク質の活性部位周辺に候補構造の基となる構造を配置し、それを成長させることでレセプターに適応した新規薬物候補構造を自動的に生成する手法である。X 線結晶構造解析などによりタンパク質の 3 次元構造が明らかになっている場合には、ポケットに適した有力な構造を得ることが可能である。しかし一方で、標的タンパク質のポケット情報が得られていなければ候補構造を作成することができないという問題がある。

そこで本研究室では、このような場合でも新規薬物候補構造の創出を行うことが可能なシステムとして LigConstructor[5]を提案した。LigConstructor では、CoMFA モデルによる予測活性値を評価基準として用いて構造を進化させることで、活性が高いと期待される薬物候補構造を生成する。そのためタンパク質の 3 次元構造が明らかになっていない場合においても適用可能である。LigBuilder と同様に、はじめに候補構造の基となる構造を入力し、進化的計算手法により構造を進化させる。入力された構造に対して、構造結合と突然変異の操作を行い新規な候補構造を複数生成する。そしてそれらの構造の活性値を CoMFA モデルによって予測し、それをもとに淘汰、選択を行う。これらの操作を繰り返し行うことによって、最終的に高い活性を示すと思われる候補構造が生成される。

### GARGS 法[6]

CoMFA 法においては構造記述子を計算するため、対象とする分子群をすべて含むような直方体を定義する必要がある。しかし多くの場合、活性値を決定するのに重要な役割を果たしているのはその一部であり、その他の領域については情報量が少ないかあるいはまったく活性と関係のないものである。PLS モデリングにこのような記述子を含めることは、モデルの精度や予測性を悪化させる要因となる。そこで本研究室では、活性を説明するのに重要な役割を果たしている領域を自動的に選択するための手法として、GARGS 法を提案した。

GARGS 法では、CoMFA フィールドをいくつかの小領域に分割し、GA (遺伝的アルゴリズム) を用いて、活性を説明するために有効な領域の探索を行う。各小領域を遺伝子の各ビットに割り当て、交差、突然変異、評価、淘汰、選択を繰り返すことによって最適化を行う。評価値としては、遺伝子において選択されている領域の記述子のみを用いた PLS モデリングの結果として得られる  $Q^2$  値を使用している。GARGS 法を用いて領域を選択することによって、シンプルで予測性の高い構造活性相関モデルを構築することができ、より確度の高い分子設計が可能となる。

## 3. まとめ

分子設計に関連する解析を共通の GUI から行うことのできる統合システムを目指し、分子設計トータルシステム ToMoCo の開発を行っている。現在のところ本システムには CoMFA による構造活性相関解析、分子構造の重ね合わせ、GARGS 法による領域選択、LigConstructor による新規薬物候補構造の自動創出機能などが実装されている。つまり活性の分かっている化合物群の 3 次元構造を入力することによって、重ね合わせ、CoMFA、領域選択、新規薬物候補構造の自動創出という一連の解析を本システム上で行うことが可能である。

今後も本研究室で開発された手法を中心に機能の拡張を行い、さらなる利用性の向上を目指して開発を行っていく予定である。現在は、分子表面上の静電ポテンシャル、疎水ポテンシャルを用いた 3 次元構造活性相関解析手法[7]について ToMoCo への組み込み作業を行っているところである。

## 参考文献

- [1] 溝渕創一郎, 鈴川哲大, 荒川正幹, 船津公人, 分子設計トータルシステムの開発, 第 26 回情報化学討論会 講演要旨集, 115-118 (2003)
- [2] R. D. Cramer, D. E. Patterson, J. D. Bunce, Journal of the American Chemical Society, 110, 5959-5967 (1988)
- [3] M. Arakawa, K. Hasegawa, K. Funatsu, Journal of Computer Aided Chemistry, 2, 29-36 (2001)
- [4] R. Wang et al., Journal of Molecular Modeling, 6, 498-516 (2000)
- [5] 竹内英憲, 荒川正幹, 船津公人, 薬物候補構造の自動創出プログラムの開発, 第 25 回情報化学討論会 講演要旨集, 75-78 (2002)
- [6] K. Hasegawa, T. Kimura, K. Funatsu, Journal of Chemical Information and Computer Sciences, 39, 112-120 (1999)
- [7] K. Hasegawa, S. Matsuoka, M. Arakawa, K. Funatsu, Computational Biology and Chemistry, 27, 381-386 (2003)