

○荒川 正幹、船津 公人

東京大学大学院 工学系研究科

〒113-8656 東京都文京区本郷 7 丁目 3 番 1 号

はじめに

効率的な医薬品開発・分子設計を行うためには、コンピュータを用いた高度な解析が必要不可欠であり、ケモインフォマティクス、ケモメトリクスの重要性はますます高まってきているといえる。そこで我々は、有機化合物の合成経路設計、各種スペクトルからの構造推定、反応予測、構造活性相関解析などを対象として、新規なケモインフォマティクス技術の開発とその応用に取り組んでいる。また、これらに関する解析を行うためのソフトウェアの開発にも力を入れている。ToMoCo (The total system for molecular designs by the computer chemistry laboratory)はこの一環として開発を進めているシステムであり、主に分子設計・構造活性相関解析に関する機能を集約したものとなっている。

本発表ではToMoCoに実装されている手法のうち、分子構造重ね合わせ機能、新規薬物候補構造自動創出機能、CoMFA領域選択機能について説明し、これらの手法を用いてCOX-2阻害剤の解析を行った結果を示す。COX-2阻害剤 54 サンプルに対して、構造重ね合わせを行うことにより活性配座を推定しCoMFA法による3D-QSAR解析を行った結果、 $R^2=0.922$ 、 $Q^2=0.653$ のCoMFAモデルが得られた。そして、そのモデルを評価基準として用いて新規薬物候補構造の自動創出を行い、予測活性値 7.11 から 7.87 までの 200 個の構造を得た。またCoMFA領域選択を行った結果、変数の数が 6 分の 1 以下に削減され、 $R^2=0.938$ 、 $Q^2=0.751$ というシンプルで予測的なQSARモデルが構

築された。

以下、ToMoCo に実装されている主な機能について示す。

CoMFA 法

CoMFA 法は 3 次元構造活性相関解析の代表的な手法である。一定間隔に配置したプローブ原子と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て活性値との間で PLS (partial least squares) 法による回帰を行い、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予測や、活性に大きな影響を与えると思われる領域の特定が可能となる。ToMoCo を用いることで、CoMFA 法による一連の解析を効率的に行うことが可能である。

分子構造の重ね合わせ

CoMFA法などの 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、対象とする化合物群について適切な重ね合わせを行うことが必要となる。しかしながら、対象構造の基本骨格が大きく異なる場合には適切な重ね合わせを見つけることは非常に困難である。そこで我々はこの問題を解決するため、Hopfield neural network (HNN)を利用した新しい分子構造重ね合わせ手法を提案した[1]。重ね合わせを行う 2 つの分子上に、疎水性、水素結合受容性などの化学プロパティを定義し、HNNを用いてそれらの対応付けを行うことによって適切な重ね合わせの探索を行う。

LigConstructor

進化的計算手法を用いて新規な薬物候補構造の自動創出を行うことを目的としたシステムとしてLigBuilder[2]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、その構造を変化させることでレセプターに適応した薬物候補構造を自動的に生成する手法である。そのため、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、活性部位に適した有力な構造を得ることが可能であるが、標的タンパク質の構造に関する情報が未知の場合には候補構造を提案することができないという問題がある。

そこで我々は、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[3]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することが可能である。

GARGS 法

CoMFA法においては構造記述子を計算するために、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と考えられる。PLSモデル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。そこで我々は、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[4]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。

まとめ

分子設計や構造活性相関解析に関連する操作を共通の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法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おり、一連の解析を本システム上で行うことが可能となっている。今後も新たに提案する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。

参考文献

- [1] M. Arakawa, K. Hasegawa, and K. Funatsu, *Journal of Computer Aided Chemistry*, 2, 29-36 (2001)
- [2] R. Wang, Y. Gao, and L. Lai, *Journal of Molecular Modeling*, 6, 498-516 (2000)
- [3] 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)
- [4] K. Hasegawa, T. Kimura, and K. Funatsu, *Journal of Chemical Information and Computer Sciences*, 39, 112-120 (1999)