

薬物動態特性の *In Silico* 予測に関する研究

橋田 充

In Silico Prediction of Pharmacokinetic Properties

Mitsuru HASHIDA

Department of Drug Delivery Research, Graduate School of Pharmaceutical Sciences, Kyoto University,
46-29 Yoshidashimoadachi-cho, Sakyo-ku, Kyoto 606-8501, Japan

(Received July 22, 2005)

In silico methods for predicting pharmacokinetic properties range from data-based approaches such as quantitative structure-activity relationships (QSARs), similarity searches, and 3-dimensional QSAR, to structure-based methods such as ligand-protein docking and pharmacophore modelling. Data-based modelling approaches are effective for many drug absorption, distribution, metabolism, and excretion (ADME) processes such as passive membrane permeation, where their molecular mechanism is barely delineated. Therefore QSAR approaches have been applied to simulate the relationships between ADME parameters and molecular structure and properties. In the present investigation, we describe the application of the genetic algorithm-combined partial least-squares (GA-PLS) method to QSAR modelling of various ADME properties. By selecting an appropriate set of molecular descriptors automatically using the genetic algorithm, many ADME properties could be well explained by simple molecular descriptors derived from the 2-dimensional chemical structure.

Key words—*in silico*; ADME; QSAR; genetic algorithm; partial least squares; multivariate analysis

1. はじめに

1980年代の後半より、コンビナトリアルケミストリーやハイスループットスクリーニングが注目を集めるようになり、創薬シード化合物の創出は非常なスピードで行われるようになってきた。しかしながら、ターゲットとの結合性のみを重視するとしばしば難溶性化合物となり、創薬研究プロジェクト全体の進行が遅延したり、研究開発を途中で断念せざるを得なくなる。したがって最近では、薬物動態、物性、安全性に関する評価を、探索研究の早期から実施し、莫大な時間やコストがかかる医薬品開発段階での成功確率を上昇させる試みがなされるようになってきた。¹⁻⁴⁾ 例えば、濁度法による溶解度試験、^{5,6)} Caco-2細胞⁷⁻⁹⁾やMDCK細胞、¹⁰⁾あるいは人工膜^{11,12)}を使った膜透過性試験、ヒト肝ミクロソーム^{13,14)}や代謝酵素発現系¹⁵⁾を使った代謝試験な

ど、スループットの高い *in vitro* 評価系が利用されている。しかしながら、依然として化合物の合成、評価は時間、コスト、労力を要するため、*in silico* による仮想スクリーニングにより有望な化合物を絞って実際の評価を行うことが望まれる。また、創薬初期におけるリード化合物の選定は、その後の展開を大きく左右するため、リード化合物から派生する化合物群の特徴を想定しながら確率的に望ましい方向性を探ることも重要である。そのためには、化合物の分子構造に基づいて薬物の体内動態特性や毒性などを予測する方法の開発が必要である。¹⁶⁻¹⁸⁾ 本稿では、*in silico* 薬物動態予測法について簡単に概説した上で、われわれがこれまで取り組んできた研究の成果について紹介したいと思う。

2. *In Silico* 薬物動態研究の方法論

薬物動態特性を予測する方法は大きく2つに分類され、分子構造に基づく機構論的モデリング法と実験データに基づく経験論的モデリング法がある。¹⁸⁾ 前者の方法は、分子動力学/分子軌道計算を駆使して、タンパク質結晶構造を元に低分子化合物とのドッキングを行う方法である。薬物代謝酵素チトクロ

京都大学大学院薬学研究科薬品動態制御学分野(〒606-8501 京都市左京区吉田下阿達町 46-29)
e-mail: hashidam@pharm.kyoto-u.ac.jp
本総説は、平成17年度日本薬学会賞の受賞を記念して記述したものである。

ム P450 (CYP) に関しては、バクテリア由来の CYP の X 線結晶構造からホモロジーモデリングにより構造推定を行う試みがなされ、CYP 1A2, 2C9, 2C19, 2D6, 3A4 など主要な薬物代謝酵素の結合シミュレーションが行われている。¹⁹⁾ 最近、ヒトの CYP 2C9 と 3A4 の X 線構造決定がなされ、²⁰⁻²³⁾ 分子シミュレーションによる解析研究の飛躍的な進歩が期待されている。また、間接的な Structure-based アプローチとして、代謝酵素²⁴⁻²⁷⁾ やトランスポータ²⁸⁻³⁰⁾ のファーマコフォアモデルも提案されてきた。これらの機構論的モデルは、比較的ロバストネスが高く、論理的であるため、合理的な分子設計指針を提示できる。しかしながら、代謝酵素による薬物代謝やトランスポータによる輸送など唯一の分子機構によって説明できる現象に対してしかモデル化を行うことができない。薬物動態では、受動拡散による膜透過のように高次の複雑系での現象を対象とすることが多く、機構論的なアプローチが適用できない場合も多い。このような場合には、実験データを基本とした推計学や統計学的モデルを利用することができる。¹⁸⁾ 構造活性相関 (QSAR) モデリングは、化合物の化学構造を何らかの形で分子記述子として数量化し、多変量解析を行う方法であり、薬物動態特性の予測モデルとしてしばしば利用されてきた。また、その変法として、化合物データをデータベースとして管理し、化合物構造の類似性/非類似性解析に基づいて目的特性値を推定する方法もある。^{31,32)} これらの方法は、実験データを基本とするので本質的には内挿的な予測しかできないが、多様な化合物に対するデータの蓄積によって問題を解決できるという期待があり、分子機構が不明なあらゆる事象に対しても適用可能であるという利点は極めて大きい。

3. 構造活性相関に基づく薬物動態特性予測

定量的構造活性相関 (Quantitative Structure-Activity Relationship, QSAR) は、生物活性と分子構造との間を定量的に解析する分野であり、薬物動態の分野でも古くから研究がなされてきた。基本的には 2 つのステップからなり、まず分子構造を構造記述子として数量的に表現した上で、基準変数 (生物活性) との関係が多変量解析法により解析する。³³⁾ 構造記述子の表現法として、化合物部分構造からなるフラグメント辞書化したフラグメントサブ

ストラクチャー法、原子間の結合パターンをグラフ理論に基づいて数値化したトポロジカルインデックス法、分子の 2 次元あるいは 3 次元構造式から分子の電氣的・立体的性質や脂溶性 (疎水性) などを計算するグローバルプロパティ法などがある。一方、多変量解析法として用いられる方法には、重回帰分析や部分最小二乗法^{34,35)} などの線形モデルがしばしば用いられるが、最近では人工ニューラルネットワーク³⁶⁾ などの非線形モデリング法も利用されるようになってきた。また、非常に多様な化合物群に対して適用可能なモデルを開発するためには効果的なモデリング技法が必要であり、われわれが行ってきたように、^{37,38)} 説明変数である分子記述子の選択において遺伝的アルゴリズムやシミュレイトドアニーリングのような組み合わせ最適化アルゴリズムも使用されるようになってきた。

4. 薬物動態特性の *In Silico* 予測法の開発

4-1. 遺伝的アルゴリズムを用いた構造活性相関モデリング法の開発 遺伝的アルゴリズムは、自然淘汰のメカニズムを利用した最適化方法である。Figure 1 は QSAR モデリングに対して遺伝的アルゴリズムを応用する方法を示したものである。すべての変数の中でモデル構築に必要な説明変数を“1”、そうでない変数を“0”と定義すると、モデル構築に必要な説明変数の組み合わせをバイナリ配列として表現することができる。例えば、総数が 5 つの変数セットに対して“01001”というバイナリ配列を与えた場合には、2 番目と 5 番目の変数がモデル化に必要ということになる。この定義に従えば、1 つのバイナリ配列 (遺伝型) に対して対応するモデル式が確定し、そのモデル式の予測性 (表現型) を評価できる。遺伝的アルゴリズムは、表現型の優劣に従い最適な遺伝型を決定する方法である。まず、第 1 世代として一定数のランダムな個体 (遺伝型と対応する表現型を持つ) を発生させる。次に、優れた表現型を持つ個体ができるだけ高確率で選択されるようにしながら 2 つの親を選び、それらの遺伝子配列を交叉、変異させて次世代の子供を作る。各世代において、子供の数が最初に設定した個体数になるまでこの操作を繰り返す。このように優れた遺伝子が受け継がれていくので、少なくとも平均的には子孫の世代は親の世代よりも優れた個体で構成されることになる。われわれは、変数選択の最適化

	genotype	fitness
Individual 1	0 1 1 0 1 0 0 1	$Y = a_2X_2 + a_3X_3 + a_5X_5 + a_8X_8$
Individual 2	1 0 1 1 0 1 0 1	$Y = a_1X_1 + a_3X_3 + a_4X_4 + a_6X_6 + a_8X_8$
Individual 3	0 1 1 1 1 1 0 0	$Y = a_2X_2 + a_3X_3 + a_4X_4 + a_5X_5 + a_6X_6$
⋮	⋮	⋮

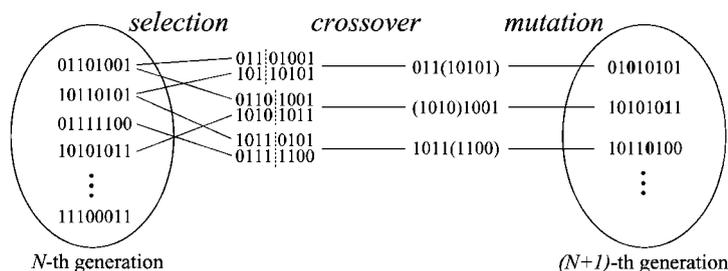


Fig. 1. Genetic Algorithm-based Optimization of QSAR Models

にはこの遺伝的アルゴリズムを利用し、多変量解析の方法としては部分最小二乗法を組み合わせた方法 (Genetic algorithm-combined partial least squares method, GA-PLS 法) を開発した.^{37,38)} 部分最小二乗法は、潜在変数ベクタから説明変数の総合特性値を計算して回帰する方法の1つであり、目的変数との間の共分散ができるだけ大きくなるように潜在変数ベクタを決定することに特徴がある.^{34,35)} 詳しくは別の総説^{34,35)}を参考いただきたい。

4.2. トポロジカル記述子からの溶解度の予測

医薬品の水への溶解度は、薬物動態学や製剤学的特性に関係する重要なパラメータの1つである。例えば、難溶性医薬品を経口投与した場合には、消化管内で溶出が起こらないためにバイオアベイラビリティが低くなるのが問題となる。そこで、薬物の分子構造から水への溶解度を予測する QSAR モデルの開発を行った。211 個の薬物あるいは開発化合物の水への溶解度に関するデータを収集し、構造記述子としては Molconn-Z Ver. 3.50 (Hall Associated Consulting, Quincy, MA) を使ってトポロジカル記述子を計算した。³⁷⁾ Molconn-Z によって得られる 220 個の記述子のうち、極端に偏った分布を示す記述子を取り除き、88 個に絞って GA-PLS 法による QSAR モデリングを行った。Figure 2(A) は遺伝的アルゴリズムによるモデル最適化の過程を示したものであり、QSAR モデルの優劣を評価するための適合度としては、

$$\text{fitness} = \frac{2}{\frac{1}{R^{*2}} + \frac{1}{r_{\text{pred}}^2}} \quad (1)$$

ここで、

$$R^{*2} = 1 - \frac{n-1}{n-c-1} \frac{\sum (y_i - y_{\text{cal}})^2}{\sum (y_i - \bar{y})^2} \quad (2)$$

(訓練用データセット)

$$r_{\text{pred}}^2 = 1 - \frac{\sum (y_i - y_{\text{pred}})^2}{\sum (y_i - \bar{y})^2} \quad (3)$$

(テスト用データセット)

であり、 n は化合物数、 c は PLS 主成分数、 \bar{y} は個々のデータ (y_i) の平均値、 y_{cal} と y_{pred} はそれぞれ訓練用、テスト用データセットにおける理論値を表す。自由度補正回帰係数 (R^{*2}) は、訓練用データに対する過学習を防ぐために導入した。 r_{pred}^2 値は、ランダムに分割したデータセットを用いて leave-some-out 法によりクロスバリデーションしたときの予測相関係数である。Figure 2(A) に示されるように、適合度の平均値は世代数とともに大きくなる傾向がみられ、約 200 世代でプラトーに到達した。最終的には、500 世代目における最良個体を選んだ。遺伝的アルゴリズムにより選択された 34 個の Molconn-Z 記述子のうち、非常に小さな補正 PLS 回帰係数を示す記述子は除いた。最終的に残った 29 個の記述子の 19 主成分で構成される PLS モデルが、水への溶解度を予測するための最良モデルであると判断した。Figure 2(B) はその QSAR モデルのバリデーション結果を示しているが、予測相

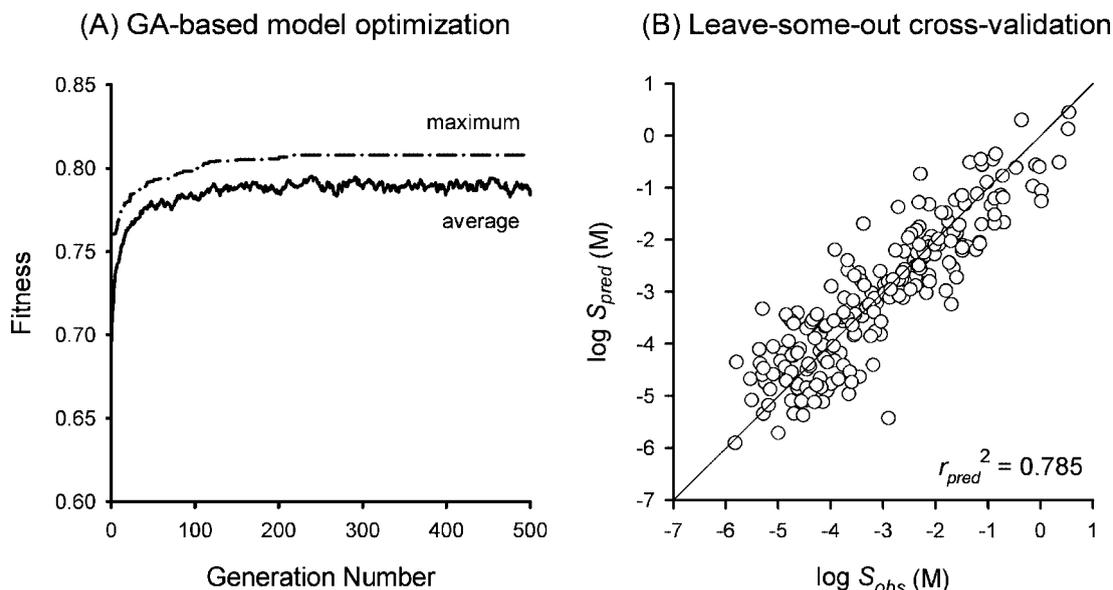


Fig. 2. QSAR Modelling of Aqueous Solubility of Compounds by Using Genetic Algorithm-combined Partial Least Squares (GA-PLS) Method

(A) shows trajectory of genetic algorithm-based model optimization. Fitness was defined as Eq. 1. (B) shows prediction of aqueous solubility evaluated by leave-some-out cross-validation procedure. An entire data set was randomly divided into 7 subsets for the leave-some-out cross-validation. The QSAR model was constructed using Molconn-Z descriptors as explanatory variables.

関係数 (r_{pred}^2) が 0.785 の良好な予測性を示すことが分かった。

神経伝達の仕組みを数理モデル化した人工ニューラルネットワークを用いると、水に対する溶解度がうまくシミュレーションできることがこれまでに報告されている。³⁹⁻⁴¹⁾ 人工ニューラルネットワークは説明変数と目的変数の複雑な関係をうまく説明できるものの、モデルの自由度が極めて大きくなるために得られたモデルの頑健性がしばしば問題となる。これに対して、PLS モデルの場合には、説明変数から抽出された総合特性値で目的変数のモデル化を行うのでモデルの自由度が小さいという特徴がある。比較のために、Huuskonen ら³⁹⁾ が報告しているデータを GA-PLS 法で解析したところ、得られたモデルの予測精度 (訓練データ: $r^2=0.85$, $s=0.57$, テストデータ: $r^2=0.89$, $s=0.47$) は、Huuskonen ら³⁹⁾ の階層的ニューラルネットワーク (訓練データ: $r^2=0.90$, $s=0.46$, テストデータ: $r^2=0.86$, $s=0.53$) とほぼ同等であった。したがって、モデルの頑健性も考慮すると、PLS モデルは非常に有効と考えられる。

4-3. トポロジカル記述子からの受動的膜透過性の予測 経口的に投与される医薬品のほとんどは受動拡散によって吸収される。その膜透過に関して

QSAR 研究が古くから行われてきたが、多くの研究は、比較的少数の薬物を対象としたものがほとんどであり、構造的に多様な薬物の膜透過を精度よく予測する方法は開発されていない。大腸がん由来の細胞株 Caco-2 細胞は、培養後小腸上皮細胞様に分化することから、薬物吸収評価にしばしば用いられている。⁷⁻⁹⁾ 比較的多くの透過性データが文献的にも入手できるので、その膜透過性に関する QSAR モデルを構築することは膜透過予測モデルとして有効と考えられる。そこで、溶解度に対する予測モデルを開発したときと同様に Molconn-Z トポロジカル記述子を使い、Caco-2 細胞透過性の予測モデルの開発を試みた。まず、文献より Caco-2 細胞透過性のデータを収集し、分子量の大きな化合物 ($MW > 1000$) や特殊輸送担体に認識される化合物のデータを除き、最終的に 73 個の化合物で総数 110 の透過性データを得た。³⁸⁾ みかけの透過係数の対数値 ($\log P_{app}$ (cm/s)) の範囲は、 -6.96 から -3.88 であった。GA-PLS 法による QSAR モデリングを行った結果、予測相関係数 (r_{pred}^2) は 0.683 の良好なモデルが得られた (Fig. 3)。このモデルは 24 個の Molconn-Z 記述子の 12 主成分によって構成され、すべてのデータを用いてフィッティングを行うと、次の線形予測式が得られた。

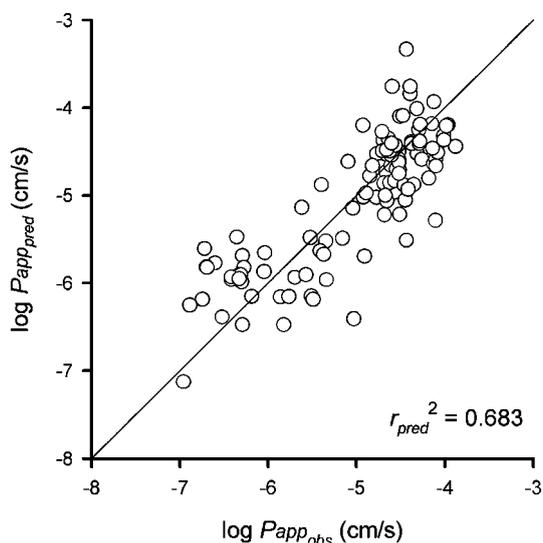


Fig. 3. QSAR Modelling of Caco-2 Cell Permeability for Compounds by Using Genetic Algorithm-combined Partial Least Squares (GA-PLS) Method

This figure shows prediction of Caco-2 cell permeability evaluated by leave-some-out cross-validation procedure. An entire data set was randomly divided into 6 subsets for the leave-some-out cross-validation. The QSAR model was constructed using Molconn-Z descriptors as explanatory variables.

$$\begin{aligned} \log P_{app} = & 0.0223 fw - 0.2482^2 \chi - 0.1481^4 \chi_p \\ & - 0.2510^5 \chi_p - 0.2082^6 \chi_p - 0.0302^3 \chi_{vp} \\ & - 0.0688^4 \chi_{vp} - 0.0995^5 \chi_{vp} - 0.0822^6 \chi_{vp} \\ & - 0.1561^3 \chi_c - 0.1831^3 \kappa_\alpha + 0.0164 \text{totop} \\ & + 0.2970 \text{sumDELI} - 0.0412 \text{SHsOH} \\ & - 0.1597 \text{SHother} - 0.1298 \text{Gmin} \\ & - 0.1220 \text{SsCH3} - 0.1119 \text{SssCH2} \\ & - 0.2093 \text{SsssN} + 0.0840 \text{SsOH} \\ & - 0.0502 \text{SssO} + 0.2543 \text{SHCsats} \\ & - 0.3990 \text{SHHBd} - 0.1694 \text{SHHBa} \\ & - 3.3156 \end{aligned} \quad (4)$$

$(r^2=0.785, s=0.389, n=110)$

一方、分子軌道計算と人工ニューラルネットワークを組み合わせて、Caco-2細胞の透過性を予測するモデルについても検討を行った。⁴²⁾この場合には、半経験的分子軌道法MOPAC97を用いて化合物の分子構造を決定し、双極子モーメント、分極率、水素結合受容可能な窒素原子の電荷の総和、酸素原子の電荷の総和、及び水素結合供与可能な水素原子の電荷の総和の計5つの分子記述子を得た。5-5-1の構造を持つ階層型ニューラルネットワークを用いて学習を行ったところ、leave-some-out法により評価される予測相関係数(r_{pred}^2)は0.623であった。⁴²⁾これは、先のMolconn-Zトポロジカル記

述子を用いたPLSモデルとほぼ同等であった。トポロジカル記述子が化合物の2次元構造式から簡単に計算できることを考慮すると、トポロジカル記述子を用いたPLSモデルのほうが大規模なバーチャルスクリーニングに適すると考えられる。

4-4. ペプチドトランスポータを介した輸送の予測：2次元QSARと3次元QSARの比較 これまでに薬物を認識する担体が数多く報告され、薬物の体内動態や薬効と密接に関係することが知られている。担体による基質の分子認識は輸送活性を決定する重要な因子であるが、担体タンパク質の結晶構造が明らかにされているケースはほとんどなく、その物理化学的特徴に関してほとんど情報がない。3次元QSAR法は基質の3次元構造からレセプター側をマッピングする方法であり、Craimerらによって提案されたcomparative molecular field analysis (CoMFA)はその最も代表的な手法である。⁴³⁾最近、3次元QSAR法を利用して担体介在性輸送を解析する研究が報告されている。^{28-30,44,45)}しかしながら、3次元QSARでは、基質となる化合物構造のコンフォメーションやアラインメントにおいて自由度が問題であり、計算に多くの時間を要するため、かならずしも大規模なスクリーニングには適していない。そこで、このような担体介在性輸送に対して単純な2次元QSARが適用できるかどうかを検討し、3次元QSARとの比較を行った。

この研究では、Caco-2細胞における20種類のβラクタム系抗生物質の取り込みについて解析した。⁴⁶⁾βラクタム系抗生物質は、Caco-2細胞に発現するペプチドトランスポータPEPT1によって能動的に取り込まれることが知られている。3次元QSARの代表例であるCoMFA法は基質の3次元アラインメントを通じて化合物の輸送活性との相関から仮想的結合部位の静電的及び立体的エネルギーポテンシャルを推定する方法である。⁴³⁾SYBYL/CoMFAを利用してCaco-2細胞でのβラクタム系抗生物質の取り込みを解析したところ、leave-one-out法による予測相関係数(r_{pred}^2)が0.759の良好なモデルが得られた(Fig. 4(A))。一方、Molconn-Zトポロジカル記述子を構造記述子とし、GA-PLS法によりモデル最適化を行った場合には、予測相関係数(r_{pred}^2)が0.923の非常によいモデルが得られた(Fig. 4(B))。⁴⁶⁾

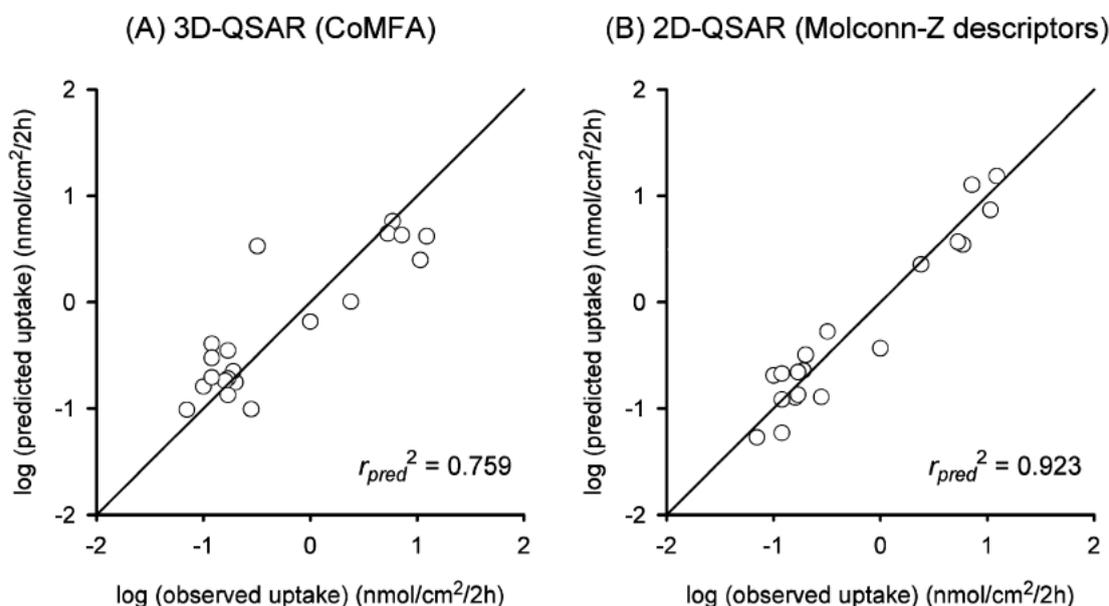


Fig. 4. QSAR Modelling of Uptake of β -Lactam Antibiotics in Caco-2 Cells

(A) shows the predictability of CoMFA model, and (B) shows that of 2D-QSAR model using Molconn-Z descriptors optimized by GA-PLS method. The predictability was evaluated by a leave-one-out cross-validation procedure.

CoMFA による解析は、予測性において若干 2 次元 QSAR に劣ったものの、“仮想”結合部位をグラフィカルにシミュレーションでき、薬物設計のための方向性を合理的に示すことができるという特徴がある。Figure 5 は、 β ラクタム系抗生物質の Caco-2 細胞への取り込みデータ解析から得られた CoMFA フィールドマップであり、ペプチドトランスポータによって良好に輸送される cephadroxil の 3 次元構造とともに示したものである。⁴⁶⁾ 領域②及び④で示されるように、3 位に大きくかつ求電子性ではない官能基を持つ化合物（例えば cefotaxime や cephalothin）の場合には取り込み速度が低下することが理解される。また、ペニシリンの 6 位あるいはセファロsporin の 7 位の近くに領域①及び③が認められ、 α 炭素原子の置換基の位置によって輸送活性が左右され、求核性基があると輸送が高くなることも推測される。

一方、トポロジカル記述子を用いた 2 次元 QSAR では、CoMFA のように合理的薬物設計のための情報を提供しないが、計算が単純で高速であるという利点がある。したがって、バーチャルに合成されたコンビナトリアルライブラリーを大規模にスクリーニングする際には、2 次元 QSAR は非常に有効である。創薬研究においては化合物ライブラリーが絞られるにつれ、ランダムスクリーニングか

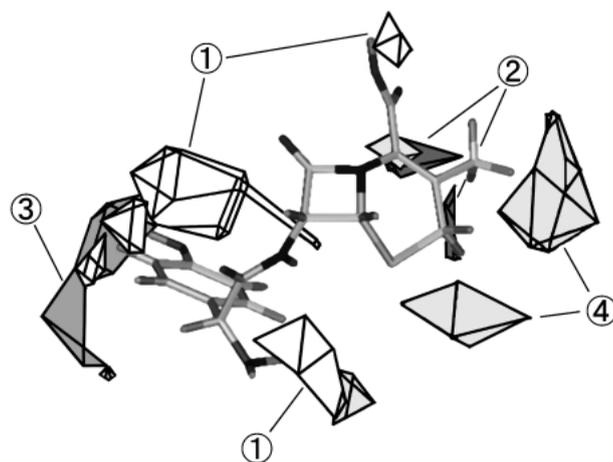


Fig. 5. Contour Diagram of CoMFA Steric and Electrostatic Fields with Cefadroxil

Sterically favored regions (contribution level of 85%) are represented by region 1, whereas sterically disfavored regions (contribution level of 15%) are represented by region 2. Negative charge-favored regions (contribution level of 85%) are represented by region 3, whereas positive charge-favored regions (contribution level of 15%) are represented by region 4.

ら合理的薬物設計へと移っていくので、*in silico* 研究においてもこの変遷に対応して QSAR アプローチを適切に選択することが必要である。

4-5. 薬物代謝酵素 CYP3A4 に対する阻害効果の予測 CYP3A4 は、薬物代謝酵素チトクロム P450 の中で、ヒト肝臓に最も多く存在するアイソフォームであり、既知薬物の約 50% の薬物代謝に関与することが知られている。⁴⁷⁾ また、CYP3A4 は

小腸にも分布し、経口投与された薬物の消化管吸収性を制限していることも報告されている。⁴⁸⁾ 併用された薬物によって CYP3A4 が阻害され、致死的な薬物間相互作用が引き起こされる事例も報告されており、⁴⁹⁾ 現在創薬研究において CYP3A4 の阻害剤となり得るかどうかを早期に予測して臨床での相互作用を未然に防ぐことが強く望まれている。そこでわれわれは、CYP3A4 による薬物代謝を予測する QSAR モデルの開発を行った。⁵⁰⁾

まず、7-benzyloxy-4-trifluoromethyl coumarin (BFC) を基質とし、リコンビナント CYP3A4 発現系マイクロソームを用いた 96 穴マイクロプレート法によって構造的に多様な 53 種の薬物の代謝阻害効果を測定した。⁵⁰⁾ その結果、評価で得られた IC₅₀ 値は薬物によって大きく異なり、溶解度の制限により評価できなかった例を除くと、その範囲は 9 nM から 2 mM であった。

溶解度の制限あるいは BFC 代謝の活性化により評価できなかった薬物を除き、44 種の薬物の IC₅₀ 値データを用いて QSAR モデルの開発を行った。⁵⁰⁾ 35 化合物をランダムに抽出して訓練用データとし、残りの 9 化合物を外部バリデーション用データとした。GA-PLS 法を用いてトポロジカル記述子からの予測モデルの開発を行った結果、Fig. 6 にも示されるように、訓練用データに対する leave-one-out 予測の予測相関係数 (r_{pred}^2) は 0.569 であり、外部バリデーション用データに対する予測精度もほ

ぼ同等であった ($r_{pred}^2=0.553$)。また、Fig. 7 に示されるように、PLS 回帰モデルの第 1 主成分は分子量を非常によく相関することが分かった ($r^2=0.801$)。第 1 主成分の寄与は回帰全体の 25% 程度であるが、過去の報告と併せて考えると、^{51,52)} 薬物の分子量が薬物と CYP3A4 との相互作用に重要であることが示された。⁵⁰⁾

Ekins らは、代表的な 3 次元 QSAR ソフトウェアである Catalyst を用いて CYP3A4 による代謝の予測を行っている。²⁷⁾ そこで、2 次元 QSAR と 3 次元 QSAR を比較する目的で、Ekins らのデータセ

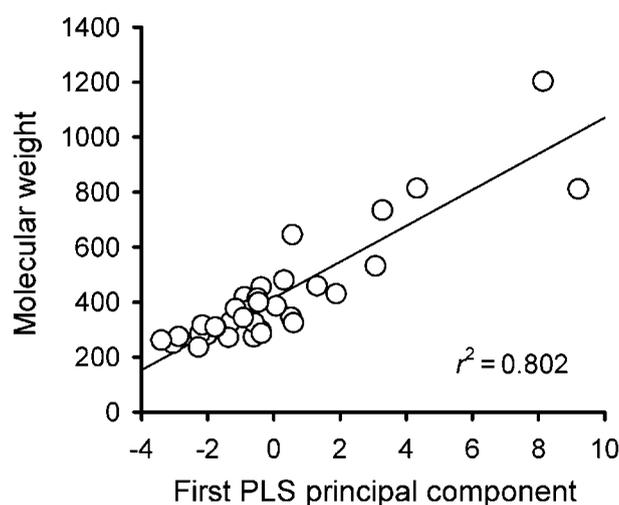


Fig. 7. Relationship between the First PLS Principal Component of QSAR Model for CYP3A4 Activity and Molecular Weight of Drugs
The QSAR model was the same as Fig. 6.

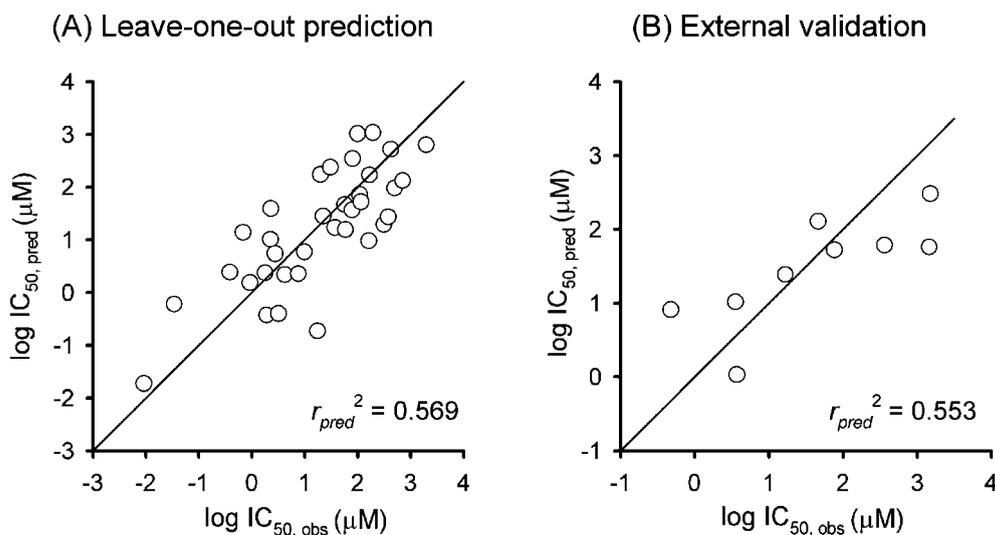


Fig. 6. QSAR Modelling of Inhibition of CYP3A4 Activity by Structurally Diverse Drugs

(A) and (B) show predictability of the model evaluated by leave-one-out cross-validation and external validation, respectively. The QSAR model was constructed using Molconn-Z descriptors as explanatory variables and optimized by GA-PLS method.

ット²⁷⁾をわれわれの方法で解析した。Ekinsらは、ヒト肝ミクロソームでの代謝を評価した38個の化合物のKm値をCatalystによりファーマコフォアモデルを構築し、その際の予測相関係数(r^2)が0.45($n=38$)であり、12個のテスト用データセットに対する予測残差が1 log unit以下($s=0.55$, $n=12$)であることを報告している。²⁷⁾一方、われわれの2次元QSAR解析では、予測相関係数(r^2)が0.76であった。また、12個のテスト用データセットに対する予測残差もEkinsらの値とほぼ同じであり($s=0.57$, $n=12$)、2次元QSARモデルは3D-QSARモデルと同等の予測性を持つことが示された。⁵⁰⁾

5. おわりに

前述したように、薬物動態特性を予測する*in silico*法には分子シミュレーションに基づく機構論的な方法と実験データに基づく経験的な方法がある。われわれは、基本的に後者の方法を主に採用し、解析を行ってきた。この方法は、実験の結果のみに着目してモデルを構築するため、複雑系で理論的な取り扱いができない場合にも予測モデルを構築できるという利点がある。われわれは、今回の研究を通じて、化合物の2次元構造式から簡便に計算可能なトポロジカル記述子によって様々な薬物動態関連パラメータがモデル化できることを明らかにした。その精度は、より分子論的な性格を持つ3次元QSARとほぼ同等であり、特に大規模なバーチャルスクリーニングにおいて有効と考えられる。しかしながら、今回十分には検証できなかったが、このアプローチはあくまで実験結果に基づく統計学的モデルであり、既存のパターンとは大きく異なる化合物の結果を予測することは困難である。また、精度だけを問題とし、実験データに対する過剰な学習モデル化を進めると、予測に対する汎化能力を失うことも多々ある。本研究ではクロスバリデーション法により汎化能力の損失を防いだが、これらのモデルの妥当性を確認するためにはさらなるデータの蓄積と検証が必要である。本研究で開発したGA-PLS法はQSARモデルを最適化するために有効な方法であり、特に近年ハイスループットスクリーニングにより収集される大規模なデータを取り扱う際に能力を発揮すると期待される。以上、本研究で得られた知見は、創薬研究を支援するためのアプローチとして

有益な基礎情報を提供するものと思われる。

謝辞 本稿で紹介した研究は、京都大学大学院薬学研究科助教授 山下富義博士らとともに推進しているものであり、この場をお借りして厚く御礼申し上げます。

REFERENCES

- 1) Li A. P., *Drug Discov. Today*, **6**, 357-366 (2001).
- 2) Lin J., Sahakian D. C., de Morais S. M., Xu J. J., Polzer R. J., Winter S. M., *Curr. Top. Med. Chem.*, **3**, 1125-1154 (2003).
- 3) Thompson T. N., *Curr. Drug Metab.*, **1**, 215-241 (2000).
- 4) Eddershaw P. J., Beresford A. P., Bayliss M. K., *Drug Discov. Today*, **5**, 409-414 (2000).
- 5) Lipinski C. A., Lombardo F., Dominy B. W., Feeney P. J., *Adv. Drug Deliv. Rev.*, **23**, 3-26 (1997).
- 6) Bevan C. D., Lloyd R. S., *Anal. Chem.*, **72**, 1781-1787 (2000).
- 7) Delie F., Rubas W., *Crit. Rev. Ther. Drug Carrier Syst.*, **14**, 221-286 (1997).
- 8) Yamashita S., Furubayashi T., Kataoka M., Sakane T., Sezaki H., Tokuda H., *Eur. J. Pharm. Sci.*, **10**, 195-204 (2000).
- 9) Ingels F. M., Augustijns P. F., *J. Pharm. Sci.*, **92**, 1545-1558 (2003).
- 10) Irvine J. D., Takahashi L., Lockhart K., Cheong J., Tolan J. W., Selick H. E., Grove J. R., *J. Pharm. Sci.*, **88**, 28-33 (1999).
- 11) Kansy M., Senner F., Gubernator K., *J. Med. Chem.*, **41**, 1007-1010 (1998).
- 12) Sugano K., Nabuchi Y., Machida M., Aso Y., *Int. J. Pharm.*, **257**, 245-251 (2003).
- 13) Hengstler J. G., Utesch D., Steinberg P., Platt K. L., Diener B., Ringel M., Swales N., Fischer T., Biefang K., Gerl M., Bottger T., Oesch F., *Drug Metab. Rev.*, **32**, 81-118 (2000).
- 14) Hakura A., Suzuki S., Satoh T., *Mutat. Res.*, **438**, 29-36 (1999).
- 15) Miller V. P., Stresser D. M., Blanchard A. P., Turner S., Crespi C. L., *Ann. N. Y. Acad. Sci.*, **919**, 26-32 (2000).
- 16) van de Waterbeemd H., Gifford E., *Nat. Rev. Drug Discov.*, **2**, 192-204 (2003).

- 17) Ekins S., Waller C. L., Swaan P. W., Cruciani G., Wrighton S. A., Wikel J. H., *J. Pharmacol. Toxicol. Methods*, **44**, 251–272 (2000).
- 18) Yamashita F., Hashida M., *Drug Metab. Pharmacokinet.*, **19**, 327–338 (2004).
- 19) Kemp C. A., Marechal J. D., Sutcliffe M. J., *Arch. Biochem. Biophys.*, **433**, 361–368 (2005).
- 20) Wester M. R., Yano J. K., Schoch G. A., Yang C., Griffin K. J., Stout C. D., Johnson E. F., *J. Biol. Chem.*, **279**, 35630–35637 (2004).
- 21) Williams P. A., Cosme J., Ward A., Angove H. C., Vinkovic D. M., Jhoti H., *Nature*, **424**, 464–468 (2003).
- 22) Williams P. A., Cosme J., Vinkovic D. M., Ward A., Angove H. C., Day P. J., Vonnrhein C., Tickle I. J., Jhoti H., *Science*, **305**, 683–686 (2004).
- 23) Yano J. K., Wester M. R., Schoch G. A., Griffin K. J., Stout C. D., Johnson E. F., *J. Biol. Chem.*, **279**, 38091–38094 (2004).
- 24) Ekins S., Bravi G., Ring B. J., Gillespie T. A., Gillespie J. S., Vandenbranden M., Wrighton S. A., Wikel J. H., *J. Pharmacol. Exp. Ther.*, **288**, 21–29 (1999).
- 25) Ekins S., Bravi G., Binkley S., Gillespie J. S., Ring B. J., Wikel J. H., Wrighton S. A., *Drug Metab. Dispos.*, **28**, 994–1002 (2000).
- 26) Ekins S., Bravi G., Binkley S., Gillespie J. S., Ring B. J., Wikel J. H., Wrighton S. A., *Pharmacogenetics*, **9**, 477–489 (1999).
- 27) Ekins S., Bravi G., Wikel J. H., Wrighton S. A., *J. Pharmacol. Exp. Ther.*, **291**, 424–433 (1999).
- 28) Ekins S., Kim R. B., Leake B. F., Dantzig A. H., Schuetz E. G., Lan L. B., Yasuda K., Shepard R. L., Winter M. A., Schuetz J. D., Wikel J. H., Wrighton S. A., *Mol. Pharmacol.*, **61**, 974–981 (2002).
- 29) Ekins S., Kim R. B., Leake B. F., Dantzig A. H., Schuetz E. G., Lan L. B., Yasuda K., Shepard R. L., Winter M. A., Schuetz J. D., Wikel J. H., Wrighton S. A., *Mol. Pharmacol.*, **61**, 964–973 (2002).
- 30) Bednarczyk D., Ekins S., Wikel J. H., Wright S. H., *Mol. Pharmacol.*, **63**, 489–498 (2003).
- 31) Shen M., Xiao Y., Golbraikh A., Gombar V. K., Tropsha A., *J. Med. Chem.*, **46**, 3013–3020 (2003).
- 32) Niwa T., *J. Chem. Inf. Comput. Sci.*, **43**, 113–119 (2003).
- 33) Yamashita F., Hashida M., *Drug Deliv. Syst.*, **18**, 22–29 (2003).
- 34) Lindberg W., Persson J. A., Wold S., *Anal. Chem.*, **55**, 643–648 (1983).
- 35) Geladi P., Kowalski B. R., *Anal. Chim. Acta*, **185**, 1–17 (1986).
- 36) Ichikawa H., *Adv. Drug Deliv. Rev.*, **55**, 1119–1147 (2003).
- 37) Wanchana S., Yamashita F., Hashida M., *Pharmazie*, **57**, 127–129 (2002).
- 38) Yamashita F., Wanchana S., Hashida M., *J. Pharm. Sci.*, **91**, 2230–2239 (2002).
- 39) Huuskonen J., Salo M., Taskinen J., *J. Chem. Inf. Comput. Sci.*, **38**, 450–456 (1998).
- 40) Huuskonen J., Rantanen J., Livingstone D., *Eur. J. Med. Chem.*, **35**, 1081–1088 (2000).
- 41) Yan A., Gasteiger J., *J. Chem. Inf. Comput. Sci.*, **43**, 429–434 (2003).
- 42) Fujiwara S., Yamashita F., Hashida M., *Int. J. Pharm.*, **237**, 95–105 (2002).
- 43) Cramer R. D., Patterson D. E., Bunce J. D., *J. Am. Chem. Soc.*, **110**, 5959–5967 (1988).
- 44) Swaan P. W., Szoka Jr. F. C., Oie S., *J. Comput. Aided Mol. Des.*, **11**, 581–588 (1997).
- 45) Swaan P. W., Koops B. C., Moret E. E., Tucker J. J., *Receptors Channels*, **6**, 189–200 (1998).
- 46) Wanchana S., Yamashita F., Hara H., Fujiwara S., Akamatsu M., Hashida M., *J. Pharm. Sci.*, **93**, 3057–3065 (2004).
- 47) Guengerich F. P., *Adv. Pharmacol.*, **43**, 7–35 (1997).
- 48) Paine M. F., Khalighi M., Fisher J. M., Shen D. D., Kunze K. L., Marsh C. L., Perkins J. D., Thummel K. E., *J. Pharmacol. Exp. Ther.*, **283**, 1552–1562 (1997).
- 49) Honig P. K., Wortham D. C., Zamani K., Conner D., Mullin J. C., Cantilena L. R., *JAMA*, **269**, 1513–1518 (1993).
- 50) Wanchana S., Yamashita F., Hashida M., *Pharm. Res.*, **20**, 1401–1408 (2003).
- 51) Austin R. P., Barton P., Cockroft S. L., Wenlock M. C., *Drug Metab. Dispos.*, **30**, 1497–1503 (2002).
- 52) Leo A., Hansch C., Jow P. Y. C., *J. Med. Chem.*, **19**, 611–615 (1976).