Review

In Silico Approaches for Predicting ADME Properties of Drugs

Fumiyoshi YAMASHITA* and Mitsuru HASHIDA
Department of Drug Delivery Research, Graduate School of Pharmaceutical Sciences,
Kyoto University, Kyoto, Japan

Summary: Combinatorial chemistry and high-throughput screening have increased the possibility of finding new lead compounds at much shorter time periods than conventional medicinal chemistry. However, too much promising drug candidates often fail because of unsatisfactory ADME properties. In silico ADME studies are expected to reduce the risk of late-stage attrition of drug development and to optimize screening and testing by looking at only the promising compounds. To this end, many in silico approaches for predicting ADME properties of compounds from their chemical structure have been developed, ranging from data-based approaches such as quantitative structure-activity relationship (QSAR), similarity searches, and 3-dimensional QSAR, to structure-based methods such as ligand-protein docking and pharmacophore modelling. In addition, several methods of integrating ADME properties to predict pharmacokinetics at the organ or body level have been studied. In this article, we briefly summarize in silico ADME approaches.

Key words: in silico; ADME; QSAR; mechanism-based approach; empirical approach; physiologically based pharmacokinetics

Introduction

In the last decade, much attention has been paid for combinatorial chemistry and high throughput screening in drug discovery setting. Combinatorial chemistry makes it possible to synthesize large series of libraries of compounds using the same chemical reaction. Then, the libraries are run through the high throughput screening mostly based on binding assays to target biomolecules. Such new technologies have increased the possibility of finding new lead compounds at much shorter time periods than conventional medicinal chemistry. However, too much promising drug candidates often fail because of unsatisfactory ADME properties. In applying such new technologies, it is more important to make effort to improve the rate of success in the more costly downstream stages of drug development.

To this end, ADME/Tox evaluations have been shifted into early discovery stages, such as lead identification or optimization, to be conducted in parallel with pharmacological activity assays. Early ADME/Tox studies would also minimize time-, cost-, and labor-intensiveness of screening and testing by looking at only the promising compounds. Such ADME/Tox evaluations became possible by developing high-throughput solubility assays (turbidometry, laser nephelometry), permeability assays (Caco-2 cells, MDCK cells, PAMPA), and metabolism assays (hepatocytes, S9 fractions, recombinant enzymes). Moreover, accumulations of these data are expected to provide us with modelling the processes and relating these to calculated physicochemical and structural features of the compounds.

Recently, in silico modelling of ADME properties have been performed using different approaches. These methods ranges from data-based approaches such as quantitative structure-activity relationship (QSAR), similarity searches, and 3-dimensional QSAR, to structure-based methods such as ligand-protein docking and pharmacophore modelling. The ultimate goal of in silico ADME is to predict disposition behavior of compounds in the whole body by assembling all kinetic processes in one global model. In this article, we review in silico methods for predicting ADME properties of compounds from chemical structures and integrating them to simulate the kinetics at the organ or body levels.
Structure-Based Approaches for Predicting ADME Properties

Structure-based approaches include the three-dimensional molecular modelling of ligands and proteins such as cytochrome P450s. Quantum mechanical methods are generally used to estimate the interaction of small molecules with a target protein. If no information on the crystallographic protein structure is available, the three-dimensional structure can be built by homology modelling of known proteins. An alternative way of assessing the potential of a small molecule to interact with a particular protein is to use pharmacophore models. The models are built from a superposition of known substrates of the protein.

Molecular modelling of drug metabolizing enzymes: To understand the interaction of drugs with drug metabolizing enzymes, their three-dimensional structures have been determined by homology modelling. As an example of these studies, De Rienzo et al. built three-dimensional models of the cytochromes P450 1A2, 2D6 and 3A4 by means of restraint-based comparative modelling using the X-ray crystallographic structures of bacterial cytochromes P450s (CAM, BM-3, TERP and ERYF) as templates. Sequence identity percentages between each of the target P450s and the templates P450s were close to 20%. Fifty specific and non-specific substrates were docked into the active site of their metabolizing enzyme by the automatic rigid body docking program DOCK. They found that substrate binding into the active site of 2D6 is mainly favored by hydrogen bonding and electrostatic interactions between the substrate and two residues (D301 and S304); in particular, D301 performs charge-reinforced hydrogen-bonding interactions with the protonated nitrogen atom that characterizes the 2D6 substrates. On the contrary, van der Waals attractive interactions mainly contribute to stabilize the complexes involving both 1A2 and 3A4. They also proposed several molecular and intermolecular-interaction descriptors, in order to translate qualitative structural information of substrate/cytochrome interaction models into semi-quantitative models of substrate specificity.

Recently, the crystal structure of a human cytochrome P450 2C9, in the absence and presence of warfarin, was determined. More accumulation of their metabolizing enzyme by the automatic rigid body docking program DOCK. They found that substrate binding into the active site of 2D6 is mainly favored by hydrogen bonding and electrostatic interactions between the substrate and two residues (D301 and S304); in particular, D301 performs charge-reinforced hydrogen-bonding interactions with the protonated nitrogen atom that characterizes the 2D6 substrates. On the contrary, van der Waals attractive interactions mainly contribute to stabilize the complexes involving both 1A2 and 3A4. They also proposed several molecular and intermolecular-interaction descriptors, in order to translate qualitative structural information of substrate/cytochrome interaction models into semi-quantitative models of substrate specificity.

Recently, the crystal structure of a human cytochrome P450 2C9, in the absence and presence of warfarin, was determined. More accumulation of their metabolizing enzyme by the automatic rigid body docking program DOCK. They found that substrate binding into the active site of 2D6 is mainly favored by hydrogen bonding and electrostatic interactions between the substrate and two residues (D301 and S304); in particular, D301 performs charge-reinforced hydrogen-bonding interactions with the protonated nitrogen atom that characterizes the 2D6 substrates. On the contrary, van der Waals attractive interactions mainly contribute to stabilize the complexes involving both 1A2 and 3A4. They also proposed several molecular and intermolecular-interaction descriptors, in order to translate qualitative structural information of substrate/cytochrome interaction models into semi-quantitative models of substrate specificity.

Pharmacophore models for analysis of substrate specificity: Essential functionalities of a molecule required for its pharmacological activity are called pharmacophores. Pharmacophoric groups can be identified from the structure-activity data by comparing the structure of active compounds with one another. In particular, three-dimensional pharmacophoric geometry is important in determining the ability of ligands to bind to a receptor, since the binding cavities are spatially constrained regions with specific steric requirements. Such an indirect structure-based approach is effective even when the crystal structure of the protein of interest is not known. Pharmacophore modelling can be done by (i) identification of chemically equivalent atoms or groups in terms of their physical or chemical properties, (ii) estimation of the relative 3-dimensional position of the possible pharmacophoric groups in allowed low-energy conformations of the molecules, and (iii) weighing similar pharmacophoric groups when there are multiple choices. Swaan and Tukker performed pharmacophore mapping for the intestinal peptide transporter by conformational search and molecular alignment programs implemented in the SYBYL molecular modelling package (Tripos, Inc., St. Louis, MO). Penicillins, cephalosporins, ACE inhibitors, and few other drugs that were known as substrates of the peptide transporter were subjected to the pharmacophore analysis. They concluded that affinity for the peptide transporter can be diminished or abolished in either of three ways: esterification of the free carboxylic acid moiety; introduction of a second negative group; and intramolecular steric hindrance of the free carboxylic acid by either side chains with a positively charged nitrogen function or groups capable of hydrogen bond formation.

Catalyst is an integrated commercially available software environment (Accelrys, Inc., San Diego, CA) that produces pharmacophores, commonly referred to as ‘hypotheses’, with the HypoGen algorithm. In order to maximize conformational space occupied by these molecules, multiple conformers were generated for all molecules subjected. ‘Hypotheses’ are generated using the conformers and the activity values after selecting chemical features likely to be important (e.g., hydrogen bond donors and acceptors). Catalyst has often been used for pharmacophore modelling drug metabolizing enzymes such as CYP2B6, CYP2C9, CYP2D6 and CYP3A4, and transporters such as P-glycoprotein and organic cation transporter. Recently, Ekins et al. developed pharmacophore models for inhibitory effects of 14 inhibitors on 7-benzoyloxy-4-trifluoromethylcoumarin (BFC) metabolism by recombinant CYP3A4, CYP3A5 and CYP3A7, and directly compared potential structural features and positioning differences for each enzyme. The CYP3A4 pharmacophore was characterized by the large distance (14.3 Å) between the furthest ring aromatic (hydrophobic) feature and the hydrogen-bond acceptor, whereas CYP3A5 and CYP3A7 contained the hydrogen-bond acceptor feature and a compact arrangement of three hydrophobic features. This indicated that compounds...
that inhibit the metabolism of BFC by CYP3A4 are likely to have key pharmacokinetic interactions further from the hydrogen-bond acceptor than in those molecules that inhibit CYP3A5 and CYP3A7. Patel et al.\textsuperscript{34} compared three commercially available pharmacophore generation programs, Catalyst/ HipHop, DISCO and GASP on their ability to generate known pharmacophores deduced from different protein-ligand complexes (thrombin, cyclin dependent kinase 2, dihydrofolate reductase, HIV reverse transcriptase and thermolysin). They reported that GASP and Catalyst outperformed DISCO at reproducing the five target pharmacophores.

**Data-Based Approaches for Predicting ADME Properties**

Data-based modeling approaches are effective for many of ADME processes such as passive membrane permeation, where their molecular mechanism is hardly delineated. In fact, QSAR methods have been widely used for predicting ADME properties of compounds from their chemical structure. These methods primarily include two steps: (1) to obtain molecular descriptors based on the chemical structure of the compounds and (2) to relate the target property to the descriptors through multivariate statistical analyses. So far, various kinds of quantitative descriptors based on 2-dimensional or 3-dimensional molecular structures have been proposed, including fragment descriptors, topological descriptors, and global physicochemical descriptors. On the other hand, linear methods (e.g., multiple linear regression and partial least squares) and non-linear methods (e.g., feed-forward artificial neural network) have been applied for multivariate analysis.

In contrast, there are also several data-based methods that do not use an explicit mathematical model. These are based on molecular similarity/dissimilarity, including k-nearest neighbor method\textsuperscript{35} and stochastic artificial neural network.\textsuperscript{36} The property of unknown compound can be predicted from those of compounds that are registered in a data base and similar in chemical structure.

Such data-based methods are empirical, so their application would be limited: for example, data errors decrease performance in the prediction model and extrapolative prediction is basically difficult. By contrast, structure-based theoretical approaches would provide a concrete interpretation of the logic, while their application is limited by several problems such as long computation time and uncertainty of approximations. At this moment, therefore, data-based approaches would be useful for prediction of complex physicochemical and biological phenomena.

**Prediction of solubility by QSAR:** Solubility of drugs in water closely affects their biological activity: for example, it is one of important factors determining oral bioavailability. Various algorithms to predict aqueous solubility have been proposed, including both physicochemical and empirical approaches.

Yalkowsky et al.\textsuperscript{37,38} developed a thermodynamics-based equation, assuming that the solubility of a solid solute in water is the product of the ideal solubility and the solubility of the liquid in water. The former determines the entropy of melting and the melting point, while the latter can be derived from the octanol/water partition coefficient. Accordingly, the solubility of a solid solute in water ($S_w$) is approximated to:

$$\log S_w = 0.5 - 0.01(mp - 25) - \log K_{ow}$$

where $mp$ and $K_{ow}$ express the melting point and octanol/water partition coefficient, respectively. They demonstrated that the equation successfully predicts the aqueous solubilities of a set of 664 organic compounds, with an average absolute error of 0.45 log units and a root mean square error of 0.62 log units.\textsuperscript{39} On the other hand, Abraham et al. proposed an amended solvation energy relationship in prediction of the solubility of solutes in water:

$$\log S_w = 0.510 - 1.020R_t + 0.813\pi^d + 2.124\Sigma\sigma^d + 4.187\Sigma\rho^d - 3.337\Sigma\sigma^d\Sigma\rho^d - 3.986V_s$$

where $R_t$ is an excess molar refraction; $\pi^d$ is the dipolarity/polarizability; $\Sigma\sigma^d$ is the overall or summation hydrogen-bond acidity; $\Sigma\rho^d$ is the overall or summation hydrogen-bond basicity; and $V_s$ is the McGowan characteristic volume. Yang et al.\textsuperscript{40} demonstrated that Abraham’s and Yalkowsky’s equations are similar in their predictability. They also stated advantages and disadvantages of both approaches: the former is based on multiple linear regression analysis of a large training set that may or may not contain the required structural fragments, whereas the latter is simpler and more user-friendly; however, the latter requires knowledge of either an experimentally determined or an estimated melting of the solute.

On the other hand, applicability of these physicochemical models to structurally complex compounds has not yet been investigated. Huuskonen et al.\textsuperscript{41} developed an empirical method for predicting the aqueous solubility of drugs and related compounds based on topological indices and artificial neural network (ANN) modeling. A feed-forward ANN that was trained using 160 compounds gave a reasonable predictive $r^2$ of 0.86 for 51 compounds. We developed a model of predicting the solubility of the same compounds that was optimized by genetic algorithm-combined partial least squares methods.\textsuperscript{42} In the leave-one-out prediction procedure, the model gave the cross-validated predictive
prediction of intestinal permeability by QSAR:

Many researchers have investigated the quantitative structure/property relationships (QSPR) involving oral absorption of candidate compounds. Oral absorption or intestinal membrane permeability is too complicated a process to be predicted based on their molecular mechanisms. Therefore, prediction of such transport at the tissue or cellular level must rely on data-based approaches at this moment. In these studies, several kinds of quantitative descriptors based on 2-dimensional or 3-dimensional molecular structures have been proposed, including fragment descriptors,\(^{42}\) hydrophobicity (logP),\(^{43–45}\) hydrogen bonding descriptors,\(^{43–45}\) topological indices,\(^{46–51,52}\) polar surface area,\(^{47–49,50,53–56}\) and quantum chemical parameters.\(^{46,49,51–57,59}\) In addition, to study the relationship between these molecular descriptors and oral absorption, multiple linear regression,\(^{52–46,50,55–57}\) partial least squares,\(^{52,57,59,60}\) and artificial neural network\(^{31,39,60}\) have been used. Recently, there have been many QSAR studies regarding permeability across the Caco-2 cell monolayers, which have been widely used as a model of intestinal epithelium.\(^{46,49,51–57,59,60}\)

We developed a model of predicting Caco-2 cell permeability from 2-dimensional topological descriptors, which was optimized by a genetic algorithm-combined partial least squares method.\(^{53}\) The data set analyzed included Caco-2 cell permeability for 73 structurally diverse compounds. The final partial least squares model consisting of 12 principal components of 29 topological descriptors gave a correlation coefficient (r) of 0.886 for the entire dataset and a predictive correlation coefficient (r\(_{pred}\)) of 0.825. The predictability of this model was comparable to that of the artificial neural network model where the same permeability data were predicted from quantum chemical descriptors.\(^{59}\) Prediction models based on topological descriptors are useful for screening of a large number of compounds, since they can be derived easily and fast from their two-dimensional structure.

It has been pointed out that considerable inter- and intra-laboratory variability exists in Caco-2 cell permeability measurements.\(^{41,42}\) Such variability makes it difficult to develop a QSAR model applicable to a wide range of the compounds. We found that each two of the data sets reported by various investigators were highly correlated.\(^{50}\) Based on the finding, they proposed the ‘latent membrane permeability’ concept and concomitant computational algorithm for simultaneously analyzing the Caco-2 permeability data from different sources.\(^{40}\) This concept is based on an assumption that all Caco-2 permeability data sets share a hidden, common relationship between the membrane permeability and physicochemical properties of the compounds. When five data sets from different sources were analyzed, there was no statistical difference in predictability between the absence and the presence of the hidden, common relationship. Thus, we successfully extracted the essence of information on membrane permeability from inter-laboratorially variable Caco-2 cell permeability measurements.

**Prediction of blood-brain barrier permeability by QSAR:**

Permeability across blood-brain barriers is a crucial factor determining effectiveness of central nervous system (CNS) drugs or CNS side effects of various drugs. Abraham et al.\(^{63}\) related the distribution of solutes between blood and brain (log BB) to the excess molar refraction, dipolarity/polarizability, hydrogen-bond acidity and basicity, and molecular volume (r = 0.952, s.d. = 0.197, n = 57, R = 99.2). Solute descriptors were obtained through fragment schemes. They found that solute size leads to an increase in log BB whereas solute dipolarity/polarizability, hydrogen-bond acidity, hydrogen-bond basicity all leads to a decrease in log BB. Lombardo et al.\(^{66}\) related log BB to calculated solvation free energy in water through molecular orbital calculation (r = 0.82, s.d. = 0.41, n = 55, F = 108.3). On the other hand, empirical approaches to predicting blood-brain barrier (BBB) permeability with Molsurf quantum chemical descriptors\(^{59}\) and topological descriptors\(^{60}\) have also been investigated, demonstrating a reasonable accuracy of prediction.

The models of discriminating BBB-permeable drugs from BBB-impermeable compounds have also been proposed. Van de Waterbeemd\(^{67}\) investigated the influence of physicochemical properties including lipophilicity, hydrogen-bonding capacity and molecular size and shape descriptors on BBB permeability. They concluded that compounds with molecular weight of less than 450 Da and polar surface area of less than 100 Å\(^2\) are BBB-permeable. In addition, various discrimination models regarding BBB permeability have been proposed with principal component analysis,\(^{61}\) Bayesian neural network,\(^{62}\) and support vector machine.\(^{63}\)

**Prediction of active transport processes using comparative molecular field analysis:** Comparative molecular field analysis (CoMFA) is a representative 3-dimensional QSAR approach, which has been developed by Cramer et al.\(^{69}\) CoMFA explains the gradual changes in observed biological properties by evaluating the electrostatic (coulombic interactions) and
steric (van der Waals interactions) fields at regularly spaced grid points surrounding a set of mutually aligned ligands. The steric and electrostatic interactions are calculated by placing a probe atom at grid points, and then tabulated for each molecule in the series. The resulting matrix is analyzed with partial least squares regression, yielding an equation that relates the CoMFA field values to the activity. One of typical features of CoMFA is to give contour maps that reflect the physicochemical environment around the structure of active compounds and are able to differentiate structures with high affinity from structures with low affinity. Swaan et al. examined the structure-affinity relationship for the small intestinal oligopeptide carrier (PepT1) using comparative molecular field analysis (CoMFA). The data set included 10 peptidomimetics; cephalosporins, penicillins, and ACE inhibitors. The model obtained showed a high correlation between the carrier permeability and the steric (76.3% contribution) and electrostatic (23.7% contribution) molecular fields with a cross-validated $R^2$ of 0.754. Besides, CoMFA has been applied to modelling the intestinal bile acid carrier, and P-glycoprotein. CoMFA has also been used for modelling inhibitors of the following enzymes: CYP1A2, CYP2A5, CYP2A6, CYP2C9, CYP2D6, rat MAO A and B. CoMFA is to give contour maps that reflect the physicochemical environment around the structure of active compounds and are able to differentiate structures with high affinity from structures with low affinity. 

Prediction of oral bioavailability using stochastic neural network: Generalized regression neural network (GRNN) is a nonparametric estimator that calculates a weighted average of the target values of training patterns by the probability density function. Gaussian kernel function is used for estimating the probability density function. The greatest advantage of GRNN is that interpretation of the output is easy because the output is probabilistic. Niwa investigated feasibility of GRNN to predict human intestinal absorption (HIA) of compounds from their two-dimensional topological descriptors. The root-mean square errors were 6.5, 27.7, and 22.8%HIA units for the training set, the test set, and the external prediction set, respectively. The predictive power was a little worse than that of a back-propagation artificial neural network model developed by Wessel et al. Niwa et al. used only 2-dimensional descriptors that were much simpler than the descriptors of Wessel et al. In addition, one of the advantageous features is that GRNN trains 100,000 times faster than back-propagation neural network. 

Prediction of metabolic stability using k-nearest neighbor method: Prediction by k-nearest neighbor method is done by calculating the weighted average property of k-nearest compounds in a database. Shen et al. analyzed metabolic turnover rate for compounds in human S9 homogenate for 631 diverse chemicals proprietary to Glaxo Smith Kline using this method. Molecular descriptors for molecular similarity search were selected from topological molecular descriptors such as molecular connectivity indices or atom pairs by simulated annealing-based variable selection optimization. The training set models were characterized by reasonable accuracy with leave-one-out cross-validated $R^2$ values ranging between 0.5 and 0.6. They also found that the test set compounds were correctly classified as stable or unstable in S9 assay with an accuracy of more than 85%.

Integrated Pharmacokinetic Models

The ultimate goal is to predict pharmacokinetics of compounds in the whole body by assembling all kinetic processes in one global model. Physiologically based pharmacokinetic modelling have been developed and used in the last 30 years to understand drug disposition behaviors systematically with physiological and biochemical parameters. The physiologically based nature of the modelling approach allows us to address mechanistic aspects of the pharmacokinetics as well as the concentration-time profiles. Application of physiologically based pharmacokinetic modelling for new chemical entities remained very limited, primarily due to the need for resources demanding input parameters to characterize ADME properties. However, progress of in vitro or in silico based technologies corresponding to each ADME process would facilitate physiologically based pharmacokinetic modelling in early stages of drug discovery and development.

Prediction of oral absorption by physiologically based approaches: Absorption of drugs from the gastrointestinal tract is complex and can be influenced by many factors. The factors would fundamentally be classified into three categories; i.e., physicochemical factors (pKa, solubility, stability, diffusivity, lipophilicity, and salt forms), physiological factors (gastrointestinal pH, gastric emptying, small and large bowel transit times, active transport and efflux, and gut wall metabolism), and formulation factors (drug particle size and crystal form, and dosage forms such as solution, tablet, capsule, suspension, emulsion, gel, and modified release). In recent years, despite such complexity, mass balance approaches have been developed to estimate oral drug absorption. Yu et al. collected more than 400 small intestinal transit time data and demonstrated that the transit flow profile can be well characterized by a seven-compartment transit model as well as a dispersion model. The seven compartments may correspond to the first half of the first compartment representing the duodenum, the second half of the first compartment, along with the second and third representing the jejunum, and the rest representing the ileum. The corresponding transit times in the duodenum, jejunum, and ileum are 14, 71, and 114 min. Based on such transit time distribution for the
compartmental absorption and transit (CAT) model to simulate the rate and extent of drug absorption. The CAT model is described by a set of differential equations that considers simultaneous movement of a drug in solution through the gastrointestinal tract and absorption of the dissolved material from each compartment into the portal vein. When the rate constant for absorption from each compartment was based on measured human intestinal permeability, the fraction absorbed was well estimated by the CAT model for various drugs.89

GastroPlus332 (Simulations Plus, Lancaster, CA)71 and IDEA88 (NaviCyte, Inc., San Diego, CA)89 are simulation software products based on advanced CAT models, in which physicochemical concepts, such as solubility and lipophilicity, are more readily incorporated than physiological aspects involving transporters and metabolism. Parrott and Lave59 have assessed IDEA ed than physiological aspects involving transporters and solubility and lipophilicity, are more readily incorporat-

models, in which physicochemical concepts, such as

inconvenient handling of multiple compound batches.89) is currently restricted by the limited functionality and widespread use with minimal training, but its usefulness is currently restricted by the limited functionality and inconvenient handling of multiple compound batches.89

Applications of physiologically-based intestinal absorption models to the analysis of oral drug absorption have been done by many investigators. Dressman et al.90 developed a two-tank perfect-mixing tank model to stimulate the effect of drug parameters (pK alpha, solubility, and intrinsic wall permeability) and system parameters (pH profile, volume of intestinal contents, and intestinal flow rate) on intestinal drug absorption. When absorption of chlorothiazide was simulated, good agreement between simulated and experimental data was observed, being led to the conclusion that the physical characteristics of chlorothiazide, rather than a saturable transport mechanism at the intestinal wall, may be responsible for the dose-dependent absorption observed for this drug. Sawamoto et al.91 developed a gastrointestinal absorption and transit model and successfully predicted plasma concentration profiles of several drug in rats from the gastrointestinal transit profile of a non-absorbed dye and the absorption rate constants of the drugs in different parts of the gut. Kalampokis et al.92 proposed a biased random walk model in the heterogeneous tube model having probabilistic dissolution and absorption processes. They concluded that their approach can be used to predict the fraction absorbed for drugs with various solubility and permeability characteristics provided that probability factors for dissolution and absorption are available. Ito et al.93 have developed a pharmacokinetic model for drug absorption that includes metabolism by CYP3A4 inside the epithelial cells, P-gp-mediated efflux into the lumen, intracellular diffusion from the luminal side to the basal side, and subsequent permeation through the basal membrane. They demonstrated that the fraction absorbed was synergistically elevated by simultaneous inhibition of both CYP3A4 and P-gp.

On the other hand, Cong et al.94 proposed a physiologically based, segregated-flow model to explain administration route-dependent intestinal metabolism. In contrast to the traditional physiological model where the intestine is subdivided into the vascular (intestinal blood), cellular (tissue), and luminal subcompartments, the model of Cong et al. further recognizes the subtle demarcation of tissue layers and distributions in blood supply. The segregated-flow model describes the majority of the intestinal blood flow to the nonabsorptive and nonmetabolizing serosal and submucosal regions, and only partial flow to the absorptive and metabolizing, enterocyte region at the villus tips of the mucosa95 where the metabolic enzymes and the P-glycoprotein reside. This model explained well that little intestinal metabolism occurs after systemic dosing but notable metabolism exists after oral dosing.96–99

Prediction of hepatic metabolism: Kinetic parameters for drug metabolism such as Vmax, Km, and intrinsic clearance (CLint) are obtained from in vitro experiments using liver microsomes and hepatocytes. These parameters should be used for predicting the in vivo situation. While empirical and allometric approaches for scaling in vitro metabolic data to in vivo have been developed100,101 physiologically based modelling approaches have been conventionally used.102,103 Houston102 compared the intrinsic clearance obtained from in vitro experiments using rat liver microsomes and isolated rat hepatocytes with that obtained in vivo for 25 drugs. Although there was a tendency to underestimate in vivo clearance when using liver microsome data, predictability from in vitro data was fairly reasonable. In particular, isolated hepatocyte data may be scaled to provide a good prediction of in vivo clearance ranging over four orders of magnitude. Iwatsubo et al.103 successfully predicted in vivo CLint in humans from in vitro CLint estimated in human liver microsomes or hepatocytes, except for several compounds. Raising several reasons for the deviations, they suggested that it is desirable to correct interindividual variability of in vitro preparations by using a scaling factor estimated from the metabolism of typical substrates. Iwatsubo et al.103 also indicated that predicted
values of hepatic extraction ratio and clearance vary depending on mathematical models especially in high-clearance drugs, suggesting the use of dispersion or distribution model for the prediction of in vivo clearance from in vitro data. Ito et al. proposed a method to predict pharmacokinetic alterations caused by drug-drug interactions that is based on in vitro metabolic inhibition studies using human liver microsomes or human enzyme expression systems.

**Mechanistic models of tissue distribution:** Poulin et al. developed tissue composition-based equations for calculating tissue-plasma partition coefficients ($P_{t:p}$):

$$P_{t:p} = \frac{P_{o:w} (V_{ss} + 0.3V_{plp}) + (V_{ss} + 0.7V_{ph})}{P_{o:w} (V_{plp} + 0.3V_{ph}) + (V_{plp} + 0.7V_{ph})} \times \frac{f_{iu}}{f_{iu}}$$

$$P_{t:p} = \frac{D_{o:w} (V_{ss} + 0.3V_{plp}) + (V_{ss} + 0.7V_{ph})}{D_{o:w} (V_{plp} + 0.3V_{ph}) + (V_{plp} + 0.7V_{ph})} \times \frac{f_{iu}}{f_{iu}}$$

where $P_{o:w}$ is the n-octanol:buffer partition coefficient of the non-ionized species at pH 7.4, $D_{o:w}$ is the olive oil:buffer partition coefficient of both the nonionized and ionized species at pH 7.4, $V$ is the fractional tissue volume content of neutral lipids ($n$), phospholipids ($p$), and water ($w$), $f_i$ is tissue, $p$ is plasma, and $f_{iu}$ is unbound fraction. These equations are based on assumption that each tissue and plasma is a mixture of lipids, water, and plasma proteins in which the drug can be homogeneously distributed. The first term of these equations is based on the drug lipophilicity in accordance with the lipophilicity-hydrophilicity balance of tissues and plasma due to their lipid and water contents, while the second term of the equations considers the binding to common proteins present in plasma and tissue interstitial space. The steady-state distribution volume ($V_{ss}$) is estimated to be the plasma volume in addition to the sum of each tissue: the $P_{t:p}$ multiplied with its corresponding tissue volume. Poulin and Theil predicted rat and human $V_{ss}$ for 123 structurally unrelated compounds by using the tissue composition-based equations. They found that 80% of all predicted values were within a factor of two of the corresponding experimental values. The $V_{ss}$ for some cationic amphiphilic bases were underestimated, suggesting additional relevant processes (e.g., ionic interactions with charged lipids of cell membranes and active transport processes). Since determination of $V_{ss}$ in vivo by using $P_{t:p}$ data is a time consuming process, this approach would be very useful. In their subsequent article, Poulin and Theil illustrated physiologically based pharmacokinetic models to estimate a priori the overall plasma and tissue kinetic behaviors under in vivo conditions. Tissue distribution of two lipophilic bases (diazepam, propranolol) and one neutral more hydrophilic drug (ethoxybenzamide) were predicted by this approach. The results indicated that most of the simulated concentration-time profiles of plasma and 10 tissues were in reasonable agreement with the corresponding experimental data determined in vivo (less than a factor of two).

**Conclusive Remarks**

The first generation of predictive ADME models are now commercially available and other models are published and implemented. These tools should allow chemists and drug-metabolism scientists to concentrate on compounds with the highest chances of meeting the required ADME criteria, and should contribute to a reduction in late-stage compound attrition. At this moment, unfortunately, such data-based models are rather limited for use. These models should be continuously refined through iterative learning/modelling. On the other hand, descriptors that are being used are not necessarily easily understood by the chemists who should translate into better molecular structures. Therefore, a new generation of mechanism-based models is required to provide better understanding and better predictability of ADME properties. As for physiological based pharmacokinetic modelling, there is still much to be done. The difficulty in modelling pharmacokinetic systems does not appear to be the application of the modelling techniques, but to do it successfully with limited data. The models to date are oversimplified due to missing information on pharmacokinetics. In the field of systems biology, dynamic cell simulations (e.g., comprehensive metabolic pathway, signaling pathway, cell physiology) are being extensively studied. Multi-scale modelling in several organs is also being investigated: it aims to understand the physiological function of intact organs in terms of the properties and behavior of the cells and tissues within the organ. Physiologically based pharmacokinetic models available only concerns the transport and metabolism of drugs, but not dynamics of the body after administration of the drugs or their formulations. The study on in silico pharmacokinetics has a long way to go to reach its goal.

**Acknowledgement:** This study was supported in part by a Grant-in-Aid for Scientific Research and a Leading Project for Biosimulation from the Ministry of Education, Culture, Sports, Science and Technology, Japan.

**References**

334  Fumiyoshi YAMASHITA and Mitsuru HASHIDA

57 Norinder, U., Osterberg, T. and Artursson, P.: Theoretical calculation and prediction of Caco-2 cell permeability using MolSurf parameterization and PLS
336 Fumiyoshi YAMASHITA and Mitsuru HASHIDA


113) Uegachi-Tanaka, M., Fujisawa, Y., Kobayashi, M.,


