PH.D. THESES

PREDICTION OF PHARMACOKINETIC PARAMETERS BY COMPUTATIONAL METHODS

by László Molnár

Supervisor: Dr. György Miklós Keserű
Gedeon Richter Ltd., Dept. of Computer Assisted Drug Discovery
BME, Dept. of Chemical Information Technology

Gedeon Richter Ltd.

CADD Dept.

2003

PREFACE

In November 2001 Tufts Center for the Study of Drug Development published the results of a research entitled "How New Drugs Move through the Development and Approval Process". The announcement stated that the average cost to develop a new prescription drug increased from $231 million in 1987 (study published by Tufts in 1991) to $802 million in 2000, a 250% increase, adjusted for inflation, over 11 years. The calculated cost contains the overall costs of failed lead compounds and opportunity cost (the amount of money that could be earned by a comparable alternative investment). Concerning the results of Tufts the development and approval of a new pharmaceutical takes between 10 and 15 years. The protection of industrial properties takes 20 years, however, the patent application usually takes place at the end of preclinical R&D activities, well before placing it on the market.

The huge R&D costs and relatively short protection of industrial property force pharmaceutical companies to reduce their R&D costs and time as much as possible. To achieve this goal one of the most common principles of the modern drug discovery is the "fail fast" approach, which means that unviable leads has to fail as early as possible in order to save associated costs and resources.

Filtering algorithms, which are able to predict pharmacokinetic (PK) properties of virtual libraries can save significant amount of money and time related to the synthesis of molecules with undesired PK profile.

AIMS

In our work we envisaged to develop of new *in silico* approaches to predict PK parameters in the distribution related and the metabolic field of ADME (Absorption, Distribution, Metabolism, Excretion).

The distribution of compounds in the human body is regulated by several bio-membranes. Regarding drugs act on the central nervous system (CNS), blood-brain barrier (BBB) is clearly the most important membrane system of this type. The penetration through the BBB is a prerequisite for CNS activity but on the other hand, BBB has to keep undesirable substances away from the brain. We have developed two

approaches to predict the BBB penetrating ability of molecules. One of them is a qualitative method which is able to do fast discrimination between penetrating and non-penetrating structures. The other one is a quantitative; thermodynamics based method to predict logBB (the logarithm of the ratio of concentrations measured in the brain and blood).

Cytochrome P450 3A4 is one of the major polymorphic izoenzyme being responsible for the metabolism of almost 50% of known drugs in human. Inhibitors of this izoenzyme might cause drug-drug interactions because of decreasing the clearance of other drugs metabolized by 3A4. Early identification of potential 3A4 inhibitors is therefore needed to minimize the risk of clinically relevant interactions. This prompted us to develop a fast *in silico* method to predict 3A4 inhibitory activity.

In the lead optimization phase the *in vitro* metabolic profile can be desirable or undesirable for the further optimization. If it is desirable, the goal is to find metabolically similar structures. If the metabolic profile is undesirable, the main challenge is to find metabolically different compounds. Neither goals are achievable with classical structural descriptors because they do not contain metabolic information. In the last part of our work we present a new metabolic descriptor whereby it is possible to discriminate between molecules with different metabolic behavior.

METHODS

The molecular mechanics and descriptor computation was carried out by Tripos' SYBYL molecular modeling suite. To create and train of artificial neural networks, I have used Stuttgart Neural Network Simulator (SNNS). The potential metabolic pathways of molecules were predicted by MetabolExpert. I have used the program package STATISTICA for the statistical analysis of results. I have written all computational and helper program source code in C, Perl, Tcl/Tk or SPL, respectively under SGI IRIX or Linux operating systems.

THESES

1. A virtual high throughput screening test for the identification of potentially CNS active drugs has been developed. Discrimination was based on the knowledge available in databases containing CNS active (Cipsline from Prous Science) and inactive compounds (Chemical Directory from Sigma Aldrich). Molecular structures were represented using 2D Unity fingerprints and a feedforward neural network was trained to classify molecules regarding their CNS activity. The neural net recognizes at least 89 % of CNS active compounds, which suggests using this methodology in our virtual screening protocol.

2. An *in silico* screening tool for potentially CNS active compounds has been developed based on the correlation of solvation free energies and blood-brain partitioning (logBB) data available from experimental sources. Utilizing a thermodynamic approach solvation free energies were calculated by the fast and efficient GB/SA continuum solvation model which enabled us to evaluate more than 10 compounds/min.

3. A virtual high throughput screening test for the identification of potential CP450 3A4 inhibitors has been developed. Molecular structures of inhibitors and non-inhibitors available in the Genetest database were represented using 2D Unity fingerprints and a feedforward neural network was trained to classify molecules regarding their inhibitory activity. Validation tests revealed that our neural net recognizes at least 89 % of 3A4 inhibitors and suggest using this methodology in our virtual screening protocol

4. METAPRINT, a metabolic fingerprint, has been developed by predicting metabolic pathways and corresponding potential metabolites. Calculated drug-likeness parameters (log P and MW) were incorporated into METAPRINT to allow the encoding of metabolic diversity within a chemical library. The application of METAPRINT in the design of cassette dosing experiments was

demonstrated using a library of α-1a antagonists synthesized at Glaxo Wellcome. Results obtained by Ward's clustering algorithm suggest that METAPRINTs are able to discriminate between low- and high-clearance compounds. Cassette design was performed by maximizing the intra-cassette Euclidean distances between compounds in METAPRINT space, using Monte Carlo simulated annealing approach. Calculated distances in METAPRINT space were in accordance with experimental data.

SUMMARY

The importance of *in silico* virtual screening has been significantly increased in the past few years of drug discovery. These virtual approaches have several advantages: the lack of any material (reagents, cells, etc.) reduces the overall costs, the increased speed comparing with the traditional *in vitro* techniques accelerates the preclinical R&D activities. The ADME related *in silico* methods give the opportunity to investigate ADME parameters in an early phase of drug discovery, since *in vivo* ADME tests are normally time and resource consuming and positioned generally at the final part of preclinical research. In the present work I have demonstrated four ADME related *in silico* methods, which were developed to solve real problems in real research projects at our environment. These approaches were implemented in Tripos' SYBYL program suite to in-house use and available for the drug discovery staff at Gedeon Richter Ltd.

Theses were published in the following scientific papers:

1. G. M. Keserű, L. Molnár, I. Greiner, A Neural Network Based Virtual High Throughput Screening Test for the Prediction of CNS Activity, *Comb. Chem. HTS* **3**: 535-540 (2000)

2. G. M. Keserű, L. Molnár, High-Throughput Prediction of Blood-Brain Partitioning: A Thermodynamic Approach, *J. Chem. Inf. Comput. Sci.* **41**: 120-128 (2000)

3. L. Molnár, G. M. Keserű, A Neural Network Based Virtual Screening of Cytochrome P450 3A4 Inhibitors, *Bioorg. Med. Chem. Lett.* **12**: 419-421 (2002)

4. G. M. Keserű, L. Molnár, METAPRINT: A Metabolic Fingerprint. Application to Cassette Design for High-Throughput ADME Screening, *J. Chem. Inf. Comput. Sci.* **42**: 437-444 (2002)

Scientific paper related but not used in the present work:

1. M. Lobell, L. Molnár, G. M. Keserű, Recent advances in the prediction of blood-brain partitioning from molecular structure *J. Pharm. Sci.* **92**: 360-370 (2003)

Oral presentation partially related to the present work:

1. Molnár L., Vágó I., A quasi platform independent chemical and biological relational database (Egy kvázi platformfüggetlen kémiai és biológiai web-alapú relációs adatbázis), Gyógyszerkémiai és Gyógyszertechnológiai Szimpózium 2001, 24-25 of September 2001, Visegrád (in hungarian)

2. L. Molnár, G. M. Keserű, Web based medicinal chemistry solutions at Gedeon Richter Ltd., TRIPOS seminar, 19 of March 2002, KKK, Budapest