Drug Metabolism and Pharmacokinetics 30 (2015) 347–351

Contents lists available at ScienceDirect

Drug Metabolism and Pharmacokinetics

journal homepage: http://www.journals.elsevier.com/drug-metabolism-and-pharmacokinetics

Regular article

Computational classification models for predicting the interaction of compounds with hepatic organic ion importers

Hwan You [a], Kyungro Lee [a], Sangwon Lee [a], Sung Bo Hwang [a], Kwang-Yon Kim [b], Kwang-Hwi Cho [c], Kyoung Tai No [a, b, *]

[a] Department of Biotechnology, Yonsei University, Seoul, Republic of Korea
[b] Bioinformatics & Molecular Design Research Center, Seoul, Republic of Korea
[c] Department of Bioinformatics and CAMDRC, Soong Sil University, Republic of Korea

ARTICLE INFO

Article history:
Received 20 March 2015
Received in revised form
22 May 2015
Accepted 17 June 2015
Available online 26 June 2015

Keywords:
Importers
Drug–drug interaction (DDI)
Classification
Hepatocyte

ABSTRACT

Hepatic transporters, a major determinant of pharmacokinetics, have been used to profile drug properties like efficacy. Among hepatic transporters, importers alter the concentration of the drug by facilitating the transport of a drug into a cell. Despite vast pharmacokinetic studies, the interacting mechanisms of the importers with its substrates or inhibitors are not well understood. Hence, we developed compound binary classification models of whether a compound is binder or nonbinder to a hepatic transporter with experimental data of 284 compounds for four representative hepatic importers, OATP1B1, OATP1B3, OAT2, and OCT1. Support Vector Machine (SVM) along with Genetic Algorithm (GA) was used to construct the classification models of binder versus nonbinder for each target importer. To construct the models, we prepared two data sets, a training data set from Fujitsu database (284 compounds) and an external validation data set from ChEMBL database (1738 compounds). Since an experimental classification criterion between binder and nonbinder has some ambiguity, there is an intrinsic limitation to expect high predictability of the binary classification models developed with the experimental data. The predictability of the classification models calculated with external validation sets were obtained as 77.72%, 84.31%, 84.21%, and 76.38 for OATP1B1, OATP1B3, OAT2, and OCT1, respectively.

1. Introduction

Membrane transporters enable the translocation of chemicals into and out of cells with active and passive transporting mechanisms. Some drugs and endogenous compounds with poor membrane permeability utilize the transporters for efficient translocation [1]. The transporters play an important role in drug metabolism and pharmacokinetics in the liver. Moreover, drug–-drug or drug–food interactions caused by the perturbation of the transporters function are relevant to pathogenesis factors. Inhibition of transporting activity by certain drugs or food ingredients may raise pharmacokinetic problems with coexisting drug [2–4].

There were many cases where an increase in statin plasma concentration has been observed following the co-administration of cyclosporine in clinical studies. Cyclosporine inhibits the OATPs (Organic Anion Transporting Polypeptide) which are related with statin uptake. This is an example of transporter-mediated DDI (Drug–Drug Interaction). DDI in hepatocyte affects the PK (Pharmacokinetics) and PD (Pharmacodynamics) of multiple drugs. They also have potential to affect the efficacy and toxicity of drugs [5]. Theses DDIs are mediated by various transporters in hepatocyte as follows.

The transporters are composed of several influx transporters including the Solute Carrier (SLC) superfamily and some efflux transporters including the ATP-binding cassette (ABC) superfamily. The transporters locate themselves in hepatocytes, intestinal epithelia, kidney proximal tubules and brain capillary endothelial cells [1]. In hepatocytes in particular, the importers in the sinusoidal membrane are Na^+-Taurocholate Cotransporting Polypeptide (NTCP), Organic Anion Transporting Polypeptide (OATP) family (OATP1B1, OATP1B3 and OATP2B1), Organic Anion Transporter (OAT) family (OAT2 and OAT7), and Organic Cation Transporter 1(OCT1) [1]. Among them, OATP1B1, OATP1B3, OAT2, and OCT1 are dominant in hepatocytes [6].

Several researchers have studied those four transporters. Meier-Abt et al. suggested that the positively charged binding pocket in

* Corresponding author. Department of Biotechnology, Yonsei University, Seoul, Republic of Korea. Tel.: +82 2 393 9551; fax: +82 2 393 9554.
E-mail address: ktno@yonsei.ac.kr (K.T. No).

http://dx.doi.org/10.1016/j.dmpk.2015.06.004
1347-4367/

OATP1B3 provides selectivity of substrates [7]. Mutational study of three amino acids (R57, K361, and R580) in the OATP1B1 binding pocket supports the hypothesis about the role of positively charged amino acids [8]. Chang et al. tried a meta-pharmacophore approach using limited data composed of diverse or unequal sources and conditions [9]. The meta-model for OATP1B1 produced a hydrophobic feature located at the center and hydrogen bond at the edge.

CoMFA (Comparative Molecular Field Analysis) in rat model revealed that the substrate selectivity of OATs was influenced by hydrophobicity and electrostatic interaction [10]. Ahlin et al. suggested that features of transport activity in OCT1 may be related with hydrophobicity, low molecular mass, and positive charge and that the properties of inhibitors may include net positive charge and high lipophilicity [11]. These studies explained the property of each transporters' binding site well yet a systematic approach is essential to precisely predict DDIs in hepatocyte.

In this study, we developed the physicochemical property based classification models for four hepatic ion importers, i.e., OATP1B1, OATP1B3, OAT2, and OCT1. Through the combination of classification results from four hepatic ion importer models, we could make an informed decision for DDI in hepatocyte. The classification models were constructed using a support vector machine with physicochemical descriptors selected by genetic algorithm.

2. Methods

2.1. Preparing data set for the classification models development

Both substrates and inhibitors data set was built for each of OATP1B1, OATP1B3, OATP2B1, OAT2 and OCT1. The 284 compounds were taken from ADME database (http://jp.fujitsu.com/group/kyushu/en/services/admedatabase) [12] and 3D structures generated with MMFF forcefield. Since ADME database were constructed from literature, the data were generated with various cell lines and culture conditions. If the compounds for the data sets were only taken from the experiments with the same cell type, such as HEK293 cell, we cannot secure enough data set to build models. For this reason, regardless of cell line type, data was established by compounds with target importers originated from human genome. Despite the alleviation of the data selection conditions, for OATP2B1, only 12 compounds are selected for the training set, OATP2B1 model was not developed.

To test the predictability of the models, external validation sets were prepared with the compounds taken from ChEMBL DB (https://www.ebi.ac.uk/chembl) [13], 1804 compounds for OATP1B1, 1709 for OATP1B3, 26 for OAT2, and 245 for OCT1. In OATP1B1 and OATP1B3 data set, most of the compounds came from Tom De Bruyn et al. [14]. Tom De Bruyn et al. used CHO cell line and sodium fluorescein as substrate in inhibition experiment for OATP1B1 and OATP1B3. The inhibitory potential of all 2000 compounds from The Spectrum Collection library was tested at an equimolar substrate-inhibitor concentration of 10 mM. When the inhibition activity was higher than 50%, compounds were considered inhibitors and used as binder class for external validation. On the other hand, the compounds represent lower than 30% inhibition activity were deemed nonbinder class. The compounds with activity from 30 to 50% were excluded in the external validation set due to the ambiguity. 50(OATP1B1), 59(OATP1B3), 7(OAT2), and 46(OCT1) compounds were duplicated with compounds in ADME DB. Some compounds classified into nonbinder in ADME DB were changed to binder class and trained to build the model because they are reported for substrate or inhibitor in ChEMBL DB (3 compounds in OATP1B1, 4 in OATP1B3, 1 in OAT2, and 2 in OCT1, respectively). After preprocessing, 1360 compounds for OATP1B1, 1530 for OATP1B3, 19 for OAT2 and 199 for OCT1 were remained. They were also converted to 3D structure with MMFF forcefield.

For model stability, we chose the more balanced data set, ADME DB as training set. Even though the size of the data in ChEMBL DB was larger than ADME DB, ChEMBL DB was unbalanced in ratio between OATPs and the others.

The substrate and inhibitor composition of each transporter is summarized in Table 1(Training set) and Table 2(External validation set). The number of binders is not equal to the number of substrates plus inhibitors since some molecules can be both substrate and inhibitor by experiment design.

2.2. Classification model development

2.2.1. Definition of binder and nonbinder

The classification criteria between the substrate and inhibitor are hard to simplify into a binary model in that even a single importer shows various binding patterns with multiple binding sites. In addition, an inhibition event of hepatic importers, unlike that of other enzymes, is not well-defined as an experimental protocol. Therefore, a realistic approach classifies whether a compound binds to a target transporter or not. We defined both substrates and inhibitors of a specific transporter as “binders of the transporter”. As a complementary class of “binders of the transporter”, we introduced “nonbinders of the transporter” where the member compounds have no positive experimental data with a specific importer. For example, Compound A, which is inhibitor, substrate, weak inhibitor, and untested compound for OATP1B1, OATP1B3, OAT2, and OCT1, respectively, is defined as binder, binder, nonbinder, and nonbinder for each importer. It shows that about 90% of nonbinders were correctly assumed by curating overlapping compounds between ADME DB and ChEMBL DB.

2.2.2. Descriptor selection

In the present work, 102 molecular descriptors were generated with preADMET software [15] and 16 molecular descriptors were calculated using the Solvation Free Energy Density (SFED) model [16]. Descriptors consist of constitutional (3 + 0), geometrical (26 + 1), electrostatic (65 + 12), and physicochemical (8 + 3) descriptors (the first value in the parenthesis is originated from preADMET and the second value from the SFED model in parentheses).

Descriptors were selected to build the model by two ways, manual selection and automatic selection. The first way of manually selecting 83 descriptors was based on the previous studies

Table 1
The composition details of binder (substrate and inhibitor) and nonbinder data set (ADME DB).

Transporter		OATP1B1		OATP1B3		OAT2		OCT1	
Binder	Substrate	131	66	81	59	41	24	101	33
	Inhibitor		94		45		30		94
Nonbinder		153		203		243		183	
Sum		284		284		284		284	

The number of binder and nonbinder reflect the change by ChEMBL DB.

Table 2
The composition details of binder (substrate and inhibitor) and nonbinder data set (ChEMBL DB).

Transporter		OATP1B1		OATP1B3		OAT2		OCT1	
Binder	Substrate	247	29	163	21	16	8	121	25
	Inhibitor		238		147		11		99
Nonbinder		1113		1367		3		78	
Sum		1360		1530		19		199	