Toxicology and Applied Pharmacology 269 (2013) 195–204

Contents lists available at SciVerse ScienceDirect

Toxicology and Applied Pharmacology

journal homepage: www.elsevier.com/locate/ytaap

In silico modeling to predict drug-induced phospholipidosis ☆

Sydney S. Choi, Jae S. Kim, Luis G. Valerio Jr. *, Nakissa Sadrieh

Science and Research Staff, Office of Pharmaceutical Science, Center for Drug Evaluation and Research, U.S. Food and Drug Administration, 10903 New Hampshire Avenue, Silver Spring, MD 20993-0002, USA

ARTICLE INFO

Article history:
Received 8 December 2012
Revised 1 March 2013
Accepted 3 March 2013
Available online 27 March 2013

Keywords:
Phospholipidosis
Drug safety
In silico toxicology

ABSTRACT

Drug-induced phospholipidosis (DIPL) is a preclinical finding during pharmaceutical drug development that has implications on the course of drug development and regulatory safety review. A principal characteristic of drugs inducing DIPL is known to be a cationic amphiphilic structure. This provides evidence for a structure-based explanation and opportunity to analyze properties and structures of drugs with the histopathologic findings for DIPL. In previous work from the FDA, *in silico* quantitative structure–activity relationship (QSAR) modeling using machine learning approaches has shown promise with a large dataset of drugs but included unconfirmed data as well. In this study, we report the construction and validation of a battery of complementary *in silico* QSAR models using the FDA's updated database on phospholipidosis, new algorithms and predictive technologies, and in particular, we address high performance with a high-confidence dataset. The results of our modeling for DIPL include rigorous external validation tests showing 80–81% concordance. Furthermore, the predictive performance characteristics include models with high sensitivity and specificity, in most cases above ≥80% leading to desired high negative and positive predictivity. These models are intended to be utilized for regulatory toxicology applied science needs in screening new drugs for DIPL.

Published by Elsevier Inc.

Introduction

Drug-induced phospholipidosis (DIPL) is recognized as a recurring preclinical finding in pharmaceutical drug development. The functional consequence and the mechanisms of DIPL are not clear and it poses a challenge for regulators and the pharmaceutical industry because of unknown clinical consequences. DIPL is characterized by accumulation of drugs and phospholipids in lysosomes. Histopathologically, DIPL manifests itself as foamy macrophages or cytoplasmic vacuoles in various tissues of both animals and humans. These findings show multilamellar inclusion bodies in electron micrographs. In addition, Cationic Amphiphilic Drugs (CADs), which contain a hydrophobic ring structure and a hydrophilic amine moiety are well-known structural features associated with induction of phospholipidosis (Fig. 1). As lipophilicity and the ionization state of a molecule affect the membrane permeability of the molecule, the physicochemical properties of CADs have drawn the attention of many researchers, as a possible mechanism underlying the manifestation of DIPL. *In vitro* screening models have been developed using fluorescence (Mesens et al., 2010), EC50 (Morelli et al., 2006), and gene expression methods (Nioi et al., 2007) to detect and predict DIPL. There are a few *in silico* DIPL prediction models including the Ploemen model (Ploemen et al., 2004), Tomizawa model (Tomizawa et al., 2006), Pelletier model (Pelletier et al., 2007), and Hanumegowda model (Hanumegowda et al., 2010). These models were based on the physicochemical properties of compounds using parameters which represent the lipophilicity (log P) or ionization state (pKa) of drugs.

Over the past several years, FDA has been carefully building a database of drugs containing more than 700 DIPL positive and negative drugs to investigate the structural similarities for discriminating between compounds that either induce or do not induce phospholipidosis. This work is intended to help identify compounds with the liability to induce phospholipidosis independent of whether the compound has a CAD structural moiety or not. In 2008, FDA had developed predictive quantitative structure–activity relationship (QSAR) models for DIPL using a preliminary database (Kruhlak et al., 2008). The models were built with various software using 190 DIPL positive compounds including 76 transmission electron microscopy (EM) confirmed and 114 unconfirmed ones and 393 DIPL negative compounds, which were assumed negative compounds taken from a marketed database. Recently, FDA reported DIPL QSAR models using an updated FDA DIPL database with 743 compounds (Orogo et al., 2012). The database contained 385 DIPL positives and 358 negatives. Many of the DIPL positive drugs were not confirmed to induce phospholipidosis using EM. However, in this study, we constructed and validated a new battery of predictive

☆ The content of this publication does not necessarily reflect the views or policies of the Department of Health and Human Services, nor does mention of trade names, commercial products, or organizations imply endorsement by the U.S. Government. This paper reflects the current thinking and experience of the authors.

* Corresponding author at: Science and Research Staff, Office of Pharmaceutical Science, Center for Drug Evaluation and Research, U.S. Food and Drug Administration, White Oak 51 Room 4128, 10903 New Hampshire Ave., Silver Spring, MD 20993-0002, USA. Fax: +1 301 796 9997.
E-mail address: luis.valerio@fda.hhs.gov (L.G. Valerio).

0041-008X/$ – see front matter. Published by Elsevier Inc.
http://dx.doi.org/10.1016/j.taap.2013.03.010

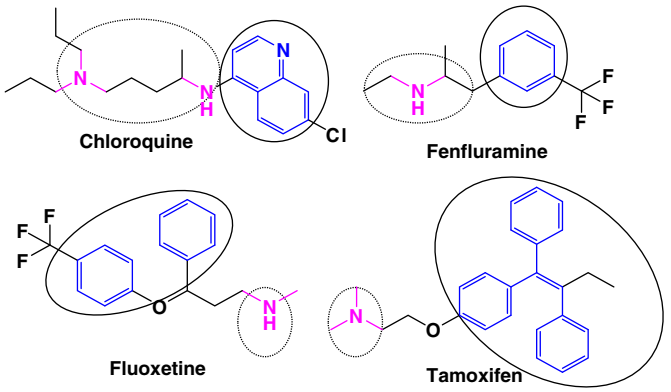

Fig. 1. Typical CAD compounds of DIPL, which contain hydrophobic ring structure (circle; blue) and hydrophilic amine moiety (dotted circle; magenta).

QSAR models using a higher quality data set derived from FDA's updated DIPL database. It is noteworthy that there are other high quality data recently published (Muehlbacher et al., 2012) and that fluorescence labeling of phospholipids is another method to monitor phospholipidosis (Mesens et al., 2009). The high quality data we obtained were from the FDA/CDER submissions in which the positive drugs were confirmed to induce phospholipidosis with EM as determined by the authors or sponsors of the studies, and the negative drugs did not show evidence of phospholipidosis induction throughout the submission file.

Materials and methods

Data source for computational modeling

Drugs that induce phospholipidosis have been identified as part of the FDA CDER's Working Group on Phospholipidosis (FDA PLWG) where drugs from various sources were collected including proprietary data derived from FDA/CDER archives as well as the public scientific literature, as previously described (Orogo et al., 2012). Briefly, these sources include information derived from Investigational New Drug and New Drug Application reviews at FDA that span a diversity of species and tissue types commonly found to be targets of DIPL (*e.g.*, liver, lung, kidney, *etc*). The histopathological manifestations of DIPL include foamy alveolar macrophages with cytoplasmic vacuolations which can then be confirmed with transmission electron microscopy (EM). The EM in this report is able to confirm the presence of multilamellar inclusion bodies, a hallmark of phospholipidosis.

The FDA DIPL database contains over 700 compounds, and all are organic molecules. No inorganics, polymers, mixtures or biologic drugs were included in this DIPL modeling study because the computational software programs we employed were not designed to analyze these types of substances. Of the more than 700 compounds in FDA DIPL database, 147 were EM-confirmed DIPL positive compounds and 232 were DIPL negative compounds, with no evidence of phospholipidosis in the data source analyzed. Thus, the total number of compounds we relied upon with high quality and modelable data from the FDA DIPL database was 379 (147 positives and 232 negatives). We utilized these 379 compounds to build a model we refer to herein as DIPL Model 1. By using the entire 379 compound data sets to build the DIPL Model 1, we could not test the predictive performance of the model using an external validation data set, since all the high quality data available to us are used in building this model. The practice of using all available compounds to construct a training set for a model is common in modeling as it has the advantage of increasing the applicability domain (*i.e.* prediction space) of a newly constructed model. However, the disadvantage of using all compounds to construct the model is that external validation is not possible, unless enough additional compounds become available to achieve sufficient statistical power. External validation is recognized as the most rigorous and true test of the predictive performance of a computational toxicology model (Valerio, 2009), and there are other valuable methods including bootstrapping (Svetnik et al., 2003). Therefore, to accomplish the goal of performing an external validation test and best learn about the performance of the majority of our data, we took steps to build another model with a smaller training set compared to DIPL Model 1. We refer herein to this smaller model as DIPL Model 2 (Fig. 2). DIPL Model 2 contains a balanced training set of 198 compounds composed of 98 positives and 100 negatives randomly chosen from the data set of DIPL Model 1 (Table 1 and Fig. 2). By using a smaller training set for DIPL Model 1, we were able to leave a hold-out set of 49 positives and 50 negatives (a total of 99 compounds) that were used for balanced external validation testing (Table 2 and Fig. 2). Thus, all compounds in DIPL Model 1 and DIPL Model 2 derive from the original aforementioned 379 data set. Another important feature of the DIPL Model 2 is that the external validation test set of 99 compounds has sufficient statistical power to judge predictive performance. We note that both computational software tools described below use their own built-in algorithms to validate a model, however, the number of compounds that were withheld by the software's validation test set was smaller than our user-defined hold-out set of 99 compounds.

The molecular structures of drugs used in modeling were collected as universal simplified molecular input line entry system (SMILES) codes. SMILES were converted into a standard ASCII text file, Quantitative Molecular Data (QMD) file, for the ADMET Predictor™ software. For the Symmetry® program, the molecular structures were converted into structure data files (SDF) which enable the software to read the drug molecular structures for the *in silico* analysis. Upon importing a SDF, Symmetry® washes the ions and salts creating a computational form of the data set suitable for modeling. The resulting structures were used to construct and test the predictive DIPL QSAR models.

Computational platforms and algorithms

Symmetry®. An in-house installation of Symmetry® 1.0.3 (Prous Institute for Biomedical Research, Barcelona, Spain) was deployed as the computational platform to perform the QSAR analysis and predictions for DIPL in this study. The software was obtained through a FDA/CDER agency-approved Research Collaboration Agreement with the Prous Institute for Biomedical Research. Symmetry® employs a bi-functional system that includes a module for performing predictions of test set chemicals, and a module for building computational QSAR-based models. This computational platform has been previously tested for predictive accuracy and model building of other preclinical and clinical endpoints by FDA (Valerio et al., 2011; Valerio et al., 2012). For descriptor calculations, Symmetry® relies upon the Mold2 descriptor package version 2.0. Mold2 is freely available to the public from the FDA National Center for Toxicology Research (Hong et al., 2008). The Mold2 descriptor package is integrated with Symmetry® and is enabled during the model building process. Mold2 calculates a diverse set of 777 two-dimensional molecular descriptors related to chemical structure information. The list and definitions of available descriptors from Mold2 can be found by query of FDA's website (www.fda.gov). The DIPL model building settings employed with Symmetry® were a combined probabilistic and similarity algorithm involving logistic regression and molecular descriptor similarity. All available Mold2 descriptors were tested in both algorithms and descriptors were range normalized (min. to max.). In DIPL Model 1, for each training class and model, outliers were eliminated at >3 interpercentile range from lower and upper percentiles. Descriptors were eliminated which correlated 100%. No restriction was set on the number of descriptors selected during model building, and a genetic algorithm for descriptor selection was enabled (number of generations = 540, population = 2000, cross-over probability = 0.6, mutation probability = 0.03). Wrapper