Food and Chemical Toxicology 107 (2017) 150–166

Contents lists available at ScienceDirect

Food and Chemical Toxicology

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

Computational prediction of immune cell cytotoxicity

Anna K. Schrey [a], Janette Nickel-Seeber [a], Malgorzata N. Drwal [a, 1], Paula Zwicker [b], Nadin Schultze [b], Beate Haertel [b], Robert Preissner [a, *]

[a] Charité – University Medicine Berlin, Institute for Physiology and ECRC, Berlin, Germany
[b] University of Greifswald, Institute of Pharmacy, Greifswald, Germany

ARTICLE INFO

Article history:
Received 10 January 2017
Received in revised form
11 May 2017
Accepted 19 May 2017
Available online 27 May 2017

Keywords:
Toxicity prediction
In silico toxicology
Immune cells
Cytotoxicity
Molecular similarity

ABSTRACT

Immunotoxicity, defined as adverse effects of xenobiotics on the immune system, is gaining increasing attention in the approval process of industrial chemicals and drugs. In-vivo and ex-vivo experiments have been the gold standard in immunotoxicity assessment so far, so the development of in-vitro and in-silico alternatives is an important issue.

In this paper we describe a widely applicable, easy-to use computational approach which can serve as an initial immunotoxicity screen of new chemical entities. Molecular fingerprints describing chemical structure were used as parameters in a machine-learning approach based on the Naïve-Bayes learning algorithm.

The model was trained using blood-cell growth inhibition data from the NCI database and validated externally with several in-house and literature-derived data sets tested in cytotoxicity assays on different types on immune cells. Both cross-validations and external validations resulted in areas under the receiver operator curves (ROC/AUC) of 75% or higher.

The classification of the validation data sets occurred with excellent specificities and fair to excellent selectivities, depending on the data set. This means that the probability of actual immunotoxicity is very high for compounds classified as immunotoxic, while the fraction of false negative predictions might vary. Thus, in a multistep immunotoxicity screening scheme, the classification as immunotoxic can be accepted without additional confirmation, while compounds classified as not immunotoxic will have to be subjected to further investigation.

1. Introduction

Immunotoxicity, defined as the adverse effects of xenobiotics (immunotoxicants) on the immune system includes two main types: immunosuppression (decreased immunocompetence) and inappropriate immunostimulation. An altered immune function can have various effects, ranging from increased incidence/severity of infections or tumors to hypersensitivity/autoimmune reactions and allergies. The term “immunotoxicants” is used to refer to both natural and synthetic compounds of various origins and applications, such as food products, food additives, and environmental pollutants, which may be either pharmaceuticals or other chemicals. Identifying immunotoxicants is an important but difficult task. In particular, reducing the need for animal experiments wherever possible is a major consideration. It has been suggested that these issues can be solved by developing appropriate *in vitro* assays and *in silico* methods and combining them in a decision-tree-like manner (Combes et al., 2008; Macela et al., 1989).

Computer-based (*in silico*) toxicity prediction has been the topic of research for more than 20 years (Hansch et al., 1995; Muster et al., 2008). However, when it comes to immunotoxicity, most of the commercial software packages such as DEREK (Barratt and Langowski, 1999; Zinke et al., 2002) and TOPKAT (Enslein et al., 1990) and published *in silico* methods focus on skin or respiratory sensitization (Karol et al., 1996) or on the interaction with one specific protein (Turabekova et al., 2014; Yuan et al., 2013) Only the HazardExpert package (Smithing and Darvas, 1992), a knowledge rule-based expert system, includes an immunotoxicity endpoint.

One promising *in silico* approach to predicting general immunotoxicity for a wide range of chemical substances is based on the assumption and observation that similar molecules exhibit similar

* Corresponding author. Charité – University Medicine Berlin, Institute for Physiology and ECRC, Philippstr. 12, 10115 Berlin, Germany.
E-mail address: robert.preissner@charite.de (R. Preissner).
[1] Current address: University of Strasbourg, Faculty of Pharmacy, Illkirch, France.

http://dx.doi.org/10.1016/j.fct.2017.05.041
0278-6915/

biological effects. Various biological problems such as the prediction of targets (Campillos et al., 2008), therapeutic indications (Nickel et al., 2014) and side effects (Lounkine et al., 2012) could be successfully solved using similarity-based approaches.

In combination with machine-learning methods such as k-nearest neighbors, naïve Bayes models, support vector machines, random forests or ensembles of different classification methods, molecular similarity defined by the molecular structure and properties can be used to screen new chemical entities virtually, e.g. for various toxicological endpoints (Drwal et al., 2014; Gadaleta et al., 2014; Li et al., 2014; Liu et al., 2015). Drwal et al. ranked among the top submissions in the Tox21 challenge using a combination of three fingerprints representing the 2D molecular structure and a naïve Bayes predictor (Drwal et al., 2015).

Here we introduce a similar approach based on the combination of two molecular fingerprints describing structural features of the molecules in which we test the working hypothesis that immunotoxicity can be estimated from immune cell cytotoxicity.

The training sets were built from B- and T-cell growth inhibition data of ca 45,000 compounds taken from the National Cancer Institute's (NCI) data base. This is the largest publicly available dataset for cytotoxic effects of chemicals containing more than 265,000 compounds and data on growth inhibition in more than 60 human cell lines, including one B-cell and 2 T-cell lines.

Our models were evaluated using one in-house set containing chemical compounds with known immunosuppressive, immunostimulating, immunomodulatory or allergenic properties, and three data sets compiled from the literature. We demonstrate the applicability of the models in predicting the results of different experimental setups evaluating immunotoxicity.

2. Methods

2.1. Computational methods

2.1.1. Data sets

2.1.1.1. Training data. The U.S. National Cancer Institute's (NCI) public database was used to build predictive models for immune cell cytotoxicity. Growth inhibition data were downloaded from https://wiki.nci.nih.gov/display/NCIDTPdata/NCI-60+Growth+Inhibition+Data and linked to molecular structures downloaded from http://cactus.nci.nih.gov/download/nci/index.html. Growth inhibition was determined by the NCI using the sulforhodamine B cytotoxicity assay and is given in form of GI_{50} values, which describe the drug concentration resulting in a 50% reduction of the net protein increase as compared to a control experiment.

The NCI database was filtered to include only measurements given in the molar unit and, in cases where multiple concentration ranges had been tested, the lowest was kept. GI_{50} values from three different cell lines were investigated, the B-cell line RPMI-8226 and the two T-cell lines MOLT-4 and CCRF-CEM. Compounds with GI_{50} values below 10 μM were defined as toxic. In the case of T-cells, compounds were defined as active or inactive only if their GI_{50} values were below or above 10 μM in both cell lines. Compounds with inconclusive results from the two cell lines were discarded. As a third, "immunotoxic" property was defined for compounds showing either B-cell or T-cell activity according to the definition given above. This approach afforded a total of 47 565 compounds which were subjected to further filtering and preparation steps (see Section "Data Preparation").

2.1.1.2. External validation data. To determine the goodness of the predictive model, we used four data sets, one obtained in-house and three from the literature. Each compound data set was associated with activity data from two or three different assays. Two smaller datasets (in-house and Markovic) consisted of manually selected immunotoxic and inert compounds, while different subsets of compounds included in the Acutoxbase (Kinsner-Ovaskainen et al., 2009) were used for the two larger data sets (Sjöström and Kooijman).

2.1.1.3. In-house Jurkat, THP1, and PBMC data sets. The cytotoxic activity on immune cells was determined for 14 selected compounds (Table 1, Fig. S1) using a MTT cytotoxicity assay on Jurkat, THP-1, and peripheral blood mononuclear cells (see the Cytotoxicity Assays section for details).

2.1.1.4. Markovic PBMC and LCL data sets (Markovič et al., 2014). This data set contained cytotoxicity values from two different cell lines – peripheral blood mononuclear cells (PMBCs) and lymphoblastoid cell lines (LCLs) – and consisted of seven compounds.

2.1.1.5. Sjöström NRU and estimated LC_{50} data sets (Sjöström et al., 2008). The activity data included a 3T3 NRU cytotoxicity assay, which measures the 50% inhibition (IC_{50}) of dye-neutral red uptake in the lysosomes (90 compounds), and human 50% lethal concentrations (LC_{50}) estimated using time-related human sub-lethal and lethal blood concentrations from the Acutoxbase (80 compounds).

2.1.1.6. Kooijman TNFalpha, IFNgamma, and IL-5 data sets (Kooijman et al., 2010). Three different cytokine assays, that measure the (IC_{50}) of cytokine production by cultured human PBMCs stimulated with phytohaemagglutinin (PHA) (58 compounds).

2.1.2. Data preparation

Compound structures were standardized using JChem version 15.10.26.0 (Chemaxon) with the following settings: Water and salts were removed, explicit hydrogens were added, structures were aromatized and their three-dimensional coordinates were cleaned. Canonical smiles and InChIKeys were calculated using Discovery Studio 4.1 (BIOVIA).

In the training set, Duplicates were identified based on canonical smiles and eliminated.

This approach afforded 44 615 compounds in the "immunotoxic" training set and 41 883, or 37 198 compounds in the B-cell, and T-cell training set, respectively.

2.1.3. Calculation and concatenation of molecular fingerprints

Different fingerprints were calculated for each compound for use as a feature input for the predictive model and in determining the best fingerprint for cytotoxicity predictions on immune cells. Various molecular fingerprints (Morgan, FeatMorgan, RDKit fingerprint, Layered, AtomPair, Torsion, Avalon, MACCS keys) were calculated with KNIME analytics platform 2. 12. 0 (KNIME.com AG, Switzerland) using RDKit Nodes (http://www.rdkit.org) and ToxPrint fingerprints (Yang et al., 2015) were calculated using the ChemoTyper software (https://chemotyper.org/). Combinations of fingerprints were created by concatenating two RDKit fingerprints or concatenating the ToxPrint fingerprint with one RDKit fingerprint. All RDKit fingerprints with different parameters and all possible combinations of fingerprints with optimized parameters were built, resulting in bit lengths of 1024 or 2048 for the RDKit fingerprints, 729 for the ToxPrint fingerprint, and 1753 or 2777 for the combinations.

2.1.4. Model development and validation

For immunotoxicity prediction based on RDKit and ToxPrint fingerprints, models were built applying the Naïve Bayes algorithm (Nidhi et al., 2006) using the Fingerprint Naïve Bayes Nodes implemented in the KNIME Analytics Platform.