Chemosphere 210 (2018) 312–319

Contents lists available at ScienceDirect

Chemosphere

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

A combination of ternary classification models and reporter gene assays for the comprehensive thyroid hormone disruption profiles of 209 polychlorinated biphenyls

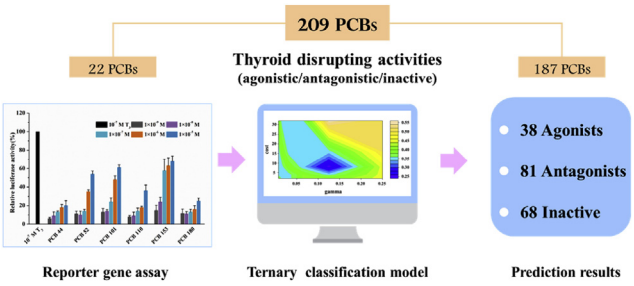
Xiaoxia Bai [a, 1], Lu Yan [b, 1], Chenyang Ji [b], Quan Zhang [b], Xiaowu Dong [c, **], An Chen [d], Meirong Zhao [b, *]

[a] Women's Hospital, School of Medicine, Zhejiang University, Hangzhou, Zhejiang, 310006, China
[b] Key Laboratory of Microbial Technology for Industrial Pollution Control of Zhejiang Province, College of Environment, Zhejiang University of Technology, Hangzhou, Zhejiang, 310032, China
[c] College of Pharmaceutical Sciences, Zhejiang University, Hangzhou, Zhejiang, 310058, China
[d] College of Information Engineering, China Jiliang University, Hangzhou, Zhejiang, 310018, China

HIGHLIGHTS

- The thyroid-disrupting mechanisms of 209 PCBs were determined.
- The reporter gene assay was performed to examine the hormone activities of 22 PCBs *via* TR.
- The remaining 187 PCB congeners' hormone activities were predicted using the SVM model.

GRAPHICAL ABSTRACT



ARTICLE INFO

Article history:
Received 30 March 2018
Received in revised form
24 June 2018
Accepted 5 July 2018
Available online 6 July 2018

Handling Editor: Jian-Ying Hu

Keywords:
Endocrine disrupting chemicals
Polychlorinated biphenyls
Thyroid receptor
Support vector machines

ABSTRACT

Computational toxicology is widely applied to screen tens and thousands of potential environmental endocrine disruptors (EDCs) for its great efficiency and rapid evaluation in recent years. Polychlorinated biphenyls (PCBs) with 209 congeners have not been comprehensively tested for their ability to interact with the thyroid receptor (TR), which is one of the most extensively studied targets related to the effects of EDCs. In this study, we aimed to determine the thyroid-disrupting mechanisms of 209 PCBs through the combination of a novel computational ternary classification model and luciferase reporter gene assay. The reporter gene assay was performed to examine the hormone activities of 22 PCBs *via* TR to classify their profiles as agonistic, antagonistic or inactive. Thus, six agonists, eleven antagonists and seven inactive chemicals against TR were identified in *in vitro* assays. Then, six relevant variables, including molecular structural descriptors and molecular docking scores, were selected for describing compounds. Machine learning methods (i.e., linear discriminant analysis (LDA) and support vector machines (SVM)) were used to build classification models. The optimal model was obtained by using SVM, with an accuracy of 88.24% in the training set, 80.0% in the test set and 75.0% in 10-fold cross-validation. The remaining 187 PCB congeners' hormone activities were predicted using the obtained models. Out of these PCBs, the SVM model predicted 38 agonists and 81 antagonists. The findings revealed that higher

* Corresponding author.
** Corresponding author.
E-mail addresses: dongxw@zju.edu.cn (X. Dong), zhaomr@zjut.edu.cn (M. Zhao).
[1] These authors contributed equally to this work.

https://doi.org/10.1016/j.chemosphere.2018.07.023

chlorinated PCBs tended to have TR-antagonistic activities, whereas lower chlorinated PCBs were agonists. This study provided a reference for further exploring PCBs' thyroid effect.

1. Introduction

Endocrine disrupting chemicals (EDCs) comprise a wide range of chemicals that are thought to cause adverse health effects and to pose an ecological threat even in trace amounts (Foster, 2001; Hayes et al., 2002; Sumpter and John, 2005). Because the number of EDCs has expanded rapidly, predictive models, computational toxicology and high-throughput *in vitro* screening assays will inherently provide fundamental information for the accurate assessment of endocrine disruption (Jacobs, 2004; Murk et al., 2013).

Due to its greater efficiency for rapid evaluation, computational toxicology, (Vuorinen et al., 2013) which fuses molecular toxicology and chemistry with computational science, has markedly expanded in recent years and provides advantageous alternative methods for evaluating environmental hazards, especially for virtual compounds that lack standards (Tice et al., 2013). Several institutions, including the US Environmental Protection Agency (EPA) and the Helmholtz Centre for Environmental Research (UFZ) in Germany, (Scholz et al., 2013; Browne et al., 2015; Mansouri et al., 2016) have developed computational toxicology models for testing endocrine disruption *via* the estrogen receptor (ER) or thyroid receptor (TR), and such models will greatly improve the efficiency of EDC assessment. Moreover, increases in *in vitro* rapid high-throughput screening (HTS) data will accelerate the development of a variety of toolboxes based on predictive quantitative structure–activity relationship (QSAR) models built from data mining and machine learning methods (DiMaggio et al., 2010; Zhang et al., 2017).

However, most current models are binary discriminant analysis models. In other words, they are applied to determine possible effects without clarifying their agonistic and antagonistic effects, and such models have failed to accurately reveal the endocrine disruption effects. Among the divergent pathways that EDCs affect, the most well-known pathways interfere with the ER, and others, such as the TR, have also recently been under the spotlight. Meanwhile, when compared to the specific signaling pathways involved in biological toxicity, endocrine-disrupting actions through nuclear receptors function with two distinct effects, specifically as agonists and antagonists (Roig et al., 2013). Therefore, it is critical to distinguish the agonistic and antagonistic effects of EDCs for their specific receptors. For example, Chen et al. (Li et al., 2009) not only identified the thyroid hormone activities of hydroxylated polybrominated diphenyl ethers (HO-PBDEs) for the first time using a recombinant yeast two-hybrid assay but also constructed a regression model to illustrate the quantitative relationships between the activities and molecular structural descriptors to reveal a promising strategy for the profound realization of the thyroid hormone activities of HO-PBDEs. However, the model could not effectively distinguish the agonistic, antagonistic or inactive properties of HO-PBDEs. In fact, these three properties are commonly observed in *in vitro* biological screening for potential ecological and environmental harm.

In addition, the collective data used to date for constructing computational toxicology models have come from published studies with diverse assessment systems, and the studies were processed and selected with strong personal preferences (Browne et al., 2015). The accuracy of a computational model is expected to improve if the input data come from similar evaluation systems. In this context, we envisioned that a combination of an *in vitro* biological screening assay for EDCs and computational ternary classification methods would be a feasible approach for profiling and distinguishing endocrine-disrupting effects (i.e., agonistic, antagonistic or inactive) in a set of widely used environmental chemicals.

Based on this vision, we proposed the construction of computational ternary classification models to fully assess the thyroid-disrupting effects of 209 polychlorinated biphenyls (PCBs). PCBs are a class of highly persistent and toxic organic chemicals which are broadly distributed and may cause great ecological and health risks to our environment (Giera et al., 2011; Gilbert et al., 2012; Zhang et al., 2014c). About 95% of serum samples of 1800 individuals from the U.S. has detectable levels of PCBs in their blood, (Zhang et al., 2014c) likely exposed through water, industrial waste, house dust, ingestion of contaminated foods (e.g., fish, meat, and dairy products) (Freels et al., 2007; Herrick, 2010; Norström et al., 2010; Meeker et al., 2011). Experimental evidence suggests that exposure to PCBs may interfere with the endocrine system and specifically with thyroid function (Brouwer et al., 1998; Jacobson et al., 2017).

In this study, 22 commercial chemical PCB standards were tested for their TR activities using our sensitive dual-fluorescence reporter system (Zhang et al., 2014a). Moreover, machine learning methods, such as linear discriminant analysis (LDA) and support vector machines (SVM), were employed to develop ternary classification models that could classify PCBs as agonistic, antagonistic or inactive disruptors *via* TR. The most promising classification model was used to predict the classification of the remaining 187 PCB congeners. This study is the first to explore the TR-disrupting effect of 209 PCBs by a combination of experimental data and computational modeling that not only offers a promising method for detecting potential environmental harm from PCBs but also provides an advantageous alternative method in computational toxicology.

2. Materials and methods

2.1. Chemicals

The 22 different PCB congeners that were tested in this study were purchased from Dr. Ehrenstorfer GmbH (Augsburg, Germany) and were all greater than 95% pure. Liothyronine (T_3, 95%) from J&K (Beijing, China) served as a positive control. All of the stock solutions were prepared in dimethyl sulfoxide as a vehicle (DMSO; Sigma-Aldrich) at concentrations of 10^{-3} M for T_3 and 10^{-1} M for the PCBs, respectively; the stock solutions were then diluted to the desired concentrations for testing. To avoid cytotoxicity to CHO-K1 cells, the final volume of DMSO did not exceed 0.1% of the total experimental medium volume (Du et al., 2010).

2.2. Transient gene expression assay in CHO-K1 cells

The plasmids pGal4-L-hTRβ and pUAS-tk-Luc were used for transient co-expression in the thyroid receptor (TR) reporter gene assay and were kindly provided by Dr. Ronald M. Evans (Gene