Ecotoxicology and Environmental Safety 165 (2018) 194–201

Contents lists available at ScienceDirect

Ecotoxicology and Environmental Safety

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

Elucidation of endocrine – disrupting polychlorinated biphenyls binding potency with steroidogenic genes: Integration of *in silico* methods and ensemble docking approaches

K. Kranthi Kumar[1], B. Uma Devi[1], P. Neeraja*

Department of Zoology, Sri Venkateswara University, Tirupati, 517502 A.P., India

ARTICLE INFO

Keywords:
Polychlorinated biphenyls
QSAR modeling
Self-Organizing map
Clustering and classification schemes
Steroidogenesis pathway genes
Ensemble docking

ABSTRACT

A myriad of polychlorinated biphenyls (PCBs) may have potential to reproductive axis and endocrine disruptions. PCBs mostly breach the cholesterol biotransformation in mitochondria through interfering with steroidogenic genes and lead to adverse consequences in steroidogenesis; however, studies are scanty. In this examination, the combinations of quantitative structure–activity relationship (QSAR) modeling and ensemble docking approaches was performed to envisage structural properties of PCBs that influence the developmental toxicity, estrogen receptor – mediated impacts, and to provide a better comprehension of binding levels between PCBs, steroidogenic acute regulatory protein (StAR) and cholesterol side – chain cleavage enzyme (CYP11A1). Prognostic QSAR models were illustrated with good robustness and predictive ability. Models provide extensive data on applicability domain and similarities between PCBs and training set compounds was used to implement for clustering and classification of toxic PCBs by employing Self – Organizing Maps. Docking and interaction profiles interpretations provided various insights into the structural features of PCBs that influence the cholesterol binding to StAR and CYP11A1 domains. The non – polar, atomic π – stacking and halogen bonds of PCBs with novel hotspots residues of StAR and CYP11A1 was indicated subtle conformational changes that barrier to cholesterol binding and/or locks to cholesterol ejection from Ω1 – loop of StAR, and inhibits cholesterol to pregnenolone biosynthesis by CYP11A1; thus, these are probably revealed as block – cluster mechanisms. These outcomes are potential to be useful to predict developmental toxicity, endocrine disruption potencies and anti-steroidogenic activities of other environmental pollutants and provided sorted pinpoints for further evaluation of interaction mechanisms of PCBs with other sterodogenic genes.

1. Introduction

Polychlorinated biphenyls (PCBs) are a gamut of environmental pollutants. They contain 209 structurally similar congeners, which contrast with the number and position of the chlorine atoms on the biphenyl molecule (Syed et al., 2013). Due to high stability nature, PCBs are used as extreme temperature, pressure and friction reducers in capacitors, transformers, heat transfer fluids, plasticizers and hydraulic fluids (Dalefield, 2017). PCBs are released into the atmosphere through improper handling, spills and leaks of electrical equipments; consequently they bind strongly to soil and sediments, and tend to persist in the environment. Thus, they are usually found in both biological and ecological media, including water and food. The major route of PCBs exposure at significant levels to humans are overwhelmingly by eating or drinking contaminated food, inhalation and skin contact (Robertson and Hansen, 2015; IARC, 2016).

PCBs are more polar owing their lipophilic nature and undergo xenobiotic biotransformation through various cytochrome P – 450 (CYP) enzymes superfamily, and prompt the generation of a variety of hydroxylated PCBs (HO – PCBs). These HO – PCBs are showing a wide-spectrum of endocrine – defects (Gutleb et al., 2010). According to previous in vitro studies one can signify that, PCBs directly act on Leydig cells to decrease testosterone production by moderating the steroidogenic genes (i.e., 3β and 17β hydroxysteroid dehydrogenase (HSD) and cytochrome P – 450scc (CYP11A1)) expression (Murugesan et al., 2007; Kjeldsen et al., 2013). Consequently, they have been strongly associated with male reproductive health problems includes reduction in testosterone biosynthesis and spermatogenesis lead to reduced weight of testis and accessory sex organs, decreased sperm counts and seminiferous tubules consisting of cellular contents, etc.

* Corresponding author.
[1] Equal contribution.

https://doi.org/10.1016/j.ecoenv.2018.08.112
Received 15 May 2018; Received in revised form 6 August 2018; Accepted 31 August 2018

(Chao et al., 2017).

The rate−limiting step of cholesterol biotransformation is assumed as a complex process due to the impermeable nature of inner mitochondrial membrane. Thusly, lipophilic cholesterol cannot be permeable unless several proteins assisted for this achievement. The steroidogenic acute regulatory (StAR) protein performs these functions, and is an indispensable part possibly implicated in steroidogenesis that supports testosterone and other steroid-hormones biosynthesis. Albeit, any possible disruptions occurring during cholesterol biotransformation (by StAR) can result in adverse steroidogenesis impacts (Kumar et al., 2017, 2018). Significant restraint of StAR and CYP11A1 genes are reported by PCBs exposure (Murugesan et al., 2007; Kjeldsen et al., 2013). Therefore, PCBs are probably paying a little attention in the StAR and CYP11A1 binding and adverse steroidogenesis effects. Unfortunately, PCBs and their molecular basis evaluation and binding patterns in the presence of steroidogenic genes information are scarce in the literature. Despite reliable knowledge of the adverse health effects of PCBs, it is unclear how they influence developmental toxicity, endocrine disruption and adverse reproductive effects. According to these perspectives PCBs have serious concerns. In this circumstance, toxicities and molecular interactions of individual PCBs are urgently required to examine against steroidogenic genes, which provides awareness to society and researchers know about the environmental contaminants (PCBs) and their reproductive health risks.

Because there are an enormous number of PCB congeners, the experimental studies have a time−consuming and a financial burden; in this contest, the prediction and analysis of developmental toxicity and steroidogenic genes mediated effects of individual PCBs are unfeasible. With the advent of increased systems biology tools, there have been considerable improvements in the structure activity relationship (SAR) and quantitative−structure activity relationship (QSAR) regions. Hence, these models have been utilized by some regulatory authorities such as agrochemical, academic, food and pharmaceutical industries to address the toxicological examination of an intrinsic chemical with an understandable annotation. These legislative SAR/QSAR programs are assuming an essential role in the assurance of toxicological information about indistinct chemicals (Chen et al., 2016). Moreover, they can define as proficient core prediction tools in systems biology. Thus, they can offer an approach to extend our understanding of the unsafe effects of PCBs.

The intent of this study was carried out to envisage developmental toxicity, nuclear hormone receptor−mediated abilities (for estrogen receptors), StAR and CYP11A1 proteins binding properties of PCBs using computational systems biology applications. A total of 210 PCBs was collected to investigate developmental toxicity and estrogen−-mediated activity endpoints using robust QSAR modeling strategies. To derive the toxicity potentials conformation for developing QSAR models, two different methods (representation space analysis, clustering and classification) were performed. The predictive ability of QSAR models was assessed by several rule−based rationale approaches. In addition, virtual screening and three−replica docking simulations were also utilized to evaluate the interaction ability of these compounds against the binding pockets of the StAR and CYP11A1 proteins. This sort of toxicogenomics studies between PCBs and steroidogenic genes has not yet been defined. This examination could provide a promising way to reach environmental pollutants (PCBs), and provide some insights into the binding mode of PCBs to steroidogenic genes, which potential to impede the involvement of developmental and endocrine disruption possibilities.

2. Materials and methods

2.1. Data set

The examined data set is a group of 210 (209 congeners and one biphenyl molecule) compounds obtained from the US EPA (https://www.epa.gov/) database. Simplified molecular input line entry system (SMILES) and 3D−structures of 210 chemicals was retrieved from the PubChem database for further exploration (https://pubchem.ncbi.nlm.nih.gov/search/search.cgi) (See Table 1S).

2.2. QSAR models

PCBs can disrupt the reproductive axis and behaved as anti-estrogens. However, lack of developmental toxicity confirmations and ER−mediated activities for most of the PCB congeners, this study was outlined here to predict toxicity regarded endpoints using government−authorized programs. The QSAR models were described as below.

2.2.1. Developmental toxicity (CAESAR)

The developmental toxicity model is explained in the literature (Cassano et al., 2010). The model comprises an Arena data set of 292 compounds. The data set was sorted by the FDA for qualitative prediction of developmental toxicity, and the model was performed by the Random Forest Method and run with WEKA (Waikato Environment for Knowledge Analysis) open source libraries. The prediction performed was binary, 'non−toxicant' or 'toxicant', and the FDA category was defined as A or B and C or D, respectively.

2.2.2. Developmental and reproductive toxicity virtual library model (PG)

The developmental toxicity (PG) model is described in the literature (Wu et al., 2013). The model is built on a set of 641 compounds. Accordingly, the efficient 25 chemical categories are implemented in the compound library that helps to find a perfect match between the given chemicals and any of the virtual compounds in this model. If possible matches are found; the predictions are expressed as binary 'non−-toxicant' or 'toxicant'.

2.2.3. Estrogen receptor relative binding affinity (ER_RBA)

The ER_RBA model is reported in the literature (Roncaglioni et al., 2008). The model is built on a set of 806 compounds. The model is a QSAR classification model and works based on a classification and regression tree (CART) algorithm. The qualitative prediction of estrogen−mediated activity is binary, defined as 'active' and 'inactive'.

Altogether, CAESAR and PG models are not equivocal, but they are more subjective, i.e., most of the training set compounds belong to toxic class. Although, ER_RBA model is deferentially distributed, i.e., most of the training set compounds belong to a non−toxic class (inactive). The QSAR models are performed according to the OECD principles for evaluation of QSAR models along with US EPA standards (OECD, 2004). These portrayed models were implicated in the VEGA v1.1.4 modeling package (http://www.vega−qsar.eu/), which permits QSAR models to study the toxicity, ecotoxicity, environmental toxicity and physicochemical properties of a group of chemical substances. The detailed VEGA classifications were described in the Supplementary Data. Overall, the resulting data are represented as a potential source providing awareness on the prediction of the QSAR model and contributing to the decision−making process about a chemical substance.

2.3. The representation space analysis, clustering and classification

In our classification scheme or clustering, the compounds were indicated by vectors in the optimization space. The representation space was built from similarity sets. It is an association of all of the similarity sets related to a particular endpoint demonstrated from the relevant QSAR models for each endpoint prediction. In fact, it is a part of the model's training set and it is evaluated from a set of data that has been investigated as part of the training set. For further investigation, a molecule reflects a multi−dimensional vector; each vector component shows a chemical from the representation space.

Self−Organizing Maps (SOM) or the Kohonen neural network is an