



Elucidating the aryl hydrocarbon receptor antagonism from a chemical-structural perspective

E. Goya-Jorge¹, T.Q. Doan², M.L. Scippo³, M. Muller⁴, R.M. Giner⁵, S.J. Barigye⁶ and R. Gozalbes¹

Q1

¹CEEI (Centro Europeo de Empresas Innovadoras), ProtoQSAR SL, Parque Tecnológico de Valencia, Valencia, Spain; ²Departament de Farmacologia, Facultat de Farmàcia, Universitat de València, Valencia, Spain;

Q2

³Laboratory of Food Analysis, FARA-H-Veterinary Public Health, ULiège, Liège, Belgium; ⁴Laboratory for Organogenesis and Regeneration, GIGA-Research, ULiège, Liège, Belgium; ⁵MolDrug AI Systems SL, Valencia, Spain

ABSTRACT

The aryl hydrocarbon receptor (AhR) plays an important role in several biological processes such as reproduction, immunity and homeostasis. However, little is known on the chemical-structural and physicochemical features that influence the activity of AhR antagonistic modulators. In the present report, in vitro AhR antagonistic activity evaluations, based on a chemical-activated luciferase gene expression (AhR-CALUX) bioassay, and an extensive literature review were performed with the aim of constructing a structurally diverse database of contaminants and potentially toxic chemicals. Subsequently, QSAR models based on Linear Discriminant Analysis and Logistic Regression, as well as two toxicophoric hypotheses were proposed to model the AhR antagonistic activity of the built dataset. The QSAR models were rigorously validated yielding satisfactory performance for all classification parameters. Likewise, the toxicophoric hypotheses were validated using a diverse set of 350 decoys, demonstrating adequate robustness and predictive power. Chemical interpretations of both the QSAR and toxicophoric models suggested that hydrophobic constraints, the presence of aromatic rings and electron-acceptor moieties are critical for the AhR antagonism. Therefore, it is hoped that the deductions obtained in the present study will contribute to elucidate further on the structural and physicochemical factors influencing the AhR antagonistic activity of chemical compounds.

ARTICLE HISTORY

Received 9 November 2019
Accepted 19 December 2019

KEYWORDS

AhR antagonist; QSAR; toxicophore; CALUX; POP; linear discriminant analysis

Introduction

The aryl hydrocarbon receptor (AhR) is an evolutionary conserved transcription factor member of the basic helix-loop-helix (bHLH) family of receptors [1]. The AhR acts as a cytoplasmatic chemical sensor mediating intracellular and cellular signals [2] with its main transcriptional regulatory function being the up-regulation of cytochrome P450 family 1 (CYP1) of metabolizing enzymes [3]. The AhR Ligand Binding Domain (LBD) has been suggested to be one of its two PER-ARNT-SIM (PAS) domains, the PAS-B [4].

CONTACT R. Gozalbes ✉ rgozalbes@protoqsar.com

Supplementary material for this article can be accessed at: <https://doi.org/10.1080/1062936X.2019.1708460>.

© 2019 Informa UK Limited, trading as Taylor & Francis Group

However, comprehensive understanding of the structural and functional profile of the AhR has been limited by the unavailability of an experimentally determined AhR structure co-crystallized with the corresponding functional domains. Consequently, studies aimed at elucidating the interaction modes involved in the AhR signal transduction pathway have relied on homologous systems [5].

AhR owes its first discovery to 2,3,7,8-tetrachlorodibenzo-*p*-dioxin (TCDD), which is one of the most toxic synthetic compounds known to date and whose effects in biological systems are attributed to the AhR binding and activation [3]. Along with TCDD, several dioxin and dioxin-like compounds have been studied for their AhR agonistic activity and associated toxicological effects [6,7]. However, the ability of AhR to interact with a structurally diverse spectrum of xenobiotic and endogenous ligands suggests that it is highly promiscuous [8].

The induction of the AhR/CYP1A1 axis, as a consequence of AhR agonism, has been reported to enhance oestrogenic detoxification [9], regulate the oxidative balance and to propagate the metabolism of proinflammatory and tumour-promoting metabolites, among other effects [10]. Moreover, AhR plays an important role in the physiological functions of reproductive organs, immune system, liver and vascular development, cardiac function, cell growth, differentiation, homeostasis and circadian rhythms [2,11–13]. On the other hand, several toxicants and high concern substances have been reported as antagonists of AhR transcriptional activity [7,14]. However, in contrast to AhR agonism, little attention has been paid to the analysis of the biochemical consequences of AhR antagonism and even less on the subsequent toxicological implications [15–17].

Computational models, coupled with *in vitro* assays, have in the recent times gained increasing utility as alternative tools for providing insight on the pharmacological and/or toxicological effects of chemical compounds [18]. Indeed when proper interpretations are feasible, these techniques jointly provide solid understanding on the structural and functional characteristics relevant for studied bioactivity profiles [19], in addition to their inherent ethical, economic and predictive advantages [20].

It should be noted that most of the AhR computational models reported in the literature have been based on the binding affinity as the endpoint, which is an important drawback since it does not discriminate between agonists and antagonists [21,22]. In addition, as an experimentally determined AhR structure including the respective functional domains is unavailable, *in silico* initiatives have relied on homology models of the PAS-B domain to provide insights on the possible ligand binding modes [5]. The differences observed in these studies between the agonistic and antagonistic interaction modes suggested greater distortions in the structure of the LBD in case of the latter [5,23]. Nonetheless, little is still known on the chemical structural and functional features that favour the binding of AhR antagonists [24].

Therefore, further analyses of the agonist and/or antagonist responses following AhR binding are needed to gain better understanding of role of the AhR in both toxicological and pharmacological contexts [25]. Certainly, informative experimental bioassays will be necessary to obtain greater insight on the AhR-ligand interaction modes, and particularly to discriminate between agonistic and antagonistic modulations. One of the most popularized assays is the *in vitro* AhR Chemically Activated Luciferase eXpression (CALUX) bioassay, which has been reported to be useful in the detection of the AhR antagonistic and agonistic effects of dioxin and dioxin-like chemical compounds [26,27]. This assay is

based on the analysis of the induced luciferase response in a recombinant cell line driven by several CYP1A1 dioxin response elements (DREs) as a direct reflection of AhR-mediated transcriptional activity.

The goal of the present manuscript is to analyse the factors, from a chemical perspective, that influence the AhR antagonism of chemical compounds using a combination of *in vitro* and *in silico* methods. Firstly, a CALUX reporter gene assay for the AhR antagonistic activity for a set of selected xenobiotics with known toxicity profiles was performed. Then, based on the results obtained from the *in vitro* assays and an extensive literature review, QSAR and toxicophoric models were built to examine the chemical structural and physicochemical features that modulate AhR antagonism. It should be noted that while QSAR methods have been employed to model and predict the AhR binding capacity [28,29], this is the first time that *in vitro*, QSAR and toxicophoric approaches are collectively employed to analyse AhR antagonism.

Materials and methods

In vitro evaluation

Chemicals and reagents

A set of 68 chemical compounds was tested *in vitro* for their AhR antagonistic activity due to their suspected or suggested toxic effects. Twenty-nine of them correspond to Persistent Organic Pollutants (POPs) prevalent in Scandinavian human blood [30] and the remaining 39 corresponded to a chemical library belonging to the Laboratory of Food Analysis of the University of Liège. The tested compounds were mostly obtained from Sigma-Aldrich (Missouri, USA), for details see Supporting Information SI-1. All the chemicals were dissolved in dimethyl sulphoxide (DMSO) (Acros Organics, Molinons, France) as the stock solution of 20 mg/mL and kept in -20°C . The standard solution of 2,3,7,8-tetrachlorodibenzodioxin (TCDD) (purity > 98%) in DMSO was supplied from Wellington laboratories (Guelph, Canada). The reagent 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide (MTT) was obtained from Sigma Aldrich (Missouri, USA).

AhR CALUX assay

The AhR-CALUX bioassay was developed based on a stably transfected dioxin response (DR) rat hepatoma (H4IIE) cell line, consistent with previous reports in the literature [6,31]. These cells were obtained from BioDetection System (BDS) (Amsterdam, The Netherlands) transfected with four native dioxin response elements from the upstream region of the mouse *cyp1a1* gene leading the Mouse Mammary Tumour Virus (MMTV) promoter and controlling expression of the luciferase gene were stably integrated into the cell's genome. The AhR transactivation activity of the compounds was reported as the expression of the inserted luciferase and measured by light production.

The cells were maintained in α -MEM (ThermoFisher, Massachusetts, USA) with 10% v/v foetal bovine serum (Greiner, Kremsmünster, Austria) and 50 IU/mL penicillin and 50 $\mu\text{g/mL}$ streptomycin (Sigma Aldrich, USA) and were incubated at 37°C in a water saturated atmosphere injected with 5% CO_2 .

The AhR antagonistic tests for dioxin responsive chemical-activated luciferase gene expression (AhR-CALUX) bioassays were performed as indicated by the provider BDS.

Briefly, after reaching about 90% of confluence in the culture flask, the cells were trypsinized and seeded homogenously in white clear-bottomed 96 well microplates (Greiner, Kremsmünster, Austria). The cells were then incubated for 24 h and afterwards treated with the test compounds for another 24 h. The experiment was terminated by cell lysis using lysis solution containing Triton X100 (Sigma Aldrich, Missouri, USA). After adding luciferin (Promega, Wisconsin, United States) and ATP (Roche Diagnostics, Rotkreuz, Switzerland) to the cell lysate, the plates were read by a luminometer (ORION II, Berthold Detection System, Pforzheim, Germany).

For antagonistic tests, the cells were co-exposed to the test compounds and 15 pM TCDD corresponding to TCDD EC₅₀ in DR-H4IIE. All the experiments were repeated at least three times independently. The maximum concentration tested to determine the AhR antagonistic activity was 40 μM with some exceptions (see SI-1), and the final concentration of DMSO in the culture medium was 0.4%. To ensure the adequacy of the test method and provide a basis for comparisons, a reference curve of the positive control TCDD, as well as concurrent negative and solvent controls were added on each plate.

Cell viability assay

An analysis of the MTT cell viability as well as a visual inspection of the cell morphology and attachment were performed to detect the cytotoxic compounds. After exposure to the test compounds, 25 μL of MTT solution 5 mg/mL was added into each well. The plates were then incubated for 4 h at 37°C to allow for the formation of the purple metabolite formazan from the tetrazolium dye MTT via the activity of the mitochondrial succinate dehydrogenase. Later, the formazan crystals were dissolved during 2 h by adding 100 μL of isopropanol. The MTT formazan absorbance was read at 550/630 nm using a microplate spectrophotometer (ELX800™ BioTek Inc., Winooski, USA).

Results for both reporter gene and MTT assays were presented as relative responses, either as the percentage of the cell response to the tested compound compared to the cell response to TCDD EC₅₀ on the same plate for CALUX assays, or to solvent control DMSO for MTT test. A compound was considered an AhR antagonist when it was able to reduce the activity of the TCDD EC₅₀ from 100% to at least 70%, while reductions in cell population greater than 15% were considered cytotoxic [32].

QSAR models

Based on both the experimental in vitro results obtained herein and those compiled from an extensive review of the literature, binary classifiers of AhR antagonism were built using statistical modelling methods. The molecular characterization was conducted by means of simple and interpretable chemical structural descriptors.

Structures, descriptors and activity

A dataset of 116 chemicals was built to develop the QSAR classification models. This set comprised of 68 compounds tested in the present report for their AhR antagonistic activity and 48 compounds retrieved from literature. From this dataset, compounds with undetermined or inconsistent activity were excluded from the analysis and therefore minimizing potential error sources (for details see SI-1). A binary scale of activity values was considered in the sense that the compounds observed or reported in the literature to

induce the AhR antagonistic effect, based on the AhR-CALUX method, were labelled as active while those that did not induce any effect were considered as inactive.

The calculated molecular descriptors were implemented in-house, based on definitions obtained from DRAGON [33] and PADEL [34] software. Consistent with the parsimony principle, only the simplest characterizations of the molecular structures were considered. A total of 1929 DRAGON descriptors were calculated corresponding to: constitutional descriptors, information indices, functional group counts, atom-centred fragments, 2D binary fingerprints and 2D frequency fingerprints. From this set, constant variables or those with pair-correlations greater than 0.9 were removed, retaining only 171 descriptors. As for the PADEL software, the Estate and MACCS fingerprints were computed yielding 245 descriptors from which 15 constants were removed. Globally, a total of 401 molecular descriptors were retained and subsequently used as input variables for the chemoinformatic modelling.

Variables and QSAR methods

The selection of molecular descriptors for modelling the AhR antagonistic activity was carried out using the information gain filter as provided by IMMAN software [35]. The information gain defined as the reduction in variable entropy (or uncertainty) given that another variable (response variable in this context) is known. This measure allows variables to be ranked based on their capacity to discriminate cases that belong to different classes and consequently filter out informative variables.

A Principal Components Analysis (PCA) was performed on the selected set of variables and 10 principal components (PCs), explaining more than 90% of the total variance, were obtained. Subsequently these PCs were used as input variables for modelling the AhR antagonistic activity. The PCs were rotated using the orthogonal rotation scheme varimax to maximize the sum of variances of the squared loadings. For the PCA the MATLAB software [36] was employed.

Linear Discriminant Analysis (LDA) and Logistic Regression (LReg) methods were employed for the classification model building. The metrics considered to evaluate the performance of the built classification models were the accuracy, precision, sensitivity and specificity (see below) where: TA = true active; FA = false active; TI = true inactive and FI = false inactive.

$$\text{Accuracy} = \frac{(TA + TI)}{(TA + FI + FA + TI)} \times 100\%$$

$$\text{Precision} = \frac{TA}{(TA + FA)} \times 100\%$$

$$\text{Sensitivity} = \frac{TA}{(TA + FI)} \times 100\%$$

$$\text{Specificity} = \frac{TI}{(TI + FA)} \times 100\%$$

Validation methods & applicability domain

Internal and external validation experiments were designed to assess the robustness and predictiveness of the built classification models. To this end, a cluster analysis was performed to rationally divide the dataset in training and test sets, based the following procedure: first, with hierarchical clustering (based on the square-Euclidean distance as the similarity criterion) an optimal number of clusters (k) was established using the graph of the amalgamation schedule. Then, k-Means method was employed to stratify the structures into clusters according to their similarity. An appropriate distribution of active and inactive compounds was considered during the clustering. With this procedure, two subsets of compounds were generated with 75% of the data comprising the training set and 25% test set. Internal leave-group-out cross validation (25% of the training set was held out for validation for each fold) was performed to check for possible overfitting during the model building procedure. Ten different combinations of training and test sets were created using cluster analysis and a 10-fold external validation was performed to evaluate the average predictive capacity of the models. This approach has been recommended as the most stringent validation protocol for QSAR models [19]. For the cluster analysis, the STATISTICA [37] software was employed.

The applicability domain (AD) of the built QSAR models was established by using the Euclidean distances and Tanimoto coefficients as the structural similarity measures and the k-nearest neighbour algorithm (k-NN) as the feature space approximator. For this AD analysis the AmbitDiscovery software [38] was used.

Toxicophoric modelling

A diverse set of active AhR antagonistic compounds was selected from the dataset built in the present report for mapping the toxicophoric hypothesis. The selection criterion was based on the hierarchical and k-means clustering previously described. Additionally, a Tanimoto matrix score provided by the PubChem [39] platform was considered. For this study, the PHASE module of Schrödinger [40] software was used.

The three-dimensional structures of the selected set of active AhR antagonistic compounds were generated and employed to build toxicophoric models following a ligand-dependent approach exclusively. The mapped space was unrestricted by considering all volumes and shapes, but with locations and orientations in 3D-spaces allowing the enantiomeric discrimination by absolute coordinated distances. All the available features included in PHASE were analysed at distances of 1.8Å and with a tolerance of 2.0 Å to match the hypothesis. These features were represented in the toxicophoric models as (A) acceptor, (D) donor, (H) hydrophobic, (N) negative ionic, (P) positive ionic, and (R) aromatic ring. The best alignment was searched considering 50 conformations for each compound. The generated hypotheses were accepted if they fitted at least 50% of the training set. Three to seven features were considered as optimum for hypothesis acceptability.

The goodness of fit of the toxicophoric hypotheses was evaluated using the Boltzmann-enhanced discrimination of receiver operating characteristic (BEDROC), which is a generalization of the ROC metric to deal with the early recognition problem. The difference criterion between toxicophoric models was set to 0.5, while the Phase Hypo Score was employed as the scoring function. The mathematical definitions for all

these metrics are detailed elsewhere and the performance is measured from 0–1 values where 1 represents perfect fitness [41]. 245

Q3 Finally, an external set of decoy structures was generated using as seeds the diverse set of active AhR antagonistic compounds employed in the toxicophore hypothesis building. For decoy structure generation, the Directory of Useful Decoys (DUD•E) online platform was employed. DUD•E constitutes the largest and most comprehensive public dataset for benchmarking virtual screening programs and was employed here to validate the predictive capacity of the toxicophoric hypotheses. DUD•E is designed to provide sets of structures with similar physicochemical properties but dissimilar topology to the active compounds being, therefore, unlikely binders [42]. 250

Results and discussion 255

In vitro results

Out of the 68 test compounds evaluated with the AhR-CALUX assay, 24 of these induced AhR antagonistic transactivity, suppressing the cell response to the spike-in control 15 pM TCDD from 100% to at least 70%. Three of these, two insecticide isomers (α -endosulfan and β -endosulfan) and the fungicide thirame, reduced the cell viability by more than 15%. Moreover, the former two showed cytotoxic effects in more than 60% of the cell population at a concentration of 40 μ M while the latter showed a similar effect even at 20 μ M. These three chemicals were therefore excluded from the analysis of AhR antagonism (see SI-1). The remaining 41 chemicals did not show any AhR antagonistic potential and were thus considered as inactive. 260 265

QSAR modelling

Database of structures and ahr antagonism

A dataset of 116 chemical compounds with defined AhR antagonistic activity profiles based on the DR-H4IIE cell line was constructed. This database contained organohalogenated (chlorinated, brominated, fluorinated), polyaromatic hydrocarbon, flavonoid, quinone, imidazole, bisphenol and pyrethroid chemical classes. Although a few antibiotics and phytochemicals were included, most of the modelled structures are or have been used as industrial chemicals, pesticides, fungicides, algaecides, insecticides, rodenticides and herbicides. Many of these profiled compounds are reported as potential endocrine disruptors or have been already banned due to their toxic spectrum through different mechanisms. 270 275

During the construction of the dataset, interest was placed on the homogeneity of the results obtained from the experimental protocols, in order to obtain logical and interpretable conclusions. Therefore, compounds with ambiguous and/or unclear experimental determinations were discarded, especially for those results extracted from literature. Such stringent selectiveness has been strongly supported by several reports in the literature [19]. 280

The hexachlorocyclohexane (HCH) conformational isomers (α -, β -, γ -, δ -HCH) were excluded since they possessed contrasting activity profiles (the descriptors used in the present study are insensitive to stereochemical differences). For the same reason, two 285

polycyclic aromatic hydrocarbons were excluded. The hexabromocyclododecane (HBCD) in vitro activity was the only one considered, while the rest of the isomers (α -, β -, γ -HBCD) reported in the literature were discarded. Finally, one structural outlier (chlormequat chloride) was removed during the modelling due to disconnected fragments in the molecular structure. 290

Moreover, BDE-49 (i.e. 2,2',4,5'-tetrabromodiphenyl ether) [43] and nine polychlorinated biphenyls (PCBs) [7] were excluded as their activity profiles were not clearly defined in the literature (see SI-1). Three cytotoxic compounds in the MTT assay were not considered (see above). 295

A standard protocol was followed to guarantee the homogeneity of the final dataset presented in this study that included 93 compounds, 40 active AhR antagonists and 53 inactive. Therefore, this dataset may be considered of acceptable quality and is, to the best of our knowledge, the largest for AhR antagonistic activity reported so far. 300

Selection and transformation of molecular descriptors

The calculated molecular descriptors were all based on simple and interpretable features of the chemical compounds. Due to the large number of generated descriptors, dimensionality reduction was deemed necessary. 305

From the calculated 401 molecular descriptors, a set of 96 descriptors was retrieved using the aforementioned information gain filter, and subsequently used to generate orthogonal descriptors using the PCA method. Ten PCs were determined to be the optimum number of PCs based on the eigenvalues of the generated covariance matrix and these described 90% of the total variance in the dataset. The matrix of the PCs coefficients obtained for the set of molecular descriptors and the scores for all the PCs are provided as supplementary information (SI-2). 310

QSAR models obtained

For the QSAR modelling of the AhR antagonistic activity, simple and interpretable classification algorithms were preferred. Hence, five models were selected following commonly established criteria [19], three corresponding to LDA and two corresponding to LReg methods. The training metrics obtained for these models are shown in Table 1 and additional parameters are provided as SI-3. One of the significant advantages offered by these statistical models, compared to most machine learning algorithms such as random forest, support vector machine or artificial neural networks, is the straightforward 315

Table 1. Evaluation metrics of the QSAR models.

Model ^a	PCs	Accuracy		Precision		Sensitivity		Specificity	
		Train ^b	Ext.V ^c	Train ^b	Ext.V ^c	Train ^b	Ext.V ^c	Train ^b	Ext.V ^c
LDA	10	88.51	81.48	87.02	73.49	88.01	83.56	89.21	79.95
LDA	9	83.02	81.94	87.02	73.58	88.12	83.33	78.11	79.62
LDA	6	78.01	74.07	76.04	65.33	63.23	77.58	89.30	68.95
LReg	7	88.50	73.61	88.51	64.89	88.20	74.78	89.22	72.14
LReg	6	93.51	78.24	87.01	70.32	88.14	74.89	99.03	79.00

^a QSAR models based on linear discriminant analysis (LDA) and logistic regression (LReg).

^b Value of the classification metrics obtained during the training of the models (Additional parameters in SI-3).

^c Average of the classification metrics made during the 10-fold external validation procedure.

interpretation and consequently applicability in the rationalization of the influence of the different chemical structural features on the modelled property [44].

Moreover, both the internal and external validation procedures yielded adequate statistical parameters, demonstrating the robustness and predictive capacity of the five QSAR models obtained herein. Indeed, all the selected models showed good external validation performance, yielding percentages higher than 60% for accuracy, precision, sensitivity and specificity (Table 1).

These models are the first QSARs for AhR antagonistic activity of which we are aware. Previous QSAR models reported in the literature focused on AhR agonism or non-specific AhR-ligand binding [28,45]. Moreover, the models presented herein were built over a diverse dataset of chemicals and were observed to be consistent with the OECD validation principles [46]. Therefore, these models can reliably predict the AhR antagonistic potential of untested chemicals in their AD.

Toxicophoric mapping

Toxicophoric (or Pharmacophoric) models are defined as an ensemble of steric and electronic features that are necessary to ensure the optimal supramolecular interactions with a specific biological target and to trigger (or block) its biological response [47]. Toxicophoric modelling aims to identify a set of features or restrictions in the molecules and/or their receptors that determine their capacity to interact in any mode and consequently cause toxic effects. These models may be built following structure-based or ligand-based approaches depending on whether the information of the target receptor is employed or not [48,49].

Herein, as a fully crystalized structure of the AhR protein (containing the PAS-B ligand binding domain) is not yet available [25], a ligand-based toxicophoric modelling was performed. The cluster analysis performed during the QSAR modelling was employed to select the most dissimilar compounds for the toxicophoric mapping.

Seven most dissimilar structures were selected (Figure 1) and used to explore toxicophoric hypothesis of AhR antagonism. These chemical compounds are considered to be an environmental and health threat and they showed strong antagonistic activity in the in vitro AhR-CALUX assay.

A total of 8 toxicophoric hypotheses were obtained and their performance assessed based on the internal validation enrichment metrics. Consequently, two toxicophoric models were selected as they showed adequate performance and contributed dissimilar structural information.

Figure 2(a,b) show the spatial representations of the selected Toxicophoric-AhR antagonistic hypotheses for the first: Hydrophobic_Hydrophobic_Hydrophobic_ Aromatic Ring (HHHR) and the second: Acceptor_Hydrophobic_Hydrophobic_ Aromatic Ring (AHHR) models, respectively. Both arrangements included a common segment of two hydrophobic constraints and an aromatic ring as critical features. On the other hand, the latter model included an electro-acceptor group while the former possessed an additional hydrophobic constraint.

An external validation of the two selected AhR-antagonistic toxicophoric models was conducted as recommended [42], using a diverse dataset of 350 decoy structures generated from the seven active molecules employed in the hypotheses exploration.

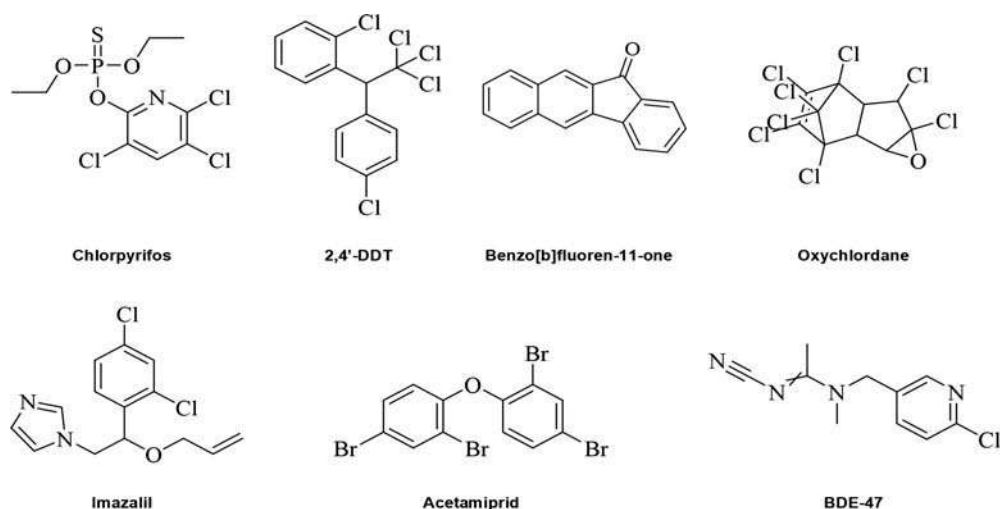


Figure 1. Active structures selected for toxicophoric modelling (For details see SI-1).

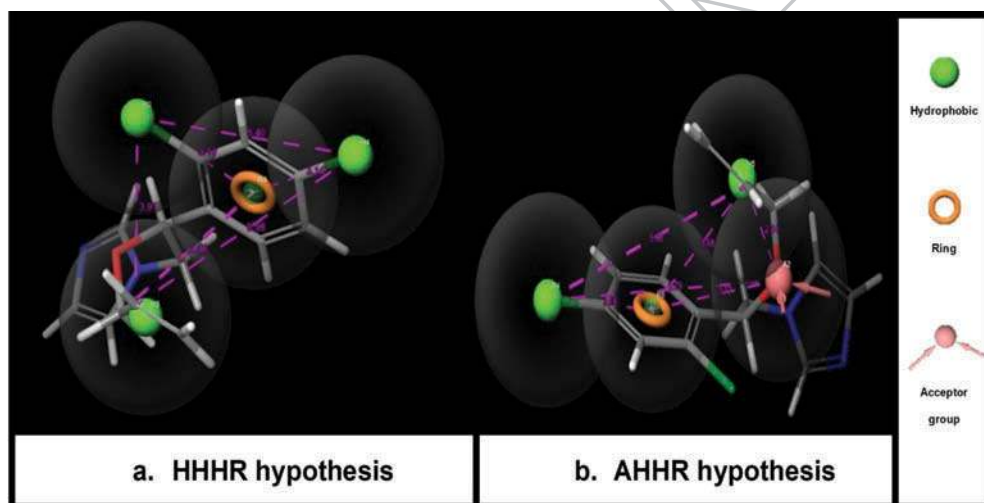


Figure 2. Representation of the toxicophoric models designed in Schrödinger software. Hydrophobic (green-ball), ring (orange ring), acceptor group (pink ball with signalling arrows). a. HHR: Hydrophobic_Hydrophobic_Hydrophobic_Aromatic Ring. b. AHR: Acceptor_Hydrophobic_Hydrophobic_Aromatic Ring.

The training and the corresponding external validation results of the HHR and AHR toxicophoric hypotheses are shown in Table 2. Overall, the built hypotheses showed good performance as all validation metrics were greater than 0.5 and thus demonstrating their robustness and predictive capacity.

Several pharmacophoric models considering nuclear receptors ER α and ER β as drug targets [50,51] have been reported in the literature. However, limited distinctions between agonist and antagonistic ligands have been suggested [52]. Furthermore, the few toxicophoric considerations available are generally restricted to drug-safety contexts

Table 2. Results of the training and the external validation for HHHR and AHHR toxicophoric hypotheses.

Hypothesis	Training performance		External validation	
	PHASE ^c	BEDROC ^d	PHASE ^c	BEDROC ^d
HHHR ^a	0.93	0.68	0.93	0.98
AHHR ^b	0.82	0.69	0.82	0.92

^a HHHR: Hydrophobic hydrophobic hydrophobic aromatic ring.

^b AHHR: Acceptor hydrophobic hydrophobic aromatic ring.

^c Scoring function phase hypo score [41].

^d Boltzmann-enhanced discrimination of receiver operating characteristic (BEDROC) [41].

[48,49]. Nevertheless, some successful and experimentally validated studies on enzymatic-mediated endocrine disruptive mechanisms have been reported based on the inhibition of 17 β -hydroxysteroid dehydrogenases [53].

The first pharmacophoric/toxicophoric estimations based on the blocking capacity of chemicals over AhR are herein proposed. Moreover, considering the good predictive power offered by these computational models, they could be a useful virtual screening tool for chemical compounds with potential AhR antagonistic activity.

Structural and physicochemical interpretations

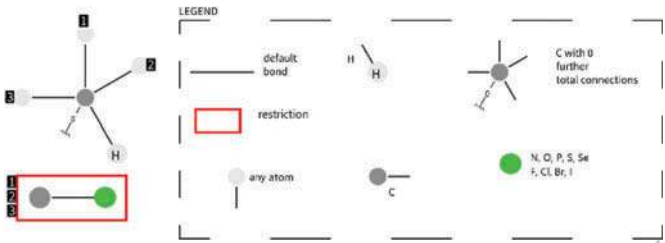
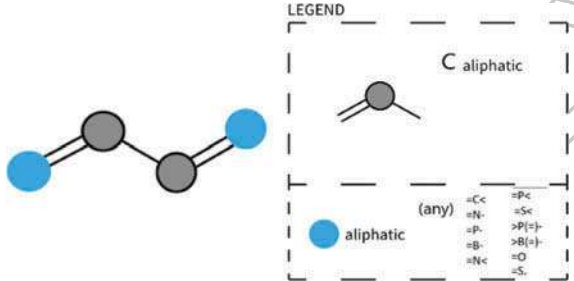
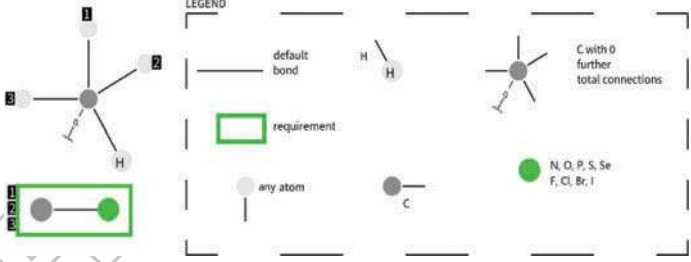
As an essential receptor whose cellular implications are still being discovered, the disruption of the AhR functions could lead to toxic effects that are yet to be fully understood [3,8,11,12]. Hence, the mapping of chemical structural features critical for the AhR agonistic/antagonistic potential could enhance the identification of specific modulators with possible toxic profiles. In this sense, based on the developed QSAR and toxicophore-based models, analyses were performed to further understand how different chemical-structural features influence the AhR antagonistic activity.

The interpretation of computational models is highly recommended by the OECD guidelines to enhance the utility of the QSAR predictive results [46]. Here, only simple 0-2D descriptors were calculated and used for the model building, therefore the elucidation of their contribution to the AhR antagonistic activity in chemical structural terms is straightforward. On the other hand, the toxicophore-based methods provide information on the steric and electrostatic features deemed critical for a given toxicity (or activity) profile. Hence both QSAR and toxicophore models may serve as complementary approaches providing a more complete outlook of the structural and physicochemical factors that influence the AhR antagonistic activity.

The molecular descriptors with relevant contribution were determined based on the PCA factor loadings (SI-2). It has been previously reported that coefficients greater than 0.7 are considered remarkably influential to the generated PCA components [54]. Therefore, the same cut-off was considered to identify the most influential descriptors on the modelled AhR antagonism. Definitions and SMARTS [55] representations of the top five variables obtained are provided in Table 3.

From the toxicophoric models, hydrophobic interactions were identified as an important constraint in active AhR antagonistic molecules. Moreover, the presence of an aromatic ring and an electron acceptor group was suggested as relevant for AhR

Table 3. Molecular descriptors most relevant for the AhR antagonistic activity of the analysed chemical compounds.

ID	Molecular descriptors*
H-046	
nCconj	<p>H attached to C (sp³) no X attached to next C</p>  <p>Sum of the number of non-aromatic conjugated C(sp²) belonging to any conjugated system, excluding aromatic rings. Including the following atom types: =C<, =N<, =P<, =B<, =N<, =P<, =S<, >P(=), >B(=), =O, =S</p>
H-052	 <p>H attached to C(sp³) with 1X attached to next C</p>
nBM	Multiple bonds
nCIR	Number of circuits

* Brief definitions/representations of the descriptors using the SMARTs [55] formats of the codified features are provided. The range of values obtained was: H-046 [0–10]; nCconj [0–6]; H-052 [0–12]; nBM [0–26]; nCIR [0–28].

antagonism. These features will now be analysed in detail and the possible complementarity with the QSAR model variables also explored.

Hydrophobic constraints

The hydrophobic features comprising the toxicophoric model and the QSAR model descriptors codifying information on hydrophobicity (i.e. hydrocarbon chains without attached heteroatoms (H-046) and closed hydrophobic circuits (nCIR) of saturated cyclic

hydrocarbons) suggest that non-polar structural segments probably play an important role in AhR-ligand interactions. Indeed, early studies on TCDD where AhR was simply considered as an intracellular, soluble and unknown receptor protein, already recognized its hydrophobic character [56].

Some other reports in the literature have also highlighted the importance of hydrophobicity in AhR-LBD interactions e.g. a strong correlation has been found between the suppression of AhR on TCDD-induced activity and the hydrophobicity of curcumin derivatives [17]. It has also been suggested that the affinity of stilbene derivatives of resveratrol to AhR is enhanced by substituents with high hydrophobicity, among other factors [57]. Moreover, an analysis of the AhR ligand binding pocket revealed that it is comprised predominantly of hydrophobic residues with some polar segments located near the medial positions of the ligands [58]. Therefore, structural fragments with strong hydrophobic character should probably be prioritized in the design of ligands with possible AhR antagonistic activity.

Ring feature

Similarly, both the toxicophoric hypotheses and the QSAR models included ring features [circuit number (nClR), in the case of the latter] as critical for AhR antagonism. Bearing in mind that the name AhR is attributed to the polycyclic aromatic hydrocarbons (PAHs), which were in fact the very first ligands found to interact with this receptor, it is unsurprising that ring features are suggested to play a critical role in AhR-mediated interactions probably through non-covalent interactions (i.e. between electron-rich and electron-deficient aromatic moieties) [59]. These interactions, also named aromatic donor-acceptor interactions, have been analysed for several decades in the AhR context [60,61]. Firstly, early studies hypothesized that π - π -interactions with phenyl and tyrosyl groups of the AhR protein occurred in the TCDD-mode of action [56]. Also, these interactions have been evoked to rationalize the effect of PBDEs and PCBs upon contact with dissolved organic matter in the environment [28], and have been suggested to play an important role in endocrine disruptive mechanisms e.g. in ER disruption [62].

Nonetheless, it is important to highlight that there ligands without the neutral, hydrophobic and non-polar molecular characteristics of PAHs that have also been shown to have effects on AhR, demonstrating that these interactions depend on the contribution of multiple factors [2,63].

Acceptor-group presence

The charge transfer between molecules and biological systems commonly influences the potential toxicity of the former [64]. Herein, one of the proposed toxicophoric hypothesis suggested the presence of an acceptor group as a relevant structural feature for AhR antagonistic activity. Likewise, the QSAR models comprise various key molecular descriptors that could be related to this toxicophoric feature. For example, heteroatoms attached to sp³ carbons (H-046 and H-052), non-aromatic conjugated systems (nConj) or multiple bonds (nBM) in general could all contribute to electron deficient fragments in the ligand's structure whose interactions may potentially influence the AhR antagonistic effects.

Altogether, the results obtained herein suggest that the presence of electron acceptor groups, connected ring systems (preferably aromatic or delocalized moieties), and most importantly hydrophobic groups seems to be critical for the potential AhR antagonistic effects.

Conclusions

Structural and physicochemical determinants of the AhR antagonistic capacity were assessed using a combination of in vitro (AhR-CALUX bioassay) and in silico (QSAR and toxicophoric mapping) methods. In the former, several chemical compounds were evaluated, and their antagonistic activity profiles determined. In the latter, QSAR algorithms based on LDA and LReg were obtained and validated. These models were shown to possess adequate robustness and predictive power, based on the quality of the obtained statistical validation parameters. Moreover, toxicophoric models were derived and the structural features potentially favouring the AhR antagonistic potential analysed. Interpretations of the QSAR and toxicophore models revealed that electron acceptor groups, aromatic or delocalized ring systems, as well hydrophobic moieties probably favour the AhR antagonistic activity, demonstrating the complementarity of the two approaches notwithstanding their dissimilar conceptual basis. The models proposed herein could be useful in the prediction of the AhR antagonistic capacity of chemical compounds with toxicological and pharmacological applications.

Disclosure statement

Q4 No potential conflict of interest was reported by the authors.

Funding

This work was developed by the Innovative Training Network 'PROTECTED': PROTECTION against Endocrine Disruptors; Detection, mixtures, health effects, risk assessment and communication. The project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie actions (MSCA) with grant agreement No. 722634.

ORCID

E. Goya-Jorge  <http://orcid.org/0000-0003-0515-989X>

References

- Q5** [1] D.W. Nebert, *Aryl hydrocarbon receptor (AHR): "Pioneer member" of the basic-helix/loop/helix per-Arnt-sim (bHLH/PAS) family of "sensors" of foreign and endogenous signals*, *Prog. Lipid Res.* 67 (2017), pp. 38–57. doi:10.1016/j.plipres.2017.06.001.
- [2] K. Kawajiri and Y. Fuji-Kuriyama, *The aryl hydrocarbon receptor: A multifunctional chemical sensor for host defense and homeostatic maintenance*, *Exp. Anim.* 66 (2016), pp. 75–89. doi:10.1538/expanim.16-0092.
- [3] K.W. Bock, *From TCDD-mediated toxicity to searches of physiologic AHR functions*, *Biochem. Pharmacol.* 155 (2018), pp. 419–424. doi:10.1016/j.bcp.2018.07.032.
- Q6** [4] R. Pohjanvirta, *The AH Receptor in Biology and Toxicology*, John Wiley and Sons, 2011.
- [5] L. Bonati, D. Corrada, S. Giani Tagliabue, and S. Motta, *Molecular modeling of the AhR structure and interactions can shed light on ligand-dependent activation and transformation mechanisms*, *Curr. Opin. Toxicol.* 2 (2017), pp. 42–49. doi:10.1016/j.cotox.2017.01.011.
- [6] T. Hamers, J.H. Kamstra, P.H. Ceniijn, K. Pencikova, L. Palkova, P. Simeckova, J. Vondracek, P. L. Andersson, M. Stenberg, and M. Machala, *In vitro toxicity profiling of ultrapure non-dioxin-*

- like polychlorinated biphenyl congeners and their relative toxic contribution to PCB mixtures in humans, *Toxicol. Sci.* 121 (2011), pp. 88–100. doi:10.1093/toxsci/kfr043.
- [7] P. Brenerová, T. Hamers, J.H. Kamstra, J. Vondráček, S. Strapáčová, P.L. Andersson, and M. Machala, *Pure non-dioxin-like PCB congeners suppress induction of AhR-dependent endpoints in rat liver cells*, *Environ. Sci. Pollut. Res.* 23 (2016), pp. 2099–2107. doi:10.1007/s11356-015-4819-6. 495
- [8] L. Stejskalova, Z. Dvorak, and P. Pavek, *Endogenous and exogenous ligands of aryl hydrocarbon receptor: Current state of art*, *Curr. Drug Metab.* 12 (2011), pp. 198–212. doi:10.2174/138920011795016818. 500
- [9] G.A. Reed, K.S. Peterson, H.J. Smith, J.C. Gray, D.K. Sullivan, M.S. Mayo, J.A. Crowell, and A. Hurwitz, *A phase I study of indole-3-carbinol in women: Tolerability and effects*, *Cancer Epidemiol. Biomarkers Prev.* 14 (2005), pp. 1953–1960. doi:10.1158/1055-9965.EPI-05-0121.
- [10] O. Hankinson, *The role of AHR-inducible cytochrome P450s in metabolism of polyunsaturated fatty acids*, *Drug Metab. Rev.* 48 (2016), pp. 342–350. doi:10.1080/03602532.2016.1197240. 505
- [11] A. Puga, Y. Xia, and C. Elferink, *Role of the aryl hydrocarbon receptor in cell cycle regulation*, *Chem. Biol. Interact.* 141 (2002), pp. 117–130. doi:10.1016/S0009-2797(02)00069-8.
- [12] C. Esser, A. Rannug, and B. Stockinger, *The aryl hydrocarbon receptor in immunity*, *Trends Immunol.* 30 (2009), pp. 447–454. doi:10.1016/j.it.2009.06.005.
- [13] C. Duval, E. Blanc, and X. Coumoul, *Aryl hydrocarbon receptor and liver fibrosis*, *Curr. Opin. Toxicol.* 8 (2018), pp. 8–13. doi:10.1016/j.cotox.2017.11.010. 510
- [14] K. Pěničková, L. Svrčková, S. Strapáčová, J. Neča, I. Bartoňková, Z. Dvořák, M. Hýžďalová, J. Pivnička, L. Pálková, H.J. Lehmler, X. Li, J. Vondráček, and M. Machala, *In vitro profiling of toxic effects of prominent environmental lower-chlorinated PCB congeners linked with endocrine disruption and tumor promotion*, *Environ. Pollut.* 237 (2018), pp. 473–486. doi:10.1016/j.envpol.2018.02.067. 515
- [15] S. Safe, S.O. Lee, and U.H. Jin, *Role of the aryl hydrocarbon receptor in carcinogenesis and potential as a drug target*, *Toxicol. Sci.* 135 (2013), pp. 1–16. doi:10.1093/toxsci/kft128.
- [16] Z. Xue, D. Li, W. Yu, Q. Zhang, X. Hou, Y. He, and X. Kou, *Mechanisms and therapeutic prospects of polyphenols as modulators of the aryl hydrocarbon receptor*, *Food Funct.* 8 (2017), pp. 1414–1437. doi:10.1039/C6FO01810F. 520
- [17] R. Nakai, S. Fukuda, M. Kawase, Y. Yamashita, and H. Ashida, *Curcumin and its derivatives inhibit 2,3,7,8-tetrachloro-dibenzo-p-dioxin-induced expression of drug metabolizing enzymes through aryl hydrocarbon receptor-mediated pathway*, *Biosci. Biotechnol. Biochem.* 82 (2017), pp. 616–628. doi:10.1080/09168451.2017.1386086. 525
- [18] H. Kandárová and S. Letaáiová, *Alternative methods in toxicology: Pre-validated and validated methods*, *Interdiscip. Toxicol.* 4 (2011), pp. 107–113. doi:10.2478/v10102-011-0018-6.
- [19] A. Cherkasov, E.N. Muratov, D. Fourches, A. Varnek, I.I. Baskin, M. Cronin, J. Dearden, P. Gramatica, Y.C. Martin, R. Todeschini, V. Consonni, V.E. Kuz'min, R. Cramer, R. Benigni, C. Yang, J. Rathman, L. Terfloth, J. Gasteiger, A. Richard, and A. Tropsha, *QSAR modeling: Where have you been? Where are you going to?* *J. Med. Chem.* 57 (2014), pp. 4977–5010. doi:10.1021/jm4004285. 530
- [20] L.G. Valerio, C. Yang, K.B. Arvidson, and N.L. Kruhlak, *A structural feature-based computational approach for toxicology predictions*, *Expert Opin Drug Metab. Toxicol.* 6 (2010), pp. 505–518. doi:10.1517/17425250903499286. 535
- [21] D. Szöllösi, Á. Erdei, G. Gyimesi, C. Magyar, and T. Hegedüs, *Access path to the ligand binding pocket may play a role in xenobiotics selection by AhR*, *PLoS ONE* 11 (2016), pp. 1–22. doi:10.1371/journal.pone.0146066.
- [22] A.D. Şahin and M.T. Saçan, *Understanding the toxic potencies of xenobiotics inducing TCDD/TCDF-like effects*, *SAR QSAR Environ. Res.* 29 (2018), pp. 117–131. doi:10.1080/1062936X.2017.1414075. 540
- [23] M.K. Gadhwal, S. Patil, P. D'mello, and A. Joshi, *Homology modeling of aryl hydrocarbon receptor and docking of agonists and antagonists*, *Int. J. Pharm. Pharm. Sci.* 5 (2013), pp. 76–81.

- [24] D. Dolciami, M. Gargaro, B. Cerra, G. Scalisi, L. Bagnoli, G. Servillo, M.A.D. Fazia, P. Puccetti, F. J. Quintana, F. Fallarino, and A. Macchiarulo, *Binding mode and structure–Activity relationships of ITE as an aryl hydrocarbon receptor (AhR) agonist*, ChemMedChem 13 (2018), pp. 270–279. doi:10.1002/cmdc.201700669. 545
- [25] S.-H. Seok, W. Lee, L. Jiang, K. Molugu, A. Zheng, Y. Li, S. Park, C.A. Bradfield, and Y. Xing, *Structural hierarchy controlling dimerization and target DNA recognition in the AHR transcriptional complex*, Proc. Natl. Acad. Sci. 114 (2017), pp. 5431–5436. doi:10.1073/pnas.1617035114. 550
- [26] G. Suzuki, M. Nakamura, C. Michinaka, N.M. Tue, H. Handa, and H. Takigami, *Dioxin-like activity of brominated dioxins as individual compounds or mixtures in in vitro reporter gene assays with rat and mouse hepatoma cell lines*, Toxicol. Vitro 44 (2017), pp. 134–141. doi:10.1016/j.tiv.2017.06.025. 555
- [27] S. Sciuto, M. Prearo, R. Desiato, C. Bulfon, E.A.V. Burioli, G. Esposito, C. Guglielmetti, L. Dell’atti, G. Ru, D. Volpatti, P.L. Acutis, and F. Martucci, *Dioxin-like compounds in lake fish species: Evaluation by DR-CALUX bioassay*, J. Food Prot. 81 (2018), pp. 842–847. doi:10.4315/0362-028XJFP-17-476. 560
- [28] A.L.J. Nuerla, X.L. Qiao, J. Li, D.M. Zhao, X.H. Yang, Q. Xie, and J.W. Chen, *Effects of substituent position on the interactions between PBDEs/PCBs and DOM*, Chin. Sci. Bull. 58 (2013), pp. 884–889. doi:10.1007/s11434-012-5464-9. 560
- [29] R. Panda, A.S.S. Cleave, and P. Suresh, *In silico predictive studies of mAHR congener binding using homology modelling and molecular docking*, Toxicol. Ind. Health 30 (2014), pp. 765–776. doi:10.1177/0748233712463774. 565
- [30] H.F. Berntsen, V. Berg, C. Thomsen, E. Ropstad, and K.E. Zimmer, *The design of an environmentally relevant mixture of persistent organic pollutants for use in in vivo and in vitro studies*, J. Toxicol. Environ. Health Part A 24 (2017), pp. 1002–1016. doi:10.1080/15287394.2017.1354439. 570
- [31] A.J. Murk, J. Legler, M.S. Denison, J.P. Giesy, C. Van de Guchte, and A. Brouwer, *Chemical-activated luciferase gene expression (CALUX): A novel in vitro bioassay for Ah receptor active compounds in sediments and pore water*, Fundam. Appl. Toxicol. 33 (1996), pp. 149–160. doi:10.1006/faat.1996.0152. 570
- [32] OECD, *Test No. 455: Performance-Based Test Guideline for Stably Transfected Transactivation in Vitro Assays to Detect Estrogen Receptor Agonists and Antagonists*, 2016. 575
- [33] R. Todeschini and V. Consonni, *Molecular Descriptors for Chemoinformatics*, WILEY-VCH, 2009. 575
- [34] C.W. Yap, *PaDEL-descriptor: An open source software to calculate molecular descriptors and fingerprints*, J. Comput. Chem. 32 (2011), pp. 1466–1474. doi:10.1002/jcc.v32.7. 575
- [35] R.W.P. Urias, S.J. Barigye, Y. Marrero-Ponce, C.R. García-Jacas, J.R. Valdes-Martini, and F. Perez-Gimenez, *IMMAN: Free software for information theory-based chemometric analysis*, Mol. Divers. 19 (2015), pp. 305–319. doi:10.1007/s11030-014-9565-z. 580
- [36] *MATLAB MathWorks Inc R2017b 9.3.0.713579*, MathWorks Inc, USA, 2017. software. Available at <https://www.mathworks.com/>.
- [37] *STATISTICA StatSoft Inc v8.0*. StatSoft Inc, USA, 2007. software. Available at <http://www.statsoft.com>.
- [38] J. Jaworska and N. Nikolova-Jeliazkova, *How can structural similarity analysis help in category formation?* SAR QSAR Environ. Res. 18 (2007), pp. 195–207. doi:10.1080/10629360701306050. 585
- [39] S. Kim, P.A. Thiessen, E.E. Bolton, J. Chen, G. Fu, A. Gindulyte, L. Han, J. He, S. He, B. A. Shoemaker, J. Wang, B. Yu, J. Zhang, and S.H. Bryant, *PubChem substance and compound databases*, Nucleic Acids Res. 44 (2016), pp. D1202–D1213. doi:10.1093/nar/gkv951. 590
- [40] S.L. Dixon, A.M. Smondryev, and S.N. Rao, *PHASE: A novel approach to pharmacophore modeling and 3D database searching*, Chem. Biol. Drug Des. 67 (2006), pp. 370–372. doi:10.1111/j.1747-0285.2006.00384.x.
- [41] J.F. Truchon and C.I. Bayly, *Evaluating virtual screening methods: Good and bad metrics for the “early recognition” problem*, J. Chem. Inf. Model. 47 (2007), pp. 488–508. doi:10.1021/ci600426e. 595

- [42] M.M. Mysinger, M. Carchia, J.J. Irwin, and B.K. Shoichet, *Directory of useful decoys, enhanced (DUD-E): Better ligands and decoys for better benchmarking*, *J. Med. Chem.* 55 (2012), pp. 6582–6594. doi:10.1021/jm300687e.
- [43] T. Hamers, J.H. Kamstra, E. Sonneveld, A.J. Murk, M.H.A. Kester, P.L. Andersson, J. Legler, and A. Brouwer, *In vitro profiling of the endocrine-disrupting potency of brominated flame retardants*, *Toxicol. Sci.* 92 (2006), pp. 157–173. doi:10.1093/toxsci/kfj187.
- [44] C. Sammut and G. Webb (eds.), *Encyclopedia of Machine Learning and Data Mining*, 2nd ed., Vol. 32, Springer, New York, NY, 2018.
- [45] M. Ghorbanzadeh, K.I. Van Ede, M. Larsson, M.B.M. Van Duursen, L. Poellinger, S. Lücke-Johansson, M. Machala, K. Pěňčíková, J. Vondráček, M. van den Berg, M.S. Denison, T. Ringsted, and P.L. Andersson, *In vitro and in silico derived relative effect potencies of ah-receptor-mediated effects by PCDD/Fs and PCBs in rat, mouse, and guinea pig CALUX cell lines*, *Chem. Res. Toxicol.* 27 (2014), pp. 1120–1132. doi:10.1021/tx5001255.
- [46] OECD, *Guidance Document on the Validation of (Quantitative) Structure-Activity Relationship [(Q)SAR] Models*, 2014. OECD Series on Testing and Assessment/OECD.
- [47] O.F. Güner and J.P. Bowen, *Setting the record straight: The origin of the pharmacophore concept*, *J. Chem. Inf. Model.* 54 (2014), pp. 1269–1283. doi:10.1021/ci5000533.
- [48] D.P. Williams, *Toxicophores: Investigations in drug safety*, *Toxicology* 226 (2006), pp. 1–11. doi:10.1016/j.tox.2006.05.101.
- [49] P.K. Singh, A. Negi, P.K. Gupta, M. Chauhan, and R. Kumar, *Toxicophore exploration as a screening technology for drug design and discovery: Techniques, scope and limitations*, *Arch. Toxicol.* 90 (2016), pp. 1785–1802. doi:10.1007/s00204-015-1587-5.
- [50] L. Chen, D. Wu, H.P. Bian, G.L. Kuang, J. Jiang, W.H. Li, G.X. Liu, S. Zou, J. Huang, and Y. Tang, *Selective ligands of estrogen receptor β discovered using pharmacophore mapping and structure-based virtual screening*, *Acta Pharmacol. Sin.* 35 (2014), pp. 1333–1341. doi:10.1038/aps.2014.69.
- [51] S.S. Ashtekar, N.M. Bhatia, and M.S. Bhatia, *Development of leads targeting ER- α in breast cancer: An in silico exploration from natural domain*, *Steroids* 131 (2018), pp. 14–22. doi:10.1016/j.steroids.2017.12.016.
- [52] N. Lagarde, S. Delahaye, A. Jérémie, N. Ben Nasr, H. Guillemain, C. Empereur-Mot, V. Laville, T. Labib, M. Réau, F. Langenfeld, J.F. Zagury, and M. Montes, *Discriminating agonist from antagonist ligands of the nuclear receptors using different chemoinformatics approaches*, *Mol. Inf.* 36 (2017), pp. 1–16. doi:10.1002/minf.201700020.
- [53] R.T. Engeli, S.R. Rohrer, A. Vuorinen, S. Herdinger, T. Kaserer, S. Leugger, D. Schuster, and A. Odermatt, *Interference of paraben compounds with estrogen metabolism by inhibition of 17 β -hydroxysteroid dehydrogenases*, *Int. J. Mol. Sci.* 18 (2017). doi:10.3390/ijms18092007.
- [54] I.T. Jolliffe, *Principal Component Analysis*, 2nd ed., Vol. 98, Springer-Verlag, New York, NY, 2002.
- [55] K. Schomburg, H. Ehrlich, K. Stierand, and M. Rarey, *From structure diagrams to visual chemical patterns*, *J. Chem. Inf. Model.* 50 (2010), pp. 1529–1535. doi:10.1021/ci100209a.
- [56] L. Poellinger and D. Gullberg, *Characterization of the hydrophobic properties of the receptor for 2,3,7,8-tetrachlorodibenzo-p-dioxin*, *Mol. Pharmacol.* 27 (1985), pp. 271–276.
- [57] T. Tripathi and A.K. Saxena, *2D-QSAR studies on new stilbene derivatives of resveratrol as a new selective aryl hydrocarbon receptor*, *Med. Chem. Res.* 17 (2008), pp. 212–218. doi:10.1007/s00044-007-9055-2.
- [58] Y. Xing, M. Nukaya, K.A. Satyshur, L. Jiang, V. Stanevich, E.N. Korkmaz, L. Burdette, G. D. Kennedy, Q. Cui, and C.A. Bradfield, *Identification of the Ah-receptor structural determinants for ligand preferences*, *Toxicol. Sci.* 129 (2012), pp. 86–97. doi:10.1093/toxsci/kfs194.
- [59] K. Bekki, H. Takigami, G. Suzuki, N. Tang, and K. Hayakawa, *Evaluation of toxic activities of polycyclic aromatic hydrocarbon derivatives using in vitro bioassays*, *J. Heal. Sci.* 55 (2009), pp. 601–610. doi:10.1248/jhs.55.601.
- [60] C. Gu, X. Ju, X. Jiang, K. Yu, S. Yang, and C. Sun, *Improved 3D-QSAR analyzes for the predictive toxicology of polybrominated diphenyl ethers with CoMFA/CoMSIA and DFT*, *Ecotoxicol. Environ. Saf.* 73 (2010), pp. 1470–1479. doi:10.1016/j.ecoenv.2009.11.003.

- [61] C.R. Martinez and B.L. Iverson, *Rethinking the term "pi-stacking,"* *Chem, Sci* 3 (2012), pp. 2191–2201. 650
- [62] R.C. Kolanczyk, J.S. Denny, B.R. Sheedy, P.K. Schmieder, and M.A. Tapper, *Estrogenic activity of multicyclic aromatic hydrocarbons in rainbow trout (Oncorhynchus mykiss) in vitro assays,* *Aquat. Toxicol.* (2019), pp. 43–51. doi:10.1016/j.aquatox.2018.11.023.
- [63] C. Yun, J.A. Weiner, D.S. Chun, J. Yun, R.W. Cook, M.S. Schallmo, A.S. Kannan, S.M. Mitchell, R. D. Freshman, C. Park, W.K. Hsu, and E.L. Hsu, *Mechanistic insight into the effects of Aryl hydrocarbon receptor activation on osteogenic differentiation,* *Bone Rep.* 6 (2017), pp. 51–59. 655
doi:10.1016/j.bonr.2017.02.003.
- [64] C.A. McQueen (ed.), *Comprehensive Toxicology*, 3rd ed., Vol. 35, Elsevier, Oxford, UK, 2018.

PROOF ONLY