Journal of Hazardous Materials 190 (2011) 106–112

Contents lists available at ScienceDirect

Journal of Hazardous Materials

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

QSAR classification models for the prediction of endocrine disrupting activity of brominated flame retardants

Simona Kovarich, Ester Papa*, Paola Gramatica

QSAR Research Unit in Environmental Chemistry and Ecotoxicology, DBSF, University of Insubria, Via J.H. Dunant 3, 21100 Varese, Italy

ARTICLE INFO

Article history:
Received 16 November 2010
Received in revised form 1 March 2011
Accepted 2 March 2011
Available online 9 March 2011

Keywords:
QSAR
Brominated flame retardants
Endocrine disruptors
SVHC
REACH

ABSTRACT

The identification of potential endocrine disrupting (ED) chemicals is an important task for the scientific community due to their diffusion in the environment; the production and use of such compounds will be strictly regulated through the authorization process of the REACH regulation. To overcome the problem of insufficient experimental data, the quantitative structure–activity relationship (QSAR) approach is applied to predict the ED activity of new chemicals. In the present study QSAR classification models are developed, according to the OECD principles, to predict the ED potency for a class of emerging ubiquitary pollutants, viz. brominated flame retardants (BFRs). Different endpoints related to ED activity (i.e. aryl hydrocarbon receptor agonism and antagonism, estrogen receptor agonism and antagonism, androgen and progesterone receptor antagonism, T4-TTR competition, E2SULT inhibition) are modeled using the *k*-NN classification method. The best models are selected by maximizing the sensitivity and external predictive ability. We propose simple QSARs (based on few descriptors) characterized by internal stability, good predictive power and with a verified applicability domain. These models are simple tools that are applicable to screen BFRs in relation to their ED activity, and also to design safer alternatives, in agreement with the requirements of REACH regulation at the authorization step.

1. Introduction

Increasing concern is being shown by the scientific community, regulators and the public about endocrine-disrupting chemicals (EDCs) that, in the environment, are adversely affecting human and wildlife health. There are different mechanisms through which these chemicals can exert their effects on the endocrine system: (i) agonistic effect by binding to the cellular receptor of a hormone, activating normal cell response at the wrong time or to an excessive extent; (ii) antagonistic effect by binding to the receptor, preventing natural hormonal binding and activation of the receptor; (iii) alteration of hormonal blood levels by binding to hormone transport proteins; (iv) interference with metabolic processes by affecting the synthesis, or elimination rate, of hormones. All these can lead to alterations in the maintenance of homeostasis, and in the reproduction, development and behaviour of the organism [1]. In the EU REACH regulation [2], endocrine disrupting chemicals are included in Title VII (Article 57-f), which deals with the authorization of substances of very high concern (SVHC).

Among the suspected EDCs, brominated flame retardants (BFRs) are an emerging class of ubiquitary pollutants that can act as endocrine disrupters.

BFRs are industrial products incorporated into combustible materials, such as plastics, wood and textiles, to increase their fire resistance. Brominated flame retardants include a structurally heterogeneous group of chemicals, and, among these, the most commercialized are polybrominated diphenyl ethers (PBDEs), tetrabromobisphenol A (TBBPA) and hexabromocyclododecane (HBCD). The wide dispersion of BFRs in the environment, their high lipophilicity, persistence and bioaccumulation potential, has led

Abbreviations: ED, endocrine disrupting; REACH, Registration, Evaluation, Authorization and Restriction of Chemicals; QSAR, quantitative structure–activity relationship; OECD, Organization for Economic Cooperation and Development; BFRs, brominated flame retardants; T4-TTR, thyroxin-transthyretin; E2SULT, estradiolsulfotransferase; *k*-NN, *k*-nearest neighbor; SVHC, substances of very high concern; EDC, endocrine disrupting chemicals; PBDEs, polybrominated diphenyl ethers; TBBPA, tetrabromobisphenol A; HBCD, hexabromocyclododecane; AhR, Aryl hydrocarbon receptor; RBA, AhR relative binding affinity; OH-PBDE, hydroxylated PBDE; 246-TBP, 2,4,6-tribromophenol; TBBPA-DBPE, tetrabromobisphenol-A-bis(2,3)dibromopropyl ether; DR_{ag}, AhR agonism; DR_{ant}, AhR antagonism; ER_{ag}, estrogen receptor agonism; ER_{ant}, estrogen receptor antagonism; AR_{ant}, androgen receptor antagonism; PR_{ant}, progesterone receptor antagonism; $T4\text{-}TTR_{comp}$, T4-TTR competing potency; $E2SULT_{inh}$, E2SULT inhibiting potency; CH_3O-PBDE, methoxylated PBDE; DBDE, decabromodiphenyl ethane; EBTPI, ethylene bistetrabromo phthalimide; TBE, 1,2-bis(2,4,6-tribromophenoxy) ethane; NER, non error rate; Sn, sensitivity; Sp, specificity; TP, true positive; TN, true negative; FP, false positive; FN, false negative; NER_{EXT}, external non error rate; AD, applicability domain; TSET, training set; PSET, prediction set; PBP, pentabromophenol; TCDD, 2,3,7,8-tetrachloro-dibenzo-p-dioxin; PCB, polychlorinated biphenyls; E2, estradiol; DHT, dihydrotestosterone; MPA, medroxyprogesterone acetate; MLR, multilinear regression.

* Corresponding author.
E-mail address: ester.papa@uninsubria.it (E. Papa).

0304-3894/$ – see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.jhazmat.2011.03.008

to increasing concentrations in wildlife and humans [3–6]. Thus, a better understanding of the risk represented by these emerging pollutants is required.

Experimental evidence shows that BFRs are endocrine-active compounds with the potential to interfere with thyroid hormone homeostasis, as well as to interact with steroid receptors (e.g. estrogens, androgens) and aryl hydrocarbon receptors (dioxin-like-activity) [7–12].

Parallel with experimental studies, *in silico* strategies like QSARs (quantitative structure–activity relationships) represent an important tool to fill the gap of information on BFRs. In fact, QSAR models, recommended for use under REACH regulation, can be applied to predict lacking experimental data and to screen and prioritize chemicals, thus reducing costs and the number of tested animals. Furthermore, QSAR approaches can be successfully applied in procedures of "safe Chemical Design" as in the Drug Design process. In fact, safe molecule design is the earliest phase in the long process of placement of new safe substances onto the market. To date, several QSARs and 3D-QSARs predicting ED potency of BFRs have been published, most of them being regression models (linear and non-linear) for AhR relative binding affinity (RBA), anti-androgenic and anti-estrogenic activity [13–19].

Furthermore, the development and application of *in silico* approaches is being financially supported by the European Commission, through the 7th Framework Programme for Research, in order to predict lacking experimental data as well as to perform risk assessment of four classes of compounds of interest, including, among others, BFRs (CADASTER FP 7 PROJECT [20]). In this context, the present study has developed, according to the OECD principles [21], classification QSARs for different endpoints related to brominated flame retardant ED activity. The models were built on small and heterogeneous data sets, and were applied to predict the activity of 243 BFRs, including three alternatives to BFRs, listed in the EU-regulations, for which no experimental data are yet available.

2. Materials and methods

2.1. Data sets and classes

The experimental data sets, obtained from two studies of Hamers and co-workers [22,23], include a heterogeneous group of 29 brominated flame retardants, in particular some PBDEs and hydroxy-BDE congeners (OH-PBDEs), TBBPA, 2,4,6-tribromophenol (246-TBP), HBCDγ, and tetrabromobisphenol A-bis(2,3)dibromopropyl ether (TBBPA-DBPE). The modeled endpoints are Aryl hydrocarbon (dioxin) Receptor agonism (DR_{ag}) and antagonism (DR_{ant}), Estrogen Receptor agonism (ER_{ag}) and antagonism (ER_{ant}), Androgen Receptor antagonism (AR_{ant}), Progesterone Receptor antagonism (PR_{ant}), T4-TTR Competing Potency ($T4\text{-}TTR_{comp}$) and E2SULT Inhibiting Potency ($E2SULT_{inh}$).

The homogeneous data sets used in our study are the result of an extended literature search specifically focused on ED properties of PBDEs and BFRs. Taking into account the complexity of the endpoints considered in this study, the decision to use only experimental data measured by one research group was made in order to guarantee a better quality and homogeneity of the input data, which were used for the development of our QSARs. In fact it is known that, mainly in case of small data sets, the use of heterogeneous experimental data from different sources and laboratories can affect the quality of QSAR models, by increasing the noise in the modeled response.

The definition of the classes of activity was based on the classification criteria proposed by Hamers and collaborators [22]. Due to the limited amount of data available for the levels of potency, from low to very high, suggested in literature [22], only binary classification models could be developed for the endpoints DR_{ag}, DR_{ant}, ER_{ag}, ER_{ant}, AR_{ant} and PR_{ant}, whose experimental data were available for 24 compounds (Class 1 = inactive (no ED potency) and Class 2 = active (any evidence of ED potency)). Three classes of ED potency were modeled for the endpoints $T4\text{-}TTR_{comp}$ and $E2SULT_{inh}$, for which a higher number of experimental data ($n_{obj} = 29$) were available (Class 1 = inactive (no ED potency), Class 2 = moderately active (low/moderate ED potency) and Class 3 = very active (high/very high ED potency)) (Table 1).

The developed models were then applied to predict the unknown ED potency for the remaining 209 PBDE congeners, several PBDE metabolites (OH-PBDEs and CH_3O-PBDEs), brominated phenols, brominated bisphenol A compounds (TBBPA analogs) and other BFRs on the market, including three alternative compounds to decaBDE, already listed in other regulations (i.e. decabromodiphenyl ethane – DBDE; ethylene bistetrabromo phthalimide – EBTPI; 1,2-bis(2,4,6-tribromophenoxy) ethane – TBE) [19]. The predicted classes of ED potency for all the BFRs considered in this study are available as Supplementary Data (Table S1).

2.2. Calculation of molecular descriptors

The chemical structures of BFRs were drawn using the Semi-empirical method AM1 in the HYPERCHEM program (ver. 7.03 for Windows, 2002) and were used as input files for descriptor calculations. The molecular descriptors, which lead to information on the mono-, bi- and tri-dimensional structure of the chemicals, were computed by the software DRAGON [24]. In a preliminary step, constant or near-constant values and descriptors with a high pair-wise correlation were excluded to reduce redundant and non-useful information. At the end of this procedure a final set of 701 descriptors was used as input variables in the model development.

2.3. QSAR modeling

Classification models quantify the relationship between one or more independent variables (the molecular descriptors) and a qualitative response variable, each representing the class of the corresponding sample (here the classes of ED potency). The classification model predicts the assignment of new compounds, for which the class is unknown, to one of the *a priori* defined classes. The *k*-nearest neighbor (*k*-NN) method was applied to predict the classes of ED potency. This classification method, based on the similarity of objects (chemicals), searches for the *k* nearest neighbors of each object in the data set. The assignment of a compound to a class is based on the class of the *k* most similar compounds, where similarity is defined by calculating the Euclidean distances between the descriptor vectors. The *k*-NN method was then applied to autoscaled data and the *a priori* probability of belonging to a class was set as proportional to the number of chemicals in the *a priori* classes of ED potency. The predictive power of the model was checked for *k* values between 1 and 10.

Due to the small dimensions of the training sets, we decided to take into account only models based on a maximum of two descriptors. Thus, all the mono- and bi-dimensional models from the 701 calculated molecular descriptors (all the possible combinations by the *All Subset Models* selection method, using in-house software) were explored by maximizing the overall percentage of correct assignments (percentage of non error rate – NER%) and the population of the best 100 models was analysed for each modeled endpoint. To compare the performances of the *k*-NN models selected in the population, NER% was also calculated separately for each class of activity [25].

Moreover, parameter sensitivity (Sn) and specificity (Sp) were calculated for the endpoints DR_{ag}, DR_{ant}, ER_{ag}, ER_{ant}, AR_{ant} and PR_{ant}

Table 1
Classification criteria proposed by Hamers and collaborators [19] for BFRs and classes modeled in this work for each end-point.

Hamers class [22]	Criteria	DR_{ag}, DR_{ant}, ER_{ag}, ER_{ant}, AR_{ant}, PR_{ant}	T4-TTR_{comp}, E2$SULT_{inh}$
No potency	Response < 20% of control at 10 μM	Class 1 (inactive)	Class 1 (inactive)
Low potency	$E(I)C_{50}$ > 10 μM and response > 20% of control	Class 2 (active)	Class 2 (moderately active)
Moderate potency	1.0 μM < $E(I)C_{50}$ < 10 μM	Class 2 (active)	Class 2 (moderately active)
High potency	0.1 μM < $E(I)C_{50}$ < 1.0 μM	Class 2 (active)	Class 3 (very active)
Very high potency	0.01 μM < $E(I)C_{50}$ < 0.1 μM	Class 2 (active)	Class 3 (very active)

(for which 2 classes of activity were defined according to Table 1):

$$Sn = \frac{TP}{TP + FN} \qquad Sp = \frac{TN}{TN + FP}$$

where TP (true positive) is the number of compounds correctly classified as active, TN (true negative) is the number of compounds correctly classified as inactive, FN (false negative) is the number of active compounds classified as inactive, and FP (false positive) is the number of inactive compounds classified as active [26].

Parameters Sn and Sp were calculated as reported above also for the endpoints T4-TTR_{comp} and E2$SULT_{inh}$, for which 3 classes of activity were defined according to Table 1. Calculations were performed after grouping into the class "active" chemicals belonging to Class 2 (moderately active) and 3 (very active).

As a precautionary principle, the misclassification of active compounds as inactive (false negative) should be considered a much greater error than the misclassification of inactive compounds as active (false positive). For this reason the number of false negatives was minimized for the selection of the best models.

2.4. External validation

In order to verify the real predicting power of the proposed classification models, external validation was performed. For each endpoint the available experimental data set was preliminarily split into a training set, which was used to develop the model, and a prediction set, which was used only to validate the model. The splitting of the data sets was performed by random selection of the test objects (nearly 30%) within each class.

Models developed on the training set were then tested for their predictive capability by calculating prediction accuracy on the prediction set (NER_{EXT}%).

2.5. Chemical applicability domain

The developed classification models were applied to predict ED potency for 243 brominated flame retardants without experimental data. The definition of the applicability domain (AD) of a model allows for the evaluation of the degree of extrapolation in prediction, especially when few chemicals are used to develop the model. In this work two approaches were used for the evaluation of the structural applicability domain. One approach was based on the range of descriptors selected in each model; the other was based on compound similarity to the training set.

In accordance with the first method, chemicals with descriptor values within the range of those of the training set compounds were considered as being inside the AD of the model. Compounds falling outside the descriptors' space were considered as structural outliers (beyond the AD of the model).

The second method is based on the calculation of Euclidean distance and it was performed by the software ToxMatch [27]. For each class, compounds having a Euclidean distance higher than the training set were considered as structural outliers (beyond the AD of the model).

Predictions of compounds lying outside the structural domain of the proposed models were considered as extrapolations, thus less reliable.

3. Results and discussion

The *k*-NN method was applied to model the proposed classes of ED potency for brominated flame retardants. Models were first developed on training sets (TSET), generated by random splitting (30%) of the available experimental data, and then validated for their external predictivity on the prediction sets (PSET). The models proposed here were selected from among a population of 100 models generated by the *All Subset* method, on the basis of accuracy in prediction, external predictivity, and interpretability of descriptors. The best variables selected for each end-point were finally used to model all the available experimental data (Full models). This procedure, which considers all the available structural and experimental information, led to an extension of the applicability domain of most of the models. The results of the classification models developed for the studied end-points are shown in Table 2 (Split models) and Table 3 (Full models). Details of the calculation of the theoretical molecular descriptors selected in the proposed models are reported as Supplementary Data (Table S2).

From an analysis of the Split models we were able to verify the external predictivity of the developed models, which were characterized by NER_{EXT}% values in the range of 87–100%.

As can be observed in Table 3, the QSARs (Full models) developed for the endpoints DR_{ag}, DR_{ant}, ER_{ag}, ER_{ant} and AR/PR_{ant} show high classification performance, with a sensitivity of always 100% (no FN) and a specificity ranging from 87 to 100%. According to the models developed for T4-TTR competition and E2SULT inhibition, the classification accuracy within each class is high, 83–100%. Note that more importance was given to the ability of the models to correctly classify compounds with high ED potency. The behaviour of sensitivity and specificity in dependence of *k* index was verified for all the Full Models. Results of this analysis and the respective ROC graphs have been added as Supplementary Data (Figure S1).

3.1. Classification models for dioxin- and anti-dioxin-like activity

The modeling variables for DR_{ag} (dioxin-like activity through the aryl hydrocarbon receptor – AhR) were F04[O–Br], a bidimensional descriptor that counts the frequency of bonds O–Br at topological distance 04 (number of Br in *meta* position) [24], and RDF055v, a tridimensional descriptor from Radial Distribution Function descriptors weighted by atomic van der Waals volumes [28]. Applying the model in prediction to all the BFRs considered in the study, the majority of mono- to hexa-PBDEs with one, two or three Br in *meta* position were predicted as active, as well as PBDE metabolites with Br in *meta* and pentabromophenol (PBP). The presence of one to three bromine substituents in *meta* position (1 < F04[O–Br] < 3) was identified as a relevant condition for BFRs to interact and activate the AhR. Indeed, higher brominated diphenyl ethers and other BFRs, such as TBBPA analogs and deca-BDE alternatives, without Br in *meta* position or with all the *meta* positions occupied (F04[O–Br] = 0 or 4), were predicted as inactive. Although the simple descriptor F04[O–Br] was, alone, able to classify the dioxin-like activity of most of the compounds (BFRs with F04[O–Br] = 0 or 4 were all inactive, BFRs with F04[O–Br] = 1 were all active), but the tridimensional information added by RDF055v was

Table 2
Results of the Split models developed for the studied end-points.

Endpoint	Descriptors	*k*	Real class	N° compounds	Assigned class			NER_{class}%	NER%	Sn	Sp
					1	2	3				
DR agonism											
TSET	F04[O–Br] RDF055v	4	1	10	8	2		80	81.3	0.83	0.8
			2	6	1	5		83.3			
PSET	F04[O–Br] RDF055v	4	1	5	5	0		100	100	1	1
			2	3	0	3		100			
DR antagonism											
TSET	Jhetm BEHm7	1	1	10	8	2		80	81.3	0.83	0.8
			2	6	1	5		83.3			
PSET	Jhetm BEHm7	1	1	5	4	1		80	87.5	1	0.8
			2	3	0	3		100			
ER agonism											
TSET	Ms BEHv7	1	1	11	10	1		90.9	93.8	1	0.91
			2	5	0	5		100			
PSET	Ms BEHv7	1	1	5	5	0		100	100	1	1
			2	3	0	3		100			
ER antagonism											
TSET	QW nArOH	1	1	11	10	1		90.9	87.5	0.8	0.91
			2	5	1	4		80			
PSET	QW nArOH	1	1	5	4	1		80	87.5	1	0.8
			2	3	0	3		100			
AR/PR antagonism											
TSET	GGI8	1	1	3	3	0		100	100	1	1
			2	13	0	13		100			
PSET	GGI8	1	1	2	2	0		100	100	1	1
			2	6	0	6		100			
T4-TTR competition											
TSET	DISPe nArOH	3	1	8	7	1	0	87.5	95	1	0.87
			2	6	0	6	0	100			
			3	6	0	0	6	100			
PSET	DISPe nArOH	3	1	4	4	0	0	100	88.9	0.8	1
			2	3	1	2	0	66.7			
			3	2	0	0	2	100			
E2SULT inhibition											
TSET	Mor21v qnmax	1	1	6	5	1	0	83.3	75	0.86	0.83
			2	8	2	5	1	62.5			
			3	6	0	1	5	83.3			
PSET	Mor21v qnmax	1	1	2	2	0	0	100	100	1	1
			2	4	0	4	0	100			
			3	3	0	0	3	100			

Table 3
Results of the Full models developed for the studied end-points.

Endpoint	Descriptors	*k*	Real class	N° compounds	Assigned class			NER_{class}%	NER%	Sn	Sp
					1	2	3				
DR agonism	F04[O–Br] RDF055v	4	1	15	14	1		93.3	95.8	1	0.93
			2	9	0	9		100			
DR antagonism	Jhetm BEHm7	1	1	15	13	2		86.7	91.7	1	0.87
			2	9	0	9		100			
ER agonism	Ms BEHv7	1	1	16	15	1		93.8	95.8	1	0.94
			2	8	0	8		100			
ER antagonism	QW nArOH	1	1	16	15	1		93.8	95.8	1	0.94
			2	8	0	8		100			
AR/PR antagonism	GGI8	1	1	5	5	0		100	100	1	1
			2	19	0	19		100			
T4-TTR competition	DISPe nArOH	3	1	12	10	2	0	83.3	89.7	0.94	0.83
			2	9	1	8	0	88.9			
			3	8	0	0	8	100			
E2SULT inhibition	Mor21v qnmax	1	1	8	8	0	0	100	89.7	0.95	1
			2	12	1	10	1	83.3			
			3	9	0	1	8	88.9			

needed to discriminate among BFRs having two and three bromines in the *meta* position for active and non active compounds.

Two bidimensional descriptors were selected as modeling variables for DR_{ant} (anti-dioxin-like activity through AhR in the presence of TCDD): Jhetm (Balaban-type index from mass weighted distance matrix), a topological descriptor related to both the number of Br and the substitution pattern (values of Jhetm increase with increasing bromination degree and with the presence of substituents in *ortho* position) [29], and BEHm7 (highest eigenvalue n. 7 of Burden matrix, weighted by atomic masses), among the Burden eigenvalues descriptors [30,31]. The trend of activity predicted by the DR_{ant} model was mainly based on the descriptor Jhetm: BFRs with low descriptor values, which were also low/moderately brominated with substituents in *meta*/*para* positions, were classified as active. Thus, the developed QSAR classified, as active, the mono- to penta-BDEs with Br in *meta*/*para*, OH-PBDEs with OH– in *meta*/*para*, small bisphenols, HBCD and the alternative TBE. This is in agreement with what has already been observed in the literature [11,12,19] where lower brominated congeners with few or without *ortho* substituents, analogous to dioxins and coplanar PCBs, show higher binding affinity with the dioxin receptor AhR.

From these results we can say that the ability to interact with AhR, as agonist or antagonist, is associated with low/moderate brominated diphenyl ethers and their metabolites with substituents in the *meta* or *meta*/*para* positions. Thus, in many cases the same compound shows both DR agonism and antagonism. This reflects the activity trend of the experimental data [22].

3.2. Classification models for estrogenic and anti-estrogenic activity

The *k*-NN model developed for ER_{ag} (estrogenic activity through estrogen receptor ER) was based on two bidimensional descriptors: Ms (mean electropological state), a constitutional descriptor whose values increase with bromination degree (constant values within the same homologous group of PBDEs) [24], and BEHv7 (highest eigenvalue n. 7 of Burden matrix) from the Burden eigenvalue descriptors, weighted by atomic van der Waals volumes, which is mainly related to dimension and bromination degree [30,31]. The model classification of active compounds was assigned mainly to the lower brominated diphenyl ethers (di- to tetra-BDEs) with BEHv values in the range of 1.6–1.74. Among these congeners, the bromine substitution pattern [2,2′,6] and [2,2′,4] occurred frequently, confirming the experimental evidence of Hamers [22]. In contrast to results by Meerts [9], the OH-PBDEs and TBBPA analogs were classified as inactive, as also was the well-known environmental estrogen bisphenol A. This can be explained by the experimental information included in our dataset for OH-PBDEs and brominated bisphenol A analogs, which was limited to 6OH-BDE-47 and TBBPA, both measured by Hamers as inactive compounds [22]. This information was different from the basic experimental data published by Meerts [9], where more OH-BFRs were reported as active. However it should be noted that our predictions for OH-PBDEs are in line with what was observed for OH-PCBs, whose minimal estrogenic activity has been reported in the literature [32].

Two very simple descriptors were selected as modeling variables for ER_{ant} (antiestrogenic activity through ER in the presence of estradiol E2): (i) QW (quasi-Wiener index – Kirchhoff number), a topological descriptor related to chemical dimension, degree of substitution and, for PBDEs, to the distance of bromine substituents from the oxygen ether [33]; (ii) nArOH, which counts the number of aromatic hydroxyls. The BFRs predicted as anti-estrogens were the highly brominated diphenyl ethers (hepta- and octa-PBDEs), OH-PBDEs, OH-bromo phenols and HBCD. Our predictions for OH-PBDEs were again in line with what has been observed for OH-PCBs, for which antiestrogenic activity has been found in several *in vitro* bioassays [32].

According to experimental observations [9,22], estrogenic activity is mainly associated with lower-brominated PBDEs, and antiestrogenic activity with higher-brominated PBDEs.

3.3. Classification models for anti-androgenic and anti-progestagenic activity

The BFRs tested in the study of Hamers [22] had identical class assignment for both AR_{ant} (anti-androgenic activity through androgen receptor, AR, in the presence of dihydrotestosterone, DHT) and PR_{ant} (anti-progestagenic activity through progesterone receptor, PR, in the presence of medroxyprogesterone acetate, MPA). Thus, the classification model was identical for both the responses. An analysis of the experimental data set revealed that both AR and PR antagonism were the most sensitive endpoints to BFR activity. Among the 24 tested compounds, 19 BFRs were observed to have endocrine disrupting activity ranging from low to very high.

The QSAR proposed here for the two endpoints is a very simple classification model based on one descriptor, GGI8 (topological charge index of order 8), which is a bidimensional descriptor related to both the number and the position of the Br substituents on the phenyl rings [34]. For all the studied BFRs, the values of this descriptor increases according to the degree of bromination and, for the PBDEs, also on the basis of the distance of the bromine substituents from the oxygen ether (maximum values for *meta*- and *para*-substitutions).

The model predicted almost all the BFRs considered in this study as potential endocrine disruptors (GGI8 < 0.21), the exception being the more stretched molecules with all the *meta* and *para* positions substituted (i.e. PBDEs with [3,3′,4,4′,5,5′] substitution pattern, TBBPA-derivates and deca-BDE alternatives). The fact that *meta* and *para* Br substitutions were unfavorable for AR antagonism by PBDEs was also confirmed in the molecular docking study of Yang and coworkers [17].

3.4. Classification models for T4-TTR competing potency and E2SULT inhibition potency

The availability of more experimental data for the endpoints T4-TTR competition (displacement of thyroid hormone T4, thyroxine, from its plasma transport protein TTR, transthyretin) and E2SULT inhibition (inhibited sulfation of E2, estradiol) allowed us to develop *k*-NN models based on the three classes of ED potency reported in Table 1: (1) inactive, (2) moderately active, (3) very active.

The modeling variables selected for the T4-TTR competition were: DISPe (d COMMA2 value, weighted by atomic Sanderson electronegativities), a geometrical descriptor that describes structural symmetry (lower values for more symmetrical molecules and vice versa) [35], and nArOH, which counts the number of aromatic hydroxyls. On analyzing the predictions obtained for all the BFRs, 75% of the PBDEs were classified as moderately active (most of them were characterized by an asymmetric distribution of Br substituents in the phenyl rings), while all the BFRs containing an aromatic –OH group (i.e. OH-PBDEs, brominated phenols and bisphenols A compounds) were classified as very active. The higher TTR binding affinity of hydroxylated BFRs, even exceeding that of the natural ligand T_4, has already been documented in the literature [8,22,23] and can be explained by their structural resemblance to the hormone T_4.

Two tridimensional descriptors were used to model the inhibition of the enzyme E2SULT: Mor21v (3D-MoRSE – signal 21, weighted by atomic van der Waals volumes), among the 3D-

MoRSE descriptors, which is partially related to bromination degree (higher values for higher brominated diphenyl ethers) [36], and qnmax (maximum negative charge), a charge descriptor influenced by polarity difference induced by aromatic –OH groups and C=O groups (low values of the descriptor) [35]. Applying the model to all the studied BFRs, we could observe that nearly 50% of the PBDEs were predicted as moderately active, even though the class assignment according to their substitution pattern was not very clear. Moderate activity was also predicted for methoxylated BFRs (i.e. CH_3O-PBDEs and 2,4,6-tribromoanisole), TBBPA-diallyl ether and the alternative TBE. As for T4-TTR competition, E2SULT inhibiting potency is higher for aromatic hydroxylated compounds (i.e. OH-PBDEs, brominated phenols and bisphenols A compounds), but also for the alternative EBTPI, characterized by the presence of four carbonyl groups.

3.5. Analysis of the applicability domain

Two different approaches, one based on the range of descriptors selected in each model, and the other based on compounds similarity, were applied, and compared to evaluate the applicability domain of the proposed models. This procedure verifies the reliability of predictions, calculated by the models developed on specific training sets and structural domains, for all the studied 243 BFRs (209 PBDE congeners in addition to 34 structurally heterogeneous BFRs, including several OH- and CH_3O-BDEs, TBBPA analogs, bromo-phenols and three decaBDE alternatives: DBDE, EBTPI and TBE). Values predicted for compounds identified as distant from the structural domain of the training sets were considered to be extrapolations, and hence less reliable.

From this analysis we could verify that the majority of the studied BFRs fell in the AD of the models (FULL models, Table 3), with a percentage of reliable predictions ranging from 87 to 99%.

More in detail, the following compounds fell outside the structural AD of individual models: BDE-5, BDE-154, 6-CH_3O-BDE-47, 4-bromophenol, PBP and TBBPA-diallylether (DR_{ag} model); mono-BDEs and di-BDEs (with Br in *meta*- and *para*-positions), bisphenol-A, mono- and di-bromobisphenol-A, 4-bromophenol, pentabromophenol, 4-phenoxyphenol, hexabromobenzene, EBTPI and TBE (DR_{ant} model); some di-BDEs, BDE-121, 4-bromophenol and EBTPI (ER_{ag} model); 4-bromophenol (ER_{ant} model); EBTPI (AR/PR_{ant} model); 16 PBDEs (from tri- to hepta-BDEs), tribromobisphenol-A, hexabromobenzene and DBDE (T4-TTR_{comp} model); finally, 22 PBDEs (from di- to octa-BDEs), 2′OH-BDE-68, 4-phenoxyphenol, hexabromobenzene, bisphenol-A and monobromobisphenol-A, TBBPA-diallylether, EBTPI and DBDE ($E2SULT_{inh}$ model). Additional information is given as Supplementary Data (Tables S3, S4; Figures S2–S7).

3.6. Activity profile of deca-BDE alternatives and final considerations

Great attention was paid to the predictions obtained for the three decaBDE alternatives considered in this study. While DBDE showed no activity for all the studied endpoints, TBE was predicted as active for DR_{ag} and moderately active for $E2SULT_{inh}$, and EBTPI was predicted as very active for $E2SULT_{inh}$. These findings are in reasonable agreement with results obtained by MLR models recently published by our research group [19].

The variability of the interactions of the studied compounds with different receptors prevented us from defining a general ranking based on their ED potency. However, for each endpoint, the most dangerous compounds or important structural alerts that could increase ED activity were identified, such as the presence of Br substitutes in *meta*/*para* positions (F04[O–Br]), that induced dioxin and anti-dioxin-like activity, and the presence of aromatic hydroxyl group (nArOH), that greatly increased T4-TTR competition, E2SULT inhibition as well as anti-estrogenic activity. The fact that the aromatic hydroxyl group increases endocrine disrupting potency has already been documented in the literature [37–39].

The classification models proposed here, developed according to the OECD principles for QSAR validation for regulatory purposes, are simple tools for the screening of BFRs in relation to their ED activity; they are useful to draw up priority lists or to suggest safer alternatives, reducing costs and time, and avoiding animal testing. This is in agreement with the requirements of the REACH regulation (Title VII, Chapter 1, Article 57-f), and for the design of safe chemicals according to the Green chemistry approach [40].

Acknowledgements

Financial support by the European Union through the project CADASTER (FP7-ENV-2007-212668) is gratefully acknowledged.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.jhazmat.2011.03.008.

References

[1] Endocrine Disrupters Website, http://ec.europa.eu/environment/endocrine/definitions/endodis_en.htm.
[2] REACH; European Chemicals Agency: Helsinki, Finland; http://ec.europa.eu/environment/chemicals/reach/reach_intro.htm (accessed January 2011).
[3] C.A. de Wit, An overview of brominated flame retardants in the environment, Chemosphere 46 (2002) 583–624.
[4] R.C. Hale, M.J. La Guardia, E. Harvey, M.O. Gaylor, T.M. Mainor, Brominated flame retardant concentrations and trends in abiotic media, Chemosphere 64 (2006) 181–186.
[5] B.C. Kelly, M.G. Ikonomou, J.D. Blair, F.A.P.C. Gobas, Bioaccumulation behaviour of polybrominated diphenyl ethers (PBDEs) in a Canadian Arctic marine food web, Sci. Total Environ. 401 (2008) 60–72.
[6] K. Hooper, T.A. McDonald, The PBDEs: an emerging challenge and another reason for breast-milk monitoring programs, Environ. Health Perspect. 108 (2000) 387–392.
[7] J. Legler, A. Brouwer, Are brominated flame retardants endocrine disruptors? Environ. Int. 29 (2003) 879–885.
[8] I.A.T.M. Meerts, J.J. van Zanden, E.A.C. Luijks, I. van Leeuwen-Bol, G. Marsh, E. Jakobsson, A. Bergman, A. Brouwer, Potent competitive interactions of some brominated flame retardants and related compounds with human transthyretin in vitro, Toxicol. Sci. 56 (2000) 95–104.
[9] I.A.T.M. Meerts, R.J. Letcher, S. Hoving, G. Marsh, A. Bergman, J.G. Lemmen, B. van der Burg, A. Brouwer, In vitro estrogenicity of polybrominated diphenyl ethers, hydroxylated PBDEs, and polybrominated bisphenol A compounds, Environ. Health Perspect. 109 (2001) 399–407.
[10] H. Lilienthal, A. Hack, A. Roth-Harer, S.W. Grande, C.E. Talsness, Effects of developmental exposure to 2,2′,4,4′,5-pentabromodiphenyl ether (PBDE-99) on sex steroids, sexual development, and sexually dimorphic behavior in rats, Environ. Health Perspect. 114 (2006) 194–201.
[11] G. Chen, A.D. Konstantinov, B.G. Chittim, E.M. Joyce, N.C. Bols, N.J. Bunce, Synthesis of polybrominated diphenyl ethers and their capacity to induce CYP1A by the Ah Receptor mediated pathway, Environ. Sci. Technol. 35 (2001) 3749–3756.
[12] A.K. Peters, S. Nijmeijer, K. Gradin, M. Backlund, Å. Bergman, L. Poellinger, M.S. Denison, M. Van den Berg, Interactions of polybrominated diphenyl ethers with the aryl hydrocarbon receptor pathway, Toxicol. Sci. 92 (2006) 133–142.
[13] Y. Wang, H. Liu, C. Zhao, H. Liu, Z. Cai, G. Jiang, Quantitative structure–activity relationship models for prediction of the toxicity of polybrominated diphenyl ether congeners, Environ. Sci. Technol. 39 (2005) 4961–4966.
[14] H.-Y. Xu, J.-W. Zou, Q.-S. Yu, Y.-H. Wang, J.-Y. Zhang, H.-X. Jin, QSPR/QSAR models for prediction of the physicochemical properties and biological activity of polybrominated diphenyl ethers, Chemosphere 66 (2007) 1998–2010.
[15] G. Zheng, M. Xiao, X. Lu, Quantitative structure–activity relationship study on the Ah Receptor Binding Affinities of polybrominated diphenyl ethers using a support vector machine, QSAR Comb. Sci. 26 (2007) 536–541.
[16] M. Harju, T. Hamers, J.H. Kamstra, E. Sonneveld, J.P. Boon, M. Tysklind, P.L. Andersson, Quantitative structure–activity relationship modeling on in vitro endocrine effects and metabolic stability involving 26 selected brominated flame retardants, Environ. Toxicol. Chem. 26 (2007) 816–826.
[17] W. Yang, Y. Mu, J.P. Giesy, A. Zhang, H. Yu, Anti-androgen activity of polybrominated diphenyl ethers determined by comparative molecular similarity indices and molecular docking, Chemosphere 75 (2009) 1159–1164.

[18] W. Yang, X. Liu, H. Liu, Y. Wu, J.P. Giesy, H. Yu, Molecular docking and comparative molecular similarity indices analysis of estrogenicity of polybrominated diphenyl ethers and their analogues, Environ. Toxicol. Chem. 29 (2010) 660–668.
[19] E. Papa, S. Kovarich, P. Gramatica, QSAR modeling and prediction of the endocrine disrupting potencies of brominated flame retardants, Chem. Res. Toxicol. 23 (2010) 946–954.
[20] CADASTER FP7 PROJECT: www.cadaster.eu (accessed January 2011).
[21] OECD principles for the validation, for regulatory purposes, of (quantitative) structure–activity relationship models. Available online at: http://www.oecd.org/dataoecd/33/37/37849783.pdf (accessed January 2011).
[22] T. Hamers, J.H. Kamstra, E. Sonneveld, A.J. Murk, M.H.A. Kester, P.L. Andersson, J. Legler, A. Brouwer, In vitro profiling of the endocrine-disrupting potency of brominated flame retardants, Toxicol. Sci. 92 (2006) 157–173.
[23] T. Hamers, J.H. Kamstra, E. Sonneveld, A.J. Murk, T.J. Visser, M.J.M. Van Velzen, A. Brouwer, A. Bergman, Biotransformation of brominated flame retardants into potentially endocrine-disrupting metabolites, with special attention to 2,2′,4,4′-tetrabromodiphenyl ether (BDE-47), Mol. Nutr. Food Res. 52 (2008) 284–298.
[24] R. Todeschini, V. Consonni, A. Mauri, M. Pavan, DRAGON 5.5 for Windows, Talete srl, Milan, Italy, 2008.
[25] SCAN Software for Chemometric Analysis, rel. 1.1 for Windows, Minitab Inc., USA, 1995.
[26] J.A. Cooper, R. Saracci, P. Cole, Describing the validity of carcinogen screening tests, Br. J. Cancer 39 (1979) 87–89.
[27] ToxMatch, v. 1.06, Developed by Ideaconsult Ltd. for European Chemicals Bureau, 2008.
[28] M.C. Hemmer, V. Steinhauer, J. Gasteiger, Deriving the 3D structure of organic molecules from their infrared spectra, Vib. Spectrosc. 19 (1999) 151–164.
[29] J. Devillers, A.T. Balaban, Topological Indices and Related Descriptors in QSAR and Drug Design, Gordon&Breach, Amsterdam, 2000.
[30] F.R. Burden, Molecular identification number for substructure searches, J. Chem. Inform. Comput. Sci. 29 (1989) 225–227.
[31] F.R.A. Burden, Chemically intuitive molecular index based on the eigenvalues of a modified adjacency matrix, Quant. Struct.-Act. Relat. 16 (1997) 309–314.
[32] M. Moore, M. Mustain, K. Daniel, I. Chen, S. Safe, T. Zacharewski, B. Gillesby, A. Joyeux, P. Balaguer, Antiestrogenic activity of hydroxylated polychlorinated biphenyl congeners identified in human serum, Toxicol. Appl. Pharmacol. 142 (1997) 160–168.
[33] B. Mohar, D. Babic, N.J. Trinajstic, A novel definition of the wiener index for trees, J. Chem. Inf. Comput. Sci. 33 (1993) 153–154.
[34] J. Galvez, R. Garcia, M.T. Salabert, R. Soler, Charge indexes—new topological descriptors, J. Chem. Inf. Comput. Sci. 34 (1994) 520–525.
[35] R. Todeschini, V. Consonni, in: R. Mannhold, H. Kubinyi, H. Timmerman (Eds.), Handbook of Molecular Descriptors, vol. 11, Wiley-VCH, Weinheim, 2000.
[36] J.H. Schuur, P. Selzer, J. Gasteiger, The coding of the three-dimensional structure of molecules by molecular transforms and its application to structure spectra correlations and studies of biological activity, J. Chem. Inf. Comput. Sci. 36 (1996) 334–344.
[37] H. Liu, E. Papa, J.D. Walker, P. Gramatica, In silico screening of estrogen-like chemicals based on different nonlinear classification models, J. Mol. Graph. Model. 26 (2007) 135–144.
[38] A. Roncaglioni, N. Piclin, M. Pintore, E. Benfenati, Binary classification models for endocrine disrupter effects mediated through the estrogen receptor, SAR QSAR Environ. Res. 19 (2008) 697–733.
[39] J. Li, P. Gramatica, Classification and virtual screening of androgen receptor antagonists, J. Chem. Inf. Model. 50 (2010) 861–874.
[40] P. Anastas, J. Warner, Green Chemistry: Theory and Practice, Oxford University Press, New York, 1998.