UNDERSTANDING THE TOXIC POTENCIES OF XENOBIOTICS INDUCING TCDD/F-LIKE EFFECTS

by

Ayşe Defne Şahin

BS. in Chem., Istanbul Technical University, 2013

Submitted to the Institute of Environmental Sciences in partial fulfillment of the requirements for the degree of

Master of Science

in

Environmental Sciences

Boğaziçi University

2017

UNDERSTANDING THE TOXIC POTENCIES OF XENOBIOTICS INDUCING TCDD/F-LIKE EFFECTS

APPROVED BY:

Prof. Dr. Melek Türker Saçan . . . . . .
(Thesis Supervisor)

Prof. Dr. Meral Birbir . . . . . .

Prof. Dr. Ferhan Çeçen . . . . . .

DATE OF APPROVAL (31/07/2017)

ACKNOWLEDGMENTS

I would like to thank everyone who has either physically contributed or encouraged me through the writing process of this thesis. I feel so grateful to have Prof. Dr. Melek Türker Saçan as my thesis advisor, who has always been guiding me and supporting me through every aspect of my life. It is a privilege to work with her. I would also like to thank to the jury members; Prof. Dr. Meral Birbir and Prof. Dr. Ferhan Çeçen for their valuable time and suggestions. I'd like to offer my sincere gratitude to Elif Kahraman, Serli Önlü, Dr. Gülçin Tuğcu and Merve Eminoğlu. It was a pleasure working with you.

I also gratefully appreciate the financial supports of the Scientific and Technological Research Council of Turkey (TUBITAK) (Project No: 214Z225).

A very special thanks to my gourmet friends Ece Özön and Dışeps Apiş, those teas at Kennedy Lodge surely helped me get through this thesis. I am also grateful for my friends from the Institute of Environmental Sciences.

Lastly, I would like to express my deepest love for my friends Korcan Yakşi, Eylül Baran Gürcan, Ece Gülşan, Serkan Durmuş, and Eren Baştanoğlu. Finally I would like to thank my mom and dad for their endless support and love. I –literally- could not be here if it weren't for them.

UNDERSTANDING THE TOXIC POTENCIES OF XENOBIOTICS INDUCING TCDD/F-LIKE EFFECTS

In the present study, the toxic potencies of xenobiotics such as halogenated aromatic hydrocarbons inducing 2,3,7,8-tetrachlorodibenzo-*p*-dioxin/2,3,7,8-tetrachlorodibenzofuran (TCDD/TCDF)-like effects were investigated by quantitative structure-toxicity relationships using their aryl hydrocarbon receptor (AhR) binding affinity data. The descriptor pool was created using SPARTAN 10, DRAGON 6.0 and ADMET 8.0 software and the descriptors were selected using QSARINS (v.2.2.1) software. Data sets were divided into training and test sets. The training sets were comprised of 81% of the complete data set for both models. The generated models for AhR of chemicals with TCDD/F-like effects were internally and externally validated in line with the Organization of Economic Co–operation and Development principles. TCDD-based model had six descriptors from DRAGON 6.0 and ADMET 8.0. TCDF-based model had seven descriptors from DRAGON 6.0 These descriptors were from various blocks including Weighted Holistic Invariant Molecular, Moriguchi Descriptors, and 2D and 3D descriptors blocks. The predictive ability of the generated models was tested for about 1000 diverse group of chemicals from polychlorinated/brominated biphenyls, dioxins/furans, ethers, polyaromatic hydrocarbons with fused heterocyclic rings (i.e. phenoxathiins, thianthrenes and dibenzothiophenes), and polyaromatic hydrocarbons (i.e. halogenated napthalenes and phenanthrenes) with no AhR data. For the external set chemicals, the structural coverage of the generated models was 95.55% and 89.37% for TCDD/F-like chemicals, respectively.

TCDD/F BENZERİ ETKİ GÖSTEREN KSENOBİYOTİKLERİN TOKSİK ETKİLERİNİ ANLAMA

Bu çalışmada, 2,3,7,8-tetraklorodibenzo-*p*-dioxin/2,3,7,8-tetraklorodibenzofuran (TCDD/TCDF)-benzeri etkiler gösteren halojenlenmiş aromatik hidrokarbonların toksik etkileri kantitatif yapı-toksisite ilişkileri ile incelenmiştir. Bu doğrultuda, bu kimyasalların Aril Hidrokarbon Reseptörü'ne (AhR) bağlanma afinitelerinden yararlanılmıştır. Modellerde kullanılan tanımlayıcılar SPARTAN 10, DRAGON 6.0 ve ADMET 8.0 yazılımları kullanılarak oluşturulmuş ve tanımlayıcı seçimi QSARINS (v.2.2.1) yazılımı kullanılarak yapılmıştır. Veri setleri, eğitim ve test setlerine ayrılmıştır. Eğitim setleri, her iki modelde de veri setlerinin % 81'inden oluşmuştur. Oluşturulan modeller, Ekonomik İşbirliği ve Kalkınma Örgütü'nün belirlediği ilkelere uygun olacak şekilde dâhili ve harici olarak doğrulanmıştır. TCDD'ye göre normalize edilen veri setinin modeli, DRAGON 6.0 ve ADMET 8.0'dan elde edilen altı tanımlayıcılıdır. TCDF'e göre normalize edilmiş veri setinin modeli ise DRAGON 6.0 yazılımından elde edilen yedi tanımlayıcıdan oluşmuştur. Bu tanımlayıcılar, Ağırlıklı Bütünsel Yapıya Bağlı Değişkenlik gösteren tanımlayıcılar, Moriguchi tanımlayıcıları, ve de iki ve üç boyutlu tanımlayıcı bloklarını içeren çeşitli bloklardan oluşmuştur. Elde edilen modellerin tahmin performansı, naftalin, fenantren gibi poliaromatik hidrokarbonlar (PAH), poliklorlu/bromlu bifeniller, dioksinler/furanlar, eterler, phenoksatinler, tiantrenler ve dibenzotiofenler gibi heterosiklik halka içeren bileşikleri de kapsayan çok çeşitli kimyasal gruba ait yaklaşık 1000 adet kimyasal ile test edilmiştir. Harici veri seti olarak kullanılan bu kimysalların AhR verileri bulunmamaktadır. TCDD/F benzeri kimyasallar için geliştirilen modeller harici set kimyasallarını sırasıyla % 95.55 ve % 89.37 oranında yapısal olarak kapsamıştır.

TABLE OF CONTENTS

ACKNOWLEDGMENT iii
ABSTRACT iv
ÖZET v
TABLE OF CONTENTS vi
1. INTRODUCTION 4
1.1 Aim of the Study 6
2. THEORETICAL BACKGROUND 6
2.1 Chlorinated and Brominated Biphenyls 7
2.1.1. Commercial Uses of Halogenated Biphenyls 7
2.1.2. Exposure and Effects on the Environment 8
2.1.3. Health Concerns Related to Halogenated Biphenyls 9
2.1.4. Polychlorinated Biphenyl Derivatives 10
2.2. Brominated and Chlorinated Diphenyl Ethers (PBDEs & PCDEs) 11
2.2.1. Commercial Uses of Halogenated Diphenyl Ethers 11
2.2.2. Exposure and Effects on the Environment 12
2.2.3. Health Concerns Related to Halogenated Ethers 13
2.2.4. Halogenated Diphenyl ether derivatives 14
2.3. Polychlorinated dibenzo-*p*-dioxins (PCDDs) and polychlorinated dibenzofurans (PCDFs) 14
2.3.1. Sources of PCDDs and PCDFs 15
2.3.2. Exposure and Effects on the Environment 15
2.3.3. Health Concerns Related to PCDD/Fs 16
2.4. Polyhalogenated Naphthalenes 17
2.4.1. Sources of Naphthalene 17
2.4.2. Exposure and Effects on the Environment 18
2.4.3. Health Concerns 18

2.5. Indolocarbazoles and Derivatives 19
2.5.1. Sources and Commercial Usage 19
2.5.2. Health Related Issues 20
2.6. Polyaromatic Hydrocarbons with Fused Heterocyclic Rings 21
2.6.1. Exposure and Effects on the Environment 22
2.6.2. Health Related Issues 22
2.7. Quantitative Structure-Activity Relationships (QSARs) 22
2.7.1. QSAR Model Validation 23
2.8. Aryl Hydrocarbon Receptor and Its Relevance to Xenobiotics 24
2.9. Studies on Existing QSAR Models for TCDD and TCDF-Normalized AhR 25
3. MATERIALS AND METHODS 28
3.1. Data Set 28
3.2. QSA/TR Model development 29
3.3. Structure Optimization and Descriptor Selection 30
3.4. Training and Test Set Divisions 31
3.5. Model Development and Validation 32
3.6. Internal Validation Parameters 33
3.6.1. R^2 (Coefficient of determination) 33
3.6.2. R^2_{adj} (Adjusted R^2) 33
3.6.3. *F* (Variance ratio) and *s* (standard error of estimate) 34
3.6.4. Leave-one-out (LOO) cross-validation (Q^2_{LOO}) 34
3.6.5. *Y*-scrambling 35
3.7. External Validation Parameters 35
3.7.1. Predictive squared correlation coefficients (Q^2_{F1}, Q^2_{F2} and Q^2_{F3}) 36
3.7.2. Concordance Correlation Coefficient (CCC) for test set 37
3.7.3. The r_m^2 37
3.7.4. Golbraikh and Tropsha method 38
3.7.5. Mean Absolute Error (*MAE*) based criteria 38
3.8. Applicability Domain (AD) 39

3.9. Insubria Graph 40
3. RESULTS AND DISCUSSION 41
4.1. Model Development 41
4.2. QSTR Model of pIC_{50} using TCDD-normalized Data Set 46
4.3. QSTR Model of pIC_{50} using TCDF-normalized Data Set 53
4.4. Applicability Domains of the Selected Models 60
4.4.1. Applicability domain for the QSTR model of the TCDD-normalized data set 60
4.4.2. Applicability domain for the QSTR model of the TCDF-normalized data set 74
4.4.3. Comparison and further discussion on the models of TCDD and TCDF-normalized data sets 95
4.4.4. Comparison of the QSTR models from the present study with the previously published models 99
5. CONCLUSION 101
REFERENCE 103
APPENDIX A1 114
APPENDIX A2 155

LIST OF FIGURES

Figure 2.1. General structure of PCBs. 7

Figure 2.2. General structure of PBDEs 11

Figure 2.3. General structure of PCDDs and PCDFs. 14

Figure 2.4. Structures of 2,3,7,8 tetrachlorodibenzo-*p*-dioxin and 2,3,7,8-tetrachlorodibenzofuran 15

Figure 2.5. General structure of naphthalene 17

Figure 2.6. General structure of some of the polyaromatic hydrocarbons with fused heterocyclic rings and chlorinated diphenyl thioethers (Mostrag et al., 2010). 21

Figure 3.1. Data set ranges of TCDD and TCDF-like chemicals. Data Set 1 refers to TCDD-normalized data set, and Data Set 2 refers to TCDD-normalized data set. 29

Figure 3.2. Flowchart of QSTR Model Development 30

Figure 4.1. Predicted pIC_{50} from Eq. 4.1 vs. experimental pIC_{50} for the training and test sets of the TCDD-normalized data set; with training set chemicals in yellow color and test set chemicals in blue. 47

Figure 4.2. Relative frequency of descriptors appeared in the model Equation 4.1. 52

Figure 4.3. Predicted pIC_{50} from Eq. 4.2 vs. experimental pIC_{50} for the training and test sets of the TCDF-normalized data set; with training set chemicals in yellow color and test set chemicals in blue. 54

Figure 4.4. Relative frequency of descriptors appeared in the TCDF-normalized model. 59

Figure 4.5. Williams plot for the QSTR model (Eq.4.1.) generated by using TCDD-normalized data set, with training set in yellow and test set in blue. 61

Figure 4.6. Insubria graph of the QSTR model generated using TCDD-normalized data set; hat values and predicted pIC_{50} values of training, test and external sets chemicals; training set in yellow, test set in blue and external set in red. 68

Figure 4.7. Insubria Graph of Eq.4.1 including PAHs and their derivatives with fused heterocyclic rings as an external set. Predicted pIC_{50} values of training, test and external set chemicals from Eq.4.1 and their hat values, where the critical hat value (h^*) is 0.241. C79 in Figure 4.7 is 2,3,6,7-tetrabromonaphthalene. 69

Figure 4.8. Insubria graph indicating the predicted pIC_{50} values of chemicals from Eq. 4.1 for training, test and external (PBBs, PCBs and PCB derivatives) sets. 71

Figure 4.9. Insubria Graph of Eq.4.1 including PCDDs, PBDDs and PCDFs as an external set. Predicted pIC_{50} values of training, test and external set chemicals from Eq.4.1 and their hat values, where the critical hat value (h^*) is 0.241. 72

Figure 4.10. Insubria Graph of Eq.4.1 including PCDEs, PBDEs and PBDE derivatives as an external set. Predicted pIC_{50} values of training, test and external set chemicals from Eq.4.1 and their hat values, where the critical hat value (h^*) is 0.241. 73

Figure 4.11. Williams plot for the QSTR model (Eq.4.2) generated by using TCDF-normalized data set, with training set in yellow and test set in blue. 74

Figure 4.12. Insubria graph of the QSTR model generated using TCDF-normalized data set; hat values and predicted pIC_{50} values of training, test and external sets chemicals; training set in yellow, test set in blue and external set in red. 82

Figure 4.13. Insubria Graph of Eq.4.2 including PAHs and indocarbazoles 84

Figure 4. 14. Insubria Graph of Eq.4.2 including PCPTs, PCTAs, PCDTs and PCDPSs as external sets. Predicted pIC50 values of training, test and external set chemicals from Eq.4.2 and their hat values, where the critical hat value (*h**) is 0.329. 84

Figure 4.15. Insubria Graph of Eq.4.2 including PCDDs and PCDFs and. Predicted pIC_{50} values of training, test and external set chemicals from Eq.4.2 and their hat values, where the critical hat value (*h**) is 0.329 86

Figure 4.16. Insubria Graph of Eq.4.2 including PCBs, PCB derivatives and PBBs. Predicted pIC50 values of training, test and external set chemicals from Eq.4.2 and their hat values, where the critical hat value (h*) is 0.329. 87

Figure 4.17. Insubria Graph of Eq.4.2 including PCDE, PBDE, and derivatives of PBDE as an external set. Predicted pIC_{50} values of training, test and external set chemicals from Eq.4.2 and their hat values, where the critical hat value (*h**) is 0.329 88

Figure 4.18. Chemical structures of the compounds that have a higher pIC_{50} value than TCDD. 89

Figure 4.19. Structures of 2-HO-BDE 68 and 2-MeO-BDE68, respectively. 93

Figure 4.20. Chemical structures of hydroxy substituted polychlorinated biphenyls. 94

Figure 4.21. Structural coverage of interpolated predictions of each chemical group showing the percentage of compounds that exceed the critical hat values and below the data set range for TCDF-normalized data. 95

Figure 4.22. Structural coverage of interpolated predictions of each chemical group showing the percentage of compounds that exceed the critical hat values and below the data set range for TCDF-normalized data. 96

LIST OF TABLES

Table 4.1. Test set chemicals and their experimental pIC_{50} values used in the QSTR model generated for the TCDD-normalized data set. 41

Table 4.2. Test set chemicals and their experimental pIC_{50} values used in the QSTR model generated for the TCDF-normalized data set. 43

Table 4.3. Developed models for AhR using TCDF and TCDD-normalized data sets, and their fit, internal and external parameters. 44

Table 4.4. List of descriptors appeared in Eq. 4.1. 48

Table 4.5. List of descriptors appeared in the Eq. 4.2. 55

Table 4.6. Chemicals that are used to model TCDD-normalized data set, their experimental and predicted pIC_{50} values, hat values and descriptor values. 62

Table 4.7. Chemicals that are used to model TCDF-normalized data set, their experimental and predicted pIC_{50} values, hat and descriptor values. 75

Table 4.8. Predicted pIC_{50} values of chemicals that show higher binding affinity compared to TCDF. 89

Table 4.9. Range of descriptors appeared in Eq.4.1 and Eq.4.2, respectively. 97

Table 4.10. Fitting criteria, internal and external validation metrics and descriptors for TCDF and TCDD-based models. 98

Table 4.11. Comparison of the statistical parameters of generated models to those of the previously published models. 100

LIST OF SYMBOLS/ABBREVIATIONS

Symbol	Explanation
AhR	Aryl hydrocarbon Receptor
CMR	Carcinogenic, Mutagenic and Reproductive toxicity
CoMFA	Comparative Molecular Field Analysis
CoMSIA	Comparative Molecular Similarity Indices Analysis
CYP1A1	Cytochrome P450 1A1
ECHA	European Chemical Agency
B04[O-Cl]	Presence/absence of O - Cl at topological distance 4
F04[Cl-Cl]	Frequency of Cl - Cl at topological distance 4
F09[C-Br]	Frequency of C - Br at topological distance 9
HAH	Halogenated Aromatic Hydrocarbon
ICZ	Indolocarbazoles
OECD	Organisation for Economic Co-operation and Development
LOC	Lopping Centric İndex
MATS5m	Moran autocorrelation of lag 5 weighted by mass
MATS5v	Moran autocorrelation of lag 5 weighted by van der Waals volume
MATS5s	Moran autocorrelation of lag 5 weighted by Intinsic state
MLR	Multiple Linear Regression
Mor03v	Signal 03 / weighted by van der Waals volume
M_RNG	Indicator variable for the presence of ring structures
nHACC	Number of acceptor atoms for H-bonds (N,O,F)
PAH	Polyaromatic Hydrocarbon
PBDE	Polybrominated Diphenyl Erher
PBN	Polybrominated Naphthalene
PBT	Persistence, Bioaccumulation and ecoToxicology
PCB	Polychlorinated Biphenyl
PCDD	Polychlorinated Dibenzo-p-dioxin
PCDE	Polychlorinated Diphenyl Erher
PCDF	Polychhlorinated Dibenzofuran
PCDPS	Polychlorinated diphenyl sulphides

PCDT	Polychlorinated Dibenzothiophenes
PCN	Polychlorinated Naphthalene
PCPT	Polychlorinated Phenoxanthiins
PCTA	Polychlorinated Thianthrenes
POP	Persistent Organic Pollutants
QSAR	Quantitative Structure-Activity Relationship
QSTR	Quantitative Structure-Toxicity Relationship
RgGrav__3D	Gravitational radius of gyration
REACH	Registration, Evaluation, Authorisation and Restriction of Chemicals
TCDD	2,3,7,8-tetrachloro dibenzo-p-dioxin
TCDF	2,3,7,8-tetrachlorodibenzofuran
TEF	Toxic Equivalency Factor
Tm	T total size index / weighted by mass
XRE	Xenobiotic-Responsive Elements

1. INTRODUCTION

Xenobiotics are considered an emerging group of persistent organic pollutants (POPs). Some of these chemicals' (i.e. polychlorinated biphenyls) production has stopped many years ago, the remains can still be found in different parts of the environment (i.e. water, soil, air). Moreover, scientific research have proven their adverse health effects (EPA, 1996).

In 2006, Registration, Evaluation, Authorisation and Restriction of Chemicals (REACH) was established by the European Council and the European Parliament. REACH states the need of evaluation of chemicals that are imported or produced in quantities greater than 1 tonne per annum (tpa) for the assessment of toxic effects by 2018. Of the minimum required data set for the assessment of environmental and human hazards PBT (Persistence, Bioaccumulation and ecoToxicology) properties are of major concern together with CMR (Carcinogenic, Mutagenic and Reproductive toxicity) properties. However this information is not available for the majority of the existing chemicals.

Some xenobiotics exert toxic effects via binding the Aryl hydrocarbon Receptor (AhR). Aryl hydrocarbon Receptor is a ligand-activated transcription factor that mediates transcription of many downstream target genes, including cytochrome P450 metabolizing enzymes. Therefore, binding affinity of chemicals to AhR can be accepted as a toxicological endpoint. Aryl Nuclear receptors are crucial in cellular processes like metabolic processes and cell growth (Gronemeyer et al., 2004). Aryl hydrocarbon receptor is one of these receptors. It is a member of the basic helix-loop-helix transcription family and located in the cytoplasm. Ligand binding to AhR is thought to lead a conformational change. These ligands can be either synthetic or natural. Polyhalogenated aromatic hydrocarbons such as biphenyls, dibenzofurans and dioxin-like chemicals are classified among synthetic ligands (Denison et al., 2002). Mechanism suggests that, activated AhR translocates into the nucleus, then it forms a heterodimer by binding to the AhR nuclear translocator protein (Arnt). Heterodimer, later, binds to xenobiotic-responsive elements (XRE) (Poland and Knutson, 1982).

AhR ligands such as polychlorinated biphenyls (PCBs), polychlorinated dibenzofurans (PCDFs) and polyhalogenated dibenzo-p-dioxins (PHDDs) and naphthalenes are classified among the persistent environmental pollutants and can be found widespread in the environment (EPA, 1997). Halogenated Aromatic Hydrocarbons (HAHs) have great adverse effects on health due to their carcinogenic, teratogenic and mutagenic effects (Mendel, 2005). For risk assessment purposes, the concept of toxic equivalent factors (TEFs) has been created for HAHs, which shows the binding affinity toward AhR. TEFs have been assigned to individual dioxins and furans based on a comparison of toxicity to 2,3,7,8-tetrachlorodibenzodioxin (2,3,7,8-TCDD). Although 2,3,7,8-TCDF has been shown to be approximately one-tenth as toxic as 2,3,7,8-TCDD in animal tests, and its toxic equivalent value is 0.1, this is not always the case (Van den Berg et al., 2006).

However, assessing binding affinities are generally time and money consuming procedures. In addition to that, they are somewhat harmful to the environment and they do not exactly promote green chemistry as the wastes often end up in the sewage system. In order to avoid all these problems in vitro and in silico methods can be employed for assessing the binding affinities.

Structure-activity relationships (SARs) and quantitative structure-activity relationships (QSARs) are useful in this aspect. Usage of QSAR is recommended and supported by European Chemical Agency (ECHA). ECHA has updated its Practical Guide on How to use and report (Q)SARs with further advice and examples on using (quantitative) structure-activity relationships for registering under REACH in Helsinki, 17 March 2016.

Quantitative structure-activity/toxicity relationship (QSA/TR) generates a quantitative relationship between the compound’s structure and its chemical, physical and biological properties. Many reliable models have been produced by this method, and it was also employed in the present study.

1.1. Aim of the Study

The aims of this study were to investigate the toxic potencies of xenobiotics (i.e. PCBs, PCDDs, PCDFs, naphthalenes and indolocarbazoles) inducing TCDD/TCDF-like effects using their AhR binding affinities and structures with QSTR, to develop robust QSTR models which complies with the OECD principles for both end points, to indicate the reliability of the predicted AhR binding affinities of test set chemicals in each of the data set comprised of TCDD and TCDF-like chemicals regarding the applicability domain of the developed models, to predict AhR binding affinities of about 1000 diverse group of chemicals from polychlorinated/brominated biphenyls, dioxins, ethers, furans, phenoxathiins, thianthrenes and dibenzothiophenes, and polyaromatic hydrocarbons (PAHs) (i.e. naphthalene, phenanthrene, anthracene, acridine) with no AhR data, and to compare the predicted toxic potencies of HAHs inducing TCDF-like effects with those of the predicted toxic potencies of HAHs inducing TCDD-like effects. In addition to the aims mentioned above contributing to REACH data need regarding the AhR is one of the major aims. The present study seeks to reach the final aim by predicting the pIC_{50} values of chemicals that do not have experimental values yet through from the developed QSTR models.

2. THEORETICAL BACKGROUND

2.1. Chlorinated and Brominated Biphenyls

The general formula for PCBs is $C_{12}H_{10-n}Cl_x$, where n is a number of chlorine atoms from 1 to 10. There are 10 different homologues dependent on the number of chlorines and 209 different theoretical congeners dependent on the number and the position of chlorines (Breivik et al., 2002). 103 of these 209 congeners are most likely to found in commercial use and industrial PCBs are usually mixtures of 50 different congeners.

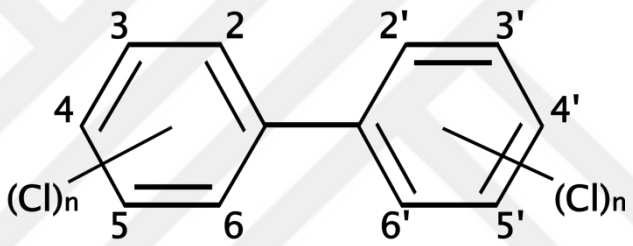


Figure 2.1. General structure of PCBs.

Bromine analogues of PCBs are the polybrominated biphenyls (PBBs). Commercial mixtures of PBBs are found a small amount compared to those of PCBs (Sundström et al., 1976).

2.1.1. Commercial uses of halogenated biphenyls

PCBs were produced from 1929 until 1979 when production was banned by the United States government after its classification as a persistent organic pollutant (POP). After that, in 2001 PCBs were banned by the Stockholm Convention on Persistent Organic Pollutants. The research of de Voogt and Brinkman (1998) estimates the total production to be around 1.5 million tons. The United States is the largest producer with 600,000 tons of production. However, the true cumulative production is expected to be higher since factories in Poland, East Germany and Austria produced unknown amounts of PCB. Even though the production was banned at 1979, 40% of the overall produced PCBs are estimated to be still in use. PBB

production was banned in the US in 1973 (US EPA, 2014). PBB production continued until 1977 in the United Kingdom, 1980s in Germany and 2000 in France. Total production estimation is around 11.000 tons; however some countries are not in this estimation. Approximately 6000 tons of PBBs were produced in the US between 1970 and 1976. FireMaster FF-1 and FireMaster BP-6 account for the 98% of the total production (Hardy, 2002).

PCBs and PBBs were highly used in industry due to their outstanding chemical stability, electrical insulating properties, low flammability and high boiling point properties. Commercial uses of PCBs are categorized as open, partially close and close applications. Close applications basically state the type where chemicals are held in the equipment. Electrical transformers and capacitors, electrical equipment such as voltage regulators, switches and electromagnets, and fluorescent light ballasts are products that are among close applications. PCBs are not directly exposed to the environment in partially closed applications. Hydraulic systems and vacuum pumps are among some examples of partially closed applications. Open applications include products such as plasticizers in paints, flame retardants, wood floor finishes and waterproofing compounds where PCBs are in direct contact with the environment (EPA, 1996; UNEP, 1999).

PBBs were mainly used as flame retardant additives in plastic. They were then made into furniture, textiles, electronics and many other household products (US EPA, 2012).

2.1.2. Exposure and effects on the environment

Thermal and chemical stability makes halogenated biphenyls very suitable to be used in industrial products yet; the same properties make them very hazardous for the environment. Some of the congeners tend to be “dioxin-like” thus very stable and resistant to biodegradation. These types of congeners were released to the environment due to accidents and inappropriate disposal and they are still thought to be present in the environment (UNEP, 1999). Their low vapour pressure and low water solubility allow them to partition between water and the atmosphere. Once released into the environment, PCBs adsorb to organic matter and sediments. It is important to point out that the composition of PCB mixtures changes once they are released into the environment.

Low-chlorinated PCBs tend to be more water soluble, have lower boiling points and more easily biodegradable. Because of these, their concentrations in sediments were found to be lower than the other higher-chlorinated PCBs which also have high boiling point and more resistant to biodegradation. More chlorinated PCBs were also found to bioaccumulate in higher concentrations in wildlife tissues. Volatilized PCBs mostly ended up in remote areas as a result of entering a global biogeochemical cycle and transported (Muir, 2001; Risebrough, 1968; Tanabe 1988).

European Food Safety Authority found PBBs in seafood, meat and meat products and in dairy products. Shen et al. (2008) tested 10 congeners on fish (BB-15, BB-49, BB-52, BB-77, BB-80, BB-101, BB-126, BB-153 and BB-209) and 9 of them (except BB-209) were found in the fat tissues. They also reported the increase in PBB contamination with the increase in fat content. Another study done on human milk from Finland and Denmark proved the presence of BB-153 in that medium (EFSA 2010).

Human and wildlife exposure can happen through contact with contaminated air, sediments, water or diet. For humans exposure mainly happens through contaminated food as a result of bioaccumulation.

2.1.3. Health concerns related to halogenated biphenyls

Both PCBs and PBBs are proven to cause adverse health effects. They have severe effects on immune system, reproductive system, endocrine system, etc. In addition to that they were demonstrated to show carcinogenic effects in animals as well as humans.

The most carcinogenic PCB mixtures were found to be the ones that bind to sediments and bioaccumulate in animals. People, who consume or are in contact with PCB-contaminated products, are under the risk of being exposed to PCBs which may be more toxic compared to their initial state before being released into the environment.

Rhesus monkeys were used for the purpose of assessing PCBs adverse effects on the immune system, as their immune system is very similar to humans' (U.S. EPA, 2013). Results showed significant reduction in the size of thymus gland, decreased resistance to

some viruses and infections and reduction in the response and development of the immune system. Effects on the reproductive system were assessed using mostly Rhesus monkeys as well. PCB exposures were found to cause several adverse effects including decrease in sperm counts and reducing birth weight. These effects were also found to be long lasting even after years after the PCB exposure. Mixtures of PCBs mostly found in breast milk were tested on monkeys for neurological effects. Results proved that PCBs were causing significant decrease in neurological development, learning and short-term memory in new-borns. Finally, PCBs were tested for their endocrine effects and they have demonstrated to increase thyroid hormone levels in humans and animals (EPA, 1996).

PBBs were proven to reach the fetus by passing through the placenta. Shen et al. (2008) investigated that of the 13 PBB congeners, the congeners BB-153 and BB-155 were detected most commonly in the samples. Wang and colleagues (2010) correlated thyroid hormone levels and PBB serum levels in the people who lived near electronic waste dismantling and recycling sites. BB-77, BB-103 and BB-209 were found in those samples. The results were compared to serum from people who lived far away from those sites (Wang et al. 2010).

2.1.4. Polychlorinated biphenyl derivatives

Hydroxylated polychlorinated biphenyls (HO-PCBs) are of concern as well since they are binding to blood proteins, they interfere with physiological development by interfering receptors and pass the placenta. Some of the mentioned derivaties, such as OH-CBs- are produced as a result of oxidative metabolism of PCBs (Morse et al., 1995). Upon binding the 4’- positioned carbon of the PCB structure; they enhance electronegativity, lipophilicity and hydrogen bonding characteristics of the chemical. These enhanced properties result in enhanced binding affinity to related receptors (Parkinson et al., 1988). HO-PCBs have capacity to bind to the AhR and show dioxin-like effects (Cao et al., 2013).

2.2. Brominated and Chlorinated Diphenyl Ethers (PBDEs & PCDEs)

There are 209 PBDE and PCDE congeners in total; differing with the number and position attached to the general structure. Even though these two groups have the same skeleton, they are different. PCDEs are structurally more close to PCBs and polychlorinated dibenzo furans (PCDFs) (Domingo, 2006).

Deca-BDE, octa-BDE and penta-BDE are the three types of commercial PBDEs. Among these decaBDE is the most common homologue. No natural sources are known except for a few marine organisms that may produce PBDEs. They are hydrophobic. They are known to release bromine radicals which reduce combustion rate and dispersion of fire at elevated temperatures which makes them great flame retardants (Hooper and McDonald, 2000; EPA, 2009).

Among 209 possible PCDE congeners, only 106 can be synthesized and 103 of them are currently available (Domingo, 2006).

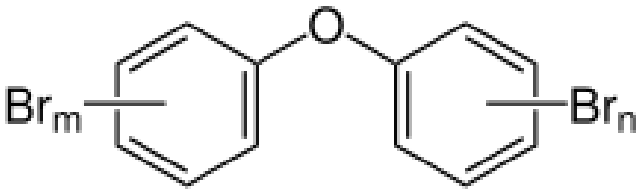


Figure 2.2. General structure of PBDEs

2.2.1. Commercial uses of halogenated diphenyl ethers

PBDEs have been widely used in the United States since 1970s. They are famously used as flame retardants, and with their introduction a significant decrease in the amount of fires were observed. Penta-BDE was used in couches, car seats and chairs with polyurethane foams.

Deca-BDE is generally used in textile products and electronic devices such as television sets and computers. Octa-BDE was used for circuit boards. Commercial mixtures may

contain different homologues. In return, their bioaccumulation properties and toxicity differ (Costa and Giordano, 2007). Penta and octa-BDE's application were banned in 2004 (Ward et al., 2008).

PCDEs are essentially by-products of chlorophenols and chlorinated phenoxyacetic acids. In the industry, they were used as plasticizers, lubricants, electric insulators and flame retardants like PBDEs. In addition, some were used as biocides and herbicides (Koistinen et al., 1996).

2.2.2. Exposure and effects on the environment

PBDEs do not form chemical binds with the polymer product and they can leak to the environment easily. They were detected in air, water and sediment previously. In addition to that they were detected indoors as dust (Darnerud et al., 2001). Their release to the environment may happen through emissions and/or volatilization from production processes. In addition to that inappropriate dispersal methods and leaching from waste may be other possible ways to enter the environment. Congeners with less bromine are found to be more persistent in the environment and bioaccumulate more compared to high brominated congeners. Due to their homophobic properties they do not easily dissolve in water and they bind to sediment. This in return reduces their mobility in the soil. If, however, they are attached to airborne matter, their mobility increases. Congeners with high number of bromine were shown to have the lowest volatilization compared to homolog with low and moderate bromine atoms (ATSDR, 2004; EPA, 2009). Although compounds like PCBs are shown to be decreasing in the environment, PBDEs stayed persistent and increased until 2001 (Darnerud et al., 2001).

Compared to other chlorinated organic pollutants like furans and biphenyls, PCDEs are more scarce in the environment (Domingo, 2006).

2.2.3. Health concerns related to halogenated ethers

PBDEs have been detected in various human tissues in addition to blood and breast milk. BDE-47, -99, -100, -153 and -154 were predominantly found in human tissues and accounting for 90% of the total exposure (McDonald, 2005).

Diet, inhalation and direct contact are some of the human exposure pathways to PBDEs. Regarding the dietary products, fish had the highest amount and this followed by meat. However, meat was thought to be the most significant intake source. This theory was later strengthened when a group of vegans' serum was tested for the PBDE content and the levels were found around one third of the general U.S. population (Schecter et al., 2004, 2006).

In addition to that, high levels of PBDE were detected in breast milk. The highest amounts were generally found in North America. A study estimates a baby's daily PBDE intake would be around 306 ng/kg, whereas the same amount would be 1 ng/kg for adults. Even though the amount of other persistent organic pollutants like biphenyls and dioxins were decreasing in human tissues between 1973 and 2003, levels of PBDE have increased (Schecter et al., 2006). Right after PBDE applications were banned in 2003, their levels in women serum started decreasing (Ward et al., 2008; Zota et al., 2011). Although the decrease is not as sharp as it was in women serum, levels of PBDE in breast milk is also decreasing (Guo et al., 2016).

House dust may be another important PBDE source for small children. Studies suggest that house dust is responsible for over 80% of the overall PBDE exposure for a toddler (Wilford et al., 2005). EPA suggests a toddler would be exposed to 100 mg PBDE daily. For adults levels due to house dust is half of the amount that was found in the toddlers' serum (1997). PBDE exposure was not found related to body mass, however highest levels were found in babies and toddlers due to house dust and breast milk exposure (Schecter et al., 2007).

PBDEs have LD_{50}s > 5 g/kg oral toxicity. Liver, kidney and thyroid gland are the common target organs. Deca-BDE was shown to be less toxic compared to other lower brominated congeners (Darnerud et al., 2001). Some PBDEs have adverse effects on

reproduction. Studies have shown that exposure to BDE-99 decreased sperm counts in male rats and affect the shape of ovary cells in female rats (Kuriyama et al. 2007) (Talsness et al., 2004).

Finally, PBDEs may have adverse effects on developmental neurotoxicity. PBDE exposures may result with adverse changes in motor activity and cognitive behaviour (Eriksson et al., 2001).

2.2.4. Halogenated diphenyl ether derivatives

Hydroxylated polybrominated diphenyl ethers (OH-PBDEs) and methoxylated polybrominated diphenyl ethers (MeO-PBDEs) were found in animal and human tissues. They are most likely the biotransformation products of PBDEs. MeO-PBDEs are produced by the marine organisms whereas HO-PBDEs may occur via biotransformation of PBDEs or biotransformation via naturally occurring MeO-PBDEs (Su et al., 2012).

2.3. Polychlorinated dibenzo-*p*-dioxins (PCDDs) and polychlorinated dibenzofurans (PCDFs)

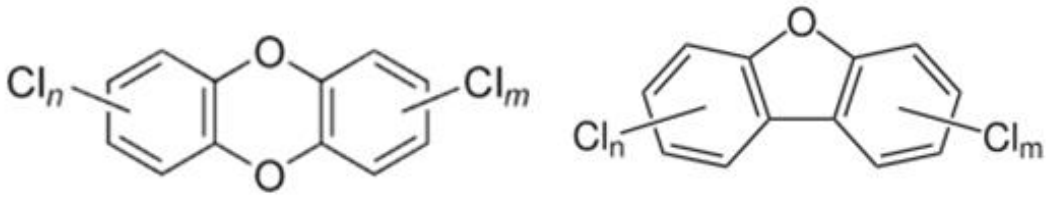


Figure 2.3. General structure of PCDDs and PCDFs.

Unlike other persistent organic pollutants (POPs), PCDFs and PCDDs were not produced on purpose. Instead, they were produced as by-products of industrial combustion processes. There are 75 congeners and 8 homologues for PCDD and 135 congeners for PCDF. They are planar tricyclic ethers. The 2,3,7,8 tetrachlorodibenzo-*p*-dioxin is the most toxic anthropogenic chemical. Studies have shown that many of the 2,3,7,8 substituted congeners are highly toxic. They have low vapour pressure, low water solubility and high binding affinity to soil and sediments (Fiedler, 2003).

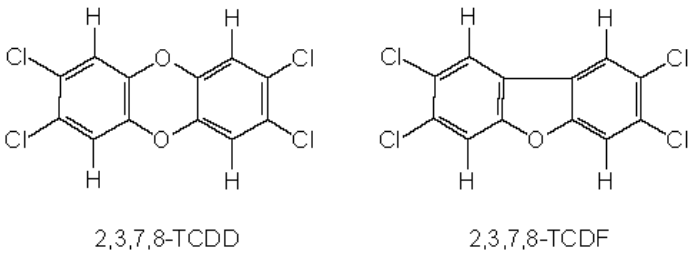


Figure 2.4. Structures of 2,3,7,8 tetrachlorodibenzo-*p*-dioxin and 2,3,7,8-tetrachlorodibenzofuran

2.3.1. Sources of PCDDs and PCDFs

Primary sources of PCDD and PCDFs were mainly from paper and pulp production industry. These chemicals were obtained under high temperature, UV-light, alkaline media and occurrence of radicals. Congeners were found both in the end products (paper and pulp) and in the sludge. Decrease in production happened as the industrial processes changed in time and with advance treatment techniques the amounts found in sludge was reduced as well (Hutzinger and Fiedler, 1993). Residuals of PCDDs and PCDFs were also detected in solid waste incinerators in the Netherlands (Olie et al., 1997).

2.3.2. Exposure and effects on the environment

PCDDs and PCDFs have low water solubility and high lipophilicity which makes them accumulate in fatty tissues. At elevated temperatures (spring and summer seasons) less chlorinated congeners are mostly stay in the vapour phase. In this phase PCDDs and PCDFs can go under dechlorination reactions and the end products may be more toxic especially if the mother compound is degraded to tetra- and penta- congeners. On the other hand, same photochemical reaction may result with non-toxic compounds, if the end products are less chlorinated congeners.

PCDD/Fs can accumulate on plant surfaces via several ways. Congeners with low chlorine numbers are shown to accumulate via dry gaseous deposition whereas; congeners with higher chlorine tend to accumulate via dry particle-bound deposition process. Zucchini

and cucumber were proven to be contaminated by PCDDs and PCDFs previously (Fiedler et al., 2000; 2003).

Studies on carry-over rates proved that rates increase with decreasing number of chlorine atoms on the structure due to the hydrophobic nature. Approximately 30% of the highly toxic PCDD and PCDF congeners were found in cow's milk, after their ingestion through grazing (Welsch-Pausch and McLachlan, 1998).

2.3.3. Health concerns related to PCDD/Fs

Human exposure happens through inhalation, direct dermal intake and diet related contamination. According to WHO, 90% of the contamination results from food consumption (1999).

PCDD/F exposure in humans may lead to skin lesions, liver problems, weakness related to weight loss and endocrine disruptions. 2,3,7,8-PCDD is found to be the cause of liver tumours in animals. In addition to that exposure to it may result in many cell growth inhibition and even cell death (Fiedler, 2003).

Increased occurrence of diabetes and mortality due to diabetes is another toxic effect related to these compounds. Like in PCBs, PCDD/PCDFs are found to be responsible for decreased neurologic development in infants.

2.4. Polyhalogenated Naphthalenes

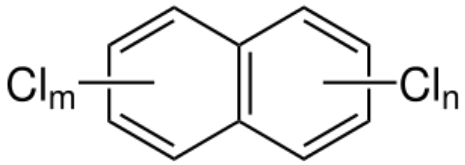


Figure 2.5. General structure of naphthalene

The general formula of polyhalogenated naphthalenes is $C_{10}H_{8-n}X_n$. There are 75 polybrominated naphthalenes (PBN) and polychlorinated naphthalenes (PCN) depending on the degree of halogenation. Unlike chloronaphthalenes there is much less information on sources and environmental and health effects of brominated naphthalenes (Falandysz et al., 2014).

The Br-C bonds' strength and polarizability are much greater compared to Cl-C bonds in PCNs. Chloronaphthalenes are known to have more or less like the same effects as 2,3,7,8-tetrachlorodibenzo-p-dioxin. PCNs are fat-loving, stable and persistent in the environment. They are known to be accumulating in the environment. PBNs may have different degree of planarity due to the size and mass of the bromine atom. In return, this might affect their toxic similarity to TCDD (Falandysz, 1998).

2.4.1. Sources of naphthalene

There are no known natural sources of naphthalene production. PCNs are mostly produced and widely used between 1910s and 1970s. They were also found to be occurring as impurities of PCBs. Penta- and hexa- bromonaphthalenes were also found as impurities in flame retardants. Firemaster BP-6 is the most significant one. Trace amounts of BPNs were found as the end products of pyrolysis process of brominated flame retardants.

Approximately 150.000 tonnes of PCNs are thought to be produced. However, just like in PCBs, the actual amount that has been produced is unknown. The amount of production in the former Soviet Union and China is unknown (Birnbaum et al. 1983).

2.4.2. Exposure and effects on the environment

Exposures to chloronaphthalenes generally happen through food and diet. Accumulation highly depends on the degree of halogenation. In addition to that, structures of PCNs may undergo some changes as they are transferred to vegetables and fruits, so accumulation of each food source may be different. Moreover, accumulation in animal tissues depends on different parameters, including the exposure route, duration, metabolic capacity of the animal and finally the composition of the chemical being exposed to (Domingo et al., 2003; Wyrzykowska et al., 2007; Falandysz et al., 2014).

Environmental contamination and food resource contamination due to PCNs has been reported previously. However, by comparing the breast milk collected in 1972 and 1992 show that the amount of PCNs present in the environment is declining. In fact, Norén and Meironyté, (2000) reported that the amount found in breast milk was halved during this period. The research that has been done in Baltic Sea also proves that the amounts of PCNs are reduced in the environment (Haglund et al., 2010).

No trace of PBNs were found in the environment, however, Firemaster BP-6 was unintentionally added to cattle feed which lead to a food contamination in the USA in 1973 (Birnbaum et al., 1983).

2.4.3. Health concerns

A study that was performed with cows showed that presence of PCNs in the system affect the cow's milk production. There was a significant decrease in cow's amount of milk after being exposed to PCNs. The same study also proved that calves were exposed to PCNs since these chemicals are excreting from the milk (Weistrand and Norén, 1998).

More research on rats prove that accumulation of different congeners were mainly found in the adipose tissue, the heart, the liver and intestines, and then made their way through the skin and to the adipose tissue, if it wasn't the first place of accumulation (Weistrand and Norén, 1998).

Toxicology tests with calves, pigs, rats and rabbits lead to some understanding about the effects of PCNs. Liver necrosis was among the symptoms. In some cases, weight loss and oedema were also observed. Reductions in post-natal development due to PCNs were also observed (Weistrand and Norén, 1998).

2.5. Indolocarbazoles and Derivatives

Indolocarbazoles (ICZ) was first isolated in 1977 from cultures of *Streptomyces staurosporeus.* They have been isolated from bacteria, fungi, invertebrates etc. ever since. In 1990, they have been isolated from cyanobacteria. They are heterocyclic compounds with an indoles attached to benzoid rings. The present five isomers are named indolo[2,3-a]carbazole, indolo[3,2-a]carbazole, indolo[2,3-b]carbazole and indolo [2,3-c]carbazole. Some of the ICZ has been tested for their possible usage against cancer (Sánchez et al., 2006).

2.5.1. Sources and commercial usage

Indolocarbazoles have a dietary origin and it is present in gastrointestinal tract of humans and rodents. They can also be produced in vitro under acidic conditions from indoles. In addition to that, they are found in cruciferous vegetables (Waller and McKinney 1995). Their planar structure allows stable conjugation and this property may be beneficial when used in electronic products. An organic field effect transistor with a layer of indolo[3,2-b]carbazoles was constructed to prove these claims. Moreover, organic thin-film transistors were manufactured from N-alkylated indolo[3,2-b]carbazoles. They might also be used as transporting material for diodes. In addition to that halogenated indolo[3,2-b]carbazoles are used in the production of polymeric materials. 6-formylindolo[3,2-b] carbazole, which is formed from tryphtophan, during sunlight exposure to indoors (Janosik et al., 2008).

2.5.2. Health related issues

ICZ shows high affinity for the Aryl hydrocarbon receptor. However, as the dehalogenated compounds are less lipophilic, they are less likely to accumulate in fatty tissues. In return they are thought to show different effect compared to TCDD.

TCDD was proven to be approximately thousand times more active than ICZ in inducing CYP1A1 cells. Moreover, research proves that TCDD shows at least 2 times higher affinity in binding to AhR (Chen et al., 1995). 6-formylindolo[3,2-b] carbazole promotes CYP1A induction in chick embryo hepatocytes and it was identified as a AhR-dependent initiator of the UVB stress response. In addition to that, indolo[3,2-b] carbazole and 6-formylindolo[3,2-b] carbazole were proven to inhibit an estrogen receptor through binding to AhR (Janosik, Wahlström, and Bergman 2008).

However, research that has been made in the most recent years (Janosik et al., 2008) proves that ICZ has great potential to be used as anticancer drug. Some types of ICZ products have proven to be inhibiting of protein kinases. In addition to that indolocarbazoles have demonstrated that they were able to inhibit human DNA topoisomerase. Therefore authors concluded that ICZ may have great potential in drug development (Tamaoki and Nakano, 1990; Yamashita et al., 1992).

2.6. Polyaromatic Hydrocarbons with Fused Heterocyclic Rings

General structures of some of the polyaromatic hydrocarbons with fused heterocyclic rings and chlorinated diphenyl thioethers were shown in Figure 2.6. (Mostrag et al., 2010).

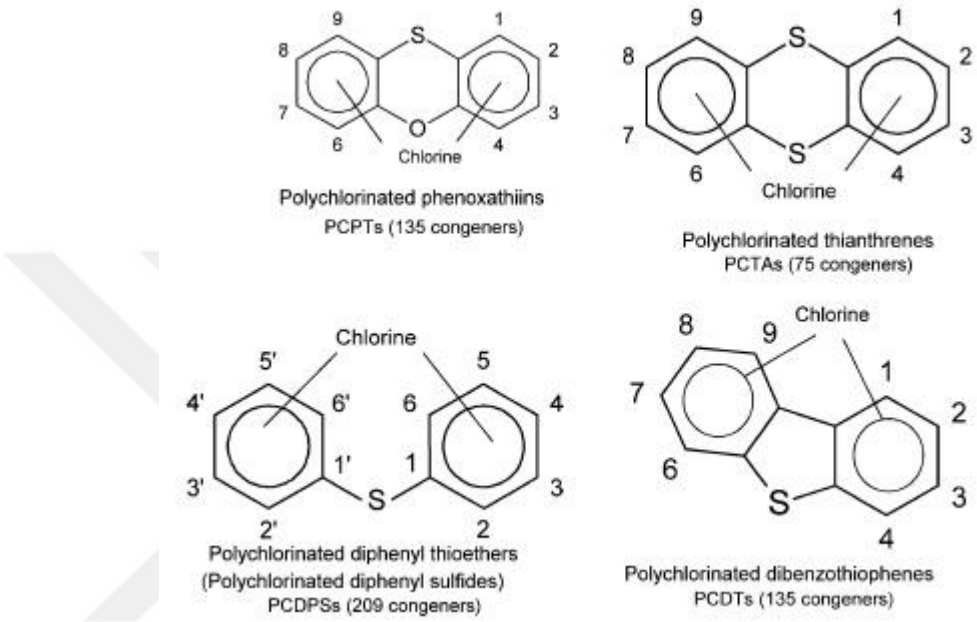


Figure 2.6. General structure of some of the polyaromatic hydrocarbons with fused heterocyclic rings and chlorinated diphenyl thioethers.

Chlorinated polycyclic aromatic compounds (CPACs) have been seen in many environmental media. Some of these include polychlorinated phenoxanthiins (PCPTs), polychlorinated thianthrenes (PCTAs), polychlorinated dibenzothiophenes (PCDTs) and polychlorinated diphenyl sulphides (PCDPSs).

These compounds are mostly by-products of industrial processes of PCDDs, PCDFs etc. They are formed as a result of incomplete combustion of other xenobiotics; they are formed in sulfurized coal combustion with presence of halogens under specific conditions (Grzybek et al., 2002). PCDTs, PCTAs and PCDPSs are sulphur analogues of PCDFs, PCDDs and PCDEs, respectively (Sinkkonen, 1997). Since these compounds are by-products of such environmentally significant chemicals, they are of interest in this work.

2.6.1. Exposure and effects on the environment

PCDTs are known to be less toxic than planar PCBs and their corresponding oxygen compounds (Mantyla et al., 1992).

In the environment, they were found in many media including sediments, aquatic organisms and pine needles. In addition to that, some were detected in the air samples and their sources were identified as pulp mills (Sinkkonen et al., 1995a; Sinkkonen et al., 1995b; Sinkonen, 1997). Their occurrence in wastewater may result in emission into atmosphere which may result in acid rain (Abalos et al., 2002).

2.6.2. Health related issues

The above mentioned compounds have been proven for their tendency to bioaccumulate in lipids. Thus, they are able to accumulate in human tissues. Furthermore, since they have high structural similarity with their oxygen-containing analogues, it is highly possible for them to show similar carcinogenic effects (Fenner et al., 2005).

2.7. Quantitative Structure-Activity Relationships (QSARs)

Quantitative structure-activity relationships have been constructed and used for a variety of purposes for more than half a century, with first developed in 1962 by Hansch and colleagues (Hansch et al., 1962).

QSARs aim to derive quantitative relationships between the molecular structures and the biological activity. In addition to that, they aim to discover the mechanism lying behind the biological activity and the structure. Moreover, QSAR models save a lot of time and money by eliminating the experimental part, they replace animal tests in many cases and by eliminating experiments they endorse greener chemistry (Cronin, 2010).

QSAR models are developed by building a relationship between chemical structures and toxicity. This relationship is obtained via molecular descriptors. Molecular descriptors, as described by Consonni and Todeschini (2010), are “the final result of a logic and

mathematical procedure which transforms chemical information encoded within a symbolic representation of a molecule into a useful number or the result of some standardized experiment". There are mainly two groups of molecular descriptors: experimental measurements (i.e. physico-chemical properties) and theoretical molecular descriptors. The one significant contrast between these two groups is that the lack of contribution of experimental mistakes in theoretical molecular descriptors (Consonni and Todeschini, 2010).

There are many types of descriptors with some being as simple as total atom counts or molecular weight, number of bonds etc. Others based on more complex algorithms are topological or 2D-descriptors. Another type of descriptors are 3D-descriptors, geometrical descriptors, they were derived from the spatial coordinates. Overall, there are more than 5000 descriptors and the number keeps increasing. This increase results in broadened research of the quantitative relationship, as each descriptor explains one slight part of the chemical structure (Consonni and Todeschini, 2010).

One of the many ways to build a QSAR model is through Multiple Linear Regression (MLR). The following formula explains the Multiple Linear Regression approach (Eq. 2.1):

$$y_i = b_0 + \sum_{j=1}^{n} b_j x_{ij} + e_i \quad (2.1)$$

where y_i represents the response to be modelled and x_{ij} are the values of selected descriptors and e_i is the random error (Gramatica et al., 2013).

2.7.1. QSAR model validation

According to Organisation for Economic Co-operation and Development (OECD) a valid QSAR model must have five features. These are (1) a defined endpoint, (2) an unambiguous algorithm, (3) a defined domain of applicability, (4) appropriate measures of goodness of fit, robustness and predictivity and (5) a mechanistic interpretation, if possible (OECD, 2007).

A defined endpoint is a physicochemical, biological or environmental property. A valid model should refer what endpoint it used to model. Moreover, it has to be determined by same protocol for every data and the unit of the endpoint must be used accurately (Dearden et al., 2009).

An unambiguous algorithm is needed to ensure that the model can be used and further validated by other people. To have an unambiguous algorithm every method and information that were used to build the model must be published. This includes software packages, descriptors etc. (Zvinavashe et al., 2008).

The applicability domain (AD) was explained as "the response and chemical structure space in which the model makes predictions with a given reliability" by Netzeva et al. (2005). A defined applicability domain also defines the limits of the descriptors that were used to build the model, and therefore very important (Dearden et al., 2009).

The robustness and predictivity, a model should both be externally and internally validated. There are many parameters such as goodness of fit (R^2), variance ratio (F), standard error of the estimate (s) (Zvinavashe et al., 2008). These properties will be explained later in detail.

Finally a mechanistic interpretation is needed if it is possible. According to OECD if descriptors have a physicochemical interpretation with a logical mechanism and if the proposed mechanism can be supported with the work from literature, only then it may be accepted as a high level of confidence interpretation (2007).

2.8. Aryl Hydrocarbon Receptor and Its Relevance to Xenobiotics

Cytochrome P450 1A1 (CYP1A1) is a xenobiotic metabolizing enzyme. In humans it is encoded by the AhR gene. Polycyclic aromatic hydrocarbons activate the transcription of CYP1A, and this leads to increased activity of CYP1A1 activity. The aryl hydrocarbon receptor (AhR) (also referred as the dioxin receptor) is an intracellular receptor and it controls CYP1A1 gene's induction. The proposed mechanism suggests that, activated AhR translocates into the nucleus, and then it forms a heterodimer by binding to the AhR nuclear

translocator protein (Arnt). Heterodimer later binds to xenobiotic-responsive elements (XRE). 2,3,7,8-tetrachlorodibenzo-p-dioxin (TCDD) was shown to have the maximum affinity towards the AhR (Mimura and Fujii-Kuriyama, 2003; Poland and Knutson, 1982).

For risk assessment purposes, the concept of toxic equivalent factors (TEFs) has been created for HAHs. Earlier, hazard and risk assessment was focused on TCDD as it shows maximum affinity towards the AhR. Nonetheless, it has been soon realized that other poly halogenated organic compounds are present in the environment and have higher concentrations than TCDD. This lead to a broadened hazard and risk assessment procedure and TEF approach was adopted (Van Den Berg et al., 1998; Safe, 1997).

Poland and Knutson (1982) found a rank order interaction between halogenated aromatic hydrocarbons' structure and their AhR binding affinities. Developed structure-activity relationships for PCDDs and PCDFs proved substituted lateral 2,3,7,8 positions were the most toxic compounds. As it shows the maximum affinity towards the AhR, TCDD was given a TEF of 1.0. (Safe et al., 1985).

Toxic equivalency factors were used to determine toxic or TCDD equivalents (TEQs). Following equation gives the relationship between TEFs and TEQs in a mixture:

$$TEQ = \sum[PCDD_i] \,.\, TEF_i + \sum[PCDF_i] . TEF_i + \sum[PCB_i] \,.\, TEF_i + \ldots \qquad (2.2)$$

Where i is the concentration and TEF for a congener in a mixture of halogenated aromatic hydrocarbons (Safe, 1997) .

2.9. Studies on Existing QSAR Models for TCDD and TCDF-Normalized AhR

Many QSARs were developed to predict halogenated aromatic hydrocarbons' binding affinities towards AhR.

In 1992, Waller and McKinney used comparative molecular field analysis (CoMFA) to build a quantitative structure-activity relationship for dioxin-like compounds. Their data set was focused on PCDFs, PCDDs, PBDDs and PCBs. The built model was good in terms of predicting dibenzofurans, however, it was poor in predicting dioxins and biphenyls. Therefore, in 1995 they conducted a new study to further validate their previous study. In addition to previous compounds this time they included naphthalenes and indolocarbazoles to have a more diverse training set. Uniquely they used 2, 3, 7, 8-tetrachlorodibenzofuran (TCDF) as an internal standard and normalized all the data to a value of 8.444, the pIC_{50} value for TCDF (Waller and McKinney, 1992; 1995) . Later in 2006, Lo Piparo et al. (2006) used the same data to further validate the model. In this study, the group used Volsurf, Hologram QSAR (HQSAR) in addition to CoMFA and hybrid models, and obtained $R^2 > 0.82$ and $q^2 > 0.62$ for their models. In 2006, Ashek et al. employed comparative molecular field analysis (CoMFA) and comparative molecular similarity indices analysis (CoMSIA) on similar compound groups to predict their AhR ligand binding affinity. Again, in most of these studies, authors compared their studies with Waller and McKinney's (1995) work and further validated the model (Ashek et al., 2006).

In 2010, Diao and co-workers developed a model to predict dioxin and furan affinities towards the AhR receptor. In this study, quantum chemical descriptors were employed to predict the AhR binding affinity. Many descriptors were used to build the model including the energy of the highest occupied and lowest unoccupied molecular orbital (E_{HOMO} and E_{LUMO}), dipole moment etc. (Diao et al., 2010).

In 2011, Li and others performed a docking (a 3D-QSAR) study with a data set including PCBs, PCDDs and PCDFs. Their docking studies showed a hydrogen-bonding and hydrophobic interactions between selected compounds and the AhR. Moreover, based on their QSAR model they concluded that molecular size, shape, polarizability and

electrotopological states were important parameters for the AhR binding affinity (Li et al., 2011).

In 2010, Papa et al. developed a model for predicting the endocrine-disrupting potencies of PBDEs. They developed the model using multiple linear regression (MLR) method and they validated their model according to OECD principles, which was explained previously in this thesis. Later, in 2012, Gu et al. derived a relationship between PBDEs and AhR binding affinity using partial least square (PLS) analysis derived QSAR method. They compared their study's predictive ability with Papa et al. (2010) and concluded that they obtained a moderate (R^2 = 0.68) correlation.

Lately, Yuan et al. (2013) studied the binding affinities of PCDDs, PCDFs and PCBs towards the aryl hydrocarbon receptor. For this purpose they employed the docking approach in addition to 3D-QSAR methods and CoMFA. They compared their results to those of other studies and concluded that docking-based CoMFA models had shown better results compared to the other CoMFA models. One year later, in 2014, the group further validated their previous model and used molecular similarity indices analysis (CoMSIA) approach (Yuan et al., 2014).

Finally, in 2013, Ruffa used multiple linear regressions (MLR) to predict the AhR binding affinity of a large group of halogenated aromatic hydrocarbons including PBDEs PCDD/Fs, PCBs and PBDDs.

3. MATERIALS AND METHODS

3.1. Data Set

Two different data sets were used in this study. First, data were taken from Ruffa (2013) which was put together from studies of Safe (1990), Chen et al., (2001) Waller and McKinney (1995) and Safe et al., (1985). This data set has 107 AhR ligands including 25 dibenzo-*p*-dioxins, 35 dibenzofurans, 18 diphenylethers, 14 biphenyls and 15 different biphenyl derivatives. The binding affinities were calculated as negative logarithm of the concentration needed to remove 50% of radiolabeled 2,3,7,8- tetrachlorodibenzo-*p*-dioxin (TCDD) from the aryl hydrocarbon receptor (AhR). Chemicals were tested on the cytochrome *P*-450 isoenzymes purified from rat liver cytosol.

The second data set was obtained from Waller and McKinney's work in 1995. This data set was eliminated by Lo Piparo and co-workers (2006). Initially, they removed five compounds from the original data set, which had 99 compounds, since the exact binding data was not readily available. In addition to that, one more compound was eliminated, since it was the duplicate of another compound in the data set. In this study, three more compounds were eliminated since they had the exact same structures and binding affinities (Lo Piparo et al., 2006; Waller and McKinney, 1995). The remaining 90 chemicals comprise 25 dibenzo-*p*-dioxins, 35 dibenzofurans, 14 biphenyls, 5 naphthalenes, 7 indolocarbazoles and 4 indolocarbazole derivatives. Again, the binding affinities (pIC_{50}) were calculated as negative logarithm of the concentration needed to displace 50% of radiolabeled 2,3,7,8-tetrachlorodibenzo-*p*-dioxin (TCDD) from the aryl hydrocarbon receptor (AhR). However, these compiled data were from different laboratories that used 2,3,7,8-tetrachlorodibenzofuran (TCDF) as an internal standard to eliminate laboratory variations. Therefore, all binding affinities were normalized to a value of 8.444 for TCDF. Data set ranges of TCDD and TCDF-like chemicals are given in Figure 3.2.

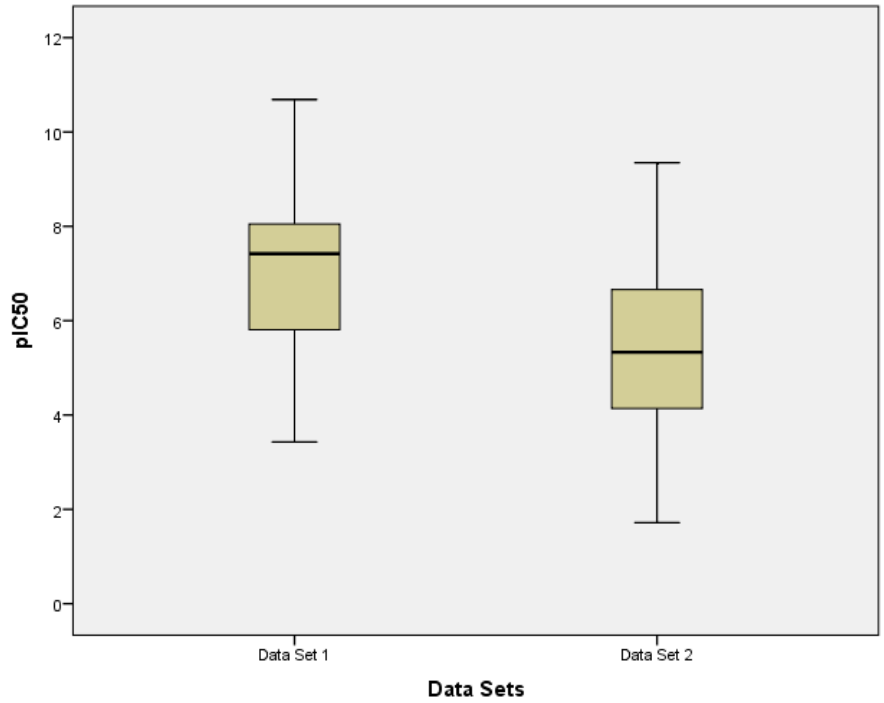


Figure 3.1. Data set ranges of TCDD and TCDF-like chemicals. Data Set 1 refers to TCDD-normalized data set, and Data Set 2 refers to TCDD-normalized data set.

TCDD-normalized values of pIC_{50} ranged from 1.72 to 9.35 mean being a value of 5.49 whereas, TCDF-normalized values of pIC_{50} ranged from 3.429 to 10.687 with mean being a value of 7.03.

3.2. QSA/TR Model development

Flowchart of the model development procedure was given in Figure 3.1. Modelling was done by following dataset preparation, geometry optimization and descriptor calculation, data splitting, descriptor selection, model selection, testing internal and external validation of the model steps. Finally, the final model was tested for its predictive capacity by testing it with compounds outside of the initial data set.

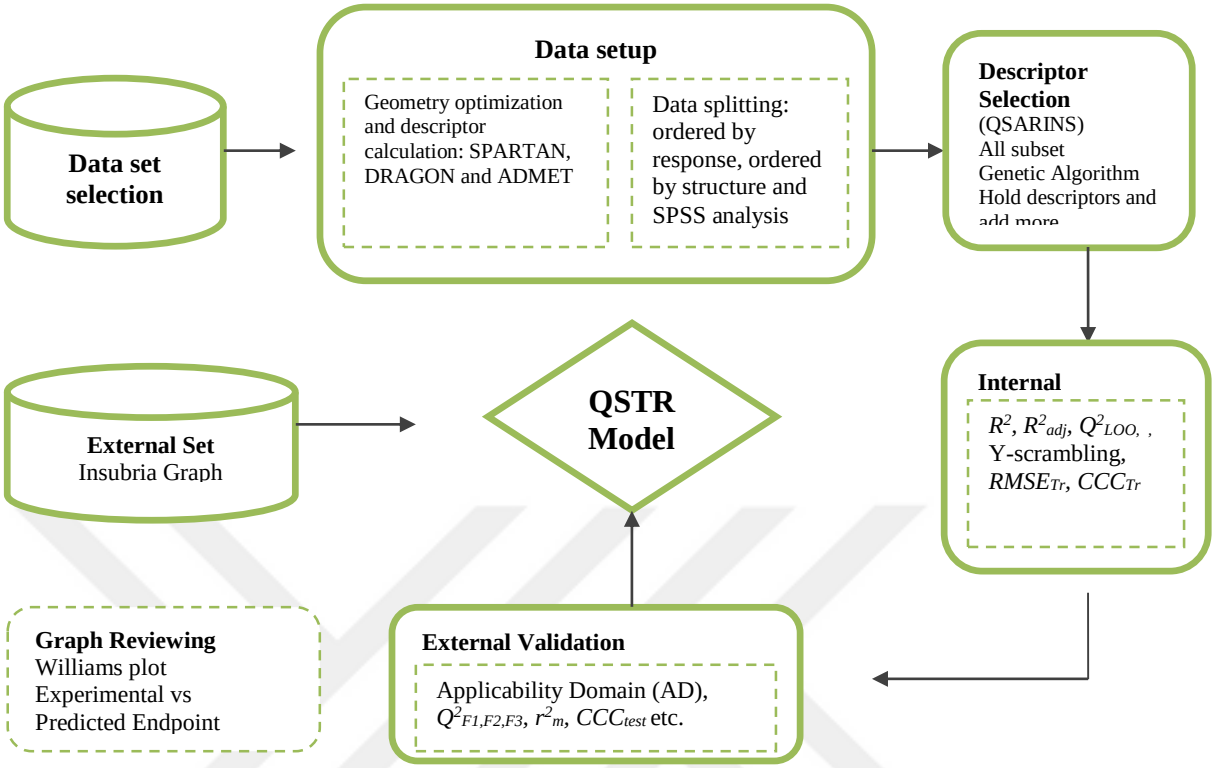


Figure 3.2. Flowchart of QSTR Model Development

3.3. Structure Optimization and Descriptor Selection

Molecular descriptors were calculated using Spartan 10 (Wavefunction, 2010), Dragon 6.0 (Talete, 2014) and Admet 8.0 software packages. Structures were drawn in Spartan 10 software package, conformers of each chemical were searched by using the Semi-Empirical PM6 method and geometry optimization was done with Semi-Empirical PM6 method again. Aqueous-phase energy (E_{aq}) values were calculated for each conformers and the conformer with the lowest E_{aq} value was selected. For some compounds calculations with Semi-Empirical PM6 method was not possible. Those compounds were calculated with Molecular Mechanics (MMF) and later their geometries were optimized using PM6. However, for some cases Spartan 10 was unable do MMF calculation as well, in that case those compounds were calculated using Spartan 14 instead.

The lowest energy conformers of the molecules were used for the descriptor calculations. Molecular weight (MW), dipole moment (μ), the energy of the lowest unoccupied molecular orbital (E_{LUMO}), the energy of the highest unoccupied molecular

orbital (E_{HOMO}), gas-phase energy (E), aqueous-phase energy (E_{aq}), the logarithm of the octanol/water partitioning coefficient (log P), space-filling (CPK) volume and area values were obtained from Spartan 10.

Spartan files were then saved as .mol2 files and loaded to DRAGON 6.0. In addition, .mol2 files were saved as .mol files, and loaded to ADMET 8.0 Software package. DRAGON 6.0 and ADMET 8.0 software packages were used to calculate the exact number of descriptors. 2760 Dragon 6.0 descriptors and 411 ADMET 8.0 descriptors were calculated for the TCDF-normalized data set. 2806 Dragon 6.0 descriptors and 414 ADMET 8.0 descriptors were calculated for TCDD-normalized data set.

Finally, descriptors were saved as a text file which was then loaded into QSARINS (v.2.2.1) software package together with the dependent variable (pIC_{50}) values. Descriptors calculated from Spartan 10 were added to this text file as well. In addition to these descriptors 4 other descriptors, E_{HOMO}-E_{LUMO} gap, hardness (η), softness (S) and electrophilicity (ω), were calculated as described by (Lopachin et al., 2007) and added to the text file.

3.4. Training and Test Set Divisions

As stated previously, training and test set divisions are crucial to have a good QSAR/QSTR. In this study, three training/test set divisions were created. Divisions were made by selecting compounds from clusters created in SPSS 22 by using between-groups linkage and squared Euclidean distance method, response order in QSARINS (v.2.2.1) software and principle component analysis tool which uses structural features of chemicals in QSARINS (v.2.2.1) software. Test set chemicals comprises 20% of the total compounds.

3.5. Model Development and Validation

Models were developed using Genetic Algorithm (GA), all subset and by holding model and adding selecting variables. All subset can be employed to calculate models with small dimensions, as the combinations grow exponentially when higher dimensions are chosen and that process requires too much time. When all subset method does not give the desired results genetic algorithm method can be employed. This method acts like the natural selection, as the best results eliminate the least successful ones. The selection in GA is mostly random. Completely random selection is avoided by selecting the best descriptors from all subset procedure. The selected descriptors are then used by the GA to build models. GA tool in QSARINS allows modifications in population size, the mutation rate and the number of generations for genetic algorithm. In addition to those methods, one can also add descriptors one by one to an already existing model. Improvements in Q^2 and R^2 can be investigated closely by doing so (Gramatica et al., 2013). Improvements in these two parameters are important in terms of obtaining a valid model; however, the increase should be no less than 0.02 to lead us to the conclusion that the descriptor is actually making a significant contribution to the model.

At this point, it is also important to point out the problems that might occur due to excess number of descriptors used in a model. As the number of descriptors increase in a model, it gets harder to interpret every descriptor and their contribution to the overall equation. Topliss ratio suggests that at least 5 training compounds should be represented with one descriptor. For instance, a training set that has 25 training set compounds should not have more than 5 descriptors (Topliss and Edwards, 1979).

There are usually huge numbers of descriptors that are transferred to QSARINS software. The number of models increases as the number of available descriptors increase. A very huge amount of models that were ever created tend to be useless in the end. It is possible to eliminate some of these unwanted models even before the program starts creating them. This is possible with the QUIK rule. Gramatica et al. (2013) explains "QUIK rule tests whether the total correlation among the block of descriptors (K_{XX}) is higher than the correlation among them and responses (K_{XY}), that is, a model is excluded if $K_{XY} - K_{XX} < \delta_K$ is a user defined threshold value." This threshold value was set to 0.05, as suggested in the

manual of the software. After the model calculation is completed, the selected model should be internally and externally validated. To do so, the model should be tested for some specific parameters which are explained in the following section.

3.6. Internal Validation Parameters

Internal validation parameters are: squared correlation coefficient (R^2), the adjusted (for degrees of freedom) squared correlation coefficient (R^2_{adj}), variance ratio (F), standard error (s), cross validation leave-one-out (Q^2_{LOO}), leave-many-out (Q^2_{LMO}), Y-scrambling and the root mean squared error ($RMSE$) of training set.

3.6.1. R^2 (Coefficient of determination)

R^2 is the coefficient of determination between observed and predicted values in a regression. The value of R^2 can be predicted from the following equation (Eq. 3.1):

$$R^2 = 1 - \frac{\sum(Y_{obs} - Y_{calc})^2}{\sum(Y_{obs} - \overline{Y_{obs}})^2} \quad (3.1)$$

where, Y_{obs} is the observed response value, Y_{calc} is the predicted response and $\overline{Y_{obs}}$ is the average of the observed response values. The ideal model is the model in which the sum of the squared residuals being 0 and therefore, the value of R^2 is 1. If R^2 is 0, then there is no relationship between the response and the descriptor. In case where $R^2 > 0.5$ the explained part of the model is greater than the unexplained part (Roy et al., 2015).

3.6.2. R^2_{adj} (Adjusted R^2)

As explained before, there should be a maximum number of descriptors in a model. Obviously, as the descriptor number increases in a model the R^2 value will increase. On the other hand, in some cases, this increase could result with a statistical reliability. So a high value of R^2 does not necessarily mean a robust model. In order to explain the fraction of the data variance explained by the model R^2_{adj} parameter was created. This parameter can be calculated with the following equation (Eq. 3.2):

$$R^2_{adj} = \frac{(N-1)\times R^2 - p}{N-1-p} \quad (3.2)$$

where, p is the number of descriptors and N is the number of points in a given model (Roy et al., 2015).

3.6.3. *F* (Variance ratio) and *s* (standard error of estimate)

Variance ratio investigates the significance of the regression coefficient and is represented by the following equation (Eq. 3.3):

$$F = \frac{\frac{\sum(Y_{calc} - \overline{Y})^2}{p}}{\frac{\sum(Y_{obs} - Y_{calc})^2}{N-p-1}} \quad (3.3)$$

Standard error of the estimate can be calculated by the following equation (Eq. 3.4):

$$s = \sqrt{\frac{(Y_{obs} - Y_{calc})^2}{N-p-1}} \quad (3.4)$$

For a robust model, the value of variance ratio should be high and the value of standard estimate should be low (Roy et al., 2015).

3.6.4. Leave-one-out (LOO) cross-validation (Q^2 LOO)

Cross-validation is the use of a statistical technique in which varying numbers of compounds are removed from the training set (leave-one-out and leave-many-out). The model is then employed to predict the affinity of the eliminated compounds. This helps to calculate the predictive ability of the model. The threshold value is 0.5 (Cruciani et al., 1992; Eriksson et al., 2003).

In a QSAR/QSTR model, leave-one-out procedure continues until every compound from the training set has been excluded and predicted once. The predicted residual sum of

squares (PRESS) value is used to determine the model's predictive capacity. The equation of Q^2_{LOO} is given below (Eq. 3.5-6-7):

$$PRESS = \sum(Y_{obs} - Y_{pred})^2 \tag{3.5}$$

$$SDEP = \sqrt{\frac{PRESS}{n}} \tag{3.6}$$

$$Q^2LOO = 1 - \frac{\sum(Y_{obs(train)} - Y_{pred(train)})^2}{\sum(Y_{obs(train)} - \overline{Y}_{training})^2} = 1 - \frac{PRESS}{\sum(Y_{obs(train)} - \overline{Y}_{training})^2} \tag{3.7}$$

where SDEP is the value of standard deviation of error of prediction, Y_{obs} and Y_{pred} are the observed and leave-one-out predicted activity, *n* is the number of repetitions, $Y_{obs(train)}$ isthe observed activity and $Y_{pred(train)}$ is the predicted activity.

3.6.5. *Y*-scrambling

Y-scrambling (also known as metrics for chance correlation) is done to understand whether if the model was simply by chance. This method is done by scrambling the response values (Y matrix) while keeping the X matrix as it is. No correlation is expected between the response and the new assigned descriptors (Roy et al., 2015).

3.7. External Validation Parameters

Internal validation on its own is not enough to examine a model's validity as it only validates the compounds that are used to build the model. For external validation the data set is generally divided into training and test tests. Training set is then employed to build the model whereas test set is used to check the external validation. External set should be a representative subgroup of the overall data set to ensure the correct validation (Roy et al., 2007).

Parameters of external validation are predictive squared correlation coefficients (Q^2_{F1}, Q^2_{F2} and Q^2_{F3}), Golbraikh and Tropsha (2002) method, Concordance Correlation Coefficient (CCC) and the r_m^2 metric.

3.7.1. Predictive squared correlation coefficients (Q^2_{F1}, Q^2_{F2} and Q^2_{F3})

Q^2_{F1} shows the degree of correlation between the experimental and predicted activity of the data set (Shi et al., 2001).

$$Q^2_{F1} = 1 - \frac{\sum(Y_{obs(test)} - Y_{pred(test)})^2}{\sum(Y_{obs(test)} - \overline{Y}_{training})^2} \quad (3.8)$$

where, $Y_{obs(test)}$ and $Y_{pred(test)}$ are the experimental and predicted activity for the data respectively and, $\overline{Y}_{training}$ is the mean experimental activity of the training set molecules.

The Q^2_{F2} parameter was described by Schüürmann et al. (2008). The main difference between Q^2_{F1} and Q^2_{F2} is that the mean experimental activity is replaced in Q^2_{F2} with the mean predicted activity.

$$Q^2_{F2} = 1 - \frac{\sum(Y_{obs(test)} - Y_{pred(test)})^2}{\sum(Y_{obs(test)} - \overline{Y}_{test})^2} \quad (3.9)$$

Finally Consonni and co-workers proposed the Q^2_{F3} parameter (2010).

$$Q^2_{F3} = 1 - \frac{\left[\sum(Y_{obs(test)} - Y_{pred(test)})^2\right]/n_{ext}}{\left[\sum(Y_{obs(test)} - \overline{Y}_{train})^2\right]/n_{tr}} \quad (3.10)$$

where, n_{tr} is the number of compounds in the training set. Even though Q^2_{F3} measures the model's predictivity it is sensitive to training set selection and it criticizes the data set when they are very homogeneous.

3.7.2. Concordance Correlation Coefficient (CCC) for test set

The CCC parameter can be calculated with the following equation:

$$\overline{\rho}_c = \frac{2\sum(x_{obs(test)}-\overline{x_{obs(test)}})(y_{pred(test)}-\overline{y}_{pred(test)})}{\sum_{i=1}^{n}(x_{obs(test)}-\overline{x_{obs(test)}})^2+\sum_{i=1}^{n}(y_{pred(test)}-\overline{y_{pred(test)}})^2+n(\overline{x_{obs(test)}}-\overline{y_{pred(test)}})} \quad (3.11)$$

where, $\overline{\rho}_c$ refers to the Concordance Correlation Coefficient of test set (CCC_{Test}), $x_{obs(test)}$ and $y_{pred(test)}$ refer to experimental and predicted values of the test compounds, n is the number of compounds, $\overline{x_{obs(test)}}$ and $\overline{y_{pred(test)}}$ refer to the average of the experimental and predicted values for the test compounds. Ideally, CCC should have a value of 1. Concordance Correlation Coefficient measures both the distance of observations to the fitting line and the distance which the regression line deviates from slope 1 passing through the origin. Thus, CCC value is often result smaller than its ideal value of 1 (Chirico and Gramatica, 2011).

3.7.3. The r_m^2

The r_m^2 predicts the relationship between the order of the experimental activity and the predicted activity. Ideally, the difference between these two is expected to be 0 in the case where the experimental and predicted values fit each other (Ojha et al., 2011). The r_m^2 can be calculated with the following equation:

$$r_m^2 = r^2(1-\sqrt{r^2-r_0^2}) \quad (3.12)$$

where r_0^2 is the squared correlation coefficient between the experimental and predicted values of the test set compounds as the intercept is fixed to 0. The value of r^2 is always greater than the value of r_0^2. In a robust model the value of r_m^2 is greater than 0.5 (Ojha et al., 2011).

3.7.4. Golbraikh and Tropsha method

Golbraikh and Tropsha (2002) set some criteria for external prediction. If all these criteria are met, they say, and then the proposed model has a degree of validation:

i. $Q^2_{Tr} > 0.5$

ii. $R^2_{\text{Test}} > 0.6$

iii. $\frac{r^2 - r_0^2}{r^2} < 0.1$ and $0.85 \leq k \leq 1.15$ or

$\frac{r^2 - r_0^2}{r^2} < 0.1$ and $0.85 \leq k' \leq 1.15$ or

iv. $|r_0^2 - r'^2_0|$

In addition to those external validity can be determined from the root mean square error (RMSE) in prediction. It can be calculated by using the following equation:

$$RMSE = \sqrt{\frac{\sum(y_{obs(test)} - y_{pred(test)})^2}{n_{ext}}} \quad (3.13)$$

where, n_{ext} refers to the number of test set chemicals (Golbraikh and Tropsha, 2002).

3.7.5. Mean Absolute Error (*MAE*) based criteria

Error based metrics like *PRESS*, *RMSE* and *MAE* can be useful in times where Q^2_{ext} metrics may be untrustworthy. Among these, *RMSE* is thought to be more complex than *MAE*. Squaring the high prediction errors will have more effect compared to low prediction errors. On the other hand, when calculated with *MAE* both prediction errors have the same weight which makes it a simpler and more reliable method.

$$MAE = \frac{1}{n_{ext}} \sum \left[Y_{obs} - Y_{pred} \right] \quad (3.14)$$

i. Good predictions:

From a general notation, an error of 10% of the training set range should be acceptable while an error value more than 20% of the training set should be a very high error. Thus, the criteria for good predictions should be the following:

$MAE \leq 0.1$ x training set range and $MAE + 3\delta \leq 0.2$ x training set range

Where, the δ value refers to the standard deviation of the absolute error values for the test set data. Considering a normal distribution pattern, mean ± 3δ covers 99.7% of the data points.

ii. Bad predictions:

A value of *MAE* more than 15% of the training set range should be high while an error more than 25% of the training set is considered very high. Hence, the predictions could be considered very high. Hence, the predictions could be considered when:

$MAE > 0.15$ x training set range or $MAE + 3\delta > 0.25$ x training set range.

The predictions which do not fall under either of the above two conditions may be considered as of moderate quality. The mentioned criteria should be used in cases where there are more than 10 data points in the test set (Roy et al., 2016).

3.8. Applicability Domain (AD)

Applicability domain was defined as "the theoretical region in the chemical space constructed by both the model descriptors and modelled response" by Roy et al. (2015). It is a very important parameter for external validation. AD is one of the five OECD criteria. It estimates the similarity of individual compound's to the rest of the data set.

The plot of standardised residuals versus leverages (hat values, *h*) gives the response outliers (*Y*-outliers) as well as the structure outliers (*X*-outliers). Standardised outliers are given in the *Y*-axis of the graph and the leverage values are given in the *X*-axis.

The leverage value of the Applicability Domain is $3 \times p/n$; where n is the number of training compounds and *p* being the number of descriptors. Compounds with high leverage generally remain outside of the AD, thus, their predictions are not reliable. On the other

hand, the response outliers are taken into consideration with the standardised residuals are greater than three standard deviation units (> 3σ) (Gramatica, 2007; Gramatica et al., 2013).

3.9. Insubria Graph

Insubria Graph is used to predict the unknown endpoints from the model equation regarding the applicability domain of the developed model. Internal and external validation are done to find the most robust and valid model among many others. Once the model is obtained, it can be used to predict the modelled endpoints for compounds with no relevant data. In this study, AhR values (pIC_{50}) of approximately 1000 compounds were predicted from the generated models and the Insubria graphs were presented in the Results and Discussion section to investigate if the predicted values fall in the applicability domain of the relevant models. Compounds in the external set are selected based on the fact that their environmental occurrence or they were reported as dioxin-like AhR receptors previously.

4. RESULTS AND DISCUSSION

4.1. Model Development

Both data sets were divided into two groups: test and training sets. Training sets were comprised of 81% of the entire data set for both of the models. Compounds for the test set were made considering many criteria. First, pIC_{50} values of both data sets (TCDD and TCDF-normalized) were listed in increasing order, compounds with minimum and maximum pIC_{50} values were left in the training set. Further splitting was made using the tool in QSARINS 2.2.1 software. Different models were created using the response and structure splitting setups. In addition, SPSS 22 software was employed to create new splitting. Cluster analysis was done using the group linkage and square Euclidian distance methods. The test sets which resulted with the most robust models are given in Table 4.1. and Table 4.2. for TCDD and TCDF-normalized data sets.

Table 4.1. Test set chemicals and their experimental pIC_{50} values used in the QSTR model generated for the TCDD-normalized data set.

CAS R.N	Name	pIC_{50}	References
189084-61-5	2,3',4,4'-tetrabromodiphenyl ether	2.70	Chen et al., 2001
5436-43-1	2,2',4,4'-tetrabromodiphenyl ether	3.25	Chen et al., 2001
51230-49-0	2-chlorodibenzofuran	3.55	Safe 1990
189084-62-6	2,4',5',6-tetrabromodiphenyl ether	3.87	Chen et al., 2001
67651-34-7	4'-hydroxy-2,3,4,5-tetrachlorobiphenyl	4.05	Safe 1985
25074-67-3	3-chlorodibenzofuran	4.38	Safe 1990
82845-24-7	4'-fluoro-2,3,4,5-tetrachlorobiphenyl	4.60	Safe 1985
70424-68-9	2,3,3',4',5-pentachlorobiphenyl	4.85	Safe 1985
64126-87-0	1,2,4,8-tetrachlorodibenzofuran	5.00	Safe 1990

Table 4.1. Continued.

CAS R.N	Name	pIC_{50}	References
38380-08-4	2,3,3',4,4',5-hexachlorobiphenyl	5.15	Safe 1985
88966-76-1	4'-acetyl-2,3,4,5-tetrachlorobiphenyl	5.17	Safe 1985
38380-08-4	2,3,3'4,4',5'-hexachlorobiphenyl	5.33	Safe 1985
83704-39-6	1,3,6-trichlorodibenzofuran	5.36	Safe 1990
88966-68-1	4'-ethyl-2,3,4,5-tetrachlorobiphenyl	5.46	Safe 1985
30746-58-8	1,2,3,4-tetrachlorodibenzo-*p*-dioxin	5.89	Safe 1990
83704-53-4	1,2,3,7,9-pentachlorodibenzofuran	6.40	Safe 1990
58802-16-7	1,3,4,7,8-pentachlorodibenzofuran	6.70	Safe 1990
34816-53-0	1,2,7,8-tetrachlorodibenzo-*p*-dioxin	6.80	Safe 1990
57117-41-6	1,2,3,7,8-pentachlorodibenzofuran	7.13	Safe 1990
51207-31-9	2,3,7,8-tetrachlorodibenzofuran	7.39	Safe 1990

Table 4.2. Test set chemicals and their experimental pIC_{50} values used in the QSTR model generated for the TCDF-normalized data set.

CAS R.N	Chemicals	pIC_{50}*
1746-01-6	2,3,6,7-tetrachlorodibenzo-*p*-dioxin	4.405
33857-28-2	2,3,7-trichlorodibenzo-*p*-dioxin	4.689
38964-22-6	2,8-dichlorodibenzo-*p*-dioxin	5.482
82306-65-8	1,3,7,8-tetrabromodibenzo-*p*-dioxin	5.715
39073-07-9	2,7-dibromodibenzo-*p*-dioxin	6.057
105906-36-3	2-bromodibenzo-*p*-dioxin	6.281
83704-45-4	2,6,7-trichlorodibenzofuran	6.857
83704-32-9	2,3,4,8-tetrachlorodibenzofuran	7.255
64126-87-0	1,2,4,8-tetrachlorodibenzofuran	7.464
70648-26-9	1,2,3,4,7,8-hexachlorodibenzofuran	7.587
38380-08-4	2,3,3',4,4',5-hexachlorobiphenyl	7.657
52663-72-6	2,3',4,4',5,5'-hexachlorobiphenyl	7.768
35065-27-1	2,2',4,4',5,5'-hexachlorobiphenyl	7.996
33284-53-6	2,3,4,5-tetrachlorobiphenyl	8.171
33649-67-1	1,2,3,5,6,7-hexabromonaphthalene	8.482
241-34-9	Benzo[1,2-b:4,5-b']bis[1]benzothiophene	8.927
57-97-6	7,12-dimethylbenz[a]anthracene	9.943

*Data were taken from Waller and McKinney, 1995.

Models were created using All Subsets and Genetic Algorithm (GA) options in QSARINS 2.2.1 software. QUICK Rule was set to 0.05 before starting to scan for models in order to eliminate the models with intercorrelated descriptors.

Models with descriptor numbers varying from 1 to 7 were created. Best models were chosen through a process of elimination. Selection was made on models' internal and external validation criteria. Moreover, models were tested for their external prediction ability and further elimination was done regarding the number of compounds left outside of the applicability domain. In this thesis, the best three models and some of their internal and external validation parameters are presented in Table 4.3. The highlighted models represent the most valid and robust model for each data set.

Table 4.3. Developed models for AhR using TCDF and TCDD-normalized data sets, and their fit, internal and external parameters.

Model No	Number of Variables	Variables	Fitting Criteria and Internal Validation Parameters							External Validation Parameters						
			R^2	R^2_{adj}	Q^2_{LOO}	$RMSE_{Tr}$	s	F	CCC_{Tr}	R^2_{Test}	Q^2_{F1}	Q^2_{F2}	Q^2_{F3}	CCC_{Test}	$RMSE_{Test}$	MAE_{Test}
TCDF_1	6	MPC09 SpAbs_Dz(p) MATS5s Tm B04[O-Cl] F04[Cl-Cl]	0.84	0.82	0.80	0.67	0.70	56.35	0.91	0.94	0.93	0.93	0.95	0.96	0.38	0.31
TCDF_2	6	Tm B04[O-Cl] F04[Cl-Cl] TPSA(NO) M_POL N_Rings	0.84	0.82	0.80	0.66	0.70	56.85	0.91	0.84	0.81	0.81	0.86	0.88	0.63	0.53
TCDF_3	**7**	**RFD MATS5s Tm nHAcc B04[O-Cl] F04[Cl-Cl] LOC**	**0.85**	**0.83**	**0.82**	**0.64**	**0.68**	**52.49**	**0.92**	**0.91**	**0.89**	**0.89**	**0.92**	**0.94**	**0.48**	**0.37**
TCDD_2	6	MATS5m RDF065s F09[C-Br] M_RNG RgGrav__3D MATS5v	0.85	0.84	0.83	0.64	0.66	77.84	0.92	0.80	0.78	0.77	0.84	0.89	0.66	0.54
TCDD_3	6	MATS5m MATS5v RDF070m F09[C-Br] M_RNG RgGrav__3D	0.85	0.84	0.82	0.65	0.68	74.20	0.92	0.82	0.82	0.81	0.87	0.91	0.61	0.50
TCDD_4	**6**	**MATS5m MATS5v F09[C-Br] M_RNG RgGrav__3D Mor03v**	**0.84**	**0.83**	**0.81**	**0.67**	**0.70**	**70.02**	**0.91**	**0.91**	**0.91**	**0.90**	**0.94**	**0.95**	**0.41**	**0.32**

The highlighted models in Table 4.3. TCDD_4 and TCDF_3 give the following equations, Eq. 4.1. and Eq 4.2. for AhR (pIC_{50}) predictions of TCDD and TCDF-like chemicals, respectively.

pIC_{50}, TCDD = - 3.605(±0.744) – 3.930(±0.816) MATS5m + 4.812(±0.844) MATS5v – 1.237(±0.160) F09[C-Br] + 2.018(±0.197) M_RNG + 2.692 (±0.247) RgGrav__3D + 0.863(±0.223) Mor03v (Eq. 4.1)

pIC_{50}, TCDF = - 1.468(±0.595) + 3.392(±1.038) RFD – 1.450(±0.489) MATS5s + 0.635(±0.039) Tm – 0.609(±0.149) nHAcc + 1.408 (±0.245) B04[O-Cl] – 0.535(±0.110) F04[Cl-Cl] – 1.360(±0.510) LOC (Eq. 4.2)

In these equations, numbers in parenthesis indicate the standard deviation of the coefficient of descriptors.

By looking at descriptors regression coefficients we can discuss that MATS5v was the most significant descriptor for the TCDD-based model. It was followed by MATS5m, RgGrav__3D, M_RNG, F09[C-Br] and Mor03v. Among these MATS5v, M_RNG, Mor03v and RgGrav__3D had a positive effect on AhR which means as the value of these descriptors increase for a chemical, so does the pIC_{50} value. On the other hand MATS5m and F09[C-Br] had a negative effect on AhR which means as the value of these descriptors increase for a chemical, pIC_{50} value decreases.

By looking at descriptors regression coefficients we can discuss that RFD was the most significant descriptor for the TCDD-based model. It was followed by MATS5s, B04[O-Cl], LOC, nHAcc and F04[O-Cl]. Among these RFD, Tm and B04[O-Cl] had a positive effect on AhR which means as the value of these descriptors increase for a chemical, so does the pIC_{50} value. On the other hand MATS5s, nHAcc, F04[O-Cl] and LOC had a negative effect on AhR which means as the value of these descriptors increase for a chemical, pIC_{50} value decreases.

4.2. QSTR Model of pIC_{50} using TCDD-normalized Data Set

By looking at some parameters one can claim that both QSTR models generated for AhR using TCDD and TCDF-normalized data sets are valid and robust. For instance, for the QSTR model of TCDD-normalized data set (Eq. 4.1) Q^2_{LOO} and R^2 values are high (0.8102 and 0.8400, respectively) which indicates that the model has a satisfying internal validation metrics. In addition, Q^2_{Yscr} and R^2_{Yscr} values affirm that the model was not build by chance, but instead it is quite robust. Likewise, R^2_{Test} and $RMSE_{Test}$ values are 0.9103 and 0.4065, respectively which points out the strength of its external predictive ability. For further investigation we can test whether if the model's parameters pass the Golbraikh and Tropsha criteria (2002).

For Eq.4.1, R^2_{Tr} and R^2_{Test} are 0.840 and 0.910, k and k' values are 1.028 and 0.968, and r_0^2 and $r_0'^2$ values are 0.910 and 0.902, respectively. With these parameters' value, the model satisfies the Golbraikh and Tropsha (2002) criteria.

The r^2_m value for the mentioned model is 0.869. This value is very close to the r_0^2 value which indicates its good external prediction.

Lastly, testing Roy's (2016) *MAE* criteria for external prediction which is; $MAE \leq 0.1$ x training set range and $MAE + 3\delta \leq 0.2$ x training set range. *MAE* (95% of the data) is 0.292 and $MAE + 3\delta$ value is 0.988 for the QSTR model of TCDD-like chemicals and the training set range is 7.630. Therefore, the model complies both of the criteria given above that proves its good external prediction ability once more.

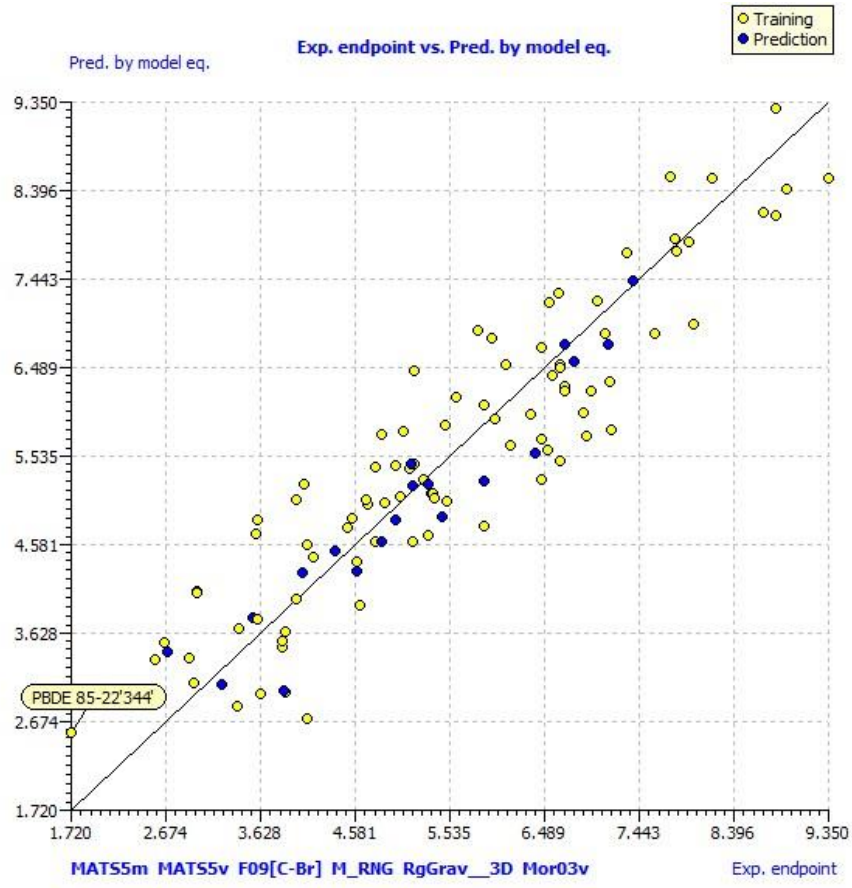


Figure 4.1. Predicted pIC_{50} from Eq. 4.1 vs. experimental pIC_{50} for the training and test sets of the TCDD-normalized data set; with training set chemicals in yellow color and test set chemicals in blue.

Figure 4.1 shows the plot of experimental and predicted pIC_{50} values from the model equation (Eq 4.1). This model was made using 6 descriptors from different blocks from different software packages. Two descriptors (MATS5m and MATS5v) from 2D Autocorrelations block, one (Mor03v) from 3D-MoRSE descriptors block and one (F09[C-Br]) from 2D Atom Pairs block were calculated using DRAGON 6.0 software. The remaining two descriptors were calculated with ADMET 8.0 software and one of the descriptors (RgGrav__3D) was from the 3D Descriptors block, and the final descriptor (M_RNG) was a Moriguchi Descriptor. The relevant blocks of descriptors appeared in Eq 4.1 and their descriptions were given in Table 4.4.

Table 4.4. List of descriptors appeared in Eq. 4.1.

Abbreviation of Descriptor	Description	Block
MATS5m	Moran autocorrelation of lag 5 weighted by mass	2D autocorrelations
MATS5v	Moran autocorrelation of lag 5 weighted by van der Waals volume	2D autocorrelations
Mor03v	signal 03 / weighted by van der Waals volume	3D-MoRSE descriptors
F09[C-Br]	Frequency of C - Br at topological distance 9	2D Atom Pairs
M_RNG	Indicator variable for the presence of ring structures except benzene and its condensed rings	Moriguchi Descriptors for MlogP
RgGrav__3D	Gravitational radius of gyration	3D Descriptors

DRAGON organizes names of descriptors as follows: the number in the name of the descriptor refers to the number of computation, autocorrelation vector of lag n with n being the number of bonds in the unit. The last character usually refers to the physiochemical property regarding its weighting. For instance e indicates Sanderson electronegativity, m indicates atomic mass, s indicates I-state, p indicates polarizability and v indicates van der Waals volume (Kier et al., 1991).

Two of the descriptors used in building the model for chemicals with TCDD-like effects (Eq. 4.1) were from 2D Autocorrelations block of DRAGON 6.0 software. These descriptors A(d), in general, are calculated using the following function (Eq 4.3):

$$A(d) = \sum_{j=1}^{a}\sum_{i=1}^{a}\sigma(d_{ij} - d)p_i p_j$$

$$\sigma = \begin{cases} 1(d_{ij} = d \\ O(d_{ij} \neq d \end{cases} \quad (4.3)$$

Where d refers to a topological distance which can take a number between 1 and the maximum distance in a given molecule, σ is a function of d_{ij}, which is the topological distance between atoms *i* and *j*, a refers to the amount of atoms in the given molecule and $p_i\, p_j$ are the properties of atoms i and j, respectively. The Moreau-Broto function, Moran and Geary functions belong to 2D Autocorrealations descriptor block; however there are slight differences between them. The latter two, calculate the real autocorrelation by taking mean and standard deviation in account for the given property. The Moran coefficient is calculated with the given equation (Eq. 4.4):

$$\boldsymbol{I_k} = \frac{\frac{1}{\Delta_k}\sum_{i=1}^{A}\sum_{j=1}^{A}(w_i - \bar{w})(w_j - \bar{w})\delta(d_{ij};k)}{\frac{1}{A}\sum_{i=1}^{A}(w_i - \bar{w})^2} \quad (4.4)$$

where $\boldsymbol{w_i}$ is any atomic property, $\bar{\boldsymbol{w}}$ is average value of the molecule, *A* is the number of atoms, *k* is the lag, and $\boldsymbol{d_{ij}}$ is the topological distance between the atoms *i* and *j*, is the Kronecker delta which is equal to 1, if $\boldsymbol{d_{ij}}$ =*k* zero and $\Delta_{\boldsymbol{k}}$ is the number of vertex pairs at distance equal to *k*. Moran coefficient generally changes from -1 to +1 (Moran, 1949 as indicated in Consonni and Todeschini, 2010).

Two descriptors, namely MATS5m and MATS5v, from the 2D-autocorrelation descriptor group appeared in Eq 4.1 which were weighted by atomic mass and van der Waals volume, respectively. The contribution of this descriptor to AhR binding affinity seem to be compound specific as the sign of the descriptor values are positive or negative for chemicals in the same group (i.e. PBDE).

3D- Molecule Representation of Structures based on Electron diffraction (3-D MoRSE descriptors) descriptors obtain theoretical scattering curves by using the data gathered from 3D atomic coordinates. Descriptors are calculated with the following equation (Eq. 4.5):

$$I(s) = \sum_{i=1}^{A-1} \sum_{j=i+1}^{A} w_i w_j \frac{\sin(s.r_{ij})}{s.r_{ij}}$$

$$s = 4\pi.\sin(\tfrac{\upsilon}{2})/\lambda \qquad (4.5)$$

where, I(s) is the intensity of scattered radiation, λ is the wavelength and υ is the scattering angle (Soltzberg and Wilkins, 1977 as indicated in Consonni and Todeschini, 2010).

3D-MoRSE (3D-Molecular Representation of Structure based on Electron diffraction) descriptors describe the distribution of the atoms in three-dimensional geometry of molecules and hence can reveal the skeleton and substituent information for a molecule.

When atomic properties appear as weighting factor; these descriptors encode the distribution of the atomic properties in molecules, as such, atomic van der Waals volume weighted Mor03v appeared in TCDD-based model reflects the importance of substituents in the molecule together with their atomic van der Waals volume.

In the QSTR model generated for the prediction of AhR of chemicals with TCDD-normalized data (Eq. 4.1) Mor03v (Morse signal no 03 calculated by weighted van der Waals volume) was used.

The last DRAGON descriptor used in Eq 4.1 was F09[C-Br] which belongs to 2D Atom Pairs descriptor block. These are substructure descriptors, which are vectorial descriptors gathering numbers of occurrences of predefined structural traits, i.e. atom pairs, in molecules or binary variables indicating their occurrence or absence (Lynch et al., 1970). F09[C-Br] descriptor calculates any given pair of atoms and bonds types that connect them and calculated as follows:

AP = [*i*th atom description][separation][*j*th atom description]

i and j atoms should not be directly connected to each other and the separation should be between topological distance (Carhart et al., 1985).

F09[C-Br] is the frequency of C–Br at topological distance of nine in the structure of a molecule and only PBDE group chemicals in the TCDD-normalized data set have value for this descriptor ranging from 1 to 2. It is obvious that only PBDE group has bromine as substituent in their skeleton. Therefore, this descriptor is accepted as indicator variable in Eq. 4.1. As the frequency of C–Br at topological distance of nine in the skeleton of a xenobiotic increases, its binding affinity to AhR decreases.

The remaining two descriptors in this model were calculated by ADMET 8.0 software. One of these descriptors was M_RNG which belong to the Moriguchi Descriptors for MlogP block. The descriptor value could be either 0 or 1 depending on the occurrence of ring structure that is not benzene and its condensed rings.

M_RNG is an indicator variable for the presence of ring structures except benzene and its condensed rings, and its value is 0 for PCBs and their derivatives and most of the PBDE in the training set (Eq. 4.1). It seems that xenobiotics with this kind of ring structure have a higher binding affinity to AhR.

The final descriptor that has been selected for Eq.4.1 was calculated by ADMET 8.0 software as well. The RgGrav__3D descriptor is a 3D descriptor. It calculates the gravitational radius of gyration which is a measure of molecular compactness. Descriptor would get a small value if the majority of atoms in the compound are close to the center of mass. For planar molecules $I_C = 0$ it (where I is inertia) can be calculated from the equation given below (Eq. 4. 6):

$$R_G = \sqrt{\frac{(I_A I_B)^{1/2}}{MW}} \qquad (4.\ 6)$$

And for non planar molecules (Eq. 4. 7):

$$R_G = \sqrt{\frac{2\pi(I_A I_B I_C)^{1/3}}{MW}} \tag{4.7}$$

Since small values are obtained when most of the atoms are close to the center of mass, PCDD, PCDF and PCBs seem to have compact structure compared to most of the PBDEs, PBDDs and derivatives of PCBs, PBCDDs. Regarding the positive sign of RgGrav__3D in Eq 4.1 as the degree that the structure spreads out from its center, the binding affinity of this structure to AhR increases.

Figure 4.2 shows the relative frequency of descriptors appeared in TCDD-normalized model. Many of the descriptors and the descriptors in the same blocks in this model have been selected in other QSAR studies before.

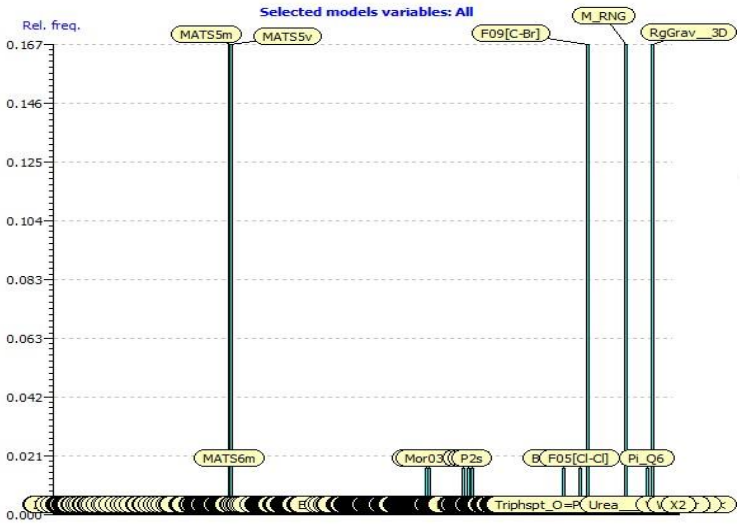


Figure 4.2. Relative frequency of descriptors appeared in the model Equation 4.1.

Mechanistic explanation for binding of xenobiotics to AhR regarding the definitions of descriptors appeared in Eq 4.1. can be done as follows.

The Eq. 4.1 demonstrates the effect of 3D structures of xenobiotics (conformation of a molecule) as encoded by both Mor03v and RgGrav_3D descriptors. Structural connectivity, compactness of the molecule (RgGrav_3D), skeleton and substituent information of the

molecule (Mor03v), atomic van der Waals volume (MATS5v and Mor3v), frequency and specific position of C-Br group (F09[C-Br]), ring structure type (M_RNG), atomic mass (MATS5m) played important roles in AhR binding affinity of TCDD-like chemicals. MATS5v played the most significant role among these descriptors. It was followed by MATS5m, RgGrav__3D, M_RNG, F09[C-Br] and Mor03v.

F09[C-Br] appeared in the QSTR model constructed by Ruffa (2013). Ruffa used exactly the same data set that we used in the present study. In addition to that she used Mor30p, which belongs to 3D-MoRSE descriptor block. Papa et al. (2010) selected many descriptors including RGyr (radius of gyration), Mor08e, MATS6v and Mor22u to develop different QSAR models in order to predict endocrine-disrupting potencies of brominated flame retardants (brominated diphenyl ethers and their hydroxylated derivatives). In addition to those, Li et al. (2011) also developed a QSAR model with a data set comprised of PCBs, PCDDs and PCDFs to predict their binding affinities on the Ah receptor and their model included RGyr and Mor14u descriptors. Finally, Tugcu et al. (2012) used Mor32u as they were investigating the toxic effects of pharmaceuticals on fish. And more recently, they selected Mor09m descriptor to represent phenolic compounds in their QSAR model (Tugcu et al., 2017).

4.3. QSTR Model of pIC_{50} Using TCDF-normalized Data Set

Q^2_{LOO}, and R^2 values for the QSTR model generated by using the TCDF-normalized data set (Eq. 4.2) are high (0.815 and 0.850, respectively) and they are very close to the values of Eq. 4.1. These high values of the model indicate the strength of model's fit and internal validation. Moreover, Q^2_{Yscr} and R^2_{Yscr} values affirm that the model was not build by chance but instead it is quite robust. R^2_{Test} and $RMSE_{Test}$ values for this QSTR model are 0.9133 and 0.476, respectively. Just like in the QSTR model (Eq. 4.1), this model (Eq. 4.2) has also a good external predictive ability. $RMSE_{Test}$ of Eq. 4.2 is slightly higher compared to that of Eq.4.1, and their R^2_{Test} values are close to each another.

This model (Eq. 4.2.) also passes the Golbraikh and Tropsha criteria as R^2_{Tr} and R^2_{Test} are 0.850 and 0.913, k and k' values are 1.027 and 0.970, and r_0^2 and r'^2_0 values are 0.910 and 0.890, respectively.

The r^2_m value for the TCDF-based model is 0.798. This value is very close to the r_0^2 value which indicates its good external prediction.

Finally, *MAE* (95% of the data) is 0.326 and *MAE*+ 3δ is 1.160 for the TCDF-based model and the training set range is 7.258. Therefore, the model complies both of the criteria given above that proves its good external prediction ability once again.

The plot of experimental and predicted pIC_{50} values from Eq. 4.2 is given in Figure 4.3. This model includes 7 descriptors from different blocks of DRAGON 6.0 software. MATS5v is from 2D Autocorrelations block, RFD from Ring descriptors block, Tm from WHIM descriptors block, nHAcc from Functional group counts, B04[O-Cl] and F04[Cl-Cl] from 2D Atom Pairs block and finally LOC from topological indices block. The relevant blocks of descriptors appeared in Eq. 4.2 and their description were given in Table 4.5.

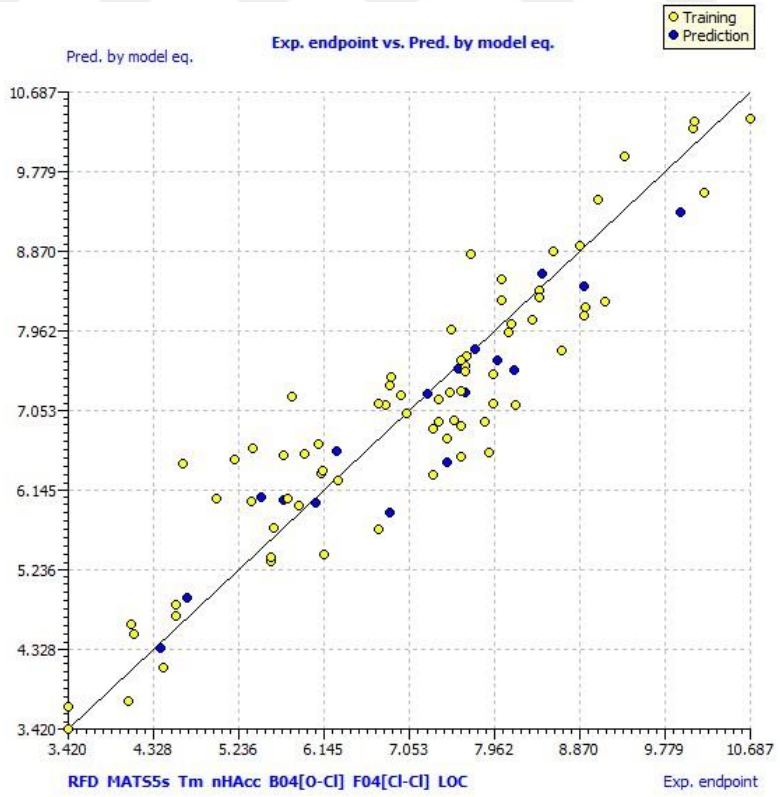


Figure 4.3. Predicted pIC_{50} from Eq. 4.2 vs. experimental pIC_{50} for the training and test sets of the TCDF-normalized data set; with training set chemicals in yellow color and test set chemicals in blue.

Table 4.5. List of descriptors appeared in the Eq. 4.2.

Abbreviation of Descriptor	Description	Block
RFD	ring fusion density	Ring descriptors
MATS5s	Moran autocorrelation of lag 5 weighted by I-state	2D autocorrelations
Tm	T total size index / weighted by mass	WHIM descriptors
nHAcc	number of acceptor atoms for H-bonds (N,O,F)	Functional group counts
B04[O-Cl]	Presence/absence of O - Cl at topological distance 4	2D Atom Pairs
F04[Cl-Cl]	Frequency of Cl - Cl at topological distance 4	2D Atom Pairs
LOC	lopping centric index	Topological indices

The first descriptor used in the TCDF-normalized model is RDF (ring fusion density) from Ring descriptors block (Table 4.5.). Todeschini and Consonni (2009) states that ring descriptors provide information about the abundance of rings in a molecule. It can be calculated through calculating the number of ring sytems (NRS) as following (Eq. 4.8):

$$NRS = (B - B_R) - (A - A_R) + 1 \tag{4.8}$$

where B and A are the total numbers of bonds and atoms, respectively and B_R and A_R are the number of atoms and bonds belonging to rings, respectively. From the equation above RFD can be calculated by dividing the cyclomatic number to the NRS as following (Eq. 4.9):

$$RFD = \frac{C}{NRS} \tag{4. 9}$$

Ring fusion density made the most significant contribution, as it can be seen in Eq.4.2. Moreover, the descriptor had a positive regression coefficient which means it shows a positive effect on pIC_{50} values of the chemicals. For the compounds in the TCDF-normalized

data, this descriptor ranged between 0-0.400. Its value for PCBs was zero as there are no ring fusion in the biphenyl structure. On the other hand 0.4 value was obtained for carbazole derivatives as the mentioned compounds has more than one fused rings. In addition to that PBDE and PBDD compounds had the same RFD value of 0.286 and dibenzofurans had a higher value of 0.308.

One other descriptor in this model is MATS5s. It belongs to 2D Autocorrelations block which is weighed by I-state. A detailed explanation of this block is provided in the descriptor section (4.2) of QSTR model on pIC_{50} for chemicals with TCDD-like effects. The contribution of this descriptor to AhR binding affinity seem to be compound specific as the sign of the descriptor values are positive or negative for chemicals in the same group (i.e. PCBs, PCDDs, PCDFs).

The third descriptor is, Tm, from the Weighted Holistic Invariant Molecular descriptors (WHIM) block (Table 4.5.). These descriptors are based on statistical indices calculated on the projections of the atoms along principal axes. They provide 3D data on properties of molecules such as molecular size, shape, symmetry etc. WHIM descriptors, in general, can be calculated using the following equation (Eq. 4.10):

$$s_{jk} = \frac{\sum_{i=1}^{A} w_i(q_{ij}-\bar{q}_j)(q_{ik}-\bar{q}_k)}{\sum_{i=1}^{A} w_i} \qquad (4.10)$$

where, s_{jk} is the weighted covariance between the *j*th and *k*th atomic coordinates. A is the number of atoms, w_i is the weight of the *i*th atom, q_{ij} and q_{ik} represent the *j*th and *k*th coordinate (j,k = x, y, z) of the *i*th atom respectively, and $\bar{q}$ is the average value (Lasagni et al., 1994).

Since WHIM descriptors reflects the whole molecular structure and size in 3D, Tm weighted with mass from WHIM group indicates the importance of holistic structure of xenobiotics for their binding to AhR. The increase in the size of the molecule increases its binding affinity.

Another descriptor, nHAcc, describes the number of acceptor atoms for H-bonds and belongs to functional group counts block. This descriptor explains the hydrogen-bonding capacity of a molecule expressed as number of possible hydrogen-bond donors. It is calculated by summing the hydrogen atoms bonded to any nitrogen and oxygen without negative charge in the molecule.

The value of this descriptor is zero for PAH and PCB group in the TCDD-normalized data set, since these groups don't have hydrogen atoms bonded to any nitrogen and oxygen with no negative charges in the molecule. This descriptor can be assigned as an indicator variable regarding the range of their values (0-2).

B04[O-Cl] and F04[Cl-Cl] belong to 2D Atom Pairs descriptor block. A detailed explanation of this block is provided in the descriptor section (4.2) of QSTR model on pIC_{50} for chemicals with TCDD-like effects. B04[O-Cl] and F04[Cl-Cl] specifically explain the presence or absence of O-Cl bond at topological distance at 4 and frequency of Cl-Cl bond at topological distance 4, respectively. B04[O-Cl] descriptor had values ranging from 0 to 1, and the value it got varied within the same group. This descriptor had a positive regression coefficient which means it shows a positive effect on pIC_{50} values of the chemicals. In return, as the value for this descriptor increases the binding affinity increases as well. F04[O-Cl] on the other hand, had values between 0-4, and again the values of chemicals varied within the same group. This descriptor had a negative regression coefficient which means it shows a negative effect on pIC_{50} values of the chemical. A high value for this descriptor would decrease the binding affinity.

Finally, lopping centric index (LOC) descriptor was used in developing the TCDF-based model. It belongs to topological indices descriptor block. These are numerical quantifiers of molecular topology that are mathematically derived from the structural graph of a molecule. They can be receptive to structural properties including size, shape, symmetry, branching and cyclicity. They can also provide information regarding atom type and bond multiplicity (Todeschini and Consonni, 2010). LOC specifically, is an index defined as the mean information content derived from the pruning partition of a graph and can be calculated with the following equation (Eq. 4.11):

$$\bar{I}_B = -\sum_{k=1}^{R} \frac{n_k}{A} \log_2 \frac{n_k}{A} \quad (4.11)$$

where, n_k is the number of terminal vertices removed at the *k*th step, A the number of graph vertices, and R the number of steps to remove all graph vertices (Balaban, 1979).

LOC considers the branching and flexibility of substituents. Thus, PCDF0 and PAHs in the TCDF-normalized data set had the value of 0 for this descriptor. LOC had a negative regression coefficient which means it shows a negative effect on pIC_{50} values of the chemicals. As the descriptors' value increases chemicals' binding affinity towards AhR decreases.

Figure 4.4 shows the relative frequency of descriptors appeared for the TCDF-normalized model. Many of the descriptors mentioned above or descriptors from the same block have been used in past research regarding Ah receptor and xenobiotics as well. Information about the mentioned work and the descriptors that have been used is provided in the TCDD-normalized model's descriptor explanation section.

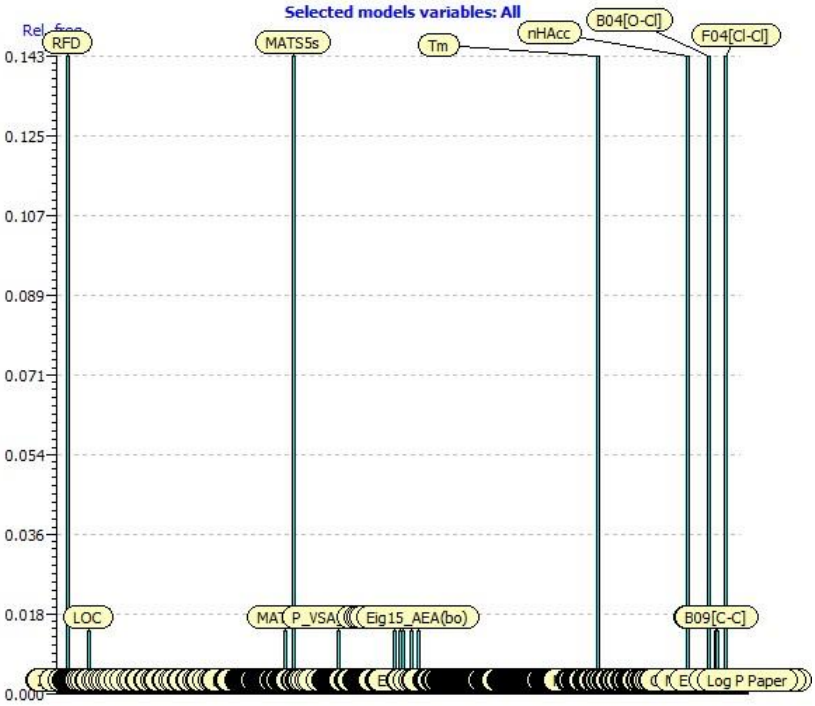


Figure 4.4. Relative frequency of descriptors appeared in the TCDF-normalized model.

Mechanistic explanation for binding of xenobiotics to AhR regarding the definitions of descriptors appeared in Eq 4.2. can be done as follows.

The Eq. 4.2 Branched and flexibility (LOC), size of the molecule (Tm), the number of acceptor atoms for H-bonds in the structure (nHAcc) , frequency and specific position of O-Cl and Cl-Cl groups (B04[O-Cl] and F04[Cl-Cl], respectively), ring fusion density (RFD), I-state of the molecules (MATS5s) played important roles in AhR binding affinity of TCDF-like chemicals. Among these descriptors RFD was the most significant one. It was followed by MATS5s, B04[O-Cl], LOC, Tm, nHAcc and F04[O-Cl].

Descriptors that appeared in Eq.4.1 and Eq.4.2 had many descriptors from common descriptor blocks. So most of the uses from literature which have been stated in section 4.2 apply here as well. For instance, Ruffa (2013) used descriptors (F09[C-Br]) from 2D-Atom Pairs block in her QSTR model. Descriptors from this descriptor block (B04[O-Cl] and F04[O-Cl]) also appeared in Eq.4.2. Papa et al. (2010) selected many descriptors including MATS6v to develop different QSAR models in order to predict endocrine-disrupting

potencies of brominated flame retardants (brominated diphenyl ethers and their hydroxylated derivatives). Tugcu et al., (2017) selected Tm descriptor to represent phenolic compounds in their QSAR model.

4.4. Applicability Domains of the Selected Models

To investigate models' external predictive ability further an external data set with many compounds related to the ones in the original data set that were used to build the model. Later on, pIC_{50} values for these compounds were calculated using the selected models. The number of chemicals which fell within applicability domain of generated models was counted. The structural coverage of each model for external set chemicals was discussed.

4.4.1. Applicability domain for the QSTR model of the TCDD-normalized data set

Williams plot of the QSTR model generated for chemicals with TCDD-like effects was given in Figure 4.5.

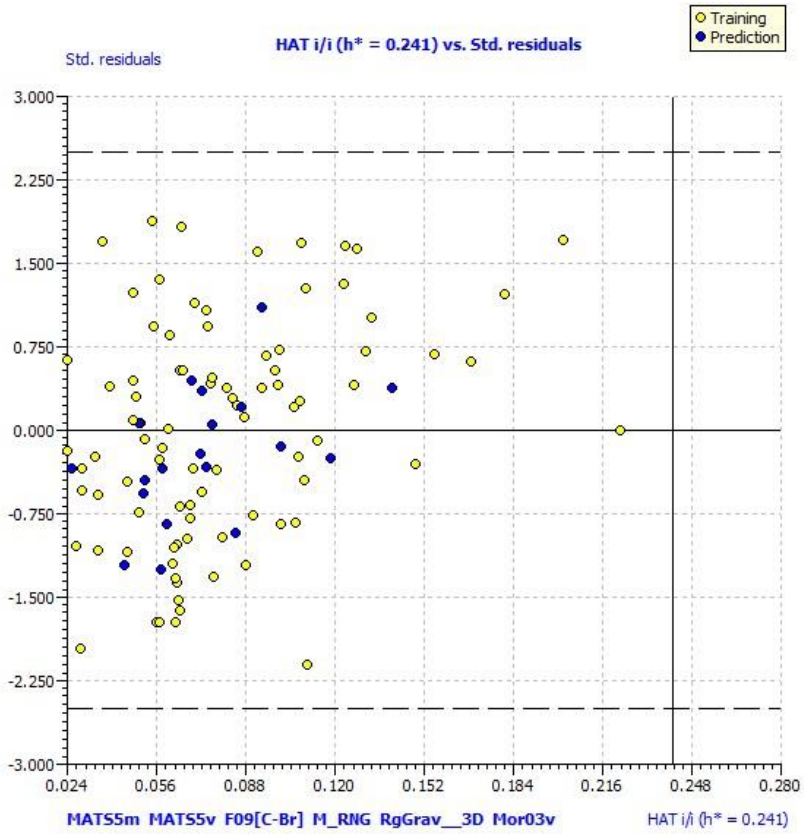


Figure 4.5. Williams plot for the QSTR model (Eq.4.1.) generated by using TCDD-normalized data set, with training set in yellow and test set in blue.

The first thing one can spot immediately in Figure 4.5 is that hat values of all the chemicals in the TCDD-normalized data set are lower than the critical hat value (h^*=0.241). In addition to that, there are no response outliers; the response outlier limit was set to 3δ = 3. Those two facts show that pIC_{50} values for all of the chemicals were well predicted by the model equation (Eq 4.1). Experimental and predicted pIC_{50} from Eq. 4.1, and descriptor values of training and test set chemicals in TCDD-normalized data set are given in Table 4.6.

Table 4.6. Chemicals that are used to model TCDD-normalized data set, their experimental and predicted pIC_{50} values, hat values and descriptor values.

Chemicals	Status	Exp. pIC_{50}	Pred. pIC_{50} by model Eq. 4.1	HAT i/i (h^*=0.241)	RgGrav__3D	M_RNG	MATS5m	MATS5v	F09[C-Br]	Mor03v
4'-acetyl-PCB061	Prediction	5.170	5.223	0.051	3.827	0	0.158	0.210	0	-2.161
4'-br-PCB061	Training	5.600	6.186	0.061	4.034	0	0.046	0.183	0	-2.048
4'-cyano-PCB061	Training	5.270	5.285	0.060	3.631	0	0.062	0.181	0	-1.753
4'ethyl-PCB061	Prediction	5.460	4.896	0.060	3.636	0	0.199	0.236	0	-1.902
4'-fluoro-PCB061	Prediction	4.600	4.302	0.052	3.442	0	0.068	0.134	0	-2.012
4'hydroxy-PCB061	Prediction	4.050	4.291	0.073	3.426	0	0.123	0.224	0	-2.227
4'iodo-PCB061	Training	5.820	6.894	0.093	4.256	0	0.033	0.186	0	-1.997
4'isopropyl-PCB061	Training	5.890	4.795	0.065	3.817	0	0.194	0.174	0	-2.260
4'-methoxy-PCB061	Training	4.800	4.630	0.057	3.650	0	0.176	0.187	0	-2.084
4'-methyl-PCB061	Training	4.510	4.773	0.081	3.417	0	0.186	0.316	0	-1.866
4'-n-butyl-PCB061	Training	5.130	5.401	0.100	4.213	0	0.178	0.171	0	-2.849
4'-nitro-PCB061	Training	4.850	5.778	0.057	3.840	0	0.040	0.182	0	-1.937
4'-phenyl-PCB061	Training	5.180	5.454	0.127	4.374	0	0.125	0.066	0	-2.943
4'-t-butyl-PCB061	Training	5.170	4.615	0.101	3.959	0	0.183	0.125	0	-2.687

Table 4.6. Continued

Chemicals	Status	Exp. pIC_{50}	Pred. pIC_{50} by model Eq. 4.1	HAT i/i (*h**=0.241)	RgGrav__3D	M_RNG	MATS5m	MATS5v	F09[C-Br]	Mor03v
4'trifluoromethyl-PCB061	Training	6.460	6.714	0.094	4.029	0	-0.014	0.249	0	-2.063
PBCDD076	Training	8.830	8.144	0.064	4.243	1	0.044	0.090	0	-2.261
PBCDD077	Training	9.350	8.543	0.088	4.382	1	0.044	0.090	0	-2.231
PBCDD078	Training	7.950	7.850	0.058	4.171	1	0.081	0.092	0	-2.219
PBDD002	Training	6.530	5.613	0.064	3.326	1	0.054	-0.021	0	-1.666
PBDD011	Training	7.810	7.889	0.088	4.188	1	0.053	0.031	0	-2.013
PBDD020	Training	8.930	8.425	0.091	4.339	1	0.030	0.065	0	-2.159
PBDD035	Training	8.700	8.173	0.068	4.373	1	-0.099	-0.032	0	-2.600
PBDD045	Training	8.820	9.286	0.131	4.356	1	-0.002	0.083	0	-2.328
PBDD053	Training	7.770	8.547	0.070	4.531	1	0.025	0.176	0	-2.711
PBDD061	Training	8.180	8.544	0.099	4.634	1	-0.137	-0.065	0	-2.653
PBDE003	Training	3.890	2.997	0.063	3.477	0	0.082	0.022	1	-1.511
PBDE015	Training	3.420	3.693	0.040	3.857	0	-0.002	-0.034	1	-1.959
PBDE017	Training	3.640	2.985	0.067	3.322	0	-0.114	-0.004	1	-1.790
PBDE028	Training	2.920	3.372	0.096	4.212	0	0.009	0.010	2	-2.201
PBDE047	Prediction	3.250	3.089	0.119	3.799	0	-0.095	-0.042	2	-1.425
PBDE049	Training	4.170	4.458	0.076	3.773	0	0.110	0.154	1	-1.352
PBDE054	Training	7.030	7.224	0.084	4.250	1	-0.233	-0.213	0	-2.918
PBDE066	Prediction	2.700	3.436	0.094	4.237	0	-0.061	-0.084	2	-1.998
PBDE071	Prediction	3.870	3.025	0.058	3.535	0	-0.186	-0.122	1	-2.076
PBDE075	Training	3.400	2.855	0.106	3.981	0	-0.095	-0.042	2	-2.264

Table 4.6. Continued

Chemicals	Status	Exp. pIC_{50}	Pred. pIC_{50} by model Eq. 4.1	HAT i/i (*h**=0.241)	RgGrav__3D	M_RNG	MATS5m	MATS5v	F09[C-Br]	Mor03v
PBDE077	Training	2.660	3.527	0.124	4.511	0	-0.027	-0.126	2	-2.358
PBDE085	Training	1.720	2.564	0.110	3.816	0	-0.169	-0.179	2	-1.656
PBDE099	Training	3.850	3.557	0.109	4.027	0	0.086	0.116	2	-1.651
PBDE100	Training	4.110	2.723	0.111	3.813	0	-0.210	-0.147	2	-1.828
PBDE119	Training	2.960	3.103	0.106	4.212	0	-0.169	-0.179	2	-2.268
PBDE126	Training	2.570	3.348	0.181	4.660	0	-0.088	-0.242	2	-2.661
PBDE153	Training	4.600	4.407	0.149	4.283	0	0.212	0.283	2	-1.822
PBDE154	Training	4.640	3.932	0.063	3.586	0	-0.044	0.049	1	-1.492
PBDE183	Training	3.600	3.780	0.108	4.208	0	0.043	0.113	2	-2.137
PCB047	Training	3.890	3.655	0.078	3.442	0	-0.026	-0.067	0	-2.068
PCB060	Training	4.550	4.877	0.076	3.585	0	-0.102	-0.009	0	-1.766
PCB061	Training	3.850	3.483	0.073	3.187	0	0.059	0.088	0	-1.950
PCB077	Training	6.150	5.654	0.050	3.844	0	0.087	0.224	0	-2.116
PCB105	Training	5.370	5.141	0.069	3.755	0	-0.176	-0.020	0	-2.267
PCB107	Prediction	4.850	4.627	0.058	3.657	0	-0.051	0.070	0	-2.491
PCB114	Training	5.390	5.084	0.046	3.677	0	0.066	0.174	0	-2.072
PCB118	Training	5.040	5.106	0.048	3.775	0	0.066	0.174	0	-2.353
PCB126	Training	6.890	6.011	0.077	3.937	0	0.018	0.257	0	-2.490
PCB153	Training	4.102	4.590	0.101	3.705	0	0.207	0.270	0	-2.625
PCB156	Prediction	5.150	5.458	0.069	3.842	0	-0.046	0.160	0	-2.584
PCB157	Prediction	5.330	5.241	0.101	3.826	0	-0.275	-0.066	0	-2.567
PCB167	Training	4.800	5.432	0.075	3.861	0	-0.046	0.160	0	-2.672

Table 4.6. Continued

Chemicals	Status	Exp. pIC_{50}	Pred. pIC_{50} by model Eq. 4.1	HAT i/i (*h**=0.241)	RgGrav__3D	M_RNG	MATS5m	MATS5v	F09[C-Br]	Mor03v
PCB168	Training	4.000	4.002	0.223	3.654	0	-0.479	-0.409	0	-2.480
PCDD001	Training	4.000	5.071	0.128	2.787	1	-0.236	-0.086	0	-1.570
PCDD012	Training	5.500	5.873	0.065	3.529	1	0.134	0.015	0	-1.836
PCDD014	Training	4.890	5.040	0.085	3.167	1	0.204	0.143	0	-2.069
PCDD019	Training	6.660	6.539	0.024	3.809	1	0.005	-0.010	0	-2.384
PCDD020	Training	7.150	6.339	0.062	3.756	1	0.176	0.058	0	-2.053
PCDD035	Training	6.100	6.538	0.024	3.809	1	0.005	-0.010	0	-2.385
PCDD040	Prediction	5.890	5.279	0.085	3.274	1	0.182	0.166	0	-2.353
PCDD045	Training	8.000	6.969	0.064	3.999	1	0.191	0.095	0	-2.220
PCDD046	Prediction	6.800	6.570	0.026	3.780	1	0.005	-0.010	0	-2.258
PCDD053	Training	5.960	6.813	0.048	3.832	1	0.149	0.182	0	-2.557
PCDD059	Training	5.190	6.471	0.055	3.713	1	0.149	0.182	0	-2.582
PCDD061	Training	7.100	6.873	0.030	3.976	1	-0.001	-0.015	0	-2.517
PCDD069	Training	6.550	7.192	0.055	3.967	1	0.103	0.187	0	-2.776
PCDD075	Training	5.000	5.441	0.156	3.982	1	-0.182	-0.313	0	-3.355
PCDF000	Training	3.000	4.073	0.202	2.468	1	-0.348	-0.258	0	-1.281
PCDF002	Prediction	3.550	3.802	0.141	2.908	1	0.121	-0.115	0	-1.631
PCDF003	Prediction	4.380	4.523	0.087	2.962	1	-0.006	-0.108	0	-1.581
PCDF004	Training	3.000	4.093	0.124	2.719	1	-0.178	-0.226	0	-1.447
PCDF012	Training	5.330	4.690	0.080	3.119	1	0.150	0.023	0	-1.900
PCDF014	Training	3.610	4.854	0.065	3.190	1	0.005	-0.113	0	-1.833
PCDF016	Training	3.590	4.703	0.108	3.296	1	0.217	0.004	0	-2.026

Table 4.6. Continued

Chemicals	Status	Exp. pIC_{50}	Pred. pIC_{50} by model Eq. 4.1	HAT i/i (*h**=0.241)	RgGrav__3D	M_RNG	MATS5m	MATS5v	F09[C-Br]	Mor03v
PCDF028	Prediction	5.360	5.140	0.074	3.251	1	-0.296	-0.282	0	-2.118
PCDF030	Training	4.070	5.236	0.037	3.358	1	-0.124	-0.152	0	-2.283
PCDF036	Training	4.720	5.027	0.048	3.137	1	0.013	0.008	0	-2.106
PCDF039	Training	6.000	5.947	0.053	3.544	1	0.185	0.137	0	-2.246
PCDF046	Training	6.350	5.987	0.030	3.471	1	0.013	0.008	0	-2.033
PCDF049	Training	6.660	5.498	0.057	3.496	1	-0.276	-0.250	0	-2.554
PCDF058	Prediction	5.000	4.864	0.072	3.390	1	0.086	-0.027	0	-2.556
PCDF065	Training	6.960	6.246	0.027	3.572	1	-0.139	-0.053	0	-2.399
PCDF071	Training	6.460	5.293	0.058	3.383	1	-0.276	-0.250	0	-2.440
PCDF072	Training	6.660	6.498	0.035	3.679	1	0.012	0.095	0	-2.581
PCDF077	Training	6.460	5.726	0.036	3.381	1	-0.139	-0.053	0	-2.407
PCDF079	Training	7.600	6.857	0.046	3.629	1	-0.003	0.144	0	-2.352
PCDF080	Training	6.700	6.305	0.035	3.578	1	0.012	0.095	0	-2.491
PCDF081	Prediction	7.390	7.426	0.077	3.818	1	0.149	0.292	0	-2.415
PCDF088	Training	5.510	5.060	0.069	3.525	1	-0.143	-0.193	0	-2.866
PCDF089	Training	6.700	6.239	0.065	3.653	1	-0.321	-0.193	0	-2.708
PCDF091	Training	7.170	5.820	0.029	3.565	1	-0.141	-0.118	0	-2.519
PCDF093	Prediction	6.700	6.746	0.050	3.700	1	0.002	0.138	0	-2.647
PCDF095	Training	5.890	6.102	0.049	3.645	1	0.000	0.062	0	-2.804
PCDF098	Training	4.700	5.068	0.066	3.477	1	0.040	-0.043	0	-2.709
PCDF103	Training	6.920	5.761	0.063	3.524	1	0.000	0.062	0	-2.822
PCDF104	Prediction	7.130	6.748	0.052	3.744	1	-0.181	-0.013	0	-2.771

Table 4.6. Continued.

Chemicals	**Status**	**Exp. pIC_{50}**	**Pred. pIC_{50} by model Eq. 4.1**	**HAT i/i (*h**=0.241)**	**RgGrav__3D**	**M_RNG**	**MATS5m**	**MATS5v**	**F09[C-Br]**	**Mor03v**
PCDF107	Prediction	6.400	5.579	0.045	3.561	1	-0.141	-0.118	0	-2.787
PCDF112	Training	7.820	7.760	0.114	3.850	1	-0.038	0.243	0	-2.707
PCDF117	Training	5.080	5.813	0.074	3.692	1	-0.251	-0.182	0	-3.068
PCDF123	Training	6.570	6.416	0.108	3.813	1	-0.411	-0.262	0	-3.028
PCDF128	Training	7.330	7.730	0.169	3.875	1	-0.258	0.105	0.00	-3.049
PCDF134	Training	6.640	7.306	0.133	3.796	1	-0.098	0.186	0.00	-3.020

Eq 4.1 was employed to test 964 external chemicals to predict their pIC_{50} values (Figure 4.6). These group chemicals consisted of 194 PCBs, 50 PCB derivatives, 163 PBBs, 193 PBDEs, 42 PBDE derivatives, 31 PCDEs, 73 PBDD/PCDDs, 107 PCDF/PBDFs, 9 PCPTs, 9 PCTAs, 9 PCDTSs, 19 CDPSs and 65 PAHs that are environmentally significant. Moreover, up to 90 percent in the external set do not have an experimental AhR value.

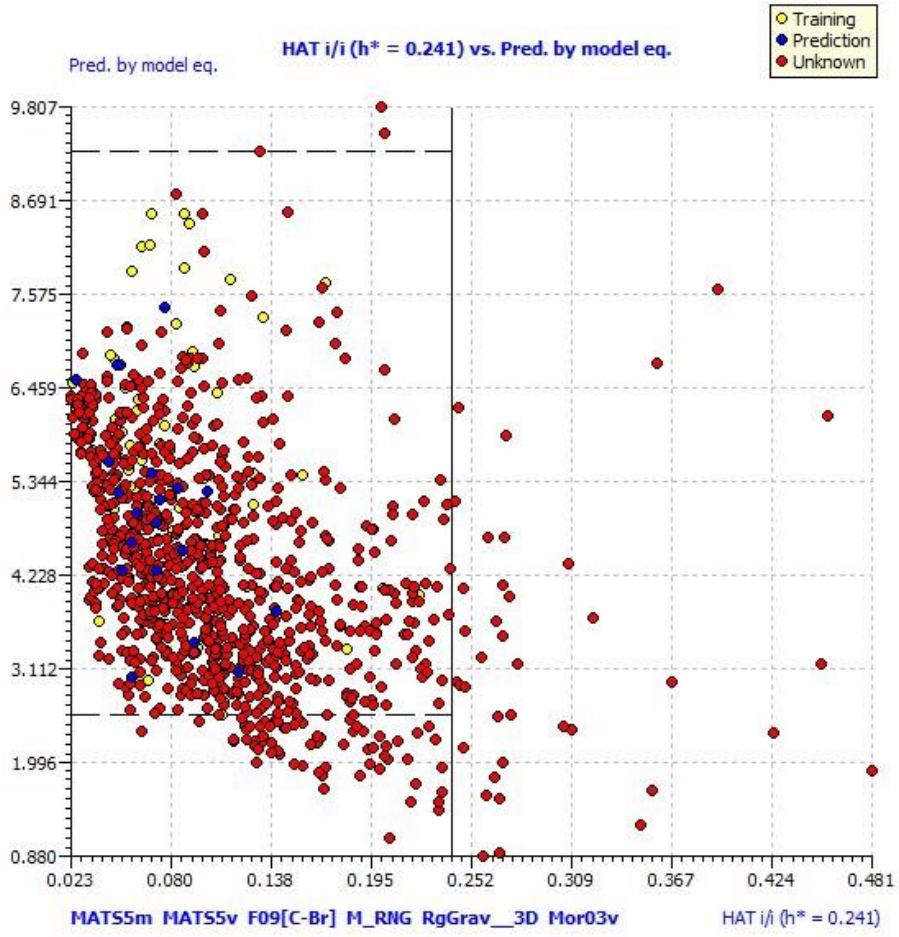


Figure 4.6. Insubria graph of the QSTR model generated using TCDD-normalized data set; hat values and predicted pIC_{50} values of training, test and external sets chemicals; training set in yellow, test set in blue and external set in red.

Among 961 compounds, 38 were out of the structural applicability domain as their hat values exceeded the critical hat value (Figure 4.7). Hat values of 29 chemicals were between 0.241 and 0.341, so they can still consider being reliable. However, 9 chemicals had hat values higher than 0.341, which makes their predictions unreliable. In addition to that, 80 chemicals fell out of the response range. The remaining 843 chemicals were in the applicability domain thus; their predicted pIC_{50} values were accepted as reliable. Predicted

pIC_{50} from Eq. 4.1 and descriptor values of external set chemicals with no TCDD-normalized AhR data are given in External Set 1 in Appendix A1. This model has 95.55% structural coverage. It is of our interest to examine the predictive performance of Eq. 4.1 for chemicals in the external set in more detail.

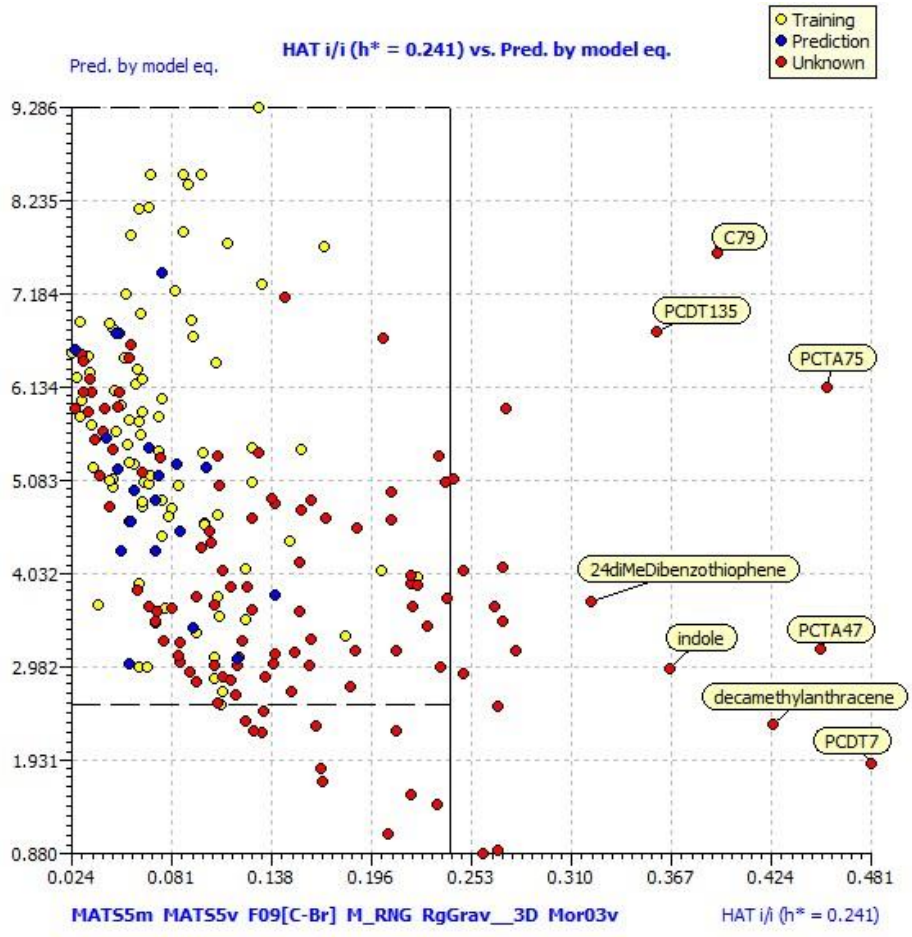


Figure 4.7. Insubria Graph of Eq.4.1 including PAHs and their derivatives with fused heterocyclic rings as an external set. Predicted pIC_{50} values of training, test and external set chemicals from Eq.4.1 and their hat values, where the critical hat value (h^*) is 0.241. C79 is 2,3,6,7-tetrabromonaphthalene.

Many of the compounds that fell outside of the applicability domain belonged to polyaromatic hydrocarbon, polychlorinated thianthrene and dibenzothiophene groups (Figure 4.7). That is fairly understandable considering that these chemicals are quite different from the chemicals in the TCDD-normalized data set. However, the TCDF-normalized data set contained polyaromatic hydrocarbons; therefore, these compounds were included in the external set to test the predictive performance of Eq 4.1. The reason behind testing these irrelevant compounds in this model is surely to see how well Eq. 4.1 can predict pIC_{50} of these chemicals. It can be observed that Eq.4.1 can do reliable predictions for this group of chemicals, with few exceptions, although there are no structurally similar chemicals in the TCDD-normalized data set.

The Insubria graph in Figure 4.8 shows the predictions of pIC_{50} of different congeners and derivatives of PBBs, PCBs and PCB derivatives. Among 400 compounds in these groups, 13 were out of the structural applicability domain. Moreover, only hat values of PBB 54 (2,2',6,6'- tetrabromobiphenyl) and PCB 54 (2,2',6,6'- tetrachlorobiphenyl) were higher than 0.341, which makes the remaining 11 predictions still highly reliable. In addition to that, the predicted pIC_{50} values of PBB 54 and PCB 54 which were 1.677 and 1.269, respectively was below the minimum experimental pIC_{50} value (1.72) of TCDD-normalized data set.

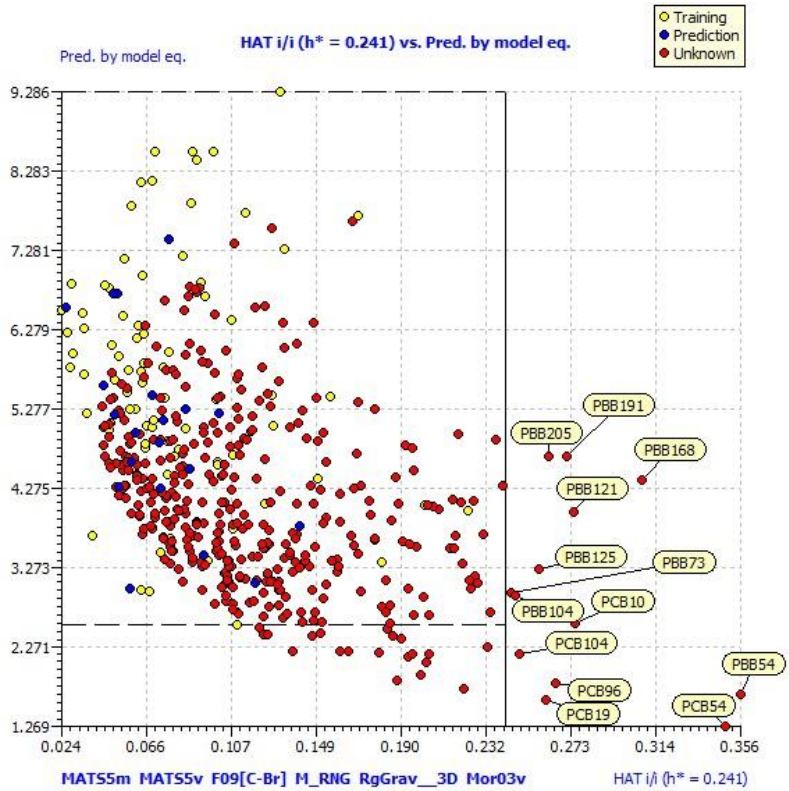


Figure 4.8. Insubria graph indicating the predicted pIC_{50} values of chemicals from Eq. 4.1 for training, test and external (PBBs, PCBs and PCB derivatives) sets.

The Insubria graph in Figure 4.9 shows the predictions of pIC_{50} value of different congeners of PCDDs and PCDFs from Eq.4.1. Among 107 chemicals, only the hat value of Octabromo-dibenzo-p-dioxin (OBDD) (0.245) is slightly higher than the critical hat value (h^*= 0.241). Since this value is between 0.24 and 0.341 it can be considered reliable. TCDD-normalized data set did not include many PBDD compounds and therefore descriptors in the model are not exact representatives of this group. Due to this, the model may not been able to predict a compound that is OBDD, which is highly substituted with bromine atoms, 8 bromine atoms to be exact.

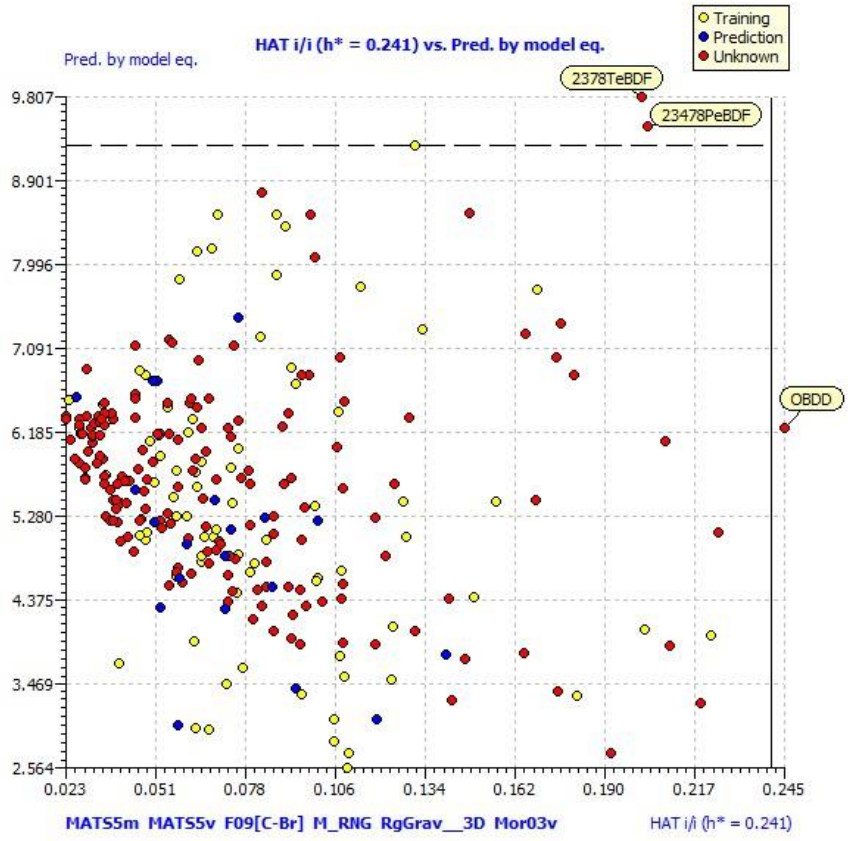


Figure 4.9. Insubria Graph of Eq.4.1 including PCDDs, PBDDs and PCDFs as an external set. Predicted pIC_{50} values of training, test and external set chemicals from Eq.4.1 and their hat values, where the critical hat value (*h**) is 0.241.

Finally, pIC_{50} values of 266 chemicals from PCDE and PBDE groups and hydroxylated and methoxylated derivatives of PBDEs were predicted from Eq. 4.1 (Figure 4.10). This model was reliable for predicting the pIC_{50} values of the ether groups with a few exceptions. Hat values of 261 congeners were lower than the critical hat value of 0.241. Hat values of the remaining 5 were between 0.241 and 0.341, so their predictions can be still reliable. The reason for that these compounds fell out of this model's applicability domain could be due the high amount of halogen groups. In addition to that, methoxy group does not occur in the data set, so the descriptors selected during model building step may not be representatives of methoxy groups. 30 compounds in the diphenyl ether group were out of the response range. The predicted pIC_{50} values for two of the PBDE derivatives (6-methoxy BDE-137 and 5-chloro-6-methoxy BDE 47) were below (1.5284 and 1.658 respectively) the minimum pIC_{50} value of the data set, which is 1.72.

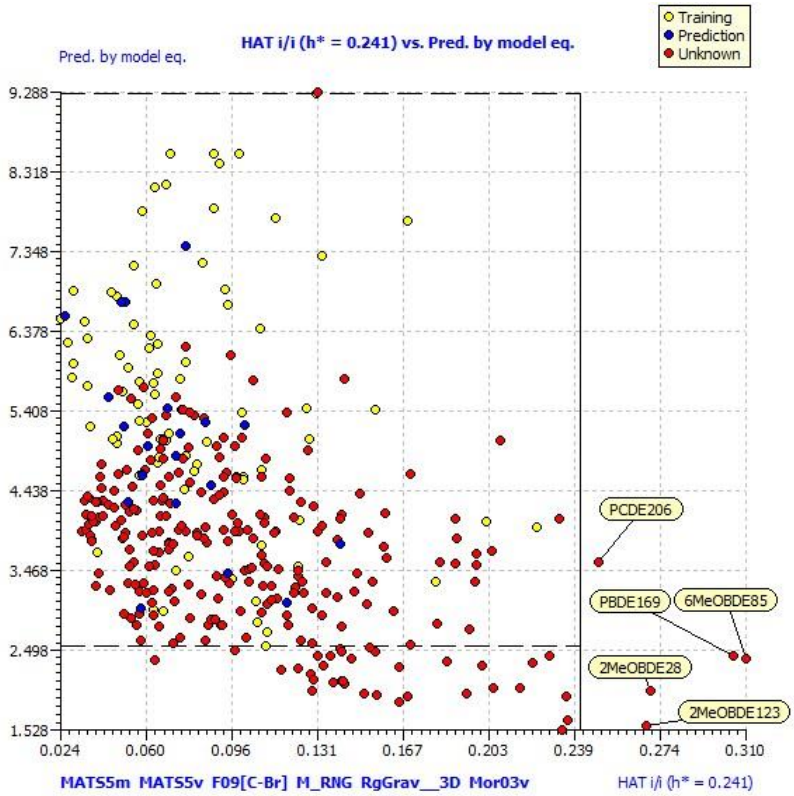


Figure 4.10. Insubria Graph of Eq.4.1 including PCDEs, PBDEs and PBDE derivatives as an external set. Predicted pIC_{50} values of training, test and external set chemicals from Eq.4.1 and their hat values, where the critical hat value (h^*) is 0.241.

4.4.2. Applicability domain for the QSTR model of the TCDF-normalized data set

Williams plot of the QSTR model generated for chemicals with TCDF-like effects was given in Figure 4.11.

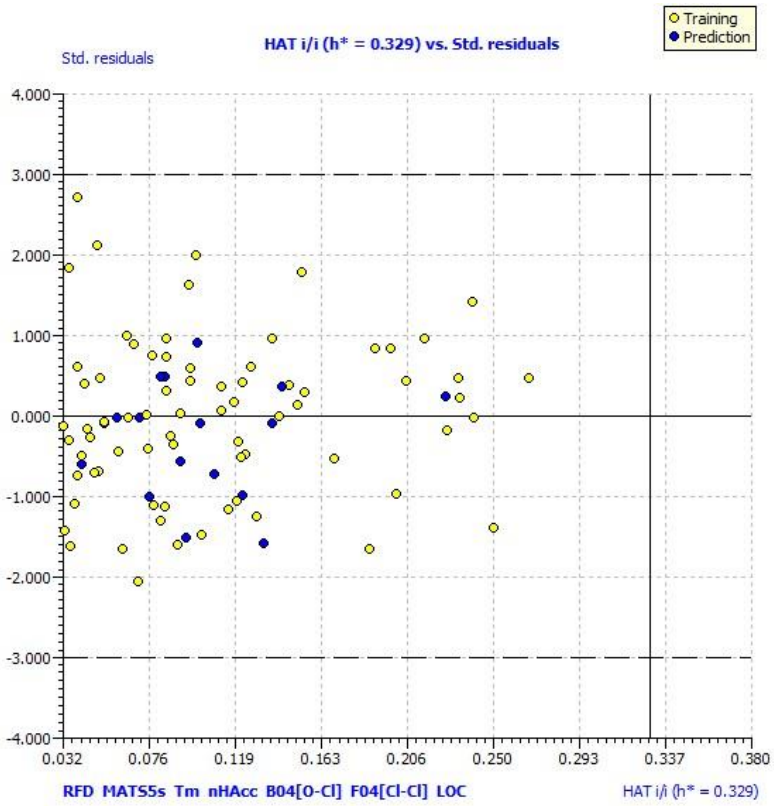


Figure 4.11. Williams plot for the QSTR model (Eq.4.2) generated by using TCDF-normalized data set, with training set in yellow and test set in blue.

Just like it was in Eq.4.1, hat values of all the chemicals in the TCDF-normalized data set are lower than the critical hat value 0.329. In addition to that, none of the chemicals are response outlier, where the response outlier limit was set to $3\delta = 3$. Those two facts show that pIC_{50} values for all of the chemicals were well predicted by the model equation (Eq. 4.2). Experimental and predicted pIC_{50} from Eq. 4.2, and descriptor values of training and test set chemicals in TCDF-normalized data set are given in Table 4.7.

Table 4.7. Chemicals that are used to model TCDF-normalized data set, their experimental and predicted pIC_{50} values, hat and descriptor values.

Chemicals	Status	Exp. pIC_{50}	Pred. pIC_{50} from Eq. 4.2	HAT i/i (*h**=0.3288)	MATS5s	RFD	Tm	nHAcc	B04[O-Cl]	F04[Cl-Cl]	LOC
1,2,3,4,6,7-hexabromonaphthalene	Training	7.608	7.639	0.092	-0.183	0.200	14.524	0	0	0	0.781
1,2,3,5,6,7-hexabromonaphthalene	Prediction	7.996	7.641	0.092	-0.183	0.200	14.527	0	0	0	0.781
1,2,4,6,7-pentabromonaphthalene	Training	7.465	6.748	0.078	0.091	0.200	13.764	0	0	0	0.789
2,3,6,7-tetrabromonaphthalene	Training	7.668	7.690	0.075	0.043	0.200	15.134	0	0	0	0.788
2,3,6,7-tetrachloro-dibenzo-p-dioxin	Prediction	7.768	7.762	0.071	0.180	0.286	14.644	2	1	0	0.714
2,3,6-tetrachloro-dibenzo-p-dioxin	Training	7.610	6.891	0.084	0.085	0.286	12.986	2	1	0	0.682
2,3-dibromonaphthalene	Training	5.616	5.729	0.119	0.263	0.200	12.425	0	0	0	0.730
4-methylindolo[3,2-b]carbazole	Training	7.721	8.843	0.153	0.134	0.400	15.510	0	0	0	0.517
5,11-diacetylindolo[3,2-b]carbazole	Training	7.951	7.140	0.250	-0.060	0.400	15.082	2	0	0	0.881
5,11-diethylindolo[3,2-b]carbazole	Training	8.051	8.321	0.206	-0.210	0.400	14.485	0	0	0	0.789

Table 4.7. Continued.

Chemical	Status	Exp. pIC_{50}	Pred. pIC_{50} from Eq. 4.2	HAT i/i (*h**=0.3288)	MATS5s	RFD	Tm	nHAcc	B04[O-Cl]	F04[Cl-Cl]	LOC
5,11-dimethylindolo[3,2-b]carbazole	Training	8.921	8.139	0.130	-0.166	0.400	13.717	0	0	0	0.517
7,12-dimethylbenz[a]anthracene	Prediction	6.857	5.892	0.095	-0.175	0.333	10.583	0	0	0	0.548
Benz[a]anthracene	Training	7.319	6.324	0.188	0.019	0.333	10.532	0	0	0	0.000
Benzo[1,2-b:4,5-b']bis[1]benzothiophene	Prediction	8.482	8.629	0.226	-0.130	0.400	13.462	0	0	0	0.000
Benzo[1,2-b:4,5-b']bisbenzofuran	Training	7.538	6.957	0.201	0.066	0.400	13.195	2	0	0	0.000
Dibenz[a,h]anhracene	Training	8.602	8.886	0.232	-0.029	0.364	14.290	0	0	0	0.000
Indolo[3,2-b]carbazole	Training	8.444	8.350	0.226	0.050	0.400	13.434	0	0	0	0.000
PBCDD076	Training	10.093	10.367	0.123	0.442	0.286	19.342	2	1	0	0.714
PBCDD077	Training	10.687	10.394	0.124	0.442	0.286	19.385	2	1	0	0.714
PBCDD078	Training	9.074	9.465	0.097	0.430	0.286	17.895	2	1	0	0.714
PBDD002	Prediction	7.464	6.470	0.134	-0.080	0.286	13.820	2	0	0	0.521
PBDD011	Prediction	8.927	8.473	0.109	0.115	0.286	17.641	2	0	0	0.625
PBDD020	Training	10.209	9.547	0.121	0.299	0.286	19.873	2	0	0	0.682
PBDD028	Training	8.038	8.557	0.198	-0.155	0.286	17.347	2	0	0	0.714
PBDD035	Prediction	9.943	9.329	0.124	0.210	0.286	19.396	2	0	0	0.714
PBDD053	Training	8.881	8.937	0.112	0.312	0.286	19.045	2	0	0	0.730
PBDD061	Training	9.350	9.966	0.138	0.383	0.286	20.827	2	0	0	0.730
PBDE045	Training	10.086	10.281	0.154	0.480	0.286	21.511	2	0	0	0.714

Table 4.7. Continued.

Chemical	Status	Exp. pIC_{50}	Pred. pIC_{50} from Eq. 4.2	HAT i/i (*h**=0.3288)	MATS5s	RFD	Tm	nHAcc	B04[O-Cl]	F04[Cl-Cl]	LOC
PCB047	Training	4.442	4.121	0.123	-0.021	0.000	12.042	0	0	2	0.750
PCB061	Prediction	4.405	4.354	0.138	0.100	0.000	12.684	0	0	2	0.750
PCB076	Training	5.584	5.387	0.121	-0.262	0.000	12.642	0	0	1	0.750
PCB077	Training	7.028	7.034	0.142	0.056	0.000	15.118	0	0	0	0.750
PCB081	Training	5.204	6.497	0.100	0.053	0.000	15.108	0	0	1	0.750
PCB105	Training	6.134	6.371	0.113	-0.192	0.000	14.372	0	0	1	0.760
PCB114	Training	6.157	5.421	0.116	0.081	0.000	14.343	0	0	2	0.760
PCB118	Training	5.762	6.051	0.097	0.035	0.000	14.387	0	0	1	0.760
PCB126	Training	7.871	6.936	0.103	0.025	0.000	15.757	0	0	1	0.760
PCB153	Prediction	4.689	4.931	0.143	0.171	0.000	13.776	0	0	2	0.760
PCB156	Prediction	6.057	6.011	0.102	-0.029	0.000	15.020	0	0	2	0.760
PCB167	Prediction	5.482	6.072	0.100	-0.074	0.000	15.014	0	0	2	0.760
PCB168	Training	4.577	4.853	0.268	-0.547	0.000	13.700	0	0	4	0.760
PCB189	Training	5.885	5.978	0.151	-0.144	0.000	15.535	0	0	3	0.754
PCDD01	Training	4.572	4.712	0.233	-0.295	0.286	8.344	2	1	0	0.521
PCDD012	Prediction	6.281	6.604	0.084	0.233	0.286	12.751	2	1	0	0.625
PCDD014	Training	5.585	5.332	0.076	0.144	0.286	11.510	2	1	1	0.682
PCDD019	Training	6.975	7.244	0.043	0.180	0.286	14.670	2	1	1	0.714
PCDD020	Prediction	8.171	7.525	0.076	0.343	0.286	14.574	2	1	0	0.682
PCDD040	Training	6.728	5.709	0.090	0.222	0.286	13.192	2	1	2	0.714
PCDD045	Training	9.144	8.305	0.082	0.421	0.286	16.048	2	1	0	0.714
PCDD053	Training	6.811	7.125	0.051	0.284	0.286	14.755	2	1	1	0.730
PCDD054	Training	8.118	7.948	0.047	0.249	0.286	15.971	2	1	1	0.730
PCDD059	Training	5.937	6.567	0.085	0.281	0.286	14.712	2	1	2	0.730

Table 4.7. Continued.

Chemical	Status	Exp. pIC_{50}	Pred. pIC_{50} from Eq. 4.2	HAT i/i (*h**=0.3288)	MATS5s	RFD	Tm	nHAcc	B04[O-Cl]	F04[Cl-Cl]	LOC
PCDD066	Training	7.490	7.268	0.089	0.327	0.286	15.935	2	1	2	0.737
PCDD075	Training	5.715	6.559	0.239	0.045	0.286	15.853	2	1	4	0.733
PCDF000	Training	3.429	3.420	0.240	-0.259	0.308	6.417	1	0	0	0.000
PCDF002	Training	4.061	3.742	0.169	0.233	0.308	9.213	1	0	0	0.544
PCDF003	Training	5.003	6.056	0.096	-0.152	0.308	9.760	1	1	0	0.544
PCDF004	Training	3.429	3.678	0.147	-0.140	0.308	8.260	1	0	0	0.544
PCDF012	Training	6.088	6.673	0.068	0.138	0.308	11.616	1	1	0	0.648
PCDF014	Training	4.125	4.517	0.127	0.138	0.308	10.439	1	0	0	0.648
PCDF016	Training	4.103	4.619	0.190	0.413	0.308	11.227	1	0	0	0.648
PCDF028	Training	6.123	6.339	0.084	-0.353	0.308	10.929	1	1	1	0.703
PCDF029	Training	4.653	6.460	0.040	-0.133	0.308	11.623	1	1	1	0.703
PCDF036	Training	5.396	6.625	0.035	0.052	0.308	12.304	1	1	1	0.703
PCDF039	Training	6.858	7.354	0.078	0.280	0.308	13.130	1	1	0	0.703
PCDF046	Prediction	7.255	7.248	0.060	0.056	0.308	12.452	1	1	0	0.703
PCDF049	Training	7.610	6.532	0.063	-0.278	0.308	12.310	1	1	2	0.732
PCDF057	Prediction	5.715	6.043	0.082	0.257	0.308	11.918	1	1	1	0.732
PCDF065	Training	7.954	7.468	0.040	-0.128	0.308	13.283	1	1	1	0.732
PCDF071	Training	7.379	6.938	0.051	-0.211	0.308	12.259	1	1	1	0.732
PCDF072	Training	7.610	7.287	0.042	0.162	0.308	13.659	1	1	1	0.732
PCDF077	Training	7.379	7.182	0.036	-0.100	0.308	12.896	1	1	1	0.732
PCDF079	Training	8.689	7.749	0.033	-0.016	0.308	13.981	1	1	1	0.732
PCDF080	Prediction	7.657	7.262	0.042	0.159	0.308	13.614	1	1	1	0.732
PCDF081	Training	8.444	8.434	0.065	0.179	0.308	14.662	1	1	0	0.732

Table 4.7. Continued.

Chemical	Status	Exp. pIC_{50}	Pred. pIC_{50} from Eq. 4.2	HAT i/i (*h**=0.3288)	MATS5s	RFD	Tm	nHAcc	B04[O-Cl]	F04[Cl-Cl]	LOC
PCDF003	Training	5.003	6.056	0.096	-0.152	0.308	9.760	1	1	0	0.544
PCDF004	Training	3.429	3.678	0.147	-0.140	0.308	8.260	1	0	0	0.544
PCDF012	Training	6.088	6.673	0.068	0.138	0.308	11.616	1	1	0	0.648
PCDF014	Training	4.125	4.517	0.127	0.138	0.308	10.439	1	0	0	0.648
PCDF016	Training	4.103	4.619	0.190	0.413	0.308	11.227	1	0	0	0.648
PCDF028	Training	6.123	6.339	0.084	-0.353	0.308	10.929	1	1	1	0.703
PCDF029	Training	4.653	6.460	0.040	-0.133	0.308	11.623	1	1	1	0.703
PCDF036	Training	5.396	6.625	0.035	0.052	0.308	12.304	1	1	1	0.703
PCDF039	Training	6.858	7.354	0.078	0.280	0.308	13.130	1	1	0	0.703
PCDF046	Prediction	7.255	7.248	0.060	0.056	0.308	12.452	1	1	0	0.703
PCDF049	Training	7.610	6.532	0.063	-0.278	0.308	12.310	1	1	2	0.732
PCDF057	Prediction	5.715	6.043	0.082	0.257	0.308	11.918	1	1	1	0.732
PCDF065	Training	7.954	7.468	0.040	-0.128	0.308	13.283	1	1	1	0.732
PCDF071	Training	7.379	6.938	0.051	-0.211	0.308	12.259	1	1	1	0.732
PCDF072	Training	7.610	7.287	0.042	0.162	0.308	13.659	1	1	1	0.732
PCDF077	Training	7.379	7.182	0.036	-0.100	0.308	12.896	1	1	1	0.732
PCDF079	Training	8.689	7.749	0.033	-0.016	0.308	13.981	1	1	1	0.732
PCDF080	Prediction	7.657	7.262	0.042	0.159	0.308	13.614	1	1	1	0.732
PCDF081	Training	8.444	8.434	0.065	0.179	0.308	14.662	1	1	0	0.732
PCDF088	Training	6.297	6.258	0.054	-0.014	0.308	12.509	1	1	2	0.745
PCDF091	Training	8.194	7.130	0.036	-0.102	0.308	12.838	1	1	1	0.745
PCDF093	Training	7.657	7.577	0.032	-0.014	0.308	13.743	1	1	1	0.745
PCDF095	Training	6.728	7.136	0.040	0.136	0.308	13.391	1	1	1	0.745

Table 4.7. Continued.

Chemicals	Status	Exp. pIC_{50}	Pred. pIC_{50} from Eq. 4.2	HAT i/i (*h**=0.3288)	MATS5s	RFD	Tm	nHAcc	B04[O-Cl]	F04[Cl-Cl]	LOC
PCDF098	Training	5.371	6.026	0.065	0.051	0.308	12.292	1	1	2	0.745
PCDF028	Training	6.123	6.339	0.084	-0.353	0.308	10.929	1	1	1	0.703
PCDF029	Training	4.653	6.460	0.040	-0.133	0.308	11.623	1	1	1	0.703
PCDF036	Training	5.396	6.625	0.035	0.052	0.308	12.304	1	1	1	0.703
PCDF039	Training	6.858	7.354	0.078	0.280	0.308	13.130	1	1	0	0.703
PCDF046	Prediction	7.255	7.248	0.060	0.056	0.308	12.452	1	1	0	0.703
PCDF049	Training	7.610	6.532	0.063	-0.278	0.308	12.310	1	1	2	0.732
PCDF057	Prediction	5.715	6.043	0.082	0.257	0.308	11.918	1	1	1	0.732
PCDF065	Training	7.954	7.468	0.040	-0.128	0.308	13.283	1	1	1	0.732
PCDF071	Training	7.379	6.938	0.051	-0.211	0.308	12.259	1	1	1	0.732
PCDF072	Training	7.610	7.287	0.042	0.162	0.308	13.659	1	1	1	0.732
PCDF077	Training	7.379	7.182	0.036	-0.100	0.308	12.896	1	1	1	0.732
PCDF079	Training	8.689	7.749	0.033	-0.016	0.308	13.981	1	1	1	0.732
PCDF080	Prediction	7.657	7.262	0.042	0.159	0.308	13.614	1	1	1	0.732
PCDF081	Training	8.444	8.434	0.065	0.179	0.308	14.662	1	1	0	0.732
PCDF088	Training	6.297	6.258	0.054	-0.014	0.308	12.509	1	1	2	0.745
PCDF091	Training	8.194	7.130	0.036	-0.102	0.308	12.838	1	1	1	0.745
PCDF093	Training	7.657	7.577	0.032	-0.014	0.308	13.743	1	1	1	0.745
PCDF095	Training	6.728	7.136	0.040	0.136	0.308	13.391	1	1	1	0.745
PCDF098	Training	5.371	6.026	0.065	0.051	0.308	12.292	1	1	2	0.745
PCDF049	Training	7.610	6.532	0.063	-0.278	0.308	12.310	1	1	2	0.732
PCDF057	Prediction	5.715	6.043	0.082	0.257	0.308	11.918	1	1	1	0.732

Table 4.7. Continued.

Chemicals	Status	Exp. pIC_{50}	Pred. pIC_{50} from Eq. 4.2	HAT i/i (*h**=0.3288)	MATS5s	RFD	Tm	nHAcc	B04[O-Cl]	F04[Cl-Cl]	LOC
PCDF065	Training	7.954	7.468	0.040	-0.128	0.308	13.283	1	1	1	0.732
PCDF071	Training	7.379	6.938	0.051	-0.211	0.308	12.259	1	1	1	0.732
PCDF072	Training	7.610	7.287	0.042	0.162	0.308	13.659	1	1	1	0.732
PCDF077	Training	7.379	7.182	0.036	-0.100	0.308	12.896	1	1	1	0.732
PCDF079	Training	8.689	7.749	0.033	-0.016	0.308	13.981	1	1	1	0.732
PCDF080	Prediction	7.657	7.262	0.042	0.159	0.308	13.614	1	1	1	0.732
PCDF081	Training	8.444	8.434	0.065	0.179	0.308	14.662	1	1	0	0.732
PCDF088	Training	6.297	6.258	0.054	-0.014	0.308	12.509	1	1	2	0.745
PCDF091	Training	8.194	7.130	0.036	-0.102	0.308	12.838	1	1	1	0.745
PCDF093	Training	7.657	7.577	0.032	-0.014	0.308	13.743	1	1	1	0.745
PCDF095	Training	6.728	7.136	0.040	0.136	0.308	13.391	1	1	1	0.745
PCDF098	Training	5.371	6.026	0.065	0.051	0.308	12.292	1	1	2	0.745
PCDF103	Training	7.911	6.583	0.071	0.129	0.308	13.347	1	1	2	0.745
PCDF104	Training	8.147	8.053	0.045	-0.111	0.308	14.270	1	1	1	0.745
PCDF107	Training	7.313	6.856	0.048	-0.132	0.308	13.181	1	1	2	0.745
PCDF112	Training	8.943	8.233	0.038	0.065	0.308	14.955	1	1	1	0.745
PCDF114	Training	7.657	7.500	0.087	-0.346	0.308	13.706	1	1	2	0.745
PCDF117	Training	5.808	7.218	0.050	-0.128	0.308	13.768	1	1	2	0.749
PCDF123	Training	7.508	7.987	0.085	-0.295	0.308	14.598	1	1	2	0.749
PCDF128	Training	8.376	8.094	0.061	-0.117	0.308	15.173	1	1	2	0.749
PCDF134	Prediction	7.587	7.537	0.053	0.016	0.308	14.599	1	1	2	0.749

Eq. 4.2 was employed to test 977 external compounds to predict their pIC_{50} values (Figure 4.12). These group of chemicals consisted of 195 PCBs, 59 PCB derivatives, 163 PBBs, 206 PBDEs, 42 PBDE derivatives, 31 PCDEs, 68 PBDD/PCDDs, 108 PCDF/PBDFs, 9 PCPTs, 9 PCTAs, 9 PCDTSs, 18 CDPSs and 60 PAHs that are environmentally significant. Moreover, up to 90 percent in the external set do not have an experimental pIC_{50} value.

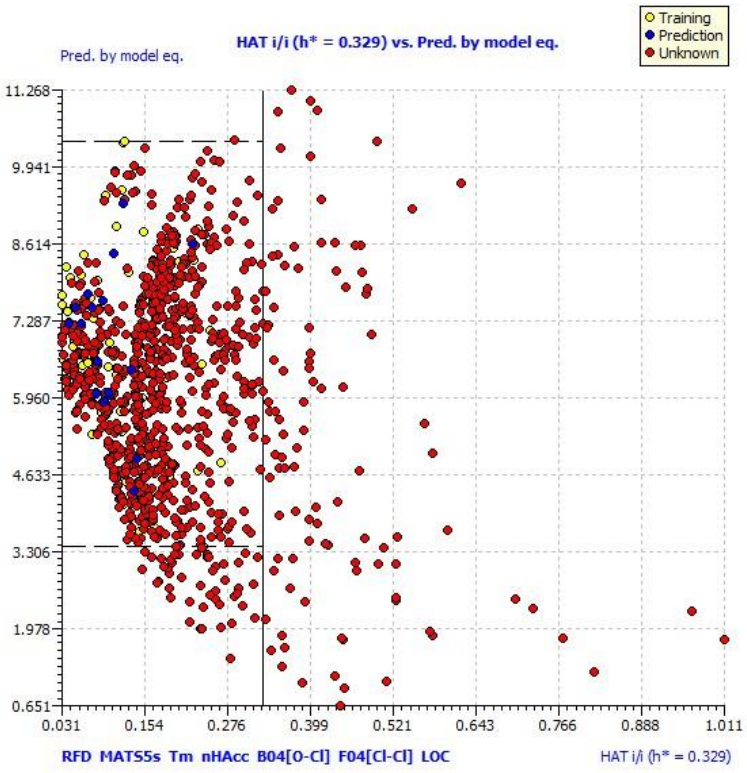


Figure 4.12. Insubria graph of the QSTR model generated using TCDF-normalized data set; hat values and predicted pIC_{50} values of training, test and external sets chemicals; training set in yellow, test set in blue and external set in red.

Of the 977 external set chemicals, 106 were out of the structural applicability domain as their hat values exceeded the critical hat value (h^*= 0.329). 64 of these compounds had hat values between 0.329 and 0.429, so they can still be considered as reliable. However, the 42 remaining chemicals had hat values higher than 0.429, which makes their predictions unreliable. In addition to that, pIC_{50} values of 87 chemicals predicted from Eq.4.2 were out of the response range of the model. Predicted pIC_{50} values of these compounds were below the minimum pIC_{50} value of the training set which was 3.429. The remaining 784 were within the applicability domain thus; their pIC_{50} value predictions are accepted as reliable.

Predicted pIC_{50} from Eq. 4.2 and descriptor values of external set chemicals with no TCDF-normalized AhR data are given in External Set 2 in Appendix A2. This model had 89.37% structural coverage. It is of our interest to examine the predictive performance of Eq. 4.2 for chemicals in the external set in more detail.

This model was better in terms of predicting polyaromatic hydrocarbons and indolocarbazoles (Figure 4.13). That is understandable because there are representatives of these groups in the TCDF-normalized data set and thus the descriptors appeared in the model equation, Eq.4.2, well represent these groups. However, the model was not reliable for predicting chemicals like indole, carbazole and less substituted naphthalenes. In the TCDF-normalized data set almost all polyaromatic hydrocarbons and indolocarbazoles are somewhat substituted and, the chemicals which fell out of the response range are unsubstituted PAHs and indolocarbazoles. This might explain why they are out of the response range with hat values higher than the critical hat value (h^*=0.329).

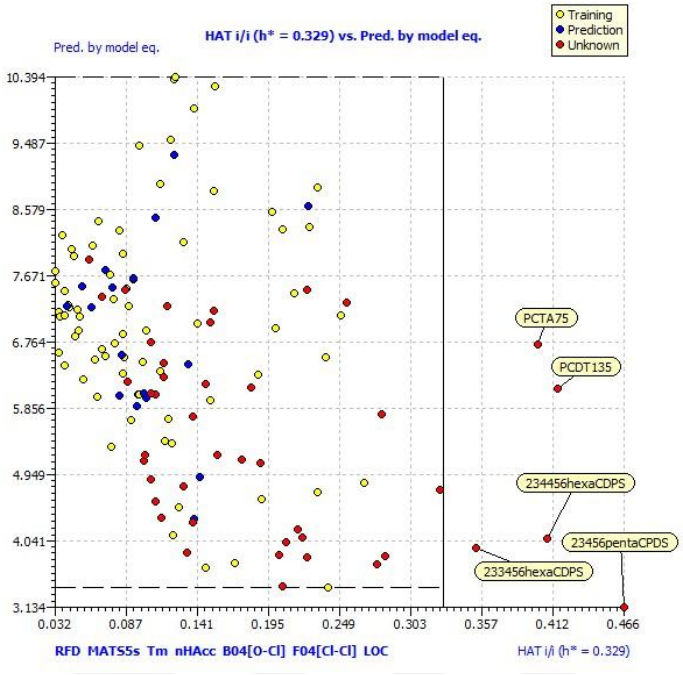


Figure 4.13. Insubria Graph of Eq.4.2 including PAHs and indocarbazoles

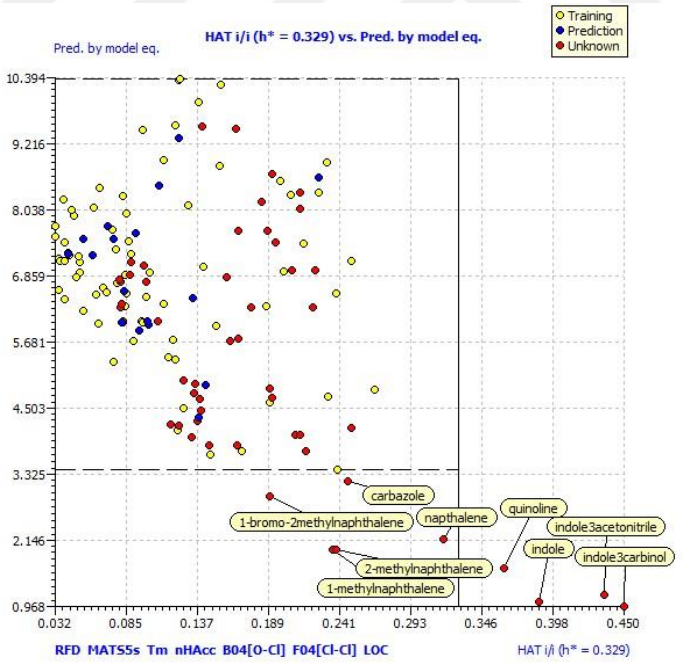


Figure 4. 14. Insubria Graph of Eq.4.2 including PCPTs, PCTAs, PCDTs and PCDPSs as external sets. Predicted pIC50 values of training, test and external set chemicals from Eq.4.2 and their hat values, where the critical hat value (*h**) is 0.329.

One of the pentachlorinated diphenyl sulfides in the external data set has a high hat value, its pIC_{50} prediction was also below the minimum pIC_{50} value of the TCDF-normalized data set which makes its prediction not reliable Figure 4.14. Two of the hexachlorinated diphenyl sulfides also had high hat values, but their hat values were between 0.329 and 0.429 which make their prediction reliable up to some degree. However, pIC_{50} values of other heptachlorinated diphenyl sulfides in the external set were reliable.

The Insubria graph in Figure 4.15 indicates pIC_{50} value predictions of different congeners of PCDDs and PCDFs. Among 107 chemicals, only dibenzo-*p*-dioxin had a hat value that is (0.374) slightly higher than the critical hat value of 0.329. Since this value is below 0.429 it can be considered to be highly reliable. Unlike in the Eq.4.1, in which some of the highly brominated compounds were response or structural outliers, this model, was good to predict the highly halogenated congeners of PCDDs and PCDFs. Two of the descriptors in this model equation, Eq. 4.2, is directly related to the relationship of Cl-Cl and O-Cl bonds in the compound, which might explain this trend. If this is the case it would also explain why PCDD0 was a structural outlier. Since the compound does not any halogen group attached to it, two out of 7 descriptors that explain the model not applicable for this compound.

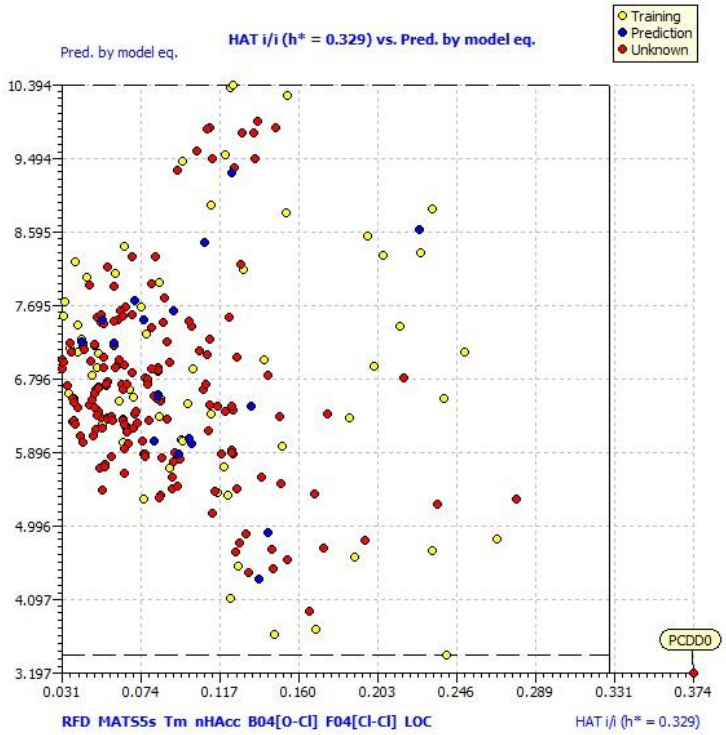


Figure 4.15. Insubria Graph of Eq.4.2 including PCDDs and PCDFs and. Predicted pIC_{50} values of training, test and external set chemicals from Eq.4.2 and their hat values, where the critical hat value (h^*) is 0.329

The Insubria graph in Figure 4.16 shows predicted pIC_{50} value of different congeners and derivatives of PBBs and PCBs. This model equation, Eq.4.2, is not as reliable as the model equation, Eq.4.1, to predict the pIC_{50} values of PBB and PCB congeners. More than 20 were outside of the structural applicability domain and approximately another 20 were out of response range as some of their predicted pIC_{50} values were below the minimum data point and, pIC_{50} prediction for two PCB derivatives (4'-phenyl-2,3,4,5-tetrachlorobiphenyl and 4'-n-butyl-2,3,4,5-tetrachlorobiphenyl) exceeded the maximum experimental data point in the TCDF-normalized data set. The labeled compounds in Figure 4.16 that have unacceptably high leverage ($h^* > 0.429$) have high numbers of halogen atoms attached to the biphenyl structure. One can argue that the descriptors in this model are not representing PCB and PBBs and their derivatives very well, even though training set contains couple of PCB congeners.

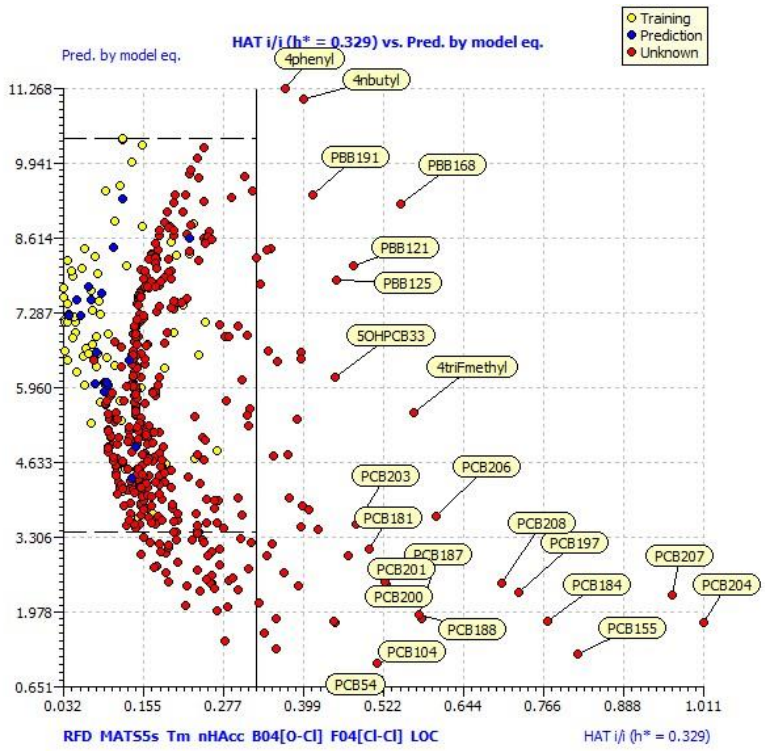


Figure 4.16. Insubria Graph of Eq.4.2 including PCBs, PCB derivatives and PBBs. Predicted pIC50 values of training, test and external set chemicals from Eq.4.2 and their hat values, where the critical hat value (h*) is 0.329.

Finally, 280 compounds from PCDE and PBDE groups and hydroxylated and methoxylated derivatives of PBDEs' pIC_{50} values have been predicted from Eq. 4.2 (Figure 4.17). This model was not as reliable as the TCDD-based model for predicting the ether groups' pIC_{50} values. Hat values of 46 compounds were higher than the critical hat value of 0.329. Hat values of 32 of them were between 0.329 and 0.429, so their predictions can be still reliable. The reason that these compounds fell out of the applicability domain of this model could be due the high number of halogen groups they contain. In addition to that, methoxy group does not occur in the data set, so the descriptors selected during building steps of this model do not represent methoxy groups. Furthermore, only a few compounds fell out of response range as their predicted pIC_{50} values below the minimum pIC_{50} value of the TCDF-normalized data set.

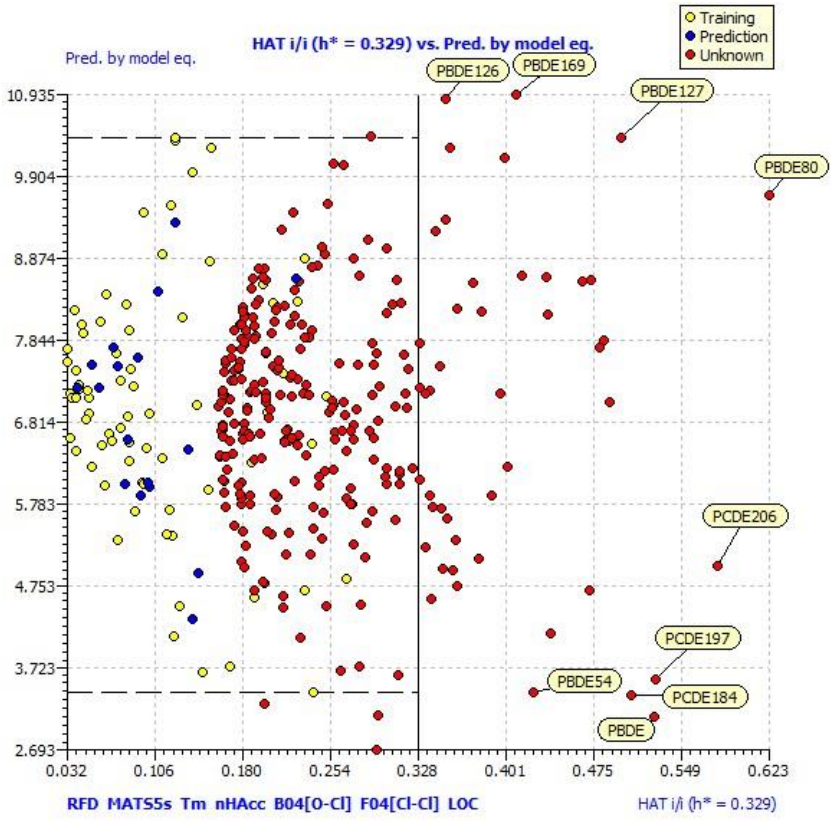


Figure 4.17. Insubria Graph of Eq.4.2 including PCDE, PBDE, and derivatives of PBDE as an external set. Predicted pIC_{50} values of training, test and external set chemicals from Eq.4.2 and their hat values, where the critical hat value (*h**) is 0.329

It has been suggested that para and meta positions of bromine substituted PBDEs result in higher AhR binding affinity. On the other hand, ortho substitution of bromines are thought be disfavored by the AhR and the binding affinity decreases (Papa et al., 2010; Gu et al., 2012).

Regarding the TCDD-normalized data; pIC_{50} values of seven chemicals were higher than TCDD's pIC_{50} value of 8.00. Among these seven compounds, 2,3,7,8-tetrabromo dibenzofuran had the highest pIC_{50} value, 9.807. It was followed by 2,3,4,7,8-pentabromo dibenzofuran (pIC_{50} =9.499), 2,2',3,6-tetrabromodiphenyl ether (pIC_{50} =9.288), 1,2,3,4,7,8-hexabromo dibenzo-*p*-dioxin (pIC_{50} =8.784), 1,2,3,4,7,8-hexabromo dibenzofuran (pIC_{50} =8.560), 1,2,3,7,8-pentabromo dibenzo-*p*-dioxin (pIC_{50} =8.543) and 1,2,3,7,8-pentabromo dibenzofuran (pIC_{50} =8.088). Chemical structures of these compounds are given in Figure 4.18. None of these compounds' hat value were higher than the critical hat value (*h** =

0.241), which suggest their predictions are reliable. In general, dioxins and furans that had bromine atoms in 2,3,7 and 8 positions showed high binding affinity towards AhR.

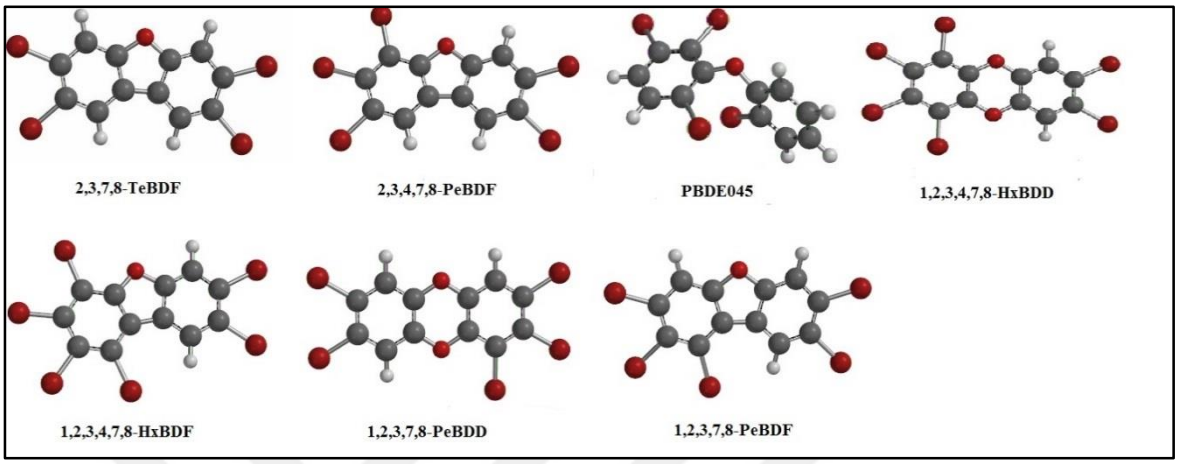


Figure 4.18. Chemical structures of the compounds that have a higher pIC_{50} value than TCDD.

On the contrary, PBDEs, their methoxylated and hydroxylated derivates, PCBs and PAHs (chemicals with $h^* < 0.241$ are taken into consideration) showed low binding affitinity towards TCDD.

Regarding the TCDF-normalized data; pIC_{50} values of 30 compounds were higher than of TCDD and 63 had higher pIC_{50} values than of TCDF. The mentioned compounds and their predicted pIC_{50} obtained from Eq.4.1 and Eq.4.2 are given in Table 4.8.

Table 4.8. Predicted pIC50 values of chemicals that show higher binding affinity compared to TCDF.

No	Name	Experimental pIC_{50}	Pred. pIC_{50} from Eq.4.2	Pred. pIC_{50} from Eq.4.1	Corrected $pIC5_0$ value for TCDF*	HAT i/i (*h**=0.329)
	TCDF	**8.444**				
1	PBDE124		8.482	4.018	4.592	0.224
2	PBDE074		8.500	3.435	3.926	0.187
3	PBB180		8.531	6.650	7.601	0.173
4	PBB193		8.540	4.506	5.150	0.251
5	PBB128		8.592	5.282	6.038	0.251
6	PBDE120		8.597	4.145	4.737	0.228
7	PBB192		8.601	4.105	4.692	0.260
8	PBDE123		8.620	2.441	2.790	0.200
9	PBDE158		8.624	3.492	3.991	0.189
10	PBB122		8.628	5.013	5.730	0.247
11	PBDE194		8.644	2.481	2.836	0.198
12	PBB158		8.654	5.373	6.141	0.254
13	PCDE126		8.675	4.655	5.321	0.278
14	PBB060		8.684	6.066	6.933	0.202
15	PBB206		8.689	4.882	5.580	0.223
16	PBB108		8.753	5.314	6.074	0.256
17	PBDE205		8.755	2.568	2.936	0.194
18	PBDE191		8.757	2.313	2.644	0.199
19	PBB114		8.765	6.539	7.474	0.179
20	PBB190		8.773	5.719	6.537	0.201
21	PBDE107		8.782	4.650	5.315	0.238
22	PBDE012		8.800	2.955	3.377	0.244
23	PBB162		8.812	5.744	6.566	0.194
24	PBB159		8.846	5.529	6.319	0.197
25	PBB170		8.860	5.670	6.481	0.228
26	PBDE038		8.881	2.386	2.727	0.274
27	PBB118		8.910	6.833	7.810	0.187
28	PBDE013		8.930	4.968	5.679	0.249
29	PBDE164		9.011	3.472	3.969	0.302
30	PBDE159		9.031	2.048	2.340	0.248
31	PBB194		9.086	6.119	6.994	0.215
32	PBB079		9.093	6.707	7.666	0.193
33	PBDE039		9.110	4.362	4.985	0.286

Table 4.8. Continued.

No	Name	Experimental pIC_{50}	Pred. pIC_{50} by model Eq.4.2	Pred. pIC_{50} by model Eq.4.1	Corrected pIC_{50} value for TCDF*	HAT i/i (*h**=0.329)
	TCDD	**9.144**				
34	PBDE167		9.246	3.158	3.610	0.213
35	PBB156		9.258	6.820	7.795	0.202
36	PBB105		9.270	6.378	7.290	0.249
37	PBDE054		9.313	4.109	4.696	0.164
38	PBB123		9.358	6.027	6.889	0.290
39	1,2,3,4,7,8-HxBDF		9.365	8.560	9.784	0.094
40	PBB127		9.387	6.579	7.520	0.204
41	PBB167		9.391	6.795	7.767	0.220
42	OBDD		9.397	6.235	7.126	0.125
43	PBDE157		9.456	3.493	3.993	0.223
44	PBB205		9.471	4.687	5.357	0.322
45	Tjipanazole		9.490	6.480	7.407	0.166
46	1,2,3,4,7,8-HxBDD		9.501	8.784	10.040	0.137
47	OBDF		9.504	6.104	6.977	0.113
48	3-methyldibenzo[ah]anthracene		9.531	4.515	5.161	0.141
49	PBDE189		9.561	2.350	2.686	0.253
50	1,2,3,7,8-PeBDF		9.601	8.088	9.245	0.105
51	PBB189		9.701	6.571	7.511	0.240
52	PBB157		9.721	6.105	6.978	0.310
53	1,2,3,4,6,7,8-HpBDF		9.816	6.817	7.792	0.130
54	1,2,3,6,7,8-HxBDD		9.820	7.368	8.422	0.136
55	PBB077		9.839	7.382	8.438	0.228
56	2,3,4,7,8-PeBDF		9.862	9.499	10.857	0.110
57	2,3,7,8-TeBDF		9.878	9.807	11.209	0.112
58	1,2,3,7,8,9-HxBDD		9.888	7.251	8.288	0.148
59	1,2,3,7,8-PeBDD		9.966	8.543	9.765	0.138

Table 4.8. Continued

No	Name	Experimental pIC_{50}	Pred. pIC_{50} from Eq.4.2	Pred. pIC_{50} from Eq.4.1	Corrected $pIC5_0$ value for TCDF*	HAT i/i (*h**=0.329)
60	PBB126		10.054	7.572	8.655	0.238
61	PBDE037		10.068	3.943	4.507	0.257
62	PBB169		10.228	7.666	8.762	0.247
63	PBDE081		10.428	4.153	4.747	0.288

*Original binding affinity data were normalized to TCDF using a scaling factor of 1.143 (Waller and McKinney, 1995).

The predicted values from Eq.4.2 indicated that 4 other chemicals had pIC_{50} values which were higher than the pIC_{50} value of TCDF. These chemicals, 4'-isopropyl-2,3,4,5-tetrachlorobiphenyl, 4'-t-butyl-2,3,4,5-tetrachlorobiphenyl, 2,3',4,4'-tetrabromodiphenyl ether, 3,3',4,4'-tetrabromodiphenyl ether and 2,3,7,8-tetrabromo dibenzo-*p*-dioxin, have their experimental values which enables us to make a comparison between experimental and predicted values. The experimental and predicted pIC_{50} value of PBDD congener was highly comparable. Predictive ability of Eq 4.2. for PBDD congeners is high as it can be seen in Figure 4.15. On the other hand, mentioned diphenyl ether and bulky substituted biphenyl compounds' predicted pIC_{50} were not consistent with the experimental values. This result is coherent with Figure 4.15 and Figure 4.16, as they each show the structural coverage of halogenated biphenyls and halogenated diphenyl ethers, respectively. The low predictive ability of BDEs could be explained by the fact that the TCDF-normalized did not have any representitives of BDE group, and therefore this model did not have any descriptors that would explain this group's properties. For instance F09[C-Br] in Eq 4.1 specifically represents diphenyl ether groups. Additioanally, predicted pIC_{50} values of PBBs, PCDFs, PCDDs from two models are close to one another, which indicate the reliability of predictive ability of the models for these chemical groups (Table 4.8).

Su and colleagues (2012b) claimed HO-PBDEs have greater AhR binding affinity in comparison to MeO-PBDEs, and both of these BDE derivatives show greater potencies to induce AhR. This assumption was supported by comparing predicted pIC_{50} values of

hydroxylated and methoxylated derivatives of BDE-123 (with 2-MeO-BDE and 2-OH-BDE), BDE-28 (with 2-MeO-BDE and 2-OH-BDE) and BDE-68 (with 2-MeO-BDE and 2-OH-BDE). Examples for these groups are given in Figure 4.19.

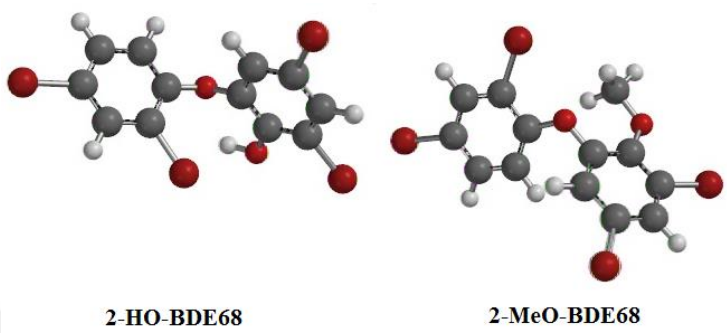


Figure 4.19. Structures of 2-HO-BDE 68 and 2-MeO-BDE68, respectively.

It is widely known that PCB congeners that lack ortho substitution are the most potent to the AhR. Due to their planar structure they can easily fit into the binding site of the receptor, and PCB 126 (3,3',4,4'-penta-CB) shows the most dioxin-like effect among the PCB congeners. Compounds that have ortho substitution tend to be more bulky, and therefore, do not have a planar configuration. These congeners show lower binding affinity to the receptor (Lindén et al., 2010). Both of the generated models in the present study are able to confirm this trend. Many planar PCB congeners in the external set, including PCB 169 (3,3',4,4',5,5'-hexa-CB) and PCB 123 (2,3', 4,4',5'-penta-CB), had high predicted pIC_{50} values.

Cao and colleagues (2013) published a paper on hydroxylated and methoxylated PCBs and their activity on AhR. They discuss that substitution at 3-position of the benzene ring in PCB would result in increased activity and they supported this claim by comparing activities of 4'OH-BDE-35 (3,3',4) and 4- OH-BDE-33 (2,3,4') to 6'OH-BDE-31(2,4',5) and 2'OH-BDE-30 (2,4,6). Both TCDD and TCDF-based models were able to confirm their claims as predicted pIC_{50} values of 4'OH-BDE-35 and 4-OH-BDE-33 were significantly higher compared to 6'OH-BDE-31(2,4',5) and 2'OH-BDE-30(2,4,6). Structures of some of the mentioned chemicals are given in Figure 4.20.

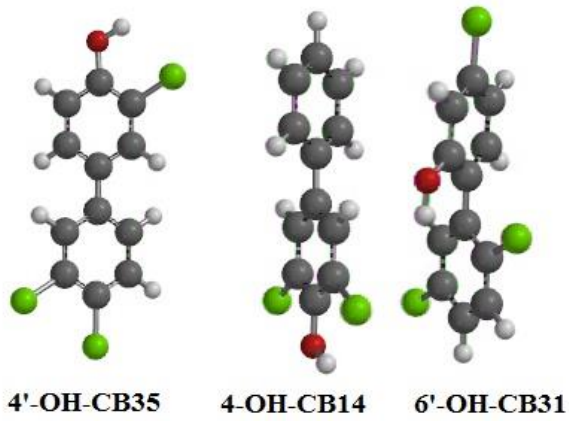


Figure 4.20. Chemical structures of hydroxy substituted polychlorinated biphenyls.

They also claimed that substitution of electropositive groups at the 5- position, and substitution of electronegative groups near the 2'-position would increase the activity. However, none of the models in the present study are able to support this claim. Finally, they state that substitution of hydrophobic groups like methoxy and chlorine to the 4-position carbon on the benzene ring would improve the activity. They tested this claim by comparing activities of 3'-OH-CB-31 and 2'-OH-CB-9 (2, 5) and predicting that 3'-OH-CB-31's activity would be higher. Again, both TCDD and TCDF-based models were able to support this claim.

Three of the polychlorinated diphenyl sulfides (PCDPSs) (2,2',3,3',4,5,6-hepta-CDPS, 2,2',3',4,5-penta-CDPS and 2,4,4',5-tetra-CDPS) are thought to have comparable or higher binding affinities towards AhR compared to many mono and ortho substituted PCBs (Zhang et al., 2016). This claim however, could not be supported with either of the models that have been proposed in the present study.

For polyaromatic hydrocarbons, adding halogens like chlorine and bromine or groups like methyl enhance binding affinity towards the AhR (Lee et al., 2015). Both of the models in this thesis were able to affirm this trend.

4.4.3. Comparison and further discussion on the models of TCDD and TCDF-normalized data sets

Overall, the TCDD-based model had a better structural coverage when compared to TCDF-based model. Hat values of the predicted compounds were mostly lower than the model's critical hat value (h^*=0.241), few compounds had hat values between 0.241 and 0.341 which was still considered to be reliable. Even fewer compounds had hat values that were above the reliable prediction limit. On the other hand, TCDF-based model had more compounds those hat value was over the critical hat value (h^*=0.329). Structural coverage of TCDF-based model and TCDD-based model are given in Figure 4.21 and Figure 4.22, respectively.

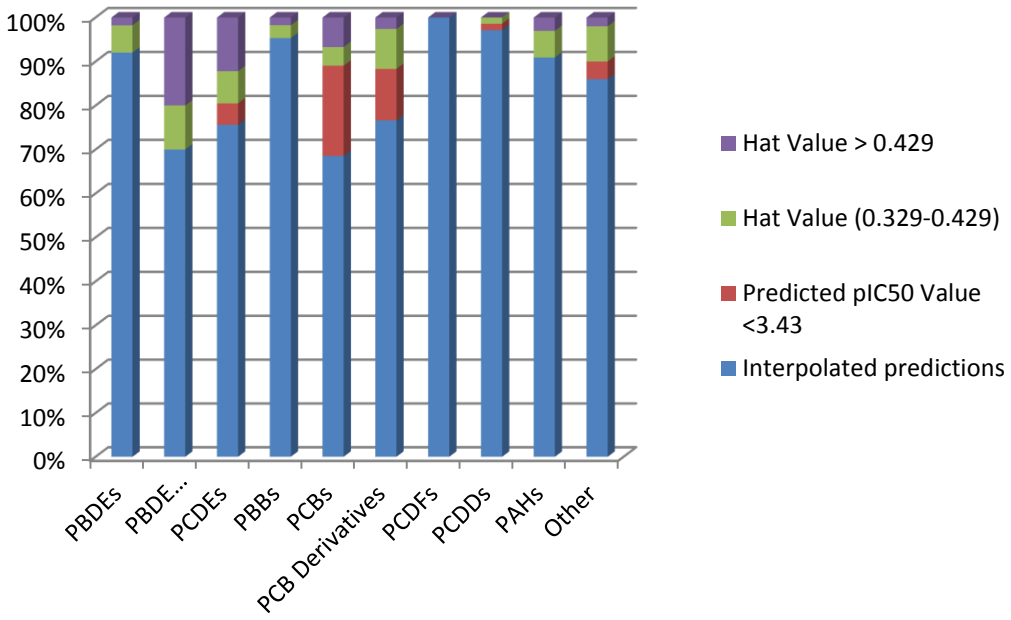


Figure 4.21. Structural coverage of interpolated predictions of each chemical group showing the percentage of compounds that exceed the critical hat values and below the data set range for TCDF-normalized data.

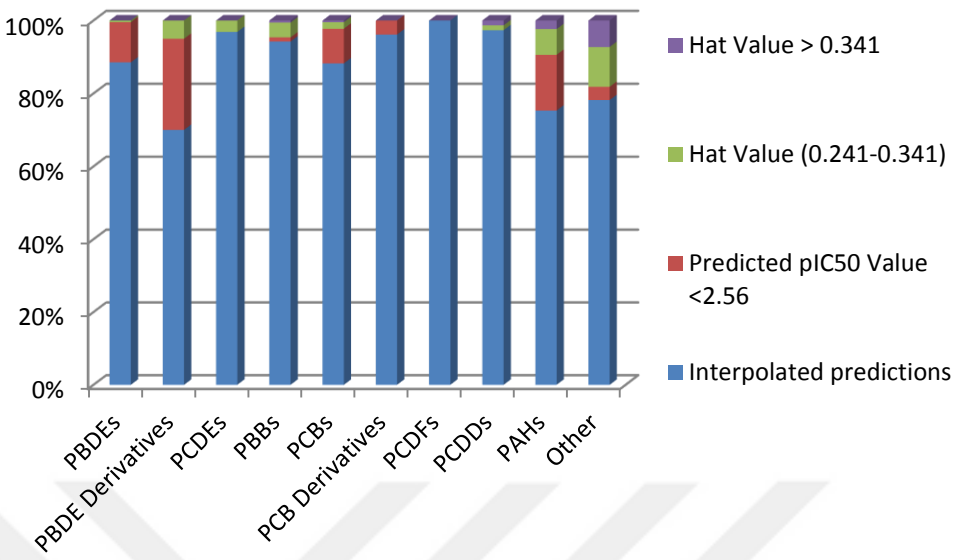


Figure 4.22. Structural coverage of interpolated predictions of each chemical group showing the percentage of compounds that exceed the critical hat values and below the data set range for TCDF-normalized data.

Both models were extremely reliable in predicting PCDDs and PCDFs. The model for TCDD-normalized data was not exceptionally satisfying in predicting PBDE congeners and derivatives as most of the predicted pIC_{50} values of selected compounds were out of the response range of the model. However, it was more reliable in predicting these compounds compared to TCDF-based model since predictions that are made by TCDD-based model either fall within the structural applicability domain, or they have hat values higher than the critical hat value yet within acceptable limits.

TCDD-based model was again more reliable in predicting PCBs and PCB derivatives as the TCDF-based model was not extremely reliable in predicting these compounds. Nevertheless, the TCDF-based model was particularly satisfactory in predicting the pIC_{50} values of selected polyaromatic hydrocarbons.

Statistical parameters of the models developed for TCDD-normalized and TCDF-normalized data sets are compared in Table 4.10. Both models had consisted descriptors from 2D Autocorrelations and 2D Atom Pairs blocks. Tm, from Equation 4.2 is mass related

parameter. TCDD-based model had 2 and, TCDF-based model had 3 indicating variables. For instance M_RNG in Equation 4.1 descriptor has a range between 0-1, F09[C-Br], another descriptor from Equation 4.1, has a range between 0-2. In addition to that, this descriptor just indicates the presence of PBDE groups in the model. F04[Cl-Cl] descriptor in Equation 4.2, on the other hand, can get a value between 0-4. All of these descriptors enhance the statistical metrics of the model and are present in the model due some groups of chemicals. Minimum and maximum values of each descriptor in both models are given in Table 4.9.

Table 4.9. Range of descriptors appeared in Eq.4.1 and Eq.4.2, respectively.

TCDD-based Model (Eq. 4.1)			**TCDF-based Model (Eq. 4.2)**		
Descriptor	**Minimum Value**	**Maximum Value**	**Descriptor**	**Minimum Value**	**Maximum Value**
MATS5m	-0.747	0.482	RFD	0	0.400
MATS5v	-0.640	0.384	MATS5s	-0.547	0.480
F09[C-Br]	0	3	Tm	6.417	21.511
M_RNG	0	1	nHAcc	0	2
RgGrav__3D	1.806	4.660	B04[O-Cl]	0	1
Mor03v	-4.541	-0.511	F04[Cl-Cl]	0	4
			LOC	0	0.881

Table 4.10. Fitting criteria, internal and external validation metrics and descriptors for TCDF and TCDD-based models.

Model	**Descriptors**
TCDF_3	RFD MATS5s Tm nHAcc B04[O-Cl] F04[Cl-Cl] LOC
TCDD_4	MATS5m MATS5v F09[C-Br] M_RNG RgGrav__3D Mor03v

Fitting Parameters

Model	R^2	R^2_{adj}	R^2-R^2_{adj}	K_{xx}	Delta K	$RMSE_{Tr}$	MAE_{Tr}	CCC_{Tr}	s	F
TCDF_3	0.850	0.834	0.016	0.412	0.019	0.638	0.503	0.919	0.677	29.746
TCDD_4	0.840	0.828	0.012	0.413	0.050	0.671	0.563	0.913	0.700	52.493

Internal Validation Metrics

Model	Q^2_{LOO}	R^2-Q^2_{LOO}	RMS_{CV}	MAE_{CV}	$PRESS_{CV}$	CCC_{CV}	R^2_{Yscr}	$RMSE_{AV}$ $_{Yscr}$	Q^2_{Yscr}
TCDF_3	0.815	0.035	0.708	0.563	36.619	0.900	0.097	1.564	-0.144
TCDD_4	0.810	0.030	0.731	0.613	46.477	0.897	0.071	1.617	-0.101

External Validation Metrics

Model	$RMSE_{Test}$	MAE_{Test}	R^2_{Test}	Q^2_{F1}	Q^2_{F2}	Q^2_{F3}	CCC_{Test}	$r^2_{m\ aver.}$	Δr^2_m	k'	k
TCDF_3	0.476	0.365	0.913	0.893	0.893	0.917	0.940	0.817	0.089	0.970	1.027
TCDD_4	0.407	0.320	0.910	0.908	0.898	0.941	0.946	0.869	0.080	0.968	1.028

4.4.4. Comparison of the QSTR models from the present study with the previously published models

To compare our models with the studies that have been published previously is important, since the comparison points out the strengths and weaknesses of the present study. Even though an exact comparison is not possible, as each author use different methods, software and a unique data set with different compounds, it is still vital to compare to contrast. A comparison of some features and parameters of selected models are given in the Table 4.11.

To start with, many of the previous work lack the necessary validation parameters. R^2, Q^2, *s* and *F* values are usually provided, however, these are not enough to assume that the model is valid and robust. In addition to that, data sets in previous studies include various groups of compounds yet, authors chose to model each compound group separately in many occasions. This surely limits the developed model's applicability domain. On the other hand, some authors chose to develop their model with a small number of compounds that belong to the same group. They obtained good R^2 and Q^2 parameters; yet again, their applicability domain is extremely limited as the variables in the model could only represent a small number of compound and not more.

The healthiest comparison that can be made between this study and the study of Ruffa (2013) as the method and the software that has been used are almost the same and the time gap is relatively short (regarding the technological advances in the field, especially regarding the software). In this work, in addition to DRAGON 6.0 software, ADMET 8.0 was employed for descriptor calculation. Added descriptors from the latter software may have enhanced the internal and external validation parameters. In addition to that a newer version of QSARINS software was used in the present study. In general, both models developed in this work seem superior to the previous models.

Table 4.11. Comparison of the statistical parameters of generated models to those of the previously published models.

Chemical groups	**Method**	***N****	R^2	Q^2	Q^2_{LOO}	R^2_{Test}	$RMSE_{Test}$	**Reference**
dioxins, furans, biphenyls, napthalenes, carbazole derivatives	CoMFA	99	0.824	0.453	N/A	N/A	N/A	**Waller an McKinney 1995**
dioxins, furans, biphenyls, napthalenes	CoMFA/ CoMSIA	95	0.9/0.873	0.631/0.711	N/A	N/A	N/A	**Ashek et al., 2006**
dioxins, furans, biphenyls, napthalenes, carbazole derivatives	CoMFA	91	0.910	0.620	0.620	N/A	N/A	**LoPiparo et al., 2006**
BFR	MLR		0.900		0.790	0.730	0.420	**Papa et al., 2010**
dioxins and furans	PLS	60	0.549	0.603	N/A	N/A	N/A	**Diao et. Al, 2010**
dioxins, furans and biphenyls	PLS	65	0.992	0.907	N/A	N/A	0.446	**Li et al., 2011**
diphenyl ethers	PLS	18	0.932	0.894	N/A	N/A	N/A	**Gu et al., 2012**
dioxins, furans and biphenyls	CoMFA	78	0.858	0.684	N/A	N/A	N/A	**Yuan et al., 2014**
dioxins, furans, biphenyls, diphenyl ethers	MLR	109	0.822	N/A	0.792	0.813	0.678	**Ruffa, 2013**
dioxins, furans, biphenyls, diphenyl ethers	MLR	108	0.850	N/A	0.815	0.913	0.476	**Present work (TCDD-based)**
dioxins, furans, biphenyls, diphenyl ethers, napthalenes and carbazole derivatives	MLR	90	0.840	N/A	0.810	0.910	0.407	**Present work (TCDF-based)**

*N** refers to the number of chemicals in the data sets.

5. CONCLUSION

In this study, two QSTR models, one for TCDD-normalized data and one for TCDF-normalized data were developed and validated both internally and externally. Both of the models fully comply with OECD criteria.

TCDD-based model had 6 descriptors and TCDF-based model had 7 descriptors to represent complex data set compounds, comprised of halogenated dioxins, dibenzofurans, biphenyls, biphenyl derivatives, diphenyl ethers, naphthalene and polyaromatic hydrocarbons, on their AhR binding affinity. Descriptors that represented the entire data were quite complex, which prove that AhR binding affinity is too complicated to be explained with a simple pathway. Nevertheless, it is likely to interpret the mechanism of AhR binding affinity of TCDD/F-like chemicals using the information gathered from the definition and sign of the descriptors appeared in the generated models.

The effects of conformation, structural connectivity, compactness, holistic structure, atomic van der Waals volume, frequency and specific position of C-Br group, ring structure type and atomic mass on the AhR binding affinity have been demonstrated by the model equation (Eq. 4.1) generated for TCDD-like chemicals.

The effects of branching and flexibility, size of the molecule, the number of acceptor atoms for H-bonds in the structure, frequency and specific position of O-Cl and Cl-Cl groups, ring fusion density and I-state of the molecules on the AhR binding affinity have been demonstrated by the model equation (Eq. 4.2) generated for TCDF-like chemicals.

The present study was able to affirm some of the claims from previous work. Both models supported the claim that states HO-PBDEs have greater AhR binding affinity in comparison to MeO-PBDEs, and both of these BDE derivatives show greater potencies to induce AhR. Moreover, assumptions which suggest that ortho substituted PCBs show lower affinity towards AhR, that substitution of hydrophobic groups like methoxy and chlorine to the 4-position carbon on the benzene ring in PCB would improve the activity, that substitution at 3-position of the benzene ring in PCB would result in increased activity and

that halogenation of PAHs enhanced binding affinity towards the AhR were supported as well.

However, the proposed models were unable to confirm the claim which suggests that substitution of electropositive groups at the 5- position, and substitution of electronegative groups near the 2'-position would increase the activity.

Generated models were superior in terms of compliance with internal and external validation metrics. In addition to that, unlike many works in the literature, this work provided all of the internal and external metrics. Moreover, this work employed many software tools that provided thousands of descriptors in total to make sure built models are covered by the most appropriate descriptors.

Models were externally tested with approximately 900 compounds which are structurally close to the compounds in the data sets. The TCDD- and TCDF-based models had 95.55% and 89.37% structural coverage, respectively. Both models were reliable in terms of predicting brominated biphenyl, halogenated dibenzo-*p*-dioxin and halogenated dibenzofuran groups. TCDF-based model was more reliable in terms of predicting PAHs, whereas TCDD-based model had a better predictive ability for substituted diphenyl ethers and substituted biphenyls, especially when the subsitiuents are bulky groups.

Seven chemicals showed higher binding towards AhR compared to the binding affinity of TCDD and TCDF. These seven chemicals included polybrominated dibenzofurans. They are very persistent in the environment and they can be found in different media. In addition to that, they have adverse effects on human health. These chemicals did not have any experimental or predicted AhR binding affinity data; this work provides reliable predicted pIC_{50} values for these chemicals. These values could turn into mg/L unit and they can be used in REACH database.

REFERENCES

Abalos, M., Prieto, X., Bayona, J.M. 2002. Determination of volatile alkyl sulfides in wastewater by headspace solid-phase micro-extraction followed by gas chromatography-mass spectrometry. Journal of Chromotography A, 963, 249 – 257.

Ashek, A., Cheolju, L., Hyunsung, P., and Seung, J.C. 2006. 3D QSAR Studies of Dioxins and Dioxin-like Compounds Using CoMFA and CoMSIA. Chemosphere, 65(3), 521–529.

Annalisa R. 2013. In Silico Modeling of Aryl Hydrocarbon Receptor Binding Affinities of a Series of Mixed Halogenated Aromatic Compounds, M.S Thesis, University of Pisa.

Balaban, A.T. 1979 Chemical graphs. XXXII. Five new topoligcal indices for the branching of tree-like graphs. Theoretica Chimica Acta, 53, 355-375.

Van den Berg, M., Birnbaum, L.S., Denison, M., de Vito, M., Farland, W., Feeley, M., Fiedler, H., Hakansson, H., Hanberg, A., Haws, L., Rose, M, Safe, S., Schrenk, D., Tohyama, C., Tritscher, A., Tuomisto, J., Tysklind, M., Walker, N., Peterson, R.E. 2006. "The 2005 World Health Organization Reevaluation of Human and Mammalian Toxic Equivalency Factors for Dioxins and Dioxin-like Compounds. Toxicological Sciences, 93(2), 223–241.

Van Den Berg, M., Birnbaum, L., Bosveld, A.T.C., Brunström, B., Cook P., Feeley, M., Giesy, J.P., Hanberg, A., Hasegawa, R., Kennedy, S.W., Kubiak, T., Larsen, J.C., van Leeuwen, F.X.R., Kiem, A.K.D., Nolt, C., Peterson, R. E., Poellinger, L., Safe, S., Schrenk, D., Tillit, D., Tysklind, M., Younes, M., Waern, F., Zacharewski, T.1998. Toxic Equivalency Factors (TEFs) for PCBs, PCDDs, PCDFs for Humans and Wildlife. Environmental Health Perspectives, 106(12), 775–792.

Breivik, K., A. Sweetman, J. M. Pacyna, Jones K.C. 2002. Towards a Global Historical Emission Inventory for Selected PCB Congeners -- a Mass Balance Approach: 1. Global Production and Consumption. Science of the Total Environment, 290, 1–3.

Cai, Z., Giblin, D.E., Sadagopa Ramanujam, V.M., Gross, M.L. 1994. Mass-profile monitoring in trace analysis: Identification of polychlorodinezothiophenes in crab tissues collected from the Newark/Raritan Bay system. Environmental Science and Technology, 28, 1535 - 1538.

Cao, F., Li, X., Xie, L., Wang, Y., Shi, W., Qian, X., Zhu, Y., Yu, H. 2013. Molecular Docking , Molecular Dynamics Simulation , and Structure-Based 3D-QSAR Studies on the Aryl Hydrocarbon Receptor Agonistic Activity of Hydroxylated Polychlorinated Biphenyls. Environmental Toxicology and Pharmacology, 36, 626–635.

Chen, G. Konstantinov, A.D., Chittim, B.G., Joyce, E.M., Bols, N.C., Bunce, N.J. 2001. Synthesis of Polybrominated Diphenyl Ethers and Their Capacity to Induce CYP1A by the Ah Receptor Mediated Pathway.Environmental Science and Technology, 35(18), 3749–3756.

Chen, Y., Riby, J., Srivastava, P., Bartholomew, J., Denison, M., Bjeldanes, L. 1995. Regulation of CYP1A1 by Indolo [3,2-b] Carbazole in Murine Hepatoma Cells. Biochemistry, 270(38), 22548–22555.

Chirico, N. and Gramatica, P. 2011. Real External Predictivity of QSAR Models: How to Evaluate It? Comparison of Different Validation Criteria and Proposal of Using the Concordance Correlation Coefficient. Journal of Chemical Information and Modeling, 51(9), 2320–2335.

Consonni, V., Ballabio, D., Todeschini, R. 2010. Evaluation of Model Predictive Ability by External Validation Techniques. Journal of Chemometrics, 24, 194–201.

Costa, L.G. and Giordano, G. 2007. Developmental Neurotoxicity of Polybrominated Diphenyl Ether (PBDE) Flame Retardants. NeuroToxicology, 28(6), 1047–1067.

Cruciani, G., Baroni, M., Clementi, S., Costantino, G., Riganelli, D., Skagerberg, B. 1992. Predictive Ability of Regression Models. Part I: Standard Deviation of Prediction Errors (SDEP). Journal of Chemometrics, 6, 335–346.

Darnerud, P.O, Eriksen, G.S., Johannesson, T. Larsen P.B., Viluksela, M. et al. 2001. Polybrominated Diphenyl Ethers : Occurrence , Dietary Exposure , and Toxicology. Environmental Health Perspectives, 109, 49–68.

Dearden, J. C., Cronin, M. T. D, Kaiser, K.L.E. 2009. How Not to Develop a Quantitative Structure-Activity or Structure-Property Relationship (QSAR/QSPR). SAR and QSAR in Environmental Research, 20, 241–266.

Diao, J. Li, Y. Shi, S., Sun, Y., Sun Y. 2010. QSAR Models for Predicting Toxicity of Polychlorinated Dibenzo-P-Dioxins and Dibenzofurans Using Quantum Chemical Descriptors. Bulletin of Environmental Contamination and Toxicology, 85(2), 109–115.

Domingo, J.L., Falco G., Llobet, J.M., Casas, C., Teixido, A., Müller, L. 2003. Polychlorinated Naphthalenes in Foods: Estimated Dietary Intake by the Population of Catalonia, Spain. Environmental Science and Technology, 37(11), 2332–2335.

Domingo, L. 2006. "Polychlorinated Diphenyl Ethers (PCDEs): Environmental Levels , Toxicity and Human Exposure A Review of the Published Literature. Environment International, 32, 121–27.

European Chemical Agency (ECHA). https://echa.europa.eu/view-article/-/journal_content/title/further-advice-on-how-to-use-qsars-for-reach-registration. (accessed July 2017).

EPA. 1996. PCBs : Cancer Dose-Response Assessment and Application to Environmental Mixtures. U.S. Environmental Protection Agency ,September, 74.

EPA. 2013. Public Health Implications of Exposure to Polychlorinated Biphenyls (PCBs). https://www.epa.gov/sites/production/files/2015-01/documents/pcb99.pdf. accessed July 2017.

Eriksson, L., Jaworska, J. Worth, A.P., Cronin M.T.D., McDowell R.M., Gramatica P. 2003.

Methods for Reliability and Uncertainty Assessment and for Applicability Evaluations of Classification- and Regression-Based QSARs. Environmental Health Perspectives, 111(10), 1361–1375.

Eriksson, P., Jakobsson, E., Fredriksson, A. 2001. Brominated Flame Retardants: A Novel Class of Developmental Toxicants in Our Environment? Environmental Health Perspectives 10(9), 903–908.

Falandysz, J. 1998. Polychlorinated Naphthalenes: An Environmental Update. Environmental Pollution, 101(1), 77–90.

Falandysz, J., Fernandes, A., Gregoraszczuk, E., Rose, M. 2014. The Toxicological Effects of Halogenated Naphthalenes: A Review of Aryl Hydrocarbon Receptor-Mediated (Dioxin-like) Relative Potency Factors. Journal of environmental science and health, 32(3), 239–272.

Fenner, K., Scheringer, M., MacLeod, M., Matthies, M., McKone, T., Strobe, M., Beyer, A., Bonnell, M., Le Gall, A.C., Klasmeier, J., Mackay, D., Van De Meent D., Pennington, D., Scharenberg, B., Suzuki, N., Wania, F. 2005. Comparing estimates of persistence and long-range transport potential among multimedia models. Environmental Science and Technology, 39, 1932 - 1942.

Fiedler, H. 2003. Dioxins and Furans (PCDD / PCDF). The Handbook of Environmental Chemistry, 3, 123–201.

Golbraikh, A., and Tropsha A. 2002. Beware of q2! Journal of Molecular Graphics and Modelling, 20, 269–276.

Gramatica, P. 2007. Principles of QSAR Models Validation: Internal and External. QSAR and Combinatorial Science, 26(5), 694–701.

Gramatica, P., Chirico, N., Papa, E., Cassani, S., Kovarich, S. 2013. QSARINS: A New Software for the Development, Analysis, and Validation of QSAR MLR Models. Journal of Computational Chemistry, 34(24), 2121–2132.

Grzybek, T., Pietrzak, R., Wachowska, H. 2002. X-ray photoelectron spectroscopy study of oxidized coals with different sulfur content. Fuel Processing Technology, 77, 1-7.

Gu, C., Goodarzi M., Yang, X., Bian, Y., Sun, C., Jiang X. 2012. Predictive Insight into the Relationship between AhR Binding Property and Toxicity of Polybrominated Diphenyl Ethers by PLS-Derived QSAR. Toxicology Letters, 208(3), 269–274.

Guo, W., Holden, A., Smith S.C., Gephart, R., Petras M., Park, J. 2016. PBDE Levels in Breast Milk Are Decreasing in California. Chemosphere, 150, 505–513.

Hardy, M L. 2002. A Comparison of the Properties of the Major Commercial PBDPO / PBDE Product to Those of Major PBB and PCB Products. Chemosphere, 46, 717–728.

Horii, Y., Khim, J.S., Ohura, T. 2009. Relative Potencies of Individual Chlorinated and Brominated Polycyclic Aromatic Hydrocarbons for Induction of Aryl Hydrocarbon Receptor-Mediated Responses. Environmental Science and Technology, 150, 2159–2165.

Janosik, T., Wahlström N., Bergman, J. 2008. Recent Progress in the Chemistry and Applications of Indolocarbazoles. Tetrahedron, 64(39), 9159–9180.

Koistinen, J., Sanderson, J. T., Giesy P., Nevalainen, T., Passivirta J. 1996. ETHOXYRESORUFIN- O -DEETHYLASE INDUCTION POTENCY OF POLYCHLORINATED DIPHENYL ETHERS IN H4IIE RAT HEPATOMA CELLS. Environmental Toxicology, 15(11), 2028–2034.

Kuriyama, S.N., Wanner A., Fidalgo-Neto, A.A., Talsness, C.E., Koerner, W., Chahoud, I. 2007. Developmental Exposure to Low-Dose PBDE-99: Tissue Distribution and Thyroid Hormone Levels. Toxicology, 242, 80–90.

Lee, S., Shin, W., Hong, S., Kang, H. Jung, D., Yim, U.H., Shim W.J., Khim, J.S., Seok, C., Giesy, J.P., Choi K. 2015. Measured and Predicted Affinities of Binding and Relative Potencies to Activate the AhR of PAHs and Their Alkylated Analogues. Chemosphere, 139,

23–29.

Li, F., Li, X., Zhang, L., You, L., Zhao, J., Wu, H. 2011. Docking and 3D-QSAR Studies on the Ah Receptor Binding Affinities of Polychlorinated Biphenyls (PCBs), Dibenzo-P-Dioxins (PCDDs) and Dibenzofurans (PCDFs). Environmental Toxicology and Pharmacology, 32, 478–485.

Lindén, J., Lensu S., Tuomisto, J., Pohjanvirta, R.. 2010. the Aryl Hydrocarbon Receptor and the Central Regulation of Energy Balance. Frontiers in Neuroendocrinology Dioxins , 31, 452–478.

Lopachin, R.M., Gavin, T., Geohagen, B.C., Das, S. 2007. Neurotoxic Mechanisms of Electrophilic Type-2 Alkenes : Soft – Soft Interactions Described by Quantum Mechanical Parameters. Toxicological Sciences, 98, 561–570.

Mantyla, E., Ahotupa, M., Nieminen, L., Paasivirta, J., Sinkkonen, S. 1992. Polychlorinated dibenzothiophenes: toxicological evaluation in mice. Organohalogen Compounds, 10, 161 - 163.

McDonald, T.A. 2005. Polybrominated Diphenylether Levels among United States Residents : Daily Intake and Risk of Harm to the Developing Brain and Reproductive Organs. Integrated Environmental Assessment and Management, 1(4), 343–54.

Morse, D. C., Klasson-Wehler, E., van de Pas, M., de Bie, A. T. H. J., van Bladeren, P. J., Brouwer, A. 1995. Metabolism and biochemical effects of 3,30,4,40-tetrachlorobiphenyl in pregnant and fetal rats. Chemico-Biological Interactions, 95, 41 - 56.

Mimura, J., and Fujii-Kuriyama, Y.. 2003. Functional Role of AhR in the Expression of Toxic Effects by TCDD. Biochimica et Biophysica Acta, 1619,263–268.

Moran, P. 1949. Notes on Continuous Stochastic Phenomena. Biometrika, 2, 17–23.

Mostrag, A., Puzyn, T., Haranczyk, M. 2010. Modeling the overall persistence and

environmental mobility of sulfur-containing polychlorinated organic compounds. Environmental Science and Pollution Research, 17, 470 - 477.

Netzeva, T.I., Worth, A.P., Aldenberg, T., Bengini, R., Cronin, M.T.D, Gramatica, P., Jaworska, J.S., Kahn, S., Klopman, G., Marchant, C.A., Myatt, G., Nikolova-Jeliazkova, N., Patlwicz, G.Y., Perkins, R., Roberts, D.W., Schultz, T.W., Stanton, D.T., van de Sandt, J.J.M, Tong, W., Veith, G., Yang, C. Current Status of Methods for Defining the Applicability Domain of (Quantitative) Structure-Activity Relationships. ATLA: Alternatives to Laboratory Animals, 33(2), 155–173.

Ojha, P.K., Mitra, I., Das, R.N., Roy, K. 2011. Further Exploring rm2 Metrics for Validation of QSPR Models. Chemometrics and Intelligent Laboratory Systems, 107(1), 194–205.

Papa, E., Kovarich, S., Gramatica, P. 2010. QSAR Modeling and Prediction of the Endocrine-Disrupting Potencies of Brominated Flame Retardants. Chemical Research in Toxicology, 23, 946–954.

Parkinson A, Thomas, P. E., Ryan, D.E., Levin, W, Fujita, T., Safe, S. 1988. Induction of rat liver microsomal cytochrome P-450 isozymes and eposide hydrolase by a series of 4'-substituted-2,3,4,5-tetrachlorobiphenyls. Toxicology, 53, 289 - 300.

Piparo, E., Koehler, K., Chana, A., Benfanti, E. 2006. Virtual Screening for Aryl Hydrocarbon Receptor Binding Prediction Cytoplasm , Associated with Heat Shock Proteins . Ligand Binding AhR Protein , Causing the Translocation of the Whole Complex For This Reason We Created a Homology Model of the Ligand. Journal of Medicinal Chemistry, 49, 5702–5709.

Poland, A., Knutson, J.C. 1982. 2,3,7,8-Tetrachlorodibenzo- P -Dioxin and Related Halogenated Aromatic Hydrocarbons: Examination of the Mechanism of Toxicity. Annual Review of Pharmacology and Toxicology, 22, 517–554.

Roy, K., Roy P., Leonard, J. 2007. On Some Aspects of Validation of Predictive QSAR Models. Chemistry Central Journal, 2, 11–13.

Roy, K., Das, R.D., Ambure, P., Aher. R.B. 2016. Be Aware of Error Measures. Further Studies on Validation of Predictive QSAR Models. Chemometrics and Intelligent Laboratory Systems, 152, 18–33.

Roy, K., Kar, S., Das, R.N.. 2015. A Primer on QSAR/QSPR Modeling. In Statistical Methods in QSAR/QSPR, New York: Springer, 37–59.

Safe, S., Bandiera, S., Sawyer, T., Zmudzka, B., Mason, G., Romkes, M., Denomme, M.A., Sparling, J., Okey, A.B., Fujita, T. 1985. Effects of Structure on Binding to the 2,3,7,8-TCDD Receptor Protein and AHH Induction--Halogenated Biphenyls. Environmental Health Perspectives, 61, 21–33.

Safe, S.H. 1997. Limitations of the Toxic Equivalency Factor Approach for Risk Assessment of TCDD and Related Compounds. Teratogenesis, Carcinogenesis, and Mutagenesis, 17, 285–304.

Sánchez, C., Méndez C., Salas J. 2006. Indolocarbazole Natural Products: Occurrence, Biosynthesis, and Biological Activity. Natural Product Reports, 23(6), 1007–1045.

Schecter, A., Johnson-Welch, S., Tung, K.C., Harris, T.R., Papke, O., Rosen, R. 2004. Polybrominated Diphenyl Ethers Contamination of United States Food. Environmental Science and Technology, 38(20), 5306–5311.

Schecter, A., Harris, T.R., Papke, O., Tung, K.C., Musumba, A. 2006. Polybrominated Diphenyl Ether (PBDE) Levels in the Blood of Pure Vegetarians (Vegans). Toxicological & Environmental Chemistry, 88(1), 107–112.

Schecter, A., Johnson-Welch, S., Tung, K.C., Harris, T.R., Papke, O., Rosen, R.. 2007. Polybrominated Diphenyl Ether (PBDE) Levels in Livers of U . S . Human Fetuses and Newborns Polybrominated Diphenyl Ether (PBDE) Levels in Livers of U . S . Journal of Toxicology and Environmental Health, 70, 1–6.

Shen, H., Main, K.M., Andersson, A., Damgaard, I.N., Virtanen, H.E., Skakkebaek, N.E., Toppari, J., Schramm K. 2008. Concentrations of Persistent Organochlorine Compounds in Human Milk and Placenta Are Higher in Denmark than in Finland. Human Reproduction 23(1), 201–210.

Shi, L.M., Fang, H., Tong W., Wu, J., Perkins R., Blair, R.M., Branham, W.S., Dial, S.L., Moland, C.L., Sheehan, D.M. 2001. QSAR Models Using a Large Diverse Set of Estrogens. Journal of Chemical Information and Computer Sciences, 41(1), 186–195.

Sinkkonen, S., Makela, R., Vesterinen, R., Lahtipera, M. 1995a. Chlorinated dioxins and dibenzothiophenes in fly ash samples from combustion of peat, wood chips, refuse derived fuel and liquid packaging boards. Chemosphere, 31(2), 2629 - 2635.

Sinkkonen S., Rantio, T., Vattulainen, A., Aittola, J.P., Paasivirta J., Lahtipera, M. 1995b. Chlorohydrocarbons, PCBcongeners, polychlorodioxins, furans and dibenzothiophenes in pine needle in the vicinity of a metal reclamation plant. Chemosphere, 30(12), 2227 - 2239.

Sinkkonen, S. 1997. PCDTs in the environment. Chemosphere, 34(12), 2585 - 2594.

Su, G., Xia, J., Liu, H., Lam, M.H.W., Yu, H., Giesy, J.P., Zhang, X. 2012a. Dioxin-like Potency of HO- and MeO- Analogues of PBDEs ' the Potential Risk through Consumption of Fish from Eastern China. Environmental Science and Technology, 46, 10781–10788.

Su, G., Xia, J., Liu, H., Lam M.H.W., Yu, H., Giesy, J.P., Zhang, X. 2012b. Dioxin-like Potency of HO- and MeO- Analogues of PBDEs ' the Potential Risk through Consumption of Fish from Eastern China. Environmental Science and Technology, 46, 10781–10788.

Tugcu, G. Sacan, M.T., Vracko, M., Novic, M., Minovski N. 2012. QSTR Modelling of the Acute Toxicity of Pharmaceuticals to Fish. SAR and QSAR in environmental research, 23(3–4), 297–310.

Tugcu,G., Erturk, M.D., Sacan, M.T. 2017. "On the Aquatic Toxicity of Substituted Phenols to Chlorella Vulgaris : QSTR with an Extended Novel Data Set and Interspecies

Models. Journal of Hazardous Materials, 339, 122–130.

Topliss J.G., Edwards, R. P. 1979. Chance factors in studies of quantitative structure-activity relationships. Journal of Medicinal Chemistry, 22(10), 1238-1244.

UNEP. 1999. Guidelines for the Identification of PCBs and Materials Containing PCBs. United Nations Environmental Programme (August), 40.

Waller, C.L., McKinney, J.D. 1995. Three-Dimensional Quantitative Structure-Activity Relationships of Dioxins and Dioxin-like Compounds: Model Validation and Ah Receptor Characterization. Chemical Research in Toxicology, 8(6), 847–858.

Waller, C.L., McKinney, J.D. 1992. Comparative Molecular Field Analysis of Polyhalogenated Dibenzo-P-Dioxins, Dibenzofurans, and Biphenyls. Journal of Medicinal Chemistry, 35(3), 3660–3666.

Wang, M., Chen, S., Huang, K., Lai Y., Chang-Chien,G., Tsai, J., Lin, W., Chang, K., Lee, J. 2010. Chemosphere Determination of Levels of Persistent Organic Pollutants (PCDD / Fs , PBDD / Fs , PBDEs , PCBs , and PBBs) in Atmosphere near a Municipal Solid Waste Incinerator. Chemosphere, 80(10), 1220–1226.

Ward, J., Mohapatra, S.P., Mitchell, A. 2008. An Overview of Policies for Managing Polybrominated Diphenyl Ethers (PBDEs) in the Great Lakes Basin. Environment International, 34(8), 1148–1156.

Weistrand, C., Norén, K. 1998. Polychlorinated Naphthalenes and Other Organochlorine Contaminants in Human Adipose and Liver Tissue. Journal of Toxicology and Environmental Health, Part A ,53(4), 293–311.

Wilford, B.H., Shoeib, M., Harner, T., Zhu, J., Jones, K.C. 2005. Polybrominated Diphenyl Ethers in Indoor Dust in Ottawa, Canada: Implications for Sources and Exposure. Environmental Science and Technology, 39(18), 7027–7035.

Wyrzykowska, B., Hanari, N., Orlikowska, A., Bochentin, I., Rostkowski, P., Falandysz, J., Taniyasu, S., Horii, Y., Jiang, Q., Yamashita, N. 2007. Polychlorinated Biphenyls and -Naphthalenes in Pine Needles and Soil from Poland - Concentrations and Patterns in View of Long-Term Environmental Monitoring. Chemosphere, 67(9), 1877–1886.

Yuan, J., Pu Y.,Yin, P. 2013. Docking-Based Three-Dimensional Quantitative Structure-Activity Relationship (3D-QSAR) Predicts Binding Affinities to Aryl Hydrocarbon Receptor for Polychlorinated Dibenzodioxins, Dibenzofurans, and Biphenyls. Environmental Toxicology and Chemistry, 32(7), 1453–1458.

Yuan, J., Pu, Y., Yin, L. 2014. Prediction of Binding Affinities of PCDDs, PCDFs and PCBs Using Docking-Based Comparative Molecular Similarity Indices Analysis. Environmental Toxicology and Pharmacology, 38(1), 1–7.

Zhang, J., Zhang, X., Xia, P., Zhang, R., Wu, Y., Xia, J., Su, G., Zhang, J., Giesy, J.P., Wang, Z., Villeneuve, D.L., Yu, H. 2016. Chemosphere Activation of AhR-Mediated Toxicity Pathway by Emerging Pollutants Polychlorinated Diphenyl Sulfides. Chemosphere, 144, 1754–1762.

Zota, A.R., Park, J., Wang, Y., Petreas, M., Zoeller, R.T., Woodruff, T.J. 2011. Polybrominated Diphenyl Ethers (PBDEs), Hydroxylated PBDEs (OH-PBDEs), and Measures of Thyroid Function in Second Trimester Pregnant Women in California. Environmental Science and Technology, 45, 7896–7905.

Zvinavashe, E., Murk, A.J., Rietjens, I.M.C.M. 2008. Promises and Pitfalls of Quantitative Structure-Activity Relationship Approaches for Predicting Metabolism and Toxicity. Chemical Research in Toxicology, 21(12), 2229–2236.

APPENDIX A1

Table A1. Hat, descriptor and predicted pIC_{50} values from the TCDD-based model Equation 4.1 for external set chemicals.

Chemical	Pred. pIC_{50} from Eq.4.1	HAT i/i (*h**=0.241)	MATS5m	MATS5v	F09[C-Br]	M_RNG	RgGrav_3D	Mor03v
1,2,3,4,6,7,8-HpBDF	6.817	0.180	-0.406	-0.356	0	1	4.303	-3.543
1,2,3,4,6,7-hexabromonaphthalene	4.841	0.141	-0.256	-0.027	0	0	3.794	-3.062
1,2,3,4,7,8-HxBDD	8.784	0.084	-0.050	0.127	0	1	4.473	-2.871
1,2,3,4,7,8-HxBDF	8.560	0.148	-0.216	0.111	0	1	4.265	-3.149
1,2,3,5,6,7-hexabromonaphthalene	4.884	0.139	-0.256	-0.027	0	0	3.810	-3.060
1,2,3,6,7,8-HxBDD	7.368	0.176	-0.284	-0.347	0	1	4.510	-3.045
1,2,3,7,8,9-HxBDD	7.251	0.165	-0.284	-0.347	0	1	4.441	-2.964
1,2,3,7,8-PeBDD	8.543	0.099	-0.137	-0.065	0	1	4.531	-2.654
1,2,3,7,8-PeBDF	8.088	0.100	-0.289	-0.062	0	1	4.229	-2.947
1,2,4,6,7-pentabromonaphthalene	5.028	0.109	0.028	0.244	0	0	3.694	-2.754
1,2,6,9-tetramethylphenanthrene	2.942	0.092	-0.039	-0.039	0	0	3.130	-2.136
1,2,6-trimethylphenanthrene	3.037	0.086	0.019	0.019	0	0	3.078	-1.922
1,2,9-trimethylhenanthrene	2.388	0.124	-0.070	-0.070	0	0	2.958	-2.210
1,2-dimethylphenanthrene	2.582	0.108	-0.004	-0.004	0	0	2.890	-1.840
1,3,6-trimethylchrysene	3.161	0.152	-0.251	-0.251	0	0	3.470	-2.725

Table A1. Continued.

Chemical	Pred. pIC_{50} from Eq.4.1	HAT i/i (*h**=0.241)	MATS5m	MATS5v	F09[C-Br]	M_RNG	RgGrav_3D	Mor03v
1-bromo-4-methylnaphthalene	1.109	0.206	-0.059	-0.175	0	0	2.524	-1.702
1-methylbenz[a]anthracene	2.880	0.111	-0.128	-0.128	0	0	3.219	-2.392
1-methylchrysene	3.015	0.106	-0.135	-0.135	0	0	3.269	-2.385
1-methylnaphthalene	0.880	0.259	-0.221	-0.221	0	0	2.165	-1.327
2,2',3,3',4,5,6-heptaCDPS	3.104	0.121	-0.057	-0.147	0	0	3.453	-2.434
2,2,3,3-tetrachloroDPS	3.455	0.228	-0.420	-0.314	0	0	3.047	-1.480
2,2',3,4,5-pentaCDPS	3.649	0.081	0.031	-0.036	0	0	3.427	-1.941
2,2,3-trichloroDPS	3.147	0.141	-0.320	-0.183	0	0	2.999	-1.966
2,2,4,4',5pentaCDPS	3.892	0.125	0.314	0.190	0	0	3.616	-2.222
2,2,4,5-tetrachloroDPS	2.880	0.135	0.173	0.088	0	0	3.271	-2.391
2,3,3',4,4',5,6-heptaCDPS	3.765	0.239	0.288	0.187	0	0	3.842	-3.174
2,3,3',4,5,6-hexaCDPS	2.994	0.236	0.157	0.057	0	0	3.590	-3.153
2,3,3-trichloroDPS	3.640	0.128	-0.315	-0.231	0	0	3.315	-2.088
2,3,4,4',5,6-hexaCDPS	3.681	0.267	0.441	0.333	0	0	3.679	-2.884
2,3,4,5,6-pentaCPDS	2.556	0.268	0.312	0.200	0	0	3.292	-2.824
2,3,4,5-tetrachloroDPS	4.082	0.111	0.191	0.046	0	0	3.759	-2.204
2,3,4,7,8-PeBDF	9.499	0.203	-0.208	0.192	0	1	4.385	-2.849
2,3,6,7-tetrabromonaphthalene	7.648	0.394	-0.360	0.250	0	0	3.885	-2.114
2,3,6,7-tetrachloro-dibenzo-p-dioxin	6.570	0.026	0.005	-0.010	0	1	3.780	-2.258
2,3,6-tetrachloro-dibenzo-p-dioxin	5.866	0.027	-0.016	-0.023	0	1	3.459	-2.096
2,3,7,8-TeBDF	9.807	0.201	-0.072	0.298	0	1	4.396	-2.501

Table A1. Continued.

Chemical	Pred. pIC_{50} from Eq.4.1	HAT i/i (*h**=0.241)	MATS5m	MATS5v	F09[C-Br]	M_RNG	RgGrav_3D	Mor03v
2,3-dichloro-DPS	3.033	0.140	-0.241	-0.139	0	0	2.870	-1.582
2.4.4.5-tetrachloroDPS	3.924	0.219	0.482	0.238	0	0	3.742	-2.081
2.4.5-trichloroDPS	3.616	0.155	0.332	0.131	0	0	3.531	-1.864
2.4.6-trichloroDPS	3.287	0.078	-0.011	-0.015	0	0	3.247	-2.107
2.4.7-trimethyldibenzothiophene	4.072	0.248	-0.645	-0.568	0	1	2.890	-2.221
2.4-dimethyldibenzothiophene	3.727	0.322	-0.747	-0.640	0	1	2.770	-2.310
2346TeCDE	3.904	0.082	-0.186	-0.076	0	0	3.301	-2.018
2-bromo-1-methylnaphthalene	1.842	0.168	0.016	-0.092	0	0	2.671	-1.430
2-hydroxy-BDE007	3.103	0.072	-0.004	0.084	1	0	3.289	-1.541
2-hydroxy-BDE028	3.325	0.095	-0.046	-0.049	2	0	4.232	-2.238
2-hydroxy-BDE-066	3.296	0.109	-0.100	-0.131	2	0	4.348	-2.422
2-hydroxy-BDE068	4.708	0.060	-0.063	0.049	1	0	4.233	-2.699
2-hydroxy-BDE123	3.066	0.152	-0.110	-0.122	2	0	4.375	-2.867
2'-hydroxy-CB005	2.442	0.123	-0.100	-0.112	0	0	2.849	-1.708
2'-hydroxy-CB009	1.925	0.200	0.245	0.029	0	0	2.927	-1.768
2'-hydroxy-CB012	3.333	0.078	0.016	0.010	0	0	3.165	-1.814
2-hydroxy-CB025	3.803	0.077	-0.068	-0.059	0	0	3.336	-1.800
2-hydroxy-CB030	3.005	0.147	-0.123	-0.160	0	0	3.000	-1.364
2-hydroxy-CB035	4.519	0.054	-0.014	0.043	0	0	3.529	-1.898
2-hydroxy-CB036	3.965	0.059	0.021	0.072	0	0	3.426	-2.221
2-hydroxy-CB036	4.074	0.062	-0.101	-0.008	0	0	3.436	-2.234
2-hydroxy-CB039	4.581	0.053	0.108	0.124	0	0	3.590	-1.913

Table A1. Continued.

Chemical	Pred. pIC_{50} from Eq.4.1	HAT i/i (*h**=0.241)	MATS5m	MATS5v	F09[C-Br]	M_RNG	RgGrav_3D	Mor03v
2-hydroxy-CB056	4.084	0.110	-0.217	-0.164	0	0	3.410	-1.798
2-hydroxy-CB061	3.081	0.105	0.111	0.099	0	0	3.204	-2.294
2-hydroxy-CB079	4.931	0.048	0.043	0.150	0	0	3.714	-2.335
2-hydroxy-CB080	4.448	0.079	-0.053	0.082	0	0	3.630	-2.691
2-methyl-BDE028	2.010	0.271	0.023	0.046	3	0	4.113	-2.177
2-methyl-BDE068	2.410	0.146	-0.156	-0.257	2	0	4.207	-2.559
2-methyl-BDE123	1.581	0.269	-0.149	-0.246	3	0	4.346	-2.552
2-methyl-dibenzothiophene	3.789	0.096	-0.094	-0.143	0	1	2.701	-1.825
2-methylnaphthalene	1.545	0.218	-0.076	-0.076	0	0	2.288	-1.089
2-methyl-phenanthrene	2.260	0.133	-0.142	-0.142	0	0	2.806	-1.810
3.3-diindoymethane	6.084	0.052	0.062	0.031	0	1	3.446	-1.749
3.4-dichloroDPS	3.014	0.120	0.097	-0.040	0	0	3.375	-2.192
3-hydroxy-BDE007	3.330	0.052	-0.019	0.036	1	0	3.461	-1.612
3-hydroxy-BDE028	3.388	0.093	0.011	0.027	2	0	4.076	-1.844
3-hydroxy-BDE047	3.068	0.111	-0.080	-0.009	2	0	3.803	-1.577
3-hydroxy-BDE100	3.198	0.135	-0.134	0.034	2	0	3.826	-1.984
3-hydroxy-BDE154	3.943	0.153	0.007	0.189	2	0	4.081	-2.139
3-hydroxy-CB009	2.646	0.133	0.164	0.020	0	0	3.084	-1.739
3-hydroxy-CB028	4.412	0.064	-0.056	-0.021	0	0	3.541	-1.891
3-hydroxy-CB030	3.391	0.159	-0.334	-0.267	0	0	3.155	-1.763
3-hydroxy-CB031	2.646	0.133	0.164	0.020	0	0	3.084	-1.739
3-hydroxy-CB061	3.604	0.077	0.050	0.086	0	0	3.342	-2.324
3-hydroxy-CB065	3.005	0.093	0.019	-0.041	0	0	3.184	-1.955
3-hydroxy-CB066	4.696	0.056	-0.103	-0.009	0	0	3.655	-2.200
3-hydroxy-CB068	4.362	0.080	-0.205	-0.081	0	0	3.591	-2.449

Table A1. Continued.

Chemical	Pred. pIC_{50} from Eq.4.1	HAT i/i (*h**=0.241)	MATS5m	MATS5v	F09[C-Br]	M_RNG	RgGrav_3D	Mor03v
3-methyl-BDE100	2.440	0.228	-0.298	-0.398	2	0	3.848	-1.263
3-methylbenz[a]anthracene	3.672	0.068	-0.016	-0.016	0	0	3.422	-2.223
3-methylchrysene	3.275	0.087	-0.075	-0.075	0	0	3.320	-2.306
3-methyldibenzo[ah]anthracene	4.515	0.103	-0.038	-0.038	0	0	3.963	-2.913
3-methylphenanthrene	2.267	0.128	-0.067	-0.067	0	0	2.752	-1.709
4.4-dichloroDPS	3.507	0.271	0.445	0.111	0	0	3.678	-1.824
4-hydroxy-BDE017	3.308	0.056	-0.145	-0.031	1	0	3.515	-2.007
4-hydroxy-BDE042	3.579	0.055	-0.205	-0.132	1	0	3.783	-2.237
4-hydroxy-BDE049	4.360	0.036	0.073	0.128	1	0	3.989	-2.162
4-hydroxy-BDE090	4.085	0.050	-0.027	0.014	1	0	4.078	-2.576
4-hydroxy-CB001	3.002	0.131	-0.139	-0.065	0	0	2.821	-1.411
4-hydroxy-CB002	3.294	0.093	0.097	0.089	0	0	3.042	-1.548
4-hydroxy-CB009	3.124	0.120	0.232	0.117	0	0	3.165	-1.671
4-hydroxy-CB014	3.667	0.074	0.066	0.139	0	0	3.175	-1.951
4-hydroxy-CB020	4.027	0.081	-0.170	-0.014	0	0	3.264	-2.033
4-hydroxy-CB025	4.509	0.057	-0.078	0.040	0	0	3.466	-1.985
4-hydroxy-CB026	3.698	0.064	0.031	0.045	0	0	3.358	-2.121
4-hydroxy-CB031	4.249	0.064	0.122	0.099	0	0	3.511	-1.847
4-hydroxy-CB033	4.231	0.078	-0.171	-0.056	0	0	3.378	-1.920
4-hydroxy-CB035	5.271	0.053	0.107	0.232	0	0	3.679	-1.997
4-hydroxy-CB036	4.931	0.057	0.015	0.178	0	0	3.611	-2.298
4-hydroxy-CB070	4.682	0.046	0.032	0.113	0	0	3.633	-2.213
4-hydroxy-CB079	5.779	0.076	0.052	0.278	0	0	3.841	-2.423

Table A1. Continued.

Chemical	Pred. pIC_{50} from Eq.4.1	HAT i/i (*h**=0.241)	MATS5m	MATS5v	F09[C-Br]	M_RNG	RgGrav_3D	Mor03v
4-hydroxy-CB106	5.061	0.094	0.026	0.242	0	0	3.660	-2.605
4-methoxy-BDE017	2.985	0.163	-0.013	0.152	2	0	3.699	-1.944
4-methoxy-BDE049	2.821	0.119	0.018	0.020	2	0	4.066	-2.400
4-methoxy-BDE090	2.372	0.185	-0.191	-0.338	2	0	3.922	-1.424
4-methylbenz[a]anthracene	3.501	0.072	-0.031	-0.031	0	0	3.346	-2.169
4-methylindolo[3.2-b]carbazole	6.513	0.031	0.013	-0.026	0	1	3.894	-2.555
4-N-acetylamino	4.807	0.107	0.068	-0.028	0	0	4.081	-2.516
5.11-diacetylindolo[3.2-b]carbazole	5.398	0.132	-0.227	-0.278	0	1	3.843	-3.373
5.11-diethylindolo[3.2-b]carbazole	5.376	0.108	-0.228	-0.261	0	1	3.758	-3.232
5.11-dimethylindolo[3.2-b]carbazole	5.345	0.075	-0.104	-0.145	0	1	3.657	-3.038
5-chloro-6-hydroxy-BDE047	2.619	0.124	-0.125	-0.136	2	0	4.113	-2.559
5-chloro-6-methoxy-BDE047	1.658	0.236	-0.046	-0.060	3	0	3.890	-1.609
5-hydroxy-BDE047	3.576	0.135	-0.027	0.105	2	0	3.847	-1.522
5-hydroxy-CB002	3.101	0.099	-0.090	-0.070	0	0	3.019	-1.664
5-hydroxy-CB025	3.882	0.059	-0.020	0.010	0	0	3.403	-2.086
5-hydroxy-CB033	3.366	0.132	-0.278	-0.230	0	0	3.194	-1.866
5-hydroxy-CB034	3.544	0.075	-0.112	-0.044	0	0	3.266	-2.166
5-hydroxy-CB066	4.515	0.081	-0.195	-0.119	0	0	3.654	-2.211
5-hydroxy-CB066	4.785	0.045	0.008	0.083	0	0	3.690	-2.214
5-hydroxy-CB068	4.443	0.059	-0.094	0.010	0	0	3.610	-2.419
5-methoxy-BDE047	4.613	0.098	-0.024	0.193	1	0	3.965	-2.600

Table A1. Continued.

Chemical	Pred. pIC_{50} from Eq.4.1	HAT i/i (*h**=0.241)	MATS5m	MATS5v	F09[C-Br]	M_RNG	RgGrav_3D	Mor03v
6-chloro-2-hydroxy-BDE068	4.514	0.066	-0.025	0.089	1	0	4.154	-2.725
6-chloro-2-hydroxy-BDE7	2.817	0.056	-0.082	-0.030	1	0	3.445	-2.075
6-chloro-2-methoxy-BDE068	2.521	0.141	-0.109	-0.130	2	0	4.131	-2.690
6-ethylchrysene	2.714	0.150	-0.157	-0.157	0	0	3.275	-2.729
6-formylindolo[3,2-b]carbazole	5.556	0.038	-0.013	-0.091	0	1	3.592	-2.477
6-hydroxy-BDE017	2.930	0.063	-0.157	-0.060	1	0	3.409	-2.006
6-hydroxy-BDE047	2.652	0.102	-0.116	-0.177	2	0	4.050	-2.055
6-hydroxy-BDE082	3.404	0.123	-0.294	-0.340	1	0	3.988	-2.324
6-hydroxy-BDE085	2.326	0.134	-0.214	-0.253	2	0	4.060	-2.487
6-hydroxy-BDE087	4.153	0.035	-0.010	0.022	1	0	4.030	-2.316
6-hydroxy-BDE090	4.210	0.056	0.007	0.087	1	0	4.050	-2.596
6-hydroxy-BDE099	4.179	0.160	0.143	0.278	2	0	4.075	-1.727
6-hydroxy-BDE137	2.494	0.142	-0.073	-0.136	2	0	4.177	-2.668
6-hydroxy-BDE140	1.881	0.165	-0.322	-0.407	2	0	3.872	-2.048
6-hydroxy-BDE157	1.940	0.235	-0.240	-0.449	2	0	4.261	-2.582
6-hydroxy-CB106	4.025	0.107	0.108	0.190	0	0	3.508	-2.670
6-hydroxy-CB26	3.208	0.104	-0.086	-0.134	0	0	3.217	-1.781
6-hydroxy-CB31	3.237	0.105	0.122	-0.002	0	0	3.360	-1.984
6-hydroxy-CB35	4.287	0.054	0.108	0.124	0	0	3.543	-2.106
6-hydroxy-CB36	3.948	0.063	0.021	0.072	0	0	3.454	-2.326
6-hydroxy-CB58	3.578	0.126	-0.314	-0.233	0	0	3.366	-2.304

Table A1. Continued.

Chemical	Pred. pIC_{50} from Eq.4.1	HAT i/i (*h**=0.241)	MATS5m	MATS5v	F09[C-Br]	M_RNG	RgGrav_3D	Mor03v
6-hydroxy-CB70	3.857	0.076	-0.081	-0.092	0	0	3.477	-2.052
6-methoxy-BDE017	1.948	0.169	-0.055	0.049	2	0	3.481	-2.081
6-methoxy-BDE085	2.403	0.310	-0.051	0.103	3	0	3.839	-1.520
6-methoxy-BDE090	2.505	0.097	-0.049	-0.095	2	0	3.839	-1.721
6-methoxy-BDE137	1.528	0.233	-0.073	-0.100	3	0	3.969	-1.906
7,12-dimethylbenz[a]anthracene	2.269	0.211	-0.244	-0.244	0	0	3.196	-2.910
7-methylbenz[a]anthracene	2.836	0.116	-0.134	-0.134	0	0	3.216	-2.427
8,9,11-trimethylbenz[a]anthracene	3.480	0.108	-0.047	-0.047	0	0	3.522	-2.726
9-methylbenz[a]anthracene	3.614	0.073	-0.016	-0.016	0	0	3.431	-2.320
Acridine	4.675	0.128	-0.087	-0.023	0	1	2.671	-1.341
Anthracene	2.491	0.134	0.062	0.062	0	0	2.699	-1.418
Benz[a]acridine	5.139	0.040	-0.050	-0.039	0	1	3.174	-2.116
Benz[a]anthracene	3.116	0.085	0.003	0.003	0	0	3.191	-2.166
Benz[b]anthracene	3.858	0.062	0.083	0.083	0	0	3.378	-1.974
Benz[c]acridine	5.445	0.048	-0.087	-0.010	0	1	3.151	-2.020
Benzo[1,2-b:4,5-b']bis[1]benzothiophene	4.761	0.155	0.093	-0.070	0	1	3.634	-3.165
Benzo[1,2-b:4,5-b']bisbenzofuran	5.869	0.034	-0.058	-0.113	0	1	3.593	-2.199
Carbazole	3.915	0.222	-0.334	-0.268	0	1	2.401	-1.136
Chrysene	2.683	0.118	-0.130	-0.130	0	0	3.131	-2.347
Decamethylanthracene	2.351	0.426	-0.187	-0.187	0	0	3.534	-3.927
Dibenz[a,h]anhracene	3.901	0.115	-0.030	-0.030	0	0	3.733	-2.915

Table A1. Continued.

Chemical	Pred. pIC_{50} from Eq.4.1	HAT i/i (*h**=0.241)	MATS5m	MATS5v	F09[C-Br]	M_RNG	RgGrav_3D	Mor03v
dibenz[ac]acridine	5.188	0.065	-0.123	-0.086	0	1	3.404	-2.845
Dibenz[ac]anthracene	3.001	0.160	-0.078	-0.078	0	0	3.429	-2.959
Dibenz[ah]acridine	6.082	0.036	-0.061	-0.027	0	1	3.702	-2.788
Dibenz[aj]acridine	5.644	0.042	-0.035	-0.047	0	1	3.605	-2.760
Dibenz[aj]anthracene	3.683	0.106	-0.030	-0.030	0	0	3.609	-2.781
Dibenz[al]acridine	6.430	0.032	0.029	0.018	0	1	3.849	-2.683
Dibenz[ch]acridine	6.088	0.031	-0.089	-0.005	0	1	3.557	-2.577
Dibenzothiphene	3.307	0.162	-0.332	-0.340	0	1	2.492	-1.713
Indole	2.972	0.367	0.033	0.056	0	1	1.806	-0.511
Indole3acetonitrile	4.117	0.271	-0.025	0.008	0	1	2.267	-0.621
Indole3carbinol	3.676	0.220	-0.098	-0.112	0	1	2.294	-0.878
Indolo[3,2-b]carbazole	5.911	0.027	0.010	-0.032	0	1	3.621	-2.379
Indolo23ccarbazole	4.795	0.047	-0.051	-0.081	0	1	3.169	-2.269
Indolo32bcarbazole	6.226	0.035	-0.069	-0.056	0	1	3.539	-1.986
Napthalene	0.921	0.268	-0.018	-0.018	0	0	2.036	-1.088
OBDD	6.235	0.245	-0.222	-0.402	0	1	4.429	-3.516
OBDF	6.104	0.208	-0.308	-0.363	0	1	4.250	-3.719
PBB040	3.298	0.146	-0.191	-0.247	0	0	3.286	-1.741
PBB041	3.410	0.133	-0.103	-0.153	0	0	3.202	-1.474
PBB042	4.313	0.147	-0.103	-0.153	0	0	3.551	-1.515
PBB043	3.345	0.092	0.043	-0.028	0	0	3.248	-1.725
PBB044	3.858	0.091	0.043	-0.028	0	0	3.432	-1.704
PBB045	2.682	0.204	0.156	-0.060	0	0	3.050	-1.182
PBB046	2.286	0.232	-0.078	-0.279	0	0	3.010	-1.358

Table A1. Continued.

Chemical	Pred. pIC_{50} from Eq.4.1	HAT i/i (*h**=0.241)	MATS5m	MATS5v	F09[C-Br]	M_RNG	RgGrav_3D	Mor03v
PBB048	3.972	0.102	0.131	0.066	0	0	3.349	-1.436
PBB049	4.825	0.111	0.131	0.066	0	0	3.681	-1.483
PBB050	2.913	0.201	0.010	-0.185	0	0	3.198	-1.343
PBB051	3.185	0.227	0.010	-0.185	0	0	3.247	-1.180
PBB052	4.369	0.099	0.276	0.192	0	0	3.557	-1.670
PBB053	2.812	0.185	0.156	-0.060	0	0	3.157	-1.366
PBB054	1.677	0.356	0.269	-0.092	0	0	2.809	-0.902
PBB055	5.105	0.098	-0.216	-0.120	0	0	3.804	-2.084
PBB056	5.390	0.121	-0.216	-0.120	0	0	3.851	-1.900
PBB057	4.766	0.053	-0.070	0.005	0	0	3.751	-2.345
PBB058	4.411	0.120	-0.304	-0.214	0	0	3.724	-2.515
PBB059	3.706	0.140	-0.191	-0.247	0	0	3.492	-1.909
PBB060	6.066	0.133	-0.128	-0.026	0	0	4.003	-1.715
PBB062	3.430	0.117	-0.103	-0.153	0	0	3.265	-1.647
PBB063	5.976	0.085	0.018	0.099	0	0	3.977	-1.772
PBB064	4.829	0.192	-0.103	-0.153	0	0	3.692	-1.355
PBB065	3.485	0.100	0.043	-0.028	0	0	3.252	-1.576
PBB066	5.389	0.121	-0.216	-0.120	0	0	3.851	-1.901
PBB067	5.401	0.048	0.018	0.099	0	0	3.870	-2.105
PBB068	5.440	0.102	-0.216	-0.120	0	0	3.991	-2.279
PBB069	4.026	0.209	-0.336	-0.372	0	0	3.664	-2.038
PBB070	5.694	0.065	0.018	0.099	0	0	3.910	-1.890
PBB071	3.708	0.231	-0.336	-0.372	0	0	3.422	-1.650
PBB072	4.855	0.058	-0.070	0.005	0	0	3.836	-2.507

Table A1. Continued.

Chemical	Pred. pIC_{50} from Eq.4.1	HAT i/i (*h**=0.241)	MATS5m	MATS5v	F09[C-Br]	M_RNG	RgGrav_3D	Mor03v
PBB073	2.963	0.244	-0.424	-0.466	0	0	3.382	-2.264
PBB074	6.485	0.099	0.106	0.193	0	0	4.108	-1.715
PBB075	4.971	0.219	-0.248	-0.278	0	0	3.875	-1.725
PBB077	7.382	0.109	-0.008	0.225	0	0	4.365	-2.176
PBB079	6.707	0.086	-0.095	0.131	0	0	4.274	-2.545
PBB080	5.818	0.122	-0.183	0.037	0	0	4.170	-3.126
PBB082	4.275	0.151	-0.213	-0.239	0	0	3.627	-1.817
PBB083	3.969	0.085	-0.105	-0.127	0	0	3.601	-2.223
PBB084	3.235	0.159	-0.026	-0.183	0	0	3.354	-1.630
PBB085	5.266	0.142	-0.119	-0.117	0	0	3.871	-1.684
PBB086	3.914	0.070	-0.011	-0.005	0	0	3.354	-1.769
PBB087	4.889	0.078	-0.011	-0.005	0	0	3.722	-1.788
PBB088	3.439	0.136	0.068	-0.061	0	0	3.283	-1.426
PBB089	3.119	0.224	-0.134	-0.295	0	0	3.291	-1.434
PBB090	5.133	0.081	-0.011	-0.005	0	0	3.832	-1.847
PBB091	4.316	0.171	0.068	-0.061	0	0	3.568	-1.298
PBB092	4.596	0.051	0.096	0.108	0	0	3.693	-2.183
PBB093	3.419	0.120	0.175	0.052	0	0	3.245	-1.476
PBB094	3.419	0.120	0.175	0.052	0	0	3.245	-1.476
PBB095	3.829	0.108	0.175	0.052	0	0	3.449	-1.637
PBB096	2.717	0.234	0.255	-0.004	0	0	3.092	-1.134
PBB097	4.901	0.079	-0.011	-0.005	0	0	3.724	-1.781
PBB098	3.536	0.214	-0.134	-0.295	0	0	3.511	-1.639
PBB099	5.866	0.096	0.083	0.117	0	0	3.959	-1.648

Table A1. Continued.

Chemical	Pred. pIC_{50} from Eq.4.1	HAT i/i (*h**=0.241)	MATS5m	MATS5v	F09[C-Br]	M_RNG	RgGrav_3D	Mor03v
PBB102	3.715	0.138	0.068	-0.061	0	0	3.390	-1.441
PBB103	4.110	0.127	0.068	-0.061	0	0	3.603	-1.647
PBB104	2.926	0.246	0.147	-0.117	0	0	3.228	-1.177
PBB105	6.378	0.133	-0.198	-0.061	0	0	4.198	-2.086
PBB107	6.079	0.071	-0.091	0.051	0	0	4.098	-2.260
PBB108	5.314	0.126	-0.292	-0.183	0	0	4.052	-2.609
PBB109	4.254	0.131	-0.213	-0.239	0	0	3.732	-2.169
PBB110	4.717	0.169	-0.213	-0.239	0	0	3.793	-1.823
PBB111	4.911	0.113	-0.185	-0.071	0	0	3.996	-3.041
PBB112	4.131	0.086	-0.105	-0.127	0	0	3.669	-2.249
PBB113	3.622	0.191	-0.307	-0.361	0	0	3.713	-2.587
PBB114	6.539	0.084	0.004	0.173	0	0	4.079	-1.917
PBB115	5.427	0.149	-0.119	-0.117	0	0	3.937	-1.701
PBB116	3.794	0.069	-0.011	-0.005	0	0	3.322	-1.809
PBB117	5.480	0.125	-0.011	-0.005	0	0	3.863	-1.542
PBB118	6.833	0.087	0.004	0.173	0	0	4.239	-2.073
PBB119	4.898	0.237	-0.320	-0.352	0	0	3.976	-2.038
PBB120	5.863	0.067	-0.091	0.051	0	0	4.127	-2.600
PBB121	3.987	0.275	-0.415	-0.474	0	0	3.885	-2.562
PBB122	5.013	0.117	-0.292	-0.183	0	0	3.867	-2.382
PBB123	6.027	0.105	-0.198	-0.061	0	0	4.118	-2.244
PBB124	5.553	0.057	-0.091	0.051	0	0	3.943	-2.384
PBB125	3.261	0.258	-0.415	-0.474	0	0	3.481	-2.143
PBB126	7.572	0.127	-0.076	0.229	0	0	4.437	-2.513

Table A1. Continued.

Chemical	Pred. pIC_{50} from Eq.4.1	HAT i/i (*h**=0.241)	MATS5m	MATS5v	F09[C-Br]	M_RNG	RgGrav_3D	Mor03v
PBB127	6.579	0.124	-0.170	0.107	0	0	4.305	-2.998
PBB128	5.282	0.178	-0.227	-0.226	0	0	3.991	-1.921
PBB129	4.569	0.084	-0.146	-0.124	0	0	3.752	-2.204
PBB130	4.996	0.094	-0.146	-0.124	0	0	3.923	-2.240
PBB131	3.918	0.154	-0.091	-0.211	0	0	3.621	-1.811
PBB132	4.126	0.175	-0.091	-0.211	0	0	3.644	-1.644
PBB133	4.542	0.082	-0.066	-0.023	0	0	3.863	-2.782
PBB134	3.733	0.101	-0.011	-0.109	0	0	3.551	-2.014
PBB135	3.727	0.101	-0.011	-0.109	0	0	3.603	-2.184
PBB136	3.645	0.139	0.226	0.066	0	0	3.373	-1.458
PBB137	5.773	0.079	-0.045	0.035	0	0	3.972	-1.921
PBB138	6.029	0.093	-0.045	0.035	0	0	4.057	-1.891
PBB139	5.053	0.136	0.010	-0.051	0	0	3.825	-1.567
PBB140	4.319	0.240	-0.172	-0.312	0	0	3.808	-1.735
PBB141	5.323	0.044	0.036	0.137	0	0	3.819	-2.166
PBB142	3.895	0.080	0.091	0.051	0	0	3.353	-1.638
PBB143	3.371	0.144	-0.091	-0.211	0	0	3.415	-1.804
PBB144	4.640	0.073	0.091	0.051	0	0	3.688	-1.818
PBB145	3.457	0.169	0.146	-0.035	0	0	3.320	-1.310
PBB146	5.752	0.049	0.036	0.137	0	0	3.988	-2.196
PBB147	5.053	0.107	0.091	0.051	0	0	3.748	-1.529
PBB148	4.088	0.147	-0.091	-0.211	0	0	3.758	-2.042
PBB149	4.853	0.089	0.091	0.051	0	0	3.712	-1.647
PBB150	3.951	0.176	0.146	-0.035	0	0	3.513	-1.341
PBB151	4,460	0,061	0,171	0,152	0	0	3,620	-2,016

Table A1. Continued.

Chemical	Pred. pIC_{50} from Eq.4.1	HAT i/i (*h**=0.241)	MATS5m	MATS5v	F09[C-Br]	M_RNG	RgGrav_3D	Mor03v
PBB152	3.358	0.139	0.226	0.066	0	0	3.271	-1.473
PBB153	6.762	0.091	0.137	0.296	0	0	4.116	-1.853
PBB154	5.029	0.118	0.010	-0.051	0	0	3.873	-1.744
PBB155	4.062	0.206	0.065	-0.137	0	0	3.660	-1.468
PBB156	6.820	0.092	-0.100	0.121	0	0	4.262	-2.345
PBB157	6.105	0.139	-0.282	-0.140	0	0	4.250	-2.507
PBB158	5.373	0.169	-0.227	-0.226	0	0	4.084	-2.107
PBB159	5.529	0.112	-0.201	-0.038	0	0	4.141	-3.034
PBB160	4.455	0.083	-0.146	-0.124	0	0	3.778	-2.416
PBB161	4.131	0.226	-0.328	-0.386	0	0	3.979	-2.783
PBB162	5.744	0.099	-0.201	-0.038	0	0	4.162	-2.852
PBB163	5.304	0.110	-0.146	-0.124	0	0	3.992	-2.098
PBB164	4.135	0.215	-0.328	-0.386	0	0	3.856	-2.396
PBB165	3.885	0.185	-0.248	-0.284	0	0	3.894	-3.010
PBB166	5.879	0.094	-0.045	0.035	0	0	3.972	-1.799
PBB167	6.795	0.090	-0.100	0.121	0	0	4.303	-2.502
PBB168	4.383	0.308	-0.409	-0.487	0	0	4.032	-2.463
PBB169	7.666	0.167	-0.155	0.207	0	0	4.522	-2.904
PBB170	5.670	0.110	-0.176	-0.112	0	0	4.111	-2.250
PBB171	4.790	0.196	-0.140	-0.241	0	0	3.939	-1.850
PBB172	5.263	0.074	-0.116	-0.020	0	0	4.028	-2.703
PBB173	4.108	0.104	-0.080	-0.149	0	0	3.693	-2.114
PBB174	4.271	0.108	-0.080	-0.149	0	0	3.751	-2.105
PBB175	4.417	0.109	-0.080	-0.149	0	0	3.869	-2.305
PBB176	4.421	0.109	0.125	0.029	0	0	3.616	-1.571

Table A1. Continued.

Chemical	Pred. pIC_{50} from Eq.4.1	HAT i/i (*h**=0.241)	MATS5m	MATS5v	F09[C-Br]	M_RNG	RgGrav_3D	Mor03v
PBB177	4.652	0.124	-0.080	-0.149	0	0	3.849	-1.968
PBB178	4.100	0.093	-0.021	-0.057	0	0	3.787	-2.659
PBB179	4.169	0.076	0.185	0.121	0	0	3.547	-1.888
PBB180	6.650	0.075	-0.006	0.196	0	0	4.159	-2.210
PBB181	5.570	0.085	0.030	0.067	0	0	3.883	-1.718
PBB182	4.789	0.196	-0.140	-0.241	0	0	3.939	-1.851
PBB183	5.732	0.080	0.030	0.067	0	0	3.988	-1.856
PBB184	4.625	0.154	0.066	-0.063	0	0	3.762	-1.543
PBB185	5.060	0.045	0.090	0.159	0	0	3.745	-2.118
PBB186	3.820	0.102	0.125	0.029	0	0	3.392	-1.568
PBB187	5.605	0.054	0.090	0.159	0	0	3.899	-1.968
PBB189	6.571	0.119	-0.211	0.018	0	0	4.362	-2.876
PBB190	5.719	0.111	-0.176	-0.112	0	0	4.145	-2.300
PBB191	4.688	0.271	-0.345	-0.420	0	0	4.169	-2.618
PBB192	4.105	0.220	-0.286	-0.328	0	0	4.034	-3.117
PBB193	4.506	0.197	-0.286	-0.328	0	0	4.068	-2.759
PBB194	6.119	0.087	-0.155	0.003	0	0	4.235	-2.664
PBB195	4.980	0.152	-0.136	-0.199	0	0	4.020	-2.097
PBB196	4.924	0.144	-0.136	-0.199	0	0	4.050	-2.255
PBB197	5.170	0.111	0.046	-0.006	0	0	3.885	-1.707
PBB199	4.673	0.095	-0.094	-0.117	0	0	3.954	-2.516
PBB200	4.679	0.060	0.089	0.076	0	0	3.688	-1.922
PBB201	4.966	0.063	0.089	0.076	0	0	3.801	-1.942
PBB202	4.576	0.061	0.131	0.158	0	0	3.721	-2.412

Table A1. Continued.

Chemical	Pred. pIC_{50} from Eq.4.1	HAT i/i (*h**=0.241)	MATS5m	MATS5v	F09[C-Br]	M_RNG	RgGrav_3D	Mor03v
PBB203	6.348	0.065	0.027	0.196	0	0	4.058	-2.096
PBB204	4.989	0.100	0.046	-0.006	0	0	3.829	-1.742
PBB205	4.687	0.262	-0.318	-0.392	0	0	4.249	-2.902
PBB206	4.882	0.150	-0.156	-0.214	0	0	4.138	-2.588
PBB207	5.430	0.073	0.008	0.025	0	0	3.966	-2.004
PBB208	5.091	0.048	0.035	0.095	0	0	3.877	-2.387
PBB209	5.204	0.080	-0.051	-0.051	0	0	4.058	-2.395
PBDD028	7.165	0.056	-0.196	-0.147	0	1	4.131	-2.814
PBDE000	3.213	0.147	-0.098	0.030	0	0	2.739	-1.256
PBDE001	3.298	0.113	-0.062	0.056	0	0	2.860	-1.515
PBDE002	4.350	0.088	-0.076	-0.022	0	0	3.387	-1.568
PBDE004	3.449	0.148	-0.141	0.039	0	0	2.784	-1.370
PBDE005	3.813	0.086	-0.127	-0.025	0	0	3.165	-1.714
PBDE006	3.344	0.103	-0.127	-0.025	0	0	2.964	-1.632
PBDE007	2.899	0.054	-0.016	0.030	1	0	3.398	-1.869
PBDE008	3.952	0.033	-0.016	0.030	1	0	3.800	-1.903
PBDE009	3.338	0.155	0.380	0.243	0	0	3.272	-1.785
PBDE010	3.335	0.126	-0.141	0.039	0	0	2.849	-1.702
PBDE011	3.959	0.077	-0.112	-0.088	0	0	3.422	-1.928
PBDE012	2.955	0.051	-0.002	-0.034	1	0	3.497	-1.692
PBDE013	4.968	0.078	-0.002	-0.034	1	0	4.301	-1.866
PBDE014	3.974	0.074	-0.112	-0.088	0	0	3.458	-2.021
PBDE016	3.947	0.114	-0.210	-0.070	0	0	3.140	-1.609
PBDE018	4.405	0.149	0.157	0.199	0	0	3.157	-0.961
PBDE019	3.852	0.140	-0.236	-0.018	0	0	2.976	-1.615

Table A1. Continued.

Chemical	Pred. pIC_{50} from Eq.4.1	HAT i/i (*h**=0.241)	MATS5m	MATS5v	F09[C-Br]	M_RNG	RgGrav_ _3D	Mor03v
PBDE020	5.408	0.119	-0.183	-0.122	0	0	3.925	-1.950
PBDE021	2.632	0.058	-0.087	-0.056	1	0	3.375	-1.950
PBDE022	4.325	0.042	-0.087	-0.056	1	0	4.041	-2.064
PBDE024	3.942	0.081	0.157	0.199	0	0	3.198	-1.626
PBDE025	3.442	0.040	-0.087	-0.056	1	0	3.757	-2.204
PBDE026	4.327	0.067	0.183	0.147	0	0	3.604	-2.036
PBDE027	4.592	0.098	-0.210	-0.070	0	0	3.437	-1.786
PBDE029	2.990	0.100	0.279	0.213	1	0	3.444	-1.585
PBDE030	2.631	0.085	-0.114	-0.004	1	0	3.308	-2.154
PBDE031	4.437	0.083	0.279	0.213	1	0	3.991	-1.616
PBDE032	3.526	0.053	-0.114	-0.004	1	0	3.527	-1.801
PBDE033	3.221	0.057	-0.087	-0.056	1	0	3.477	-1.585
PBDE034	3.870	0.085	-0.183	-0.122	0	0	3.379	-2.028
PBDE035	4.995	0.097	-0.061	-0.108	1	0	4.403	-2.010
PBDE036	4.610	0.108	-0.157	-0.174	0	0	3.883	-2.334
PBDE037	3.943	0.114	0.035	-0.042	2	0	4.498	-2.021
PBDE038	2.386	0.064	-0.061	-0.108	1	0	3.440	-2.029
PBDE039	4.362	0.067	-0.061	-0.108	1	0	4.291	-2.393
PBDE040	3.872	0.190	-0.276	-0.203	0	0	3.148	-1.278
PBDE041	2.790	0.066	-0.186	-0.122	1	0	3.383	-1.875
PBDE042	3.470	0.101	-0.186	-0.122	1	0	3.463	-1.335
PBDE043	3.902	0.059	0.020	0.074	0	0	3.356	-2.089
PBDE044	5.363	0.068	0.020	0.074	0	0	3.803	-1.790
PBDE045	9.288	0.132	-0.002	0.083	0	1	4.634	-2.325
PBDE046	4.117	0.122	-0.310	-0.161	0	0	3.325	-1.936

Table A1. Continued.

Chemical	Pred. pIC_{50} from Eq.4.1	HAT i/i (*h**=0.241)	MATS5m	MATS5v	F09[C-Br]	M_RNG	RgGrav_3D	Mor03v
PBDE048	2.813	0.086	0.110	0.154	1	0	3.386	-2.050
PBDE050	3.139	0.115	-0.220	-0.081	1	0	3.224	-1.356
PBDE051	3.261	0.071	-0.220	-0.081	1	0	3.482	-2.022
PBDE052	5.804	0.143	0.316	0.350	0	0	3.729	-1.243
PBDE053	4.944	0.128	-0.014	0.116	0	0	3.330	-1.192
PBDE054	4.109	0.233	-0.345	-0.119	0	0	2.954	-1.181
PBDE055	3.945	0.066	-0.152	-0.164	1	0	3.972	-1.981
PBDE056	3.824	0.070	-0.152	-0.164	1	0	3.877	-1.826
PBDE057	5.362	0.080	0.054	0.032	0	0	3.993	-1.994
PBDE058	4.255	0.130	-0.242	-0.244	0	0	3.705	-2.186
PBDE059	5.331	0.062	0.020	0.074	0	0	3.814	-1.863
PBDE060	2.819	0.091	-0.061	-0.084	2	0	3.981	-1.917
PBDE061	2.661	0.074	0.144	0.112	1	0	3.384	-1.831
PBDE062	2.895	0.079	0.110	0.154	1	0	3.386	-1.955
PBDE063	4.283	0.045	0.144	0.112	1	0	4.031	-1.970
PBDE064	4.314	0.040	0.110	0.154	1	0	3.894	-1.894
PBDE065	3.951	0.132	0.316	0.350	0	0	3.334	-2.159
PBDE067	4.701	0.052	0.144	0.112	1	0	4.173	-1.927
PBDE068	3.694	0.066	-0.152	-0.164	1	0	4.038	-2.478
PBDE069	3.939	0.052	-0.186	-0.122	1	0	3.940	-2.281
PBDE070	4.825	0.062	0.144	0.112	1	0	4.166	-1.763
PBDE071	3.708	0.063	-0.186	-0.122	1	0	3.687	-1.760
PBDE072	5.083	0.067	0.054	0.032	0	0	3.986	-2.297
PBDE073	4.125	0.110	-0.276	-0.203	0	0	3.554	-2.249
PBDE074	3.435	0.094	-0.061	-0.084	2	0	4.237	-2.000

Table A1. Continued.

Chemical	Pred. pIC_{50} from Eq.4.1	HAT i/i (*h**=0.241)	MATS5m	MATS5v	F09[C-Br]	M_RNG	RgGrav_3D	Mor03v
PBDE076	2.951	0.058	-0.152	-0.164	1	0	3.623	-2.045
PBDE078	3.182	0.083	-0.117	-0.206	1	0	3.947	-2.394
PBDE079	3.967	0.102	-0.117	-0.206	1	0	4.291	-2.557
PBDE080	5.055	0.208	-0.208	-0.286	0	0	4.306	-2.746
PBDE081	4.153	0.142	-0.027	-0.126	2	0	4.641	-2.038
PBDE082	2.934	0.119	-0.258	-0.277	1	0	3.501	-1.536
PBDE083	4.429	0.092	-0.092	-0.081	0	0	3.548	-1.722
PBDE084	4.963	0.066	-0.133	-0.049	0	0	3.780	-2.193
PBDE086	2.597	0.071	-0.003	0.017	1	0	3.443	-2.228
PBDE087	4.333	0.034	-0.003	0.017	1	0	4.026	-2.033
PBDE088	3.227	0.070	-0.044	0.049	1	0	3.330	-1.511
PBDE089	2.884	0.087	-0.299	-0.245	1	0	3.573	-2.186
PBDE090	4.080	0.039	-0.003	0.017	1	0	3.851	-1.780
PBDE091	3.857	0.038	-0.044	0.049	1	0	3.741	-2.064
PBDE092	5.585	0.072	0.163	0.214	0	0	3.812	-1.694
PBDE093	4.663	0.094	0.122	0.245	0	0	3.273	-1.439
PBDE094	4.135	0.066	-0.133	-0.049	0	0	3.508	-2.303
PBDE095	5.783	0.105	0.122	0.245	0	0	3.675	-1.395
PBDE096	4.593	0.120	-0.173	-0.018	0	0	3.295	-1.464
PBDE097	3.860	0.050	-0.003	0.017	1	0	3.703	-1.576
PBDE098	3.416	0.090	-0.299	-0.245	1	0	3.861	-2.465
PBDE099	3.557	0.109	0.086	0.116	2	0	4.027	-1.651
PBDE101	5.086	0.093	0.252	0.312	1	0	3.983	-1.513

Table A1. Continued.

Chemical	Pred. pIC_{50} from Eq.4.1	HAT i/i (*h**=0.241)	MATS5m	MATS5v	F09[C-Br]	M_RNG	RgGrav_3D	Mor03v
PBDE102	3.931	0.063	-0.044	0.049	1	0	3.586	-1.493
PBDE103	4.241	0.047	-0.044	0.049	1	0	3.769	-1.704
PBDE104	3.116	0.113	-0.339	-0.214	1	0	3.397	-1.722
PBDE105	2.269	0.117	-0.129	-0.210	2	0	4.016	-2.268
PBDE106	4.053	0.098	-0.117	-0.206	1	0	4.273	-2.404
PBDE107	4.650	0.059	0.037	-0.014	1	0	4.280	-2.105
PBDE108	3.280	0.122	-0.218	-0.308	1	0	4.062	-2.531
PBDE109	3.976	0.035	-0.003	0.017	1	0	3.969	-2.271
PBDE110	4.610	0.041	-0.003	0.017	1	0	4.106	-1.962
PBDE111	4.834	0.110	-0.052	-0.112	0	0	4.076	-2.546
PBDE112	5.568	0.054	0.163	0.214	0	0	3.932	-2.087
PBDE113	4.673	0.074	-0.092	-0.081	0	0	3.813	-2.268
PBDE114	2.271	0.117	-0.129	-0.210	2	0	4.016	-2.265
PBDE115	2.934	0.133	0.086	0.116	2	0	4.004	-2.302
PBDE116	2.759	0.195	0.252	0.312	1	0	3.406	-2.411
PBDE117	4.549	0.079	0.252	0.312	1	0	3.970	-2.095
PBDE118	3.197	0.126	0.126	0.084	2	0	4.240	-2.370
PBDE120	4.145	0.068	0.037	-0.014	1	0	4.259	-2.625
PBDE121	3.210	0.122	-0.258	-0.277	1	0	4.012	-2.812
PBDE122	2.972	0.108	-0.218	-0.308	1	0	3.837	-2.185
PBDE123	2.441	0.132	-0.129	-0.210	2	0	4.158	-2.512
PBDE124	4.018	0.047	0.037	-0.014	1	0	4.101	-2.278
PBDE125	2.875	0.089	-0.258	-0.277	1	0	3.673	-2.142
PBDE127	3.675	0.198	-0.177	-0.340	1	0	4.430	-2.854

Table A1. Continued.

Chemical	Pred. pIC_{50} from Eq.4.1	HAT i/i (*h**=0.241)	MATS5m	MATS5v	F09[C-Br]	M_RNG	RgGrav_3D	Mor03v
PBDE128	1.959	0.156	-0.246	-0.368	2	0	3.883	-1.863
PBDE129	3.090	0.063	-0.107	-0.165	1	0	3.693	-1.891
PBDE130	3.553	0.064	-0.107	-0.165	1	0	3.905	-2.018
PBDE131	3.329	0.059	-0.153	-0.144	1	0	3.644	-1.791
PBDE132	3.794	0.069	-0.153	-0.144	1	0	3.792	-1.712
PBDE133	5.152	0.061	0.032	0.038	0	0	3.916	-2.132
PBDE134	4.839	0.067	-0.014	0.059	0	0	3.554	-1.694
PBDE135	5.430	0.075	-0.014	0.059	0	0	3.791	-1.746
PBDE136	5.397	0.079	-0.061	0.079	0	0	3.659	-1.699
PBDE137	2.811	0.092	-0.017	-0.042	2	0	3.993	-1.996
PBDE138	3.133	0.090	-0.017	-0.042	2	0	4.087	-1.918
PBDE139	2.793	0.142	-0.063	-0.022	2	0	4.095	-2.655
PBDE140	2.106	0.143	-0.293	-0.347	2	0	3.863	-1.961
PBDE141	4.477	0.041	0.122	0.161	1	0	3.954	-1.878
PBDE142	3.309	0.078	0.076	0.181	1	0	3.376	-1.748
PBDE143	3.239	0.063	-0.153	-0.144	1	0	3.586	-1.713
PBDE144	4.658	0.049	0.076	0.181	1	0	3.894	-1.803
PBDE145	3.197	0.060	-0.200	-0.123	1	0	3.525	-1.904
PBDE146	4.762	0.042	0.122	0.161	1	0	4.091	-1.975
PBDE147	4.439	0.047	0.076	0.181	1	0	3.833	-1.865
PBDE148	3.732	0.053	-0.153	-0.144	1	0	3.870	-2.028
PBDE149	4.939	0.057	0.076	0.181	1	0	3.964	-1.694
PBDE150	3.523	0.061	-0.200	-0.123	1	0	3.622	-1.827
PBDE151	6.092	0.095	0.215	0.384	0	0	3.786	-1.737

Table A1. Continued.

Chemical	Pred. pIC_{50} from Eq.4.1	HAT i/i (h^*=0.241)	MATS5m	MATS5v	F09[C-Br]	M_RNG	RgGrav_3D	Mor03v
PBDE152	4.835	0.075	-0.061	0.079	0	0	3.437	-1.660
PBDE155	2.014	0.129	-0.223	-0.257	2	0	3.887	-2.326
PBDE156	2.134	0.142	-0.118	-0.273	2	0	4.085	-2.238
PBDE157	3.493	0.115	0.029	-0.063	2	0	4.454	-2.318
PBDE158	3.492	0.103	-0.017	-0.042	2	0	4.366	-2.371
PBDE159	2.048	0.205	-0.200	-0.388	2	0	4.246	-2.571
PBDE160	4.272	0.052	0.122	0.161	1	0	4.039	-2.380
PBDE161	3.506	0.104	-0.107	-0.165	1	0	4.153	-2.846
PBDE162	4.196	0.053	0.122	0.161	1	0	4.012	-2.384
PBDE163	3.724	0.070	-0.107	-0.165	1	0	4.085	-2.381
PBDE164	3.472	0.123	-0.061	-0.186	1	0	4.227	-2.788
PBDE165	5.060	0.068	0.032	0.038	0	0	4.024	-2.576
PBDE166	3.337	0.197	0.212	0.283	2	0	4.059	-2.366
PBDE167	3.158	0.140	0.029	-0.063	2	0	4.434	-2.643
PBDE168	1.984	0.194	-0.246	-0.368	2	0	4.167	-2.721
PBDE169	2.433	0.305	-0.154	-0.409	2	0	4.637	-3.016
PBDE170	2.134	0.142	-0.118	-0.273	2	0	4.086	-2.241
PBDE171	2.288	0.124	-0.170	-0.265	2	0	4.030	-2.168
PBDE172	3.890	0.051	0.002	-0.053	1	0	4.086	-2.321
PBDE173	3.283	0.039	-0.050	-0.046	1	0	3.675	-2.017
PBDE174	4.118	0.041	-0.050	-0.046	1	0	3.970	-1.972
PBDE175	3.910	0.037	-0.050	-0.046	1	0	3.908	-2.019
PBDE176	4.038	0.036	-0.102	-0.038	1	0	3.903	-2.136
PBDE177	4.079	0.038	-0.050	-0.046	1	0	4.021	-2.177

Table A1. Continued.

Chemical	Pred. pIC_{50} from Eq.4.1	HAT i/i (*h**=0.241)	MATS5m	MATS5v	F09[C-Br]	M_RNG	RgGrav_3D	Mor03v
PBDE178	5.664	0.049	0.070	0.174	0	0	3.903	-2.085
PBDE179	5.706	0.059	0.018	0.181	0	0	3.782	-1.935
PBDE180	3.660	0.108	0.095	0.106	2	0	4.274	-2.206
PBDE181	3.015	0.129	0.043	0.113	2	0	3.951	-2.222
PBDE182	2.211	0.129	-0.170	-0.265	2	0	4.040	-2.289
PBDE184	2.114	0.138	-0.223	-0.257	2	0	3.976	-2.488
PBDE185	4.862	0.090	0.163	0.333	1	0	3.909	-2.065
PBDE186	3.231	0.045	-0.102	-0.038	1	0	3.568	-2.026
PBDE187	5.335	0.084	0.163	0.333	1	0	4.075	-2.035
PBDE188	3.840	0.038	-0.102	-0.038	1	0	3.856	-2.219
PBDE189	2.350	0.221	-0.065	-0.280	2	0	4.440	-2.813
PBDE190	2.835	0.182	0.095	0.106	2	0	4.135	-2.729
PBDE191	2.313	0.200	-0.118	-0.273	2	0	4.353	-2.867
PBDE192	3.522	0.111	0.002	-0.053	1	0	4.150	-2.947
PBDE193	4.231	0.055	0.002	-0.053	1	0	4.206	-2.301
PBDE194	2.481	0.156	-0.020	-0.180	2	0	4.294	-2.560
PBDE195	2.155	0.130	-0.079	-0.191	2	0	4.053	-2.394
PBDE196	2.584	0.130	-0.079	-0.191	2	0	4.227	-2.440
PBDE197	2.443	0.137	-0.138	-0.202	2	0	4.160	-2.602
PBDE198	4.148	0.042	0.027	0.064	1	0	4.044	-2.430
PBDE199	4.381	0.035	0.027	0.064	1	0	4.091	-2.308
PBDE200	4.138	0.037	-0.032	0.053	1	0	3.934	-2.307
PBDE201	4.309	0.039	-0.032	0.053	1	0	4.023	-2.385
PBDE202	6.197	0.077	0.074	0.307	0	0	3.898	-2.177

Table A1. Continued.

Chemical	Pred. pIC_{50} from Eq.4.1	HAT i/i (*h**=0.241)	MATS5m	MATS5v	F09[C-Br]	M_RNG	RgGrav_3D	Mor03v
PBDE203	4.104	0.189	0.124	0.284	2	0	4.226	-2.402
PBDE204	1.984	0.151	-0.138	-0.202	2	0	3.999	-2.630
PBDE205	2.568	0.170	-0.020	-0.180	2	0	4.368	-2.691
PBDE206	2.534	0.153	-0.009	-0.118	2	0	4.259	-2.687
PBDE207	2.304	0.166	-0.075	-0.163	2	0	4.192	-2.794
PBDE208	4.620	0.069	0.020	0.173	1	0	4.056	-2.561
PBDE209	2.048	0.216	-0.029	-0.168	2	0	4.233	-2.981
PCB001	2.214	0.179	-0.216	-0.153	0	0	2.563	-1.378
PCB002	2.818	0.108	0.028	0.006	0	0	2.911	-1.541
PCB003	3.169	0.106	0.095	0.032	0	0	3.075	-1.487
PCB004	1.751	0.221	-0.018	-0.156	0	0	2.581	-1.053
PCB005	2.896	0.152	-0.199	-0.133	0	0	2.800	-1.363
PCB006	2.782	0.126	-0.199	-0.133	0	0	2.872	-1.719
PCB007	3.047	0.104	-0.118	-0.092	0	0	3.002	-1.677
PCB008	3.302	0.107	-0.118	-0.092	0	0	3.057	-1.554
PCB009	2.225	0.161	0.186	0.022	0	0	2.918	-1.621
PCB010	2.571	0.275	-0.404	-0.310	0	0	2.625	-1.140
PCB011	3.770	0.065	0.005	0.045	0	0	3.255	-1.834
PCB012	3.548	0.078	0.087	0.086	0	0	3.172	-1.687
PCB013	4.288	0.066	0.087	0.086	0	0	3.453	-1.707
PCB014	3.385	0.075	0.005	0.045	0	0	3.132	-1.898
PCB015	4.718	0.084	0.168	0.126	0	0	3.631	-1.618
PCB016	2.234	0.138	-0.123	-0.174	0	0	2.879	-1.803
PCB017	2.528	0.121	-0.030	-0.117	0	0	3.031	-1.831

Table A1. Continued.

Chemical	Pred. pIC_{50} from Eq.4.1	HAT i/i (*h**=0.241)	MATS5m	MATS5v	F09[C-Br]	M_RNG	RgGrav_3D	Mor03v
PCB018	1.853	0.189	0.185	-0.012	0	0	2.921	-1.877
PCB019	1.611	0.261	-0.006	-0.214	0	0	2.643	-1.032
PCB020	3.373	0.106	-0.239	-0.135	0	0	3.129	-2.006
PCB021	3.449	0.101	-0.146	-0.078	0	0	3.070	-1.630
PCB022	4.013	0.085	-0.146	-0.078	0	0	3.326	-1.776
PCB023	2.818	0.099	0.069	0.027	0	0	3.068	-1.962
PCB024	2.594	0.152	-0.123	-0.174	0	0	2.888	-1.412
PCB025	3.854	0.076	-0.146	-0.078	0	0	3.335	-1.989
PCB026	3.265	0.082	0.069	0.027	0	0	3.249	-2.010
PCB027	3.019	0.224	-0.430	-0.336	0	0	2.942	-1.582
PCB028	4.407	0.069	-0.053	-0.020	0	0	3.514	-1.806
PCB029	3.040	0.103	0.162	0.084	0	0	3.154	-1.869
PCB030	3.127	0.183	-0.337	-0.279	0	0	3.008	-1.558
PCB031	3.788	0.085	0.162	0.084	0	0	3.401	-1.772
PCB032	3.528	0.217	-0.337	-0.279	0	0	3.081	-1.320
PCB033	3.485	0.087	-0.146	-0.078	0	0	3.153	-1.848
PCB034	3.302	0.107	-0.239	-0.135	0	0	3.120	-2.063
PCB035	4.609	0.048	0.045	0.124	0	0	3.532	-1.985
PCB036	4.321	0.056	-0.048	0.066	0	0	3.478	-2.251
PCB037	5.151	0.055	0.138	0.181	0	0	3.726	-1.857
PCB038	3.775	0.069	0.045	0.124	0	0	3.238	-2.035
PCB039	4.837	0.045	0.045	0.124	0	0	3.644	-2.071
PCB040	2.967	0.123	-0.234	-0.222	0	0	3.160	-2.066
PCB041	2.863	0.105	-0.130	-0.144	0	0	3.121	-2.026

Table A1. Continued.

Chemical	Pred. pIC_{50} from Eq.4.1	HAT i/i (*h**=0.241)	MATS5m	MATS5v	F09[C-Br]	M_RNG	RgGrav_3D	Mor03v
PCB042	3.312	0.094	-0.130	-0.144	0	0	3.291	-2.036
PCB043	2.441	0.125	0.031	-0.048	0	0	3.082	-2.197
PCB044	2.695	0.111	0.031	-0.048	0	0	3.157	-2.136
PCB045	2.090	0.203	0.108	-0.106	0	0	2.885	-1.316
PCB046	2.194	0.204	-0.158	-0.280	0	0	2.873	-1.396
PCB048	2.669	0.123	0.136	0.030	0	0	3.151	-2.106
PCB049	3.111	0.106	0.136	0.030	0	0	3.320	-2.120
PCB050	2.510	0.183	-0.053	-0.203	0	0	2.999	-1.376
PCB051	2.691	0.196	-0.053	-0.203	0	0	3.029	-1.258
PCB052	2.575	0.185	0.297	0.127	0	0	3.219	-2.235
PCB053	2.161	0.194	0.108	-0.106	0	0	2.940	-1.405
PCB054	1.269	0.349	0.184	-0.164	0	0	2.686	-0.975
PCB055	4.003	0.082	-0.206	-0.086	0	0	3.399	-2.243
PCB056	4.179	0.080	-0.206	-0.086	0	0	3.425	-2.120
PCB057	3.644	0.075	-0.045	0.011	0	0	3.373	-2.387
PCB058	3.953	0.114	-0.311	-0.164	0	0	3.376	-2.272
PCB059	3.196	0.125	-0.234	-0.222	0	0	3.183	-1.873
PCB062	3.033	0.108	-0.130	-0.144	0	0	3.099	-1.760
PCB063	4.295	0.051	0.059	0.088	0	0	3.538	-2.103
PCB064	3.829	0.123	-0.130	-0.144	0	0	3.322	-1.535
PCB065	2.761	0.109	0.031	-0.048	0	0	3.053	-1.735
PCB066	4.671	0.057	-0.102	-0.009	0	0	3.617	-2.106
PCB067	2.669	0.123	0.136	0.030	0	0	3.151	-2.106
PCB068	3.111	0.106	0.136	0.030	0	0	3.320	-2.120

Table A1. Continued.

Chemical	Pred. pIC_{50} from Eq.4.1	HAT i/i (*h**=0.241)	MATS5m	MATS5v	F09[C-Br]	M_RNG	RgGrav_3D	Mor03v
PCB069	3.636	0.177	-0.396	-0.319	0	0	3.307	-1.947
PCB070	4.196	0.053	0.059	0.088	0	0	3.510	-2.131
PCB071	3.550	0.198	-0.396	-0.319	0	0	3.193	-1.688
PCB072	3.814	0.068	-0.045	0.011	0	0	3.435	-2.382
PCB073	3.091	0.228	-0.500	-0.396	0	0	3.132	-2.074
PCB074	4.615	0.057	0.164	0.166	0	0	3.641	-2.011
PCB075	4.215	0.160	-0.291	-0.241	0	0	3.451	-1.682
PCB076	3.682	0.089	-0.206	-0.086	0	0	3.255	-2.165
PCB078	4.858	0.059	-0.017	0.147	0	0	3.616	-2.369
PBDE078	3.182	0.083	-0.117	-0.206	1	0	3.947	-2.394
PBDE079	3.967	0.102	-0.117	-0.206	1	0	4.291	-2.557
PBDE080	5.055	0.208	-0.208	-0.286	0	0	4.306	-2.746
PBDE081	4.153	0.142	-0.027	-0.126	2	0	4.641	-2.038
PBDE082	2.934	0.119	-0.258	-0.277	1	0	3.501	-1.536
PBDE083	4.429	0.092	-0.092	-0.081	0	0	3.548	-1.722
PBDE084	4.963	0.066	-0.133	-0.049	0	0	3.780	-2.193
PBDE086	2.597	0.071	-0.003	0.017	1	0	3.443	-2.228
PBDE087	4.333	0.034	-0.003	0.017	1	0	4.026	-2.033
PBDE088	3.227	0.070	-0.044	0.049	1	0	3.330	-1.511
PBDE089	2.884	0.087	-0.299	-0.245	1	0	3.573	-2.186
PBDE090	4.080	0.039	-0.003	0.017	1	0	3.851	-1.780
PBDE091	3.857	0.038	-0.044	0.049	1	0	3.741	-2.064
PBDE092	5.585	0.072	0.163	0.214	0	0	3.812	-1.694
PBDE093	4.663	0.094	0.122	0.245	0	0	3.273	-1.439

Table A1. Continued.

Chemical	Pred. pIC_{50} from Eq.4.1	HAT i/i (*h**=0.241)	MATS5m	MATS5v	F09[C-Br]	M_RNG	RgGrav_3D	Mor03v
PBDE094	4.135	0.066	-0.133	-0.049	0	0	3.508	-2.303
PBDE095	5.783	0.105	0.122	0.245	0	0	3.675	-1.395
PBDE096	4.593	0.120	-0.173	-0.018	0	0	3.295	-1.464
PBDE097	3.860	0.050	-0.003	0.017	1	0	3.703	-1.576
PBDE098	3.416	0.090	-0.299	-0.245	1	0	3.861	-2.465
PBDE099	3.557	0.109	0.086	0.116	2	0	4.027	-1.651
PBDE101	5.086	0.093	0.252	0.312	1	0	3.983	-1.513
PBDE102	3.931	0.063	-0.044	0.049	1	0	3.586	-1.493
PBDE103	4.241	0.047	-0.044	0.049	1	0	3.769	-1.704
PBDE104	3.116	0.113	-0.339	-0.214	1	0	3.397	-1.722
PCB106	4.300	0.067	-0.051	0.070	0	0	3.539	-2.500
PCB108	4.589	0.104	-0.293	-0.124	0	0	3.640	-2.499
PCB109	3.621	0.109	-0.245	-0.207	0	0	3.404	-2.202
PCB110	4.116	0.103	-0.168	-0.034	0	0	3.582	-2.800
PCB111	3.346	0.088	-0.120	-0.117	0	0	3.339	-2.253
PCB112	3.297	0.160	-0.362	-0.311	0	0	3.369	-2.420
PCB113	5.084	0.046	0.066	0.174	0	0	3.677	-2.072
PCB115	4.331	0.086	-0.128	-0.103	0	0	3.545	-1.867
PCB116	3.108	0.084	-0.003	-0.013	0	0	3.172	-2.054
PCB117	4.108	0.068	-0.003	-0.013	0	0	3.475	-1.842
PCB119	4.276	0.168	-0.370	-0.297	0	0	3.573	-2.036
PCB120	4.602	0.066	-0.051	0.070	0	0	3.684	-2.604
PCB121	3.685	0.217	-0.487	-0.401	0	0	3.492	-2.423
PCB122	4.439	0.103	-0.293	-0.124	0	0	3.539	-2.357

Table A1. Continued.

Chemical	Pred. pIC_{50} from Eq.4.1	HAT i/i (*h**=0.241)	MATS5m	MATS5v	F09[C-Br]	M_RNG	RgGrav_3D	Mor03v
PCB123	4.891	0.069	-0.176	-0.020	0	0	3.706	-2.405
PCB124	4.448	0.061	-0.051	0.070	0	0	3.583	-2.468
PCB125	3.339	0.221	-0.487	-0.401	0	0	3.276	-2.148
PCB127	5.431	0.100	-0.099	0.153	0	0	3.836	-2.798
PCB129	3.608	0.102	-0.153	-0.098	0	0	3.486	-2.665
PCB130	3.467	0.136	-0.056	-0.014	0	0	3.518	-2.953
PCB131	3.109	0.127	-0.129	-0.215	0	0	3.342	-2.028
PCB132	3.263	0.130	-0.129	-0.215	0	0	3.359	-1.905
PCB133	3.467	0.136	-0.056	-0.014	0	0	3.518	-2.953
PCB134	2.841	0.116	-0.032	-0.131	0	0	3.276	-2.163
PCB135	2.854	0.118	-0.032	-0.131	0	0	3.308	-2.247
PCB136	2.377	0.190	0.221	-0.021	0	0	3.124	-1.690
PCB137	4.214	0.070	-0.022	0.043	0	0	3.627	-2.593
PCB138	4.384	0.061	-0.022	0.043	0	0	3.669	-2.528
PCB139	3.841	0.094	0.002	-0.074	0	0	3.485	-1.818
PCB140	3.564	0.167	-0.227	-0.300	0	0	3.478	-1.898
PCB141	3.775	0.113	0.075	0.128	0	0	3.504	-2.751
PCB142	2.938	0.105	-0.003	-0.013	0	0	3.216	-2.391
PCB143	2.808	0.131	-0.129	-0.215	0	0	3.225	-2.015
PCB144	3.399	0.089	0.099	0.011	0	0	3.380	-2.035
PCB145	2.429	0.187	0.124	-0.106	0	0	3.119	-1.581
PCB146	4.198	0.075	0.075	0.128	0	0	3.602	-2.567
PCB147	3.632	0.086	0.099	0.011	0	0	3.416	-1.877
PCB148	3.210	0.125	-0.129	-0.215	0	0	3.422	-2.164

Table A1. Continued.

Chemical	Pred. pIC_{50} from Eq.4.1	HAT i/i (*h**=0.241)	MATS5m	MATS5v	F09[C-Br]	M_RNG	RgGrav_3D	Mor03v
PCB149	3.552	0.086	0.099	0.011	0	0	3.397	-1.912
PCB150	2.724	0.183	0.124	-0.106	0	0	3.232	-1.590
PCB151	3.136	0.114	0.197	0.096	0	0	3.315	-2.167
PCB152	2.184	0.196	0.221	-0.021	0	0	3.064	-1.724
PCB154	3.849	0.090	0.002	-0.074	0	0	3.515	-1.903
PCB155	3.020	0.194	0.026	-0.190	0	0	3.343	-1.570
PCB158	4.453	0.103	-0.251	-0.183	0	0	3.684	-2.272
PCB159	4.710	0.112	-0.178	0.018	0	0	3.737	-2.930
PCB160	3.668	0.091	-0.153	-0.098	0	0	3.473	-2.553
PCB161	3.650	0.177	-0.382	-0.325	0	0	3.597	-2.735
PCB162	4.857	0.096	-0.178	0.018	0	0	3.752	-2.807
PCB163	4.226	0.073	-0.153	-0.098	0	0	3.606	-2.320
PCB164	3.692	0.166	-0.382	-0.325	0	0	3.528	-2.474
PCB165	3.377	0.157	-0.285	-0.240	0	0	3.523	-2.855
PCB166	4.555	0.048	-0.022	0.043	0	0	3.609	-2.142
PCB169	6.378	0.148	-0.071	0.277	0	0	4.032	-2.878
PCB170	4.461	0.097	-0.178	-0.062	0	0	3.757	-2.836
PCB171	3.758	0.126	-0.175	-0.233	0	0	3.598	-2.184
PCB172	4.096	0.140	-0.103	0.018	0	0	3.678	-3.117
PCB173	3.165	0.112	-0.100	-0.153	0	0	3.422	-2.430
PCB174	3.307	0.106	-0.100	-0.153	0	0	3.462	-2.389
PCB175	3.380	0.112	-0.100	-0.153	0	0	3.533	-2.526
PCB176	2.993	0.131	0.128	-0.043	0	0	3.339	-1.947
PCB177	3.533	0.100	-0.100	-0.153	0	0	3.522	-2.315

Table A1. Continued.

Chemical	Pred. pIC_{50} from Eq.4.1	HAT i/i (*h**=0.241)	MATS5m	MATS5v	F09[C-Br]	M_RNG	RgGrav_3D	Mor03v
PCB178	3.099	0.131	-0.025	-0.074	0	0	3.461	-2.728
PCB179	2.723	0.148	0.202	0.037	0	0	3.275	-2.168
PCB180	4.905	0.115	0.046	0.220	0	0	3.771	-2.921
PCB181	4.165	0.056	0.049	0.049	0	0	3.559	-2.147
PCB182	3.598	0.125	-0.175	-0.233	0	0	3.578	-2.308
PCB183	4.311	0.055	0.049	0.049	0	0	3.627	-2.191
PCB184	3.285	0.141	0.053	-0.122	0	0	3.449	-1.851
PCB185	3.721	0.087	0.124	0.128	0	0	3.453	-2.432
PCB186	2.596	0.142	0.128	-0.043	0	0	3.200	-1.972
PCB187	3.065	0.130	0.128	-0.043	0	0	3.381	-1.993
PCB188	3.065	0.130	0.128	-0.043	0	0	3.381	-1.993
PCB189	5.643	0.131	-0.181	0.109	0	0	3.933	-2.984
PCB190	4.665	0.078	-0.178	-0.062	0	0	3.759	-2.605
PCB191	4.095	0.189	-0.402	-0.343	0	0	3.777	-2.772
PCB192	3.576	0.195	-0.327	-0.264	0	0	3.669	-3.137
PCB193	3.858	0.161	-0.327	-0.264	0	0	3.695	-2.891
PCB194	5.197	0.117	-0.142	0.094	0	0	3.853	-2.991
PCB195	3.843	0.110	-0.159	-0.181	0	0	3.682	-2.566
PCB196	3.824	0.116	-0.159	-0.181	0	0	3.702	-2.651
PCB197	3.596	0.100	0.050	-0.063	0	0	3.564	-2.193
PCB198	3.444	0.143	-0.104	-0.106	0	0	3.612	-2.980
PCB199	3.583	0.122	-0.104	-0.106	0	0	3.624	-2.855
PCB200	3.175	0.112	0.106	0.013	0	0	3.419	-2.398
PCB201	3.365	0.106	0.106	0.013	0	0	3.492	-2.405

Table A1. Continued.

Chemical	Pred. pIC_{50} from Eq.4.1	HAT i/i (h^*=0.241)	MATS5m	MATS5v	F09[C-Br]	M_RNG	RgGrav_3D	Mor03v
PCB202	3.073	0.154	0.161	0.088	0	0	3.423	-2.695
PCB203	4.905	0.073	0.068	0.212	0	0	3.707	-2.574
PCB204	3.463	0.103	0.050	-0.063	0	0	3.527	-2.229
PCB205	4.064	0.203	-0.368	-0.300	0	0	3.861	-3.156
PCB206	3.846	0.152	-0.178	-0.163	0	0	3.790	-3.088
PCB207	3.830	0.091	0.018	-0.011	0	0	3.648	-2.620
PCB208	3.579	0.134	0.055	0.061	0	0	3.574	-2.911
PCDD000	5.118	0.225	-0.338	-0.097	0	1	2.603	-1.345
PCDD002	5.022	0.070	0.026	-0.036	0	1	3.076	-1.616
PCDD003	5.211	0.056	-0.078	-0.049	0	1	3.076	-1.796
PCDD004	5.258	0.046	-0.078	-0.049	0	1	3.133	-1.922
PCDD005	4.563	0.109	0.214	0.114	0	1	2.956	-1.754
PCDD006	5.584	0.109	-0.290	-0.112	0	1	3.024	-1.818
PCDD007	5.671	0.043	-0.078	-0.049	0	1	3.272	-1.877
PCDD008	5.564	0.047	-0.078	-0.049	0	1	3.216	-1.826
PCDD009	5.276	0.119	-0.290	-0.112	0	1	2.901	-1.791
PCDD010	5.177	0.066	0.134	0.015	0	1	3.275	-1.848
PCDD011	5.913	0.063	0.134	0.015	0	1	3.554	-1.868
PCDD013	5.295	0.036	-0.016	-0.023	0	1	3.254	-2.119
PCDD015	5.881	0.050	-0.208	-0.103	0	1	3.323	-2.081
PCDD016	6.176	0.028	-0.016	-0.023	0	1	3.572	-2.090
PCDD017	6.088	0.031	-0.016	-0.023	0	1	3.513	-2.006
PCDD018	5.599	0.058	-0.208	-0.103	0	1	3.193	-2.002
PCDD023	6.112	0.025	-0.016	-0.023	0	1	3.571	-2.161

Table A1. Continued.

Chemical	Pred. pIC_{50} from Eq.4.1	HAT i/i (*h**=0.241)	MATS5m	MATS5v	F09[C-Br]	M_RNG	RgGrav_3D	Mor03v
PCDD024	5.680	0.048	-0.208	-0.103	0	1	3.279	-2.178
PCDD025	5.322	0.055	0.012	0.063	0	1	3.112	-1.998
PCDD027	6.235	0.031	-0.182	-0.115	0	1	3.639	-2.472
PCDD028	6.149	0.031	-0.182	-0.115	0	1	3.624	-2.524
PCDD029	5.913	0.035	-0.004	0.062	0	1	3.430	-2.370
PCDD030	5.702	0.093	0.174	0.238	0	1	3.241	-2.200
PCDD031	5.718	0.040	-0.004	0.062	0	1	3.349	-2.346
PCDD032	5.670	0.041	-0.004	0.062	0	1	3.315	-2.296
PCDD033	6.177	0.052	0.182	0.166	0	1	3.605	-2.345
PCDD034	6.161	0.052	0.182	0.166	0	1	3.590	-2.315
PCDD036	6.283	0.032	-0.182	-0.115	0	1	3.625	-2.371
PCDD037	6.179	0.051	0.182	0.166	0	1	3.593	-2.306
PCDD038	6.154	0.031	-0.182	-0.115	0	1	3.578	-2.374
PCDD039	5.933	0.034	-0.004	0.062	0	1	3.398	-2.247
PCDD041	5.866	0.033	-0.182	-0.115	0	1	3.474	-2.385
PCDD042	6.364	0.023	0.005	-0.010	0	1	3.736	-2.361
PCDD043	6.330	0.023	0.005	-0.010	0	1	3.706	-2.307
PCDD044	5.712	0.035	-0.182	-0.115	0	1	3.409	-2.359
PCDD047	6.365	0.033	-0.182	-0.115	0	1	3.641	-2.328
PCDD048	6.969	0.064	0.191	0.095	0	1	3.999	-2.220
PCDD049	6.318	0.034	-0.040	0.041	0	1	3.663	-2.675
PCDD050	6.301	0.034	-0.040	0.041	0	1	3.650	-2.654
PCDD051	6.254	0.090	0.110	0.238	0	1	3.459	-2.532
PCDD052	6.331	0.038	-0.189	-0.157	0	1	3.800	-2.661

Table A1. Continued.

Chemical	Pred. pIC_{50} from Eq.4.1	HAT i/i (*h**=0.241)	MATS5m	MATS5v	F09[C-Br]	M_RNG	RgGrav_3D	Mor03v
PCDD054	6.874	0.030	-0.001	-0.015	0	1	3.976	-2.516
PCDD055	6.365	0.030	-0.040	0.041	0	1	3.651	-2.584
PCDD056	6.166	0.034	-0.040	0.041	0	1	3.579	-2.589
PCDD057	6.338	0.027	-0.040	0.041	0	1	3.613	-2.496
PCDD058	5.801	0.046	-0.040	0.041	0	1	3.452	-2.615
PCDD060	6.394	0.037	-0.189	-0.157	0	1	3.801	-2.591
PCDD062	6.269	0.035	-0.189	-0.157	0	1	3.728	-2.508
PCDD063	6.815	0.098	0.036	0.220	0	1	3.695	-2.853
PCDD064	6.812	0.096	0.036	0.220	0	1	3.685	-2.824
PCDD065	6.347	0.045	-0.095	-0.021	0	1	3.785	-2.927
PCDD066	7.192	0.055	0.103	0.187	0	1	3.967	-2.776
PCDD067	6.515	0.035	-0.095	-0.021	0	1	3.815	-2.824
PCDD068	6.530	0.109	0.036	0.220	0	1	3.584	-2.839
PCDD070	6.325	0.076	-0.225	-0.263	0	1	3.989	-2.828
PCDD071	6.402	0.035	-0.095	-0.021	0	1	3.766	-2.804
PCDD072	6.243	0.074	-0.225	-0.263	0	1	3.950	-2.804
PCDD073	6.997	0.108	-0.058	0.153	0	1	3.828	-3.112
PCDD074	6.150	0.074	-0.174	-0.203	0	1	3.972	-3.083
PCDE047	4.394	0.091	-0.026	-0.006	0	0	3.432	-1.519
PCDE066	4.025	0.075	-0.012	-0.057	0	0	3.599	-2.119
PCDE077	4.152	0.108	0.001	-0.107	0	0	3.792	-2.236
PCDE085	4.159	0.096	-0.142	-0.127	0	0	3.486	-1.812
PCDE099	4.011	0.085	0.164	0.118	0	0	3.633	-2.417
PCDE101	4.624	0.092	0.312	0.274	0	0	3.590	-1.772

Table A1. Continued.

Chemical	Pred. pIC_{50} from Eq.4.1	HAT i/i (*h**=0.241)	MATS5m	MATS5v	F09[C-Br]	M_RNG	RgGrav_3D	Mor03v
PCDE102	3.801	0.071	-0.021	0.075	0	0	3.323	-2.296
PCDE105	3.874	0.105	-0.117	-0.172	0	0	3.654	-2.303
PCDE118	4.060	0.109	-0.117	-0.172	0	0	3.766	-2.437
PCDE126	4.655	0.171	-0.091	-0.217	0	0	4.122	-2.486
PCDE128	3.765	0.159	-0.267	-0.305	0	0	3.560	-2.074
PCDE137	3.956	0.099	0.023	-0.022	0	0	3.747	-2.699
PCDE138	4.288	0.070	0.023	-0.022	0	0	3.692	-2.140
PCDE140	4.104	0.141	-0.304	-0.265	0	0	3.740	-2.635
PCDE147	4.495	0.054	0.113	0.187	0	0	3.502	-2.066
PCDE153	4.987	0.089	0.313	0.261	0	0	3.832	-2.029
PCDE154	3.972	0.093	-0.014	0.018	0	0	3.649	-2.765
PCDE157	3.710	0.204	-0.231	-0.345	0	0	3.845	-2.642
PCDE167	4.306	0.125	0.060	-0.062	0	0	3.980	-2.628
PCDE170	3.573	0.183	-0.127	-0.244	0	0	3.841	-2.879
PCDE180	4.448	0.120	0.157	0.103	0	0	3.958	-2.875
PCDE181	4.329	0.063	0.109	0.137	0	0	3.633	-2.406
PCDE182	3.633	0.160	-0.175	-0.210	0	0	3.766	-2.984
PCDE184	3.954	0.097	-0.223	-0.176	0	0	3.553	-2.354
PCDE190	4.026	0.134	0.157	0.103	0	0	3.807	-2.892
PCDE194	3.539	0.198	-0.011	-0.188	0	0	3.893	-2.863
PCDE196	3.569	0.189	-0.072	-0.162	0	0	3.857	-3.140
PCDE197	3.837	0.094	-0.133	-0.137	0	0	3.626	-2.526
PCDE203	5.094	0.100	0.214	0.312	0	0	3.843	-2.673

Table A1. Continued.

Chemical	Pred. pIC_{50} from Eq.4.1	HAT i/i (*h**=0.241)	MATS5m	MATS5v	F09[C-Br]	M_RNG	RgGrav_3D	Mor03v
PCDE206	3.572	0.249	0.013	-0.119	0	0	3.984	-3.384
PCDF001	3.807	0.165	-0.351	-0.345	0	1	2.605	-1.547
PCDF005	3.904	0.096	-0.140	-0.249	0	1	2.879	-1.862
PCDF006	4.483	0.075	-0.206	-0.231	0	1	2.988	-1.933
PCDF007	3.394	0.175	0.135	-0.151	0	1	2.853	-1.667
PCDF008	4.399	0.141	-0.351	-0.367	0	1	2.909	-1.684
PCDF009	4.823	0.076	-0.206	-0.231	0	1	3.081	-1.830
PCDF010	4.048	0.087	-0.140	-0.249	0	1	2.961	-1.950
PCDF011	2.732	0.192	-0.010	-0.287	0	1	2.675	-1.780
PCDF013	4.361	0.073	0.005	-0.113	0	1	3.023	-1.883
PCDF015	5.472	0.066	0.150	0.023	0	1	3.403	-1.878
PCDF017	4.786	0.067	-0.061	-0.095	0	1	2.994	-1.700
PCDF018	5.234	0.052	-0.061	-0.095	0	1	3.191	-1.796
PCDF019	5.996	0.047	0.084	0.041	0	1	3.455	-1.836
PCDF020	4.489	0.096	-0.206	-0.231	0	1	2.903	-1.661
PCDF021	4.547	0.055	-0.124	-0.152	0	1	3.082	-2.221
PCDF022	3.903	0.119	0.135	-0.076	0	1	3.037	-2.070
PCDF023	4.791	0.085	-0.260	-0.313	0	1	3.203	-2.034
PCDF024	5.431	0.040	-0.124	-0.152	0	1	3.368	-2.089
PCDF025	4.567	0.059	-0.088	-0.184	0	1	3.229	-2.311
PCDF026	3.310	0.142	-0.002	-0.237	0	1	2.927	-2.139
PCDF027	4.319	0.084	0.100	-0.045	0	1	3.066	-2.010
PCDF029	5.236	0.037	-0.124	-0.152	0	1	3.358	-2.283
PCDF031	3.966	0.093	-0.037	-0.205	0	1	3.071	-2.168

Table A1. Continued.

Chemical	Pred. pIC_{50} from Eq.4.1	HAT i/i (*h**=0.241)	MATS5m	MATS5v	F09[C-Br]	M_RNG	RgGrav_3D	Mor03v
PCDF032	4.220	0.093	-0.037	-0.205	0	1	3.065	-1.855
PCDF033	4.908	0.067	0.100	-0.045	0	1	3.280	-1.996
PCDF034	4.367	0.102	0.135	-0.076	0	1	3.221	-2.107
PCDF035	3.268	0.219	0.221	-0.129	0	1	2.961	-1.883
PCDF037	5.699	0.029	0.013	0.008	0	1	3.401	-2.149
PCDF038	6.601	0.044	0.150	0.168	0	1	3.634	-2.100
PCDF040	4.906	0.044	-0.124	-0.152	0	1	3.209	-2.200
PCDF041	5.067	0.042	-0.124	-0.152	0	1	3.247	-2.133
PCDF042	5.915	0.026	0.013	0.008	0	1	3.490	-2.178
PCDF043	5.367	0.039	0.048	-0.023	0	1	3.437	-2.314
PCDF044	5.372	0.048	-0.159	-0.121	0	1	3.199	-1.963
PCDF045	6.344	0.034	-0.022	0.039	0	1	3.485	-1.998
PCDF047	5.104	0.087	-0.296	-0.282	0	1	3.174	-1.919
PCDF048	4.488	0.082	-0.260	-0.313	0	1	3.135	-2.174
PCDF050	4.726	0.058	-0.066	-0.175	0	1	3.326	-2.380
PCDF051	5.638	0.035	-0.125	-0.102	0	1	3.534	-2.651
PCDF052	5.015	0.040	-0.080	-0.127	0	1	3.283	-2.244
PCDF053	4.687	0.057	-0.066	-0.175	0	1	3.283	-2.291
PCDF054	6.280	0.027	-0.139	-0.053	0	1	3.612	-2.485
PCDF055	5.993	0.030	0.057	0.070	0	1	3.507	-2.286
PCDF056	5.567	0.037	0.071	0.021	0	1	3.476	-2.345
PCDF057	4.864	0.072	0.086	-0.027	0	1	3.390	-2.556
PCDF059	5.228	0.039	-0.080	-0.127	0	1	3.446	-2.505
PCDF060	6.117	0.058	-0.291	-0.201	0	1	3.535	-2.299

Table A1. Continued.

Chemical	Pred. pIC_{50} from Eq.4.1	HAT i/i (*h**=0.241)	MATS5m	MATS5v	F09[C-Br]	M_RNG	RgGrav_3D	Mor03v
PCDF061	4.986	0.071	-0.262	-0.299	0	1	3.419	-2.569
PCDF062	5.246	0.038	-0.080	-0.127	0	1	3.334	-2.137
PCDF063	5.437	0.042	0.071	0.021	0	1	3.447	-2.405
PCDF064	4.531	0.092	0.131	-0.052	0	1	3.299	-2.313
PCDF066	5.053	0.061	-0.276	-0.250	0	1	3.274	-2.378
PCDF067	6.180	0.028	-0.139	-0.053	0	1	3.541	-2.381
PCDF068	4.645	0.073	0.071	0.021	0	1	3.146	-2.386
PCDF069	5.458	0.038	-0.125	-0.102	0	1	3.457	-2.619
PCDF070	3.896	0.210	0.341	0.023	0	1	3.148	-2.040
PCDF073	5.685	0.029	-0.125	-0.102	0	1	3.506	-2.509
PCDF074	4.314	0.097	0.131	-0.052	0	1	3.189	-2.222
PCDF075	5.774	0.062	-0.276	-0.250	0	1	3.489	-2.212
PCDF076	3.742	0.147	0.145	-0.101	0	1	3.120	-2.331
PCDF078	4.672	0.062	-0.066	-0.175	0	1	3.336	-2.475
PCDF082	6.555	0.044	-0.154	-0.004	0	1	3.518	-2.217
PCDF083	7.131	0.044	-0.003	0.144	0	1	3.714	-2.301
PCDF084	4.050	0.131	0.145	-0.101	0	1	3.217	-2.278
PCDF085	4.183	0.081	-0.066	-0.175	0	1	3.163	-2.502
PCDF086	3.923	0.109	-0.051	-0.224	0	1	3.189	-2.543
PCDF087	5.631	0.039	-0.141	-0.118	0	1	3.558	-2.717
PCDF090	5.631	0.039	-0.141	-0.118	0	1	3.558	-2.717
PCDF092	5.464	0.039	-0.141	-0.118	0	1	3.470	-2.635
PCDF094	5.633	0.080	-0.323	-0.268	0	1	3.628	-2.922
PCDF096	5.670	0.042	0.042	0.032	0	1	3.534	-2.602

Table A1. Continued.

Chemical	Pred. pIC_{50} from Eq.4.1	HAT i/i (*h**=0.241)	MATS5m	MATS5v	F09[C-Br]	M_RNG	RgGrav_3D	Mor03v
PCDF097	5.028	0.096	0.222	0.108	0	1	3.366	-2.425
PCDF099	5.247	0.046	-0.141	-0.118	0	1	3.396	-2.656
PCDF100	5.819	0.029	-0.141	-0.118	0	1	3.565	-2.520
PCDF101	6.511	0.061	-0.181	-0.013	0	1	3.690	-2.877
PCDF102	6.505	0.054	0.002	0.138	0	1	3.599	-2.611
PCDF105	5.688	0.069	-0.323	-0.268	0	1	3.585	-2.724
PCDF106	6.460	0.064	-0.321	-0.193	0	1	3.690	-2.568
PCDF108	5.163	0.053	0.040	-0.043	0	1	3.466	-2.566
PCDF109	7.135	0.075	-0.179	0.062	0	1	3.704	-2.607
PCDF110	4.929	0.069	0.040	-0.043	0	1	3.417	-2.685
PCDF111	4.401	0.108	0.038	-0.118	0	1	3.381	-2.774
PCDF113	4.527	0.085	0.040	-0.043	0	1	3.256	-2.648
PCDF114	4.851	0.074	-0.143	-0.193	0	1	3.444	-2.853
PCDF114	6.239	0.065	-0.321	-0.193	0	1	3.653	-2.708
PCDF115	5.633	0.091	-0.251	-0.182	0	1	3.661	-3.179
PCDF116	6.554	0.067	-0.234	-0.059	0	1	3.725	-2.922
PCDF118	6.407	0.092	0.103	0.226	0	1	3.600	-2.757
PCDF119	5.702	0.077	0.086	0.102	0	1	3.567	-2.856
PCDF120	4.855	0.122	0.069	-0.021	0	1	3.513	-3.060
PCDF121	6.562	0.062	-0.234	-0.059	0	1	3.700	-2.833
PCDF122	6.042	0.107	-0.240	-0.121	0	1	3.772	-3.341
PCDF124	5.987	0.067	-0.074	0.022	0	1	3.620	-2.972
PCDF125	5.387	0.097	0.086	0.102	0	1	3.460	-2.888
PCDF126	6.184	0.055	-0.074	0.022	0	1	3.676	-2.921

Table A1. Continued.

Chemical	Pred. pIC_{50} from Eq.4.1	HAT i/i (*h**=0.241)	MATS5m	MATS5v	F09[C-Br]	M_RNG	RgGrav_3D	Mor03v
PCDF127	5.289	0.087	-0.091	-0.101	0	1	3.605	-3.125
PCDF129	5.189	0.080	-0.091	-0.101	0	1	3.538	-3.035
PCDF130	5.780	0.080	-0.251	-0.182	0	1	3.696	-3.116
PCDF131	6.355	0.129	-0.387	-0.221	0	1	3.850	-3.331
PCDF132	6.996	0.175	-0.061	0.225	0	1	3.711	-3.164
PCDF133	5.640	0.125	-0.094	-0.020	0	1	3.663	-3.367
PCDF135	5.461	0.169	-0.292	-0.264	0	1	3.828	-3.626
PCDT000	4.643	0.207	-0.153	0.083	0	1	2.492	-1.713
PCDT002	3.176	0.187	-0.040	-0.340	0	1	2.974	-2.039
PCDT007	1.907	0.481	0.338	-0.214	0	1	2.873	-2.178
PCDT046	4.336	0.098	-0.133	-0.288	0	1	3.354	-2.594
PCDT076	4.166	0.155	0.053	-0.186	0	1	3.456	-2.834
PCDT104	5.914	0.051	-0.132	-0.043	0	1	3.586	-2.855
PCDT125	4.559	0.188	-0.078	-0.215	0	1	3.663	-3.456
PCDT132	4.966	0.208	-0.161	-0.145	0	1	3.652	-3.719
PCDT135	6.772	0.359	-0.306	0.066	0	1	3.819	-3.986
PCPhX000	4.871	0.162	-0.207	0.007	0	1	2.640	-1.732
PCPhX003	3.173	0.279	0.215	-0.211	0	1	3.110	-2.027
PCPhX010	4.674	0.170	0.157	-0.168	0	1	3.567	-2.215
PCPhX095	7.164	0.146	-0.350	-0.017	0	1	3.676	-2.824
PCPhX128	6.619	0.059	-0.165	-0.029	0	1	3.825	-3.009
PCPhX133	5.110	0.243	0.032	-0.067	0	1	3.868	-3.783
PCPhX135	5.366	0.235	-0.084	-0.169	0	1	4.008	-3.883
PCTA001	2.772	0.184	-0.067	-0.315	0	1	2.842	-2.357

Table A1. Continued.

Name	Pred. by model eq.	HAT i/i (h*=0.2414)	MATS5m	MATS5v	F09[C-Br]	M_RNG	RgGrav__3D	Mor03v
PCTA004	3.179	0.210	-0.051	-0.343	0	1	3.171	-2.684
PCTA015	4.388	0.104	-0.151	-0.232	0	1	3.356	-2.935
PCTA047	3.187	0.453	0.161	-0.333	0	1	3.613	-3.142
PCTA055	6.691	0.202	-0.284	0.007	0	1	3.788	-3.555
PCTA069	5.077	0.239	-0.141	-0.173	0	1	3.825	-3.884
PCTA073	5.906	0.273	-0.225	-0.110	0	1	3.974	-4.122
PCTA075	6.136	0.457	-0.219	0.015	0	1	3.979	-4.541
Pentamethylantracene	2.326	0.165	-0.277	-0.277	0	0	3.042	-2.328
Phenanthrene	1.699	0.168	-0.139	-0.139	0	0	2.560	-1.696
Quinoline	2.911	0.249	-0.140	-0.125	0	1	2.019	-1.025
Tetramethylanthracene	2.820	0.095	-0.030	-0.030	0	0	3.041	-2.009
Tjipanazole	6.480	0.057	0.146	0.083	0	1	3.943	-2.748

APPENDIX A2

Table A2. Hat, descriptor and predicted pIC_{50} values from the TCDF-based model Equation 4.2 for external set chemicals

Name	Pred. pIC_{50} from Eq. 4.2	HAT i/i (*h**=0.329)	RFD	MATS5s	Tm	nHAcc	B04[O-Cl]	F04[Cl-Cl]	LOC
1,2,3,4,6,7,8-HpBDF	9.816	0.130	0.308	-0.053	18.554	1	0	0	0.746
1,2,3,4,7,8-HxBDD	9.501	0.137	0.286	0.485	20.343	2	0	0	0.737
1,2,3,4,7,8-HxBDF	9.365	0.094	0.308	0.244	18.529	1	0	0	0.749
1,2,3,6,7,8-HxBDD	9.820	0.136	0.286	0.27	20.354	2	0	0	0.737
1,2,3,7,8,9-HxBDD	9.888	0.148	0.286	0.177	20.25	2	0	0	0.737
1,2,3,7,8-PeBDD	9.966	0.138	0.286	0.383	20.827	2	0	0	0.73
1,2,3,7,8-PeBDF	9.601	0.105	0.308	0.048	18.444	1	0	0	0.745
1,2,6,9-tetramethylphenanthrene	5.019	0.127	0.286	-0.008	10.195	0	0	0	0.714
1,2,6-trimethylphenanthrene	4.787	0.135	0.286	0.045	9.883	0	0	0	0.682
1,2,9-trimethylhenanthrene	4.469	0.140	0.286	-0.028	9.216	0	0	0	0.682
1,2-dimethylphenanthrene	4.285	0.137	0.286	-0.012	8.84	0	0	0	0.625
1,3,6-trimethylchrysene	7.059	0.098	0.333	-0.209	12.474	0	0	0	0.61
1-bromo-2methylnaphthalene	2.952	0.190	0.2	-0.186	7.029	0	0	0	0.73
1-bromo-4-methylnaphthalene	3.997	0.133	0.2	-0.124	8.815	0	0	0	0.73
1-methylbenz[a]anthracene	6.050	0.083	0.333	-0.116	10.749	0	0	0	0.447
1-methylchrysene	6.311	0.081	0.333	-0.149	11.085	0	0	0	0.447
1-methylnaphthalene	1.990	0.237	0.2	-0.281	5.081	0	0	0	0.629

Table A2. Continued.

Name	Pred. pIC_{50} from Eq. 4.2	HAT i/i (*h**=0.329)	RFD	MATS5s	Tm	nHAcc	B04[O-Cl]	F04[Cl-Cl]	LOC
2,2,3,3-tetrachloroDPS	4.089	0.221	0	-0.315	9.597	0	0	0	0.732
2,2',3,4,5-pentaCDPS	4.901	0.106	0	-0.01	12.441	0	0	1	0.745
2,2,3-trichloroDPS	3.852	0.203	0	-0.211	9.399	0	0	0	0.703
2,2,4,4',5pentaCDPS	4.805	0.130	0	0.108	13.403	0	0	2	0.745
2,2,4,5-tetrachloroDPS	4.306	0.137	0	0.132	11.801	0	0	1	0.732
2,3,3',4,4',5,6-heptaCDPS	4.745	0.326	0	0.149	15.089	0	0	4	0.746
2,3,3',4,5,6-hexaCDPS	3.950	0.353	0	0.177	13.908	0	0	4	0.749
2,3,3-trichloroDPS	5.160	0.175	0	-0.207	11.467	0	0	0	0.703
2,3,4,4',5,6-hexaCDPS	4.078	0.408	0	0.292	14.371	0	0	4	0.749
2,3,4,5,6-pentaCPDS	3.134	0.466	0	0.332	12.968	0	0	4	0.745
2,3,4,5-tetrachloroDPS	6.052	0.110	0	0.145	14.58	0	0	1	0.732
2,3,4,6-TeCDE	5.167	0.237	0	-0.055	11.472	1	1	1	0.732
2,3,4,7,8-PeBDF	9.862	0.110	0.308	0.293	19.414	1	0	0	0.745
2,3,7,8-TeBDF	9.878	0.112	0.308	0.32	19.474	1	0	0	0.732
2,3-dichloro-DPS	3.423	0.206	0	-0.098	8.864	0	0	0	0.648
2,4,4,5-tetrachloroDPS	5.759	0.138	0	0.268	14.399	0	0	1	0.732
2,4,5-trichloroDPS	5.122	0.189	0	0.31	12.587	0	0	0	0.703
2,4,6-trichloroDPS	3.897	0.133	0	0.004	10.803	0	0	1	0.703
2,4,7-trimethyldibenzothiophene	4.858	0.190	0.308	-0.456	8.778	0	0	0	0.703
2,4-dimethyldibenzothiophene	4.706	0.192	0.308	-0.499	8.323	0	0	0	0.648

Table A2. Continued.

Name	Pred. pIC_{50} from Eq. 4.2	HAT i/i (*h**=0.329)	RFD	MATS5s	Tm	nHAcc	B04[O-Cl]	F04[Cl-Cl]	LOC
2-methyl-BDE068	7.757	0.481	0	-0.283	17.836	2	0	0	0.954
2-methyl-BDE123	8.660	0.415	0	-0.118	19.632	2	0	0	0.953
2-methyl-dibenzothiophene	3.842	0.166	0.308	0.036	7.961	0	0	0	0.544
2-methylnaphthalene	1.998	0.239	0.2	-0.038	5.648	0	0	0	0.629
2-methyl-phenanthrene	4.222	0.117	0.286	-0.11	8.295	0	0	0	0.521
2-OH-BDE007	4.608	0.339	0	-0.022	12.937	2	0	0	0.703
2-OH-BDE028	7.817	0.329	0	0.015	18.135	2	0	0	0.732
2-OHBDE-066	8.568	0.375	0	-0.041	19.218	2	0	0	0.745
2-OH-BDE068	7.549	0.290	0	0.121	17.984	2	0	0	0.745
2-OH-BDE123	8.616	0.310	0	0.153	19.745	2	0	0	0.749
2'-OH-CB005	2.836	0.273	0	-0.02	9.241	1	0	0	0.725
2'-OH-CB009	2.537	0.287	0	0.151	9.161	1	0	0	0.725
2'-OH-CB012	4.775	0.237	0	-0.177	11.935	1	0	0	0.725
2-OHCB025	3.668	0.175	0	-0.035	11.412	1	0	1	0.75
2-OH-CB030	1.334	0.358	0	0.147	9.839	1	0	3	0.75
2-OH-CB035	5.458	0.233	0	-0.191	13.032	1	0	0	0.75
2-OH-CB036	4.079	0.158	0	0.035	12.22	1	0	1	0.75
2-OH-CB036	4.493	0.197	0	-0.262	12.194	1	0	1	0.75
2-OH-CB039	4.537	0.145	0	0.104	13.098	1	0	1	0.75
2-OH-CB056	6.426	0.360	0	-0.178	12.39	1	1	0	0.76
2-OH-CB061	3.820	0.161	0	0.009	12.617	1	0	2	0.76
2-OH-CB079	5.194	0.134	0	0.018	13.958	1	0	1	0.76
2-OH-CB080	4.322	0.149	0	-0.056	13.258	1	0	2	0.76
3,3-diindoymethane	6.983	0.223	0.222	0.039	12.208	0	0	0	0

Table A2. Continued.

Name	Pred. pIC_{50} from Eq. 4.2	HAT i/i (*h**=0.329)	RFD	MATS5s	Tm	nHAcc	B04[O-Cl]	F04[Cl-Cl]	LOC
3,4-dichloroDPS	5.224	0.156	0	-0.055	11.797	0	0	0	0.648
3-methyl-BDE100	5.750	0.340	0	-0.032	15.247	2	0	0	0.953
3-methylbenz[a]anthracene	6.759	0.080	0.333	0	12.13	0	0	0	0.447
3-methylchrysene	6.365	0.081	0.333	-0.033	11.435	0	0	0	0.447
3-methyldibenzo[ah]anthracene	9.531	0.141	0.364	-0.036	16.131	0	0	0	0.393
3-methylphenanthrene	3.852	0.146	0.286	0.012	7.991	0	0	0	0.521
3-OH-BDE007	5.604	0.353	0	-0.103	14.321	2	0	0	0.703
3-OH-BDE028	7.260	0.329	0	-0.018	17.183	2	0	0	0.732
3-OH-BDE047	5.906	0.338	0	-0.073	14.954	2	0	0	0.745
3-OH-BDE100	5.556	0.285	0	0.126	14.866	2	0	0	0.749
3-OH-BDE154	6.786	0.273	0	0.157	16.866	2	0	0	0.746
3-OH-CB009	3.270	0.247	0	0.03	10.038	1	0	0	0.725
3-OH-CB028	4.604	0.154	0	-0.076	12.792	1	0	1	0.75
3-OH-CB030	2.496	0.259	0	-0.309	10.628	1	0	3	0.75
3-OH-CB031	3.270	0.247	0	0.03	10.038	1	0	0	0.725
3-OH-CB061	4.438	0.145	0	-0.03	13.501	1	0	2	0.76
3-OH-CB065	3.046	0.189	0	-0.074	11.209	1	0	2	0.76
3-OH-CB066	5.323	0.154	0	-0.134	13.813	1	0	1	0.76
3-OH-CB068	4.403	0.164	0	-0.203	13.051	1	0	2	0.76
4,4-dichloroDPS	6.201	0.147	0	0.102	13.693	0	0	0	0.648
4'-acetyl-PCB061	6.883	0.147	0	-0.005	17.861	1	0	2	0.972
4'-bromo-PCB061	6.690	0.121	0	0.096	16.375	0	0	2	0.76
4'-cyano-PCB061	5.101	0.158	0	-0.096	14.858	1	0	2	0.977
4'-ethyl-PCB061	7.269	0.150	0	0.191	17.968	0	0	2	0.977

A2. Continued

Name	Pred. pIC_{50} from Eq. 4.2	HAT i/i (*h**=0.329)	RFD	MATS5s	Tm	nHAcc	B04[O-Cl]	F04[Cl-Cl]	LOC
4'-fluoro-PCB061	4.233	0.151	0	0.04	13.338	1	0	2	0.76
4-hydroxy-BDE017	5.279	0.274	0	0.266	14.713	2	0	0	0.732
4-hydroxy-BDE042	5.867	0.267	0	0.22	15.562	2	0	0	0.745
4-hydroxy-BDE049	6.223	0.257	0	0.307	16.321	2	0	0	0.745
4-hydroxy-BDE090	6.614	0.258	0	0.26	16.837	2	0	0	0.749
4-hydroxy-CB001	2.377	0.300	0	-0.039	8.364	1	0	0	0.673
4-hydroxy-CB002	3.432	0.238	0	0.065	10.261	1	0	0	0.673
4-hydroxy-CB009	3.388	0.238	0	0.137	10.468	1	0	0	0.725
4-hydroxy-CB014	3.630	0.174	0	0.062	11.521	1	0	1	0.725
4-hydroxy-CB020	4.380	0.224	0	-0.095	11.554	1	0	0	0.75
4-hydroxy-CB025	4.418	0.156	0	-0.068	12.518	1	0	1	0.75
4-hydroxy-CB026	5.607	0.318	0	0.01	11.508	1	1	0	0.75
4-hydroxy-CB031	6.109	0.306	0	0.074	12.444	1	1	0	0.75
4-hydroxy-CB033	6.613	0.346	0	-0.123	12.789	1	1	0	0.75
4-hydroxy-CB035	5.892	0.170	0	0.078	14.329	1	0	0	0.75
4-hydroxy-CB036	4.829	0.138	0	0.051	13.436	1	0	1	0.75
4-hydroxy-CB070	6.904	0.315	0	0.009	13.569	1	1	0	0.76
4-hydroxy-CB079	5.735	0.126	0	0.063	14.911	1	0	1	0.76
4-hydroxy-CB106	5.026	0.136	0	0.029	14.56	1	0	2	0.76
4'-hydroxy-PCB061	4.538	0.148	0	0.097	13.947	1	0	2	0.76
4'-iodo-PCB061	8.062	0.152	0	0.101	18.546	0	0	2	0.76
4'-isopropyl-PCB061	8.679	0.180	0	0.171	20.131	0	0	2	0.972
4-methoxy-BDE017	6.691	0.258	0	0.358	17.604	2	0	0	0.946
4-methoxy-BDE049	6.919	0.266	0	0.25	17.734	2	0	0	0.954
4-methoxy-BDE090	6.091	0.329	0	-0.008	15.839	2	0	0	0.953

Table A2. Continued

Name	Pred. pIC_{50} from Eq. 4.2	HAT i/i (*h**=0.329)	RFD	MATS5s	Tm	nHAcc	B04[O-Cl]	F04[Cl-Cl]	LOC
4'-methoxy-PCB061	5.919	0.146	0	0.161	16.732	1	0	2	0.977
4-methylbenz[a]anthracene	6.461	0.080	0.333	-0.031	11.59	0	0	0	0.447
4'-methyl-PCB061	5.950	0.142	0	0.208	15.466	0	0	2	0.76
4-N-acetylamino	7.810	0.334	0	-0.033	20.636	2	0	2	1.169
4'-n-butyl-PCB061	11.086	0.399	0	0.196	24.817	0	0	2	1.364
4'-nitro-PCB061	5.290	0.316	0	-0.196	15.875	2	0	2	0.972
4'-phenyl-PCB061	11.269	0.372	0	0.121	23.404	0	0	2	0.65
4'-tributyl-PCB061	9.776	0.226	0	0.151	21.791	0	0	2	0.962
4'-trifluoromethyl-PCB061	5.539	0.569	0	-0.235	17.115	3	0	2	0.962
5-chloro-6-hydroxy-BDE047	8.614	0.473	0	-0.052	17.057	2	1	0	0.749
5-chloro-6-methoxy-BDE047	7.862	0.484	0	-0.197	15.979	2	1	0	0.953
5-methoxy-BDE047	6.470	0.267	0	0.26	17.049	2	0	0	0.954
5-OH-BDE047	5.800	0.272	0	0.185	15.376	2	0	0	0.745
5-OH-CB002	4.770	0.354	0	-0.027	9.942	1	1	0	0.673
5-OH-CB0025	3.879	0.167	0	0.128	12.117	1	0	1	0.75
5-OH-CB033	6.159	0.448	0	-0.382	11.482	1	1	0	0.75
5-OH-CB034	3.610	0.175	0	0.066	11.551	1	0	1	0.75
5-OH-CB066	5.087	0.133	0	0.115	14.011	1	0	1	0.76
5-OH-CB066	6.875	0.279	0	-0.235	13.809	1	1	1	0.76
5-OH-CB068	4.145	0.154	0	0.045	13.21	1	0	2	0.76
6-chloro-2-hydroxy-BDE007	7.073	0.489	0	-0.177	14.31	2	1	0	0.732
6-chloro-2-hydroxy-BDE068	8.647	0.435	0	0.052	17.346	2	1	0	0.749
6-chloro-2-methoxy-BDE068	8.593	0.466	0	-0.118	17.31	2	1	0	0.953

Table A2. Continued

Name	Pred. pIC_{50} from Eq. 4.2	HAT i/i (*h**=0.329)	RFD	MATS5s	Tm	nHAcc	B04[O-Cl]	F04[Cl-Cl]	LOC
6-ethylchrysene	6.071	0.108	0.333	-0.131	11.178	0	0	0	0.648
6-formylindolo[3,2-b]carbazole	6.771	0.100	0.4	0.051	13.212	1	0	0	0.608
6-hydroxy-BDE017	5.343	0.359	0	-0.118	13.937	2	0	0	0.732
6-hydroxy-BDE047	7.030	0.309	0	0.031	16.934	2	0	0	0.732
6-hydroxy-BDE082	7.189	0.335	0	-0.04	17.058	2	0	0	0.749
6-hydroxy-BDE085	7.160	0.290	0	0.098	17.327	2	0	0	0.749
6-hydroxy-BDE087	7.484	0.319	0	0.018	17.654	2	0	0	0.749
6-hydroxy-BDE090	6.720	0.287	0	0.095	16.628	2	0	0	0.749
6-hydroxy-BDE099	7.071	0.265	0	0.217	17.459	2	0	0	0.749
6-hydroxy-BDE137	7.697	0.293	0	0.12	18.217	2	0	0	0.746
6-hydroxy-BDE140	6.245	0.323	0	-0.033	15.581	2	0	0	0.746
6-hydroxy-BDE157	8.324	0.314	0	0.104	19.167	2	0	0	0.746
6-hydroxy-CB026	5.414	0.390	0	-0.257	10.595	1	1	0	0.75
6-hydroxy-CB031	5.494	0.314	0	0.057	11.437	1	1	0	0.75
6-hydroxy-CB035	5.181	0.181	0	0.063	13.175	1	0	0	0.75
6-hydroxy-CB036	4.247	0.155	0	-0.008	12.386	1	0	1	0.75
6-hydroxy-CB058	5.731	0.282	0	-0.284	11.897	1	1	1	0.76
6-hydroxy-CB070	6.595	0.397	0	-0.269	12.448	1	1	0	0.76
6-hydroxy-CB106	4.204	0.162	0	0.144	13.529	1	0	2	0.76
6-methoxy-BDE017	5.898	0.390	0	-0.159	15.176	2	0	0	0.946
6-methoxy-BDE085	6.346	0.290	0	0.123	16.539	2	0	0	0.953
6-methoxy-BDE090	6.256	0.404	0	-0.186	15.693	2	0	0	0.953
6-methoxy-BDE137	7.218	0.338	0	-0.028	17.555	2	0	0	0.947
7-methylbenz[a]anthracene	6.066	0.083	0.333	-0.142	10.715	0	0	0	0.447

Table A2. Continued.

Name	Pred. pIC_{50} from Eq. 4.2	HAT i/i (*h**=0.329)	RFD	MATS5s	Tm	nHAcc	B04[O-Cl]	F04[Cl-Cl]	LOC
8,9,11-trimethylbenz[a]anthracene	7.115	0.089	0.333	-0.055	12.914	0	0	0	0.61
9-methylbenz[a]anthracene	6.803	0.080	0.333	0	12.199	0	0	0	0.447
Acridine	3.740	0.217	0.286	-0.073	7.463	1	0	0	0
Anthracene	4.168	0.250	0.286	0.125	7.631	0	0	0	0
Benz[a]acridine	5.712	0.161	0.333	-0.028	10.42	1	0	0	0
Benz[b]anthracene	6.965	0.206	0.333	0.116	11.762	0	0	0	0
Benz[c]acridine	5.763	0.167	0.333	-0.131	10.265	1	0	0	0
Carbazole	3.217	0.248	0.308	-0.258	6.1	1	0	0	0
Chrysene	6.307	0.177	0.333	-0.134	10.156	0	0	0	0
Decamethylanthracene	6.885	0.087	0.286	-0.115	12.895	0	0	0	0.717
Dibenz[a,c]acridine	6.839	0.159	0.364	-0.073	11.926	1	0	0	0
Dibenz[a,c]anthracene	7.476	0.195	0.364	-0.012	12.109	0	0	0	0
Dibenz[a,h]acridine	8.191	0.185	0.364	-0.076	14.047	1	0	0	0
Dibenz[a,j]acridine	7.680	0.168	0.364	-0.028	13.352	1	0	0	0
Dibenz[a,j]anthracene	8.358	0.213	0.364	-0.061	13.386	0	0	0	0
Dibenz[a,l]acridine	8.695	0.192	0.364	0.065	15.162	1	0	0	0
Dibenz[c,h]acridine	7.683	0.189	0.364	-0.186	12.996	1	0	0	0
Dibenzothiphene	4.031	0.209	0.308	-0.153	6.662	0	0	0	0
Indole	1.065	0.388	0.222	-0.082	3.573	1	0	0	0
Indole3-acetonitrile	1.176	0.436	0.222	-0.17	5.566	1	0	0	0.943
Indole-3-carbinol	0.968	0.450	0.222	-0.04	5.536	1	0	0	0.943
Indolo[2,3-c]carbazole	6.321	0.222	0.4	0.116	10.39	0	0	0	0

Table A2. Continued.

Name	Pred. pIC_{50} from Eq. 4.2	HAT i/i (*h**=0.329)	RFD	MATS5s	Tm	nHAcc	B04[O-Cl]	F04[Cl-Cl]	LOC
Indolo[3,2-b]carbazole	8.077	0.212	0.4	-0.016	12.854	0	0	0	0
Napthalene	2.172	0.318	0.2	-0.077	4.487	0	0	0	0
OBDD	9.397	0.125	0.286	0.242	19.617	2	0	0	0.733
OBDF	9.504	0.113	0.308	-0.034	18.094	1	0	0	0.74
PBB040	5.826	0.174	0	-0.22	12.587	0	0	0	0.75
PBB041	6.383	0.157	0	-0.135	13.657	0	0	0	0.75
PBB042	6.484	0.158	0	-0.135	13.816	0	0	0	0.75
PBB043	5.490	0.144	0	0.025	12.616	0	0	0	0.75
PBB044	5.506	0.144	0	0.035	12.664	0	0	0	0.75
PBB045	3.840	0.201	0	0.158	10.323	0	0	0	0.75
PBB046	4.300	0.188	0	-0.193	10.246	0	0	0	0.75
PBB047	7.141	0.150	0	-0.041	15.065	0	0	0	0.75
PBB048	6.061	0.141	0	0.12	13.732	0	0	0	0.75
PBB049	6.161	0.140	0	0.12	13.89	0	0	0	0.75
PBB050	4.812	0.157	0	-0.002	11.487	0	0	0	0.75
PBB051	4.770	0.158	0	0.007	11.442	0	0	0	0.75
PBB052	5.178	0.180	0	0.289	12.728	0	0	0	0.75
PBB053	3.982	0.184	0	0.061	10.325	0	0	0	0.75
PBB054	2.162	0.333	0	0.292	7.986	0	0	0	0.75
PBB055	8.053	0.199	0	-0.171	16.205	0	0	0	0.75
PBB056	8.012	0.195	0	-0.162	16.16	0	0	0	0.75
PBB057	6.927	0.140	0	0.105	15.061	0	0	0	0.75
PBB058	7.395	0.202	0	-0.247	14.995	0	0	0	0.75
PBB059	5.906	0.174	0	-0.22	12.712	0	0	0	0.75
PBB060	8.684	0.202	0	-0.086	17.391	0	0	0	0.75

Table A2. Continued.

Name	Pred. pIC_{50} from Eq. 4.2	HAT i/i (*h**=0.329)	RFD	MATS5s	Tm	nHAcc	B04[O-Cl]	F04[Cl-Cl]	LOC
PBB062	6.370	0.144	0	-0.047	13.837	0	0	0	0.75
PBB063	7.595	0.151	0	0.181	16.286	0	0	0	0.75
PBB064	6.385	0.142	0	-0.029	13.902	0	0	0	0.75
PBB065	5.292	0.163	0	0.219	12.748	0	0	0	0.75
PBB066	8.012	0.195	0	-0.162	16.16	0	0	0	0.75
PBB067	7.684	0.151	0	0.084	16.206	0	0	0	0.75
PBB068	8.236	0.240	0	-0.277	16.251	0	0	0	0.75
PBB069	7.070	0.300	0	-0.496	13.915	0	0	0	0.75
PBB070	7.637	0.150	0	0.093	16.152	0	0	0	0.75
PBB071	6.880	0.285	0	-0.477	13.66	0	0	0	0.75
PBB072	7.058	0.144	0	0.008	15.047	0	0	0	0.75
PBB073	6.492	0.397	0	-0.678	12.59	0	0	0	0.75
PBB074	8.348	0.167	0	0.169	17.445	0	0	0	0.75
PBB075	7.546	0.222	0	-0.295	15.123	0	0	0	0.75
PBB077	9.839	0.228	0	0.133	19.71	0	0	0	0.75
PBB079	9.093	0.193	0	0.154	18.583	0	0	0	0.75
PBB080	8.324	0.167	0	0.184	17.441	0	0	0	0.75
PBB082	7.370	0.201	0	-0.248	14.974	0	0	0	0.76
PBB083	6.570	0.152	0	-0.098	14.058	0	0	0	0.76
PBB084	5.282	0.155	0	-0.092	12.043	0	0	0	0.76
PBB085	7.923	0.188	0	-0.15	16.069	0	0	0	0.76
PBB086	6.951	0.144	0	-0.009	14.86	0	0	0	0.76
PBB087	7.053	0.145	0	0	15.041	0	0	0	0.76
PBB088	5.607	0.144	0	0.1	12.994	0	0	0	0.76
PBB089	6.024	0.178	0	-0.233	12.89	0	0	0	0.76

Table A2. Continued.

Name	Pred. pIC_{50} from Eq. 4.2	HAT i/i (*h**=0.329)	RFD	MATS5s	Tm	nHAcc	B04[O-Cl]	F04[Cl-Cl]	LOC
PBB090	7.127	0.146	0	-0.009	15.137	0	0	0	0.76
PBB091	5.658	0.144	0	0.109	13.094	0	0	0	0.76
PBB092	6.251	0.141	0	0.15	14.122	0	0	0	0.76
PBB093	4.814	0.180	0	0.241	12.067	0	0	0	0.76
PBB094	4.814	0.180	0	0.241	12.067	0	0	0	0.76
PBB095	4.965	0.162	0	0.156	12.111	0	0	0	0.76
PBB096	3.516	0.237	0	0.256	10.057	0	0	0	0.76
PBB097	7.054	0.145	0	0	15.043	0	0	0	0.76
PBB098	6.149	0.181	0	-0.242	13.066	0	0	0	0.76
PBB099	7.608	0.148	0	0.097	16.136	0	0	0	0.76
PBB100	6.522	0.143	0	-0.032	14.133	0	0	0	0.76
PBB101	7.608	0.148	0	0.097	16.136	0	0	0	0.76
PBB102	5.710	0.142	0	0.015	12.961	0	0	0	0.76
PBB103	5.833	0.141	0	0.006	13.135	0	0	0	0.76
PBB104	4.201	0.197	0	0.218	11.05	0	0	0	0.76
PBB105	9.270	0.249	0	-0.165	18.155	0	0	0	0.76
PBB107	8.255	0.164	0	0.088	17.134	0	0	0	0.76
PBB108	8.753	0.256	0	-0.263	17.118	0	0	0	0.76
PBB109	7.477	0.207	0	-0.257	15.123	0	0	0	0.76
PBB110	7.402	0.199	0	-0.239	15.045	0	0	0	0.76
PBB111	7.766	0.158	0	-0.001	16.162	0	0	0	0.76
PBB112	6.500	0.140	0	-0.003	14.164	0	0	0	0.76
PBB113	7.092	0.271	0	-0.439	14.101	0	0	0	0.76
PBB114	8.765	0.179	0	0.167	18.118	0	0	0	0.76
PBB115	7.859	0.167	0	-0.056	16.182	0	0	0	0.76

Table A2. Continued.

Name	Pred. pIC_{50} from Eq. 4.2	HAT i/i (*h**=0.329)	RFD	MATS5s	Tm	nHAcc	B04[O-Cl]	F04[Cl-Cl]	LOC
PBB116	6.785	0.142	0	0.17	15.007	0	0	0	0.76
PBB117	6.886	0.144	0	0.188	15.208	0	0	0	0.76
PBB118	8.910	0.187	0	0.082	18.153	0	0	0	0.76
PBB119	8.448	0.351	0	-0.492	16.115	0	0	0	0.76
PBB120	8.424	0.179	0	-0.015	17.166	0	0	0	0.76
PBB121	8.145	0.477	0	-0.702	15.158	0	0	0	0.76
PBB122	8.628	0.247	0	-0.254	16.941	0	0	0	0.76
PBB123	9.358	0.290	0	-0.269	18.057	0	0	0	0.76
PBB124	8.296	0.173	0	-0.006	16.985	0	0	0	0.76
PBB125	7.892	0.451	0	-0.684	14.802	0	0	0	0.76
PBB126	10.054	0.238	0	0.161	20.134	0	0	0	0.76
PBB127	9.387	0.204	0	0.176	19.118	0	0	0	0.76
PBB128	8.592	0.251	0	-0.269	16.851	0	0	0	0.76
PBB129	7.788	0.181	0	-0.134	15.893	0	0	0	0.76
PBB130	7.866	0.183	0	-0.134	16.015	0	0	0	0.76
PBB131	6.744	0.163	0	-0.148	14.217	0	0	0	0.76
PBB132	6.710	0.160	0	-0.139	14.184	0	0	0	0.76
PBB133	7.134	0.145	0	0.009	15.19	0	0	0	0.76
PBB134	6.018	0.141	0	-0.013	13.383	0	0	0	0.76
PBB135	6.158	0.151	0	-0.107	13.389	0	0	0	0.76
PBB136	4.576	0.183	0	0.219	11.642	0	0	0	0.76
PBB137	8.261	0.176	0	-0.032	16.871	0	0	0	0.76
PBB138	8.275	0.175	0	-0.023	16.913	0	0	0	0.76
PBB139	7.040	0.141	0	0.066	15.172	0	0	0	0.76
PBB140	7.496	0.213	0	-0.274	15.114	0	0	0	0.76

Table A2. Continued.

Name	Pred. pIC_{50} from Eq. 4.2	HAT i/i (*h**=0.329)	RFD	MATS5s	Tm	nHAcc	B04[O-Cl]	F04[Cl-Cl]	LOC
PBB141	7.467	0.146	0	0.112	15.949	0	0	0	0.76
PBB142	6.197	0.145	0	0.183	14.112	0	0	0	0.76
PBB143	6.761	0.187	0	-0.242	14.03	0	0	0	0.76
PBB144	6.426	0.138	0	0.098	14.278	0	0	0	0.76
PBB145	5.110	0.162	0	0.187	12.41	0	0	0	0.76
PBB146	7.546	0.147	0	0.112	16.073	0	0	0	0.76
PBB147	6.319	0.145	0	0.192	14.325	0	0	0	0.76
PBB148	6.950	0.193	0	-0.251	14.307	0	0	0	0.76
PBB149	6.392	0.138	0	0.107	14.246	0	0	0	0.76
PBB150	5.187	0.160	0	0.187	12.531	0	0	0	0.76
PBB151	5.700	0.157	0	0.233	13.444	0	0	0	0.76
PBB152	4.565	0.182	0	0.21	11.604	0	0	0	0.76
PBB153	7.951	0.160	0	0.224	16.967	0	0	0	0.76
PBB154	7.179	0.149	0	-0.027	15.178	0	0	0	0.76
PBB155	5.795	0.147	0	0.164	13.436	0	0	0	0.76
PBB156	9.258	0.202	0	0.076	18.686	0	0	0	0.76
PBB157	9.721	0.310	0	-0.264	18.64	0	0	0	0.76
PBB158	8.654	0.254	0	-0.269	16.948	0	0	0	0.76
PBB159	8.846	0.197	0	-0.027	17.802	0	0	0	0.76
PBB160	7.743	0.163	0	-0.049	16.016	0	0	0	0.76
PBB161	8.419	0.344	0	-0.483	16.089	0	0	0	0.76
PBB162	8.812	0.194	0	-0.018	17.769	0	0	0	0.76
PBB163	7.754	0.161	0	-0.031	16.074	0	0	0	0.76
PBB164	8.272	0.327	0	-0.465	15.9	0	0	0	0.76

Table A2. Continued

Name	Pred. pIC_{50} from Eq. 4.2	HAT i/i (*h**=0.329)	RFD	MATS5s	Tm	nHAcc	B04[O-Cl]	F04[Cl-Cl]	LOC
PBB165	7.514	0.201	0	-0.237	15.227	0	0	0	0.76
PBB166	8.047	0.158	0	0.156	16.963	0	0	0	0.76
PBB167	9.391	0.220	0	-0.018	18.682	0	0	0	0.76
PBB168	9.238	0.549	0	-0.711	16.858	0	0	0	0.76
PBB169	10.228	0.247	0	0.183	20.457	0	0	0	0.76
PBB170	8.860	0.228	0	-0.162	17.504	0	0	0	0.754
PBB171	7.915	0.198	0	-0.188	15.956	0	0	0	0.754
PBB172	8.186	0.173	0	-0.031	16.741	0	0	0	0.754
PBB173	7.207	0.156	0	-0.075	15.1	0	0	0	0.754
PBB174	7.319	0.175	0	-0.161	15.08	0	0	0	0.754
PBB175	7.420	0.180	0	-0.17	15.218	0	0	0	0.754
PBB176	5.928	0.144	0	0.151	13.603	0	0	0	0.754
PBB177	7.247	0.155	0	-0.066	15.183	0	0	0	0.754
PBB178	6.746	0.145	0	-0.039	14.457	0	0	0	0.754
PBB179	5.436	0.153	0	0.168	12.866	0	0	0	0.754
PBB180	8.531	0.173	0	0.088	17.556	0	0	0	0.754
PBB181	7.434	0.147	0	0.149	15.968	0	0	0	0.754
PBB182	7.915	0.198	0	-0.188	15.956	0	0	0	0.754
PBB183	7.586	0.149	0	0.063	16.012	0	0	0	0.754
PBB184	6.478	0.140	0	0.133	14.427	0	0	0	0.754
PBB185	6.877	0.143	0	0.176	15.154	0	0	0	0.754
PBB186	5.849	0.144	0	0.142	13.458	0	0	0	0.754
PBB187	6.918	0.144	0	0.185	15.239	0	0	0	0.754
PBB189	9.701	0.240	0	-0.033	19.122	0	0	0	0.754

Table A2. Continued.

Name	Pred. pIC_{50} from Eq. 4.2	HAT i/i (*h**=0.329)	RFD	MATS5s	Tm	nHAcc	B04[O-Cl]	F04[Cl-Cl]	LOC
PBB190	8.773	0.201	0	-0.067	17.583	0	0	0	0.754
PBB191	9.386	0.415	0	-0.508	17.542	0	0	0	0.754
PBB192	8.601	0.260	0	-0.291	16.801	0	0	0	0.754
PBB193	8.540	0.251	0	-0.273	16.746	0	0	0	0.754
PBB194	9.086	0.215	0	-0.063	18.066	0	0	0	0.745
PBB195	8.247	0.193	0	-0.12	16.616	0	0	0	0.745
PBB196	8.397	0.226	0	-0.219	16.625	0	0	0	0.745
PBB197	7.064	0.142	0	0.094	15.242	0	0	0	0.745
PBB199	7.769	0.173	0	-0.098	15.913	0	0	0	0.745
PBB200	6.556	0.139	0	0.097	14.449	0	0	0	0.745
PBB201	6.622	0.139	0	0.097	14.553	0	0	0	0.745
PBB202	6.171	0.140	0	0.109	13.87	0	0	0	0.745
PBB203	7.901	0.156	0	0.141	16.667	0	0	0	0.745
PBB204	7.055	0.142	0	0.085	15.207	0	0	0	0.745
PBB205	9.471	0.322	0	-0.324	18.076	0	0	0	0.745
PBB206	8.689	0.223	0	-0.164	17.185	0	0	0	0.733
PBB207	7.588	0.153	0	0.032	15.899	0	0	0	0.733
PBB208	7.183	0.146	0	0.029	15.254	0	0	0	0.733
PBB209	8.085	0.176	0	-0.051	16.464	0	0	0	0.72
PBDD045	10.282	0.154	0.286	0.48	21.512	2	0	0	0.714
PBDE000	3.113	0.527	0	-0.148	7.832	1	0	0	0
PBDE001	3.701	0.263	0	-0.114	10.001	1	0	0	0.544
PBDE002	6.850	0.294	0	-0.253	14.64	1	0	0	0.544
PBDE003	6.680	0.195	0	0.142	15.274	1	0	0	0.544

Table A2. Continued.

Name	Pred. pIC_{50} from Eq. 4.2	HAT i/i (*h**=0.329)	RFD	MATS5s	Tm	nHAcc	B04[O-Cl]	F04[Cl-Cl]	LOC
PBDE004	3.135	0.294	0	-0.204	9.127	1	0	0	0.648
PBDE005	6.706	0.264	0	-0.235	14.678	1	0	0	0.648
PBDE006	3.753	0.279	0	-0.226	10.049	1	0	0	0.648
PBDE007	6.432	0.172	0	0.146	15.116	1	0	0	0.648
PBDE008	6.812	0.173	0	0.146	15.714	1	0	0	0.648
PBDE009	4.807	0.199	0	0.015	12.259	1	0	0	0.648
PBDE010	2.693	0.293	0	-0.109	8.647	1	0	0	0.648
PBDE011	6.040	0.313	0	-0.361	13.342	1	0	0	0.648
PBDE012	8.800	0.244	0	0.02	18.557	1	0	0	0.648
PBDE013	8.930	0.249	0	0.02	18.761	1	0	0	0.648
PBDE014	7.524	0.346	0	-0.361	15.678	1	0	0	0.648
PBDE015	7.329	0.197	0	0.02	16.24	1	0	0	0.648
PBDE016	5.709	0.289	0	-0.327	13.016	1	0	0	0.703
PBDE017	5.447	0.181	0	0.04	13.441	1	0	0	0.703
PBDE018	4.119	0.230	0	-0.086	11.063	1	0	0	0.703
PBDE019	3.646	0.311	0	-0.311	9.804	1	0	0	0.703
PBDE020	8.243	0.361	0	-0.351	16.951	1	0	0	0.703
PBDE021	8.279	0.215	0	0.015	17.842	1	0	0	0.703
PBDE022	7.945	0.205	0	0.015	17.317	1	0	0	0.703
PBDE024	4.823	0.197	0	-0.003	12.361	1	0	0	0.703
PBDE025	6.673	0.180	0	0.024	15.334	1	0	0	0.703
PBDE026	5.718	0.208	0	-0.11	13.525	1	0	0	0.703
PBDE027	5.367	0.247	0	-0.226	12.708	1	0	0	0.703
PBDE028	7.909	0.185	0	0.39	18.115	1	0	0	0.703

Table A2. Continued.

Name	Pred. pIC_{50} from Eq. 4.2	HAT i/i (*h**=0.329)	RFD	MATS5s	Tm	nHAcc	B04[O-Cl]	F04[Cl-Cl]	LOC
PBDE029	6.733	0.164	0	0.256	15.958	1	0	0	0.703
PBDE030	5.071	0.180	0	0.131	13.057	1	0	0	0.703
PBDE031	6.789	0.164	0	0.256	16.047	1	0	0	0.703
PBDE032	5.003	0.181	0	0.131	12.95	1	0	0	0.703
PBDE033	7.196	0.186	0	0.024	16.158	1	0	0	0.703
PBDE034	6.245	0.301	0	-0.343	13.823	1	0	0	0.703
PBDE035	10.274	0.355	0	-0.102	20.716	1	0	0	0.703
PBDE036	8.171	0.437	0	-0.477	16.549	1	0	0	0.703
PBDE037	10.068	0.257	0	0.273	21.247	1	0	0	0.703
PBDE038	8.881	0.274	0	-0.102	18.523	1	0	0	0.703
PBDE039	9.110	0.286	0	-0.102	18.883	1	0	0	0.703
PBDE040	4.760	0.361	0	-0.459	11.282	1	0	0	0.732
PBDE041	7.161	0.212	0	-0.092	15.901	1	0	0	0.732
PBDE042	5.405	0.205	0	-0.092	13.136	1	0	0	0.732
PBDE043	6.137	0.245	0	-0.222	13.992	1	0	0	0.732
PBDE044	7.032	0.256	0	-0.222	15.4	1	0	0	0.732
PBDE046	5.735	0.347	0	-0.439	12.863	1	0	0	0.732
PBDE047	6.074	0.165	0	0.274	15.025	1	0	0	0.732
PBDE048	5.760	0.167	0	0.144	14.233	1	0	0	0.732
PBDE049	5.937	0.165	0	0.144	14.512	1	0	0	0.732
PBDE050	4.632	0.215	0	-0.082	11.942	1	0	0	0.732
PBDE051	5.456	0.202	0	-0.082	13.239	1	0	0	0.732
PBDE052	5.791	0.180	0	0.014	13.985	1	0	0	0.732
PBDE053	4.503	0.250	0	-0.203	11.463	1	0	0	0.732

Table A2. Continued.

Name	Pred. pIC_{50} from Eq. 4.2	HAT i/i (*h**=0.329)	RFD	MATS5s	Tm	nHAcc	B04[O-Cl]	F04[Cl-Cl]	LOC
PBDE054	9.313	0.164	0.286	0.009	18.945	2	0	0	0.73
PBDE054	3.427	0.426	0	-0.528	9.027	1	0	0	0.732
PBDE055	7.971	0.238	0	-0.111	17.132	1	0	0	0.732
PBDE056	7.895	0.236	0	-0.111	17.012	1	0	0	0.732
PBDE057	7.816	0.290	0	-0.25	16.571	1	0	0	0.732
PBDE058	7.184	0.398	0	-0.478	15.056	1	0	0	0.732
PBDE059	6.737	0.218	0	-0.132	15.142	1	0	0	0.732
PBDE060	7.607	0.168	0	0.255	17.395	1	0	0	0.732
PBDE061	7.343	0.171	0	0.117	16.663	1	0	0	0.732
PBDE062	6.394	0.161	0	0.226	15.419	1	0	0	0.732
PBDE063	7.478	0.174	0	0.117	16.876	1	0	0	0.732
PBDE064	6.408	0.161	0	0.226	15.44	1	0	0	0.732
PBDE065	5.515	0.173	0	0.087	13.718	1	0	0	0.732
PBDE066	8.500	0.187	0	0.263	18.818	1	0	0	0.732
PBDE067	8.357	0.194	0	0.125	18.279	1	0	0	0.732
PBDE068	7.753	0.229	0	-0.103	16.807	1	0	0	0.732
PBDE069	6.995	0.185	0	0.006	15.862	1	0	0	0.732
PBDE070	8.296	0.192	0	0.125	18.182	1	0	0	0.732
PBDE071	6.804	0.182	0	0.006	15.562	1	0	0	0.732
PBDE072	7.555	0.277	0	-0.241	16.18	1	0	0	0.732
PBDE073	6.050	0.302	0	-0.351	13.56	1	0	0	0.732
PBDE074	8.500	0.187	0	0.263	18.819	1	0	0	0.732
PBDE075	6.592	0.170	0	0.364	16.045	1	0	0	0.732
PBDE076	7.881	0.233	0	-0.103	17.009	1	0	0	0.732

Table A2. Continued.

Name	Pred. pIC_{50} from Eq. 4.2	HAT i/i (*h**=0.329)	RFD	MATS5s	Tm	nHAcc	B04[O-Cl]	F04[Cl-Cl]	LOC
PBDE077	10.053	0.265	0	0.146	20.996	1	0	0	0.732
PBDE078	9.233	0.343	0	-0.229	18.85	1	0	0	0.732
PBDE079	9.378	0.351	0	-0.229	19.078	1	0	0	0.732
PBDE080	9.680	0.623	0	-0.604	18.697	1	0	0	0.732
PBDE081	10.428	0.288	0	0.146	21.586	1	0	0	0.732
PBDE082	6.215	0.250	0	-0.234	14.114	1	0	0	0.745
PBDE083	6.172	0.313	0	-0.37	13.737	1	0	0	0.745
PBDE084	6.999	0.319	0	-0.356	15.07	1	0	0	0.745
PBDE085	6.719	0.163	0	0.139	15.759	1	0	0	0.745
PBDE086	7.213	0.187	0	0.003	16.226	1	0	0	0.745
PBDE087	8.095	0.208	0	0.003	17.615	1	0	0	0.745
PBDE088	5.860	0.180	0	0.008	14.108	1	0	0	0.745
PBDE089	6.956	0.253	0	-0.22	15.314	1	0	0	0.745
PBDE090	6.671	0.181	0	0.003	15.373	1	0	0	0.745
PBDE091	6.060	0.179	0	0.008	14.423	1	0	0	0.745
PBDE092	6.583	0.216	0	-0.134	14.923	1	0	0	0.745
PBDE093	5.151	0.217	0	-0.128	12.682	1	0	0	0.745
PBDE094	6.142	0.301	0	-0.347	13.742	1	0	0	0.745
PBDE095	5.874	0.209	0	-0.12	13.838	1	0	0	0.745
PBDE096	4.986	0.348	0	-0.44	11.709	1	0	0	0.745
PBDE097	6.603	0.181	0	0.003	15.267	1	0	0	0.745
PBDE098	7.086	0.256	0	-0.22	15.518	1	0	0	0.745
PBDE099	7.119	0.170	0	0.375	16.928	1	0	0	0.745
PBDE100	6.101	0.163	0	0.144	14.798	1	0	0	0.745

Table A2. Continued.

Name	Pred. pIC_{50} from Eq. 4.2	HAT i/i (*h**=0.329)	RFD	MATS5s	Tm	nHAcc	B04[O-Cl]	F04[Cl-Cl]	LOC
PBDE101	7.025	0.161	0	0.239	16.47	1	0	0	0.745
PBDE102	5.924	0.178	0	0.017	14.230	1	0	0	0.745
PBDE103	6.021	0.177	0	0.017	14.382	1	0	0	0.745
PBDE104	5.118	0.284	0	-0.313	12.208	1	0	0	0.745
PBDE105	7.871	0.179	0	0.125	17.541	1	0	0	0.745
PBDE106	10.149	0.400	0	-0.229	20.292	1	0	0	0.732
PBDE107	8.782	0.238	0	-0.02	18.644	1	0	0	0.745
PBDE108	8.307	0.306	0	-0.248	17.376	1	0	0	0.745
PBDE109	7.353	0.174	0	0.092	16.650	1	0	0	0.745
PBDE110	7.792	0.182	0	0.092	17.342	1	0	0	0.745
PBDE111	8.218	0.381	0	-0.393	16.905	1	0	0	0.745
PBDE112	7.184	0.200	0	-0.052	16.055	1	0	0	0.745
PBDE113	6.722	0.271	0	-0.271	14.829	1	0	0	0.745
PBDE114	7.870	0.179	0	0.125	17.539	1	0	0	0.745
PBDE115	6.974	0.184	0	0.456	16.884	1	0	0	0.745
PBDE116	6.649	0.163	0	0.312	16.044	1	0	0	0.745
PBDE117	6.763	0.163	0	0.312	16.223	1	0	0	0.745
PBDE118	8.181	0.181	0	0.361	18.567	1	0	0	0.745
PBDE119	7.974	0.173	0	0.237	17.959	1	0	0	0.745
PBDE120	8.597	0.228	0	-0.011	18.373	1	0	0	0.745
PBDE121	7.360	0.226	0	-0.127	16.161	1	0	0	0.745
PBDE122	8.189	0.301	0	-0.248	17.19	1	0	0	0.745
PBDE123	8.620	0.200	0	0.133	18.739	1	0	0	0.745
PBDE124	8.482	0.224	0	-0.011	18.192	1	0	0	0.745

Table A2. Continued.

Name	Pred. pIC_{50} from Eq. 4.2	HAT i/i (*h**=0.329)	RFD	MATS5s	Tm	nHAcc	B04[O-Cl]	F04[Cl-Cl]	LOC
PBDE125	7.156	0.222	0	-0.127	15.84	1	0	0	0.745
PBDE126	10.898	0.352	0	0.012	22.048	1	0	0	0.745
PBDE127	10.393	0.499	0	-0.369	20.384	1	0	0	0.745
PBDE128	7.286	0.191	0	-0.007	16.327	1	0	0	0.749
PBDE129	7.190	0.231	0	-0.152	15.846	1	0	0	0.749
PBDE130	7.285	0.233	0	-0.152	15.995	1	0	0	0.749
PBDE131	6.589	0.219	0	-0.143	14.919	1	0	0	0.749
PBDE132	6.629	0.220	0	-0.143	14.982	1	0	0	0.749
PBDE133	7.270	0.295	0	-0.297	15.64	1	0	0	0.749
PBDE134	5.993	0.271	0	-0.288	13.65	1	0	0	0.749
PBDE135	6.613	0.273	0	-0.279	14.647	1	0	0	0.749
PBDE136	6.206	0.313	0	-0.37	13.798	1	0	0	0.749
PBDE137	7.571	0.166	0	0.233	17.324	1	0	0	0.749
PBDE138	7.607	0.166	0	0.233	17.38	1	0	0	0.749
PBDE139	7.462	0.164	0	0.233	17.152	1	0	0	0.749
PBDE140	6.716	0.182	0	0.002	15.451	1	0	0	0.749
PBDE141	7.528	0.177	0	0.088	16.926	1	0	0	0.749
PBDE142	6.230	0.167	0	0.088	14.882	1	0	0	0.749
PBDE143	6.570	0.216	0	-0.134	14.911	1	0	0	0.749
PBDE144	6.948	0.168	0	0.097	16.032	1	0	0	0.749
PBDE145	6.200	0.249	0	-0.234	14.099	1	0	0	0.749
PBDE146	7.586	0.178	0	0.088	17.016	1	0	0	0.749
PBDE147	6.383	0.167	0	0.088	15.122	1	0	0	0.749
PBDE148	6.708	0.218	0	-0.134	15.128	1	0	0	0.749

Table A2. Continued.

Name	Pred. pIC_{50} from Eq. 4.2	HAT i/i (*h**=0.329)	RFD	MATS5s	Tm	nHAcc	B04[O-Cl]	F04[Cl-Cl]	LOC
PBDE149	6.949	0.168	0	0.097	16.034	1	0	0	0.749
PBDE150	5.687	0.248	0	-0.234	13.292	1	0	0	0.749
PBDE151	6.344	0.190	0	-0.048	14.751	1	0	0	0.749
PBDE152	5.586	0.309	0	-0.37	12.822	1	0	0	0.749
PBDE153	7.916	0.192	0	0.473	18.415	1	0	0	0.749
PBDE154	5.924	0.178	0	0.017	14.23	1	0	0	0.745
PBDE155	6.675	0.186	0	-0.018	15.334	1	0	0	0.746
PBDE156	8.030	0.190	0	0.077	17.684	1	0	0	0.746
PBDE157	9.456	0.223	0	0.215	20.251	1	0	0	0.749
PBDE158	8.624	0.189	0	0.324	19.189	1	0	0	0.749
PBDE159	9.031	0.248	0	-0.016	19.054	1	0	0	0.749
PBDE160	8.045	0.178	0	0.171	17.929	1	0	0	0.749
PBDE161	8.159	0.224	0	-0.052	17.599	1	0	0	0.749
PBDE162	7.074	0.162	0	0.171	16.4	1	0	0	0.749
PBDE163	8.044	0.220	0	-0.052	17.418	1	0	0	0.749
PBDE164	9.011	0.302	0	-0.169	18.673	1	0	0	0.749
PBDE165	7.558	0.261	0	-0.206	16.302	1	0	0	0.749
PBDE166	7.536	0.206	0	0.547	17.986	1	0	0	0.749
PBDE167	9.246	0.213	0	0.224	19.94	1	0	0	0.749
PBDE168	8.135	0.189	0	0.102	17.913	1	0	0	0.749
PBDE169	10.935	0.410	0	-0.133	21.784	1	0	0	0.749
PBDE170	8.034	0.190	0	0.077	17.69	1	0	0	0.746
PBDE171	7.489	0.178	0	0.081	16.842	1	0	0	0.746
PBDE172	8.060	0.230	0	-0.08	17.373	1	0	0	0.746

Table A2. Continued.

Name	Pred. pIC_{50} from Eq. 4.2	HAT i/i (*h**=0.329)	RFD	MATS5s	Tm	nHAcc	B04[O-Cl]	F04[Cl-Cl]	LOC
PBDE173	6.881	0.202	0	-0.076	15.526	1	0	0	0.746
PBDE174	7.473	0.210	0	-0.067	16.479	1	0	0	0.746
PBDE175	6.977	0.203	0	-0.076	15.677	1	0	0	0.746
PBDE176	7.158	0.236	0	-0.166	15.757	1	0	0	0.746
PBDE177	7.512	0.211	0	-0.067	16.540	1	0	0	0.746
PBDE178	7.003	0.256	0	-0.224	15.381	1	0	0	0.746
PBDE179	6.628	0.290	0	-0.314	14.585	1	0	0	0.746
PBDE180	8.274	0.181	0	0.325	18.634	1	0	0	0.746
PBDE181	7.166	0.165	0	0.32	16.879	1	0	0	0.746
PBDE182	7.516	0.177	0	0.09	16.905	1	0	0	0.746
PBDE183	7.747	0.171	0	0.329	17.814	1	0	0	0.746
PBDE184	7.195	0.192	0	-0.018	16.153	1	0	0	0.746
PBDE185	7.172	0.163	0	0.172	16.55	1	0	0	0.746
PBDE186	6.534	0.227	0	-0.166	14.774	1	0	0	0.746
PBDE187	7.227	0.164	0	0.172	16.637	1	0	0	0.746
PBDE188	6.664	0.228	0	-0.166	14.979	1	0	0	0.746
PBDE189	9.561	0.253	0	0.065	20.067	1	0	0	0.746
PBDE190	7.744	0.180	0	0.41	17.994	1	0	0	0.746
PBDE191	8.757	0.199	0	0.181	19.066	1	0	0	0.746
PBDE192	8.269	0.210	0	0.015	17.919	1	0	0	0.746
PBDE193	8.235	0.209	0	0.015	17.866	1	0	0	0.746
PBDE194	8.644	0.198	0	0.158	18.823	1	0	0	0.74
PBDE195	7.632	0.172	0	0.157	17.227	1	0	0	0.74
PBDE196	8.145	0.182	0	0.166	18.056	1	0	0	0.74

Table A2. Continued.

Name	Pred. pIC_{50} from Eq. 4.2	HAT i/i (*h**=0.329)	RFD	MATS5s	Tm	nHAcc	B04[O-Cl]	F04[Cl-Cl]	LOC
PBDE197	7.879	0.191	0	0.056	17.386	1	0	0	0.74
PBDE198	7.699	0.199	0	-0.006	16.960	1	0	0	0.74
PBDE199	7.709	0.200	0	-0.006	16.977	1	0	0	0.74
PBDE200	7.381	0.221	0	-0.107	16.229	1	0	0	0.74
PBDE201	7.424	0.221	0	-0.107	16.297	1	0	0	0.74
PBDE202	6.977	0.276	0	-0.27	15.221	1	0	0	0.74
PBDE203	7.829	0.182	0	0.417	18.131	1	0	0	0.74
PBDE204	7.370	0.180	0	0.056	16.584	1	0	0	0.74
PBDE205	8.755	0.194	0	0.258	19.226	1	0	0	0.74
PBDE206	8.217	0.180	0	0.243	18.323	1	0	0	0.73
PBDE207	7.980	0.183	0	0.129	17.689	1	0	0	0.73
PBDE208	7.602	0.211	0	-0.055	16.675	1	0	0	0.73
PBDE209	8.045	0.179	0	0.204	17.939	1	0	0	0.719
PCB000	2.985	0.468	0	-0.019	6.967	0	0	0	0
PCB001	2.502	0.259	0	-0.168	7.085	0	0	0	0.569
PCB002	3.622	0.208	0	0.061	9.371	0	0	0	0.569
PCB003	4.423	0.180	0	0.026	10.551	0	0	0	0.569
PCB004	2.078	0.283	0	0.062	7.164	0	0	0	0.673
PCB005	3.608	0.201	0	-0.137	9.119	0	0	0	0.673
PCB006	3.736	0.209	0	-0.22	9.132	0	0	0	0.673
PCB007	3.726	0.136	0	-0.132	10.159	0	0	1	0.673
PCB008	4.240	0.180	0	-0.128	10.135	0	0	0	0.673
PCB009	3.067	0.260	0	0.25	9.15	0	0	0	0.673
PCB010	2.097	0.220	0	-0.325	7.154	0	0	1	0.673

Table A2. Continued.

Name	Pred. pIC_{50} from Eq. 4.2	HAT i/i (*h**=0.329)	RFD	MATS5s	Tm	nHAcc	B04[O-Cl]	F04[Cl-Cl]	LOC
PCB011	4.681	0.163	0	0.056	11.249	0	0	0	0.673
PCB012	5.287	0.151	0	0.056	12.203	0	0	0	0.673
PCB013	5.346	0.150	0	0.056	12.296	0	0	0	0.673
PCB014	4.147	0.132	0	0.051	11.239	0	0	1	0.673
PCB015	6.000	0.144	0	0.06	13.335	0	0	0	0.673
PCB016	3.414	0.205	0	-0.09	9.032	0	0	0	0.725
PCB017	3.306	0.155	0	0.004	9.92	0	0	1	0.725
PCB018	2.984	0.255	0	0.213	9.047	0	0	0	0.725
PCB019	1.489	0.280	0	0.105	7.289	0	0	1	0.725
PCB020	4.691	0.183	0	-0.216	10.756	0	0	0	0.725
PCB021	4.571	0.113	0	-0.129	11.608	0	0	1	0.725
PCB022	5.144	0.161	0	-0.126	11.673	0	0	0	0.725
PCB023	3.656	0.164	0	0.148	10.799	0	0	1	0.725
PCB024	2.805	0.175	0	-0.025	9.065	0	0	1	0.725
PCB025	4.718	0.116	0	-0.194	11.691	0	0	1	0.725
PCB026	4.298	0.174	0	0.087	10.828	0	0	0	0.725
PCB027	3.406	0.202	0	-0.457	9.025	0	0	1	0.725
PCB028	5.159	0.103	0	-0.1	12.6	0	0	1	0.725
PCB029	4.166	0.151	0	0.174	11.661	0	0	1	0.725
PCB030	2.698	0.190	0	-0.302	9.949	0	0	3	0.725
PCB031	4.730	0.171	0	0.177	11.714	0	0	0	0.725
PCB032	3.724	0.152	0	-0.295	9.895	0	0	1	0.725
PCB033	5.177	0.171	0	-0.19	11.58	0	0	0	0.725
PCB034	4.259	0.138	0	-0.284	10.762	0	0	1	0.725

Table A2. Continued.

Name	Pred. pIC_{50} from Eq. 4.2	HAT i/i (*h**=0.329)	RFD	MATS5s	Tm	nHAcc	B04[O-Cl]	F04[Cl-Cl]	LOC
PCB035	5.994	0.140	0	0.047	13.407	0	0	0	0.725
PCB036	4.960	0.107	0	0.025	12.571	0	0	1	0.725
PCB037	6.553	0.139	0	0.07	14.34	0	0	0	0.725
PCB038	5.405	0.103	0	0.044	13.316	0	0	1	0.725
PCB039	5.508	0.102	0	0.044	13.478	0	0	1	0.725
PCB040	4.462	0.188	0	-0.214	10.453	0	0	0	0.75
PCB041	4.241	0.119	0	-0.119	11.165	0	0	1	0.75
PCB042	4.285	0.118	0	-0.119	11.234	0	0	1	0.75
PCB043	3.546	0.149	0	0.039	10.431	0	0	1	0.75
PCB044	4.084	0.179	0	0.041	10.439	0	0	0	0.75
PCB045	2.365	0.241	0	0.181	8.896	0	0	1	0.75
PCB046	2.790	0.172	0	-0.129	8.858	0	0	1	0.75
PCB048	3.881	0.153	0	0.136	11.18	0	0	1	0.75
PCB049	3.927	0.151	0	0.136	11.252	0	0	1	0.75
PCB050	2.018	0.268	0	0.023	9.674	0	0	3	0.75
PCB051	2.535	0.193	0	0.026	9.653	0	0	2	0.75
PCB052	3.740	0.237	0	0.297	10.483	0	0	0	0.75
PCB053	2.448	0.221	0	0.126	8.901	0	0	1	0.75
PCB054	0.651	0.443	0	0.324	7.367	0	0	2	0.75
PCB055	5.365	0.112	0	-0.207	12.733	0	0	1	0.75
PCB056	5.881	0.170	0	-0.204	12.71	0	0	0	0.75
PCB057	4.588	0.113	0	0.011	12.007	0	0	1	0.75
PCB058	5.015	0.131	0	-0.299	11.972	0	0	1	0.75
PCB059	3.929	0.135	0	-0.214	10.457	0	0	1	0.75

Table A2. Continued.

Name	Pred. pIC_{50} from Eq. 4.2	HAT i/i (*h**=0.329)	RFD	MATS5s	Tm	nHAcc	B04[O-Cl]	F04[Cl-Cl]	LOC
PCB060	5.746	0.100	0	-0.112	13.55	0	0	1	0.75
PCB062	3.128	0.194	0	-0.068	11.214	0	0	3	0.75
PCB063	4.968	0.115	0	0.103	12.816	0	0	1	0.75
PCB064	4.205	0.120	0	-0.062	11.238	0	0	1	0.75
PCB065	2.877	0.208	0	0.148	10.469	0	0	2	0.75
PCB066	5.830	0.105	0	-0.167	13.556	0	0	1	0.75
PCB067	3.881	0.153	0	0.136	11.18	0	0	1	0.75
PCB068	3.927	0.151	0	0.136	11.252	0	0	1	0.75
PCB069	3.682	0.177	0	-0.433	11.253	0	0	3	0.75
PCB070	5.550	0.143	0	0.052	12.773	0	0	0	0.75
PCB071	4.668	0.170	0	-0.427	11.134	0	0	1	0.75
PCB072	4.674	0.109	0	-0.043	12.019	0	0	1	0.75
PCB073	3.893	0.200	0	-0.583	10.401	0	0	2	0.75
PCB074	5.408	0.114	0	0.144	13.602	0	0	1	0.75
PCB075	3.958	0.157	0	-0.277	12.043	0	0	3	0.75
PCB078	6.025	0.097	0	0.018	14.285	0	0	1	0.75
PCB079	6.070	0.097	0	0.018	14.357	0	0	1	0.75
PCB080	5.097	0.106	0	-0.016	13.589	0	0	2	0.75
PCB082	5.079	0.118	0	-0.234	12.243	0	0	1	0.76
PCB083	4.463	0.114	0	-0.108	11.56	0	0	1	0.76
PCB084	3.487	0.144	0	-0.052	10.152	0	0	1	0.76
PCB085	4.857	0.102	0	-0.134	12.964	0	0	2	0.76
PCB086	4.167	0.123	0	-0.01	12.161	0	0	2	0.76
PCB087	4.744	0.109	0	-0.007	12.234	0	0	1	0.76

Table A2. Continued.

Name	Pred. pIC_{50} from Eq. 4.2	HAT i/i (*h**=0.329)	RFD	MATS5s	Tm	nHAcc	B04[O-Cl]	F04[Cl-Cl]	LOC
PCB088	2.631	0.259	0	0.097	10.829	0	0	3	0.76
PCB089	3.528	0.130	0	-0.176	10.776	0	0	2	0.76
PCB090	4.249	0.121	0	-0.01	12.29	0	0	2	0.76
PCB091	3.191	0.179	0	0.1	10.875	0	0	2	0.76
PCB092	4.144	0.140	0	0.119	11.577	0	0	1	0.76
PCB093	2.564	0.251	0	0.221	10.165	0	0	2	0.76
PCB094	3.017	0.151	0	-0.104	10.135	0	0	2	0.76
PCB095	3.182	0.193	0	0.175	10.189	0	0	1	0.76
PCB096	1.607	0.341	0	0.279	8.79	0	0	2	0.76
PCB097	4.744	0.109	0	-0.007	12.234	0	0	1	0.76
PCB098	3.050	0.182	0	-0.179	10.859	0	0	3	0.76
PCB099	4.539	0.132	0	0.093	12.982	0	0	2	0.76
PCB100	2.755	0.316	0	-0.024	11.592	0	0	4	0.76
PCB101	4.441	0.153	0	0.22	12.275	0	0	1	0.76
PCB102	3.224	0.166	0	0.051	10.816	0	0	2	0.76
PCB103	2.746	0.238	0	0.048	10.899	0	0	3	0.76
PCB104	1.082	0.512	0	0.207	9.485	0	0	4	0.76
PCB106	5.112	0.105	0	-0.017	13.632	0	0	2	0.76
PCB107	5.659	0.097	0	-0.015	13.656	0	0	1	0.76
PCB108	5.519	0.113	0	-0.293	13.644	0	0	2	0.76
PCB109	4.029	0.155	0	-0.237	12.268	0	0	3	0.76
PCB110	4.837	0.103	0	-0.113	12.98	0	0	2	0.76
PCB111	3.875	0.125	0	-0.06	11.587	0	0	2	0.76
PCB112	4.326	0.135	0	-0.384	11.558	0	0	2	0.76

Table A2. Continued.

Name	Pred. pIC_{50} from Eq. 4.2	HAT i/i (*h**=0.329)	RFD	MATS5s	Tm	nHAcc	B04[O-Cl]	F04[Cl-Cl]	LOC
PCB113	5.421	0.116	0	0.081	14.343	0	0	2	0.76
PCB115	4.265	0.161	0	-0.085	12.987	0	0	3	0.76
PCB116	2.990	0.345	0	0.085	12.21	0	0	4	0.76
PCB117	4.109	0.143	0	0.09	12.298	0	0	2	0.76
PCB119	4.717	0.164	0	-0.409	12.959	0	0	3	0.76
PCB120	5.220	0.101	0	-0.066	13.69	0	0	2	0.76
PCB121	3.977	0.269	0	-0.564	12.283	0	0	4	0.76
PCB122	5.996	0.129	0	-0.29	13.558	0	0	1	0.76
PCB123	5.885	0.109	0	-0.244	14.331	0	0	2	0.76
PCB124	5.697	0.097	0	-0.063	13.605	0	0	1	0.76
PCB125	4.931	0.186	0	-0.559	12.111	0	0	2	0.76
PCB127	6.019	0.102	0	-0.022	15.049	0	0	2	0.76
PCB129	4.927	0.101	0	-0.145	13.049	0	0	2	0.76
PCB130	4.423	0.113	0	-0.042	12.491	0	0	2	0.76
PCB131	3.557	0.173	0	-0.112	11.811	0	0	3	0.76
PCB132	4.078	0.116	0	-0.109	11.795	0	0	2	0.76
PCB133	4.423	0.113	0	-0.042	12.491	0	0	2	0.76
PCB134	3.557	0.141	0	-0.011	11.198	0	0	2	0.76
PCB135	3.623	0.133	0	-0.055	11.202	0	0	2	0.76
PCB136	2.420	0.264	0	0.232	9.962	0	0	2	0.76
PCB137	4.670	0.161	0	-0.04	13.727	0	0	3	0.76
PCB138	5.208	0.102	0	-0.038	13.736	0	0	2	0.76
PCB139	3.226	0.321	0	0.043	12.485	0	0	4	0.76
PCB140	3.572	0.253	0	-0.21	12.453	0	0	4	0.76

Table A2. Continued.

Name	Pred. pIC_{50} from Eq. 4.2	HAT i/i (*h**=0.329)	RFD	MATS5s	Tm	nHAcc	B04[O-Cl]	F04[Cl-Cl]	LOC
PCB141	4.635	0.124	0	0.063	13.064	0	0	2	0.76
PCB142	4.167	0.123	0	-0.01	12.161	0	0	2	0.76
PCB143	3.564	0.168	0	-0.156	11.721	0	0	3	0.76
PCB144	3.277	0.231	0	0.097	11.847	0	0	3	0.76
PCB145	1.797	0.448	0	0.178	10.544	0	0	4	0.76
PCB146	4.662	0.123	0	0.063	13.107	0	0	2	0.76
PCB147	3.223	0.249	0	0.142	11.864	0	0	3	0.76
PCB148	3.118	0.271	0	-0.159	11.855	0	0	4	0.76
PCB149	3.799	0.155	0	0.099	11.83	0	0	2	0.76
PCB150	1.834	0.445	0	0.178	10.602	0	0	4	0.76
PCB151	3.276	0.206	0	0.198	11.232	0	0	2	0.76
PCB152	1.877	0.358	0	0.229	9.943	0	0	3	0.76
PCB154	3.295	0.304	0	-0.002	12.492	0	0	4	0.76
PCB155	1.250	0.819	0	0.126	11.25	0	0	6	0.76
PCB157	6.345	0.121	0	-0.282	14.969	0	0	2	0.76
PCB158	4.981	0.146	0	-0.246	13.746	0	0	3	0.76
PCB159	5.208	0.147	0	-0.134	14.359	0	0	3	0.76
PCB160	3.820	0.264	0	-0.103	13.088	0	0	4	0.76
PCB161	4.273	0.242	0	-0.401	13.121	0	0	4	0.76
PCB162	5.728	0.099	0	-0.132	14.34	0	0	2	0.76
PCB163	4.898	0.102	0	-0.098	13.11	0	0	2	0.76
PCB164	5.276	0.134	0	-0.396	13.026	0	0	2	0.76
PCB165	4.192	0.152	0	-0.251	12.493	0	0	3	0.76
PCB166	4.018	0.301	0	0.05	13.749	0	0	4	0.76

Table A2. Continued.

Name	Pred. pIC_{50} from Eq. 4.2	HAT i/i (*h**=0.329)	RFD	MATS5s	Tm	nHAcc	B04[O-Cl]	F04[Cl-Cl]	LOC
PCB169	6.795	0.111	0	-0.018	16.28	0	0	2	0.76
PCB170	5.280	0.146	0	-0.171	14.376	0	0	3	0.754
PCB171	3.981	0.253	0	-0.154	13.212	0	0	4	0.754
PCB172	4.791	0.153	0	-0.088	13.795	0	0	3	0.754
PCB173	3.475	0.279	0	-0.075	12.596	0	0	4	0.754
PCB174	4.062	0.161	0	-0.115	12.586	0	0	3	0.754
PCB175	3.573	0.267	0	-0.117	12.654	0	0	4	0.754
PCB176	2.461	0.392	0	0.14	11.49	0	0	4	0.754
PCB177	4.033	0.167	0	-0.073	12.636	0	0	3	0.754
PCB178	3.625	0.184	0	-0.034	12.082	0	0	3	0.754
PCB179	2.588	0.292	0	0.176	10.929	0	0	3	0.754
PCB180	5.005	0.170	0	0.026	14.392	0	0	3	0.754
PCB181	3.106	0.500	0	0.084	13.22	0	0	5	0.754
PCB182	3.502	0.397	0	-0.199	13.198	0	0	5	0.754
PCB183	3.717	0.306	0	0.043	13.246	0	0	4	0.754
PCB184	1.825	0.773	0	0.103	12.09	0	0	6	0.754
PCB185	3.209	0.352	0	0.122	12.627	0	0	4	0.754
PCB186	1.883	0.580	0	0.137	11.416	0	0	5	0.754
PCB187	1.952	0.577	0	0.137	11.524	0	0	5	0.754
PCB188	1.952	0.577	0	0.137	11.524	0	0	5	0.754
PCB190	4.696	0.249	0	-0.129	14.394	0	0	4	0.754
PCB191	5.090	0.246	0	-0.411	14.371	0	0	4	0.754
PCB192	4.021	0.378	0	-0.288	13.812	0	0	5	0.754
PCB193	5.065	0.148	0	-0.283	13.781	0	0	3	0.754

Table A2. Continued.

Name	Pred. pIC_{50} from Eq. 4.2	HAT i/i (*h**=0.329)	RFD	MATS5s	Tm	nHAcc	B04[O-Cl]	F04[Cl-Cl]	LOC
PCB194	5.039	0.250	0	-0.122	14.931	0	0	4	0.745
PCB195	3.817	0.409	0	-0.124	13.846	0	0	5	0.745
PCB196	3.881	0.398	0	-0.165	13.852	0	0	5	0.745
PCB197	2.349	0.730	0	0.067	12.813	0	0	6	0.745
PCB198	3.452	0.422	0	-0.101	13.323	0	0	5	0.745
PCB199	3.974	0.263	0	-0.098	13.31	0	0	4	0.745
PCB200	2.483	0.527	0	0.089	12.232	0	0	5	0.745
PCB201	2.515	0.525	0	0.089	12.282	0	0	5	0.745
PCB202	2.682	0.370	0	0.113	11.757	0	0	4	0.745
PCB203	3.561	0.480	0	0.067	13.878	0	0	5	0.745
PCB204	1.805	1.011	0	0.065	12.795	0	0	7	0.745
PCB205	4.783	0.376	0	-0.314	14.932	0	0	5	0.745
PCB206	3.699	0.602	0	-0.155	14.406	0	0	6	0.733
PCB207	2.293	0.963	0	0.016	13.426	0	0	7	0.733
PCB208	2.496	0.703	0	0.026	12.925	0	0	6	0.733
PCDD000	3.197	0.374	0.286	-0.296	7.058	2	0	0	0
PCDD002	5.459	0.126	0.286	0.056	10.322	2	1	0	0.521
PCDD003	5.889	0.125	0.286	-0.059	10.959	2	1	0	0.625
PCDD004	5.376	0.085	0.286	-0.059	10.993	2	1	1	0.625
PCDD005	4.723	0.145	0.286	0.028	9.321	2	1	0	0.625
PCDD006	5.266	0.236	0.286	-0.347	9.32	2	1	0	0.625
PCDD007	5.931	0.124	0.286	-0.059	11.025	2	1	0	0.625
PCDD008	5.899	0.124	0.286	-0.055	10.984	2	1	0	0.625
PCDD009	5.340	0.278	0.286	-0.433	9.24	2	1	0	0.625

Table A2. Continued.

Name	Pred. pIC_{50} from Eq. 4.2	HAT i/i (*h**=0.329)	RFD	MATS5s	Tm	nHAcc	B04[O-Cl]	F04[Cl-Cl]	LOC
PCDD010	6.541	0.085	0.286	0.233	12.651	2	1	0	0.625
PCDD011	6.617	0.084	0.286	0.233	12.771	2	1	0	0.625
PCDD013	6.329	0.053	0.286	0.081	12.936	2	1	1	0.682
PCDD015	6.349	0.150	0.286	-0.176	11.538	2	1	0	0.682
PCDD016	6.950	0.084	0.286	0.081	13.07	2	1	0	0.682
PCDD017	6.910	0.084	0.286	0.085	13.016	2	1	0	0.682
PCDD018	6.382	0.176	0.286	-0.246	11.43	2	1	0	0.682
PCDD023	6.404	0.051	0.286	0.085	13.063	2	1	1	0.682
PCDD024	5.893	0.118	0.286	-0.249	11.496	2	1	1	0.682
PCDD025	5.401	0.169	0.286	-0.183	10.029	2	1	0	0.682
PCDD027	6.200	0.070	0.286	-0.059	13.324	2	1	2	0.714
PCDD028	6.296	0.079	0.286	-0.131	13.311	2	1	2	0.714
PCDD029	5.895	0.076	0.286	-0.079	11.955	2	1	1	0.714
PCDD030	5.603	0.140	0.286	-0.095	10.617	2	1	0	0.714
PCDD031	5.888	0.077	0.286	-0.082	11.937	2	1	1	0.714
PCDD032	5.862	0.077	0.286	-0.079	11.903	2	1	1	0.714
PCDD033	6.329	0.056	0.286	0.225	13.333	2	1	1	0.714
PCDD034	6.318	0.056	0.286	0.225	13.316	2	1	1	0.714
PCDD035	7.244	0.043	0.286	0.18	14.67	2	1	1	0.714
PCDD036	6.733	0.064	0.286	-0.062	13.314	2	1	1	0.714
PCDD037	6.827	0.077	0.286	0.229	13.283	2	1	0	0.714
PCDD038	6.798	0.078	0.286	-0.127	13.268	2	1	1	0.714
PCDD039	6.413	0.121	0.286	-0.079	11.929	2	1	0	0.714
PCDD041	6.684	0.065	0.286	-0.062	13.237	2	1	1	0.714

Table A2. Continued.

Name	Pred. pIC50 from Eq. 4.2	HAT i/i (*h**=0.329)	RFD	MATS5s	Tm	nHAcc	B04[O-Cl]	F04[Cl-Cl]	LOC
PCDD042	7.237	0.043	0.286	0.176	14.65	2	1	1	0.714
PCDD043	7.211	0.043	0.286	0.18	14.618	2	1	1	0.714
PCDD044	6.739	0.078	0.286	-0.127	13.174	2	1	1	0.714
PCDD045	8.305	0.082	0.286	0.421	16.048	2	1	0	0.714
PCDD046	7.762	0.071	0.286	0.18	14.644	2	1	0	0.714
PCDD047	7.288	0.112	0.286	-0.065	13.338	2	1	0	0.714
PCDD049	6.239	0.066	0.286	-0.011	13.529	2	1	2	0.73
PCDD050	6.225	0.066	0.286	-0.008	13.514	2	1	2	0.73
PCDD051	6.013	0.068	0.286	-0.033	12.28	2	1	1	0.73
PCDD052	6.959	0.061	0.286	0.018	14.729	2	1	2	0.73
PCDD055	6.769	0.056	0.286	-0.011	13.521	2	1	1	0.73
PCDD056	6.721	0.056	0.286	-0.008	13.452	2	1	1	0.73
PCDD057	6.736	0.056	0.286	-0.008	13.476	2	1	1	0.73
PCDD058	6.164	0.067	0.286	-0.011	13.411	2	1	2	0.73
PCDD060	7.504	0.053	0.286	0.014	14.735	2	1	1	0.73
PCDD061	7.948	0.047	0.286	0.249	15.971	2	1	1	0.73
PCDD062	7.532	0.062	0.286	-0.043	14.649	2	1	1	0.73
PCDD063	6.308	0.065	0.286	0.008	13.697	2	1	2	0.737
PCDD064	6.300	0.065	0.286	0.008	13.684	2	1	2	0.737
PCDD065	6.418	0.124	0.286	0.038	14.78	2	1	3	0.737
PCDD067	6.945	0.060	0.286	0.041	14.775	2	1	2	0.737
PCDD068	6.258	0.065	0.286	0.008	13.618	2	1	2	0.737
PCDD069	7.268	0.089	0.286	0.327	15.935	2	1	2	0.737
PCDD070	7.635	0.063	0.286	0.072	15.932	2	1	2	0.737

Table A2. Continued.

Name	Pred. pIC_{50} from Eq. 4.2	HAT i/i (*h**=0.329)	RFD	MATS5s	Tm	nHAcc	B04[O-Cl]	F04[Cl-Cl]	LOC
PCDD071	6.940	0.061	0.286	0.038	14.759	2	1	2	0.737
PCDD072	7.687	0.066	0.286	0.014	15.881	2	1	2	0.737
PCDD073	6.464	0.124	0.286	0.033	14.842	2	1	3	0.737
PCDD074	7.066	0.126	0.286	0.072	15.879	2	1	3	0.737
PCDE047	3.275	0.198	0	0.157	12.036	1	0	2	0.732
PCDE066	6.084	0.224	0	0.178	13.447	1	1	1	0.732
PCDE077	7.676	0.316	0	0.129	14.998	1	1	0	0.732
PCDE085	5.265	0.183	0	0.023	12.674	1	1	2	0.745
PCDE099	5.432	0.220	0	0.294	13.556	1	1	2	0.745
PCDE101	5.749	0.240	0	0.302	13.231	1	1	1	0.745
PCDE102	4.700	0.190	0	0.007	11.749	1	1	2	0.745
PCDE105	6.620	0.222	0	0.047	14.019	1	1	1	0.745
PCDE118	6.637	0.186	0	0.05	14.896	1	1	2	0.745
PCDE126	8.675	0.278	0	0.01	17.17	1	1	1	0.745
PCDE128	5.794	0.187	0	-0.104	13.225	1	1	2	0.749
PCDE137	5.920	0.208	0	0.149	14.844	1	1	3	0.749
PCDE138	5.903	0.187	0	0.149	13.974	1	1	2	0.749
PCDE140	5.492	0.240	0	-0.134	14.367	1	1	4	0.749
PCDE147	4.497	0.214	0	0.115	12.527	1	1	3	0.749
PCDE153	6.027	0.245	0	0.402	14.747	1	1	2	0.749
PCDE154	4.532	0.279	0	0.119	13.433	1	1	4	0.749
PCDE157	7.072	0.200	0	-0.073	15.309	1	1	2	0.749
PCDE167	7.137	0.199	0	0.18	15.988	1	1	2	0.749
PCDE170	6.369	0.196	0	0.01	15.227	1	1	3	0.746

Table A2. Continued.

Name	Pred. pIC_{50} from Eq. 4.2	HAT i/i (*h**=0.329)	RFD	MATS5s	Tm	nHAcc	B04[O-Cl]	F04[Cl-Cl]	LOC
PCDE180	6.414	0.234	0	0.252	15.85	1	1	3	0.746
PCDE181	4.172	0.440	0	0.212	13.915	1	1	5	0.746
PCDE182	4.958	0.357	0	-0.024	14.613	1	1	5	0.746
PCDE184	3.388	0.507	0	-0.118	12.77	1	1	6	0.746
PCDE190	5.259	0.334	0	0.301	14.986	1	1	4	0.746
PCDE194	5.785	0.272	0	0.103	15.35	1	1	4	0.74
PCDE196	5.111	0.379	0	0.063	15.04	1	1	5	0.74
PCDE197	3.578	0.528	0	-0.033	13.251	1	1	6	0.74
PCDE203	4.707	0.473	0	0.297	14.939	1	1	5	0.74
PCDE206	5.011	0.579	0	0.137	15.873	1	1	6	0.73
PCDF001	4.597	0.154	0.308	-0.271	7.191	1	1	0	0.544
PCDF005	5.470	0.091	0.308	0.017	9.445	1	1	0	0.648
PCDF006	5.646	0.066	0.308	-0.254	9.947	1	1	1	0.648
PCDF007	4.793	0.128	0.308	0.155	8.695	1	1	0	0.648
PCDF008	5.522	0.151	0.308	-0.347	8.696	1	1	0	0.648
PCDF009	6.177	0.111	0.308	-0.249	9.951	1	1	0	0.648
PCDF010	5.596	0.091	0.308	-0.072	9.441	1	1	0	0.648
PCDF011	4.432	0.133	0.308	0.038	7.86	1	1	0	0.648
PCDF013	3.957	0.166	0.308	0.133	10.388	1	0	1	0.648
PCDF015	6.748	0.067	0.308	0.138	11.733	1	1	0	0.648
PCDF017	6.561	0.083	0.308	-0.137	10.812	1	1	0	0.648
PCDF018	6.591	0.082	0.308	-0.132	10.87	1	1	0	0.648
PCDF019	7.441	0.081	0.308	-0.133	12.206	1	1	0	0.648
PCDF020	4.477	0.146	0.308	-0.23	9.536	1	0	0	0.648

Table A2. Continued.

Name	Pred. pIC_{50} from Eq. 4.2	HAT i/i (*h**=0.329)	RFD	MATS5s	Tm	nHAcc	B04[O-Cl]	F04[Cl-Cl]	LOC
PCDF021	6.290	0.037	0.308	-0.06	11.521	1	1	1	0.703
PCDF022	5.156	0.113	0.308	0.265	10.478	1	1	1	0.703
PCDF023	6.313	0.088	0.308	-0.144	10.523	1	1	0	0.703
PCDF024	6.888	0.070	0.308	-0.056	11.629	1	1	0	0.703
PCDF025	6.384	0.071	0.308	0.081	11.149	1	1	0	0.703
PCDF026	5.497	0.095	0.308	0.08	9.75	1	1	0	0.703
PCDF027	5.715	0.052	0.308	0.055	10.879	1	1	1	0.703
PCDF030	6.460	0.040	0.308	-0.133	11.623	1	1	1	0.703
PCDF031	5.448	0.054	0.308	-0.053	10.211	1	1	1	0.703
PCDF032	5.791	0.092	0.308	-0.105	9.79	1	1	0	0.703
PCDF033	6.273	0.073	0.308	0.063	10.933	1	1	0	0.703
PCDF034	5.818	0.096	0.308	0.196	10.519	1	1	0	0.703
PCDF035	4.915	0.131	0.308	0.195	9.096	1	1	0	0.703
PCDF037	7.209	0.060	0.308	0.06	12.4	1	1	0	0.703
PCDF038	7.947	0.060	0.308	0.067	13.577	1	1	0	0.703
PCDF040	6.927	0.080	0.308	-0.129	11.524	1	1	0	0.703
PCDF041	4.689	0.126	0.308	-0.032	11.282	1	0	1	0.703
PCDF042	6.717	0.035	0.308	0.056	12.458	1	1	1	0.703
PCDF043	4.738	0.174	0.308	0.269	12.046	1	0	1	0.703
PCDF044	7.150	0.106	0.308	-0.241	11.619	1	1	0	0.703
PCDF045	7.793	0.087	0.308	-0.153	12.832	1	1	0	0.703
PCDF047	6.845	0.143	0.308	-0.353	10.884	1	1	0	0.703
PCDF048	5.857	0.058	0.308	-0.22	10.475	1	1	1	0.703
PCDF050	6.115	0.041	0.308	-0.068	11.289	1	1	1	0.732

Table A2. Continued.

Name	Pred. pIC_{50} from Eq. 4.2	HAT i/i (*h**=0.329)	RFD	MATS5s	Tm	nHAcc	B04[O-Cl]	F04[Cl-Cl]	LOC
PCDF051	4.831	0.196	0.308	0.076	12.656	1	0	2	0.732
PCDF052	6.483	0.047	0.308	-0.177	11.62	1	1	1	0.732
PCDF053	6.026	0.043	0.308	-0.005	11.294	1	1	1	0.732
PCDF054	7.561	0.051	0.308	-0.192	13.283	1	1	1	0.732
PCDF055	6.930	0.032	0.308	-0.023	12.676	1	1	1	0.732
PCDF056	6.453	0.049	0.308	0.149	12.317	1	1	1	0.732
PCDF058	6.043	0.082	0.308	0.257	11.918	1	1	1	0.732
PCDF059	6.183	0.050	0.308	-0.135	12.087	1	1	2	0.732
PCDF060	7.452	0.102	0.308	-0.383	12.676	1	1	1	0.732
PCDF061	6.565	0.038	0.308	-0.103	11.918	1	1	1	0.732
PCDF062	7.015	0.090	0.308	-0.173	11.624	1	1	0	0.732
PCDF063	7.040	0.063	0.308	0.089	12.262	1	1	0	0.732
PCDF064	5.759	0.055	0.308	0.075	11.055	1	1	1	0.732
PCDF066	6.984	0.066	0.308	-0.278	12.179	1	1	1	0.732
PCDF067	7.240	0.036	0.308	-0.1	12.988	1	1	1	0.732
PCDF068	5.838	0.086	0.308	0.142	12.176	1	1	2	0.732
PCDF069	7.039	0.031	0.308	-0.02	12.854	1	1	1	0.732
PCDF070	5.437	0.115	0.308	0.215	10.025	1	1	0	0.732
PCDF073	7.573	0.065	0.308	-0.017	12.859	1	1	0	0.732
PCDF074	5.737	0.055	0.308	0.072	11.014	1	1	1	0.732
PCDF075	7.508	0.101	0.308	-0.211	12.313	1	1	0	0.732
PCDF076	5.347	0.084	0.308	0.177	10.639	1	1	1	0.732
PCDF078	6.306	0.038	0.308	-0.03	11.677	1	1	1	0.732
PCDF082	8.216	0.129	0.308	-0.271	13.291	1	1	0	0.732

Table A2. Continued.

Name	Pred. pIC_{50} from Eq. 4.2	HAT i/i (*h**=0.329)	RFD	MATS5s	Tm	nHAcc	B04[O-Cl]	F04[Cl-Cl]	LOC
PCDF083	8.306	0.070	0.308	-0.013	14.022	1	1	0	0.732
PCDF084	5.902	0.093	0.308	0.18	10.677	1	1	0	0.732
PCDF085	6.247	0.039	0.308	-0.03	11.584	1	1	1	0.732
PCDF086	6.406	0.072	0.308	0.078	11.238	1	1	0	0.732
PCDF087	6.679	0.050	0.308	-0.163	12.831	1	1	2	0.745
PCDF089	7.500	0.087	0.308	-0.346	13.706	1	1	2	0.745
PCDF090	6.679	0.050	0.308	-0.163	12.831	1	1	2	0.745
PCDF092	7.160	0.043	0.308	-0.16	12.753	1	1	1	0.745
PCDF094	7.079	0.054	0.308	-0.196	13.386	1	1	2	0.745
PCDF096	6.366	0.051	0.308	-0.035	12.631	1	1	2	0.745
PCDF097	6.113	0.048	0.308	0.087	11.668	1	1	1	0.745
PCDF099	6.544	0.048	0.308	-0.105	12.751	1	1	2	0.745
PCDF100	7.130	0.036	0.308	-0.102	12.838	1	1	1	0.745
PCDF101	7.224	0.049	0.308	-0.02	14.016	1	1	2	0.745
PCDF102	6.940	0.054	0.308	0.041	13.707	1	1	2	0.745
PCDF105	7.592	0.053	0.308	-0.196	13.351	1	1	1	0.745
PCDF106	7.971	0.080	0.308	-0.285	13.744	1	1	1	0.745
PCDF108	6.560	0.038	0.308	0.051	12.29	1	1	1	0.745
PCDF109	8.169	0.056	0.308	-0.17	14.318	1	1	1	0.745
PCDF110	6.526	0.038	0.308	0.054	12.243	1	1	1	0.745
PCDF111	6.201	0.054	0.308	0.139	11.925	1	1	1	0.745
PCDF113	5.954	0.066	0.308	0.048	12.171	1	1	2	0.745
PCDF114	7.012	0.033	0.308	-0.044	12.785	1	1	1	0.745
PCDF115	6.684	0.108	0.308	-0.131	13.764	1	1	3	0.749

Table A2. Continued

Name	Pred. pIC_{50} from Eq. 4.2	HAT i/i (*h**=0.329)	RFD	MATS5s	Tm	nHAcc	B04[O-Cl]	F04[Cl-Cl]	LOC
PCDF116	7.597	0.070	0.308	-0.26	14.063	1	1	2	0.749
PCDF118	6.627	0.049	0.308	-0.034	13.052	1	1	2	0.749
PCDF119	6.345	0.058	0.308	0.037	12.77	1	1	2	0.749
PCDF120	6.042	0.075	0.308	0.111	12.463	1	1	2	0.749
PCDF121	7.518	0.060	0.308	-0.205	14.064	1	1	2	0.749
PCDF122	7.106	0.111	0.308	-0.135	14.412	1	1	3	0.746
PCDF124	6.480	0.112	0.308	-0.072	13.577	1	1	3	0.749
PCDF125	6.296	0.059	0.308	0.037	12.694	1	1	2	0.749
PCDF126	7.031	0.048	0.308	-0.069	13.609	1	1	2	0.749
PCDF127	6.714	0.052	0.308	0.004	13.276	1	1	2	0.749
PCDF129	6.697	0.051	0.308	0.001	13.243	1	1	2	0.749
PCDF130	7.475	0.054	0.308	-0.163	14.093	1	1	2	0.749
PCDF131	7.554	0.122	0.308	-0.251	14.853	1	1	3	0.746
PCDF132	6.738	0.110	0.308	-0.099	13.915	1	1	3	0.746
PCDF133	6.470	0.116	0.308	-0.037	13.634	1	1	3	0.746
PCDF135	6.821	0.217	0.308	-0.193	14.661	1	1	4	0.74
PCDT000	4.031	0.209	0.308	-0.153	6.662	0	0	0	0
PCDT002	5.152	0.100	0.308	-0.04	9.851	0	0	0	0.544
PCDT007	3.840	0.284	0.308	0.338	8.87	0	0	0	0.648
PCDT046	6.070	0.106	0.308	-0.133	11.424	0	0	0	0.703
PCDT076	6.298	0.116	0.308	0.053	12.269	0	0	0	0.732
PCDT104	6.490	0.115	0.308	-0.132	13.02	0	0	1	0.745
PCDT125	6.159	0.182	0.308	-0.078	13.473	0	0	2	0.749
PCDT132	5.780	0.282	0.308	-0.161	13.523	0	0	3	0.746

Table A2. Continued.

Name	Pred. pIC_{50} from Eq. 4.2	HAT i/i (*h**=0.329)	RFD	MATS5s	Tm	nHAcc	B04[O-Cl]	F04[Cl-Cl]	LOC
PCDT135	6.138	0.416	0.308	-0.306	14.586	0	0	4	0.74
PCPhX000	3.820	0.225	0.286	-0.207	7.284	1	0	0	0
PCPhX003	4.590	0.109	0.286	0.215	10.574	1	0	0	0.521
PCPhX010	7.394	0.068	0.286	0.157	12.862	1	1	0	0.625
PCPhX095	7.486	0.086	0.286	-0.35	13.726	1	1	2	0.714
PCPhX128	7.908	0.059	0.286	-0.165	14.846	1	1	2	0.73
PCPhX133	7.261	0.118	0.286	0.032	15.135	1	1	3	0.737
PCPhX135	7.492	0.225	0.286	-0.084	16.068	1	1	4	0.733
PCTA000	4.202	0.218	0.286	-0.018	7.357	0	0	0	0
PCTA001	4.367	0.113	0.286	-0.067	8.621	0	0	0	0.521
PCTA004	5.228	0.101	0.286	-0.051	11.078	0	0	1	0.625
PCTA015	6.225	0.088	0.286	-0.151	11.699	0	0	0	0.682
PCTA047	6.769	0.106	0.286	0.161	13.337	0	0	0	0.714
PCTA055	7.204	0.154	0.286	-0.284	14.725	0	0	2	0.73
PCTA069	7.045	0.151	0.286	-0.141	14.816	0	0	2	0.737
PCTA073	7.310	0.255	0.286	-0.225	15.885	0	0	3	0.737
PCTA075	6.738	0.400	0.286	-0.219	15.832	0	0	4	0.733
Pentamethylantracene	4.954	0.135	0.286	-0.22	9.644	0	0	0	0.73
Phenanthrene	4.039	0.213	0.286	-0.105	6.903	0	0	0	0
Quinoline	1.664	0.362	0.2	-0.181	4.408	1	0	0	0
Tetramethylanthracene	4.682	0.139	0.286	-0.025	9.626	0	0	0	0.714
Tjipanazole	9.490	0.166	0.4	0.121	16.286	0	0	0	0.418