QSPR PREDICTION OF PHYSICO-CHEMICAL PROPERTIES

OF

POLYHALOGENATED DIPHENYL ETHERS

by

NİHAN TAŞDİZEN

BS. in Chem., Yıldız Technical University, 2006

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

2011

QSPR PREDICTION OF PHYSICO-CHEMICAL PROPERTIES

OF

POLYHALOGENATED DIPHENYL ETHERS

APPROVED BY:

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

Prof. Dr. Ferhan Çeçen ……

Prof. Dr. Safiye Erdem ……

DATE OF APPROVAL: 30 September 2011

ACKNOWLEDGEMENTS

Completing this project would not have been possible without the help and support of many people. I would like to briefly acknowledge their contributions.

I wish to express my deep gratitude and appreciation to my advisor, Prof. Dr. Melek Türker Saçan, for her excellent support, providing exceptional advice during the evolution of this research and supporting me through a difficult period in my life. Her insightful advice, patience, understanding, and support have made a lasting impression. She is a mentor from whom I learned not only science but the appreciation of life.

I would like to thank Ünal Taşdizen for writing a program for calculations of *CRI* descriptors of the compounds.

Grateful acknowledgments are also made to Doğa Ertürk and Gülçin Tuğcu for their valuable advice, time and assistance.

I would like to express my sincere gratitude to Prof. Dr. Ferhan Çeçen and Prof. Dr. Safiye Erdem for serving on my dissertation committee for their time, insight, and recommendations.

Last but not least, I would like to thank my parents and brother, who have done and given so much to enrich my life. Without their permanent encouragement and love for me during the last years, I would never achieve my current position.

ABSTRACT

Polyhalogenated compounds (PHCs) such as polybrominated and polychlorinated diphenyl ethers (PBDEs/PCDEs) are an important chemical group because of their environmental persistence, high hydrophobicity, and bioaccumulation in humans. Their physico-chemical properties such as *n*-octanol/air partition coefficient (log K_{oa}) and *n*-octanol/water partition coefficient (log K_{ow}) and toxicities are of fundamental importance to gain a better understanding of the environmental fate and behavior of these compounds.

In this study, several Quantitative Structure-Activity Relationship/Quantitative Structure-Property Relationship (QSAR/QSPR) models were developed on log K_{oa} of PBDEs, log K_{ow} of PBDEs and PCDEs, and the aryl hydrocarbon receptor relative binding affinity (log *RBA*) of PBDEs by employing Heuristic Method (HM) and Multiple Linear Regression (MLR). Descriptors used were from DRAGON 5.4, SPARTAN 06, and CODESSA 2.2 software and the Characteristic Root Index *(CRI)* program. All the best models were internally validated for their performance using the leave-one-out procedure and scrambling of the responses. External validation was provided by splitting the data sets into training and test sets either using random division in terms of property modeled or Kohonen network considering the size of the data sets.

Of the models developed log K_{oa} and log K_{ow} models were validated externally by using test sets. log *RBA* model could not be validated externally because of a lack of *RBA* data. E_{HOMO} and E_{aq} from SPARTAN, and the *CRI* appeared to be significant descriptors for the developed log K_{ow} models of PBDEs/PCDEs. The *CRI* also appeared to be an important parameter in modeling log K_{oa} of PBDEs. The statistical quality of all the models for polyhalogenated diphenyl ethers is compared to those of the previously published models using the same experimental data and found to be superior to those models. All the QSAR/QSPR models were developed taking into account the OECD principles for validation, for regulatory purposes, of QSAR. This implied internal and external validations, the analysis of the applicability domain (AD) and, when possible, a mechanistic interpretation of the models.

ÖZET

Polibromlu ve poliklorlu difenil eterler (PBDEs/PCDEs) gibi Polihalojenli bileşikler çevredeki dayanıklıkları, yüksek hidrofob özellikleri ve canlılardaki biyobirikimlerinden dolayı önemli bir kimyasal gruplardır. *n*-oktanol/hava (log K_{oa}), *n*-oktanol/su (log K_{ow}) dağılım katsayısı gibi fizikokimyasal ve toksisite gibi özellikleri, bu bileşiklerin çevresel akıbeti ve davranışlarını daha iyi anlayabilmek için önem taşımaktadır.

Bu çalışmada, PBDE için log K_{oa} ve log *RBA*, PBDE ve PCDE için log K_{ow} modelleri Heuristic yöntem ve Çoklu Doğrusal Regresyon (MLR) kullanılarak Kantitatif yapı-aktivite/Kantitatif yapı-özellik ilişkisi (QSAR/QSPR) ile oluşturuldu. Tanımlayıcılar DRAGON 5.4, SPARTAN 06 ve CODESSA 2.2 programları ve *CRI* programından elde edilmiştir. Tüm modellerin içsel validasyonları, birini dışarda bırak (leave-one-out) ve bağımlı değişkenin rastgele karıştırılması (scrambling of the responses) yöntemleri kullanılarak yapıldı. Dışsal validasyon için veri setini eğitim ve test setlerine ayırmada modellenen özelliğe ait veri setinin boyutu göz önünde bulundurularak rastgele ayırma veya Kohonen ağları kullanıldı.

log K_{oa} ve log K_{ow} modelleri test set kullanılarak dışsal olarak valide edildi. Log *RBA* modeli, *RBA* verisi eksikliğinden dolayı dışsal olarak valide edilemedi. PBDEs/PCDEs için oluşturulan log K_{ow} modelinde, SPARTAN programı ile hesaplanan E_{HOMO} ve E_{aq}, ve *CRI* tanımlayıcıları en önemli tanımlayıcı olarak yer aldılar. *CRI* ayrıca PBDEler için oluşturulan log K_{oa} modelinde de önemli bir tanımlayıcı olarak ortaya çıktı. Polihalojenli difenil eterler için oluşturulan modeller, literatürde daha önce aynı deneysel veri için yayınlanmış olan modeller ile kıyaslandığında, bu çalışmadaki modellerin literatürdeki modellere göre istatistiksel kalite bakımından daha üstün oldukları bulundu. Bütün QSAR/QSPR modelleri yönetmelik amaçlı kullanılabilecek şekilde OECD tarafından yayınlanan prensipler göz önünde bulundurularak oluşturuldu. Bu prensipler, modellerin içsel ve dışsal validasyonunu, uygulanabilirlik alanının (AD) analizini ve mümkünse mekanistik yorumunu içermektedir.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF SYMBOLS/ABBREVIATIONS

Symbol	Explanation	Units Used
AhR	Aryl Hydrocarbon Receptor	
AD	Applicability Domain	
BDE	Brominated Diphenyl Ether	
BFRs	Brominated Flame Retardants	
CAS	Chemical Abstracts Service	
CRI	Characteristic Root Index	
CV	Cross-validation	
DBDPE	Decabromodiphenyl Ether	
E	Energy	(eV)
EDCs	Endocrine Disrupting Chemicals	
E_{HOMO}	Energy of the highest occupied molecular orbital	(eV)
E_{LUMO}	Energy of the lowest unoccupied molecular orbital	(eV)
EU	European Union	
F	Fisher statistic	
G	Graph as a square matrix with the entries	
h	Number of hydrogen atoms bound to the same atom	
h^*	Critical hat value	
H	Henry's law constant	
HM	Heuristic Method	
IUPAC	International Union of Pure and Applied Chemistry	
K_{ow}	Octanol-Water Partition Coefficient	
LOO	Leave-one-out statistical procedure	
LMO	Leave-many-out statistical procedure	
MLR	Multiple Linear Regression	
n	Number of compounds	
OECD	Organisation for Economic Cooperation and Development	
PBBEs	Polybrominated Biphenyl Ethers	
PBBOs	Polybrominated Biphenyl Oxides	

PBDEs	Polybrominated Diphenyl Ethers
PBDOs	Polybrominated Diphenyl Oxides
PCBs	Polychlorinated Biphenyls
PCDEs	Polychlorinated Diphenyl Ethers
PCDDs	Polychlorinated Dibenzo-p-dioxins
PCDFs	Polychlorinated Dibenzofurans
PCNs	Polychlorinated Naphthalenes
QSAR	Quantitative Structure - Activity Relationship
QSPR	Quantitative Structure - Property Relationship
QSTR	Quantitative Structure - Toxicity Relationship
RBT	Ready Biodegradability Tests
REACH	Registration, Evaluation, Authorization and Restrictions of Chemicals
R^2	Coefficient of determination – measure of goodness-of-fit
RBA	Relative Binding Affinity
RMSE	Root Mean Squared Error
RoHS	Restriction of Hazardous Substances Directive
SE	Standard Error
SMILES	Simplified Molecular Line Entry system
SOM	Self-Organizing Maps
SR	Standardized Residual
SSE	Sum of Squared Errors
SVM	Support Vector Machines
TLSER	Theoretical Linear Solvation Energy Relationship
VIF	Variance Inflation Factor
PBDDs	Polybrominated dibenzo-*p*-dioxins
PLS	Partial Least Square regression
PRESS	Predictive Residual Error Sum of Squares
RBFNs	Radial Basis Function Neural Networks
R^2_{cv}	The cross-validated squared correlation coefficient
R^2_{pred}	The squared correlation coefficient of external validation

1. INTRODUCTION

The European Union (EU) Registration, Evaluation, Authorization and Restrictions of Chemicals (REACH) regulation encourages the use of alternative *in vitro* and *in silico* methods in order to minimize animal testing, costs and time. In this context the use of Quantitative Structure-(Activity/Property/Toxicity) Relationships (QSAR/QSPR/QSTR) becomes particularly useful to predict unknown activities/properties for existing or even not yet synthesized chemicals.

In the EU REACH regulation, the chemicals with particularly harmful behaviors, such as endocrine disruptors (EDs), are subject to authorization, and the identification of safer alternatives to these chemicals is required. In this context, the use of QSAR and QSPR becomes particularly useful to fill the data gap due to the very small number of experimental data available to characterize the environmental and toxicological profiles of new and emerging pollutants with ED behavior such as polybrominated diphenyl ethers (PBDEs) and polychlorinated diphenyl ethers (PCDEs).

Brominated flame retardants including PBDEs are man-made chemicals that are widely used to decrease the flammability of plastics, electronic appliances, textiles, and polyurethanes (Pirard et al., 2003; Birnbaum and Staskal, 2004; Richardson, 2004; Keum and Li, 2005). Because of their relatively low cost and effective flame retardation capabilities, more than 67,400 tons of PBDEs were produced annually throughout the world, including penta-BDE, octa-BDE, and deca-BDE (Birnbaum and Staskal, 2004; Richardson, 2004). Unlike more reactive fire retardants such as tetrabromobisphenol-A, PBDEs can be readily incorporated into polymers without making any covalent bonds to adjacent materials. Therefore, they can easily migrate to the environment by improper disposal of used products that contain these substances. Their widespread presence and rapid appearance in mammalian tissues through bioaccumulation have evoked concern over their toxicity and potential ability to cause endocrine disruption (Hakk and Letcher, 2003; Birnbaum and Staskal, 2004). Recently, PBDEs with 4 to 7 bromines were newly listed as persistent organic pollutants (POPs) in the Stockholm Convention.

PCDEs have been found as by-products in chlorinated phenols and chlorinated phenoxyacetic acids, where they have been identified at a level of 100-1000 mg/kg (Nilsson and Renberg, 1974; Kurz and Ballschmiter, 1995a; 1995b). The extensive use of the chlorinated phenols, especially pentachlorophenol as an herbicide and in wood preserving formulations, leads to a ubiquitous appearance of the PCDEs in the environment. Another source of PCDEs is all processes of incomplete combustion e.g., municipal waste incinerators (Paasivirta et al., 1986; Kurz and Ballschmiter, 1995a). PCDEs have been detected in a wide range of environmental samples. PCDEs have been determined in sediments and fish from the North American Great Lakes (Coburn and Comba, 1981; Coburn and Comba, 1985; Niimi et al., 1994), in bird tissue and eggs (Stafford, 1983), in white-tailed eagle muscles (Paasivirta et al., 1986; Koistinen et al., 1993a), in salmon (Koistinen et al., 1993b), in cod liver oils (Kurz and Ballschmiter, 1995a; 1995b), and in human adipose tissue (Stonley et al., 1990; Williams et al., 1991). Besides the knowledge of toxicological and analytical data, the availability of physico-chemical data of the PCDEs is important. From the knowledge of physico-chemical properties (e.g., vapour pressure, aqueous solubility) statements can be made about the distribution, and the regional and global transport in the different environmental compartments.

Predicting the fate and transport of these compounds remains tenuous due to the lack of physical-chemical data. Physico-chemical properties of micropollutants, especially those partitioning properties such as *n*-octanol/air partition coefficient (K_{oa}) and *n*-octanol/water partition coefficient (K_{ow}) play a major role in their transport and mobility in the global environment. This is particularly true for those Persistent Organic Pollutants (POPs) chemicals like PBDEs and PCDEs. The distribution of these substances between the particle and gas phases in the atmosphere is a key factor which governs their deposition to water bodies and vegetation and controls their atmospheric lifetimes which are dependent on removal by gas-phase reactions. Other significant partition coefficient is K_{ow} and describes partitioning of the substance between water and environmental organic phases. They are mathematically defined as a quotient of *n*-octanol/air and *n*-octanol/water concentration at the equilibrium state (Xiao and Wania, 2003; Wania and Su, 2004).

Besides physico-chemical properties the toxicology of PBDEs is currently under investigation. But, evidences emerge that PBDEs may be developmental neurotoxicants, and may cause neurochemical and hormonal deficits (Eriksson et al., 2001; Viberg et al., 2002; Zhou et al., 2002; Branchi et al., 2003). EU has banned the use in all applications of penta-BDE and octa-BDE in the EU market. However, limited toxicological data are obtained for only the most prevalent individual PBDE congeners (e.g., BDE-47, BDE-99, and BDE-100) (Chen et al., 2001; Rahman et al., 2001). Additionally, Aryl hydrocarbon Receptor (AhR) binding affinities of some PBDE congeners have been determined. The AhR is a ligand-activated transcription factor. AhR plays a very important role in the detoxification of endo- and xenobiotics. The large number of PBDE congeners complicates toxicological studies further.

Maximum half of the congeners of PBDEs and PCDEs have reported log K_{ow} and only 22 and 18 congeners of PBDEs have reported log K_{oa} and log *RBA* values, respectively. Although no experimental data concerning the physico-chemical and toxic properties of the remaining PBDE and PCDE congeners are currently available, it has been suggested that QSPR can be an alternative approach for estimating these physico-chemical properties. As a common and successful research approach, QSAR/QSPR studies are applied extensively to chemometrics, pharmacodynamics, pharmacokinetics, toxicology and so on.

Multiple Linear Regression (MLR) is one of the earliest methods used for constructing QSAR/QSPR models, but it is still one of the most commonly used ones to date. The advantage of MLR is its simple form and easily interpretable mathematical expression. Although utilized to great effect, MLR is vulnerable to descriptors which are correlated to one another, making it incapable of deciding which correlated sets may be more significant to the model. Some new methodologies based on MLR have been developed and reported in recent papers aimed at improving this technique. Heuristic Method (HM)**,** an advanced algorithm based on MLR, is popular for building linear QSAR/QSPR equations because of its convenience and high calculation speed. The advantage of HM is totally based on its unique strategy of selecting variables.

HM is commonly used in linear QSAR and QSPR studies, and also as an excellent tool for descriptor selection before a linear or nonlinear model is built (Luan et al., 2006; Xia et al., 2009). The advantages of HM are the high speed and the absence of software restrictions on the size of the data set. HM can either quickly give a good estimation about what quality of correlation to expect from the data, or derive several best regression models. HM usually produces correlations 2-5 times faster than other methods with comparable quality. Additionally, the maximum number of parameters in the resulting model can be fixed in accordance with the situation so as to save time. As a method inherited from MLR, HM is also limited in linear models.

In this work, several QSPR models were used to predict some important physico-chemical properties for all the 209 PBDE and PCDE congeners along with the diphenyl ether. The studied end points were: *n*-octanol/air partition coefficient (log K_{oa}), *n*-octanol/water partition coefficient (log K_{ow}) and Ah receptor binding affinity (log *RBA*). In previous studies, the Characteristic Root Index (*CRI*), which comprises all possible orders of path-type valence connectivity indices, has been demonstrated to correlate many physico-chemical and biological properties of organic compounds including solubility (Saçan and İnel, 1993), *n*-octanol/water partition coefficient (Saçan and İnel, 1995; Saçan et al., 2005), soil sorption coefficient (K_{oc}) (Saçan and Balcıoğlu, 1996), vapor pressure (Saçan and Balcıoğlu, 1998), bioconcentration factor (Saçan et al., 2003), and toxicity of aromatic compounds to algae (Saçan et al., 2007). Hence, it is logical to examine the relationship between the *CRI* and physico-chemical properties and toxicities of polyhalogenated diphenyl ethers. The *CRI* was calculated by using a program written in Delphi 2007 by Ünal Taşdizen for Polyhalogenated Diphenyl Ethers. Other descriptors were calculated using DRAGON 5.4, CODESSA 2.2 (Comprehensive Descriptors for Structural and Statistical Analysis) and SPARTAN 06 software. The latter descriptors were used to increase the predictive ability of the *CRI*-based models.

A total of 1164 from DRAGON 5.4, 163 from CODESSA 2.2, and 8 from SPARTAN 06 software and *CRI* descriptors were retained and subsequently used in model development. The predictive ability of the resulting models is compared in order to determine whether or not the more complex descriptors are necessary to predict the property or activity of interest. For comparative purposes, predictive models based on each

descriptor class, alone, were also developed. To compare the predictive performance of all the models developed for training and test sets in this study with those of the literature models R^2_{cv} and R^2_{pred} were compared.

1.1. Purpose of the Study

The aim of the present study is (1) to calculate the theoretical molecular descriptors representing the studied molecular structures from DRAGON 5.4, SPARTAN 06 and CODESSA 2.2 software and the *CRI* program; (2) to select descriptors with HM of CODESSA 2.2; (3) to explore QSPR models relating to log K_{oa} of PBDEs, log K_{ow} of PBDEs and PCDEs, and log *RBA* of PBDEs by employing Multiple Linear Regression (MLR) method. Additionally, the purpose of the study is to build models according to the following scheme: (1) a split of the chemicals into a training set and test set in terms of the property to be modeled for small data set and using Kohonen network for large data set; (2) selection of a model satisfactory for the training set; (3) external validation of the model with the test set; (4) using the leave-one/many-out cross-validation and *Y*-scrambling procedures for the internal validation of all the best models; (5) using the leverage approach for verifying the applicability domain by highlighting both the response-outliers and the structural influential chemicals (Williams graph). Another purpose of the study is to predict log K_{ow} values of PBDE/PCDE congeners, and log K_{oa} and log *RBA* values of PBDE congeners outside the sample set with the best developed models and to compare the best models with the reported literature models.

We hope that the resulting QSPRs are to be both descriptive (pinpointing the key descriptors) and predictive (able to predict the physico-chemical properties of compounds which are not included in the QSPR determinations).

2. THEORETICAL BACKGROUND

2.1. Polybrominated Diphenyl Ethers (PBDEs)

Polybrominated diphenyl ethers (PBDEs) are chemical relatives of Polychlorinated biphenyls (PCBs), another family of highly persistent and bioaccumulative toxicants with a structure very much like PCBs that came to the attention after a huge quantity of them had been released into the environment (Hites, 2004).

2.1.1. Structure of PBDEs

PBDEs are a class of structurally similar brominated hydrocarbons, in which 2-10 bromine atoms are attached to the diphenyl ether molecule. Monobrominated structures (i.e., one bromine atom attached to the molecule) are often included when describing PBDEs.

PBDE is the common name for the 209 different Brominated Diphenyl Ether (BDE) congeners that can theoretically be obtained in complex mixtures via bromination of diphenyl ether or as individual, pure, compounds via specific routes for their synthesis. The common structure of PBDE congeners is depicted in Figure 2.1.

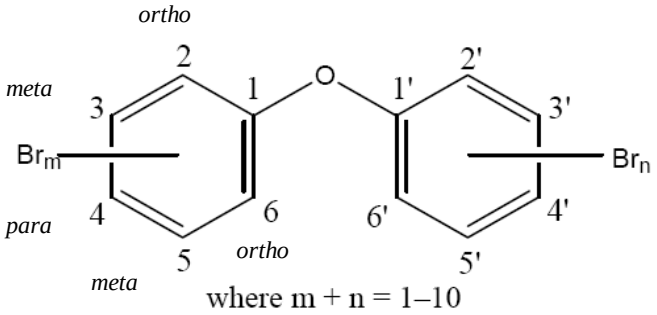


Figure 2.1 Common structure (General structural formula) and substitution positions of the PBDEs. The numbers denote the various bromine atoms numbering by carbon position.

Various synonyms and abbreviations of polybrominated diphenyl ethers are used in the literature. In this thesis - in accordance with the monograph from IPCS (1994) - the chemical name polybrominated diphenyl ethers is used. To indicate that it is a group of compounds the abbreviation PBDEs is used instead of the more widespread PBDE. The same group may as well be named polybrominated biphenyl ethers (PBBEs), polybrominated biphenyl oxides (PBBOs), or polybrominated diphenyl oxides (PBDOs) in the literature.

The 209 possible compounds for PBDEs are called “congeners”. The term “homolog” is used to refer to all PBDEs with the same number of bromines (e.g., tribromodiphenyl ether refers to PBDEs containing only three bromine atoms). Based on the number of bromine substituents, there are 10 homologous groups of PBDEs (monobrominated through decabrominated). Each homologous group contains one or more congeners. The mono-, di-, tri-, tetra-, penta-, hexa-, hepta-, octa-, nona-, and decabromo-congeners can exist in 3, 12, 24, 42, 46, 42, 24, 12, 3, and 1 forms, respectively. Homologs with different substitution patterns are referred to as isomers. For example, the group of dibromodiphenyl ether homologs contains 12 isomers. There are important three-dimensional differences in their structures within congeners due to the ether linkage and location/number of halogen atoms.

The ether linkage in the diphenyl oxide molecule introduces a high barrier to rotation and prevents the two aromatic rings from assuming a planar configuration. In addition, the *ortho* positions of the aromatic rings must be nonhalogen-substituted for a biphenyl or diphenyl oxide (diphenyl ether) molecule to assume a planar or near planar configuration.

Halogen substitution of the diphenyl ether molecule in the *ortho* positions (2,2’,6,6’) will force the aromatic rings orthogonal to one another. Decabromodiphenyl ether (DBDPE), fully substituted at all ring positions, exists with its two aromatic rings orthogonal to one another. In addition, the ether bridge in the DBDPE molecule introduces a 120° bend in the alignment of the biphenyl rings. The benzene rings of non-*ortho* substituted PBDEs may assume a small dihedral angle (in which the dihedral angle is small, but >0°) or “near” planar configuration. These molecules are referred to as planar or

coplanar congeners. The molecular weights of the PBDEs range from 249.103 g/mol to 959.167 g/mol based on the natural abundance of carbon, hydrogen, and bromine. The PBDEs and polychlorinated diphenyl ethers (PCDEs) are numbered as proposed for the PCBs by Ballschmiter and Zell (1980) and Schulte and Malisch (1983) by using the abbreviation BDE and their corresponding number (which are also referred to as congener, IUPAC, or BZ numbers) (Table 2.1.).

Table 2.1 Systemic numbers of the PBDEs and PCDEs[1].

No	Structure	No	Structure	No	Structure	No	Structure	No	Structure
1	2	43	2,2',3,5	85	2,2',3,4,4'	127	3,3',4,5,5'	169	3,3',4,4',5,5'
2	3	44	2,2',3,5'	86	2,2',3,4,5	128	2,2',3,3',4,4'	170	2,2',3,3',4,4',5
3	4	45	2,2',3,6	87	2,2',3,4,5'	129	2,2',3,3',4,5	171	2,2',3,3',4,4',6
4	2,2'	46	2,2',3,6'	88	2,2',3,4,6	130	2,2',3,3',4,5'	172	2,2',3,3',4,5,5'
5	2.3	47	2,2',4,4'	89	2,2',3,4,6'	131	2,2',3,3',4,6	173	2,2',3,3',4,5,6
6	2,3'	48	2,2',4,5	90	2,2',3,4',5	132	2,2',3,3',4,6'	174	2,2',3,3',4,5,6'
7	2.4	49	2,2',4,5'	91	2,2',3,4',6	133	2,2',3,3',5,5'	175	2,2',3,3',4,5',6
8	2,4'	50	2,2',4,6	92	2,2',3,5,5'	134	2,2',3,3',5,6	176	2,2',3,3',4,6,6'
9	2.5	51	2,2',4,6'	93	2,2',3,5,6	135	2,2',3,3',5,6'	177	2,2',3,3',4,5',6'
10	2.6	52	2,2',5,5'	94	2,2',3,5,6'	136	2,2',3,3',6,6'	178	2,2',3,3',5,5',6
11	3,3'	53	2,2',5,6'	95	2,2',3,5',6	137	2,2',3,4,4',5	179	2,2',3,3',5,6,6'
12	3.4	54	2,2',6,6'	96	2,2',3,6,6'	138	2,2',3,4,4',5'	180	2,2',3,4,4',5,5'
13	3,4'	55	2,3,3',4	97	2,2',3,4',5'	139	2,2',3,4,4',6	181	2,2',3,4,4',5,6
14	3.5	56	2,3,3',4'	98	2,2',3,4',6'	140	2,2',3,4,4',6'	182	2,2',3,4,4',5,6'
15	4,4'	57	2,3,3',5	99	2,2',4,4',5	141	2,2',3,4,5,5'	183	2,2',3,4,4',5',6
16	2,2',3	58	2,3,3',5'	100	2,2',4,4',6	142	2,2',3,4,5,6	184	2,2',3,4,4',6,6'
17	2,2',4	59	2,3,3',6	101	2,2',4,5,5'	143	2,2',3,4,5,6'	185	2,2',3,4,5,5',6
18	2,2',5	60	2,3,4,4'	102	2,2',4,5,6'	144	2,2',3,4,5',6	186	2,2',3,4,5,6,6'
19	2,2',6	61	2,3,4,5	103	2,2',4,5',6	145	2,2',3,4,6,6'	187	2,2',3,4',5,5',6
20	2,3,3'	62	2,3,4,6	104	2,2',4,6,6'	146	2,2',3,4',5,5'	188	2,2',3,4',5,6,6'
21	2,3,4	63	2,3,4',5	105	2,3,3',4,4'	147	2,2',3,4',5,6	189	2,3,3',4,4',5,5'
22	2,3,4'	64	2,3,4',6	106	2,3,3',4,5	148	2,2',3,4',5,6'	190	2,3,3',4,4',5,6
23	2,3,5	65	2,3,5,6	107	2,3,3',4',5	149	2,2',3,4',5',6	191	2,3,3',4,4',5',6
24	2,3,6	66	2,3',4,4'	108	2,3,3',4,5'	150	2,2',3,4',6,6'	192	2,3,3',4,5,5',6
25	2,3',4	67	2,3', 4,5	109	2,3,3',4,6	151	2,2',3,5,5',6	193	2,3,3',4',5,5',6
26	2,3',5	68	2,3',4,5'	110	2,3,3',4',6	152	2,2',3,5,6,6'	194	2,2',3,3',4,4',5,5'
27	2,3',6	69	2,3',4,6	111	2,3,3',5,5'	153	2,2',4,4',5,5'	195	2,2',3,3',4,4',5,6
28	2,4,4'	70	2,3',4',5	112	2,3,3',5,6	154	2,2',4,4',5,6'	196	2,2',3,3',4,4',5,6'
29	2,4,5	71	2,3',4',6	113	2,3,3',5',6	155	2,2',4,4',6,6'	197	2,2',3,3',4,4',6,6'
30	2,4,6	72	2,3',5,5'	114	2,3,4,4',5	156	2,3,3',4,4',5	198	2,2',3,3',4,5,5',6
31	2,4',5	73	2,3',5',6	115	2,3,4,4',6	157	2,3,3',4,4',5'	199	2,2',3,3',4,5,5',6'
32	2,4',6	74	2,4,4',5	116	2,3,4,5,6	158	2,3,3',4,4',6	200	2,2',3,3',4,5,6,6'
33	2,3',4'	75	2,4,4',6	117	2,3,4',5,6	159	2,3,3',4,5,5'	201	2,2',3,3',4,5',6,6'
34	2,3',5'	76	2,3',4',5'	118	2,3',4,4',5	160	2,3,3',4,5,6	202	2,2',3,3',5,5',6,6'
35	3,3',4	77	3,3',4,4'	119	2,3',4,4',6	161	2,3,3',4,5',6	203	2,2',3,4,4',5,5',6
36	3,3',5	78	3,3',4,5	120	2,3',4,5,5'	162	2,3,3',4',5,5'	204	2,2',3,4,4',5,6,6'
37	3,4,4'	79	3,3',4,5'	121	2,3',4,5',6	163	2,3,3',4',5,6	205	2,3,3',4,4',5,5',6
38	3,4,5	80	3,3',5,5'	122	2,3,3',4',5'	164	2,3,3',4',5',6	206	2,2',3,3',4,4',5,5',6
39	3,4',5	81	3,4',4,5	123	2,3',4,4',5'	165	2,3,3',5,5',6	207	2,2',3,3',4,4',5,6,6'
40	2,2',3,3'	82	2,2',3,3',4	124	2,3',4',5,5'	166	2,3,4,4',5,6	208	2,2',3,3',4,5,5',6,6'
41	2,2',3,4	83	2,2',3,3',5	125	2,3',4',5',6	167	2,3',4,4',5,5'	209	2,2',3,3',4,4',5,5',6,6'
42	2,2',3,4'	84	2,2',3,3',6	126	3,3',4,4',5	168	2,3',4,4',5',6		

[1]Ballschmiter and Zell, 1980; Schulte and Malisch, 1983.

Structural difference between the chlorinated and brominated molecules relates to molecular geometries. Bromine atoms occupy a considerably larger volume than chlorine atoms, and as a consequence, brominated molecules have a larger molecular volume than do molecules containing a similar number of chlorine atoms (Hardy, 2002).

2.1.2. Uses

PBDEs are used as flame retardants with amounts ranging from 5 to 30% added to products to protect them from catching on fire. The products may be polymers, resins, rubbers or textiles (WHO, 1994).

PBDEs are one of the Brominated Flame Retardants (BFRs) which are a diverse group of chemicals that are used to increase fire safety. They are incorporated into a wide range of products, such as TVs, computers, household appliances, textiles and upholstery. BFRs have led to both scientific and public concern since they have been found to accumulate in man and wildlife. BFRs are linked to adverse physiological effects both in vitro and in vivo (e.g., interference in neurobehavioral development, fetal health and thyroid function). Adequate data on the effects are currently still insufficient to fully understand their toxicology.

The PBDEs are used as additives, which mean that they are moulded with the polymer or rubber to form the final product. Since the PBDEs are not covalently bound with the product, they can migrate (diffuse) out of the products during their lifetime. A study has shown the migration of PBDEs from electronic devices into air, from recycling of electronic equipment, followed by uptake in humans (Sjödin et al., 2001).

2.1.3. Physico-chemical Properties of PBDEs

PBDEs have high octanol-water partitioning coefficient (log K_{ow}) values, all above 5, which is an indication of possible bioaccumulation. The high log K_{ow} values generate problems when performing tests in aqueous media, which is one reason for them not passing the ready biodegradability tests (RBT tests). For BDE-209, solubility in water is extremely low and its abundance for possible microbial degradation in any water based

RBT is therefore an inappropriate methodology for assessment of persistence. PBDEs also have low vapour pressures at room temperature, changing in relation to the bromine content of the congener with lower values for the more brominated congeners. Vapour pressure for BDE-47 has been measured to be $2.5x10^{-4}$ Pa (Wong et al., 2001), while BDE-183 has a vapour pressure of $4.7x10^{-7}$ Pa (Tittlemier et al., 2002). The chemical characteristics of PBDEs influence their behaviour in the abiotic environment and in biota. In other words, these characteristics are critical for modeling aspects such as the transport and fate, persistence, bioconcentration, and biological activity of the congeners.

2.1.4. PBDE Human Levels: A Cause of Concern

Synthetic halogenated compounds including chlorinated dioxins, dibenzofurans, and PCBs have been identified as global environmental and human contaminants over the past 30 years. Their harmful effects on wildlife and humans have been extensively reviewed, although their mechanism(s) of action remains unclear. Continuous monitoring of human samples has revealed that PBDE levels have exponentially increased in the last decades. For example, an analysis of human breast milk between 1972 and 1997 showed a 60-fold increase in the PBDE levels in Swedish women (Meironyte et al., 1999), and recent studies have reported much higher levels in human breast adipose tissue (She et al., 2002) and human breast milk (Schecter et al., 2003). In the last years, a large number of articles have been published reporting very high levels of this compound in milk, blood and adipose tissue. Like environmental levels, an exponential increase with a doubling time of ~5 years has also been observed in human tissue, and the levels have risen by a factor of ~100 over the last 30 years according to Hites (2004). Another interesting factor is that North American samples are always above the regression line by a factor of >10 (when the data from all countries are plotted together) and the Japanese always below by a factor of ~5 with Europe in the middle range. European and Japanese demand for PBDEs are mostly deca-BDE mixtures, whereas there is a massive demand for penta-BDE mixtures in the U.S., accounting for 98% of total world production (BSEF, 2004). Deca-BDE congeners in the environment vary from nondetectable to low detectable levels, while tetra- and penta-BDE congeners are very persistent (WHO, 1994). Thus, it has been suggested that the “quality” of PBDEs has a direct impact on human and environmental

levels. Moreover, widespread use of manufacture products with flame retardancy mandated by federal law may account for the high levels in the U.S.

Congener-specific analysis in human samples should not be taken literally into account because only a few congeners have been measured thus far. However, it gives a rough, but extremely valuable, estimation of congener distribution. The most predominant PBDE congener is the tetra-BDE 47 (accounting for more than 50% of total PBDE), followed by penta-BDE 99 (the subject of the present investigation), hexa-BDE 153 and penta-BDE 100 (Krupp et al., 1988; de Wit, 2002; Kalantzi et al., 2004; Sjödin et al., 2004). A recent study estimated the half-life of the most predominant congeners, PBDE 47, PBDE 99, PBDE 100, PBDE 154 and PBDE 153, which reported a very high half-life of 1.8, 2.9, 1.6, 3.3 and 6.5 years, respectively (Geyer et al., 2004). The predominance of these congeners with long half-lives in commercial mixtures is the key factor for their persistence in environmental and human samples. Given that PBDE are used in plastics and in electronic products, one would expect that workers involved in assembling or disassembling these products would have a higher burden. In a few publications, PBDE serum concentrations from occupationally exposed employees (involved in dismantling electronics like computers) were compared to those of non-exposed employees in the same facility. Although the data for occupational exposure are not as complete as those for environmental exposure, a congener-specific analysis has shown that blood levels of exposed workers were twice those of the controls (Sjödin et al., 1999; Thomsen et al., 2001; Lado-Abeal et al., 2003).

While there is a considerable amount of information on PBDE levels in human samples, no data has thus far been reported on their toxicological effects in humans. Several studies have only recently published experimental evidence from animal studies suggesting that PBDEs may interfere with normal developmental and physiological processes sometimes at dose levels close to human exposure. Moreover, congener-specific potency and a broad scope of toxic effects have been demonstrated for PBDEs similar to PCB-mediated toxic effects.

2.1.5. Metabolism and Toxicological Effects of PBDEs

The metabolism of Brominated Flame Retardants (BFRs) was recently reviewed by Hakk and Letcher (2003). Congener specific PBDE uptake and metabolism has been investigated for several PBDE congeners, in the rat and mouse. The uptake of ^{14}C-labelled BDE-47 is similar to PCBs in the rodents and the compounds are retained in the body as observed after five days (Klasson Wehler et al., 1996; Örn and Klasson Wehler, 1998). The uptake of BDE-99 in orally dosed rats is also extensive, and comparable to that of BDE-47 (Hakk et al., 1999). The lower brominated PBDEs, BDE-47 and BDE-99, are distributed to lipid rich tissues as expected (Örn and Klasson Wehler, 1998; Hakk et al., 1999). BDE-209 that is orally dosed to rats has shown to be absorbed to at least 10% (Mörck et al., 2003), while previous studies have indicated poor uptake (Norris et al., 1975; El Dareer et al., 1987). The previously low reported absorption of BDE-209 following an oral exposure may display a dosing vehicle dependency (Hakk and Letcher, 2003).

From an exposure point of view it is interesting to see how PBDEs are transformed in animal experiments *in vivo*. Six metabolites of BDE-47 have been identified in rat and mouse excreta after orally dosing. Most of these metabolites were hydroxylated PBDEs and it was shown that mice are metabolizing BDE-47 rather fast, which is in contrast to rats (Örn and KlassonWehler, 1998). The metabolism of BDE-209 was studied by Mörck et al. (2003) indicating the formation of a large number of metabolites. Another metabolism study was performed on a mixture of PBDE congeners in order to determine whether PBDEs may form OH-PBDEs that are retained in rat blood (Malmberg, 2004a; 2004b).

The toxicity of PBDEs is not as well understood as that of Polychlorinated dibenzo-*p*-dioxins/Polychlorinated dibenzofurans (PCDDs/PCDFs) and PCBs; however, several studies reported their ability to cause neurotoxic effects in mice and dioxin-like endocrine disruption (Branchi et al., 2002; Viberg et al., 2002; 2003b). During incineration of materials containing PBDEs, highly toxic compounds such as polybrominated dibenzo-*p*-dioxins and dibenzofurans, which exhibit similar toxicities as their chlorinated analogues

(PCDDs/PCDFs), can be formed at low combustion temperatures (Birnbaum and Staskal, 2004).

The present knowledge of PBDE toxicity was recently summarised by Gill et al. (2004) and by Birnbaum and Staskal (2004). Some previous reviews including PBDE toxicity have also been published (Darnerud et al., 2001; McDonald, 2002; Darnerud, 2003). PBDEs have toxic effects at relevant concentrations to consider in the risk assessment work. Among those are the endocrine effects, relating both to thyroidogenic and estrogenic effects (Vos et al., 2003). Developmental neurotoxicity has been shown for a number of PBDEs, including BDE-209 (Eriksson et al., 2001; 2002; Viberg et al., 2003a; 2003b; 2004a; 2004b) and the latter compound has been indicated to cause cancer (El Dareer et al., 1987). PBDEs are non-dioxin like compounds even though contradicting results have been presented but that was probably due to contamination of polybrominated dibenzofurans in the technical PBDE products.

PBDEs have several toxic effects. Neurobehavioral effects of PBDEs include primarily impairment in motor behavior, reduced learning and reduction in memory process (Branchi et al., 2003). Reproductive effects e.g., delay in onset of puberty, decrease in the sperm count and reduction in weight of gonads have been seen in male rats. (Birnbaum and Staskal, 2004). PBDEs closely resemble thyroid hormones (T3 & T4) and bind competitively to thyroid hormone transfer proteins thus acting as endocrine disrupters (McDonald, 2002). Chronic exposure to deca-BDE resulted in hepatic and pancreatic adenomas in rats, whereas, a combined incidence of hepatocellular adenomas and carcinomas was seen in mice (NTP, 1986).

2.1.6. Bioaccumulation of PBDEs

Bioaccumulation of individual PBDE congeners depends on their solubility and bioavailability. As the highly brominated deca-BDE has, owing to its high molecular weight, a limited tendency to bioaccumulate, especially the less brominated diphenyl ethers, such as penta-BDEs, which tend to strongly bioaccumulate and have been restricted in several countries. However, in the environment, both microbial reductive dehalogenation and photocatalytic debromination by UV light can produce these critical

congeners from higher brominated diphenyl ethers like deca-BDE or nona-BDEs. Lower brominated diphenyl ethers (Rayne et al., 2003; Söderström et al., 2004) are now commonly detected in the soil around landfills, sludge, and wastewater treatment plants (Richardson, 2004).

In spite of the rapid emergence and accumulation of these contaminants in the environment, the remediation technology for these persisting PBDEs has not been investigated sufficiently. Recent studies showed the possibility of abiotic degradation of PBDEs using zero valent iron through reductive debromination and photolytic debromination using UV light (Keum and Li, 2005). However, in most cases, this results in the production of lower brominated diphenyl ethers being generated from deca-BDE or octa-BDE (Rayne et al., 2003; Söderström et al., 2004; Cupples et al., 2005). With respect to biological degradation, the ether bonds in organic compounds are usually considered to be difficult to be degraded via enzymatic reaction because of their thermodynamic stability (White et al., 1996). Previously, however, the aerobic bacterial degradation and transformation of mono- and dihalogenated diphenyl ethers by *Sphingomonas* sp. SS3 and SS31 was reported (Schmidt et al., 1992; 1993). Further, the fungal transformation was reported for diphenyl ether and mono halogenated diphenyl ethers by the white rot fungus Trametes versicolor (Hundt et al., 1999) demonstrating an oxidative transformation of diphenyl ether and the halogenated derivatives. Schematic representation of environmental behavior of BFRs was given in Figure 2.2.

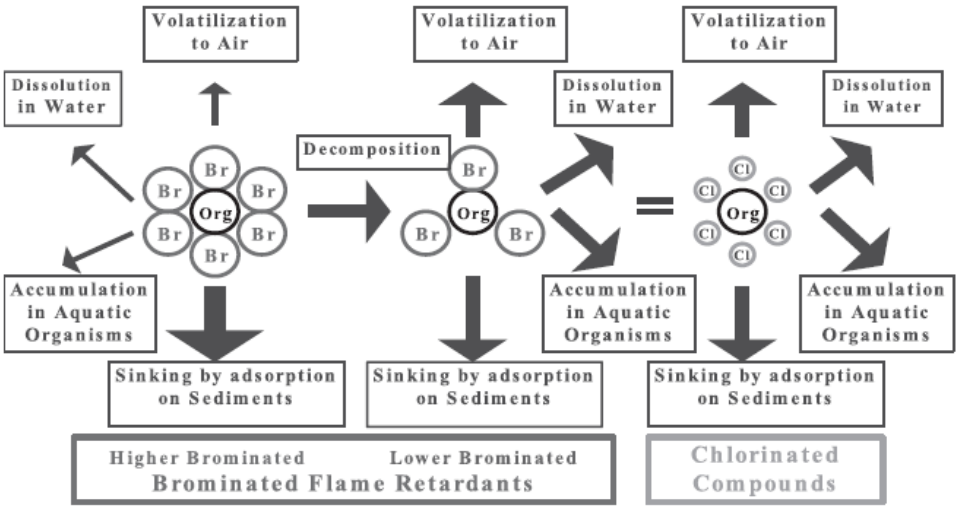


Figure 2.2 Schematic representation of environmental behavior of brominated flame retardants (Watanabea and Sakai, 2003).

2.1.7. Experimental Studies on Toxicology of PBDEs

Most experimental toxicological studies have used commercial PBDE mixtures, but a few investigators did use a single congener. The main criticism of using commercial mixtures is the comparably low purity of the mixtures and lack of knowledge about the possible interference of other compounds (e.g., dioxins). The main toxicological effects and possible modes of action of PBDEs are summarized below.

In rodents, penta-BDE showed a low acute toxicity displaying an LD-50 (lethal dose necessary to kill 50%) in the range of 0.5-5 g/kg body weight (BW). Clinical signs included reduced growth, diarrhea, piloerection, reduced activity, forelimb tremors, reddening around the eyes and nose and continuous chewing. The porphyrin concentration increased considerably after oral dosing with the commercial penta-BDE mixture, DE-71 (mixture of tetra- penta- and hexa-BDE), at 100 mg/kg BW for 13 weeks. No mutagenic activity was observed in the Ames test using several Salmonella strains under induced and noninduced microsomal activation conditions (WHO, 1994). Immunological effects were suggested in mice after exposure to DE-71; suppression of the anti-SRBC response was observed as well as decreased thymus weight (Fowles et al., 1994). The congener PBDE 47 markedly reduced the number of splenocytes in mice (C57BL) after daily oral administration for 14 days (Darnerud and Thuvander, 1999).

Both commercial penta-BDE and single congeners have been known to affect thyroid hormone homeostasis by reducing serum thyroxin levels in rats and mice (WHO, 1994; Hallgren et al., 2001; Zhou et al., 2001; 2002; Hallgren and Darnerud, 2002). This system seems to be one of the most sensitive to PBDE exposure. Effects on thyroxin levels were observed already at single dose of 0.8 mg/kg (Fowles et al., 1994). Tetra-BDE congeners reduced thyroid hormone levels in female rats following 14-day oral administration (18 mg/kg) (Hallgren and Darnerud, 2002). In another study, commercial technical mixtures were administered for 4 days at dosages of 0, 0.3, 1, 3, 10, 30, 100 and 300 mg/kg BW. Dose-related reductions in serum T4 levels were observed for penta- and octa-BDE but not for deca-BDE mixtures. When pregnant rats were treated orally with 0, 1, 10 and 30 mg penta-BDE mixture/kg BW from gestational day 6 to postnatal day 21, the fetuses and dams on gestational day 21 and the offspring on postnatal day 4 and 14 had

lower serum concentrations of T4 in all doses tested (Zhou et al., 2002). Two mechanisms have thus far been proposed to explain the PBDE-induced hypothyroidism. PBDE metabolites are thought to bind to the thyroxin-transporting protein TTR, thereby decreasing the thyroxin levels in blood (Meerts et al., 2000). Moreover, PBDEs may induce the phase II enzyme uridine diphosphoglucuronosyl transferase (UDPGT), the enzyme involved in the metabolism of T4 that increases the rate of elimination (Zhou et al., 2001). However, insufficient information and a lack of consistency in the set of data reported thus far warrant further investigation on the mode of action in thyroid disruption.

Technical PBDE mixtures as well as some congeners are able to induce both phase I and phase II detoxification enzymes in the liver, which are involved in the metabolism and/or metabolic activation of xenobiotics. Regarding cytochrome P450-mediated phase I metabolism, CYP 1A and CYP 2B families were induced as demonstrated by the increasing activity of liver microsomal ethoxyresurufin-O-deethylase (EROD), methoxyresorufin-O-demethylase (MROD) and penthoxyresorufin-O-despenthylase (PROD) after exposure to Bromkal 70, DE-71, DE-79 (pentaBDE mixtures) and PBDE 47 (Hallgren et al., 2001; Zhou et al., 2001; 2002; Hallgren and Darnerud, 2002). Other enzymes used as indicators of microsomal phase I activity were also induced by PBDEs, including benzphetamine N-demethylation, p-nitroanisole demethylase, aryl hydrocarbon hydroxylase (AHH) and benzo(a)pyrene hydroxylase (Carlson, 1980a; 1980b; Trainor et al., 2003). In a 14-day study, penta- and octa-BDE mixtures, but not deca-BDE, induced long-lasting UDPGT activity in rats. In two studies, Hallgren et al. (2001) and Hallgren and Darnerud (2002) demonstrated that Bromkal 70 and PBDE 47 also increased the UDPGT activity but to a lesser degree (Hallgren et al., 2001; Hallgren and Darnerud, 2002).

Some studies have also shown that PBDE is a potent neurotoxicant displaying effects similar to PCBs. Based on the available data; this system seems to be the most sensitive to PBDE-induced toxicity. For example, mice neonatally and perinatally exposed to PBDEs have shown impaired sensorimotor development, altered locomotor activity and delayed development of spontaneous behavior (Branchi et al., 2002; 2003; Viberg et al., 2003a; 2003b). When a single dose of PBDE 47 (0.7 or 10.5 mg/kg) or PBDE 99 (0.8 or 12 mg/kg) was administered to mice on postnatal day 10, these animals exhibited

permanent aberrations in motor behavior. Moreover, learning and memory deficits were also observed in PBDE 99-treated mice (Eriksson et al., 2001; 2002). Two studies suggest that cholinergic nicotinic receptors may also be a target of PBDEs, since a decrease in alpha-bungarotoxin binding in the hippocampus was found in mice neonatally exposed to PBDE 153 and the response of the cholinergic agent nicotine was also altered in mice neonatally exposed to PBDE 99 (Viberg et al., 2002; 2003a). In vitro, PBDEs were capable of inducing cell death in cerebellar granule cells in culture (Reistad et al., 2002) and releasing arachidonic acid via the phospholipase A_2 pathway, which is associated with learning and memory (Kodavanti et al., 2002).

The most relevant data on the toxicological effects of PBDEs were given above, leading to the conclusion that a greater effort should be made in the experimental field in the coming years. This is extremely important for identifying and assessing the risks to humans from this new environmental pollutant. The available data indicate that PBDEs display variable dose-related effects, suggesting various modes of action for this compound similar to those of PCBs.

2.2. Polychlorinated Diphenyl Ethers (PCDEs)

2.2.1. Structures and Nomenclature

PCDEs are a group of halogenated aromatic compounds. PCDEs have physico-chemical properties similar to those PCBs, PCDDs and PCDFs (Kurz and Ballschmiter, 1999). Figure 2.3 shows the structure of the PCDEs, which have one to ten chlorine atoms and a nomenclature following the numbering system of PCBs (Ballschmiter and Zell, 1980).

PCDEs are also called bis (chlorophenyl) ethers and chlorodiphenyl oxides. The empirical formula of the PCDEs is $C_{12}H_{10-n}Cl_nO$, where n = 1-10. Similarly to PCBs, there are 209 possible PCDE congeners, in which the number of chlorines varies from one to ten. The numbers of possible isomers at each chlorination degree (congener groups) from mono to decachlorinated PCDEs are presented in Table 2.3. The PCDEs congeners are numbered analogously to PCBs (Paasivirta and Koistinen, 1994) following the systematic

numbering presented by Ballschmiter and Zell (1980), except that according to the numbering rules the order of a few congeners should be different (Ballschmiter et al., 1989; Kurz, 1994). The chlorine substitution and numbering of PCDEs is presented in Table 2.1.

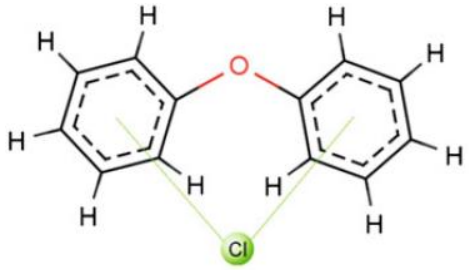


Figure 2.3 The structure of PCDEs (209 Congeners).

2.2.2. Sources and Usage

PCDEs are industrial by-products found in many ecosystems at low levels (Rosiak et al., 1997) and have been found as by-products of technical chlorophenols (Kurz and Ballschmiter, 1994) and chlorinated phenoxyacetic acids (Nevalainen and Kolehmainen, 1994). PCDEs also come from processes of incomplete combustion, fly ashes, transformer fluids, wood preservatives and flues of municipal waste incinerators (Williams et al., 1991; Koistinen et al., 1993b). PCDEs are also a group of environmental pollutants with potential toxic effects in mammals (Domingo, 2006).

PCDEs can also be formed by perchlorination of industrial or sewage effluents containing diphenyl ether (Komsta et al., 1988). Combustion is a well known source of various groups of organic compounds (Junk and Ford, 1980) and the occurrence of PCDEs in combustion wastes was reported by Paasivirta et al. (1986).

There have been many patents for PCDEs and desired applications have included use as hydraulic fluids, electric insulators, flame retardants, lubricants, and plasticizers (Koistinen, 2000).

Lower chlorinated PCDEs, mono- and di-chlorinated, have aroused interest as chemical intermediates (Hake and Rowe, 1963) and 4-mono-CDE has been desirable as a synthetic intermediate in agricultural and topical medicinal applications (Koistinen, 2000). Higher chlorinated PCDEs have been used in the electrical industry, but their toxicity hindered their use as plasticizers and in high-pressure greases (Hake and Rowe, 1963). Tetra- through hexa-CDEs was suspected of causing handling hazards, since they showed high toxicity when fed to guinea pigs.

PCDEs have also been suggested for other applications such as for pesticides and synergistics in pesticides. Dichlorodiphenyl ether was among 34 most effective materials out of 5963 tested for clothing treatments against chigger mites, but it was suspected of being unsafe for use on clothing worn for an extended period. One di-CDE showed acaricidal activity against the citrus red mite when 100 organic compounds related to DDT were tested, but no insecticidal activity against the greenhouse thrips. The toxicity of a mono-CDE was tested among 106 compounds as an insecticide to body lice, but it proved to be not very toxic. Insecticidal activity of mono- through trichlorodiphenyl ethers has not been high to house flies, but they have acted as synergists when mixed with pyrethrins. OCDE has been synergistic with insecticides such as Sumithion, DDT, malation, and bromophos (Koistinen, 2000).

PCDEs have been found nearly everywhere (Stanley et al., 1990; Huestis and Sergeant, 1992; Koistinen et al., 1995b; Koistinen et al., 1997), and the highest concentrations of these pollutants were found in fish and seafood, dairy products, meats, oils and fats, and cereals (Falco' et al., 2005). They are calling increasing attention because of their immunotoxicity similarity to PCBs (Huang et al., 2004). They can induce cytochrome P-450-dependent monoxygenase activity and have a broad spectrum of toxicity (Kurz and Ballschmiter, 1995a). PCDEs are regarded as additional persistent indicator molecules for a global pollution of the environment by organochlorine compounds (Kurz and Ballschmiter, 1999). Also, as potential persistent organic pollutants (POPs), PCDEs' risk assessment and exposure evaluation have received increasing researchers' interest.

PCDEs have been frequently detected in a wide range of environmental samples including sediments (Koistinen et al., 1995a), fishes (Koistinen et al., 1993b), white-tailed sea eagle (Koistinen et al., 1995a), foodstuffs (vegetables, tubers, pulses, cereals, fruits, fish and shellfish, meat and meat products, eggs, milk and dairy products, and fats and oils) (Bocio et al., 2004) and human adipose tissue (Stanley et al., 1991; Williams et al., 1991). The studies have shown that PCDEs have similar toxic properties to PCBs (Becker et al., 1991), and therefore they are also regarded as a type of persistent indicator molecules other than PCBs for a global environmental pollution by organochlorine compounds (Kurz and Ballschmiter, 1999). Wide distribution, high lipophilicity, and persistence of PCDEs have raised concern about their bioaccumulation, their potential biomagnification in the food webs, and their adverse effects (Domingo, 2006). Hence PCDEs have attracted great attention recently as an important type of environmental pollutants.

2.2.3. Physico-chemical Properties of PCDEs

PCDEs are similar to PCBs by their structure and physical properties. They have low water solubility and are lipophilic. PCDEs are quite resistant to degradation and are persistent in the environment. These compounds are found in sediment, mussel, fish, bird, and seal. PCDEs show biomagnification potential, since levels of PCDEs increase in species at higher trophic levels. PCDEs are also detected in human tissue. Despite the persistence and bioaccumulation, the significance of PCDEs as environmental contaminants is uncertain. The acute toxicity and Ah-receptor-mediated (aryl hydrocarbon) activity of PCDEs is low compared to those of PCDDs and PCDFs. Due to structural similarity to thyroid hormone, PCDEs could bind to thyroid hormone receptor and alter thyroid function. Furthermore, PCDEs might be metabolized to toxic metabolites. In the environment, it is possible that photolysis converts PCDEs to toxic PCDDs and PCDFs.

Properties of PCDEs, including physico-chemical ones, are not well known as the literature reviews of PCDEs have shown (Becker et al., 1991; Kurz, 1994; de Boer and Denneman, 1998). PCDEs like PCDDs, PCDFs, and related compounds, are known to be stable and resistant to breakdown by heat, hydrolysis, bases, and acids. PCBs are also quite stable to oxidation under moderate conditions (Ballschmiter et al., 1989), but there is not much data about PCDEs concerning their stability. There is some evidence that PCDEs are

resistant to bases and acids and the occurrence of PCDEs in the environment indicates that PCDEs are persistent and bioaccumulating compounds. The study of Koistinen (2000) already showed that PCDEs are quite stable, since PCDEs could be measured in chlorophenol extracts after sulfuric acid treatment. Tetra- and octa-chlorinated PCDE congeners were later proven resistant in treatment with concentrated sulfuric acid (Koistinen et al., 1993b). Chemical treatment of PCDE-containing wastes with NaClO and NaOH has not affect the content of PCDEs (Paasivirta et al., 1982). Reactions of PCDEs that are significant from the environmental point of view include photochemical reactions, thermal reactions, hydrolysis in water, and oxidation by air, but they have not been studied intensively.

Most PCDEs synthesized are solids and their vapor pressures are low (Kurz, 1994; Kurz and Ballschmiter, 1999). The vapor pressures of sub-cooled liquids (–log P_L) of PCDEs, determined from gas chromatographic (GC) data for 106 PCDE congeners at 30°C, have ranged between 0.27 and 5.80 Pa, being similar to those of PCBs (Kurz, 1994; Kurz and Ballschmiter, 1999).

Henry's law constants for PCDEs, which describe the distribution between air and water and can be calculated from vapor pressure and water solubility, increase with increasing chlorination degree (Kurz and Ballschmiter, 1999). The values of Henry's law constant for PCDEs calculated based on chromatographic data at 30°C have ranged from 2.95 Pa m^3/mol to 14125.37 Pa m^3/mol (Kurz and Ballschmiter, 1999).

Water solubility is one of the major parameters which affect the fate and distribution in the environment. Hydrophobic compounds with high octanol-water partition coefficients tend to bioaccumulate. Opperhuizen and Voors (1987) have shown that hydrophobicity of PCDEs determines the bioconcentration factor of PCDEs and that bioconcentration kinetics of PCDEs resemble those of PCBs.

The aqueous solubilities (–log S) of mono- through deca-CDEs (106 PCDE congeners) range between 4.21 mol/L and 12.95 mol/L (Kurz and Ballschmiter, 1999). These values were calculated from high performance liquid chromatography (HPLC) data at 30°C. PCDEs are more soluble in water than PCBs, which have extremely low water

solubilities. The water solubilities of PCBs range from $7x10^3$ μg/L to 0.02 μg/L (Ballschmiter et al., 1989).

Physico-chemical properties of an organic chemical compound play an important role in determining its distribution and fate in the global environment. Due to not only the time consumption and high expense, but also the unavailability of chemical standards of many PCDEs, it is very difficult to determine experimentally the physico-chemical properties and biological activity for all the PCDE congeners. Therefore, alternative approaches are needed. Many previous studies showed that it was indeed feasible to predict the properties or activities with QSAR/QSPR models for many organic compounds (Tuppurainen and Ruuskanen, 2000; Zou et al., 2005; Ghasemi and Saaidpour, 2007; Toropov et al., 2008; Puzyn et al., 2009). In fact, QSAR/QSPR studies with respect to PCDEs can also be found sporadically in recent publications (Huang et al., 2004; Chen et al., 2007; Sun et al., 2007; Zeng et al., 2007), although most of these studies aimed at only one or few properties (usually chromatographic relative retention time).

2.3. Quantitative Structure-Property Relationship (QSPR) Method

The QSPR method is based on the principle that the properties of a compound are directly related to, or at least correlated with, the molecular structure. The goal of the QSPR method is to construct a quantitative relationship between structure and one or more properties of interest based on existing data. The compounds with known property data are called the training set, and by comparing structural variation with property variation for the training set, predictive relationships can be established.

It is helpful to clarify what is meant by property and structure. The QSPR method is very general, so the properties that can be analyzed using this method include all types of bulk physical properties of compounds along with the chemical interactions they exhibit such as reactivity or biological activity. Molecular structure itself is an abstract concept, but at the most basic level it involves a symbolic representation that typically includes the chemical composition of a molecule and the connectivity (that is, what atoms are present and how are they connected to one another).

Given this general outline of the QSPR method, it can be readily seen there are two major difficulties associated with creating a functional relationship between property data and molecular structure. The first difficulty is quantifying the inherently abstract molecular structure. This problem is overcome by calculating molecular descriptors, which are developed to quantify various aspects of molecular structure. In fact, hundreds of descriptors have been created to numerically quantify a diverse array of structural features. The second difficulty is determining which structural features most influence a given property and then establishing the functional relationship that best describes the relationship between these structure descriptors and the property data. To accomplish this, a variety of statistical methods are used for descriptor selection and regression.

Once a structure-property relationship has been developed, there are two main end uses for these models. One use is to analyze the functional descriptors that are preferentially selected for the model. These descriptors indicate which structural features of a molecule most influence the property of interest and give an idea of how varying the structure can change the resulting property. This is very useful for trying to design a compound to have favorable properties. The other use is to calculate descriptors for compounds with unknown property data and use structure-property relationships for property prediction.

In the following sections, the major elements of the QSPR method are discussed. This involves a general introduction to the use of descriptors for quantifying structural features and a discussion of the statistical methods used to create structure-property models based on these descriptors.

2.3.1. Using Molecular Descriptors to Quantify Structure

In general, a descriptor is a number or set of numbers generated for a compound either by experimental determination or through a functional transformation which operates on a symbolic representation of the compound. The purpose of the descriptor is to create a numerical scale to capture the variation in some structural feature. An example of an experimentally determined descriptor is the octanol-water partition coefficient, K_{ow}, which measures the relative solubility of a compound in octanol and water phases and therefore quantifies the hydrophobicity of a compound (a very hydrophobic molecule prefers the octanol

phase and so the K_{ow} value is very high). In fact, this descriptor has found wide use in the structure-toxicity modeling community, where it has been found that toxicity is often directly correlated with the log K_{ow} parameter. An example of a computational descriptor is molecular weight, which can be calculated based solely on the composition of a molecule and quantifies the size of a compound.

As shown above, descriptors can be obtained experimentally or numerically. Since experimental descriptors can be time-consuming to determine and require that a compound be available for experimentation, much more emphasis has been placed on theoretical descriptors, which can be directly calculated using only the molecular representation of a compound's structure. Because of this, there are varying degrees of complexity for theoretical descriptors. It is useful to examine how varying levels of structural information can be used to calculate these theoretical descriptors. For instance, the lowest level of structural information would contain only the composition, such as number and type of atoms. From this information molecular weight could be calculated along with atom counts. At the next level, there might be basic connectivity information, which would allow fragment count descriptors, for example. A fragment would be a specific functional group, such as a carbonyl or methylene group, and these fragment count descriptors have found wide use in group contribution methods. By including full connectivity, the whole family of topological descriptors becomes available, which can describe general shape, size and branching of a compound. Continuing on, the three-dimensional structure of a molecule can be used for descriptor calculation, and even the electronic structure obtained from quantum calculations can be used, especially for reactivity descriptors. Therefore, it can be seen that numerical descriptors allow varying levels of structural information to be used depending upon the level of information needed.

While there is a certain amount of freedom in choosing descriptors, there are also some limitations, in the form of certain desirable characteristics that descriptors should have. If experimental descriptors are used, they should preferably be quick and inexpensive to measure. For theoretical descriptors, they should be easy to calculate and not require a lot of time for calculation. It is also preferable that descriptors be easy to interpret, meaning that a descriptor should vary in a regular way with the structural feature or features it attempts to quantify. Also, since multiple descriptors are calculated for a compound, it is also helpful if

this pool of descriptors is orthogonal, meaning the descriptors are completely uncorrelated with one another. In practice, descriptors are typically correlated with one another to varying degrees.

As shown above, there are varying levels of information that can be derived from molecular structure. There is similarly a large variety of structural characteristics that can be useful in describing different properties and correspondingly a wide array of descriptors that have been developed to quantify various structural elements. Therefore the number of descriptors that are available for calculation is extremely large. This allows a higher degree of freedom in creating structure-property relationships, but it also introduces the added need to preferentially select the most important descriptors. In the next section, the methods used to select the optimal descriptors and establish structure-property relationships are discussed.

Table 2.2 List of desirable requirements for molecular descriptors.

No.	Descriptors
1	Should have structural interpretation
2	Should have good correlation with at least one property
3	Should preferably discriminate among isomers
4	Should be possible to apply to local structure
5	Should possible to generalize to ‘higher’ descriptors
6	Descriptors should be preferably independent
7	Should be simple
8	Should not be based on properties
9	Should not be trivially related to other descriptors
10	Should be possible to construct efficiently
11	Should use familiar structural concepts
12	Should have the correct size dependence
13	Should change gradually with gradual change in structures

2.3.2. Descriptor Selection and Regression

Since the ultimate aim of a QSPR analysis is to take a set of compounds with corresponding property data and create a structure-property model, the next step after quantifying the structure for the compounds is model creation. As detailed in the previous section, it is possible to generate hundreds of descriptors to quantify the various elements of molecular structure. In many cases, these descriptors can be correlated to some degree, even being completely redundant in some cases, and many may have little to no influence over the property of interest. Therefore, methods are needed to search through this descriptor space and reduce it down to the most significant descriptors. In addition, a functional relationship between these descriptors must be developed or assumed either following descriptor selection or, more commonly, at the same time.

As a first step toward constructing a model, variable reduction techniques can be used to eliminate unnecessary descriptors. Some simple techniques include eliminating descriptors that are nearly constant for all compounds and discarding one of a pair of descriptors that are highly correlated. More involved techniques use tools such as principal component analysis or cluster analysis to eliminate descriptors while retaining the original information of the full descriptor set. Performing this initial pruning of the descriptor pool can simplify the following process of selection and model building. The selection of variables/descriptors/predictors in the model can be performed also by one of those techniques Genetic Algorithm (GA), Principal Component Analysis (PCA) or Factor Analysis (FA).

Once molecular descriptors are generated, the heuristic method in CODESSA was used to accomplish the pre-selection of the descriptors and build the linear model (Katritzky et al., 1994). Its advantages are the high speed and no software restrictions on the size of the data set. The heuristic method can either quickly give a good estimation about what quality of correlation to expect from the data, or derive several best regression models. Besides, it will demonstrate which descriptors have bad or missing values, which descriptors are insignificant (from the standpoint of a single-parameter correlation), and which descriptors are highly intercorrelated. This information will be helpful in reducing the number of descriptors involved in the search for the best QSAR/QSPR model.

First of all, all descriptors are checked to ensure: (a) that values of each descriptor are available for each structure and (b) that there is a variation in these values. Descriptors for which values are not available for every structure in the data in question are discarded. Descriptors having a constant value for all structures in the data set are also discarded. Thereafter all possible 1-parameter regression models are tested and insignificant descriptors removed. As a next step, the program calculates the pair correlation matrix of descriptors and further reduces the descriptor pool by eliminating highly correlated descriptors. All 2-parameter regression models with remaining descriptors are subsequently developed and ranked by the regression correlation coefficient, R^2. A stepwise addition of further descriptor scales is performed to find the best multi-parameter regression models with the optimum values of statistical criteria (highest values of R^2, the cross-validated R_{cv}^2, and the *F*-value).

The heuristic method usually produces correlations 2-5 times faster than other methods with comparable quality (Katritzky et al., 2001). The rapidity of calculations from the heuristic method renders it the first method of choice in practical research. Thus, in the present investigation, we used this method to select structural descriptors and build the linear model.

Typically the process of descriptor selection and model building are conducted simultaneously. To do this, a linear functional relationship between property data and descriptors is typically assumed. Given this, the goal is now to determine which subset of descriptors out of the available set best models the property data. At the same time, it is desirable to minimize the number of descriptors used and maximize the predictive ability of the model. The most direct way to do this would be to test every possible combination of descriptors; however, since the number of possible models grows exponentially with the size of the descriptor pool, this is usually impossible. For example, with a pool of 50 descriptors, there are nearly 2.5 million possible equations using one to five descriptors. In place of an exhaustive search, one popular technique is stepwise regression, where descriptors are added one by one until further addition no longer improves the model. Another popular method is the genetic algorithm approach, where a small population of equations is randomly generated and then evolved by using a set of evolution operations to finally create a population of optimal models. Finally, there are also some nonlinear methods, such as neural network algorithms,

which bypass the assumption of a linear model and perform variable selection and model development simultaneously. Overall, though, all methods focus on selecting the optimum descriptors and the optimal functional relationship using a variety of statistical techniques.

In developing QSPR/QSTR models, the approach begins with the compilation of available endpoint data sets for a variety of chemicals. If endpoint data are available for a sufficient number of chemicals, the data set is often divided into a training set used in the model development, and a test set containing chemicals not used in the derivation of the model but used to evaluate the model. The method used for splitting the data set should be clear in proposed model. Methods available include those based on similarity analysis, for example, D-optimal distance (Hasegawaa and Funatsub, 1998), Kohonen map or self-organizing map (SOM) (Vracko et al., 2006), the *k*-means cluster analysis (Caballero and Fernandez, 2006), sphere-exclusion algorithms (Golbraikh et al., 2003) or random selection through activity sampling (Gramatica, 2007).

Since structure-property models are often developed for prediction of new compounds based on available data, predictive capability is an important criterion in the creation and assessment of QSPR models. The first indicator of predictive ability is the ability of the model to predict property data for the training set.

However, since the training set data are used in the model creation, the model is optimized to predict training set data, so this gives little indication of how purely predictive the model is. One way to provide a better predictive estimate is to leave out of the data set a small set of validation compounds. After a model has been developed using the training set, the model is used to predict data for the validation set compounds. If the predictions for the validation compounds are significantly worse than for the training set, this indicates poor predictive ability.

Another method used to assess predictive ability using only the training data is the leave-one-out (LOO) cross-validation method. In this method, a reduced training set is created by eliminating one compound. A new model is developed using the reduced training set and used to predict the left out compound. This procedure is performed for every compound and the resulting predicted vs. experimental correlation is a numerical measure of

predictive ability. As with the validation set method described above, if the cross-validated predictive ability is significantly lower than for the full training set, the predictive ability is poor. For large data sets, the LOO method can be too optimistic since the reduced training set is not a significant perturbation from the original. In this case, leave-more-out techniques can be used, which remove multiple compounds to create the reduced training set. There are additional techniques which can be used to assess predictive ability. In many cases, these predictive measures can be used during model creation to not only optimize the ability to model the training set but also to optimize the predictive ability of the model.

The final result of the QSPR method is a model that uses the most important descriptors to optimally describe the relationship between property data and molecular structure for training set compounds. Additionally, this model is designed to predict property data for compounds outside of the training set. However, for compounds that are structurally very dissimilar to the training set compounds, the prediction is less likely to be accurate.

2.4. QSPR/QSTR Studies on Physico-chemical Properties and Toxicity of PBDEs/PCDEs

QSPR/QSAR/QSTR studies are expected to reduce the cost and the number of organisms used for toxicity testing and to fill the existing data gaps within the REACH regulatory framework in the EU. Many QSTR studies in environmental chemistry and ecotoxicology are carried out with different types of descriptors using statistical methods like regression analysis (Cronin et al., 2004; Saçan et al., 2007), partial least squares (PLS) (Roy and Gosh, 2006; 2007) and artificial neural network (ANN) (Roy and Roy, 2009) for diverse set of chemicals.

The octanol-air partition coefficient (K_{oa}) is useful for predicting the partitioning behavior of organic compounds between air and environmental matrices such as soil, vegetation, and aerosol particles. At present, experimentally determined K_{oa} values are available for only several hundred compounds. Therefore, the ability to estimate K_{oa} is necessary for screening level evaluation of most chemicals. Although it is possible to estimate K_{oa} from the octanol-water partition coefficient (K_{ow}) and Henry's law constant

(*H*) various concerns have been raised in prediction of this property from molecular structure.

Various models utilize K_{oa} to screen and rank chemicals for environmental persistence and long-range transport (Webster et al., 1998; Gouin et al., 2000; Bennett et al., 2001). The molecular weights of PCDDs, PCDFs, and PBDEs have been shown to correlate well with the corresponding log K_{oa} values (Xiao and Wania, 2003). Accurate QSPRs have been developed for PCBs, Polychlorinated naphthalenes (PCNs), PCDDs, PCDFs, PBDEs and other similar compound classes using combined molecular descriptors such as polarizability, dipole moment, electronic energy, core repulsion energy, and various molecular orbital descriptors (HOMO, LUMO, etc.) (Chen et al., 2001; 2002a; 2002b; 2003).

Wania et al. (2002) used relative gas chromatographic retention times on a nonpolar stationary phase to determine log K_{oa} of nonpolar semi-volatile organic compounds, including PBDEs. Additionally, Harner and Shoeib (2002) measured K_{oa} values for 13 PBDEs at temperatures ranging from 15 to 45°C. It is apparent that the data is not enough since K_{oa} values for the other 196 PBDEs are lacking. Experimental determination of K_{oa} is costly in expenditures such as equipment and time. Furthermore, because of the limited number of standard PBDE congeners, it seems impossible to measure K_{oa} values for all the other PBDEs. Thus, it is the purpose of this study to develop QSTR models for PBDEs, based on the reported experimental values.

PLS regression was used to develop models for log K_{oa} of PBDEs at different environmental temperatures (T) by Chen et al. (2003). The optimal model contains nine theoretical molecular descriptors and $1/T$ as predictor variables. Intermolecular dispersive interactions play a leading role in governing the magnitude of log K_{oa}.

Zhao et al. (2005) proposed a QSPR approach to predict log K_{oa} for 13 PBDEs using molecular connectivity indexes which are topological descriptors of molecular structure based on a count of skeletal atom groupings. The stepwise multiple linear regression was used to derive equation with correlation coefficients with 0.96 for PBDEs with descriptors, χ_p and χ_{vp} .

Xu et al. (2007) reported linear relationships between log K_{oa} and log *RBA* of PBDEs and the structural descriptors. For log K_{oa}, they obtained a model with 2 descriptors and $R^2 = 0.9761$. For log *RBA*, they obtained a model with 4 descriptors and $R^2 = 0.647$. It has been shown that structural descriptors derived from molecular electrostatic potentials together with molecular volume can be well used to express the quantitative structure – property relationships of PBDEs. As for the prediction of biological activity or toxicity, this parameter set does not offer significant advantage over other kinds of descriptors due to inherent limitations. The QSAR model for the Ah receptor binding affinity (*RBA*) is relatively poor, which can be ascribed to the complexity of factors which affect biological activity and the limitations of the present parameter set in describing steric characters of the molecule.

Wang et al. (2008) modeled log K_{oa} values of PBDEs in their study. The molecular geometries of 209 PBDEs were optimized at the B3LYP/6-31G* level with Gaussian 98 program. The calculated structural parameters were taken as theoretical descriptors to establish one novel QSPR models for predicting octanol/air partition coefficients (K_{oa}) of PBDEs based on the theoretical linear solvation energy relationship (TLSER) model. The model achieved in their work contains three variables: most negative atomic partial charge in molecule (q^-), dipole moment of the molecules (μ) and mean molecular polarizability (α), of which R^2 value is as high as 0.997, its root-mean-square errors in modeling (*RMSE*) is 0.062.

Papa et al. (2009) modeled log K_{oa} and log K_{ow} values of PBDEs. The molecular descriptors (0D, 1D, 2D, 3D) for the studied compounds were computed by the software DRAGON 5.4 (Ver. 5.4 for Windows, 2006). In addition, four quantum-chemical descriptors were included in their modeling procedure: highest occupied molecular orbital (HOMO), lowest unoccupied molecular orbital (LUMO), HOMO-LUMO gap, and ionization potential. Linear regression models were performed by ordinary least squares (OLS) regression using the software MOBYDIGS (Ver. 1.0 beta for Windows, 2004). Due to the limited dimension of the studied data sets, the All Subset Models selection method was applied in their study to select the best single descriptor-based models. All the models were internally validated by the LOO procedure (Q^2_{LOO}), and the robustness of the models were further evaluated by the bootstrap method (Q^2_{BOOT}). To verify the real predictive

power of the developed models external validation was performed in the presence of a sufficient number of experimental data (n_{obj} > 20). This procedure differs from the literally meant 'external validation', which is based on the use of 'new measured' data as prediction set. Their approach was different simply for the fact that the prediction set is chosen, before the modeling step, by splitting the available experimental data. Therefore, the 'sampled' prediction set is completely excluded from the development of QSPR models, which are performed, and internally validated, only on the basis of the complementary pool of data (training set). The real predictive power of these models is then externally checked on the prediction set chemicals. T(O...Br), was selected as the best modeling variable for three different properties which are strongly related to each other log $1/P_L$, log K_{oa}, log K_{ow}. This descriptor gives double structural information: its values increases according to both the number and the distance of bromine substituents, on each phenyl ring, from the oxygen ether. Thus, T(O...Br) takes also into account the information related to the position of the bromine atoms on the phenyl rings (R^2 for log K_{oa} 0.973; for log K_{ow} 0.964).

Of the physico-chemical parameters indispensable to the fate assessment of persistent organic pollutants, the octanol–water partition coefficient (K_{ow}) indicates the environmental behavior of hydrophobic organic compounds (Sabljic et al., 1995). The K_{ow} is defined as the ratio of a compound's concentration in octanol to its concentration in water after the partition between two phases reaches equilibrium at a specified temperature. Although other partition coefficients such as octanol–air partition coefficient were discovered to influence the Bioconcentration Factors (BCFs) of organic chemicals recently, K_{ow} still plays an important role in governing BCFs (Streets et al., 2006; Kelly et al., 2007). Thus, it is important to obtain more information about K_{ow} of PBDEs/PCDEs. Several reports have appeared on predicting log K_{ow} for PBDEs and PCDEs separately.

Two one-descriptor models were established by Wania and Dugani (2003) and Braekevelt et al. (2003) at 25°C for log K_{ow} of PBDEs. Single descriptor was molecular mass in the former model, whereas the number of Br atoms was the descriptor appearing in the latter model. Li et al. (2008) used PLS regression method for modeling the log K_{ow} values of PBDEs based on quantum molecular descriptors. The values of log K_{ow} for PBDEs are mainly governed by molecular surface area, energy of the lowest unoccupied molecular orbital and the net atomic charges on the oxygen atom. All these descriptors

have been discussed to interpret the partitioning mechanism of PBDE chemicals. The bulk property of the molecules represented by molecular surface area is the leading factor, and K_{ow} values increase with the increase of molecular surface area. In conclusion it was stated that higher energy of the lowest unoccupied molecular orbital and higher net atomic charge on the oxygen atom of PBDEs result in smaller K_{ow}. The energy of the lowest unoccupied molecular orbital and the net atomic charge on PBDEs oxygen also play important roles in affecting the partition of PBDEs between octanol and water by influencing the interactions between PBDEs and solvent molecules.

Kurz and Ballschmiter (1999) synthesized 106 PCDEs along with diphenyl ether at the laboratory and measured the log K_{ow} values by HPLC method. Yang et al. (2003) performed PLS regression to build the QSPR model for log K_{ow} values of PCDEs by using 12 quantum chemical descriptors. Sun et al. (2007) developed QSPR Model to predict log K_{ow} values of PCDEs by using MLR method with molecular electronegativity distance vector (MEDV-4) descriptor which is a structural descriptor. In their method, they did not separate the data set (107) into training and test set. Their conclusion in their study was log K_{ow} increase with the degree of chlorination in general. In another study, Chen et al. (2007) modeled log K_{ow} values of PCDEs by the method of Cl substitution position. Stepwise MLR has been used to construct the QSPR models by using six elements (the numbers of positions of Cl substitution (N_{PCS}). Xu et al. (2010) obtained model for log K_{ow} with 3 descriptors with the $R^2 = 0.974$; $R_{cv}^2 = 0.972$ without dividing data set into training and test set ($N = 107$); and $R^2 = 0.976$; $R_{cv}^2 = 0.972$ by division the data set into training and test set ($N = 72$). However, no QSPR study on log K_{ow} for the combination of these two sets (PBDEs/PCDEs) has been reported so far. It is reasonable to derive a new model combining these two chemical classes. The purpose of this study is to develop a comprehensive and explainable K_{ow} prediction model for all 209 PBDEs and PCDEs along with the diphenyl ether, based on the experimental data.

Three-dimensional quantitative structure activity relationships (3-D-QSAR) models, using comparative molecular field analysis (CoMFA) and comparative similarity indices analysis (CoMSIA), were built based on calculated structural indices and a reported experimental toxicology index (aryl hydrocarbon receptor (AhR) relative binding affinities

(*RBA*)) of 18 PBDEs congeners, to determine the factors required for the *RBA* of these PBDEs (Wang et al., 2005). The results showed clearly that the nonplanar conformations of PBDEs result in the lowest energy level and that the electrostatic index was the main factor reflecting the *RBA* of PBDEs. The two QSAR models were then used to predict the *RBA* value of 46 PBDEs for which experimental values are unavailable at present.

Wang et al. (2006) constructed QSAR models based on 406 descriptors for the log *RBA* of 18 PBDE congeners. The best regression model involved four descriptors, which were related to the conformational changes, atomic reactivity, molecular electrostatic field, and non-uniformity of mass distribution in a molecule of PBDEs.

Zheng et al. (2007) used PLS method for predicting log *RBA* of PBDEs. Using Support Vector Machines (SVM) and Radial Basis Function Neural Networks (RBFNs), the log *RBAs* of 15 PBDE congeners have been correlated with the extracted seven quantum chemical descriptors. The SVM models generalize better than the RBFN models. The good performance of the QSAR models based on net atomic charges suggests that electrostatic interactions may play important roles in the log *RBA* of PBDEs.

With quantum chemical computation of density functional theory (DFT), the electronic properties including the polarisabilities, polarisability anisotropies and quadrupole moments of a total of 209 congeners of PBDEs were evaluated (Gu et al., 2009). The electronic properties were shown to be highly dependent on the bromination pattern, i.e. their values changed sensitively with the number and sites of bromination. Some of electronic properties were found to be potent in explaining the variance of toxicity, and the potency was verified by the development of QSARs. To further improve the stability and predictability of QSARs for toxicity, two-dimensional topological indices were introduced. In QSARs, polarisability anisotropy was more significant than other polarisability tensors, indicating the implicit occurrence of dispersion interaction between the ligand and AhR. For PBDEs, the quadrupole moment was as significant as shown previously for dioxins. QSARs were developed from MLR analysis to test the validity and significance of electronic properties in correlation with the relative binding affinities of PBDEs. Polarisability anisotropy ($\Delta\alpha$), polarisability tensor (αxy), SIC which is a topological descriptor significantly entered the equation they obtained. They concluded

that with quantum chemical computation of all PBDE congeners by DFT, the variation of such important electronic properties as the polarisability, polarisability anisotropy and quadrupole moment has been clearly illustrated with a strong dependence on the bromination pattern. On the basis of bromination pattern dependence, the role of electronic properties in explaining the toxicity variance of PBDE congeners has been confirmed by the development of satisfactory QSAR models, in which either the polarisability anisotropy or the quadrupole moment tensor has proved to be more statistically significant and favourable to improving the model's stability and predictability. The superior performance of the chosen descriptors in explaining the toxicity variance of PBDEs implies that the dispersion and electrostatic interactions are important and necessary to the binding affinities of AhR.

Papa et al. (2010) used DRAGON software to calculate the descriptors and to model log *RBA* of PBDEs. In addition, four quantum-chemical descriptors, calculated by the HYPERCHEM program, were included in the modeling procedure: highest occupied molecular orbital (HOMO), lowest unoccupied molecular orbital (LUMO), HOMO-LUMO gap, and ionization potential. Multiple linear regression analysis and variable selection were performed by the software MOBY-DIGS using, respectively, the ordinary least squares (OLS) regression and the All Subset Selection method that verifies the modeling ability of all of the possible combinations of the available descriptors. To generate QSAR models that are also able to give reliable prediction for new chemicals (compounds not participating in model development), external validation was performed (Tropsha et al., 2003; Gramatica, 2007). Thus, when a sufficient number of experimental data were available ($n_{obj} > 16$), the original data sets were preliminarily split into a training set, on which the models were developed, and a prediction set, on which the developed models were verified. The descriptors involved in the equation were L1v and Mor22u with $R^2 = 0.9$.

3. MATERIALS AND METHODS

3.1. Data Sets

The general steps of constructing a QSPR model include compilation of experimental data of chemical property of interest, selection of descriptors that influence the target physico-chemical properties, calculation of the derived molecular structural descriptors of compounds, development of the QSPR model, and validation of the developed model.

Experimental data on 22 *n*-octanol/air partition coefficients (log K_{oa}) at 298.15 K were taken from Wania et al. (2002) and Harner and Shoeib (2002) and the values range from 7.24 (2-Mono-BDE) to 12.15 (2,2′,4,4′,5,5′-Hexa-BDE).

The experimental log K_{ow} values for 107 PCDEs were mainly obtained from the investigations performed by Kurz and Ballschmiter (1999). The experimentally determined log K_{ow} values for 14 PBDEs were taken from different literatures (Darnerud et al., 2001; Braekevelt et al., 2003; Hayward et al., 2006). log K_{ow} values of the PBDEs and PCDEs range from 5.51 to 10.00 and 3.97 to 8.16, respectively.

Eighteen PBDE congeners that were synthesised and tested for their AhR binding affinities (log *RBA*) in rat hepatocytes were taken from the literature (Chen et al., 2001) and used for the structure based QSAR study. The binding affinities were calculated as the ratios of EC_{50} values for AhR binding of individual congeners to that of the reference compound, namely [^{3}H]-TCDD in 1.0 nM, where EC_{50} was the molar concentration of chemical necessary to inhibit 50% of the specific binding of radio-labelled TCDD. To facilitate the analysis, the ratios could be expressed as the negative of logarithm range -log *RBA*, i.e. p*RBA*. It is obvious that the selected data span a range of at least 3-log units, which guarantees a broad and homogenous dataset (McKinney et al., 2000).

3.2. Calculation and Selection of Molecular Descriptor

The main step in every QSAR/QSPR/QSTR study is to calculate and select the structural descriptors in which they encode numerical parameters representing the chemical structures.

The descriptors were calculated from the energy-minimized three-dimensional conformations. Semi-empirical PM3 method (Stewart, 1989a; 1989b) was used to optimize the conformations by including the effect of an aqueous solvent in SPARTAN 06 software package (Wavefunction, 2006).

Aqueous solvation energies of all the conformers were estimated by using the SM 5.4 model (Chambers et al., 1996) and added to the gas phase energies. Based on these optimized geometries, for all 209 PCDEs and 209 PBDEs along with diphenyl ether calculations of semi-emprical molecular descriptors were performed for the conformer having the lowest total energy in aqueous phase. The molecular descriptors obtained from SPARTAN software are the energy values of the highest occupied molecular orbital (E_{HOMO}) - eV, the energy values of the lowest unoccupied molecular orbital (E_{LUMO}) - eV, gas-phase energy (E) - eV, aqueous-phase energy (E_{aq}) - eV, dipole moment (μ) – debye and standard heat of formation (ΔH_f). Using these quantum-chemical parameters additional variables such as E_{LUMO}-E_{HOMO} gap; quantum chemical indices of hardness (η); softness (S); electro-negativity (χ); and electrophilicity (ω) were calculated according to the equations proposed by Thanikaivelan et al., (2000).

Optimized SPARTAN structures were saved as mol2 file. The descriptor values from SPARTAN and DRAGON were saved as text file to be loaded to CODESSA 2.2 software package. Molecular structures with mol2 extension were converted into MDL mol files with ChemBio3D Ultra 12.0 software (CambridgeSoft, 2010) to be loaded into CODESSA software for the calculation of CODESSA descriptors (Katritzky, 1994).

DRAGON 5.4 software (Talete, 2006) was used to calculate the theoretical molecular descriptors, belonging to 20 different types of theoretical descriptors for each molecule.

Constitutional descriptors are basically related to the number of atoms and bonds in each molecule. Topological descriptors include valence and non-valence molecular connectivity indices calculated from the hydrogen-suppressed formula of the molecule, encoding information about the size, composition and the degree of branching of a molecule. The topological descriptors describe the atomic connectivity in the molecule. The geometrical descriptors describe the size of the molecule and require 3D-co-ordinates of the atoms in the given molecule. The electrostatic descriptors reflect characteristics of the charge distribution of the molecule. The quantum chemical descriptors include information about binding and formation energies, partial atom charge, dipole moment, and molecular orbital energy levels (Lu et al., 2007).

DRAGON 5.4 can calculate a total of 1664 descriptors; however, this number decreases depending on the structure of the molecules. In this study, the number of descriptors calculated was 1162 and 1172 for PCDEs and PBDEs, respectively.

Additionally, the Characteristic Root Index (*CRI*) which is a hybrid of path-type valence connectivity index and distance matrix was included in descriptor pool. The *CRI* is calculated from the structure of the compound. The calculation of the *CRI* starts from the hydrogen suppressed skeleton of a molecule. First, each non-hydrogen atom is assigned a delta value, which is calculated from their electronic configuration by equation 3.1.

$$\delta^{\upsilon} = \frac{(Z^{v} - h)}{(Z - Z^{v} - 1)} \qquad (3.1)$$

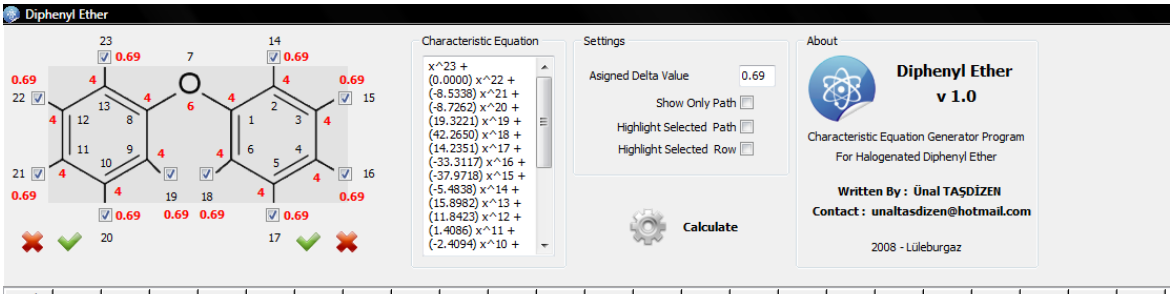
Z^{v} is the number of valance electrons in an atom, Z is its atomic number and h is the number of hydrogen atoms bound to the same atom.

Secondly, the structural graph, G, of a hydrogen suppressed skeleton of a molecule is drawn. Graph, G, is defined as a square matrix with the entries, $w_{ij,}$ representing the weighted distance traversed in moving from vertex *i*, to vertex *j* in G and calculated using atomic δ^{υ} values reported by Hall and Kier (2001). Thus, the constructed matrix comprises all possible orders of path-type valence connectivity index for a molecule, except zero order.

$$w_{ij} = (\delta_i^{\upsilon} . \delta_j^{\upsilon} \ldots\ldots \delta_n^{\upsilon})^{-1/2} \qquad (3.2)$$

where i,j,……,n correspond to consecutive non-hydrogen atoms.

Each non-hydrogen atom is assigned a value which is equal to the number of non-hydrogen atoms or is the difference between outer shell electrons and the number of hydrogen atoms attached to that atom. The entries, w_{ij}, of the matrix are calculated by considering the shortest path to any other non-hydrogen atoms. In case of equal paths ($w_{ij} = w_{ji}$) clockwise direction is chosen. Using Bocher's formula (Istefanopulos, 1981) which states that the sum of the diagonal elements of a square matrix is equal to the sum of its characteristic values, a computer program written in Delphi 2007 by Ünal TAŞDİZEN for a personal computer was used to convert the constructed matrix into characteristic polynomial form (Figure 3.1). Then the *CRI* was obtained by summing up the positive characteristic roots of the polynomial calculated with Scientific Workplace 3.0 (MacKichan Software, Inc.).



From/To	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
1	0	0.2500	0.1250	0.0625	0.1250	0.2500	0.2041	0.1021	0.0510	0.0255	0.0128	0.0255	0.0510	0.3010	0.1505	0.0752	0.1505	0.3010	0.0614	0.0307	0.0154	0.0307	0.0614
2	0.2500	0	0.2500	0.1250	0.0625	0.1250	0.1021	0.0510	0.0255	0.0128	0.0064	0.0128	0.0255	0.6019	0.3010	0.1505	0.0752	0.1505	0.0307	0.0154	0.0077	0.0154	0.0307
3	0.1250	0.2500	0	0.2500	0.1250	0.0625	0.0510	0.0255	0.0128	0.0064	0.0032	0.0064	0.0128	0.3010	0.6019	0.3010	0.1505	0.0752	0.0154	0.0077	0.0038	0.0077	0.0154
4	0.0625	0.1250	0.2500	0	0.2500	0.1250	0.0255	0.0128	0.0064	0.0032	0.0016	0.0032	0.0064	0.1505	0.3010	0.6019	0.3010	0.1505	0.0077	0.0038	0.0019	0.0038	0.0077
5	0.1250	0.0625	0.1250	0.2500	0	0.2500	0.0510	0.0255	0.0128	0.0064	0.0032	0.0064	0.0128	0.0752	0.1505	0.3010	0.6019	0.3010	0.0154	0.0077	0.0038	0.0077	0.0154
6	0.2500	0.1250	0.0625	0.1250	0.2500	0	0.1021	0.0510	0.0255	0.0128	0.0064	0.0128	0.0255	0.1505	0.0752	0.1505	0.3010	0.6019	0.0307	0.0154	0.0077	0.0154	0.0307
7	0.2041	0.1021	0.0510	0.0255	0.0510	0.1021	0	0.2041	0.1021	0.0510	0.0255	0.0510	0.1021	0.1229	0.0614	0.0307	0.0614	0.1229	0.1229	0.0614	0.0307	0.0614	0.1229
8	0.1021	0.0510	0.0255	0.0128	0.0255	0.0510	0.2041	0	0.2500	0.1250	0.0625	0.1250	0.2500	0.0614	0.0307	0.0154	0.0307	0.0614	0.3010	0.1505	0.0752	0.1505	0.3010
9	0.0510	0.0255	0.0128	0.0064	0.0128	0.0255	0.1021	0.2500	0	0.2500	0.1250	0.0625	0.1250	0.0307	0.0154	0.0077	0.0154	0.0307	0.6019	0.3010	0.1505	0.0752	0.1505
10	0.0255	0.0128	0.0064	0.0032	0.0064	0.0128	0.0510	0.1250	0.2500	0	0.2500	0.1250	0.0625	0.0154	0.0077	0.0038	0.0077	0.0154	0.3010	0.6019	0.3010	0.1505	0.0752
11	0.0128	0.0064	0.0032	0.0016	0.0032	0.0064	0.0255	0.0625	0.1250	0.2500	0	0.2500	0.1250	0.0077	0.0038	0.0019	0.0038	0.0077	0.1505	0.3010	0.6019	0.3010	0.1505
12	0.0255	0.0128	0.0064	0.0032	0.0064	0.0128	0.0510	0.1250	0.0625	0.1250	0.2500	0	0.2500	0.0154	0.0077	0.0038	0.0077	0.0154	0.0752	0.1505	0.3010	0.6019	0.3010
13	0.0510	0.0255	0.0128	0.0064	0.0128	0.0255	0.1021	0.2500	0.1250	0.0625	0.1250	0.2500	0	0.0307	0.0154	0.0077	0.0154	0.0307	0.1505	0.0752	0.1505	0.3010	0.6019
14	0.3010	0.6019	0.3010	0.1505	0.0752	0.1505	0.1229	0.0614	0.0307	0.0154	0.0077	0.0154	0.0307	0	0.3623	0.1812	0.0906	0.1812	0.0370	0.0185	0.0092	0.0185	0.0370
15	0.1505	0.3010	0.6019	0.3010	0.1505	0.0752	0.0614	0.0307	0.0154	0.0077	0.0038	0.0077	0.0154	0.3623	0	0.3623	0.1812	0.0906	0.0185	0.0092	0.0046	0.0092	0.0185
16	0.0752	0.1505	0.3010	0.6019	0.3010	0.1505	0.0307	0.0154	0.0077	0.0038	0.0019	0.0038	0.0077	0.1812	0.3623	0	0.3623	0.1812	0.0092	0.0046	0.0023	0.0046	0.0092
17	0.1505	0.0752	0.1505	0.3010	0.6019	0.3010	0.0614	0.0307	0.0154	0.0077	0.0038	0.0077	0.0154	0.0906	0.1812	0.3623	0	0.3623	0.0185	0.0092	0.0046	0.0092	0.0185
18	0.3010	0.1505	0.0752	0.1505	0.3010	0.6019	0.1229	0.0614	0.0307	0.0154	0.0077	0.0154	0.0307	0.1812	0.0906	0.1812	0.3623	0	0.0370	0.0185	0.0092	0.0185	0.0370
19	0.0614	0.0307	0.0154	0.0077	0.0154	0.0307	0.1229	0.3010	0.6019	0.3010	0.1505	0.0752	0.1505	0.0370	0.0185	0.0092	0.0185	0.0370	0	0.3623	0.1812	0.0906	0.1812
20	0.0307	0.0154	0.0077	0.0038	0.0077	0.0154	0.0614	0.1505	0.3010	0.6019	0.3010	0.1505	0.0752	0.0185	0.0092	0.0046	0.0092	0.0185	0.3623	0	0.3623	0.1812	0.0906
21	0.0154	0.0077	0.0038	0.0019	0.0038	0.0077	0.0307	0.0752	0.1505	0.3010	0.6019	0.3010	0.1505	0.0092	0.0046	0.0023	0.0046	0.0092	0.1812	0.3623	0	0.3623	0.1812
22	0.0307	0.0154	0.0077	0.0038	0.0077	0.0154	0.0614	0.1505	0.0752	0.1505	0.3010	0.6019	0.3010	0.0185	0.0092	0.0046	0.0092	0.0185	0.0906	0.1812	0.3623	0	0.3623
23	0.0614	0.0307	0.0154	0.0077	0.0154	0.0307	0.1229	0.3010	0.1505	0.0752	0.1505	0.3010	0.6019	0.3010	0.0185	0.0092	0.0185	0.0370	0.1812	0.0906	0.1812	0.3623	0

Figure 3.1 The output of characteristic root equation generator program for Polyhalogenated Diphenyl Ether written by our group.

The *CRI* has also been reviewed by Todeschini et al. (2009) as below.

The sum of the positive eigenvalues λ of the → χ matrix, based on the path connectivities calculated by the → valence vertex degree δ^v of the atoms in the path:

$$CRI = \sum_m \lambda m \qquad \text{for} \quad \lambda m > 0 \tag{3.3}$$

The *CRI* descriptor encodes information about all connectivities in the H-depleted molecular graph and is sensitive to the presence of heteroatoms in the molecule (Saçan and Inel, 1993; Saçan and Inel, 1995; Saçan and Balcioglu, 1996).

In the present work, the heuristic method (HM) algorithm running in CODESSA 2.2 was used for the selection of the most relevant descriptors from the descriptors pool. The program calculates all correlations between individual descriptors and property (log K_{ow}) and eliminates descriptors. The HM for descriptor selection proceeds with a pre-selection of descriptors by sequentially eliminating descriptors that do not match any of the following criteria: (i) Fisher *F*-criteria (F-test) greater than 1.0; (ii) R^2 value less than a value defined at the start (0.1); (iii) the Student's *t*-criterion (t-test) less than a defined value (0.1); and (iv) duplicate descriptors having a higher squared inter-correlation coefficient than 0.8 "a predetermined level" (retaining the descriptor with higher R^2 with reference to the property). The next step involves correlation of the given property with (i) the top descriptor in the above list with each of the remaining descriptors, and (ii) the next one with each of the remaining descriptors, etc. The best pairs, as evidenced by the highest F-values in the two parameter correlations, are chosen and used for further inclusion of descriptors in a similar manner.

The goodness of the correlation is tested by the correlation coefficient (R^2), the F-test (*F*), and the squared standard error (s^2). The stability of the correlations was tested against the cross-validated coefficient, R^2_{cv} , or adjusted squared correlation coefficient, R^2_{adj}, these two parameters can avoid the over training of the model. The heuristic method's advantages are the high speed and no restrictions on the size of the data set. The heuristic method can either quickly give a good estimation about what quality of correlation to expect from the data, or derive several best regression models. Besides, it will demonstrate

descriptors which have bad or missing values, which are insignificant and which are highly inter-correlated. This information will be helpful in reducing the number of descriptors involved in the search for the best QSAR/QSPR model. The heuristic method usually produces correlations 2-5 times faster than other methods with comparable quality. The rapidity of calculations from the heuristic method makes it the first choice of method for building the best QSAR models.

The descriptors that remain are then listed in decreasing order of correlation coefficients when used in global search for 2-parameter correlations. Each significant 2-parameter correlation by *F*-criteria is recursively expanded to an *n*-parameter correlation till the normalized *F*-criteria remains greater than the startup value. The best correlations up to 5-descriptor sorted by R^2, as well as by *F*-criterion, are saved.

Heuristic methods successfully solve the initial selection problem by reducing the number of pairs of descriptors in the "starting set". The major limitations of these methods are the pairwise selection on the first step and the low consistence of the presentation of the upper (according to the selected criteria) segment of the search (N in both cases is 400) due to the small size of the correlation selection.

A flowchart of QSPR method employed in this thesis for log K_{ow} constant is given in Figure 3.2 as a representation.

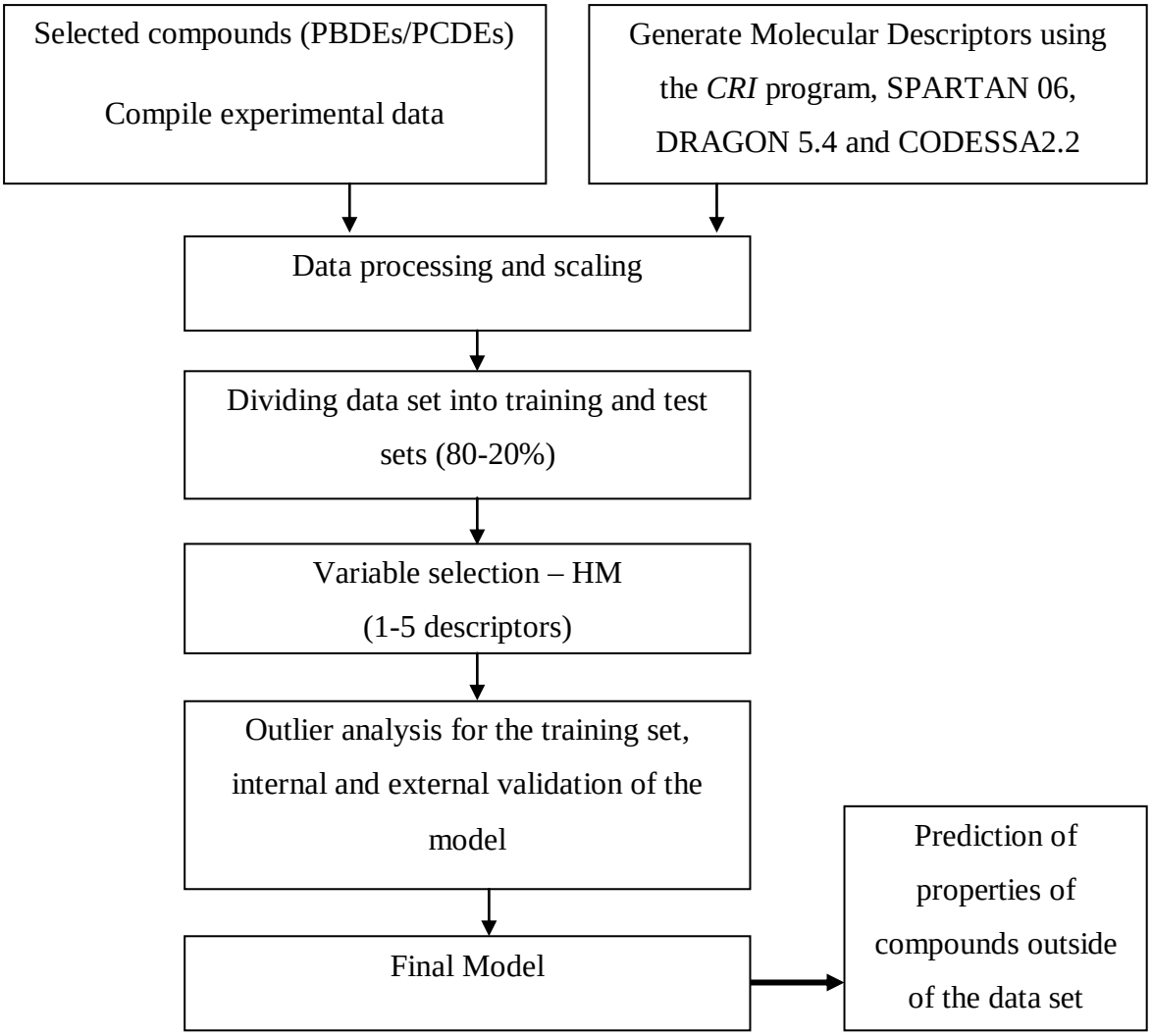


Figure 3.2 Flowchart of QSPR method

3.3. Model Development and Validation

QSPR models for each physico-chemical parameter for PBDEs and PCDEs were developed using MLR considering the Organisation for Economic Cooperation and Development (OECD) principles.

The OECD principles of QSAR validation give five basic elements for a reliable model.

1. a defined endpoint;
2. an unambiguous algorithm;
3. a defined domain of applicability;
4. appropriate measures of goodness-of-fit, robustness, and predictivity;
5. a mechanistic interpretation, if possible.

The OECD principles of QSAR validation give four basic elements for a reliable model. According to Principle 4, a QSAR model should have appropriate measures of goodness-of-fit, robustness, and predictivity. While the internal performance of a model determined by using a training set, the predictivity is determined by using an appropriate test set (OECD, 2007). Therefore, before performing modeling, the data set was divided into training and test set. The data sets were split into training, for model development, and test (prediction) sets, for predictivity check, in two different ways: a) by random selection of response values (for log K_{oa}), and b) by structural similarity verified by using Self-organized map (SOM, i.e., Kohonen Neural Network) in case of large data set (for log K_{ow}). Kohonen networks are able to select a meaningful training set and a representative test set. Kohonen networks have been adequately explained by Zupan and Gasteiger (1999) and Devillers (1996). Kohonen networks project multi-dimensional space onto 2D array of neurons. The projection, which is called learning of network, runs in two steps. In the first step, an object (represented by a vector) is presented to all neurons and the algorithm selects the neuron that is most similar to it. The selected neuron is called "winning neuron". In the second step, the weights of the winning neuron are modified to the vector values and in the same time the neighboring neurons are modified to become similar to it (Vracko, 2005). We used different networks for log K_{ow} model and approximately 80% of the data set was allocated for training set. For log *RBA* we did not use division of the data set into training and test set, since the data set is small ($n < 20$).

3.3.1. Multiple Linear Regression

Once descriptors were generated, a forward stepwise regression method was used to develop the linear model of the property of interest, which takes the form below:

$$Y = b_0 + b_1X_1 + b_2X_2 + ... + b_nX_n \quad (3.4)$$

In this equation, Y is the property, that is, the dependent variable, X_1–X_n represent the specific descriptor, while b_1–bn represent the coefficients of those descriptors, and b_0 the intercept of this equation. The LOO cross-validation method was used to evaluate the modeling ability of the model.

MLR analysis was carried out using the stepwise strategy in the Statistical Package for Social Scientists (SPSS® 17.0) for Windows (SPSS Inc., 2008). For each regression, the following descriptive information was provided: number of observations used in the analysis (n), standard error of estimate (SE), square of the correlation coefficient adjusted for degrees of freedom (R^2) and Fisher's criterion (F).

The quality of the model was considered as statistically satisfactory on the basis of the number of compounds (n), squared correlation coefficient (R^2), standard error of estimate (SE), t-values (to compare the importance of the descriptors appearing in the model) and F values when all the variables in the final model are significant at the 95% confidence level. Colinearity between variables in the model was tested by variance inflation factor (VIF) using SPSS® 17.0. The VIF is defined as $1/(1-r^2)$, where r is the correlation coefficient of one independent variable against others. Large VIF values imply strong correlation between variables in the model. To safeguard against chance correlations, another recommended criterion is that the ratio of training set compounds to descriptors in the QSPR should be at least 5:1 (Topliss and Costello, 1972).

Furthermore, to compare the predictive performance of the all models developed for training and test sets in this study with those of the literature models, both the root mean square error (*RMSE*) values was compared. Formula for *RMSE* is given in equation 3.5.

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(\hat{y}_i - y_i)^2}{n}} \quad (3.5)$$

where, n is the number of compounds, $\hat{y}_i$ is predicted and y_i is observed physico-chemical value.

Internal validation of all models was tested with the LOO cross validation R^2_{cv} procedure. According to the literature (R^2_{cv}) should be higher than 0.50 for obtaining a validated model (Golbraikh and Tropsha, 2002). Additionally, leave-many-out (LMO) procedure was performed for MLR models using Weka 3.6.1 (2009) software.

The robustness of MLR models was also tested by response randomization (*Y*-scrambling) procedure (Wold et al., 1995). This technique consisted of scrambling the physico-chemical property values in such a way that they did not correspond to the respective compounds.

For model randomization, the dependent variables of the training set are shuffled and new correlation coefficients are calculated. The process is repeated several times. Kiralj and Ferreira (2009) demonstrated that it is sufficient to perform 10-25 *Y*-randomization runs for a model validation. The significantly low correlation coefficients of the new models indicate that the originally proposed model was not obtained by chance correlation. *Y*-scrambling procedure was run in MDM 2010.2 (Molegro Data Modeler, 2007-2010).

3.3.2. Validation

The best ways to validate a model is through an external test set, i.e. a set of compounds that have been left out from modeling, to assess the validity of the established model. However all molecules are often needed for the model building in QSAR and it might be too time consuming to generate new ones. When building a model with all molecules available the model still has to be evaluated.

We selected the molecules composing the training and test series as a previous step to the model search, and this was done in such a way that both sets shared similar qualitative structure-property characteristics for the compounds of the sets.

Consonni et al. (2009) formulated a novel external correlation coefficient (Q^2_{F3}) for the test set based on sum of squares (SS) referring to mean deviations of observed values from the training set mean over the training set instead of the external evaluation set. They concluded that correlation coefficients using either training set activity mean or test set activity mean have drawbacks. Therefore, the external predictive ability of the models should have information about the whole data set. They proposed equation 3.6 to test the external predictive ability of the models,

$$Q^2_{F3} = 1 - \frac{\left[\sum_{i=1}^{n_{test}} (\hat{y}_i - y_i)^2\right] / n_{test}}{\left[\sum_{i=1}^{n_{tr}} (y_i - \bar{y}_{tr})^2\right] / n_{tr}} \qquad (3.6)$$

where, $\hat{y}_i$ is the predicted test set compound, y_i is the observed value, $\bar{y}_{tr}$ is the mean of training set compounds, n_{test} is the number of compounds in test set and n_{tr} is the number of compounds in training set.

We adopted the criteria of Golbraikh et al. (2003) which correspond to OECD principle no 4 (OECD, 2007). Models were considered acceptable, if they satisfied all of the following conditions:

I. $R^2_{cv} > 0.5$

II. $R^2 > 0.6$,

III. R^2_0 or R'^2_0 close to R^2.

i.e.: **(a)** $(R^2 - R^2_0) / R^2 < 0.1$ and $0.85 \leq k \leq 1.15$ or

(b) $(R^2 - R'^2_0) / R^2 < 0.1$ and $0.85 \leq k' \leq 1.15$

IV. $\left| R_0^2 - R_0'^2 \right| < 0.3,$

where R^2 is predicted vs. observed, R'^2 is observed vs. predicted, *k* and *k*' are slopes, R_0^2 and $R_0'^2$ are squared correlation coefficients (without intercept).

(i) correlation coefficient *R* between the predicted and observed activities; (ii) coefficients of determination (predicted versus observed activities R_0^2, and observed versus predicted activities $R_0'^2$, for regressions through the origin); (iii) slopes *k* and *k'* of regression lines through the origin.

R_{cv}^2 is the cross-validated correlation coefficient calculated for the training set, but all other criteria are calculated for the test set.

3.3.3. Applicability Domain

The QSAR models are derived empirically from the analysis of a training set of chemicals, whose biological activity is known. The QSAR analysis is aimed at discovering the properties, or features of the molecules that correlate with the biological activity. In order to attain the best results, a QSPR analysis should focus on a well defined set of congeneric chemicals, i.e., chemicals with similar structure that act through the same mechanism of action (Franke 1984; Hansch and Leo 1995). Thus when the QSPR model is applied to new chemicals to predict their biological activity, it is crucial that the chemicals to be predicted have the same characteristics of the training set. These characteristics are called Applicability Domain of the model, and are typical of each individual model.

The Applicability Domain of the models contained in this expert system is defined in terms of structural characteristics of the chemical classes to which they apply. This expert system applies the models only to chemicals that respect such constraints.

The concept of AD concerns the predictive use of QSAR/QSPR models and, then, is closely related to the concept of model validation. In other words, the AD is a concept related to the quality of the QSAR/QSPR model predictions and prevention of the potential

misuse of model's results. A key component of the quality prediction is to define when a QSAR/QSPR model is suitable to predict a property/activity of a new compound, that is, define model's AD (Eriksson et al., 2003; Tropsha et al., 2003; Dimitrov et al., 2005; Nikolova-Jeliazkova and Jaworska, 2005; Jaworska et al., 2005a, 2005b; Tetko et al., 2006). A model will yield reliable predictions when model assumptions are fulfilled and unreliable predictions when they are violated. In particular, for QSAR/QSPR models, based on statistical mining techniques, the training set and the model prediction space are the basis for estimation of chemical space where predictions are reliable.

Two basic approaches were proposed to evaluate the AD. The first approach to AD evaluation is the analysis of the training set, which has its grounds in statistics, because the interpolated prediction results are more reliable than extrapolated. Extrapolation is not a problem in principle, because extrapolated results from theoretical well-founded models can often be reliable. However, QSAR/QSPR models are usually based on empirical and limited experimental evidence and/or are only locally valid; therefore, extrapolation always results in higher uncertainty and usually in unreliable predictions.

Different approaches to estimate interpolation regions in a multivariate space were evaluated by Jaworska (Netzeva et al., 2005; Jaworska et al., 2005) based on (1) ranges of the descriptor space; (2) distance-based methods, using Euclidean, Manhattan, and Mahalanobis distances, Hotelling T^2 method, and leverage values; and (3) probability density distribution methods based on parametric and nonparametric approaches. Both ranges and distance-based methods were also evaluated in the principal component space.

One of the common tools used to visualize the AD of a QSAR model is the plot of standardized residuals in prediction (r_i) versus leverage values (h_i) for each *i*th sample. This plot, called Williams plot, allows an immediate and simple graphical detection of both the response outliers (i.e., compounds with standardized residuals in prediction greater than three standard deviation units, $r_i > 3\sigma$) and structurally influential chemicals in a model ($h_i > h^*$), where h^* is a threshold value, usually 2 or 3 times the average leverage value. h^* is generally fixed at 3p/n, where n is the number of training compounds and p the number of model parameters, whereas x = 2 or 3, lying outside this area (vertical lines) the outliers and (horizontal lines) influential chemicals. In effect, when the leverage value of a

compound is lower than the critical value h^*, the probability of accordance between predicted and actual values is as high as that for the training set chemicals. Conversely, a high leverage chemical is structurally distant from the other chemicals; thus, it can be considered outside the AD of the model.

In any case, regardless of the specific method chosen for AD evaluation, this is always a very important task in order to avoid unreliable predictions and a misuse of the results.

The chemical domain of the studied chemicals in the models was verified by the leverage approach to verify prediction reliability. The plot of standardised residuals versus leverages (hat diagonals), i.e. the Williams graph verified the presence of response outliers (i.e. compounds with cross-validated standardized residuals greater than two-three standard deviation units, ($\pm 2\sigma$-3σ) and chemicals very structurally influential in determining model parameters. Also the data predicted by the models were verified for reliability by their leverage, so that only predicted data for chemicals belonging to the chemical domain of the training set would be proposed.

4. RESULTS AND DISCUSSION

4.1. Modeling log K_{oa} for PBDEs

Wania et al. (2002) used relative gas chromatographic retention times on a nonpolar stationary phase to determine the octanol/air partition coefficients (K_{oa}) of nonpolar semi-volatile organic compounds, including PBDEs. Harner and Shoeib (2002) measured K_{oa} values for 13 PBDEs at temperatures ranging from 15 to 45°C. It is apparent that the data are not enough since K_{oa} values for the other 196 PBDEs are lacking as stated before. Experimental determination of K_{oa} is costly in expenditures such as equipment and time. Furthermore, because of the limited number of standard PBDE congeners, it seems impossible to measure K_{oa} values for all the other PBDEs. Thus it is the purpose of this study to develop QSPR models for PBDEs, based on the experimental values.

Of the 22 log K_{oa} values of PBDEs reported by Wania et al. (2002), 3 of them were selected as test set compounds. The data (22 compounds) were divided into training/test set compounds in terms of property (log K_{oa}) value using an 80/20 split. Additionally, 8 log K_{oa} values of PBDEs taken from Harner and Shoeib (2002) were included in the test set. Descriptors were selected from 3 software (DRAGON 5.4, SPARTAN 06 and CODESSA 2.2) and the *CRI* program separately and combined using Heuristic Method (HM).

The main target of any QSAR modeling is that the developed model should be robust enough to be capable of making accurate and reliable predictions of biological activities of new compounds (Leonard and Roy, 2006; Roy, 2007; Roy et al., 2008). So, QSAR models that are developed from a training set should be validated using new chemical entities for checking the predictive capacity of the developed models. The validation strategies check the reliability of the developed models for their possible application on a new set of data, and confidence of prediction can thus be judged (Roy and Mandal, 2008). In many cases, enough new chemicals being unavailable for prediction purpose, the original data set is divided into a training set and a test set. For the present work, the compounds were ranked in terms of the log K_{oa} values and 11 of them were selected as a test set.

A number of QSPR models were built to assess the predictive power of QSPR models for log K_{oa}. Although analysis is done with various models where the number of descriptors is increased from 1 to 5, it is interesting to note that in most cases one descriptor-based models are adequate. Two one-descriptor models were obtained with the highest predictive ability obtained for the prediction of log K_{oa} using a training set of 19 compounds (Equations 4.1 and 4.2). Eq. 4.1 includes one descriptor from DRAGON and Eq. 4.2 includes descriptor from SPARTAN-*CRI* group. Both models have no response outlier.

$$\log K_{oa} = -3.533\ (\pm 0.231) + 0.126\ (\pm 0.002)\ \mathrm{D/Dr06} \qquad (4.1)$$

(Model 1)

$n = 19, \quad R^2 = 0.995, \quad F = 3246.33, \quad SE = 0.101$

$$\log K_{oa} = 4.086\ (\pm 0.175) + 0.966\ (\pm 0.030)\ CRI \qquad (4.2)$$

(Model 2)

$n = 19, \quad R^2 = 0.984, \quad F = 1043.50, \quad SE = 0.178$

It is interesting to note that both descriptors appeared in Equations 4.1 and 4.2 have topological characteristics. Each of the descriptors, D/Dr06 and the *CRI* from topological group explained 99 and 98% of variance for log K_{oa} value of PBDEs, respectively. D/Dr06 is the distance/detour ring index of order 6. The *CRI* which is an eigenvalue-based descriptor has been shown to be effective in modeling various properties of chemicals including the toxicity (Saçan et al., 2007).

In Table 4.1, the descriptors selected by HM are given together with the statistical parameters for DRAGON-based and SPARTAN-*CRI*-based models for log K_{oa} of PBDEs. The model obtained using all (1339) descriptors from DRAGON 5.4 (1167), SPARTAN 06-*CRI* (9) and CODESSA 2.2 (163) resulted in the same model as Eq. 4.1.

Table 4.1 Statistical parameters for one-parameter models developed for log K_{oa} of PBDEs.

log K_{oa} = ax + b									
Model No	Descriptors (x)	a	b	R^2	*F*	R^2_{cv}	*t*-test	R^2_{pred}	*RMSE*
DRAGON Model									
1	D/Dr06	0.126 (±0.002)[#]	-3.533 (±0.231)	0.995	3246	0.993	a: 56.98 b: -15.28	0.861	0.096
SPARTAN-*CRI* Model									
2	*CRI*	0.966 (±0.030)	4.086 (±0.175)	0.984	1043	0.980	a: 32.30 b: 23.41	0.857	0.173

[#] The numbers in the parentheses are 95% confidence intervals.

The *CRI*-based model (Eq. 4.2) was highlighted instead of DRAGON-based model (Eq. 4.1), because the *CRI* encodes global molecular properties such as size, volume, and surface area (Saçan and Balcıoğlu, 1996; 1998; Saçan et al., 2003), as well as hydrophobicity (log K_{ow}) (Saçan and İnel, 1995) whereas D/Dr06 explained as the distance/detour ring index of order 6. Although it was used in Partial Least Square (PLS) modeling of transfer protein inhibitors for a series of cholesteryl ester compounds by Castilho et al. (2007), it is difficult to explain its mechanistic relationship with log K_{oa}.

The *CRI*, which comprises all possible orders of path-type molecular indices, encodes local structural properties and possible long-range interactions described by path-type molecular indices. The *CRI* reflects the change in substituent pattern in PBDEs. In other words, it gives double structural information: its value changes according to both the number and the distance of bromine substituent, on each phenyl ring, from the oxygen ether. Thus, it takes also into account the information related not only molecular dimension but also the presence and the position of Br substitution on the phenyl rings like T(O...Br) reported by Papa et al. (2009). Considering Eq. 4.2, the higher the *CRI* was, the higher the log K_{oa} was. The dependence of the log K_{ao} of PBDEs on the compound size is reasonable,

since their log K_{oa} is related chiefly to their ability to stay in organic phase rather than to escape to the air phase.

Although we highlighted Eq. 4.2 in terms of mechanistic interpretation, internal validation and applicability domain of both models were given in the following paragraphs. For Eq. 4.1 and 2, 9 and 10 fold cross validations were run using Weka 3.6.1 (Waikato, 2009). The overall results of random deletion study statistics are summarized in Table 4.2.

Table 4.2 Leave-many-out cross validation results for Eq. 4.1 and Eq. 4.2.

Number of compounds deleted	**Eq. 4.1**		**Eq. 4.2**	
	Average R^2_{LMO}	**Average *RMSE***	**Average** R^2_{LMO}	**Average *RMSE***
2	0.997	0.098	0.990	0.182
2	0.996	0.120	0.988	0.198

Applicability domain of the proposed models is defined by the following limits given in Table 4.3.

Table 4.3 Boundaries of the proposed models for log K_{oa}.

Compound Name	**log K_{oa}**		**D/Dr06**		***CRI***	
	min	**max**	**min**	**max**	**min**	**max**
2-2,2',4,4',5,5'-	7.24	12.15	85.829	125.388	3.495	8.559

The observed values, predicted values from Eq. 4.2 are plotted in Figure 4.1. It shows a good alignment of the studied PBDE congeners along the optimal line.

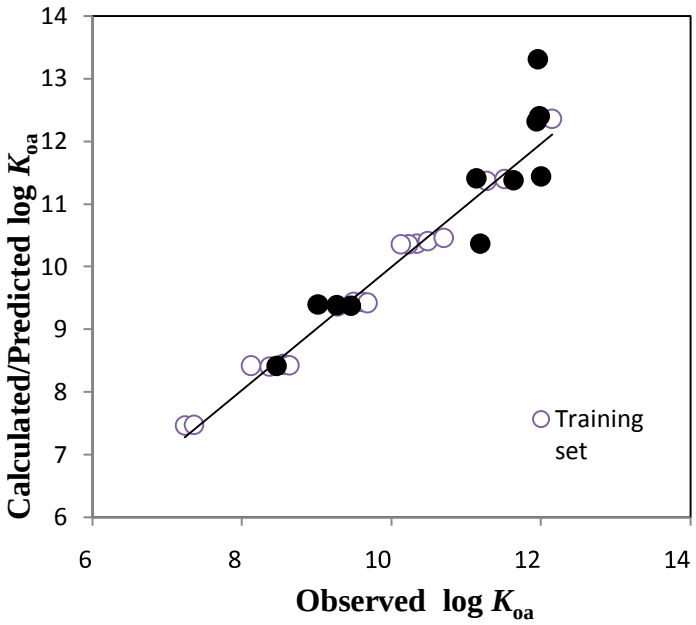


Figure 4.1 Plot of calculated/predicted vs. observed values of log K_{oa} for the training/test set compounds by Eq. 4.2.

The $[(R^2-R_0^2)/R^2]$ and k values for Eq. 4.2 are found to be within the acceptable range with values being equal to 0.002 and 1.01, respectively. External validation set of Eq. 4.2 with 11 compounds yielded good prediction statistics as indicated below.

(i) Correlation coefficient R^2 between the predicted and observed activities is 0.864 which is higher than 0.6.

(ii) R_0^2 or $R_0'^2$ close to R^2. R_0^2 and $R_0'^2$ are 0.862 and 0.830, respectively.

(a) $(R^2 - R_0^2) / R^2 < 0.1$ and $0.85 \leq k \leq 1.15$

$(R^2 - R_0^2) / R^2 = 2.31\text{x}10^{-3}$ and $k = 1.01$

(b) $(R^2 - R_0'^2) / R^2 < 0.1$ and $0.85 \leq k' \leq 1.15$

$(R^2 - R_0'^2) / R^2 = 0.039$ and $k' = 0.987$

(iii) $\left| R_0^2 - R_0'^2 \right| < 0.3$,

$\left| R_0^2 - R_0'^2 \right| = 0.032$

The Eq. 4.2 was also subjected to a randomisation test. In this test, the log K_{oa} values (Y) are randomly permuted keeping the descriptor matrix intact, followed by a MLR run. Each randomisation and subsequent MLR analysis generates a new set of R^2 values, which are plotted against the correlation coefficient between the original Y values and the permuted Y values. Random shuffling of response was repeated several times (15) for Eq. 4.2 and the average R^2 was 0.1057. The results confirm that the proposed model is well founded and not just the result of a chance correlation.

4.1.1. Analysis of the Applicability Domain

The AD of Eq. 4.2 exploited by Williams plot is shown in Figure 4.2. It can be seen that h_i values of all compounds are lower than the warning value (h^* = 0.315). For all the compounds in the training and test sets, their standardized residuals are smaller than two standard deviation units, i.e., there are no outliers for the developed QSPR model. Standardized residuals are symmetrically distributed around zero with no specific trend for Eq. 4.2 as shown in Figure 4.2. Therefore, the Eq. 4.2 can be employed to predict log K_{oa} for PBDEs.

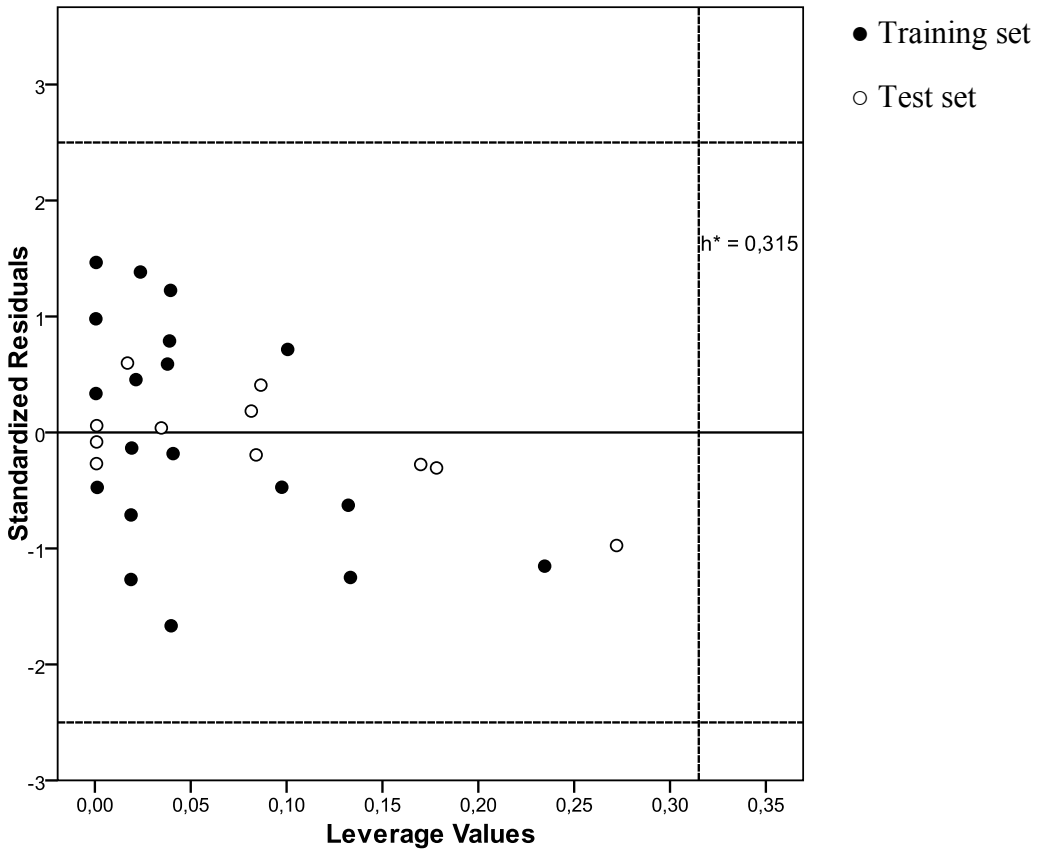


Figure 4.2 Williams plot for the Eq. 4.2. The log K_{oa} values for the training and test chemicals are labeled differently. The dotted lines are the 2.5σ limit and the warning value of hat (h^* = 0.315), respectively.

The observed and predicted values of PBDEs for training and test sets are listed together with the residuals and literature range for PBDE congeners in Table 4.4. New reliable and predictive MLR-QSPR models were proposed for the prediction of the log K_{oa} values of PBDEs. The best model chosen to predict the data set was evaluated for its predictivity by external validation in agreement with the OECD principles for QSAR validation. The predictive ability of the model was high (Q^2_{F3} = 0.857).

Because of its high statistical significance, the externally and internally validated Eq. 4.2 has been used to predict the log K_{oa} of those compounds where there are no experimental measurements (Appendix B, Table B.1). It is important to notice that the

log K_{oa} values of compounds/congeners outside of the applicability domain of Eq. 4.2 were indicated (Figure 4.3).

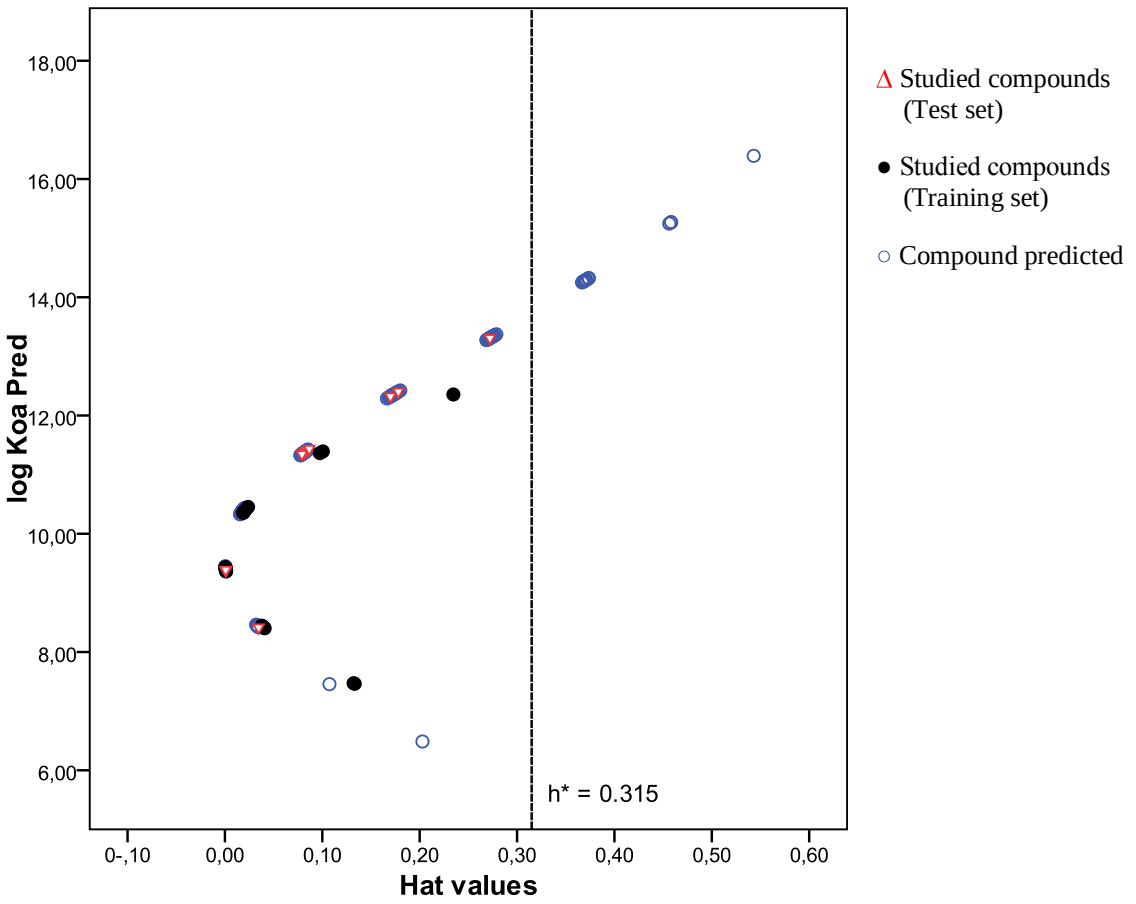


Figure 4.3 Plot of *hat* values vs. log K_{oa} predicted values of PBDEs (Eq. 4.2, Table 4.1).

Chemicals with the leverage values of higher than h^* = 0.315 reported in Figure 4.3 are considered to be out of the domain of the respective model (values marked with an asterisk (*) given in Appendix B (Table B.1). The model predictions for the chemicals with a leverage value higher than h^* should be considered potentially unreliable, since they are extrapolated.

7.65% of 209 PBDEs fell outside the AD. Those 16 PBDEs congeners belong to 8-10 bromination degree. Approximately, 93% of log K_{oa} values of PBDEs can be predicted with Eq. 4.2.

Table 4.4 The Compound names, descriptor, observed, predicted and reported literature values of log K_{oa} (25°C) values obtained from MLR (Eq. 4.2).

Congener Number	Compound Name	*CRI*	Observed value of log K_{oa}	Predicted value From Eq. 4.2	res	log K_{oa} (25°C) Literature Range	
						Minimum	Maximum
1	2-	3.495	7.24	7.46	-0.22	7.24[a]	7.4[c]
2	3-	3.504	7.36	7.47	-0.11	7.17[c]	7.39[b]
7	2,4-	4.468	8.37	8.40	-0.03	8.37[a,c]	8.38[b]
8	2,4'-*	4.481	8.47	8.42	0.05	8.47[a]	8.85[c]
10	2,6-	4.483	8.12	8.42	-0.30	8.08[b]	8.52[c]
12	3,4-	4.511	8.55	8.44	0.11	8.48[c]	8.57[b]
13	3,4'-	4.495	8.57	8.43	0.14	8.53[c]	8.57[a,b]
15	4,4'-	4.487	8.64	8.42	0.22	8.58[b]	8.7[c]
17	2,2',4-*	5.484	9.27	9.38	-0.11	9.39[e]	9.4[f]
21	2,3,4-	5.531	9.49	9.43	0.06	9.07[c]	9.49[a]
28	2,4,4'-*	5.479	9.46	9.38	0.08	9.62[f]	9.7[e]
30	2,4,6-*	5.491	9.02	9.39	-0.37	9.02[a,c]	9.08[b]
32	2,4',6-	5.463	9.28	9.36	-0.08	9.18[b]	9.28[a]
35	3,3',4-	5.536	9.61	9.44	0.17	9.38[c]	9.61[a]
37	3,4,4'-	5.519	9.68	9.42	0.26	9.47[c]	9.68[a]
47	2,2',4,4'-	6.497	10.34	10.36	-0.02	10.1[b]	10.63[f]
66	2,3',4,4'-	6.544	10.49	10.41	0.08	10.25[b]	10.77[d,e]
69	2,3',4,6-	6.490	10.23	10.36	-0.13	10.2[b]	10.38[c]

Table 4.4 (continued).

Congener Number	Compound Name	*CRI*	Observed value of log K_{oa}	Predicted value From Eq. 4.2	res	log K_{oa} (25°C) Literature Range	
						Minimum	Maximum
75	2,4,4',6-	6.489	10.13	10.36	-0.23	10.11[b]	10.25[c]
77	3,3',4,4'-	6.590	10.70	10.45	0.25	10.41[c]	10.83[d]
82	2,2',3,3',4-*	7.577	11.14	11.41	-0.27	11.14[a]	10.94[b]
85	2,2',3,4,4'-*	7.546	11.63	11.38	0.25	11.35[f]	11.55[e]
99	2,2',4,4',5-	7.533	11.28	11.36	-0.08	11.18[f]	11.49[e]
100	2,2',4,4',6-*	7.515	11.19	11.35	-0.16	11.14[e]	11.7[f]
119	2,3',4,4',6-	7.562	11.52	11.39	0.13	11.17[b]	11.52[a]
126	3,3',4,4',5-*	7.608	12.00	11.44	0.56	11.65[f]	11.99[e]
153	2,2',4,4',5,5'-	8.559	12.15	12.36	-0.21	11.73[f]	12.32[c]
154	2,2',4,4',5,6'-*	8.526	11.94	12.32	-0.38	11.89[e]	12.26[f]
156	2,3,3',4,4',5-*	8.610	11.98	12.40	-0.42	11.96[f]	12.63[e]
183	2,2',3,4,4',5',6-*	9.546	11.96	13.31	-1.35	11.97[f]	12.68[e]

[a] Wania et al., 2002; [b]Wang et al., 2008; [c] Xu et al., 2007; [d]Harner and Shoeib, 2002; [e]Chen et al., 2003; [f]Zhao et al., 2005. *the ones marked by an asterisk are the PBDE congeners in the test set.

4.1.2. Comparison with the Reported Methods

The studied physico-chemical property, log K_{oa}, has already been modeled by different authors (Chen et al., 2003; Wania et al., 2003; Zhao et al., 2005; Xu et al., 2007; Wang et al., 2008; Papa et al., 2009). However, none of the other already published QSPRs were externally validated and checked for their applicability domain, with the only exception of log K_{oa} model by Wang et al. (2008) and Papa et al. (2009). This fact is important to note since, as demonstrated in this study, predicted values calculated by models based on few experimental data are extrapolated, and possibly unreliable, for compounds that fall out of their structural applicability domain.

Due to different amount of studied compounds or differences in the development of the models, it is not possible to perform strict comparisons among our proposed models and already published QSPRs for PBDEs. However, some general considerations can still be made. As shown in Table 4.5, our models had comparable or even better performances in comparison to other QSPRs at the same level of complexity (one descriptor) or higher. The *CRI*-based model is superior to the models reported in Table 4.5 in terms of number of variables and statistical parameters. It is important to note that the high values of the calculated statistical parameters of the existing models, shown in Table 4.5, could give overoptimistic idea of the predictive ability of the models.

Normally the ratio of observations to variables should be as high as possible and at least 5:1 (Topliss and Costello, 1972). Of the results reported in Table 4.5, models reported by Zhao et al. (2005) and Xu et al. (2007) have higher standard errors than the *CRI*-based model. It must be also noted that training and test set division are missing in four of the six reported models. The model developed by Wang et al. (2008) for the property log K_{oa} was the result of a stepwise multiple linear regression based on three theoretical descriptors. Predictions were given for all the 209 PBDEs for these three models, but no parameter was calculated to verify the AD of the model and/or to evaluate the reliability of predictions.

On the other hand, the models by Wania and Dugani (2003) as well as by Papa et al. (2009) were built by using only very simple variables such as molar mass and sum of topological distances between oxygen and bromine atoms T(O..Br), as molecular

descriptors, respectively. In the study of Papa et al. (2009), BDE183 was identified as strong response outlier in the training set (residual > 3.5σ) for the property log K_{oa}. When molar mass is used as descriptor its value is constant for isomeric compounds and calculates constant predictions for each group of PBDE congeners. Even though this QSPR is very simple, it is not sensitive at all to variations in the responses and the AD of this model was not investigated.

Our model is also superior to the T(O…Br)-based model proposed by Papa et al. (2009) in many aspects, including statistical parameter and mechanistic interpretation of the descriptor appearing in model e.g., information encoded by the *CRI* which is explained well previously.

Table 4.5 Statistical performance comparison of different QSPR models of log K_{oa} for PBDEs.

Number of descriptors	Descriptor type	Training/test set	R^2	SE	F	N*	R^2_{pred}	References
1	*CRI*	19/11	0.984	0.178	1043.50	30	0.82	Current study
9	1/T; *CAA*; *CSEV*; α; *TE*; M_w; *CMA*; ΔH_f; *CCR*; E_{LUMO}	—	0.979	0.150	—	36	—	Chen et al. (2003)
1	M_M	—	0.994	—	—	9	—	Wania and Dugani (2003)
2	χ_p; χ_{vp}	—	0.927	0.275	64	13	—	Zhao et al. (2005)
2	V_{mc}; $\sum V_s^+$	—	0.976	0.218	388.94	22	—	Xu et al. (2007)
3	q^-; μ; α	—	0.997	—	2176.64	22	0.98	Wang et al. (2008)
1	T(O...Br)	24/6	0. 960	—	—	30	0.98	Papa et al. (2009)

*Here *N* donates the total number of chemicals in the data set.

4.2. Modeling log K_{ow} for PBDEs and PCDEs

The splitting of the data on PBDEs and PCDEs into training and test sets was realized by applying Self Organized Maps (SOM) Kohonen Artificial Neural Networks (K-ANN) in the software CODESSA 2.2. The training set was formed by Kohonen network after HM analysis on 1186 DRAGON, 163 CODESSA and 8 SPARTAN and the *CRI*. The training data set were used to develop 1-5 descriptor QSPR models for log K_{ow} of PBDEs/PCDEs. Different network size and epochs (iterations) were tried to obtain a training/test set ratio as 80/20 (Table 4.6). At the end of the iterative learning process of the map, it is assumed that similar chemicals, which carry the same structural information, fall within the same neuron (cell in the top map). Chemicals sampled as a test set (about 20% of overall data) was those with the minimal distance from the centroid of each cell of the top map. We wrote in bold the selected 13x13 network and 500 epochs combination in Table 4.6.

Table 4.6 Kohonen division trials of log K_{ow} data set of PBDEs/PCDEs.

Architecture	**Network size**	**Epochs**	**Number of compounds in training / test set**
1	11x11	100	79 / 42
2	11x11	500	76 / 45
3	11x11	800	71 / 50
4	12x12	100	76 / 45
5	12x12	400	80 / 41
6	13x13	100	84 / 37
7	13x13	200	81 / 40
8	13x13	400	88 / 33
9	**13x13**	**500**	**93 / 28**
10	13x13	800	89 / 32
11	13x13	1000	88 / 33

Descriptors appeared in 1-5 variable models for log K_{ow} of the training set were given in Table 4.7. DRAGON-based models included mostly GETAWAY (GEometry, Topology and Atom-Weights AssemblY) descriptors (R1e, H2e, HATS8p and HIC), constitutional (nBr), 3D MoRSE (Mor02u) and WHIM (G2p) descriptors.

GETAWAY descriptors are 3D Descriptors or Geometrical Descriptors which are obtained by using double-weighted autocorrelation functions, where one weighting scheme is the leverage and the other is an atomic property (e.g., atomic mass).

The GETAWAY class of descriptors represents recently proposed group of descriptors, which are based on a leverage matrix similar to that defined in statistics and usually used for regression diagnostics. These molecular descriptors match the three dimensional molecular geometry provided by the molecular influence matrix and atom relatedness by molecular topology, with chemical information by using various atomic weight schemes (Consonni et al., 2002a; 2002b). Therefore, this class of descriptors is highly sensitive to the 3-dimensional molecular structure. Combined with appropriate weighting schemes the GETAWAY descriptors are used to compare molecules or even conformers taking into account their molecular shape, size symmetry and atom distribution, which are 'scaled' using specific atomic property.

WHIM descriptors are the molecular descriptors based on statistical indices calculated on the projections of the atoms along principal axes (Todeschini and Gramatica, 1997; Todeschini and Consonni, 2000). They are built in such a way as to capture relevant molecular 3-dimensional information regarding molecular size, shape, symmetry, and atom distribution with respect to invariant reference frames.

One-descriptor MLR model have been developed to calculate the log K_{ow} of PBDEs and PCDEs from GETAWAY descriptor, namely, R1e. Data set on 121 compounds (107 PCDE, 14 PBDE) is divided into a training set and a test set (93 and 28) by SOM Method (13x13 neuron, 500 epochs).

R1e is based on the relationship between electronegativity and atomic size, and calculated from the reciprocal of the atomic volume. A preliminary MLR model of log K_{ow}

led to a model with one DRAGON descriptor (R1e) having R^2 value of 0.971 with four outliers: 2,2',4-trichloro-DPE, 3,3',4,4'-tetrabromo-DPE, 2,2',4-tribromo-DPE and 3,3',4-tribromo-DPE. After removal of outliers, the following equation was found (Model 4 in Table 4.7).

$$\log K_{ow} = 23.539\ (\pm 0.320) - 15.168\ (\pm 0.280)\ \text{R1e} \qquad (4.3)$$

(Model 4)

$$n_{training} = 89, \quad R^2 = 0.971, \quad F = 2936.62, \quad SE = 0.151, \quad R^2_{cv} = 0.9767, \quad R^2_{pred} = 0.973$$

The 98% confidence intervals are given in parentheses. All the β-coefficients are significant at 95% level. All compounds were estimated within a three standard deviation range for Eq. 4.3. R1e is based on the relationship between electronegativity and atomic size, and calculated from the reciprocal of the atomic volume. Although R1e seems to be a prevailing descriptor in modeling log K_{ow} of the combined dataset, it is not an important descriptor for the PCDE subset. For the PCDE subset containing 107 compounds separate correlations with another GETAWAY descriptors H2e and Mor08v (3D-MoRSE descriptor) significantly improve the quality of the model and give SE = 0.090 (Eq. 4.7).

Two-descriptor MLR model developed for training set without an outlier is given in Eq. 4.4 (Model 5; Table 4.7). The *t*-values for partial correlation coefficients in Eq. 4.4 are 69.1915 and 58.0169 for the H2e and nBr, respectively. On the basis of the *t*-values, it can be concluded that H2e explains the log K_{ow} significantly more than nBr.

$$\log K_{ow} = -19.108\ (\pm 0.3641) + 14.967\ (\pm 0.2163)\ \text{H2e} + 0.413\ (\pm 0.007)\ \text{nBr} \qquad (4.4)$$

(Model 5)

$$n_{training} = 93, \quad R^2 = 0.9856, \quad F = 3089.35, \quad SE = 0.011, \quad R^2_{cv} = 0.9846$$

CODESSA descriptor, XY shadow, alone explained more than 90% of variances of log K_{ow} of PBDEs/PCDEs with an outlier (2,2',4-trichlorinated diphenyl ether) (Table 4.7). The model was stabilized after removal of this outlier with R^2 = 0.965. Increasing the

number of descriptors, replaced XY Shadow with the other descriptors from CODESSA. Those descriptors were given in Table 4.7. The XY shadow (XY) is the projection of the molecules onto the XY-plane oriented in space along the axes of inertia and represents the size of the molecule along the longest axis.

It can be seen from Table 4.7 that the performance of the models changed with the number of selected descriptors in the models. The values of R^2 and R^2_{cv} increase gradually with the increase in the number of descriptors (m). However, when m reaches four, its succeeding increase will have only trivial influence on all these values, which suggesting that the addition of variable number has not significant difference in statistical quantities and the simpler is always the better option. Additionally, some of these descriptors (nBr) resulted too simple (constant predictions for different classes of PBDEs and PCDEs) and not sensitive at all to the variations in the substitution on the aromatic rings (same predicted value for isomers); other variables were of too complex interpretation (i.e. 3D descriptors) and thus not mechanistically informative. Among these variables appearing in models listed in Table 4.7 the *CRI* was the one which gave a clear additional contribution to mechanistic interpretation by taking into account not only molecular dimension but also the presence, and the position of Br and Cl substitution. Additionally, the *CRI* along with other quantum chemical descriptors (E_{HOMO} and E_{aq}) is a more powerful model (Eq. 4.5; model 16). Therefore, it was highlighted in Table 4.7 and is regarded as the best model.

Table 4.7 Comparative analysis of QSPR models based on Heuristic method for log K_{ow} of PBDEs and PCDEs.

Model No	**Number of descriptor (*m*)**	**Variables**	R^2	F	SE	R^2_{cv}
		DRAGON-based models				
3	1	Mor02u	0.956	1959	0.034	0.953
4	1	R1e	0.956	1954	0.034	0.954
5	2	H2e, nBr	0.986	3089	0.011	0.985
6	3	H2e, nBr, HATS8p	0.988	2418	0.009	0.987
7	5	H2e, nBr, HATS8p, G2p, HIC	0.990	1679	0.008	0.988
		CODESSA-based models				
8	1	S_{XY}	0.961	2250	0.030	0.960
9	2	M; WNSA3	0.971	1487	0.023	0.969
10	3	WNSA3; N_{Br}; P	0.980	1424	0.016	0.977
11	4	${}^0\chi^v$; WNSA3; Q_{max} for a O atom; P	0.986	1549	0.011	0.984
12	5	${}^0\chi^v$; WNSA3; Q_{max} for a O atom; Q_{max} for a C atom; I_k	0.987	1278	0.011	0.985
		SPARTAN-*CRI* based models				
13	1	E_{LUMO}	0.806	377	0.152	0.792
14	1	*CRI*	0.762	292	0.186	0.740
15	2	*CRI*, E_{aq}	0.953	912	0.037	0.950
16	**3**	***CRI*, E_{aq}, E_{HOMO}**	**0.970**	**954**	**0.024**	**0.967**
17	4	*CRI*, E_{aq}, E_{HOMO}, μ	0.977	935	0.019	0.974
18	5	*CRI*, E_{aq}, E_{LUMO}, μ, η	0.977	741	0.019	0.974

As it can be observed from Eq. 4.5, the non-cross-validated R^2 = 0.970 and F = 49.749 indicates that log K_{ow} model is highly satisfactory. Next, the R^2_{cv} is highly significant and calculated root mean square error of cross-validation ($RMSE_{cv}$ = 0.090) was low which suggest the good stability of the model. The plot of calculated/predicted versus observed values of log K_{ow} (see Figure 4.4) also confirmed the linear character of the model. The *CRI* from SPARTAN-*CRI* group alone explained more than 75% of the variance in log K_{ow} of PBDEs/PCDEs.

$$\log K_{ow} = -6.341\ (\pm 1.262) + 0.765\ (\pm 0.015)\ CRI - 0.625\ (\pm 0.029)\ E_{aq} - 0.958\ (\pm 0.136)\ E_{HOMO} \quad (4.5)$$

(Model 16)

$n_{\text{training}} = 93,\quad R^2 = 0.970,\quad R^2_{adj} = 0.969,\quad R^2_{cv} = 0.967,$

$F = 49.749,\quad SE = 0.155,\quad RMSE = 0.151$

$n_{\text{test}} = 28,\quad R^2 = 0.9472,\quad R_0^2 = 0.9471,\quad R_0'^2 = 0.946,\quad Q^2_{F3} = 0.940$

The *t*-values are 52,633, -21.399, and -7.053 for the *CRI*, E_{aq} and E_{HOMO}, respectively. The *t*-values indicate that each parameter is highly significant ($p < 0.05$). This model explains more than 95% of the variance in the experimental log K_{ow} values for combined set. The greater the *t*-value, the more contribution to the regression equation. VIF values for descriptors the *CRI*, E_{aq} and E_{HOMO} appeared in Eq. 4.5 are 1.436, 3.055 and 2.625, respectively, indicating that these variables are not intercorrelated.

In this study, the *CRI* relates hydrophobicity since the bigger the molecules, the greater the *CRI* and thus the greater the hydrophobicity. The *CRI* represents many molecular properties besides hydrophobicity as explained before. The *CRI* term was introduced in Eq. 4.5 because the larger PBDE and PCDE molecule would yield stronger dispersion-type interaction between each other (lowering the volatility and liquid vapour pressure) or with the *n*-octanol molecule (i.e. the log K_{ow} value becomes larger), and tend to be excluded from water (i.e. water solubility becomes smaller).

In the octanol-water system, octanol molecules are easier to release electrons than water molecules, since octanol have smaller E_{HOMO} values. The lower the E_{HOMO} values, the greater is the tendency of chemicals to donate electrons in intermolecular interactions, the greater is the intermolecular interactions between octanol molecules and PCDEs/PBDEs, the higher the log K_{ow} values. Similarly, in a study reported by Reddy and Locke (1996) E_{HOMO} was found to be significant rather than E_{LUMO} in describing the log K_{ow} of 90 herbicides. In other words, E_{HOMO} seems particularly related to the tendency to dissolve in octanol phase.

The lower the E_{aq} values, the greater the tendency of PBDE/PCDE molecules to stay in water phase the greater the intermolecular interactions between PBDE/PCDE and water molecules, and thus the lower the log K_{ow} value. For training set of 93 compounds in Eq. 4.5, 31, 15 and 10 fold cross validations were run using Weka 3.6.1 (Waikato, 2009). The overall results of random deletion study statistics are summarized in Table 4.8.

Table 4.8 Leave-many-out cross validation results for Eq. 4.5.

Number of compounds deleted	**Average** R^2_{LMO}	**Average** ***RMSE***
3	0.983	0.160
6	0.983	0.159
9	0.983	0.159

The predicted vs. observed log K_{ow} values of the training and test set compounds obtained from Eq. 4.5 are shown in Figure 4.4. The $[(R^2\text{-}R_0^2)/R^2]$ and k values for Eq. 4.5 are found to be within the acceptable range with values being equal to 0.0001 and 0.994, respectively. External validation set of Eq. 4.5 with 28 compounds yielded good prediction statistics as indicated below.

(i) Correlation coefficient R^2 between the predicted and observed activities is 0.9472 which is higher than 0.6.

(ii) R_0^2 or $R_0'^2$ close to R^2. R_0^2 and $R_0'^2$ are 0.9471 and 0.946, respectively.

(a) $(R^2 - R_0^2)/R^2 < 0.1$ and $0.85 \leq k \leq 1.15$
$(R^2 - R_0^2)/R^2 = 1.05\text{x}10^{-4}$ and $k = 0.994$

(b) $(R^2 - R_0'^2)/R^2 < 0.1$ and $0.85 \leq k' \leq 1.15$
$(R^2 - R_0'^2)/R^2 = 1.266\text{x}10^{-3}$ and $k' = 1.004$

(iii) $| R_0^2 - R_0'^2 | < 0.3$,
$| R_0^2 - R_0'^2 | = 0.001$

Moreover, the model robustness was also checked by response randomization (*Y*-scrambling). The log K_{ow} values were shuffled randomly between the molecules and regression models were developed. Random shuffling of response was repeated several times (15) for Eq. 4.5. R^2 values were between 0.0016 and 0.094, and the average R^2 was 0.0382. The results confirm that the proposed model is well founded and not just the result of a chance correlation.

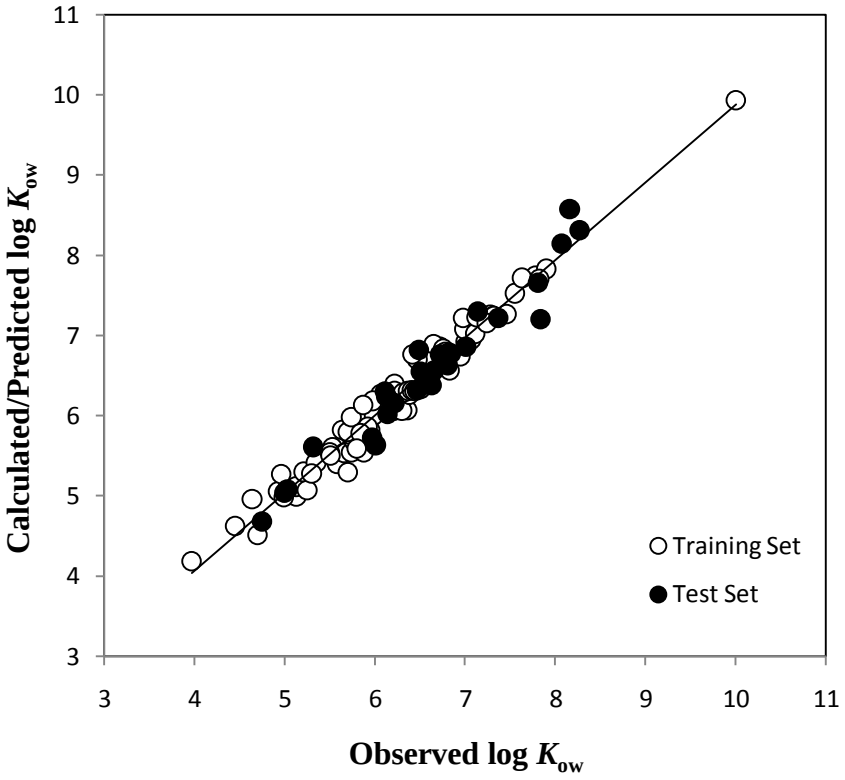


Figure 4.4 Plot of calculated/predicted vs. observed values of log K_{ow} for the training/test set compounds for PBDEs/PCDEs (Eq. 4.5).

As far as our knowledge there is no reported model in the literature for log K_{ow} of the combined data set. We applied the same procedure, to generate a QSPR models for the compounds of the PBDE and PCDE subsets. The *CRI* also appears to be the most significant descriptor in modeling log K_{ow} values of PBDE and PCDE subsets. The QSPR equation in terms of the *CRI* turned out to be basically the same in both combined set and PBDE subset models (Eq. 4.6).

$$\log K_{ow} = -5.691\ (\pm 2.313) + 0.506\ (\pm 0.024)\ CRI - 0.939\ (\pm 0.254)\ E_{HOMO} \qquad (4.6)$$

$$n = 13, \qquad R^2 = 0.993, \qquad F = 728.289, \qquad SE = 0.113, \qquad R^2_{cv} = 0.993$$

We obtained above equation within 2.0σ standard deviation limit with one outlier (2,2',4-tri-BDE). The *t*-values are 21.213 and -3.698 for the *CRI* and E_{HOMO}, respectively. The *t*-values indicate that each parameter is highly significant (p<0.05). The *CRI* explained more than 95% variances in log K_{ow} values of PBDEs.

Data set on 107 PCDE is divided into a training set and a test set by SOM Method (11x11 neuron, 100 epochs). The 79 PCDE as the training set is used to generate the log K_{ow} models. There is only one outlier (2,2',4-trichloro diphenyl ether (PCDE-17) with a slightly larger predictive error of 0.45. The same compound was detected as an outlier in a structure-log K_{ow} correlation reported by Yang et al. (2003) and Sun et al. (2007). Yang et al. (2003) also reported that PCDE molecules with great molecular size or high degree of chlorination may have high strong intermolecular dispersive interactions, which result in high log K_{ow} values. PCDE molecules with high molecular size (volume) need more free energy of enthalpy input to overcome solvent-solvent cohesive interactions to provide a suitable sized cavity for solutes. Their model has one outlier (PCDE-17) similar to our study. Generally, log K_{ow} values should increase with increasing molecular size, which implies that higher chlorinated PCDE congeners have higher log K_{ow} values. However, observing the log K_{ow} values reported by Kurz and Ballschmiter (1999), it can be found that the log K_{ow} values of the three dichlorinated diphenyl ethers (PCDE-13, 14, 15) are larger than the log K_{ow} value of PCDE-17, a trichlorinated diphenyl ether. As can be found from Table B.2, the log K_{ow} value of PCDE-17 is the smallest among the 16 trichlorinated diphenyl ethers. Therefore, the log K_{ow} value of PCDE-17 reported by Kurz and Ballschmiter (1999) may have relevant experimental errors. However, if we discard this compound from the training set, for the remaining 78 PCDE, a two-parameter model including the H2e and Mor08v is obtained along with the statistical parameters were shown in the following equation (Eq. 4.7).

$$\log K_{ow} = -19.600\ (\pm 0.346) + 15.132\ (\pm 0.187)\ \text{H2e} - 0.441\ (\pm 0.175)\ \text{Mor08v} \qquad (4.7)$$

$$n = 78, \quad R^2 = 0.989, \quad SE = 0.090, \quad F = 6.330, \quad R^2_{cv} = 0.9939, \quad R^2_{pred} = 0.976$$

A comparison of calculated or predicted log K_{ow} values and experimental data for 107 PCDEs is shown in Figure 4.5. One can observe that calculated or predicted log K_{ow} values agree well with experimental data, and the plot shows no obviously observable pattern. Therefore, the final model (Eq. 4.7) represents an excellent QSPR model judging from the statistics and the plot in Figure 4.5.

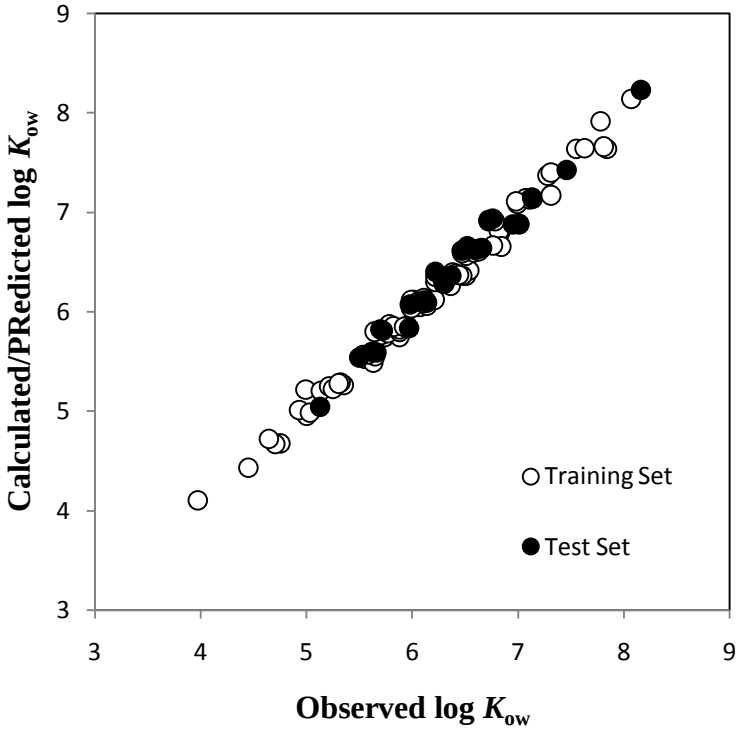


Figure 4.5 Plot of calculated/predicted vs. observed values of log K_{ow} for the training/test set compounds for PCDEs.

Table 4.9 The Compound names, descriptors, observed and predicted log K_{ow} values, and residuals obtained from MLR (Eq. 4.5).

Compound name	**Congener number**	***CRI***	E_{HOMO}	E_{aq}	**log K_{ow} obs**	**log K_{ow} pred from Eq. 4.5**	**res**
Training Set							
2-	1	3.04	-9.36	0.51	4.45	4.63	-0.18
4-	3	2.96	-9.26	0.45	4.7	4.52	0.18
2,4-	7	3.44	-9.32	0.25	4.93	5.06	-0.13
2,5-	9	3.46	-9.23	0.25	5.13	5.00	0.13
2,6-	10	3.43	-9.26	0.32	4.64	4.96	-0.32
3,4-	12	3.42	-9.23	0.22	4.99	4.99	0.00
3,4'-	13	3.42	-9.35	0.17	5.13	5.12	0.01
3,5-	14	3.46	-9.52	0.20	5.21	5.30	-0.09
4,4'-	15	3.40	-9.31	0.17	5.25	5.08	0.17
2,2',4-	17	3.90	-9.04	0.04	4.96	5.27	-0.31
2,3,4-	21	3.82	-9.36	0.05	5.55	5.52	0.03
2,3,4'-	22	3.85	-9.40	-0.01	5.63	5.62	0.01
2,3,5-	23	3.86	-9.26	0.04	5.62	5.46	0.16
2,3,6	24	3.84	-9.27	0.10	5.35	5.42	-0.07
2,3',4-	25	3.90	-9.38	-0.04	5.65	5.66	-0.01

Table 4.9 (continued).

Compound name	Congener number	*CRI*	E_{HOMO}	E_{aq}	log K_{ow} obs	log K_{ow} pred from Eq. 4.5	res
2,4,4'-	28	3.88	-9.36	-0.03	5.53	5.61	-0.08
2,4,5-	29	3.86	-9.21	0.04	5.58	5.41	0.17
2,4',5-	31	3.86	-9.30	-0.03	5.66	5.54	0.12
2,4',6-	32	3.87	-9.07	0.04	5.3	5.29	0.01
2,3',4'-	33	3.89	-9.30	-0.01	5.5	5.55	-0.05
3,3',4-	35	3.86	-9.28	-0.06	5.74	5.55	0.19
3,4,4'-	37	3.86	-9.28	-0.06	5.88	5.55	0.33
3,4,5-	38	3.84	-9.08	0.00	5.7	5.30	0.40
3,4',5-	39	3.86	-9.42	-0.08	5.77	5.68	0.09
2,2',3,4-	41	4.32	-9.08	-0.17	5.72	5.77	-0.05
2,2',3,4'-	42	4.33	-9.15	-0.20	5.88	5.86	0.02
2,2',4,4'-	47	4.37	-9.07	-0.22	5.95	5.83	0.12
2,2',4,5'-	49	4.36	-9.09	-0.22	5.78	5.84	-0.06
2,3,3',4-	55	4.33	-9.41	-0.24	6.07	6.14	-0.07
2,3,3',4'-	56	4.32	-9.35	-0.25	5.99	6.08	-0.09
2,3,4,4'-	60	4.32	-9.39	-0.23	6.14	6.11	0.03

Table 4.9 (continued).

Compound name	Congener number	*CRI*	E_{HOMO}	E_{aq}	log K_{ow} obs	log K_{ow} pred from Eq. 4.5	res
2,3,4,6-	62	4.31	-9.30	-0.10	5.88	5.92	-0.04
2,3,4’,5-	63	4.34	-9.32	-0.24	6.21	6.06	0.15
2,3,4’,6-	64	4.34	-9.12	-0.18	5.64	5.83	-0.19
2,3,5,6-	65	4.25	-9.22	-0.10	5.82	5.80	0.02
2,3’,4,4’-	66	4.35	-9.35	-0.27	6.13	6.12	0.01
2,3’,4’,5-	70	4.37	-9.33	-0.27	6.11	6.10	0.01
2,3’,4’,6-	71	4.35	-9.07	-0.20	5.7	5.80	-0.10
2,4,4’,5-	74	4.34	-9.26	-0.24	5.99	6.01	-0.02
2,4,4’,6-	75	4.37	-9.12	-0.20	5.92	5.87	0.05
3,3’,4,4’-	77	4.33	-9.31	-0.30	6.36	6.07	0.29
3,4’,4,5-	81	4.30	-9.34	-0.28	6.3	6.07	0.23
2,2’,3,3’,4-	82	4.72	-9.16	-0.41	6.3	6.30	0.00
2,2’,3,4,4’-	85	4.75	-9.12	-0.43	6.28	6.29	-0.01
2,2’,3,4,5’-	87	4.80	-9.13	-0.42	6.51	6.34	0.17
2,2’,3,4,6’-	89	4.75	-9.10	-0.38	6.11	6.25	-0.14
2,2’,3,4’,5-	90	4.79	-9.17	-0.45	6.54	6.38	0.16

Table 4.9 (continued).

Compound name	Congener number	*CRI*	E_{HOMO}	E_{aq}	log K_{ow} obs	log K_{ow} pred from Eq. 4.5	res
2,2’,3,4’,6-	91	4.77	-9.10	-0.39	6.06	6.27	-0.21
2,2’,3,4’,5’-	97	4.75	-9.22	-0.44	6.22	6.40	-0.18
2,2’,4,4’,5-	99	4.79	-9.05	-0.44	6.38	6.27	0.11
2,2’,4,5,5’-	101	4.82	-9.07	-0.44	6.22	6.31	-0.09
2,2’,4,5,6’-	102	4.79	-9.00	-0.39	5.98	6.19	-0.21
2,3,3’,4’,5-	107	4.76	-9.35	-0.49	6.52	6.57	-0.05
2,3,3’,4,5’-	108	4.78	-9.47	-0.48	6.58	6.69	-0.11
2,3,3’,4’,6-	110	4.75	-9.11	-0.43	6.31	6.29	0.02
2,3,4,5,6-	116	4.71	-9.27	-0.27	6.37	6.31	0.06
2,3,4’,5,6-	117	4.77	-9.16	-0.38	6.41	6.32	0.09
2,3’,4,4’,5-	118	4.76	-9.30	-0.49	6.6	6.51	0.09
2,3’,4,4’,6-	119	4.77	-9.12	-0.44	6.44	6.32	0.12
3,3’,4,4’,5-	126	4.74	-9.35	-0.52	6.83	6.57	0.26
2,2’,3,3’,4,6’-	132	5.20	-9.15	-0.60	6.47	6.78	-0.31
2,2’,3,4,4’,5-	137	5.16	-9.25	-0.64	6.72	6.87	-0.15
2,2’,3,4,4’,5’-	138	5.21	-9.26	-0.66	7.01	6.93	0.08

Table 4.9 (continued).

Compound name	Congener number	*CRI*	E_{HOMO}	E_{aq}	log K_{ow} obs	log K_{ow} pred from Eq. 4.5	res
2,2’,3,4,4’,6’-	140	5.34	-9.16	-0.61	6.65	6.90	-0.25
2,2’,3,4’,5,5’-	146	5.24	-9.15	-0.66	6.76	6.84	-0.08
2,2’,3,4’,5,6-	147	5.17	-9.14	-0.59	6.76	6.74	0.02
2,2’,3,4’,5’,6-	149	5.23	-9.04	-0.61	6.47	6.71	-0.24
2,3,3’,4,4’,5-	156	5.13	-9.35	-0.67	7.07	6.96	0.11
2,3,3’,4,4’,5’-	157	5.20	-9.42	-0.68	6.99	7.09	-0.10
2,3,3’,4’,5,6-	163	5.21	-9.15	-0.62	6.78	6.80	-0.02
2,3,4,4’,5,6-	166	5.15	-9.19	-0.55	6.95	6.75	0.20
2,3’,4,4’,5,5’-	167	5.22	-9.33	-0.70	7.11	7.03	0.08
2,2’,3,3’,4,4’,5-	170	5.66	-9.14	-0.84	7.28	7.27	0.01
2,2’,3,3’,4,5,6’-	174	5.68	-9.10	-0.80	6.98	7.22	-0.24
2,2’,3,4,4’,5,5’-	180	5.68	-9.12	-0.85	7.46	7.28	0.18
2,2’,3,4,4’,5,6-	181	5.63	-9.17	-0.75	7.31	7.22	0.09
2,2’,3,4’,5,5’,6-	187	5.71	-9.08	-0.81	7.13	7.23	-0.10
2,3,3’,4,4’,5,5’-	189	5.67	-9.38	-0.88	7.55	7.53	0.02
2,3,3’,4,4’,5,6-	190	5.62	-9.17	-0.79	7.31	7.25	0.06

Table 4.9 (continued).

Compound name	Congener number	*CRI*	E_{HOMO}	E_{aq}	log K_{ow} obs	log K_{ow} pred from Eq. 4.5	res
2,2’,3,3’,4,4’,5,5’-	194	6.10	-9.16	-1.03	7.78	7.75	0.03
2,2’,3,3’,4,5,5’,6’-	199	6.13	-9.14	-1.00	7.63	7.72	-0.09
4,4’-	315	4.49	-9.68	1.37	5.51	5.51	0.00
2,2’,4-	317	5.48	-9.61	1.71	5.74	5.99	-0.25
2,4,4’-	328	5.48	-9.46	1.80	5.84	5.79	0.05
2,4’,6-	332	5.46	-9.36	1.94	5.8	5.59	0.21
3,3’,4-	335	5.54	-9.73	1.73	5.87	6.14	-0.27
3,3’,4,4’-	377	6.59	-9.78	2.09	6.42	6.77	-0.35
2,2’,4,4’,5-	399	7.53	-9.75	2.43	7.32	7.25	0.07
2,2’,4,4’,6-	400	7.51	-9.78	2.60	7.24	7.16	0.08
2,2’,4,4’,5,5’-	453	8.56	-9.78	2.79	7.9	7.84	0.06
2,2’,4,4’,5,6’-	454	8.53	-9.81	2.99	7.82	7.71	0.11
2,2’,3,3’,4,4’,5,5’,6,6’-	509	12.74	-9.79	4.55	10	9.94	0.06
diphenyl ether	600	2.48	-9.48	0.72	3.97	4.19	-0.22

Table 4.9 (continued).

Compound name	Congener number	*CRI*	E_{HOMO}	E_{aq}	log K_{ow} obs	log K_{ow} pred from Eq. 4.5	res
Test Set							
3-	2	2.99	-9.41	0.45	4.75	4.68	0.07
2,3-	5	3.40	-9.34	0.27	5	5.04	-0.04
2,4'-	8	3.41	-9.35	0.23	5.03	5.08	-0.05
2,4,6-	30	3.86	-9.45	0.09	5.32	5.61	-0.29
2,2',4,5-	48	4.36	-9.00	-0.19	5.97	5.73	0.24
2,3,4,5-	61	4.24	-9.02	-0.15	6.01	5.64	0.37
2,3', 4,5-	67	4.36	-9.26	-0.25	6.14	6.03	0.11
2,3',4,5'-	68	4.39	-9.44	-0.28	6.13	6.24	-0.11
3,3',4,5'-	79	4.37	-9.36	-0.32	6.22	6.16	0.06
2,2',4,4',6-	100	4.79	-9.11	-0.40	6.11	6.30	-0.19
2,3,3',4,4'-	105	4.73	-9.37	-0.48	6.51	6.55	-0.04
2,3,4,4',5-	114	4.74	-9.31	-0.43	6.61	6.48	0.13
2,3,4,4',6-	115	4.77	-9.16	-0.38	6.47	6.32	0.15
2,3',4,5,5'-	120	4.81	-9.31	-0.50	6.66	6.57	0.09
2,3',4,4',5'-	123	4.78	-9.16	-0.48	6.63	6.39	0.24
2,2',3,3',4,4'-	128	5.19	-9.15	-0.63	6.82	6.79	0.03
2,2',3,3',4,5'-	130	5.21	-9.20	-0.65	7.01	6.86	0.15

Table 4.9 (continued).

Compound name	**Congener number**	***CRI***	E_{HOMO}	E_{aq}	**log K_{ow} obs**	**log K_{ow} pred from Eq. 4.5**	**res**
2,2’,3,4,4’,6-	139	5.22	-9.14	-0.59	6.84	6.78	0.06
2,2’,4,4’,5,5’-	153	5.23	-9.08	-0.66	6.72	6.77	-0.05
2,2’,4,4’,5,6’-	154	5.36	-9.05	-0.63	6.49	6.82	-0.33
2,2’,3,3’,4,5’,6’-	177	5.67	-9.19	-0.79	7.14	7.30	-0.16
2,2’,3,3’,4,4’,5,6-	195	5.38	-9.22	-0.96	7.84	7.21	0.63
2,2’,3,4,4’,5,5’,6-	203	6.11	-9.10	-0.98	7.81	7.66	0.15
2,2’,3,3’,4,4’,5,5’,6-	206	6.52	-9.16	-1.17	8.07	8.15	-0.08
2,2’,3,3’,4,4’,5,5’,6,6’-	209	6.91	-9.26	-1.21	8.16	8.58	-0.42
2,2’,4,4’-	347	6.50	-9.70	2.06	6.81	6.64	0.17
2,2’,3,4,4’-	385	7.55	-9.75	2.48	7.37	7.22	0.15
2,2’,3,4,4’,5’,6-	483	9.55	-9.90	3.40	8.27	8.32	-0.05

4.2.1. Analysis of the Applicability Domain

The AD of the developed model for the combined data set exploited by Williams plot is shown in Figure 4.6. It can be seen that h_i values of all compounds are lower than the warning value (h^* = 0.129) except for 2,2',3,3',4,4',5,5',6,6'-BDE whose maximum descriptor values (the *CRI* and E_{HOMO}) and maximum log K_{ow} values within the data set leads to it far from the centroid of the descriptor space. However, its standardized residual is nearly zero, thus it stabilizes the model and makes the model more precise. 2,2',3,4,4',5',6-BDE whose E_{HOMO} value falls out of E_{HOMO} values range of the training set. This compound is from test set. For all the compounds in the training and test sets, their standardized residuals are smaller than three standard deviation units, i.e., there is no response outlier for the developed QSPR model. Therefore, the developed QSPR model was used to predict log K_{ow} for PBDEs and PCDEs.

Due to different amount of studied compounds or differences in the development of the models, it is not possible to perform strict comparisons among our proposed models and already published QSPRs for PBDEs and PCDEs. However, some general considerations can still be made. As shown in Table 4.11, our models had comparable or even better performances in comparison to other QSPRs at the same level of complexity (three descriptor) or higher.

Table 4.10 Boundaries of the proposed models log K_{ow} values of PBDEs/PCDEs (Eq. 4.5).

	Training set		**Test set**	
	min	**max**	**min**	**max**
log K_{ow}	3.97 (diphenyl ether)	10 (2,2',3,3',4,4',5,5',6,6'-BDE)	4.75 (3-CDE)	8.27 (2,2',3,4,4',5',6-BDE)
CRI	2.48 (diphenyl ether)	12.74 (2,2',3,3',4,4',5,5',6,6'-BDE)	2.99 (3-CDE)	9.55 (2,2',3,4,4',5',6-BDE)
E_{aq}	-1.03 (2,2',3,3',4,4',5,5'-CDE)	4.55 (2,2',3,3',4,4',5,5',6,6'-BDE)	-1.21 (2,2',3,3',4,4',5,5',6,6'-CDE)	3.40 (2,2',3,4,4',5',6-BDE)
E_{HOMO}	-9.81 (2,2',4,4',5,6'-BDE)	-9.00 (2,2',4,5,6'-CDE)	-9.90 (2,2',3,4,4',5',6-BDE)	-9.00 (2,2',4,5-CDE)

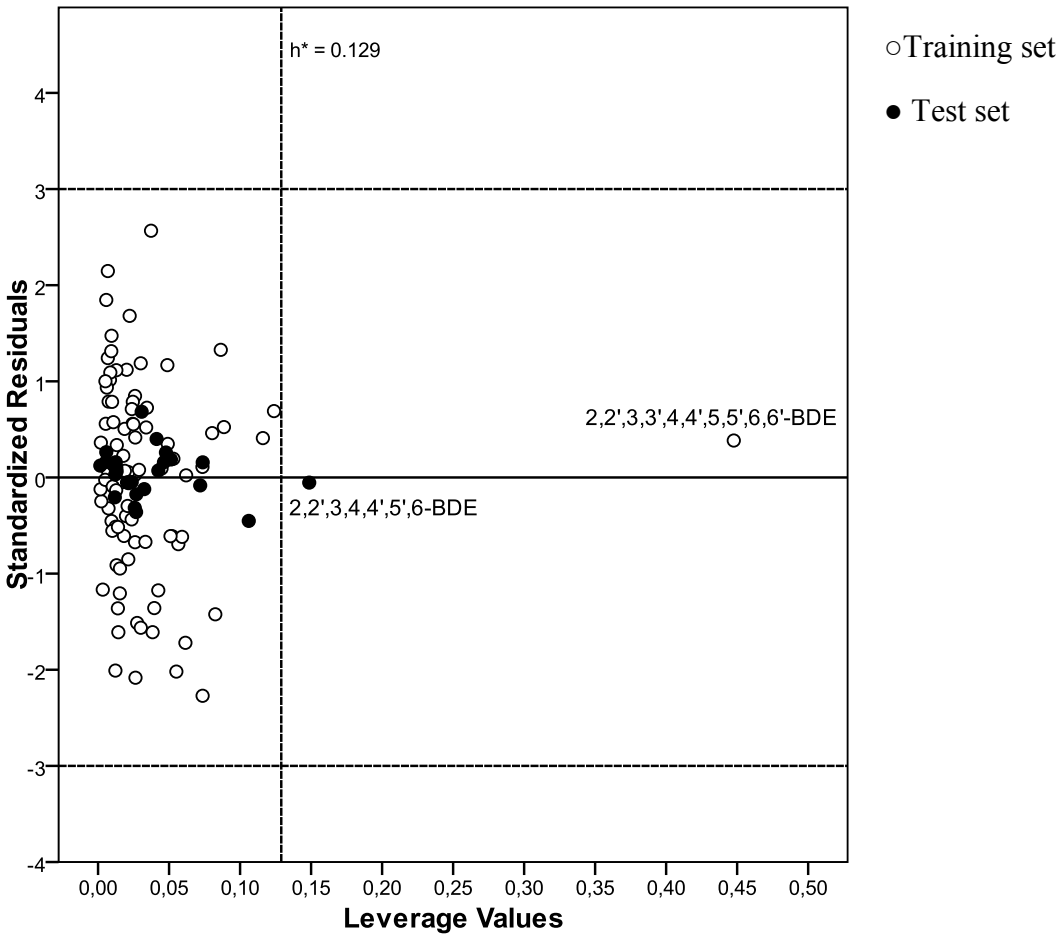


Figure 4.6 Williams plot for the Eq. 4.5. The log K_{ow} values for the training and test chemicals are labeled differently. The dotted lines are the 3.0σ limit and the warning value of hat (h^* = 0.129), respectively.

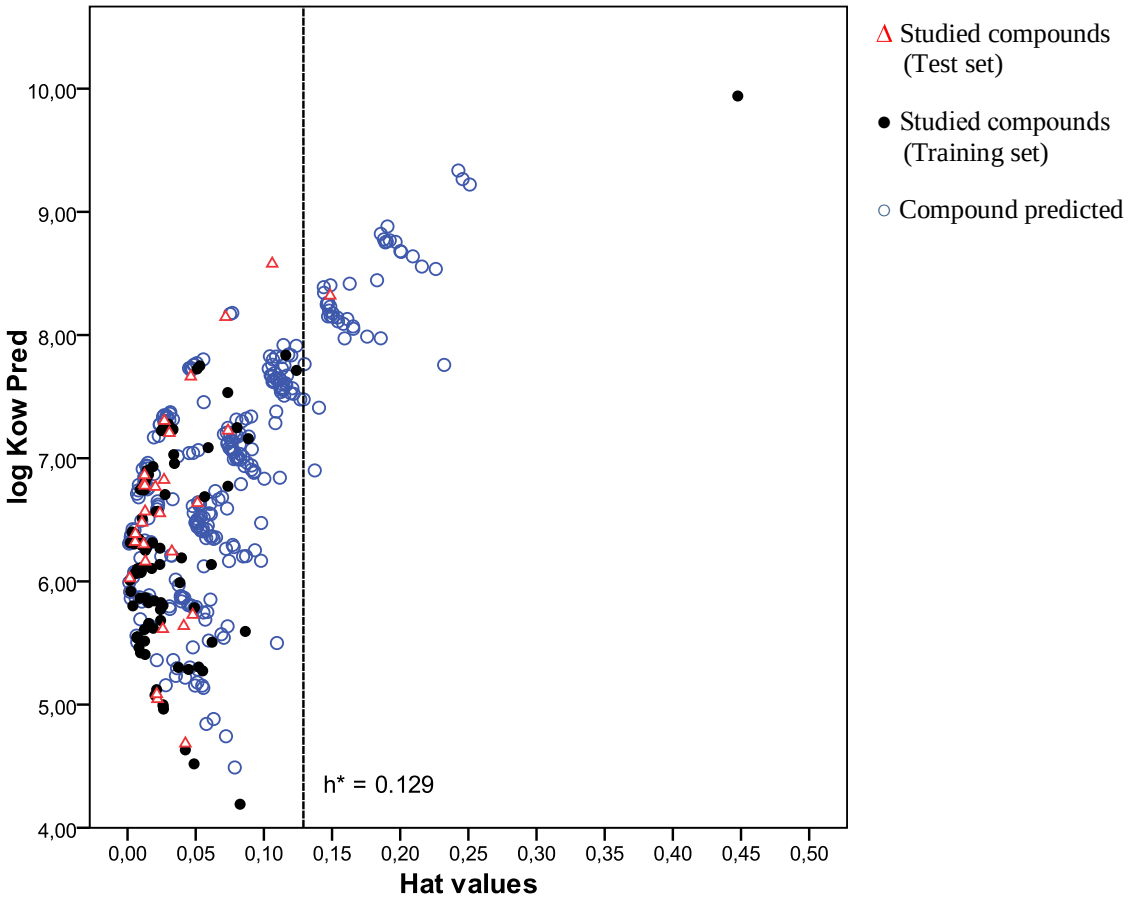


Figure 4.7 Plot of *hat* values vs. log K_{ow} predicted values of PBDEs/PCDEs (Eq. 4.5).

The model predictions for the chemicals with the leverage values of higher than h^* = 0.129 reported in Figure 4.7 are considered to be out of the domain of the respective model and predicted values of those chemicals should be considered potentially unreliable, since they are extrapolated (congener number marked with an asterisk (*) in Appendix B, Table B.2). Bromination degrees of extrapolated PBDE congeners are from 4 to 9.

21% of PBDE congeners for log K_{ow} predictions fell outside the AD of Eq. 4.5 and should be considered as extrapolations (Figure 4.7). The reason for this can be related to the limited number of chemicals available for this endpoint (log K_{ow} of PBDEs). On the other hand, the Eq. 4.5 covers a large part of the log K_{ow} prediction of PBDEs (79%).

4.2.2. Comparison with the Reported Methods

There is no reported model for the combined data set. Therefore, we applied the same computational procedure to generate a QSPR models for PBDEs and PCDEs separately. Hence, we compared each model obtained for PBDE and PCDE with the reported models.

It is important to note that the one-parameter models for PBDE by Wania and Dugani (2003), and Braekevelt (2003) were built by using only very simple properties (molar mass and number of bromine substituents) as molecular descriptors. These descriptors, whose value is constant for isomeric compounds, calculate constant predictions for each group of PBDE congeners. Even though these models are very simple, they are not sensitive at all to variations in the responses, which are related to other structural properties such as the position of the bromine atoms on the phenyl ring.

Papa et al. (2009) obtained a QSPR model for log K_{ow} values of BFRs with T(O...Br) (R^2 = 0.97). No outliers were detected in the log K_{ow} model response domain. Yang et al. (2003) used MOPAC software for computing 12 quantum chemical descriptors. PLS regression was performed to build the QSPR model for log K_{ow} values of PCDEs. Sun et al. (2007) developed QSPR Model to predict log K_{ow} values of PCDEs by using MLR method with molecular electronegativity distance vector (MEDV-4) descriptor. In their method, they did not separate the data set (107) into training and test set. They concluded that log K_{ow} increases with the degree of chlorination in general. Chen et al. (2007) modeled log K_{ow} values of PCDEs by the method of Cl substitution position. Stepwise MLR has been used to construct the QSPR models by using six elements (the numbers of positions of Cl substitution (N_{PCS})).

Xu et al. (2010) modeled the structural descriptors of PCDEs and the physico-chemical properties and biological activity by stepwise linear regression analysis. They randomly split these data into calibration and test sets in the ratio of 2:1 (72 vs. 35) to evaluate further the predictive power of the log K_{ow} model. By using stepwise linear regression analysis, the relationship has been established for the calibration set. They obtained model for log K_{ow} with 3 descriptors, namely N_{Cl}, N^{+}_{V}, $V^{-}_{s,av.}$ With the R^2 = 0.974;

R^2_{cv} = 0.972 without dividing data set into training and test set (N = 107); and R^2 = 0.976; R^2_{cv} = 0.972 by dividing the data set into training and test set (N = 72).

Li et al. (2008) used PLS regression method for modeling the log K_{ow} values of PBDEs based on quantum molecular descriptors. Theoretical molecular descriptors were tested against log K_{ow} values for PBDEs using the PLS regression method which can be used to analyze data with many variables and few observations. The values of log K_{ow} for PBDEs are mainly governed by molecular surface area, energy of the lowest unoccupied molecular orbital and the net atomic charges on the oxygen atom. All these descriptors have been discussed to interpret the partitioning mechanism of PBDE chemicals. The bulk property of the molecules represented by molecular surface area is the leading factor, and K_{ow} values increase with the increase of molecular surface area. In conclusion it was stated that higher energy of the lowest unoccupied molecular orbital and higher net atomic charge on the oxygen atom of PBDEs result in smaller K_{ow}.

The models developed for each of PBDEs and PCDEs are superior to the literature models in terms of mechanistic interpretation of descriptors appearing in the relevant models.

Table 4.11 Statistical performance comparison of different QSPR models of log K_{ow} for PBDEs/PCDEs.

Chemical Group	Number of descriptors	Descriptor type	Training/ test set	Method[#]	R^2	*SE*	*F*	*N**	R^2_{pred}	References
PBDEs/PCDEs	**3**	***CRI*; E_{aq}; E_{HOMO}**	**93/28**	**MLR**	**0.970**	**0.155**	**49.749**	**121**	**0.940**	**Current study**
PCDE	**2**	**H2e; Mor08v**	**79/28**	**MLR**	**0.989**	**0.090**	**6.330**	**107**	**0.976**	**Current study**
PBDE	**2**	***CRI*; E_{HOMO}**	**—**	**MLR**	**0.993**	**0.113**	**728.289**	**14**	**—**	**Current study**
PBDEs	3	SAG; E_{LUMO}; q^-_o	—	PLS	0.989	—	—	9	—	Li et al., (2008)
PBDEs	1	M_w	—	LRTP	0.975	—	—	6	—	Wania and Dugani, 2003
PBDEs	1	nBr	—		0.970	—	—	9	—	Braekvelt et al., (2003)
PCDEs	3	ΔH_f; α; M_w	—	PLS	0.976	—	—	107	—	Yang et al., (2003)
PCDEs	4	(MEDV-4 type) M_{11}; M_{13}; M_{22}; M_{33}	—	MLR	0.984	0.101	—	107	—	Sun et al., (2007)
PCDEs	3	$N_{2(6)}$; $N_{3(5)}$; N_4	72/35	MLR	0.984	0.10	—	107	—	Chen et al., (2007)
PCDEs	3	$^0\chi_A$; $^1\chi_A$; $^2\chi_A$	—	MCI	0.894	0.18	—	107	—	Huang et al., (2004)
PCDEs	3	N_{Cl}, N^+_V, $V^-_{s,av}$	72/35	MLR	0.976	—	—	107	0.972	Xu et al., (2010)

*Here *N* donates the total number of chemicals in the data set. [#]MCI: Molecular Connectivity Indices; PLS: Partial Least-Squares; LRTP: Long-Range Transport Potential assessment methods.

4.3. Modeling log *RBA* for PBDEs

MLR models were developed for log *RBA* of 18 PBDE congeners. The heuristic correlations were performed for the whole set provided the optimal equations for different numbers of descriptors in the range of 1-4. To avoid the over-parameterization of the model, an increase of the R^2 value of less than 0.02 was chosen as the breakpoint criterion. From the viewpoint of statistics, the number of samples should be 4-5 times of that of variables for linear regression. Plotting of R^2 and R^2_{cv} values against the number of descriptors (Figure 4.8) which provide guidance regarding the number of descriptors to retain in the models suggested that the best model contained four parameters. Our approach was based on the parsimony principle, which implies, inter alia, that the ratio of observations to variables should be as high as possible and at least 5:1 (Topliss and Costello, 1972; Cronin and Schultz, 2003). This approach is efficient to reduce the chance of overfitting, which increases with the increase in the number of the variables included in the models and gives an overoptimistic idea of their predictive ability.

The experimental versus calculated log *RBA* of PBDEs and the residuals of the best regression model (Table 4.12) indicated that the best regression model (containing four descriptors, Eq. 4.8; Model 30) was very satisfactory.

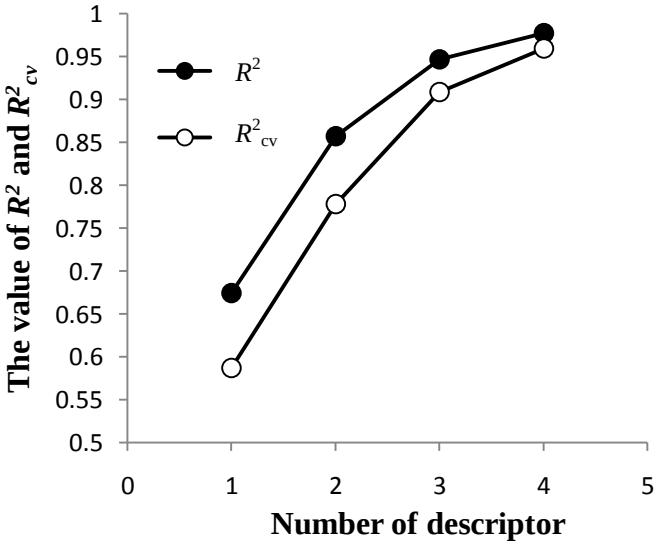


Figure 4.8 Correlation coefficients (R^2, R^2_{cv}) versus number of descriptors.

The descriptors involved in these correlations and the regression coefficients were also listed in Table 4.12.

$$\log RBA = -1.751\ (\pm 0.078) - 0.392\ (\pm 0.027)\ \text{RDF075p} + 0.725\ (\pm 0.049)\ \text{Mor28m} + 8.327\ (\pm 1.253)\ \text{XY-S/R} + 0.482\ (\pm 0.098)\ \text{RDF090u} \quad (4.8)$$

(Model 30)

$n = 18, \quad R^2 = 0.977, \quad F = 139.97, \quad SE = 0.017, \quad R^2_{cv} = 0.959$

The *t*-values are -14.5097, 14.7810, 6.6425 and 4.9089 for RDF075p, Mor28m, XY-S/R and RDF090u, respectively. The *t*-values indicate that each parameter is highly significant ($p<0.05$). Mor28m is the most important descriptor for predicting log *RBA* considering the *t*-values. This model explains more than 97% of the variance in the experimental log *RBA* values for PBDEs.

3D-MoRSE descriptors (Schuur et al., 1996) are based on the idea of obtaining information from the 3D atomic coordinates by the transform used in electron diffraction studies for preparing theoretical scattering curves. The derived expression is shown in Eq. 4.9.

$$I(s) = \sum_{i=1}^{A-1} \sum_{j=i+1}^{A} w_i w_j \frac{\sin(s r_{ij})}{s r_{ij}} \quad (4.9)$$

where $I(s)$ is the scattered electron intensity, w an atomic property (e.g., the atomic number), r_{ij} the interatomic distance between the ith and the jth atoms, and A the number of atoms. Radial distribution function (RDF) descriptors (Raevsky et al., 2000) are based on the distance distribution in the geometrical representation of a molecule and constitute a RDF code that shows certain characteristics in common with the 3D-MoRSE descriptors. RDF075p means Radial distribution function – 7.5/weighted by atomic polarizabilities. RDF090u means Radial Distribution Function – 9.0 / unweighted. Mor28m means – signal 28 / weighted by atomic masses.

Descriptors that belong to the class of RDF descriptors (Hemmer et al., 1999) are based on the distance distribution in the geometrical representation of the molecule. In addition to interatomic distances in the entire molecule, the RDF also provides valuable information about bond distances, ring types, planar and non-planar systems, atom types and other important structural motifs. By using different weighting schemes, which include atom types, electronegativity, atom mass or van der Waals radii, RDF can be adjusted to select among those atoms of molecule, which give rise to an important descriptor in deriving an appropriate QSAR.

Formally, the RDF of a molecule of A atoms can be interpreted as the probability distribution of finding an atom in a spherical volume of radius R. In this sense, according to our model a spherical molecular volume with these dimensions could have certain restrictions to the addition of bulky substituents.

XY-S/R is the projection of the molecules onto the XY-plane oriented in space along the axes of inertia and represents the size of the molecule along the longest axis. XY Shadow appeared in the model proposed by Colombo et al., 2008. Colombo et al. (2008) studied the toxicity of 568 industrial organic compounds and their acute toxicity data expressed as lethal concentration (LC_{50}) in 96-flow-through exposures for the juvenile stage of the Fathead minnow.

The Compound names, descriptor values, observed and predicted log *RBA* of PBDEs, and residuals obtained from Eq. 4.8 are given in Table 4.13.

Table 4.12 Comparative statistics of QSTR models based on Heuristic method for log *RBA* of PBDEs.

Model Number	Number of descriptors	Variables	R^2	*F*	*SE*	R^2_{cv}
		DRAGON-based models				
19	1	RDF075u	0.673	32.98	0.200	0.586
20	2	QYYp; Mor28p	0.858	45.18	0.093	0.779
21	3	QYYp; Mor28p; SPAM	0.935	67.21	0.045	0.877
22	4	QYYp; Mor28p; Mor09m; Mor26m	0.975	127.47	0.019	0.948
		CODESSA-based models				
23	1	WPSA1	0.452	13.20	0.336	0.319
24	2	P''; Information content (order 2)	0.672	15.34	0.215	0.584
25	3	P''; Information content (order 2); S_{YZ}	0.811	19.98	0.133	0.683
26	4	P''; Information content (order 2); Yz-S/R; WPSA3	0.873	22.41	0.096	0.722
		DRAGON-CODESSA-SPARTAN-*CRI* based models				
27	1	RDF075u	0.674	33.08	0.200	0.587
28	2	QYYp; Mor28p	0.857	45.04	0.093	0.778
29	3	RDF075p; Mor28m; V_{XYZ}	0.947	82.68	0.037	0.909
30	**4**	**RDF075p; RDF090u; Mor28m; XY-S/R**	**0.977**	**139.97**	**0.017**	**0.959**

Table 4.13 The Compound names, descriptors, observed and predicted log *RBA* of PBDEs obtained from Eq. 4.8.

Congener number	Compound name	RDF090u	RDF075p	Mor28m	XY-S/R	Observed value of log *RBA*	Predicted value from Eq. 4.8	res
3	4-	1.268	1.124	0.223	0.642	3.89	3.93	-0.04
15	4,4'-	0.437	3.529	0.672	0.708	3.42	3.45	-0.03
17	2,2',4-	1.509	2.398	0.315	0.619	3.64	3.42	0.22
28	2,4,4'-	1.122	3.283	0.193	0.629	2.92	2.88	0.04
47	2,2',4,4'-	1.312	3.331	0.265	0.675	3.25	3.39	-0.13
49	2,2',4,5'-	2.076	1.679	0.227	0.657	4.17	4.23	-0.05
66	2,3',4,4'-	0.377	5.104	0.964	0.661	2.70	2.63	0.07
71	2,3',4',6-	1.245	4.198	1.410	0.688	3.87	3.96	-0.09
75	2,4,4',6-	1.122	3.283	1.066	0.630	3.40	3.52	-0.12
77	3,3',4,4'-	0.776	8.058	1.100	0.761	2.66	2.60	0.06
85	2,2',3,4,4'-	0.919	6.482	-0.198	0.697	1.72	1.81	-0.09
99	2,2',4,4',5-	1.195	2.929	0.968	0.647	3.85	3.76	0.09
100	2,2',4,4',6-	1.312	3.331	1.384	0.655	4.11	4.03	0.08
119	2,3',4,4',6-	0.377	5.104	1.651	0.653	2.96	3.06	-0.10
126	3,3',4,4',5-	0.657	7.705	2.011	0.680	2.57	2.66	-0.09
153	2,2',4,4',5,5'-	1.029	2.459	1.793	0.652	4.60	4.51	0.09
154	2,2',4,4',5,6'-	1.195	2.929	2.069	0.671	4.64	4.76	-0.13
183	2,2',3,4,4',5',6-	0.753	6.011	1.555	0.720	3.60	3.38	0.23

4.3.1. Comparison with the Reported Methods

Since data set of log *RBA* was small, validation steps proposed by OECD principles 4 were not performed. AD analysis with Williams plot was not performed either. Only comparison with literature was done to compare our log *RBA* model.

A strict comparison between our models and the already published QSARs for log *RBA* (Wang et al., 2005; 2006; Zheng et al., 2007; Xu et al., 2007; Papa et al., 2010) was not possible either, due to the different amounts of studied compounds or differences in the development of the models; however, some general considerations can be made. Table 4.11 shows that our log *RBA* model had comparable or, in some cases, better performances than other more complex QSTRs.

Papa et al. (2010) studied the endocrine-disrupting potencies of BFRs including PBDEs. They modeled log *RBA* endpoint using L1v and Mor22u DRAGON descriptors by MLR method. L1v is WHIM descriptor (Todeschini and Gramatica, 1997), Mor22u is 3D-MoRSE group (Schuur et al., 1996). L1v is defined as the first component size directional WHIM index weighted by atomic van der Waals volumes and provides information about the distribution of the molecular size along the first principal direction of the molecule. Mor22u (Morse signal no. 22 unweighted) and Mor08e (Morse signal no. 08 weighted by the Sanderson electronegativity) provide 3D information related to the weights of the atoms in the structure, as viewed by an angular scattering function. The values of these functions are calculated at 32 evenly distributed values of scattering angles of 0-32 Å-1 from the 3D atomic coordinates of a molecule (Schuur et al., 1996).

The general tendency that we observed in other existing models with only a few exceptions was to use more than two variables and complex modeling techniques. A comparison of the models proposed here with other QSTRs demonstrated that our models have comparable or higher fitting performance than the existing ones, which are in general more complex in terms of number and type of variables or modeling methods such as PLS even for small data sets (Wang et al., 2007).

This study and the study done by Papa et al. (2010) resulted in common 3D descriptor in the models obtained for log *RBA* values of PBDEs. The type of the descriptors is different because the data set are different in these two studies. While we modeled only for PBDEs, Papa et al. (2010) modeled log *RBA* values for BFRs.

Table 4.14 Comparison of the performances of the QSAR models developed in this study for the end point log *RBA* for PBDEs with other models existing in literature.

Chemical Group	no. of variables or PLS components	Descriptor type	Training/ test set	Method	R^2	*SE*	*F*	*N**	References
PBDEs	**4**	**RDF075p; RDF090u; Mor28m; XY-S/R**	**—**	**MLR**	**0.977**	**0.017**	**139.97**	**18**	**Current study**
BFRs	2	L1v; Mor22u	—	MLR	0.82	*RMSE*: 0.31	—	18	Papa et al. (2010)
PBDEs	4	σ^2_{tot}; N_v; $\overline{V_s}$; $V_{s,min}$	—	MLR	0.647	0.515	5.97	18	Xu et al. (2007)
PBDEs	4	DIP; D (ZX shadow); ER(C–C); B (Balaban index)	—	MLR	0.90	0.270	30.20	18	Wang et al. (2006)
PBDEs	6	—	—	CoMFA (PLS)	0.995	0.057	—	18	Wang et al. (2005)
PBDEs	6	—	—	CoMSIA (PLS)	0.982	0.105	—	18	Wang et al. (2005)

*Here *N* donates the total number of chemicals in the data set.

5. CONCLUSIONS

In this study, new predictive QSAR/QSPR models were developed for several physico-chemical properties of the 209 PBDEs and 209 PCDEs congeners along with the diphenyl ether molecule.

We calculated for the first time a large number of descriptors for PBDE and PCDE using DRAGON 5.4, CODESSA 2.2-SPARTAN 06 software and the *CRI* program. We used an efficient variable selection procedure like Heuristic Method; and a training/test set splitting methodology like Kohonen networks. The obtained structural parameters were taken as theoretical descriptors to correlate three QSPR models for predicting log K_{oa} and log *RBA* of PBDEs and log K_{ow} of PBDEs and PCDEs. We compared the outputs of linear modeling applied to these data sets with the previously published models.

The contribution of this study is to develop new, robust, comprehensive and validated MLR model for the log K_{ow} of PBDE/PCDE along with the diphenyl ether. In other words, this study highlights modeling log K_{ow} values of combined sets of PBDEs/PCDEs which were never done before. The models developed for log K_{oa} and log K_{ow} indicated that the topology based *CRI* was the most important parameter in modeling these parameters. The highlighted three-descriptor log K_{ow} model included the *CRI*, E_{aq} and E_{HOMO} with R^2 values as high as 0.970, whereas one-descriptor log K_{oa} model for PBDE included only the *CRI*. The *CRI* represents the size of the molecule, branching and global molecular properties, such as size, volume and surface area are important in the prediction of 97hysic-chemical properties. The *CRI* relates hydrophobicity since the bigger the molecules, the greater the *CRI* and thus the greater the hydrophobicity.

In the octanol–water system, octanol molecules are easier to release electrons than water molecules, since octanol have smaller E_{HOMO} values. The lower the E_{HOMO} values, the greater is the tendency of chemicals to donate electrons in intermolecular interactions, the greater is the intermolecular interactions between octanol molecules and PBDEs/PCDEs the higher the log K_{ow} values. In other words, E_{HOMO} seems particularly related to the tendency to dissolve in octanol phase.

The lower the E_{aq} values, the greater the tendency of PBDE/PCDE molecules to stay in water phase the greater the intermolecular interactions between PBDE/PCDE and water molecules, and thus the lower the log K_{ow} value. The cross validated R^2_{cv} values for models are all high, indicating a good predictive ability and stability. The proposed models have been proved to fulfill the fundamental points set down by OECD principles for regulatory QSAR acceptability.

A particular attention was paid to validation of models obtained, also external validation if it is applicable, and to the definition of their applicability domain. Reliable predictions for the studied endpoints were provided for all the PBDEs/PCDEs belonging to the chemical domain of the models, which was higher than 90% for log K_{oa} values of PBDEs and 79% for log K_{ow} values of PBDEs/PCDEs.

The comparison of the proposed models with other QSAR/QSPR models demonstrated that our models have comparable or higher fitting performance than the existing ones, which are in general more complex in terms of variables or modeling method.

Depending on the predictive power of the *CRI* based models which were built in this study, log K_{oa} and log K_{ow} of some other chemicals not being used in the data set and having environmental importance can be predicted through these models. These compounds might include PBDE and PCDE congeners – with unknown properties – and PCBs (polychlorinated biphenyls), polychlorinated benzenes, etc. that fall into the AD of the models. Therefore, these models can be used to fill data gaps according to the new REACH regulation, facilitating the screening and prioritization of chemicals and for the identification of more problematic compounds even before their synthesis. The information from these models could be also used for the design of safer alternatives to dangerous PBDEs which is a subgroup of brominated flame retardants and PCDEs which are quite resistant to degradation, persistent in the environment, and bioaccumulate in aquatic media.

Those properties estimated in this study can be used to determine the distribution of PBDEs and PCDEs in the environment, thus are used in modeling the environmental fate of these compounds.

Although a four-parameter log *RBA* model for PBDE has been developed, it couldn't be validated according to the OECD principles. The descriptors appeared in this model were from DRAGON and CODESSA. Therefore, this study also highlights the urgent need to have more experimental data on log *RBA* and toxicity of PBDEs.

Compared with the direct measured methods, the developed models need only chemical structure data which can be easily calculated using chemical software.

REFERENCES

Ballschmiter, K., Zell, M., 1980. Analysis of polychlorinated biphenyls (PCB) by glass capillary gas chromatography. Composition of technical Aroclor- and Clophen-PCB mixtures. Fresenius' Journal of Analytical Chemistry, 302, 20-31.

Ballschmiter, K., Rappe, C., Buser, H. R., 1989. Chemical properties, analytical and environmental levels of PCBs, PCTs, PCNs and PCBs. In "halogenated biphenyls, terphenyls, naphthalenes, dibenzodioxins and related products" (R. D. Kimbrough and A. A. Jensen, Eds.), 47-69. Elsevier-North Holland, Amsterdam.

Becker, M., Phillips, T., Safe, S., 1991. Polychlorinated diphenyl ethers – a review. Toxicological and Environmental Chemistry, 33, 189-200.

Bennett, D. H., Scheringer, M., McKone, T. E., Hungerbühler, K., 2001. Predicting long-range transport potential: A systematic evaluation of two multimedia transport models. Environmental Science & Technology, 35, 1181-1189.

Birnbaum, L. S., Staskal, D. F., 2004. Brominated flame retardants: Cause for concern? Environmental Health Perspectives, 112, 9-17.

Bocio, A., Llobet, J. M., Domingo, J. L., 2004. Human exposure to polychlorinated diphenyl ethers through the diet in Catalonia, Spain. Journal of Agricultural and Food Chemistry, 52, 1769-1772.

Braekevelt, E., Tittlemier, S. A., Tomy, G. T., 2003. Direct measurement of octanol-water partition coefficients of some environmentally relevant brominated diphenyl ether congeners. Chemosphere, 51, 563-567.

Branchi, I., Alleva, E., Costa, L. G., 2002. Effects of perinatal exposure to a polybrominated diphenyl ether (PBDE 99) on mouse neurobehavioural development. Neurotoxicology, 23, 375-384.

Branchi, I., Capone, F., Alleva, E., Costa, L. G., 2003. Polybrominated diphenyl ethers: neurobehavioral effects following developmental exposure. Neurotoxicology, 24, 449-462.

BSEF, 2004. Bromine Science and Environmental Forum, www.bsef.com (Accessed August 2011).

Buser, H. R., 1986. Polybrominated dibenzofurans and dibenzo-p-dioxins: Thermal reaction products of polybrominated diphenyl ether flame retardants. Environmental Science & Technology, 20, 404-408.

Buser, H. R., 1987. Brominated and brominated/chlorinated dibenzodioxins and dibenzofurans: Potential environmental contaminants. Chemosphere, 16, 713-732.

Caballero, J., Fernandez, M., 2006. Linear and nonlinear modeling of antifungal activity of some heterocyclic ring derivatives using multiple linear regression and Bayesian-regularized neural networks. Journal of Molecular Modeling, 12, 168-181.

Carlson, G. P., 1980a. Induction of xenobiotic metabolism in rats by brominated diphenyl ethers administered for 90 days. Toxicology Letters, 6, 207-212.

Carlson, G. P., 1980b. Induction of xenobiotic metabolism in rats by short-term administration of brominated diphenyl ethers. Toxicology Letters, 5, 19-25.

Castilho, M. S., Guido, R. V. C., Andricopulo, A. D., 2007. 2D Quantitative structure-activity relationship studies on a series of cholesteryl ester transfer protein inhibitors. Bioorganic & Medicinal Chemistry, 15, 6242-6252.

Chambers, C. C., Hawkins, G. D., Cramer, C. J., Truhlar, D. G., 1996. Model for aqueous solvation based on class IV atomic charges and first solvation shell effects. The Journal of Physical Chemistry, 100, 16385-16398.

ChemBio3D Ultra 12.0, CambridgeSoft Life Science Enterprise Solutions, Cambridge, MA (2010).

Chen, G. S., 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 & Technology, 35, 3749-3756.

Chen, J. W., Quan, X., Zhao, Y. Z., Yang, F. L., Schramm, K. W., Kettrup, A., 2001. Quantitative structure–property relationships for octanol–air partition coefficients of PCDD/Fs. Bulletin of Environmental Contamination and Toxicology, 66, 755-761.

Chen, J. W., Xue, X. Y., Schramm, K. W., Quan, X., Yang, F., Kettrup, A., 2002a. Quantitative structure–property relationships for octanol–air partition coefficients of polychlorinated biphenyls. Chemosphere, 48, 535-544.

Chen, J. W., Harner, T., Schramm, K. W., Kettrup, A., 2002b. Quantitative predictive models for octanol–air partition coefficients of persistent organic pollutants at different environmental temperatures. Organohalogen Compounds, 57, 467-470.

Chen, J. W., Harner, T., Yang, P., Quan, X., Chen, S., Schramm, K. W., Kettrup, A., 2003. Quantitative predictive models for octanol–air partition coefficients of polybrominated diphenyl ethers at different temperatures. Chemosphere, 51, 577-584.

Chen, S. D., Zeng, X. L., Wang, Z. Y., Liu, H. X., 2007. QSPR modeling of *n*-octanol/water partition coefficients and water solubility of PCDEs by the method of Cl substitution position. Science of the Total Environment, 382, 59-69.

Coburn, J. A., Comba, M., 1981. Identification of polychlorinated diphenyl ethers in Whitby Harbour sediments. Association of Analytical Chemist's, Spring Workshop, Ottawa, Ontario.

Coburn, J. A., Comba, M., 1985. The development of an analytical method for polychlorinated diphenyl ethers in fish, final contract report. National Water Research Institute, Burlington, Ontario.

CODESSA 2.20 (1994-1996), Semichem Inc., Shawnee Mission, USA.

Colombo, A., Benfenati, E., Karelson, M., Maran, U., 2008. The proposal of architecture for chemical splitting to optimize QSAR models for aquatic toxicity. Chemosphere, 72, 772-780.

Consonni, V., Todeschini, R., Pavan, M., 2002a. Structure/response correlations and similarity/diversity analysis by GETAWAY descriptors. 1. Theory of the novel 3D molecular descriptors. Journal of Chemical Information and Computer Sciences., 42, 682-692.

Consonni, V., Todeschini, R., Pavan, M., Gramatica, P., 2002b. Structure/response correlations and similarity/diversity analysis by GETAWAY descriptors. 2. Application of the novel 3D molecular descriptors to QSAR/QSPR studies. Journal of Chemical Information and Computer Sciences, 42, 693-705.

Consonni, V., Ballabio, D., Todeschini, R., 2009. Comments on the definition of the Q^2 parameter for QSAR validation. The Journal of Chemical Information and Modeling, 49, 1669-1678.

Cronin, M. T. D., Schultz, T. W., 2003. Pitfalls in QSAR. Journal of Molecular Structure, THEOCHEM, 622, 39-51.

Cronin, M. T. D., Netzeva, T. I., Dearden, J. C., Edwards, R., Worgan, A. D. P., 2004. Assessment and modeling of the toxicity of organic chemicals to *Chlorella vulgaris*: development of a novel database. Chemical Research in Toxicology, 17, 545-554.

Cupples, A. M., Sanford, R. A., Sims, G. K., 2005. Dehalogenation of the herbicides bromoxynil (3,5-dibromo-4-hydroxybenzonitrile) and ioxynil (3,5-diiodino-4-hydroxybenzonitrile) by *Desulfitobacterium chlororespirans*. Applied and Environmental Microbiology, 71, 3741-3746.

Darnerud, P. O., Thuvander, A., 1999. Effects of polybrominated diphenyl ethers (PBDE) and polychlorinated biphenyl (PCB) on some immunological parameters after oral exposure in rats and mice. Toxicological and Environmental Chemistry, 70, 229-242.

Darnerud, P. O., Eriksen, G., Jóhannesson, T., Larsen, P., Viluksela, M., 2001. Polybrominated diphenyl ethers: occurrence, dietary exposure, and toxicology. Environmental Health Perspectives, 109, 49-68.

Darnerud, P. O., 2003. Toxic effects of brominated flame retardants in man and in wildlife. Environment International, 29, 841-853.

De Boer, J., Denneman, M., 1998. Polychlorinated diphenyl ethers: origin, analysis, distribution, and toxicity in the marine environment. Reviews of Environmental Contamination and Toxicology, 157, 131-144.

Devillers, J. (Ed.), 1996. Neural networks in QSAR and drug design, Academic Press Limited, London, UK.

De Wit, C. A., 2002. An overview of brominated flame retardants in the environment. Chemosphere, 46, 583-624.

Dimitrov, S., Dimitrova, G., Pavlov, T., Dimitrova, N., Patlewicz, G., Niemela, J., Mekenyan, O., 2005. A stepwise approach for defining the applicability domain of SAR and QSAR models. Journal of Chemical Information and Modeling, 45, 839-849.

Domingo, J. L., 2006. Polychlorinated diphenyl ethers (PCDEs): environmental levels, toxicity and human exposure a review of the published literature. Environment International, 32, 121-127.

Donnelly, J. R., Grange, A. H., Nunn, N. J., Sovocool, G. W., Brumley, W. C., Mitchum, R. K., 1989. Analysis of thermoplastic resins for brominated dibenzofurans. Biomedical & Environmental Mass Spectrometry, 18, 884-896.

El Dareer, S. M., Kalin, J. R., Tillery, K. F., Hill, D. L., 1987. Disposition of decabromobiphenyl ether in rats dosed intravenously or by feeding. Journal of Toxicology and Environmental Health, 22, 405-415.

Eriksson, P., Jakobsson, E., Fredriksson, A., 2001. Brominated flame retardants: A novel class of developmental neurotoxicants in our environment. Environmental Health Perspectives, 109, 903-908.

Eriksson, P., Viberg, H., Jakobsson, E., Örn, U., Fredriksson, A., 2002. A brominated flame retardant, 2,2',4,4',5-pentabromodiphenyl ether: Uptake, retention, and induction of neurobehavioral alterations in mice during a critical phase of neonatal brain development. Toxicological Sciences, 67, 98-103.

Eriksson, L., Jaworska, J. S., 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, 1361-1375.

Falco' G., Bocio, A., Llobet, J. M., Domingo, J. L., 2005. Health risks of dietary intake of environmental pollutants by elite sportsmen and sportswomen. Food and Chemical Toxicology, 43, 1713-1721.

Fowles, J. R., Fairbrother, A., Baecher-Steppan, L., Kerkvliet, N. I., 1994. Immunologic and endocrine effects of the flame-retardant pentabromodiphenyl ether (DE-71) in C57BL/6J mice. Toxicology, 86, 49-61.

Franke, R., 1984. Theoretical drug design methods. Elsevier, Amsterdam, 7, 412.

Geyer, H. J., Schramm, K. W., Darnerud, P. O., Aune, M., Feicht, E. A., Fried, K. W., Henkelmann, B., Lenoir, D., Schmid, P., McDonald, T. A., 2004. Terminal elimination half-lives of the brominated flame retardants TBBPA, HBCD, and lower brominated PBDEs in humans. Organohalogen Compounds, 66, 3867-3872.

Ghasemi, J., Saaidpour, S., 2007. Quantitative structure–property relationship study of *n*-octanol/water partition coefficients of some of diverse drugs using multiple linear regression. Analytica Chimica Acta, 604, 99-106.

Gill, U., Chu, I., Ryan, J. J., Feeley, M., 2004. Polybrominated diphenyl ethers: Human tissue levels and toxicology. Reviews of Environmental Contamination and Toxicology, 182, 55-96.

Golbraikh, A., Tropsha, A., 2002. Beware of q^2! Journal of Molecular Graphics and Modelling, 20, 269-276.

Golbraikh, A., Shen, M., Xiao, Z., Xiao, Y., Lee, K., Tropsha, A., 2003. Rational selection of training and test sets for the development of validated QSAR models. Journal of Computer-Aided Molecular Design, 17, 241-253.

Gouin, T., Mackay, D., Webster, E., Wania, F., 2000. Screening chemicals for persistence in the environment. Environmental Science & Technology, 34, 881-884.

Gramatica, P., 2007. Principles of QSAR models validation: Internal and external. QSAR and Combinatorial Science, 26, 694-701.

Gu, C. G., Ju, X. H., Jiang, X., Wang, F., Yang, S. G., Sun, C., 2009. DFT study on the bromination pattern dependence of electronic properties and their validity in quantitative structure–activity relationships of polybrominated diphenyl ethers. SAR and QSAR in Environmental Research, 20, 287-307.

Hake, C. L., Rowe, V. K., 1963. Chlorinated phenyl ethers. In: Patty FA (ed) Industrial hygiene and toxicology, 2nd rev Edn., Vol II. Interscience, New York, 1706.

Hakk, H., Larsen, G., Klasson Wehler, E., Örn, U., Bergman, Å., 1999. Tissue disposition, excretion, and metabolism of 2,2',4,4',5-pentabromodiphenyl ether (BDE-99) in male Sprague-Dawley rats. Organohalogen Compounds, 40, 337-340.

Hakk, H., Letcher, R. J., 2003. Metabolism in the toxicokinetics and fate of brominated flame retardants – a review. Environment International, 29, 801-828.

Hall, L. H., Kier, L. B., 2001. Issues in representation of molecular structure the development of molecular connectivity. Journal of Molecular Graphics and Modelling, 20, 4-18.

Hallgren, S., Sinjari, T., Hakansson, H., Darnerud, P. O., 2001. Effects of polybrominated diphenyl ethers (PBDEs) and polychlorinated biphenyls (PCBs) on thyroid hormone and vitamin A levels in rats and mice. Archives of Toxicology, 75, 200-208.

Hallgren, S., Darnerud, P. O., 2002. Polybrominated diphenyl ethers (PBDEs), polychlorinated biphenyls (PCBs) and chlorinated paraffins (CPs) in rats-testing interactions and mechanisms for thyroid hormone effects. Toxicology, 177, 227-243.

Hansch, C., Leo, A., 1995. Exploring QSAR. 1. Fundamentals and applications in chemistry and biology. Washington, DC: American Chemical Society.

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.

Harner, T., Shoeib, M., 2002. Measurements of octanol-air partition coefficients (K_{oa}) for polybrominated diphenyl ethers (PBDEs): predicting partitioning in the environment. The Journal of Chemical & Engineering Data, 47, 228-232.

Hasegawaa, K., Funatsub, K., 1998. GA strategy for variable selection in QSAR studies: GAPLS and D-optimal designs for predictive QSAR model. Journal of Molecular Structure (Theochem), 425, 255-262.

Hayward, S. J., Lei, Y. D., Wania, F., 2006. Comparative evaluation of three high-performance liquid chromatography–based K_{ow} estimation methods for highly hydrophobic organic compounds: polybrominated diphenyl ethers and hexabromocyclododecane. Environmental Toxicology and Chemistry, 25, 2018-2027.

Hemmer, M. C., Steinhauer, V., Gasteiger, J., 1999. Deriving the 3D structure of organic molecules from their infrared spectra. Vibrational Spectroscopy, 19, 151-164.

Hites, R. A., 2004. Polybrominated diphenyl ethers in the environment and in people: a meta-analysis of concentration (critical review). Environmental Science & Technology, 38, 945-956.

Huang, J., Yu, G., Yang, X., Zhang, Z. L., 2004. Predicting physico-chemical properties of polychlorinated diphenyl ethers (PCDEs): potential persistent organic pollutants (POPs). Journal of Environmental Sciences, 16, 204-207.

Huestis, S. Y., Sergeant, D. B., 1992. Removal of chlorinated diphenyl ether interferences for analyses of PCDDs and PCDFs in fish. Chemosphere, 24, 537-545.

Hundt, K., Jonas, U., Hammer, E., Schauer, F., 1999. Transformation of diphenyl ethers by Trametes versicolor and characterization of ring cleavage products. Biodegradation, 10, 279-286.

IPCS. 1994. Brominated diphenyl ethers. Environmental Health Criteria 162. WHO, Geneva.

Istefanopulos, Y., State Variables and Linear Control Systems, Boğaziçi University Publications, 61-63, 1981.

Jaworska, J., Aldenberg, T., Nikolova, N., 2005a. Review of methods for assessing the applicability domains of SARS and QSARS. Paper 1: Review of methods for QSAR applicability domain estimation by the training set. Ispra, Italy: The European Commission – Joint Research Centre, Institute for Health & Consumer Protection – ECVAM, 28.

Jaworska, J. S., Nikolova-Jeliazkova, N., Aldenberg, T., 2005b. QSAR Applicability domain estimation by projection of the training set in descriptor space: A Review. ATLA, 33, 445-459.

Junk, G. A., Ford, C. S., 1980. A review of organic emissions from selected combustion processes. Chemosphere, 9, 187-230.

Kalantzi, O. I., Martin, F. L., Thomas, G. O., Alcock, R. E., Tang, H. R., Drury, S. C., Carmichael, P. L., Nicholson, J. K., Jones, K. C., 2004. Different levels of polybrominated diphenyl ethers (PBDEs) and chlorinated compounds in breast milk from two U.K. Regions. Environmental Health Perspectives, 112, 1085-1091.

Katritzky, R., Lobanov, V. S., Karelson, M., 1994. CODESSA: Comprehensive Descriptors for Structural and Statistical Analysis, version 2.2.1. Reference Manual. University of Florida, Gainesville, Florida, U.S.A.

Katritzky, A. R., Petrukhin, R., Jain, R., Karelson, M., 2001. QSPR analysis of flash points. Journal of Chemical Information and Computer Sciences, 41, 1521-1530.

Kelly, B. C., Ikonomou, M. G., Blair, J. D., Morin, A. E., Gobas, F. A. P. C., 2007. Food web-specific 106hysic106r106c106ation106 of persistent organic pollutants. Science, 317, 236-239.

Keum, Y. S., Li, Q. X., 2005. Reductive debromination of polybrominated diphenyl ethers by zerovalent iron. Environmental Science & Technology, 39, 2280-2286.

Kiralj, R., Ferreira, M. M. C., 2009. Basic validation procedures for regression models in QSAR and QSPR studies: Theory and application. Journal of the Brazilian Chemical Society, 20, 770-787.

Klasson Wehler, E., Jakobsson, E., Örn, U., 1996. Metabolism of polychlorinated naphtalenes and tetrabrominated diphenyl ether. Organohalogen Compounds, 28, 495-499.

Kodavanti, P. R., Derr-Yellin, E. C., 2002. Differential effects of polybrominated diphenyl ethers and polychlorinated biphenyls on [3H]arachidonic acid release in rat cerebellar granule neurons. Toxicological Sciences, 68, 451-457.

Koistinen, J., Koivusaari, J., Nuuja, I., Paasivirta, J., 1993a. Levels of polychlorinated diphenyl ethers, PCBs, PCDDs and PCDFs in the 106hysic white-tailed sea eagle. Organohalogen Compounds, 1, 329-332.

Koistinen, J., Vuorinen, P. J., Paasivirta, J., 1993b. Contents and origin of polychlorinated diphenyl ethers (PCDE) in salmon from the Baltic Sea, Lake Saimaa and the Tenojoki river in Finland. Chemosphere, 27, 2365-2380.

Koistinen, J., Paasivirta, J., Suonpera, M., 1995a. Contamination of pike and sediments from Kymijoki River by PCDEs, PCDDs, and PCDFs: contents and patterns compared to pike and sediment from the Bothnian Bay and seals from Lake Saimaa. Environmental Science & Technology, 29, 2541-2547.

Koistinen, J., Koivusaari, J., Nuuja, I., Paasivirta, J., 1995b. PCDEs, PCBs, PCDDs and PCDFs in black guillemots and white-tailed sea eagles from the Baltic Sea. Chemosphere, 30, 1671-1684.

Koistinen, J., Stenman, O., Haahti, H., Suonperä, M., Paasivirta, J., 1997. Polychlorinated diphenyl ethers, dibenzo-p-dioxins, dibenzofurans and biphenyls in seals and sediment from the Gulf of Finland. Chemosphere, 35, 1249-1269.

Koistinen, J., 2000. Polychlorinated Diphenyl Ethers (PCDE). In Otto Hutzinger, Jaakko Paasivirta (Eds.): The Handbook of Environmental Chemistry. Berlin, Heidelberg: Springer Berlin Heidelberg, 157-201.

Komsta, E., Chu, I., Villeneuve, D. C., Benoit, F. M., Murdoch, D., 1988. Tissue distribution metabolism and excretion of 2,2',4,4',5-pentachlorodiphenyl ether in the rat. Archives of Toxicology, 62, 258-262.

Krupp, P. P., Lee, K. P., 1988. The effects of dietary iodine on thyroid ultrastructure. Tissue & Cell, 20, 79-88.

Kurz, J., 1994. Chemie und Analytik der polychlorierten Diphenylether (PCDE). Ph. D. Thesis. University of Ulm, Germany.

Kurz, J., Ballschmiter, K., 1994. Relationship between structure and retention of polychlorinated diphenyl ethers (PCDE) in HRGC in comparison with other groups of halogenated aromatic compounds. Fresenius' Journal of Analytical Chemistry, 349, 533-537.

Kurz, J., Ballschmiter, K., 1995a. Isomer specific determination of 79 polychlorinated diphenyl ethers (PCDE) in cod liver oils, chlorophenols and a fly ash. Fresenius' Journal of Analytical Chemistry, 351, 98-109.

Kurz, J., Ballschmiter, K., 1995b. Analysis of polychlorinated diphenyl ethers (PCDE) in environmental samples. Organohalogen Compounds, 24, 485-488.

Kurz, J., Ballschmiter, K., 1999. Vapour pressures, aqueous solubilities, Henry's law constants, partition coefficients between gas/water (K_{gw}), *n*-octanol/water (K_{ow}) and gas/*n*-octanol (K_{go}) of 106 polychlorinated diphenyl ethers (PCDE). Chemosphere, 38, 573-586.

Lado-Abeal, J., Molinaro, E., DeValk, E., Pacini, F., Refetoff, S., 2003. The effect of short-term treatment with recombinant human thyroid-stimulating hormones on leydig cell function in men. Thyroid, 13, 649-652.

Leonard, J. T., Roy, K., 2006. On selection of training and test sets for the development of predictive QSAR models. QSAR & Combinatorial Science, 25, 235-251.

Li, L., Xie, S., Cai, H., Bai, X., Xue, Z., 2008. Quantitative structure-property relationships for octanol-water partition coefficients of polybrominated diphenyl ethers. Chemosphere, 72, 1602-1606.

Lu, W., Chen, Y., Liu, M., Chen, X., Hu, Z., 2007. QSPR prediction of *n*-octanol/water partition coefficient for polychlorinated biphenyls. Chemosphere, 69, 469-478.

Luan, F., Ma, W. P., Zhang, X. Y., Zhang, H. X., Liu, M. C., Hu, Z. D., Fan, B. T., 2006. QSAR study of polychlorinated dibenzodioxins, dibenzofurans, and biphenyls using the heuristic method and support vector machine. QSAR & Combinatorial Science, 25, 46-55.

Malmberg, T., 2004a. Identification and characterisation of hydroxylated PCB and PBDE metabolites in blood. Congener specific synthesis and analysis, Ph.D. Thesis, Department of Environmental Chemistry, Stockholm University.

Malmberg, T., Athanasiadou, M., Marsh, G., Brandt, I., Bergman, Å., 2004b. Hydroxylated PBDE metabolites in rat blood after exposure to a mixture of PBDE congeners, Organohalogen Compounds, 66, 3836-3841.

McDonald, T. A., 2002. A perspective on the potential health risks of PBDEs. Chemosphere, 46, 745-755.

McKinney, J. D., Richard, A., Waller, C., Newman, M. C., Gerberick, F., 2000. The practice of structure activity relationships (SAR) in toxicology. Toxicological Sciences, 56, 8-17.

Meerts, I. A., van Zanden, J. J., Luijks, E. A., Leeuwen-Bol, I., Marsh, G., Jakobsson, E., Bergman, A., Brouwer, A., 2000. Potent competitive interactions of some brominated flame retardants and related compounds with human transthyretin in vitro. Toxicological Sciences, 56, 95-104.

Meironyte, D., Noren, K., Bergman, A., 1999. Analysis of polybrominated diphenyl ethers in Swedish human milk. A time-related trend study, 1972-1997. Journal of Toxicology and Environmental Health A., 58, 329-341.

Molegro Data Modeller (MDM 2010.2) 2007-2010, Molegro ApS.

Mörck, A., Hakk, H., Örn, U., Klasson Wehler, E., 2003. Decabromodiphenyl ether in the rat: absorption, distribution, metabolism, and excretion. Drug Metabolism and Disposition, 31, 900-907.

Netzeva, T. I., Worth, A. P., Aldenberg, T., Benigni, R., Cronin, M. T. D., Gramatica, P., Jaworska, J. S., Kahn, S., Klopman, G., Marchant, C. A., Myatt, G., Nikolova-Jeliazkova, N., Patlewicz, 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., 2005. Current status of methods for defining the applicability domain of (Quantitative) Structure–Activity Relationships. The report and recommendations of ECVAM Workshop 52, Ispra (VA), Italy, 29 September-1 October 2004. Alternatives to Laboratory Animals, 33, 155-173.

Nevalainen, T., Kolehmainen, E., 1994. New QSAR models for polyhalogenated aromatics. Environmental Toxicology & Chemistry, 13, 1699-1706.

Niimi, A. J., Metcalfe C. D., Huestis S. Y., 1994. Chlorinated diphenyl ethers in Great Lakes fish and their environmental implication. Environmental Toxicology & Chemistry, 13, 1133-1138.

Nikolova-Jeliazkova, N., Jaworska, J. S., 2005. An approach to determining applicability domains for QSAR Group contribution models: An analysis of SRC KOWWIN. ATLA, 33, 461-470.

Nilsson, C. A., Renberg, L., 1974. Further studies on impurities in chlorophenols. Journal of Chromatography, 89, 325-333.

Norris, J. M., Kociba, R. J., Schwetz, B. A., Rose, J. Q., Humiston, C. G., Jewett, G. L., Gehring, P. J., Mailhes, J. B., 1975. Toxicology of octabromobiphenyl and decabromodiphenyl oxide. Environmental Health Perspectives, 11, 153-161.

NTP, 1986. Toxicology and Carcinogenesis Studies of Decabromodiphenyl Oxide (CAS No 1163-19-5) in R344/N rats and B6C3F1 mice (feed studies). NTP Technical Report Series no 309. Research Triangle Park, NC: National Toxicology Program.

OECD, Guidance Document on the Validation of (Quantitative) Structure-Activity Relationships [(Q)SAR] Models. ENV/JM/MONO (2007) 2.

Opperhaizen, A., Voors, P. L., 1987. Uptake and elimination of polychlorinated aromatic ethers by fish: Chloroanisoles. Chemosphere, 16, 953-962.

Örn, U., Klasson Wehler, E., 1998. Metabolism of 2,2',4,4'-tetrabromodiphenyl ether in rat and mouse. Xenobiotica, 28, 199-211.

Paasivirta, J., Lahtiperä, M., Leskijärvi, T., 1982. In: Hutzinger O, Frei RW, Merian E, Pocchiari F (eds) Chlorinated dioxins and related compounds: Impact on the environment. Pergamon, Oxford and New York, 191.

Paasivirta, J., Tarhanen, J., Soikkeli J., 1986. Occurrence and fate of polychlorinated aromatic ethers (PCDE, PCA, PCV, PCPA and PCBA) in environment. Chemosphere, 15, 1429-1433.

Paasivirta, J., Koistinen, J., 1994. Chlorinated ethers. In: Kiceniuk, J. W., Ray, S. (Eds.), Analysis of contaminants in edible aquatic resources, VCH Publishers, New York, 411-427.

Papa, E., Kovarich, S., Gramatica, P., 2009. Development, validation and inspection of the applicability domain of QSPR models for physico-chemical properties of polybrominated diphenyl ethers. QSAR & Combinatorial Science, 28, 790-796.

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.

Pirard, C., Pauw, E. D., Focant, J. F., 2003. New strategy for comprehensive analysis of polybrominated diphenyl ethers, polychlorinated dibenzo-p-dioxins, polychlorinated dibenzofurans and polychlorinated biphenyls by gas chromatography coupled with mass spectrometry. Journal of Chromatography A, 998, 169-181.

Puzyn, T., Mostrag, A., Falandysz, J., Kholod, Y., Leszczynski, J., 2009. Predicting water solubility of congeners: chloronaphthalenes-A case study. Journal of Hazardous Materials, 170, 1014-1022.

Raevsky, O. A., Trepalin, S. V., Razdol'skii, A. N., 2000. New QSAR Descriptors Calculated from Interatomic Interaction Spectra. Pharmaceutical Chemistry Journal, 34, 646-649.

Rahman, F., Langford, K. H., Scrimshaw, M. D., Lester, J. N., 2001. Polybrominated diphenyl ether (PBDE) flame retardants. Science of the Total Environment, 275, 1-17.

Rayne, S., Ikonomou, M. G., Whale, M. D., 2003. Anaerobic microbial and photochemical degradation of 4,4′-dibromodiphenyl ether. Water Research, 37, 551-560.

Reddy, K. N., Locke, M. A., 1996. Molecular properties as descriptors of K_{ow} of herbicides. Water, Air, and Soil Pollution, 86, 389-405.

Reistad, T., Mariussen, E., Fonnum, F., 2002. The effects of polybrominated flame retardants on cell death and free radical formation in cerebellar granule cells. Organohalogen Compounds, 57, 391-394.

Richardson, S. D., 2004. Environmental mass spectrometry: emerging contaminants and current issues. Analytical Chemistry, 76, 3337-3364.

Rosiak, K., Li, M. H., Degitz, S. J., Skalla, D. W., Chu, I., Francis, B. M., 1997. Maternal and developmental toxicity of polychlorinated diphenyl ethers (PCDEs) in Swiss-Webster mice and Sprague-Dawley rats. Toxicology, 121, 191-204.

Roy, K., Gosh, G., 2006. QSTR with extended topochemical atom (ETA) indices. 8.a QSAR for the inhibition of substituted phenols on germination rate of *Cucumis sativus* using chemometric tools. QSAR & Combinatorial Sciences, 25, 846-859.

Roy, K., 2007. On some aspects of validation of predictive quantitative structure-activity relationship models. Expert Opinion on Drug Discovery, 2, 1567-1577.

Roy, K., Ghosh, G., 2007. QSTR with extended topochemical atom (ETA) indices. 9. Comparative QSAR for the toxicity of diverse functional organic compounds to *Chlorella vulgaris* using chemometric tools. Chemosphere, 70, 1-12.

Roy, P. P., Leonard, J. T., Roy, K., 2008. Exploring the impact of size of training sets for the development of predictive QSAR models. Chemometrics and Intelligent Laboratory Systems, 90, 31-42.

Roy, K., Mandal, A. S., 2008. Development of linear and nonlinear predictive QSAR models and their external validation using molecular similarity principle for anti-HIV indolyl aryl sulfones. Journal of Enzyme Inhibition and Medicinal Chemistry, 23, 980-995.

Roy, K., Roy, P. P., 2009. Comparative chemometric modeling of cytochrome 3A4 inhibitory activity of structurally diverse compounds using stepwise MLR, FA-MLR, PLS, GFA, G/PLS and ANN techniques. European Journal of Medicinal Chemistry, 44, 2913-2922.

Sabljic, A., Gusten, H., Verhaar, H., Hermens, J., 1995. QSAR modelling of soil sorption. Improvements and systematics of log K_{oc} vs log K_{ow} correlations. Chemosphere, 31, 4489-4514.

Saçan, M. T., Inel, Y., 1993. Prediction of aqueous solubility of PCBs related to molecular structure. Turkish Journal of Chemistry, 17, 188-195.

Saçan, M. T., Inel, Y., 1995. Application of the Characteristic Root Index model to the estimation of *n*-Octanol water partition coefficients. Polychlorinated biphenyls. Chemosphere, 30, 39-50.

Saçan, M. T., Balcioglu, I. A., 1996. Prediction of soil sorption coefficient of organic pollutants by the Characteristic Root Index model. Chemosphere, 32, 1993-2001.

Saçan, M. T., Balcioglu, I. A., 1998. Estimation of liquid vapor pressures for low-volatility environmental chemicals. Chemosphere, 36, 451-460.

Saçan, M. T., Erdem, S. S., Özpınar, G. A., Balcıoğlu, I. A., 2003. QSPR study on the bioconcentration factors of nonionic organic compounds in fish by characteristic root index and semiempirical molecular descriptors. Journal of Chemical Information and Computer Sciences, 44, 985-992.

Saçan, M. T., Ozkul, M., Erdem., S. S., 2005. Physico-chemical properties of PCDD/PCDFs and phthalate esters. SAR and QSAR in Environmental Research, 16, 443-459.

Saçan, M. T., Özkul, M., Erdem, S. S., 2007. QSPR analysis of the toxicity of aromatic compounds to the algae (*Scenedesmus obliquus*). Chemosphere, 68, 695-702.

Schecter, A., Pavuk, M., Papke, O., 2003. Polybrominated diphenyl ethers (PBDEs) in U.S. mothers' milk. Environmental Health Perspectives, 111, 1723-1729.

Schmidt, S., Wittich, R. M., Erdmann, D., Wilkes, H., Francke, W., Fortnagel, P., 1992. Biodegradation of diphenyl ether and its monohalogenated derivatives by *Sphingomonas* sp. Strain SS3. Applied and Environmental Microbiology, 58, 2744-2750.

Schmidt, S., Fortnagel, P., Wittich, R. M., 1993. Biodegradation and transformation of 4,4′- and 2,4-dihalodiphenyl ethers by *Sphingomonas* sp. Strain SS33. Applied and Environmental Microbiology, 59, 3931-3933.

Schulte, E., Malisch, R., 1983. Berechnung der wahren PCB-gehalte in umweltproben I. ermittlung der zusammensetzung zweier technischer PCB-gemische, Fresenius Zeitschrift fur Analytische Chemie, 314, 545-551.

Schuur, J. H., Selzer, P., Gasteiger, J., 1996. The coding of the three-dimensional structure of molecules by molecular transforms and its application to structure spectra correlations and studies of biological activity. Journal of Chemical Information and Computer Sciences, 36, 334-344.

Scientific Workplace 3.0, MacKichan Software, Inc., Windows 2000/XP.

She, J., Petreas, M., Winkler, J., Visita, P., McKinney, M., Kopec, D., 2002. PBDEs in the San Francisco Bay Area: measurements in harbor seal blubber and human breast adipose tissue. Chemosphere, 46, 697-707.

Sjödin, A., Hagmar, L., Klasson Wehler, E., Kronholm-Diab, K., Jakobsson, E., Bergman, Å., 1999. Flame retardant exposure: Polybrominated diphenyl ethers in blood from Swedish workers. Environmental Health Perspectives, 107, 643-648.

Sjödin, A., Carlsson, H., Thuresson, K., Sjölin, S., Bergman, Å., Östman, C., 2001. Flame retardants in indoor air at an electronics recycling plant and at other work environments. Environmental Science & Technology, 35, 448-454.

Sjödin, A., Jones, R. S., Focant, J. F., Lapeza, C., Wang, R. Y., McGahee III, E. E., Zhang, Y., Turner, W. E., Slazyk, B., Needham, L. L., Patterson Jr, D. G., 2004. Retrospective time-trend study of polybrominated diphenyl ether and polybrominated and polychlorinated biphenyl levels in human serum from the United States. Environmental Health Perspectives, 112, 654-658.

Söderström, G., Sellström, U., Wit, C. A., Tysklind, M., 2004. Photolytic debromination of decabromodiphenyl ether (BDE 209). Environmental Science & Technology, 38, 127-132.

SPARTAN 06 Copyright © 2006 by Wavefunction, Inc., Irvine, USA.

SPSS 17.0 for Windows (SPSS® 2008) Statistical Package for Social Scientists by SPSS, Inc.

Stafford, C. J., 1983. Halogenated diphenyl ethers identified in avian tissues and egg by GUMS, Chemosphere, 12, 1487-1495.

Stanley, J. S., Cramer, P. H., Ayling, R. E., Thornburg, K. R., Remmers, J. C., Breen, J. J., 1990. Determination of the prevalence of polychlorinated diphenyl ethers (PCDEs) in human adipose tissue samples. Chemosphere, 20, 981-985.

Stanley, J. S., Cramer, P. H., Thornburg, K. R., Remmers, J. C., Breen, J. J., Schweberger, J., 1991. Mass spectral confirmation of chlorinated and brominated diphenyl ethers in human adipose tissues. Chemosphere, 23, 1185-1195.

Stewart, J. J. P., 1989a. Optimization of parameters for semiempirical methods 1. Method. Journal of Computational Chemistry, 10, 209-220.

Stewart, J. J. P., 1989b. Optimization of parameters for semiempirical methods 2. Appl. Journal of Computational Chemistry, 10, 221-264.

Stonley, S., Crumer, P. H., Ayling, R. E., Thornburg, K. R., Remmers, J. C., Breemand, J. J., Schwemmberger, J., 1990. Determination of the prevalence of polychlorinated diphenyl ethers (PCDPEs) in human adipose tissue samples. Chemosphere, 20, 98l-985.

Streets, S. S., Henderson, S. A., Stoner, A. D., Carlson, D. L., Simcik, M. F., Swackhamer, D. L., 2006. Partitioning and bioaccumulation of PBDEs and PCBs in Lake Michigan. Environmental Science & Technology, 40, 7263-7269.

Sun, L. L., Zhou, L. P., Yu, Y., Lan, Y. K., Li, Z. L., 2007. QSPR study of polychlorinated diphenyl ethers by molecular electronegativity distance vector (MEDV-4). Chemosphere, 66, 1039-1051.

Talete srl, DRAGON for Windows (Software for Molecular Descriptor Calculations). Version 5.4 – 2006 – http://www.talete.mi.it/

Tetko, I. V., Bruneau, P., Mewes, H. W., Rohrer, D. C., Poda, G. I., 2006. Can we estimate the accuracy of ADME-Tox Predictions?. Drug Discovery Today, 11, 700-707.

Thanikaivelan, P., Subramanian, V., Rao, J. R., Nair, B. U., 2000. Application of quantum chemical descriptor in quantitative structure activity and structure property relationship. Chemical Physics Letters, 323, 59-70.

Thoma, H., Hauschulz, G., Knorr, E., Hutzinger, O., 1987. Polybrominated dibenzofurans (PBDF) and dibenzodioxins (PBDD) from the pyrolysis of neat brominated diphenylethers, biphenyls and plastic mixtures of these compounds. Chemosphere, 16, 277-285.

Thomsen, C., Lundanes, E., Becher, G., 2001. Brominated flame retardants in plasma samples from three different occupational groups in Norway. Journal of Environmental Monitoring, 3, 366-370.

Tittlemier, S. A., Halldorson, T., Stern, G. A., Tomy, G. T., 2002. Vapor pressures, aqueous solubilities, and Henry's law constants of some brominated flame retardants. Environmental Toxicology & Chemistry, 21, 1804-1810.

Todeschini, R., Gramatica, P., 1997. 3D-modelling and prediction by WHIM descriptors. Part 5. Theory development and chemical meaning of the WHIM descriptors. Quantitative Structure-Activity Relationships, 16, 113-119.

Todeschini, R., Consonni, V., 2000. Handbook of molecular descriptors. Wiley-VCH: Weinheim, Germany.

Todescini, R., Consonni, V., Mannhold, R., 2009. Molecular Descriptors for Chemoinformatics, 520 Wiley, Weinheim.

Topliss, J. G., Costello, R. J., 1972. Chance correlations in structure-activity studies using multiple regression analysis. Journal of Medicinal Chemistry, 15, 1066-1068.

Toropov, A. A., Toropova, A. P., Raska, J. I., 2008. QSPR modeling of octanol/water partition coefficient for vitamins by optimal descriptors calculated with SMILES. European Journal of Medicinal Chemistry, 43, 714-740.

Trainor, B. C., Bird, I. M., Alday, N. A., Schlinger, B. A., Marler, C. A., 2003. Variation in aromatase activity in the medial preoptic area and plasma progesterone is associated with the onset of paternal behavior. Neuroendocrinology, 78, 36-44.

Tropsha, A., Gramatica, P., Gombar, V. K., 2003. The importance of being Earnest: Validation is the absolute essential for successful application and interpretation of QSPR models. QSAR & Combinatorial Science 22, 69-77.

Tuppurainen, K., Ruuskanen, J., 2000. Electronic eigenvalue (EEVA): a new QSAR/QSPR descriptor for electronic substituent effects based on molecular orbital energies. A QSAR approach to the Ah receptor binding affinity of polychlorinated biphenyls (PCBs), dibenzo-p-dioxins (PCDDs) and dibenzofurans (PCDFs). Chemosphere, 41, 843-848.

Viberg, H., Fredriksson, A., Eriksson, P., 2002. Neonatal exposure to the brominated flame retardant 2,2′,4,4′,5-pentabromodiphenyl ether causes altered susceptibility in the cholinergic transmitter system in the adult mouse. Toxicological Sciences, 67, 104-107.

Viberg, H., Fredriksson, A., Eriksson, P., 2003a. Neonatal exposure to polybrominated diphenyl ether (PBDE 153) disrupts spontaneous behavior, impairs learning and memory, and decreases hippocampal cholinergic receptors in adult mice. Toxicology and Applied Pharmacology, 192, 95-106.

Viberg, H., Fredriksson, A., Jakobsson, E., Örn, U., Eriksson, P., 2003b. Neurobehavioral derangements in adult mice receiving decabrominated diphenyl ether (PBDE 209) during a defined period of neonatal brain development. Toxicological Sciences, 76, 112-120.

Viberg, H., Fredriksson, A., Eriksson, P., 2004a. Neonatal exposure to the brominated flame-retardant, 2,2',4,4',5-pentabromodiphenyl ether, decreases cholinergic nicotinic receptors in hippocampus and affects spontaneous behaviour in the adult mouse. Environmental Toxicology and Pharmacology, 17, 61-65.

Viberg, H., Fredriksson, A., Eriksson, P., 2004b. Investigations of strain and/or gender differences in developmental neurotoxic effects of polybrominated diphenyl ethers in mice. Toxicological Sciences, 81, 344-353.

Vos, J. G., Becher, G., van den Berg, M., de Boer, J., Leonards, P. E. G., 2003. Brominated flame retardants and endocrine disruption. Pure and Applied Chemistry, 75, 2039-2046.

Vracko, M., 2005. Kohonen artificial neural network and counter propagation neural network in molecular structure-toxicity studies. Current Computer-Aided Drug Design, 1, 73-78.

Vracko, M., Bandelj, V., Barbieri, P., Benfenati, E., Chaudhry, Q., Cronin, M., Devillers, J., Gallegos, A., Gini, G., Gramatica, P., Helma, C., Mazzatorta, P., Neagu, D., Netzeva, T., Pavan, M., Patlewicz, G., Randic, M., Tsakovska, I., Worth, A., 2006. Validation of counter propagation neural network models for predictive toxicology according to the OECD principles: a case study. SAR and QSAR in Environmental Research, 17, 265-284.

Xia, B., Liu, K., Gong, Z., Zheng, B., Zhang, X., Fan, B., 2009. Rapid toxicity prediction of organic chemicals to *Chlorella vulgaris* using quantitative structure-activity relationships methods. Ecotoxicology and Environmental Safety, 72, 787-794.

Xiao, H., Wania, F., 2003. Is vapor pressure or the octanol-air partition coefficient a better descriptor of the partitioning between gas phase and organic matter? Atmospheric Environment, 37, 2867-2878.

Xu, H. Y., Zou, J. W., Yu, Q. S., Wang, Y. H., Zhang, J. Y., Jin, H. X., 2007. QSPR/QSAR models for prediction of the physico-chemical properties and biological activity of polybrominated diphenyl ethers. Chemosphere, 66, 1998-2010.

Xu, H. Y., Zou, J. W., Hu, G. X., Wang, W., 2010. QSPR/QSAR models for prediction of the physico-chemical properties and biological activity of polychlorinated diphenyl ethers (PCDEs). Chemosphere, 80, 665-670.

Wang, Y., Liu, H., Zhao, C., Liu, H., Cai, Z., Jiang, G., 2005. Quantitative structure-activity relationship models for prediction of the toxicity of polybrominated diphenyl ether congeners. Environmental Science & Technology, 39, 4961-4966.

Wang, Y., Zhao, C., Ma, W., Liu, H., Wang, T., Jiang, G., 2006. Quantitative structure–activity relationship for prediction of the toxicity of polybrominated diphenyl ether (PBDE) congeners. Chemosphere, 64, 515-524.

Wang, Z. Y., Zeng, X. L., Zhai, Z. C., 2008. Prediction of supercooled liquid vapor pressures and *n*-octanol/air partition coefficients for polybrominated diphenyl ethers by means of molecular descriptors from DFT method. Science of the Total Environment, 389, 296-305.

Wania, F., Lei, Y. D., Harner, T., 2002. Estimating octanol–air partition coefficients of nonpolar semivolatile organic compounds from gas chromatographic retention times. Analytical Chemistry, 74, 3476-3483.

Wania, F., Dugani, C., 2003. Assessing the long-range transport potential of polybrominated diphenyl ethers: a comparison of four multimedia models. Environmental Toxicology & Chemistry, 22, 1252-1261.

Wania, F., Su, Y., 2004. Quantifying the global fractionation of polychlorinated biphenyls. Ambio, 33, 161-168.

Watanabea, I., Sakai, S. I., 2003. Environmental release and behavior of brominated flame retardants. Environment International, 29, 665-682.

Webster, E., Mackay, D., Wania, F., 1998. Evaluating environmental persistence. Environmental Toxicology & Chemistry, 17, 2148-2158.

WEKA 3.6.1: Waikato Environment for Knowledge Analysis © 1999-2009 http://www.cs.waikato.ac.nz/~ml/weka/, March 2010.

White, G. F., Russell, N. J., Tidswell, E. C., 1996. Bacterial scission of ether bonds. Microbiological Reviews, 60, 216-232.

WHO, 1994. World Health Organization. Environmental Health Criteria 162. Brominated diphenyl ethers. International Program on Chemical Safety, World Health Organization, Geneva, Switzerland.

Williams, D. T., Kennedy, B., LeBel, G. L., 1991. Chlorinated diphenyl ethers in human adipose tissue: Part 2. Chemosphere, 23, 601-608.

Wold, S., Eriksson, L., Clementi, S., 1995. Statistical validation of QSAR results. In H. van de Waterbeemd (Ed.), Chemometric methods in molecular design. Book series: Methods and principles in medicinal chemistry, 309-318). Weinheim: VCH (Chapter 5, ISBN 3-527-30044-9).

Wong, A., Lei, Y. D., Alaee, M., Wania, F., 2001. Vapor pressures of the polybrominated diphenyl ethers. The Journal of Chemical & Engineering Data, 46, 239-242.

Yang, P., Chen, J., Chen, S., Yuan, X., Schramm, A., 2003. QSPR models for physico-chemical properties of polychlorinated diphenyl ethers. Science of the Total Environment, 15, 65-76.

Zeng, X. L., Wang, Z. Y., Ge, Z. G., Liu, H. X., 2007. Quantitative structure-property relationships for predicting subcooled liquid vapor pressure (P_L) of 209 polychlorinated diphenyl ethers (PCDEs) by DFT and the position of Cl substitution (PCS) methods. Atmospheric Environment, 41, 3590-3603.

Zhao, H. X., Zhang, Q., Chen, J. P., Xue, X. Y., Liang, X. M., 2005. Prediction of octanol-air partition coefficients of semivolatile organic compounds based on molecular connectivity index. Chemosphere, 59, 1421-1426.

Zheng, G., Xiao, M., Lu, X., 2007. Quantitative Structure-Activity Relationships study on the Ah receptor binding affinities of polybrominated diphenyl ethers using a support vector machine. QSAR & Combinatorial Science, 26, 536-541.

Zhou, T., Ross, D. G., DeVito, M. J., Crofton, K. M., 2001. Effects of short-term *in vivo* exposure to polybrominated diphenyl ethers on thyroid hormones and hepatic enzyme activities in weanling rats. Toxicological Sciences, 61, 76-82.

Zhou, T., Taylor, M. M., DeVito, M. J., Crofton, K. M., 2002. Developmental exposure to brominated diphenyl ethers results in thyroid hormone disruption. Toxicological Sciences, 66, 105-116.

Zou, J. W., Jiang, Y. J., Hu, G. X., Zeng, M., Zhang, S. L., Yu, Q. S., 2005. QSPR/QSAR studies on the physico-chemical properties and biological activities of polychlorinated biphenyls. Acta Physico-Chimica Sinica, 21, 267–272 (in Chinese).

Zupan, J., Gasteiger, J., 1999. Neural Networks in Chemistry and Drug Design. WILEY-VCH, Germany.

Appendix A. The abbreviations and full names of theoretical molecular descriptors appeared in the QSAR/QSPR models selected by the HM.

Name	Description [1]	Type
DRAGON 5.4 [a]		
D/Dr06	distance/detour ring index of order 6	Topological
nBr	number of Bromine atoms	Constitutional
R1e	R autocorrelation of lag 1 / weighted by atomic Sanderson electronegativities	GETAWAY descriptors
H2e	H autocorrelation of lag 2 / weighted by atomic Sanderson electronegativities	GETAWAY descriptors
HIC	mean information content on the leverage magnitude	GETAWAY descriptors
HATS8p	leverage-weighted autocorrelation of lag 8 / weighted by atomic polarizabilities	GETAWAY descriptors
G2p	2st component symmetry directional WHIM index / weighted by atomic polarizabilities	WHIM
RDF075u	Radial Distribution Function – 7.5 / unweighted	Radial distribution function (RDF) descriptors
RDF075p	Radial distribution function – 7.5 / weighted by atomic polarizabilities	Radial distribution function (RDF) descriptors

[1] Todeschini and Consonni (2000)); [a] Talete (2006); [b] Wavefunction (2006); [c] Semichem (1996); [d] Delphi 2007

Appendix A (continued).

Name	Description [1]	Type
RDF090u	Radial Distribution Function – 9.0 / unweighted	Radial distribution function (RDF) descriptors
QYYp	quadrupole y-component value / weighted by polarizability	Geometrical descriptors
SPAM	average span R	Geometrical descriptors
Mor08v	signal 08 / weighted by van der Waals volume	3D-MoRSE
Mor28p	signal 28 / weighted by atomic polarizabilities	3D-MoRSE
Mor02u	3D-MoRSE - signal 02 / unweighted	3D MoRSE
Mor28m	signal 28 / weighted by atomic masses	3D-MoRSE
Mor09m	signal 09 / weighted by atomic masses	3D-MoRSE
Mor26m	signal 26 / weighted by atomic masses	3D-MoRSE
SPARTAN 06 [b]		
E_{HOMO}	Highest occupied molecular orbital energy	Semi-empirical quantum chemical descriptor
E_{LUMO}	Lowest unoccupied molecular orbital energy	Semi-empirical quantum chemical descriptor
E_{aq}	aqueous-phase energy	Semi-empirical quantum chemical descriptor

Appendix A (continued).

Name	Description [1]	Type
μ	dipole moment	Semi-empirical quantum chemical descriptor
η	hardness	Semi-empirical quantum chemical descriptor
CODESSA 2.2 [c]		
S_{XY}	XY Shadow	Geometrical descriptor
S_{YZ}	YZ Shadow	Geometrical descriptor
YZ-S/R	YZ Shadow / YZ Rectangle	Geometrical descriptor
XY-S/R	XY Shadow / XY Rectangle	Geometrical descriptor
V_{XYZ}	Molecular Volume/XYZBox	Geometrical descriptor
M	Relative molecular weight	Constitutional descriptor
WNSA3	Surface weighted charged partial negative charged surface area WNSA-3 Weighted PNSA (PNSA3*TMSA/1000) [Zefirov's PC]	Charged partial surface area (CPSA) descriptor
WPSA1	Surface weighted charged partial positive charged surface area WPSA-1 Weighted PPSA (PPSA1*TMSA/1000) [Zefirov's PC]	Charged partial surface area (CPSA) descriptor

Appendix A (continued).

Name	Description [1]	Type
WPSA3	Surface weighted charged partial positive charged surface area WPSA-3 Weighted PPSA (PPSA3*TMSA/1000) [Zefirov's PC]	Charged partial surface area (CPSA) descriptor
N_{Br}	Number of Br atoms	Constitutional descriptor
P	Polarity parameter (Qmax-Qmin)	Electrostatic descriptor
${}^{0}\chi^{v}$	Kier and Hall valence connectivity indices (order 0)	Topological descriptor
Q_{max} for a O atom	Max partial charge for a O atom [Zefirov's PC]	Electrostatic descriptor
Q_{max} for a C atom	Max partial charge for a C atom [Zefirov's PC]	Electrostatic descriptor
I_k	Moment of Inertia A	Geometrical descriptor
P″	Polarity parameter / square distance	Electrostatic descriptor
***CRI* Program written in Delphi[d]**		
CRI	Characteristic Root Index	Eigenvalue-based descriptor

Appendix B. Values of the descriptors appeared in the proposed models, experimental, and calculated/predicted log K_{oa} and log K_{ow} values of PBDEs/PCDEs.

Table B.1. Congener numbers and names of the PBDEs, descriptors, experimental, and calculated/predicted values from Eq. 4.2.

PBDE Congener Numbers	Compound Names	*CRI*	log K_{oa} Exp.	Calc/Pred	res
0	Diphenyl ether	2.485		6.49	
1	2-monoBDE	3.495	7.24	7.46	-0.22
2	3-monoBDE	3.504	7.36	7.47	-0.11
3	4-monoBDE	3.489		7.46	
4	2,2'-diBDE	4.505		8.44	
5	2,3-diBDE	4.529		8.46	
6	2,3'-diBDE	4.504		8.44	
7	2,4-diBDE	4.468	8.37	8.40	-0.03
8	2,4'-diBDE	4.481	8.47	8.42	0.05
9	2,5-diBDE	4.484		8.42	
10	2,6-diBDE	4.483	8.12	8.42	-0.30
11	3,3'-biBDE	4.520		8.45	
12	3,4-diBDE	4.511	8.55	8.44	0.11
13	3,4'-diBDE	4.495	8.57	8.43	0.14
14	3,5-diBDE	4.505		8.44	
15	4,4'-diBDE	4.487	8.64	8.42	0.22
16	2,2',3-triBDE	5.505		9.40	
17	2,2',4-triBDE	5.484	9.27	9.38	-0.11
18	2,2',5-triBDE	5.517		9.42	
19	2,2',6-triBDE	5.462		9.36	
20	2,3,3'-triBDE	5.519		9.42	
21	2,3,4-triBDE	5.531	9.49	9.43	0.06
22	2,3,4'-triBDE	5.504		9.40	
23	2,3,5-triBDE	5.520		9.42	
24	2,3,6-triBDE	5.509		9.41	
25	2,3',4-triBDE	5.498		9.40	
26	2,3',5-triBDE	5.530		9.43	
27	2,3',6-triBDE	5.477		9.38	
28	2,4,4'-triBDE	5.479	9.46	9.38	0.08
29	2,4,5-triBDE	5.517		9.42	
30	2,4,6-triBDE	5.491	9.02	9.39	-0.37
31	2,4',5-triBDE	5.472		9.37	
32	2,4',6-triBDE	5.463	9.28	9.36	-0.08
33	2,3',4'-triBDE	5.527		9.43	
34	2,3',5'-triBDE	5.534		9.43	
35	3,3',4-triBDE	5.536	9.61	9.44	0.17
36	2,3',5-triBDE	5.547		9.45	
37	3,4,4'-triBDE	5.519	9.68	9.42	0.26

Table B.1. (continued).

PBDE Congener Numbers	Compound Names	*CRI*	log K_{oa} Exp.	Calc/Pred	res
38	3,4,5-triBDE	5.551		9.45	
39	3,4',5-triBDE	5.500		9.40	
40	2,2',3,3'-tetraBDE	6.559		10.42	
41	2,2',3,4-tetraBDE	6.504		10.37	
42	2,2',3,4'-tetraBDE	6.527		10.39	
43	2,2',3,5-tetraBDE	6.490		10.36	
44	2,2',3,5'-tetraBDE	6.516		10.38	
45	2,2',3,6-tetraBDE	6.503		10.37	
46	2,2',3,6'-tetraBDE	6.512		10.38	
47	2,2',4,4'-tetraBDE	6.497	10.34	10.36	-0.02
48	2,2',4,5-tetraBDE	6.522		10.39	
49	2,2',4,5'-tetraBDE	6.464		10.33	
50	2,2',4,6-tetraBDE	6.493		10.36	
51	2,2',4,6'-tetraBDE	6.481		10.35	
52	2,2',5,5'-tetraBDE	6.487		10.35	
53	2,2',5,6'-tetraBDE	6.471		10.34	
54	2,2',6,6'-tetraBDE	6.464		10.33	
55	2,3,3',4-tetraBDE	6.519		10.39	
56	2,3,3',4'-tetraBDE	6.574		10.44	
57	2,3,3',5-tetraBDE	6.504		10.37	
58	2,3,3',5'-tetraBDE	6.548		10.41	
59	2,3,3',6-tetraBDE	6.519		10.38	
60	2,3,4,4'-tetraBDE	6.508		10.37	
61	2,3,4,5-tetraBDE	6.544		10.41	
62	2,3,4,6-tetraBDE	6.514		10.38	
63	2,3,4',5-tetraBDE	6.493		10.36	
64	2,3,4',6-tetraBDE	6.476		10.34	
65	2,3,5,6-tetraBDE	6.521		10.39	
66	2,3',4,4'-tetraBDE	6.544	10.49	10.41	0.08
67	2,3',4,5-tetraBDE	6.501		10.37	
68	2,3',4,5'-tetraBDE	6.502		10.37	
69	2,3',4,6-tetraBDE	6.490	10.23	10.36	-0.13
70	2,3',4',5-tetraBDE	6.507		10.37	
71	2,3',4',6-tetraBDE	6.528		10.39	
72	2,3',5,5'-tetraBDE	6.505		10.37	
73	2,3',5',6-tetraBDE	6.504		10.37	
74	2,4,4',5-tetraBDE	6.491		10.36	
75	2,4,4',6-tetraBDE	6.489	10.13	10.36	-0.23
76	2,3',4',5'-tetraBDE	6.532		10.40	
77	3,3',4,4'-tetraBDE	6.590	10.70	10.45	0.25
78	3,3',4,5-tetraBDE	6.544		10.41	
79	3,3',4,5'-tetraBDE	6.546		10.41	
80	3,3',5,5'-tetraBDE	6.515		10.38	
81	3,4,4',5-tetraBDE	6.526		10.39	

Table B.1. (continued).

PBDE Congener Numbers	Compound Names	*CRI*	log K_{oa} Exp.	Calc/Pred	res
82	2,2',3,3',4-pentaBDE	7.577	11.14	11.41	-0.27
83	2,2',3,3',5-pentaBDE	7.562		11.39	
84	2,2',3,3',6-pentaBDE	7.545		11.38	
85	2,2',3,4,4'-pentaBDE	7.546	11.63	11.38	0.25
86	2,2',3,4,5-pentaBDE	7.547		11.38	
87	2,2',3,4,5'-pentaBDE	7.529		11.36	
88	2,2',3,4,6-pentaBDE	7.520		11.35	
89	2,2',3,4,6'-pentaBDE	7.529		11.36	
90	2,2',3,4',5-pentaBDE	7.530		11.36	
91	2,2',3,4',6-pentaBDE	7.517		11.35	
92	2,2',3,5,5'-pentaBDE	7.514		11.35	
93	2,2',3,5,6-pentaBDE	7.527		11.36	
94	2,2',3,5,6'-pentaBDE	7.514		11.35	
95	2,2',3,5',6-pentaBDE	7.501		11.33	
96	2,2',3,6,6'-pentaBDE	7.495		11.33	
97	2,2',3,4',5'-pentaBDE	7.560		11.39	
98	2,2',3,4',6'-pentaBDE	7.543		11.37	
99	2,2',4,4',5-pentaBDE	7.533	11.28	11.36	-0.08
100	2,2',4,4',6-pentaBDE	7.515	11.19	11.35	-0.16
101	2,2',4,5,5'-pentaBDE	7.511		11.34	
102	2,2',4,5,6'-pentaBDE	7.512		11.34	
103	2,2',4,5',6-pentaBDE	7.515		11.35	
104	2,2',4,6,6'-pentaBDE	7.494		11.33	
105	2,3,3',4,4'-pentaBDE	7.593		11.42	
106	2,3,3',4,5-pentaBDE	7.562		11.39	
107	2,3,3',4',5-pentaBDE	7.578		11.41	
108	2,3,3',4,5'-pentaBDE	7.565		11.40	
109	2,3,3',4,6-pentaBDE	7.535		11.37	
110	2,3,3',4',6-pentaBDE	7.565		11.40	
111	2,3,3',5,5'-pentaBDE	7.578		11.41	
112	2,3,3',5,6-pentaBDE	7.542		11.37	
113	2,3,3',5',6-pentaBDE	7.561		11.39	
114	2,3,4,4',5-pentaBDE	7.541		11.37	
115	2,3,4,4',6-pentaBDE	7.522		11.35	
116	2,3,4,5,6-pentaBDE	7.528		11.36	
117	2,3,4',5,6-pentaBDE	7.512		11.34	
118	2,3',4,4',5-pentaBDE	7.576		11.41	
119	2,3',4,4',6-pentaBDE	7.562	11.52	11.39	0.13
120	2,3',4,5,5'-pentaBDE	7.575		11.40	
121	2,3',4,5',6-pentaBDE	7.538		11.37	
122	2,3,3',4',5'-pentaBDE	7.593		11.42	
123	2,3',4,4',5'-pentaBDE	7.561		11.39	
124	2,3',4',5,5'-pentaBDE	7.539		11.37	
125	2,3',4',5',6-pentaBDE	7.545		11.38	

Table B.1. (continued).

PBDE Congener Numbers	Compound Names	*CRI*	log K_{oa} Exp.	Calc/Pred	res
126	3,3',4,4',5-pentaBDE	7.608	12.00	11.44	0.56
127	3,3',4,5,5'-pentaBDE	7.576		11.41	
128	2,2',3,3',4,4'-hexaBDE	8.596		12.39	
129	2,2',3,3',4,5-hexaBDE	8.593		12.39	
130	2,2',3,3',4,5'-hexaBDE	8.579		12.38	
131	2,2',3,3',4,6-hexaBDE	8.561		12.36	
132	2,2',3,3',4,6'-hexaBDE	8.561		12.36	
133	2,2',3,3',5,5'-hexaBDE	8.562		12.36	
134	2,2',3,3',5,6-hexaBDE	8.566		12.36	
135	2,2',3,3',5,6'-hexaBDE	8.544		12.34	
136	2,2',3,3',6,6'-hexaBDE	8.525		12.32	
137	2,2',3,4,4',5-hexaBDE	8.565		12.36	
138	2,2',3,4,4',5'-hexaBDE	8.577		12.37	
139	2,2',3,4,4',6-hexaBDE	8.534		12.33	
140	2,2',3,4,4',6'-hexaBDE	8.543		12.34	
141	2,2',3,4,5,5'-hexaBDE	8.564		12.36	
142	2,2',3,4,5,6-hexaBDE	8.542		12.34	
143	2,2',3,4,5,6'-hexaBDE	8.545		12.34	
144	2,2',3,4,5',6-hexaBDE	8.532		12.33	
145	2,2',3,4,6,6'-hexaBDE	8.513		12.31	
146	2,2',3,4',5,5'-hexaBDE	8.560		12.36	
147	2,2',3,4',5,6-hexaBDE	8.540		12.34	
148	2,2',3,4',5,6'-hexaBDE	8.528		12.33	
149	2,2',3,4',5',6-hexaBDE	8.542		12.34	
150	2,2',3,4',6,6'-hexaBDE	8.506		12.30	
151	2,2',3,5,5',6-hexaBDE	8.538		12.34	
152	2,2',3,5,6,6'-hexaBDE	8.519		12.32	
153	2,2',4,4',5,5'-hexaBDE	8.559	12.15	12.36	-0.21
154	2,2',4,4',5,6'-hexaBDE	8.526	11.94	12.32	-0.38
155	2,2',4,4',6,6'-hexaBDE	8.489		12.29	
156	2,3,3',4,4',5-hexaBDE	8.610	11.98	12.40	-0.42
157	2,3,3',4,4',5'-hexaBDE	8.613		12.41	
158	2,3,3',4,4',6-hexaBDE	8.580		12.38	
159	2,3,3',4,5,5'-hexaBDE	8.587		12.38	
160	2,3,3',4,5,6-hexaBDE	8.551		12.35	
161	2,3,3',4,5',6-hexaBDE	8.559		12.36	
162	2,3,3',4',5,5'-hexaBDE	8.596		12.39	
163	2,3,3',4',5,6-hexaBDE	8.586		12.38	
164	2,3,3',4',5',6-hexaBDE	8.580		12.38	
165	2,3,3',5,5',6-hexaBDE	8.563		12.36	
166	2,3,4,4',5,6-hexaBDE	8.540		12.34	
167	2,3',4,4',5,5'-hexaBDE	8.595		12.39	
168	2,3',4,4',5',6-hexaBDE	8.562		12.36	
169	3,3',4,4',5,5'-hexaBDE	8.629		12.42	

Table B.1. (continued).

PBDE Congener Numbers	Compound Names	*CRI*	log K_{oa} Exp.	Calc/Pred	res
170	2,2',3,3',4,4',5-heptaBDE	9.594		13.36	
171	2,2',3,3',4,4',6-heptaBDE	9.564		13.33	
172	2,2',3,3',4,5,5'-heptaBDE	9.579		13.34	
173	2,2',3,3',4,5,6-heptaBDE	9.571		13.33	
174	2,2',3,3',4,5,6'-heptaBDE	9.560		13.32	
175	2,2',3,3',4,5',6-heptaBDE	9.548		13.31	
176	2,2',3,3',4,6,6'-heptaBDE	9.528		13.29	
177	2,2',3,3',4,5',6'-heptaBDE	9.569		13.33	
178	2,2',3,3',5,5',6-heptaBDE	9.554		13.32	
179	2,2',3,3',5,6,6'-heptaBDE	9.534		13.30	
180	2,2',3,4,4',5,5'-heptaBDE	9.577		13.34	
181	2,2',3,4,4',5,6-heptaBDE	9.545		13.31	
182	2,2',3,4,4',5,6'-heptaBDE	9.542		13.31	
183	2,2',3,4,4',5',6-heptaBDE	9.546	11.96	13.31	-1.35
184	2,2',3,4,4',6,6'-heptaBDE	9.510		13.27	
185	2,2',3,4,5,5',6-heptaBDE	9.542		13.31	
186	2,2',3,4,5,6,6'-heptaBDE	9.523		13.29	
187	2,2',3,4',5,5',6-heptaBDE	9.552		13.32	
188	2,2',3,4',5,6,6'-heptaBDE	9.516		13.28	
189	2,3,3',4,4',5,5'-heptaBDE	9.612		13.37	
190	2,3,3',4,4',5,6-heptaBDE	9.591		13.35	
191	2,3,3',4,4',5',6-heptaBDE	9.584		13.35	
192	2,3,3',4,5,5',6-heptaBDE	9.570		13.33	
193	2,3,3',4',5,5',6-heptaBDE	9.590		13.35	
194*	2,2',3,3',4,4',5,5'-octaBDE*	10.594		14.32	
195*	2,2',3,3',4,4',5,6-octaBDE*	10.589		14.32	
196*	2,2',3,3',4,4',5,6'-octaBDE*	10.563		14.29	
197*	2,2',3,3',4,4',6,6'-octaBDE*	10.531		14.26	
198*	2,2',3,3',4,5,5',6-octaBDE*	10.559		14.29	
199*	2,2',3,3',4,5,5',6'-octaBDE*	10.569		14.30	
200*	2,2',3,3',4,5,6,6'-octaBDE*	10.539		14.27	
201*	2,2',3,3',4,5',6,6'-octaBDE*	10.536		14.27	
202*	2,2',3,3',5,5',6,6'-octaBDE*	10.542		14.27	
203*	2,2',3,4,4',5,5',6-octaBDE*	10.557		14.29	
204*	2,2',3,4,4',5,6,6'-octaBDE*	10.521		14.25	
205*	2,3,3',4,4',5,5',6-octaBDE*	10.596		14.32	
206*	2,2',3,3',4,4',5,5',6-nonaBDE*	11.574		15.27	
207*	2,2',3,3',4,4',5,6,6'-nonaBDE*	11.551		15.25	
208*	2,2',3,3',4,5,5',6,6'-nonaBDE*	11.574		15.27	
209*	2,2',3,3',4,4',5,5',6,6'-decaBDE*	12.736		16.39	

*Compounds fall outside the applicability domain (AD) of model (Eq. 4.2) (extrapolated values).

Table B.2. Congener numbers and names of the PBDEs/PCDEs, descriptors, experimental, and calculated/predicted log K_{ow} values from Eq. 4.5).

PBDE	Compound Name	*CRI*	E_{HOMO}	E_{aq}	log K_{ow} Exp.	Calc/Pred	res	PCDE	Compound Name	*CRI*	E_{HOMO}	E_{aq}	log K_{ow} Exp.	Calc/Pred	res
0	Diphenyl ether	2.485	-9.480	0.723	3.97	4.19	-0.22	1	2-monoCDE	3.038	-9.358	0.506	4.45	4.63	-0.18
1	2-monoBDE	3.495	-9.249	1.128		4.49		2	3-monoCDE	2.989	-9.415	0.450	4.75	4.68	0.07
2	3-monoBDE	3.504	-9.602	1.051		4.88		3	4-monoCDE	2.961	-9.263	0.449	4.7	4.52	0.18
3	4-monoBDE	3.489	-9.571	1.048		4.84		4	2,2'-diCDE	3.390	-9.057	0.299		4.74	
4	2,2'-diBDE	4.505	-9.500	1.354		5.36		5	2,3-diCDE	3.403	-9.342	0.268	5	5.04	-0.04
5	2,3-diBDE	4.529	-9.382	1.494		5.18		6	2,3'-diCDE	3.415	-9.419	0.222		5.16	
6	2,3'-diBDE	4.504	-9.353	1.460		5.15		7	2,4-diCDE	3.437	-9.321	0.251	4.93	5.06	-0.13
7	2,4-diBDE	4.468	-9.389	1.471		5.15		8	2,4'-diCDE	3.413	-9.352	0.232	5.03	5.08	-0.05
8	2,4'-diBDE	4.481	-9.351	1.459		5.14		9	2,5-diCDE	3.461	-9.235	0.249	5.13	5.00	0.13
9	2,5-diBDE	4.484	-9.437	1.462		5.22		10	2,6-diCDE	3.433	-9.265	0.317	4.64	4.96	-0.32
10	2,6-diBDE	4.483	-9.617	1.631		5.28		11	3,3'-biCDE	3.422	-9.459	0.171		5.23	
11	3,3'-biBDE	4.520	-9.723	1.379		5.57		12	3,4-diCDE	3.424	-9.230	0.218	4.99	4.99	0.00
12	3,4-diBDE	4.511	-9.634	1.403		5.46		13	3,4'-diCDE	3.415	-9.350	0.173	5.13	5.12	0.01
13	3,4'-diBDE	4.495	-9.687	1.377		5.52		14	3,5-diCDE	3.465	-9.519	0.199	5.21	5.30	-0.09
14	3,5-diBDE	4.505	-9.728	1.415		5.54		15	4,4'-diCDE	3.404	-9.308	0.168	5.25	5.08	0.17
15	4,4'-diBDE	4.487	-9.677	1.372	5.51	5.51	0.00	16	2,2',3-triCDE	3.880	-9.150	0.054		5.36	
16	2,2',3-triBDE	5.505	-9.602	1.763		5.97		17	2,2',4-triCDE	3.897	-9.036	0.039	4.96	5.27	-0.31
17	2,2',4-triBDE	5.484	-9.605	1.709	5.74	5.99	-0.25	18	2,2',5-triCDE	3.910	-9.051	0.033		5.30	
18	2,2',5-triBDE	5.517	-9.600	1.702		6.01		19	2,2',6-triCDE	3.893	-9.094	0.088		5.30	
19	2,2',6-triBDE	5.462	-9.567	1.821		5.87		20	2,3,3'-triCDE	3.885	-9.391	-0.022		5.64	
20	2,3,3'-triBDE	5.519	-9.457	1.829		5.80		21	2,3,4-triCDE	3.821	-9.357	0.050	5.55	5.52	0.03
21	2,3,4-triBDE	5.531	-9.431	1.888		5.75		22	2,3,4'-triCDE	3.854	-9.401	-0.008	5.63	5.62	0.01
22	2,3,4'-triBDE	5.504	-9.446	1.828		5.78		23	2,3,5-triCDE	3.861	-9.260	0.036	5.62	5.46	0.16
23	2,3,5-triBDE	5.520	-9.548	1.856		5.87		24	2,3,6-triCDE	3.841	-9.274	0.099	5.35	5.42	-0.07
24	2,3,6-triBDE	5.509	-9.383	1.966		5.64		25	2,3',4-triCDE	3.904	-9.384	-0.037	5.65	5.66	-0.01
25	2,3',4-triBDE	5.498	-9.465	1.804		5.81		26	2,3',5-triCDE	3.893	-9.290	-0.035		5.56	
26	2,3',5-triBDE	5.530	-9.511	1.795		5.88		27	2,3',6-triCDE	3.900	-9.273	0.032		5.51	

Table B.2. (continued).

PBDE	Compound Name	*CRI*	E_{HOMO}	E_{aq}	log K_{ow} Exp.	Calc/Pred	res	PCDE	Compound Name	*CRI*	E_{HOMO}	E_{aq}	log K_{ow} Exp.	Calc/Pred	res
27	2,3',6-triBDE	5.477	-9.449	1.942		5.69		28	2,4,4'-triCDE	3.876	-9.360	-0.028	5.53	5.61	-0.08
28	2,4,4'-triBDE	5.479	-9.457	1.803	5.84	5.79	0.05	29	2,4,5-triCDE	3.857	-9.205	0.036	5.58	5.41	0.17
29	2,4,5-triBDE	5.517	-9.463	1.825		5.81		30	2,4,6-triCDE	3.863	-9.449	0.086	5.32	5.61	-0.29
30	2,4,6-triBDE	5.491	-9.282	2.008		5.50		31	2,4',5-triCDE	3.862	-9.296	-0.028	5.66	5.54	0.12
31	2,4',5-triBDE	5.472	-9.511	1.794		5.84		32	2,4',6-triCDE	3.874	-9.066	0.037	5.3	5.29	0.01
32	2,4',6-triBDE	5.463	-9.361	1.943	5.8	5.59	0.21	33	2,3',4'-triCDE	3.887	-9.302	-0.009	5.5	5.55	-0.05
33	2,3',4'-triBDE	5.527	-9.392	1.818		5.75		34	2,3',5'-triCDE	3.999	-9.353	-0.023		5.69	
34	2,3',5'-triBDE	5.534	-9.436	1.828		5.79		35	3,3',4-triCDE	3.862	-9.283	-0.063	5.74	5.55	0.19
35	3,3',4-triBDE	5.536	-9.732	1.733	5.87	6.14	-0.27	36	2,3',5-triCDE	3.916	-9.549	-0.079		5.85	
36	2,3',5-triBDE	5.547	-9.853	1.745		6.25		37	3,4,4'-triCDE	3.861	-9.283	-0.064	5.88	5.55	0.33
37	3,4,4'-triBDE	5.519	-9.729	1.729		6.12		38	3,4,5-triCDE	3.844	-9.084	0.002	5.7	5.30	0.40
38	3,4,5-triBDE	5.551	-9.502	1.832		5.87		39	3,4',5-triCDE	3.855	-9.419	-0.082	5.77	5.68	0.09
39	3,4',5-triBDE	5.500	-9.795	1.741		6.16		40	2,2',3,3'-tetraCDE	4.314	-9.226	-0.187		5.92	
40	2,2',3,3'-tetraBDE	6.559	-9.461	2.062		6.45		41	2,2',3,4-tetraCDE	4.322	-9.081	-0.170	5.72	5.77	-0.05
41	2,2',3,4-tetraBDE	6.504	-9.651	2.123		6.56		42	2,2',3,4'-tetraCDE	4.332	-9.148	-0.203	5.88	5.86	0.02
42	2,2',3,4'-tetraBDE	6.527	-9.705	2.119		6.63		43	2,2',3,5-tetraCDE	4.360	-9.116	-0.192		5.85	
43	2,2',3,5-tetraBDE	6.490	-9.521	2.090		6.44		44	2,2',3,5'-tetraCDE	4.360	-9.247	-0.220		5.99	
44	2,2',3,5'-tetraBDE	6.516	-9.700	2.113		6.62		45	2,2',3,6-tetraCDE	4.355	-9.148	-0.129		5.84	
45	2,2',3,6-tetraBDE	6.503	-9.683	2.222		6.52		46	2,2',3,6'-tetraCDE	4.330	-9.176	-0.157		5.86	
46	2,2',3,6'-tetraBDE	6.512	-9.750	2.167		6.63		47	2,2',4,4'-tetraCDE	4.366	-9.072	-0.219	5.95	5.83	0.12
47	2,2',4,4'-tetraBDE	6.497	-9.700	2.062	6.81	6.64	0.17	48	2,2',4,5-tetraCDE	4.359	-8.995	-0.187	5.97	5.73	0.24
48	2,2',4,5-tetraBDE	6.522	-9.663	2.074		6.61		49	2,2',4,5'-tetraCDE	4.359	-9.090	-0.223	5.78	5.84	-0.06
49	2,2',4,5'-tetraBDE	6.464	-9.713	2.054		6.63		50	2,2',4,6-tetraCDE	4.380	-9.157	-0.141		5.87	
50	2,2',4,6-tetraBDE	6.493	-9.741	2.265		6.55		51	2,2',4,6'-tetraCDE	4.364	-9.050	-0.172		5.78	
51	2,2',4,6'-tetraBDE	6.481	-9.703	2.235		6.52		52	2,2',5,5'-tetraCDE	4.402	-9.104	-0.221		5.89	
52	2,2',5,5'-tetraBDE	6.487	-9.705	2.048		6.64		53	2,2',5,6'-tetraCDE	4.388	-9.050	-0.177		5.80	
53	2,2',5,6'-tetraBDE	6.471	-9.738	2.221		6.55		54	2,2',6,6'-tetraCDE	4.360	-9.243	-0.019		5.86	
54	2,2',6,6'-tetraBDE	6.464	-9.597	2.319		6.35		55	2,3,3',4-tetraCDE	4.327	-9.413	-0.241	6.07	6.14	-0.07
55	2,3,3',4-tetraBDE	6.519	-9.504	2.223		6.37		56	2,3,3',4'-tetraCDE	4.320	-9.346	-0.253	5.99	6.08	-0.09
56	2,3,3',4'-tetraBDE	6.574	-9.584	2.186		6.51		57	2,3,3',5-tetraCDE	4.366	-9.310	-0.252		6.08	

Table B.2. (continued).

PBDE	Compound Name	*CRI*	E_{HOMO}	E_{aq}	log K_{ow} Exp.	Calc/Pred	res	PCDE	Compound Name	*CRI*	E_{HOMO}	E_{aq}	log K_{ow} Exp.	Calc/Pred	res
57	2,3,3',5-tetraBDE	6.504	-9.626	2.192		6.49		58	2,3,3',5'-tetraCDE	4.360	-9.446	-0.263		6.21	
58	2,3,3',5'-tetraBDE	6.548	-9.780	2.199		6.67		59	2,3,3',6-tetraCDE	4.359	-9.313	-0.187		6.03	
59	2,3,3',6-tetraBDE	6.519	-9.525	2.285		6.35		60	2,3,4,4'-tetraCDE	4.321	-9.390	-0.231	6.14	6.11	0.03
60	2,3,4,4'-tetraBDE	6.508	-9.512	2.221		6.36		61	2,3,4,5-tetraCDE	4.237	-9.023	-0.149	6.01	5.64	0.37
61	2,3,4,5-tetraBDE	6.544	-9.509	2.274		6.36		62	2,3,4,6-tetraCDE	4.305	-9.295	-0.096	5.88	5.92	-0.04
62	2,3,4,6-tetraBDE	6.514	-9.467	2.418		6.20		63	2,3,4',5-tetraCDE	4.344	-9.321	-0.244	6.21	6.06	0.15
63	2,3,4',5-tetraBDE	6.493	-9.623	2.190		6.48		64	2,3,4',6-tetraCDE	4.339	-9.116	-0.183	5.64	5.83	-0.19
64	2,3,4',6-tetraBDE	6.476	-9.475	2.286		6.26		65	2,3,5,6-tetraCDE	4.250	-9.215	-0.099	5.82	5.80	0.02
65	2,3,5,6-tetraBDE	6.521	-9.804	2.322		6.59		66	2,3',4,4'-tetraCDE	4.355	-9.347	-0.271	6.13	6.12	0.01
66	2,3',4,4'-tetraBDE	6.544	-9.505	2.163		6.42		67	2,3',4,5-tetraCDE	4.363	-9.260	-0.251	6.14	6.03	0.11
67	2,3',4,5-tetraBDE	6.501	-9.538	2.160		6.42		68	2,3',4,5'-tetraCDE	4.391	-9.444	-0.282	6.13	6.24	-0.11
68	2,3',4,5'-tetraBDE	6.502	-9.530	2.176		6.41		69	2,3',4,6-tetraCDE	4.386	-9.329	-0.200		6.08	
69	2,3',4,6-tetraBDE	6.490	-9.428	2.325		6.21		70	2,3',4',5-tetraCDE	4.367	-9.326	-0.268	6.11	6.10	0.01
70	2,3',4',5-tetraBDE	6.507	-9.559	2.154		6.45		71	2,3',4',6-tetraCDE	4.352	-9.068	-0.204	5.7	5.80	-0.10
71	2,3',4',6-tetraBDE	6.528	-9.452	2.291		6.28		72	2,3',5,5'-tetraCDE	4.451	-9.342	-0.282		6.19	
72	2,3',5,5'-tetraBDE	6.505	-9.596	2.166		6.48		73	2,3',5',6-tetraCDE	4.387	-9.451	-0.226		6.21	
73	2,3',5',6-tetraBDE	6.504	-9.634	2.297		6.43		74	2,4,4',5-tetraCDE	4.342	-9.264	-0.245	5.99	6.01	-0.02
74	2,4,4',5-tetraBDE	6.491	-9.535	2.159		6.41		75	2,4,4',6-tetraCDE	4.371	-9.124	-0.197	5.92	5.87	0.05
75	2,4,4',6-tetraBDE	6.489	-9.390	2.327		6.17		76	2,3',4',5'-tetraCDE	4.348	-9.110	-0.222		5.85	
76	2,3',4',5'-tetraBDE	6.532	-9.442	2.253		6.30		77	3,3',4,4'-tetraCDE	4.326	-9.306	-0.302	6.36	6.07	0.29
77	3,3',4,4'-tetraBDE	6.590	-9.785	2.087	6.42	6.77	-0.35	78	3,3',4,5-tetraCDE	4.333	-9.338	-0.277		6.09	
78	3,3',4,5-tetraBDE	6.544	-9.559	2.170		6.47		79	3,3',4,5'-tetraCDE	4.367	-9.355	-0.317	6.22	6.16	0.06
79	3,3',4,5'-tetraBDE	6.546	-9.846	2.099		6.79		80	3,3',5,5'-tetraCDE	4.448	-9.607	-0.333		6.47	
80*	3,3',5,5'-tetraBDE	6.515	-9.994	2.111		6.90		81	3,4,4',5-tetraCDE	4.301	-9.337	-0.282	6.3	6.07	0.23
81	3,4,4',5-tetraBDE	6.526	-9.793	2.165		6.68		82	2,2',3,3',4-pentaCDE	4.721	-9.161	-0.409	6.3	6.30	0.00
82	2,2',3,3',4-pentaBDE	7.577	-9.535	2.495		7.03		83	2,2',3,3',5-pentaCDE	4.751	-9.177	-0.429		6.35	
83	2,2',3,3',5-pentaBDE	7.562	-9.566	2.441		7.09		84	2,2',3,3',6-pentaCDE	4.742	-9.223	-0.376		6.36	
84	2,2',3,3',6-pentaBDE	7.545	-9.559	2.566		6.99		85	2,2',3,4,4'-pentaCDE	4.751	-9.116	-0.427	6.28	6.29	-0.01
85	2,2',3,4,4'-pentaBDE	7.546	-9.748	2.480	7.37	7.22	0.15	86	2,2',3,4,5-pentaCDE	4.758	-9.046	-0.381		6.20	
86	2,2',3,4,5-pentaBDE	7.547	-9.750	2.542		7.19		87	2,2',3,4,5'-pentaCDE	4.796	-9.131	-0.423	6.51	6.34	0.17

Table B.2. (continued).

PBDE	Compound Name	*CRI*	E_{HOMO}	E_{aq}	log K_{ow} Exp.	Calc/Pred	res	PCDE	Compound Name	*CRI*	E_{HOMO}	E_{aq}	log K_{ow} Exp.	Calc/Pred	res
87	2,2',3,4,5'-pentaBDE	7.529	-9.742	2.475		7.21		88	2,2',3,4,6-pentaCDE	4.768	-9.191	-0.323		6.31	
88	2,2',3,4,6-pentaBDE	7.520	-9.786	2.688		7.11		89	2,2',3,4,6'-pentaCDE	4.751	-9.104	-0.378	6.11	6.25	-0.14
89	2,2',3,4,6'-pentaBDE	7.529	-9.574	2.651		6.94		90	2,2',3,4',5-pentaCDE	4.791	-9.165	-0.446	6.54	6.38	0.16
90	2,2',3,4',5-pentaBDE	7.530	-9.836	2.438		7.32		91	2,2',3,4',6-pentaCDE	4.773	-9.098	-0.391	6.06	6.27	-0.21
91	2,2',3,4',6-pentaBDE	7.517	-9.752	2.574		7.15		92	2,2',3,5,5'-pentaCDE	4.830	-9.168	-0.444		6.42	
92	2,2',3,5,5'-pentaBDE	7.514	-9.822	2.433		7.30		93	2,2',3,5,6-pentaCDE	4.771	-9.194	-0.324		6.32	
93	2,2',3,5,6-pentaBDE	7.527	-9.803	2.575		7.20		94	2,2',3,5,6'-pentaCDE	4.789	-9.114	-0.399		6.30	
94	2,2',3,5,6'-pentaBDE	7.514	-9.603	2.588		6.99		95	2,2',3,5',6-pentaCDE	4.814	-9.094	-0.396		6.30	
95	2,2',3,5',6-pentaBDE	7.501	-9.785	2.560		7.17		96	2,2',3,6,6'-pentaCDE	4.772	-9.232	-0.243		6.31	
96	2,2',3,6,6'-pentaBDE	7.495	-9.733	2.441		7.19		97	2,2',3,4',5'-pentaCDE	4.749	-9.223	-0.436	6.22	6.40	-0.18
97	2,2',3,4',5'-pentaBDE	7.560	-9.798	2.429		7.31		98	2,2',3,4',6'-pentaCDE	4.746	-9.231	-0.389		6.38	
98	2,2',3,4',6'-pentaBDE	7.543	-9.600	2.606		7.00		99	2,2',4,4',5-pentaCDE	4.789	-9.048	-0.443	6.38	6.27	0.11
99	2,2',4,4',5-pentaBDE	7.533	-9.752	2.430	7.32	7.25	0.07	100	2,2',4,4',6-pentaCDE	4.790	-9.107	-0.404	6.11	6.30	-0.19
100	2,2',4,4',6-pentaBDE	7.515	-9.783	2.599	7.24	7.16	0.08	101	2,2',4,5,5'-pentaCDE	4.820	-9.074	-0.439	6.22	6.31	-0.09
101	2,2',4,5,5'-pentaBDE	7.511	-9.762	2.422		7.25		102	2,2',4,5,6'-pentaCDE	4.788	-8.999	-0.395	5.98	6.19	-0.21
102	2,2',4,5,6'-pentaBDE	7.512	-9.760	2.589		7.14		103	2,2',4,5',6-pentaCDE	4.832	-9.103	-0.408		6.33	
103	2,2',4,5',6-pentaBDE	7.515	-9.793	2.588		7.18		104	2,2',4,6,6'-pentaCDE	4.789	-9.217	-0.254		6.31	
104	2,2',4,6,6'-pentaBDE	7.494	-9.788	2.494		7.21		105	2,3,3',4,4'-pentaCDE	4.729	-9.372	-0.476	6.51	6.55	-0.04
105	2,3,3',4,4'-pentaBDE	7.593	-9.556	2.582		7.01		106	2,3,3',4,5-pentaCDE	4.761	-9.308	-0.436		6.49	
106	2,3,3',4,5-pentaBDE	7.562	-9.593	2.611		7.01		107	2,3,3',4',5-pentaCDE	4.761	-9.352	-0.487	6.52	6.57	-0.05
107	2,3,3',4',5-pentaBDE	7.578	-9.660	2.554		7.12		108	2,3,3',4,5'-pentaCDE	4.780	-9.466	-0.485	6.58	6.69	-0.11
108	2,3,3',4,5'-pentaBDE	7.565	-9.584	2.595		7.01		109	2,3,3',4,6-pentaCDE	4.776	-9.352	-0.384		6.51	
109	2,3,3',4,6-pentaBDE	7.535	-9.569	2.739		6.88		110	2,3,3',4',6-pentaCDE	4.752	-9.112	-0.425	6.31	6.29	0.02
110	2,3,3',4',6-pentaBDE	7.565	-9.541	2.637		6.94		111	2,3,3',5,5'-pentaCDE	4.818	-9.359	-0.498		6.62	
111	2,3,3',5,5'-pentaBDE	7.578	-9.680	2.568		7.13		112	2,3,3',5,6-pentaCDE	4.777	-9.258	-0.385		6.42	
112	2,3,3',5,6-pentaBDE	7.542	-9.924	2.647		7.28		113	2,3,3',5',6-pentaCDE	4.800	-9.368	-0.445		6.59	
113	2,3,3',5',6-pentaBDE	7.561	-9.669	2.646		7.06		114	2,3,4,4',5-pentaCDE	4.739	-9.314	-0.430	6.61	6.48	0.13
114	2,3,4,4',5-pentaBDE	7.541	-9.594	2.610		6.99		115	2,3,4,4',6-pentaCDE	4.772	-9.157	-0.381	6.47	6.32	0.15
115	2,3,4,4',6-pentaBDE	7.522	-9.529	2.740		6.83		116	2,3,4,5,6-pentaCDE	4.712	-9.272	-0.266	6.37	6.31	0.06
116	2,3,4,5,6-pentaBDE	7.528	-9.830	2.826		7.07		117	2,3,4',5,6-pentaCDE	4.771	-9.160	-0.381	6.41	6.32	0.09

Table B.2. (continued).

PBDE	Compound Name	*CRI*	E_{HOMO}	E_{aq}	log K_{ow} Exp.	Calc/Pred	res	PCDE	Compound Name	*CRI*	E_{HOMO}	E_{aq}	log K_{ow} Exp.	Calc/Pred	res
117	2,3,4',5,6-pentaBDE	7.512	-9.838	2.653		7.18		118	2,3',4,4',5-pentaCDE	4.757	-9.298	-0.487	6.6	6.51	0.09
118	2,3',4,4',5-pentaBDE	7.576	-9.572	2.521		7.05		119	2,3',4,4',6-pentaCDE	4.767	-9.120	-0.439	6.44	6.32	0.12
119	2,3',4,4',6-pentaBDE	7.562	-9.464	2.676		6.84		120	2,3',4,5,5'-pentaCDE	4.807	-9.309	-0.499	6.66	6.57	0.09
120	2,3',4,5,5'-pentaBDE	7.575	-9.604	2.534		7.07		121	2,3',4,5',6-pentaCDE	4.814	-9.503	-0.459		6.73	
121	2,3',4,5',6-pentaBDE	7.538	-9.558	2.686		6.91		122	2,3,3',4',5'-pentaCDE	4.734	-9.172	-0.462		6.36	
122	2,3,3',4',5'-pentaBDE	7.593	-9.649	2.623		7.08		123	2,3',4,4',5'-pentaCDE	4.776	-9.156	-0.481	6.63	6.39	0.24
123	2,3',4,4',5'-pentaBDE	7.561	-9.546	2.605		6.96		124	2,3',4',5,5'-pentaCDE	4.808	-9.364	-0.477		6.61	
124	2,3',4',5,5'-pentaBDE	7.539	-9.590	2.596		6.99		125	2,3',4',5',6-pentaCDE	4.774	-9.136	-0.423		6.33	
125	2,3',4',5',6-pentaBDE	7.545	-9.557	2.715		6.89		126	3,3',4,4',5-pentaCDE	4.741	-9.352	-0.517	6.83	6.57	0.26
126	3,3',4,4',5-pentaBDE	7.608	-9.848	2.524		7.34		127	3,3',4,5,5'-pentaCDE	4.804	-9.394	-0.532		6.67	
127	3,3',4,5,5'-pentaBDE	7.576	-9.922	2.536		7.38		128	2,2',3,3',4,4'-hexaCDE	5.188	-9.152	-0.631	6.82	6.79	0.03
128	2,2',3,3',4,4'-hexaBDE	8.596	-9.627	2.918		7.64		129	2,2',3,3',4,5-hexaCDE	5.190	-9.115	-0.616		6.75	
129	2,2',3,3',4,5-hexaBDE	8.593	-9.607	2.869		7.65		130	2,2',3,3',4,5'-hexaCDE	5.213	-9.195	-0.649	7.01	6.86	0.15
130	2,2',3,3',4,5'-hexaBDE	8.579	-9.624	2.874		7.65		131	2,2',3,3',4,6-hexaCDE	5.200	-9.261	-0.571		6.87	
131	2,2',3,3',4,6-hexaBDE	8.561	-9.614	3.032		7.53		132	2,2',3,3',4,6'-hexaCDE	5.205	-9.151	-0.598	6.47	6.78	-0.31
132	2,2',3,3',4,6'-hexaBDE	8.561	-9.641	3.000		7.57		133	2,2',3,3',5,5'-hexaCDE	5.239	-9.225	-0.659		6.92	
133	2,2',3,3',5,5'-hexaBDE	8.562	-9.666	2.880		7.67		134	2,2',3,3',5,6-hexaCDE	5.202	-9.263	-0.571		6.87	
134	2,2',3,3',5,6-hexaBDE	8.566	-9.687	2.927		7.67		135	2,2',3,3',5,6'-hexaCDE	5.230	-9.154	-0.618		6.82	
135	2,2',3,3',5,6'-hexaBDE	8.544	-9.675	2.940		7.63		136	2,2',3,3',6,6'-hexaCDE	5.223	-9.224	-0.467		6.78	
136	2,2',3,3',6,6'-hexaBDE	8.525	-9.555	2.977		7.48		137	2,2',3,4,4',5-hexaCDE	5.157	-9.252	-0.638	6.72	6.87	-0.15
137	2,2',3,4,4',5-hexaBDE	8.565	-9.812	2.901		7.80		138	2,2',3,4,4',5'-hexaCDE	5.210	-9.264	-0.657	7.01	6.93	0.08
138	2,2',3,4,4',5'-hexaBDE	8.577	-9.796	2.850		7.83		139	2,2',3,4,4',6-hexaCDE	5.223	-9.139	-0.587	6.84	6.78	0.06
139	2,2',3,4,4',6-hexaBDE	8.534	-9.853	3.044		7.73		140	2,2',3,4,4',6'-hexaCDE	5.337	-9.159	-0.612	6.65	6.90	-0.25
140	2,2',3,4,4',6'-hexaBDE	8.543	-9.672	3.040		7.56		141	2,2',3,4,5,5'-hexaCDE	5.204	-9.109	-0.631		6.76	
141	2,2',3,4,5,5'-hexaBDE	8.564	-9.835	2.898		7.82		142	2,2',3,4,5,6-hexaCDE	5.149	-9.222	-0.489		6.74	
142	2,2',3,4,5,6-hexaBDE	8.542	-9.773	3.078		7.64		143	2,2',3,4,5,6'-hexaCDE	5.159	-9.059	-0.583		6.65	
143	2,2',3,4,5,6'-hexaBDE	8.545	-9.653	3.033		7.55		144	2,2',3,4,5',6-hexaCDE	5.216	-9.130	-0.591		6.77	
144	2,2',3,4,5',6-hexaBDE	8.532	-9.871	3.028		7.75		145	2,2',3,4,6,6'-hexaCDE	5.172	-9.206	-0.441		6.71	
145	2,2',3,4,6,6'-hexaBDE	8.513	-9.781	2.910		7.73		146	2,2',3,4',5,5'-hexaCDE	5.236	-9.148	-0.658	6.76	6.84	-0.08
146	2,2',3,4',5,5'-hexaBDE	8.560	-9.877	2.808		7.92		147	2,2',3,4',5,6-hexaCDE	5.172	-9.142	-0.587	6.76	6.74	0.02

Table B.2. (continued).

PBDE	Compound Name	*CRI*	E_{HOMO}	E_{aq}	log K_{ow} Exp.	Calc/Pred	res	PCDE	Compound Name	*CRI*	E_{HOMO}	E_{aq}	log K_{ow} Exp.	Calc/Pred	res
147	2,2',3,4',5,6-hexaBDE	8.540	-9.895	2.935		7.84		148	2,2',3,4',5,6'-hexaCDE	5.363	-9.163	-0.631		6.94	
148	2,2',3,4',5,6'-hexaBDE	8.528	-9.704	2.978		7.62		149	2,2',3,4',5',6-hexaCDE	5.227	-9.042	-0.614	6.47	6.71	-0.24
149	2,2',3,4',5',6-hexaBDE	8.542	-9.804	2.930		7.76		150	2,2',3,4',6,6'-hexaCDE	5.357	-9.272	-0.479		6.94	
150	2,2',3,4',6,6'-hexaBDE	8.506	-9.643	3.041		7.51		151	2,2',3,5,5',6-hexaCDE	5.218	-9.132	-0.591		6.77	
151	2,2',3,5,5',6-hexaBDE	8.538	-9.648	2.913		7.62		152	2,2',3,5,6,6'-hexaCDE	5.174	-9.172	-0.444		6.68	
152	2,2',3,5,6,6'-hexaBDE	8.519	-9.667	3.046		7.54		153	2,2',4,4',5,5'-hexaCDE	5.233	-9.075	-0.657	6.72	6.77	-0.05
153	2,2',4,4',5,5'-hexaBDE	8.559	-9.784	2.795	7.9	7.84	0.06	154	2,2',4,4',5,6'-hexaCDE	5.362	-9.050	-0.628	6.49	6.82	-0.33
154	2,2',4,4',5,6'-hexaBDE	8.526	-9.811	2.994	7.82	7.71	0.11	155	2,2',4,4',6,6'-hexaCDE	5.373	-9.223	-0.490		6.91	
155	2,2',4,4',6,6'-hexaBDE	8.489	-9.878	2.878		7.82		156	2,3,3',4,4',5-hexaCDE	5.128	-9.346	-0.674	7.07	6.96	0.11
156	2,3,3',4,4',5-hexaBDE	8.610	-9.623	2.976		7.61		157	2,3,3',4,4',5'-hexaCDE	5.198	-9.418	-0.683	6.99	7.09	-0.10
157	2,3,3',4,4',5'-hexaBDE	8.613	-9.611	3.026		7.57		158	2,3,3',4,4',6-hexaCDE	5.206	-9.149	-0.624		6.80	
158*	2,3,3',4,4',6-hexaBDE	8.580	-9.587	3.093		7.48		159	2,3,3',4,5,5'-hexaCDE	5.182	-9.358	-0.683		7.02	
159	2,3,3',4,5,5'-hexaBDE	8.587	-9.640	2.991		7.60		160	2,3,3',4,5,6-hexaCDE	5.156	-9.313	-0.552		6.87	
160*	2,3,3',4,5,6-hexaBDE	8.551	-9.948	3.153		7.76		161	2,3,3',4,5',6-hexaCDE	5.189	-9.407	-0.643		7.04	
161	2,3,3',4,5',6-hexaBDE	8.559	-9.699	3.104		7.56		162	2,3,3',4',5,5'-hexaCDE	5.171	-9.386	-0.695		7.04	
162	2,3,3',4',5,5'-hexaBDE	8.596	-9.738	3.005		7.69		163	2,3,3',4',5,6-hexaCDE	5.208	-9.151	-0.624	6.78	6.80	-0.02
163	2,3,3',4',5,6-hexaBDE	8.586	-9.899	3.007		7.83		164	2,3,3',4',5',6-hexaCDE	5.215	-9.178	-0.645		6.84	
164	2,3,3',4',5',6-hexaBDE	8.580	-9.617	3.067		7.52		165	2,3,3',5,5',6-hexaCDE	4.366	-9.308	-0.643		6.32	
165*	2,3,3',5,5',6-hexaBDE	8.563	-10.057	3.001		7.97		166	2,3,4,4',5,6-hexaCDE	5.153	-9.186	-0.550	6.95	6.75	0.20
166	2,3,4,4',5,6-hexaBDE	8.540	-9.862	3.158		7.67		167	2,3',4,4',5,5'-hexaCDE	5.221	-9.332	-0.697	7.11	7.03	0.08
167	2,3',4,4',5,5'-hexaBDE	8.595	-9.614	2.966		7.59		168	2,3',4,4',5',6-hexaCDE	5.346	-9.186	-0.659		6.96	
168*	2,3',4,4',5',6-hexaBDE	8.562	-9.541	3.109		7.41		169	3,3',4,4',5,5'-hexaCDE	5.153	-9.399	-0.735		7.07	
169	3,3',4,4',5,5'-hexaBDE	8.629	-9.916	2.962		7.91		170	2,2',3,3',4,4',5-heptaCDE	5.662	-9.137	-0.837	7.28	7.27	0.01
170*	2,2',3,3',4,4',5-heptaBDE	9.594	-9.694	3.295		8.23		171	2,2',3,3',4,4',6-heptaCDE	5.674	-9.189	-0.795		7.30	
171*	2,2',3,3',4,4',6-heptaBDE	9.564	-9.685	3.467		8.09		172	2,2',3,3',4,5,5'-heptaCDE	5.683	-9.177	-0.845		7.33	
172*	2,2',3,3',4,5,5'-heptaBDE	9.579	-9.694	3.250		8.25		173	2,2',3,3',4,5,6-heptaCDE	5.617	-9.288	-0.738		7.32	
173*	2,2',3,3',4,5,6-heptaBDE	9.571	-9.629	3.423		8.07		174	2,2',3,3',4,5,6'-heptaCDE	5.677	-9.100	-0.803	6.98	7.22	-0.24
174*	2,2',3,3',4,5,6'-heptaBDE	9.560	-9.675	3.365		8.14		175	2,2',3,3',4,5',6-heptaCDE	5.714	-9.189	-0.814		7.34	
175*	2,2',3,3',4,5',6-heptaBDE	9.548	-9.716	3.406		8.15		176	2,2',3,3',4,6,6'-heptaCDE	5.709	-9.220	-0.666		7.28	
176*	2,2',3,3',4,6,6'-heptaBDE	9.528	-9.577	3.424		7.99		177	2,2',3,3',4,5',6'-heptaCDE	5.673	-9.193	-0.795	7.14	7.30	-0.16

Table B.2. (continued).

PBDE	Compound Name	*CRI*	E_{HOMO}	E_{aq}	log K_{ow} Exp.	Calc/Pred	res	PCDE	Compound Name	*CRI*	E_{HOMO}	E_{aq}	log K_{ow} Exp.	Calc/Pred	res
177*	2,2',3,3',4,5',6'-heptaBDE	9.569	-9.702	3.365		8.18		178	2,2',3,3',5,5',6-heptaCDE	5.713	-9.190	-0.814		7.34	
178*	2,2',3,3',5,5',6-heptaBDE	9.554	-9.689	3.289		8.20		179	2,2',3,3',5,6,6'-heptaCDE	5.709	-9.213	-0.669		7.27	
179*	2,2',3,3',5,6,6'-heptaBDE	9.534	-9.671	3.377		8.11		180	2,2',3,4,4',5,5'-heptaCDE	5.680	-9.124	-0.847	7.46	7.28	0.18
180*	2,2',3,4,4',5,5'-heptaBDE	9.577	-9.858	3.273		8.39		181	2,2',3,4,4',5,6-heptaCDE	5.634	-9.167	-0.755	7.31	7.22	0.09
181*	2,2',3,4,4',5,6-heptaBDE	9.545	-9.861	3.439		8.26		182	2,2',3,4,4',5,6'-heptaCDE	5.768	-9.108	-0.818		7.31	
182*	2,2',3,4,4',5,6'-heptaBDE	9.542	-9.741	3.424		8.16		183	2,2',3,4,4',5',6-heptaCDE	5.712	-9.077	-0.811		7.23	
183*	2,2',3,4,4',5',6-heptaBDE	9.546	-9.895	3.402	8.27	8.32	-0.05	184	2,2',3,4,4',6,6'-heptaCDE	5.779	-9.231	-0.679		7.35	
184*	2,2',3,4,4',6,6'-heptaBDE	9.510	-9.879	3.296		8.34		185	2,2',3,4,5,5',6-heptaCDE	5.591	-9.155	-0.758		7.18	
185*	2,2',3,4,5,5',6-heptaBDE	9.542	-9.851	3.425		8.26		186	2,2',3,4,5,6,6'-heptaCDE	5.631	-9.204	-0.613		7.17	
186*	2,2',3,4,5,6,6'-heptaBDE	9.523	-9.427	3.556		7.76		187	2,2',3,4',5,5',6-heptaCDE	5.712	-9.079	-0.811	7.13	7.23	-0.10
187*	2,2',3,4',5,5',6-heptaBDE	9.552	-9.912	3.300		8.40		188	2,2',3,4',5,6,6'-heptaCDE	5.777	-9.219	-0.681		7.34	
188*	2,2',3,4',5,6,6'-heptaBDE	9.516	-9.767	3.437		8.15		189	2,3,3',4,4',5,5'-heptaCDE	5.668	-9.379	-0.883	7.55	7.53	0.02
189*	2,3,3',4,4',5,5'-heptaBDE	9.612	-9.658	3.422		8.13		190	2,3,3',4,4',5,6-heptaCDE	5.622	-9.174	-0.794	7.31	7.25	0.06
190*	2,3,3',4,4',5,6-heptaBDE	9.591	-9.569	3.510		7.97		191	2,3,3',4,4',5',6-heptaCDE	5.700	-9.212	-0.844		7.37	
191*	2,3,3',4,4',5',6-heptaBDE	9.584	-9.666	3.527		8.05		192	2,3,3',4,5,5',6-heptaCDE	5.646	-9.362	-0.812		7.46	
192*	2,3,3',4,5,5',6-heptaBDE	9.570	-10.085	3.522		8.44		193	2,3,3',4',5,5',6-heptaCDE	5.681	-9.214	-0.843		7.36	
193*	2,3,3',4',5,5',6-heptaBDE	9.590	-9.990	3.445		8.42		194	2,2',3,3',4,4',5,5'-octaCDE	6.101	-9.163	-1.032	7.78	7.75	0.03
194*	2,2',3,3',4,4',5,5'-octaBDE	10.594	-9.736	3.743		8.76		195	2,2',3,3',4,4',5,6-octaCDE	5.379	-9.216	-0.962	7.84	7.21	0.63
195*	2,2',3,3',4,4',5,6-octaBDE	10.589	-9.694	3.861		8.64		196	2,2',3,3',4,4',5,6'-octaCDE	6.128	-9.134	-1.001		7.72	
196*	2,2',3,3',4,4',5,6'-octaBDE	10.563	-9.730	3.826		8.67		197	2,2',3,3',4,4',6,6'-octaCDE	6.153	-9.208	-0.866		7.73	
197*	2,2',3,3',4,4',6,6'-octaBDE	10.531	-9.777	3.695		8.78		198	2,2',3,3',4,5,5',6-octaCDE	6.110	-9.211	-0.981		7.77	
198*	2,2',3,3',4,5,5',6-octaBDE	10.559	-9.724	3.801		8.68		199	2,2',3,3',4,5,5',6'-octaCDE	6.127	-9.136	-1.000	7.63	7.72	-0.09
199*	2,2',3,3',4,5,5',6'-octaBDE	10.569	-9.762	3.732		8.77		200	2,2',3,3',4,5,6,6'-octaCDE	6.104	-9.250	-0.838		7.72	
200*	2,2',3,3',4,5,6,6'-octaBDE	10.539	-9.657	3.875		8.56		201	2,2',3,3',4,5',6,6'-octaCDE	6.145	-9.241	-0.868		7.76	
201*	2,2',3,3',4,5',6,6'-octaBDE	10.536	-9.769	3.619		8.82		202	2,2',3,3',5,5',6,6'-octaCDE	6.142	-9.244	-0.870		7.76	
202*	2,2',3,3',5,5',6,6'-octaBDE	10.542	-9.784	3.748		8.76		203	2,2',3,4,4',5,5',6-octaCDE	6.109	-9.100	-0.978	7.81	7.66	0.15
203*	2,2',3,4,4',5,5',6-octaBDE	10.557	-9.938	3.806		8.88		204	2,2',3,4,4',5,6,6'-octaCDE	6.139	-9.234	-0.851		7.73	
204*	2,2',3,4,4',5,6,6'-octaBDE	10.521	-9.770	3.711		8.75		205	2,3,3',4,4',5,5',6-octaCDE	6.093	-9.236	-1.013		7.80	
205*	2,3,3',4,4',5,5',6-octaBDE	10.596	-9.641	3.951		8.54		206	2,2',3,3',4,4',5,5',6-nonaCDE	6.515	-9.157	-1.168	8.07	8.15	-0.08

Table B.2. (continued).

PBDE	Compound Name	*CRI*	E_{HOMO}	E_{aq}	log K_{ow} Exp.	Calc/Pred	res	PCDE	Compound Name	*CRI*	E_{HOMO}	E_{aq}	log K_{ow} Exp.	Calc/Pred	res
206*	2,2',3,3',4,4',5,5',6-nonaBDE	11.574	-9.760	4.235		9.22		207	2,2',3,3',4,4',5,6,6'-nonaCDE	6.534	-9.252	-1.038		8.17	
207*	2,2',3,3',4,4',5,6,6'-nonaBDE	11.551	-9.753	4.125		9.27		208	2,2',3,3',4,5,5',6,6'-nonaCDE	6.533	-9.261	-1.039		8.18	
208*	2,2',3,3',4,5,5',6,6'-nonaBDE	11.574	-9.753	4.042		9.34		209	2,2',3,3',4,4',5,5',6,6'-decaCDE	6.915	-9.263	-1.209	8.16	8.58	-0.42
209*	2,2',3,3',4,4',5,5',6,6'-decaBDE	12.736	-9.787	4.550	10	9.94	0.06								

*Compounds fall outside the applicability domain (AD) of model (Eq. 4.5) (extrapolated values).