Using electrotopological state indices to model the depuration rates of polychlorinated biphenyls in mussels of *Elliptio complanata*

Lei Wang[1], Xinhui Liu[2,*], Zhengjun Shan[1], Lili Shi[1]

1. Nanjing Institute of Environmental Sciences/Key laboratory of Pesticide Environmental Assessment and Pollution Control, Ministry of Environmental Protection, Nanjing 210042, China. E-mail: wanglei@nies.org
2. State Key Laboratory of Water Environment Simulation, School of Environment, Beijing Normal University, Beijing 100875, China

Abstract

A quantitative structure-activity relationship (QSAR) model on depuration rate constants (k_d) of polychlorinated biphenyls (PCBs) in freshwater mussel *Elliptio complanata* was successfully constructed using electrotopological state indices (ESI) and partial least squares (PLS) regression. The cross validated Q^2_{cum} and the correlation coefficient R for the model were determined to be 0.845 and 0.954, respectively. The satisfactory Q^2_{cum} and R values indicated significantly high robustness and good predictive ability for the model. The model was tested and found acceptable for the prediction of $\log k_d$ (the logarithm of the depuration rate constants) by validation set. According to the model, an increase in the values of S_{aasC}, $S_{5'}$, S_4, S_5 and $S_{4'}$ led to increased $\log k_d$, and a decrease in the values of N_{Cl}, $S_{2'}$ and S_6 also resulted in increased $\log k_d$. Among these descriptors, N_{Cl}, S_{aasC}, $S_{5'}$, S_4 and S_5 made significant contributions to the value of $\log k_d$. These significant descriptors showed that the depuration of PCBs in *Elliptio complanata* may be mainly attributed to an equilibrium partitioning process among compartments with different lipid contents, while the reactivity of PCBs with enzymes or other molecules may play a subordinate role.

Key words: depuration; polychlorinated biphenyls; electrotopological state indices; QSAR; *Elliptio complanata*

DOI: 10.1016/S1001-0742(09)60287-4

Introduction

Polychlorinated biphenyls (PCBs), due to their high hydrophobicity and low degradability, have now attracted close international attention. Extensive applications of PCBs since 1929 as heat transfer fluids, organic diluents, plasticizers and paint additives have now led to their ubiquitous existences in soil, water, sediment and organisms (Mandalakis et al., 2008). Chronic exposure to PCBs may cause a wide range of toxic and biological effects such as immune deficiency, reproductive failure, teratogenesis and abnormal behaviors in both animals and humans (Iwata et al., 2004; Zhang et al., 2002).

Despite of the persistence of PCBs, they are found to be depurated to different extents in aquatic organisms (Drouillard et al., 2007; Morrison et al., 1995; O'Rourke et al., 2004; Rodríguez-Ariza et al., 2003). The occurrence of this biological process counteracts the effect of bioaccumulation more or less, and should thus be taken into account in making any ecotoxicological risk assessment. Once the contaminants are ingested by the bivalves, both bioaccumulation and depuration processes begin simultaneously until chemical equilibrium is achieved between the organisms and the surrounding environment. During the process of depuration, bivalves eliminate these xenobiotics to water, feces, or some other surrounding media as well as metabolizing by biotransformation (Drouillard et al., 2007; Van der Linde et al., 2001). In the case of biotransformation, parent compounds are modified by enzyme-mediated reactions into residual metabolites, which are more hydrophilic and, thus, possibly better excretable. It is thought that the effects of biotransformation may also be adverse, resulting in more toxic organic metabolites (Van der Linde et al., 2001). Therefore, it is necessary to understand how organisms depurate contaminants of persistent organic pollutants (POPs) like PCBs.

Since the 1970s, a few researchers have studied the depuration of PCBs in aquatic organisms. It has been generally accepted that equilibrium partitioning is the major factor determining the uptake and release rates of lipophilic pollutants like PCBs in gill-breathing aquatic animals (Drouillard et al., 2007; Van der Linde et al., 2001). Among these animals, bivalves, as surrogates for biomonitoring, have attracted great attention. Although biotransformation in bivalves is much less extensive than in higher invertebrates or vertebrates (Livingstone et al., 1992), less chlorinated congeners ($\leqslant$ 3Cl atoms) are reported to be oxidized by cytochrome P450 isozymes to quinones, that may undergo redox cycling (Mclean et al., 2000). Highly chlorinated PCBs are relatively resistant

to biotransformation, but have been found to undergo transference among different biological compartments (Rodríguez-Ariza et al., 2003; Antunes et al., 2007). In addition, the degree of chlorination and the position of chlorine substituents have been found to qualitatively and quantitatively affect the toxicological properties of PCBs. For example, in flounders and rats, PCBs with free *meta-* (*m-*) and *para-*(*p-*) positions seem to favor elimination (Goerke and Weber, 2001; Kato et al., 1980). Nevertheless, the exact bio-depuration mechanism for PCBs is still unclear.

Given the special position of mussels as biomonitors in monitoring contaminant cycling in aquatic ecosystems, it is desirable to determine the depuration rate coefficients (k_d) of PCBs to enhance the range of their use. Prior to their use, it is essential to know the time for the organism to reach equilibrium with surrounding environment. The time to achieve the equilibrium may just be calculated from the depuration rate constants of chemicals in reference organisms. Even the deployment period as a biomonitor is shorter than the time required for equilibrium, k_d values may be used in conjunction with the toxic exposure time to adjust time-dependent residue concentrations to equilibrium values (O'Rourke et al., 2004). Accordingly, it is of great importance to explore the depuration rate constants of hazardous PCBs.

In order to expediently interpret the depuration in bivalves with minimal expenses, the QSAR model is considered an effective tool. The main concept of the QSAR is to formulate a model that can express chemical/biological activities of compounds in terms of molecular descriptors. With a good QSAR model, the important aspect of the molecular structure may be screened out which significantly affects molecular activity, and the activities of chemicals can also be predicted in the absence of experimental data. The model has been applied broadly in chemical research areas, such as toxicity, reaction rate constant and estrogen activity level determinations (Liu et al., 2003; Long and Niu, 2007; Lv et al., 2008; OECD, 2009a, 2010). However, previous QSAR models of the depuration rate constants have been only constructed using octanol/water partitioning coefficients (K_{ow}) (Drouillard et al., 2007; O'Rourke et al., 2004). And K_{ow} values allow accurate description of the liquid-liquid partitioning process but are insufficient to describe partitioning process or other complicated processes in organisms (Liu et al., 2003). In comparison, the electrotopological state indices (ESI) encode the intrinsic electronic state of the atom perturbed by the electronic influence of all atoms in the molecule within the context of topological character of the molecule. Thus, ESI makes it possible to consider submolecular influences which may contribute toward intermolecular phenomena among biologically important molecules in processes as bio-depuration. Models employing ESI facilitate visualization of reaction mechanisms at both atomic and fragment levels (Hall and Kier, 1995). Using ESI, many robust QSAR models have been developed (Liu et al., 2006; Wang et al., 2007; Kar and Roy, 2010), however as yet none has been developed for the estimation of

depuration rates of PCB congeners in bivalves.

Among various QSAR model-developing techniques, Partial Least Squares (PLS) regression has been adopted in the present study due to its strength to analyze data with noisy, collinear and even incomplete variables in both the independent and dependent variables. In addition, PLS has the desirable quality of improving parameter precision with increases in the number of relevant observations (Svante et al., 2001).

In this article, a QSAR model for the depuration rates of PCB congeners in the freshwater mussel, *Elliptio complanata*, was developed using ESI and PLS regression. The speculations as to the mechanism of depuration in *E. complanata* were also made and discussed.

1 Materials and methods

1.1 Data set

The depuration rate constants (k_d) of 34 PCBs in freshwater mussels (*E. complanata*) were cited from the literature (Drouillard et al., 2007; O'Rourke et al., 2004) and then converted into the form of logk_d. The set of 34 PCB congeners considered in this study was divided into two groups, called the training set (Table 1) and the validation set (Table 2). PCB congeners included in the validation set were randomly chosen, and at least one PCB congener from each level of chlorination was included in the validation set to ensure that sampling was representative.

1.2 Descriptor generation

The E-state indices of all atoms in the compound molecular skeleton were calculated following the method given by Hall and Kier (1995). Besides, all atoms of the same

Table 1 Experimental and predictive logk_d values of PCBs for the training set

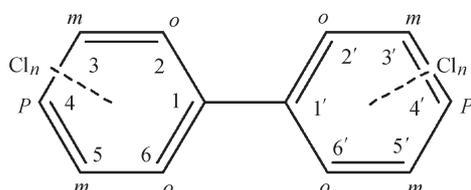| PCB congeners	Observed logk_d	Predicted logk_d	Residuals
PCB 22	−1.15	−1.18	0.03
PCB 23	−1.38	−1.26	−0.12
PCB 42	−1.35	−1.34	−0.01
PCB 66	−1.51	−1.48	−0.03
PCB 74	−1.55	−1.63	0.08
PCB 87	−1.60	−1.76	0.16
PCB 91	−1.55	−1.51	−0.04
PCB 92	−1.68	−1.70	0.02
PCB 95	−1.54	−1.58	0.04
PCB 97	−1.64	−1.70	0.06
PCB 99	−1.66	−1.80	0.14
PCB 109	−1.59	−1.63	0.04
PCB 110	−1.60	−1.51	−0.09
PCB 118	−1.85	−1.80	−0.05
PCB 128	−1.72	−1.93	0.20
PCB 136	−1.75	−1.66	−0.09
PCB 138	−2.16	−2.07	−0.09
PCB 141	−2.16	−2.13	−0.02
PCB 146	−2.00	−2.01	0.01
PCB 151	−2.00	−1.94	−0.06
PCB 156	−2.05	−2.06	0.02
PCB 157	−2.30	−2.08	−0.22
PCB 173	−2.22	−2.27	0.05
PCB 178	−2.22	−2.17	−0.05

Table 2 Experimental and predictive $\log k_d$ values of PCBs for the validation set

PCB congeners	Observed $\log k_d$	Predicted $\log k_d$	Residuals
PCB 7	−0.96	−1.13	0.17
PCB 19	−1.08	−1.09	0.01
PCB 61	−1.50	−1.67	0.17
PCB 85	−1.72	−1.70	−0.03
PCB 105	−1.82	−1.70	−0.13
PCB 130	−1.16	−1.53	0.37
PCB 134	−1.10	−1.30	0.21
PCB 137	−2.22	−2.06	−0.16
PCB 149	−1.32	−1.89	0.57
PCB 179	−2.30	−2.04	−0.26

type were grouped and their ESI values were summed to give the atom-type ESI. Likewise, the ESI values of all atoms in a functional group were summed to give the group-type E-state indices. The ESI descriptors considered in this article consisted of the 12 numbered atom ESI, named S_n ($n = 1, 2, 3, 4, 5, 6, 1', 2', 3', 5', 6'$, according to the primary structure of PCB molecule in Fig. 1), 3 atom-type ESI and 2 group-type ESI. In addition, as an important character of PCBs, the number of chlorine atoms (N_{Cl}) was also included. The molecular skeleton and ESI of IUPAC PCB-13 is exemplified in Tables 3 and 4.
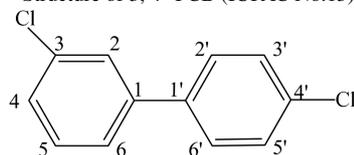
1.3 Statistical analysis

The QSAR model with quantum chemical descriptors was developed using PLS regression, as implemented in the Simca-S package (Umetrics AB, Sweden). Model

Fig. 1 Primary structure for PCBs. *o*, *m*, *p* denote *ortho*, *meta* and *para* positions, respectively.

Table 3 E-state indices for the molecular skeleton atoms of 3,4′-PCB*

Atom ID	Atom structure	Atom type	E-state symbol	Intrinsic state	E-state value
1	aaaC	aaaC	S_1	1.667	1.106
2	aCHa	aaCH	S_2	2.000	1.935
3	aCHa	aaCH	S_3	2.000	0.748
4	aCa-	aasC	S_4	1.667	1.862
5	aCHa	aaCH	S_5	2.000	1.937
6	aCHa	aaCH	S_6	2.000	2.023
1′	aaaC	aaaC	$S_{1'}$	1.667	1.123
2′	aCHa	aaCH	$S_{2'}$	2.000	1.980
3′	aCHa	aaCH	$S_{3'}$	2.000	1.873
4′	aCHa	aaCH	$S_{4'}$	2.000	0.747
5′	aCa-	aasC	$S_{5'}$	1.667	1.873
6′	aCHa	aaCH	$S_{6'}$	2.000	1.980

* Structure of 3, 4′-PCB (IUPAC No.13):

Table 4 Atom-type E-state indices and grouped ones of 3,4'-PCB

E-state indices	E-state symbol	E-state value
Atom-type E-state indices	S_{aasC}	3.724
	S_{aaCH}	15.46
	S_{sCl}	11.70
Grouped E-state indice	S_M	19.19
	S_{SUM}	30.89

dimensionality was determined by cross-validation. The robustness and predictive power of the model were assessed using Q^2_{cum} (cumulative Q^2). When Q^2_{cum} is larger than 0.5, the model is considered to have a good predictive ability. Q^2_{cum} can be calculated as Eq. (1)

$$Q^2 = 1.0 - \left[\sum_i \sum_m \left(Y_{im,obs} - Y_{im,pred} \right)^2 \right] \Big/ SS \qquad (1)$$

$$Q^2_{cum} = 1.0 - \left\{ \prod \left[\sum_i \sum_m \left(Y_{im,obs} - Y_{im,pred} \right)^2 \right] \Big/ SS \right\}_a \qquad (2)$$

where, $Y_{im,obs}$ and $Y_{im,pred}$ denote the observed and predicted $\log k_d$ values, respectively. i stands for different observations in the training set, m stands for different dependent variables ($m = 1$ for this study), SS is the residual sum of squares of the previous component, and $a = 1, 2, ..., A$ (the number of PLS principle components). In addition, the standard deviation (SD) was also adopted to assess prediction precision of model. SD can be calculated by Eq. (3):

$$SD = \sqrt{\frac{1}{n - A - 1} \sum_{i=1}^{n} \left(Y_{im,obs} - Y_{im,pred} \right)^2} \qquad (3)$$

Model adequacy was mainly measured by the number of PLS principal components (A), Q^2_{cum}, the correlation coefficient (R) between observed values and predicted values, and the significance level (p). The best PLS model was selected with respect to the statistics Q^2_{cum}, R, p and SD.
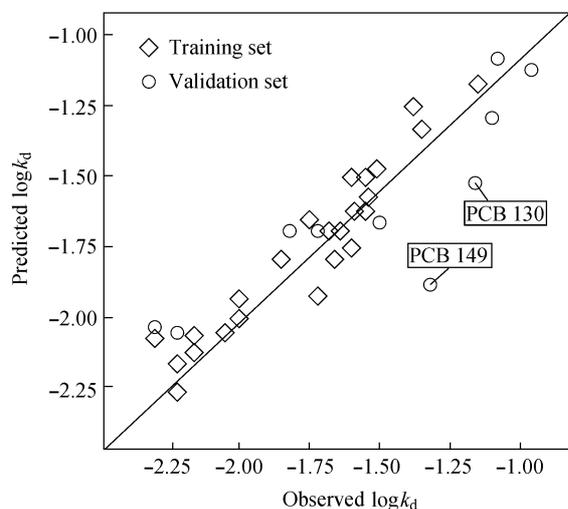
2 Results and discussion

2.1 Results

Based on the unscaled regression coefficients of the independent variables and a constant transformed from PLS regression, a QSAR regression model for the training set was obtained for *E. complanata* as follows.

$$\begin{aligned}
\log k_d = &-2.17 - 1.26 \times 10^{-1} N_{Cl} + \\
&1.58 \times 10^{-1} S_{aasC} + 1.41 \times 10^{-1} S_{5'} + \\
&1.48 \times 10^{-1} S_4 + 1.12 \times 10^{-1} S_5 - \\
&8.30 \times 10^{-4} S_{2'} + 9.22 \times 10^{-2} S_{4'} - \\
&9.74 \times 10^{-3} S_6
\end{aligned} \qquad (4)$$

The concrete results of the model are listed in Table 5. $R^2_{X,adj,cum}$ and $R^2_{Y,adj,cum}$ stand for the cumulative variance of all the independent and dependent variables, respectively, explained by the extracted components. Eigenvalue is

Table 5   Model fitting results

A	Eigenvalue	$R^2_{X,\text{adj,cum}}$	$R^2_{Y,\text{adj,cum}}$	Q^2_{cum}	SD	R	p
1st	2.869	0.234	0.769	0.728	0.098	0.954	5.518×10^{-13}
2nd	1.547	0.346	0.901	0.845			

Fig. 2   Observed versus predicted $\log k_\text{d}$ values for the model.

the eigenvalue that denotes the importance of the PLS principal components. It can be concluded that two PLS principle components were selected in the QSAR model, which explain 34.6% and 90.1% of the variance for the independent and dependent variables, respectively. The higher eigenvalue of the first principle component indicates its greater contribution to the model comparing to the second one. Q^2_{cum} value of 0.845 is much higher than 0.5, indicating the inherent stability of the model. A plot between the observed and the predicted $\log k_\text{d}$ values (Fig. 2) gives a correlation coefficient of 0.954. The standard deviation for the model is 0.098, which is about 5.51% of the average $\log k_\text{d}$ (observed) for the PCB congeners included in the training set. All the residuals for PCB congeners in the training set are below 3SD. The high correlation coefficient and the low standard deviation show that the QSAR model is reliable and it can be used for prediction.

Based on the QSAR model, $\log k_\text{d}$ for the PCB congeners in the validation set were predicted, as shown in Table 2. the external prediction capability of the model was quantified using the predictive squared correlation coefficient (q^2), average relative error ($\overline{\text{RE}}$) and the root mean square error of prediction (RMSEP) which measures the average difference between predicted and experimental values at the prediction stage. They were obtained by the following equations.

$$\overline{\text{RE}} = \frac{1}{n} \sum_{j=1}^{n} \left[\left(Y_{j,\text{obs}} - Y_{j,\text{pred}} \right) \big/ Y_{j,\text{obs}} \right] \times 100\% \tag{5}$$

$$\text{RMSEP} = \sqrt{\frac{1}{n} \sum_{j=1}^{n} \left(Y_{j,\text{obs}} - Y_{j,\text{pred}} \right)^2} \tag{6}$$

$$q^2 = 1 - \frac{\sum\limits_{j=1}^{N} \left(Y_{j,\text{obs}} - Y_{j,\text{obs}} \right)^2}{\sum\limits_{j=1}^{N} \left(Y_{j,\text{obs}} - Y_{\text{mean}}^{\text{test}} \right)^2} \tag{7}$$

where, $Y_{j,\text{obs}}$ and $Y_{j,\text{pred}}$ represent the observed and predicted $\log k_\text{d}$ values, respectively for the PCB congeners in the validation set. $Y_{\text{mean}}^{\text{text}}$ represents the average of observed $\log k_\text{d}$ values in the whole test set and j stands for different observations in the validation set. The q^2 for the validation was calculated to be 0.718, and the values of $\overline{\text{RE}}$ and RMSEP set are 0.058% and 0.248, respectively. As shown in Fig. 2, all residuals for PCB congeners in the validation set are below 3SD except two compounds (PCB 130 and PCB 149). The same deviation appearing in other modeling process of these data suggests that excessive prediction residuals of the two compounds may be derived from the original experiment (Xu et al., 2009). According to the guidance document of the validation of the model (OECD, 2009b; Schuurmann et al., 2008), QSAR model is proved to be reliable.

2.2 Discussion

In the QSAR model, variable importance in the projection (VIP) is a parameter showing the importance of a variable in a PLS model. Independent variables with large values of VIP, larger than 1, are the most relevant for explaining the dependent variable. In the developed model, $\log k_\text{d}$ is correlated to eight predictor variables N_{Cl}, S_{aasC}, $S_{5'}$, S_4, S_5, $S_{2'}$, $S_{4'}$ and S_6. Their corresponding VIP values are 1.624, 1.122, 1.000, 0.917, 0.902, 0.806, 0.645 and 0.619, respectively. The VIP value for S_4 and S_5 are very close to 1, thus descriptors of N_{Cl}, S_{aasC}, $S_{5'}$, S_4 and S_5 made great contributions to $\log k_\text{d}$. The increase in the values of S_{aasC}, $S_{5'}$, S_4, S_5 and $S_{4'}$ leads to increased $\log k_\text{d}$, while the decrease in the values of N_{Cl}, $S_{2'}$ and S_6 results in increased $\log k_\text{d}$. The values of the ESI descriptors for PCB congeners are listed in Table 6.

Previous publications have demonstrated that PCBs in organisms may act in at least three ways (McKinney and Waller, 1994). (1) Through the accumulation of highly lipid-soluble, metabolically stable PCBs in lipid-rich tissues or tissue compartments, which can be seen as an equilibrium partitioning gradient, ranging from lipid-rich compartments to water-rich compartments. (2) A reversibly binding interaction of the PCB with specific molecular sites of action such as receptors, enzymes. (3) An irreversibly covalent binding interaction between the PCB (probably a reactive metabolite) and target molecules (particularly macromolecules such as DNA and proteins). Different mechanisms involved in these actions would in turn govern the depuration property of PCBs. Moreover, because metabolic biotransformation in bivalves is much

Table 6 E-state indices values for the descriptors entered in the model

PCB congeners	S_4	S_5	S_6	$S_{2'}$	$S_{4'}$	$S_{5'}$	S_{aasC}	N_{Cl}
PCB 7	0.662	1.843	1.930	2.023	1.979	1.986	3.465	2
PCB 19	1.810	1.798	0.608	0.647	1.872	1.904	3.503	3
PCB 22	1.767	1.873	1.928	1.916	0.706	1.827	3.763	3
PCB 23	1.637	0.587	1.810	1.966	1.948	1.945	3.462	3
PCB 42	1.733	1.828	1.864	0.553	0.590	1.764	3.765	4
PCB 61	0.286	0.399	1.715	1.932	1.927	1.919	3.124	4
PCB 66	0.606	1.783	1.851	1.777	0.524	1.740	4.019	4
PCB 74	0.449	0.483	1.745	1.883	0.681	1.801	3.974	4
PCB 85	0.409	1.676	1.769	0.513	0.565	1.738	3.672	5
PCB 87	0.402	1.669	1.758	0.551	1.707	0.576	3.631	5
PCB 91	1.638	1.676	0.490	0.476	0.550	1.718	3.641	5
PCB 92	1.577	0.492	1.701	0.537	1.702	0.567	3.748	5
PCB 95	1.633	1.669	0.476	0.515	1.694	0.553	3.584	5
PCB 97	1.707	1.794	1.819	0.476	0.402	0.419	3.631	5
PCB 99	0.409	0.429	1.682	0.513	0.565	1.738	3.895	5
PCB 105	0.418	1.688	1.788	1.731	0.493	1.706	3.709	5
PCB 109	0.327	1.563	0.437	1.763	1.768	1.812	3.403	5
PCB 110	1.646	1.688	0.513	1.701	0.477	1.687	3.704	5
PCB 118	0.418	0.442	1.700	1.731	0.493	1.706	3.932	5
PCB 128	0.377	1.643	1.724	0.324	0.377	1.643	3.253	6
PCB 130	0.371	1.635	1.712	0.363	1.556	0.460	3.357	6
PCB 134	1.482	0.304	0.288	0.349	1.668	1.741	2.996	6
PCB 136	1.599	1.623	0.399	0.288	1.599	1.623	3.056	6
PCB 137	0.220	0.313	1.618	0.458	0.534	1.704	3.325	6
PCB 138	0.377	1.643	1.724	0.436	0.377	0.388	3.476	6
PCB 141	0.214	0.304	1.606	0.497	1.681	0.536	3.266	6
PCB 146	1.556	0.460	1.668	0.422	0.371	0.379	3.579	6
PCB 149	1.612	1.643	0.436	0.399	0.362	0.365	3.398	6
PCB 151	1.482	0.304	0.288	0.460	1.668	0.512	3.219	6
PCB 156	0.229	0.327	1.636	1.686	0.462	1.672	3.390	6
PCB 157	0.386	1.654	1.742	1.668	0.305	0.349	3.435	6
PCB 173	0.098	0.115	0.172	0.308	1.647	1.715	2.283	7
PCB 178	1.455	0.263	0.233	0.272	1.517	0.397	2.817	7
PCB 179	1.448	0.249	0.211	0.233	1.573	1.590	2.581	7

less extensive than in higher invertebrates or vertebrates (Livingstone et al., 1992), it has been generally accepted that equilibrium partitioning is the major factor in determining the uptake and release rates of lipophilic PCBs in bivalves (Drouillard et al., 2007; Van der Linde et al., 2001). Thus, via three sub-processes ways might the structural properties of PCBs influence the depuration rates, which can be interpreted by the QSAR model as follows.

According to the model, the chlorine-atom number (N_{Cl}) plays a significant role in PCBs depuration. This parameter possesses the largest VIP value (1.624) among all of the molecular descriptors. One negative aspect of N_{Cl} shows that as the number of Cl-atom substitutes increases, the rate of depuration decreases. This may be attributed to restricted partitioning from lipid-rich tissues to water or water-rich compartments, caused by the greater steric hindrance effect and polarization brought on by Cl-atom increases. On one hand, greater steric hindrance would prevent the transfer of bulky PCB molecules from lipid-rich tissues to water or water-rich compartments; on the other hand, increases in polarization due to Cl-atom increases would probably impart a preferred vector of polarization to entire PCB molecule (Harper et al., 1993). This could strengthen PCB-liquid binding considering that the vector of polarization of water molecules is lower compared with that of lipid molecules (McKinney and Waller, 1994). As a

result, highly lipid-soluble PCBs with greater chlorination degree are more difficult to enter the water-soluble phase to be eliminated.

With a great VIP value of 1.122, S_{aasC} contributes significantly to the model possibly representing the reactivity of the PCB molecule with metabolic enzymes. As shown in Fig. 3, the average S_{aasC} value increases with enhanced chlorination until the chlorine degree reaches 3–4 atoms, and then decreases consistently with further increase of substituted chlorines. With the positive influence of S_{aasC} indicated by the model, it might be deduced that the depuration rate rises with an increase of chlorine degree when $N_{Cl} \leqslant 3$–4, while the trend reverses when $N_{Cl} > 4$. This agrees with the previous hypothesis that highly chlorinated PCBs are relatively resistant to biotransformation, while less chlorinated congeners ($\leqslant 3$Cl atoms) are oxidized by cytochrome P450 isozymes to quinones, that may undergo redox cycling (Mclean et al., 2000).

In addition, $S_{5'}$, S_4 and S_5 are also of great importance to the model. Given the primary structure of two phenyl rings in the skeleton construction of the PCB molecule, $S_{5'}$ and S_5 represent the ESI of *meta*-carbon, while S_4 represents the ESI of *para*-carbon (Fig. 1). Due to the strength of the Cl-polarization vector, the density of the electron cloud among the C–Cl bond is greater in the vicinity in the chlorine atom, leading to a decrease in the ESI value of the carbon atom. Therefore, the introduction
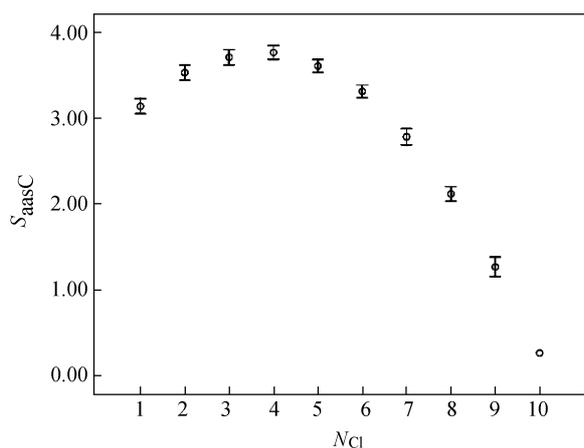
Fig. 3   Average value of S_{aasC} for PCBs with different chlorination degree.

of further chlorines would result in a decrease in the ESI of the binding carbon. Take $S_{5'}$ for example, of all three chlorinated PCBs (PCB16–PCB 39), the $S_{5'}$ value of the PCB with Cl-substitution at the 5′ position (PCB 34, $S_{5'}$ = 0.609) is much less than that of others ($S_{5'}$ is in the range of 1.766–1.964). Accordingly, the Cl^- influences of $S_{5'}$, S_4 and S_5 on the depuration rates of PCBs are shown in the negative effects of lateral (*meta-*, *para-*) Cl^- substitution. This correlates with the previously recognized role of lateral chlorination in PCB binding interactions and toxicity (McKinney et al., 1985, 1987; Rickenbacher et al., 1986). This might be attributed to three factors. First, bulky chlorine atoms in *ortho*-positions would hinder the free rotation of the phenyl rings, creating potential barriers to coplanar conformality which are higher for *ortho*-substituted PCBs than lateral-substituted PCBs (E et al., 2006; McKinney and Waller, 1994; Safe et al., 1985). Therefore, lateral-substituted PCBs are more likely to be found in coplanar form, which facilitates them "sticking together" with other planar aromatic ring systems, such as the heme system in hemoproteins or other planar metabolic intermediates (McKinney and Waller, 1994). These types of interactions would tend to prevent lateral-substituted PCBs from partitioning into water or water-rich compartments to be eliminated. Second, halogen atoms like chlorine contain many electrons and are thus highly polarizable. In PCB molecules, the most polarizable chlorines are those contained in the lateral positions (McKinney and Waller, 1994). This is, no doubt, to facilitate stacking interactions between halogen substituants and hydrophobic macromolecules, making the PCB molecule difficult to partition into the hydrophilic compartments. Third, lateral-substituted PCBs possess weak oxidative potential, displaying poor hydrophilicity and difficulties in compartmental partitions, due to limited availability of vacant *meta-*, *para-* positions that can provide sites for oxidative metabolism (Kato et al., 1980). The disruptive effects influences of $S_{2'}$ and S_6 (ortho-substitutions) as well as the facilitating influences of $S_{4'}$ (lateral-substitutions) in this model are also in accord with this hypothesis.

3 Conclusions

A robust QSAR model for analysis of the depuration rates of PCBs in *Elliptio complanata* was developed using an approach employing ESI and PLS regression. It is clear that the depuration of PCB congeners may be mainly attributed to a compartmentalized equilibrium partitioning process, possibly an isomeric gradient of varying hydro-lipophilicity. In addition, the reactivity of PCB molecules with enzymes or other molecules may play a subordinate role in depuration. Using this model, the depuration rate constants of PCB congeners in *E. complanata* were predicted accurately and conveniently using literature references. This optimal model enables an estimation of the time required for the fresh water mussel *E. complanata* to achieve self-depuration and facilitates to establish appropriate biomonitoring deployment periods. By this way, the optimal model can allow effective depletion of xenobiotics in aquaculture products, and improve food safety and aquaculture sustainability.

References

Antunes P, Gil O, Ferreira M, Vale C, Reis-Henriques M A, 2007. Depuration of PCBs and DDTs in mullet under captivity clean conditions. *Chemosphere*, 67: S58–S64.

E Y X, Wang T, Li L, Chen D Z, 2006. A theoretical study of the molecular structure of polychlorinated biphenyls. *Journal of Beijing University of Chemical Technology*, 33: 81–84.

Drouillard K G, Chan S, O'Rourke S, Haffner G D, Letcher R J, 2007. Elimination of 10 polybrominated diphenyl ether (PBDE) congeners and selected polychlorinated biphenyls (PCBs) from the freshwater mussel, *Elliptio complanata*. *Chemosphere*, 69: 362–370.

Goerke H, Weber K, 2001. Species-specific elimination of polychlorinated biphenyls in estuarine animals and its impact on residue patterns. *Marine Environmental Research*, 51: 131–149.

Hall L H, Kier L B, 1995. Electrotopological state indices for atom types: A novel combination of electronic, topological and valence state information. *Journal of Chemical Information and Modeling*, 35: 1039–1045.

Harper N, Howie L, Connor K, Arellano L, Craig A, Dickerson R et al., 1993. Immunosuppressive and monooxygenase induction activities of highly chlorinated diphenyl ether congeners in C57BL/6 and DBA/2 mice. *Fundamental and Applied Toxicology*, 20: 496–502.

Iwata H, Watanabe M, Okajima Y, Tanabe S, Amano M, Miyazaki N et al., 2004. Toxicokinetics of PCDD, PCDF and coplanar PCB congeners in Baikal seals, *pusa sibirica*: age-related accumulation, maternal transfer, and hepatic sequestration. *Environmental Science & Technology*, 38:

3505–3513.

Kar S, Roy K, 2010. QSAR modeling of toxicity of diverse organic chemicals to *Daphnia magna* using 2D and 3D descriptors. *Journal of Hazardous Materials*, 177(1-3): 344–351.

Kato S, McKinney J D, Matthews H B, 1980. Metabolism of hexachlorobiphenyls in the rat. *Toxicology and Applied Pharmacology*, 53: 389–398.

Liu X H, Hou J, Wang L, Luo W R, Cui B S, 2006. β-Cyclodextrin-enhanced solubilization of phenylsulfonyl carboxylates in water. *Bulletin of Environmental Contamination and Toxicology*, 77: 51–59.

Liu X H, Wang B, Huang Z, Han S, Wang L S, 2003. Acute toxicity and quantitative structure-activity relationships of α-branched phenylsulfonyl acetates to *Daphnia magna*. *Chemosphere*, 50: 403–408.

Livingstone D R, Donkin P, Walker C H, 1992. Pollutants in marine ecosystems. In: Persistent Pollutants in Marine Ecosystems (Walker C H, Livingstone D R, eds). Oxford Publishers, Oxford, UK. 235–263.

Long X X, Niu J F, 2007. Estimation of gas-phase reaction rate constants of alkylnaphthalenes with chlorine, hydroxyl and nitrate radicals. *Chemosphere*, 67: 2028–2034.

Lv Y Y, Yin C S, Liu H Y, Yi Z S,Wang Y, 2008. 3D-QSAR study on atmospheric half-lives of POPs using CoMFA and CoMSIA. *Journal of Environmental Sciences*, 20(12): 1433–1438.

Mandalakis M, Holmstrand H, Andersson P, Gustafsson O, 2008. Compound-specific chlorine isotope analysis of polychlorinated biphenyls isolated from Aroclor and Clophen technical mixtures. *Chemosphere*, 71: 299–305.

McKinney J D, Chae K, McConnell E E, Birnbaum L, 1985. Structure-induction versus structure-toxicity relationships for polychlorinated biphenyls and related aromatic hydrocarbons. *Environmental Health Perspectives*, 60: 57–68.

McKinney J D, Fannin R, Jordan S, Chae K, Rickenbacher U, Darden T et al., 1987. Polychlorinated biphenyls (PCBs) and related compound interactions with specific binding sites for thyroxine in rat liver nuclear extracts. *Journal of Medicinal Chemistry*, 30: 79–86.

McKinney J D, Waller L C, 1994. Polychlorinated biphenyls as hormonally active structural analogues. *Environmental Health Perspectives*, 102: 290–297.

Mclean M R, Twaroski T P, Robertson L W, 2000. Redox cycling of 2-(x'-mono, -di, -trichlorophenyl)-1,4-benzoquinones, oxidation products of polychlorinated biphenyls. *Archives of Biochemistry and Biophysics*, 376: 449–455.

Morrison H, Yankovich T, Lazar R, Haffner G D, 1995. Elimination rate constants of 36 PCBs in zebra mussels (*Dreissena polymorpha*) and exposure dynamics in the lake St. Clair-Lake Erie corridor. *Canadian Journal of Fisheries and Aquatic Sciences*, 52: 2574–2582.

OECD (Organisation for Economic Co-operation and Development), 2010. Report of the expert consultation on scientific and regulatory evaluation of organic chemistry mechanism-based structural alerts for the identification of DNA-binding chemicals. Paris, France.

OECD (Organisation for Economic Co-operation and Development), 2009a. Report of the expert consultation to evaluate an estrogen receptor binding affinity model for hazard identification. Paris, France.

OECD (Organisation for Economic Co-operation and Development), 2009b. Guidance document on the validation of (Q)SAR models. Paris, France.

O'Rourke S, Drouillard K G, Haffner G D, 2004. Determination of laboratory and field elimination rates of polychlorinated biphenyls (PCBs) in the freshwater mussel, *Elliptio complanata*. *Archives of Environmental Contamination and Toxicology*, 47: 74–83.

Rickenbacher U, McKinney J D, Oatley S J, Blake C C F, 1986. Structurally specific binding of halogenated biphenyls to thyroxine transport protein. *Journal of Medicinal Chemistry*, 29: 641–648.

Rodríguez-Ariza A, Rodríguez-Ortega M J, Marenco J L, Amezcua O, Alhama J, López-Barea J, 2003. Uptake and clearance of PCB congeners in *Chamaelea gallina*: response of oxidative stress biomarkers. *Comparative Biochemistry and Physiology*, C134: 57–67.

Safe S, Bandiera S, Sawyer T, Robertson L, Safe L, Parkinson A et al., 1985. PCBs: structure-function relationships and mechanism of action. *Environmental Health Perspectives*, 60: 47–56.

Schuurmann G, Ebert R-U, Chen J W, Wang B, Kuhne R, 2008. External validation and prediction employing the predictive squared correlation coefficient -test set activity mean vs training set activity mean. *Journal of Chemical Information and Modeling*, 48(11): 2140–2145.

Svante W, Michael S, Lennart E, 2001. PLS-regression: A basic tool of chemometrics. *Chemometrics and Intelligent Laboratory Systems*, 58: 109–130.

Van der Linde A, Hendriks A J, Sijm D T H M, 2001. Estimating biotransformation rate constants of organic chemicals from modeled and measured elimination rates. *Chemosphere*, 44: 423–435.

Wang L, Liu X H, Hou J, Cui B S, 2007. Prediction of photolysis half-lives of PCDFs with the electrotopological state indices. *Acta Chimica Sinica*, 65: 184–190.

Xu M Z, Liu X H, Wang L, Wu D, Yang Z F, Cui B S, 2009. Quantitative structure-activity relationship for the depuration rate constants of polychlorinated biphenyls in the freshwater mussel, *Elliptio complanata*. *Journal of Environmental Science and Health, Part B*, 44(3): 278–283.

Zhang C Q, Fang C G, Liu L, Xia G L, Qiao H L, 2002. Disruption effects of polychlorinated biphenyls on gonadal development and reproductive functions in chickens. *Journal of Environmental Science and Health*, A37: 509–519.